Supporting Information for

Mixing and Matching Siderophore Clusters: Structure and Biosynthesis of Serratiochelins from *Serratia sp.* V4

Mohammad R. Seyedsayamdost,[†,ǁ] Sara Cleto,[‡,§,ǁ] Gavin Carr,[†] Hera Vlamakis,[‡] Maria João Vieira,[§] Roberto Kolter,[‡,*] Jon Clardy[†,*]

†Department of Biological Chemistry and Molecular Pharmacology, and ‡Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts 02115, USA

[§]Institute for Biotechnology and Bioengineering, Center of Biological Engineering, University of Minho, 4710-057 Braga, Portugal

[ǁ]These authors contributed equally

*email: roberto_kolter@hms.harvard.edu; jon_clardy@hms.harvard.edu

General procedures. UV-visible absorbance spectra were acquired on an Amersham Biosciences Ultrospec 5300 Pro Spectrophotometer. HPLC purifications were carried out on an Agilent 1100 or 1200 Series HPLC system equipped with a photo diode array detector. HPLC-MS analysis was performed on an Agilent 1200 Series HPLC system equipped with a diode array detector and a 6130 Series ESI mass spectrometer using an analytical Phenomenex Luna C18 column (5 µm, 4.6 × 100 mm) operating at 0.7 mL/min with a gradient of 10 % MeCN in H_2O to 100 % MeCN over 25 min. High resolution mass spectrometry (HR-MS) and tandem MS (MS/MS) were performed at the University of Illinois Urbana-Champaign Mass Spectrometry Facility. ^{1}H, ^{13}C, and 2D NMR spectra were recorded in the inverse-detection probe of a Varian Inova spectrometer (600 MHz for ^{1}H, 150 MHz for ^{13}C). Chemical shifts were referenced to the residual solvent peaks in D_2O.

Bacterial strains, plasmids, and growth conditions. *Serratia sp.* V4, the conjugative strain *Escherichia coli* S17-1 λPir, *E. coli* ER1100A[1] and *E. coli* K12 were maintained on Luria Bertani agar (Novagen) at 30°C (*Serratia sp.* V4) or 37 °C (other strains, Table S5).[1-2] The *E. coli* S17 strain containing plasmid pBTK30, later modified and used for targeted mutagenesis, was maintained in LB agar supplemented with 10 µg/mL gentamicin.[2] Stocks of all strains were kept in LB with 30% glycerol at –80°C.

A minimal medium was used to promote production of serratiochelin with glucose as the sole carbon source. The minimal medium consisted of (per L): 5.96 g Na_2HPO_4, 3.0 g K_2HPO_4, 1.0 g NH_4Cl, 0.5 g NaCl, 0.058 g $MgSO_4$, and 5.0 g glucose, pH 7.0. A baffled Erlenmeyer flask containing 1 L of the minimal medium was inoculated with an overnight culture of *Serratia sp.* V4, which was incubated for 8 days with shaking at 10°C.

To examine whether serratiochelin is produced in the presence of iron, the minimal medium described above was supplemented with $FeCl_3$ (50 µM), and *Serratia sp.* V4 cultured in a similar fashion as above. After 8 days, the supernatant was analyzed by HPLC-MS (see general procedures).

Purification of Serratiochelin. After culturing *Serratia sp.* V4 in a minimal medium for 8 days, cells were removed by centrifugation and the spent medium passed through a pre-packed C18 Sep-Pak cartridge (2 g, Waters), which had been equilibrated in 10% MeCN (in H_2O). Following application of the spent medium, the cartridge was washed with 10% MeCN and bound material eluted with 50% and 100% MeCN. The 50% fraction contained serratiochelin. This fraction was further purified by HPLC on a Supelco Discovery C18 column (10 μm, 250 mm × 10 mm) using a gradient of 10% MeCN (in H_2O) to 50% MeCN (in H_2O) over 35 min. Both H_2O and MeCN contained 0.1% formic acid. Elution of serratiochelin was monitored at 316 nm.

Marfey's Analysis of Serratiochelin C. A small sample of serratiochelin C was hydrolyzed with 6 N HCl (1 mL) for 6 hours at 110^{o}C. The reaction was cooled to room temperature and dried *in vacuo*. The crude hydrolysis product was re-dissolved in water (100 μL), and to this solution was added a solution of Marfey's reagent (180 μL of a 0.9 mg/mL solution) in acetone, followed by 20 μL of 1 N $NaHCO_3$. The reaction was heated to 40^{o}C for 1 h, then acidified with 30 μL of 1 N HCl and diluted with MeOH (5 mL). Standards were prepared from 0.1 mg of D-Thr or L-Thr using the above procedure. The serratiochelin C Marfey's derivative and standards prepared from D- or L-Thr were analyzed by HPLC-MS using the following gradient: 0-4 minutes, 10% MeCN/H_2O + 0.1% formic acid; 4-24 minutes, linear gradient from 10% MeCN/H_2O + 0.1% formic acid to 50% MeCN/H_2O + 0.1% formic acid. The serratiochelin C Marfey's derivative was also co-injected with either standard. The retention times are as follows: serratiochelin C Marfey's derivative, 18.35 min; L-Thr Marfey's derivative, 18.33 min; D-Thr Marfey's derivative, 19.39 min; serratiochelin C Marfey's derivative spiked with L-Thr Marfey's derivative, 18.38 min; serratiochelin C Marfey's derivative spiked with D-Thr Marfey's derivative, 18.32 min and 19.39 min (see Figure S4).

Generation of the Serratiochelin-Ga or -Fe Complexes. The Ga- and Fe-complexes of serratiochelin A (**3**) were generated as previously described with minor modifications. Briefly, serratiochelin A (1 mg/mL, 1 mL total volume) was transferred to a 4 mL scintillation vial equipped with a stir bar. A 20-fold excess of solid $GaBr_3$ or $FeCl_3$ was slowly added to the solution over 10 min. The mixture was stirred for an additional 10 min at room temperature and the mixture incubated at 4˚C overnight. Excess $GaBr_3$ or $FeCl_3$ was removed by centrifugation followed by filtering the supernatant through a 0.2 μm syringe filter. The filtrate was then analyzed by HPLC-MS as described above (see general procedures).

Identification of the Serratiochelin Biosynthetic Cluster. To find the gene cluster responsible for serratiochelin biosynthesis, we first searched for a 2,3-dihydroxybenzoate-AMP ligase (EntE), which acylates the phosphopantetheine group of the aryl carrier protein EntB with 2,3-dihydroxybenzoate and is usually involved in the biosyntheses of siderophores that contain hydroxybenzoyl or dihydroxybenzoyl moieties.[3] By aligning the sequences for 2,3-dihydroxybenzoate-AMP ligases from *Serratia proteomaculans* 568 (Accession number NC_009832), *Serratia odorifera* DSM4582 (Accession number NZ_ADBY00000000) and *Serratia odorifera* 4Rx13SODc (Accession number NZ_ADBX00000000), a set of degenerate primers were designed: forward primer 5'-CCGCCCCACTCTTCMATCGCCTG-3' and reverse primer 5'- CTGGTGCAGYTGGGCAACGTGG-3'. These primers were used to PCR-amplify the EntE-like gene from the genome of *Serratia sp.* V4. The product was purified by gel extraction and sequenced (Dana-Farber/Harvard Cancer Center DNA Resource Core, Boston, USA). The resulting 659 base pair (bp) sequence was aligned against the genomic sequence of *Serratia sp.* V4 (GenoTech Corporation, South Korea), and the open reading frames upstream and downstream of this fragment were then characterized using bioinformatic tools.

Bioinformatic Analysis. The nucleotide and amino acid sequences of the genes in the serratiochelin cluster were analyzed using the software package DNASTAR - Lasergene® 8.

Putative functional assignments of open reading frames were made by homology searches with BLAST[4], FASTA,[5] and InterProScan,[6] and subsequent comparison of the query with sequences of characterized enzymes. Domain search and organization analysis was performed using ClustScan,[7] Web-CD Search Tool,[8-10] SMART,[11,12] NRPSpredictor,[13] NRPSpredictor2,[14] and the PKS/NRPS Analysis Web-site.[15]

Generation of Phylogenetic Trees. Homology searches of the *Serratia sp.* V4 proteins, SchA, SchB, SchF1, SchF2, SchF3, and SchH were carried out using BLAST[4] and the entire NCBI database. The top 100 hits, corresponding to the 100 sequences in the database with the highest homology to the query sequence, were used for phylogenetic comparisons. Unrooted trees displaying phylogenetic distance between the sequences were generated using NCBI's BLAST Tree View Neighbor Joining algorithm.[16] The Maximum Sequence Difference allowed was 0.85 and the evolutionary distance model according to Grishin *et al.*[17] was used. The online tool iTOL[18,19] was used to view and format the trees.

Construction of Targeted Mutants in *Serratia sp.* V4. Genes were selected for disruption by Campbell insertions, using a suicide vector with an R6K origin. This vector was derived from transposon pBTK30 as follows: the genes encoding for both the Mariner C9 transposase and ampicillin resistance were removed by digestion with restriction enzymes SpeI and StuI. In order to allow single crossover events, fragments of 600 to 1000 bp of DNA internal to the coding region of the genes were subcloned into the suicide vector. A 723 bp long SpeI-StuI fragment of *schC*, a 640 bp fragment of *schE*, a 662 bp fragment of *schB,* a 615 bp fragment of *schF0*, a 985 bp fragment of *schF2*, and a 762 bp fragment of *schH* were amplified using the primers shown in Table S6 and cloned into the vector to generate pSC30C, pSC30E, pSC30B, pSC30F0, pSC30F and pSC30H, respectively. The constructs were confirmed by PCR and by DNA sequencing.

The plasmids generated in this fashion were transformed into *E. coli* S17-1 λPir, followed by mobilization into *Serratia sp.* V4. All insertional mutants, by a single crossover event, were selected on LB agar supplemented with 10 μg/mL gentamicin. The integration of the plasmid was confirmed by PCR, using two different sets of primers (Table S7): one set with a primer homologous to the chromosomal region upstream of the plasmid insertion site and a primer homologous to the sequence of the plasmid, and another set with primers homologous to the chromosomal region upstream and downstream of the plasmid.

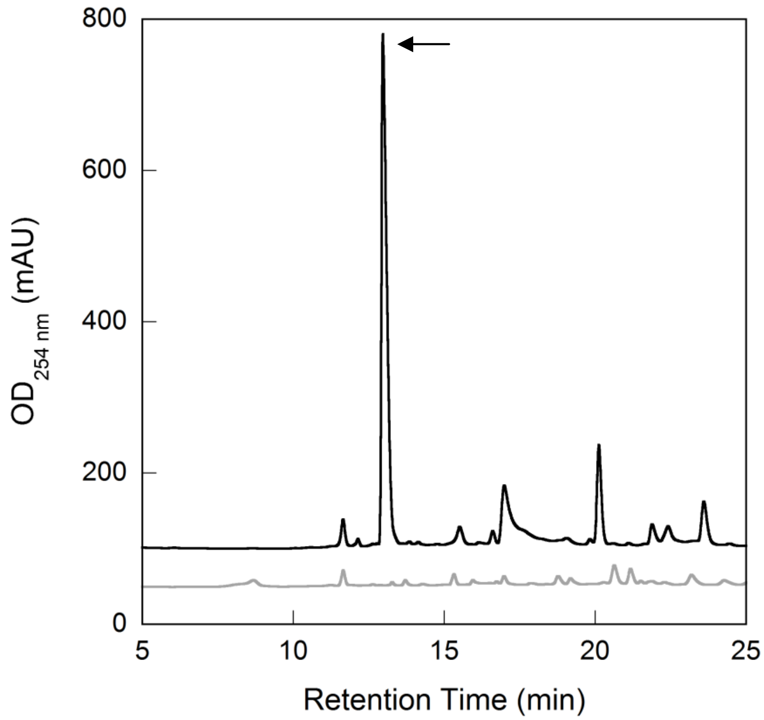


Figure S1. HPLC-MS analysis of the spent medium of *Serratia sp.* V4 grown in the absence (black) or presence (gray trace) of 50 µM $FeCl_3$. The peak corresponding to serratiochelin, only produced in the absence of iron, is highlighted with an arrow.

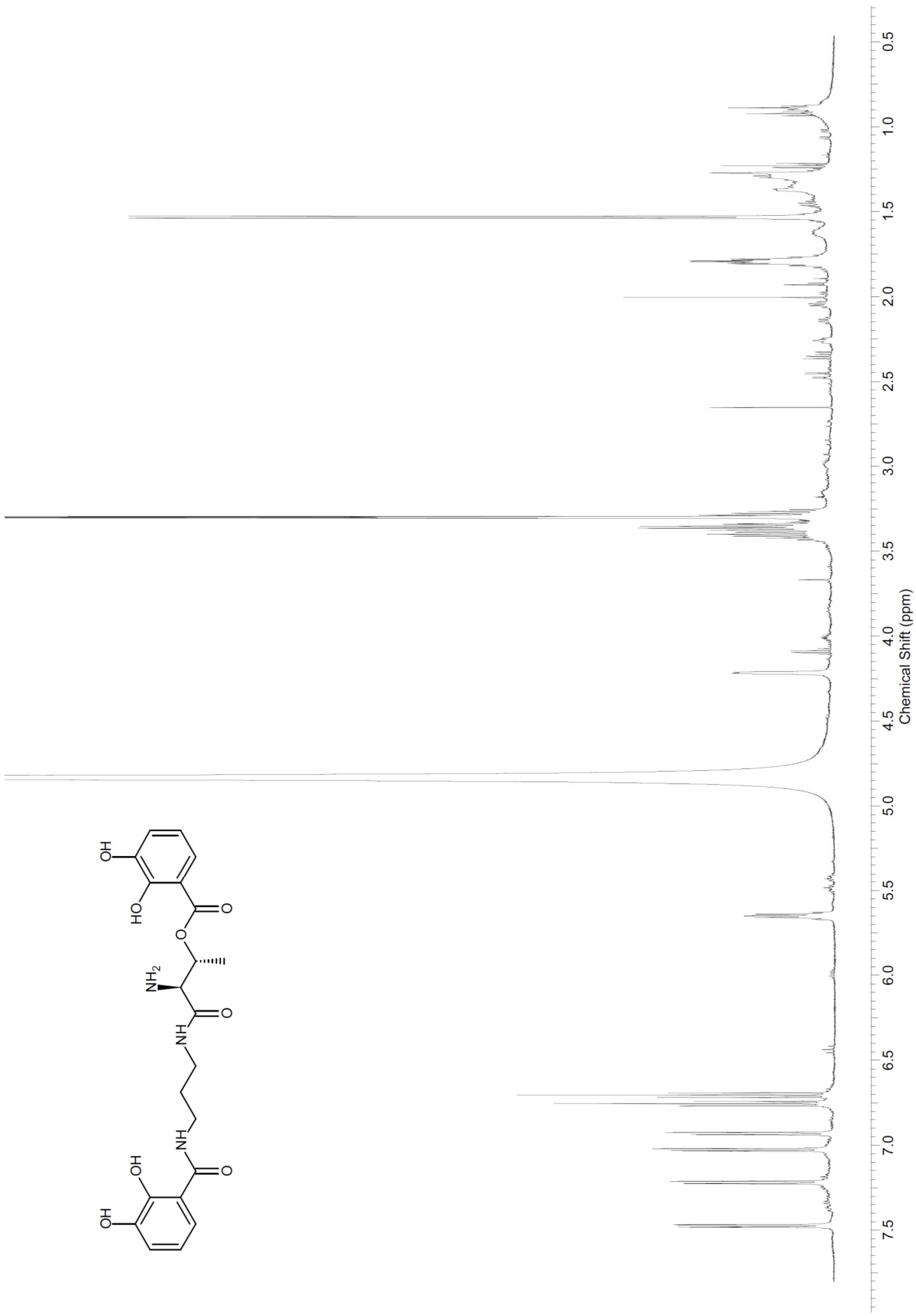
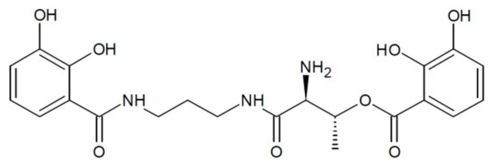
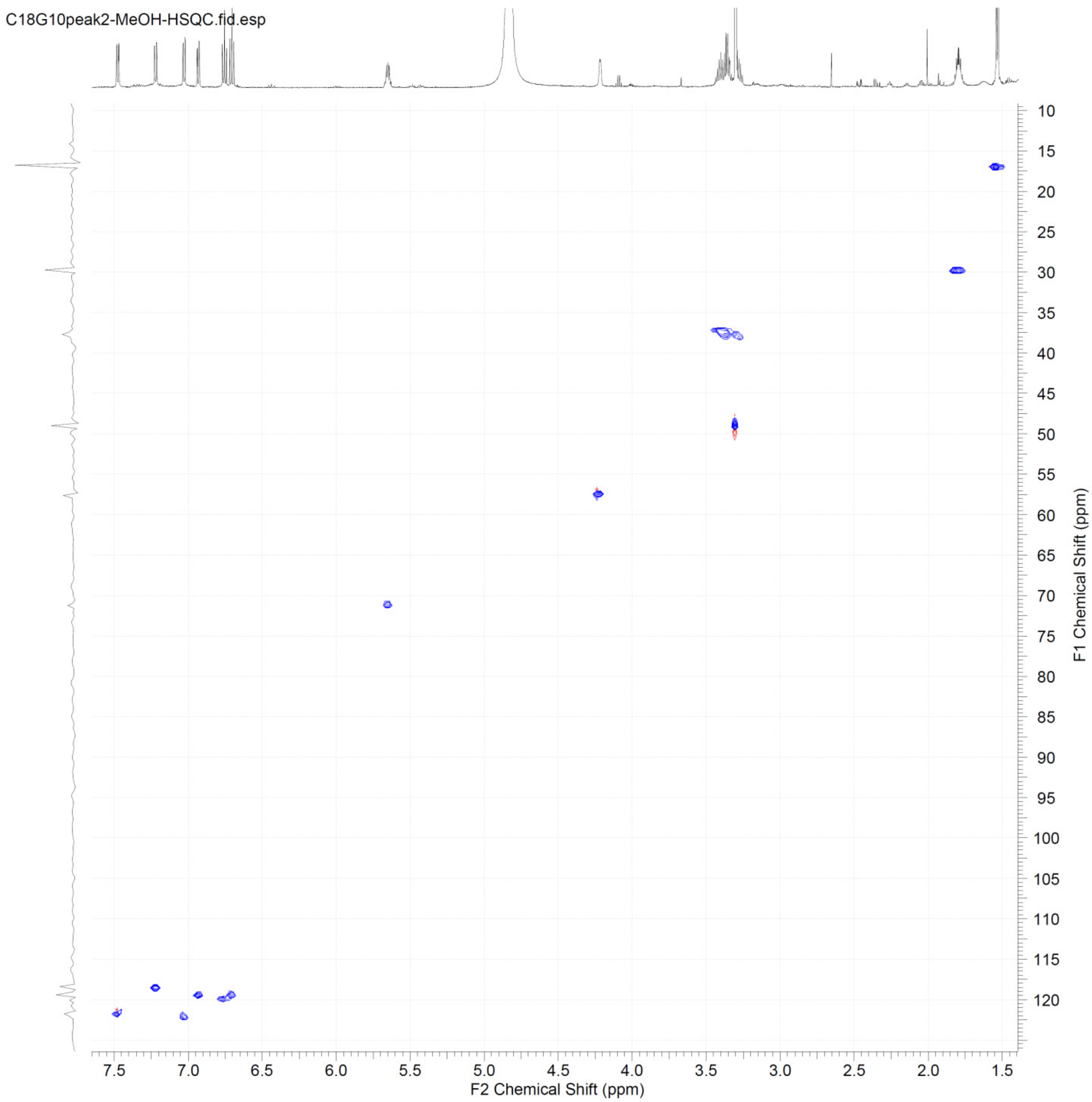
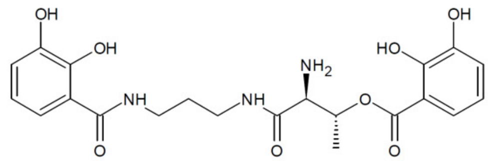


