Electronic Supplementary Material (ESI) for Organic Chemistry Frontiers.

Characterization of GvgD and GvgH encoded in the biosynthetic gene cluster of 4-formylaminooxyvinylglycine

Linlin Pang,[a] Weijing Niu,[a] Yuwei Duan,[a] Xiaoying Bian,[a] Youming Zhang[a,c] and Guannan Zhong*[a,b]

a, Helmholtz International Laboratory for AntiInfectives, State Key Laboratory of Microbial Technology, Shandong University, Qingdao 266237, China.

b, Suzhou Research Institute of Shandong University, Suzhou 215123, China.

c, CAS Key Laboratory of Quantitative Engineering Biology, Shenzhen Institute of Synthetic Biology and Faculty of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China.

* To whom correspondence should be addressed. Email: zhonggn@sdu.edu.cn

Supplementary Methods

Materials and methods

Materials, bacterial strains, and plasmids. Biochemicals and media were purchased from Sinopharm Chemical Reagent Co. Ltd. (China) or Oxoid Ltd. (UK) unless stated otherwise. Enzymes were purchased from New England Biolabs Ltd. (UK) except *ApexHF* HS DNA polymerase for high-fidelity amplification from Accurate Biology Co. Ltd. (China). Chemical compounds and reagents were purchased from Bide Pharmatech Ltd. (China), Macklin Biochemical Technology Co., Ltd. (China) and J&K Scientific Ltd. (China) unless stated otherwise. Gene synthesis and codon optimization were performed at GENEWIZ, Inc (China). Primer synthesis and DNA sequencing were performed at Shanghai Sangon Biotech Co. Ltd. (China). Bacteria strains and plasmids used in this study are listed in Table S2. Primers used in this study are summarized in Table S3.

Analysis. High performance liquid chromatography (HPLC) analysis was carried out on Thermo Fisher Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific Inc., USA). Electrospray ionization mass spectrometry (ESI-MS) was performed on Bruker AmaZon SL Ion Trap LC/MS spectrometer (Bruker Co. Ltd, Germany), and the data were analyzed using Bruker Daltonics DataAnalysis. ESI-high resolution MS (ESI-HR-MS) analysis was carried out on Bruker High Resolution Q-TOF mass spectrometry (impactHD) (Bruker Co. Ltd, Germany) and the data were analyzed using Bruker Daltonics DataAnalysis. NMR data were recorded on the Bruker Avance Neo 600 MHz NMR spectrometer (Bruker Co. Ltd, Germany).

Sequence analysis. Open reading frames (ORFs) were identified using the FramePlot 4.0beta program (http://nocardia.nih.go.jp/fp4/). The deduced proteins were compared with other known proteins in the databases using available BLAST methods (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Amino acid sequence alignments were performed using the Strap program (http://www.bioinformatics.org/strap/). Homology modeling of the GvgD structure was conducted using the I-TASSER on-line server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/).

Gene sequences of the codon optimized *gvgD*、*gvgH*、*metY*$_{WH6}$、*metZ*$_{WH6}$

gvgD:

ATGAGCCTGGTGGAAGTGCATAACGAATGGGACCCGCTGGAAGAAATGATTGTGGGCGTGGCGCA
TGGCGCGCGCGTGCCGCGCCCGGATCGCGGCCTGTTTGCGCTGGATTATAGCGAACATCATGATA
GCCCGCATGAAATTCCGAGCGGCCCGTATCCGGAACAGCTGATTGAAGAAGCGCAAGAAGATCTG
GATGCGTTTGCGGCGTTTCTGGTGGGCCATGGCGTGACCGTGCGCCGCCCGCGCGAAACCTATCA
TGCGGCGCTGTTTGGCACCGCGGATTGGAAAACCGATGGCGAATATAACTATTGCCCGCGCGATG
TGCTGCTGCCGATTGGCAAAACCATTATTGAAGCGCCGATGGCGCTGCGCAGCCGCTATTTTGAA
CCGTTTGCGTATCGCGAACATCTGCAAGCGTATTTTGCGAGCGGCGCGAACTGGATTAGCGCGCC
GAAACCGGAACTGCCGGATAGCACCTATCGCGTGAACCCGCGCGAAGGCAGCATTATTGCGAACG
ATGAACCGATTTTTGATGCGGCGAACGTGCTGCGCATGGGCCGCGATATTCTGTATCTGGTGAGC
CGCAGCGGCAACCGCCTGGGCTATGAATGGCTGAAACGCGTGCTGGGCGATCAGTATCGCGTGCA
TGCGCTGAGCGGCTTTTATGATGGCACCCATCTGGATACCACCATTACCCTGGTGCGCCCGGGCC
TGGTGGTGCTGAACCCGGAACGCATTGGCAAAGATCAAGTGCCGGATGTGTTTAAAAACTGGGAT
ATTATTTGGGCGCCGGATATGGTGGATACCGGCTATTGCTGGAGCTATCCGCGCGCGAGCATTTG
GCAAGGCATGAACTTTATTATGGTGAACCCGAGCCTGGCGGTGATTAACGATCAGCAAGTGCCGC
TGATTCGCGCGCTGGAAAAACATCATGTGGATGTGGCGCCGCTGAAAATGCGCCATGCGCGCAGC
CTGAGCGGCGGCTTTCATTGCGTGAGCGTGGATATTCGCCGTCGCGGCACCCTGGAAGATTGCCG
CTAA

gvgH:

ATGAGCAAAGTGATTTATAAAAACCTGAACACGAGCCCGATTCTGGCGGTGGGCGGCGCGGGCGT
GTGGCTGGAAGATGCGACCGGCAAACGCTATTTAGATACCTGCGGCGGCGTGGCGGTGAGCAGCC
TGGGCCATGGCCATCCGCGCATTGCGGCCGCGTTCGAACGCGAAGCGAAAAAATTAGCGTGGGCG
CATGCCGGCAGCTTCACCACGGCGGCCGCGGAGGAACTGGCGGAACGCCTGGTTGCCGCGAGTGG
TGGCCTGGCCCGCGCGCAGTTTCTGAGCGGCGGCAGCGAAGTGATGGAACTGGCGATGAAAATTG
CGTATCAGTATCAGTGCGAACGCGGCCTGCCGGAAAAAAGCCTGTTTATTAGCCGCCGTCAGAGC
TATCATGGCAGCACCCTGGGCACCCTGAGCATTAGCGGCAACCCGCAGCGTCAGAGCGTGTTTGG
TCAGCTGTTTGCGGCCGCGGAATTTGTGAGCCCGTGCTATGCGTATCGCGATCAGCGCCCGGATG
AAAGCGAAGAACAGTATGCGACCCGCCTGGCGGATGAACTGGATGCGAAAATTCGCAGCCTGGGC

AGCGAAAACGTGGCGGCGTTTTTTGCGGAAACCGTGGTGGGCAGCACCAACGGCGCGGTGCCGGC
GGTGCCGGGCTATTTTCGCAAAATTAAAGCGGTGTGCGAACGCCATGATGTGCTGCTGATTCTGG
ATGAAGTGATGGCGGGCATGGGCCGCACCGGCCAACTGTTTGCCTATGCCGATGATGGCATTGTG
CCGGATATGGTGGCGGTGGGCAAAGGCCTGGCGGCGGGCTATATGCCGATTAGCGCGCTGCTGAT
TAGCGATCAAGTGCATAGCGTGATGGCGCGCGGCAGCGGCGTGCTGGGCAACGGTCAGACCCATG
TGAACCATCCGCTGGCGTGCGCGATTGCGCTGGAAGTGCAGCAAGTGATTGAAGATGAAGATCTG
CTGGCGCAAGTGCGTCAGCGCGGCGAACAGCTGCGCGCGTGGCTGCGCGAAAGCCTGGCGGATCT
GGATATTGTGGGCGATGTGCGCGGCCGCGGCCTGTTTGTGGGCGTGGAATTTGTGGAAAGCCGCG
CGACCAAAGCGCCGTTTAAGGGCGGTGGCGCGTATGCGGCCGCGCTGAAACAAGAAGCGCTGCAG
CGCGGCCTGCTGATTTATCCGGGCAGCGGCACCGTGCAAGGCACCCATGGCAACCATGTGCTGTT
TGCGCCGCCGTTTATTACGAGCGAAGGCGAACTGGGTCAGATGGTGGAACGCTTTAACGCGGTGG
TGCGCGCGGTGGCGTATTAA

metY$_{WH6}$:
ATGAAAGATGCGACCATTGCGCTGCATCATGGCTTTAAAAGCGATCCGACCACCAAAGCGGTGGC
GGTGCCGATTTATCAGAACGTGGCGTTTGAATTTGATAACGCGCAGCATGGCGCGGATCTGTTTA
ACCTGGATGTGCCGGGCAACATTTATACCCGCATTATGAACCCGACCAACGATGTTTTAGAACAG
CGTCTGGCGGCCCTGGAAGGCGGCATTGCCGGTCTGGCCGTGAGCGCGGGCAGCGCGGCGATTCA
TTATGCGATTCAGACCCTGACCCGCGCGGGCGATAACATTGTGACCACCCCGCAGCTGTATGGCG
GCACCTATACCCTGTTTGCGCATCTGCTGCCGAGCTTTGGCGTGGATGTGCGCTTTGCGCGCGAT
GATAGCGCGGAAGCGATTGCGGAACTGATTGATGATAATACCAAACTGGTGTATTGCGAAAGCAT
TGGCAACCCGGCGGGCAACATTGTGGATCTGGAAGCGCTGGCGAACGTGGCGCATGATCGCGGCG
TGCCGCTGATGGTGGATAACACCGTGGCGACCCCGATTCTGTGCAAACCGATTCAGTTTGGCGCG
GATATTGTGGTGCATAGCGTGACCAAATATATTGGCGGCCATGGCAACAGCCTGGGCGGCGTGAT
TGTGGATAGCGGCACCTTTCCGTGGGCGCAGTATCCGGAAAAATTTCCGGGCCTGAACACCCCGG
AACCGGCGTATCATGGCGTGGTGTATACCGAAAAATTTGGCCCGGCGGCCTTTATTGCGCGTGCG
CGTACCGTTCCGCTGCGCAATACCGGCGCGGCCCTGGCGCCGATGAACGCGTTTCTGTTACTGCA
AGGCCTGGAAACCCTGGCGCTGCGCATGGAACGCCATACCGAAAACGCGCTGAAAGTGGCGCAGT
TTCTGCAAGGCCATACCGCGGTGGAATGGGTGAGCTATGCGGGCCTGCCGGATCATCCGCATCAT
GCGCTGGCGCAGAAATATCTGCAAGGCAAACCGAGCGCGATTCTGAGCTTTGGCCTGAAAGGCGG
CTTTGATGCGGGCGTGCGCTTTTATGATGCGCTGCAGATTTTTAAACGCCTGGTGAACATTGGCG
ATGCGAAAAGCCTGGCGTGCCATCCGGCGAGCACCACCCATCGTCAGATGAACGAACAAGAACAA
GCGAAAGCGGGCGTGAAACCGGAAATGATTCGCCTGAGCGTGGGCATTGAAGCGATTGAAGATCT
GATTGACGACCTGCAACAAGCGCTGGGCAGCACCCGCCTGTAA

metZ$_{WH6}$:
ATGAGCCAAGAATGGGATGCGGGCCGCCTGGATAGCGATCTGGATGGCGTGGCGTTTGATACCCT
GGCGGTGCGCGCGGGTCAGCATCGCACCCCGGAAGGCGAACATGGCGATCCGATGTTTTTTACGA
GCAGCTATGTGTTTCGCACCGCGGCGGATGCGGCCGCGCGCTTTGCGGGCGAAGTGCCGGGCAAC
GTGTATAGCCGCTATACCAACCCGACCGTGCGCGCGTTTGAAGAACGCATTGCGGCGCTGGAGGG
TGCGGAACAAGCCGTGGCCACCGCGACCGGCATGGCGGCGATTCTGGCGGTGGTGATGAGCCTGT
GCAGCGCGGGCGATCATGTGCTGGTGAGCCGCAGCGTGTTTGGCAGCACCATTAGCCTGTTTGAA
AAATATTTTAAACGCTTTGGCATTGAAGTGGATTATGTGCCGCTGGCGGATCTGAGCGGCTGGGA
TACCGCGATTAAAGCGAACACCCGCCTGCTGTTTGTGGAAAGCCCGAGCAACCCGCTGGCGGAAC

