Supporting Information for:

Nucleophilic 1,4-additions for natural product discovery

Courtney L. Cox[1-2], Jonathan I. Tietz[3], Karol Sokolowski[2-3], Joel O. Melby[2-3], James R. Doroghazi[2], and Douglas A. Mitchell[1-3]*

[1]Department of Microbiology, [2]Institute for Genomic Biology, and [3]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois, 61801, USA,* *douglasm@illinois.edu*

TABLE OF CONTENTS

I. Methods

MALDI-TOF mass spectrometric analysis. MALDI-TOF mass spectrometry was performed using a Bruker Daltonics UltrafleXtreme MALDI-TOF/TOF instrument operating in positive reflector mode. The instrument was calibrated before data acquisition using a commercial peptide calibration kit (AnaSpec – Peptide Mass Standard Kit). Analysis was carried out with Bruker Daltonics flexAnalysis software. All spectra were processed by smoothing and baseline subtraction.

Isolation of cyclothiazomycin C. WC-3908 was grown in 10 mL of ATCC 172 medium at 30 °C for 48 h. 300 µL of the culture was spread onto 15 cm plates (*ca*. 75 mL of solid ATCC medium). The plates were then incubated for 7 d at 23 °C. A razor blade was used to remove the bacterial lawn from the solid medium. The bacterial growth from 14 plates (~1 L of medium) was extracted with *n*-BuOH (500 mL) for 24 h at 23 °C. The extract was then filtered through Whatman filter paper and allowed to evaporate under nitrogen before being redissolved in 3:1 pyridine:water (*ca*. 3 mL) and transferred to a 50 mL conical tube. The resulting solution was clarified by centrifugation, to remove insoluble debris (4000 × *g*, 5 min). The supernatant was then injected onto a reverse-phase C18 silica column (TeleDyne Isco 5.5 g C18 Gold cartridge) and purified by MPLC (gradient elution from 20-95% MeOH/10 mM aq. NH_4HCO_3). Fractions containing the desired product (as determined by MALDI-TOF MS; [M+H] *m/z* = 1486) were combined and immediately concentrated by rotary evaporation. The resulting residue was dissolved in 3:1 pyridine/water (*ca*. 0.5 mL), transferred to a microcentrifuge tube, centrifuged (15000 × *g*, 5 min), filtered (0.2 µm polyethersulfone syringe filter), and further purified by HPLC. Semi-preparative HPLC employed a Thermo Scientific Betasil C18 column (100 Å; 250

× 10 mm; 5 μm particle size) operating at 4.0 mL min^{-1} on a PerkinElmer Flexar LC system using Flexar Manager software. Solvent A was 10 mM aq. NH_4HCO_3. Solvent B was MeOH. Cyclothiazomycin C was purified by isocratic elution at 72% B, typically eluting 19.5 min after initiation of the HPLC run (alternatively, the elution time was ~12 min when 75% B was used). HPLC progress was monitored by photodiode array (PDA) UV-Vis detection. Fractions corresponding to the desired product (as determined by UV-Vis and MALDI-TOF MS) were immediately concentrated under rotary evaporation or under a stream of N_2 gas. The resulting residue was suspended in water (*ca*. 1 mL), assisted by vortex mixing and sonication. The suspended product was flash-frozen in liquid N_2 and lyophilized for >24 h to give purified cyclothiazomycin C as a white to off-white powder. Purity was determined by analytical HPLC [Thermo Scientific Betasil C18 column (100 Å; 250 × 4.6 mm; 5 μm particle size) operating at 1.0 mL min^{-1} using the same solvents] and NMR. Isolated yield ranged from 10-90 μg/plate (15 cm diameter).

Isolation of cyclothiazomycin B. NRRL strain B-3306 was grown in a fashion identical isolation conditions for WC-3908. Cyclothiazomycin B ([M+H] *m/z* = 1528) was also purified in the same manner as cyclothiazomycin C, except that HPLC purification employed 75% B (retention time typically *ca*. 17 min). After lyophilization, an off-white powder was obtained. Purity was determined by analytical HPLC [Thermo Scientific Betasil C18 column (100 Å; 250 × 4.6 mm; 5 μm particle size) operating at 1.0 mL min^{-1} using the same solvents]; identity was determined by high-resolution mass spectrometry. Isolated yield was approximately 13 μg/plate (15 cm diameter).

FT-MS/MS analysis of cyclothiazomycin B and C. The purified cyclothiazomycins were dissolved in 80% aq. MeCN with 0.1% formic acid. Samples were directly infused using a 25 μL Hamilton gas-tight syringe (cyclothiazomycin C) or an Advion Nanomate 100 (cyclothiazomycin B), into a ThermoFisher Scientific LTQ-FT hybrid linear ion trap, operating at 11T (calibrated weekly). The FT-MS was operated using the following parameters: minimum target signal counts, 5,000; resolution, 100,000; *m/z* range detected, dependent on target *m/z*; isolation width (MS/MS), 5 *m/z*; normalized collision energy (MS/MS), 35; activation *q* value (MS/MS), 0.4; activation time (MS/MS), 30 ms. Data analysis was conducted using the Qualbrowser application of Xcalibur software (Thermo-Fisher Scientific).

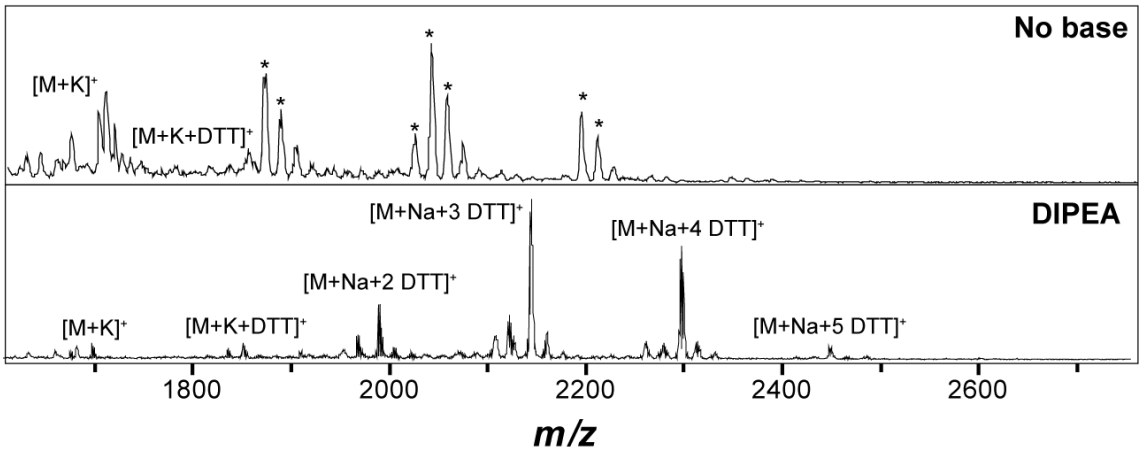
NMR spectroscopy of cyclothiazomycin C. NMR spectra were recorded on a Varian NMR System 750 MHz narrow bore magnet spectrometer (VNS750NB employing a 5 mm Varian 1H[13C/15N] PFG X, Y, Z probe) or a Varian Unity Inova 500 MHz narrow bore magnet spectrometer (UI500NB employing a 5 mm Varian 1H[13C/15N] PFG Z probe). Spectrometers were operated at 750 MHz and 500 MHz, respectively, for ^{1}H detection, and 188 MHz for indirect ^{13}C detection. Carbon resonances were assigned via indirect detection (HSQC and HMBC experiments). Resonances were referenced internally to the most downfield solvent peak (8.74 ppm, pyridine). Default Varian pulse sequences were employed for ^{1}H, COSY, DQF-COSY, TOCSY, HSQC, HMBC, and ROESY experiments. Samples were prepared by dissolving approximately 3-7 mg of cyclothiazomycin C (HPLC-purified and lyophilized) in pyridine-d_5/D_2O (3:1, 600 μL). Pyridine-d_5 (99.94% D) and D_2O (99.9% D) were obtained from Cambridge Isotope Laboratories (Andover, MA). Samples were held at 25 °C during acquisition.

Analysis of NMR data. Assigned resonances are shown in tabular form and directly on the structure within Supplemental Figure 7. Due to the solvent employed (3:1 pyridine-d_5/D_2O), exchangeable peaks (*i.e.* N-H, O-H) were not detected. The corresponding ^1H resonances of the analogous locations in cyclothiazomycin B1 (reported previously(1)) are also given in Supplemental Figure 7 for comparison. Resonances were assigned by 2D NMR spectroscopy, as well as by comparison to the reported spectra of cyclothiazomycin B1 (1).

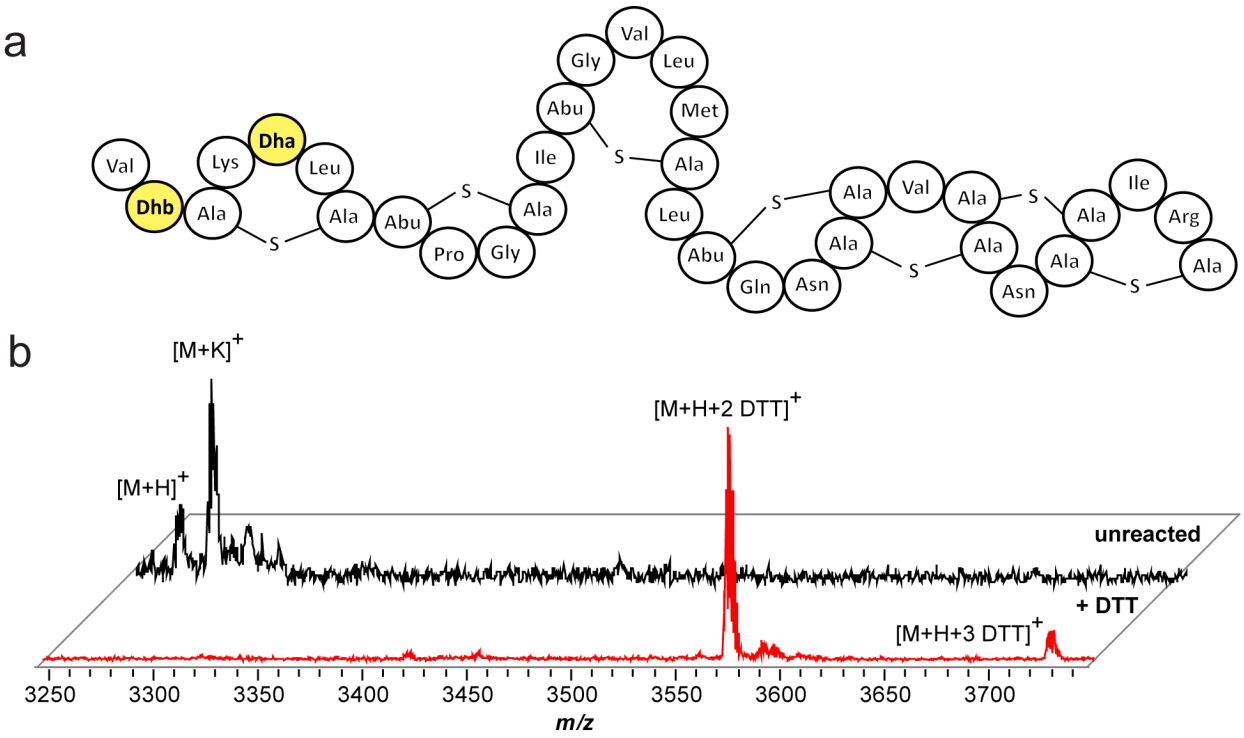
Evaluation of cyclothiazomycin B and C antibiotic activity. *Bacillus subtilis* strain 168, *Bacillus anthracis* strain Sterne, *E. coli* MC4100, and *Pseudomonas putida* KT2440 were grown to stationary phase in 10 mL of Luria-Bertani broth (LB) at 37 °C. *Staphylococcus aureus* USA300 (methicillin-resistant), *Enterococcus faecalis* U503 (vancomycin-resistant), and *Listeria monocytogenes* strain 4b F2365 were grown to stationary phase in 10 mL brain-heart infusion (BHI) medium at 37 °C. *Neisseria sicca* ATCC 29256 was grown to stationary phase in 5 mL of gonococcal broth at 37 °C. The cultures were adjusted to an OD_{600} of 0.013 in the designated medium before being added to 96-well microplates. Successive two-fold dilutions of cyclothiazomycin C or cyclothiazomycin B (standard solution: 5 mg mL^{-1} in DMSO) were added to the cultures (0.5–64 μg mL^{-1}). As a control, kanamycin was added to samples of *E. coli*, *B. subtilis*, *B. anthracis*, *P. putida*, *L. monocytogenes*, and *N. sicca* with dilutions from 1–32 μg mL^{-1}. Gentamycin was used as a control for *S. aureus* and *E. faecalis*. As a negative control, an equal volume of DMSO lacking antibiotic was used. Plates were covered and incubated at 37 °C for 12 h with shaking. The minimum inhibitory concentration (MIC) reported is the value that suppressed all visible growth.

Evaluation of cyclothiazomycin B and C antifungal activity. *Saccharomyces cerevisiae*, *Talaromyces stipitatus,* and *Aspergillus niger* were grown for 36 h in 2 mL of YPD medium (1 L contains 10 g yeast extract, 20 g Peptone and 20 g Dextrose) at 30 °C. *Fusarium virguliforme* was grown for 7 d on potato dextrose agar at 30 °C. Spores were isolated and a suspension of 10^6 spores in potato dextrose broth was added to the 96-well microplate. *S. cerevisiae* cultures were adjusted to an OD_{600} of 0.013 in the designated medium before being added to 96-well microplates. *T. stipitatus,* and *A. niger* were not diluted prior to adding to the 96-well microplate. Successive two-fold dilutions of cyclothiazomycin C and cyclothiazomycin B (standard solution: 5 mg mL^{-1} in DMSO) were added to the cultures (0.5–64 μg mL^{-1}). As a positive control, amphotericin B was added to the cultures with dilutions from 0.5-8 μg mL^{-1}. An equal volume of DMSO was used as a negative control. Plates were covered and incubated at 30 °C for 36 h for *T. stipitatus, A. niger, and S. cerevisiae* or 60 h for *F. virguliforme* with shaking. The minimum inhibitory concentration (MIC) reported is the value that suppressed all visible growth.

II. Supplemental figures

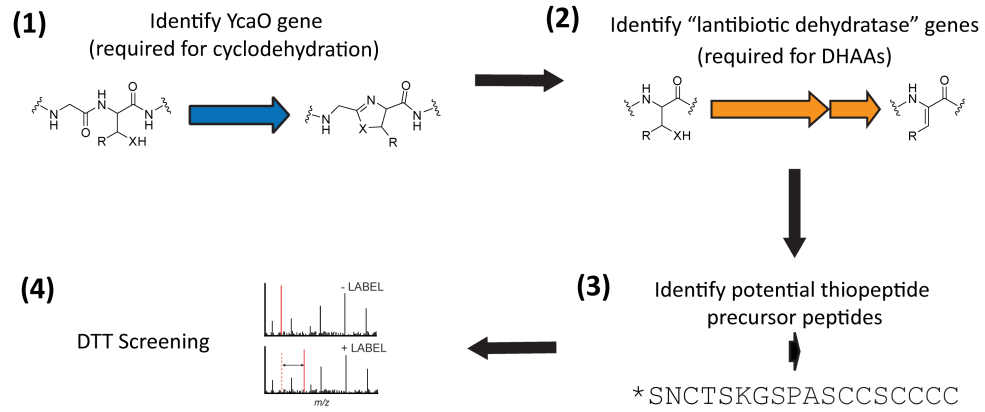


Supplemental Figure 1. Base-dependence of the DTT-labeling reaction. MALDI-TOF MS of pure (commercially-obtained) thiostrepton reacted with DTT in the presence of diisopropylethylamine (DIPEA) (top), or no base (bottom). Thiostrepton was visibly labeled with 1-5 DTT moieties. * denotes peaks not corresponding to DTT labeling.



