Supporting Information for:

Inverse stable isotope probing-metabolomics (InverSIP) identifies an iron acquisition system in a methane-oxidizing bacterial community

Jose Miguel D. Robes[1,2], Tashi C.E. Liebergesell[1,2], Delaney G. Beals[1,2#], Xinhui Yu[1,2], William J. Brazelton[3], & Aaron W. Puri[1,2*]

[1]Department of Chemistry, University of Utah, Salt Lake City, Utah, USA
[2]Henry Eyring Center for Cell and Genome Science, University of Utah, Salt Lake City, Utah, USA
[3]School of Biological Sciences, University of Utah, Salt Lake City, Utah, USA

#Current address:
Delaney G. Beals
Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

*Corresponding author:
Aaron W. Puri
315 S 1400 E Rm 2020
Salt Lake City, UT 84112
USA
(801) 213-1408
a.puri@utah.edu

This PDF file includes:

TABLE OF CONTENTS

SUPPLEMENTARY METHODS

MS/MS analysis of methylocystabactin and linear dehydrated (DHB-Gly-^{L}Ser)$_3$

Methylocystabactin-containing extracts have mass spectral features corresponding to *m/z* 841.2159 at t_R=5.9 minutes and t_R=6.2 minutes (**Figure 3D**). This indicates that one is the cyclic trimer form, methylocystabactin (**1**), and the other is a linear form of the siderophore where one serine is dehydrated to dehydroalanine (**6, see Figure S8 for structure**) as reported for the siderophore turnerbactin (1). To differentiate between the two, a custom Python script (https://github.com/YuxNP/MS2-ratio-plot/tree/main) was used to compare MS/MS peak lists for the two features. The peak lists were first normalized to remove intensity bias, and a mirror plot was then generated to visualize and sort unique fragmentations (**Figure S16**). Fragments that could indicate significant molecular structural changes were then identified using ratiometric analysis: for common fragments, if the ratio between two corresponding fragments was either less than 0.97 or greater than 1.2, the fragment was marked for further study. We also found a unique fragment present in the feature at t_R=5.9 minutes that is absent in the feature at t_R=6.2 minutes (**Figure S17**). Based on the fragment, we hypothesized that the peak at t_R=5.9 minutes is the linear dehydrated (DHB-Gly-^{L}Ser)$_3$ (**6**). We further confirmed that the peak at t_R = 6.2 minutes is methylocystabactin (**1**) by purification and structural elucidation using NMR (**Table S4**).

Identification and Analysis of Putative TonB-Dependent Receptors Encoded in the Genomes of *Methylosinus* sp. strains LW3 and LW4

The genomes of *Methylosinus* sp. strains LW3 (IMG genome ID 2517093011) and LW4 (2516493024) were searched for genes with products containing both pfam00593 (TonB dependent receptor-like, beta-barrel) and pfam07715 (TonB-dependent receptor plug domain). This returned 45 genes for *Methylosinus* sp. strain LW3 and 62 genes for *Methylosinus* sp. strain LW4.

To identify which of these predicted TonB-dependent receptors were most similar to the characterized receptors FepA and CirA from *Escherichia coli* and IroN from *Salmonella enterica* subsp. *enterica* serovar Typhimurium, the amino acid sequences for these receptors were used as BLAST queries against the genomes for *Methylosinus* sp. strains LW3 and LW4. The same gene product in the *Methylosinus* sp. strain LW3 genome was the top hit for all three characterized receptors (IMG gene ID 2517139767). This was also the case for *Methylosinus* sp. strain LW4 (IMG gene ID 2516605591) (see **Table S7**).

To identify the most similar characterized TonB-dependent receptors to the predicted methylocystabactin TonB-dependent receptors, the amino acid sequences for these receptors from *Methylosinus* sp. strains LW3 (IMG gene ID 2517142394) and LW4 (IMG gene ID 2516608097) were used as BLAST queries against the UniProtKB Swiss-Prot database. In both cases, the two best hits were FatA from *Vibrio anguillarum* and FcuA from *Yersinia enterocolitica* (see **Table S7)**.

SUPPLEMENTARY FIGURES

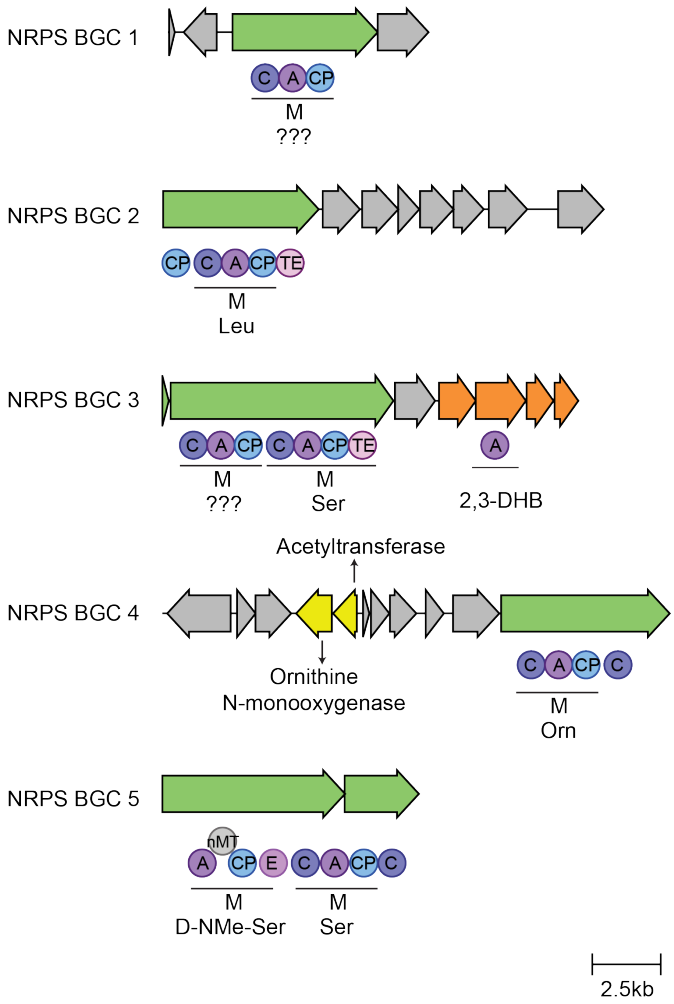


Figure S1. Top 5 transcribed NRPS BGCs identified from BiG-MAP (2) showing the NRPS gene (green) and predicted adenylation domain substrates. C: condensation domain, A: adenylation domain, CP: carrier protein, TE: thioesterase domain, E: epimerization domain, nMT: *N*-methyl transferase domain. Green: non-ribosomal peptide synthetase, orange: DHB operon, yellow: ornithine modifying enzymes.

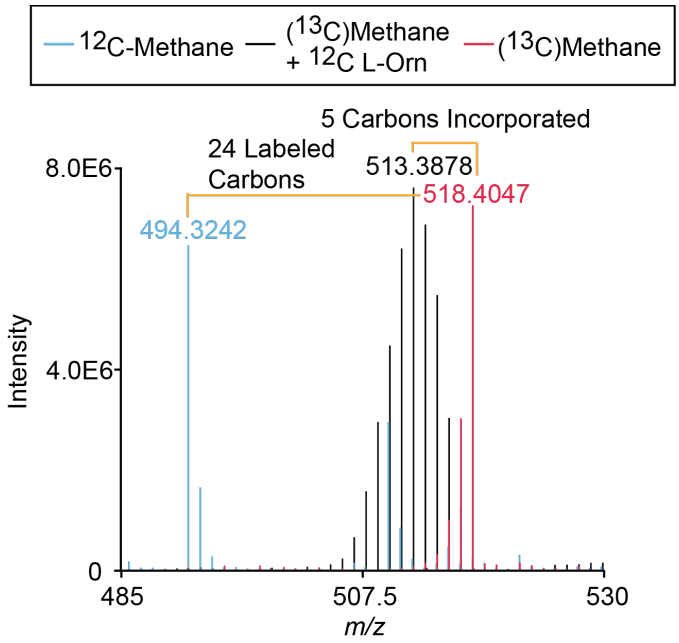


Figure S2. Overlaid mass spectra showing incorporation of ^{12}C-L-ornithine into an actively produced metabolite in the methane-oxidizing bacterial community.

