

Supplementary Information

Discovery of Type II Polyketide Synthase-like Enzymes for the Biosynthesis of Cispentacin

Genki Hibi,^a Taro Shiraishi,^{a,b} Tatsuki Umemura,^a Kenji Nemoto,^a Yusuke Ogura,^a Makoto Nishiyama,^{a,b} Tomohisa Kuzuyama^{*a,b}

^aGraduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan; ^bCollaborative Research Institute for Innovative Microbiology, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan.

Table of Contents

Supplementary Tables	1
Supplementary Table 1. Deduced function of the <i>amc</i> genes	1
Supplementary Table 2. List of primers used in this study.	2
Supplementary Table 3. Strains used in this study.	3
Supplementary Table 4. Vectors used in this study.	3
Supplementary Table 5. The characterization of chemical and biomolecular materials.	4
Supplementary Figures.....	5
Supplementary Figure 1. Comparison of type II PKS subfamilies.	5
Supplementary Figure 2. Biosynthetic pathway of CPC-ACP postulated in previous studies.	6
Supplementary Figure 3. Schematic representation of the plasmid constructs.	6
Supplementary Figure 4. X-LC and MS analysis of DNFB-derivatized cispentacin.	8
Supplementary Figure 5. Production of cispentacin by <i>S. albus</i> G153/pSEamcB-H.	9
Supplementary Figure 6. SDS–PAGE analysis of purified recombinant proteins.	10
Supplementary Figure 7. Gel filtration of AmcF–AmcG.	11
Supplementary Figure 8. Detection of AmcB bound intermediate derived from malonyl-CoA.	12
Supplementary Figure 9. Substrate specificity of AmcH.	13
Supplementary Figure 10. HPLC analysis of AmcH activity.	14
Supplementary Figure 11. LC–ESI–HRMS analysis of 2-AmcB digested with thermolysin.	15
Supplementary Figure 12. MS/MS spectrum of the PPant-eliminated ion of 2-AmcB.	16
Supplementary Figure 13. Phylogenetic analysis of KSs in the cispentacin cluster and various types of known KSs.	17
Supplementary Figure 14. Sequence analysis of AmcF and AmcG.	18
Supplementary Figure 15. AmcF–AmcG heterodimer fold prediction.	19
Supplementary Figure 16. Assessment of the dimer structure.	20
Supplementary Figure 17. Comparative analysis of CLF and cyclization factors (CYFs).	22
Supplementary Figure 18. Sequence analysis of AmcE.	24
Supplementary Figure 19. UV–vis spectra assay for AmcC.	25
Supplementary Figure 20. LC–MS analysis of cispentacin in the reconstituted reaction mixture.	26
Supplementary Figure 21. Amine donor specificity of AmcC.	27

Supplementary Figure 22. NADH-dependent reduction of 3-AmcB.....	29
Supplementary Figure 23. Validation of the enzyme that catalyzes the reduction of 3-AmcB.	32
Supplementary Figure 24. LC–MS analysis of FDLA-derivatized cispentacins.	33
Supplementary Figure 25. Organization of <i>amcG</i> homolog-containing BGC.....	35
Supplementary Figure 26. Representatives of the <i>amcG</i> homolog-containing BGCs for unknown natural products.	35
Supplementary Figure 27. FabI-catalyzed reduction.	36
Supplementary Figure 28A. ¹ H NMR spectrum of 2,3-di(carbo- <i>tert</i> -butoxy)-2-cyclopentenone in CDCl ₃ at 600 Hz.....	37
Supplementary Figure 28B. ¹³ C NMR spectrum of 2,3-di(carbo- <i>tert</i> -butoxy)-2-cyclopentenone in CDCl ₃ at 150 Hz.....	38
Supplementary Figure 28C. HSQC spectrum of 2,3-di(carbo- <i>tert</i> -butoxy)-2-cyclopentenone in CDCl ₃	39
Supplementary Figure 28D. HMBC spectrum of 2,3-di(carbo- <i>tert</i> -butoxy)-2-cyclopentenone in CDCl ₃	40
Supplementary Figure 29A. ¹ H NMR spectrum of 3-oxocyclopent-1-ene-1,2-dicarboxylic acid in DMSO- <i>d</i> ₆ at 600 Hz.	41
Supplementary Figure 29B. ¹³ C NMR spectrum of 3-oxocyclopent-1-ene-1,2-dicarboxylic acid in DMSO- <i>d</i> ₆ at 150 Hz.	42
Supplementary Figure 30. HRMS spectra of 3-oxocyclopent-1-ene-1,2-dicarboxylic acid.	43
Supplementary References	44

Supplementary Tables

Supplementary Table 1. Deduced function of the *amc* genes

ORFs	Size (aa)	Proposed function*	E-value	Homologous protein [origin]
AmcB	88	Acyl carrier protein (ACP) for PKS	0.001	Acyl carrier protein [<i>Listeria welshimeri</i> serovar 6b str. SLCC5334]
AmcC	471	Aminotransferase	1.00E-34	Putrescine aminotransferase [<i>Cronobacter sakazakii</i> ATCC BAA-894]
AmcD	155	Thioesterase	8.00E-12	1,4-dihydroxy-2-naphthoyl-CoA hydrolase [<i>Parasynecococcus marenigrum</i> WH 8102]
AmcE	167	Dehydratase	3.00E-74	Coronafacic acid dehydratase [<i>Pseudomonas savastanoi</i> pv. glycinea]
AmcF	397	Ketosynthase (KS) for PKS	2.00E-38	3-oxoacyl-[acyl-carrier-protein] synthase 2 [<i>Bacillus subtilis</i> subsp. subtilis str. 168]
AmcG	184	Hypothetical protein	-	No significant similarity found
AmcH	460	Adenylate-forming enzyme	5.00E-12	Putative acyl-CoA synthetase YngI [<i>Bacillus subtilis</i> subsp. subtilis str. 168]

The nucleotide sequence of the *amc* cluster is deposited in the DDBJ/EMBL/GenBank nucleotide sequence database under the accession number [LC389220](#). *The predicted function of ORFs. E-Value and Homologous protein [origin] represent the results obtained by a BLATP search using the UniProtKB/Swiss-Prot (swissprot) database, with scoring parameters set to default (BLOSUM62, Existence: 11 Extension: 1, Conditional compositional score matrix adjustment).

Supplementary Table 2. List of primers used in this study.

Primer	Oligonucleotide sequence	Restriction site (underlined)
<i>E. coli</i>		
AmcB-CHisfw	5'-GGAATT <u>CCATATGA</u> ACCAGACGGAAGCCGCCAT-3'	<i>NdeI</i>
AmcB-CHisrv	5'-CCC <u>AAGCTT</u> GGCCGTCTCGAGTCGGGA-3'	<i>HindIII</i>
AmcC-NHisfw	5'-CGG <u>GATCC</u> ACGTCCGCACCGGACTTCC-3'	<i>BamHI</i>
AmcC-NHisrv	5'-CCC <u>AAGCTT</u> TCATGCGTGCGTCTCGAG-3'	<i>HindIII</i>
AmcD-NHisfw	5'-CGG <u>GATCC</u> ACCGCCGCGCAGCCGGTACC-3'	<i>BamHI</i>
AmcD-NHisrv	5'-CCC <u>AAGCTT</u> TCAACGCTGTGCACGGAC-3'	<i>HindIII</i>
AmcE-NHisfw	5'-CGG <u>GATCC</u> CACAGCGTTGACCCCGAC-3'	<i>BamHI</i>
AmcE-NHisrv	5'-CCC <u>AAGCTT</u> TACCACAGCTGCTCCCCGA-3'	<i>HindIII</i>
AmcH-NHisfw	5'-CGC <u>GATCC</u> AAAGCAGCCCTCCTCCCCGA-3'	<i>BamHI</i>
AmcH-NHisrv	5'-CCC <u>AAGCTT</u> TCACTTGATACAGTCGATGGACCG-3'	<i>HindIII</i>
sFabI-CHisfw	5'-GGC <u>ATATG</u> AGTGGACTTCTCGCAGGCAAG-3'	<i>NdeI</i>
sFabI-CHisrv	5'-GGG <u>AATTC</u> CCGGCGCCGATGGCGTG-3'	<i>EcoRI</i>
eFabI-CHisfw	5'-GGC <u>ATATG</u> GTTTTTCTTCCGGTAAGCG-3'	<i>NdeI</i>
eFabI-CHisrv	5'-GGG <u>AATTC</u> CCCTTTCAGTTCGAGTTCGTTTC-3'	<i>EcoRI</i>
<i>S. albus</i>		
pSEamcB-H-N	5'-GGG <u>AAGCTT</u> GAAGGGAATGCACGATGAACC-3'	<i>HindIII</i>
pSEamcH-NHisfw	5'-GGG <u>AAGCTT</u> AGCAACGGAGGTACGGACATGCACCACCACCACCACCACCACAAAGCAGCCCTCC TCCC-3'	<i>HindIII</i>
pSEamcH-Rv	5'-GGG <u>TCTAG</u> AGTCACTTGATACAGTCGATGG-3'	<i>XbaI</i>
pSEAmcF-NHisfw	5'-CCC <u>AAGCTT</u> AGCAACGGAGGTACGGACATGCACCACCACCACCACCACCACGTGACGCCCCGCC CGCG-3'	<i>HindIII</i>
pSEAmcG-rv	5'-CCCGCT <u>CTAGA</u> GCTCACCTCCGGCCGGCCG-3'	<i>XbaI</i>

Supplementary Table 3. Strains used in this study.

Strains	Characteristic(s)	Source/Reference
DH5 α	Host for general cloning	Takara
BL21 (DE3)	Host for protein expression	Takara
BL21 (DE3)/pET26(+)-amcB/pACYCDuet-pptA2	expressing C-terminal His-tagged AmcB with pptA2	This study
BL21 (DE3)/pHis8-amcC	expressing N-terminal His-tagged AmcC	This study
BL21 (DE3)/pHis8-amcD	expressing N-terminal His-tagged AmcD	This study
BL21 (DE3)/pHis8-amcE	expressing N-terminal His-tagged AmcE	This study
BL21 (DE3)/pET26(+)-sFabI	expressing C-terminal His-tagged sFabI	This study
BL21 (DE3)/pET26(+)-eFabI	expressing C-terminal His-tagged eFabI	This study
<i>Streptomyces</i> sp. SN-C1	Cisptentacin producer	Gift from Kumiai Chemical Industry Co., Ltd.
<i>S. albus</i> G153	Host for heterologous expression	Gift from Prof. Bradley S. Moore
<i>S. albus</i> G153/pSE101	Negative control of heterologous expression	This study
<i>S. albus</i> G153/pSEamcB-H	Recombinant strain/Cisptentacin producer	This study
<i>S. albus</i> G153/pSEamcF-G	expressing N-terminal His-tagged AmcF and AmcG	This study
<i>S. albus</i> G153/pSEamcH	expressing N-terminal His-tagged AmcH	This study

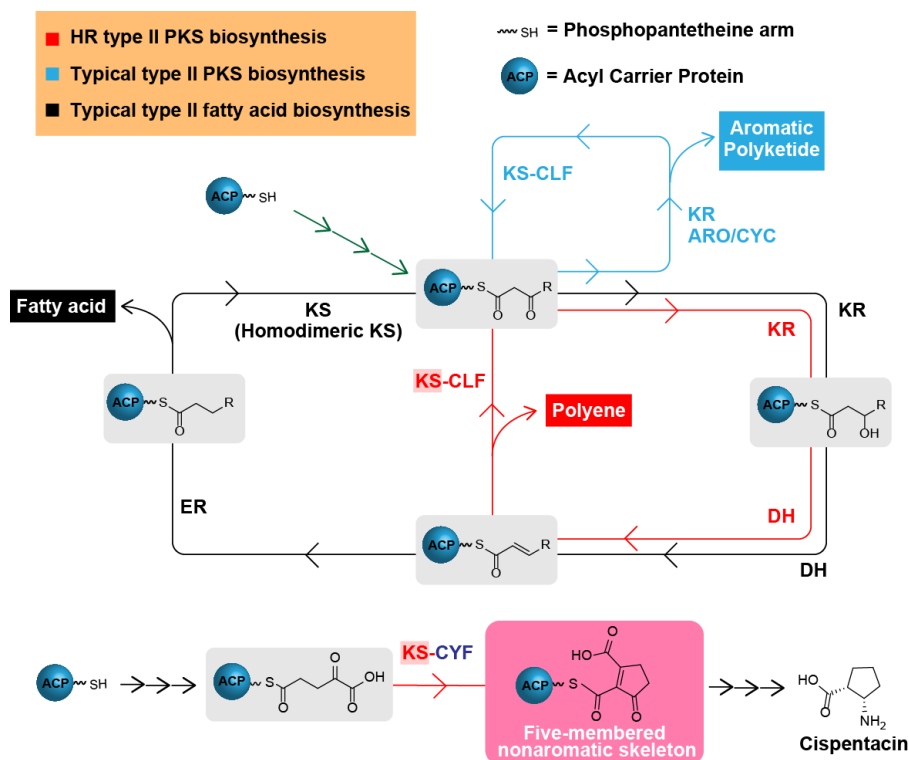
Supplementary Table 4. Vectors used in this study.

Vectors/Plasmids	Characteristic(s)	Source/Reference
pSE101	Amp ^r and Tsr ^r , gene expression in <i>S. albus</i> G153	(1)
pSEamcB-H	Using for heterologous production of cisptentacin	This study
pSEamcF-G	pSE101 derived plasmid for expressing N-terminal His-tagged AmcF and AmcG	This study
pSEamcH	pSE101 derived plasmid for expressing N-terminal His-tagged AmcH	This study
pKU503_ampr	Using as gene template	(2)
pET26(+)	Kan ^r , protein expression vector	Novagen
pET26(+)-amcB	pET26(+) derived plasmid for expressing C-terminal His-tagged AmcB	This study
pET26(+)-sFabI	pET26(+) derived plasmid for expressing C-terminal His-tagged sFabI	This study
pET26(+)-eFabI	pET26(+) derived plasmid for expressing C-terminal His-tagged eFabI	This study
pHis8	Kan ^r , protein expression vector in <i>E. coli</i> BL21(DE3)	(3)
pHis8-amcC	pHis8 derived plasmid for expressing N-terminal His-tagged AmcC	This study
pHis8-amcD	pHis8 derived plasmid for expressing N-terminal His-tagged AmcD	This study
pHis8-amcE	pHis8 derived plasmid for expressing N-terminal His-tagged AmcE	This study

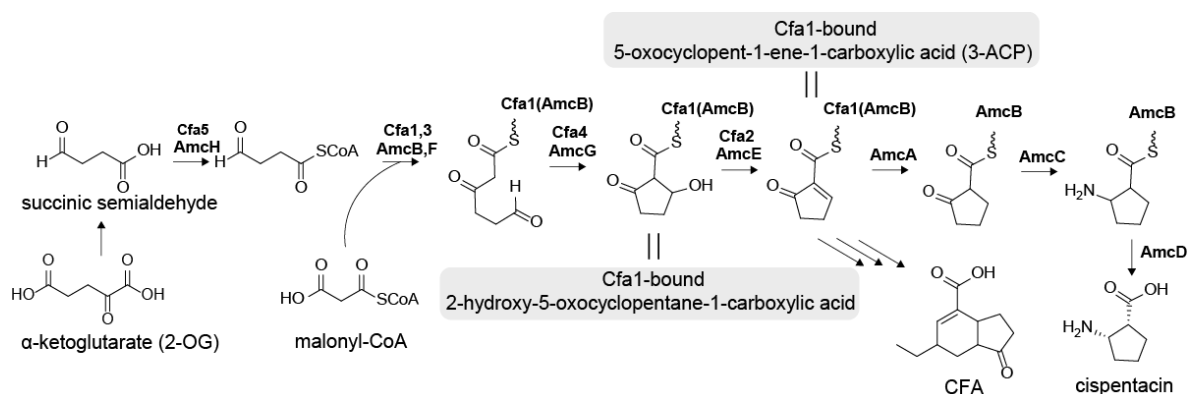
Supplementary Table 5. The characterization of chemical and biomolecular materials.