TGGTGGATATTGCGGCCCTGAGCGAAGTGGCGCATGCGAAAGGCGCGATGCTGGTGGTGGATAAC
TGCTTTTGCACCCCGGCGCTGCAGCAGCCGCTGAAACTGGGCGCGGATATTGTGGTGCATAGCGC
GACCAAATTTATTGATGGCCAAGGCCGCTGCATGGGCGGCGTGGTGGCGGGCCGCAGCGAACAGA
TGAAAGAAGTGGTGGGCTTTCTGCGCACCGCGGGCCCGACCCTGAGCCCGTTTAACGCGTGGATT
TTTCTGAAAGGCCTGGAAACCCTGGGCCTGCGCATGAAAGCGCATTGCGCGAACGCGCAAGCGCT
GGCGGAATGGCTGGAACAGCAAGATGGCATTGAAAAAGTGCATTATGCGGGCCTGAAAAGCCATC
CGCAGCATGAACTGGCGCAGCGTCAGCAGCGCGGCTTTGGCGCGGTGGTGAGCTTTGAAGTGAAA
GGCGGCAAAGAAGGCGCGTGGCGCTTTATTGATGCGACCCGCCTGATTAGCATTACCGCGAACCT
GGGCGATAGCAAAACCACCATTACCCATCCGAGCACCACGAGCCATGGCCGCCTGGCGCCGCAAG
AACGCGAAGCGGCGGGCATTCGCGATAGCCTGATTCGCGTGGCGGTGGGCCTGGAAGATGTGGCG
GATCTGCAAGCGGATCTGGCGCGCGGCCTGGCGGCGCTGTAA

Protein expression and purification

GvgD. The plasmid pFG1001 (pET28a+*gvgD*) was electrotransformed into *Escherichia coli* BL21 (DE3) for protein expression. GvgD fused to an *N*-terminal 6 x His tag was expressed at 25 °C and shook at 200 rpm for 24 h with 100 μM isopropyl-β-D-thiogalactopyranoside (IPTG) added at OD_{600}=0.6 for induction. Cells were harvested by centrifugation at 4 °C and re-suspended in lysis buffer containing 50 mM K_2HPO_4 (pH 8.0), 300 mM NaCl, 5 mM imidazole and 10% (v/v) glycerol. After disruption by a low-temperature ultra-high-pressure cell disrupter, the insoluble material was removed by centrifugation at 20000 g at 4 °C. The soluble fraction was subjected to purification using a Ni Bestarose FF column (Bestchrom, China) according to the manufacturer's protocol. The elution fraction containing the recombinant protein was desalted using a PD-10 Desalting Column (GE Healthcare, USA) into storage buffer (50 mM K_2HPO_4 (pH 8.0), 100 mM NaCl, 10% (v/v) glycerol and 1 mM 2-mercaptoethanol (2-ME)). The resulting protein was concentrated and stored at -80 °C. The purity of the protein was determined by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis, and the concentration was determined by the Bradford assay using bovine serum albumin (BSA) as the standard.

GvgH, MetY, MetZ. The plasmids pFG1002 (pET28a+*gvgH*), pFG1003 (pET28a+$metY_{WH6}$) and pFG1004 (pET28a+ $metZ_{WH6}$) were transferred into *E. coli* BL21 (DE3) for expression, respectively. The proteins were purified to homogeneity, and concentrated according to the procedures described above for GvgD. Storage buffer free of 2-mercaptoethanol (50 mM K_2HPO_4 (pH 8.0), 100 mM NaCl, 10% (v/v) glycerol) was utilized instead of the 2-mercaptoethanol containing buffer for GvgD.

Site-directed mutations of GvgD C332A, C332S, H227A, D179A, D108A, D229A and GvgH K268R. Plasmids containing site-directed mutations were generated using the primers listed in Table S2 for PCR amplification with pFG1001or pFG1002 as the template. Then, 1 μL of *Dpn*I enzyme was added to the PCR system to remove the template. The PCR system was desalted by 0.025 μm mixed cellulose ester (MCE) membrane (Merck Millipore Ltd., USA) and electrotransformed into *E. coli* BL21 (DE3). After sequencing to validate the fidelity, the resulting mutant proteins were expressed in *E. coli* BL21 (DE3), purified to homogeneity, and concentrated according to the procedures described above for the native proteins.

Site-directed mutation of GvgH K268A. Two DNA fragments containing K268A mutant (fragment 1) and the antibiotic resistance marker *kan* (fragment 2) were generated using the primers listed in Table S2 for PCR amplification with pFG1002 as the template, respectively. After purification by agarose gel electrophoresis, 600 ng fragment 1 and 400 ng fragment 2 were assembled in a 20 μL system including 10 μL 2× MultiF Samless Assembly Mix (ABclonal Technology Co.,Ltd., USA) at 50 °C for 2 h. The assembled system was desalted by

MCE membrane and electrotransformed into *E. coli* BL21 (DE3). After sequencing to validate the fidelity, the resulting mutant protein was expressed in *E. coli* BL21 (DE3), purified to homogeneity, and concentrated according to the procedures described above for the native proteins.

***In vitro* enzymatic assay for GvgD.** The assays for GvgD (total volume, 100 μL) were performed at 30 °C for 3 h in 50 mM Tris-HCl buffer (pH 7.5) containing 1 mM L-canaline (or L-ornithine, L-lysine, L-Dab), 1 mM L-arginine (or L-canavanine, L-homoarginine) in the presence of 40 μM GvgD or GvgD mutants. The assays were quenched by adding an equal volume (100 μL) of CH_3CN. After centrifugation, 4 μL of dansyl chloride (DNSC, 50 mM in CH_3CN) was added to the 200 μL of supernatant of the assays, followed by further incubation at 50 °C for 1 h to derivatize the amino group. After removal of precipitate by centrifugation, the mixtures were then subjected to HPLC-ESI-(HR)MS analysis on an Acclaim TM RSLC 120 C18 column (2.1 × 100 mm, 2.2 μm, 120 Å, Thermo Fisher Scientific Inc., USA) by gradient elution of solvent A (H_2O containing 0.1% formic acid) and solvent B (CH_3CN containing 0.1% formic acid) with a flow rate of 0.3 mL/min over a 25 min period as follows: T = 0 min, 5% B; T = 3 min, 5% B; T = 18 min, 95% B; T = 22 min, 95% B; T = 23 min, 5 % B; and T = 25 min, 5% B (mAU at 254 nm).

Kinetic studies for GvgD. To determine the apparent kinetic parameters of the forward (L-Arg + L-canaline) and reverse (L-canavanine + L-ornithine) reactions, the assays contained 50 mM Tris-HCl buffer (pH 7.5), 40 μM GvgD, and various concentrations of reactants (L-Arg + L-canaline, 0.5 mM to 2.5 mM, or L-canavanine + L-ornithine, 0.7 mM to 2.5 mM) in a final volume of 300 μL. To determine the apparent kinetic constants for L-Arg + L-Lys, the same reaction system was

constructed with L-Arg + L-Lys various from 0.7 mM to 2.5 mM. All reactions were performed at 30 °C, and terminated by addition of an equal volume of CH_3CN. After centrifugation and DNSC derivatization, the mixtures were then analyzed by HPLC-ESI-MS.

***In vitro* enzymatic assay for GvgH ($MetY_{WH6}$ or $MetZ_{WH6}$).** The assays for GvgH ($MetY_{WH6}$ or $MetZ_{WH6}$, total volume, 100 μL) were performed at 30 °C for 3 h in 100 mM K_2HPO_4 buffer (pH 7.5) containing 1 mM *O*-succinylhomoserine (or *O*-acetylhomoserine, 2-aminobut-3-enoic acid), 1 mM pyridoxal 5'-phosphate monohydrate (PLP), (and 1 mM 2-mercaptoethanol) in the presence of 50 μM GvgH (or $MetY_{WH6}$, $MetZ_{WH6}$). The assays were quenched by adding 90 μL of CH_3CN. After centrifugation, 4 μL of DNSC (50 mM in CH_3CN) and 10 uL of Li_2CO_3 (800 mM suspension in CH_3CN) were added to the supernatant of the assays, followed by further incubation at room temperature for 1 h to derivatize the amino group. The mixtures were then subjected to HPLC-ESI-(HR)MS analysis according to the procedure described above for GvgD.

Derivatization of the commercially available amino acids as standards. Commercially available amino acids were dissolved in Tris-HCl buffer (50 mM, pH 7.5) or K_2HPO_4 buffer (100 mM, pH 7.5) to obtain 1 mM aqueous solutions. Then 100 μL of the aqueous solutions were mixed with an equal volume of CH_3CN, 4 μL of DNSC (50 mM in CH_3CN) was added to each of the 200 μL amino acid aqueous solution, following by further incubation at 50 °C for 1 h to derivatize the amino group. After removal of precipitate by centrifugation, the mixtures were then subjected to HPLC-ESI-(HR)MS analysis according to the procedure described above for GvgD.

Table S1. Apparent kinetic parameters of GvgD on different substrates. All experiments were repeated three times.

Substrate	Apparent K_m (mM)	Apparent V_{max} (mM s^{-1})	Apparent K_{cat} (s^{-1})	Apparent K_{cat}/K_m (mM^{-1} s^{-1})
L-cnanaline + L-arginine	4.59	0.0962	2.41	0.524
L-ornithine + L-canavanine	6.13	0.220	5.51	0.899
L-lysine + L-arginine	13.67	0.00399	0.0998	0.00730

Table S2. Bacterial strains and plasmids.

Strain/Plasmid	Characteristic(s)	Source/ Reference
E. coli		
BL21 (DE3)	Host for protein expression	NEB
FG1001	BL21 (ED3) derivative, containing pFG1001 for producing GvgD	This study
FG1002	BL21 (ED3) derivative, containing pFG1002 for producing GvgH	This study
FG1003	BL21 (ED3) derivative, containing pFG1003 for producing $MetY_{WH6}$	This study
FG1004	BL21 (ED3) derivative, containing pFG1004 for producing $MetZ_{WH6}$	This study
FG1005	FG1001 derivative, containing pFG1005	This study
FG1006	FG1001 derivative, containing pFG1006	This study
FG1007	FG1001 derivative, containing pFG1007	This study
FG1008	FG1001 derivative, containing pFG1008	This study
FG1009	FG1001 derivative, containing pFG1009	This study
FG1010	FG1001 derivative, containing pFG1010	This study
FG1011	FG1002 derivative, containing pFG1011	This study
FG1012	FG1002 derivative, containing pFG1012	This study
Plasmids		
pET-28a(+)	Protein expression vector used in *E. coli*, encoding *N*-terminal 6× His-tag, kanamycin resistance	Novagen
pFG1001	pET28a(+) derivative containing *gvgD* (NdeI+XhoI)	This study
pFG1002	pET28a(+) derivative containing *gvgH* (NdeI+XhoI)	This study
pFG1003	pET28a(+) derivative containing $metY_{WH6}$ (NdeI+HindIII)	This study
pFG1004	pET28a(+) derivative containing $metZ_{WH6}$ (NdeI+HindIII)	This study
pFG1005	pET28a(+) derivative containing *gvgD* D108A	This study
pFG1006	pET28a(+) derivative containing *gvgD* D179A	This study
pFG1007	pET28a(+) derivative containing *gvgD* D229A	This study
pFG1008	pET28a(+) derivative containing *gvgD* H227A	This study
pFG1009	pET28a(+) derivative containing *gvgD* C332A	This study
pFG1010	pET28a(+) derivative containing *gvgD* C332S	This study
pFG1011	pET28a(+) derivative containing *gvgH* K268R	This study
pFG1012	pET28a(+) derivative containing *gvgH* K268A	This study

Table S3. Primers used in this study.