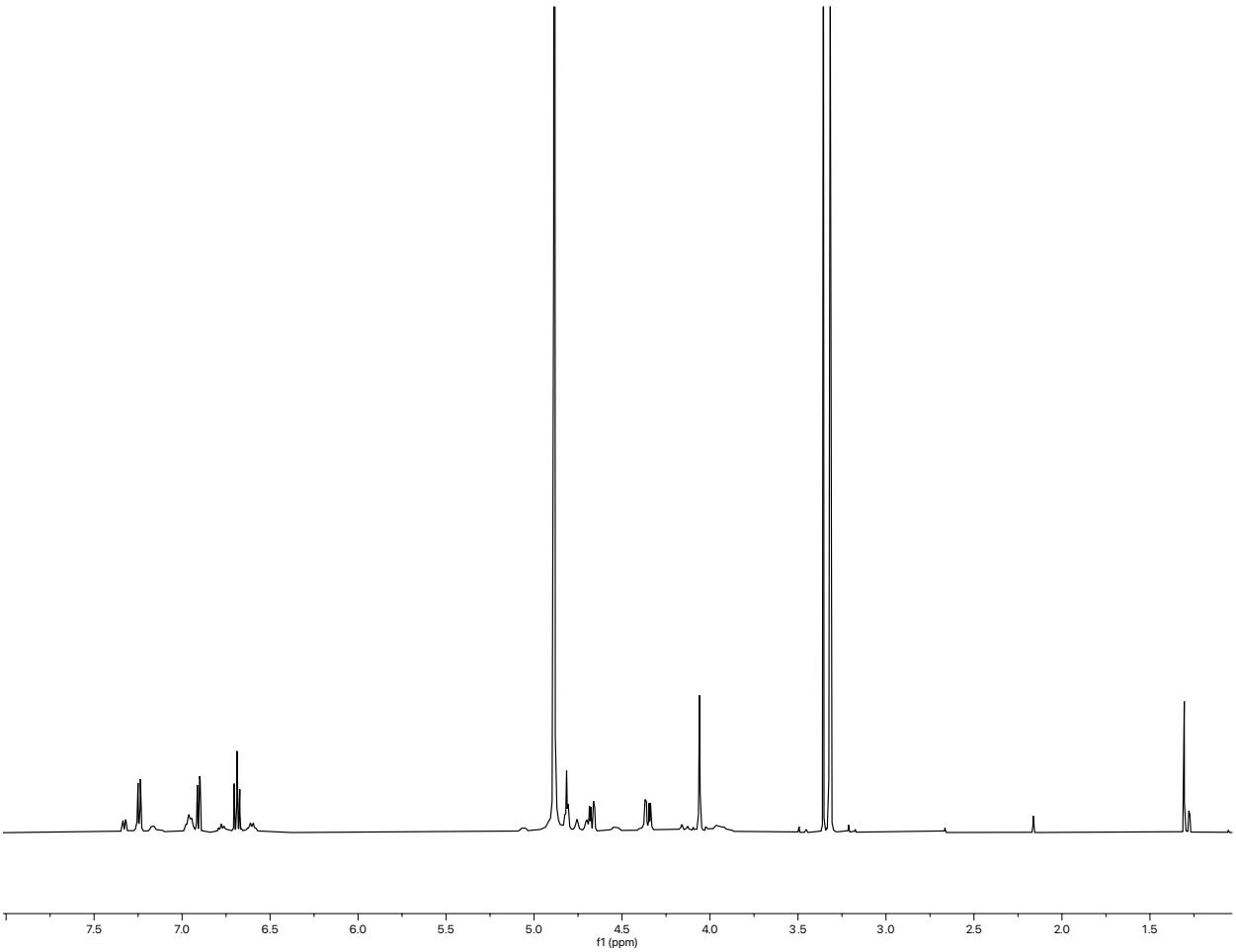


Figure S3. ^{1}H NMR spectrum of methylocystabactin in CD_3OD (500 MHz)**.**

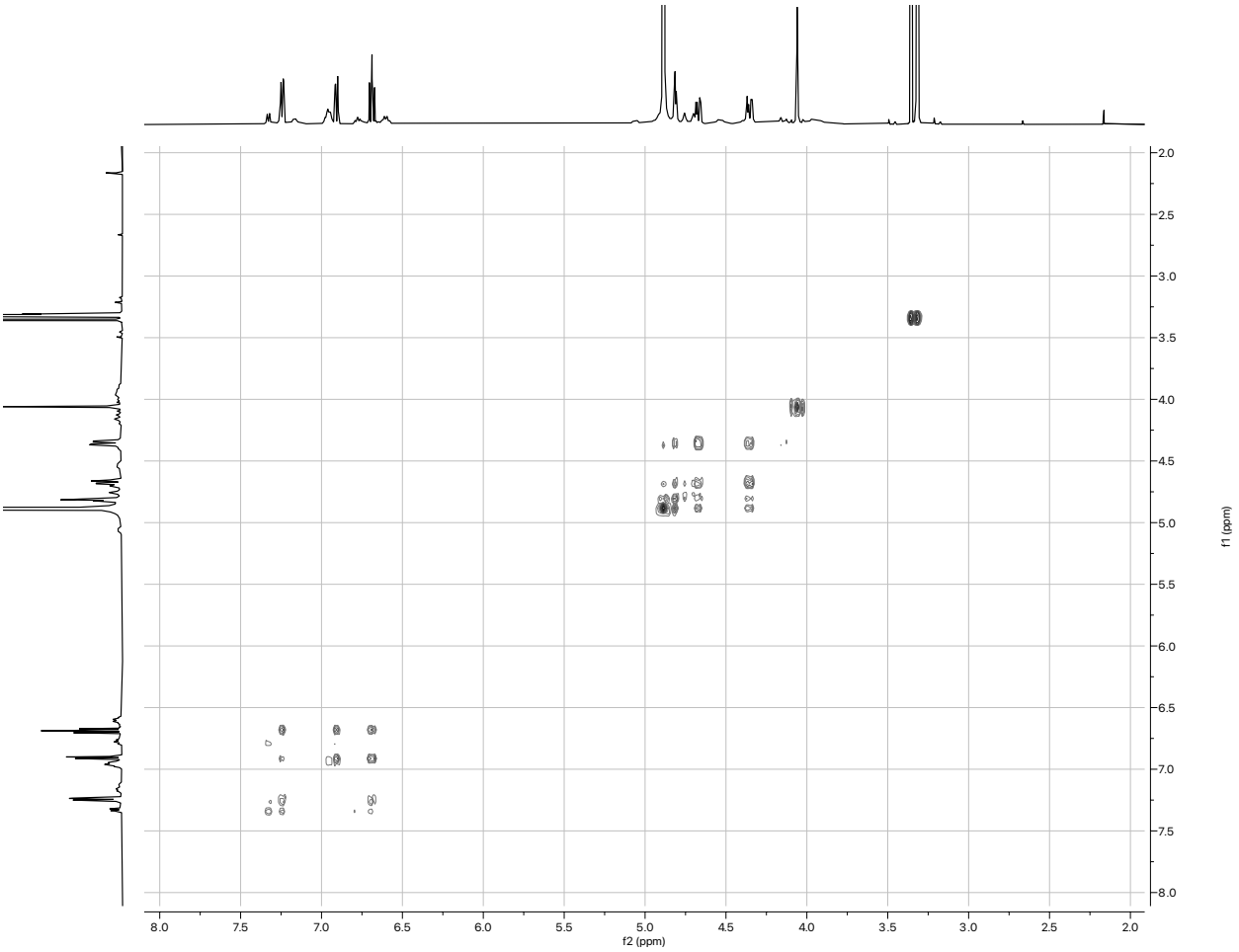


Figure S4. COSY spectrum of methylocystabactin in CD_3OD (500 MHz)**.**

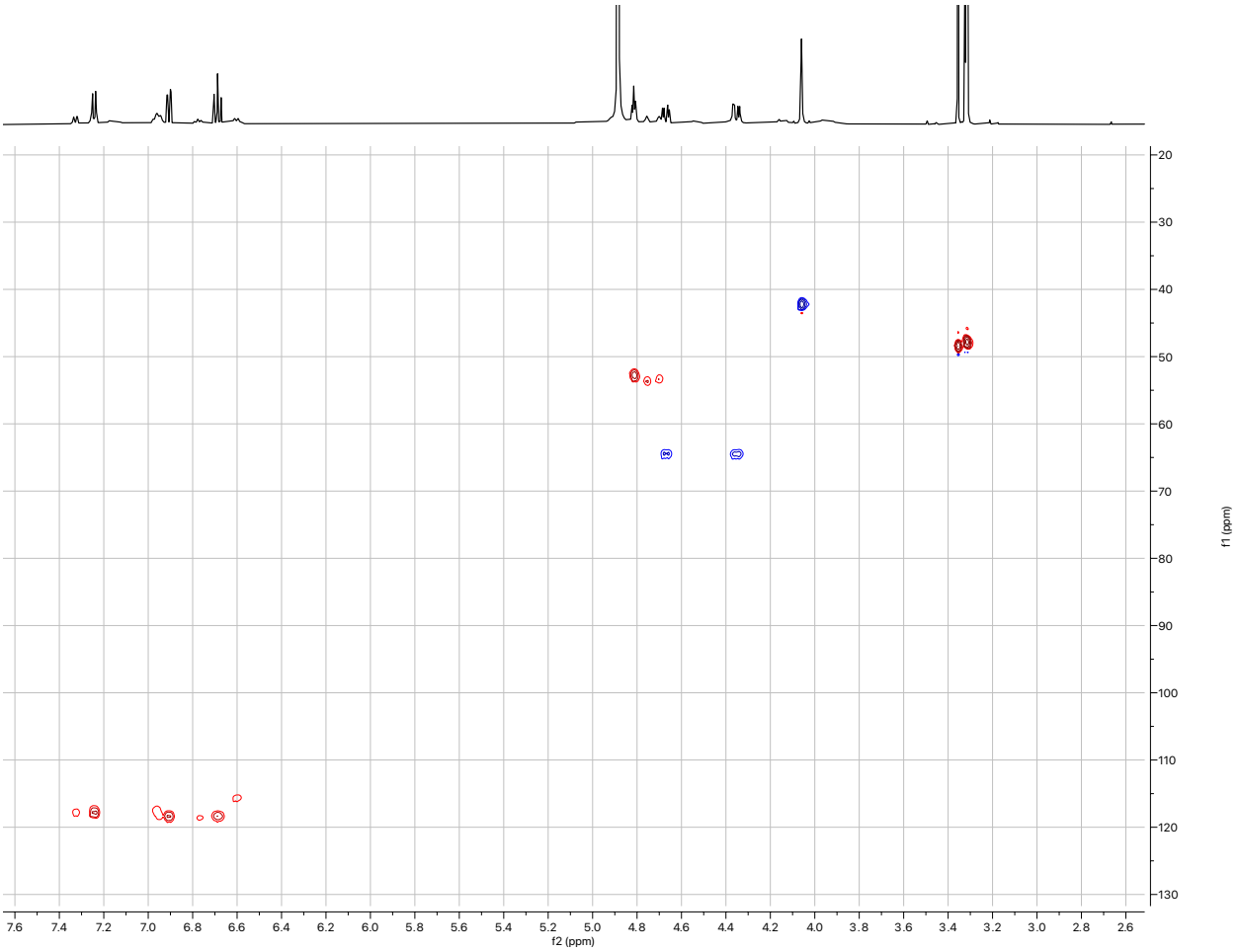


Figure S5. gHSQC spectrum of methylocystabactin in CD_3OD (500 MHz)**.**

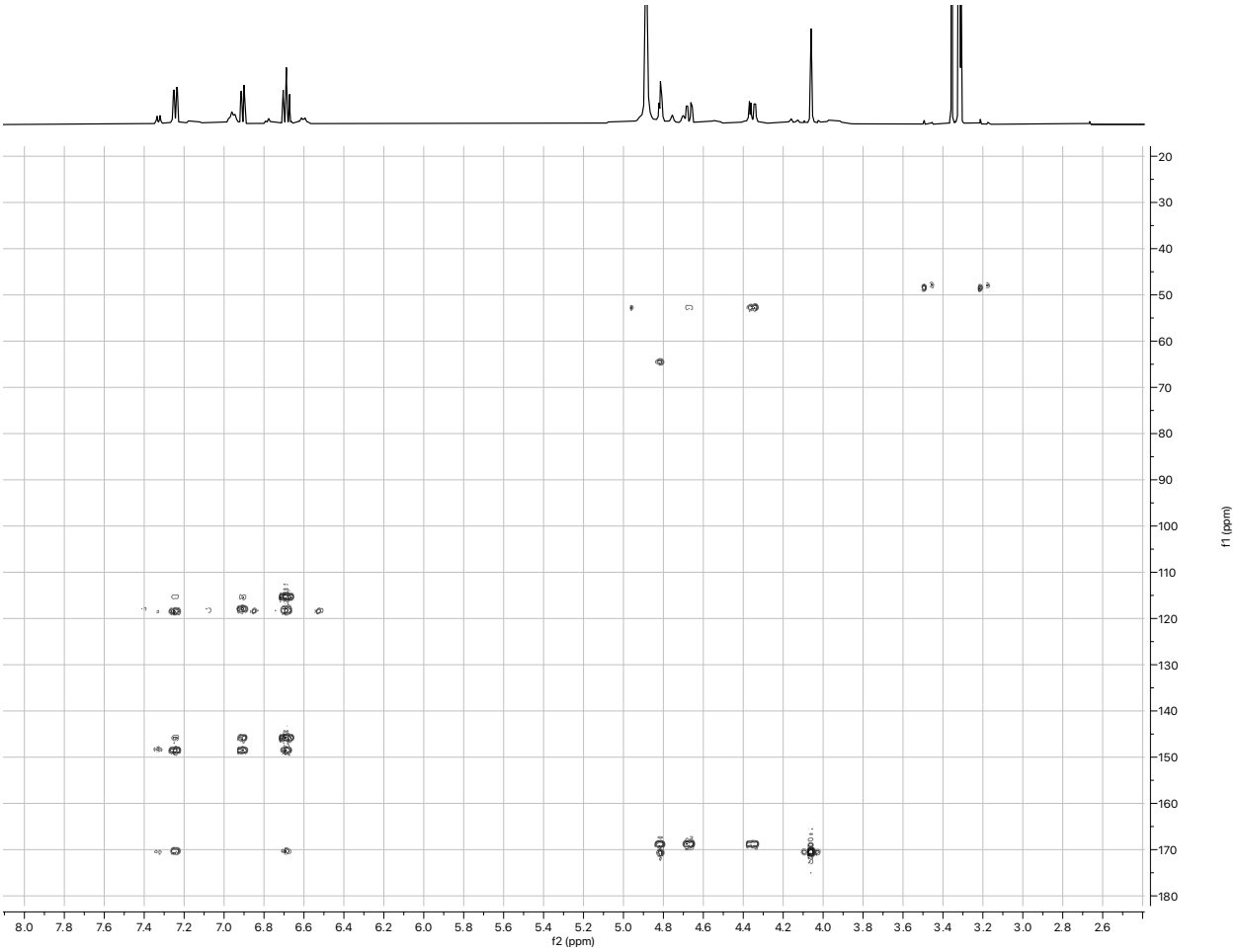