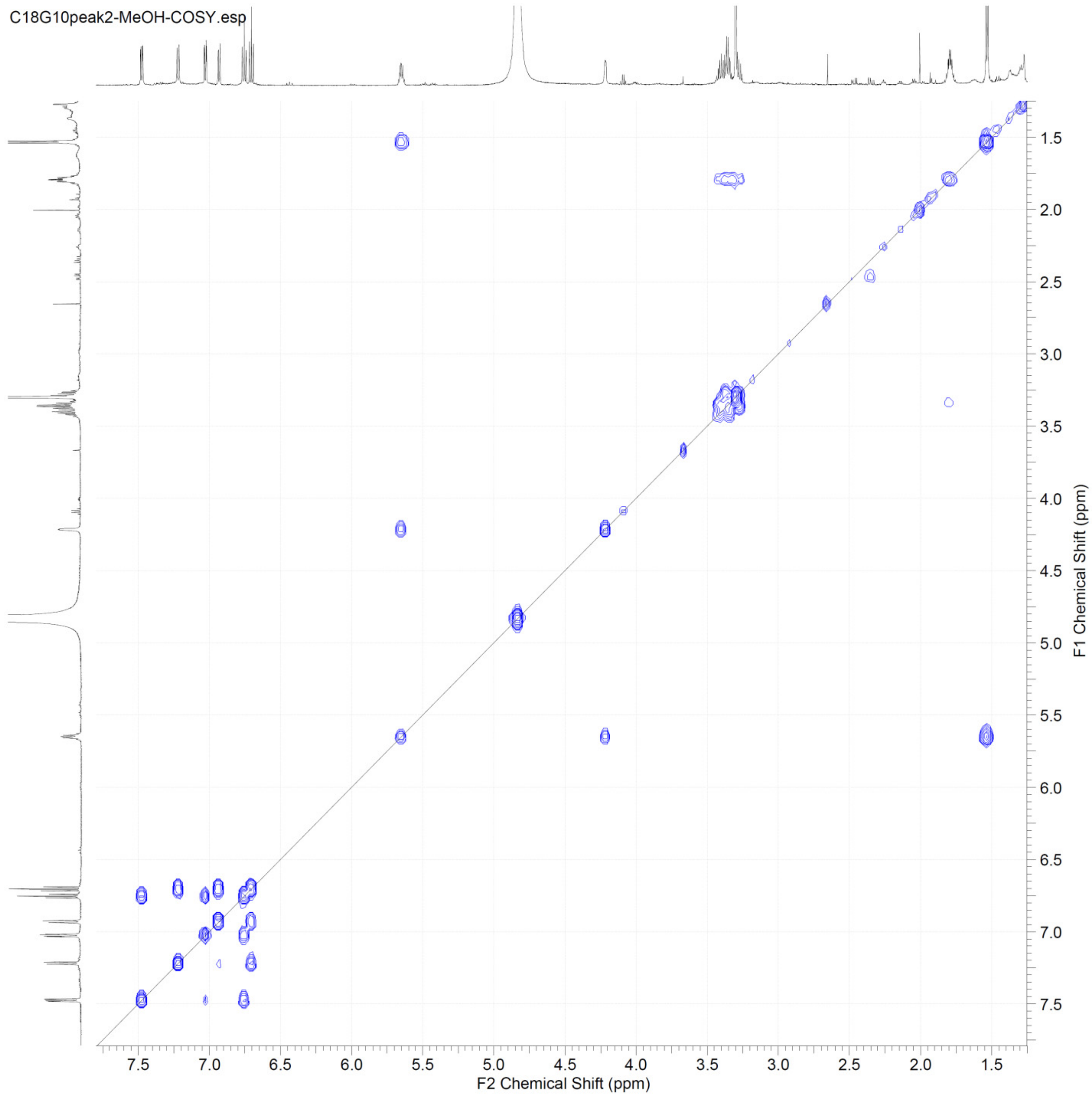
Figure S2. 1D and 2D NMR spectra for serratiochelin C (**1**) in deuterated methanol on pages S7–S11. Shown are ^{1}H (page S8), gHSQC (page S9), gCOSY (page S10), and gHMBC (page S11) spectra, and a magnified view of the downfield region of the gHMBC spectrum (page S12).

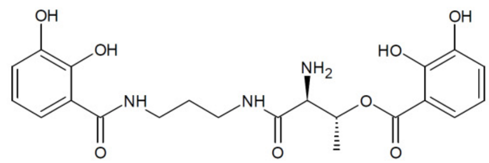


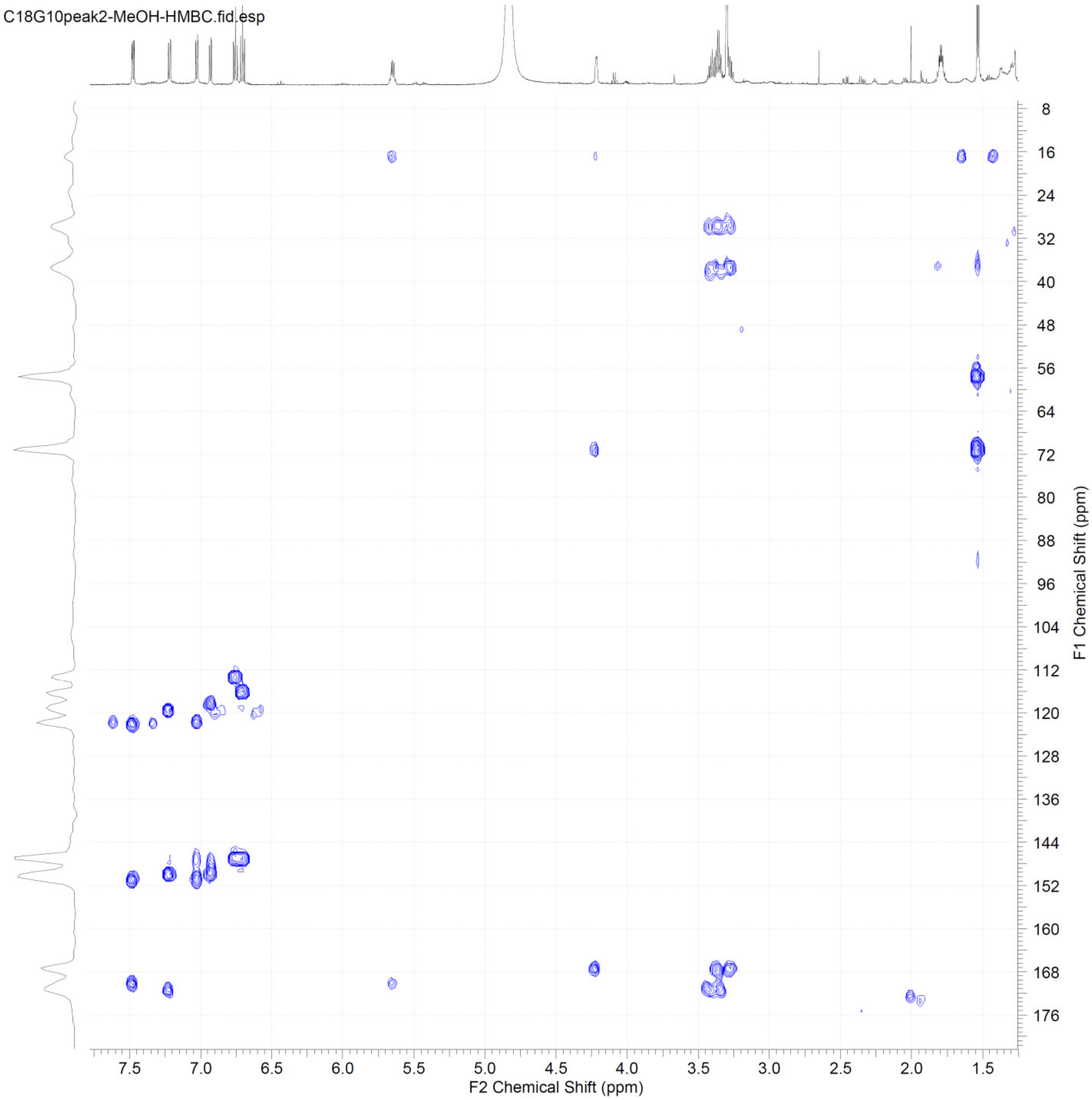


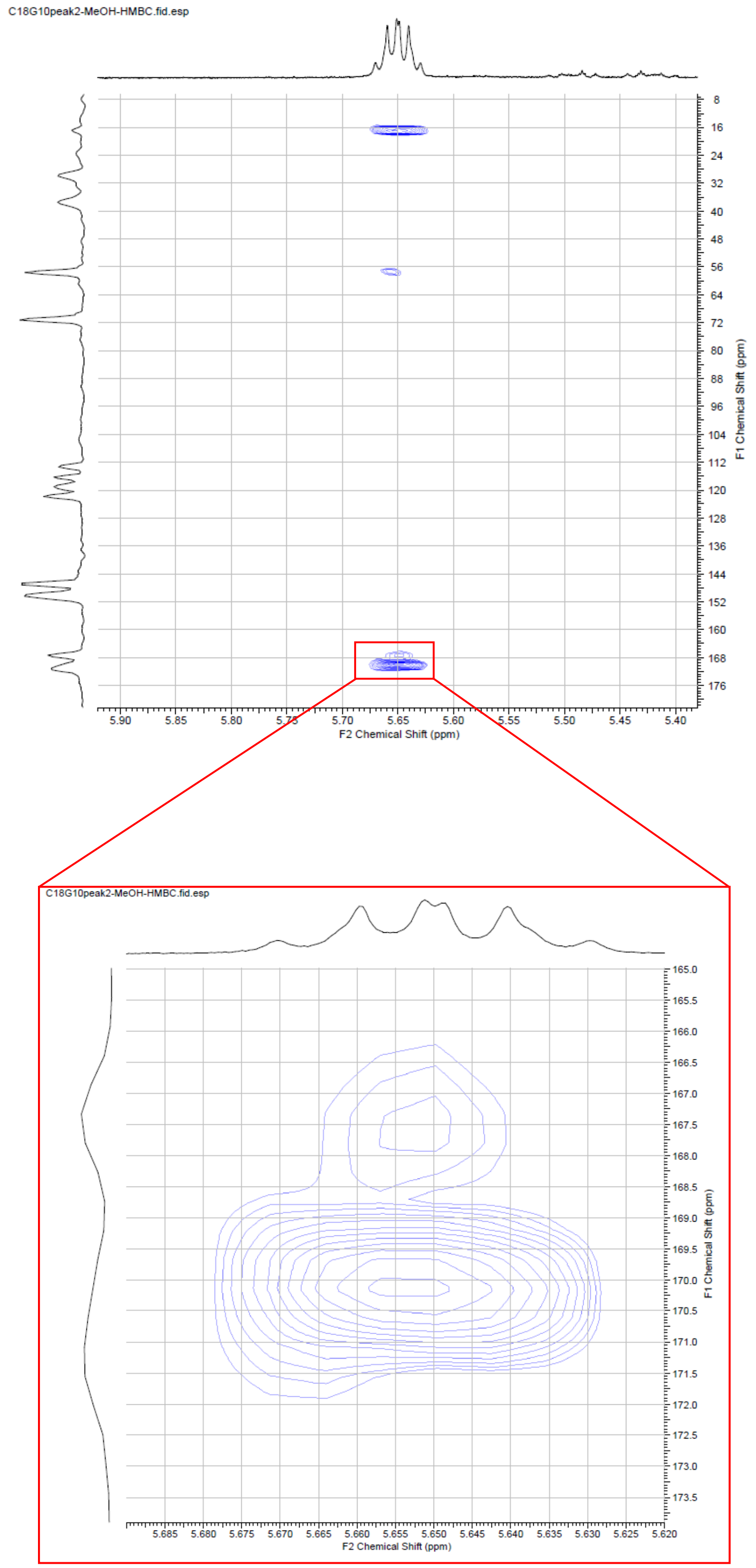
C18G10peak2-MeOH-HSQC.fid.esp
F2 Chemical Shift (ppm)
F1 Chemical Shift (ppm)


OH
OH
NH
NH
NH_2
O
HO
OH
O
O
O


C18G10peak2-MeOH-COSY.esp
F2 Chemical Shift (ppm)
F1 Chemical Shift (ppm)




C18G10peak2-MeOH-HMBC.fid.esp
F2 Chemical Shift (ppm)
F1 Chemical Shift (ppm)



C18G10peak2-MeOH-HMBC.fid.esp
F2 Chemical Shift (ppm)
F1 Chemical Shift (ppm)
C18G10peak2-MeOH-HMBC.fid.esp
F2 Chemical Shift (ppm)
F1 Chemical Shift (ppm)

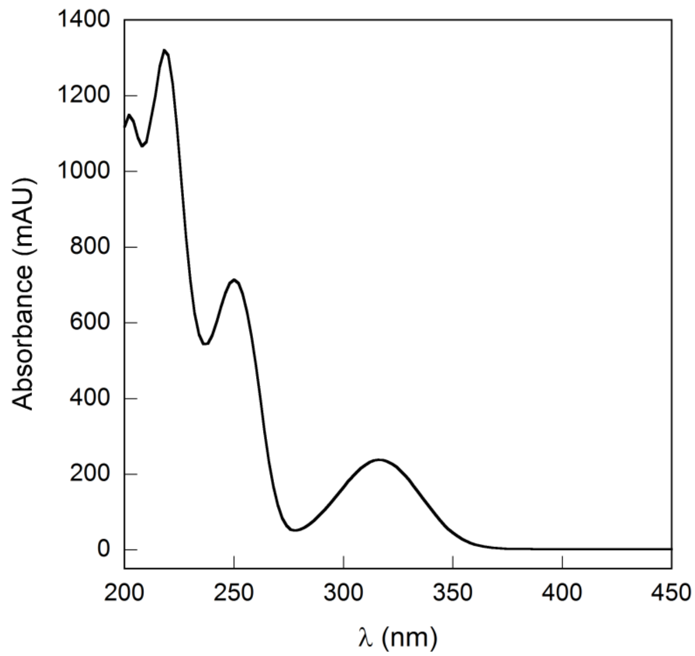


Figure S3. UV-visible spectrum of serratiochelin C (**1**) in H_2O/MeCN (1:1) containing 0.1% formic acid.

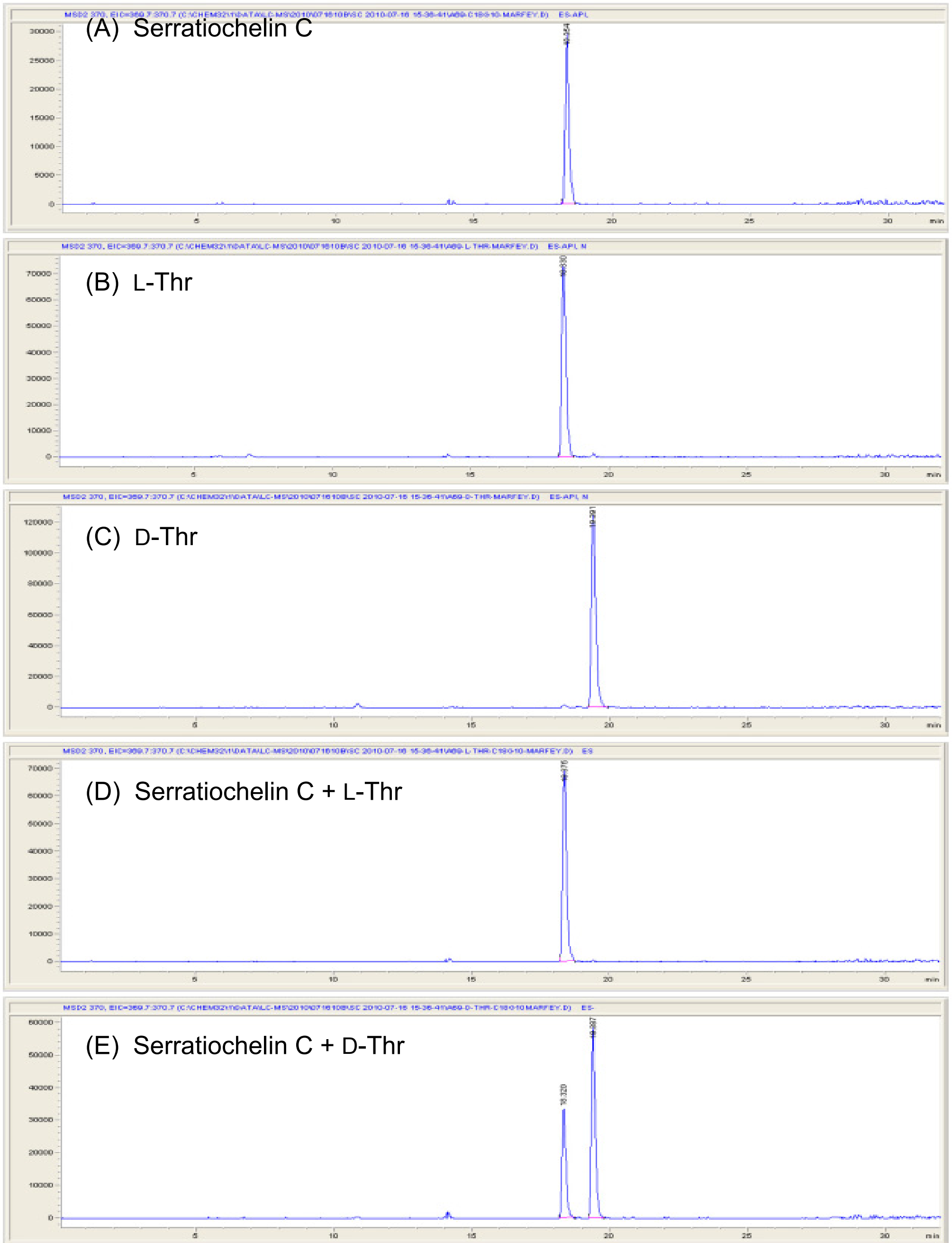


Figure S4. HPLC-MS traces for Marfey's analysis of serratiochelin C and D- or L-Thr. The mass ion for the Marfey's derivative of serratiochelin C (A), L-Thr (B), D-Thr (C), serratiochelin C spiked with the L-Thr derivative (D), and serratiochelin C spiked with the D-Thr derivative are shown. The data show that the Thr moiety in serratiochelin is L-Thr.

