EcdGHK are Three Tailoring Iron Oxygenases for Amino Acid Building Blocks of the Echinocandin Scaffold

Wei Jiang*[1], Ralph A. Cacho*[2], Grace Chiou[3], Neil K. Garg[3], Yi Tang[2,3], Christopher T Walsh[1]

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 02115

[2]Department of Chemical and Biochemical Engineering, [3]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095

* These authors contributed equally to this work.

acknowlegements (This work is supported by NIH Grants 1R01GM092217 and 1DP1GM106413 to Y. T.; 1R01GM49338 to C.T.W.; NRSA GM08496 to R.A.C. and NSF GFRP DGE-070742 to G.C.)

Table of Contents:

SUPPLEMENTARY TABLES……S3

Table S1: ^{13}C and ^{1}H NMR data for Echinocandin B **1**, D **3**, **6** and **7**……S3

Table S2: Comparison of ^{1}H NMR data of trideoxy compound from Δ*ecdH* and reported Echinocandin D **3**……S5

Table S3: Primers used for this study……S6

SUPPLEMENTARY FIGURES……S8

Figure S1: PCR screening of Δ*ecdH* ……S8

Figure S2: MS of Echinocandin B and C produced by *E. rugulos* WT……S9

Figure S3: MS of Echinocandin **1**, **2**, **7** and **8** with supplement of D_7-L-proline ……S10

Figure S4: PCR screening of Δ*ecdG* ……S11

Figure S5: ^{1}H spectrum of 3-hydroxyl-homoTyr (400 MHz) in 1N DCl ……S12

Figure S6: gCOSY spectrum of 3-hydroxyl-homoTyr (400 MHz) in 1N DCl……S13

Figure S7: PCR screening of Δ*ecdK*……S14

Figure S8: HPLC-UV (440 nm) trace of *o*-AB derivatized MeP5C……S15

Figure S9: Proposed mechanism for *o*-AB derivatization of MeP5C……S16

Figure S10: ^{1}H NMR of *o*-AB derivatized MeP5C (400 MHz) in D_2O……S17

Figure S11: HSQC spectrum of *o*-AB derivatized MeP5C (400 MHz) in D_2O……S18

Figure S12: gCOSY spectrum of *o*-AB derivatized MeP5C (400 MHz) in D_2O……S19

Figure S13: HMBC spectrum of *o*-AB derivatized MeP5C (400 MHz) in D_2O……S20

Figure S14: ^{1}H and ^{13}C parameters of *o*-AB derivatized MeP5C……S21

Table S1: ^{13}C and ^{1}H NMR data for Echinocandin B **1** (from WT), D **3** (from Δ*ecdH* variant), **6** and **7** (from Δ*ecdG* variant) in CD_3OD (600 MHz). The signals most useful for indicating structural changes are highlighted.

	WT **1**	Δ*ecdH* (tri) **3**	Δ*ecdG* (di) **6**	Δ*ecdG* (tri) **7**		WT **1**	Δ*ecdH* (tri) **3**	Δ*ecdG* (di) **6**	Δ*ecdG* (tri) **7**
4,5-dihydroxyornithine (DiOHOrn)									
C-2	51.3	52.8	51.5	52.1	2-H	4.39	4.39	4.41	4.37
C-3	34.8	27.7	34.8	26.8	3-Ha	2.07	1.55	2.07	1.65
					3-Hb	1.93	2.12	1.93	1.65
C-4	70.6	24.6	70.6	30.8	4-Ha	3.94	1.71	3.90	2.00
					4-Hb	--	1.71	--	1.73
C-5	74.2	37.8	74.4	79.3	5-Ha	5.25	2.97	5.25	5.10
					5-Hb	--	3.48	--	--
Threonine A (ThrA)									
C-2	58.4	58.2	58.5	58.5	2-H	4.95	4.91	4.92	4.96
C-3	68.1	67.9	68.5	68.8	3-H	4.51	4.48	4.45	4.46
C-4	19.5	19.5	19.2	19.2	4-H	1.20	1.23	1.23	1.22
Threonine B (ThrB)									
C-2	56.7	57.0	55.9	55.7	2-H	4.84	4.91	4.85	4.89
C-3	69.5	69.1	69.0	69.1	3-H	4.15	4.26	4.17	4.14
C-4	19.5	20.1	20.3	20.3	4-H	1.21	1.19	1.18	1.19

Table 1: ^{13}C and ^{1}H NMR data for Echinocandin B **1** (from WT), D **3** (from Δ*ecdH* variant), **6** and **7** (from Δ*ecdG* variant) in CD_3OD (600 MHz) (continued).

	WT **1**	Δ*ecdH* (tri) **3**	Δ*ecdG* (di) **6**	Δ*ecdG* (tri) **7**		WT **1**	Δ*ecdH* (tri) **3**	Δ*ecdG* (di) **6**	Δ*ecdG* (tri) **7**
4-hydroxy-proline (OHPro)									
C-2	62.3	62.2	61.8	61.9	2-H	4.58	4.65	4.55	4.53
C-3	38.3	38.8	38.4	38.1	3-Ha	2.42	2.45	2.33	2.33
					3-Hb	2.04	2.08	1.77	1.75
C-4	71.1	71.1	70.9	71.0	4-H	4.50	4.58	4.48	4.47
C-5	56.8	57.0	56.8	56.7	5-Ha	3.96	4.00	3.76	3.76
					5-Hb	3.79	3.81	3.74	3.72
3,4-dihydroxyhomotyrosine (diOHTyr)									
C-2	56.1	57.6	54.7	54.5	2-H	4.32	4.44	4.38	4.48
C-3	76.7	74.0	34.9	35.5	3-Ha	4.24	4.41	2.20	2.20
					3-Hb	--	--	2.07	1.99
C-4	75.4	40.8	32.6	32.4	4-Ha	4.30	2.65	2.59	2.59
					4-Hb	--	2.57	2.59	2.59
C-1'	132.9	129.5	132.8	132.7					
C-2'/C-6'	129.5	131.2	130.2	130.2	2'-H/6'-H	7.14	7.01	7.01	7.01
C-3'/C-5'	116.0	116.0	116.0	116.0	3'-H/5'-H	6.76	6.71	6.70	6.70
C-4'	158.3	156.9	156.4	156.5					
3-hydroxy-4-methylproline (OHMePro)									
C-2	69.3	69.7	69.3	69.3	2-H	4.34	4.32	4.37	4.28
C-3	75.4	75.8	75.2	76.0	3-H	4.19	4.14	4.21	4.17
C-4	38.8	38.8	38.8	38.8	4-H	2.52	2.48	2.55	2.56
C-5	52.7	52.7	52.6	52.7	5-Ha	3.86	3.85	3.88	3.88
					5-Hb	3.38	3.38	3.40	3.42
4-CH3	11.1	11.1	11.1	11.1	4-CH3	1.05	1.05	1.08	1.08

Table S2: ^{1}H NMR data of trideoxy compound from Δ*ecdH* (black) and reported Echinocandin D **3** (red) {Evans, 1987 #8} in CD_3OD are consistent.