Supplemental Figure 2. DTT-labeling of geobacillin I in a cell-surface extract. a) Structure of geobacillin I. Dha, dehydroalanine; Dhb, dehydrobutyrine; Ala-S-Ala, lanthionine; Abu-S-Ala, β-methyl-lanthionine. b) Nucleophilic labeling with DTT of geobacillin I within the context of the organic extract of *Geobacillus sp.* M10EXG. Mass spectra of crude unlabeled extract (black spectrum, top) and DTT-labeled material (red spectrum, bottom) are shown. Extent of labeling with DTT is indicated on the bottom spectrum (2 DTT adducts are clearly observed, with the third being a very low intensity ion).

a



b

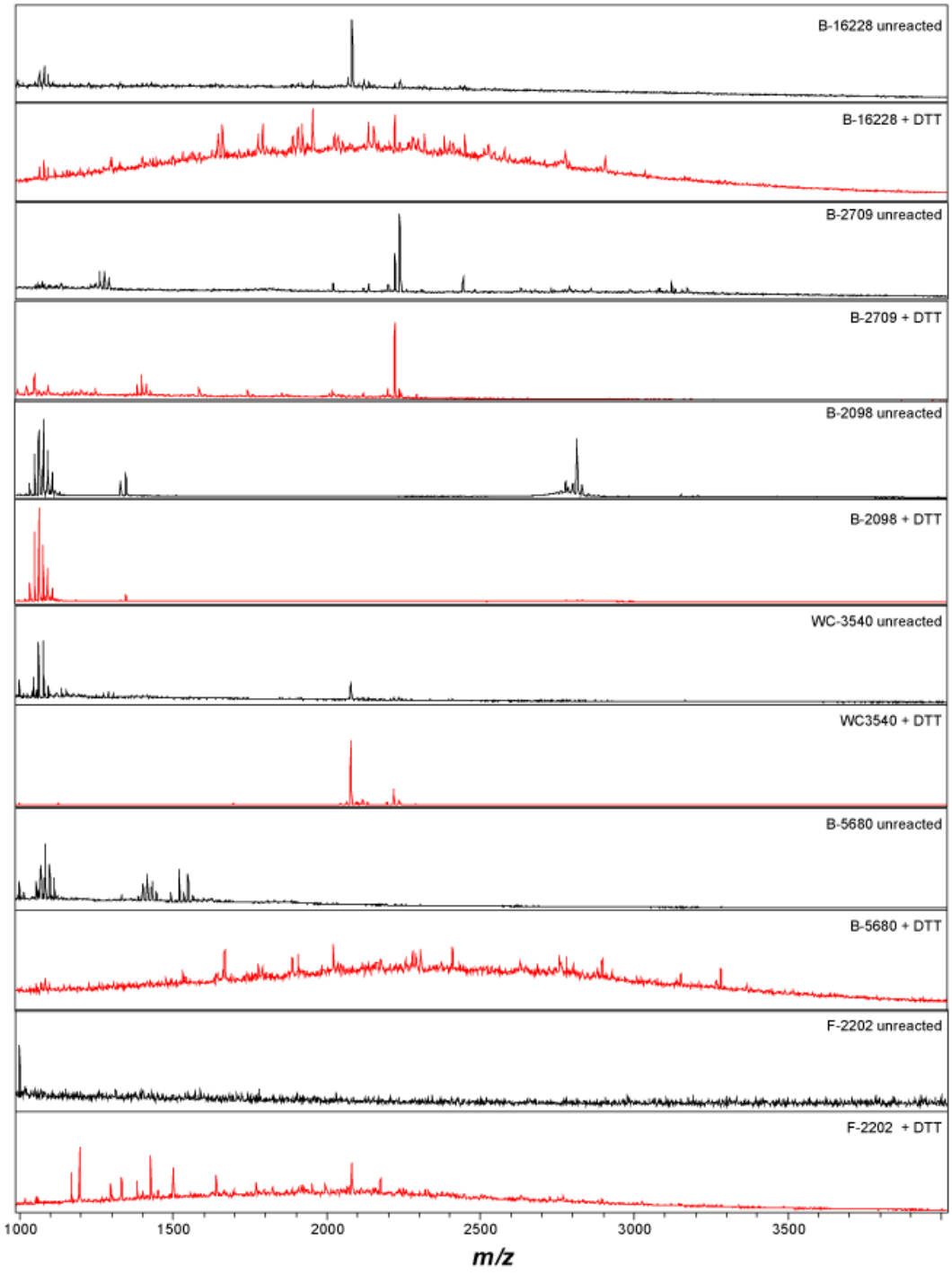
Strain	Core Peptide Prediction
WC-3908	*SNCTSKGSPASCCSCCCC
WC-3480	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
B16228	*SVIIIFGYTEDGEEILTEDSLNSEAAGSGSLSTSSSCC
	*SVSSLLSAACPVSSFSSVSNSVAEVSAE
B-2709	*STTTTGCTTSSCSSSSS
	*SCVGTACACSSTSSSS
	*STSCSSTSTCSSTTSTTSCSA
B-2098	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
	*TFLQALWPFLVVLSCRRSTST
WC-3540	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
	*CGSYCECYCPSPTDDSGRFQPCGDLMGQQCPERCCLVCVDCQCEE
B-5680	*TFSTTSTASTASCPASSAGSVMTANCTG
	*SWFSAASASTASCPGSSAATVGTASTFG
	*SFSSATSASTASCPATSASSLTSASTIG
	*SSVSTAGSVGTASCPATSASSFTSFSSYT
	*GAAAGGTAGGAPQQQAPGGGGCCGGGCCG
F-2202	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
ISP-5257	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
	*LKIGHHMRTVFEQYVEEWQGSSTGN
ISP-5461	*SNCTSTGTPASCCSCCCC
F-3273 (P021)	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC

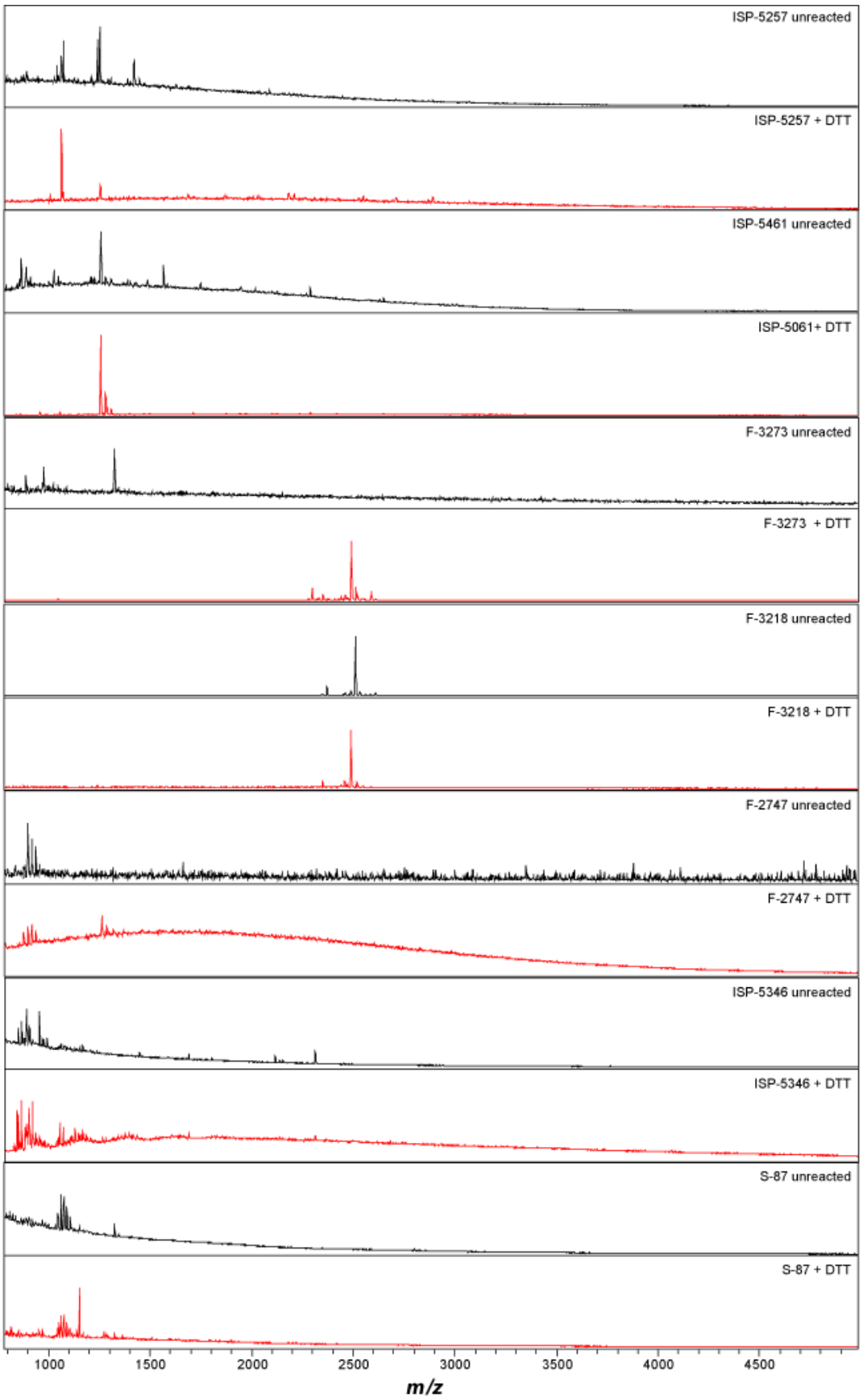
Strain	Core Peptide Prediction
F-3218	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
F-2747	*STSCSSTSTCSSTTSTTSCSA
	*SWGSCSCQGSSSCAQPQVETPTA
ISP-5346	*SSCCCCCPCCSCT
	*SSTSCDSSSCSSSTTSTTSTCA
	*STSCTSCSTTSSTTSCTSCCA
S87	*SWGSCSCQASSSCAQPHVTTPEVL
B-2988	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
B-2098	*TFLQALWPFLVVLSCRRSTST
	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
ISP-5373	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
WC-3540	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
	*CGSYCECYCPSPTDDSGRFQPCGDLMGQQCPERCCLVCVDCQCEE
F-5702	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
F-3273	*STSCSSTSTCSSTTSTTSCSA
	*NGCNCAGCVGCGLNCMSGGWKCATCGSACC
F-2747	*STSCSSTSTCSSTTSTTSCSA
	*VSVSSLLSAACPVSSFSSVSNSVAEVAA
WC-3779	*STSCSSTSTCSSTTSTTSCSA

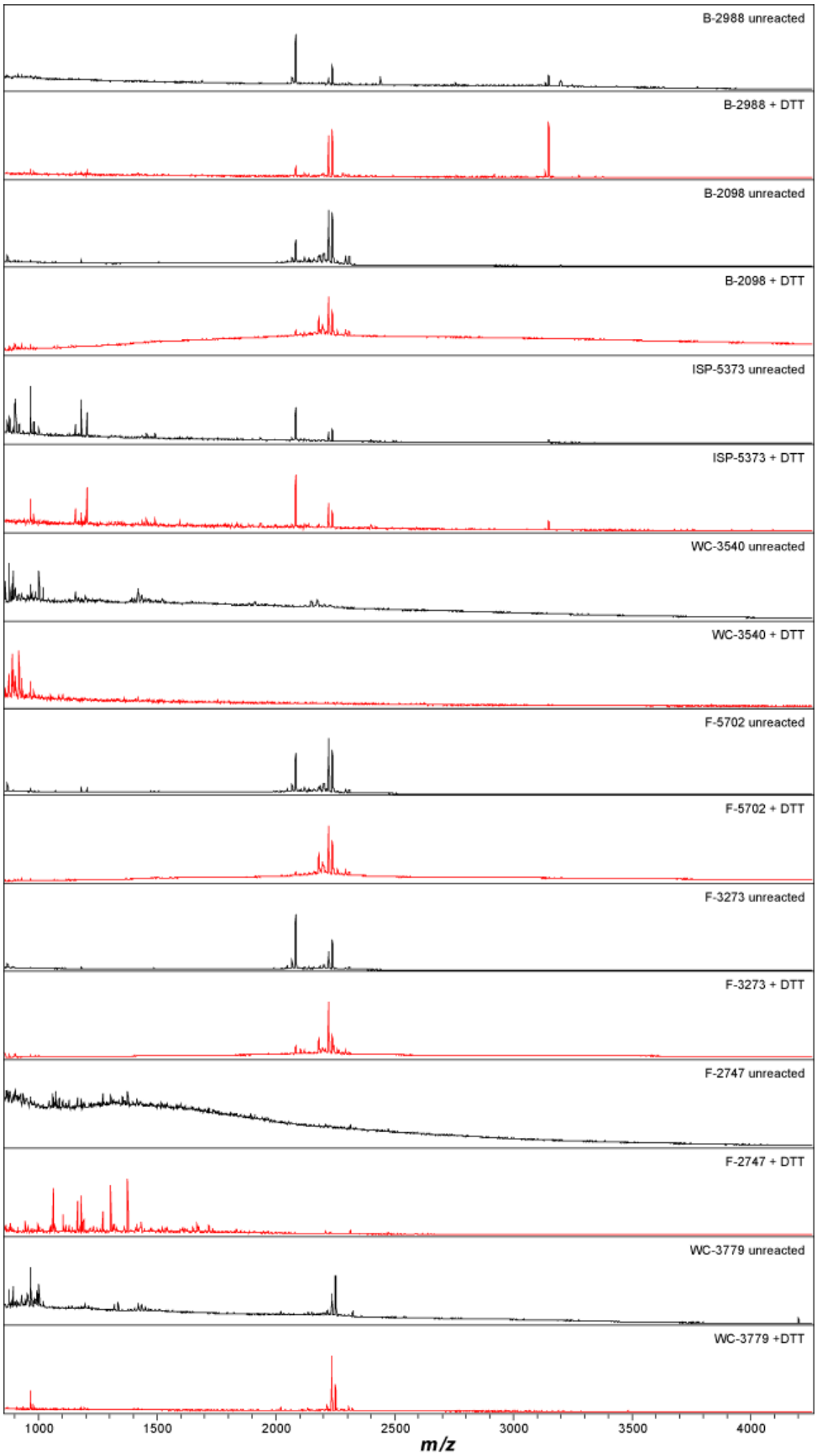
Supplemental Figure 3. Bioinformatic prioritization of strains. a) Bioinformatics prioritization schematic. 1) A list is populated with strains encoding a thiazole/oxazole-modified microcin (TOMM) cyclodehydratase "YcaO" necessary for the heterocyclization of select Cys, Ser, and Thr, residues. 2) The list of strains is then trimmed to only contain strains that also harbor a "lantibiotic" dehydratase in close proximity (within 10 open reading frames on either side) to the YcaO protien. 3) TOMM-like precursor peptides from the trimmed list are then identified, and the mass of the final natural product is predicted for use in the dereplication process. (4) If strains make it through steps 1-3, reactivity-based screening with DTT is utilized to identify natural products of interest. b) Predicted core regions of the precursor peptides