Primer	Sequence (5'-3')
GvgD D108A-for	AACTATTGCCCGCGCGCAGTGCTGCTGCCGATT
GvgD D108A-rev	AATCGGCAGCAGCACTGCGCGCGGGCAATAGTT
GvgD D179A-for	GATGAACCGATTTTTGCAGCGGCGAACGTGCTG
GvgD D179A-rev	CAGCACGTTCGCCGCTGCAAAAATCGGTTCATC
GvgD D229A-for	GATGGCACCCATCTGGCAACCACCATTACCCTG
GvgD D229A-rev	CAGGGTAATGGTGGTTGCCAGATGGGTGCCATC
GvgD H227A-for	TTTTATGATGGCACCGCACTGGATACCACCATT
GvgD H227A-rev	AATGGTGGTATCCAGTGCGGTGCCATCATAAAA
GvgD C332A-for	AGCGGCGGCTTTCATGCAGTGAGCGTGGATATT
GvgD C332A-rev	AATATCCACGCTCACTGCATGAAAGCCGCCGCT
GvgD C332S-for	AGCGGCGGCTTTCATAGCGTGAGCGTGGATATT
GvgD C332S-rev	AATATCCACGCTCACGCTATGAAAGCCGCCGCT
GvgH K268R-for	ATGGTGGCGGTGGGCCGTGGCCTGGCGGCGGGC
GvgH K268R-rev	GCCCGCCGCCAGGCCACGGCCCACCGCCACCAT
GvgH K268A-1-for	CCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAA
GvgH K268A-1-rev	GCTAATCGGCATATAGCCCGCCGCCAGGCCTGCGCCCACCGCCACCATAT CCG
GvgH K268A-2-for	GGCCTGGCGGCGGGCTATAT
GvgH K268A-2-rev	GCCAGAGTTGTTTCTGAAAC

Figure S1. Proposed inhibitive mechanism of oxyvinylglycine for the eliminase subgroup of PLP-dependent enzymes.[1]

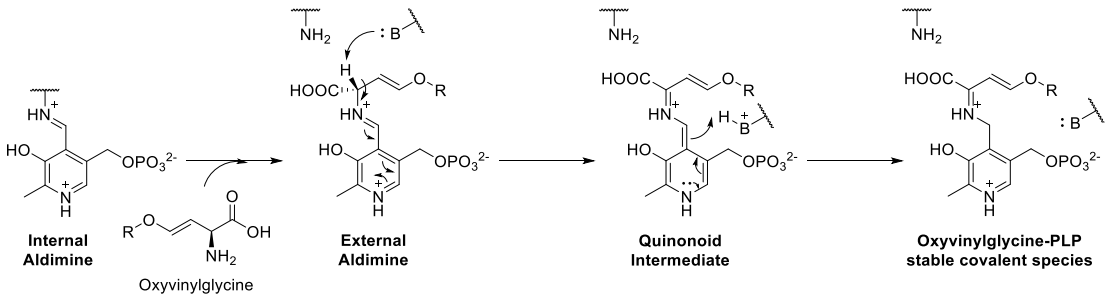


Figure S2. Distinct biosynthetic pathways of AMB (**A**), rhizobitoxine (**B**) and D-cycloserine (**C**).[2-11]

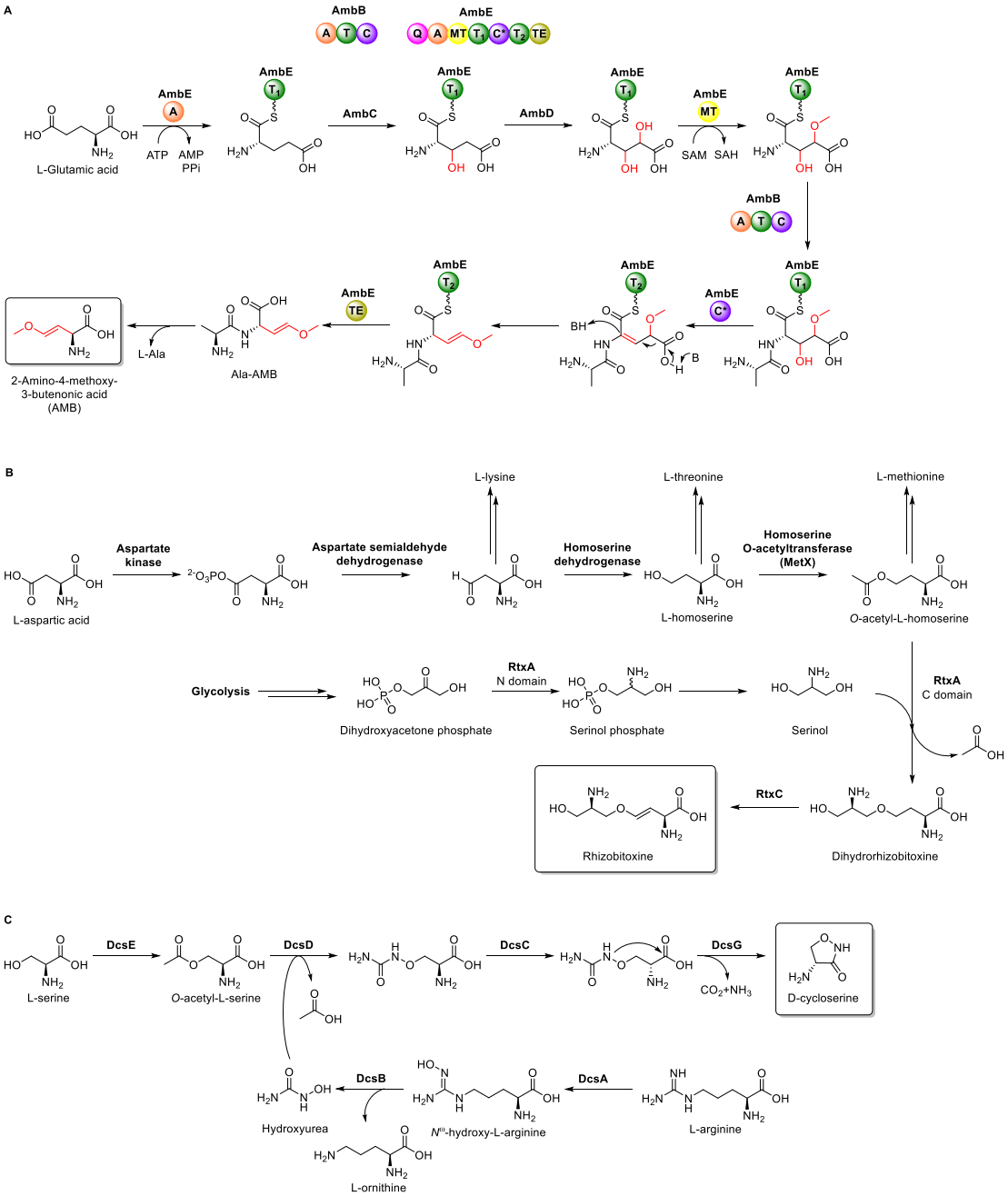
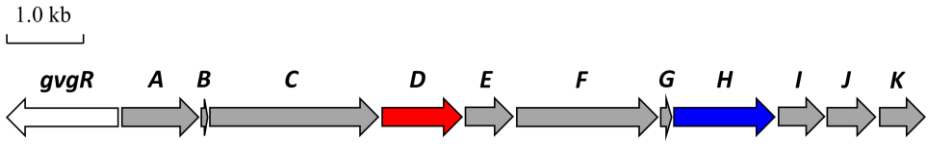


Figure S3. The biosynthetic gene cluster of FVG in *P. fluorescens* WH6 with predicted functions. *gvgD* and *gvgH* are highlighted in red and blue, respectively.[12-14]



Gene	Length (aa)	Predicted function
gvgR	484	GntR family transcriptional regulator
gvgA	335	SGNH/GDSL hydrolase family protein
gvgB	27	Hypothetical protein
gvgC	736	Iron-containing redox enzyme family protein
gvgD	347	Amidinotransferase
gvgE	207	LysE family transporter
gvgF	613	Carbamoyltransferase
gvgG	47	Hypothetical protein
gvgH	439	PLP-dependent aminotransferase
gvgI	202	Formyltransferase
gvgJ	203	LysE family transporter
gvgK	228	LysE family transporter

Figure S4. Denaturing SDS-PAGE analysis (conc. 10%) of GvgD and its mutants. Lane 1, Protein standard; Lane 2, GvgD (41.61 kDa); Lane 3, Protein standard; Lane 4, GvgD C332A; Lane 5, Protein standard; Lane 6, GvgD C332S; Lane 7, Protein standard; Lane 8, GvgD H227A; Lane 9, Protein standard; Lane 10, GvgD D179A; Lane 11, Protein standard; Lane 12, GvgD D108A.

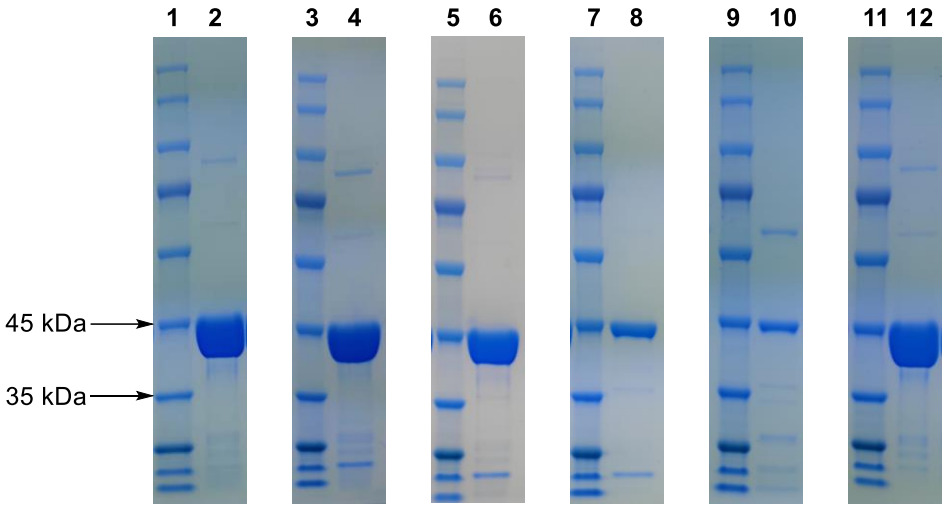


Figure S5. HPLC-ESI-HRMS of DNS-**3**, DNS-Orn, DNS-Arg and DNS-**2**, respectively.

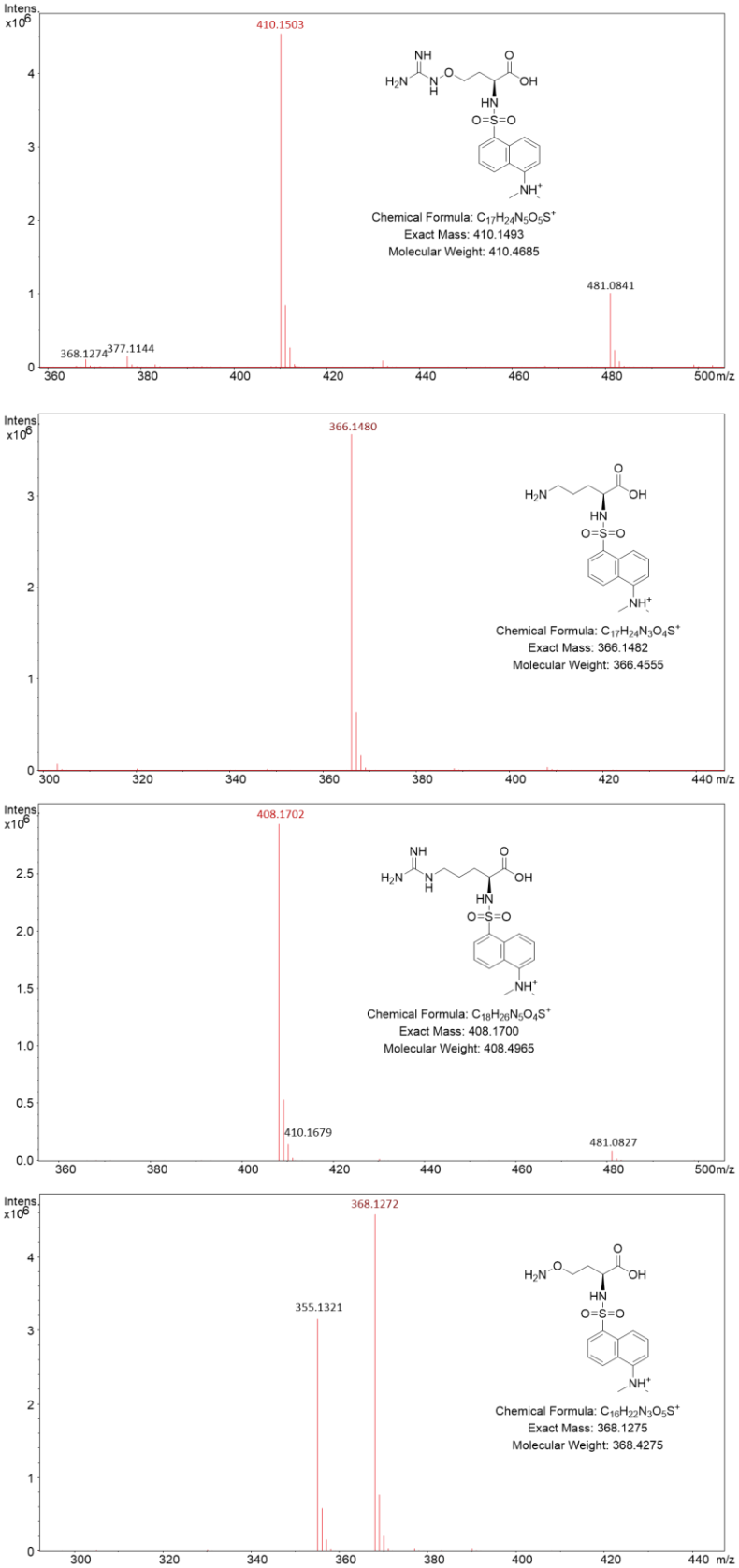


Figure S6. *In vitro* characterization of GvgD with L-Lys (**A**) or L-Dab (**B**) as amidino acceptors.