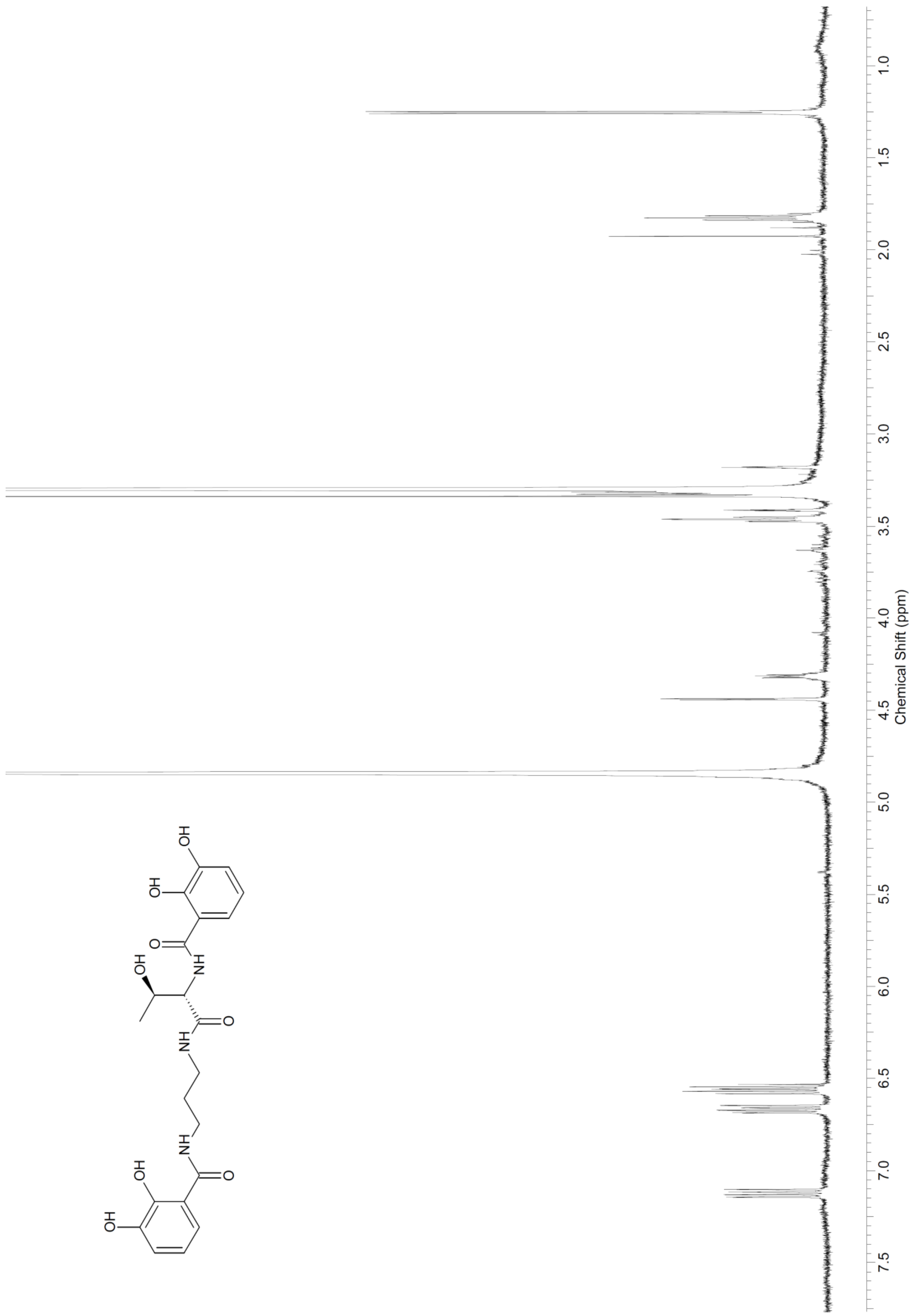
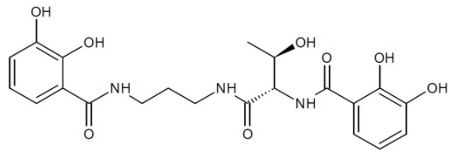
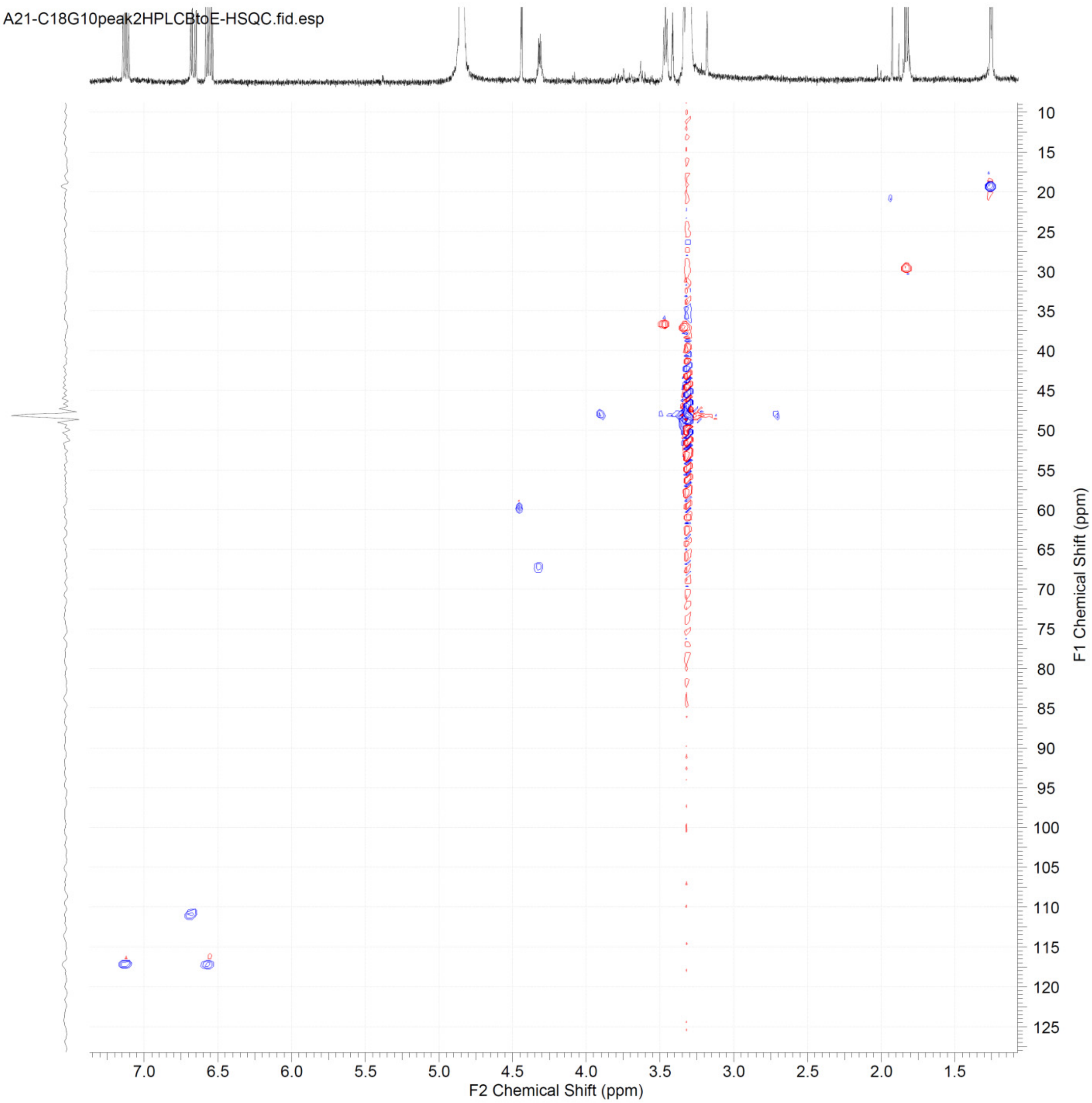
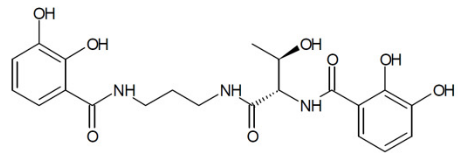
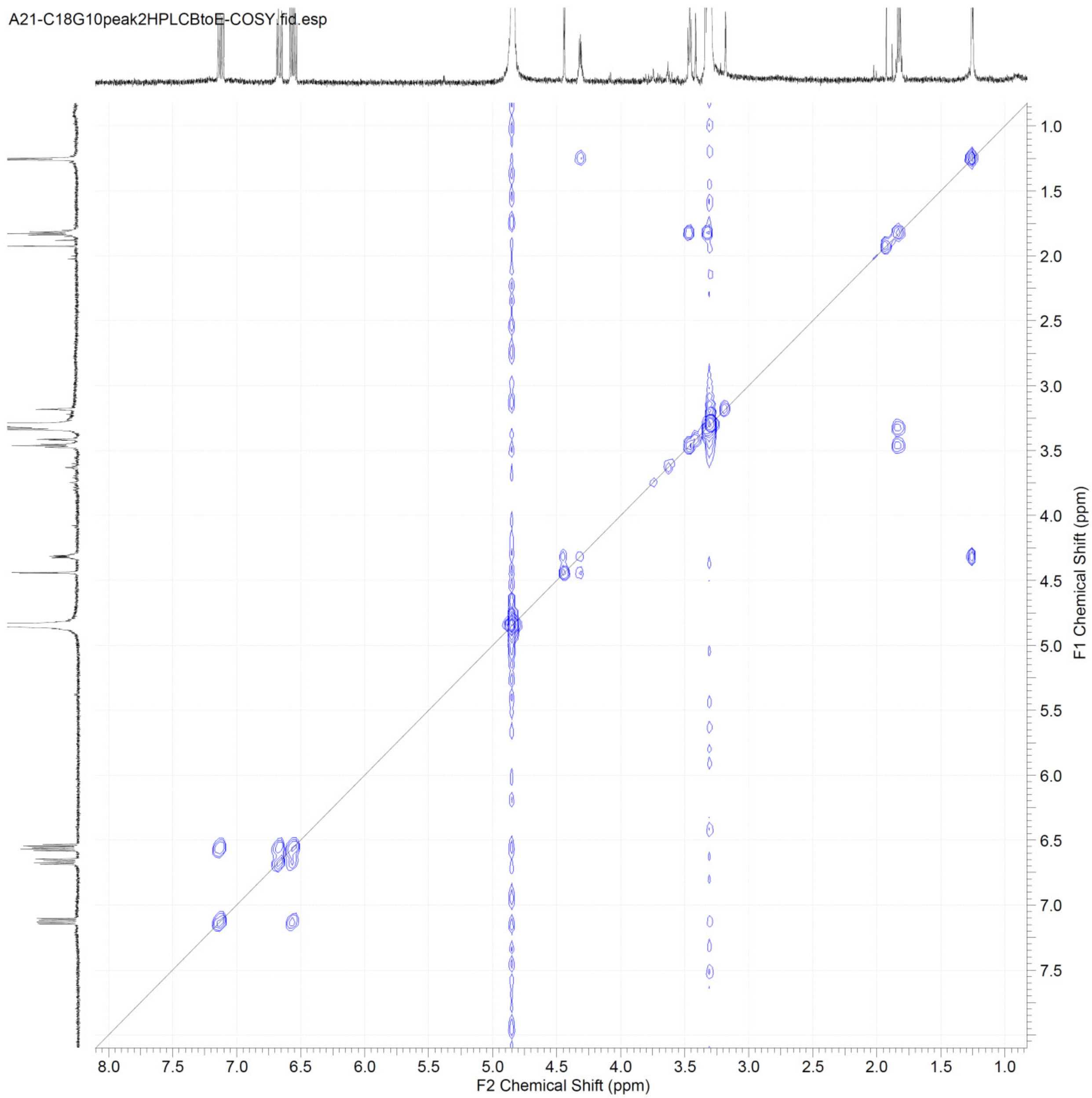


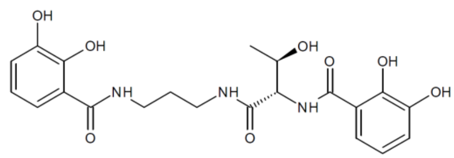
Figure S5. 1D and 2D NMR spectra for serratiochelin B (**2**) in deuterated methanol on pages S12–S15. Shown are ^{1}H (page S15), gHSQC (page S16), gCOSY (page S17), and gHMBC (page S18) spectra.


OH
OH
NH
NH
OH
O
OH
NH
OH
O
O

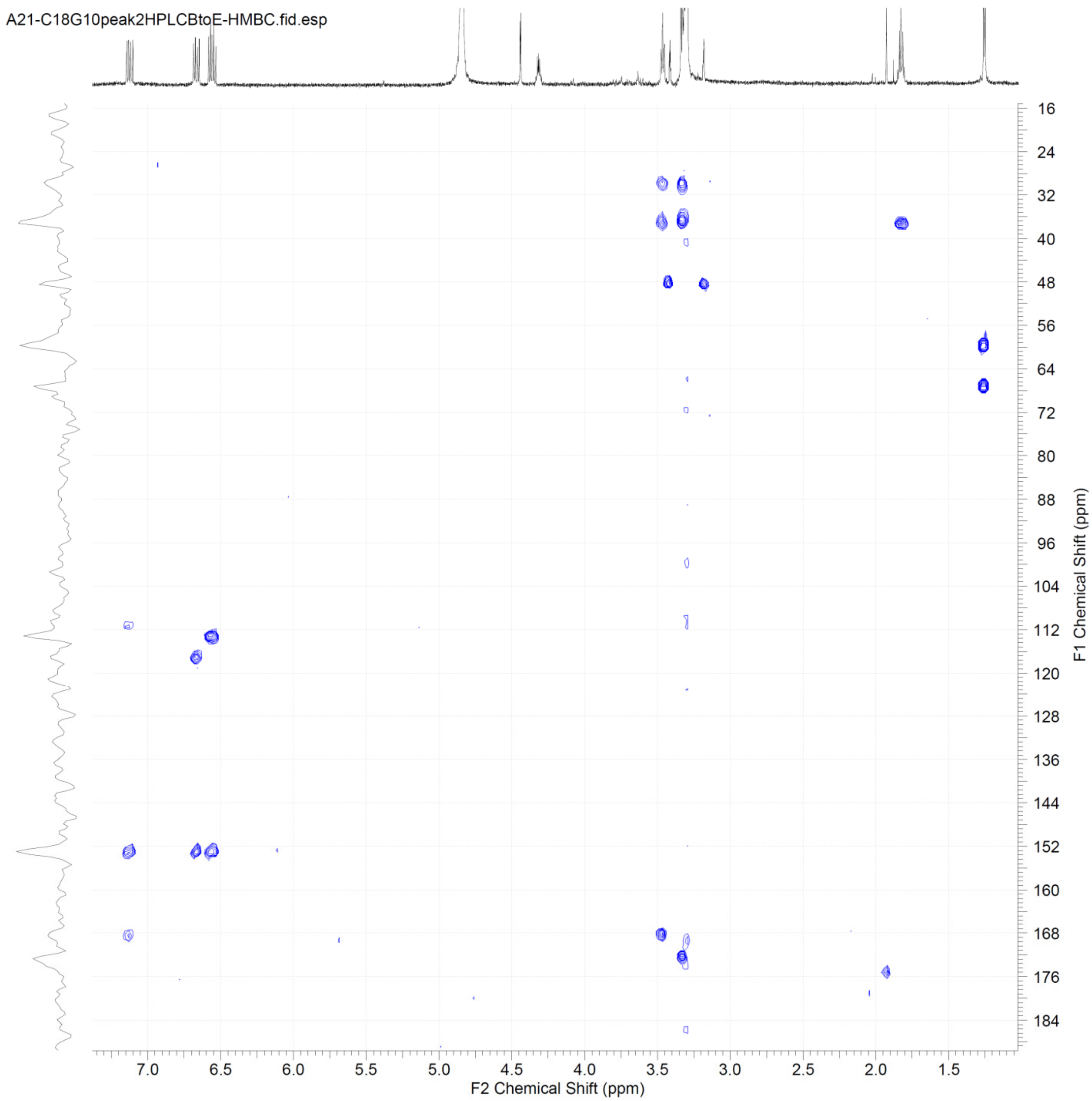

A21-C18G10peak2HPLCBtoE-HSQC.fid.esp
F2 Chemical Shift (ppm)
F1 Chemical Shift (ppm)




A21-C18G10peak2HPLCBtoE-COSY.fid.esp
F2 Chemical Shift (ppm)
F1 Chemical Shift (ppm)



OH
OH
NH
NH
OH
O
OH
NH
OH
O
O



A21-C18G10peak2HPLCBtoE-HMBC.fid.esp
F2 Chemical Shift (ppm)
F1 Chemical Shift (ppm)

