Post-genomic approach based discovery of alkylresorcinols from a cricket-associated fungus, Penicillium soppi

Akiho Kaneko[a], Yohei Morishita[a], Kento Tsukada[a] Tohru Taniguchi[b], Teigo Asai*[a]

[a]Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo

[b]Faculty of Advanced Life Science, Frontier Research Center for Post-Genome Science and Technology, Hokkaido University

Table of contents

General experimental procedure

Analytical TLC were performed on silica gel 60 F254 (Merck) and RP-18 F254 (Merck). Column chromatography was carried out on silica gel 60 (70–230 and 40-50 mesh). The 500 MHz NMR spectra were recorded on a Bruker AVANCE III 500 spectrometer (^{1}H NMR, 500 MHz; ^{13}C NMR, 126 MHz). Chemical shifts for ^{1}H and ^{13}C NMR are given in parts per million (δ) relative to tetramethylsilane (δ_H 0.00) and residual solvent signals (δ_C 77.0) for $CDCl_3$ as internal standards. Mass spectra were measured on Exactive Orbitrap Mass Spectrometer (Thermo Fischer Scientific). IR and VCD spectra were measured on a JASCO FVS-6000 spectrometer. UV spectra were recorded on a JASCO-V-730 spectrophotometer. HPLC analysis was performed on a Hitachi LaChrom Elite Pump L-2100, Hitachi LaChrom Elite Autosampler L-2200, and Hitachi LaChrom Elite Diode Array Detector L-2450 (Hitachi), which equipped with COSMOSIL Packed Column 5C18-MS-II (ϕ 4.6 mm×150 mm) (nacalai tesque). LC-MS analysis was performed on a Chrommaster 5610 MS Detector (HITACHI, Ltd), a Chromaster 5110 Pump (HITACHI, Ltd) and Chromaster 5430 Diode Array Detector (HITACHI, Ltd), which equipped with COSMOSIL Packed Column 5C18-MS-II (ϕ 4.6 mm×150 mm) (nacalai tesque).

Biosynthetic gene donating fungal material

Penicillium soppi Okera-1 was isolated from the surface-sterilized a cricket collected in 2013 in the Campus of Tohoku University. The fungus (Strain Okera-1) was identified by 28S rDNA gene D1/D2 region sequencing and species identification. The fungus was cultured in potato dextrose agar and the mycelium was ground to a fine powder in liquid N_2. Genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega), and 28S rDNA gene D1/D2 region was amplified by PCR using primers NL1 (5'-GCATATCAATAAGCGGAGGAAAAG-3') and NL4 (5'-GGTCCGTGTTTCAAGACGG-3'). PCR products were sequenced (ABI PRISMTM 310 Genetic Analyzer). The following consensus sequence were used in a BLAST search against deposited sequences.

5'-
GCATATCAATAAGCGGAGGAAAAGAAACCAACAGGGATTGCCTCAGTAACGGCGAGTGAAGC
GGCAAGAGCTCAAATTTGAAAGCTGGCTCCTTCGGGGTCCGCATTGTAATTTGCAGAGGATGCT
TCGGGAGCGGTCCCCATCTAAGTGCCCTGGAACGGGACGTCATAGAGGGTGAGAATCCCGTAT
GGGATGGGGTGTCCGCGCCCGTGTGAAGCTCCTTCGACGAGTCGAGTTGTTTGGGAATGCAGCT
CTAAATGGGTGGTAAATTTCATCTAAAGCTAAATATTGGCCGGAGACCGATAGCGCACAAGTA
GAGTGATCGAAAGATGAAAAGCACTTTGAAAAGAGAGTTAAAAAGCACGTGAAATTGTTGAAA
GGGAAGCGCTTGCGACCAGACTCGCTCGTGGGGTTCAGCCGGCATTCGTGCCGGTGTACTTCCC
CGCGAGCGGGCCAGCGTCGGTTTGGGCGGTCGGTCAAAGGCCCTCGGAAGGTAACGCCCCCTC
GGGGGCGTCTTATAGCCGAGGGTGCAATGCGACCTGCCTGGACCGAGGAACGCGCTTCGGCTC
GGACGCTGGCATAATGGTCGTAAGCGACCCGTCTTGAAACACGGACC-3'

Heterologous host strain

Asperillus oryzae NSAR1 (*niaD*$^-$, *sC*$^-$, Δ*argB*, *adeA*$^-$) was used as the host for fungal expression.

Construction of the *pspA*, *pspB* and/or *pspC* coexpression system

The genes were amplified by PrimeSTAR® MAX DNA Polymerase (TAKARA) with primers in Table S4. *Escherichia coli* DH5α were used for cloning, following standard recombinant DNA techniques. Fungal expression plasmid pUARA2 or pUAdeA2 possessing the α-amylase promoter (*amyB*) of *A. oryzae* and auxotrophic marker *argB* or *adeA* of *A. nidulans* respectively was used. The half-length *pspA* gene were amplified with the primers *pspA*_IFpUKpnI-FW and *pspA*_R1, or *pspA*_F1 and *pspA*_IFpUKpnI-FW. The *pspB* gene was amplified with the primers *pspB*_IFpUKpnI-FW and *pspB*_IFpUKpnI-RV. The *pspC* gene was amplified with the primers *pspC*_IFpUNotI-FW and *pspC*_IFpUNotI-RV. The PCR products were purified. The resultant fragment *pspA* was subcloned into pUARA2 which had been digested with KpnI to yield pUARA2-*pspA*. The *pspB* was subcloned into pUAdeA2 which had been digested with KpnI to yield pUAdeA2-*pspB*. The *pspA, psp1B* and *pspC* were subcloned into pUARA2 which had been digested with KpnI and NotI to yield pUARA2-*pspABC*. *A. oryzae* NSAR1 was transformed with pUARA2-*pspA* or/and pUAdeA2-*pspB* or pUARA2-*pspABC* to construct AO-*pspA*, AO-*pspB*, AO-*pspAB*, or AO-*pspABC* respectively.

Culture medium for AO-*pspA*, AO-*pspB*, AO-*pspAB*, AO-*pspABC* and control

CPS medium (+ adenine): 1.75% Czapek-Dox Broth, 0.17% Meat peptone, 0.17% Soy peptone, 0.17% Casein peptone, 1.0% Soluble Starch, 0.5% Maltose H, 0.015% Adenine sulfate dihydrate in 60/150 mL Distilled water.

Cultivation of the transformants in CPS medium and HPLC analysis

The transformants were cultivated on a selection agar plant at 30°C and its mycelia was inoculated in CPS medium and incubated at 30°C for 6 days. The cultured mycelia were harvested and freeze-dried. The 40 mg of the crushed mycelia was extracted with 1 mL EtOAc (10 % MeOH) for 60 min. After centrifuged at 13,500 rpm for 15 min, 500 μL was transferred to a new tube and concentrated under reduced pressure to obtain EtOAc (10 % MeOH) extract. The extract was resuspended with 100 μL MeOH, centrifuged again, and 20 μL was injected into HPLC. Flow rate; 1 mL/min, Solvent gradient system: acetonitrile and water with 0.01% TFA (0-1.5 min: 20:80, 1.5-11.5 min: 20:80 to 100: 0, 11.5-25 min: 100:0). Absorbance was monitored at 280 nm.