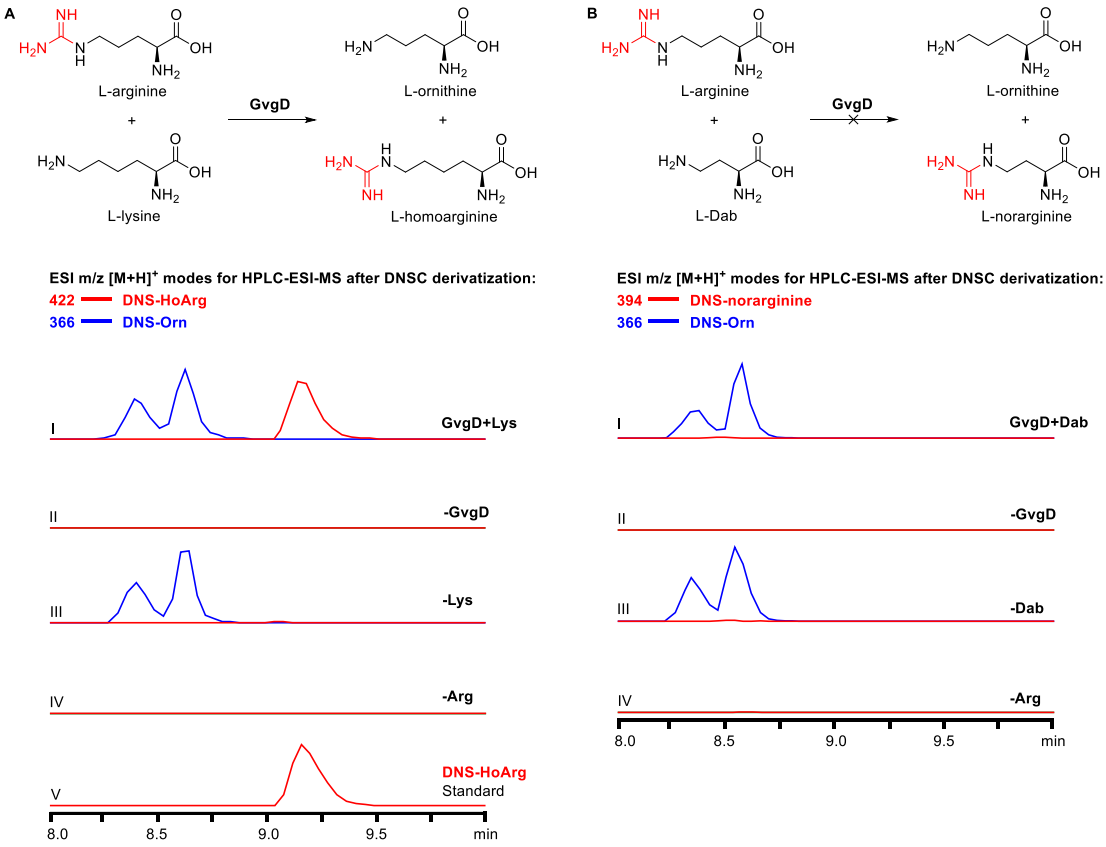


Figure S7. *In vitro* characterization of GvgD with L-HoArg as amidino donor and L-canaline (**A**) or L-Orn (**B**) as amidino acceptors.

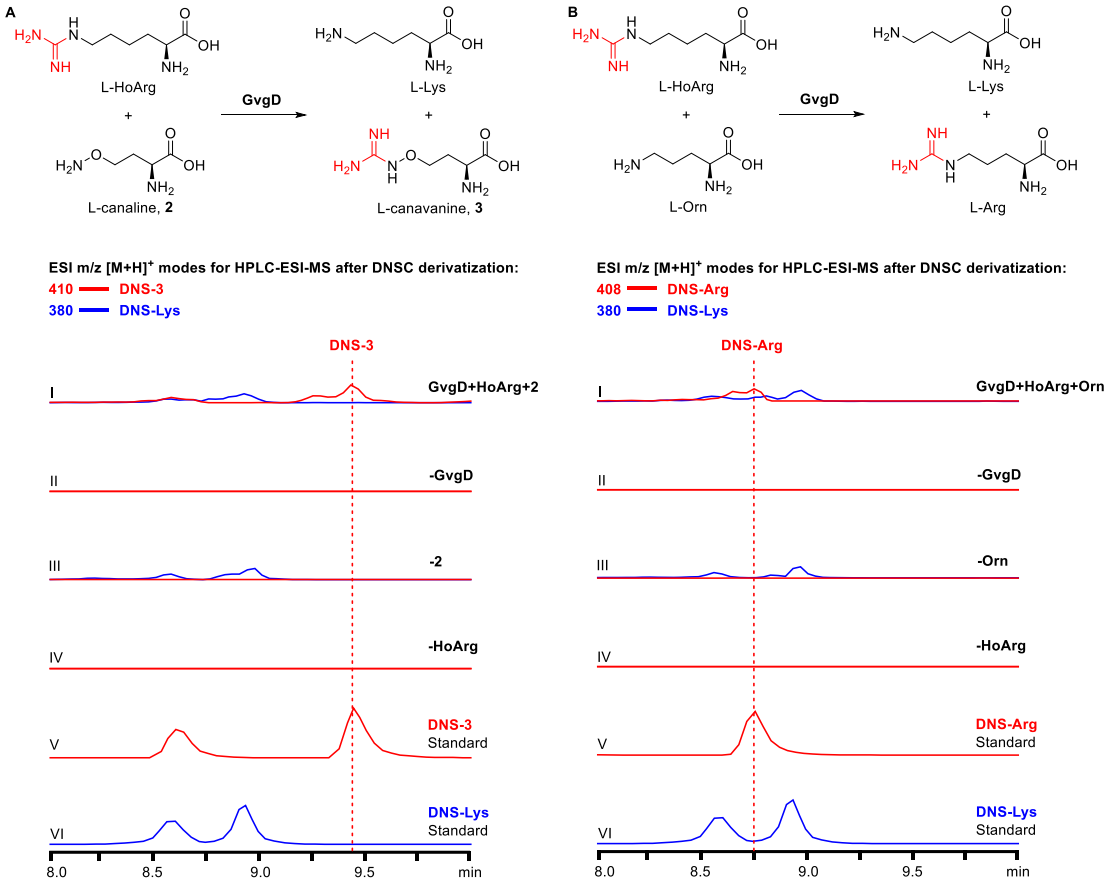


Figure S8. Homology modeling of GvgD performed by I-TASSER. (**A**) The alignment of GvgD (Green) and StxG (Cyan, Protein DataBank entry 6U1R[15]). (**B**) Partial enlarged detail of GvgD (Green) and StxG (Cyan) at the catalytic sites. The related residues (indicated in sticks) of the two proteins are close in space.

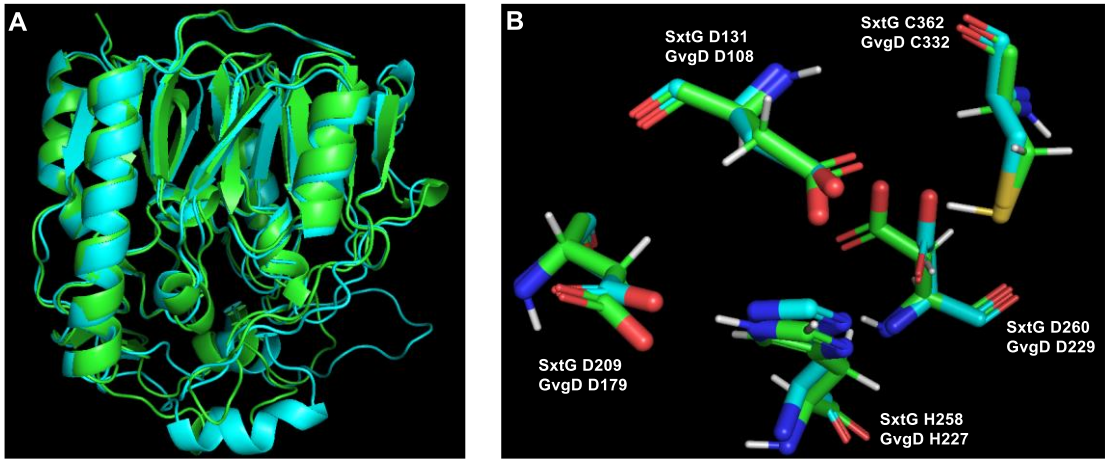


Figure S9. Partial sequence alignment of GvgD and the homologous proteins. The proposed catalytic sites are indicated by red triangles.

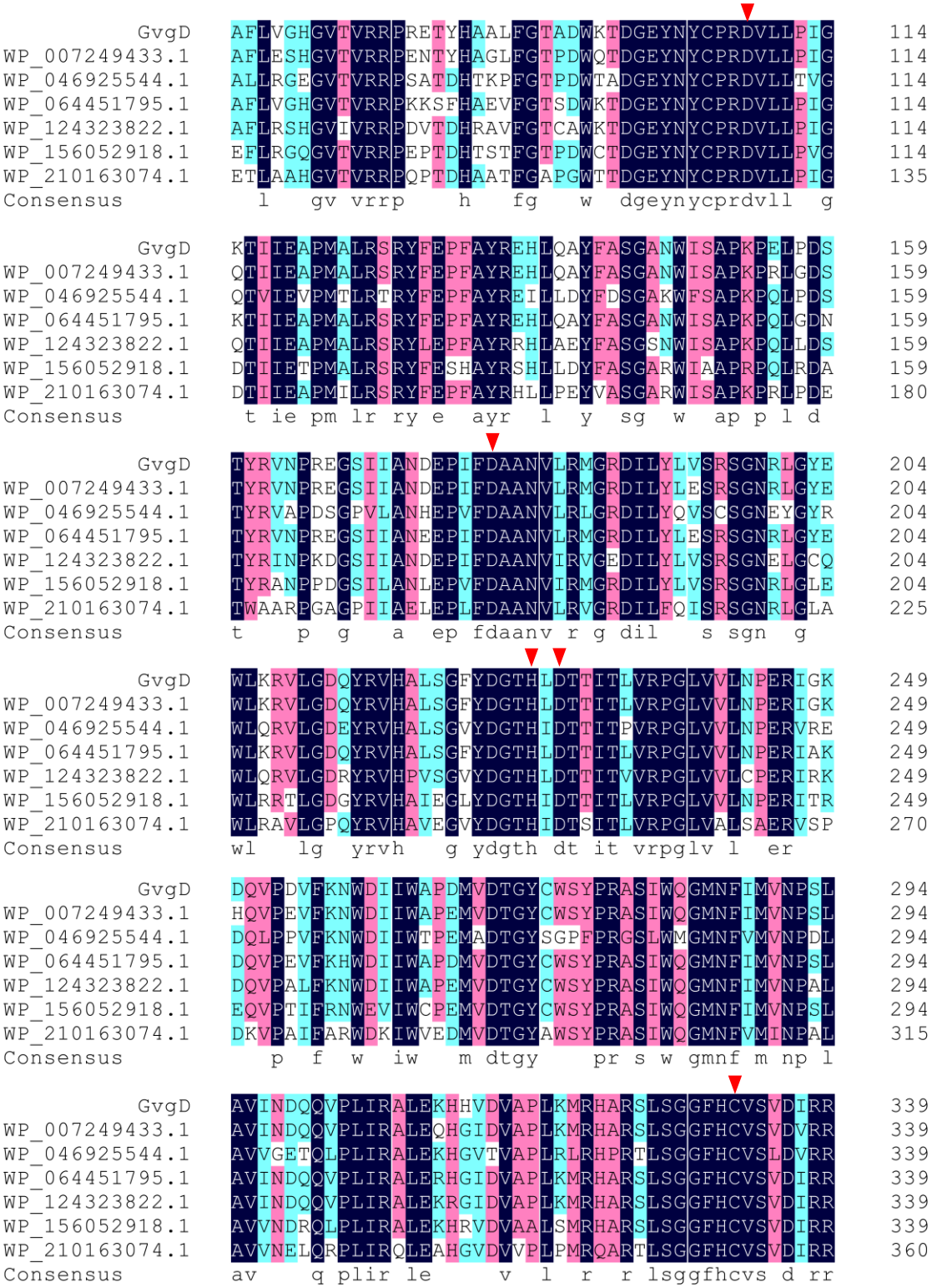


Figure S10. Evaluation of the putative catalytic sites (**A**) and the proposed catalytic mechanism of GvgD (**B**). The peaks with $[M+H]^+$=410 in **II-VII** are isotopic peaks of DNS-Arg ($[M+H]^+$=408) and are extraneous to DNS-**3**.