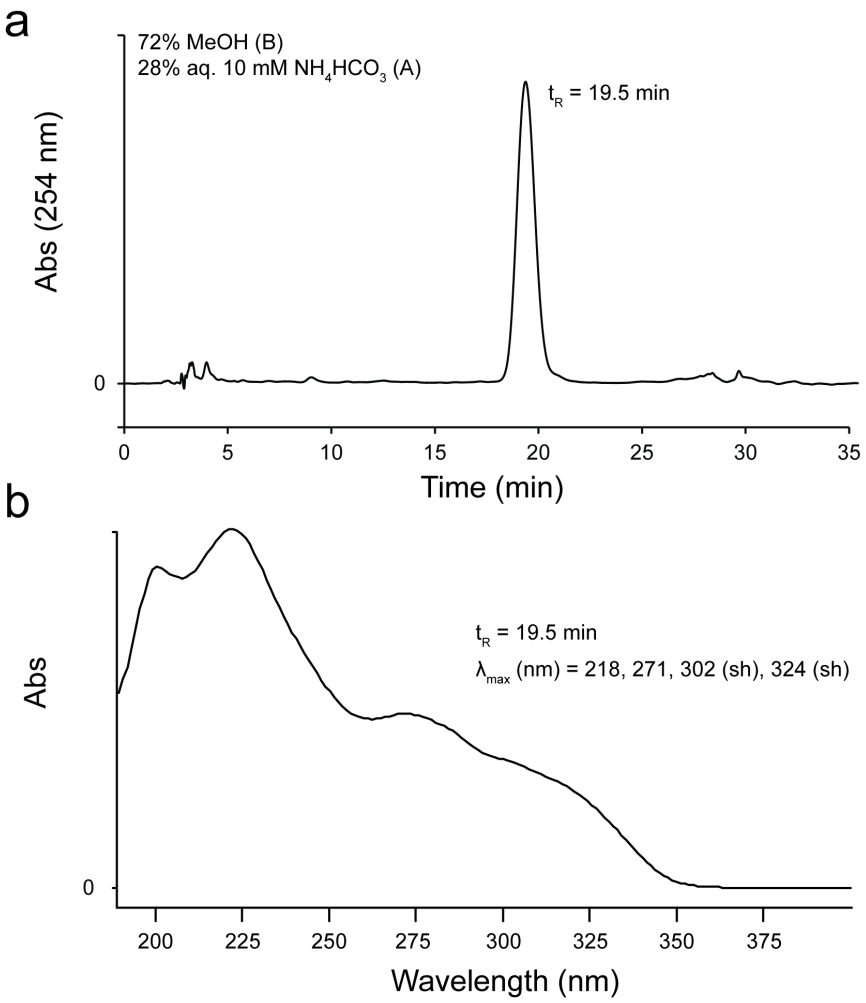
identified in the 23 strains prioritized and screened using the DTT labeling method. Highlighted in red are the precursor peptides predicted from WC-3908 (the producer of cyclothiazomycin C) and WC-3480 (the producer of grisemycin).

Supplemental Figure 4 (Below). Mass spectra of strains screened by the DTT labeling method. Mass spectrometry data (*m/z* 900 – 4200 Da) is shown for all strains screened except *Streptomyces griseus* subsp. *griseus* and WC-3908 (shown as figures 4 and 5 in the main text, respectively). The mass spectra of the unreacted organic cell-surface extracts are shown in black with the corresponding DTT-reacted extracts in red. Each spectrum is labeled according to the strain designation (NRRL identifier) and whether or not DTT/DIPEA was added. NRRL, Northern Regional Research Laboratory collection, which is curated by the Agricultural Research Service under the supervision of the U.S. Department of Agriculture (USDA/ARS).

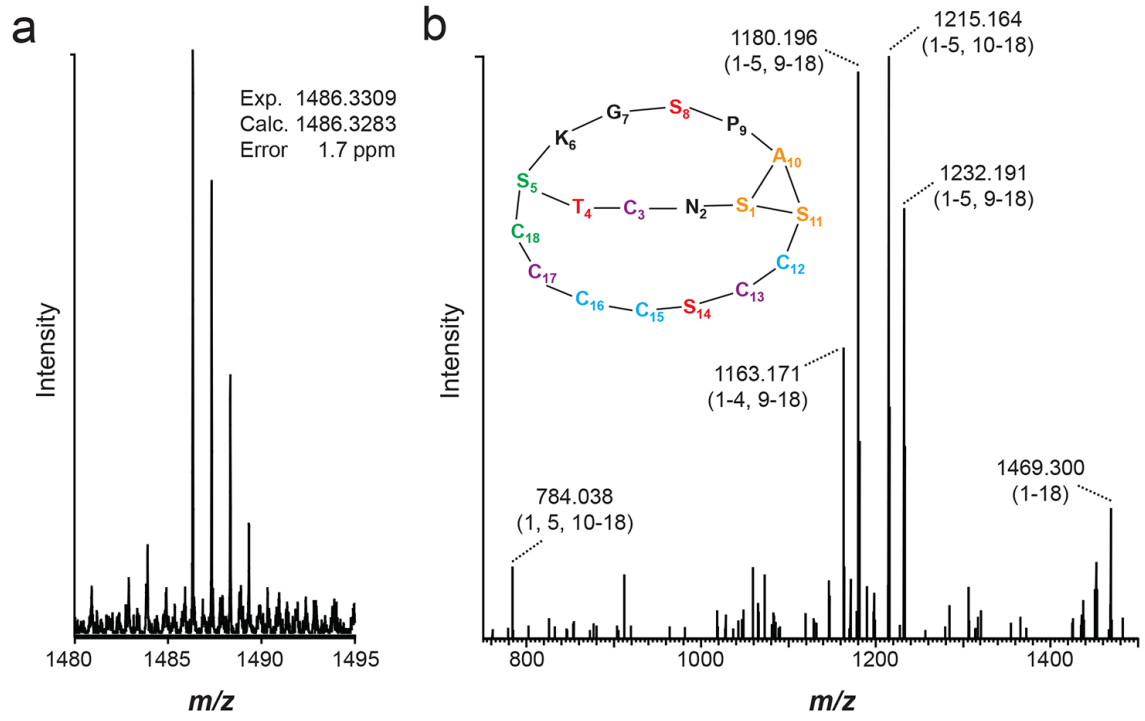

B-16228 unreacted
B-16228 + DTT
B-2709 unreacted
B-2709 + DTT
B-2098 unreacted
B-2098 + DTT
WC-3540 unreacted
WC3540 + DTT
B-5680 unreacted
B-5680 + DTT
F-2202 unreacted
F-2202 + DTT
1000
1500
2000
2500
3000
3500
m/z


ISP-5257 unreacted
ISP-5257 + DTT
ISP-5461 unreacted
ISP-5061+ DTT
F-3273 unreacted
F-3273 + DTT
F-3218 unreacted
F-3218 + DTT
F-2747 unreacted
F-2747 + DTT
ISP-5346 unreacted
ISP-5346 + DTT
S-87 unreacted
S-87 + DTT
1000
1500
2000
2500
3000
3500
4000
4500
m/z


B-2988 unreacted
B-2988 + DTT
B-2098 unreacted
B-2098 + DTT
ISP-5373 unreacted
ISP-5373 + DTT
WC-3540 unreacted
WC-3540 + DTT
F-5702 unreacted
F-5702 + DTT
F-3273 unreacted
F-3273 + DTT
F-2747 unreacted
F-2747 + DTT
WC-3779 unreacted
WC-3779 +DTT
1000
1500
2000
2500
3000
3500
4000
m/z



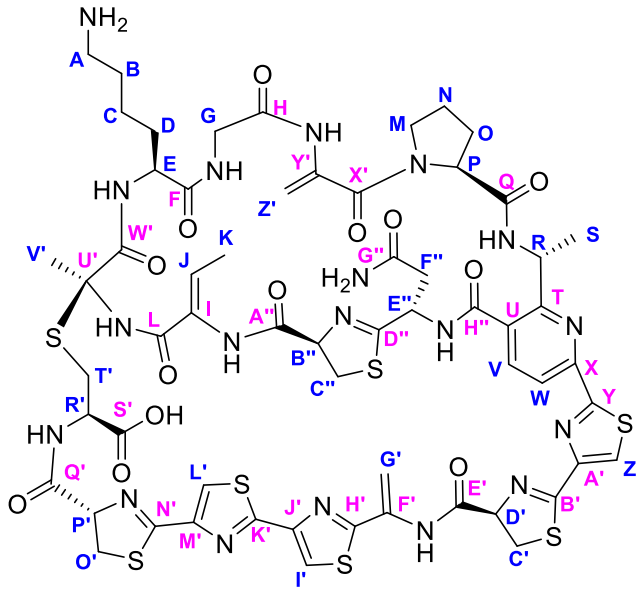
Supplem/ental Figure 5. HPLC trace and UV spectrum of cyclothiazomycin C. a) A sample (spatula tip) of purified cyclothiazomycin C was dissolved in 50% MeOH (B)/aq. 10 mM NH_4HCO_3 (A) (100 µL). An aliquot (20 µL) was analyzed by HPLC (isocratic 72% B for 35 min). Photodiode array (PDA) detection was used to monitor absorbance (abs) from 190-400 nm. A blank injection was also run and subtracted from the cyclothiazomycin C chromatogram; the resulting spectrum with UV monitoring at 254 nm is shown. b) Cyclothiazomycin C exhibits UV absorbance consistent with that reported for cyclothiazomycin A and B1/B2.(1,2) A UV spectrum (PDA) from the HPLC trace at 19.5 min is shown (sh, shoulder).



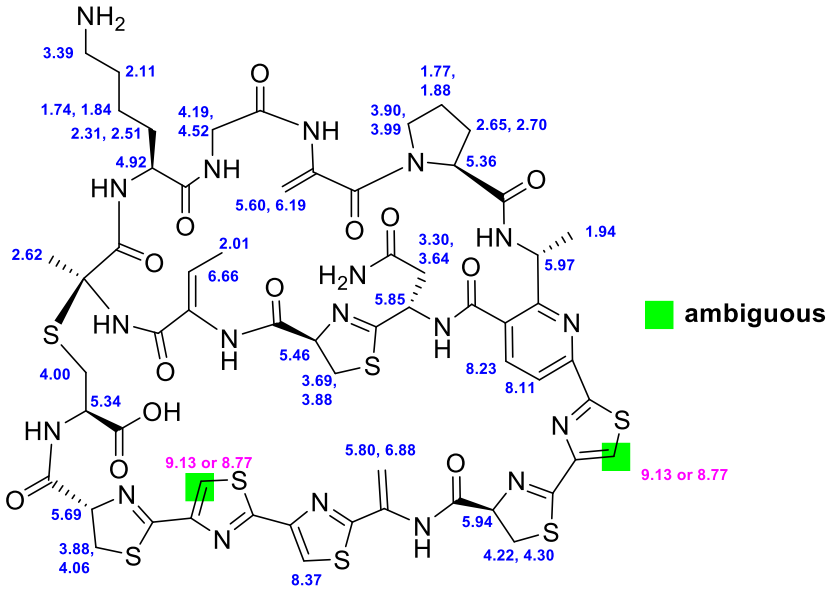
Supplemental Figure 6. High resolution Fourier transform mass spectrometry (FT-MS) analysis of cyclothiazomycin C. a) The *m/z* scan of purified cyclothiazomycin C showed an ion in the 1^+ charge state with an observed isotopic *m/z* value with < 2 ppm error from the calculated value for cyclothiazomycin C. b) CID spectrum of *m/z* 1486. The monoisotopic mass values are given for assigned peak predictions. The number ranges given below the mass values refer to a shorthand notation describing predicted fragments of cyclothiazomycin C. A key for the shorthand notation for the structure of cyclothiazomycin C is given in pictorial format using single letter codes for the amino acids, the residue's *N* to *C* position, and lines depicting molecular connectivity within the mature structure. The colors used for the shorthand notation depict the modification present at a particular residue. Purple, thiazoline moieties; green, thioether linkage; cyan, thiazole moieties; red, dehydrated amino acids; orange, pyridine moiety; black, unmodified amino acids.

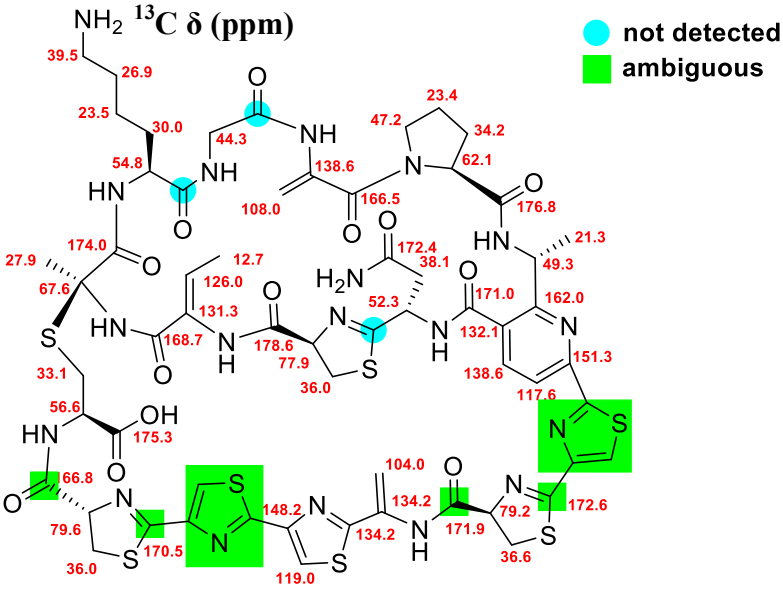
Supplemental Figure 7 (Below). NMR assignments. Assignments of ^{1}H and ^{13}C resonances are given. a) The labeling scheme below depicts the lettering system utilized in the table. b) Peak assignments are shown directly on the structure of cyclothiazomycin C. Sites where a resonance could not be unambiguously assigned or was not detected are noted. Note that the resonances corresponding to two of the thiazole systems could not be precisely assigned. c) Diagram of connectivity established via 2D correlational experiments. Observed correlations are indicated by red arrows ($^{1}H/^{13}C$ HMBC correlations) or thick black bonds (COSY or TOCSY correlations). Significant geminal $^{1}H/^{1}H$ correlations observed by COSY (blue circles) or TOCSY (green squares) are indicated. d) Table of NMR peak assignments. ^{1}H NMR shifts of analogous positions on cyclothiazomycin B1 in the same solvent system are shown in the table for comparison. Observed 2D correlations are listed.