Figure S6. HMBC spectrum of methylocystabactin in CD_3OD (500 MHz).

Chirality	π−π* transition		LMCT transition	
	λmax (nm)	Δϵ ($M^{-1}cm^{-1}$)	λmax (nm)	Δϵ ($M^{-1}cm^{-1}$)
Λ	310, 352	-2.3,+4.4	544	+1.55

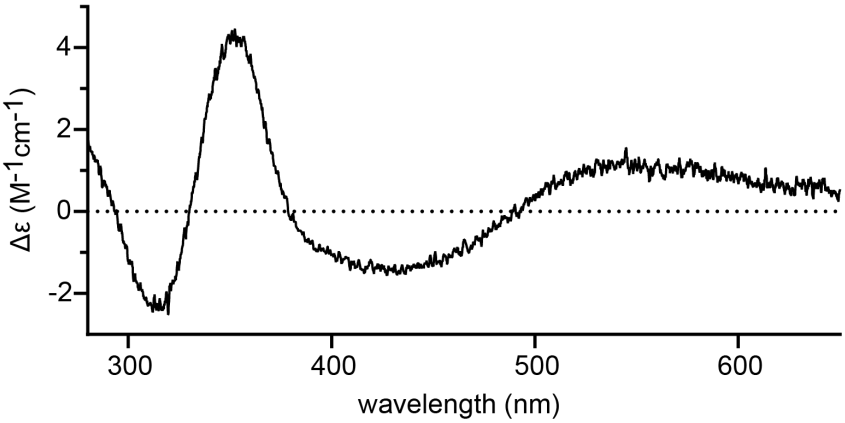


Figure S7. ECD spectra of Fe^{3+}-methylocystabactin. (Fe^{3+}-methylocystabactin = 0.1 mM in 1mM sodium phosphate, pH 7.4). LMCT: ligand-to-metal charge transfer.

