Shipworm symbiosis ecology-guided discovery of an antibiotic that kills colistin-resistant *Acinetobacter*

Bailey W. Miller,[1] Albebson L. Lim,[1] Zhenjian Lin,[1] Jeannie Bailey,[2] Kari L. Aoyagi,[3] Mark A. Fisher,[3] Louis R. Barrows,[4] Colin Manoil,[2] Eric W. Schmidt,*[1,5] and Margo G. Haygood*[1]

[1]Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah, 81112, USA
[2]Department of Genome Sciences, University of Washington, Seattle, Washington, 98195, USA
[3]Department of Pathology and ARUP Laboratories, University of Utah, Salt Lake City, Utah, 84112, USA
[4]Department of Pharmacology and Toxicology, University of Utah, Salt Lake City, Utah, 84112, USA
[5]Lead Contact: Eric W. Schmidt
Correspondence: ews1@utah.edu, margo.haygood@utah.edu

Table of Contents

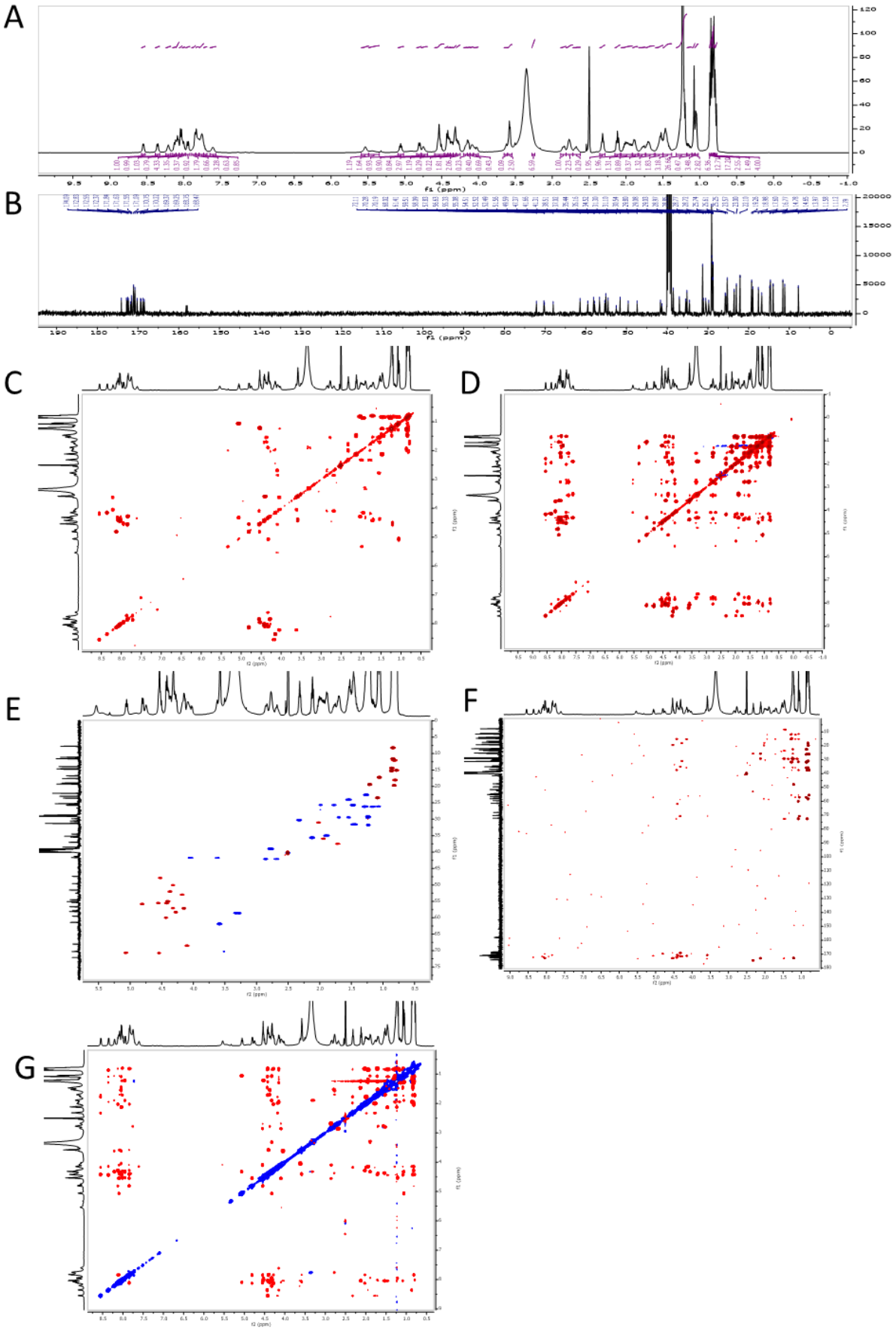


Figure S1. NMR Spectra of **1**, related to Figure 2. A) ^{1}H NMR spectrum of **1** in DMSO-d_6, 500 MHz, B) ^{13}C NMR Spectrum of **1**, DMSO-D_6, 125 MHz, C) gCOSY spectrum of **1**, DMSO-d_6, 500 MHz, D) zTOCSY spectrum of **1**, DMSO-d_6, 500 MHz, E) gHSQCAD spectrum of **1**, DMSO-d_6, 500 MHz, F) gHMBCAD spectrum of **1**, DMSO-d_6, 500 MHz, G) 2D ROESY spectrum of **1**, DMSO-d_6, 500 MHz