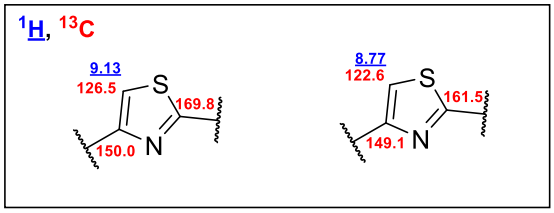
(a) Labeling scheme for NMR peak assignments



(b) Peak assignments ^{1}H δ (ppm)



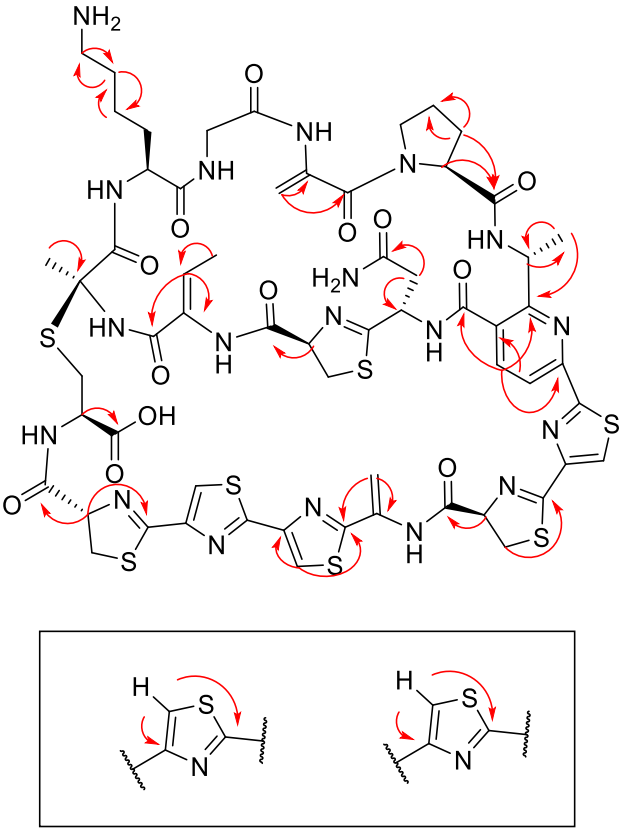




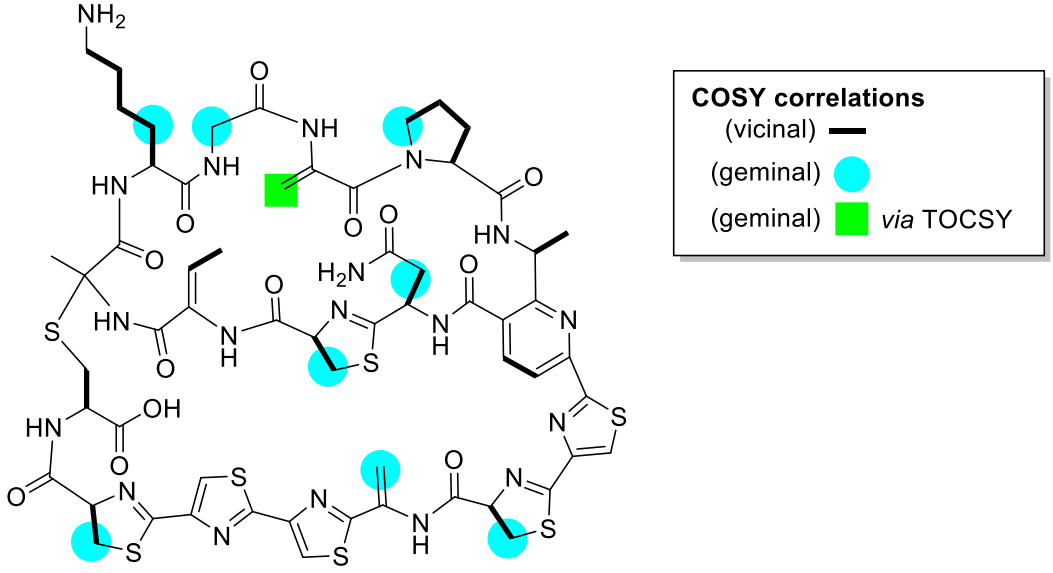
Unassigned thiazole systems

(c) Through-molecule correlations from 2D experiments

^{1}H-^{13}C HMBC connectivity



^{1}H-^{1}H COSY/TOCSY connectivity



(d) NMR assignments for cyclothiazomycin C

Position		^{1}H δ, ppm (*J*)	^{13}C δ, ppm	^{1}H δ, ppm (cycloB1)(1)	COSY	DQF-COSY	HMBC	TOCSY	ROESY
A	Lys-ε	3.39 m	39.5		B	B	B	B, C_a, C_b, D_a, D_b, E	B, C_a, C_b
B	Lys-δ	2.11 m	26.9		A, D_b	A, D_b, C_b	C, A	A, C_a, C_b, D_a, D_b, E	A
C	Lys-γ	a: 1.74 m	23.5					A, B, D_a, D_b, E	A, C_b, E
		b: 1.84 m			D_b	D_b, B			A, C_a
D	Lys-β	a: 2.31 m	30		E	D_b		A, B, C_a, C_b, E	D_b
		b: 2.51 m			C_b, B, E	D_a, D_b, B, E			B, D_a
E	Lys-α	4.92 dd (10.8, 4.2 Hz)	54.8	5.01	D_b, D_a	D_b		A, B, C_a, C_b, D_a, D_b	B, C_a, D_a
F	Lys-CO	-	n.d.		-	-	-	-	-
G	Gly-α	a: 4.19 m	44.3	4.17	G_b	G_b		G_b	G_b
		b: 4.52 d (16.5 Hz)		4.27	G_a	G_a		G_a	G_a
H	Gly-CO	-	n.d.		-	-	-	-	-
I	Dhb-α	-	131.3		-	-	-	-	-
J	Dhb-β	6.66 q (6.9 Hz)	126.0	6.68	K	K	I, L	K	K
K	Dhb-Me	2.01 d (7.0 Hz)	12.7	2.02	J	J	J	J	J
L	Dhb-CO	-	168.7		-	-	-	-	-
M	Pro-δ	a: 3.90 m	47.2	3.73	M_b, N	M_b		N, O	
		b: 3.99 m		3.93	M_a, N	M_a		N, O, P	P
N	Pro-γ	1.77 m, 1.88 m	23.4	1.50, 1.78	M_a, M_b, O	O	O	M_a, M_b, O, P	
O	Pro-β	2.65 m, 2.70 m	34.2	2.45, 2.85	P, N	P, N	N, Q	M_a, M_b, N, P	P
P	Pro-α	5.36 m	62.1	5.23	O	O		M_a, M_b, N, O	M_b, O
Q	Pro-CO	-	176.8		-	-	-	-	-
R	Pyr-Ala-α	5.97 m	49.3	6.08	S	S	S	S	S
S	Pyr-Ala-Me	1.94 d (6.4 Hz)	21.3	1.93	R	R		R	R
T	Pyr-6	-	162		-	-	-	-	-
U	Pyr-5	-	132.1		-	-	-	-	-
V	Pyr-4	8.23 d (8.0 Hz)	138.6	8.23	W	W	H'', T, X	W	
W	Pyr-3	8.11 d (8.0 Hz)	117.6	8.11	V	V	U	V	
X	Pyr-2	-	151.3		-	-	-	-	-
B'	Tzn-A-2	-	172.6**		-	-	-	-	-
C'	Tzn-A-5	a: 4.30 t (11.1 Hz)	36.6	4.28	C'$_b$	D', C'$_b$	B'	D'	
		b: 4.22 m		4.16	C'$_a$	D', C'$_a$	B'	D'	D'

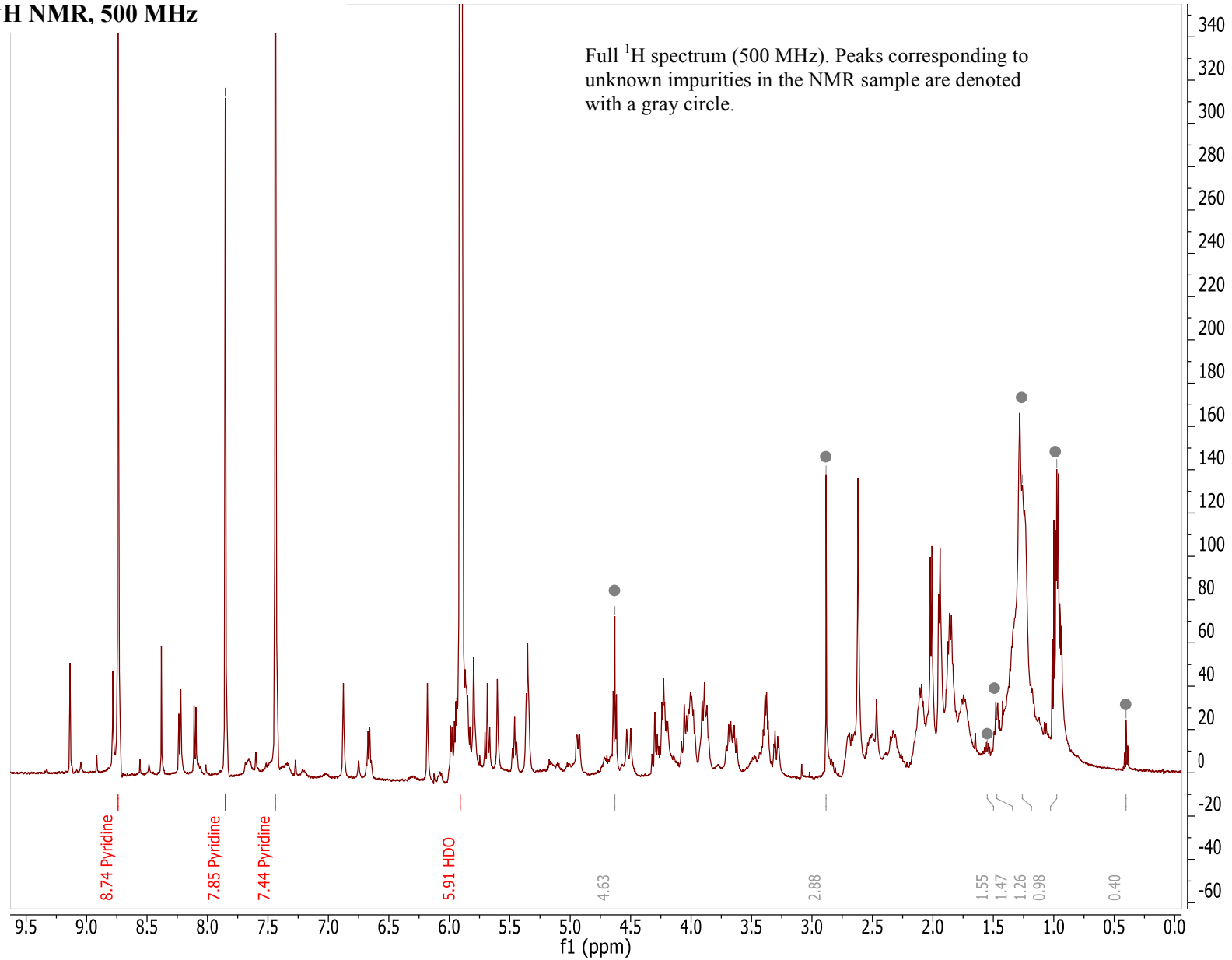
D’	Tzn-A-4	5.94 m	79.2	5.90		C'_a, C'_b		C'_a, C'_b	C'_b
E’	Tzn-A-CO	-	171.9**		-	-	-	-	-
F’	Dha-A-α	-	134.2		-	-	-	-	-
G’	Dha-A-β	a: 6.88 s	104.0	6.85	G'_b		F’	G'_b	G'_b
		b: 5.80 s		5.80	G'_a		H’	G'_a	G'_a
H’	Tzl-B-2	-	165.9		-	-	-	-	-
I’	Tzl-B-5	8.37 s	119.0	8.38			H’, J’		
J’	Tzl-B-4	-	148.2		-	-	-	-	-
N’	Tzn-B-2	-	166.8**		-	-	-	-	-
O’	Tzn-B-5	a: 4.06 t (10.3 Hz)	35.98	4.04	P’, O'_b	P’, O'_b		P’	P’, O'_b
		b: 3.88 m		3.95	P’, O'_a	P’, O'_a		P’	O'_a
P’	Tzn-B-4	5.69 t (9.8 Hz)	79.6	5.66	O'_a, O'_b	O'_a, O'_b	Q’, N’	O'_a, O'_b	O'_a
Q’	Tzn-B-CO	-	170.5**		-	-	-	-	-
R’	Cys-α	5.34 m	56.6	5.33	T’	T’		T’	T’
S’	Cys-CO	-	175.3		-	-	-	-	-
T’	Cys-β	4.00 m	33.1	3.95, 4.10	R’	R’		R’	R’, V’
U’	M-Ala-α	-	67.6		-	-	-	-	-
V’	M-Ala-Me	2.62 s	27.9	2.57			U’, W’		T’
W’	M-Ala-CO	-	174.0		-	-	-	-	-
X’	Dha-B-CO	-	166.5		-	-	-	-	-
Y’	Dha-B-α	-	138.62		-	-	-	-	-
Z’	Dha-B-β	a: 6.19 s	108				Y’ X’	Z'_b	Z'_b
		b: 5.60 s					X’	Z'_a	Z'_a
A’’	Tzn-C-CO	-	178.6**		-	-	-	-	-
B’’	Tzn-C-4	5.46 t (8.7 Hz)	77.9	5.40	C''_a, C''_b	C''_a, C''_b		C''_a, C''_b	C''_a
C’’	Tzn-C-5	a: 3.69 m	35.97	3.62	B’’, C''_b	B’’, C''_b		B’’, C''_b	B’’
		b: 3.88 m		3.75	B’’, C''_a	B’’, C''_a		B’’, C''_a	
D’’	Tzn-C-2	-	nd		-	-	-	-	-
E’’	Asn-α	5.85 m	52.3	5.80	F''_b, F''_a	F''_b		F''_a, F''_b	F''_a, F''_b
F’’	Asn-β	a: 3.30 dd (15.3, 4.9 Hz)	38.1	3.21	E’’, F''_b	F''_b	G''_a	E’’, F''_b	F''_b
		b: 3.64 dd (15.3, 10.1 Hz)		3.59	E’’, F''_a	E’’, F''_a	E''_b, G''_b	E’’, F''_a	F''_a
G’’	Asn-CO	-	172.4		-	-	-	-	-
H’’	Pyr-CO	-	171.0		-	-	-	-	-

Unassigned thiazole 1									
a	Tzl-U1-2	-	169.8		-	-	-	-	-
b	Tzl-U1-5	9.13 s	126.49	9.08			a, c		
c	Tzl-U1-4	-	150		-	-	-	-	-
Unassigned thiazole 2									
d	Tzl-U2-2	-	161.4		-	-	-	-	-
e	Tzl-U2-5	8.77 s	122.57	8.73			d, f		
f	Tzl-U2-4	-	149.1		-	-	-	-	-

Abbreviations: Dha, dehydroalanine; Dhb, dehydrobutyrene; Pyr, pyridine; Tzn, thiazoline; Tzl, thiazole; U, unknown; s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; n.d., not detected; cycloB1, cyclothiazomycin B1, ** ambiguous assignments

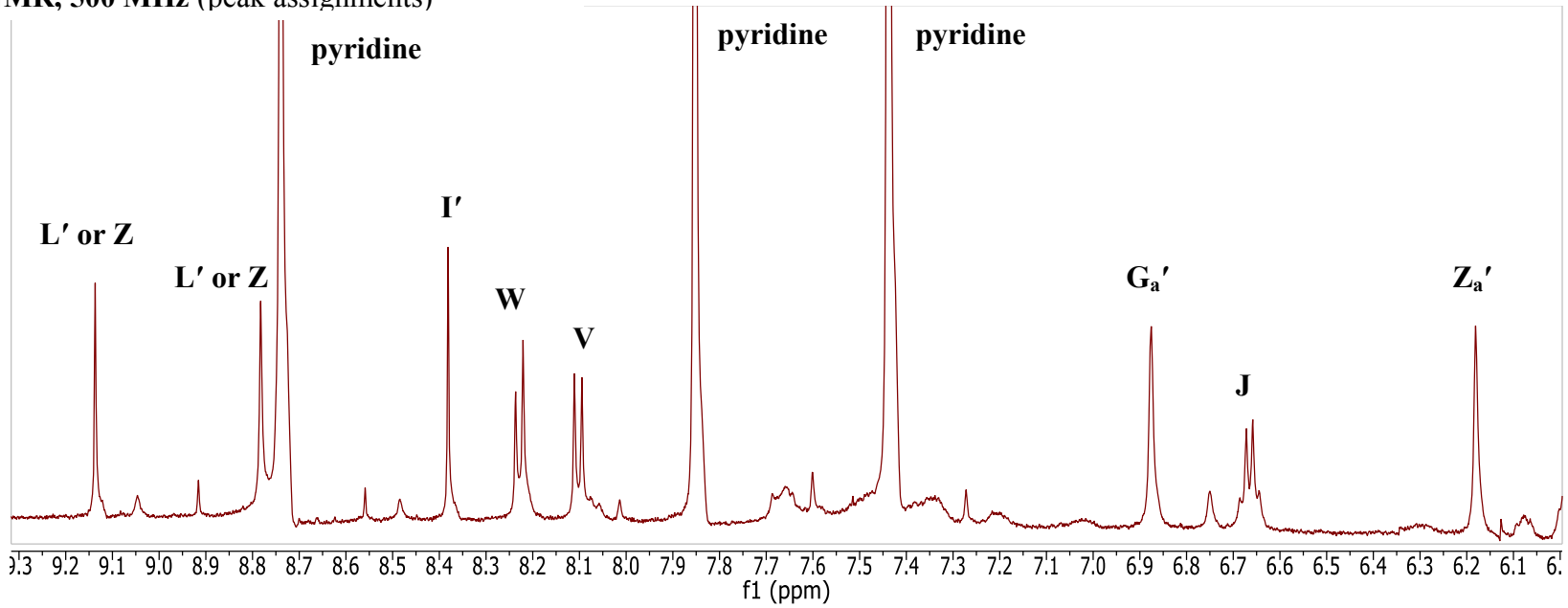
Supplemental Figure 8. NMR spectra of cyclothiazomycin C. Complete NMR spectra (^{1}H, COSY, TOCSY, HSQC, HMBC, and ROESY) are shown in the subsequent pages.