Isolation of compound 1-3

AO-*pspABC* was cultivated in CPS medium (6.3 L; 150 mL x 42) at 30ºC for 6 days. The cultured mycelia were harvested, freeze-dried, crushed to powder and extracted with EtOAc (10% MeOH). The extract was dissolved in EtOAc and the solution was washed by distilled water. After dried the EtOAc layer, the extract was dissolved in MeOH (1% water) and the solution was washed by *n*-hexane. The MeOH layer was dried to give the mixture containing **1**-**3** (350 mg). The mixture was subjected to flash silica gel column chromatography eluted with *n*-hexane-EtOAc (6:1 - 1:1) with 1% acetic acid to give 5 fractions (Fr.1 – Fr. 5). HPLC analysis of all the fractions showed that Fr. 2 contained **1** and **2**, and Fr. 3 contained **1** and Fr. 4 contained **3**. Fr. 2 (54.2 mg) was subjected to flash silica gel column chromatography eluted with $CHCl_3$-MeOH (60:1 - 20:1) to give clude **2** and **1** containing Fr. 2-3. Clude **2** (10.4 mg) was separated by PTLC

($CHCl_3$: MeOH : acetic acid = 40:1:0.2) to give **2** (6.8 mg). Fr. 2-3 (5.2 mg) was combined to Fr. 3 (8.2 mg) and the mixture was subjected to flash silica gel column chromatography eluted with $CHCl_3$-MeOH (60:1 - 20:1) to give **1** (2.2 mg). Fr.4 (22.3 mg) was subjected to flash silica gel column chromatography eluted with $CHCl_3$-MeOH (60:1 - 20:1) to clude **3** and $CHCl_3$ soluble portion was removed to give **3** (1.8 mg).

Soppiline A (1)

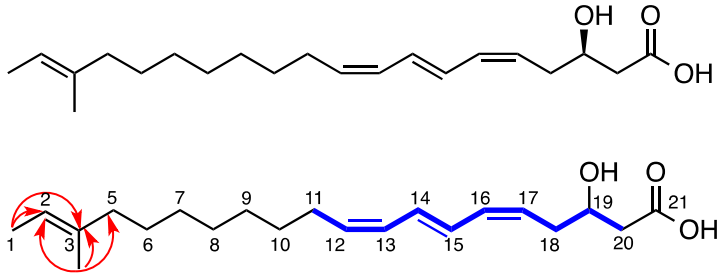
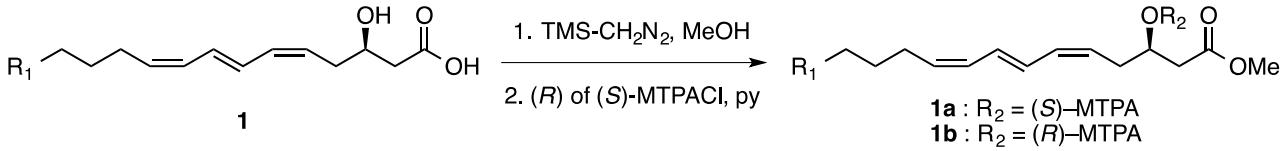


Fig. S1 Structure of 1 and 2D NMR correlation of compound 1 (^{1}H-^{1}H COSY in blue bold line and HMBC in red arrow).

Compound 1: Colorless amorphous; UV (MeOH) λ max nm (log ε) 264 (4.19), 273 (4.27), 283 (4.17); IR (KBr) ν_{max} (cm^{-1}) 3421, 2924, 2853, 1715; negative HRESIMS: *m/z* 333.2436 $[M–H]^-$ (333.2435 calcd for $C_{21}H_{33}O_3$). ^{1}H and ^{13}C NMR data, see Supplementary Table S1.

Synthesis of (*S*)- and (*R*)-MTPA esters of methylated 1.



Compound **1** (1.0 mg) was treated with 30 μL of TMS-diazomethane (ca. 10% in Hexane, ca. 0.6mol/L) and one drop of MeOH to convert corresponding methyl ester derivative of **1** (1.0 mg).

For the preparation of the (*S*)-MTPA ester **1a**, 0.5 mg of the methyl ester was dissolved in 30 μL pyridine and 2 μL (*R*)-MTPA chloride was added. The mixture was incubated at room temperature for 2 hours and then EtOAc and water were added. The water layer was extracted with EtOAc twice and concentrated. The resulting mixture was separated by PTLC (*n*-hexane-EtOAc = 4/1) to give (*S*)-MTPA ester **1a**. The (*R*)-MTPA ester **1b** was prepared in the same manner by the addition of 2 μL (*S*)-MTPA chloride.

Compound 1a: HRESIMS: *m/z* 587.2942 $[M+Na]^+$ (587.2966 calcd. for $C_{32}H_{43}O_5F_3Na$), ^{1}H NMR data, see Supplementary Table S2.

Compound 1b: HRESIMS: *m/z* 587.2948 $[M+Na]^+$ (587.2966 calcd. for $C_{32}H_{43}O_5F_3Na$), ^{1}H NMR data, see Supplementary Table S2.

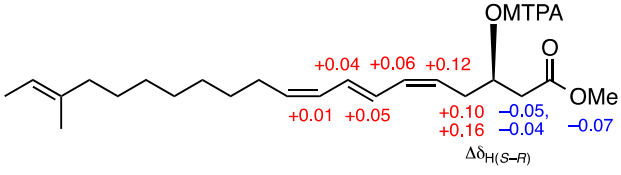
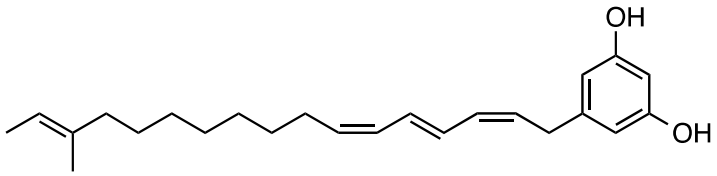


Fig. S2 $\Delta\delta_{H(S-R)}$ value calculated from 1a and 1b

Soppiline B (2)



Compound 2: Colorless amorphous; UV (MeOH) λ max nm (log ε) 200 (4.53), 269 (4.20), 276 (4.26), 287 (4.15); IR (KBr) ν_{max} (cm^{-1}) 3376, 2925, 2854, 1600, 1153, 837; negative HRESIMS: *m*/*z* 353.2484 $[M–H]^-$ (353.2486 calcd for $C_{24}H_{33}O_2$). ^1H and ^{13}C NMR data, see Supplementary Table S1.

Soppiline C (3)

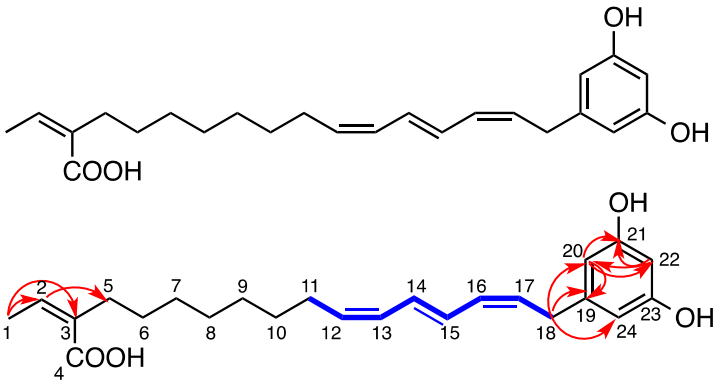


Fig. S3 Structure of 3 and 2D NMR correlation of compound 3 (^1H-^1H COSY in blue bold line and HMBC in red arrow).

Compound 3: Colorless amortphous; UV (MeOH) λ max nm (log ε) 200 (4.41), 269 (4.24), 275 (4.29), 285 (4.19); IR (KBr) ν_{max} (cm^{-1}) 3335, 2925, 2853, 1685, 1602, 1154, 839; negative HRESIMS: *m*/*z* 383.2229 $[M–H]^-$ (383.2228 calcd for $C_{24}H_{31}O_4$). ^1H and ^{13}C NMR data, see Supplementary Table S1.