Compound	Source	Purity	Identity
(1R,2S)-(-)-2-Amino-1-cyclopentanecarboxylic acid hydrochloride hemihydrate (Cispentacin)	Sigma-Aldrich	98.0-102.0% (calc. on dry substance, AT)	712205
Detection and quantification of cispentacin			
1-fluoro-2,4-dinitrobenzene (DNFB)	NACALAI TESQUE, INC.	≥ 99.0%(GC)	70-34-8
Synthesis of 3-oxocyclopent-1-ene-1,2-dicarboxylic acid			
tert-butyl 2-(triphenylphosphoranylidene)acetate	Sigma-Aldrich	98%	369799
3-chloropropionyl chloride	Tokyo Chemical Industry Co., Ltd. (TCI)	>98.0%(T)	625-36-5
OXONE®, monopersulfate compound	Sigma-Aldrich	n/a	228036
2,3-di(carbo-tert-butoxy)-2-cyclopentenone	In this study	>90.0% (by ¹ H-NMR)	
3-oxocyclopent-1-ene-1,2-dicarboxylic acid	In this study	>90.0% (by ¹ H-NMR)	
Marfey's Method			
N α -(5-Fluoro-2,4-dinitrophenyl)-L-leucinamide (L-FDLA)	Tokyo Chemical Industry Co., Ltd. (TCI)	>98.0%(HPLC)	178065-29-7
N α -(5-Fluoro-2,4-dinitrophenyl)-D-leucinamide (D-FDLA)	Tokyo Chemical Industry Co., Ltd. (TCI)	>98.0%(HPLC)	178065-30-0

Supplementary Figures

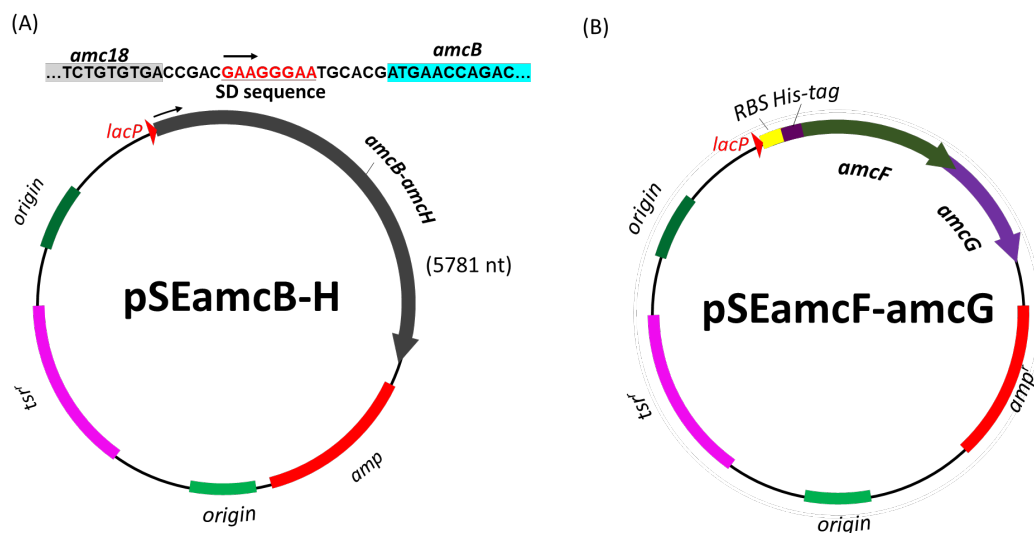


Supplementary Figure 1. Comparison of type II PKS subfamilies. Biosynthetic pathways involving highly reducing (HR) type II PKS (red arrows), typical type II PKS (blue arrows), type II fatty acid (black arrows), and the phylogenetically new subfamily of type II PKS (bottom) identified in this study. Abbreviations: ARO/CYC, aromatase/cyclase; CLF, chain length factor; DH, dehydratase; ER, enoyl-acyl carrier protein reductase; KR, ketoreductase; KS, ketosynthase; CYF, cyclization factor.

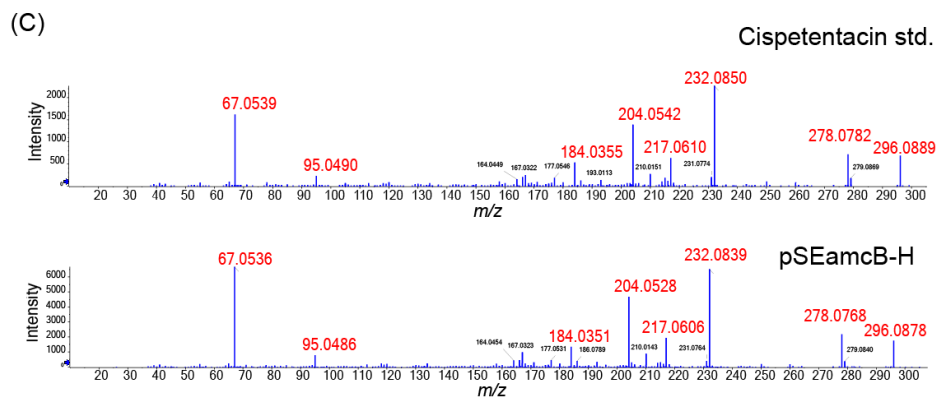
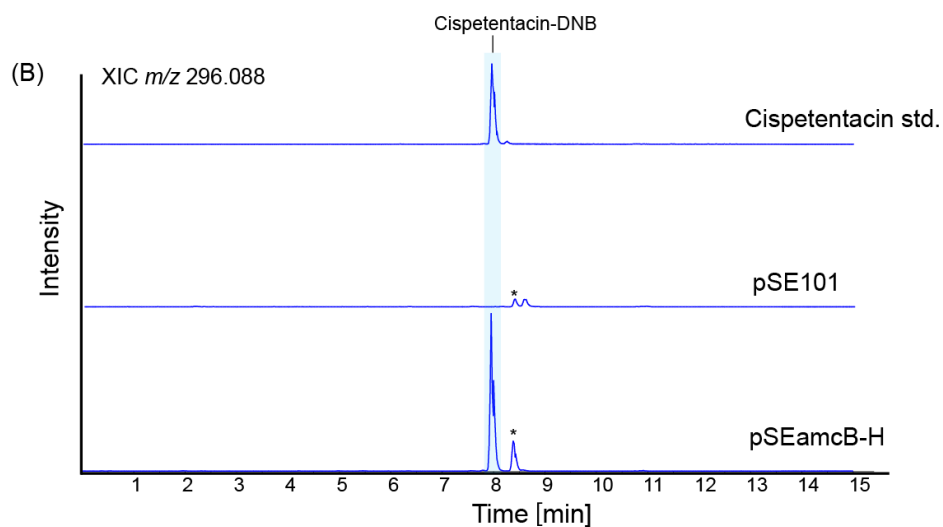
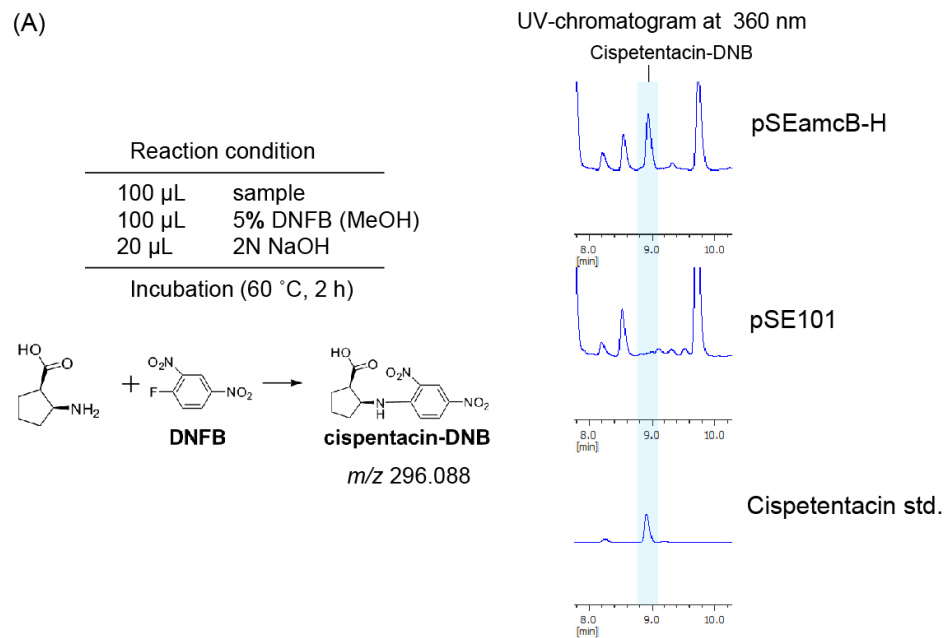


Supplementary Figure 2. Biosynthetic pathway of CPC-ACP postulated in previous studies.

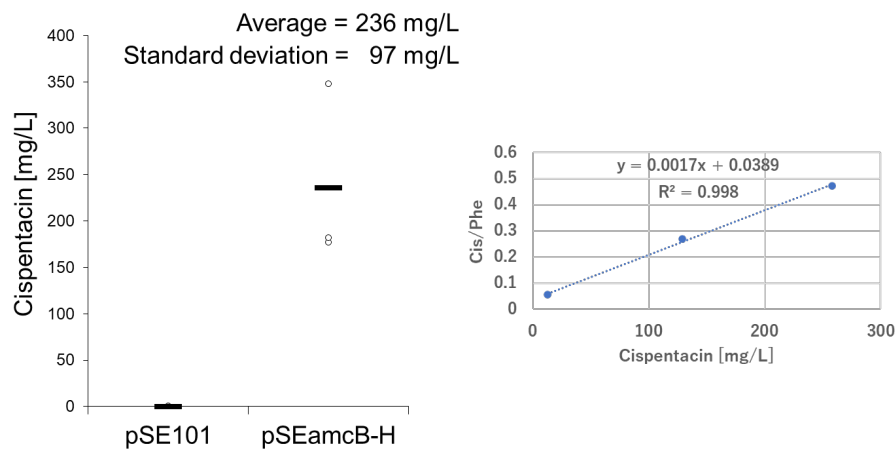
The postulated pathway begins with the conversion of succinic semialdehyde derived from 2-oxoglutarate into its CoA ester by Cfa5. Cfa1 (ACP), Cfa3 (KS), and Cfa4 are postulated to produce Cfa1-bound 2-hydroxy-5-oxocyclopentane-1-carboxylic acid, which is subsequently dehydrated by Cfa2 (DH) to give Cfa1-bound 5-oxocyclopent-1-ene-1-carboxylic acid. AmcB, E, F, G, and H show high similarity to those of Cfa1, 2, 3, 4, and 5, respectively, which suggests that the biosynthesis of CFA and cispentacin share a common route leading to 5-oxocyclopent-1-ene-1-carboxylic acid bound to either Cfa1 or AmcB.



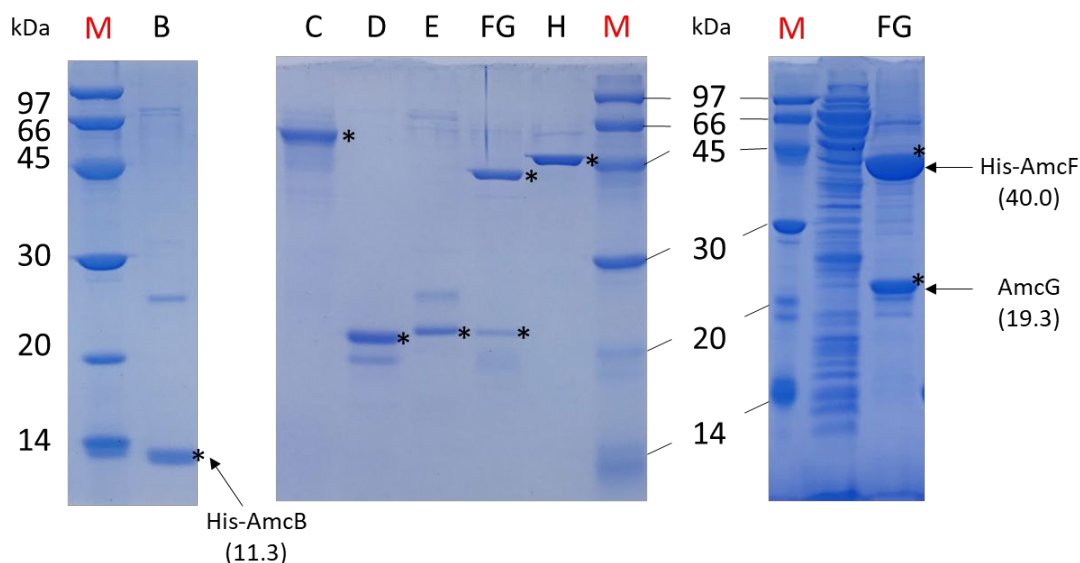
Supplementary Figure 3. Schematic representation of the plasmid constructs. (A) pSEamcB-H expresses the cispentacin biosynthetic operon (*amcB* to *amcH*) containing the SD sequence for translation initiation upstream of *amcB*. The operon is inserted downstream of the *lacP* promoter (*lacP*). We have reported previously that genes inserted downstream of *lacP* demonstrate robust expression in *S. albus* (4). (B) pSEamcF-amcG coexpresses the operon of *amcF* and *amcG* (KS-CYF) in *S. albus* G153. The operon is inserted downstream of the *lacP* promoter (*lacP*).



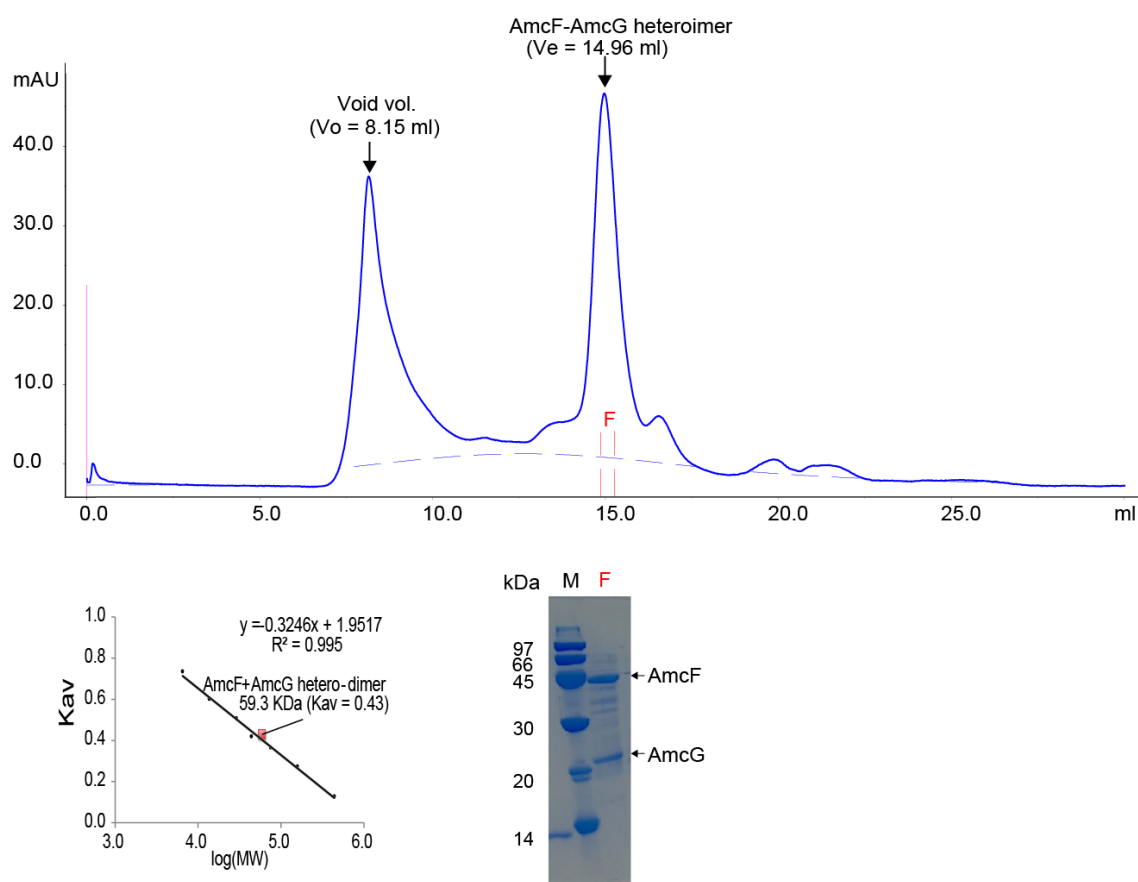
Supplementary Figure 4. X-LC and MS analysis of DNFB-derivatized cispentacin. (A) Reaction conditions for DNFB derivatization and chromatograms at 360 nm of the DNFB-derivatized culture broths from *S. albus* G153/pSE101 and *S. albus* G153/pSEamcB-H. (B) Extracted ion count (XIC) chromatograms at m/z 296.088 corresponding to the DNFB-derivatized cispentacin in the broths from *S. albus* G153/pSE101 and *S. albus* G153/pSEamcB-H. (C) MS/MS spectrum of the DNFB-derivatized cispentacin. The spectrum from the culture broth of *S. albus* G153/pSEamcB-H matches well with that of the authentic DNFB-derivatized cispentacin.



Supplementary Figure 5. Production of cispentacin by *S. albus* G153/pSEamcB-H. *S. albus* G153/pSE101 was used as a negative control transformant. Dots correspond to cispentacin production in individual cultures; the bars correspond to average values. Average and standard deviation values were calculated with *S. albus* G153 samples ($n = 3$, independent culture samples). Source data are provided as a Source Data file.



Supplementary Figure 6. SDS-PAGE analysis of purified recombinant proteins. SDS-PAGE (12%) analysis of recombinant proteins. Lane B, AmcB, calculated MW 11.3 kDa; Lane C, AmcC, calculated MW 51.1 kDa; Lane D, AmcD, calculated MW 16.7 kDa; Lane E, AmcE, calculated MW 18.8 kDa; Lane FG, AmcF–AmcG, calculated MWs 40.0 kDa for AmcF and MWs 19.3 kDa for AmcG; Lane H, AmcH, calculated MW 49.6 kDa. All these proteins except AmcG were produced as His-tagged proteins (marked with an asterisk) and purified by affinity chromatography using a Ni-NTA Superflow resin. Lane M is a molecular weight marker. In Lane B, a small amount of phosphopantetheinyl transferase (25 kDa), which is coexpressed in *E. coli* cells for phosphopantetheinyl transfer of AmcB, was also detected. In Lanes C to H, each target protein marked with an asterisk is major, although unknown trace proteins were detected. Source data are provided as a Source Data file.