	Δ*ecdH* (tri)	Echinocandin D **3**
4,5-dihydroxyornithine (DiOHOrn)		
2-H	4.39	4.40-4.36
3-Ha	1.55	1.62-1.51
3-Hb	2.12	2.15-2.04
4-Ha	1.71	1.72-1.66
4-Hb	1.71	1.72-1.66
5-Ha	2.97	2.99-2.94
5-Hb	3.48	3.49-3.43
Threonine A (ThrA)		
2-H	4.91	4.89
3-H	4.48	4.48-4.44
4-H	1.23	1.21
Threonine B (ThrB)		
2-H	4.91	4.88
3-H	4.26	4.26-4.22
4-H	1.19	1.18

	Δ*ecdH* (tri)	Echinocandin D **3**
4-hydroxy-proline (OHPro)		
2-H	4.65	4.64
3-Ha	2.45	2.49-2.43
3-Hb	2.08	2.15-2.04
4-H	4.58	4.56
5-Ha	4.00	4.00
5-Hb	3.81	3.80
3,4-dihydroxyhomotyrosine (diOHTyr)		
2-H	4.44	4.41
3-Ha	4.41	4.40-4.36
4-Ha	2.65	2.64
4-Hb	2.57	2.55
2'-H/6'-H	7.01	7.00
3'-H/5'-H	6.71	6.69
3-hydroxy-4-methylproline (OHMePro)		
2-H	4.32	4.30
3-H	4.14	4.14
4-H	2.48	2.49-2.43
5-Ha	3.85	3.83
5-Hb	3.38	3.38
4-CH3	1.05	1.04

Table S3: Primers used for this study

Primer Name	Primer Sequence (5'-> 3')	Notes
bar-f	aagtaaccatgagcccagaacga	For mutant screening
bar-st-R	ctaaatctcggtgacgggca	For knockout cassette cloning and mutant screening
PtrpC-F	cgacagaagatgatattgaa	For knockout cassette cloning
ecdG-KO-P1	atctccaacggattgcatctcgca	For knockout cassette cloning and mutant screening
ecdG-KO-P2	cgacaactgacttccaagcct	For knockout cassette cloning
ecdG-KO-P3	ggtcctgcccgtcaccgagatttagcgcagcagcccatcgta	
ecdG-KO-P4	gctccttcaatatcatcttctgtcgattcacagggttataccga	
ecdG-KO-P5	ataacgagctgctcttgagga	
ecdG-KO-P6	cttggctacttgacccgtctt	For knockout cassette cloning and mutant screening
ecdH-KO-P1	ggagtcgaaataatcatggtgccata	For knockout cassette cloning and mutant screening
ecdH-KO-P2	ccagtcagacaatgttccgaagagca	For knockout cassette cloning
ecdH-KO-P3	gtcctgcccgtcaccgagatttagcacccacaagcccatcacgcaca	
ecdH-KO-P4	aaagtgctccttcaatatcatcttctgtcgttataagacgggtcaagta	
ecdH-KO-P5	tggtgcagccaactggcaagca	
ecdH-KO-P6	ccgttgacattgcattgcactgca	For knockout cassette cloning and mutant screening
ecdK-KO-P1	tgctgtatgttccgaatgcga	For knockout cassette cloning and mutant screening
ecdK-KO-P2	ctttctcgccgacagtaagct	For knockout cassette cloning
ecdK-KO-P3	ggtcctgcccgtcaccgagatttagctcgacttcagagcctga	
ecdK-KO-P4	tgctccttcaatatcatcttctgtcgtgctctggcgtatcagca	
ecdK-KO-P5	cgactctcgctcttcggcga	
ecdK-KO-P6	cctttacacgcctccgtca	For knockout cassette cloning and mutant screening

Table S3: Primers used for this study (continued)

Primer Name	Primer Sequence (5'-> 3')	Notes
*ecd*G-F	ggtattgagggtcgcatggcatctatctcatacgat	For cloning of EcdG expression construct
*ecd*G-R	agaggagagttagagccttcaagcagcactaaccccgt	
ecdK-F	ggtattgagggtcgcatgtctgttctaactctcga	For cloning of EcdK expression construct
ecdK-R	agaggagagtctagtacacaatccctgtcctca	

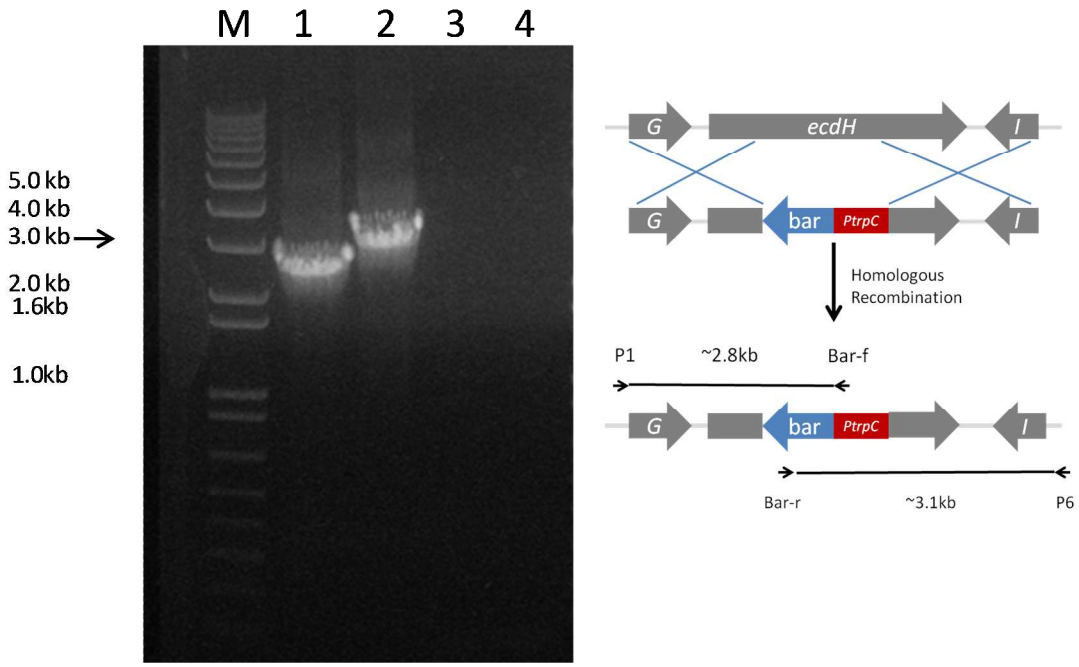


Figure S1: PCR screening of Δ*ecdH:* (M)1 kb plus ladder (Invitrogen). Amplification of Δ*ecdH* genomic DNA (gDNA) using primer pair *ecdH*-KO-P1 and bar-f showing expected ~2.8kb amplicon (lane 1). Amplification of Δ*ecdH* gDNA using primer pair *ecdH*-KO-P6 and bar-r showing expected ~3.1 kb amplicon (lane 2). Amplification of wild-type *E.rugulosa* gDNA using primer pair *ecdH*-KO-P1 and bar-f (lane 3) Amplification of wild-type *E.rugulosa* gDNA using primer pair *ecdH*-KO-P6 and bar-r (lane 4)..