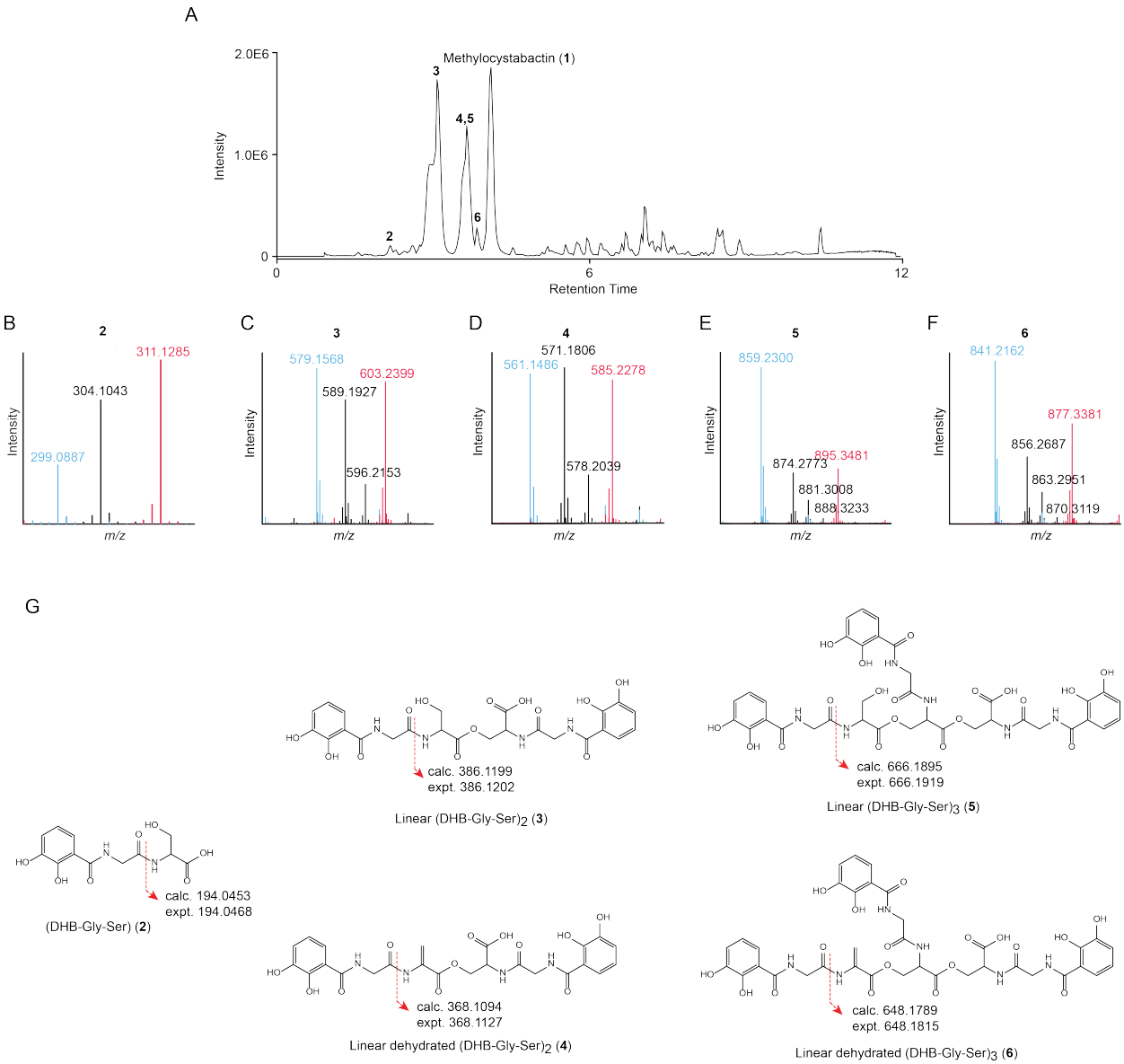


Figure S8. Linear DHB-incorporating metabolites identified through InverSIL in the supernatant of a *Methylosinus* sp. strain LW3 culture. (A) Total ion chromatogram of *Methylosinus* sp. strain LW3 crude supernatant extract. (B-F) Overlayed mass spectra of compounds **2**-**6** from *Methylosinus* sp. strain LW3 grown in ^{12}C-methane (blue), (^{13}C)methane (red), and (^{13}C)methane + ^{12}C-DHB (black). (G) Structures and important fragments of compounds **2**-**6** based on LC-HRMS/MS analysis and InverSIL.

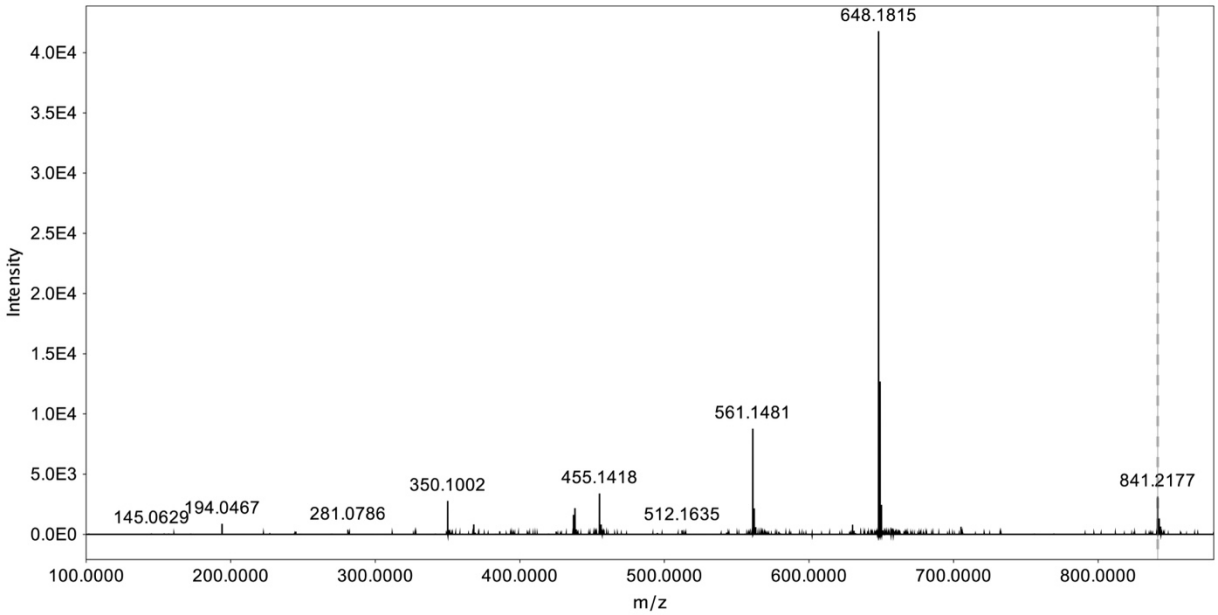


Figure S9. Positive ion MS/MS spectrum of methylocystabactin (**1**).

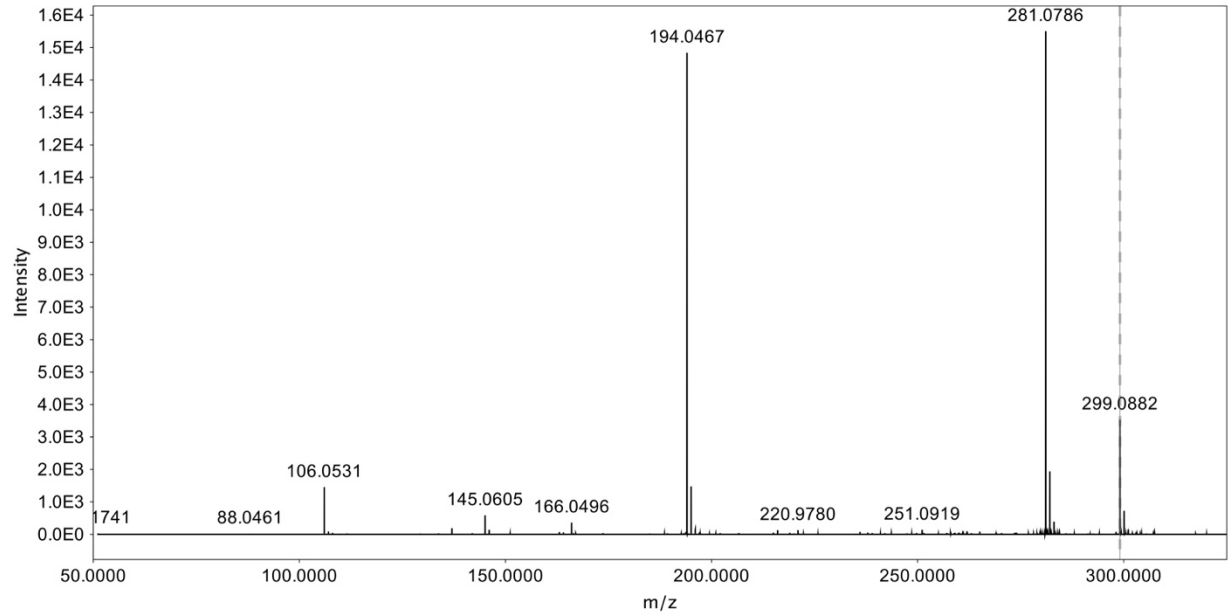


Figure S10. Positive ion MS/MS spectrum of linear (DHB-Gly-^{L}Ser) (**2**).

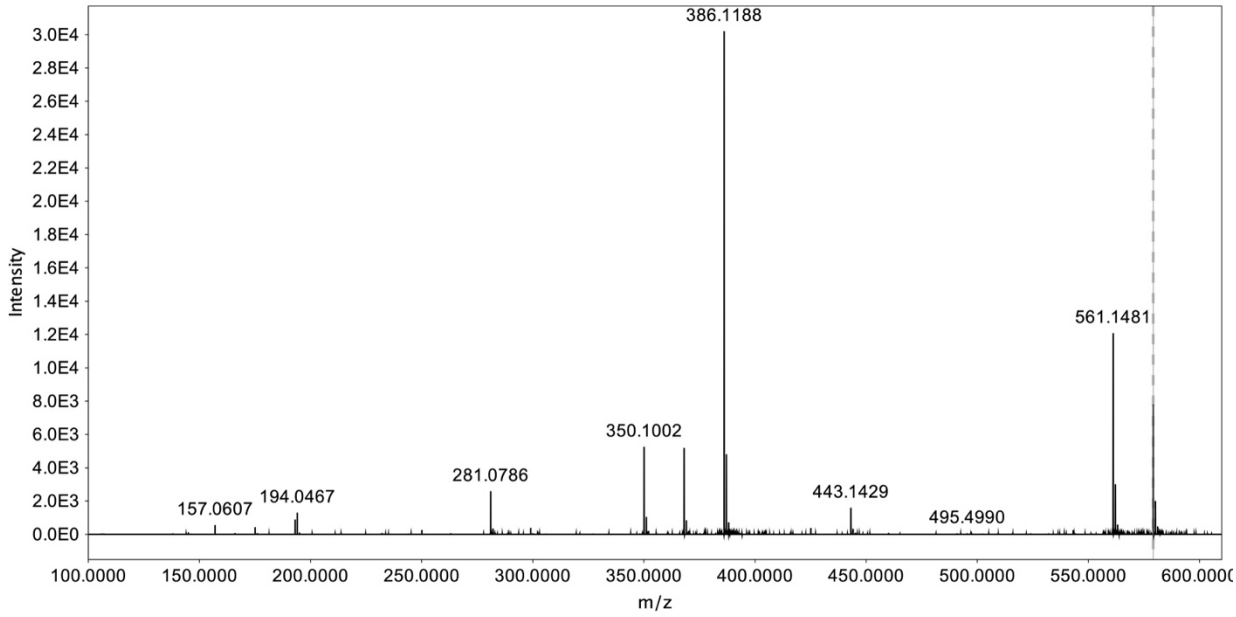


Figure S11. Positive ion MS/MS spectrum of linear (DHB-Gly-^{L}Ser)$_2$ (**3**).