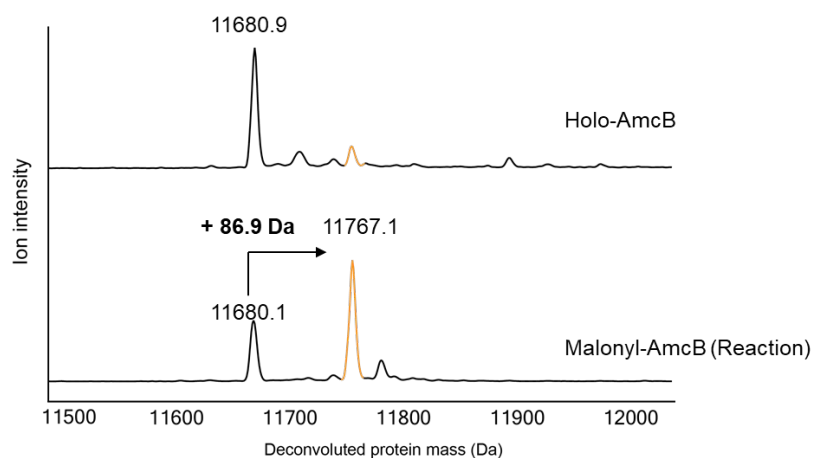
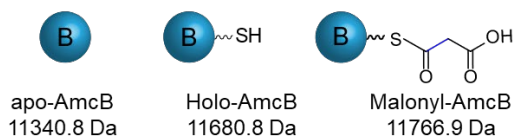


Supplementary Figure 7. Gel filtration of AmcF-AmcG. Gel filtration on a Superdex 200 Increase 10/300 GL column. AmcF-AmcG (1.5 mg) was subjected to the column. A single peak corresponding to the mass of the AmcF-AmcG heterodimer was detected. Lane F, SDS-PAGE (12%) analysis of the purified AmcF-AmcG heterodimer fraction. * $K_{av} = (V_e - V_o)/(V_c - V_o)$, V_c (geometric column volume) = 24 (mL), V_o (column void volume). Source data are provided as a Source Data file.

Reaction condition (30°C, 1h)

50 mM	Tris HCl (pH8.0)
1 mM	MgCl ₂
0.5 mM	malonyl-CoA
200 μM	holo-AmcB

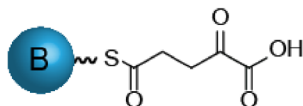
total 50 μL



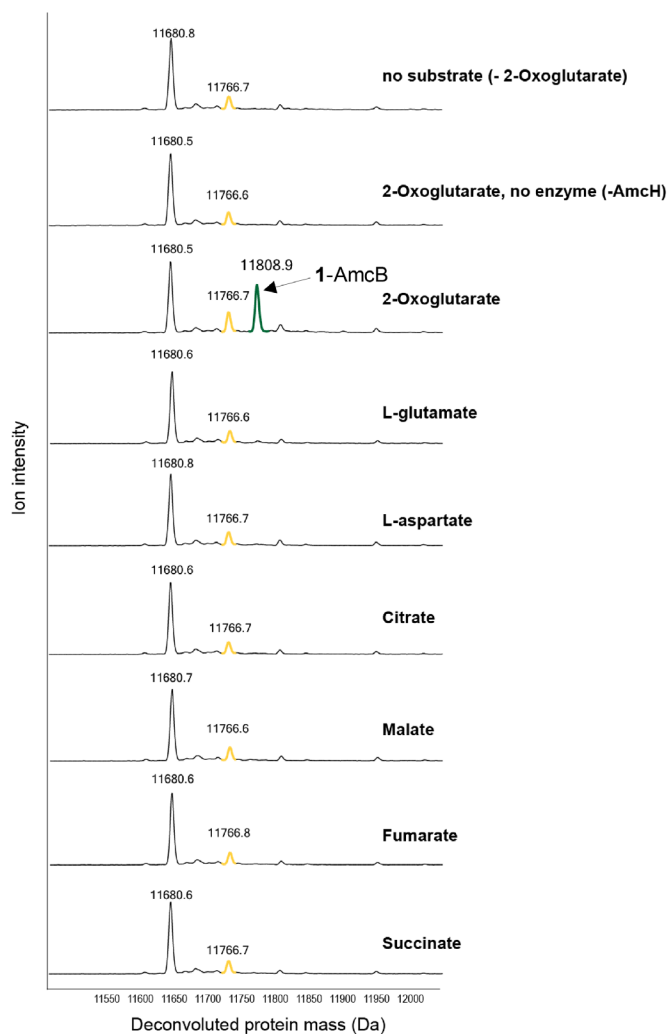
Supplementary Figure 8. Detection of AmcB bound intermediate derived from malonyl-CoA. Deconvoluted protein mass for *holo*-AmcB was detected at 11680.9 Da. The deconvoluted protein mass for malonyl-AmcB (orange line) was detected at 11767.1 Da under the reaction conditions shown. A small amount of malonyl-AmcB that formed during AmcB production in *E. coli* was detected in the purified sample of *holo*-AmcB.

Reaction condition (30°C, 1h)

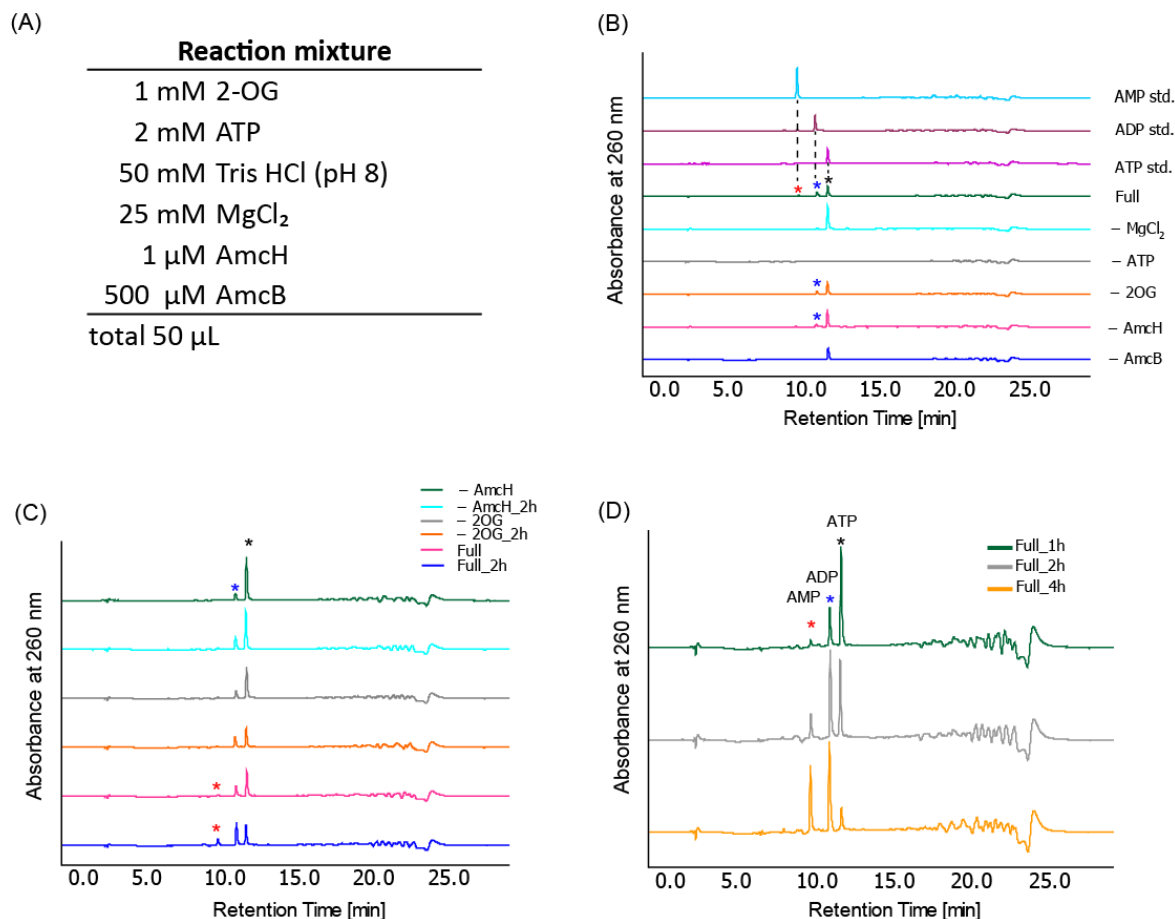
1 mM Substrate
 5 mM ATP
 50 mM Tris HCl (pH8)
 25 mM MgCl₂
 200 μM AmcB
 2 μM AmcH
 total 50 μL



1-AmcB
 11808.8 Da

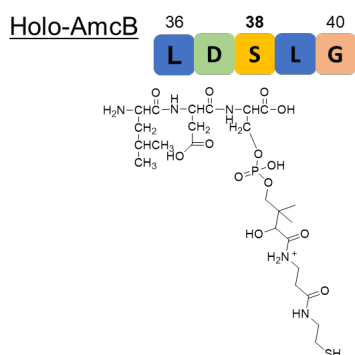


Supplementary Figure 9. Substrate specificity of AmcH. Spectra for deconvoluted protein mass of 1-AmcB, which was detected at 11808.9 Da. Substrate-tethered AmcB was formed only when 2-oxoglutarate was used as a substrate for AmcH.

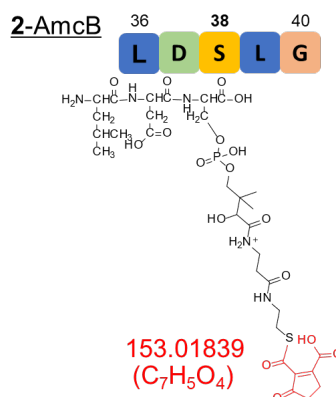


Supplementary Figure 10. HPLC analysis of AmcH activity. The reaction conditions are described in the Experimental Section. (A) Composition of the reaction mixture. (B) Each AmcH reaction was conducted for 1 h. Conversion of ATP to ADP and AMP was detected in the sample of Full. (C) Conversion of ATP to ADP was detected in the sample including AmcB. (D) The AmcH reactions were conducted for 1, 2, and 4 h. Almost all of the ATP was converted to AMP or ADP during the 4 h reaction. ADP probably formed due to spontaneous hydrolysis or hydrolysis by residual phosphatase from *E. coli*.

(A)

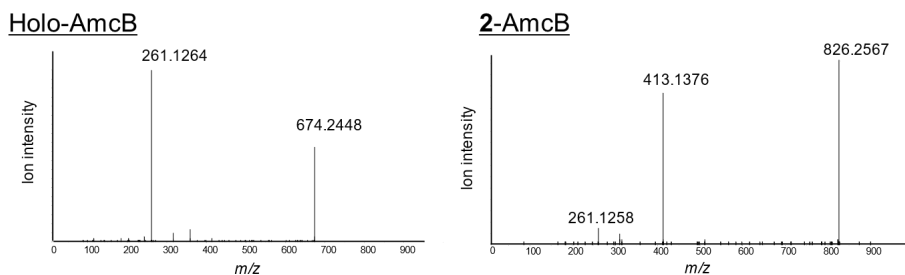


Chemical Formula: $C_{24}H_{45}N_5O_{13}PS^+$
Exact Mass: 674.24667



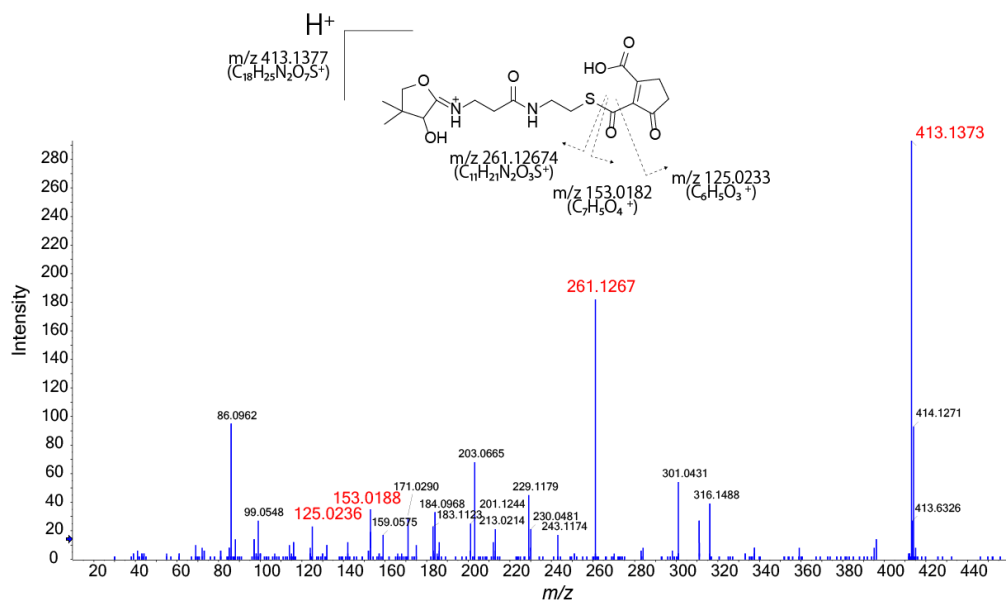
Chemical Formula: $C_{31}H_{49}N_5O_{17}PS^+$
Exact Mass: 826.25763

(B)

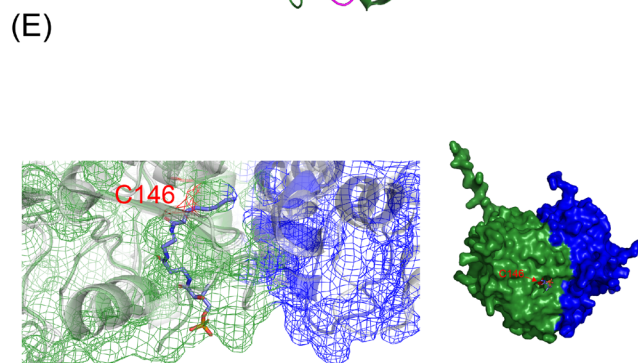
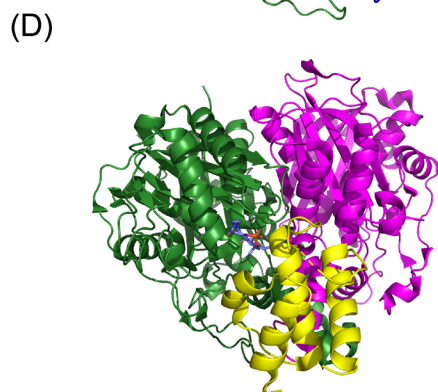
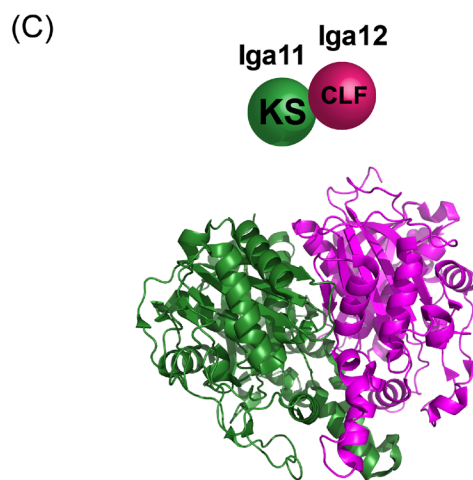
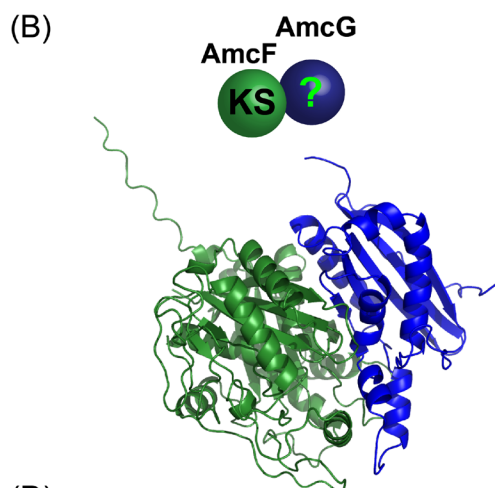
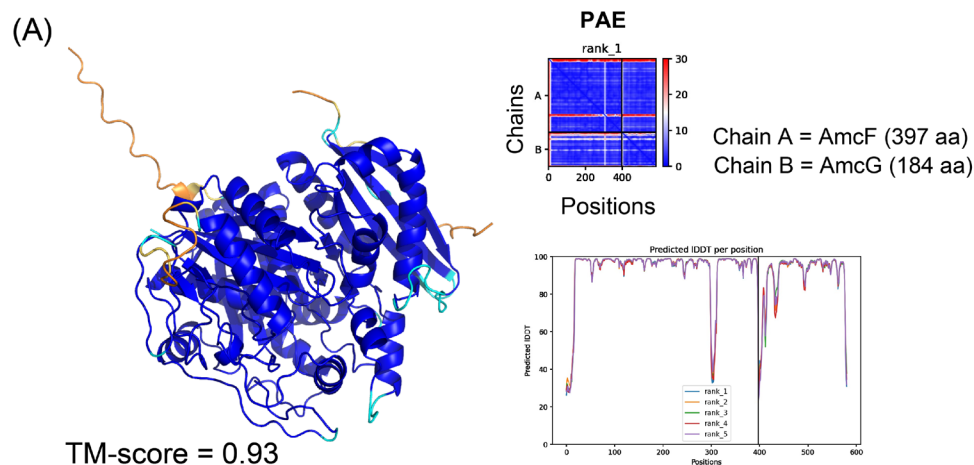


Supplementary Figure 11. LC-ESI-HRMS analysis of 2-AmcB digested with thermolysin.

(A) Schematic representation of thermolysin-digested peptide fragments of *holo*-AmcB and 2-AmcB. (B) Extracted ion chromatograms are shown for the digested peptide fragments of *holo*-AmcB (m/z 674.2448 $[M+H]^+$) and 2-AmcB (m/z 826.2567 $[M+H]^+$).