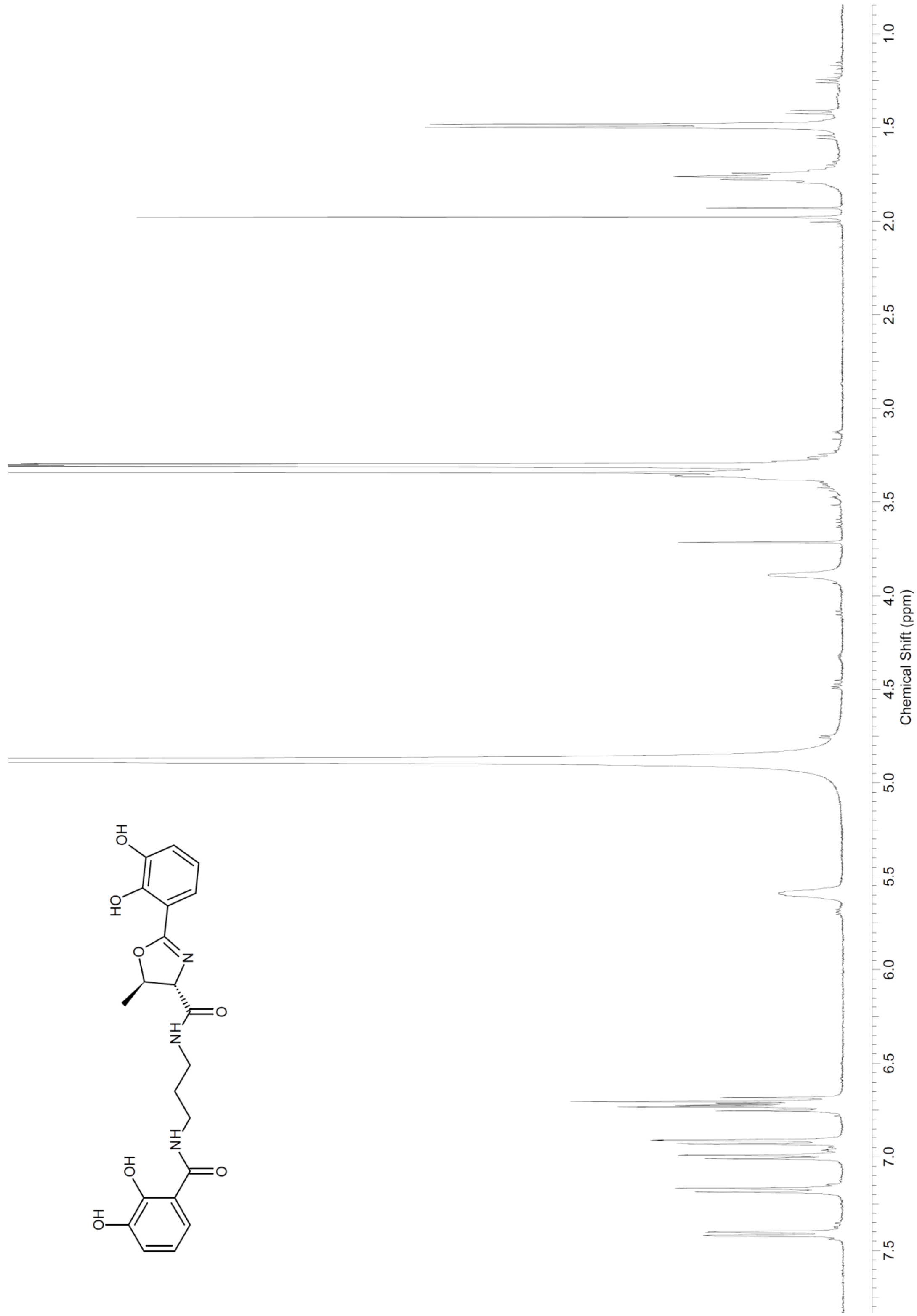


Figure S6. ^{1}H spectrum of serratiochelin A (**3**) in deuterated methanol.

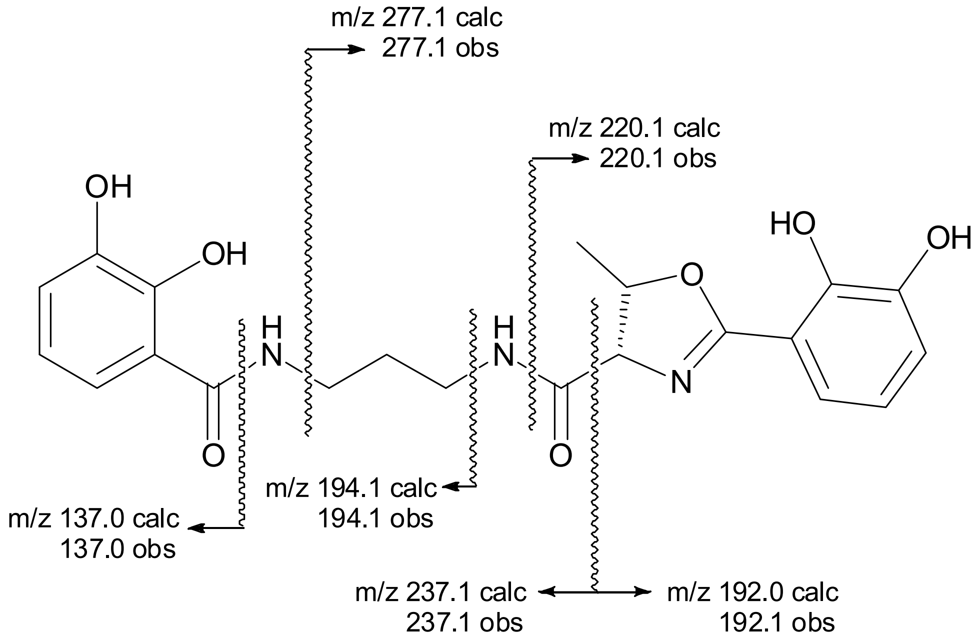


Figure S7. Results from tandem MS analysis of serratiochelin A (**3**). The fragments obtained along with the calculated and observed masses are shown.

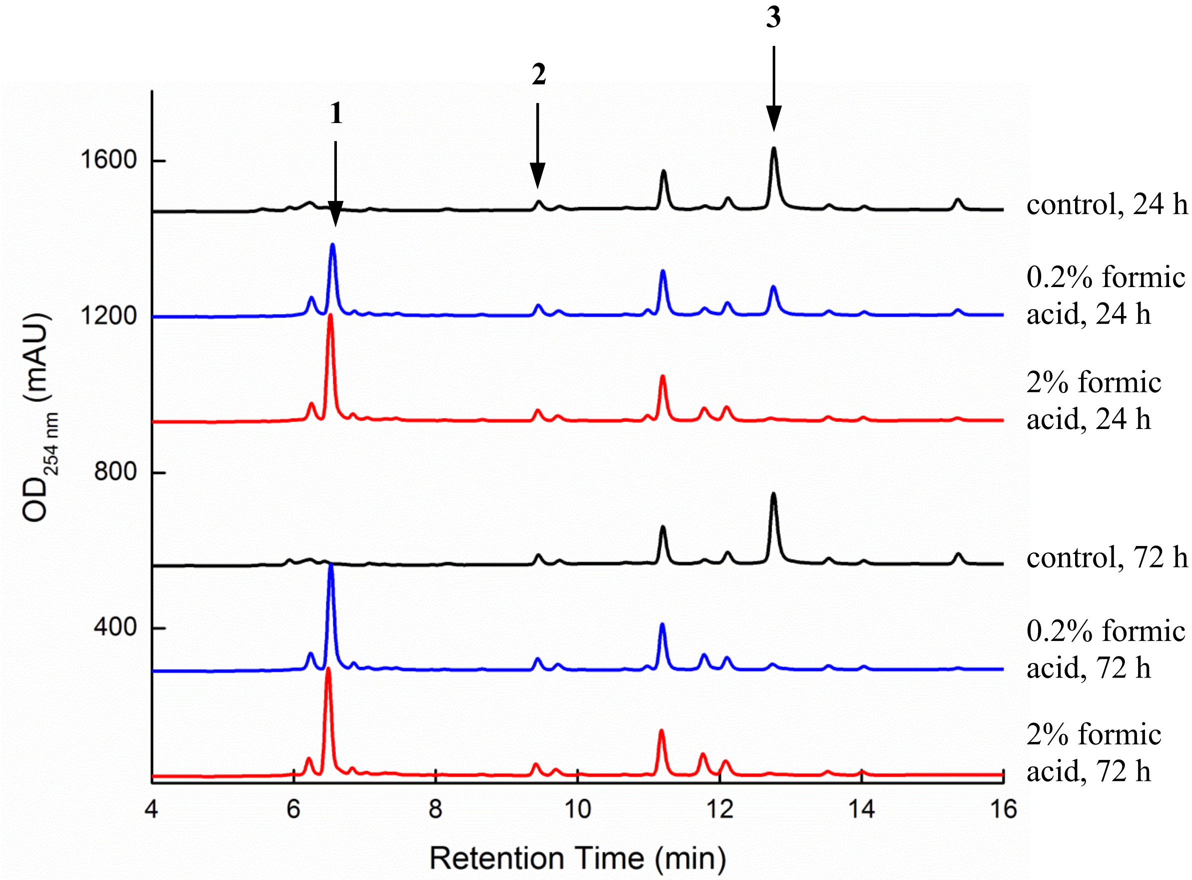


Figure S8. Stability of serratiochelins as a function of formic acid concentration and time of incubation. HPLC-MS analysis of the spent medium of *Serratia sp.* V4 cultures after treatment with no acid (black traces), 0.2% formic acid (blue traces), or 2% formic acid (red traces) for 24 h or 72 h. The peaks corresponding to serratiochelins A (**3**), B (**2**), and C (**1**) are marked. In the presence of acid, and as a function of time of incubation, **3** hydrolyzes to give **1**.

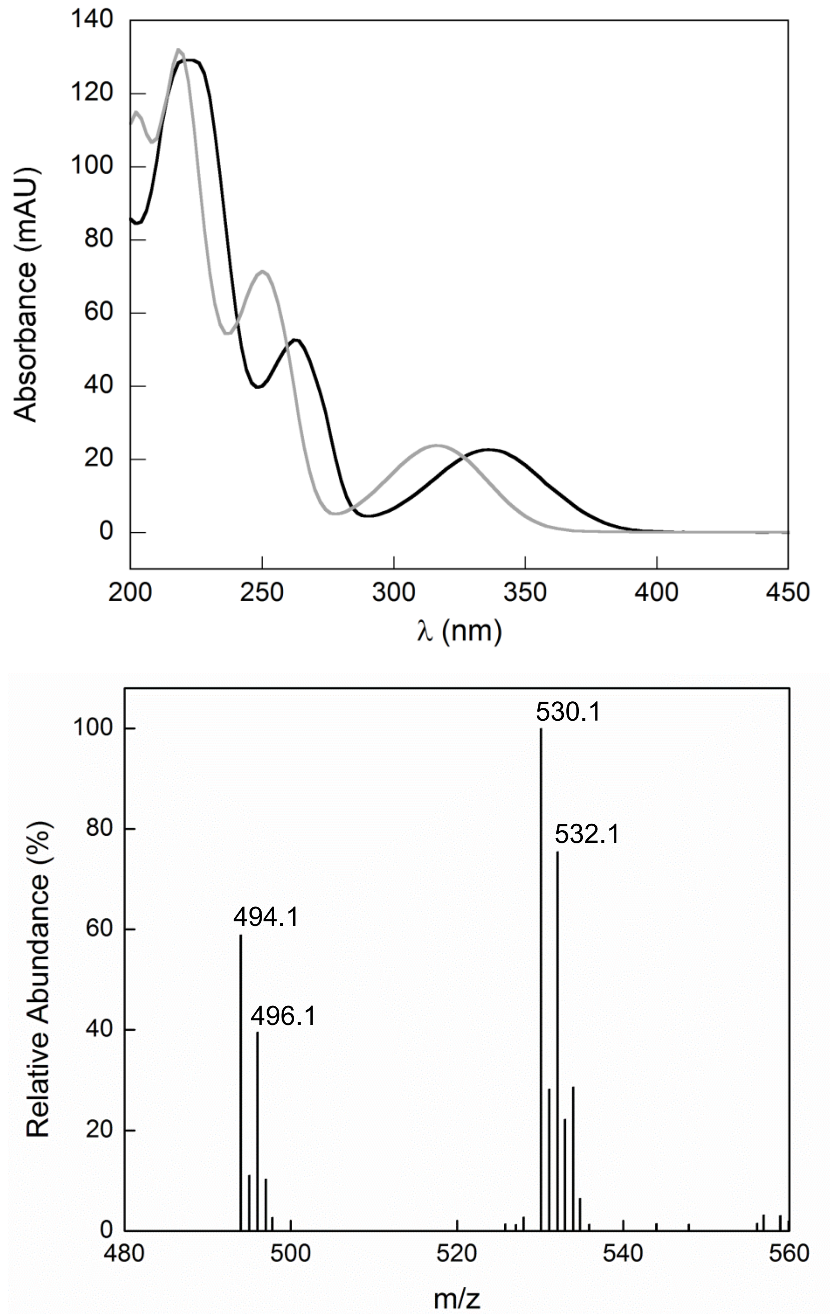


Figure S9. Characterization of Ga-serratiochelin A. (Top) Comparison of the UV-visible spectrum of serratiochelin A (**3**, gray trace) with that of Ga-serratiochelin A (black trace). The spectrum of the Ga-complex is red-shifted relative to the apo-form of **3**. (Bottom) Negative mode ESI-MS spectrum of Ga-serratiochelin A. Both the tetradentate Ga-complex, in which Ga is chelated by three catechol hydroxyl groups and the oxazoline nitrogen ($[M\text{-}H]^-$: calc. 494.0, expt. 494.1), as well as the hexadentate, bis-aqua form of the same complex ($[M\text{-}H]^-$: calc. 530.1, expt. 532.0) are observed. A similar result was obtained with Fe-serratiochelin A (data not shown).