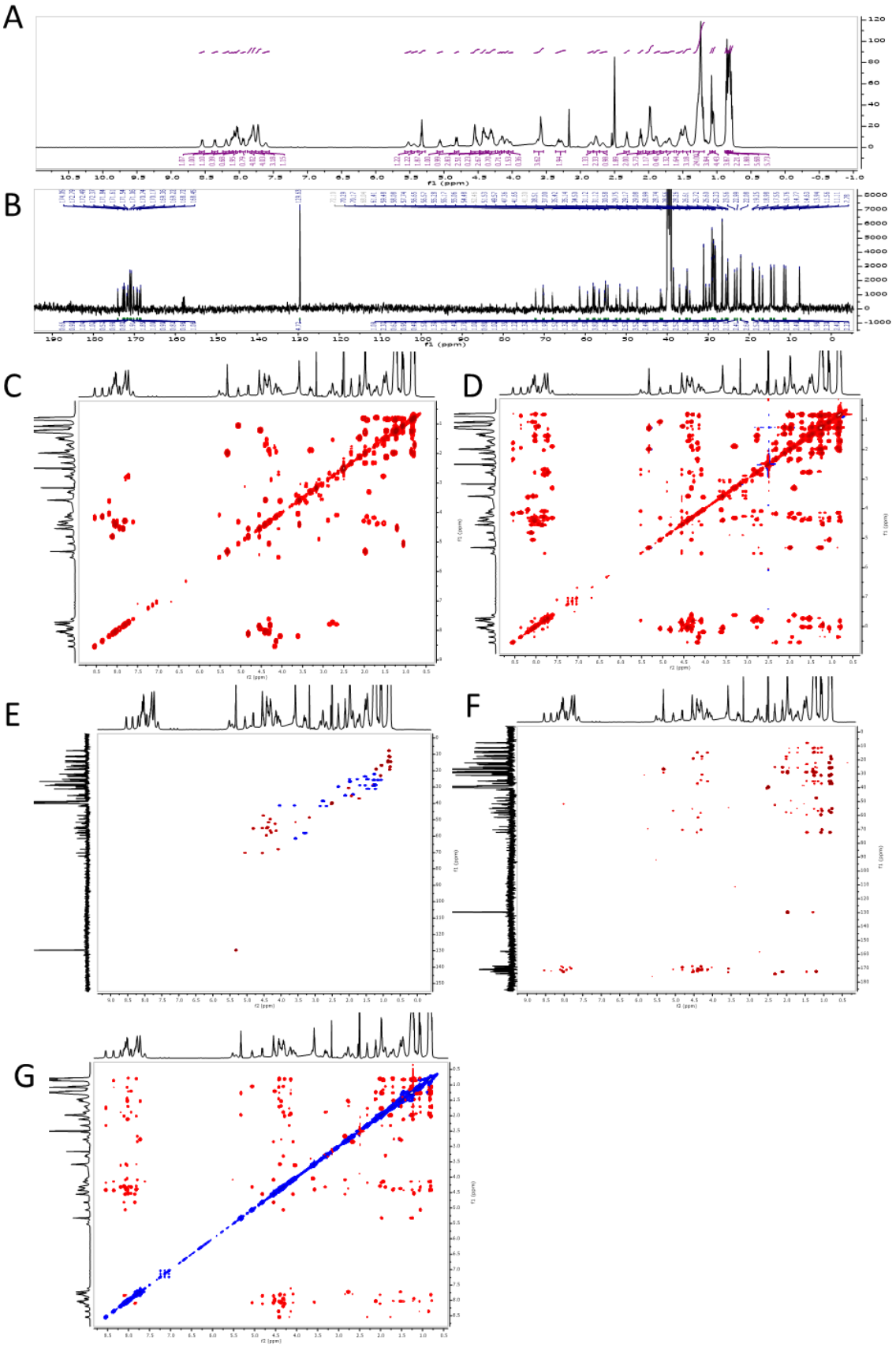


Figure S2. NMR Spectra of **2**, related to Figure 2. A) ^{1}H NMR spectrum of **2** in DMSO-d_6, 500 MHz, B) ^{13}C NMR Spectrum of **2**, DMSO-D$_6$, 125 MHz, C) gCOSY spectrum of **1**, DMSO-d_6, 500 MHz, D) zTOCSY spectrum of **2**, DMSO-d_6, 500 MHz, E) gHSQCAD spectrum of **2**, DMSO-d_6, 500 MHz, F) gHMBCAD spectrum of **2**, DMSO-d_6, 500 MHz, G) 2D ROESY spectrum of **2**, DMSO-d_6, 500 MHz