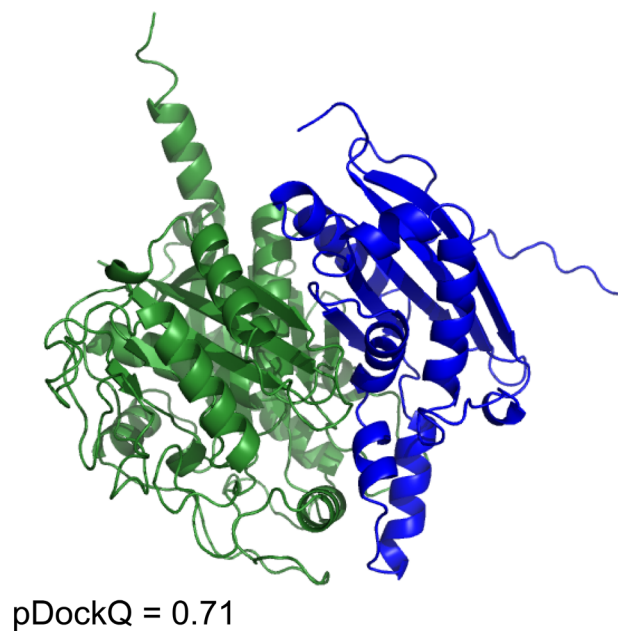
Supplementary Figure 12. MS/MS spectrum of the PPant-eliminated ion of 2-AmcB.



Executive: RMSD = 1.597 (337 to 337 atoms)

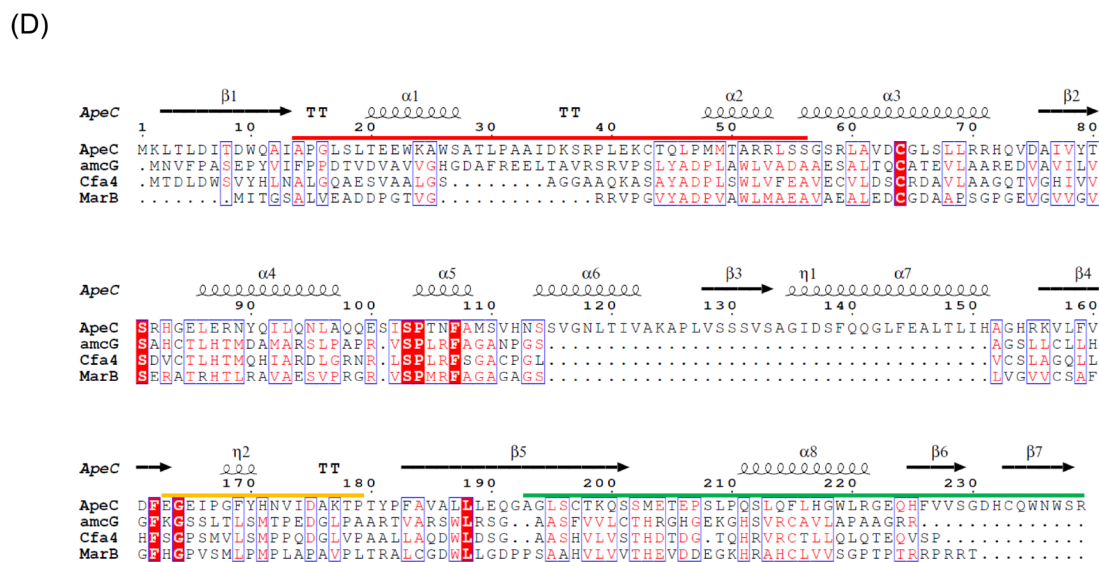
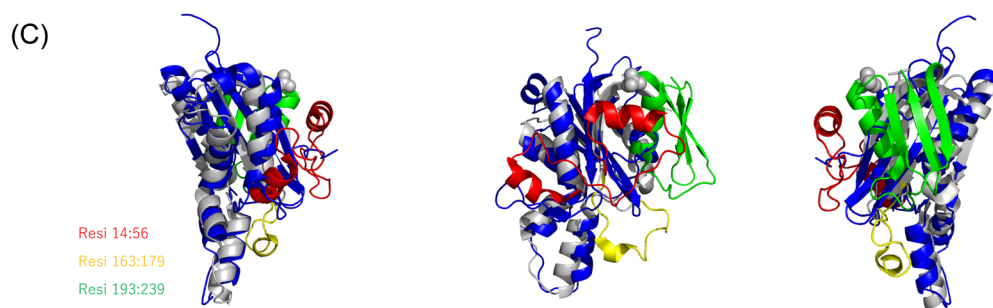
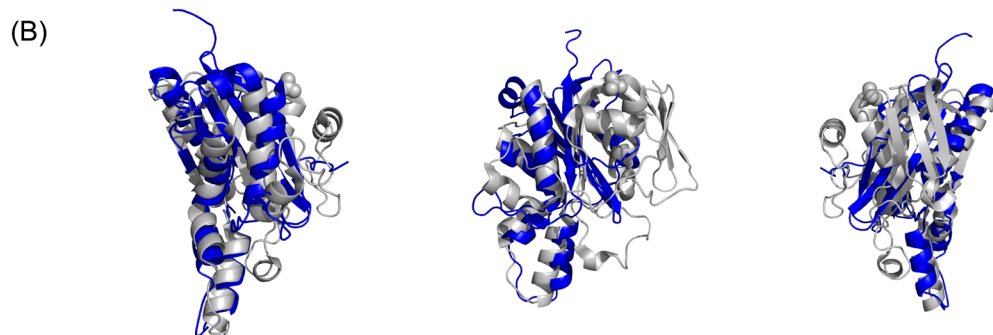
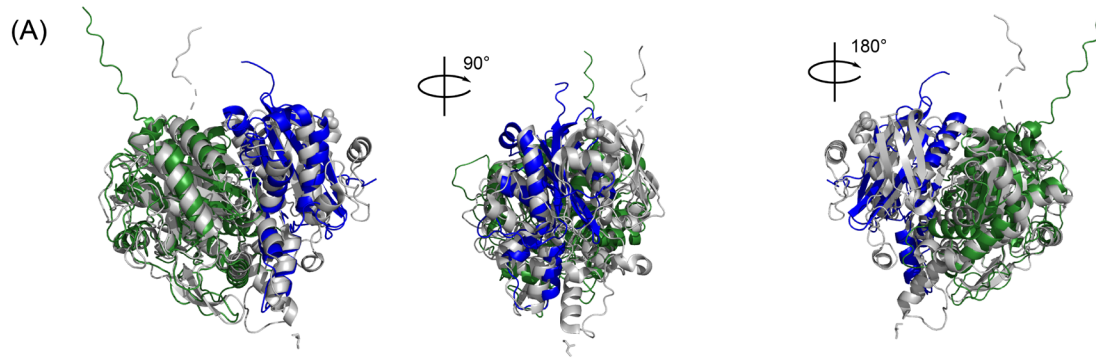
Supplementary Figure 15. AmcF–AmcG heterodimer fold prediction. (A) The predicted structure of the heterodimer AmcF–AmcG obtained by ColabFold. b=pLDDT (blue, $b < 100$), (cyan, $b < 90$), (yellow–orange, $b < 70$), (color tv_orange, $b < 50$). Predicted aligned error (PAE) and pLDDT (Rank_1 corresponds to the structure of the description) are shown as confidence measures. (B) Color-coded predicted structure of AmcF–AmcG (green, AmcF), (blue, AmcG). (C) The structure of the KS–CLF heterodimer (6KXD) (green, Iga11 (KS)), (pink, Iga12 (CLF)). (D) The structure of ACP=KS–CLF (6KXF) used to predict the ligand binding site of AmcF–AmcG

(yellow, Iga10 (ACP)). (E) Ligand binding cavity and predicted catalytic residue (green, AmcF), (blue, AmcG), (white, Iga11–12), (yellow, the substrate of Iga11–12). The RMSD value between AmcF–AmcG and Iga11–Iga12 is also shown.



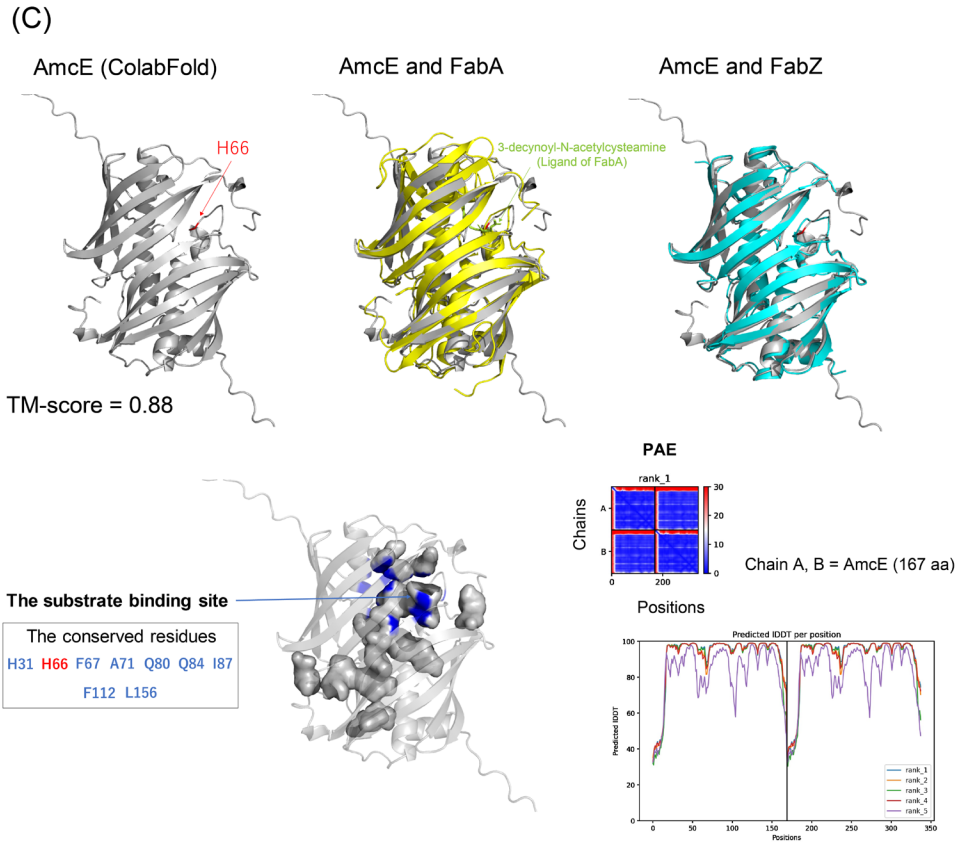
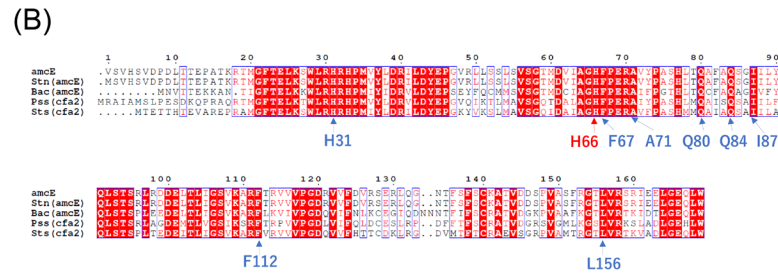
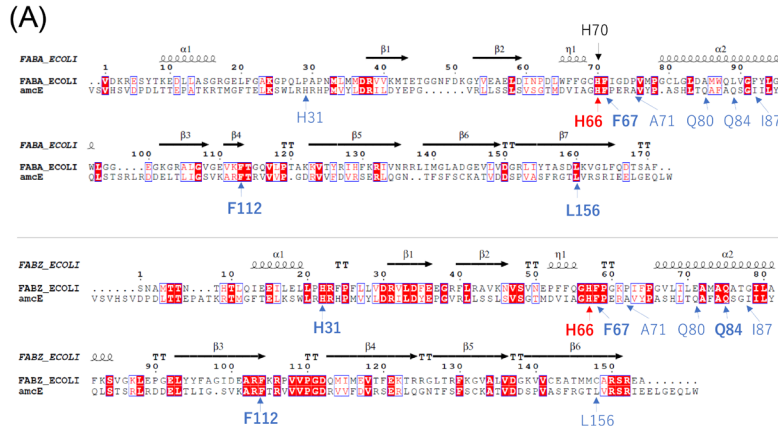
num_contacts	avg_if_plddt	pdockq
179	93.11050324	0.706735165

Supplementary Figure 16. Assessment of the dimer structure. The predicted structure of the heterodimer AmcF–AmcG and the score of the predicted DockQ score (pDockQ) were obtained by SpeedPPI.

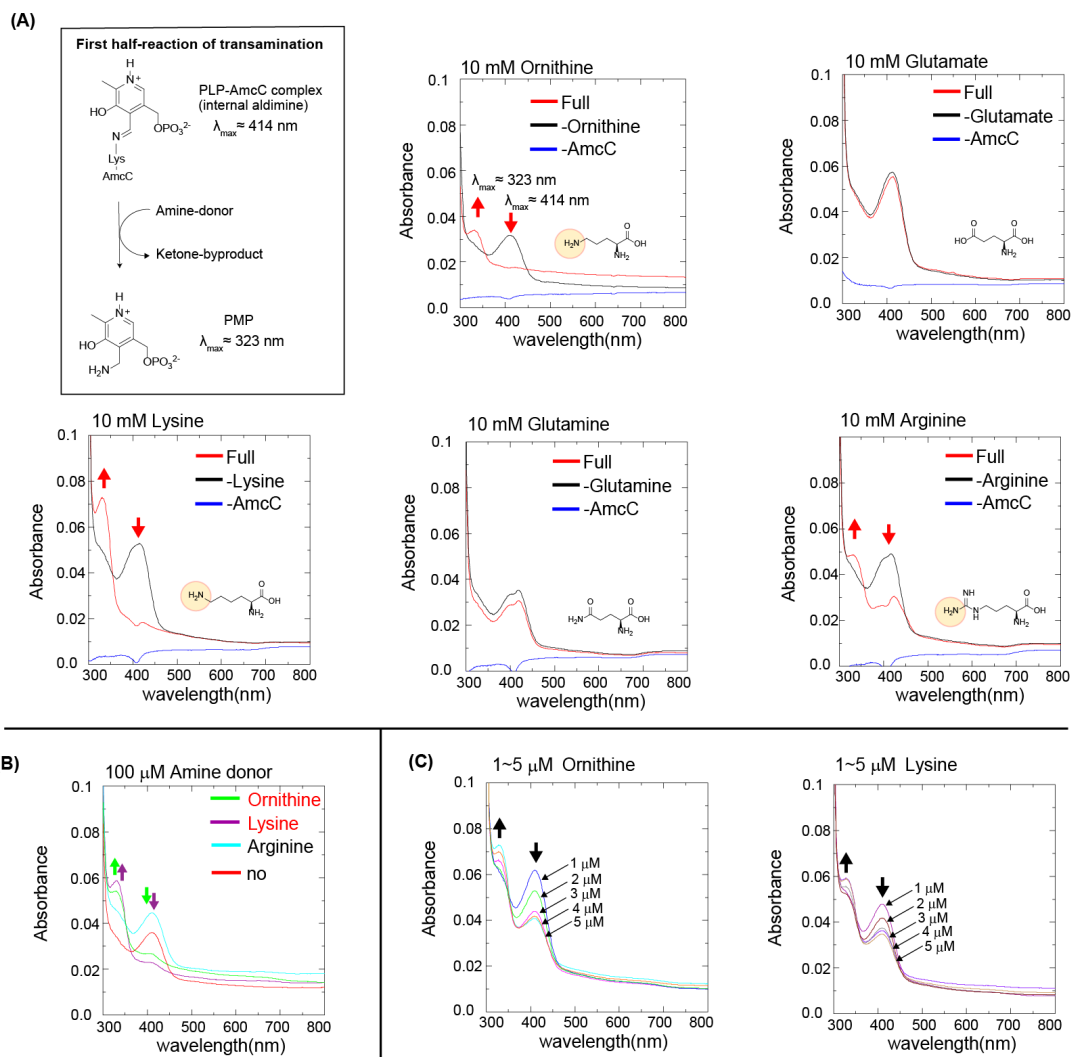


Supplementary Figure 17. Comparative analysis of CLF and cyclization factors (CYFs).