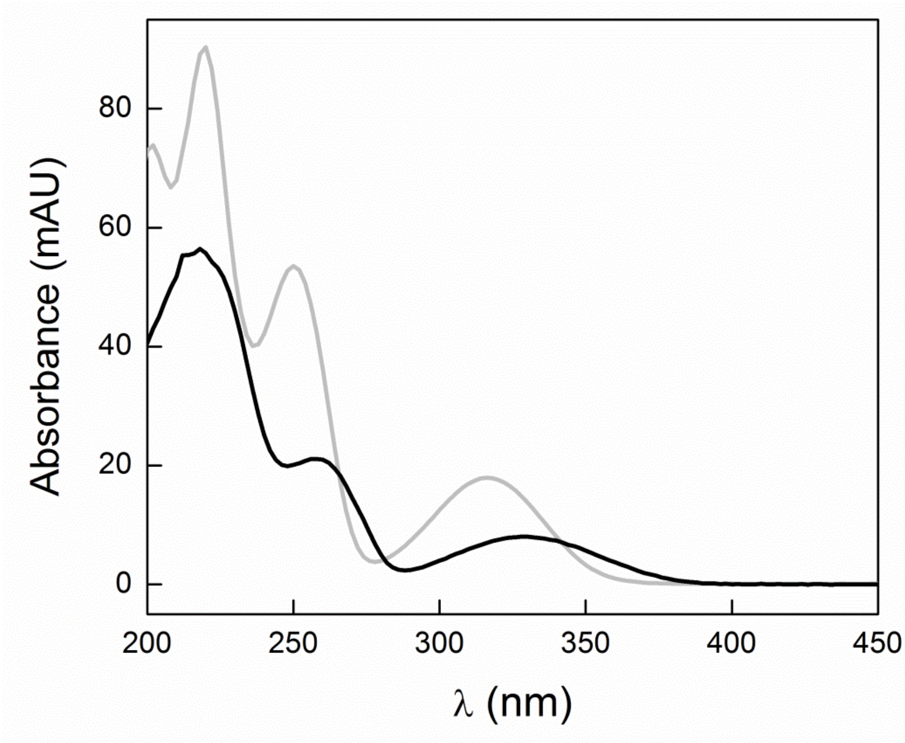
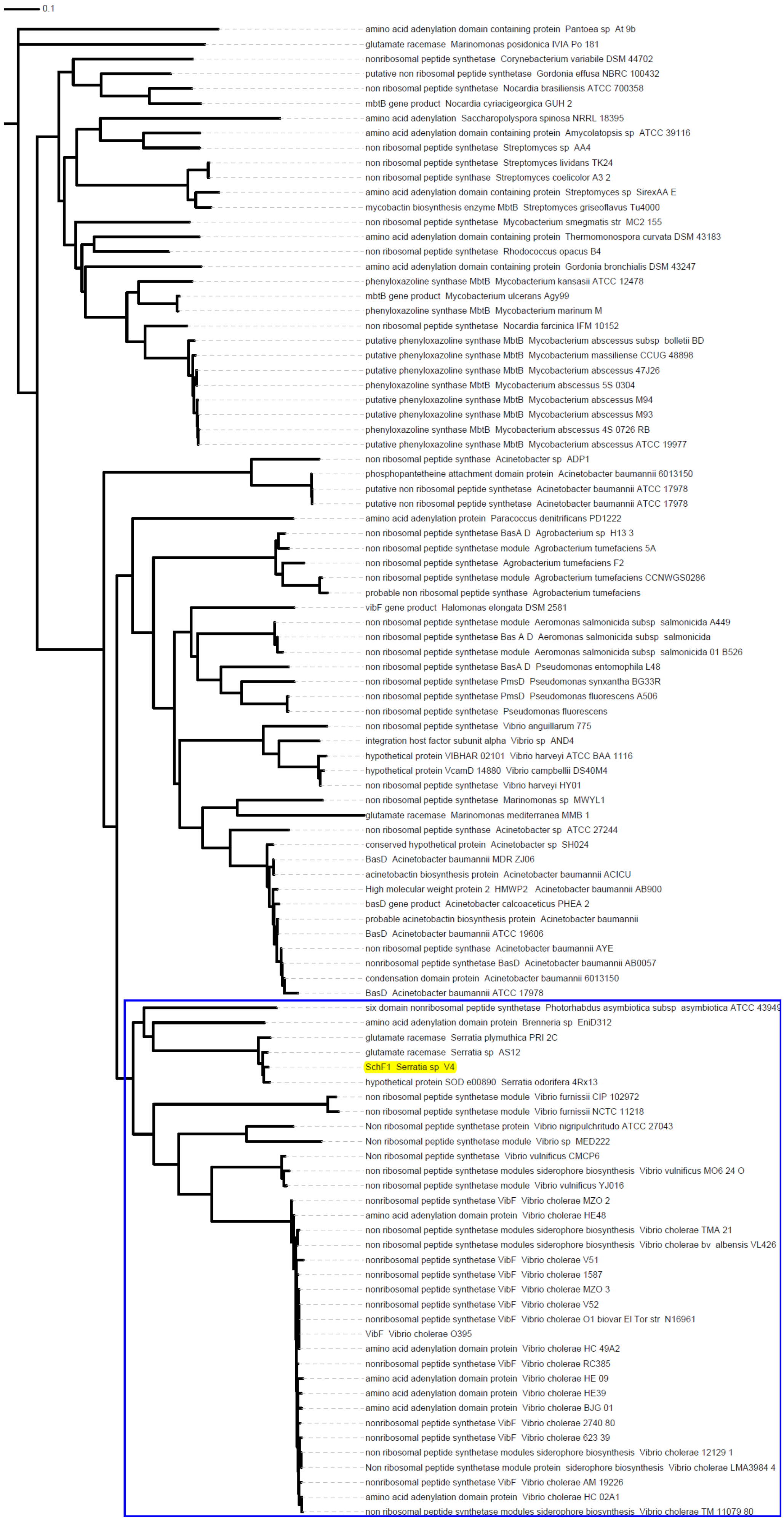
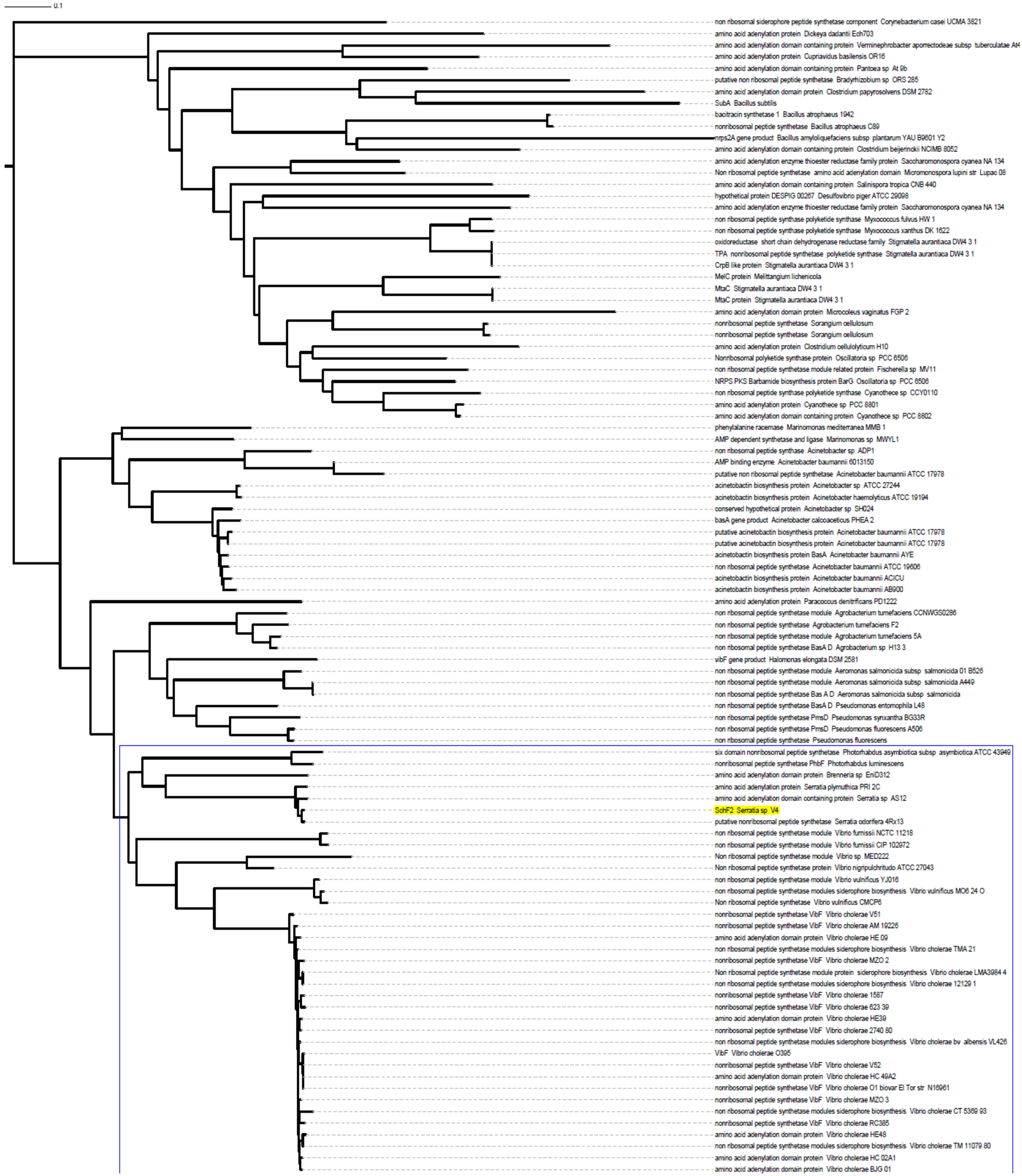


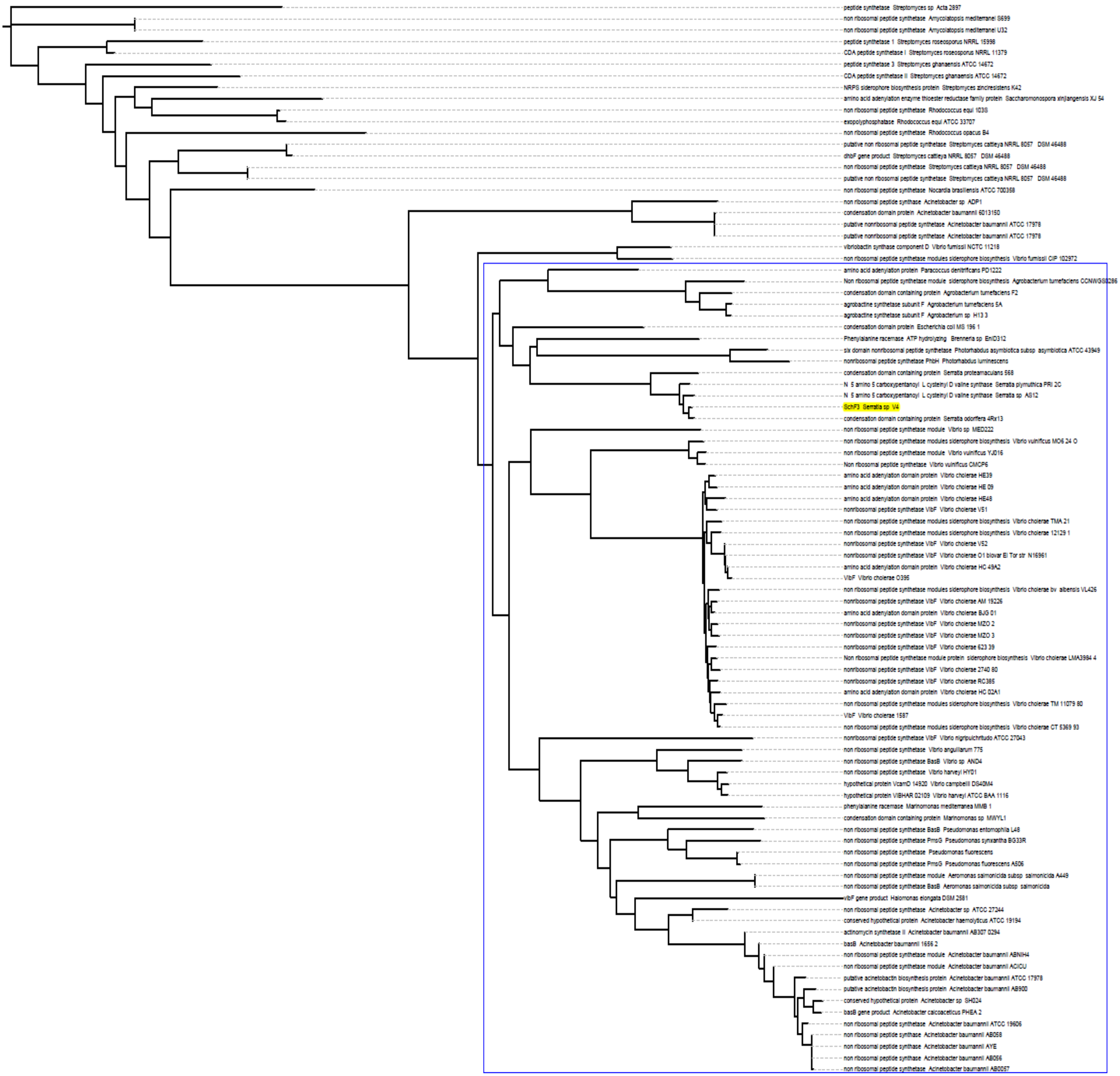
Figure S10. Characterization of Ga-serratiochelin B. Comparison of the UV-visible spectrum of serratiochelin B (**2**, gray trace) with that of Ga-serratiochelin B (black trace). The spectrum of the Ga-complex is red-shifted relative to the apo-form of **2**. Serratiochelin B has a significantly lower affinity for Ga and Fe than serratiochelin A.

Figure S11. Phylogenetic analyses of vibriobactin-like genes in the *sch* cluster. On the next four pages follow phylogenetic trees for SchF1 (page S25), SchF2 (page S26), SchF3 (page S27), and SchH (page S28). The clade incorporating *Serratia* sp. V4 and the *Vibrio spp.* common ancestor is marked in blue (*Vibrio cholera* strains) or in red (*Vibrio furnissii* and *Vibrio nigripulchritudo* strains). The *Serratia sp.* V4 query protein is highlighted in yellow. In all four cases, the proteins in the *sch* cluster grouped in the same clade as the homologous proteins involved in vibriobactin biosynthesis in a number of *Vibrio* species. Proteins SchF1, SchF2 and SchF3 share common ancestry with *Vibrio spp.* strains. SchH (page S27) is predicted to share a more recent common ancestor with its orthologs in *V. furnissii* and *V. nigripulchritudo*, but also displays a high level of homology with *V. cholera* strains, sharing the same clade at a higher level.


0.1
amino acid adenylation domain containing protein Pantoea sp At 9b
glutamate racemase Marinomonas posidonica IVIA Po 181
nonribosomal peptide synthetase Corynebacterium variabile DSM 44702
putative non ribosomal peptide synthetase Gordonia effusa NBRC 100432
non ribosomal peptide synthetase Nocardia brasiliensis ATCC 700358
mbtB gene product Nocardia cyriacigeorgica GUH 2
amino acid adenylation Saccharopolyspora spinosa NRRL 18395
amino acid adenylation domain containing protein Amycolatopsis sp ATCC 39116
non ribosomal peptide synthetase Streptomyces sp AA4
non ribosomal peptide synthetase Streptomyces lividans TK24
non ribosomal peptide synthase Streptomyces coelicolor A3 2
amino acid adenylation domain containing protein Streptomyces sp SirexAA E
mycobactin biosynthesis enzyme MbtB Streptomyces griseoflavus Tu4000
non ribosomal peptide synthetase Mycobacterium smegmatis str MC2 155
amino acid adenylation domain containing protein Thermomonospora curvata DSM 43183
non ribosomal peptide synthetase Rhodococcus opacus B4
amino acid adenylation domain containing protein Gordonia bronchialis DSM 43247
phenyloxazoline synthase MbtB Mycobacterium kansasii ATCC 12478
mbtB gene product Mycobacterium ulcerans Agy99
phenyloxazoline synthase MbtB Mycobacterium marinum M
non ribosomal peptide synthetase Nocardia farcinica IFM 10152
putative phenyloxazoline synthase MbtB Mycobacterium abscessus subsp bolletii BD
putative phenyloxazoline synthase MbtB Mycobacterium massiliense CCUG 48898
putative phenyloxazoline synthase MbtB Mycobacterium abscessus 47J26
phenyloxazoline synthase MbtB Mycobacterium abscessus 5S 0304
putative phenyloxazoline synthase MbtB Mycobacterium abscessus M94
putative phenyloxazoline synthase MbtB Mycobacterium abscessus M93
phenyloxazoline synthase MbtB Mycobacterium abscessus 4S 0726 RB
putative phenyloxazoline synthase MbtB Mycobacterium abscessus ATCC 19977
non ribosomal peptide synthase Acinetobacter sp ADP1
phosphopantetheine attachment domain protein Acinetobacter baumannii 6013150
putative non ribosomal peptide synthetase Acinetobacter baumannii ATCC 17978
putative non ribosomal peptide synthetase Acinetobacter baumannii ATCC 17978
amino acid adenylation protein Paracoccus denitrificans PD1222
non ribosomal peptide synthetase BasA D Agrobacterium sp H13 3
non ribosomal peptide synthetase module Agrobacterium tumefaciens 5A
non ribosomal peptide synthetase Agrobacterium tumefaciens F2
non ribosomal peptide synthetase module Agrobacterium tumefaciens CCNWGS0286
probable non ribosomal peptide synthase Agrobacterium tumefaciens
vibF gene product Halomonas elongata DSM 2581
non ribosomal peptide synthetase module Aeromonas salmonicida subsp salmonicida A449
non ribosomal peptide synthetase Bas A D Aeromonas salmonicida subsp salmonicida
non ribosomal peptide synthetase module Aeromonas salmonicida subsp salmonicida 01 B526
non ribosomal peptide synthetase BasA D Pseudomonas entomophila L48
non ribosomal peptide synthetase PmsD Pseudomonas synxantha BG33R
non ribosomal peptide synthetase PmsD Pseudomonas fluorescens A506
non ribosomal peptide synthetase Pseudomonas fluorescens
non ribosomal peptide synthetase Vibrio anguillarum 775
integration host factor subunit alpha Vibrio sp AND4
hypothetical protein VIBHAR 02101 Vibrio harveyi ATCC BAA 1116
hypothetical protein VcamD 14880 Vibrio campbellii DS40M4
non ribosomal peptide synthetase Vibrio harveyi HY01
non ribosomal peptide synthetase Marinomonas sp MWYL1
glutamate racemase Marinomonas mediterranea MMB 1
non ribosomal peptide synthase Acinetobacter sp ATCC 27244
conserved hypothetical protein Acinetobacter sp SH024
BasD Acinetobacter baumannii MDR ZJ06
acinetobactin biosynthesis protein Acinetobacter baumannii ACICU
High molecular weight protein 2 HMWP2 Acinetobacter baumannii AB900
basD gene product Acinetobacter calcoaceticus PHEA 2
probable acinetobactin biosynthesis protein Acinetobacter baumannii
BasD Acinetobacter baumannii ATCC 19606
non ribosomal peptide synthase Acinetobacter baumannii AYE
nonribosomal peptide synthetase BasD Acinetobacter baumannii AB0057
condensation domain protein Acinetobacter baumannii 6013150
BasD Acinetobacter baumannii ATCC 17978
six domain nonribosomal peptide synthetase Photorhabdus asymbiotica subsp asymbiotica ATCC 43949
amino acid adenylation domain protein Brenneria sp EniD312
glutamate racemase Serratia plymuthica PRI 2C
glutamate racemase Serratia sp AS12
SchF1 Serratia sp V4
hypothetical protein SOD e00890 Serratia odorifera 4Rx13
non ribosomal peptide synthetase module Vibrio furnissii CIP 102972
non ribosomal peptide synthetase module Vibrio furnissii NCTC 11218
Non ribosomal peptide synthetase protein Vibrio nigripulchritudo ATCC 27043
Non ribosomal peptide synthetase module Vibrio sp MED222
Non ribosomal peptide synthetase Vibrio vulnificus CMCP6
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio vulnificus MO6 24 O
non ribosomal peptide synthetase module Vibrio vulnificus YJ016
nonribosomal peptide synthetase VibF Vibrio cholerae MZO 2
amino acid adenylation domain protein Vibrio cholerae HE48
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae TMA 21
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae bv albensis VL426
nonribosomal peptide synthetase VibF Vibrio cholerae V51
nonribosomal peptide synthetase VibF Vibrio cholerae 1587
nonribosomal peptide synthetase VibF Vibrio cholerae MZO 3
nonribosomal peptide synthetase VibF Vibrio cholerae V52
nonribosomal peptide synthetase VibF Vibrio cholerae O1 biovar El Tor str N16961
VibF Vibrio cholerae O395
amino acid adenylation domain protein Vibrio cholerae HC 49A2
nonribosomal peptide synthetase VibF Vibrio cholerae RC385
amino acid adenylation domain protein Vibrio cholerae HE 09
amino acid adenylation domain protein Vibrio cholerae HE39
amino acid adenylation domain protein Vibrio cholerae BJG 01
nonribosomal peptide synthetase VibF Vibrio cholerae 2740 80
nonribosomal peptide synthetase VibF Vibrio cholerae 623 39
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae 12129 1
Non ribosomal peptide synthetase module protein siderophore biosynthesis Vibrio cholerae LMA3984 4
nonribosomal peptide synthetase VibF Vibrio cholerae AM 19226
amino acid adenylation domain protein Vibrio cholerae HC 02A1
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae TM 11079 80