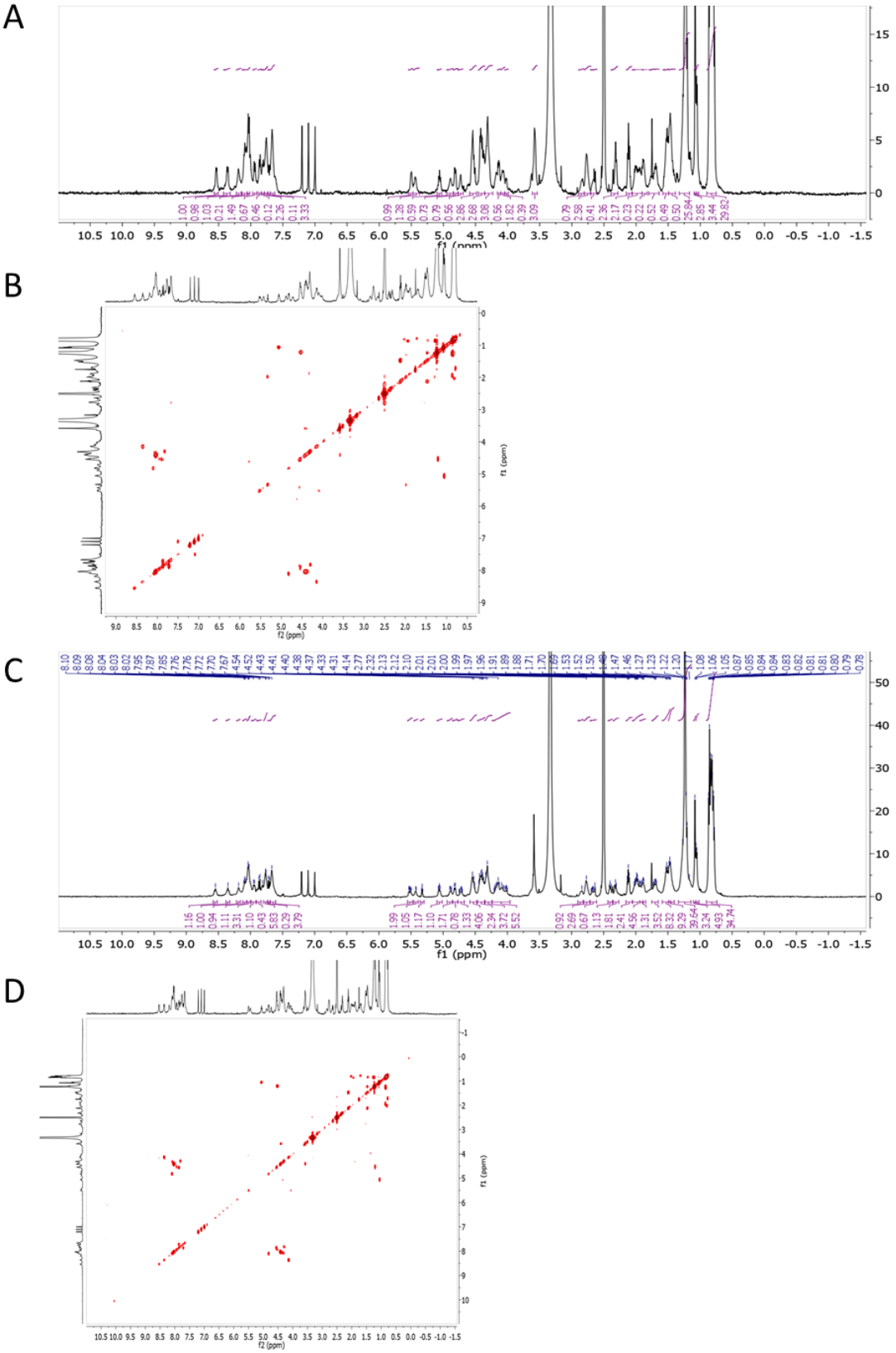
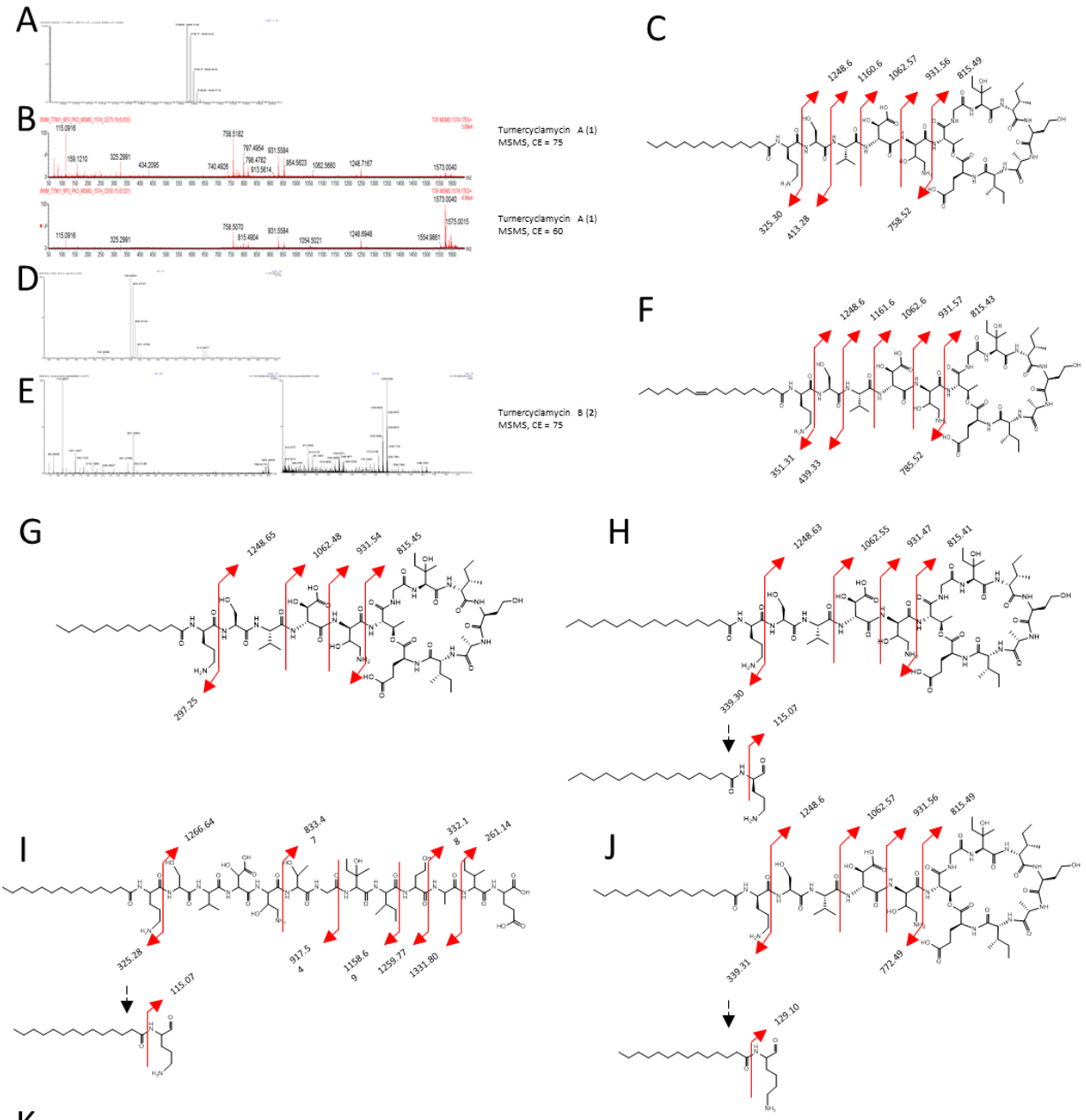