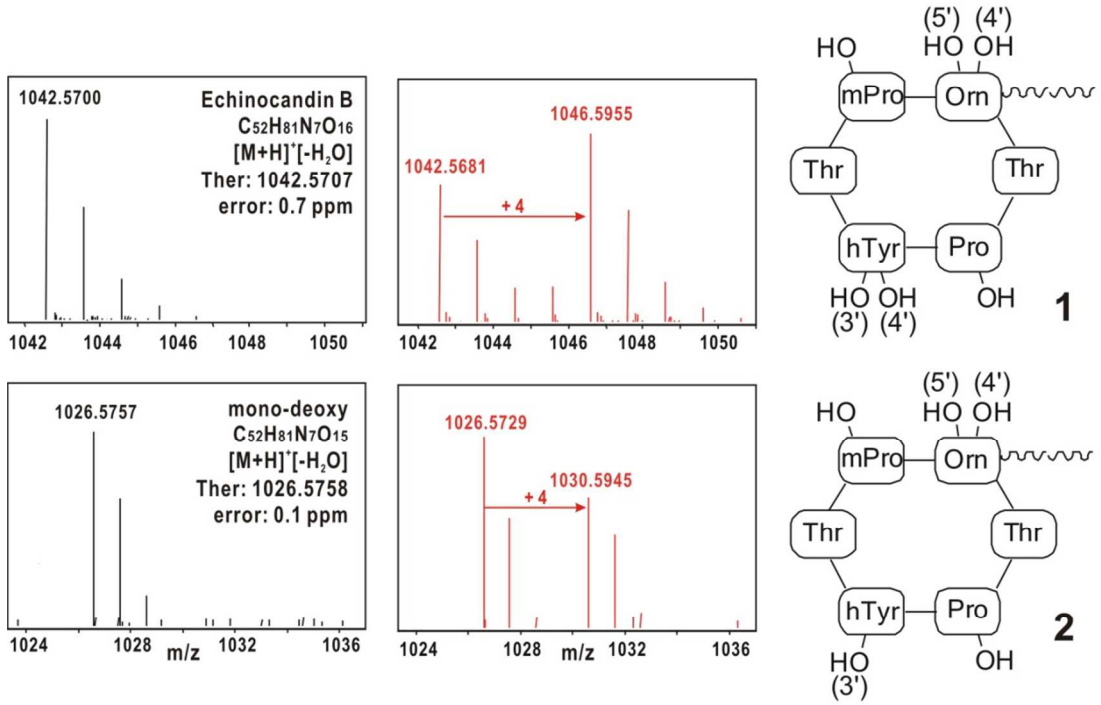


Figure S2: MS of parent ions from Echinocandin B **1** and C **2** produced by *Emericella rugulosa*, either from regular media (black) or fed with D_6-L-ornithine (red). The schematic structure of each compound is on the right.

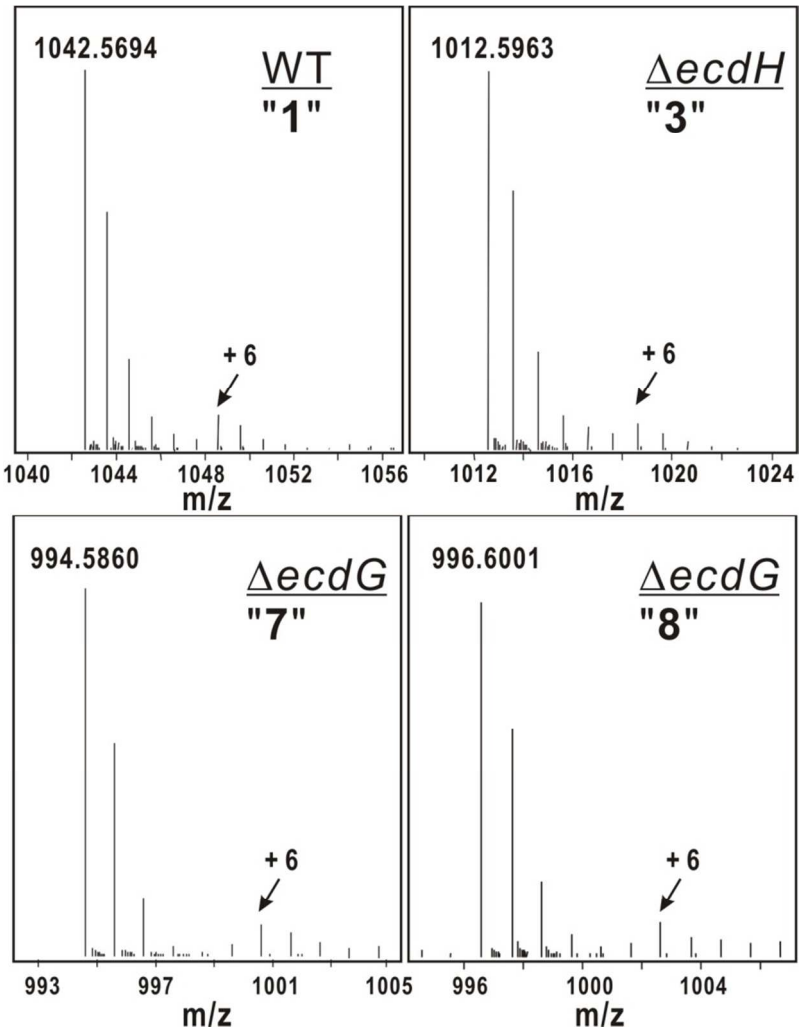


Figure S3: MS of parent ions from Echinocandin **1**, **2**, **7** and **8** produced by *Emericella rugulosa* wildtype or gene knockouts (as labeled) grown on minimal media with supplement of D_7-L-proline.