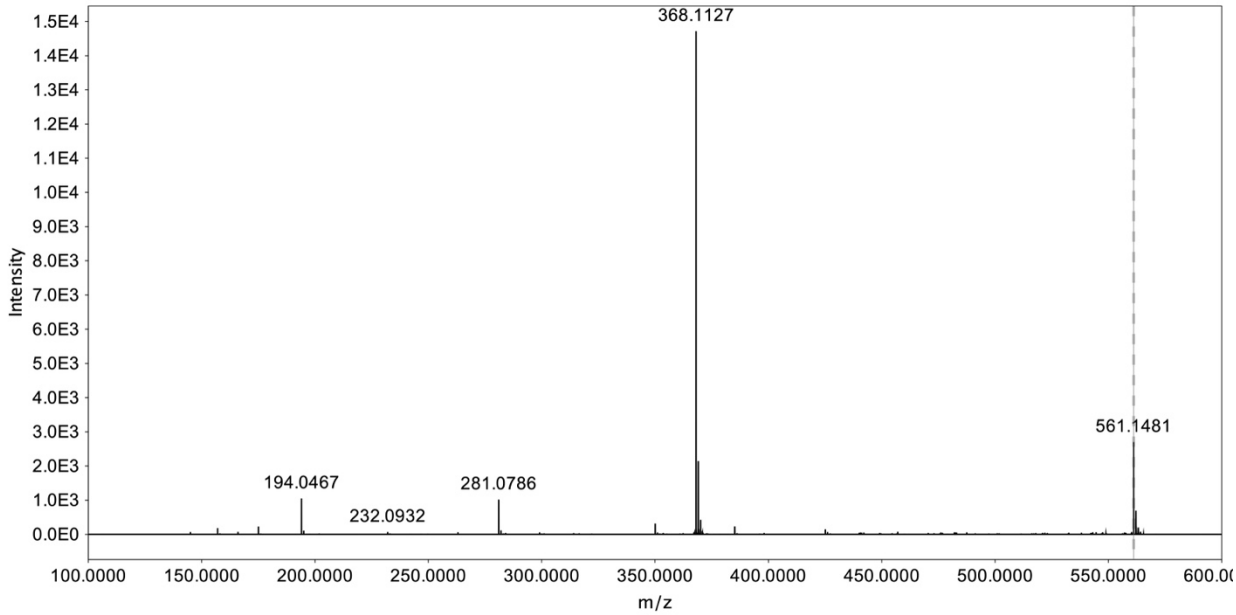


Figure S12. Positive ion MS/MS spectrum of linear dehydrated (DHB-Gly-^{L}Ser)$_2$ (**4**).

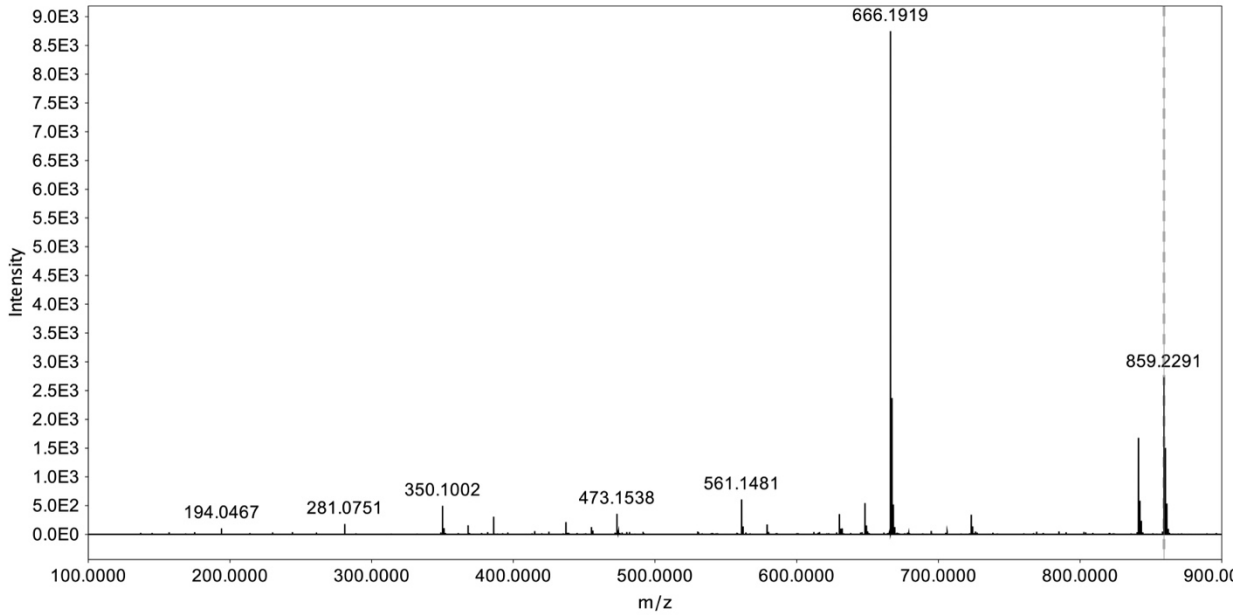


Figure S13. Positive ion MS/MS spectrum of linear (DHB-Gly-^{L}Ser)$_3$ (**5**).

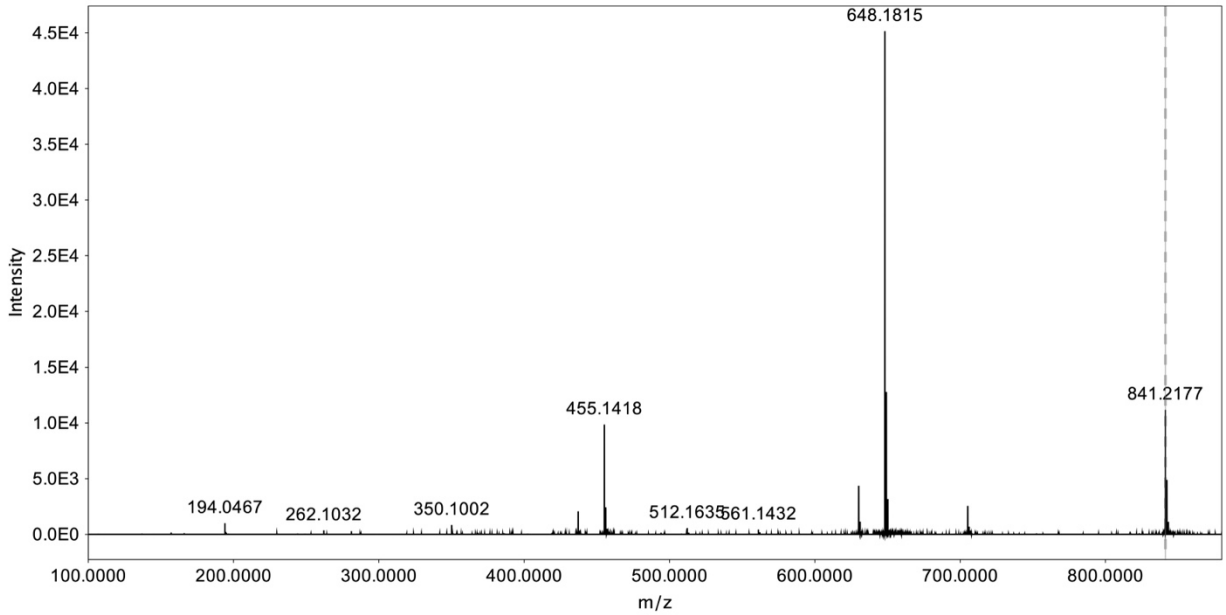


Figure S14. Positive ion MS/MS spectrum of linear dehydrated (DHB-Gly-^{L}Ser)$_3$ (**6**).