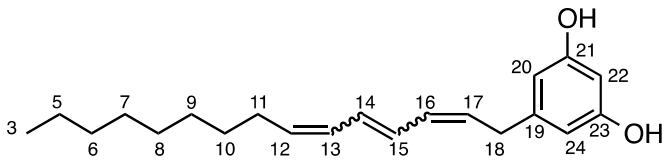
Table S1. ^{13}C (125 MHz) and ^{1}H NMR (500 MHz) data for **1**–**3**.[a]

	1 (soppiline A)[b]		**2** (soppiline B)[b]		**3** (soppiline C)[c]	
Position	^{13}C	^{1}H (multi, *J* in Hz)	^{13}C	^{1}H (multi, *J* in Hz)		
1	13.3	1.56 (3H, d, 6.6)	13.3	1.57 (3H, d, 6.6)	15.7	19.7 (3H, d, 7.1)
2	118.0	5.18 (1H, m)	118.0	5.19 (1H, m)	136.8	6.02 (1H, q, 7.1)
3	136.1	------	136.1	------	132.9	------
4	15.5	1.58 (3H, s)	15.6	1.57 (3H, s)	170.7	------
5	39.6	1.94 (2H, brt, 7.5)	39.7	1.95 (2H, brt, 7.7)	34.5	2.23 (2H, m)
6	29.6[d]	1.37 (2H, m)[e]	28.0	1.35 (2H, m)	29.5[h]	1.30 (2H, m)[i]
7	29.4[d]	1.29 (2H, m)[e]	29.2	1.30 (2H, m)	29.2[h]	1.30 (2H, m)[i]
8	29.2[d]	1.29 (2H, m)[e]	29.2[f]	1.30 (2H, m)[g]	29.1[h]	1.30 (2H, m)[i]
9	29.2[d]	1.29 (2H, m)[e]	29.4[f]	1.25 (2H, m)[g]	29.0[h]	1.25 (2H, m)[i]
10	28.0[d]	1.37 (2H, m)[e]	29.7	1.37 (2H, m)	29.0[h]	1.37 (2H, m)[i]
11	27.9	2.19 (2H, brq, 7.5)	27.9	2.21 (2H, brq, 7.7)	27.8	2.19 (2H, m)
12	133.9	5.50 (1H, dt, 11.0, 7.5)	133.6	5.48 (1H, dt, 10.9, 7.7)	133.2	5.46 (1H, dt, 10.5, 7.8)
13	128.4	6.05 (1H, t, 11.0)	128.5	6.07 (1H, t, 10.9)	128.6	6.06 (1H, t)
14	129.7	6.53 (1H, dd, 14.8, 11.0)	129.2	6.56 (1H, dd, 14.9, 10.9)	128.8	6.50 (1H, m)
15	126.9	6.42 (1H, dd, 14.8, 11.0)	127.1	6.51 (1H, dd, 14.9, 10.7)	127.4	6.56 (1H, m)
16	132.3	6.24 (1H, t, 11.0)	130.2	6.20 (1H, t, 10.7)	129.6	6.16 (1H, t, 10.4)
17	125.5	5.44 (1H, dt, 11.0, 7.7)	129.2	5.56 (1H, dt, 10.7, 7.9)	129.9	5.57 (1H, dt, 10.4, 7.9)
18	34.9	2.48 (1H, m) 2.44 (1H, m)	33.8	3.34 (2H, brd, 7.9)	33.9	3.41 (2H, brd, 7.9)
19	67.8	4.11 (1H, m)	143.6	------	143.1	------
20 (24)	40.2	2.61 (1H, dd, 16.6, 3.3) 2.52 (1H, dd, 16.6, 8.9)	108.0	6.26 (2H, d, 2.0)	107.3	6.21 (2H, brs)
21 (23)	175.7		156.8	------	156.8	------
22	8.3	2.11 (3H, s)	100.6	6.19 (1H, d, 2.0)	100.6	6.16 (1H, brs)

[a] Assignment for all compounds were based on COSY, HMQC, HMBC experiments.
[b] in $CDCl_3$.
[c] in $CDCl_3$ (10% CD_3OD).
[d-h] signals can be exchangeable each other.

Determination of the geometry of the double bonds in 2 by GIAO NMR Calculations

2



2a (12*Z*, 14*E*, 16*Z*)
2b (12*E*, 14*E*, 16*Z*)
2c (12*E*, 14*E*, 16*E*)
2d (12*Z*, 14*E*, 16*E*)

In addition to NOESY analysis, ^{13}C NMR calculations were performed to determine the double bond geometries of C-12/C-13, C-14/C-15, and C-16/C-17. To reduce the calculation cost, truncated model structures **2a-d** were constructed *in silico* and their stable conformers were generated by using MMFF conformational search on Spartan'10 software.[1] Conformers within a 12 kJ/mol energy window from the most stable ones for each molecule were optimized using DFT/B3LYP/6-31G(d)/PCM(chloroform) on Gaussian 16 program.[2] Resultant stable conformers for each molecule were submitted to GIAO ^{13}C NMR calculations using DFT/ωB97X-D/6-31G(d,p)/PCM(chloroform). The calculated shielding constants for each conformer were averaged on the basis of its Boltzmann population, and corrected using that predicted for TMS using the same level of theory. The obtained predicted ^{13}C shielding constants for each model and the observed ones for **2** are shown in Table S2. Comparison of the calculated data for **2a-d** found relatively large deviations in the chemical shift for the carbon atoms adjacent to the double bond when its geometry is changed (e.g. change in the C-12/C-13 double bond geometry resulted in a large difference in the ^{13}C signal for C-11 and C-14). When compared with the observed NMR data for C10-C24, the deviation of the calculated shifts are less than 5 ppm, except for C-11 of **2b** and **2c** and for C-18 of **2c** and **2d**. Furthermore, DP4 analysis [3] using the values in Table S1 provided 100% probability for the model **2a** and 0% for the others. These results corroborated the 12*Z*,14*E*,16*Z* geometry of **2** proposed by the NOESY analysis.

	obsd **2** [ppm]	calcd **2a** (12*Z*,14*E*,16*Z*) [ppm]	calcd **2b** (12*E*,14*E*,16*Z*) [ppm]	calcd **2c** (12*E*,14*E*,16*E*) [ppm]	calcd **2d** (12*Z*,14*E*,16*E*) [ppm]
C10	29.7	32.1	32.3	32.4	32.2
C11	27.9	30.9	36.1	36.1	30.8
C12	133.6	132.7	135.7	135.0	132.1
C13	128.5	125.6	127.5	127.6	125.6
C14	129.2	127.1	132.0	130.4	125.4
C15	127.1	124.9	123.3	128.2	129.9
C16	130.2	127.0	127.2	129.7	129.4
C17	129.2	129.4	128.7	131.5	132.3
C18	33.8	36.0	35.9	41.2	41.3
C19	143.6	142.1	142.2	142.0	141.9
C20	108.0	103.8	103.3	103.8	103.4
C21	156.8	154.3	154.5	154.2	154.5
C22	100.6	96.2	96.2	96.2	96.2
C23	156.8	154.6	154.5	154.7	154.5
C24	108.0	105.3	105.6	105.1	105.5
DP4		100%	0%	0%	0%

Table S2. Comparison of the observed ^{13}C NMR chemical shift of **2** in $CDCl_3$ and calculated values for **2a-d** at DFT/ωB97X-D/6-31G(d,p)/PCM(chloroform). The calculated values deviated from the observed ones for more than 5 ppm are highlighted in red. The DP4 probability scores are shown in the bottom line.