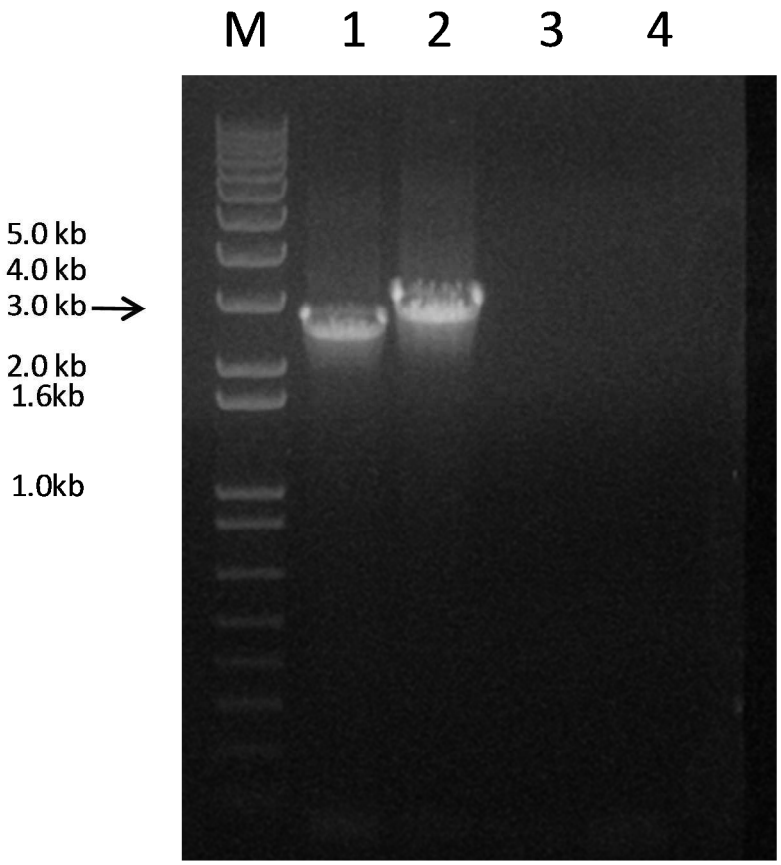


Figure S4: PCR screening of Δ*ecdG*.1 kb plus ladder (Invitrogen)(M). Amplification of Δ*ecdG* genomic DNA (gDNA) using primer pair *ecdG*-KO-P1 and bar-f showing expected ~2.8kb amplicon (lane 1). Amplification of Δ*ecdG* gDNA using primer pair *ecdH*-KO-P6 and bar-r showing expected ~3.1 kb amplicon (lane 2). Amplification of wild-type *E.rugulosa* gDNA using primer pair *ecdG*-KO-P1 and bar-f (lane 3) Amplification of wild-type *E.rugulosa* gDNA using primer pair *ecdG*-KO-P6 and bar-r (lane 4).

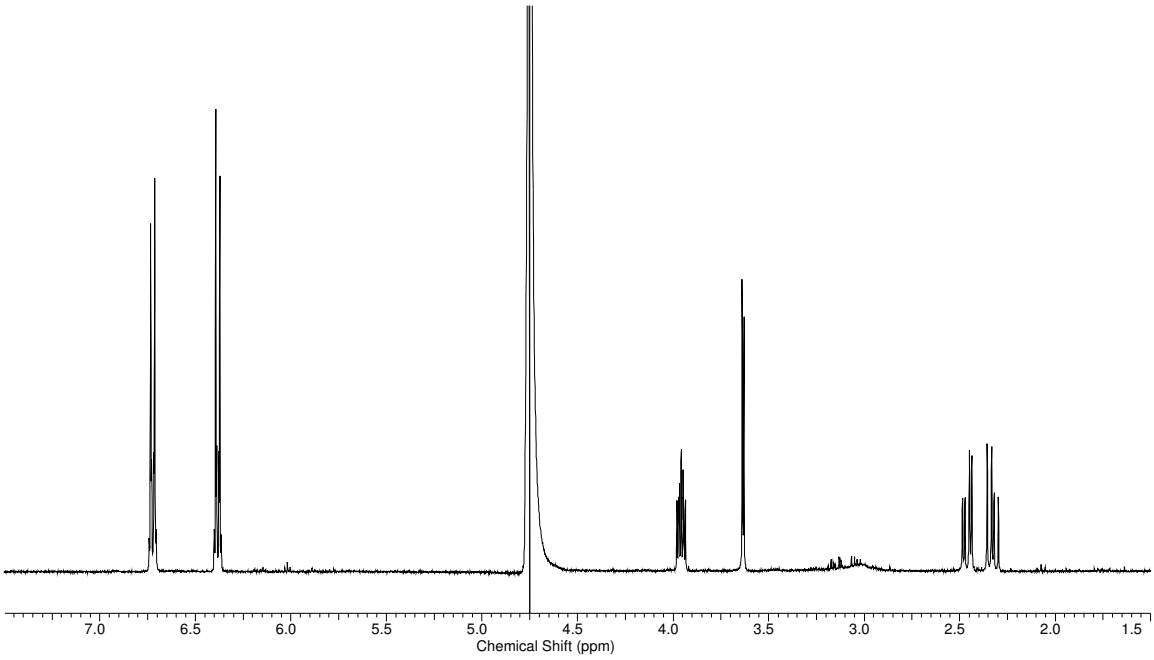


Figure S5: ^{1}H spectrum of 3-hydroxyl-homoTyr (400 MHz) in 1N DCl: δppm 6.68 - 6.76 (m, 2 H), 6.36 - 6.41 (m, 2 H), 3.96 (ddd, *J*=9.1, 5.2, 3.7 Hz, 1 H), 3.63 (d, *J*=3.9 Hz, 1 H), 2.46 (dd, *J*=14.1, 5.1 Hz, 1 H), 2.33 (dd, *J*=14.1, 9.4 Hz, 1 H)