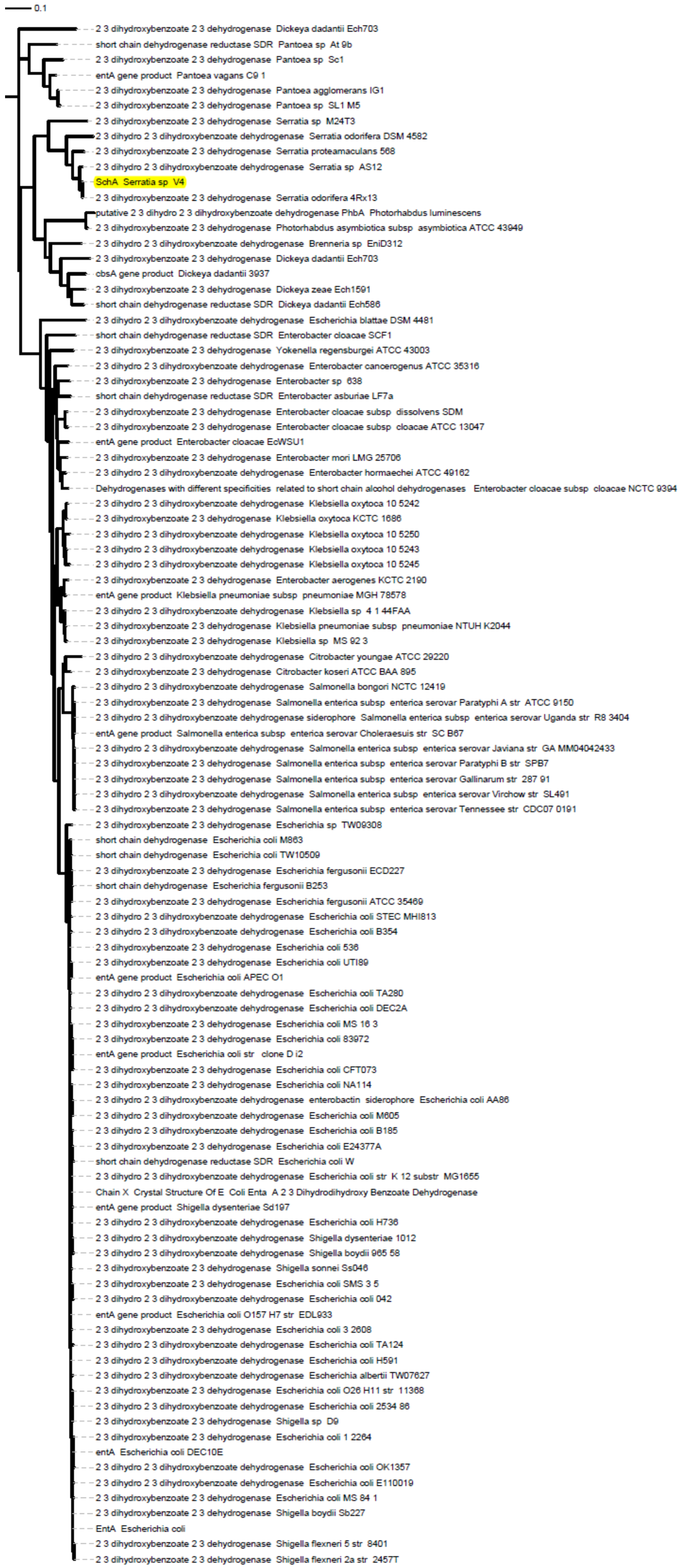
0.1
non ribosomal siderophore peptide synthetase component Corynebacterium casei UCMA 3821
amino acid adenylation protein Dickeya dadantii Ech703
amino acid adenylation domain containing protein Verminephrobacter aporrectodeae subsp tuberculatae At4
amino acid adenylation protein Cupriavidus basilensis OR16
amino acid adenylation domain containing protein Pantoea sp At 9b
putative non ribosomal peptide synthetase Bradyrhizobium sp ORS 285
amino acid adenylation domain protein Clostridium papyrosolvens DSM 2782
SubA Bacillus subtilis
bacitracin synthetase 1 Bacillus atrophaeus 1942
nonribosomal peptide synthetase Bacillus atrophaeus C89
nrps2A gene product Bacillus amyloliquefaciens subsp plantarum YAU B9601 Y2
amino acid adenylation domain containing protein Clostridium beijerinckii NCIMB 8052
amino acid adenylation enzyme thioester reductase family protein Saccharomonospora cyanea NA 134
Non ribosomal peptide synthetase amino acid adenylation domain Micromonospora lupini str Lupac 08
amino acid adenylation domain containing protein Salinispora tropica CNB 440
hypothetical protein DESPIG 00267 Desulfovibrio piger ATCC 29098
amino acid adenylation enzyme thioester reductase family protein Saccharomonospora cyanea NA 134
non ribosomal peptide synthase polyketide synthase Myxococcus fulvus HW 1
non ribosomal peptide synthase polyketide synthase Myxococcus xanthus DK 1622
oxidoreductase short chain dehydrogenase reductase family Stigmatella aurantiaca DW4 3 1
TPA nonribosomal peptide synthetase polyketide synthase Stigmatella aurantiaca DW4 3 1
CrpB like protein Stigmatella aurantiaca DW4 3 1
MelC protein Melittangium lichenicola
MtaC Stigmatella aurantiaca DW4 3 1
MtaC protein Stigmatella aurantiaca DW4 3 1
amino acid adenylation domain protein Microcoleus vaginatus FGP 2
nonribosomal peptide synthetase Sorangium cellulosum
nonribosomal peptide synthetase Sorangium cellulosum
amino acid adenylation protein Clostridium cellulolyticum H10
Nonribosomal polyketide synthase protein Oscillatoria sp PCC 6506
non ribosomal peptide synthetase module related protein Fischerella sp MV11
NRPS PKS Barbamide biosynthesis protein BarG Oscillatoria sp PCC 6506
non ribosomal peptide synthase polyketide synthase Cyanothece sp CCY0110
amino acid adenylation protein Cyanothece sp PCC 8801
amino acid adenylation domain containing protein Cyanothece sp PCC 8802
phenylalanine racemase Marinomonas mediterranea MMB 1
AMP dependent synthetase and ligase Marinomonas sp MWYL1
non ribosomal peptide synthase Acinetobacter sp ADP1
AMP binding enzyme Acinetobacter baumannii 6013150
putative non ribosomal peptide synthetase Acinetobacter baumannii ATCC 17978
acinetobactin biosynthesis protein Acinetobacter sp ATCC 27244
acinetobactin biosynthesis protein Acinetobacter haemolyticus ATCC 19194
conserved hypothetical protein Acinetobacter sp SH024
basA gene product Acinetobacter calcoaceticus PHEA 2
putative acinetobactin biosynthesis protein Acinetobacter baumannii ATCC 17978
putative acinetobactin biosynthesis protein Acinetobacter baumannii ATCC 17978
acinetobactin biosynthesis protein BasA Acinetobacter baumannii AYE
non ribosomal peptide synthetase Acinetobacter baumannii ATCC 19606
acinetobactin biosynthesis protein Acinetobacter baumannii ACICU
acinetobactin biosynthesis protein Acinetobacter baumannii AB900
amino acid adenylation protein Paracoccus denitrificans PD1222
non ribosomal peptide synthetase module Agrobacterium tumefaciens CCNWGS0286
non ribosomal peptide synthetase Agrobacterium tumefaciens F2
non ribosomal peptide synthetase module Agrobacterium tumefaciens 5A
non ribosomal peptide synthetase BasA D Agrobacterium sp H13 3
vibF gene product Halomonas elongata DSM 2581
non ribosomal peptide synthetase module Aeromonas salmonicida subsp salmonicida 01 B526
non ribosomal peptide synthetase module Aeromonas salmonicida subsp salmonicida A449
non ribosomal peptide synthetase Bas A D Aeromonas salmonicida subsp salmonicida
non ribosomal peptide synthetase BasA D Pseudomonas entomophila L48
non ribosomal peptide synthetase PmsD Pseudomonas synxantha BG33R
non ribosomal peptide synthetase PmsD Pseudomonas fluorescens A506
non ribosomal peptide synthetase Pseudomonas fluorescens
six domain nonribosomal peptide synthetase Photorhabdus asymbiotica subsp asymbiotica ATCC 43949
nonribosomal peptide synthetase PhbF Photorhabdus luminescens
amino acid adenylation domain protein Brenneria sp EniD312
amino acid adenylation protein Serratia plymuthica PRI 2C
amino acid adenylation domain containing protein Serratia sp AS12
SchF2 Serratia sp V4
putative nonribosomal peptide synthetase Serratia odorifera 4Rx13
non ribosomal peptide synthetase module Vibrio furnissii NCTC 11218
non ribosomal peptide synthetase module Vibrio furnissii CIP 102972
Non ribosomal peptide synthetase module Vibrio sp MED222
Non ribosomal peptide synthetase protein Vibrio nigripulchritudo ATCC 27043
non ribosomal peptide synthetase module Vibrio vulnificus YJ016
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio vulnificus MO6 24 O
Non ribosomal peptide synthetase Vibrio vulnificus CMCP6
nonribosomal peptide synthetase VibF Vibrio cholerae V51
nonribosomal peptide synthetase VibF Vibrio cholerae AM 19226
amino acid adenylation domain protein Vibrio cholerae HE 09
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae TMA 21
nonribosomal peptide synthetase VibF Vibrio cholerae MZO 2
Non ribosomal peptide synthetase module protein siderophore biosynthesis Vibrio cholerae LMA3984 4
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae 12129 1
nonribosomal peptide synthetase VibF Vibrio cholerae 1587
nonribosomal peptide synthetase VibF Vibrio cholerae 623 39
amino acid adenylation domain protein Vibrio cholerae HE39
nonribosomal peptide synthetase VibF Vibrio cholerae 2740 80
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae bv albensis VL426
VibF Vibrio cholerae O395
nonribosomal peptide synthetase VibF Vibrio cholerae V52
amino acid adenylation domain protein Vibrio cholerae HC 49A2
nonribosomal peptide synthetase VibF Vibrio cholerae O1 biovar El Tor str N16961
nonribosomal peptide synthetase VibF Vibrio cholerae MZO 3
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae CT 5369 93
nonribosomal peptide synthetase VibF Vibrio cholerae RC385
amino acid adenylation domain protein Vibrio cholerae HE48
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae TM 11079 80
amino acid adenylation domain protein Vibrio cholerae HC 02A1
amino acid adenylation domain protein Vibrio cholerae BJG 01



peptide synthetase Streptomyces sp Acta 2897
non ribosomal peptide synthetase Amycolatopsis mediterranei S699
non ribosomal peptide synthetase Amycolatopsis mediterranei U32
peptide synthetase 1 Streptomyces roseosporus NRRL 15998
CDA peptide synthetase I Streptomyces roseosporus NRRL 11379
peptide synthetase 3 Streptomyces ghanaensis ATCC 14672
CDA peptide synthetase II Streptomyces ghanaensis ATCC 14672
NRPS siderophore biosynthesis protein Streptomyces zincresistens K42
amino acid adenylation enzyme thioester reductase family protein Saccharomonospora xinjiangensis XJ 54
non ribosomal peptide synthetase Rhodococcus equi 103S
exopolyphosphatase Rhodococcus equi ATCC 33707
non ribosomal peptide synthetase Rhodococcus opacus B4
putative non ribosomal peptide synthetase Streptomyces cattleya NRRL 8057 DSM 46488
dhbF gene product Streptomyces cattleya NRRL 8057 DSM 46488
non ribosomal peptide synthetase Streptomyces cattleya NRRL 8057 DSM 46488
putative non ribosomal peptide synthetase Streptomyces cattleya NRRL 8057 DSM 46488
non ribosomal peptide synthetase Nocardia brasiliensis ATCC 700358
non ribosomal peptide synthase Acinetobacter sp ADP1
condensation domain protein Acinetobacter baumannii 6013150
putative nonribosomal peptide synthetase Acinetobacter baumannii ATCC 17978
putative nonribosomal peptide synthetase Acinetobacter baumannii ATCC 17978
vibriobactin synthase component D Vibrio furnissii NCTC 11218
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio furnissii CIP 102972
amino acid adenylation protein Paracoccus denitrificans PD1222
Non ribosomal peptide synthetase module siderophore biosynthesis Agrobacterium tumefaciens CCNWGS0286
condensation domain containing protein Agrobacterium tumefaciens F2
agrobactine synthetase subunit F Agrobacterium tumefaciens 5A
agrobactine synthetase subunit F Agrobacterium sp H13 3
condensation domain protein Escherichia coli MS 196 1
Phenylalanine racemase ATP hydrolyzing Brenneria sp EniD312
six domain nonribosomal peptide synthetase Photorhabdus asymbiotica subsp asymbiotica ATCC 43949
nonribosomal peptide synthetase PhbH Photorhabdus luminescens
condensation domain containing protein Serratia proteamaculans 568
N 5 amino 5 carboxypentanoyl L cysteinyl D valine synthase Serratia plymuthica PRI 2C
N 5 amino 5 carboxypentanoyl L cysteinyl D valine synthase Serratia sp AS12
SchF3 Serratia sp V4
condensation domain containing protein Serratia odorifera 4Rx13
non ribosomal peptide synthetase module Vibrio sp MED222
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio vulnificus MO6 24 O
non ribosomal peptide synthetase module Vibrio vulnificus YJ016
Non ribosomal peptide synthetase Vibrio vulnificus CMCP6
amino acid adenylation domain protein Vibrio cholerae HE39
amino acid adenylation domain protein Vibrio cholerae HE 09
amino acid adenylation domain protein Vibrio cholerae HE48
nonribosomal peptide synthetase VibF Vibrio cholerae V51
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae TMA 21
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae 12129 1
nonribosomal peptide synthetase VibF Vibrio cholerae V52
nonribosomal peptide synthetase VibF Vibrio cholerae O1 biovar El Tor str N16961
amino acid adenylation domain protein Vibrio cholerae HC 49A2
VibF Vibrio cholerae O395
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae bv albensis VL426
nonribosomal peptide synthetase VibF Vibrio cholerae AM 19226
amino acid adenylation domain protein Vibrio cholerae BJG 01
nonribosomal peptide synthetase VibF Vibrio cholerae MZO 2
nonribosomal peptide synthetase VibF Vibrio cholerae MZO 3
nonribosomal peptide synthetase VibF Vibrio cholerae 623 39
Non ribosomal peptide synthetase module protein siderophore biosynthesis Vibrio cholerae LMA3984 4
nonribosomal peptide synthetase VibF Vibrio cholerae 2740 80
nonribosomal peptide synthetase VibF Vibrio cholerae RC385
amino acid adenylation domain protein Vibrio cholerae HC 02A1
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae TM 11079 80
VibF Vibrio cholerae 1587
non ribosomal peptide synthetase modules siderophore biosynthesis Vibrio cholerae CT 5369 93
nonribosomal peptide synthetase VibF Vibrio nigripulchritudo ATCC 27043
non ribosomal peptide synthetase Vibrio anguillarum 775
non ribosomal peptide synthetase BasB Vibrio sp AND4
non ribosomal peptide synthetase Vibrio harveyi HY01
hypothetical protein VcamD 14920 Vibrio campbellii DS40M4
hypothetical protein VIBHAR 02109 Vibrio harveyi ATCC BAA 1116
phenylalanine racemase Marinomonas mediterranea MMB 1
condensation domain containing protein Marinomonas sp MWYL1
non ribosomal peptide synthetase BasB Pseudomonas entomophila L48
non ribosomal peptide synthetase PmsG Pseudomonas synxantha BG33R
non ribosomal peptide synthetase Pseudomonas fluorescens
non ribosomal peptide synthetase PmsG Pseudomonas fluorescens A506
non ribosomal peptide synthetase module Aeromonas salmonicida subsp salmonicida A449
non ribosomal peptide synthetase BasB Aeromonas salmonicida subsp salmonicida
vibF gene product Halomonas elongata DSM 2581
non ribosomal peptide synthetase Acinetobacter sp ATCC 27244
conserved hypothetical protein Acinetobacter haemolyticus ATCC 19194
actinomycin synthetase II Acinetobacter baumannii AB307 0294
basB Acinetobacter baumannii 1656 2
non ribosomal peptide synthetase module Acinetobacter baumannii ABNIH4
non ribosomal peptide synthetase module Acinetobacter baumannii ACICU
putative acinetobactin biosynthesis protein Acinetobacter baumannii ATCC 17978
putative acinetobactin biosynthesis protein Acinetobacter baumannii AB900
conserved hypothetical protein Acinetobacter sp SH024
basB gene product Acinetobacter calcoaceticus PHEA 2
non ribosomal peptide synthetase Acinetobacter baumannii ATCC 19606
non ribosomal peptide synthase Acinetobacter baumannii AB058
non ribosomal peptide synthase Acinetobacter baumannii AYE
non ribosomal peptide synthase Acinetobacter baumannii AB056
non ribosomal peptide synthetase Acinetobacter baumannii AB0057