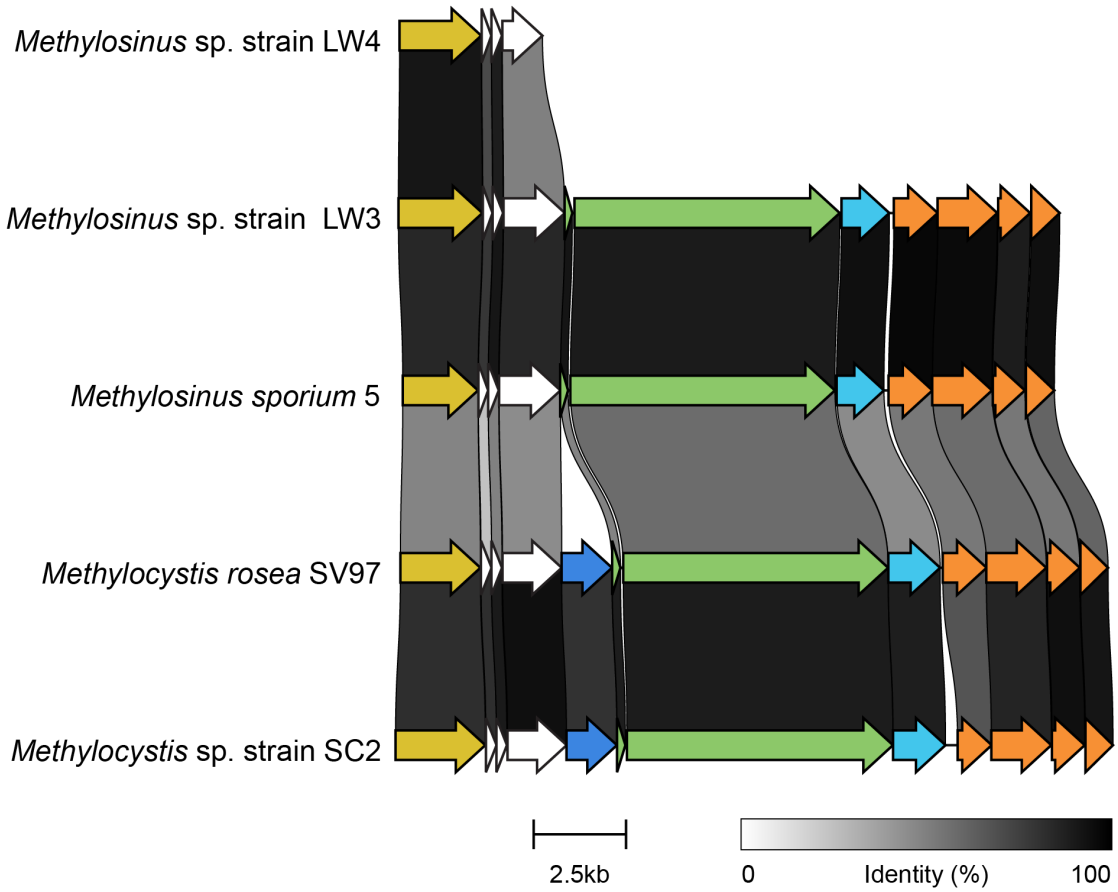


Figure S15. The methylocystibactin BGC is present in several species of the genera *Methylosinus* and *Methylocystis*. *Methylosinus* sp. strain LW4 is missing the BGC but has the putative TonB-dependent siderophore receptor. Yellow (TonB Dependent Siderophore Receptor); Green (NRPS); Orange (DHB operon); Blue (Enterochelin esterase); Cyan (Enterobactin exporter); White (hypothetical protein). Comparison visualized with Clinker (3).

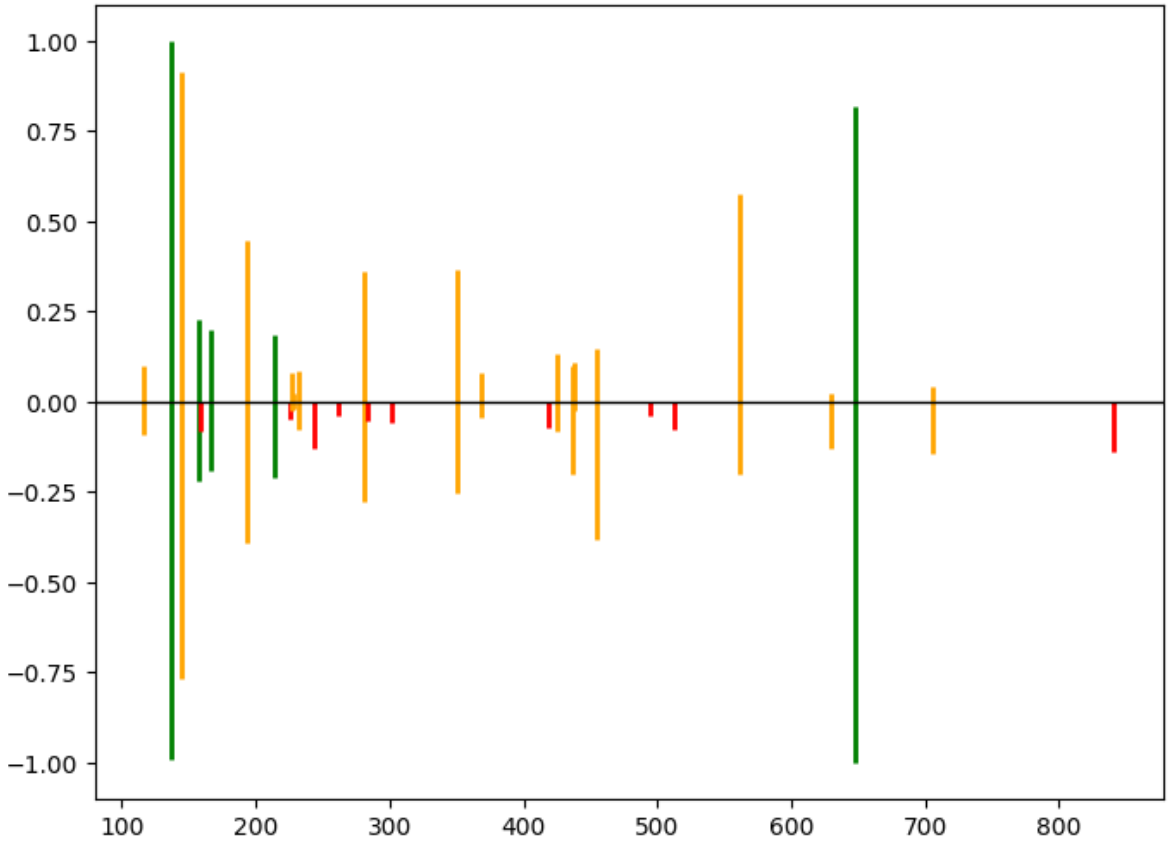


Figure S16. Mirror plot comparison of the MS/MS of the two *m/z* 841.*2159* found in the community and *Methylosinus* sp. strain LW3 extract. The upper plot represents the compound at t_R = 6.*2* minutes, while the lower plot corresponds to the compound at t_R = 5.9 minutes. Green peaks indicate common fragments with a ratio in the range of 0.97–1.2, whereas orange peaks represent common fragments with a ratio greater than 1.2 or less than 0.97. Red peaks correspond to unique fragments present in the t_R = 5.9 minute*s.*