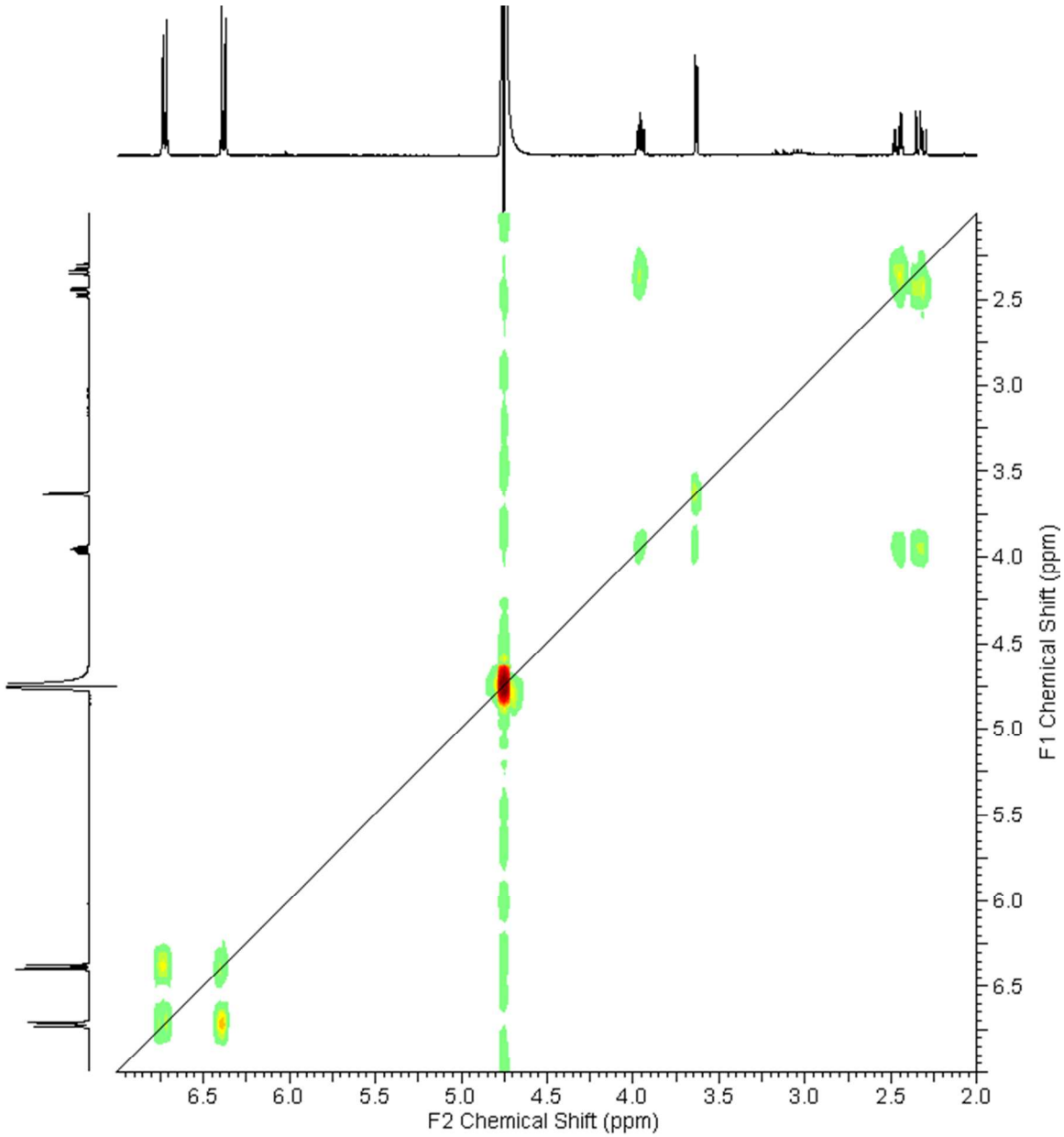


Figure S6: gCOSY spectrum of 3-hydroxyl-homoTyr (400 MHz) in 1N DCl.

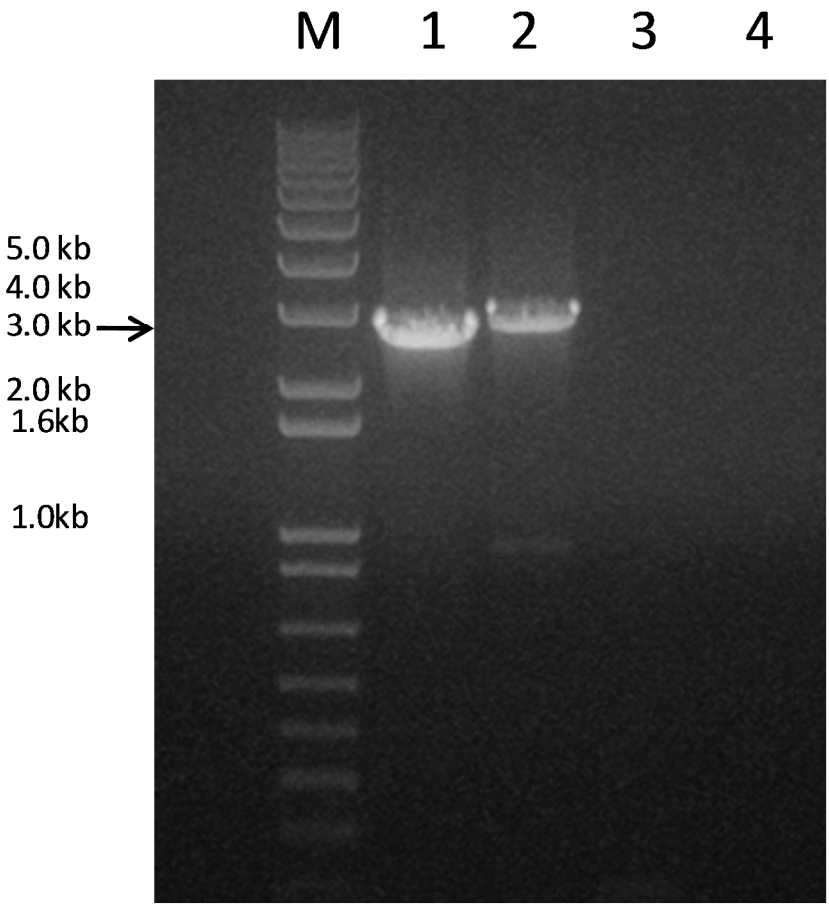


Figure S7: PCR screening of Δ*ecdK*. 1 kb plus ladder (Invitrogen)(M). Amplification of Δ*ecdK* genomic DNA (gDNA) using primer pair *ecdK*-KO-P1 and bar-f showing expected ~2.8kb amplicon (lane 1). Amplification of Δ*ecdK* gDNA using primer pair *ecdK*-KO-P6 and bar-r showing expected ~3.1 kb amplicon (lane 2). Amplification of wild-type *E.rugulosa* gDNA using primer pair *ecdG*-KO-P1 and bar-f (lane 3) Amplification of wild-type *E.rugulosa* gDNA using primer pair *ecdG*-KO-P6 and bar-r (lane 4)