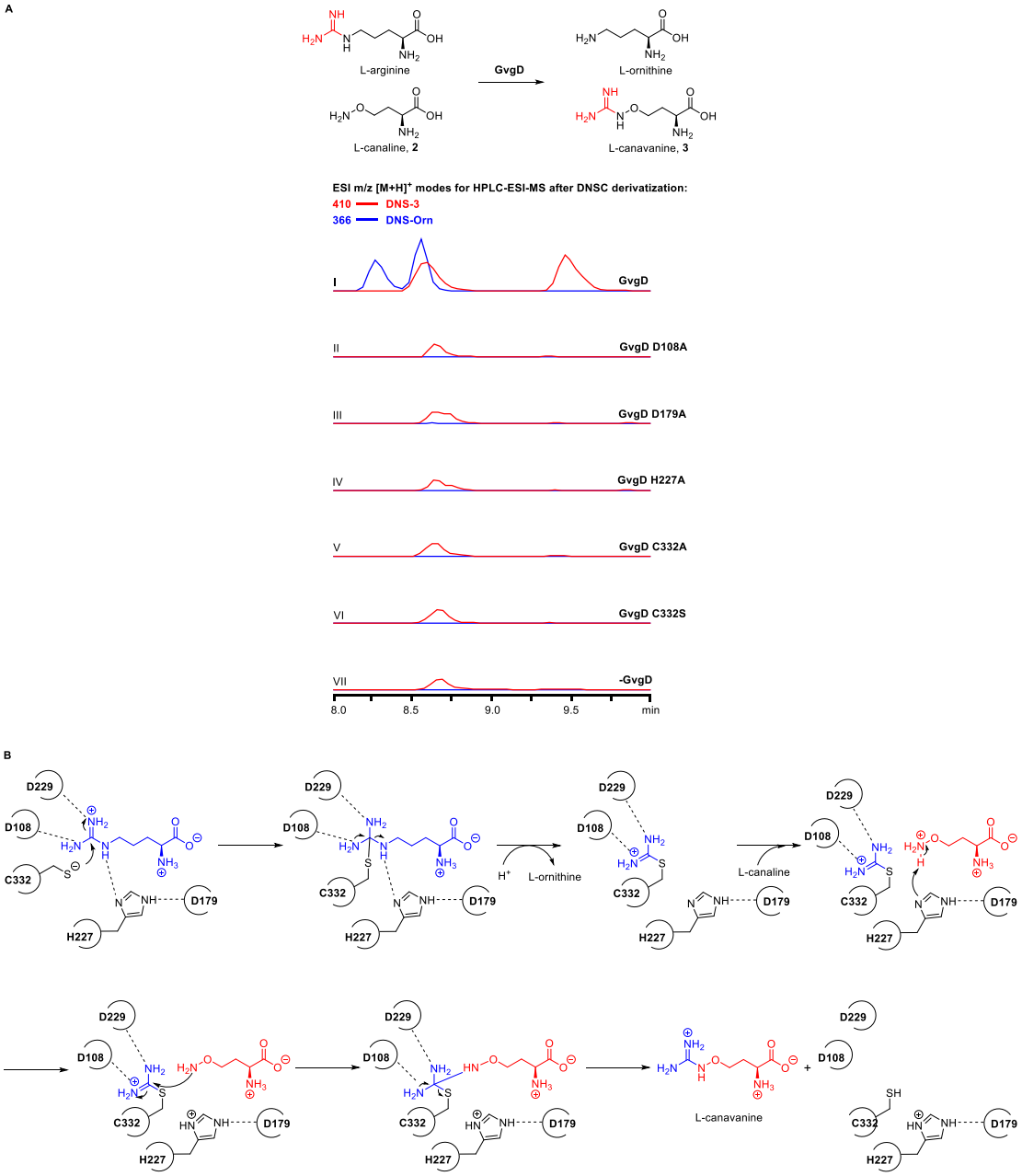


Figure S11. Denaturing SDS-PAGE analysis (conc. 10%) of GvgH, GvgH mutants, $MetY_{WH6}$ and $MetZ_{WH6}$. Lane 1, Protein standard; Lane 2, GvgH (49.04 kDa); Lane 3, Protein standard; Lane 4, GvgH K268A; Lane 5, Protein standard; Lane 6, GvgH K268R; Lane 7, Protein standard; Lane 8, $MetY_{WH6}$ (47.86 kDa); Lane 9, Protein standard; Lane 10, $MetZ_{WH6}$ (45.41 kDa).

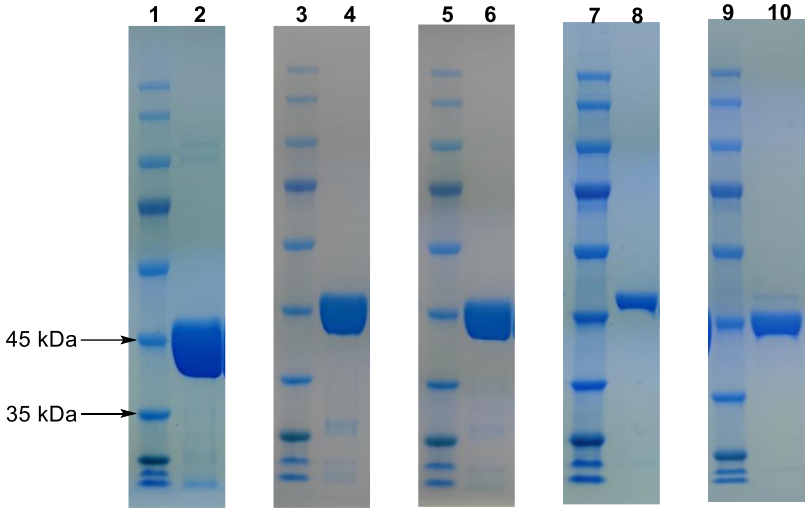


Figure S12. Visible spectra of GvgH, $MetY_{WH6}$, $MetZ_{WH6}$, and GvgH mutants.

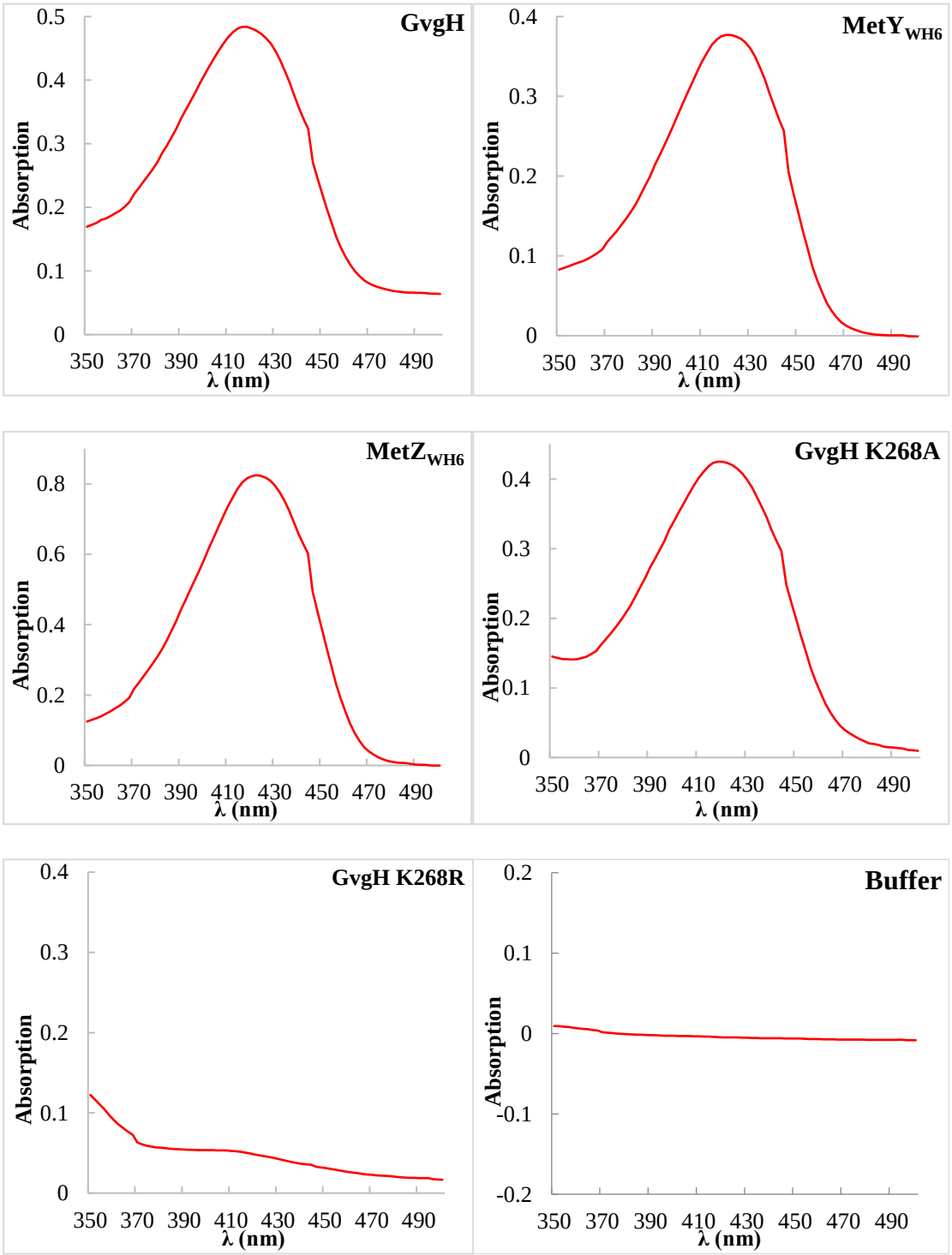


Figure S13. HPLC-ESI-HRMS of DNS-**6**, DNS-**7** and MS/MS fragmentation analysis of DNS-**7**, respectively.

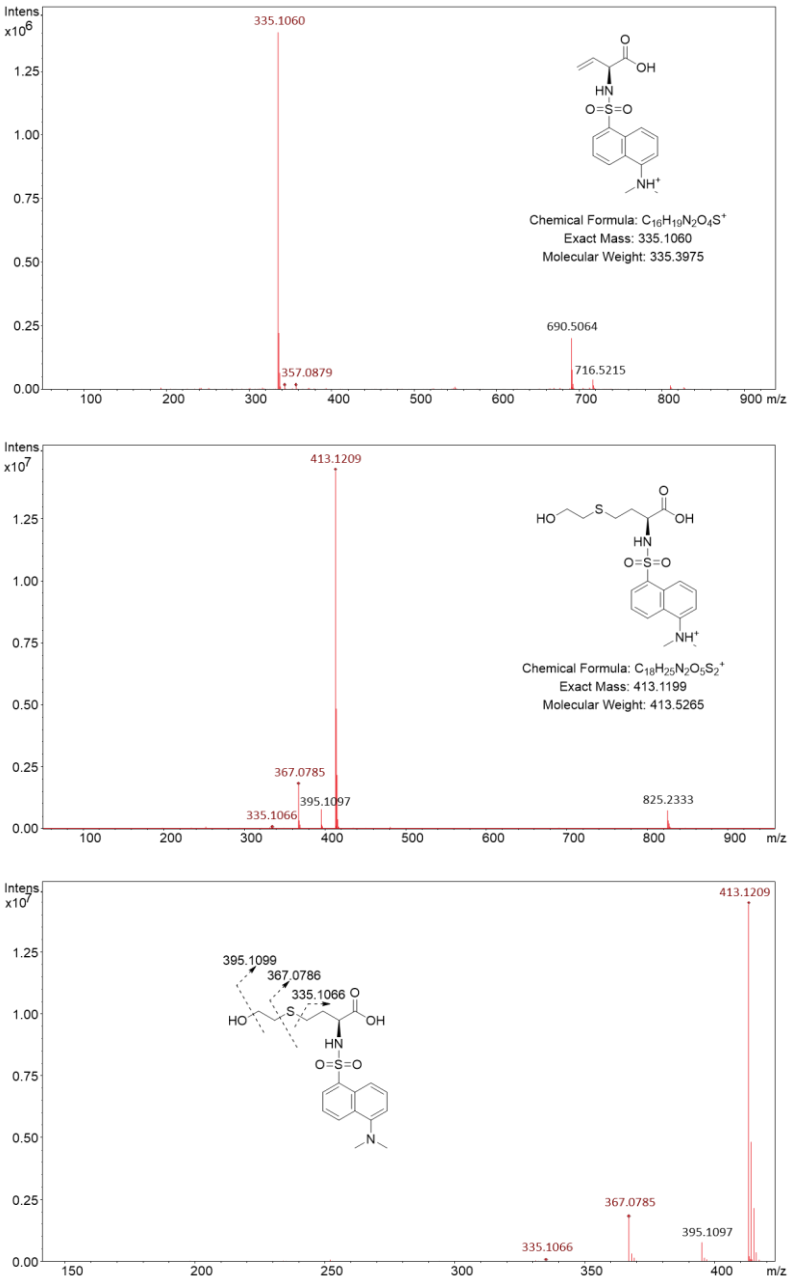


Figure S14. Characterization of the γ-elimination activity of GvgH with L-homoserine (HoSer) and L-homocysteine (HoCys).