Figure S3: NMR Spectra of **3** and **4**, related to Figure 2. A) ^{1}H spectrum of **3**, DMSO-d_6, 500 MHz, B) COSY spectrum of **3**, DMSO-d_6, 500 MHz, C) ^{1}H NMR spectrum of **4**, DMSO-d_6, 500 MHz, D) COSY spectrum of **4**, DMSO-d_6, 500 MHz



K

Parent Mass $[M+2H]^{2+}$	T_R (min)	Fatty Acid	Peptide Sequence Modifications
779.96	3.01	C13:0	Same as 1
786.96	3.02	C13:0	Orn->Lys
785.96	2.97	C14:1	Same as 1
795.96	3.11	C14:0	Ring Open
779.96	3.05	C14:0	Loss of 14 amu in cyclic portion: Ile->Val
802.96	3.13	C14:0	Ring Open + Methyl Ester (Glu)
793.96	3.18	C14:0	Orn->Lys
793.96	3.24	C14:0	+14 in cyclic portion, possible methyl ester of Glu
808.99	3.21	C16:1	Ring Open
800.96	3.46	C16:0	Same as 1
800.96	3.58	C14:0	Ring Open, Methyl Ester on Asp?
813.97	3.7	C16:1	Ring Open, Methyl Ester on Asp?

Figure S4: Figure S4: HRESI-MS and MSMS fragmentation of **1-4**, minor analogs, related to Figure 2. A) MSMS fragmentation of **3**, B) MSMS fragmentation of **4**, C) Characteristic ions from a ring open analog. While the cyclic portion does not fragment well in the intact molecule, the linearized peptide shows fragments from within the tail. All fragments that include that region also include the mass of an additional H_2O. D) Characteristic fragment ions arising from the Orn->Lys analogs. The FA and the adjacent residue can be distinguished by the subsequent fragmentation to *m/z* = 115.07 (Orn) or 129.10 (Lys). A full table of minor analogs can be found in Table S2.