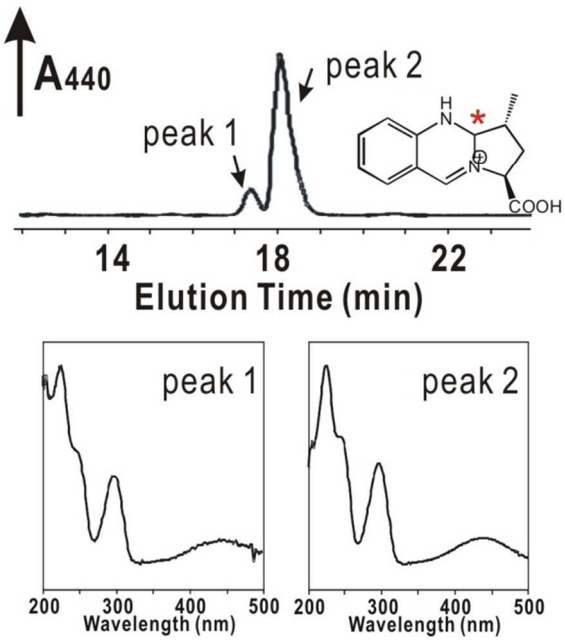
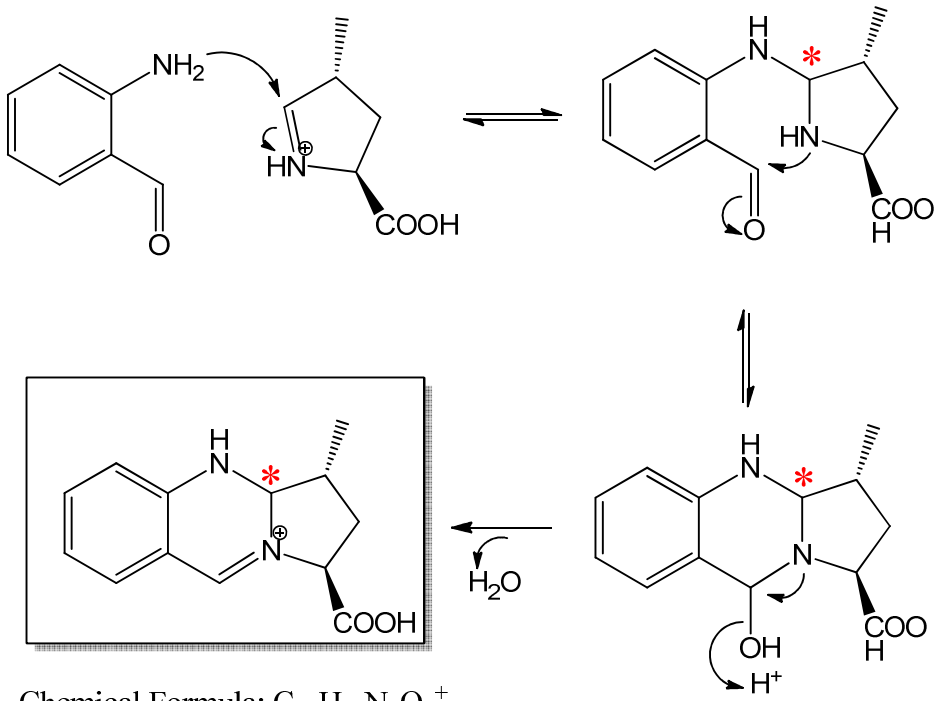


Figure S8: HPLC-UV (440 nm) trace showing two peaks of *o*-aminobenzaldehyde derivatized MeP5C (major/minor = 6:1). They have nearly identical mass and UV-Vis spectrum, indicating a pair of diastereomers. (define Me5PC)



Chemical Formula: $C_{13}H_{15}N_2O_2^+$
Exact Mass: 231.1128

Figure S9: Proposed mechanism for *o*-aminobenzaldehyde derivatization of MeP5C.

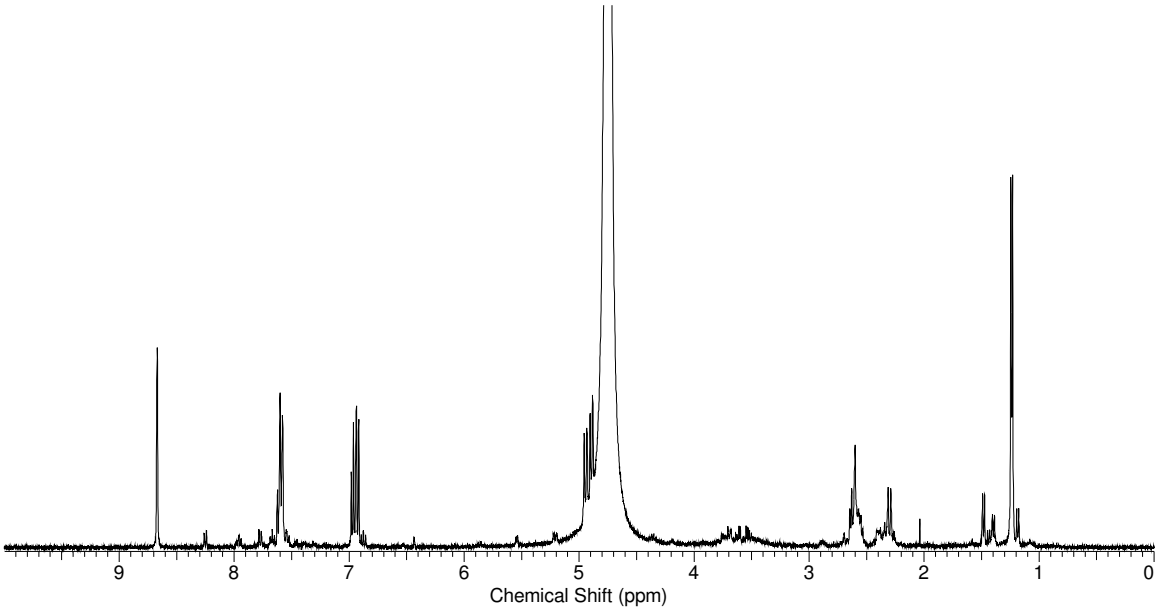


Figure S10: ^{1}H NMR spectrum of *o*-aminobenzaldehyde derivatized MeP5C (400 MHz) in D_2O: δppm 8.67 (s, 1 H), 7.56 - 7.64 (m, 2 H), 6.95 (dd, *J*=18.0, 8.0 Hz, 2 H), 4.94 (d, *J*=9.4 Hz, 1 H), 4.89 (d, *J*=8.6 Hz, 1 H), 2.52 - 2.67 (m, 3 H), 2.30 (d, *J*=9.8 Hz, 1 H), 1.24 (d, *J*=5.9 Hz, 3 H)