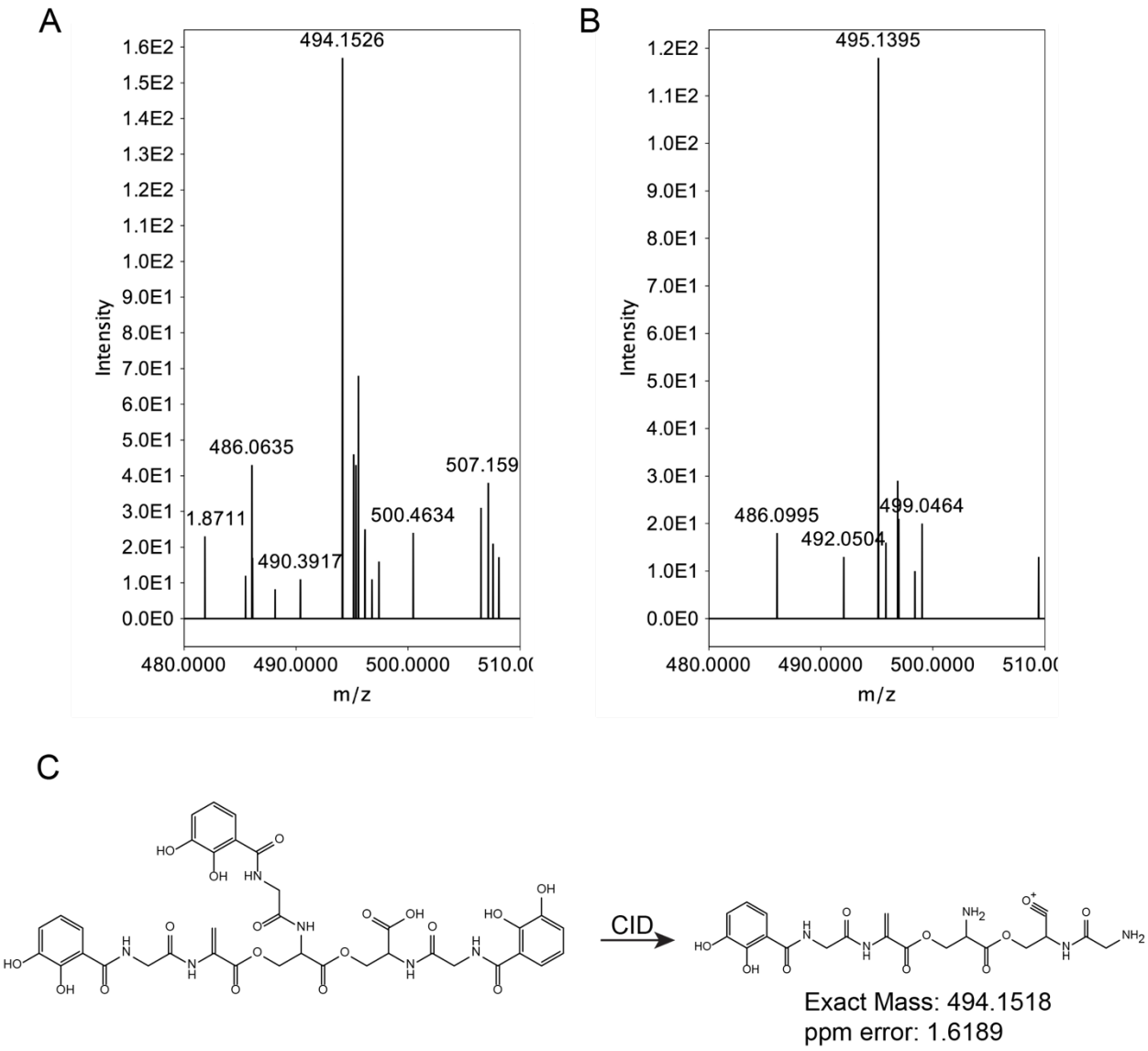


Figure S17. Focused comparison of the MS/MS spectra for the two *m/z* 841.2159 features at t_R = 5.9 min (A) and t_R = 6.2 min (B). (C) The hypothesized structure for the unique fragment *m/z* 494.1526 that can only be found in the linear dehydrated form. CID: collision-induced dissociation.

SUPPLEMENTARY TABLES

Table S1. List of bacterial families found in the methane-oxidizing bacterial community using Kaiju (4).

Family	Percent	Reads	taxon_id
Methylophilaceae	34.28	4424553	32011
Methylococcaceae	15.29	1973540	403
Comamonadaceae	12.28	1585528	80864
Flavobacteriaceae	5.18	668707	49546
Methylocystaceae	4.47	576615	31993
Chitinophagaceae	3.26	420543	563835
Burkholderiaceae	1.42	183778	119060
Pseudomonadaceae	1.39	179179	135621
Oxalobacteraceae	1.35	173718	75682
Caulobacteraceae	0.97	125040	76892
Xanthomonadaceae	0.91	117089	32033
Sulfuricellaceae	0.78	100354	2772226
Gallionellaceae	0.66	85247	90627
Bradyrhizobiaceae	0.65	83489	41294
Chromobacteriaceae	0.52	66973	1499392
Nitrosomonadaceae	0.51	66130	206379
Viruses	0.01	1100	10239
cannot be assigned to a (non-viral) family	3.94	509134	NA
belong to a (non-viral) family with less than 0.5% of all reads	12.14	1567558	NA
unclassified	40.03	8617734	NA

Table S2. Top 20 antiSMASH BGCs detected in the methane-oxidizing bacterial community metagenome and mapped to the metatranscriptome using BiG-MAP (2).

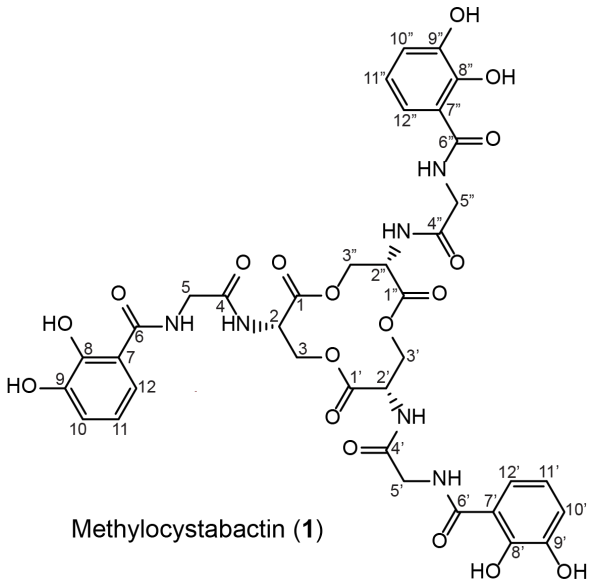
AntiSMASH BGC	BGC Number	Binning Genus	BGC-type	Length (bp)	RPKM		
					21113X1	21113X2	21113X3
NODE_103_length_219532_cov_736.217485.region001		*Ramlibacter*	redox-cofactor	219532	4344	45.4	7109.7
NODE_6487_length_9518_cov_5.987425.region001	NRPS BGC 1	*Methylosinus*	NRPS:T1PKS	9518	5477.7	0.5	596.8
NODE_10047_length_6474_cov_194.754012.region001		unbinned	NAPAA	6474	3597.5	0.4	683.9
NODE_2174_length_16094_cov_117.687948.region001	NRPS BGC 2	*Methylobacter*	NRPS	16094	1297	0.1	2910
NODE_104_length_138539_cov_6.082161.region001		*Lacibacter*	RiPP-like	138539	2601.2	24.9	1500.1
NODE_4787_length_12504_cov_116.938710.region001		unbinned	terpene	12504	568.5	0	2238.9
NODE_328_length_73047_cov_38.367465.region001		*Methylosinus*	redox-cofactor	73047	1993	0	152.3
NODE_3732_length_10596_cov_15.065933.region001		unbinned	T3PKS:hglE-KS	10596	1506.6	0.7	162.9
NODE_160_length_127501_cov_127.441826.region001		*Methylobacter*	redox-cofactor	127501	247.9	0.1	1314
NODE_474_length_39305_cov_79.198038.region001	NRPS BGC 3	*Methylosinus*	NRP-metallophore:NRPS	39305	877.3	0.2	635.1
NODE_1_length_1263261_cov_1930.356874.region003		*Methylophilus*	arylpolyene	1263261	1106.9	0	303.2
NODE_1535_length_21184_cov_6.861801.region001		*Methylotenera*	terpene	21184	899.7	8.4	398.1
NODE_1083_length_18527_cov_10.967951.region001	NRPS BGC 4	*Methylosinus*	NRP-metallophore:NRPS	18527	1259	0	25.2
NODE_6605_length_9370_cov_14.570156.region001	NRPS BGC 5	*Methylosinus*	NRPS	9370	1038.6	1.8	58.4
NODE_5343_length_7902_cov_10.079266.region001		unbinned	NRPS-like	7902	1043.3	0.3	49.8

NODE_5607_length_10826_cov_17.150868.region001		unbinned	NRPS	10826	990.6	1.1	14.8
NODE_1650_length_29586_cov_9.636687.region001		*Lacibacter*	terpene	29586	19.2	796.2	46
NODE_285_length_114100_cov_130.480258.region001		*Methylobacter*	RiPP-like	114100	287.3	0	497.3
NODE_8428_length_7547_cov_50.729578.region001		unbinned	NRPS-like	7547	633.4	0.3	114.4
NODE_1173_length_11150_cov_6.302208.region001		unbinned	terpene	11150	250.2	237.4	202

Table S3. Methylocystabactin MS/MS fragmentation comparison between crude extracts from the methane-oxidizing bacterial community and *Methylosinus* sp. strain LW3.

Community		***Methylosinus* sp. strain LW3**	
m/z	Intensity	*m/z*	Intensity
137.0231	103280	145.0607	124060
145.0607	91129	137.0233	113032
194.0445	47512	194.0447	59096
350.0982	39755	281.0768	44178
281.0779	39427	350.0977	42931
157.0603	28668	157.0606	29451
425.1320	17466	214.0826	22384
214.0828	17454	166.0505	19144
166.0496	17441	425.1289	17321
561.1496	14325	561.1469	14697

Table S4. Summary of 1H NMR data (δ in ppm, *J* in Hz) and ^{13}C NMR data (δ in ppm, type) for methylocystabactin in CD_3OD.