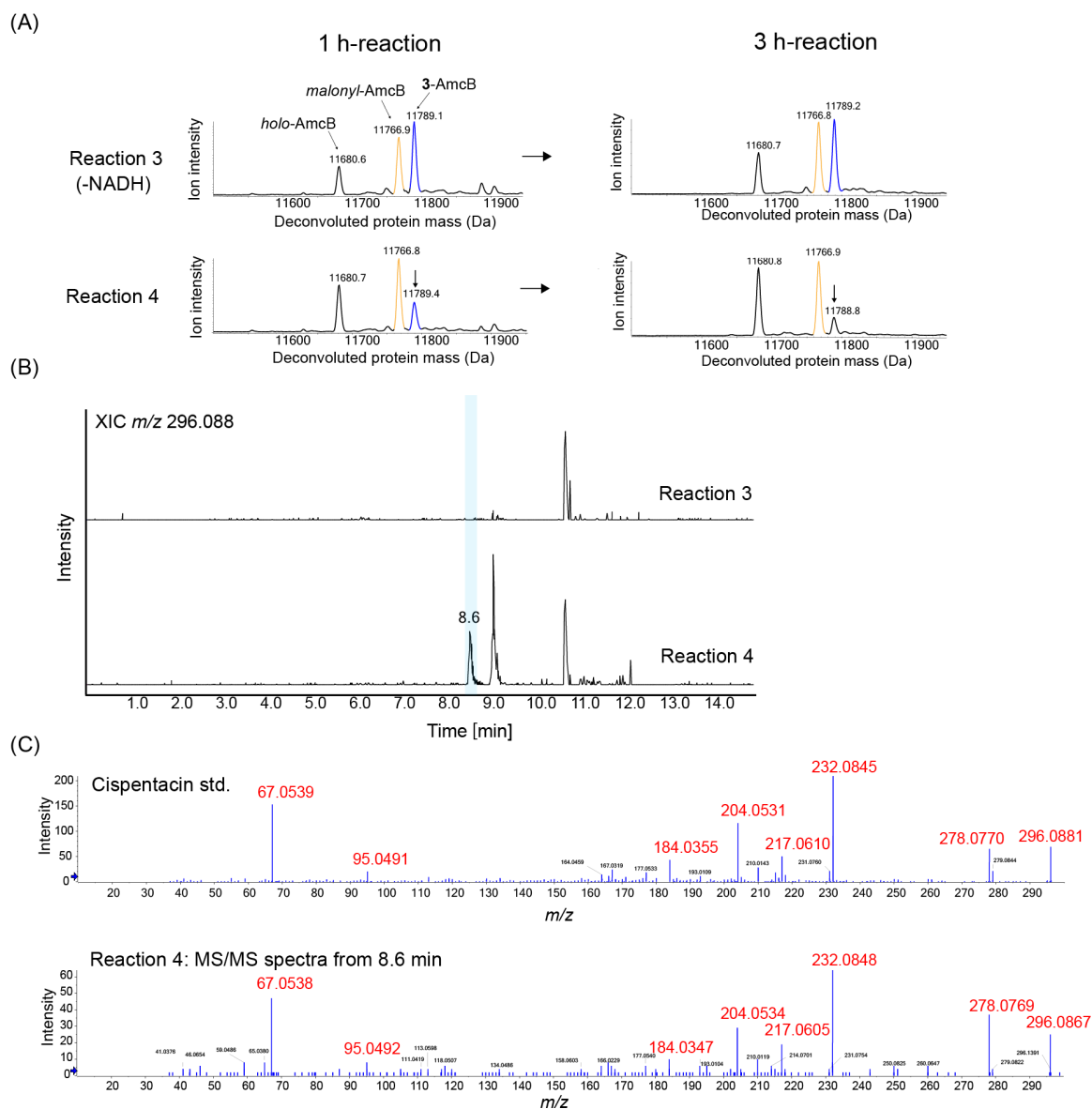
(A) Predicted structure of the heterodimer AmcF–AmcG (green, AmcF) and (blue, AmcG) compared with the structure of the KS–CLF heterodimer (6qsp) (gray, ApeO (KS) and ApeC (CLF)) from three angles. ApeC is the smallest reported structure in CLF, and the structure of AmcG (CYF) further lacks several secondary structures existing in ApeC (CLF). The RMSD value between AmcF–AmcG and ApeO–ApeC is also shown. (B) AmcG structures extracted from the heterodimeric structure compared with the ApeC structure from three angles. (C) Secondary structures of ApeC not present in AmcG (ApeC residues from 14 to 56, red) (ApeC residues from 163 to 179, orange) (ApeC residue from 193 to 239, green). (D) Alignment of AmcG homologs (Cfa4 and MarB) with HR type II PKS ApeC (CLF). The secondary structure is derived from the ApeC structure (6QSP_chainB) and indicated on the corresponding ApeC sequence. The three colored lines (red, orange, and green) correspond to the secondary structures of ApeC in (C). AmcG homologs show almost no sequence similarity to ApeC. This alignment shows no correlation between similar regions in structure and similar regions in alignment.



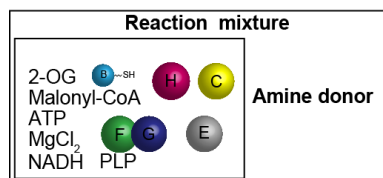
Supplementary Figure 18. Sequence analysis of AmcE. (A) Sequence alignments of AmcE with FabA (P0A6Q3) and AmcE with FabZ (P0A6Q6). The predicted catalytic residue histidine 66 (H66) is conserved. (B) Alignment of AmcE homologs in the cispentacin and CFA gene cluster. Stn, AmcE homolog of *Streptomyces novoguineensis* QHW08541.1; Bac, AmcE homolog of *Bacillus cereus* WP_098276130.1; Pss, Cfa2 homolog of *Pseudomonas syringae* WP_248842617.1; Sts, Cfa2 homolog of *Streptomyces scabies* WP_013005371.1. The residues of the active site (blue and red) are highly conserved in the AmcE homologs in contrast to FabZ. (C) The predicted structure of the homodimer of AmcE obtained by ColabFold (gray). The overall structure is similar to the crystal structure of FabA (1mka) and FabZ (6n3p) (yellow, FabA), (cyan, FabZ). The substrate binding sites and colored conserved residues are shown in the predicted structure of AmcE. Predicted aligned error (PAE) and pLDDT (Rank_1 corresponds to the structure of the description) are shown as confidence measures.



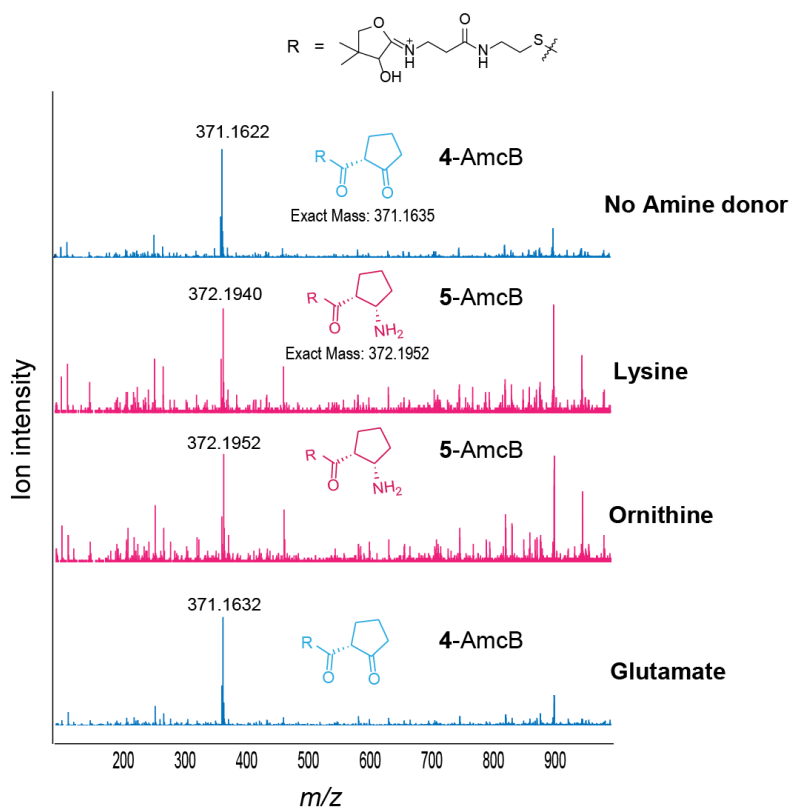
Supplementary Figure 19. UV–vis spectra assay for AmcC. (A) Analysis of the AmcC-catalyzed reaction with UV–vis spectra assay using 10 mM amine donor. The first half-reaction of transamination can be detected by monitoring the change in absorbance from 414 nm (PLP–AmcC complex) to 323 nm (PMP). A shift to 323 nm was detected when ornithine, lysine or arginine was used. (B) Analysis of the AmcC-catalyzed reaction with UV–vis spectra assay using 100 μM amine donor. A shift to 323 nm was detected when ornithine or lysine was used. (C) Analysis of the AmcC-catalyzed reaction with UV–vis spectra assay using 1–5 μM ornithine or lysine. A shift to 323 nm was detected in both cases.



Supplementary Figure 20. LC–MS analysis of cispentacin in the reconstituted reaction mixture. (A) Deconvoluted protein mass spectra from Reactions 3 and 4. The deconvoluted protein mass signal for **3**-AmcB (11789 Da, blue) was significantly decreased in Reaction 4. (B) LC–MS analysis of DNFB-derivatized reaction mixtures. XIC chromatograms at m/z 296.088 in the derivatized Reactions 3 and 4. (C) MS/MS spectra of the DNFB-derivatized cispentacin. The spectrum from reaction mixture 4 well matches that of the DNFB-derivatized cispentacin.

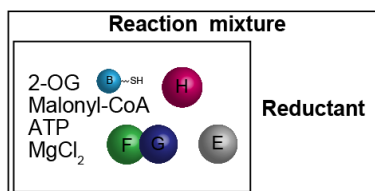


Reaction mixture
1 mM 2-Oxoglutarate
5 mM ATP
1 mM Malonyl-CoA
50 mM Tris HCl (pH8)
25 mM MgCl ₂
200 μM AmcB
0.02 μM AmcH
8 μM AmcF-G
8 μM AmcE
8 μM AmcC
10 mM Amine donor
1 mM PLP
1 mM NADH
total 50 μL



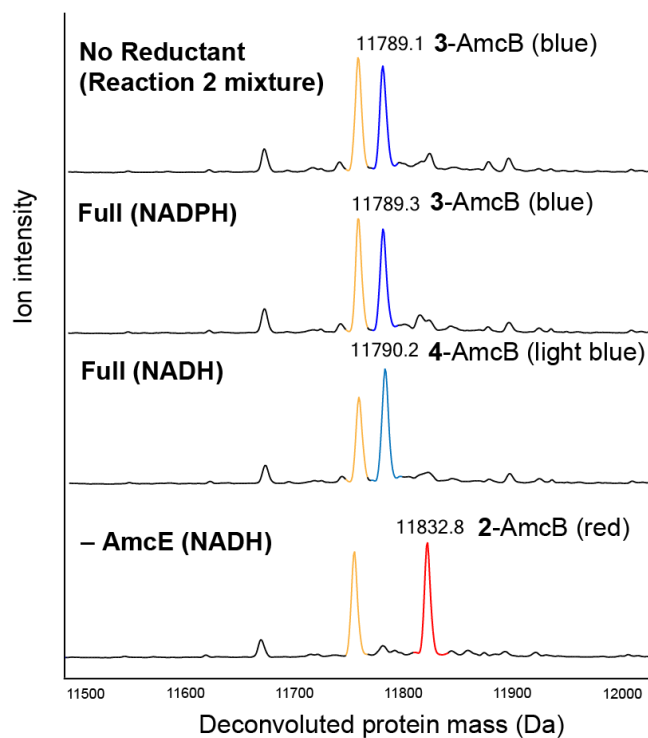
Supplementary Figure 21. Amine donor specificity of AmcC. PPant ejection assay was used for the amine donor specificity of AmcC. MS/MS fragments of the 13+ charge state of 4-AmcB and 5-AmcB were analyzed. The PPant-eliminated ion of 5-AmcB (m/z 372.1952) was detected when lysine or ornithine was added as an amine donor.

(A)

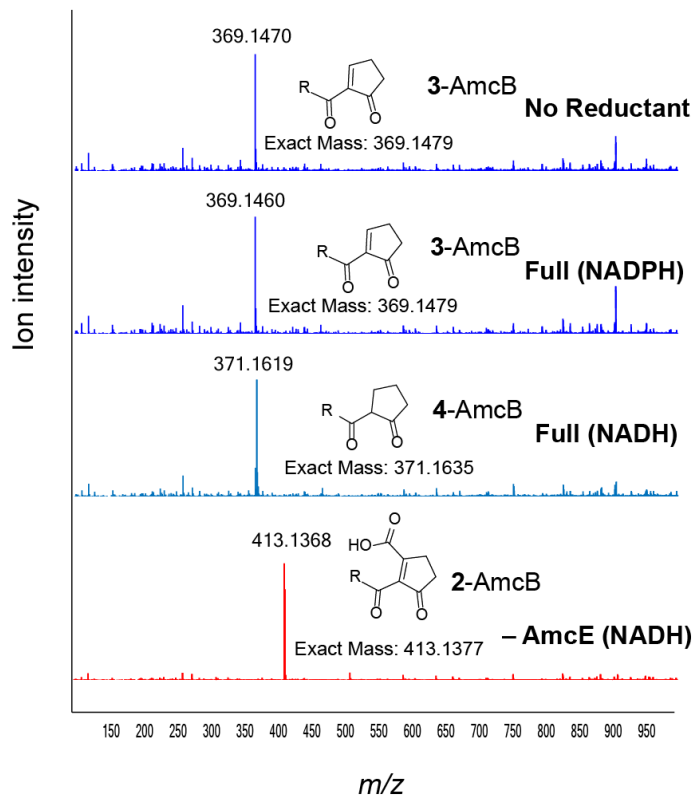
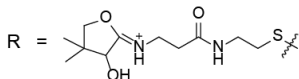


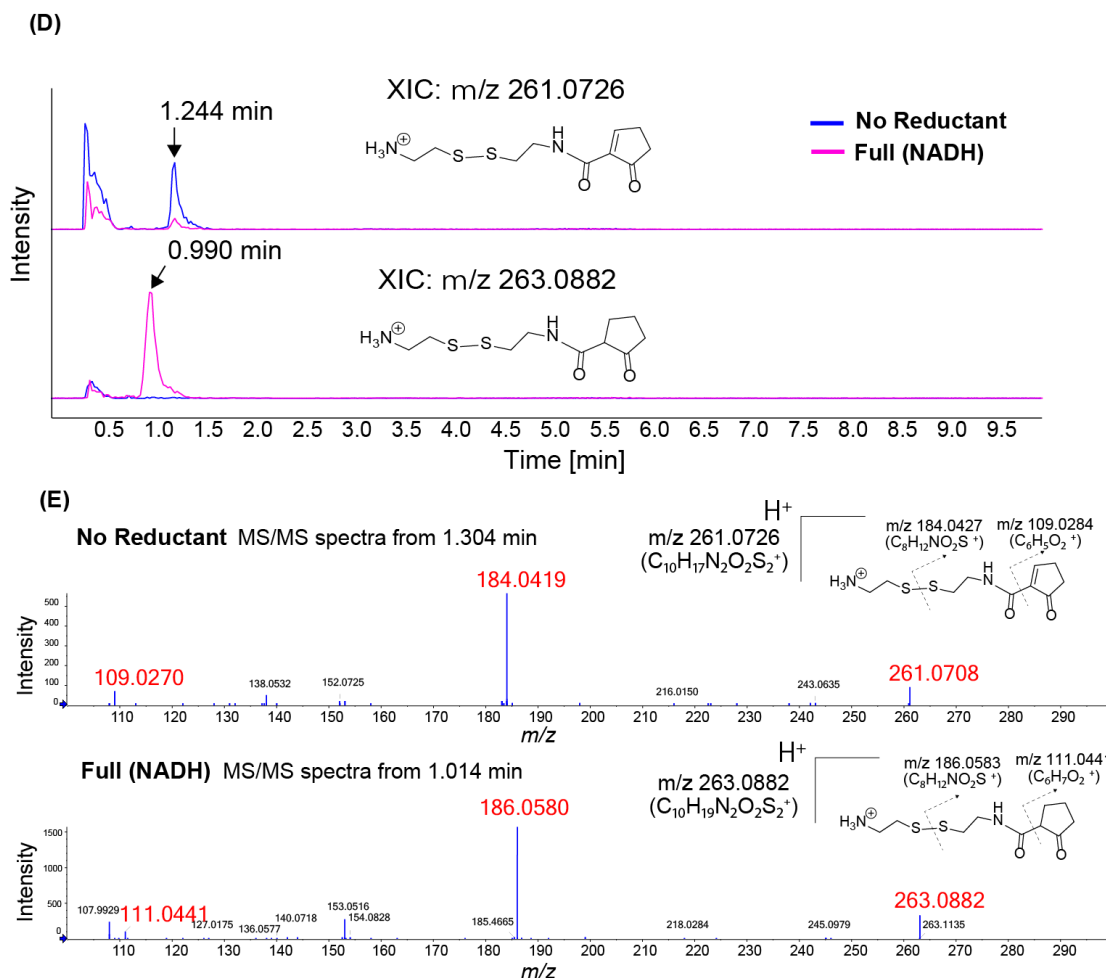
Reaction mixture
1 mM 2-Oxoglutarate
5 mM ATP
1 mM Malonyl-CoA
50 mM Tris HCl (pH8)
25 mM MgCl ₂
200 μM AmcB
0.02 μM AmcH
8 μM AmcF-G
8 μM AmcE
1 mM Reductant
total 50 μL

(B)



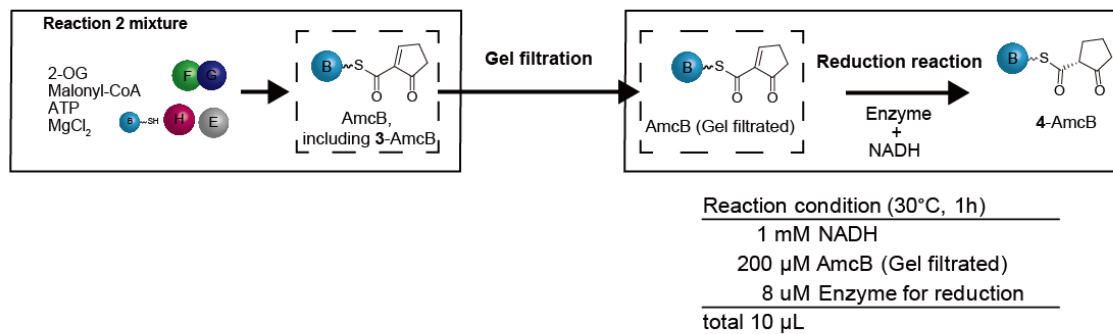
(C)