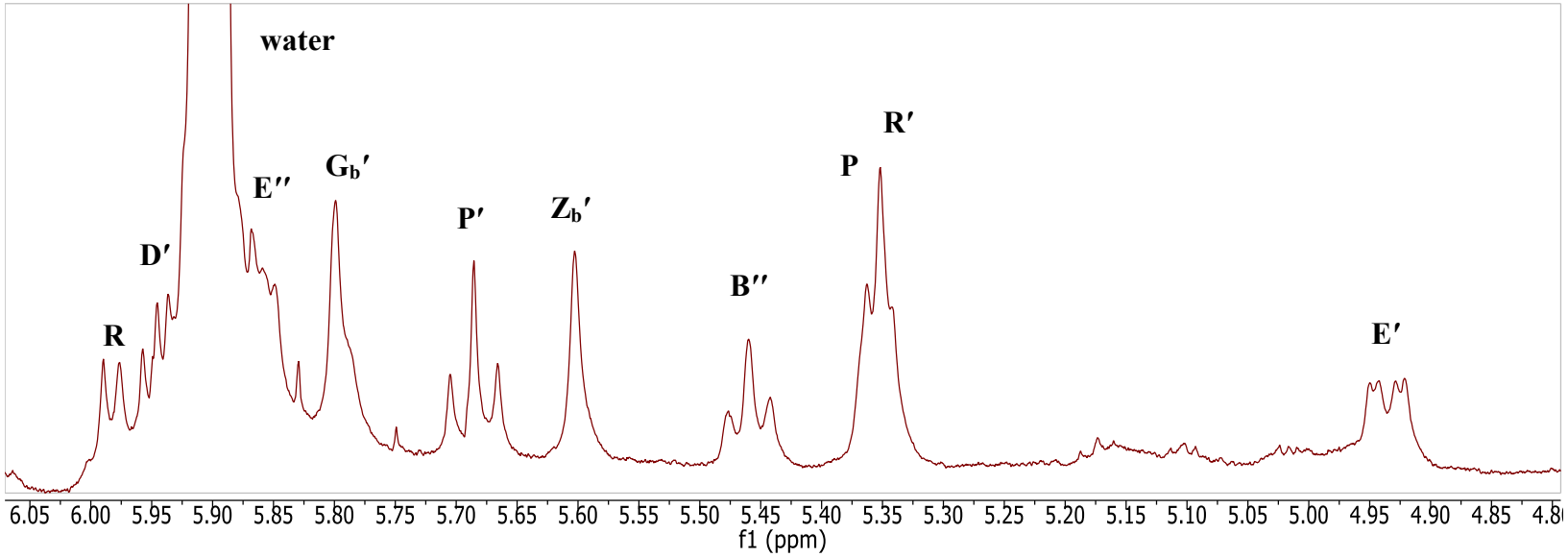
^{1}H NMR, 500 MHz

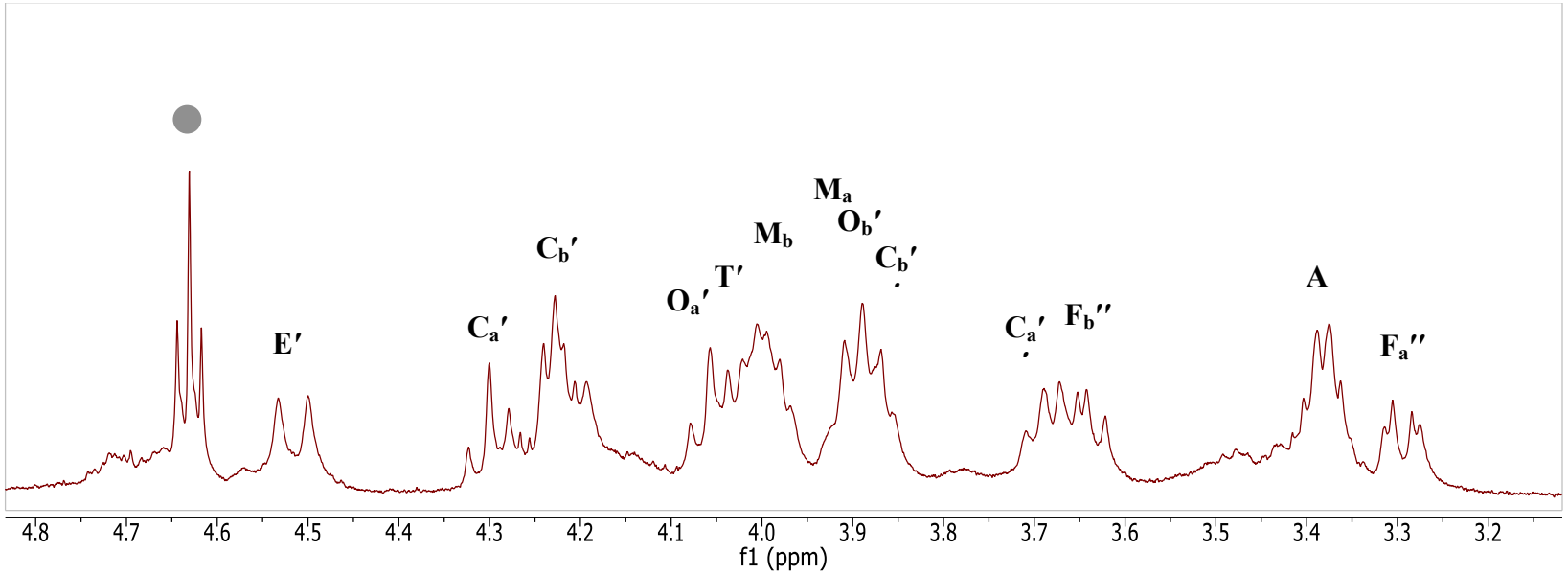


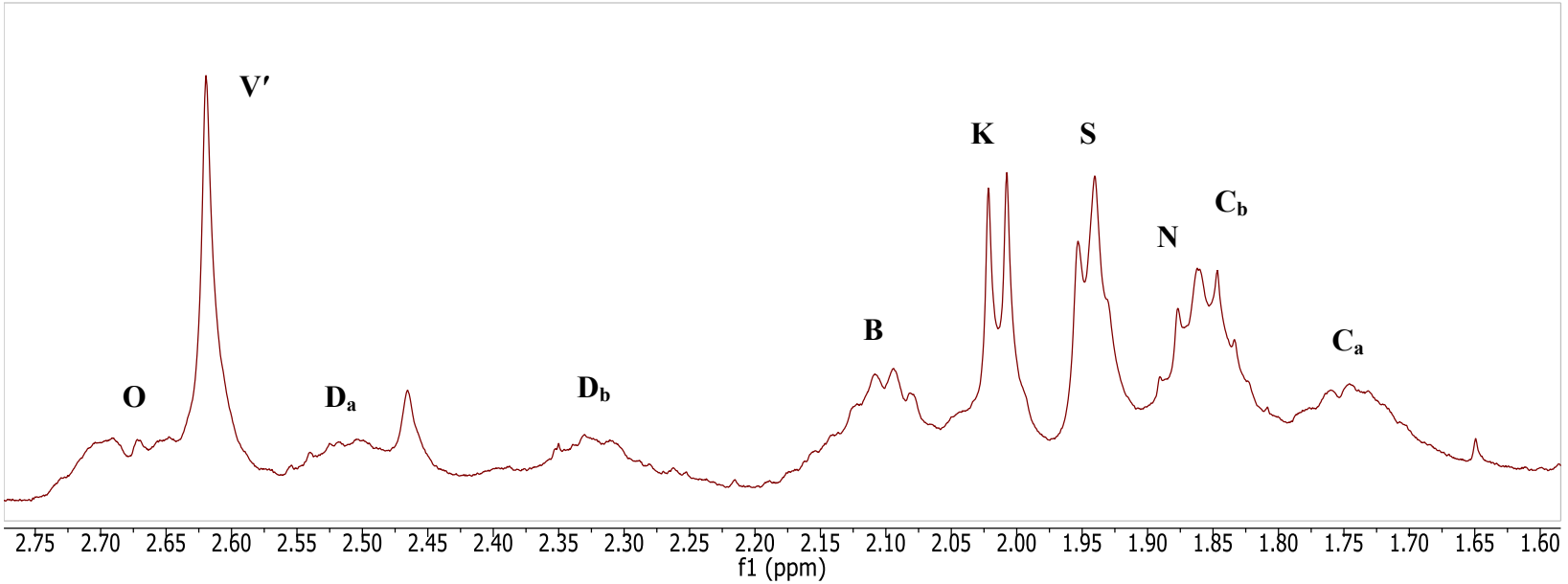
Full ^{1}H spectrum (500 MHz). Peaks corresponding to unknown impurities in the NMR sample are denoted with a gray circle.

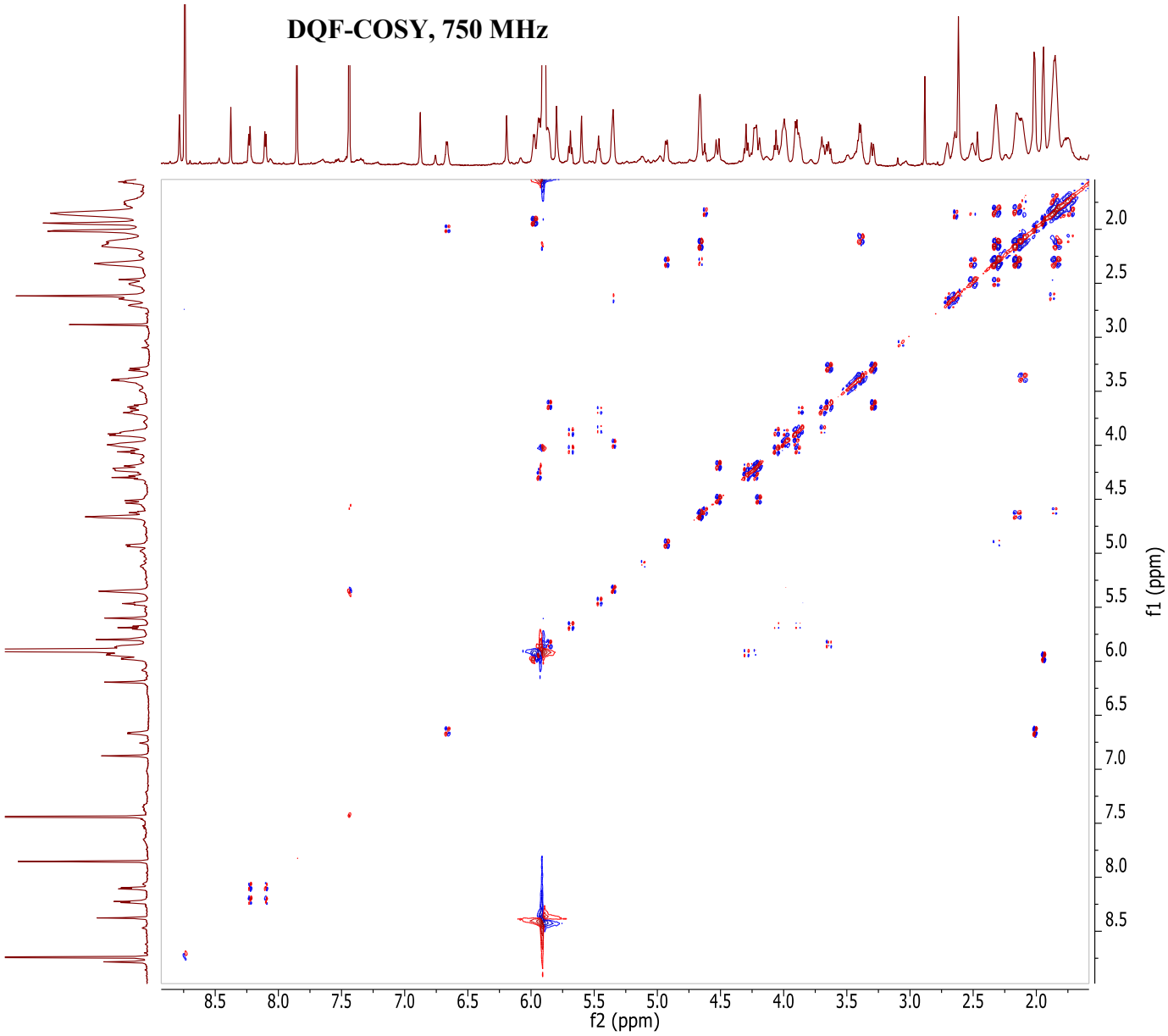
^{1}H NMR, 500 MHz (peak assignments)