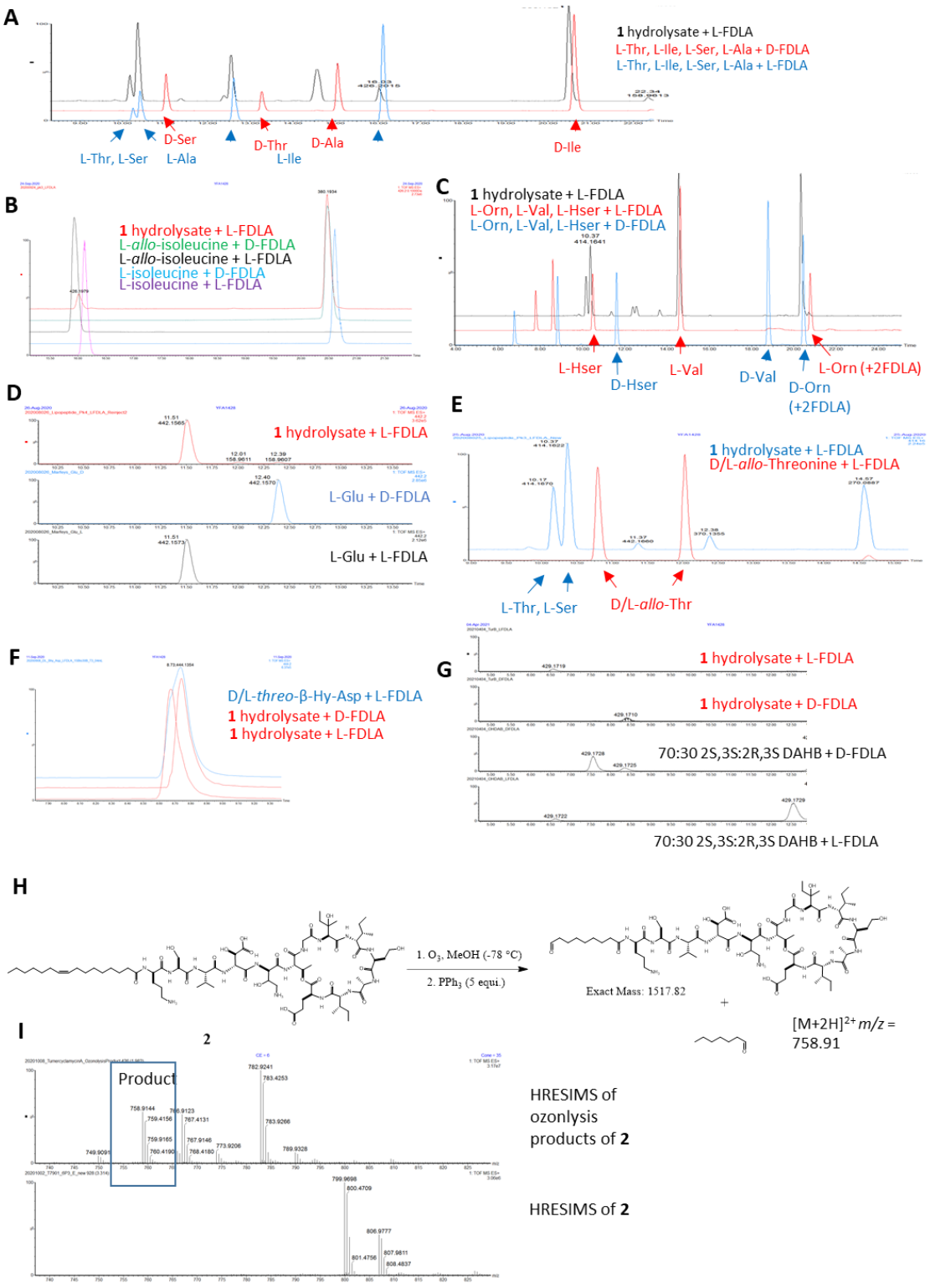


Figure S5: Advanced Marfey's for stereochemical analysis and ozonolysis for double bond placement of **2**, related to Figure 2. Panels A-E compare the L-FDLA derivitized hydrolysate of **1** to the L- or D-FDLA derivitized standards for A) L-Thr, L-Ile, L-Ser, L-Ala; B) L-*allo*-Ile and L-Ile; C) L-Orn, L-Val, L-homoSer; D) L-Glu; E) D/L-*allo*-Thr. Panel F shows EIC of L-FDLA derivitized D/L-*threo*-B-Hy-ASP compared to the L- and D-FDLA derivatized hydrolysate of **1**. Panel G shows EIC of L-FDLA or D-FDLA derivitized reaction mixtures from the synthesis of DAHB compared to EICs of the L-FDLA or D-FDLA derivatized hydrolysate of **1**. Panel H is a reaction scheme for ozolysis and reductive workup of **2**. I) HRESIMS of ozonlysis products. Only one product matched a mass that could be explained by a terminal aldehyde in the fatty acid chain. This mass matched a 9-formyl-nonoic acid containing peptide. The other masses observed match the aldehyde product with additional oxidations.

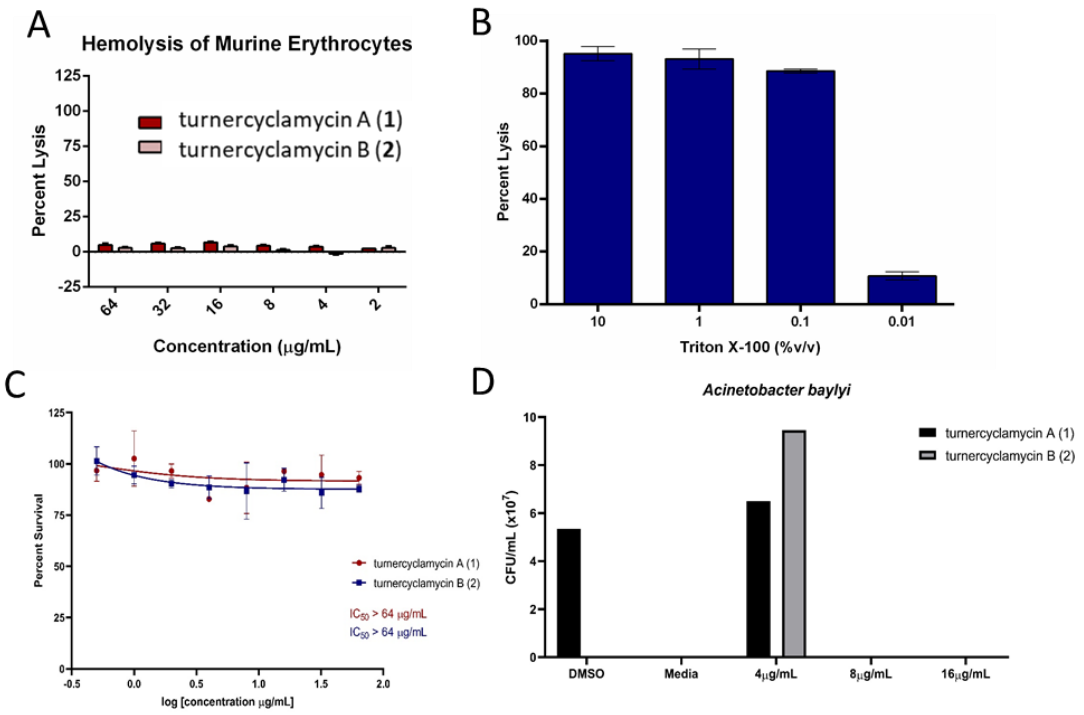


Figure S6: Additional antimicrobial and cytotoxicity experiments, related to Tables 1 and 2. A) **1** and **2** were both inactive in a cell lysis assay with murine erythrocytes, while B) Triton-X-100, the positive control, acted as expected. C) **1** and **2** were non-toxic to mammalian HEK-293 (ATCC® CRL-1573™) cells up to 64 µg/mL. D) In order to distinguish between bactericidal vs. bacteriostatic activity, *A. baylyi* was incubated in the presence of **1** (dark) and **2** (light) at 4, 8, and 16 µg/mL for 18 hours, then cultures were streaked on fresh agar plates to assess cell viability. At the MIC, no colonies formed on the plates, indicating bactericidal, rather than bacteriostatic activity