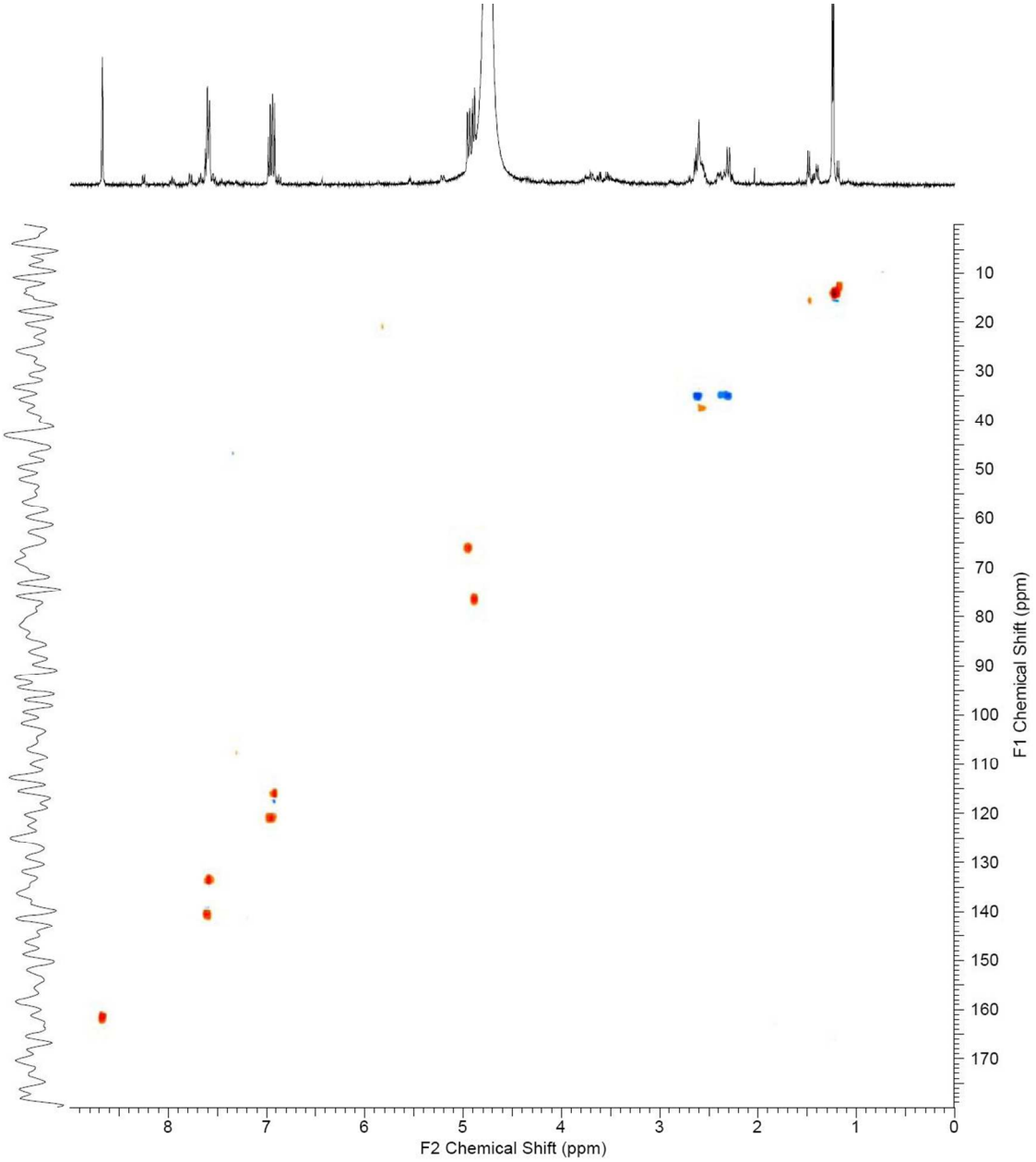


Figure S11: HSQC spectrum of *o*-aminobenzaldehyde derivatized MeP5C (400 MHz) in D_2O.