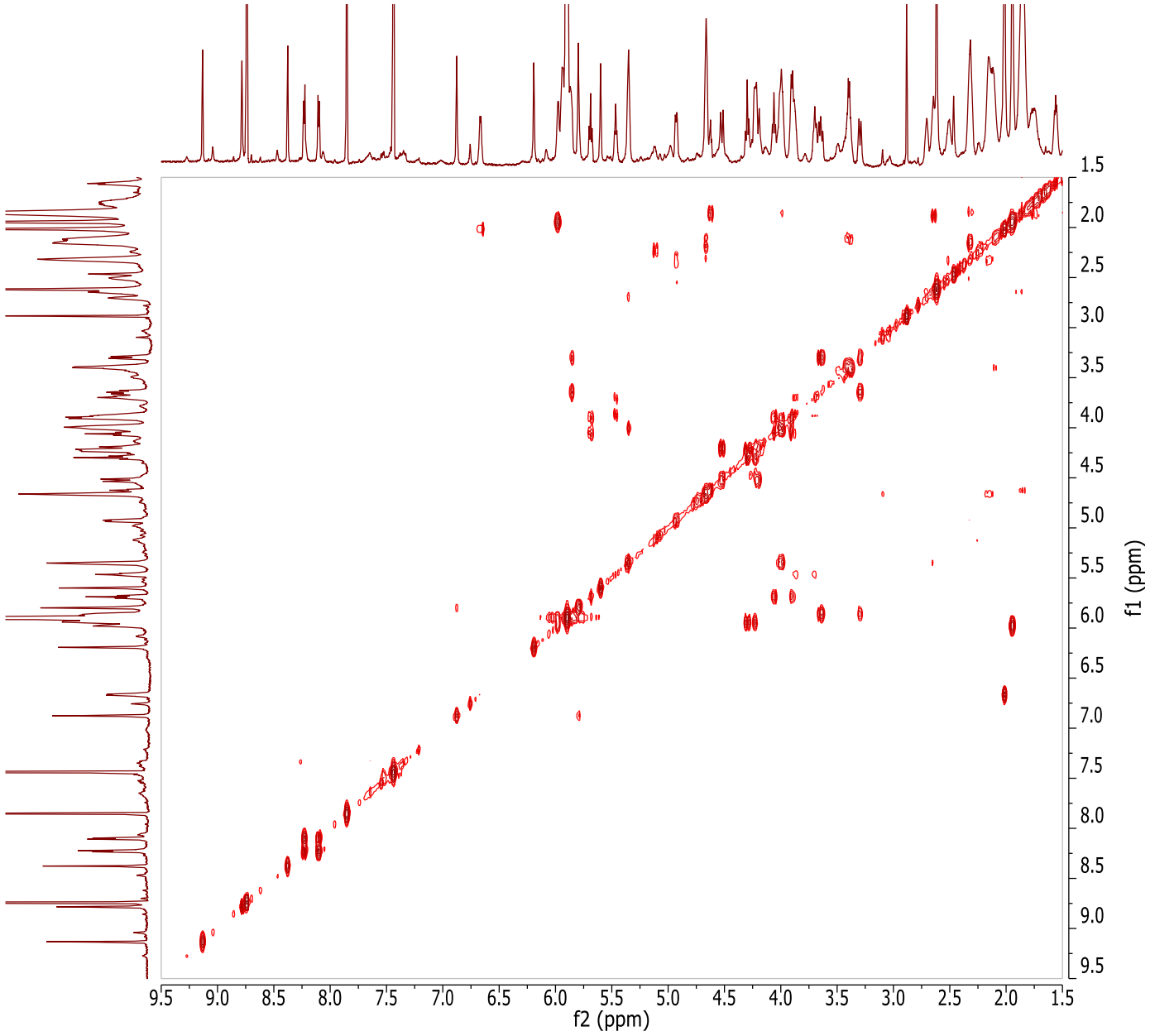



E′
C_a′
C_b′
O_a′
T′
M_b
M_a
O_b′
C_b′
C_a′
F_b″
A
F_a″
4.8 4.7 4.6 4.5 4.4 4.3 4.2 4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3 3.2
f1 (ppm)

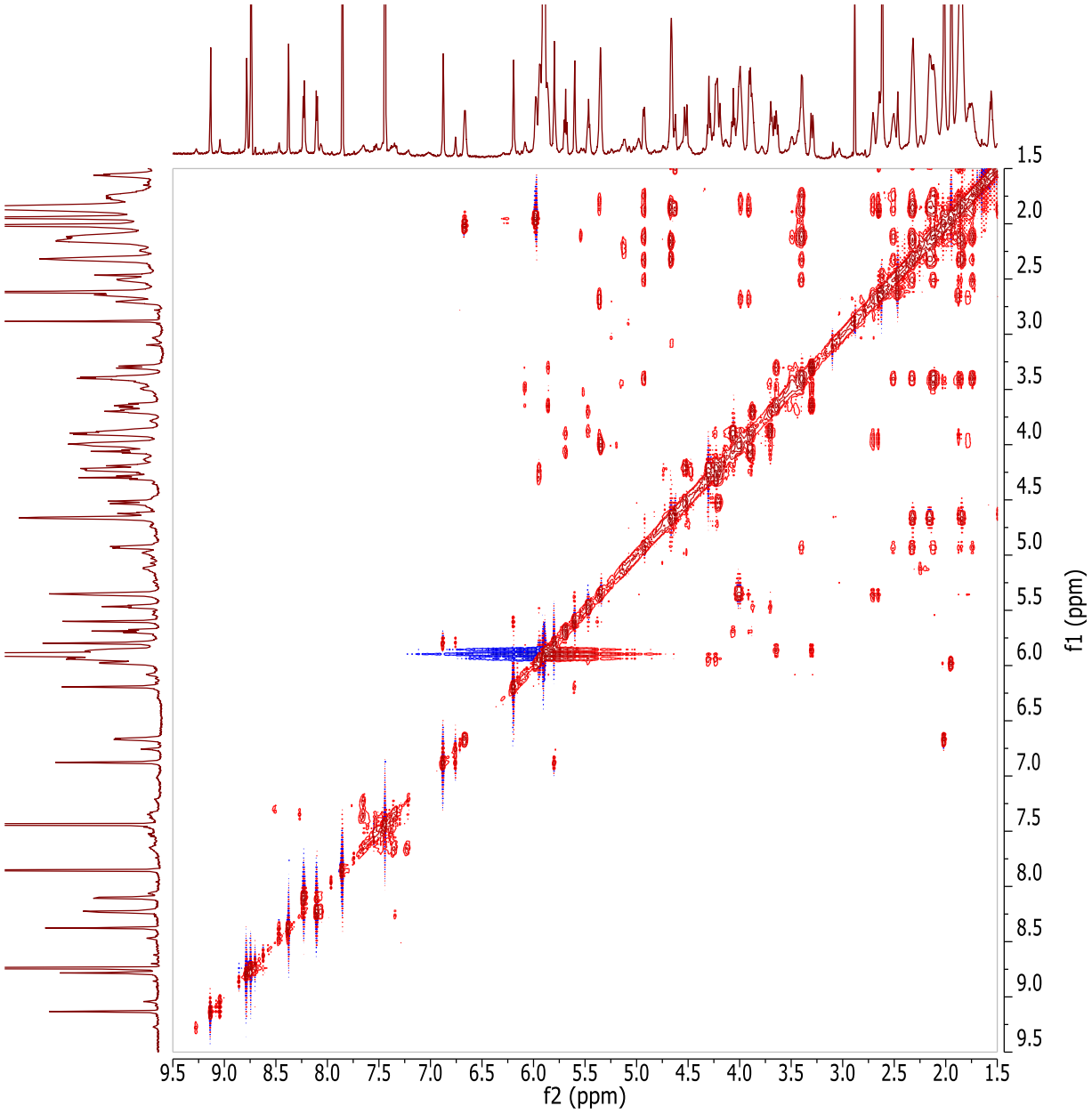

V′
O
D_a
D_b
B
K
S
N
C_b
C_a
2.75 2.70 2.65 2.60 2.55 2.50 2.45 2.40 2.35 2.30 2.25 2.20 2.15 2.10 2.05 2.00 1.95 1.90 1.85 1.80 1.75 1.70 1.65 1.60
f1 (ppm)


DQF-COSY, 750 MHz
f2 (ppm)
f1 (ppm)
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0

COSY,
750 MHz

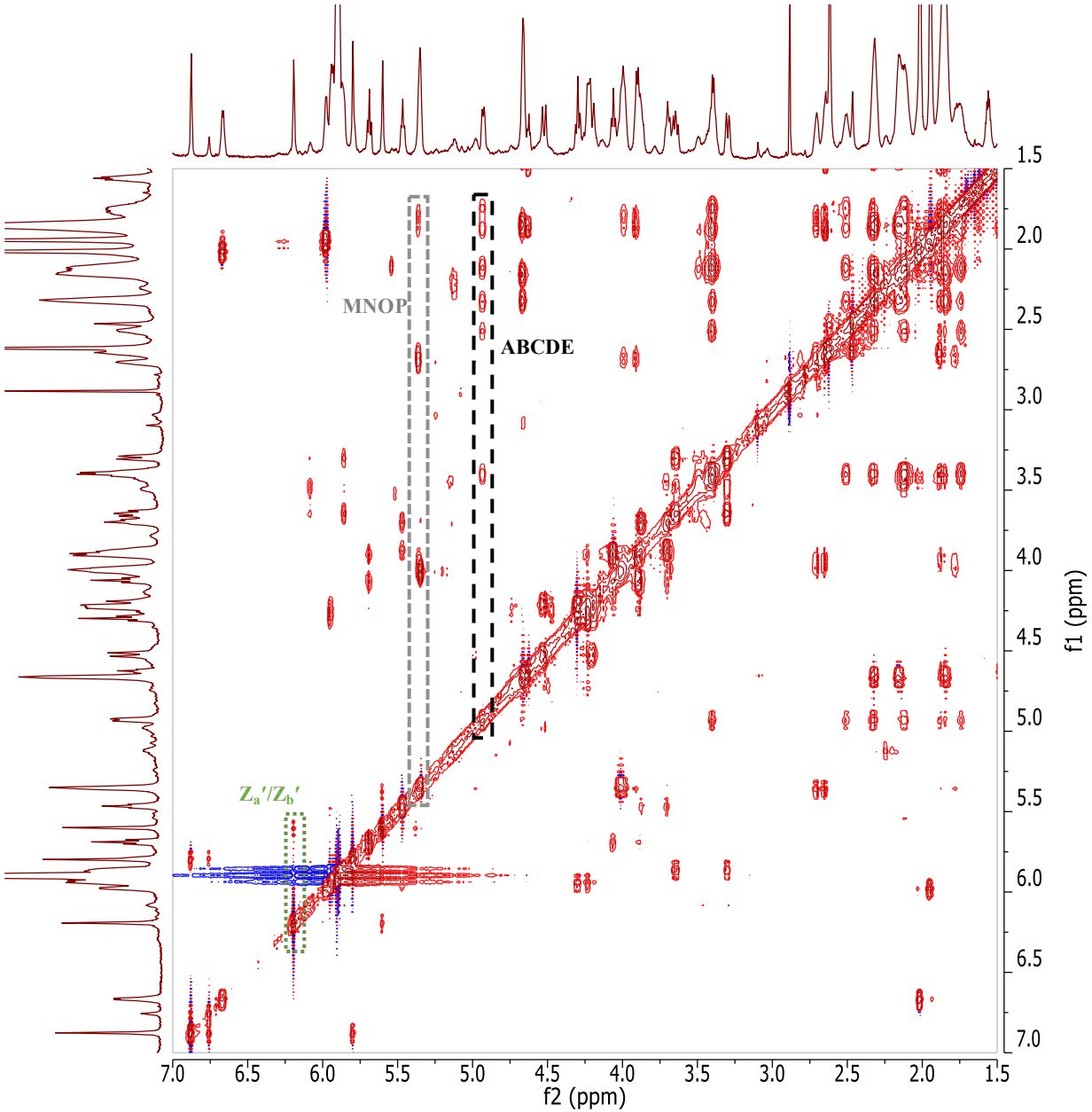

9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
f2 (ppm)
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
f1 (ppm)