0.1
amino acid adenylation domain containing protein Saccharomonospora azurea SZMC 14600
putative peptide synthetase Streptomyces hygroscopicus subsp jinggangensis 5008
non ribosomal peptide synthetase partial Streptomyces tsukubaensis NRRL18488
mxcG gene product Corallococcus coralloides DSM 2259
mxcG gene product Sorangium cellulosum So ce 56
siderophore 2 3 dihydroxybenzoate glycine threonine trimeric ester bacillibactin synthetase Paenibacillus sp Aloe 11
siderophore 2 3 dihydroxybenzoate glycine threonine trimeric ester bacillibactin synthetase Paenibacillus peoriae KCTC 3763
siderophore 2 3 dihydroxybenzoate glycine threonine trimeric ester bacillibactin synthetase Paenibacillus elgii B69
amino acid adenylation domain containing protein Paenibacillus dendritiformis C454
dimodular nonribosomal peptide synthetase Paenibacillus sp HGF5
amino acid adenylation domain containing protein Paenibacillus sp Y412MC10
DhbF Bacillus licheniformis WX 02
DhbF Bacillus licheniformis DSM 13 ATCC 14580
DhbF protein Bacillus sp BT1B CT2
amino acid adenylation domain containing protein Geobacillus thermoglucosidasius C56 YS93
amino acid adenylation protein Geobacillus sp Y4 1MC1
siderophore 2 3 dihydroxybenzoate glycine threonine trimeric ester bacillibactin synthetase Bacillus atrophaeus 1942
dhbF gene product Bacillus amyloliquefaciens subsp plantarum YAU B9601 Y2
amino acid adenylation protein Bacillus subtilis subsp spizizenii TU B 10
siderophore 2 3 dihydroxybenzoate glycine threonine trimeric ester bacillibactin synthetase Bacillus subtilis subsp spizizenii ATCC 6633
2 3 dihydroxybenzoate glycine threonine trimeric ester bacillibactin synthetase Bacillus subtilis subsp subtilis
amino acid adenylation domain protein Bacillus subtilis subsp subtilis str RO NN 1
non ribosomal peptide synthetase Bacillus sp JS
siderophore 2 3 dihydroxybenzoate glycine threonine trimeric ester bacillibactin synthetase Bacillus subtilis subsp natto BEST195
DhbF Bacillus subtilis
involved in siderophore 2 3 dihydroxybenzoate DHB synthesis Bacillus subtilis subsp subtilis str 168
siderophore 2 3 dihydroxybenzoate glycine threonine trimeric ester bacillibactin synthetase Bacillus subtilis subsp subtilis str SC 8
siderophore 2 3 dihydroxybenzoate glycine threonine trimeric ester bacillibactin synthetase Bacillus subtilis BSn5
amino acid adenylation protein Herpetosiphon aurantiacus DSM 785
enterobactin synthase subunit F Pusillimonas sp T7 7
enterobactin synthase subunit F Paracoccus sp TRP
entF gene product Paracoccus denitrificans PD1222
AMP binding enzyme Acetobacteraceae bacterium AT 5844
dimodular nonribosomal peptide synthetase Brevundimonas diminuta ATCC 11568
amino acid adenylation protein Starkeya novella DSM 506
vulnibactin synthetase amide synthase subunit Vibrio nigripulchritudo ATCC 27043
enterobactin synthetase component F Vibrio sp MED222
amide synthase component of siderophore synthetase Vibrio vulnificus MO6 24 O
vulnibactin synthetase amide synthase subunit Vibrio vulnificus YJ016
amide synthase component of siderophore synthetase Vibrio vulnificus CMCP6
vibriobactin synthetase amide synthase subunit VibH Vibrio cholerae AM 19226
amide synthase component of siderophore synthetase Vibrio cholerae bv albensis VL426
vibriobactin synthetase Vibrio cholerae RC385
amide synthase component of siderophore synthetase Vibrio cholerae TMA 21
vibriobactin synthesis protein putative Vibrio cholerae HE48
vibriobactin synthesis protein putative Vibrio cholerae HE39
vibriobactin synthetase amide synthase subunit VibH Vibrio cholerae V51
amide synthase component of siderophore synthetase Vibrio cholerae CT 5369 93
vibriobactin synthetase amide synthase subunit VibH Vibrio cholerae 623 39
vibriobactin synthesis protein putative Vibrio cholerae HE 09
vibriobactin synthetase amide synthase subunit VibH Vibrio cholerae MZO 2
vibriobactin synthetase amide synthase subunit VibH Vibrio cholerae MZO 3
VibH Vibrio cholerae
amide synthase component of siderophore synthetase Vibrio cholera CIRS 101
vibriobactin synthesis protein Vibrio cholerae O1 biovar El Tor str N16961
vibriobactin synthesis protein putative Vibrio cholerae HC 02A1
amide synthase component of siderophore synthetase Vibrio cholerae TM 11079 80
Amide synthase component of siderophore synthetase Vibrio cholerae LMA3984 4
hypothetical protein VCG 003207 Vibrio cholerae 12129 1
vibriobactin synthesis protein putative Vibrio cholerae BJG 01
condensation domain containing protein Paracoccus denitrificans PD1222
condensation domain containing protein Paracoccus sp TRP
agrobactine synthetase component F Vibrio nigripulchritudo ATCC 27043
enterobactin synthetase component F Vibrio furnissii NCTC 11218
enterobactin synthetase component f Vibrio furnissii CIP 102972
condensation domain protein Brenneria sp EniD312
photobactin biosynthesis protein phbg Photorhabdus asymbiotica subsp asymbiotica ATCC 43949
putative photobactin biosynthesis protein PhbG Photorhabdus luminescens
condensation domain containing protein Serratia proteamaculans 568
condensation domain containing protein Serratia plymuthica PRI 2C
condensation domain containing protein Serratia sp AS12
SchH Serratia sp V4
condensation domain containing protein Serratia odorifera 4Rx13
Amide synthase component of siderophore synthetase Agrobacterium tumefaciens CCNWGS0286
condensation domain containing protein Agrobacterium tumefaciens F2
agrobactine synthetase subunit F Agrobacterium sp H13 3
agrobactine synthetase subunit F Agrobacterium tumefaciens 5A
brucebactin synthetase Brucella abortus
tyrocidine synthetase 3 Brucella sp NF 2653
brucebactin synthetase Brucella sp 83 13
tyrocidine synthetase 3 Brucella sp BO1
tyrocidine synthetase 3 Brucella sp BO2
enterobactin synthetase component F Brucella abortus bv 5 str B3196
brucebactin synthetase Brucella abortus bv 9 str C68
putative enterobactin synthetase subunit F Brucella ovis ATCC 25840
hypothetical protein BSUIS B0019 Brucella suis ATCC 23445
condensation domain containing protein Brucella abortus bv 3 str Tulya
Condensation domain protein Brucella abortus S19
enterobactin synthetase subunit F Brucella suis 1330
brucebactin synthetase Brucella ceti B1 94
brucebactin synthetase Brucella ceti M13 05 1
enterobactin synthetase component F Brucella pinnipedialis B2 94
Linear gramicidin synthetase subunit B Brucella ceti str Cudo
condensation domain containing protein Brucella sp F5 99
condensation domain containing protein Brucella abortus NCTC 8038
brucebactin synthetase Brucella neotomae 5K33
enterobactin synthetase component F Brucella melitensis bv 1 str 16M
condensation domain containing protein Brucella melitensis bv 1 str 16M
Linear gramicidin synthetase subunit B Brucella melitensis M5 90
Linear gramicidin synthetase subunit B Brucella melitensis NI
Linear gramicidin synthetase subunit B Brucella melitensis ATCC 23457

Figure S12. Phylogenetic analyses of enterobactin-like genes in the *sch* cluster. On the next two pages follow phylogenetic trees for SchA (page S30) and SchB (page S31). The *Serratia sp.* V4 query protein is highlighted in yellow. In both cases, of the top 100 sequences, the majority include the corresponding enterobactin genes from *Escherichia* and related strains.


0.1
2 3 dihydroxybenzoate 2 3 dehydrogenase Dickeya dadantii Ech703
short chain dehydrogenase reductase SDR Pantoea sp At 9b
2 3 dihydroxybenzoate 2 3 dehydrogenase Pantoea sp Sc1
entA gene product Pantoea vagans C9 1
2 3 dihydroxybenzoate 2 3 dehydrogenase Pantoea agglomerans IG1
2 3 dihydroxybenzoate 2 3 dehydrogenase Pantoea sp SL1 M5
2 3 dihydroxybenzoate 2 3 dehydrogenase Serratia sp M24T3
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Serratia odorifera DSM 4582
2 3 dihydroxybenzoate 2 3 dehydrogenase Serratia proteamaculans 568
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Serratia sp AS12
SchA Serratia sp V4
2 3 dihydroxybenzoate 2 3 dehydrogenase Serratia odorifera 4Rx13
putative 2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase PhbA Photorhabdus luminescens
2 3 dihydroxybenzoate 2 3 dehydrogenase Photorhabdus asymbiotica subsp asymbiotica ATCC 43949
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Brenneria sp EniD312
2 3 dihydroxybenzoate 2 3 dehydrogenase Dickeya dadantii Ech703
cbsA gene product Dickeya dadantii 3937
2 3 dihydroxybenzoate 2 3 dehydrogenase Dickeya zeae Ech1591
short chain dehydrogenase reductase SDR Dickeya dadantii Ech586
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia blattae DSM 4481
short chain dehydrogenase reductase SDR Enterobacter cloacae SCF1
2 3 dihydroxybenzoate 2 3 dehydrogenase Yokenella regensburgei ATCC 43003
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Enterobacter cancerogenus ATCC 35316
2 3 dihydroxybenzoate 2 3 dehydrogenase Enterobacter sp 638
short chain dehydrogenase reductase SDR Enterobacter asburiae LF7a
2 3 dihydroxybenzoate 2 3 dehydrogenase Enterobacter cloacae subsp dissolvens SDM
2 3 dihydroxybenzoate 2 3 dehydrogenase Enterobacter cloacae subsp cloacae ATCC 13047
entA gene product Enterobacter cloacae EcWSU1
2 3 dihydroxybenzoate 2 3 dehydrogenase Enterobacter mori LMG 25706
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Enterobacter hormaechei ATCC 49162
Dehydrogenases with different specificities related to short chain alcohol dehydrogenases Enterobacter cloacae subsp cloacae NCTC 9394
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Klebsiella oxytoca 10 5242
2 3 dihydroxybenzoate 2 3 dehydrogenase Klebsiella oxytoca KCTC 1686
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Klebsiella oxytoca 10 5250
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Klebsiella oxytoca 10 5243
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Klebsiella oxytoca 10 5245
2 3 dihydroxybenzoate 2 3 dehydrogenase Enterobacter aerogenes KCTC 2190
entA gene product Klebsiella pneumoniae subsp pneumoniae MGH 78578
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Klebsiella sp 4 1 44FAA
2 3 dihydroxybenzoate 2 3 dehydrogenase Klebsiella pneumoniae subsp pneumoniae NTUH K2044
2 3 dihydroxybenzoate 2 3 dehydrogenase Klebsiella sp MS 92 3
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Citrobacter youngae ATCC 29220
2 3 dihydroxybenzoate 2 3 dehydrogenase Citrobacter koseri ATCC BAA 895
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Salmonella bongori NCTC 12419
2 3 dihydroxybenzoate 2 3 dehydrogenase Salmonella enterica subsp enterica serovar Paratyphi A str ATCC 9150
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase siderophore Salmonella enterica subsp enterica serovar Uganda str R8 3404
entA gene product Salmonella enterica subsp enterica serovar Choleraesuis str SC B67
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Salmonella enterica subsp enterica serovar Javiana str GA MM04042433
2 3 dihydroxybenzoate 2 3 dehydrogenase Salmonella enterica subsp enterica serovar Paratyphi B str SPB7
2 3 dihydroxybenzoate 2 3 dehydrogenase Salmonella enterica subsp enterica serovar Gallinarum str 287 91
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Salmonella enterica subsp enterica serovar Virchow str SL491
2 3 dihydroxybenzoate 2 3 dehydrogenase Salmonella enterica subsp enterica serovar Tennessee str CDC07 0191
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia sp TW09308
short chain dehydrogenase Escherichia coli M863
short chain dehydrogenase Escherichia coli TW10509
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia fergusonii ECD227
short chain dehydrogenase Escherichia fergusonii B253
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia fergusonii ATCC 35469
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli STEC MHI813
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli B354
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli 536
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli UTI89
entA gene product Escherichia coli APEC O1
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli TA280
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli DEC2A
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli MS 16 3
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli 83972
entA gene product Escherichia coli str clone D i2
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli CFT073
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli NA114
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase enterobactin siderophore Escherichia coli AA86
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli M605
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli B185
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli E24377A
short chain dehydrogenase reductase SDR Escherichia coli W
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli str K 12 substr MG1655
Chain X Crystal Structure Of E Coli Enta A 2 3 Dihydrodihydroxy Benzoate Dehydrogenase
entA gene product Shigella dysenteriae Sd197
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli H736
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Shigella dysenteriae 1012
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Shigella boydii 965 58
2 3 dihydroxybenzoate 2 3 dehydrogenase Shigella sonnei Ss046
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli SMS 3 5
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli 042
entA gene product Escherichia coli O157 H7 str EDL933
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli 3 2608
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli TA124
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli H591
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia albertii TW07627
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli O26 H11 str 11368
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli 2534 86
2 3 dihydroxybenzoate 2 3 dehydrogenase Shigella sp D9
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli 1 2264
entA Escherichia coli DEC10E
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli OK1357
2 3 dihydro 2 3 dihydroxybenzoate dehydrogenase Escherichia coli E110019
2 3 dihydroxybenzoate 2 3 dehydrogenase Escherichia coli MS 84 1
2 3 dihydroxybenzoate 2 3 dehydrogenase Shigella boydii Sb227
EntA Escherichia coli
2 3 dihydroxybenzoate 2 3 dehydrogenase Shigella flexneri 5 str 8401
2 3 dihydroxybenzoate 2 3 dehydrogenase Shigella flexneri 2a str 2457T

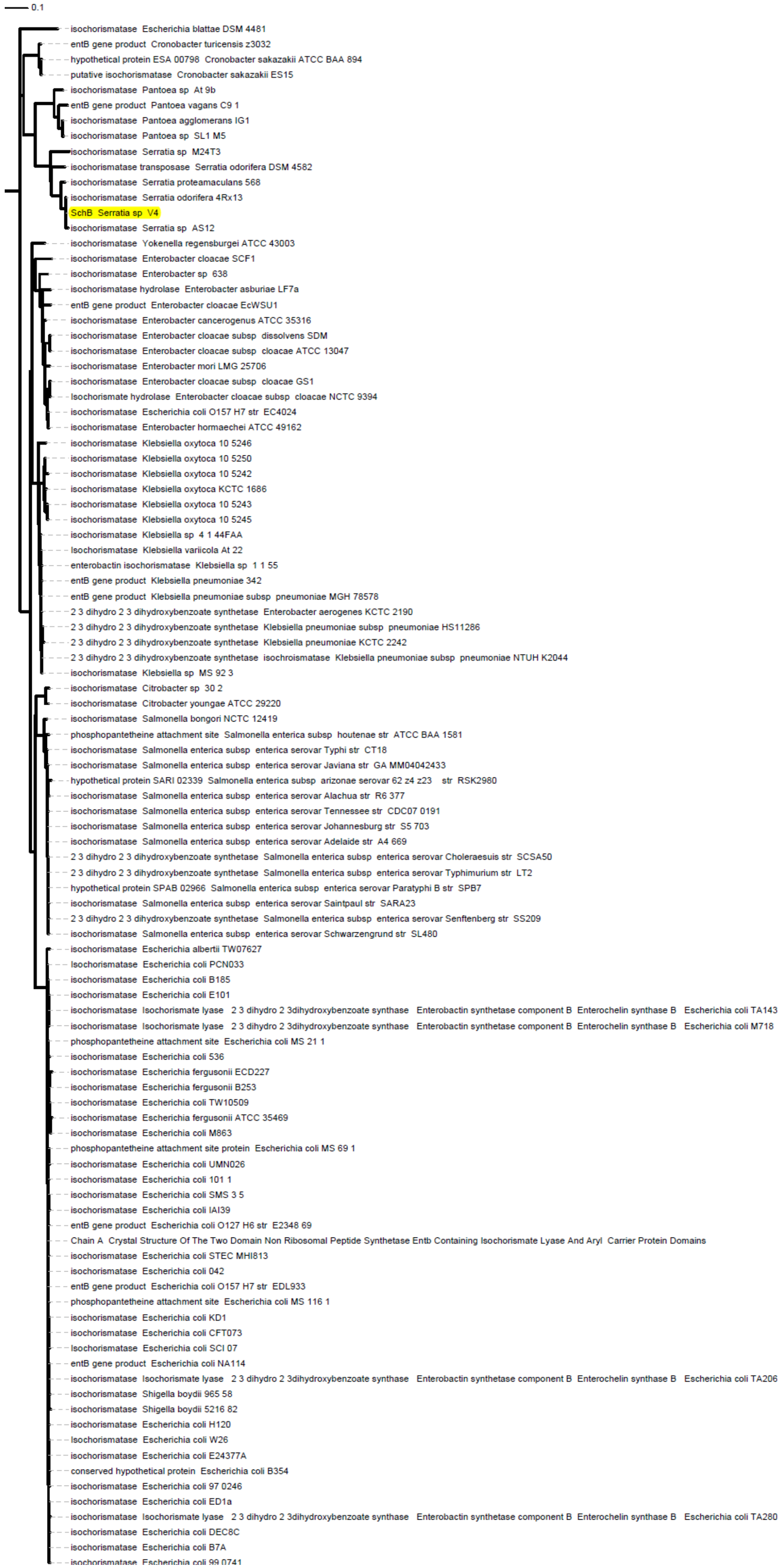
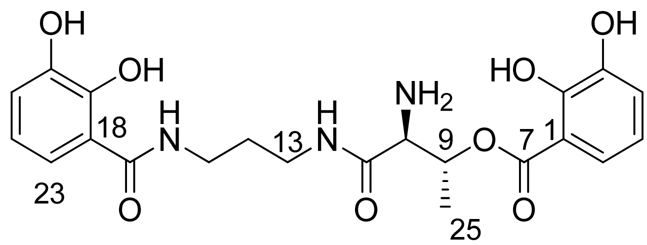
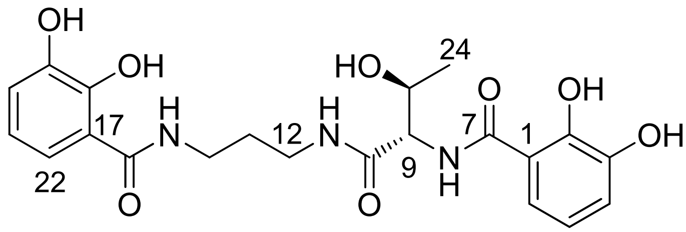