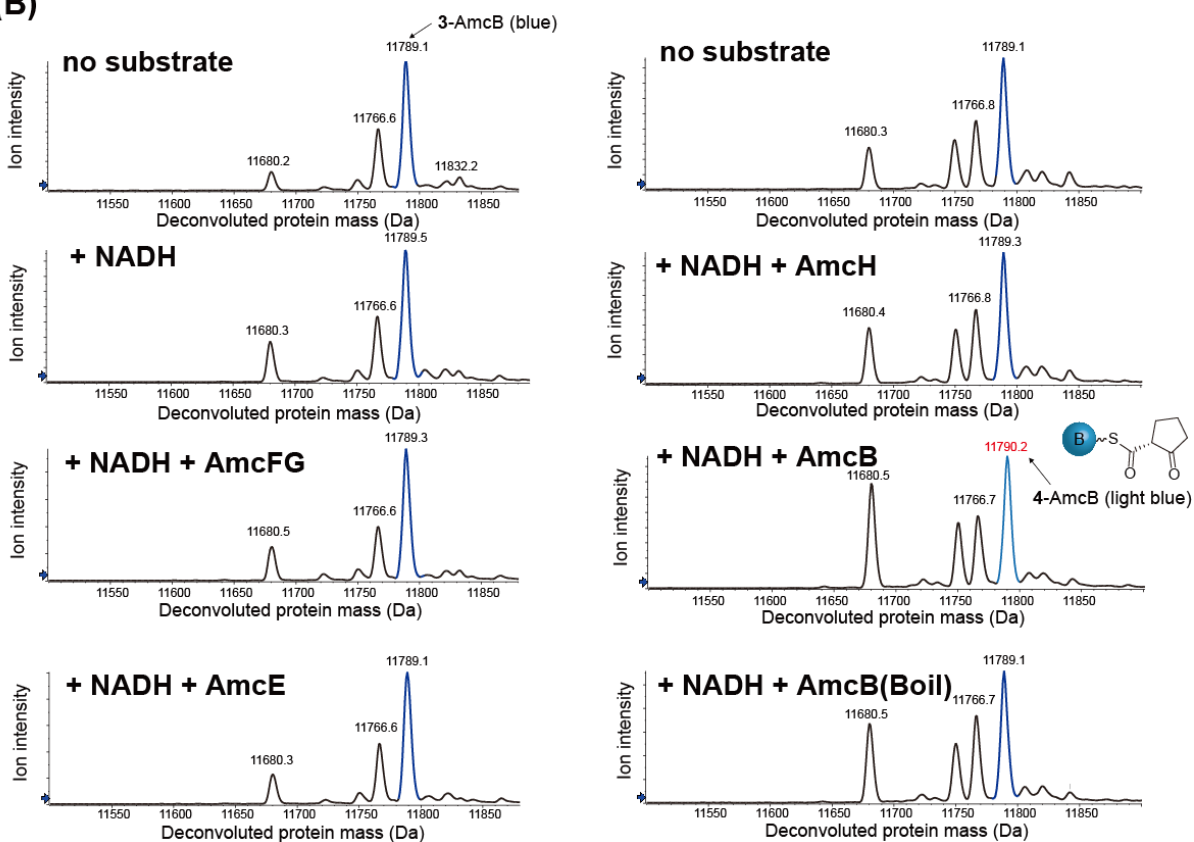


Supplementary Figure 22. NADH-dependent reduction of 3-AmcB. (A) The reaction conditions are based on Reaction 2. (B) Deconvoluted protein mass spectra for the reaction product under each condition. MS/MS fragments of the 13+ charge state of intermediate-AmcB were analyzed. The MS/MS fragment of 4-AmcB was detected only in the reaction mixture (Full (NADH)). (C) The PPANT ejection assay for the reaction product under each condition. The PPant-eliminated ion of 4-AmcB was detected only in the reaction mixture (Full (NADH)). (D) Cysteamine-promoted cleavage assay. Extracted ion count (XIC) chromatograms at m/z 261.0726 corresponding to the cysteamine adduct of **3** and m/z 263.0882 corresponding to the cysteamine adduct of **4** in the reaction mixture (No reductant) and (Full (NADH)), respectively. The formation of 4-AmcB was confirmed in the reaction mixture (Full (NADH)). (E) MS spectra for the cysteamine adduct of **3** and the cysteamine adduct of **4**.

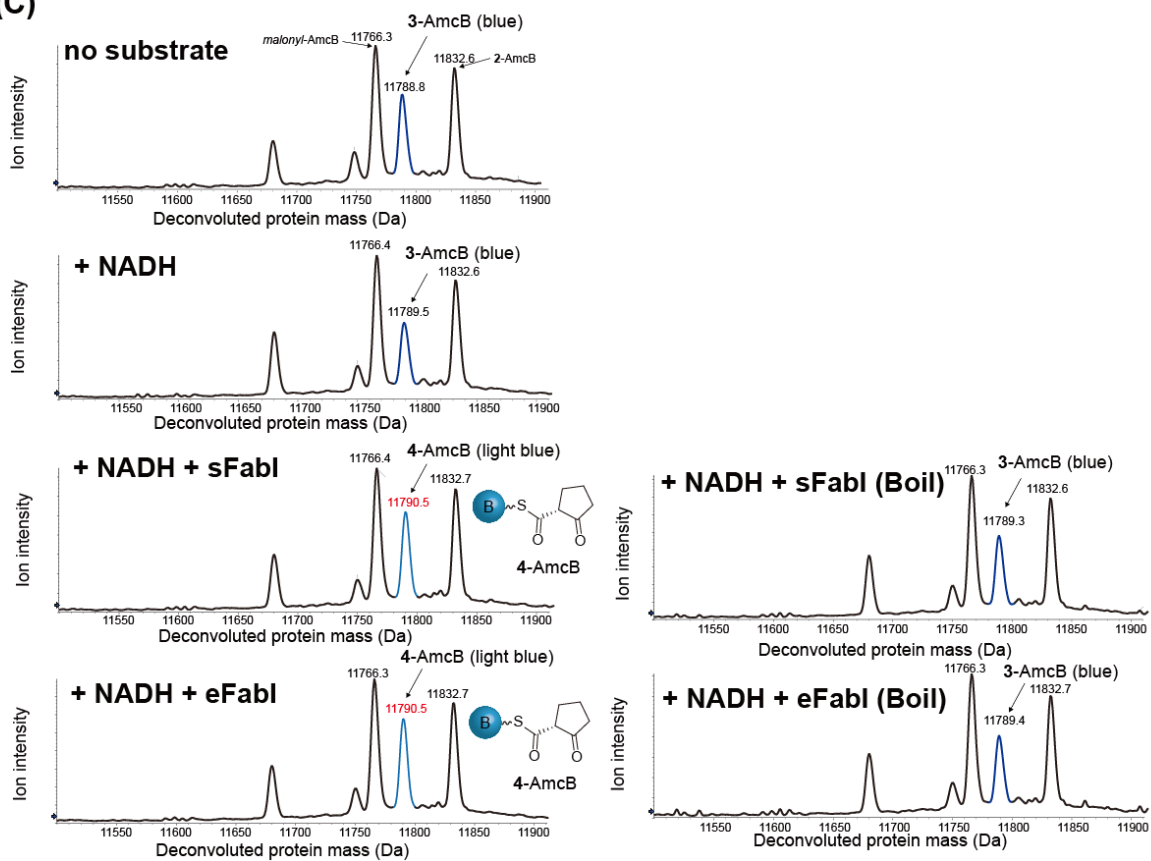
(A)

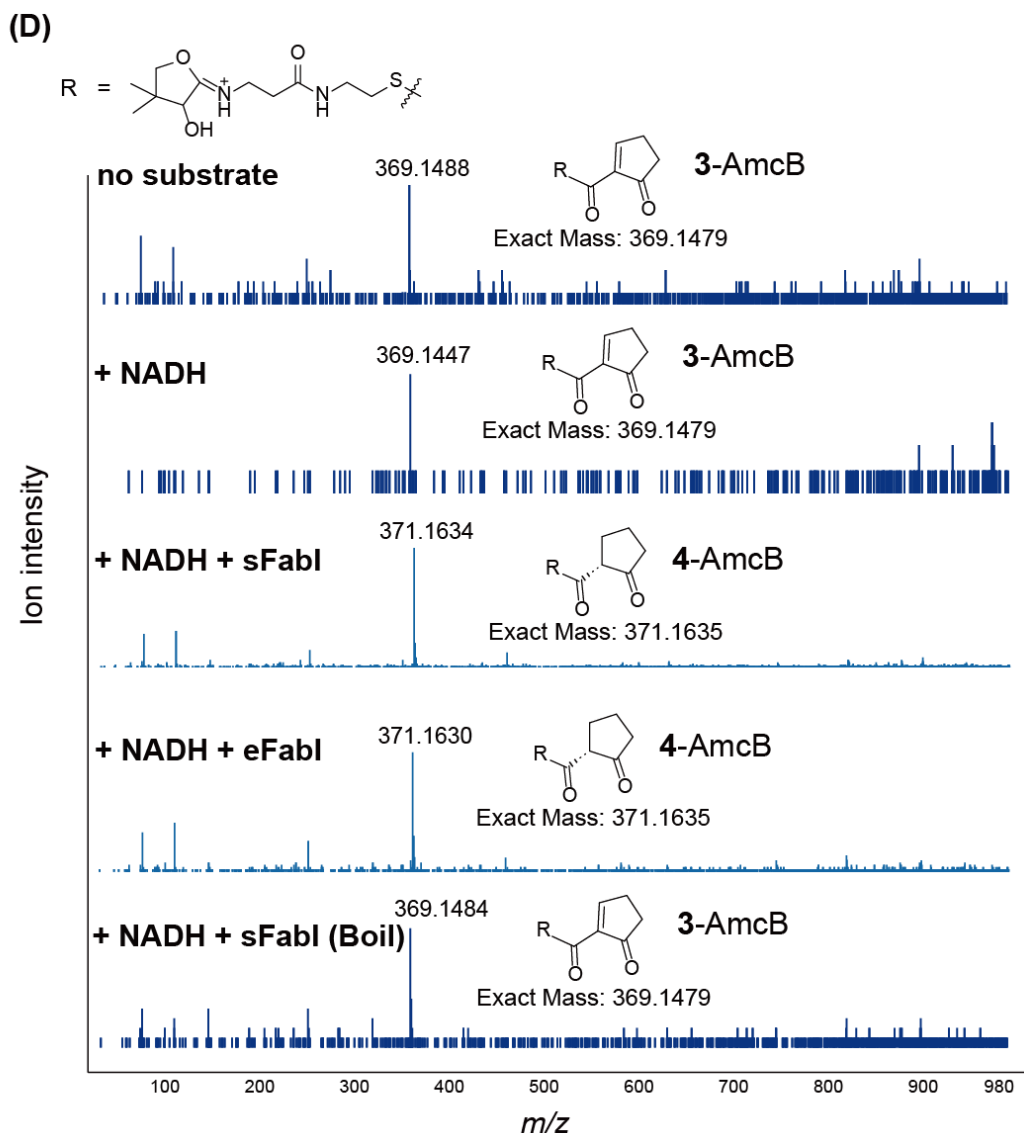


(B)



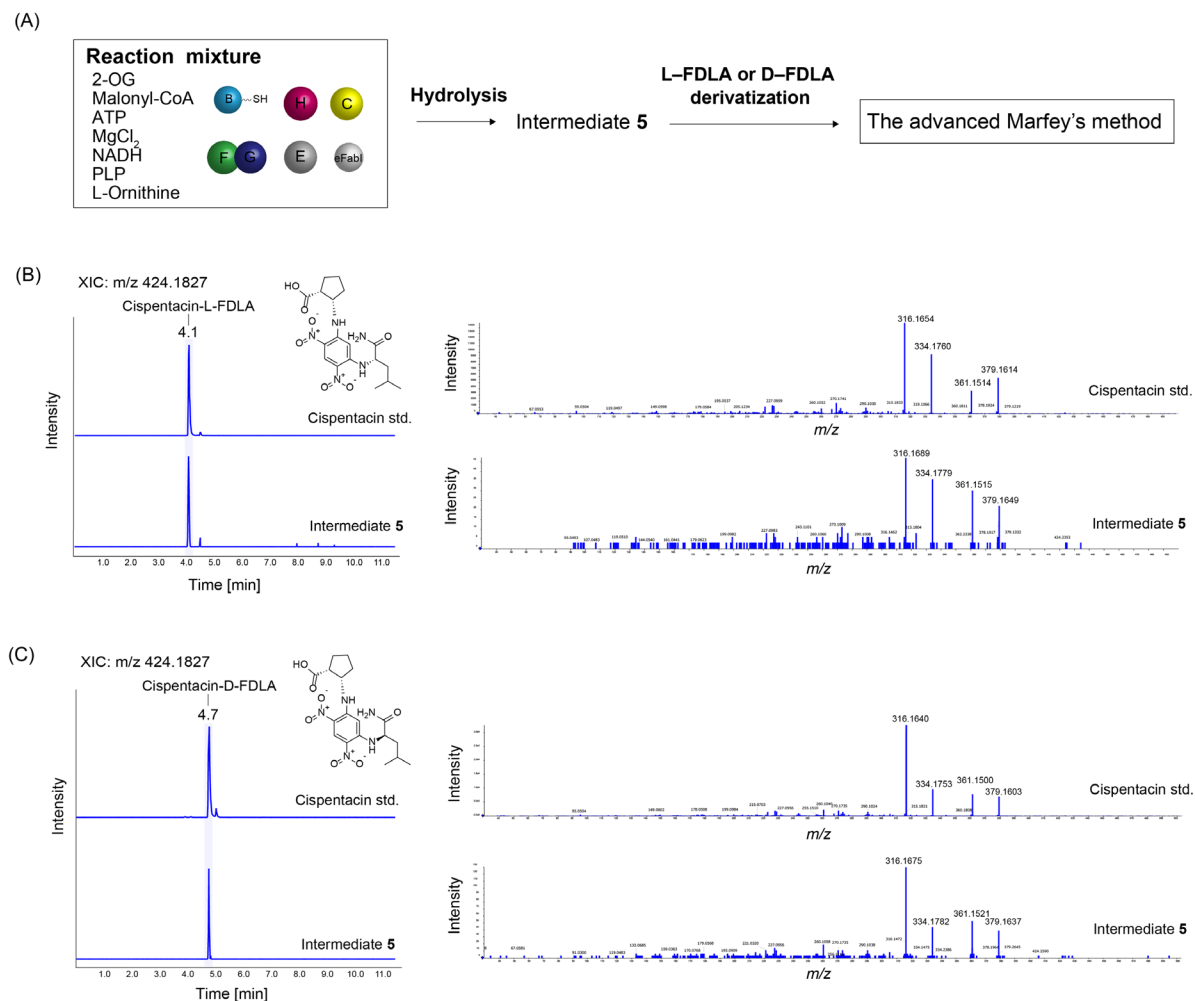
(C)





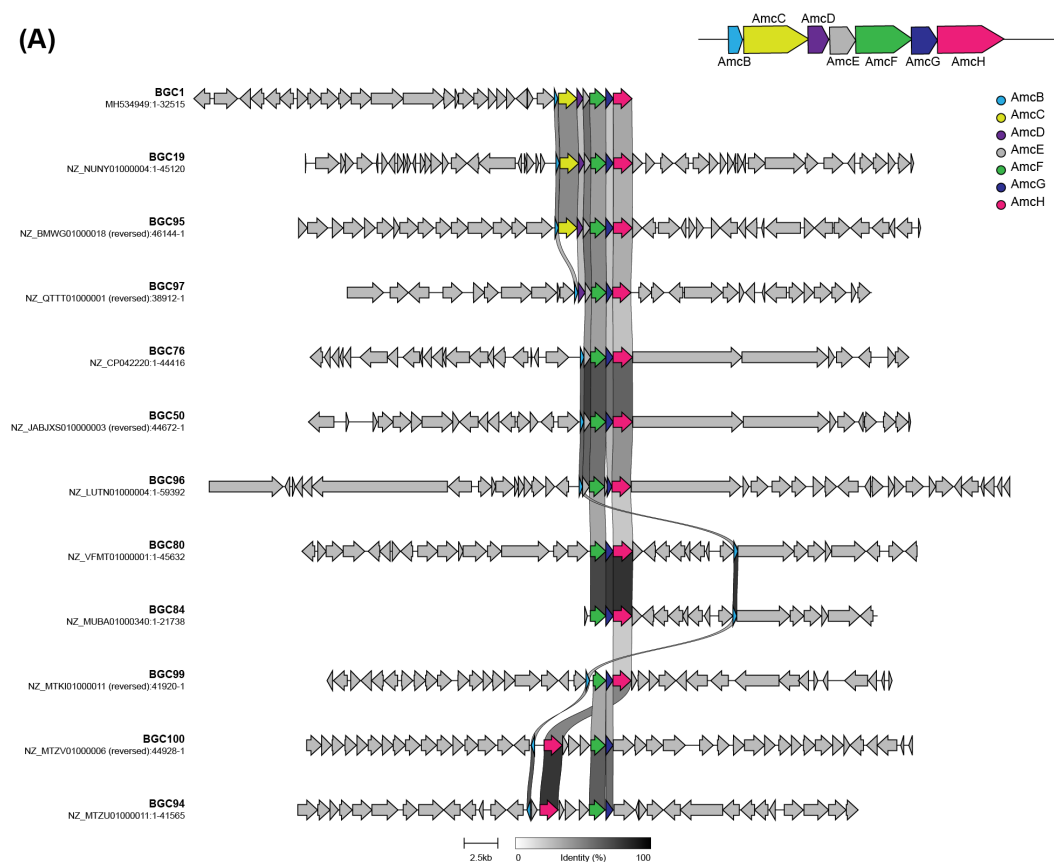
Supplementary Figure 23. Validation of the enzyme that catalyzes the reduction of 3-AmcB.