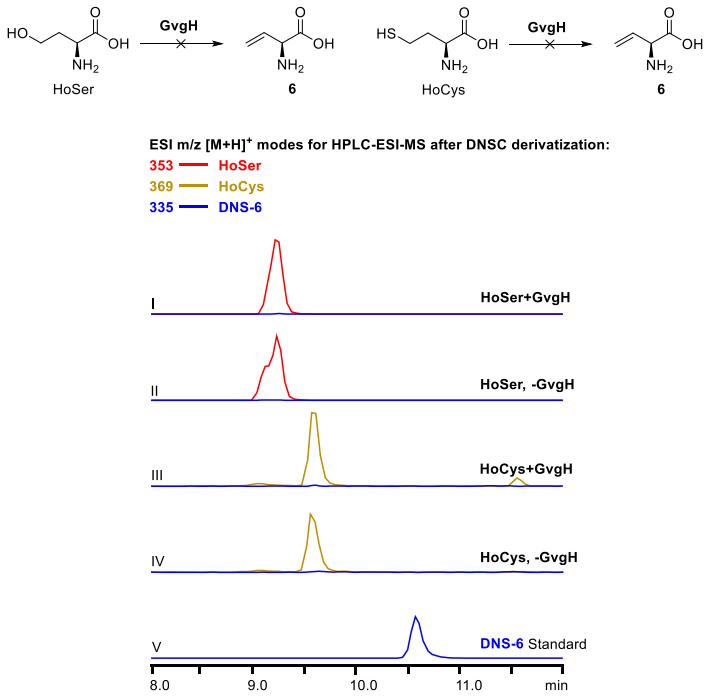
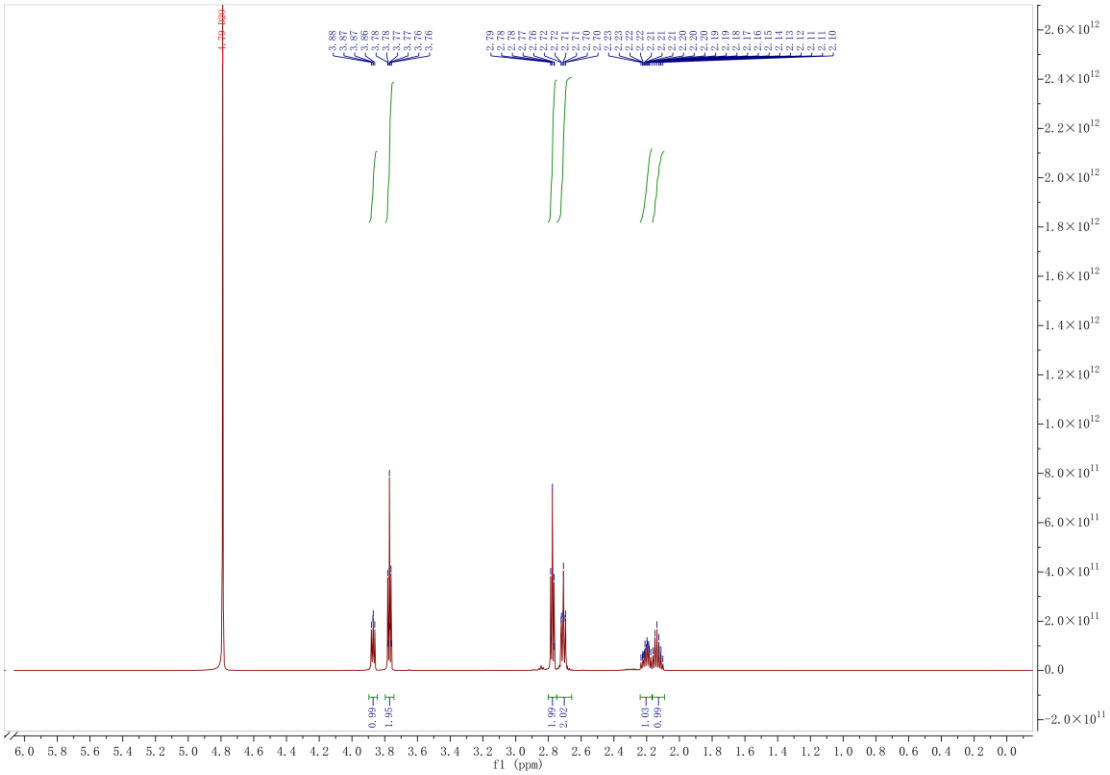
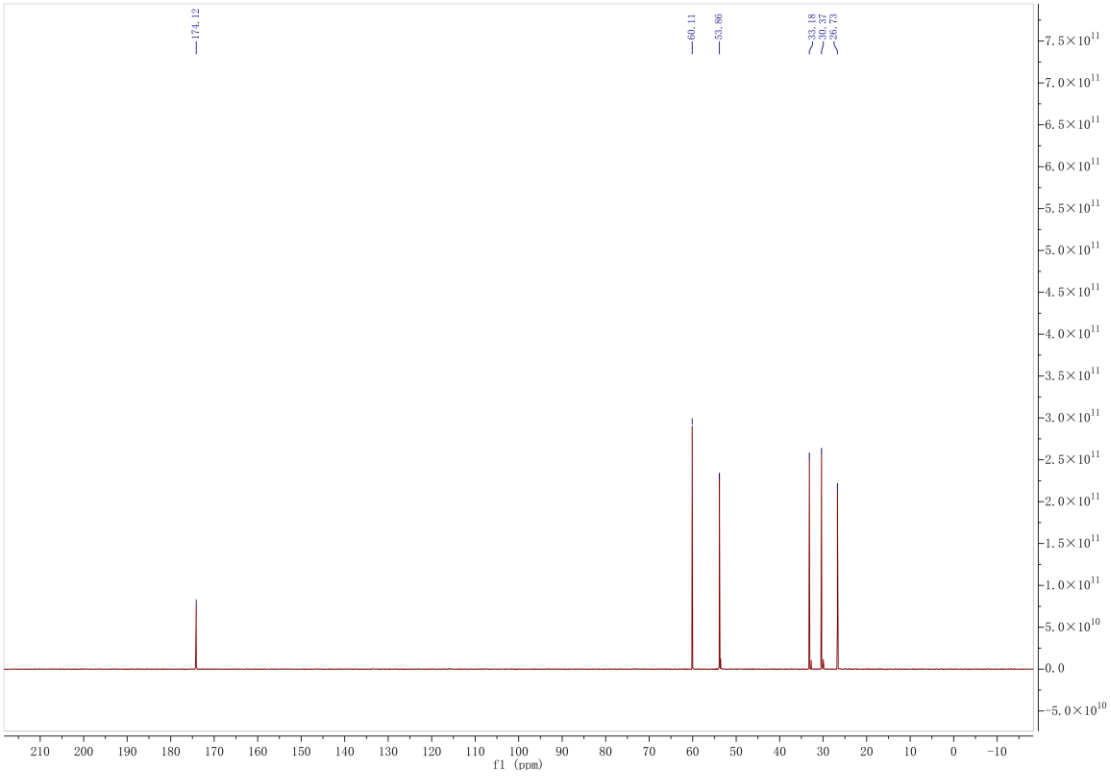


Figure S15. NMR and HPLC-ESI-HRMS spectra of chemically synthesized **7**.

A: ^{1}H-NMR spectrum (600 MHz, D_2O).



B: ^{13}C-NMR spectrum (150 MHz, D_2O)



C: HPLC-ESI-HRMS, $[C_6H_{14}NO_3S]^+$, calcd. 180.0689, found 180.0684.

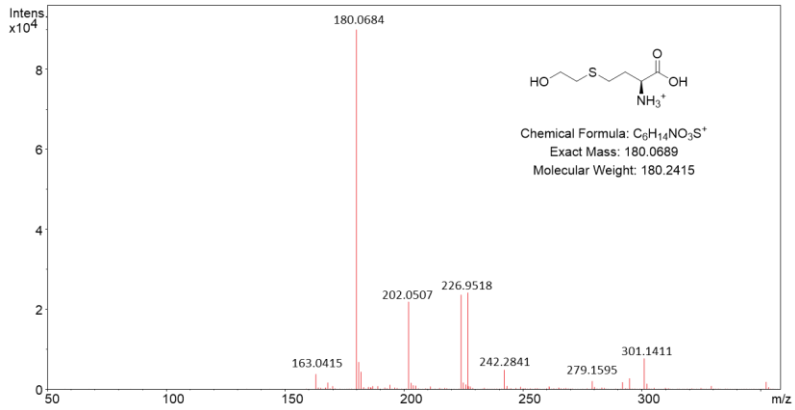


Figure S16. Characterization of the addition activity of GvgH, $MetY_{WH6}$ and $MetZ_{WH6}$ with exogenous **6** and 2-ME.

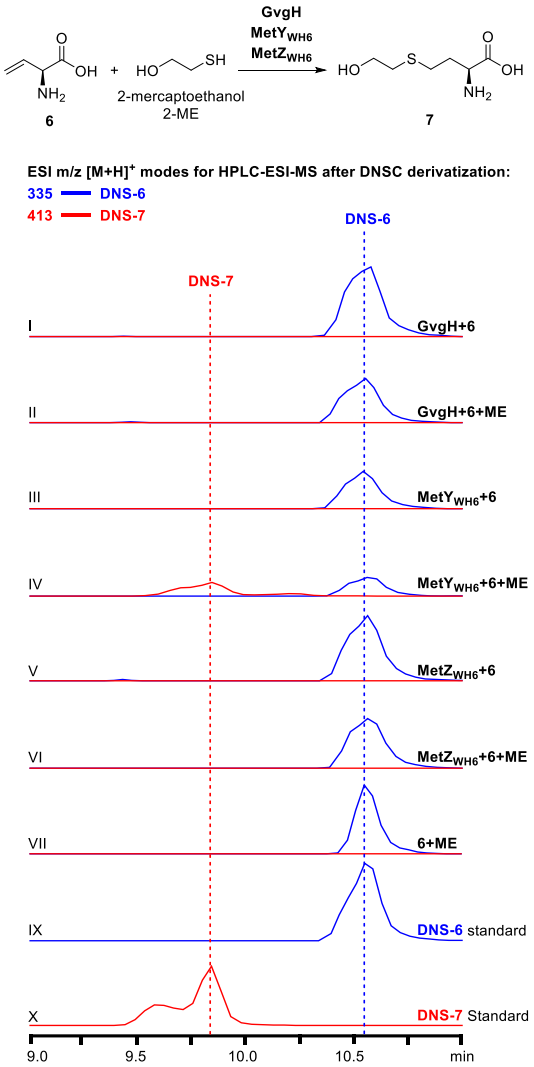


Figure S17. Characterization of the addition activity of GvgH with (**A**) hydroxylamine (**8**), (**B**) hydroxyurea (**9**) or (**C**) hydroxyguanidine (**10**). The peak of $[M+H]^+$=410 in figure **C** is extraneous to DNS-**3** as this peak was also observed in negative controls omitting GvgH and the retention time was distinct from DNS-**3** standard.

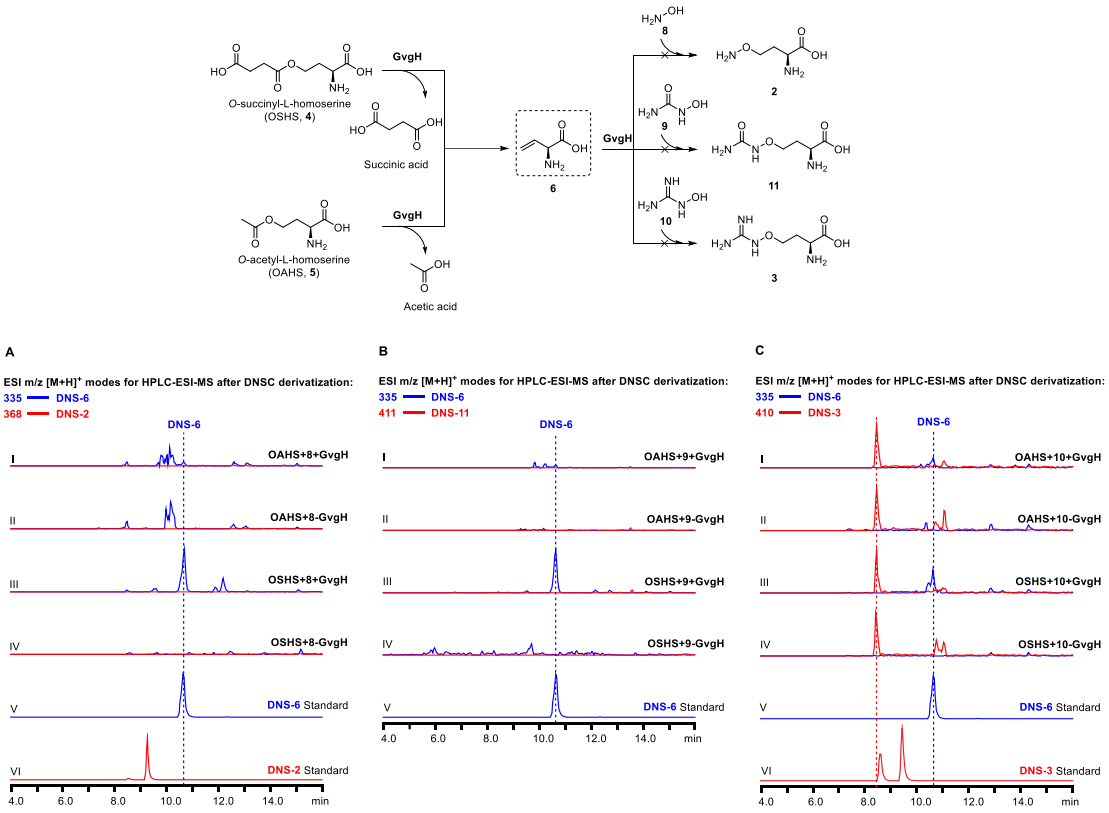


Figure S18. HPLC-ESI-HRMS analysis of PLP releasement from denatured GvgH, K268A and K268R by either boiling or adding 2% (v/v) formic acid.