[1] SPARTAN'10, Wavefunction, Inc., Irvine, CA.
[2] Gaussian 16, Revision A.03, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2016.
[3] Smith, S. G.; Goodman, J. M. *J. Am. Chem. Soc.* **2010**, *132*, 12946-12959.

Table S3. ^{13}C (125 MHz) and ^{1}H NMR (500 MHz) data for **1a** and **1b** (in $CDCl_3$).

	1a	**1b**	$\Delta\delta_{H(S-R)}$
Position	^{1}H (multi, *J* in Hz)	^{1}H (multi, *J* in Hz)	
1	1.57 (3H, d, 6.6)	1.57 (3H, d, 6.6)	0.00
2	5.18 (1H, m)	5.18 (1H, m)	0.00
4	1.58 (3H, s)	1.58 (3H, s)	0.00
5	1.95 (2H, brt, 7.6)	1.95 (2H, brt, 7.8)	0.00
6-10	1.40-1.20 (10H, m)	1.40-1.20 (10H, m)	--.---
11	2.19 (2H, brq, 7.6)	2.18 (2H, brq, 7.2)	**+0.01**
12	5.51 (1H, dt, 10.7, 7.5)	5.48 (1H, dt, 10.8, 7.5)	**+0.03**
13	6.04 (1H, t, 10.7)	6.03 (1H, brt, 10.8)	**+0.01**
14	6.53 (1H, dd, 14.6, 10.7)	6.49 (1H, dd, 14.9, 10.8)	**+0.04**
15	6.39 (1H, dd, 14.6, 10.9)	6.34 (1H, dd, 14.9, 11.0)	**+0.05**
16	6.22 (1H, t, 10.9)	6.16 (1H, t, 11.0)	**+0.06**
17	5.37 (1H, dt, 10.9, 7.5)	5.25 (1H, dt, 11.0, 8.0)	**+0.12**
18	2.71 (1H, m)	2.61 (1H, m) **+0.10**	
	2.68 (1H, m)	2.52 (1H, m)	**+0.16**
19	5.50 (1H, m)	5.52 (1H, m)	--.---
20	2.64 (1H, m)	2.69 (1H, dd, 16.3 8.5)	**–0.05**
	2.60 (1H, m)	2.64 (1H, dd, 16.3, 4.2)	**–0.04**
COOMe	3.58 (3H, s)	3.65 (3H, s)	**–0.07**
OMe	3.52 (3H, s)	3.52 (3H, s)	0.00
Ph	7.51-7.38 (5H)	7.51-7.38 (5H)	--.---

Culture medium for *Penicillium soppi* Okera-1

CPS medium (+adenine): 1.75% Czapek-Dox Broth, 0.17% Meat peptone, 0.17% Soy peptone, 0.17% Casein peptone, 1.0% Soluble Starch, 0.5% Maltose H, 0.015% Adenine sulfate dihydrate in 60 mL Distilled water.
PDB medium (+adenine): 2.4% Potato Dextrose Broth and 0.015% Adenine sulfate dihydrate in 60 mL Distilled water.
MYG medium (+adenine): 1.0 % Malt extract, 0.4% Yeast extract, 0.4% D-Glucose, 1.0 % Maltose H, 0.015% Adenine sulfate dihydrate in 60 mL Distilled water.

Cultivation of *Penicillium soppi* Okera-1 in CPS medium and HPLC analysis

Penicillium soppi Okera-1 was cultivated on a PDB agar plate at 30°C and its mycelia was inoculated in MYG, PDB or CPS medium and incubated at 30°C for 6 days. The cultured mycelia were harvested and freeze-dried. The 40 mg of the crushed mycelia was extracted with 1 mL EtOAc (10 % MeOH) for 60 min. After centrifuged at 13,500 rpm for 15 min, 500 μL was transferred to a new tube and concentrated under reduced pressure to obtain EtOAc (10 % MeOH) extract. The extract was resuspended with 100 μL MeOH, centrifuged again, and 20 μL was injected into HPLC. Flow rate; 1 mL/min, Solvent gradient system: acetonitrile and water with 0.01% TFA (0-1.5 min: 20:80, 1.5-11.5 min: 20:80 to 100: 0, 11.5-20 min: 100:0). Absorbance was monitored at 280 nm.

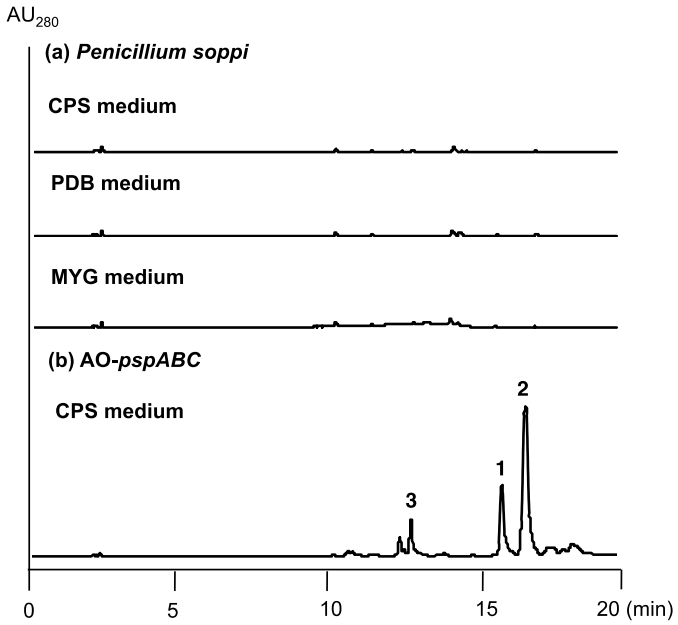


Fig. S4 Comparison of HPLC profiles of AO-*pspABC* and *P. soppi*.

HPLC profiles of mycelial EtOAc (10% MeOH) extracts of AO-*pspABC* in CPS media (a) and *P. soppi* in some laboratory culture conditions.

Table S3. *psp* cluster in a cricket -associated fungus, *Penicillium soppi* Okera-1.

Gene name	Size	Protein homologue (accession number)	identity (%)/ similarity (%)
pspA	8420 bp	Fum1 [*Fusarium verticillioides* 7600] (W7LKX1.1)	41/60
pspB	1882 bp	StlA, [*Dictyostelium discoideum*] (Q55E72.1)	27/50
pspC	1912 bp	FsdH, [*Fusarium heterosporum*] (S0ARX1.1)	44/64