TOCSY,
750 MHz

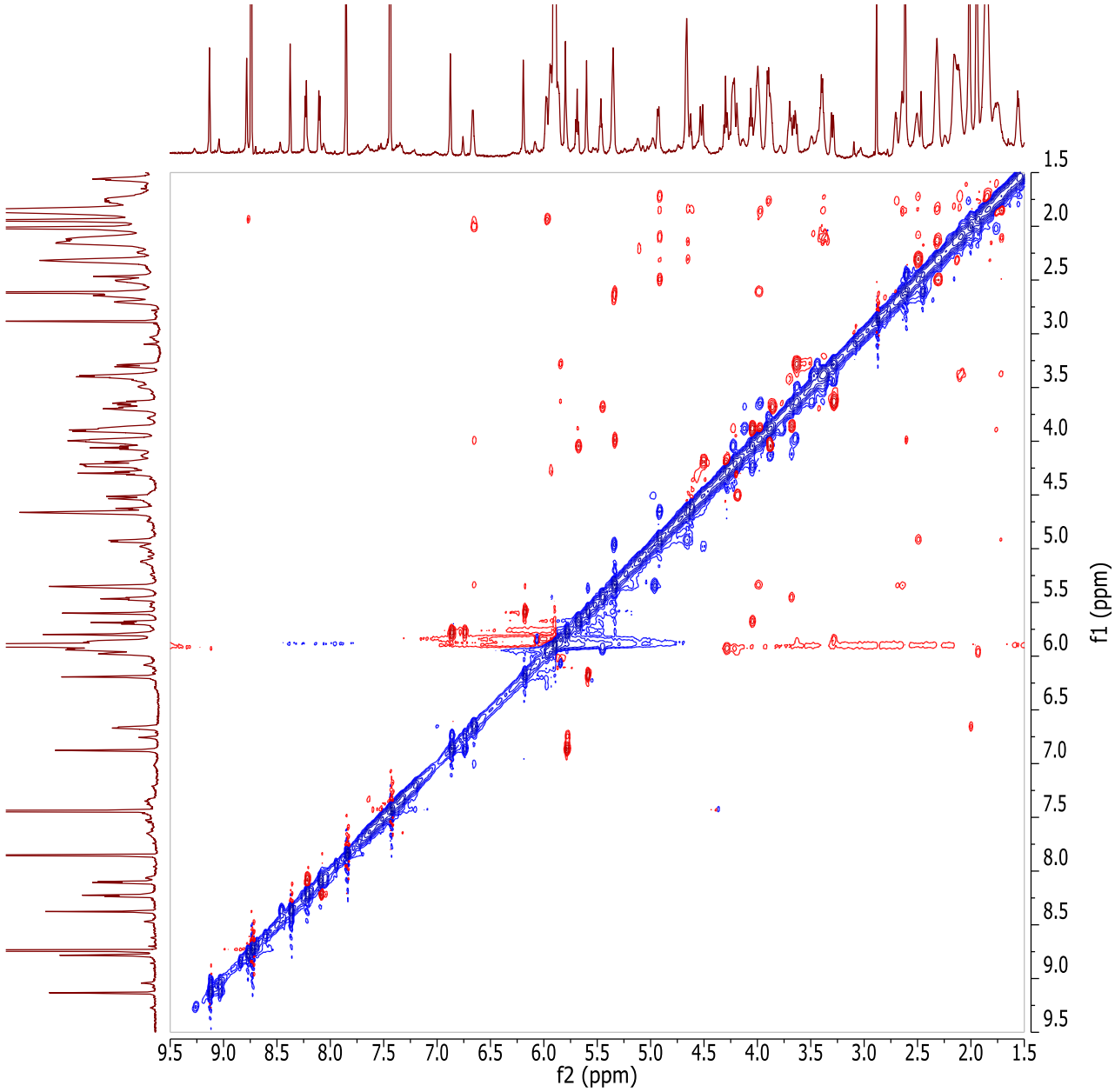


TOCSY,
750 MHz
Expansion

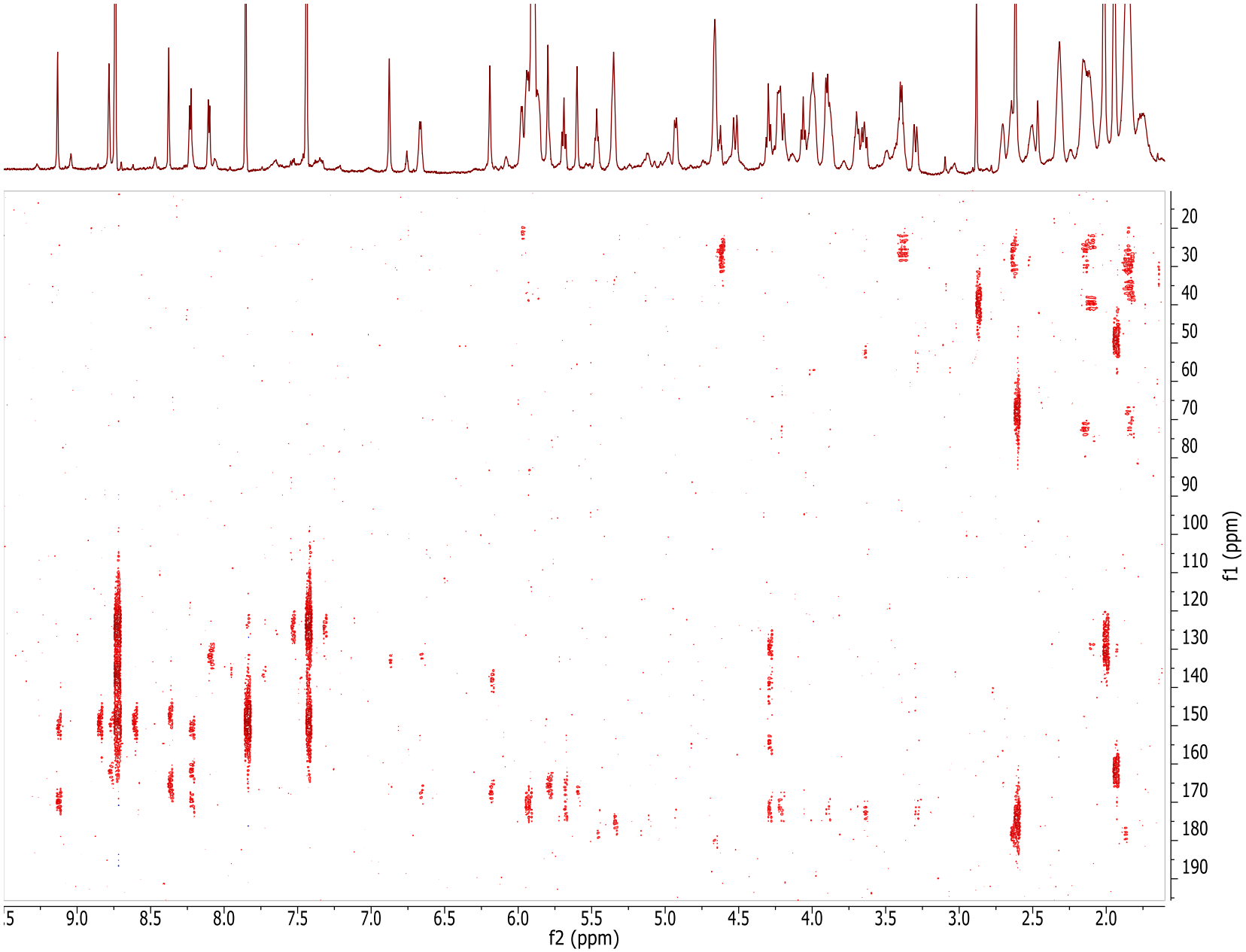
Key TOCSY correlations are enclosed in dashed rectangles.



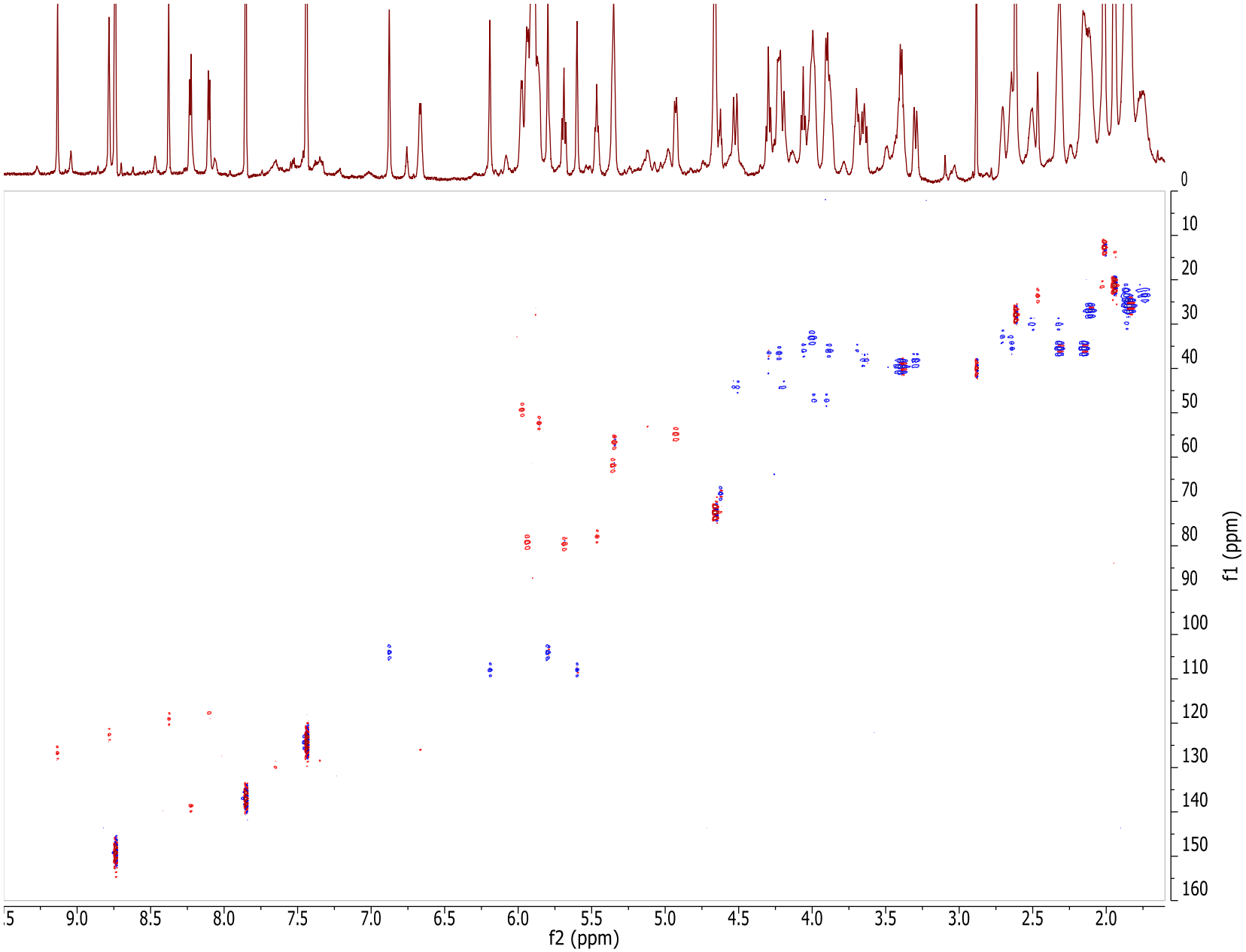
ROESY,
750 MHz



^{1}H-^{13}C HMBC, 750 MHz



^{1}H-^{13}C HSQC, 750 MHz



a

b

```
WC-3908             --SGAAATWHSIHVHHHDESTEPALVLHAVRPAFEAVRASVDGAWFGRHWLRGPHLRLNF 58
S. mobaraensis      -MSRPAPRWHSLHVHHHDADAEPALVLDAVRPAFEAVRPWTRGTWFGRHWLRGPHLRLNF 59
S. hygroscopicus    MNGTAGSSWYSLHVHHHDEEAEPELVLRAVRPAFEAVAPWAEAAWFGRHWLRGPHLRLNF 60
                      . ... *:*:****** .:** *** ********* . . .:***************

WC-3908             RTTADAWASKVRPQVVAVLETYVRDHPSTVRLDEAALAPVHERLAELEMETGPRHPWVPD 118
S. mobaraensis      RAPGPVWEARIRPVVTERIEAYLRVRPSRVRLDEAALAPVHARLAELEMETGARHPWVPD 119
S. hygroscopicus    RCRSTDWEERVRPTVTGIVTDYLRARPSAARLDEGALAPVHARLAELEMETGPRHPWVPD 120
                    *  .  *  ::** *.  :  *:* :** .****.****** **********.*******

WC-3908             NTVLEQPYDHRLPVLGSMRGSEVLADFLSDTNDLAFAMYEHLQGGGALPVLALDLMWTTA 178
S. mobaraensis      NTVLERPYDHRLPVLGSPRASELLAGFLSDTDDLAFRMYERRAAGGSTAALALDLMWTTA 179
S. hygroscopicus    NTVLERPYDHRLPVLESLRASELLAGFLSDTNGQAFRAYERVRAGGALSLLALDLMWTTT 180
                    *****:********* * *.**:**.*****:. **  **:  .**: . *********:

WC-3908             SVAAVPFDEHSAPIARGFLSLRSHADAFLSRTHDPAAYRAAFDERYRRQADALTTRLRTV 238
S. mobaraensis      AVAAVPFETGAPPIVRGYLSLRSHADAFLSRTRDPAAHRAAFDARYHRQADALGARLRAV 239
S. hygroscopicus    SVAAVPFTTGGEPIERGFLSLRSHADAFLSRTRDPVAVRAAFDDRFRRQETVLCERLRSV 240
                    :******   . ** **:***************:**.* ***** *::**  .*  ***:*

WC-3908             ESVLSAEPGS------------PEAADATVPFVREWAEAVRRCQRAAYPLLASGELTLSG 286
S. mobaraensis      QEAVDGR-------------------REDVPFVADWAAAIRRRQRIAYPQLASGAVSMRG 280
S. hygroscopicus    EAALSDGATEGDGADRSDRSDRSEAVGDVVPFVTEWAAAVRHHQRIAHPLLASGEVSMGG 300
                    : .:.                          **** :** *:*: ** *:* **** ::: *

WC-3908             AGRAPRLPTRQLSDFHRLLQTDHGHDGFLRDDHWFASFRLVVNYLYVHLNRLGLKPVDRG 346
S. mobaraensis      AGLAPRPPARQVSGFHRLLQSGHGQEDFLWTDPWFASFRLVVNYLYVHLNRLGLGPADRA 340
S. hygroscopicus    AARAPRMPTRRTSEFHAVLRSDHGHEDFVRTDDWFASFRLMMNYLYLHLNRLGLKPVDRA 360
                    *. *** *:*: * ** :*::.**::.*:  * *******::****:******* *.**.

WC-3908             LLCHLAAGTVETVHGVTAEDAFRRHVVSASPTEDTPEWSRLSKEWAEGR--------- 395
S. mobaraensis      LLCHLAASTVEGVHGVTATEVFRRHVLPPS-RDQVPEWRRLAAEWGRAAGSAPAERQG 397
S. hygroscopicus    LLCHLAALTVEAVHGVDAVGSFQRYVASVDPSSERPEWRRIGEAWAAGGAG------- 411
                    ******* *** **** *   *:*:* . .  .: *** *:.  *. .
```

c

	WC-3908	*S. mob*	*S. hyg*	TpdD	TpdD	SioL	NosO	TclM	TsrE
WC-3908 (cyclothiazomycin C)	100	67	60	34	28	22	19	15	12
S. mobaraensis (cyclothiazomycin B)	88	100	65	33	34	23	22	16	15
S. hygroscopicus (cyclothiazomycin A)	85	81	100	33	34	64	20	15	14
TpdD (thiomuracin)	58	55	56	100	38	34	20	15	15
TpdD (GE2270)	54	58	56	64	100	28	21	16	15
SioL (siomycin)	53	51	87	58	54	100	20	15	12
NosO (nosihpeptide)	47	50	51	49	51	47	100	16	15
TclM (thiocillin)	52	48	47	53	54	48	48	100	9
TsrE (thiostrepton)	46	47	46	40	39	46	46	38	100

Supplemental Figure 9. Amino acid similarity for CtmG, the putative formal [4+2] cycloaddition protein. (a) The cyclothiazomycin C biosynthetic gene cluster (strain WC-3908, NCBI accession KJ651958 apparently lacked the *ctmG* gene for the carrying out the [4+2] cycloaddition required for pyridine formation (Figure 5b). However, BLAST searching found a highly similar gene elsewhere on the WC-3908 chromosome (NCBI accession KJ690935). Interestingly, *ctmG* from WC-3908 is adjacent to *ctmF*, which appears to have been duplicated

from the rest of the cyclothiazomycin C biosynthetic gene cluster (Figure 5). (b) Amino acid alignment of the CtmG proteins from the cyclothiazomycin A (*S. hygroscopicus*), cyclothiazomycin B (*S. mobaraensis*), and cyclothiazomycin C (WC-3908) biosynthetic gene clusters. Below the aligned residues, * represents identical residues, while **:** and **.** represent highly and moderately similar residues, respectively. (c) Sequence similarity (sum of identical and similar residues / length of longest protein) and identity (identical residues / length of longest protein) between other known formal [4+2] cycloaddition proteins. The gene name and resulting thiopeptide product are given. Values in blue indicate sequence similarity, while green represent sequence identity values.

Supplemental Figure 10 (Below). Gene similarities for the cyclothiazomycin biosynthetic gene clusters. a) Genes surrounding the conserved portion of the cyclothiazomycin biosynthetic gene clusters were used as query sequences to identify homologs via BLAST searching. Genes 1-10 represent the genes upstream of the conserved cluster with 1 being the farthest from *ctmI*. *CtmI – H* are the conserved genes in the clusters (Figure 3b, NCBI accession number KJ651958) and are highlighted in gray. Red denotes *ctmG* from the cyclothiazomycin C producer that is conserved, but not present in the gene cluster, but rather elsewhere in the genome. Genes 11-20 lie downstream of the conserved region. b) BLAST results using the conserved genes from the cyclothiazomycin C gene cluster as query sequences. The best match returned by BLAST and the percent identities are given.