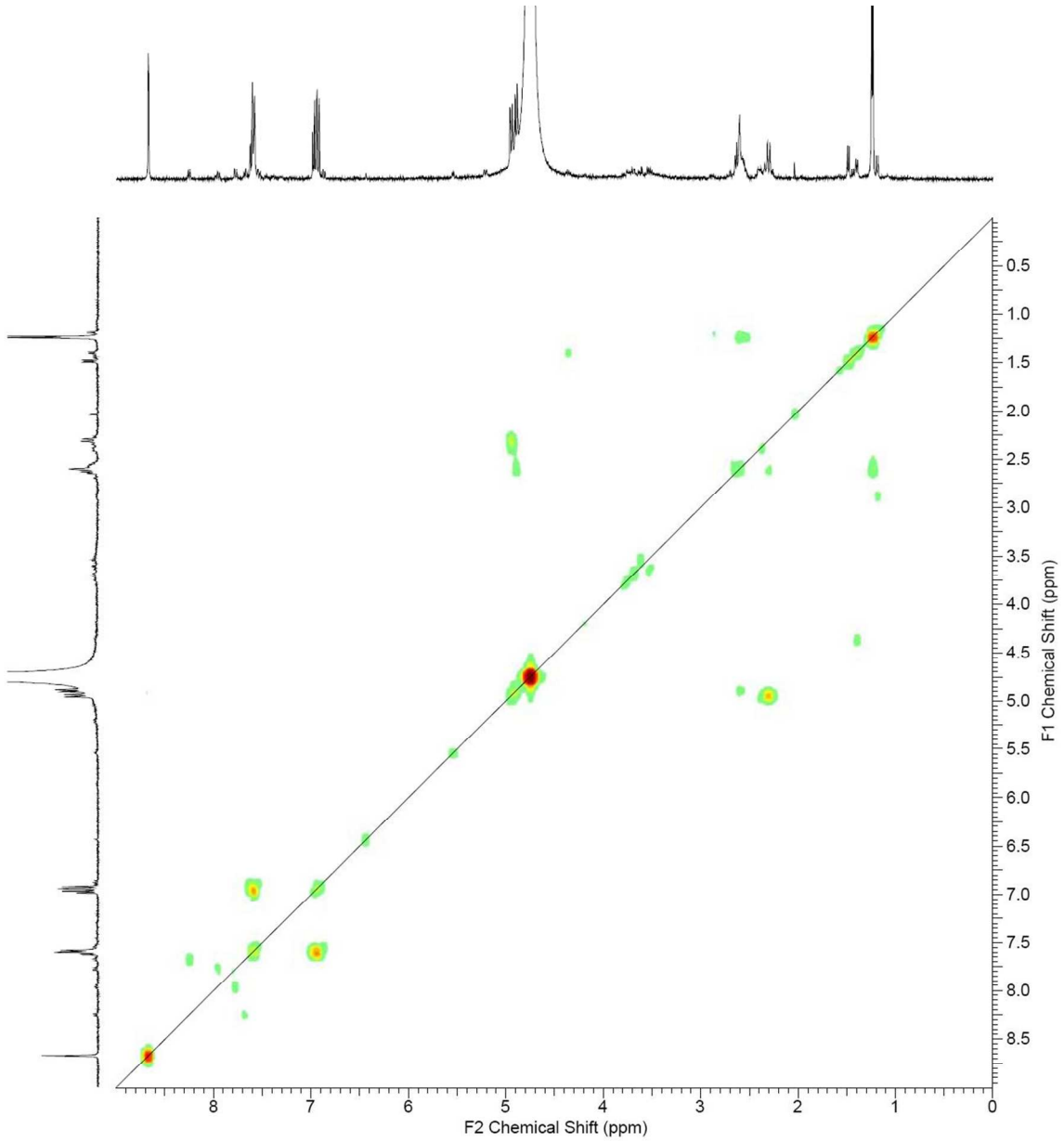


Figure S12: gCOSY spectrum of *o*-aminobenzaldehyde derivatized MeP5C (400 MHz) in D_2O.

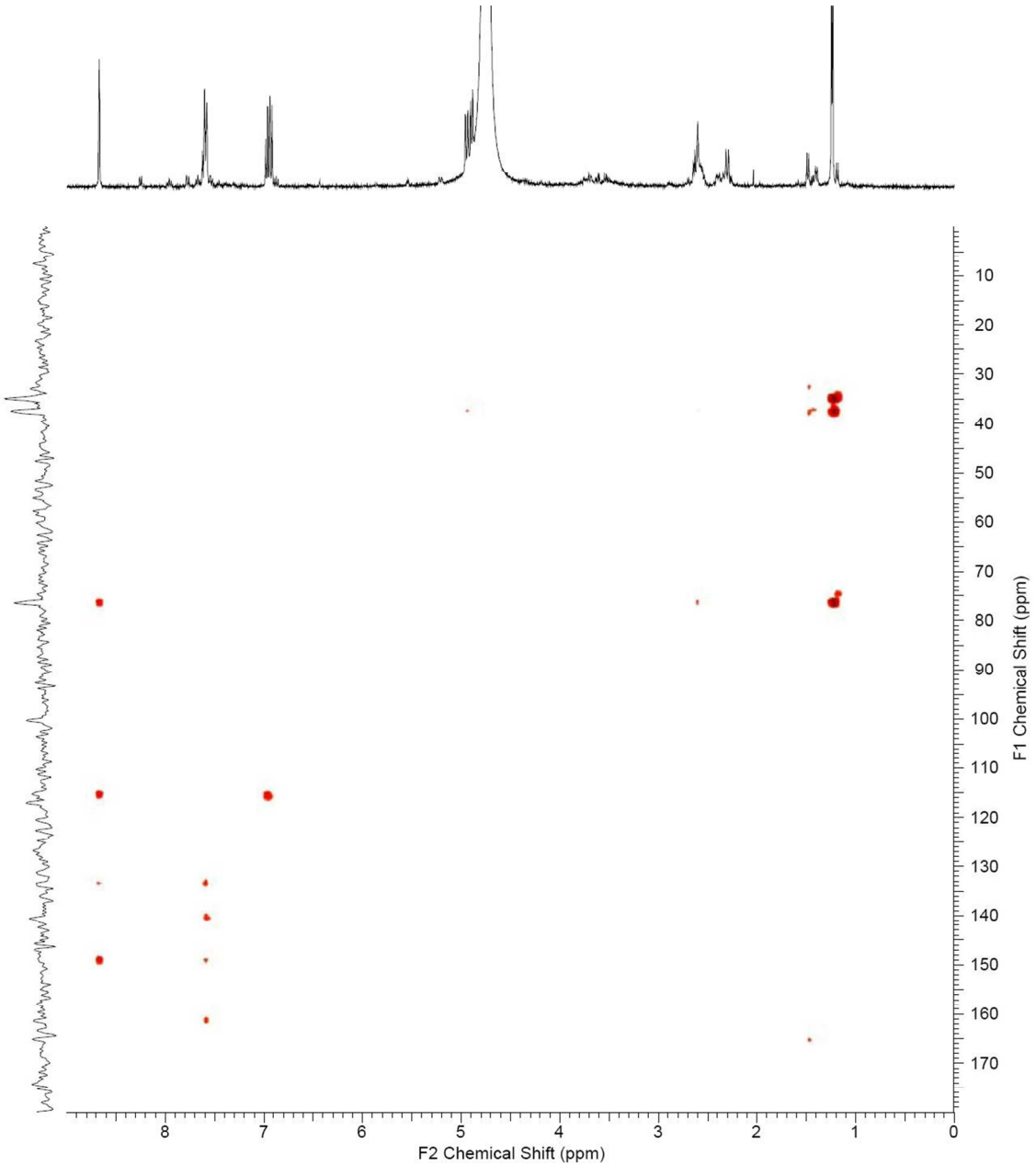


Figure S13: HMBC spectrum of *o*-aminobenzaldehyde derivatized MeP5C (400 MHz) in D_2O.

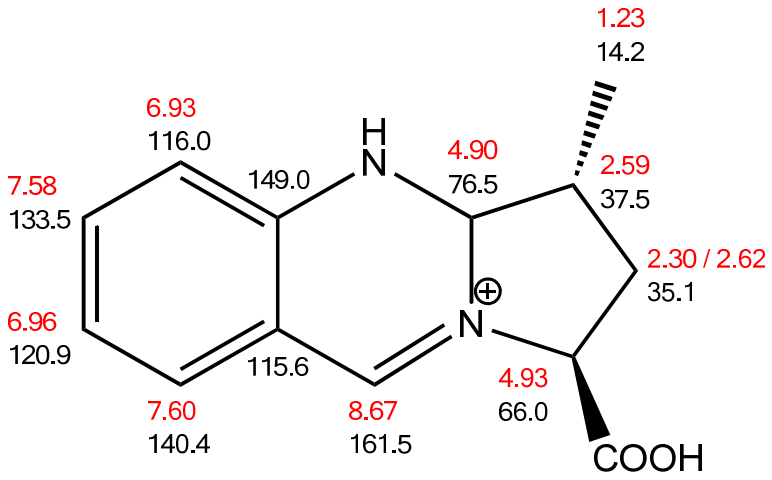


Figure S14: ^{1}H and ^{13}C parameters of *o*-aminobenzaldehyde derivatized MeP5C extracted from ^{1}D and ^{2}D-NMR (Figure S10-S13).