psp1A

Nucleotide

ATGCTTGCACAAGATGTCGAATTCGTGGACCTTCCTCCACCCGAAGCCACTGCTGGGGCCGCTACAACGGACAATGAAACATCTTCCTTCAATT
CAAATCCCGTGCCAACCCCCAGCGAGGCGAGCTCGATAGGTCCTCCCCATCAACTCCCTGTTCCCGTTCCCGACGGTGATCAACCGCCATTGGT
GGAGCCTATGGCCATTTGCGGCATGGCAATGCGTCTTCCGGGCGGCATCCGCGATGCGGAAGGCTTCTGGGATCTATTATACAACAAGCGGAGT
GGGAGATGTCGCGTCCCGAAAGACCGTTATAACGTCGAGAACTGGTTTGGACCGGGAAAGATTGGGCACGTCGCCAGTGAGTACGGCTACTTCT
TGGATGATGTAGACCTGCGAAATGCGGATGCTTCTTTTTGGTCGATGACCAAACAAGAGATTGAGGCGATGGACCCTCAGCAGCGACTTTCCCT
GGAGGTCACGTATGAGTGTCTTCAAAACGCTGGACAGCGGCCTGAAGAACTTCGCGGCCGCAAAATCGGTGTTTATCTCGGCACCTTTGAGGGA
GACTGGCTGGAACTCGATGGGAGAGATCCTCAGCACTATCATATGTACCGTCTCACCGGATATGGAGATTATATGTCTGCCAATCGCATTCACT
ACGAGTTCGGATTTATGGGTCCCAGGTAACTTCATAACTATGTTGATACCGAAGAACACTATTGATCGTTGGAAATTGCAGTGTAACAATTCGA
ACCGCCTGCTCCTCTTCTCTCACCGGTCTGTATGATGCCTGTCATGCTATCTCCGCTGGGGACTGCGACTCCGCCATCGTGGCATGTGCAAATA
TTATCTACTCTCCCCGAACCTCTATTACCATGCAAGAACAGGGTGTCATATCGCCGAGCGGATTCTGCAAGACTTTTGATGCAAATGCCGACGG
ATATGCACGTGGTGAGGCAGTTTCGGCTGTGTACGTGAAAAAGCTCAGCGATGCTATTCGTGATGGAGATCCCATTCGCTCTGTTATCCGATCA
ACATGCATCAACGCCGGCGGCAAAGCGTCAACTCTTACAGCCCCAAATACTGCTGCACACGAGACATTGATCCGTCGTGGTCACGAGTTGGCCG
GTGTCACGGATTTCTCCAAGACCGCTATGATAGAATGCCATGGTACAGGAACAGCGGTGGGTCTATTACCATTATCAATTGCGATGTCTTGCTT
TTTGGTGATCATACTAAATCTTATCTAGGTCGGCGATCCCATCGAGACAGCGGCTGTAGCTAATGTGTTTGGAGAGCACGGCATTTATATTGGA
TCGGTGAGCGACGGGCAGATATCTTCGACTCAATCGAAAGATTCTAATACCCCTTCAGGTCAAAACCAATCTGGGACATAGCGAAGGTGCATCA
GGCTTGGCTAGTATCATTAAAATGACCCTGGCGCTTGAACACAAAATAATTCCGCCCAACATCAACTTCACGACTCCCAACCCCAAGAGTGCGT
ATCGCAACCCGAACTTTTGCTGGTAATTTGTGAACTAACGACTTAATCCTAGTTCCATTTGAACGATGCAAGCTGAAAGTACCCACAGAGCCAC
TTCCCTGGCCCAAAGACCGCGCTGAGCTGGTAGGAGTCAACTCCTTCGGTATAGGTGGGTCAAACGCCCATGTGAGTATATACCATAGAGTACA
GAACTACTCTGCAAGTGACTATGTGACTGACATAGACTTAAGGTTCTGTTGGGATCCGCTGCCTCCTTTGGAATAGGGTCAGTACAGCAGAAGA
TTATTGCTAGTGAACAATCTGCAGAGGTTGCAATGACCGAGCTGACACCACGCTTGTTGCTATTTTCCGCTAAGCATCAACAATCGTTGGAGCG
GATGGTCGCTAACCACCAAGCGTATTTCCTGTCCCACCCCGAGTCTCTTGATGATATGGCCTATTCTCTTGCCCTCAAGCGAGAAGAGCTTTCC
CATCGCAGCTTCTGTGTCACAAATGGCGAAGATGACTGGGTCCCTTCGAGGACTCACCGCACCAGTGGTAGAGCACCTCCCATGTTGATATTCA
CCTTCACCGGTCAGGGTGCCCAGTGGGCACAGATGGGTAAAGCACTTATCGACCAAGTGCCTCGTTTCCGACGCAGTATCGAGAAGCTGGATCA
AGTACTTCAAGCTCTTCCAACTCCTCCTCGATGGAAACTGATAGGTAAGTCAATCAGCGTTCTGGTTCCTCATCAATTGGACTCTTCTATAATT
GGAGTCTTTGTTTCTCACGCAACGCATAGATGAAATTCGAGCTTCGAAGAAGAAAAGTCGTCTTTCCGAGGCAGAGCTTTCACAGCCCTGCTGT
ACAGCAATTCAAATCGCTCTGGTCGATATTTTGGAACATTACGGAATCCACCCAGACGCCGTAATAGGCCACTCCAGTGGTGAGATTGGTGCCG
CTTATGCCTCTCACGCCATCAGTGGAGCAGATGCAATTCAAATCGCATTTTATCGAGGCTTAGTGATGTGCAGTTTGAATCCAGCGGAGCGTCC
TGGGGGCATGGCAGCCGTTGGTCTTGGGGCTGAAGAGCTCACACCTTACTTGCGACCAGGAGTTCGCGTGGGATGTGAAAATAGTCCCAACAGT
ACAACTCTAACAGGCGATAAAGTATCATTAGAGGAGACCATGAAAGCCATCAAGGAGGCCAACCCCGACGTCTTTGTCAGAGCACTCCAGGTGG
ACCGAGCCTATCACTCGCGTATGTCATGTGATTGTAATGGATGGACTGAAAATGCTAATAGCGTTGTTTGTAGACCATATGGAGACCGTGGCTC
CTGAATATGTCGAACTTCTCACCAACCAGAGGGTCCAAGCCATGGATCCCTCAGTCAAATTCTTTTCCAGCGTTACTGGCCGTCAAGTGGATCA
AAGCAAAGAGCTCGGCCCCTTGTATTGGGCCAAAAACCTTGTCTCTCCAGTCCGCTTCTCGACGGCCATGGAGGAATTGGTGCAGTCTCTGATT
GGACCTAAGGTCTTCCTGGAAATTGGACCTCATTCTGCGCTGGCTGGGCCTATTCGGCAGATCCTCCAGCATCATAAATCAACCGATGAGTATT
TCAACACATTGACTCGGGGCAGTGACTCCCACAAGGATCTGCTGAAGGCAGTTGGTGAAATGTGGCTGCAGAATATCCCAGTGAATCTGACTGC
AGTCTTGGGCGAAGGTCGCTTCCTGCCAGACTTGCCTCTCTACCCTTGGCACTATGAGGAGCCTCTCTGGTGCGAAAGTCGTCTCTCCAAGGAA
TGGCGCTTAAGGGAATTCCCTCATCATGACATCCTCGGATCAAGGGTCCTCGAAAGCACCGATCAGAATCCTAGCTGGCGTAATATTCTCCGTC
TCGATGTGGTGCCATGGATCAAGGAACACGAAGTAGCCGGTGAAATTGTCTTCCCGGGTGTTGGTTACATATGCATGGCTGGAGAAGCCATTCG
TCAACTTACTGGAGAGACTGGGTTCACAGCCCGCCGCGTTCACATCAAGGCTGCGCTGGTCATGCATCAGGGTCAAGACGTGGAAGTGATCACT
CAATTGCAGAGAATTCCACTTACCAATGCGGCCGACTCAAAGTGGTACAACTTCACTGTTCATTCCTATAATAAGGGAATCTGGGTTAAGCACA
TCTTTGGCCAAGTGTGCGCAGGGTCAGATCGAGAGCACCAAGCTCCTTCATTGGAGTCTTTGCCACGACAGTTGTCGCGGCGTGGGTGGTATCG
TAAGATGAAAGAAATGGGACTGGAATATGGGTCGCGATTCATGGGCCTGACTGACATGACTGCTCATCCCATTGAGCGCAAGACCATCGCCACT
GTGGTTAATGATATTCGTGAGGGAGAGTCCAAATATGCGGTACACCCTGTCTCTCTGGACTGTTTGCTTCAAGCCATTGTTCCTGCAACCTTCA
ATGGACTGACAAGACGATTCCAACATCTCGGAATCCCTACCTACATGGAAGAGATCTATGTATGCCCTCCGCTACAGCCAGAGATGATTATTGA
AGCTTGTGCCGACGAGCAACCTACAGCTGCTCTCTCGGGAAGTATCATCGCAGTCTCAAATGGACATGTTACAATTGACATCAGGGGCCTGCAG
ATGTCTGCAATTGGTGATGCAGCCGATGCCTCTGGGCAAGATCCACATGCCGCGGTTGAATTGGAGTGGCGTGAAGATATCAATCTCATCAGTG
ATGCTGCCAAACTGATCCACCCAGCGAAGGACCGAACCGATCTACACCACTTGCTGGATAGGTTTGCCTCGGCCTCCATGATGGATACATCTAC
CCGCCTTCGAGGGGTTGAGCCAACCCGATCGCATTTGACTCATTACCAGAAGTGGATTGAATCGACCGCAGATCTAATAAAGCTCGGCAAATAC
CCTGGCCTTCAGCCTGAGGATGAGATTGTCGAAGTCTCGGACGCGGAGCGGGTAAACATCATCGAGAGTCTGTATCTTTCTCTCTTGGAAACAG
ATGCGGCTGCGACCGCGACAGCCATCTATCGAATCTGGAAAGAGTGTCAGGGCATCTTCACAGGTGAAACCGATGAACTTGAGCTCCTTCTAGA
GGGTGAGGTTCTCCATTCGTTGTATGACTTCATGCAAAACTCAGAGTACCGGGGTGTTCCTCGAGCTTCTCGCACACCGGAAGCCAAACCTCCGG
GTTTTGGAGATCGGTGCCGGCACAGGAGGCACTACTGCGACGGTACTTCCTGCCTTGAAATCTCTTTATGGAGAGCGAATGTACCATTCATACA
CGTATACCGATATTAGTGCCGGATTTTTCCCACAGGCAAAGAAGCGATTTGAGAATTACCCTGGATTAGAATTTGCGACGTTGGATATTTCGCA
GGATCCGCTTTCACAGGGCTTCGAGGCAGAGTCCTTTGATTTGATTATTGCCTGCAATGTGAGCTCATCAAATCAACATGTCTGAGATATTCTG
CTAACAGAAAACTAAAGGTGCTCCATGCGACTTCAACTCTTCAGGACACGCTTACGAATGTTCGCAGGTTGCTCCATCCTCAGGGTCGGCTCTT
CCTTCAAGAATTATCGCCAGGTAAGTGCAATATGTTTTTCTCCCATTCATGGATATGAACGAAATTAATATCTAATTTGTTCCACTAGCCACAA