0.1
isochorismatase Escherichia blattae DSM 4481
entB gene product Cronobacter turicensis z3032
hypothetical protein ESA 00798 Cronobacter sakazakii ATCC BAA 894
putative isochorismatase Cronobacter sakazakii ES15
isochorismatase Pantoea sp At 9b
entB gene product Pantoea vagans C9 1
isochorismatase Pantoea agglomerans IG1
isochorismatase Pantoea sp SL1 M5
isochorismatase Serratia sp M24T3
isochorismatase transposase Serratia odorifera DSM 4582
isochorismatase Serratia proteamaculans 568
isochorismatase Serratia odorifera 4Rx13
SchB Serratia sp V4
isochorismatase Serratia sp AS12
isochorismatase Yokenella regensburgei ATCC 43003
isochorismatase Enterobacter cloacae SCF1
isochorismatase Enterobacter sp 638
isochorismatase hydrolase Enterobacter asburiae LF7a
entB gene product Enterobacter cloacae EcWSU1
isochorismatase Enterobacter cancerogenus ATCC 35316
isochorismatase Enterobacter cloacae subsp dissolvens SDM
isochorismatase Enterobacter cloacae subsp cloacae ATCC 13047
isochorismatase Enterobacter mori LMG 25706
isochorismatase Enterobacter cloacae subsp cloacae GS1
Isochorismate hydrolase Enterobacter cloacae subsp cloacae NCTC 9394
isochorismatase Escherichia coli O157 H7 str EC4024
isochorismatase Enterobacter hormaechei ATCC 49162
isochorismatase Klebsiella oxytoca 10 5246
isochorismatase Klebsiella oxytoca 10 5250
isochorismatase Klebsiella oxytoca 10 5242
isochorismatase Klebsiella oxytoca KCTC 1686
isochorismatase Klebsiella oxytoca 10 5243
isochorismatase Klebsiella oxytoca 10 5245
isochorismatase Klebsiella sp 4 1 44FAA
Isochorismatase Klebsiella variicola At 22
enterobactin isochorismatase Klebsiella sp 1 1 55
entB gene product Klebsiella pneumoniae 342
entB gene product Klebsiella pneumoniae subsp pneumoniae MGH 78578
2 3 dihydro 2 3 dihydroxybenzoate synthetase Enterobacter aerogenes KCTC 2190
2 3 dihydro 2 3 dihydroxybenzoate synthetase Klebsiella pneumoniae subsp pneumoniae HS11286
2 3 dihydro 2 3 dihydroxybenzoate synthetase Klebsiella pneumoniae KCTC 2242
2 3 dihydro 2 3 dihydroxybenzoate synthetase isochroismatase Klebsiella pneumoniae subsp pneumoniae NTUH K2044
isochorismatase Klebsiella sp MS 92 3
isochorismatase Citrobacter sp 30 2
isochorismatase Citrobacter youngae ATCC 29220
isochorismatase Salmonella bongori NCTC 12419
phosphopantetheine attachment site Salmonella enterica subsp houtenae str ATCC BAA 1581
isochorismatase Salmonella enterica subsp enterica serovar Typhi str CT18
isochorismatase Salmonella enterica subsp enterica serovar Javiana str GA MM04042433
hypothetical protein SARI 02339 Salmonella enterica subsp arizonae serovar 62 z4 z23 str RSK2980
isochorismatase Salmonella enterica subsp enterica serovar Alachua str R6 377
isochorismatase Salmonella enterica subsp enterica serovar Tennessee str CDC07 0191
isochorismatase Salmonella enterica subsp enterica serovar Johannesburg str S5 703
isochorismatase Salmonella enterica subsp enterica serovar Adelaide str A4 669
2 3 dihydro 2 3 dihydroxybenzoate synthetase Salmonella enterica subsp enterica serovar Choleraesuis str SCSA50
2 3 dihydro 2 3 dihydroxybenzoate synthetase Salmonella enterica subsp enterica serovar Typhimurium str LT2
hypothetical protein SPAB 02966 Salmonella enterica subsp enterica serovar Paratyphi B str SPB7
isochorismatase Salmonella enterica subsp enterica serovar Saintpaul str SARA23
2 3 dihydro 2 3 dihydroxybenzoate synthetase Salmonella enterica subsp enterica serovar Senftenberg str SS209
isochorismatase Salmonella enterica subsp enterica serovar Schwarzengrund str SL480
isochorismatase Escherichia albertii TW07627
Isochorismatase Escherichia coli PCN033
isochorismatase Escherichia coli B185
isochorismatase Escherichia coli E101
isochorismatase Isochorismate lyase 2 3 dihydro 2 3dihydroxybenzoate synthase Enterobactin synthetase component B Enterochelin synthase B Escherichia coli TA143
isochorismatase Isochorismate lyase 2 3 dihydro 2 3dihydroxybenzoate synthase Enterobactin synthetase component B Enterochelin synthase B Escherichia coli M718
phosphopantetheine attachment site Escherichia coli MS 21 1
isochorismatase Escherichia coli 536
isochorismatase Escherichia fergusonii ECD227
isochorismatase Escherichia fergusonii B253
isochorismatase Escherichia coli TW10509
isochorismatase Escherichia fergusonii ATCC 35469
isochorismatase Escherichia coli M863
phosphopantetheine attachment site protein Escherichia coli MS 69 1
isochorismatase Escherichia coli UMN026
isochorismatase Escherichia coli 101 1
isochorismatase Escherichia coli SMS 3 5
isochorismatase Escherichia coli IAI39
entB gene product Escherichia coli O127 H6 str E2348 69
Chain A Crystal Structure Of The Two Domain Non Ribosomal Peptide Synthetase Entb Containing Isochorismate Lyase And Aryl Carrier Protein Domains
isochorismatase Escherichia coli STEC MHI813
isochorismatase Escherichia coli 042
entB gene product Escherichia coli O157 H7 str EDL933
phosphopantetheine attachment site Escherichia coli MS 116 1
isochorismatase Escherichia coli KD1
isochorismatase Escherichia coli CFT073
Isochorismatase Escherichia coli SCI 07
entB gene product Escherichia coli NA114
isochorismatase Isochorismate lyase 2 3 dihydro 2 3dihydroxybenzoate synthase Enterobactin synthetase component B Enterochelin synthase B Escherichia coli TA206
isochorismatase Shigella boydii 965 58
isochorismatase Shigella boydii 5216 82
isochorismatase Escherichia coli H120
Isochorismatase Escherichia coli W26
isochorismatase Escherichia coli E24377A
conserved hypothetical protein Escherichia coli B354
isochorismatase Escherichia coli 97 0246
isochorismatase Escherichia coli ED1a
isochorismatase Isochorismate lyase 2 3 dihydro 2 3dihydroxybenzoate synthase Enterobactin synthetase component B Enterochelin synthase B Escherichia coli TA280
isochorismatase Escherichia coli DEC8C
isochorismatase Escherichia coli B7A
isochorismatase Escherichia coli 99 0741

Table S1. NMR Data for serratiochelin C (**1**) acquired in CD_3OD at 600 MHz. The structure and numbering scheme of **1** is shown above the table.



Position	δ_H (*J* in Hz)	δ_C	COSY	HMBC
1		113.4		
2		150.9		
3		146.8		
4	7.03, dd (7.9, 1.5)	122.0	H-5	C-2, C-3, C-6
5	6.75, t (8.1)	120.1	H-4, H-6	C-1, C-3
6	7.48, dd (8.2, 1.5)	121.7	H-5	C-2, C-4, C-7
7		170.1		
9	5.65, dq (6.5, 1.5)	71.3	H-10, H-25	C-7, C-10, C-11, C-25
10	4.22, m	57.6	H-9	C-9, C-11, C-25
11		167.3		
13	3.38, m; 3.27, m	38.0	H-14	C-11, C-14, C-15
14	1.79, dtd (13.4, 6.7, 6.7)	29.7	H-13, H-15	C-13, C-15
15	3.43, m; 3.35, m	37.4	H-14	C-13, C-14, C-17
17		171.4		
18		116.3		
19		150.0		
20		147.0		
21	6.93, dd (7.9, 1.5)	119.4	H-22	C-19, C-20, C-23
22	6.71, t (8.1)	119.4	H-21, H-23	C-18, C-20
23	7.22, dd (8.1, 1.3)	118.4	H-22	C-17, C-19, C-21
25	1.53, d (6.4)	16.8	H-9	C-9, C-10

Table S2. NMR Data for serratiochelin B (**2**) acquired in CD_3OD at 600 MHz. The structure and numbering scheme of **2** is shown above the table.



Position	δ_H (*J* in Hz)	δ_C	COSY	HMBC
1		114.0		
2		153.9		
3		153.9		
4	6.68, dd (7.5, 1.3)	111.8	H-5	C-2, C-6
5	6.57, t (8.2)	118.1	H-4, H-6	C-1, C-3
6	7.14, dd (8.2, 1.2)	118.0	H-5	C-2, C-4, C-7
7		169.4		
9	4.44, d (3.2)	60.6	H-23	
10		173.2		
12	3.33, m	37.9	H-13	C-10, C-13, C-14
13	1.83, q (6.8)	30.4	H-12, H-14	C-12, C-14
14	3.46, t (7.0)	37.5	H-13	C-12, C-13, C-16
16		169.2		
17		114.0		
18		153.6		
19		153.6		
20	6.66, dd (7.5, 1.3)	111.6	H-21	C-18, C-22
21	6.55, t (7.8)	117.9	H-20, H-22	C-17, C-19
22	7.11, dd (8.4, 1.3)	117.9	H-21	C-16, C-18, C-20
23	4.31, m	68.1	H-9, H-24	
24	1.25, d (6.5)	20.1	H-23	C-9, C-23

Table S3. Annotation of the serratiochelin biosynthetic cluster A.[a]

Gene	AA length	Predicted Function	Strain (UniProtKB #)	% Identity	% Similarity
schA	252	2,3-dihydroxybenzoate-2,3-dehydrogenase	*Serratia odorifera* 4Rx13 (D1RU45)	100	100
schB	284	Isochorismatase	*Serratia odorifera* 4Rx13 (D1RU46)	99.6	100
schC	402	Isochorismate synthase	*Serratia odorifera* 4Rx13 (D1RU48)	99.5	100
schE	541	Enterobactin synthase subunit E	*Serratia odorifera* 4Rx13 (D1RU47)	98.7	100
schF0	1325	Amino acid adenylation domain protein	*Serratia sp.* AS13 (G0C7N5)	97.9	99.2
schG	561	Acetolactate synthase	*Serratia sp.* AS13 (G0C7P0)	100	100
schI	319	Iron-enterobactin transporter periplasmic binding protein	*Serratia odorifera* 4Rx13 (D1RU49)	99.7	100
schJ	422	Enterobactin exporter EntS	*Serratia odorifera* 4Rx13 (D1RU50)	99.8	99.8
schK	319	Iron-enterobactin transporter periplasmic binding protein	*Serratia odorifera* 4Rx13 (D1RU49)	99.7	100
schL	339	ABC-type transporter, integral membrane subunit	*Serratia sp.* AS9 (G0BC03)	98.8	99.7
schM	266	Iron-chelate-transporting ATPase	*Serratia sp.* AS9 (G0BC04)	99.2	99.6
schN	457	Enterobactin/ferric enterobactin esterase	*Serratia odorifera* 4Rx13 (D1RU56)	99.3	100
schO	759	TonB-dependent siderophore receptor	*Serratia sp.* AS9 (G0BC08)	98.8	99.7
schP	531	ABC-type transporter, periplasmic subunit	*Serratia* sp. AS12 (G0BTV6)	98.3	99.2

[a] Predicted functions are from FASTA, BLAST, and InterProScan searches as well as other online bioinformatic tools (see general procedures) conducted between 12/5/2011 to 4/25/2012.

Table S4. Annotation of the serratiochelin biosynthetic cluster B.[a]

Gene	AA length	Predicted Function	Strain (UniProtKB #)	% Identity	% Similarity
schF1	530	Glutamate racemase	*Serratia odorifera* AS13 (G0C5K3)	97.2	99.2
schF1	530	*N*-terminal domain of Non-ribosomal peptide synthetase module protein	*Vibrio cholerae* M66-2 VM66_2132	44	60
schF2	1030	Putative nonribosomal peptide synthetase	*Serratia odorifera* 4Rx13 (D1RVM4)	98.8	99.5
schF2	1030	Mid-protein domain of nonribosomal peptide synthetase VibF	*Vibrio cholerae* M66-2 VM66_2132	50	64
schF3	917	Condensation domain-containing protein	*Serratia odorifera* 4Rx13 (D1RVM7)	97.9	99.1
schF3	917	*C*-terminal portion of Nonribosomal peptide synthetase VibF	*Vibrio cholerae* M66-2 VM66_2132	38	56
schH	448	Condensation domain-containing protein	*Serratia odorifera* 4Rx13 (D1RVM3)	99.3	100
schH	448	Vibriobactin synthetase	*Vibrio cholerae* M66-2 VM66_0733	31	45
schQ	707	TonB-dependent siderophore receptor	*Serratia odorifera* 4Rx13 (D1RVM6)	100	100
schQ	707	Ferric vibriobactin receptor	*Vibrio cholerae* M66-2 VM66_2134	43	60
schR	532	Transporter	*Serratia odorifera* 4Rx13 (D1RVM2)	99.6	100
schR	532	Peptide ABC transporter	*Vibrio cholerae* M66-2 VM66_1620	36	56

[a] Predicted functions are from FASTA, BLAST, and InterProScan searches as well as other online bioinformatic tools (see general procedures) conducted between 12/5/2011 to 4/25/2012. For each sequence, results from a FASTA search as well as a *V. cholera*-targeted BLAST search (from img.jgi.doe.gov/cgi-bin/w/main.cgi) are shown.

Table S5. Bacterial strains and plasmids used in this study.

	Strain or Plasmid	Characteristic	Reference number	Source
Strains	***Escherichia coli***			
	E. coli S17-1 *λPir*	Tp^r Sm^r Kn^r recA, thi, pro, hsdR-M^+RP4: 2-Tc:Mu: Km Tn7 *λPir* ; RP4; cloning and conjugative strain	ZK2162	Laboratory stock
	E. coli ER1100A	Defective in the production of siderophore; *E. coli ΔEntF*	--	SI Ref. 1
	E. coli K12	--	ZK638	Laboratory stock
	***Serratia* sp.**			
	V4	Siderophore producer, wildtype	ZK4911	This study
	V4 *schB*:: Gen^r	*schB* insertional mutant	ZK4954	This study
	V4 *schE*:: Gen^r	*schE* insertional mutant	ZK4952	This study
	V4 sch*C*:: Gen^r	*schC* insertional mutant	ZK4964	This study
	V4 *schH*:: Gen^r	*schH* insertional mutant	ZK4984	This study
	V4 *schF2*:: Gen^r	*schF2* insertional mutant	ZK4987	This study
	V4 *schF0*:: Gen^r	*schF0* insertional mutant	ZK4962	This study
Plasmids	pBTK30*	With an R6K ori and Gen^r cassette; Mariner C9 and Amp^r cassette removed and renamed pSC30;	-	SI Ref. 2
	pSC30C	pSC33 with a 723 bp long SpeI-StuI fragment of *schC*	ZK4951	This study
	pSC30E	pSC33 with a 640 bp long SpeI-StuI fragment of *schE*	ZK4950	This study
	pSC30B	pSC33 with a 662 bp long SpeI-StuI fragment of *schB*	ZK4949	This study
	pSC30H	pSC33 with a 762 bp long SpeI-StuI fragment of *sch*H	ZK4958	This study
	pSC30F	pSC33 with a 985 bp long SpeI-StuI fragment of *schF2*	ZK4968	This study
	pSC30EF	pSC33 with a 615 bp long SpeI-StuI fragment of *schF0*	ZK4956	This study

*Provided by Dr. Josh Sharp, Children's Hospital Boston, USA

Table S6. List of forward (Fw) and reverse (Rv) primers used for PCR amplification of inserts and for confirmation of constructs. SpeI (Fw) and StuI (Rv) restriction sites have been added.

Primer Name	Primer Sequence
schH-Fw	5' CCG GAC TAG TCG GGC GAT AAC CCT CAC TGT CG 3'
schH-Rv	5' CCG GAG GCC TCG GCA GCG TAT TGA CCG CC 3'
schF2-Fw	5' CCG GAC TAG TCG ATC GCG GCG GGC ACT CG 3'
schF2-Rv	5' CCG GAG GCC TGC CAA GCA GGC GCA GGG CG 3'
schE-Fw	5' CCG GAC TAG TGG GCT TGC ACA GTG GCG ATA CC 3'
schE-Rv	5' CCG GAG GCC TCG GCA CCA GCG CCG TCA CG 3'
schF0-Fw	5' CCG GAC TAG TCG CCT TAC GCC AAT GCC TAT GC 3'
schF0-Rv	5' CCG GAG GCC TGG TGG ATG CGG GTG GTT GGC 3'
schB-Fw	5' CCG GAC TAG TGC GTT ACT GAT CCA CGA TAT GC 3'
schB-Rv	5' CCG G AG GCC T GC GGC CAA CTC CAT GAT ACG 3'
schC-Fw	5' CCG GCC GCG TTG TTT ATT CCC CAG C 3'
schC-Rv	5' CCG GAG GCC TCG AAC AGG CCG CGA TCA AAG G 3'
Conf-Fw	5' CCA GGG TTT TCC CAG TCA CG 3'
Conf-Rv	5' GGC ACA TCA AGG CCA AGC CC 3'

Table S7. List of primers used for the PCR confirmation of gene interruptions. Up and dw refer to annealing upstream and downstream of the gene of interest.

Primer Name	Primer Sequence
schH-up	5' GGC TGC CGT TGT CGC CGG CG 3'
schH-dw	5' CCG CGA TCG GCG GCA ATC TGC 3'
schF2-up	5' GGC CCT CAG CCT GCC GGA TCC 3'
schF2-dw	5' CGA TCA ACC CCG CCA GCG TCA GC 3'
schE-up	5' GCG GGC GTT TGG CCT GAA CC 3'
schE-dw	5' CCC TGA GTG GTG AGA ATA TGC CG 3'
schF0-up	5' CGG TTT TCG GCC CGG CGG 3'
schF0-dw	5' GCG TAC CAC CCA GCG AAT AAC 3'
schB-up	5' GGT CGA TAC GCT GCC GCT CAC C 3'
schB-dw	5' CGA AGC CAA TTA CCT CGG CAC C 3'
schC-up	5' GCG TTT CGC CCA CCG CGG G 3'
schC-dw	5' CCT TGG TTC AGG CCA AAC GCC 3'
Plasmid	5' CCA GGG TTT TCC CAG TCA CG 3'

Supporting Information References

1. Roech, E. D.; Walsh, C. T. *Biochemistry* **2003**, *42*, 1334.
2. Goodman, A. L.; Kulasekara, B.; Rietsch, A.; Boyd, D.; Smith, R. S.; Lory, S. *Dev. Cell* **2004**, *7*, 745.
3. Hider, R. C.; Kong, X. *Nat. Prod. Rep.* **2010**, *27*, 637.
4. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403.
5. Pearson, W. R.; Lipman, D. J. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 2444.
6. Zdobnov, E. M.; Apweiler, R. *Bioinformatics* **2001**, *17*, 847.
7. Starcevic, A.; Zucko, J.; Simunkovic, J.; Long, P. F.; Cullum, J.; Hranueli, D. *Nucleic Acids Res.* **2008**, *36*, 6882.
8. Marchler-Bauer, A.; Bryant, S. H. *Nucleic Acids Res.* **2004**, *32*, W327.
9. Marchler-Bauer, A. *et al.* *Nucleic Acids Res.* **2001**, *39*, D225.
10. Marchler-Bauer, A. *et. al.* *Nucleic Acids Res.* **2009**, *37*, D205.
11. Letunic, I.; Doerks, T.; Bork, P. *Nucleic Acids Res.* **2009**, *37*, D229.
12. Schultz, J. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5857.
13. Rausch, C.; Weber, T.; Kohlbacher, O.; Wohlleben, W.; Huson, D. H. *Nucleic Acids Res.* **2005**, *33*, 5799.
14. Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. *Nucleic Acids Res.* **2011**, *39*, W362.
15. Ansari, M. Z.; Yadav, G.; Gokhale, R. S.; Mohanty, D. *Nucleic Acids Res.* **2004**, *32*, W405.
16. Saitou, N.; Nei, M. *Mol. Biol. Evol.* **1987**, *4*, 406.
17. Grishin, N. V. *J. Mol. Evol.* **1995**, *41*, 675.
18. Letunic, I.; Bork, P. *Bioinformatics* **2006**, *23*, 127.
19. Letunic, I.; Bork, P. Nucleic Acid Res. **2011**, *39*, W475.