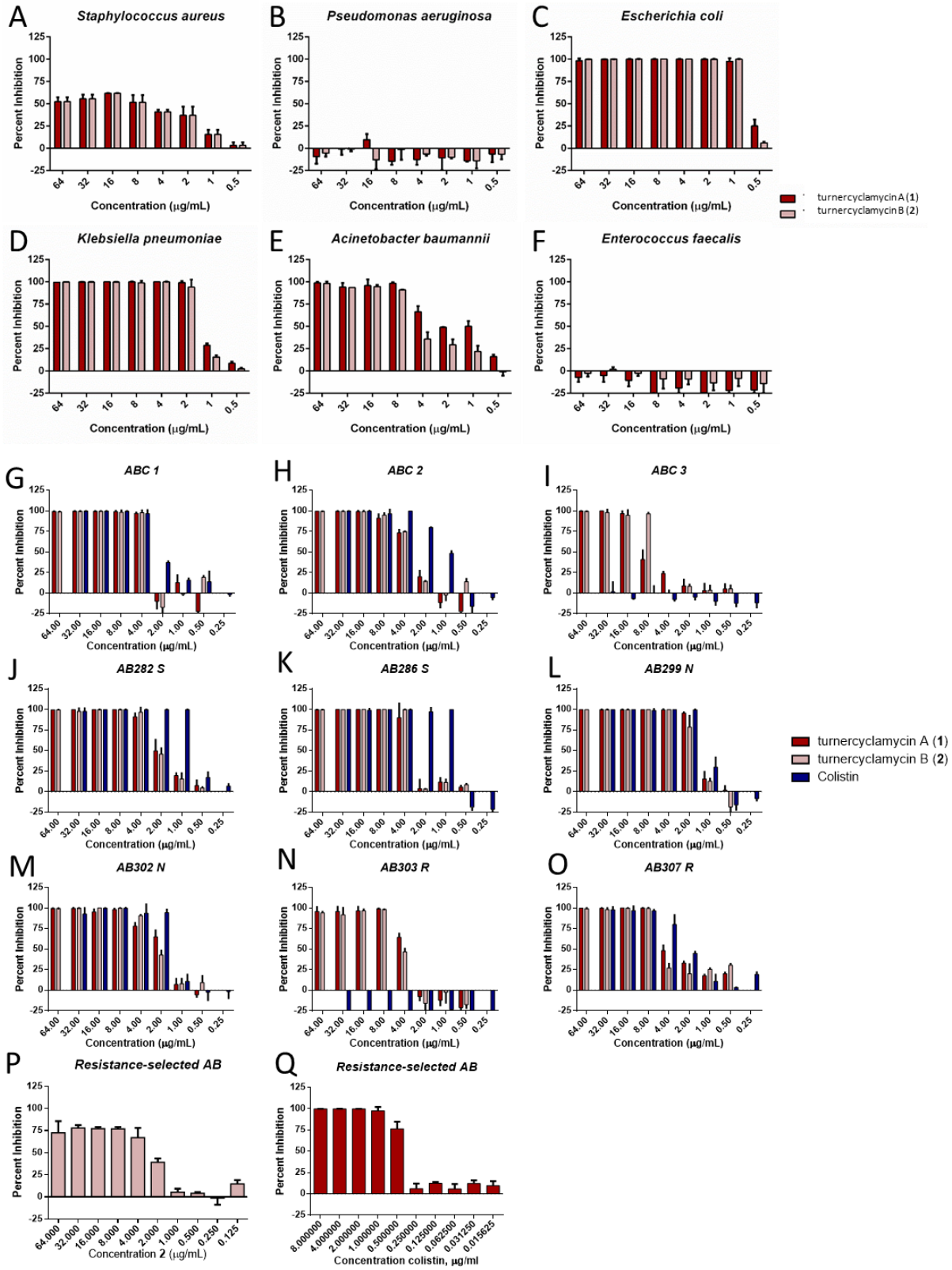


Figure S7: Broth microdilution assays for MIC determination, related to Tables 1 and 2. A-F) Antimicrobial microdilution assays for determination of minimum inhibitory concentrations of **1** (dark red), **2** (light red), and colistin (blue) against a panel of *Acinetobacter* strains. G-O) *Acinetobacter* complex (ABC) 1, 2, and 3 were obtained from ARUP laboratories. The additional six strains are all obtained from the CDC and have varying levels of resistance to the last line antibiotic colistin. Strains with the designation *S* are susceptible (colistin MIC = 1 µg/mL), *N* are normal (MIC = 2 µg/mL), and *R* are resistant strains (MIC > 8 µg/mL). Each strain was incubated in the presence of compound for 18-20 hours and viability was measured using MTT.

Table S1. ^{1}H and ^{13}C NMR assignments for **1** and **2**, related to Figure 2