AATGGATCAACTATGTCATGGGTGTTCTTCCAGGATGGTGGTTGGGGGAGCAAGACGGCCGATACCCAGAGCCTTATATCGGCATCGACCAGTG
GGACGCCCTGCTTAGCCAGAGCGGCTTCTCGGGCATTAATTTGGTGAGCCACGATGGATACCTGAACAACAACATCGTTGCCAGGCCGGCGGCA
GACACACAAAGACAAAAGCGAATCACCCTACTTCACTCGTGTGAAGACAGCGCTTCTGTCACAACTTCAATAAGCCAACTTCTCTCCTCAGCTG
GATTTGGAATCGATCTCTATGCAATTGAGAATGCGAACATACCAACCCCCACGCAGCAGGATATCGTGTCCATTTTGGATCTTGACAGGCCATT
CTTCCATGACCTCCACGAATCCTTGTTTGAGAACCTGAAGGGTCTGCTTTCACAGCTGCGAGATACGGACAGCGGAATTCTTTGGGTTACCAGA
GCAAGCCAGGTAGGCTGCAAGGACCCTAGGTATGCGATGGTCAACGGAGTTGCTCGTGTTATTCGCACCGAGCTGAACATAGACTTTGCGACAC
TGGAGCTGGAGGATTTTGAGCAGGAAACCCTGGCTCTTATACCGCAGGTGCTGGGAGAGTTCCAGCAACGTATCTCTGAACCGAACATCAACAC
GACTACAGAATGGGCCGTTGTGGGACAAAAGCCACTGATCAGCAGATACCACTACATTCAAGTTGCCGAGGAGCTGAAGAACAACGCGGTCGCT
GATAGCTCCACTGTCAAGAAGTTGGAACAGAGCAGGCCCGGTCTTGTTGATACCTTGTGCTGGAAGAGCATGCCGATTTCACATGCGTTGGATG
AAAATGACGTTCTGGTTCAAGTCAAATGTGTTGGGATGAACTTCAAGGTATGCACAGATAGAGACGTTTCGGTGGTTGAGGGACGCTGGCTAAC
GCTTTACACTTAGGATGTTCTGATCAGCACCGGAGTGATCACAGAGAAATCCTCGATTGGCCGGGGACTAGGATACGAAGGAAGTGGTCTAGTG
CTCCAAGTGGGGTCGGCCGTGCACAAGCTTTCAGTGGGAGATCGAGTTATTATGAGCTCAAGTGGATCTCTCACCACCACACAGCAGCTTGACC
AACGCCTATGTGTTAAAATGCCGGACTCCATGACATACGAAGAGGGTGCTACTATGTCCGCGGTCTATTGCACCGCGATTCACTGCCTCCTAGA
TGTTGGTGGATTGCGCAAGGGACAGTCCGTCTTGATTCATTCCGCCAGTGGAGGTGTCGGTATTGCAGCCATGTACATTGCGCAGATGGTAGGC
GCGGAGGTAAGTCTTTATCGGAACCTGTCGAGTGAGCATTAGATCGATTTAGCTGACGGAACATCTCTTAGGTATATGCTACAGTCGGGAGTGA
GGAGAAAACTCAATCACTCATGTCCACTTTCAATATTCCGCGCAATCGTATATTCAACTCTCGCACCAGTGAGTTCCTTCCGCGCATAATGGAA
GAGACGAATGGTATGGGAGTGGACGTGGTGCTCAACTCACTTTCGGGTGAGCTTTTGCACGCTTCTTGGAAATGCACAGCAGAATTCGGTACTT
TCGTCGAGATTGGACGACGTGATTTCGTCGGCCAGGGCTTGCTTGACATGCAACCATTCGAGCCTAACAGAAGCTTTGTCGGGTTTGATCTTCT
TCTTTTCTCTAACAAACGACCGGAAAGGATTGAAAGGTATGTTAGATGATTCATTTCTGCTTCGTTGAAACAATGGGAGACTAAATCATTTCTT
ATAGTATCATGACTCGTGCGATGGACTATTACCGAGCCGGTTTCATTCAACCCATTAAGCCCATGACAATGTTTGATGCTGTTTCCATTGTTGA
TGCTATTCGCTACATGCAGCGCGGGCAGCACATAGGCAAGATTGTCATCACAATGCCAGAAAACAGTACCGAGTTGTCAGCTGAACCACCCCGG
CAAGAACTGGCTCTGCGTCAAGATCGAGCCTACCTATTTGTCGGTGGACTGGGTGGCTTGGGTCGTTCGATCGCAACCTGGCTTGTCGAGCATG
GAGCCCGCCATCTCGTGTTCTTGTCACGATCGGCGGGTAATGTTCCTGATGACGATCCATTTGTTCAAGAGCTAGCAGTCCTCGGATGCACCAC
TACCAGGATATCCGGTGATGTGTCTAAGCTTGATGATGTTCTTCTGGCTATCCGGGCATCTGGAAAGCCTGTTGGAGGTGTTCTACAATCTTCC
ATGGTACTTCGGGTAAGTCATACGCGTTCATAACCACATATCCCTCGTTTCATCCCCTTTAGCAACATGCTGACAAACTCCATCGCAGGACAAC
TCATTTGTGGATATGAATTGGGATGAGTGGCTGGGTGCAGTGCAGCCTAAGGTTCTAGGAACATGGAATCTGCACAACGCTCTCCTCAGCGAGC
AGCCGGAAGAAGCCCTCGACTTCTTCTTCTTATTCAGCTCAGCCGGCGCAATGAGTGGCCAATGGGGACAAGCCAACTACAATGCTGGCAACAC
ATTCCTGGACGCTTTTGTAGCCTATCGCCATTCTCTGGGGCTTCCTGCTTCCACAGTCAATATTGGCGTGATCCAAGACATCGGATACGTTAGC
CAGAACCCAGAGATCCTGGACTCACTACGTTCAACAGCGCAGTACCTCATGCGTGAGCCTGAGCTGCTTGAGTCAATCGAGCTCATGCTCCACA
GGTCTTCGCCCGCTGAATCAGTTGTCGACCACGCTGTAGGCCGGTATGTTACGCGGTCGCAGATCGGCATCGGCATGCGGTCCACGGTCCCCAT
GGACGCCCCGAGCAATCGAACCATTTGGAGAAAAGATCCCCGTATGCTGGTCTATAGGAACCTTGAAGTCCAATCCGGGCCAGTGTCGTCTTCT
ACCGGGTCCGATCAAGTTCTCACCCAGTTCCTCCGCGAGATCGGCTCCAATATGACGATGTTGAAAGCCCCCGAGACGGCTGAACTACTGGCCG
GTGAGATTGGTCGTACTTTGTTTGGATTCTTGATGCGGGCGGATACTGAAGTGGTTGATCTTGATGCTCCACTTGCTAGTGTTGGAATTGACTC
GCTCATCAGCATCGAGCTACGCAACTGGATTCGCCGAAAGATTGGTGTAGAAGTTACTGTGTTGGAGATTGTGCGGGCAGACAGTGTGAGGGAT
TTGGGTGTGCTTGCTCAAAAGAAATTGGCGGAAAAGTATGAAGCTCGTATGTAG