Methylocystabactin (**1**)

Position	δC, type	δH (*J* in Hz)	COSY	HMBC
L-Serine				
1, 1', 1"	168.75, C			
2, 2', 2"	52.70, CH	4,81, t (3.8 Hz)	3	3, 1, 4
3, 3', 3"	64.41, CH_2	4.67, dd (11.5, 3.6 Hz)	2	2
		4.35, dd (11.5, 4.1 Hz)		
Glycine				
4, 4', 4"	170.62, C			
5, 5', 5"	42.15, CH_2	4.06		4
DHB				
6, 6', 6"	170.27, C			
7, 7', 7"	115.31, C			
8, 8', 8"	148.47, C			
9, 9', 9"	145.78, C			
10, 10', 10"	118.32, CH	6.91, dd (7.9, 1.5 Hz)	11, 12	7, 8, 9, 11, 12
11, 11', 11"	118.32, CH	6.69, t (8.0 Hz)	10, 12	6, 7, 8, 9, 10, 12
12, 12', 12"	117.93, CH	7.24, dd (8.1, 1.5 Hz)	10, 11	6, 7, 8, 9, 11, 12

Table S5. Advanced Marfey's analysis of methylocystabactin.

Sample	**Retention time (min)**
Marfey's derivatized D-Serine standard (*m/z* 358)	4.021
Marfey's derivatized L-Serine standard (*m/z* 358)	3.864
Marfey's derivatized methylocystabactin hydrolysate (*m/z* 358)	3.865

Table S6. Molecular ions and common mass fragments of DHB-containing secondary metabolites detectable in the supernatant of a *Methylosinus* sp. strain LW3 culture.

Methylocystabactin (**1**)	Linear dehydrated (DHB-Gly-^{L}Ser)$_3$ (**6**)	Linear (DHB-Gly-^{L}Ser)$_3$ (**5**)	Linear dehydrated (DHB-Gly-^{L}Ser)$_2$ (**4**)	Linear (DHB-Gly-^{L}Ser)$_2$ (**3**)	Fragment
841.2177	841.2177	859.2291	561.1481	579.1568	Parent Mass
705.1994	705.1940	723.2209	425.1309	443.1429	Loss of DHB
648.1815	648.1815	666.1919	368.1127	386.1188	Loss of DHB-Gly
194.0467	194.0467	194.0467	194.0467	194.0467	DHB-Gly

Table S7. BLAST comparisons of amino acid sequences of select predicted TonB-dependent receptors encoded in the genomes *Methylosinus* sp. strains LW3 and LW4 with characterized TonB-dependent receptors. Each cell lists the E-value, percent identity, and percent query sequence coverage separated by commas for the indicated comparison. The sequence listed in the lefthand column was the query sequence. E-values < 1e-10 are shaded yellow and < 1e-100 are shaded blue.

	LW3 Mcb[a]	**LW3 Other**[b]	**LW4 Mcb**[c]	**LW4 Other**[d]	**FatA**[e]	**FcuA**[f]	**FepA**[g]	**IroN**[h]	**CirA**[i]
LW3 Mcb[a]	0.0e+00, 100,100	1.9e-04, 40,6	0.0e+00, 97,99	1.7e-04, 40,7	4.9e-144, 39,98	1.7e-132, 36,98	1.2e-07, 31,16	3.0e-06, 47,6	6.3e-11, 29,19
LW3 Other[b]	2.0e-04, 40,6	0.0e+00, 100,100	2.1e-04, 40,6	0.0e+00, 95,100	3.1e-04, 38,6	3.2e-04, 31,24	2.0e-15, 33,24	1.3e-15, 33,25	5.3e-30, 24,85
LW4 Mcb[c]	0.0e+00, 97,90	2.4e-04, 40,6	0.0e+00, 100,100	2.0e-04, 40,7	4.5e-144, 39,89	3.8e-130, 36,89	3.0e-07, 31,15	5.2e-06, 47,5	2.8e-10, 29,17
LW4 Other[d]	1.8e-04, 40,6	0.0e+00, 95,100	1.9e-04, 40,6	0.0e+00, 100,100	2.8e-04, 38,6	9.0e-03, 38,6	1.2e-14, 36,17	2.5e-15, 32,25	3.6e-31, 24,85

[a]*Methylosinus* sp. strain LW3 predicted methylocystabactin TonB-dependent receptor; IMG gene ID 2517142394.
[b]*Methylosinus* sp. strain LW3 predicted other TonB-dependent receptor; IMG gene ID 2517139767.
[c]*Methylosinus* sp. strain LW4 predicted methylocystabactin TonB-dependent receptor; IMG gene ID 2516608097.
[d]*Methylosinus* sp. strain LW4 predicted other TonB-dependent receptor; IMG gene ID 2516605591.
[e]*Vibrio anguillarum* TonB-dependent receptor FatA; IMG gene ID 2817597447; UniProt ID P11461.
[f]*Yersinia enterocolitica* TonB-dependent receptor FcuA; IMG gene ID 8118011944; UniProt ID Q05202.
[g]*Escherichia coli* TonB-dependent receptor FepA; IMG gene ID 2600371883; UniProt ID P05825.
[h]*Salmonella enterica* subsp. *enterica* serovar Typhimurium TonB-dependent receptor IroN; IMG gene ID 637213488; UniProt ID Q8ZMN0.
[i]*Escherichia coli* TonB-dependent receptor CirA; IMG gene ID 2600368988; UniProt ID P17315.

Table S8. Strains and plasmids used in this study.

Strain or plasmid	**Puri Lab Strain Collection Number**	**Description**	**Source or reference**
Strains			
E. coli TOP10	EAWP2	F– *mcrA* Δ(*mrr-hsdRMS-mcrBC*) Φ80*lacZ*ΔM15 Δ*lacX74 recA1 araD139* Δ(*ara leu*) 7697 *galU galK rpsL* (Str^R) *endA1 nupG*	Invitrogen
E. coli S17-1 λpir	EAWP3	Donor strain. Tp^R Sm^R *recA thi pro hsd*(r^-m^+) RP4-2-Tc::Mu::Km Tn7 λ*pir*	(5)
Methylosinus sp. strain LW3	AWP131	Methane-oxidizing bacteria isolated from Lake Washington Sediment	(6)
Methylosinus sp. strain LW4	AWP221	Methane-oxidizing bacteria isolated from Lake Washington Sediment	(6)
Methylosinus sp. strain LW3 *ΔdhbA*	AWP443	*Deletion of ΔdhbA* (IMG gene ID 2517142384)	This study
Methylosinus sporium 5	AWP276	Methane-oxidizing bacteria belonging to family *Methylocystaceae*	(7)
Methylocystis rosea SV97T	AWP275	Methane-oxidizing bacteria belonging to family *Methylocystaceae*	(8)
Methylocystis sp. strain SC2	AWP237	Methane-oxidizing bacteria belonging to family *Methylocystaceae*	(9)
Plasmids	**Description**		
pAWP484	pCM433KanT containing flanks to knock out *dhbA* (IMG gene ID 2517142384)		This study

Table S9. Primers used in this study

Primer	Sequence (5' – 3')	Description
oAWP1583_pAWP484_UP_fwd	cgtcagtatggcgcccatctcgtccctcctttcgtg	For amplifying flanks to knockout *dhbA* using pAWP484 (IMG gene ID 2517142384)
oAWP1584_pAWP484_UP_rev	catcggctggatcctctagtgagctgctcattcacgacatgcagg	
oAWP1585_pAWP484_DOWN_fwd	ctgaattcagctgtacaattggtaccgcgtcgatcgccaattcatcc	
oAWP1586_pAWP484_DOWN_rev	aggagggacgagatgggcgccatactgacggcctaga	

SUPPLEMENTARY REFERENCES

1. A. W. Han, *et al.*, Turnerbactin, a Novel Triscatecholate Siderophore from the Shipworm Endosymbiont Teredinibacter turnerae T7901. *PLOS ONE* **8**, e76151 (2013).

2. V. Pascal Andreu, *et al.*, BiG-MAP: an Automated Pipeline To Profile Metabolic Gene Cluster Abundance and Expression in Microbiomes. *mSystems* **6**, 10.1128/msystems.00937-21 (2021).

3. M. van den Belt, *et al.*, CAGECAT: The CompArative GEne Cluster Analysis Toolbox for rapid search and visualisation of homologous gene clusters. *BMC Bioinformatics* **24**, 181 (2023).

4. P. Menzel, K. L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* **7**, 11257 (2016).

5. R. Simon, U. Priefer, A. Pühler, A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology* **1**, 784–791 (1983).

6. A. J. Auman, S. Stolyar, A. M. Costello, M. E. Lidstrom, Molecular Characterization of Methanotrophic Isolates from Freshwater Lake Sediment. *Appl. Environ. Microbiol.* **66**, 5259–5266 (2000).

7. R. Whittenbury, K. C. Phillips, J. F. Wilkinson, Enrichment, Isolation and Some Properties of Methane-utilizing Bacteria. *J. Gen. Microbiol.* **61**, 205–218 (1970).

8. I. Wartiainen, A. G. Hestnes, I. R. McDonald, M. M. Svenning, Methylocystis rosea sp. nov., a novel methanotrophic bacterium from Arctic wetland soil, Svalbard, Norway (78° N). *Int. J. Syst. Evol. Microbiol.* **56**, 541–547 (2006).

9. B. Dam, S. Dam, M. Kube, R. Reinhardt, W. Liesack, Complete Genome Sequence of Methylocystis sp. Strain SC2, an Aerobic Methanotroph with High-Affinity Methane Oxidation Potential. *J. Bacteriol.* **194**, 6008–6009 (2012).