Residue	Position	**1**		**2**	
		δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)
Ornithine	NH		8.03[a]		8.03[a]
	C=O	171.6		171.6	
	α-C	51.6	4.37 m	51.5	4.37 m
	β-C	29.0	1.53 m, 1.69 m	29.0	1.53 m, 1.69 m
	γ-C	23.6	1.53[a]	23.6	1.53[a]
	δ-C	38.5	2.77 br	38.5	2.77 br
	NH_3		7.72		7.72
Serine	NH		8.06[a]		8.06[a]
	C=O	170.2		170.2	
	α-C	54.5	4.39[a]	54.5	4.39[a]
	β-C	61.4	3.58 d (5.8)	61.4	3.59 d (5.8)
	β-OH		N/A		N/A
Valine	NH		7.82[a]		7.82[a]
	C=O	171.1		171.1	
	α-C	57.7	4.28	57.7	4.28
	β-C	30.5	2.02 m	30.6	2.02 m
	γ_1-C	17.6	0.80[a]	17.6	0.80[a]
	γ_2-C	19.3	0.82[a]	19.3	0.82[a]
β-OH-Aspartic acid	NH		8.12 d (9.4)		8.13 d (9.0)
	C=O	169.2		169.2	
	α-C	55.3	4.80 dd (8.9, 2.6)	55.3	4.82 dd (8.8, 2.6)
	β-C	70.3	4.55[a]	70.3	4.55[a]
	β-OH		N/A		N/A
	γ-C = O	168.5		168.5	
2,4-Diamino-3-hydroxybutyric acid	NH		7.84[a]		7.85[a]
	C=O	172.8		172.8	
	α-C	55.1	4.55[a]	55.1	4.55[a]
	β-C	68.0	4.09 m	68.0	4.08 m
	β-OH		5.54 br		5.52 br
	γ-C	41.7	2.68, 2.85 m	41.7	2.67, 2.86 m
	γ-NH_3		7.79 t (br)		7.79 t (br)
Threonine	NH		8.07[a]		8.07[a]
	C=O	168.7		168.7	
	α-C	56.6	4.32 m	56.6	4.32 m
	β-C	70.2	5.05 p (6.7)	70.2	5.06 p (6.7)
	γ-C	16.8	1.05 d (6.5)	16.8	1.05 d (6.5)
Glycine	NH		8.21 t (6.0)		8.19 t (6.1)
	C=O	169.4		169.4	
	α-C	41.3	3.6 m, 4.05 dd (16.2, 5.3)	41.3	3.6 m, 4.04 dd (16, 5.4)
3-OH-isoleucine	NH		8.03[a]		8.03[a]
	C=O	171.8		171.8	
	α-C	59.5	4.42 d (7.0)	59.5	4.42 d (7.3)
	β-C	72.1		72.1	
	γ_1-C	23.0	1.08 s	23.0	1.08 s
	γ_2-C	31.1	1.45 m	31.1	1.45 m
	δ-C	7.8	0.83[a]	7.8	0.83[a]
Isoleucine (1)	NH		8.35 d (7.7)		8.36 d (7.8)
	C=O	170.7		170.7	
	α-C	56.7	4.14 dd (7.6, 4.5)	56.7	4.14 dd (7.6, 4.6)
	β-C	35.4	1.93 m	35.4	1.93 m
	γ_1-C	14.8	0.86 d (6.8)	14.8	0.87 d (6.8

	γ_2-C	25.7	1.16 m, 1.29 m	25.7	1.16 m, 1.29 m
	δ-C	11.6	0.81[a]	11.6	0.81[a]
Homoserine	NH		7.60 d (br)		7.61 d (br)
	C=O	171.1		171.1	
	α-C	49.6	4.32[a]	49.6	4.32[a]
	β-C	34.5	1.89 m	34.5	1.89 m
	γ-C	58.1	3.31 m	58.1	3.31 m
	γ-OH				
Alanine	NH		7.94 d (8.6)		7.94 d (8.6)
	C=O	172.4		172.4	
	α-C	47.4	4.52 m	47.4	4.52 m
	β-C	19.0	1.20 d (7.1)	19.0	1.2 d (7.1)
Isoleucine (2)	NH		8.01 d (8.8)		8.01[a]
	C=O	171.5		171.5	
	α-C	55.0	4.41[a]	55.0	4.41[a]
	β-C	37.0	1.71 m	37.0	1.71 m
	γ_1-C	25.6	1.06[a], 1.25[a]	25.6	1.06[a], 1.25[a]
	γ_2-C	14.7	0.78 d (7.0)	14.6	0.78 d (7.8)
	δ-C	11.1	0.83[a]	11.1	0.83[a]
Glutamic acid	NH		8.54 d (6.7)		8.54 d (6.7)
	C=O	170.7		170.7	
	α-C	52.5	4.18 m	52.5	4.17 m
	β-C	25.1	1.77 m, 1.99 m	25.1	1.76 m, 1.99 m
	γ-C	29.8	2.32 t (7.4)	29.8	2.32 t (7.4)
	δ-C = O	174.1		174.1	
Fatty Acid	1	172.6		172.4	
	2	35.1	2.12 t (7.1)	35.1	2.12 t (7.1)
	3	25.3	1.48 m	25.3	1.47 m
	4	28.7-29.1	1.21-1.30	28.7-29.1	1.21-1.30
	5	28.7-29.1	1.21-1.30	28.7-29.1	1.21-1.30
	6	28.7-29.1	1.21-1.30	28.7-29.1	1.21-1.30
	7	28.7-29.1	1.21-1.30	29.2	1.29
	8	28.7-29.1	1.21-1.30	26.7	1.98
	9	28.7-29.1	1.21-1.30	129.6	5.32
	10	28.7-29.1	1.21-1.30	129.6	5.32
	11	28.7-29.1	1.21-1.30	26.7	1.98
	12	28.7-29.1	1.21-1.30	29.2	1.29
	13	28.7-29.1	1.21-1.30	28.7-29.1	1.21-1.30
	14	13.9	0.86	31.1	1.24
	15			22.1	1.26
	16			13.9	0.85

[a]multiplicity and coupling undetermined due to signal overlap

Table S2: Retention times for hydrolysate and amino acid standards, related to Figure 2

Amino Acid	$t_{rL\text{-}FDLA}$ (min) Analyte	t_{rL} (min) Standard	t_{rD} (min) Standard
Valine	14.58	14.63	18.7
Alanine	12.54	12.61	15.15
Serine	10.36	10.41	11.01
Threonine	10.17	10.25	13.27
allo-Threonine	N/A	10.8/12.05	10.8/12.05
Isoleucine	N/A	16.11	20.6
allo-Isoleucine	20.49	15.95	20.49
Glutamic acid	11.51	11.51	12.4
Homoserine	10.36	10.47	11.44
β-OH-Aspartic Acid	8.66	8.73	8.66
αN,δN-diFDLA Ornithine	20.4	20.82	20.45
2*S*,3*S*-β-OH-DAB	N/A	7.55	12.6
2*R*,3*S*-β-OH-DAB	8.3	8.31	6.59