(A) Intermediate-AmcB, including 3-AmcB used for the reduction, was obtained by the Reaction 2 and then purified by gel filtration. Deconvoluted protein mass spectra were obtained for the reaction product under each condition. MS/MS fragments of the 13+ charge state of intermediate-AmcB were analyzed. The MS/MS fragment of 4-AmcB was detected only under the condition (+ NADH + AmcB). 4-AmcB was not detected under boiling conditions, indicating that the reaction is enzyme dependent. (B) Analysis of the reaction solution reacted with sFabI and eFabI. Deconvoluted protein mass spectra for the reaction product under each condition. MS/MS fragments of the 13+ charge state of intermediate-AmcB were analyzed. (C) The MS/MS fragment of 4-AmcB was detected under the conditions (+ NADH + sFabI, + NADH + eFabI). 4-AmcB was not detected when boiled FabI was used, indicating that the reaction is FabI dependent.

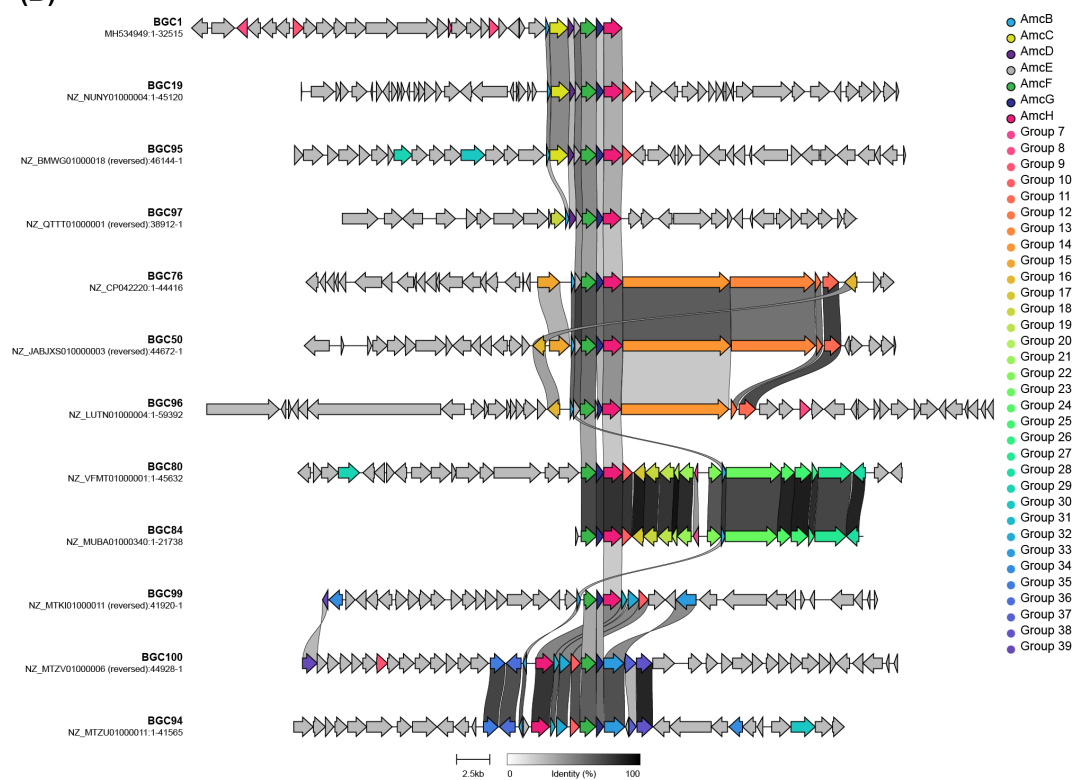


Supplementary Figure 24. LC–MS analysis of FDLA-derivatized cisptentacins. (A) 5-AmcB was prepared in a Reaction 5 mixture and then digested from AmcB via alkaline hydrolysis. The resulting 5-AmcB hydrolysates were analyzed by the advanced Marfey's method [7, 8, 9]. (B) Extracted ion count (XIC) chromatograms at m/z 424.1827 corresponding to the L-FDLA-derivatized intermediate 5 and the L-FDLA-derivatized cisptentacin standard. The retention time in LC–MS and the MS/MS spectrum of the derivatized intermediate 5 match well with those of the derivatized standard. (C) XIC chromatograms at m/z 424.1827 corresponding to the D-FDLA-derivatized intermediate 5 and the D-FDLA-derivatized cisptentacin standard. The retention time in LC–MS and the MS/MS spectrum of the derivatized intermediate 5 match well with those of the derivatized standard.

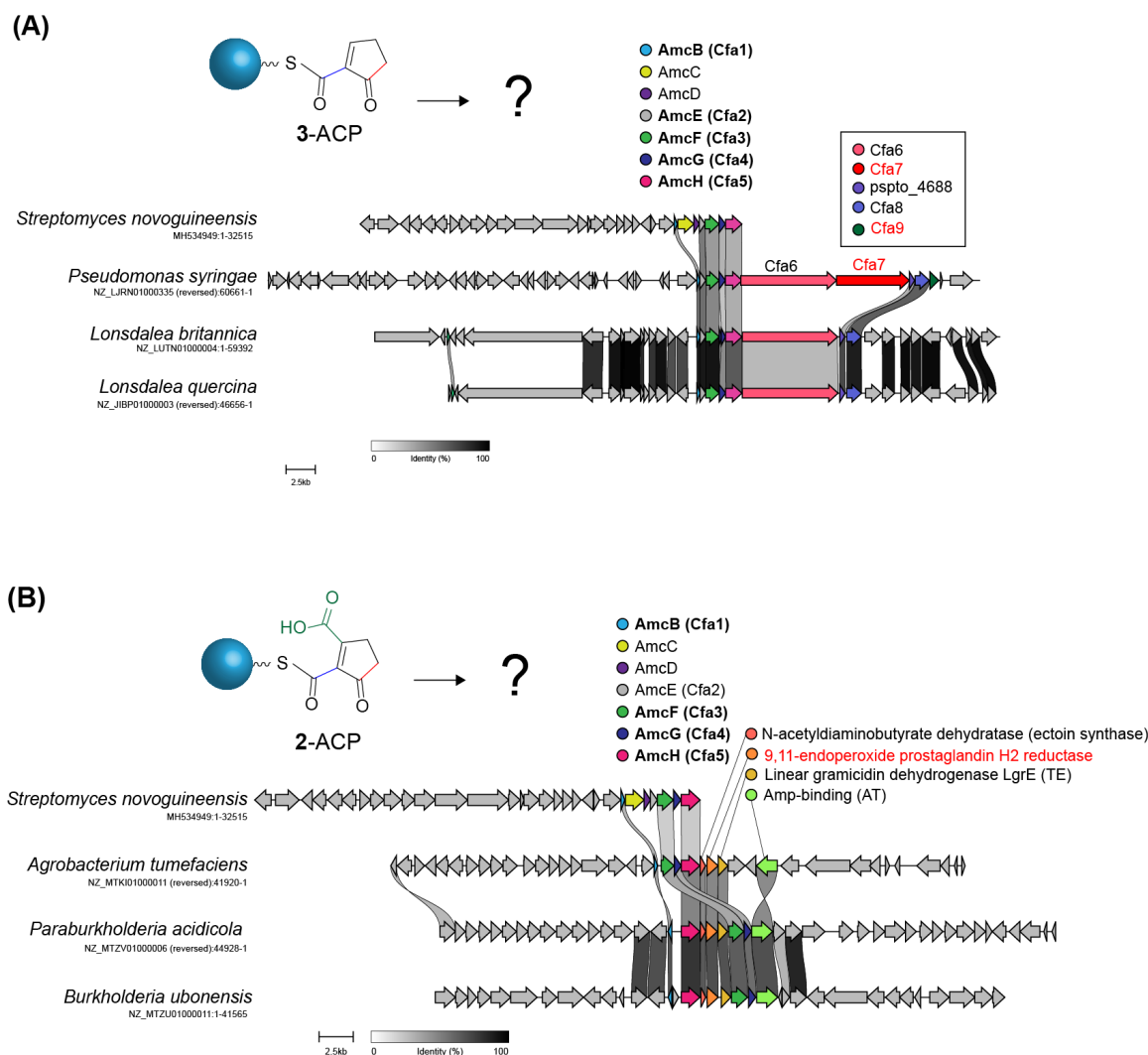
(A)



(B)



Supplementary Figure 25. Organization of *amcG* homolog-containing BGC. The figure was generated using a clinker tool [10]. The numbers for each BGC are based on the Supplementary Data 2. (A) The cispentacin homolog genes are colored. BGC1, 19, and 95 represent the cispentacin type. BGC76, 50, and 96 represent the CFA type. BGC80, 84, 99, 100, and 94 represent the minimal type. BGC97 lacking AmcC does not belong to any type. (B) In addition to the cispentacin homologous genes, the other homologous genes (identity threshold: 0.3) are also colored.

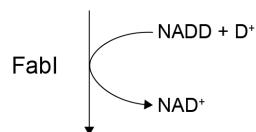
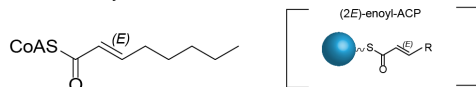


Supplementary Figure 26. Representatives of the *amcG* homolog-containing BGCs for unknown natural products. (A) CFA-type BGCs of *Lonsdalea britannica* and *Lonsdalea quercina* *cfa7* (type I PKS) and *cfa9*, both of which are CFA biosynthesis genes of *Pseudomonas syringae*, are absent in these BGCs. (B) Minimal-type BGCs of *Agrobacterium tumefaciens*, *Paraburkholderia acidicola*, and *Burkholderia ubonensis*. 9,11-Endoperoxide prostaglandin H2 reductase and several genes are conserved in these BGCs.

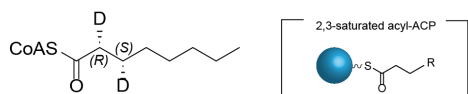
(A)

Fillgrove, K.L. & Anderson (2000)

trans-2-octenoyl-CoA



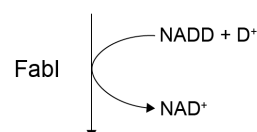
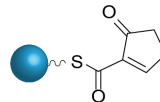
(2R,3S)-[2,3-²H₂]octanoyl-CoA



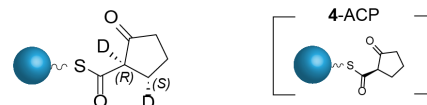
(B)

Predicted reaction mechanism

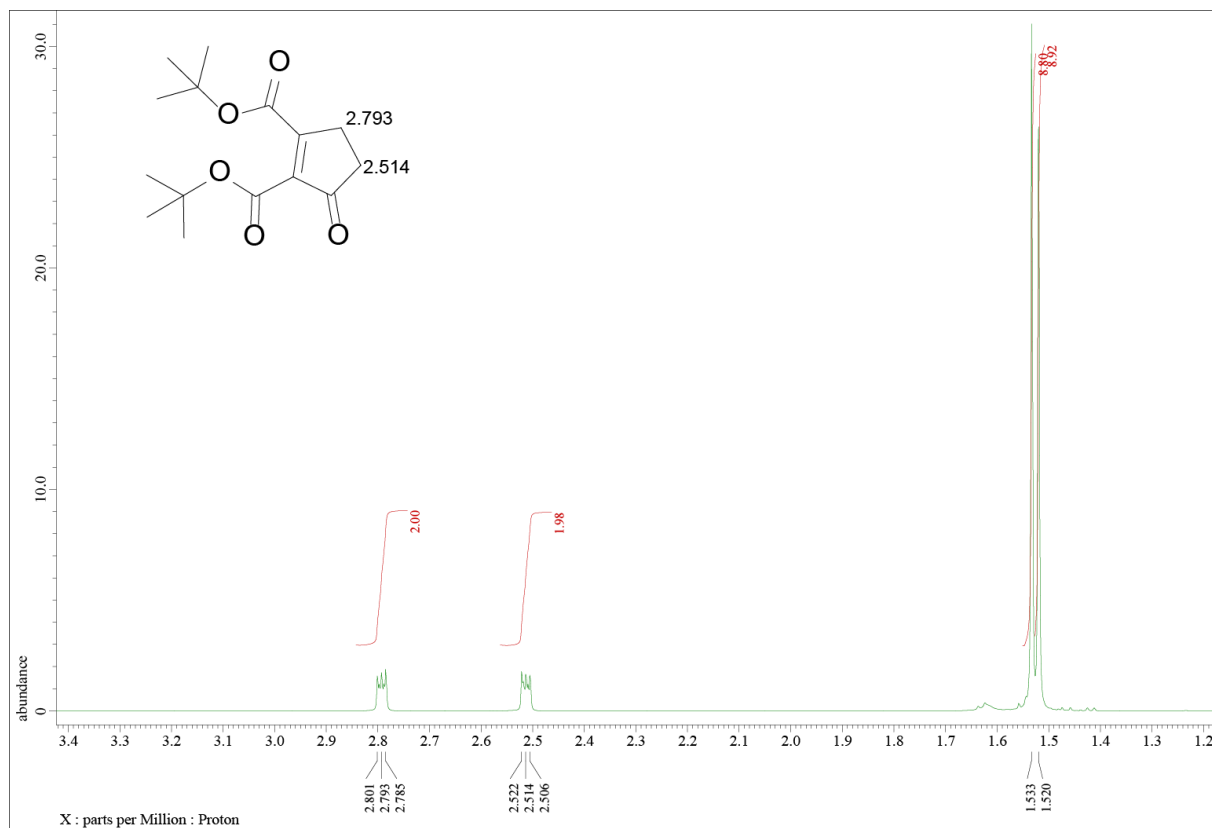
3-ACP



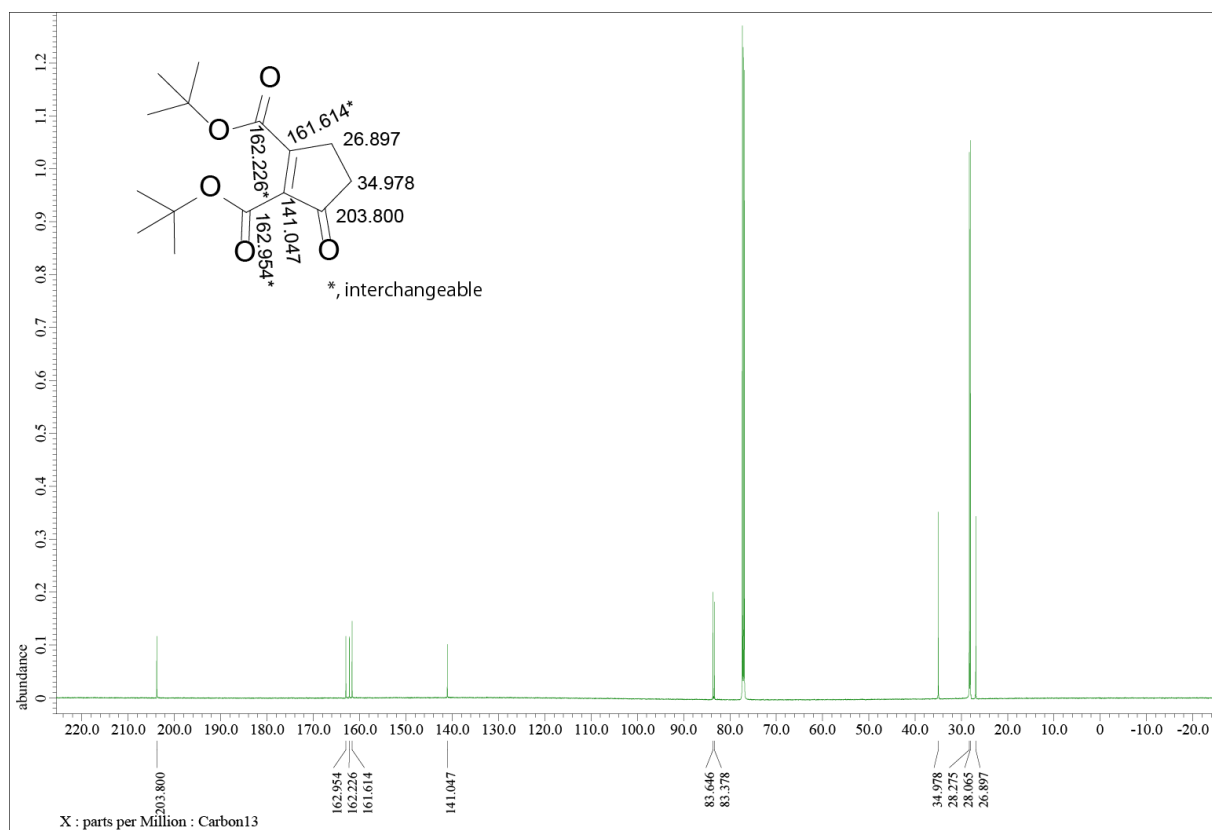
(1R,5S)-[1,5-²H₂]-2-oxocyclopentane-1-carboxyl-ACP