Amino acid

MLAQDVEFVDLPPPEATAGAATTDNETSSFNSNPVPTPSEASSIGPPHQLPVPVPDGDQPPLVEPMAICGMAMRLPGGIRDAEGFWDLLYNKRS
GRCRVPKDRYNVENWFGPGKIGHVASEYGYFLDDVDLRNADASFWSMTKQEIEAMDPQQRLSLEVTYECLQNAGQRPEELRGRKIGVYLGTFEG
DWLELDGRDPQHYHMYRLTGYGDYMSANRIHYEFGFMGPSVTIRTACSSSLTGLYDACHAISAGDCDSAIVACANIIYSPRTSITMQEQGVISP
SGFCKTFDANADGYARGEAVSAVYVKKLSDAIRDGDPIRSVIRSTCINAGGKASTLTAPNTAAHETLIRRGHELAGVTDFSKTAMIECHGTGTA
VGDPIETAAVANVFGEHGIYIGSVKTNLGHSEGASGLASIIKMTLALEHKIIPPNINFTTPNPKIPFERCKLKVPTEPLPWPKDRAELVGVNSF
GIGGSNAHVLLGSAASFGIGSVQQKIIASEQSAEVAMTELTPRLLLFSAKHQQSLERMVANHQAYFLSHPESLDDMAYSLALKREELSHRSFCV
TNGEDDWVPSRTHRTSGRAPPMLIFTFTGQGAQWAQMGKALIDQVPRFRRSIEKLDQVLQALPTPPRWKLIDEIRASKKKSRLSEAELSQPCCT
AIQIALVDILEHYGIHPDAVIGHSSGEIGAAYASHAISGADAIQIAFYRGLVMCSLNPAERPGGMAAVGLGAEELTPYLRPGVRVGCENSPNST
TLTGDKVSLEETMKAIKEANPDVFVRALQVDRAYHSHHMETVAPEYVELLTNQRVQAMDPSVKFFSSVTGRQVDQSKELGPLYWAKNLVSPVRF
STAMEELVQSLIGPKVFLEIGPHSALAGPIRQILQHHKSTDEYFNTLTRGSDSHKDLLKAVGEMWLQNIPVNLTAVLGEGRFLPDLPLYPWHYE
EPLWCESRLSKEWRLREFPHHDILGSRVLESTDQNPSWRNILRLDVVPWIKEHEVAGEIVFPGVGYICMAGEAIRQLTGETGFTARRVHIKAAL
VMHQGQDVEVITQLQRIPLTNAADSKWYNFTVHSYNKGIWVKHIFGQVCAGSDREHQAPSLESLPRQLSRRGWYRKMKEMGLEYGSRFMGLTDM
TAHPIERKTIATVVNDIREGESKYAVHPVSLDCLLQAIVPATFNGLTRRFQHLGIPTYMEEIYVCPPLQPEMIIEACADEQPTAALSGSIIAVS
NGHVTIDIRGLQMSAIGDAADASGQDPHAAVELEWREDINLISDAAKLIHPAKDRTDLHHLLDRFASASMMDTSTRLRGVEPTRSHLTHYQKWI
ESTADLIKLGKYPGLQPEDEIVEVSDAERVNIIESLYLSLLETDAAATATAIYRIWKECQGIFTGETDELELLLEGEVLHSLYDFMQNSEYRVF
LELLAHRKPNLRVLEIGAGTGGTTATVLPALKSLYGERMYHSYTYTDISAGFFPQAKKRFENYPGLEFATLDISQDPLSQGFEAESFDLIIACN
VLHATSTLQDTLTNVRRLLHPQGRLFLQELSPATKWINYVMGVLPGWWLGEQDGRYPEPYIGIDQWDALLSQSGFSGINLVSHDGYLNNNIVAR
PAADTQRQKRITLLHSCEDSASVTTSISQLLSSAGFGIDLYAIENANIPTPTQQDIVSILDLDRPFFHDLHESLFENLKGLLSQLRDTDSGILW
VTRASQVGCKDPRYAMVNGVARVIRTELNIDFATLELEDFEQETLALIPQVLGEFQQRISEPNINTTTEWAVVGQKPLISRYHYIQVAEELKNN
AVADSSTVKKLEQSRPGLVDTLCWKSMPISHALDENDVLVQVKCVGMNFKDVLISTGVITEKSSIGRGLGYEGSGLVLQVGSAVHKLSVGDRVI
MSSSGSLTTTQQLDQRLCVKMPDSMTYEEGATMSAVYCTAIHCLLDVGGLRKGQSVLIHSASGGVGIAAMYIAQMVGAEVYATVGSEEKTQSLM
STFNIPRNRIFNSRTSEFLPRIMEETNGMGVDVVLNSLSGELLHASWKCTAEFGTFVEIGRRDFVGQGLLDMQPFEPNRSFVGFDLLLFSNKRP
ERIESIMTRAMDYYRAGFIQPIKPMTMFDAVSIVDAIRYMQRGQHIGKIVITMPENSTELSAEPPRQELALRQDRAYLFVGGLGGLGRSIATWL
VEHGARHLVFLSRSAGNVPDDDPFVQELAVLGCTTTRISGDVSKLDDVLLAIRASGKPVGGVLQSSMVLRDNSFVDMNWDEWLGAVQPKVLGTW
NLHNALLSEQPEEALDFFFLFSSAGAMSGQWGQANYNAGNTFLDAFVAYRHSLGLPASTVNIGVIQDIGYVSQNPEILDSLRSTAQYLMREPEL
LESIELMLHRSSPAESVVDHAVGRYVTRSQIGIGMRSTVPMDAPSNRTIWRKDPRMLVYRNLEVQSGPVSSSTGSDQVLTQFLREIGSNMTMLK
APETAELLAGEIGRTLFGFLMRADTEVVDLDAPLASVGIDSLISIELRNWIRRKIGVEVTVLEIVRADSVRDLGVLAQKKLAEKYEARM