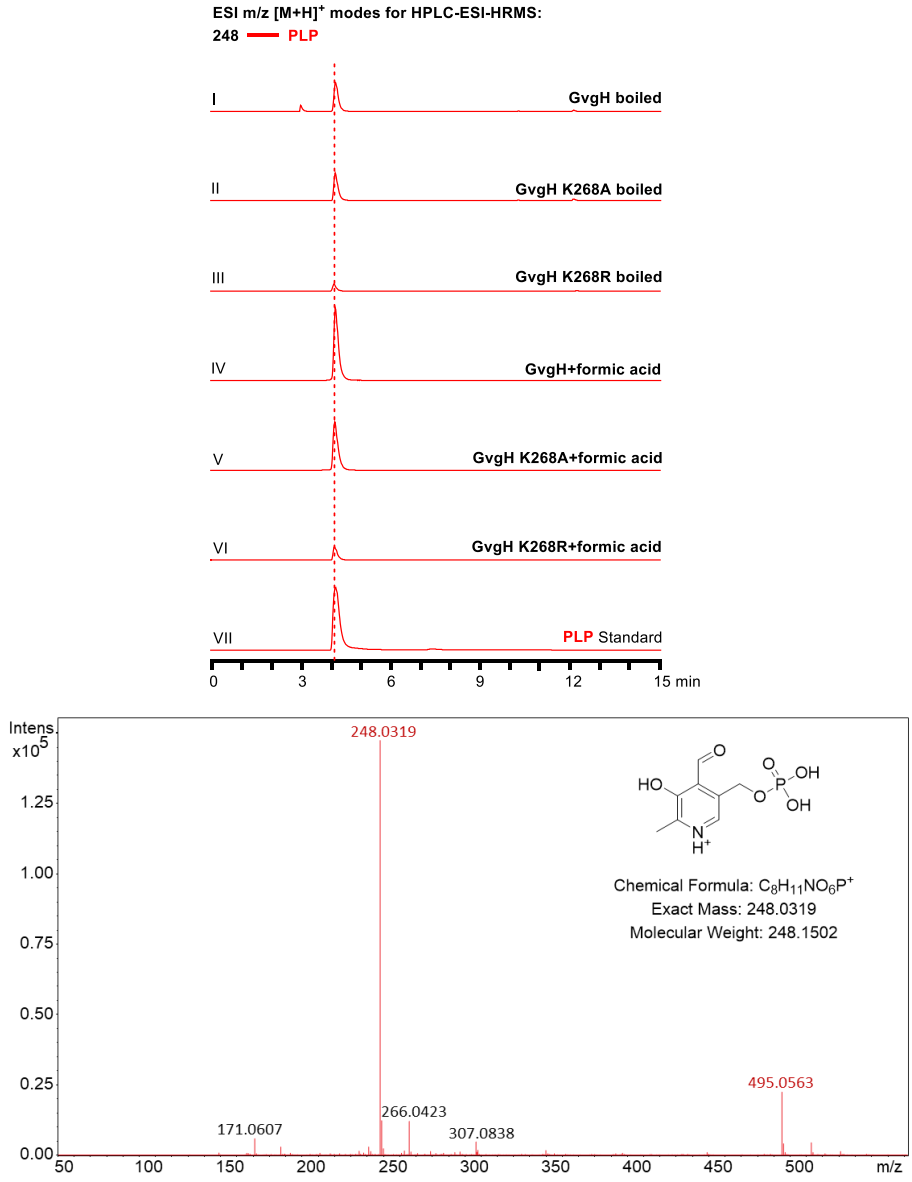


Figure S19. Evaluation of the putative catalytic K268 in GvgH.

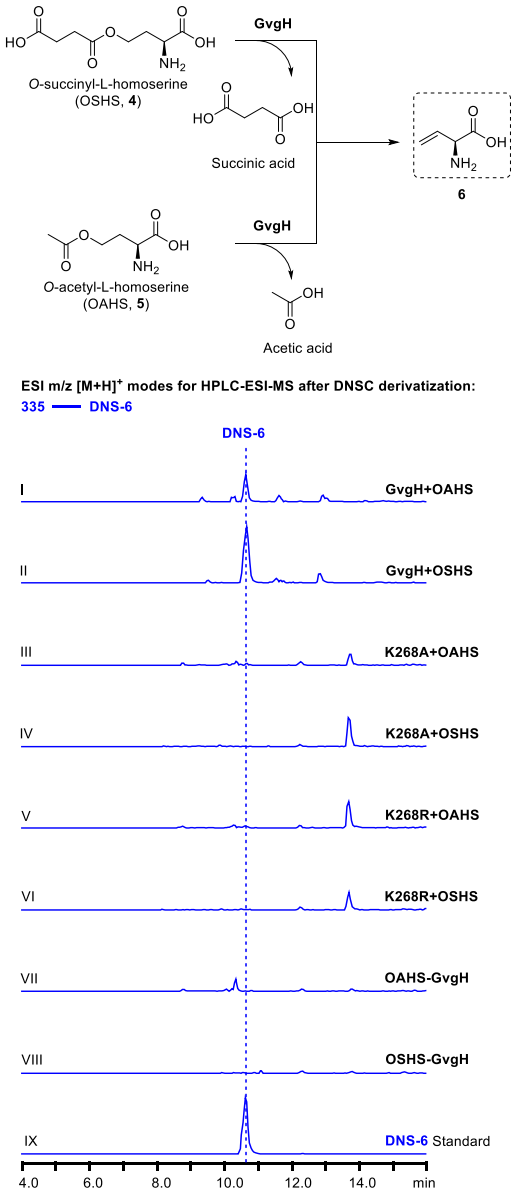


Figure 20. Reactions catalyzed by known biosynthetic enzymes belonging to HDO superfamily.[16-24]

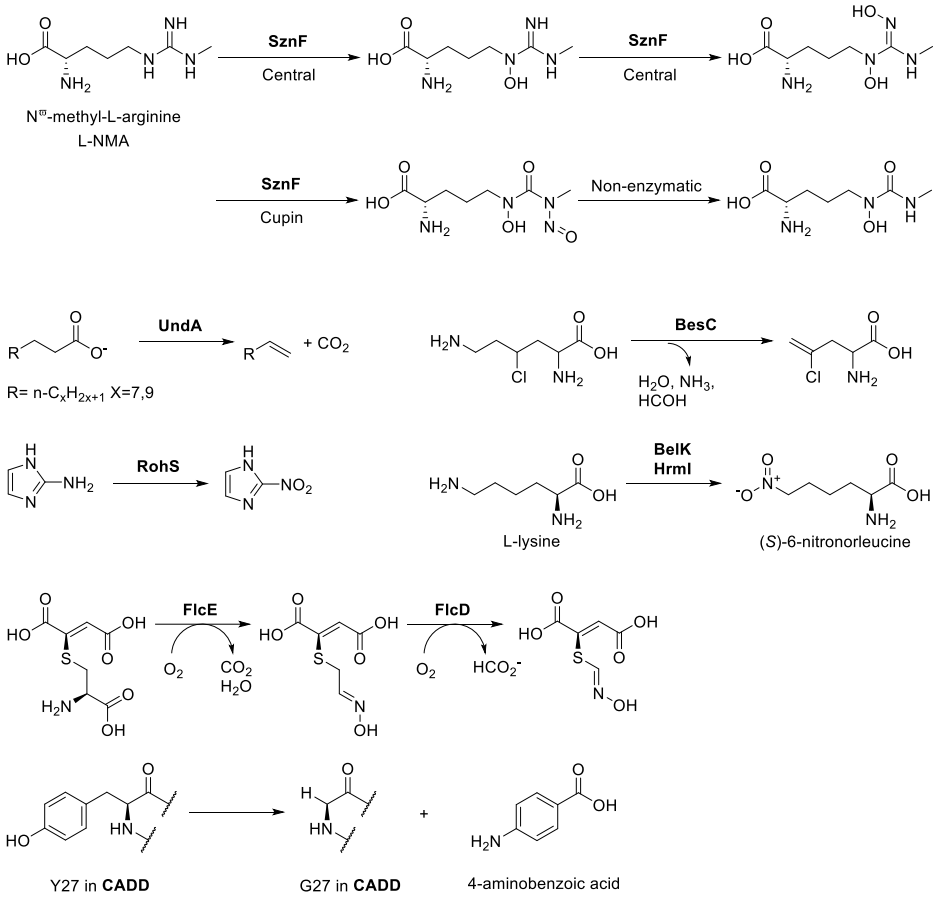


Figure 21. Plausible functions of GvgA, GvgC and GvgF.

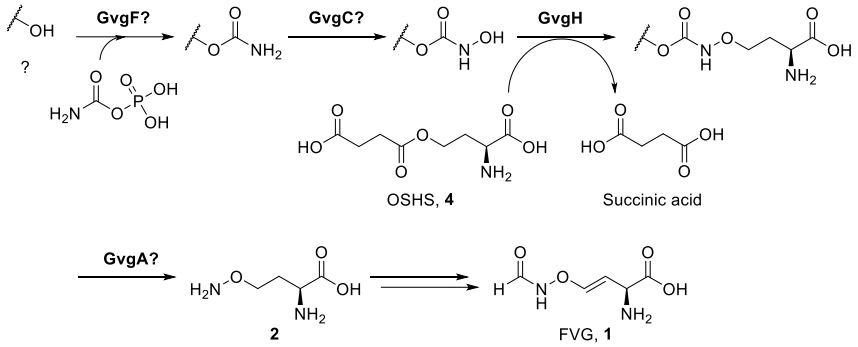


Figure 22. Chemical structures of crocagins, saxitoxin and the proposed functions of CgnI, SxtI and SxtL.[25-28]