Table S3: Gene annotations and nearest homologs of biosynthetic gene cluster, related to Figure 3

Gene	Gene name	Protein Size (AAs)	Annotation or Proposed Function	Homolog (Accession, organism)	Identity (%)
TERTU_RS10355	*tur*F	332	Fe/αKg dependent dioxygenase	WP_144695392.1, Alteromonadaceae bacterium 2753L.S.0a.02	71.9
TERTU_RS10350	*tur*G	290	jmjC-like cupin domain containing protein	WP_044617703.1, Gynuella sunshinyii	47.45
TERTU_RS10345	*tur*H	322	Fe/αKg dependent dioxygenase	WP_044619903.1, Gynuella sunshinyii	61.49
TERTU_RS10340	*tur*I	561	Cyclic peptide export ABC transporter	WP_044616742.1, Gynuella sunshinyii	68.86
TERTU_RS10335	*tur*A	3718	NRPS	WP_044616743.1, Gynuella sunshinyii	48.39
TERTU_RS10330	*tur*B	4249	NRPS	WP_044616744.1, Gynuella sunshinyii	54.07
TERTU_RS10325	*tur*C	2214	NRPS	WP_044616745.1, Gynuella sunshinyii	56.71
TERTU_RS10320	*tur*D	2685	NRPS	WP_052830181.1, Gynuella sunshinyii	62.56
TERTU_RS21365	*tur*E	4025	NRPS	WP_052830181.1, Gynuella sunshinyii	60.44
TERTU_RS10310	*tur*J	531	MBL fold metallo-hydrolase	WP_044616748.1, Gynuella sunshinyii	72.69
TERTU_RS10305	*tur*K	542	Family 43 glycosylhydrolase	WP_012488586.1, Cellvibrio japonicus	67.34
TERTU_RS10300	*tur*L	417	Lipolytic enzyme	NVK57725.1, Alteromonadaceae bacterium	68.01

Table S4: Symbiont strain and host species used in *tur* cluster analysis, related to Figure 1

Symbiont	Host		
Isolate	**Genus**	**Species**	**Location**
T7901	*Bankia*	*gouldi*	Beaufort, NC, USA
T8602	*Dicyathifer*	*manni*	Townsville, QLD, Aus.
T7902	*Lyrodus*	*pedicellatus*	Long Beach, CA, USA
T8513	*Teredo*	*navalis*	Sao Paulo, Brazil
T8412	*Lyrodus*	*bipartitus*	Jim Isl., Fort Pierce, FL
T8402	*Teredora*	*malleolus*	Floating wood
PMS-1133Y.S.0a.04	*N/A*	*N/A*	Bil-isan, Panglao, Bohol, Philippines
PMS-991H.S.0a.06	*Lyrodus*	*pedicellatus*	Danao, Iloilo City, Iloilo, Philippines

Script 1: *tur* gene cluster analysis script for generation of identity matrix, related to Figure 1

```
#!/bin/bash
	rm temp3
	rm $2
	n1=$(grep -n "FEATURES" $1 | awk -F : '{print $1}')
	n2=$(grep -n "ORIGIN" $1 | awk -F : '{print $1}')
	sed -n "$n1, $n2 p" $1 > temp1
	cut -b 1-20 temp1  | nl -b a  > temp2
	while read s1 s2; do
		if [ "$s2" != "" ]; then
			echo $s1 $s2 >> temp3
		fi
	done < temp2
	nl -b a temp3 > temp3_1
	while read line1 line2 line3; do
		if [ "$line3" = "unsure" ]; then   # replace "unsure" with "CDS" if extract NRPS genes
			line_start=$line2
			line_next=$[$line1+1]
			line_end=$[$(sed -n "$line_next p" temp3_1 | awk '{print $2}')-1]
			sed -n "$line_start, $line_end p" temp1 | sed 's/^.\{21\}//' > temp4
			grep -n "\/" temp4 | awk -F : '{print $1,$2}' | nl -b a  > temp5
			temp4_line=$(grep \^ temp4 | wc -l)
			temp5_line=$(grep \^ temp5 | wc -l)
			seq_head1=$(echo $1 | awk -F / '{print $NF}' | awk -F . '{print $1}')
			seq_head2=$(grep "/domain_id=" temp4  | awk -F \"" '{print $2}' | sed 's/ /_/g')
				# seq_head2=$(grep "/locus_tag=" temp4 | sed 's/\/locus_tag="//' | sed 's/"//g' | awk
-F _ '{print $(NF-1)"_"$NF}')  # enable this line if extract NRPS genes
			seq_head3=$(grep "/specificity=" temp4  | awk -F \"" '{print $2}' | sed 's/ /_/g')
			aSDomain=$(grep "/aSDomain=" temp4 | awk -F \"" '{print $2}' | sed 's/ /_/g')
			if [ "$aSDomain" = "AMP-binding" ]; then   # remove this if loop if extract A-donmain only
				echo \>$seq_head1"_"$seq_head2"_"$seq_head3 >> $2
				read position seq_start <<< $(grep  "translation=" temp5 | awk  '{print $1,$2}')
				echo  $seq_start
				if [ $position -eq $temp5_line ]; then
					seq_end=$temp4_line
					elif [ $position -lt $temp5_line ]; then
						seq_end=$(sed -n "$[$position+1] p " temp5 | awk '{print $2}')
						echo $seq_end
				fi
				sed -n "$seq_start, $[$seq_end-1] p" temp4 | sed 's/\/translation="//g' | sed 's/"//g'>> $2
			fi # remove this if loop if extract A-donmain only
		fi
	done < temp3_1
.
```