A. Gene Neighborhood Homologs

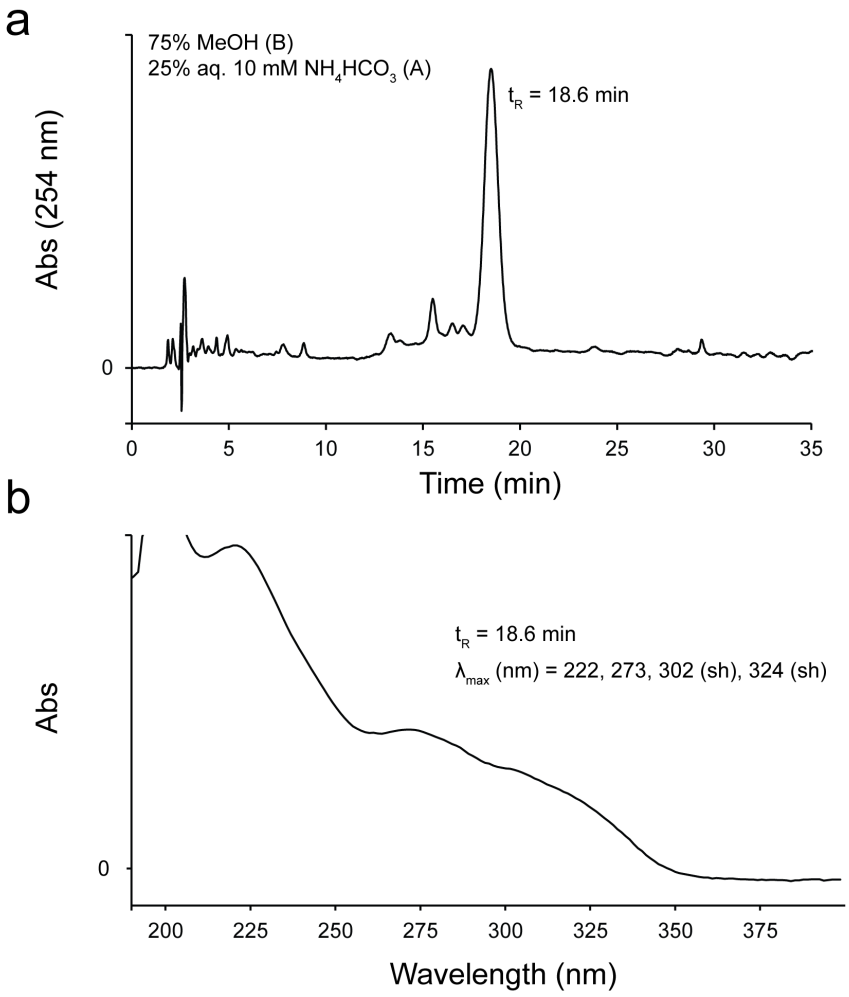
Gene	Cyclothiazomycin A (***Streptomyces hygroscopicus*** **subsp. *jinggangensis* 5008) Homolog (Species / Accession number)**	**Cyclothiazomycin B (*Streptomyces mobaraensis*) Homolog (Species / Accession number)**	**Cyclothiazomycin C Homolog (Species / Accession number)**
1	Ferredoxin reductase (*Streptomyces albulus* / WP_020929831.1)	ArsR family transcriptional regulator (*Nocardiopsis halophila* / WP_017541891.1)	Anti-sigma b factor RsbV (S*treptomyces rimosus* / WP_003980648.1)
2	Ferredoxin reductase (*Streptomyces albulus* / WP_016575819.1)	Helix-turn-helix protein (*Streptomyces purpureu*s / WP_019884370.1)	Anti-sigma b factor antagonist RsbV (S*treptomyces rimosus* / WP_003980649.1)
3	Copper domain protein (*Streptomyces* sp. FR1 / YP_008995649.1)	Secreted amidase *(Streptomyces bingchenggensis* BCW-1 / YP_004958967.1)	Anti-anti-sigma regulatory factor (*Streptomyces rimosus* / WP_003980650.1)
4	Isoprenylcysteine carboxyl methyltransferase (*Frankia* sp. CN3 / WP_007507509.1)	NAD binding lipoprotein *(Streptomyces viridosporus* / WP_016827948.1)	Hypothetical (S*treptomyces rimosus* / WP_003980651.1)
5	Hypothetical protein (*Rhodococcus* sp. DK17 / WP_016881545.1)	Hypothetical protein (*Streptomyces* sp. MspMP-M5 / WP_018538980.1)	Acyl-CoA dehydrogenase (*Streptomyces* sp. FxanaC1 / WP_018088719.1)

6	Hypothetical protein (*Rhododcoccus* sp. DK17 / WP_016881544.1)	Hypothetical protein (*Streptomyces* sp. 351MFTsu5.1 / WP_020141300.1)	PadR-like family transcriptional regulator (*Streptomyces rimosus* / WP_003980654.1)
7	Transposase (*Streptomyces coelicolor* A3(2) / NP_639869.1)	Putative oxidoreductase (*Sinorhizobium meliloti* SM11 / YP_005719396.1)	Protein serine/threonine phosphatase (*Streptomyces rimosus* / WP_003982697.1)
8	Hypothetical protein (*Streptomyces albulus* / WP_016576055.1)	MFS transporter (*Streptomyces roseochromogenes* / WP_023546609.1)	Cytochrome P450 FAS1 (*Streptomyces rimosus* / WP_003980657.1)
9	Hypothetical protein (*Streptomyces* sp. PsTaAH-124 / WP_018569508.1)	Transglycosylase-associated protein (*Streptomyces rimosus* / WP_003987332.1)	Aminoglycoside phosphotransferase (*Streptomyces cattleya* / YP_004920848.1)
10	Hypothetical protein (*Streptomyces hygroscopicus* subsp. *jinggangensis* 5008 / YP_006249961.1)	Chitinase (*Streptomyces* / WP_018524268.1)	ABC transporter (*Streptomyces rimosus* / WP_003980658.1)
CtmI	**Hypothetical protein (*Streptomyces mobaraensis* / WP_004943462.1)**	**Putative peptidase containing docking domain (*Streptomyces hygroscopicus* / ACS50133.1)**	**Putative peptidase containing docking domain (*Streptomyces hygroscopicus* / ACS50133.1)**
CtmB	**NADH oxidase (*Streptomyces mobaraensis* / WP_004943465.1)**	**NADH oxidase (*Streptomyces hygroscopicus* / ACS50132.1)**	**NADH oxidase (*Streptomyces mobaraensis* / WP_004943465.1)**
CtmA	**Hypothetical protein (*Streptomyces mobaraensis* / WP_004943468.1)**	**Hypothetical protein (*Streptomyces hygroscopicus* / YP_006249964.1)**	**Hypothetical (*Streptomyces mobaraensis* / WP_004943468.1)**
CtmD	**Cyclodehydratase (*Streptomyces mobaraensis* / WP_004943471.1)**	**Putative docking protein (*Streptomyces hygroscopicus* / ACS50130.1)**	**Hypothetical protein (*Streptomyces mobaraensis* / WP_004943471.1)**
CtmE	**Lantibiotic dehydratase (*Streptomyces mobaraensis* / WP_004943473.1)**	**Lantibiotic dehydratase (*Streptomyces hygroscopicus* / ACS50129.1)**	**Lanthionine biosynthesis protein (*Streptomyces mobaraensis* / WP_004943473.1)**
CtmF	**Lantibiotic biosynthesis protein (*Streptomyces mobaraensis* / WP_004943476.1)**	**Lantibiotic biosynthesis protein (*Streptomyces hygroscopicus* / YP_006249967.1)**	**Lantibiotic biosynthesis protein (*Streptomyces mobaraensis* / WP_004943476.1)**
CtmG	**Hypothetical protein (*Streptomyces mobaraensis* / WP_004943480.1)**	**Hypothetical protein (*Streptomyces hygroscopicus* / YP_006249968.1)**	**Hypothetical protein (*Streptomyces mobaraensis* / WP_004943480.1)**

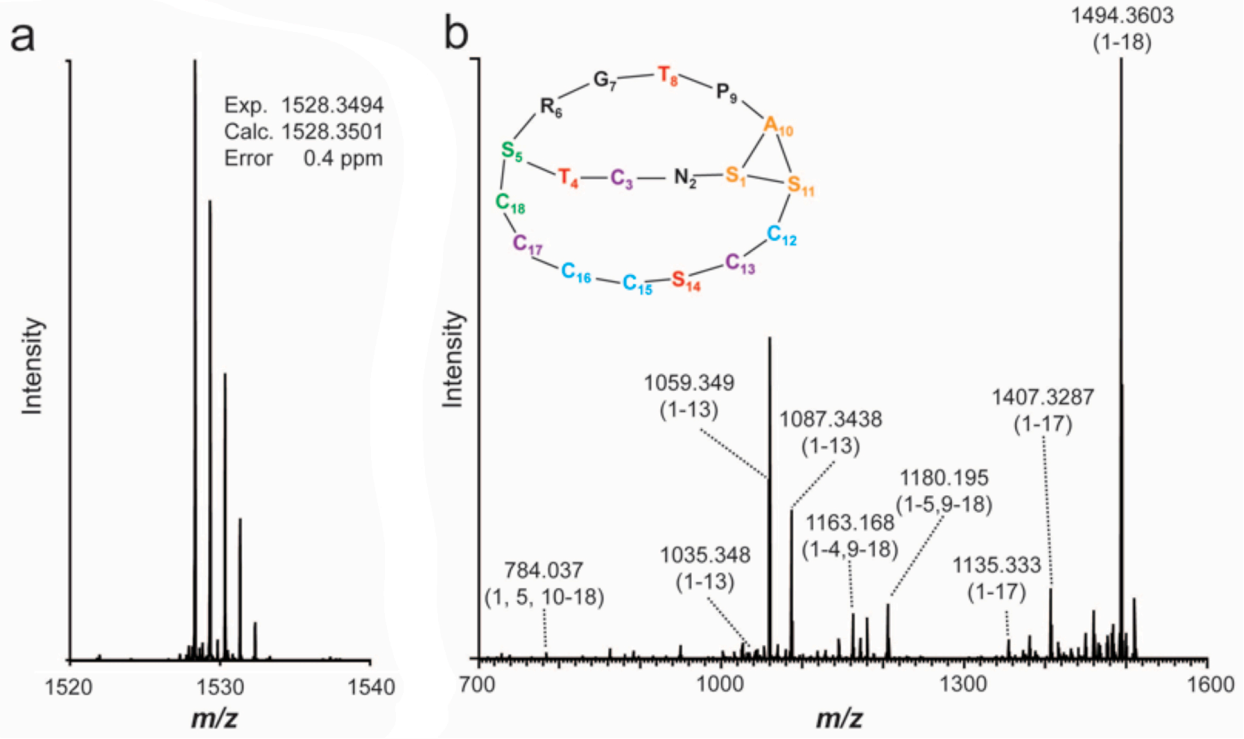
CtmJ		**Hypothetical protein (*Streptomyces auratus* / WP_006606582.1)**	
CtmK		**Hypothetical protein (S*treptomyces auratus* / WP_006606583.1)**	
CtmH	**LuxR transcriptional regulator (*Streptomyces mobaraensis* / WP_004943488.1)**	**LuxR transcriptional regulator (*Streptomyces hygroscopicus* / ACS50126.1)**	**LuxR family transcriptional regulator (*Streptomyces hygroscopicus* / ACS50126.1)**
11	Hypothetical protein (*Streptomyces* sp. FR1 / YP_008995424.1)	Integral membrane protein (*Streptomyces avermitilis* MA-4680 / NP_828391.1)	Short chain dehydrogenase (*Streptomyces auratus* / WP_006608143.1)
12	Ethyl tert-butyl ether degradation (*Streptomyces griseoaurantiacus* / WP_006139146.1)	Hypothetical protein (*Streptomyces griseoflavus* / WP_004934717.1)	Hypothetical protein (*Streptomyces rimosus* / WP_003980661.1)
13	MFS transporter (*Streptomyces griseoaurantiacus* / WP_006139145.1)	LysR family transcriptional regulator (*Streptomyces himastatinicus* / WP_009713438.1)	Hypothetical protein (S*treptomyces rimosus* / WP_003980662.1)
14	DNA binding protein (*Streptomyces griseroaurantiacus* / WP_006139144.1)	Ornithine carbamoyltransferase (*Patulibacter medicamentivorans* WP_007570660.1)	Methylated DNA protein cysteine S-methyltransferase (S*treptomyces rimosu*s / WP_003980663.1)
15	AraC family transcriptional protein (*Streptomyces albulus* / WP_020930180.1)	Hypothetical protein (*Streptomyces mobaraensis* WP_004943504.1)	Ricin superfamily (S*treptomyces rimosus* / WP_003980666.1)
16	Hypothetical protein (*Streptomyces hygroscopicus* subsp. *jinggangensis* 5008 / YP_006249975.1)	Hypothetical protein (*Amycolatopsis nigrescen*s / WP_020673411.1)	Oxidoreductase domain protein (*Streptomyces rimosus* / WP_003980667.1)
17	Transposase (*Streptomyces lividans* / WP_003974545.1)	TetR family transcriptional regulator (S*treptomyces clavuligerus* / WP_003956021.1)	ArsR family transcriptional regulator (*Streptomyces rimosus* / WP_003980668.1)
18	Transposase (*Streptomyces* sp. Mg1 / WP_008736550.1)	Hypothetical protein (S*treptomyces sulphureus* / WP_019547651.1)	30S ribosomal protein (*Thauera terpenica* / WP_021250332.1)
19	Transposase (*Streptomyces violaceusniger* Tu 4113/ YP_004800126.1)	Hypothetical protein (*Streptomyces aurantiacus* / WP_016638852.1)	Ribosome-inactivating protein (S*treptomyces scabiei* 87.22 / YP_003494282.1)
20	Oxidoreductase (*Streptomyces viridochromogenes* / WP_003995212.1)	Hypothetical protein (*Streptomyces thermolilacinus* / WP_023588110.1)	Ricin superfamily (*Streptomyces* sp. Mg1 / WP_008735505.1)

10 B. Cyclothiazomycin C protein identities

Protein (cyclothiazomycin C)	Top BLAST hit / accession number	% Identity	Residues (aligned/total)
CtmI (WC3908_03952)	*Streptomyces hygroscopicus* / ACS50133.1	60	379/634
CtmB (WC3908_03953)	*Streptomyces mobaraensis* / WP_004943465.1	61	324/531
CtmA (WC3908_03954)	*Streptomyces mobaraensis* / WP_004943468.1	78	43/55
CtmD (WC3908_03955)	*Streptomyces mobaraensis* / WP_004943471.1	72	455/636
CtmE (WC3908_03956)	*Streptomyces mobaraensis* / WP_004943473.1	58	500/858
CtmF (WC3908_03957)	*Streptomyces mobaraensis* / WP_004943476.1	57	192/334
CtmG (WC3908_00024)	*Streptomyces mobaraensis* / WP_004943480.1	68	265/397
CtmH (WC3908_03958)	*Streptomyces hygroscopicus* / ACS50133.1	46	429/924



Supplemental Figure 11. HPLC trace and UV spectrum of cyclothiazomycin B. a) A sample (spatula tip) of purified cyclothiazomycin B was dissolved in 50% MeOH (B)/aq. 10 mM NH_4HCO_3 (A) (200 µL). An aliquot (27 µL) was analyzed by HPLC (isocratic 75% B for 35 min). Photodiode array (PDA) detection was used to monitor absorbance (abs) from 200-400 nm. A blank injection was also run and subtracted from the cyclothiazomycin B chromatogram; the resulting spectrum with UV monitoring at 254 nm is shown. b) Cyclothiazomycin B exhibits UV absorbance consistent with that previously reported (1) and cyclothiazomycin C (Supplemental Figure 5). A UV spectrum (PDA) from the HPLC trace at 18.6 min is shown (sh, shoulder).



Supplemental Figure 12. High resolution Fourier transform mass spectrometry (FT-MS) of cyclothiazomycin B. a) The *m/z* scan of purified cyclothiazomycin B showed an ion in the 1^+ charge state with an observed isotopic *m/z* value with <1 ppm error from the calculated value for cyclothiazomycin B. b) CID spectrum of *m/z* 1528. The monoisotopic mass values are given for assigned peak predictions. The number ranges given below the mass values are a shorthand notation describing predicted fragments of cyclothiazomycin B. A key for the shorthand notation for the structure of cyclothiazomycin B is given in pictorial format using single letter codes for the amino acids, the residue's *N* to *C* position, and lines depicting molecular connectivity within the mature structure. The colors used for the shorthand notation depict the modification present at a particular residue. Purple, thiazoline moieties; green, thioether linkage; cyan, thiazole moieties; red, dehydrated amino acids; orange, pyridine moiety; black, unmodified amino acids.

III. References

1. Hashimoto, M., Murakami, T., Funahashi, K., Tokunaga, T., Nihei, K., Okuno, T., Kimura, T., Naoki, H., Himeno, H. (2006) An RNA polymerase inhibitor, cyclothiazomycin B1, and its isomer, *Bioorg. Med. Chem. 14*, 8259-8270.

2. Aoki, M., Ohtsuka, T., Itezono, Y., Yokose, K., Furihata, K., Seto, H. (1991) Structure of cyclothiazomycin, a unique polythiazole-containing peptide with renin inhibitory activity. Part 1. Chemistry and partial structures of cyclothiazomycin, *Tetrahedron Lett. 32*, 217-220.