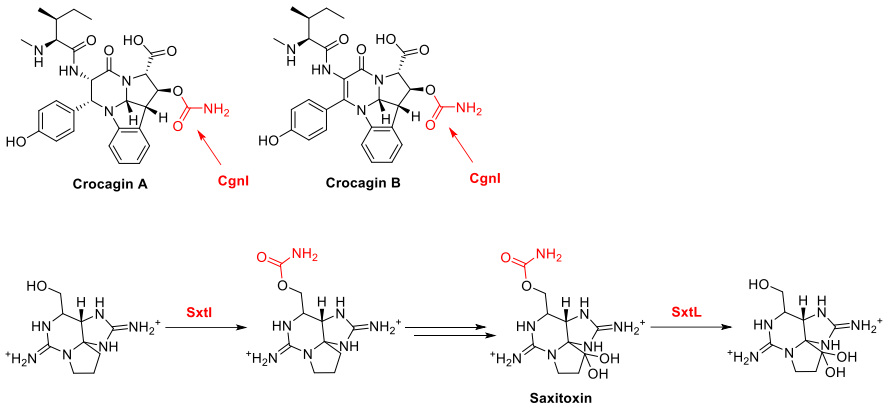
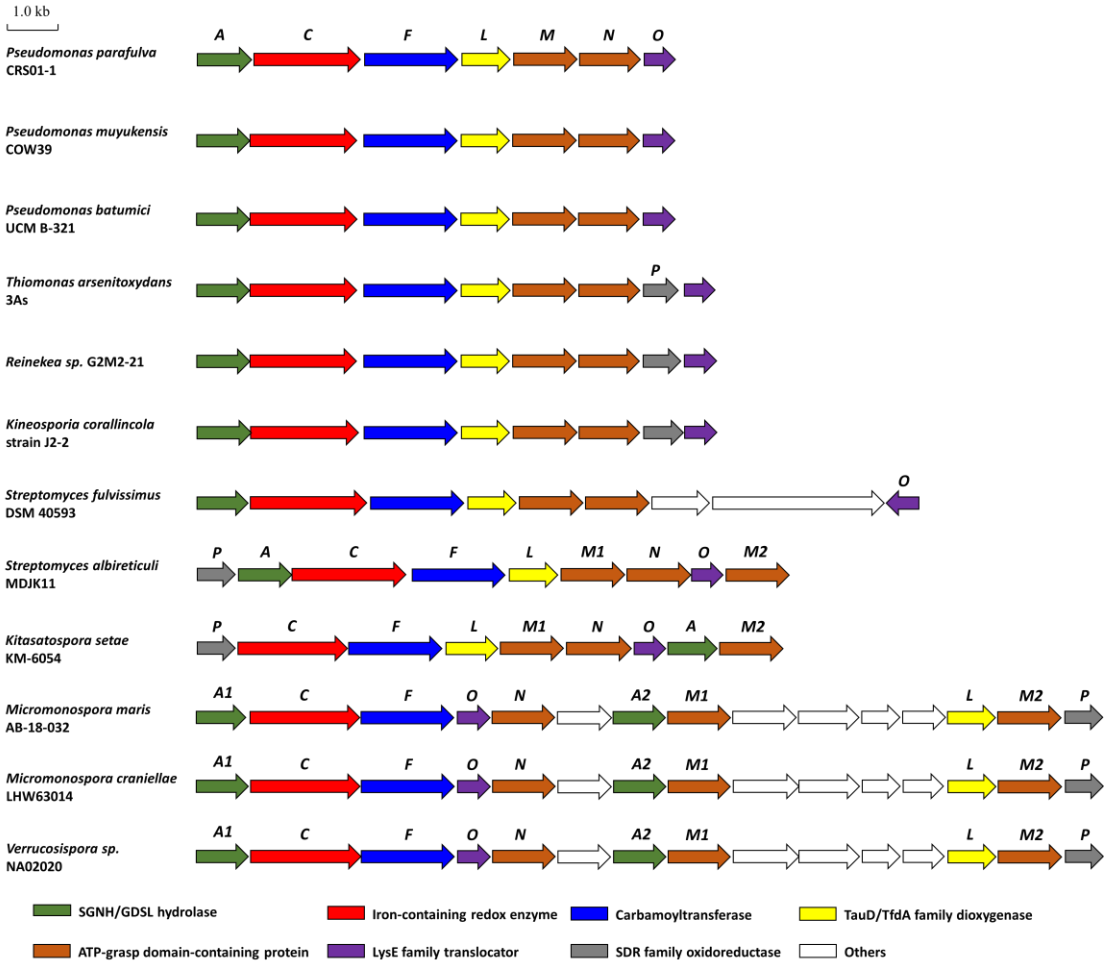
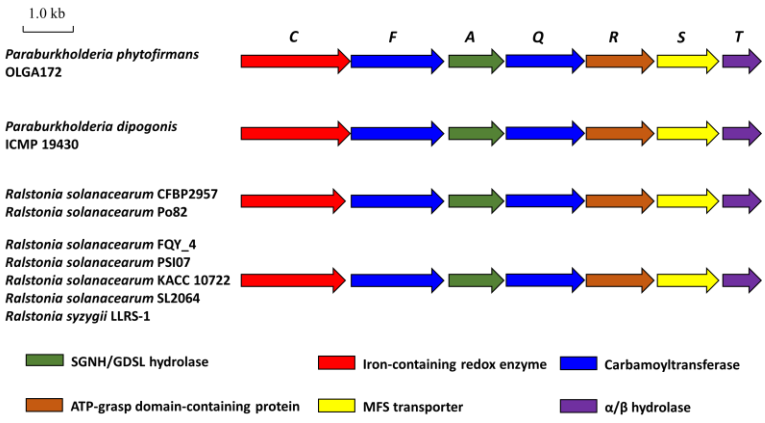


Figure 23. Other BGCs that including homologous genes of *gvgACF*.

A:



B:



Reference:

1. D. B. Berkowitz, B. D. Charette, K. R. Karukurichi and J. M. McFadden, α-Vinylic amino acids: occurrence, asymmetric synthesis, and biochemical mechanisms, *Tetrahedron: Asymmetry*, 2006, **17**, 869-882.
2. J. B. Patteson, Z. D. Dunn and B. Li, In vitro biosynthesis of the nonproteinogenic amino acid methoxyvinylglycine, *Angew. Chem. Int. Ed.*, 2018, **57**, 6780-6785.
3. J. B. Patteson, C. M. Fortinez, A. T. Putz, J. Rodriguez-Rivas, L. H. Bryant, K. Adhikari, M. Weigt, T. M. Schmeing and B. Li, Structure and function of a dehydrating condensation domain in nonribosomal peptide biosynthesis, *J. Am. Chem. Soc.*, 2022, **144**, 14057-14070.
4. T. Yasuta, S. Okazaki, H. Mitsui, K.-I. Yuhashi, H. Ezura and K. Minamisawa, DNA sequence and mutational analysis of rhizobitoxine biosynthesis genes in *Bradyrhizobium elkanii*, *Appl. Environ. Microbiol.*, 2001, **67**, 4999-5009.
5. S. Okazaki, M. Sugawara and K. Minamisawa, *Bradyrhizobium elkanii rtxC* gene is required for expression of symbiotic phenotypes in the final step of rhizobitoxine biosynthesis, *Appl. Environ. Microbiol.*, 2004, **70**, 535-541.
6. M. Sugawara, R. Haramaki, S. Nonaka, H. Ezura, S. Okazaki, S. Eda, H. Mitsui and K. Minamisawa, Rhizobitoxine production in *Agrobacterium tumefaciens* C58 by *Bradyrhizobium elkanii rtxACDEFG* genes, *FEMS Microbiol. Lett.*, 2007, **269**, 29-35.
7. T. Kumagai, Y. Koyama, K. Oda, M. Noda, Y. Matoba and M. Sugiyama, Molecular cloning and heterologous expression of a biosynthetic gene cluster for the antitubercular agent D-cycloserine produced by *Streptomyces lavendulae*, *Antimicrob. Agents Chemother.*, 2010, **54**, 1132-1139.
8. T. Kumagai, K. Takagi, Y. Koyama, Y. Matoba, K. Oda, M. Noda and M. Sugiyama, Heme protein and hydroxyarginase necessary for biosynthesis of D-cycloserine, *Antimicrob. Agents Chemother.*, 2012, **56**, 3682-3689.
9. D. Dietrich, M. J. van Belkum and J. C. Vederas, Characterization of DcsC, a PLP-independent racemase involved in the biosynthesis of D-cycloserine, *Org. Biomol. Chem.*, 2012, **10**, 2248-2254.
10. N. Uda, Y. Matoba, T. Kumagai, K. Oda, M. Noda and M. Sugiyama, Establishment of an *in vitro* D-cycloserine-synthesizing system by using *O*-ureido-L-serine synthase and D-

cycloserine synthetase found in the biosynthetic pathway, *Antimicrob. Agents Chemother.*, 2013, **57**, 2603-2612.

11. Y. Matoba, N. Uda, M. Kudo and M. Sugiyama, Cyclization mechanism catalyzed by an ATP-grasp enzyme essential for D-cycloserine biosynthesis, *FEBS J.*, 2020, **287**, 2763-2778.
12. A. Halgren, M. Maselko, M. Azevedo, D. Mills, D. Armstrong and G. Banowetz, Genetics of germination-arrest factor (GAF) production by *Pseudomonas fluorescens* WH6: identification of a gene cluster essential for GAF biosynthesis, *Microbiology*, 2013, **159**, 36-45.
13. E. W. Davis, R. A. Okrent, V. A. Manning and K. M. Trippe, Unexpected distribution of the 4-formylaminooxyvinylglycine (FVG) biosynthetic pathway in *Pseudomonas* and beyond, *PloS one*, 2021, **16**, e0247348.
14. R. A. Okrent, K. M. Trippe, M. Maselko and V. Manning, Functional analysis of a biosynthetic cluster essential for production of 4-formylaminooxyvinylglycine, a germination-arrest factor from *Pseudomonas fluorescens* WH6, *Microbiology*, 2017, **163**, 207-217.
15. A. L. Lukowski, L. Mallik, M. E. Hinze, B. M. Carlson, D. C. Ellinwood, J. B. Pyser, M. Koutmos and A. R. H. Narayan, Substrate promiscuity of a paralytic shellfish toxin amidinotransferase, *ACS Chem. Biol.*, 2020, **15**, 626-631.
16. M. J. McBride, S. R. Pope, K. Hu, C. D. Okafor, E. P. Balskus, J. M. Bollinger and A. K. Boal, Structure and assembly of the diiron cofactor in the heme-oxygenase–like domain of the *N*-nitrosourea–producing enzyme SznF, *Proc. Natl. Acad. Sci. U.S.A.*, 2021, **118**, e2015931118.
17. Z. Rui, X. Li, X. Zhu, J. Liu, B. Domigan, I. Barr, J. H. D. Cate and W. Zhang, Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 18237-18242.
18. J. B. Hedges and K. S. Ryan, In vitro reconstitution of the biosynthetic pathway to the nitroimidazole antibiotic azomycin, *Angew. Chem. Int. Ed.*, 2019, **58**, 11647-11651.
19. J. A. Marchand, M. E. Neugebauer, M. C. Ing, C. I. Lin, J. G. Pelton and M. C. Y. Chang, Discovery of a pathway for terminal-alkyne amino acid biosynthesis, *Nature*, 2019, **567**, 420-424.
20. J. B. Patteson, A. T. Putz, L. Tao, W. C. Simke, L. H. Bryant, R. D. Britt and B. Li, Biosynthesis of fluopsin C, a copper-containing antibiotic from *Pseudomonas aeruginosa*, *Science*, 2021, **374**, 1005-1009.

21. S. Shimo, R. Ushimaru, A. Engelbrecht, M. Harada, K. Miyamoto, A. Kulik, M. Uchiyama, L. Kaysser and I. Abe, Stereodivergent nitrocyclopropane formation during biosynthesis of belactosins and hormaomycins, *J. Am. Chem. Soc.*, 2021, **143**, 18413-18418.

22. X. Li, R. Shimaya, T. Dairi, W.-c. Chang and Y. Ogasawara, Identification of cyclopropane formation in the biosyntheses of hormaomycins and belactosins: sequential nitration and cyclopropanation by metalloenzymes, *Angew. Chem. Int. Ed.*, 2022, **61**, e202113189.

23. L. Pang, W. Niu, Y. Duan, L. Huo, A. Li, J. Wu, Y. Zhang, X. Bian and G. Zhong, *In vitro* characterization of a nitro-forming oxygenase involved in 3-(*trans*-2'-aminocyclopropyl)alanine biosynthesis, *Engin. Microbiol.*, 2022, **2**, 100007.

24. O. M. Manley, H. N. Phan, A. K. Stewart, D. A. Mosley, S. Xue, L. Cha, H. Bai, V. C. Lightfoot, P. A. Rucker, L. Collins, T. I. Williams, W.-C. Chang, Y. Guo and T. M. Makris, Self-sacrificial tyrosine cleavage by an Fe:Mn oxygenase for the biosynthesis of *para*-aminobenzoate in *Chlamydia trachomatis*, *Proc. Natl. Acad. Sci. U.S.A.*, 2022, **119**, e2210908119.

25. K. Viehrig, F. Surup, C. Volz, J. Herrmann, A. Abou Fayad, S. Adam, J. Köhnke, D. Trauner and R. Müller, Structure and biosynthesis of crocagins: polycyclic posttranslationally modified ribosomal peptides from *Chondromyces crocatus*, *Angew. Chem. Int. Ed.*, 2017, **56**, 7407-7410.

26. S. Adam, D. Zheng, A. Klein, C. Volz, W. Mullen, S. L. Shirran, B. O. Smith, O. V. Kalinina, R. Müller and J. Koehnke, Unusual peptide-binding proteins guide pyrroloindoline alkaloid formation in crocagin biosynthesis, *Nat. Chem.*, 2023, **15**, 560-568.

27. R. Kellmann, T. K. Mihali, Y. J. Jeon, R. Pickford, F. Pomati and B. A. Neilan, Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in Cyanobacteria, *Appl. Environ. Microbiol.*, 2008, **74**, 4044-4053.

28. T. K. Mihali, R. Kellmann and B. A. Neilan, Characterisation of the paralytic shellfish toxin biosynthesis gene clusters in *Anabaena circinalis* AWQC131C and *Aphanizomenon sp.* NH-5, *BMC Biochem.*, 2009, **10**, 8.