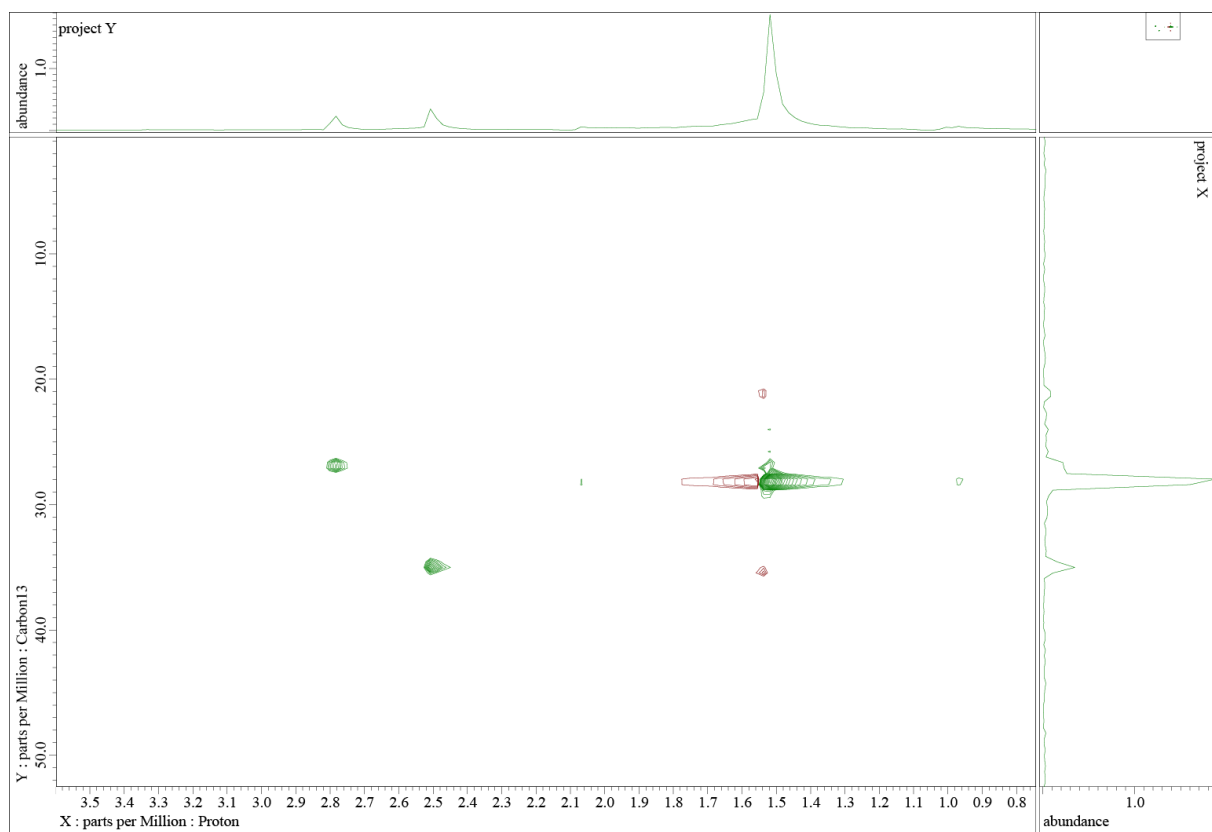
Supplementary Figure 27. FabI-catalyzed reduction. (A) FabI catalyzes the *syn* addition of 2H to the C2–C3 double bond of fatty acid intermediates. Labeling experiments with 4-²H-labeled β-nicotinamide adenine dinucleotide phosphate (NADD) in D₂O revealed that FabI catalyzes the stereoselective reduction of the C2–C3 double bond of fatty acid intermediates to yield (*R*)-C2 and (*S*)-C3 [11]. (B) By analogy with fatty acid biosynthesis, FabI should catalyze the *syn* addition of 2H to the double bond of 3-ACP to yield (1*R*)-2-oxocyclopentane-1-carboxyl-ACP (4-ACP). The stereochemistry of C1 in 4-ACP is identical to that of C1 in cispentacin.



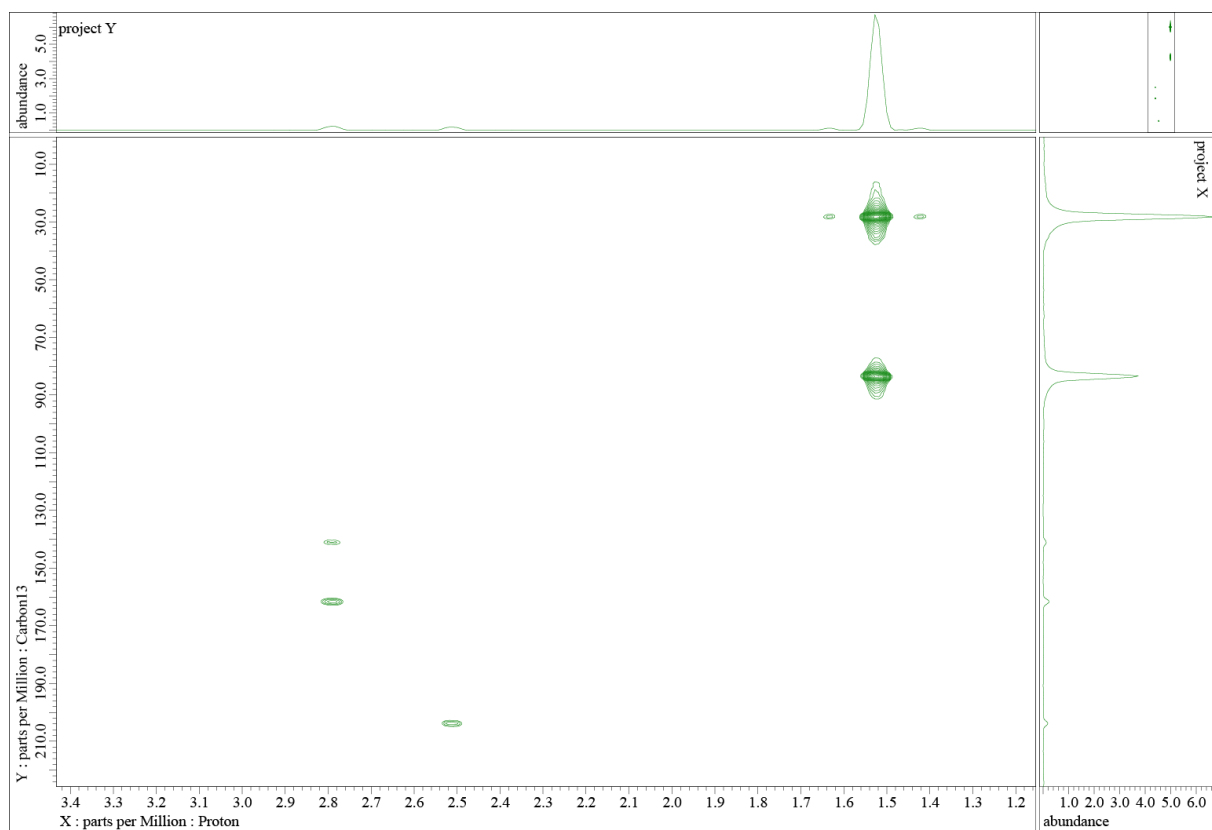
Supplementary Figure 28A. ^1H NMR spectrum of 2,3-di(carbo-*tert*-butoxy)-2-cyclopentenone in CDCl_3 at 600 Hz.



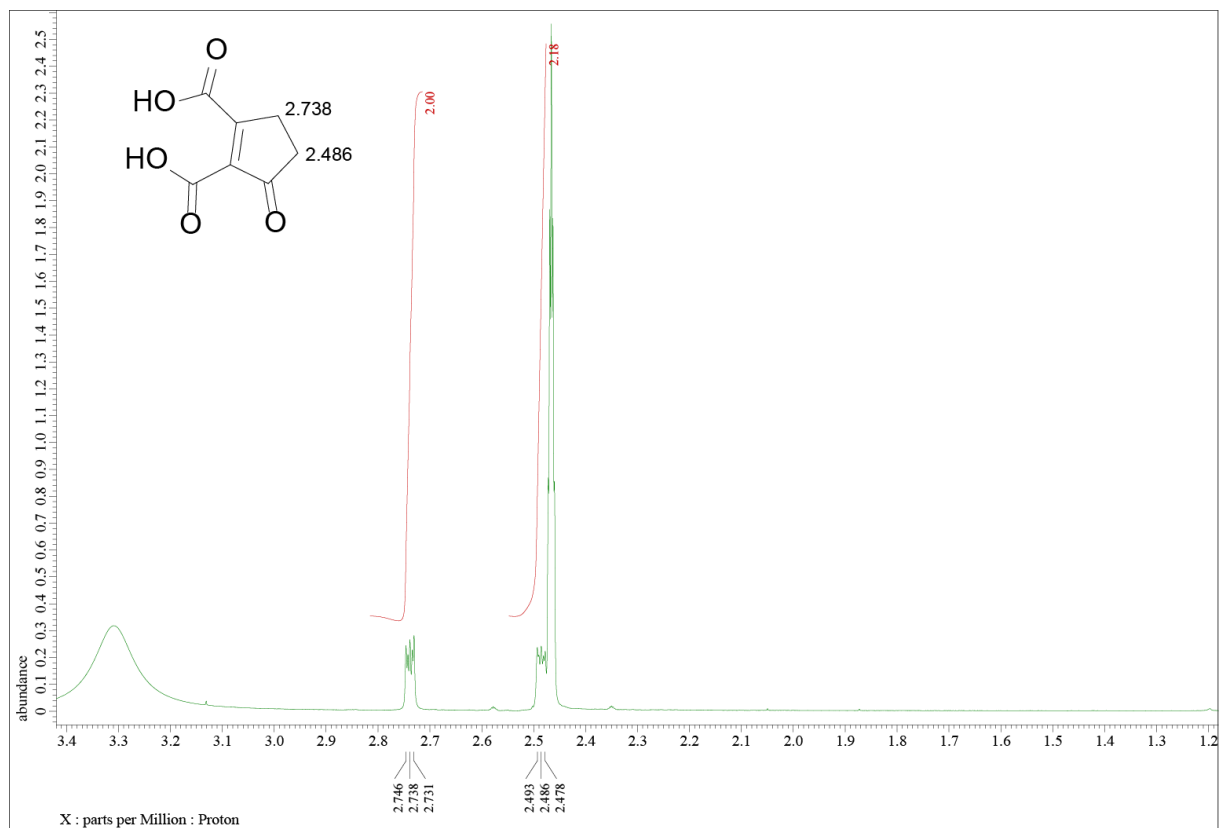
Supplementary Figure 28B. ^{13}C NMR spectrum of 2,3-di(carbo-*tert*-butoxy)-2-cyclopentenone in CDCl_3 at 150 Hz.



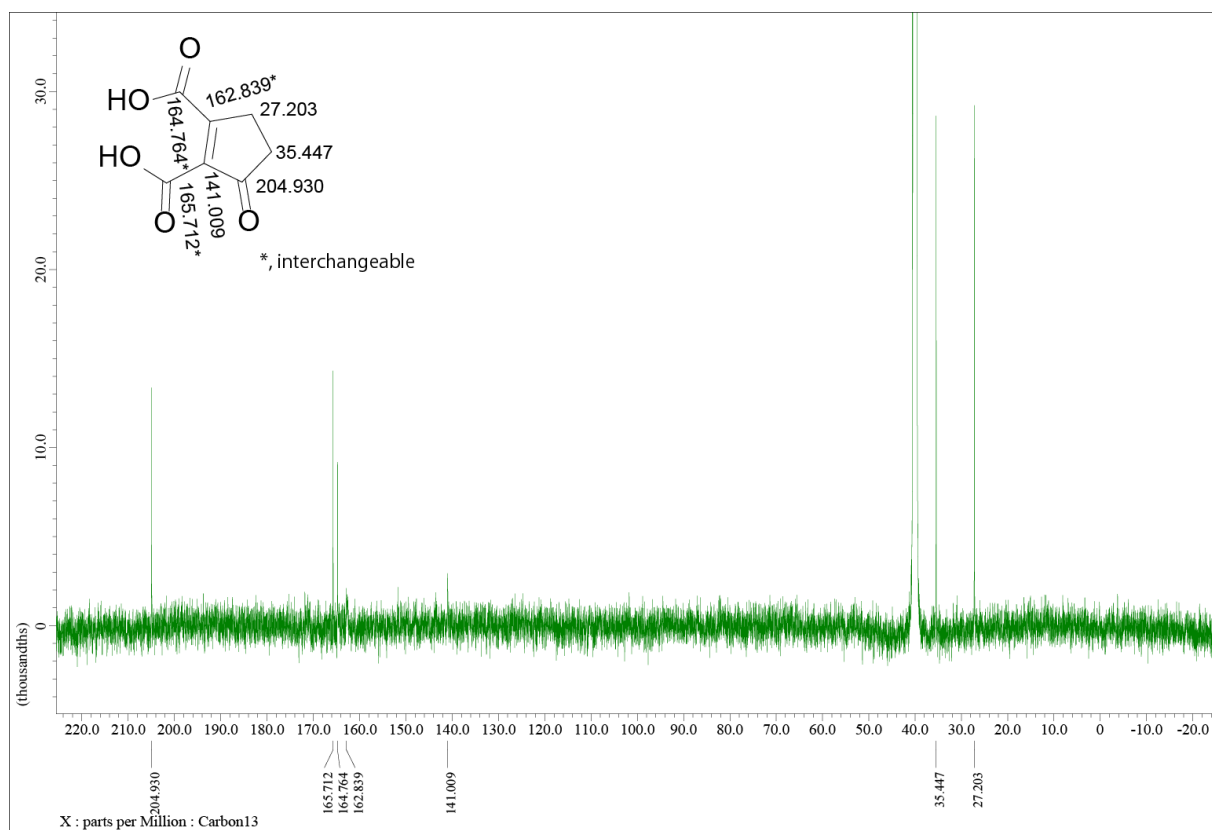
Supplementary Figure 28C. HSQC spectrum of 2,3-di(carbo-*tert*-butoxy)-2-cyclopentenone in CDCl₃.



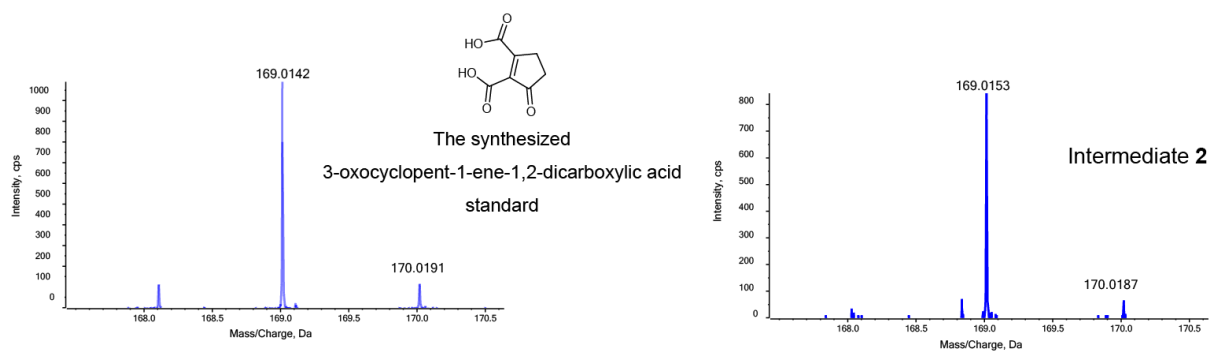
Supplementary Figure 28D. HMBC spectrum of 2,3-di(carbo-*tert*-butoxy)-2-cyclopentenone in CDCl₃.



Supplementary Figure 29A. ^1H NMR spectrum of 3-oxocyclopent-1-ene-1,2-dicarboxylic acid in $\text{DMSO}-d_6$ at 600 Hz.



Supplementary Figure 29B. ^{13}C NMR spectrum of 3-oxocyclopent-1-ene-1,2-dicarboxylic acid in $\text{DMSO}-d_6$ at 150 Hz.



Supplementary Figure 30. HRMS spectra of 3-oxocyclopent-1-ene-1,2-dicarboxylic acid.

Supplementary References

- (1) Dairi, T., Aisaka, K., Katsumata, R. & Hasegawa, M. A self-defense gene homologous to tetracycline effluxing gene essential for antibiotic production in *Streptomyces aureofaciens*. *Biosci. Biotechnol. Biochem.* **59**, 1835-1841 (1995).
- (2) Romo, A.J. et al. The amipurimycin and miharamycin biosynthetic gene clusters: unraveling the origins of 2-aminopurinyl peptidyl nucleoside antibiotics. *J. Am. Chem. Soc.* **141**, 14152-14159 (2019).
- (3) Jez, J.M., Ferrer, J.-L., Bowman, M.E., Dixon, R.A. & Noel, J.P. Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. *Biochemistry* **39**, 890-902 (2000).
- (4) Meguro, A. et al. An unusual terpene cyclization mechanism involving a carbon-carbon bond rearrangement. *Angew. Chem. Int. Ed.* **54**, 4353-4356 (2015).
- (5) Du, D., Katsuyama, Y., Shin - ya, K. & Ohnishi, Y. Reconstitution of a type II polyketide synthase that catalyzes polyene formation. *Angew. Chem.* **130**, 1972-1975 (2018).
- (6) Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425 (1987).
- (7) Marfey, P. Determination of D-amino acids. II. Use of a bifunctional reagent, 1, 5-difluoro-2, 4-dinitrobenzene. *Carlsberg Res. Commun.* **49**, 591-596 (1984).
- (8) Fujii, K., Ikai, Y., Oka, H., Suzuki, M. & Harada, K.-i. A nonempirical method using LC/MS for determination of the absolute configuration of constituent amino acids in a peptide: combination of Marfey's method with mass spectrometry and its practical application. *Anal. Chem.* **69**, 5146-5151 (1997).
- (9) Fujii, K. et al. A nonempirical method using LC/MS for determination of the absolute configuration of constituent amino acids in a peptide: Elucidation of limitations of Marfey's method and of its separation mechanism. *Anal. Chem.* **69**, 3346-3352 (1997).
- (10) Gilchrist, C.L. & Chooi, Y.-H. Clinker & clustermap. js: Automatic generation of gene cluster comparison figures. *Bioinformatics* **37**, 2473-2475 (2021).
- (11) Fillgrove, K.L. & Anderson, V.E. Orientation of Coenzyme A Substrates, Nicotinamide and Active Site Functional Groups in (Di) enoyl- coenzyme A Reductases. *Biochemistry* **39**, 7001-7011 (2000).