pspB

Nucleotide

ATGACTCTGGCACAACCCCAGCCCTACGCTTCCCCGGTCATCAGGCAATATGGGGTTTTGCTAGCAGTGACCAGCCTTATTCTGATTAGCGTCA
TGATTGATATTGAACACCGCGCCAAAATGACCAACGACGATCCCTATGCTCAAATTTTGAGACAGCGGATGGCGGGAGCGCTCGCGGTCTATCA
TCTTGCACCTACCTTTAGAGCCATTGGACGGTTGGGTGTTCATTTCCGATCTCCATTTTGGGTTGGTTTGGAACCGAGTTCAAAGGAGGTTCGC
GGAAGGACTGTGGAAAAAGGATTATCCAAAGCTATACTTTCGTCTCCCTTGGTGCATTTGATGTTGCATGGAGTGTGTTGTGCGCTGCTGGTTG
ACATGGCCCTCGGAAGAAGCTCTTTGTATTAAAGGAAGTCAAATCGCAAAGTGCTGCTGTGGTTTACTCTGGAAACTCCAAGGATTATTGATAG
ACTAAAGAAGCAGCTTACAGGAGTGAATTGAGCACACAAATATCCTATATTGCCAAACATGGGAGTACCCATCACGGAAACCTTTCCAACAATC
TTAGTCCTCATCCACCTCGCATCTCCTCAACATAGCCATTTCAATTGCTAGTCCAGGACCAAATGATGTCGCGACAACATACTCCCTCTTTCCC
AACGACCGAAGCCGATCAAGAACAATCAGGACCGTCGCACTCGATGAATTACCCCGAGTACGGTATATTTCCCGACTTGCCTGCAGCTGATCTT
CGGTCAATCCCAGAACCTGTTTTGCGCCCTCTATAATAGCTTCTCCGCCAGGGTGGAGGGCCCAGTCAAAATCACTGACGCCTAGCGACTTTGC
TACTTCCCCTACTCCCTCTCCGGAGCTAGATTGAATCTGGCTCTGATAAGAGGGAAGAAGTTTCTCGAACATTGGACCGATCGCGTGCTTAGTG
TATTGCGGAACATCACGAGTGAGAACAGTCCGGTAACCTACGCAGGAGCGAATGGTTAGTGCATATCCCGATCTCAGCAAAGAATCAACGAACC
ATCAACATCTGCAAAAAAGGCCATATGCTCAACCGTATCCGGGATCAAGGAATTGCCCCACTCGAGAAGTTGAAACAGAGGTGTGATCTCCGTT
TCAGCCATGGCATACTCGTTACAAAGAACGAATGCAGCGGCAGCATCTGAGAATAAAGCCCCTGCAATGCTAATGTTTTCCGCATTGGGAGCCT
TCTCCGCAAATGCCAGATCATGGCGGACATTGGGAGTACACAGCTCACATGCGAAGGCGAGGATGCGGACCGGCTTCCGTCGCATACTAGCACC
GCAGGCTATCTGTGCTGCAGCTCGCATAATAGCTAGTCCACCTGCACAACCAACACCATGGAGAAGCATACGATCAACATTGGCTGATAAACCA
AGTTTACGATTCACAAGTAGGTCGAATCCCGGGTTACCTTGATTGGTGCAAGTGACCCCGATAGTGTGGGTGACATGTTGAGGACTAATCTGAG
ACTCCCGCAGAGCTTTCTTACAAGCTTGAGTGGTGAGGTCGACTCCGGCTTGATGATAGAATTCAGCAAGCTCACTGATCGTTGGCGCCTCGGG
CCGAGTTGCAAATCCGCTTTCGTAGGACCGAATAGCAGATCTTGTCTCGATGCCAGAGGAGCGGTTGATTTGCAAGAGCTTTTTGAGTCTACAT
AAACTGAATTAGGAGATTCATGCTGTGAAGACAGTTAAAGAAGCGATATCGCATACCCTGGACTCTCTACATCGTAGAAACGTGCCGCGAATTC
CTCGAGCTTTTCTGGTCCCAAAAGATATGGCGGGTATTGGGACCCCAGTCCAGTGATGTACAGGCCTGGCGGTGGTGAGTGGTTGAGCACCATA
GT

Amino acid

TMVLNHSPPPGLYITGLGSQYPPYLLGPEKLEEFAARFYDVESPGLKKLLQINRSSGIETRSAIRSYESGFATRPEAPTISELAEFYHQAGVDL
TTQACKKALRESQISPQHVTHTIGVTCTNQGNPGFDLLVNRKLGLSANVDRMLLHGVGCAGGLAIMRAAAQIACGASMRRKPVRILAFACELCT
PNVRHDLAFAEKAPNAENISIAGALFSDAAAAFVLCNEYAMAETEITPLFQLLEWGNSLIPDTVEHMAFFADVDGYRTVLTRDVPQYTKHAIGP
MFEKLLPSYQSQIQSSSGEGVGEVAKSLGVSDFDWALHPGGEAIIEGAKQVLGLTEDQLQASREIYRTRGNSSSATVLIVLDRLRSLGKREYVV
ATSFGPGLAIEMAMLRRCEVDED

pspC

Nucleotide

ATGGCTTATACTTTAACGTCGATAGTTCTGGGGGTTGTGGTATTTGGCCTCACACGAATTATTTGGGGGTCCATATCCTGGAATTTGCGGTATC
GGAGATTCCACAGCAGTGATGGATGCCAACAGCTTCGTTCAGCACCGTCAAAAGATCCCATATTGGGTTTGGATCACTTCTTCCGTCTGGGCAA
AGCAGCCCACAACAGCAGATATCTCGAAGCATTTCAAGAATGGTTTCAAGCAGTCGGATCCACCTTTGGCGTCAATCTGATGGGTGACTATGTG
ATCTTCACCAACGAACCAAAAAATGTGCAGGCTGTGTTGGTCACAAAATTCAAAGATTTTGAGATCGGCCAGCGGCGTCGCGACAACTCTGCTG
AGCTGTTGGGAATTGGCGTCTTTAATGCGGACGGCCAGACATGGGAGCATGGTCGAGCCTTGGTACGGCCGAACTTTACGCGGAAACAGGTCGC
CGACTTGGGTCTCTTCGAAAAGCATGTCCAGCGTCTCTTTGAGGCGCTTCCGAAAGATGGGACCGCGGTGGATATTCAAGAATGGGTATTTCGA
TTTGTGTGTCCATTGGCCTTTCCTCGTATACATCATCCCCACTCAGCTAACTCAATTTGCTCTTGTCTAGACTTTGGACACCGGAACAGACGTT
TTGTTCAGTGAGTCGAGCGATGTGCTCTTGCCCAGCGCCACCGAGGTAGCTCGCAAGTTTGCATGGGCATTTAACCGTGGCATTGACGGGATTG
CACAACGCATTCGATTGGGCCGGTTCGCCCGATTCTACTATGATCCGCAGTATACAGCTGCCTGCAAATTTGTCCACGACTATGTTGATGAGAT
TGTTGCCAAGGCAGTTTATCGTGCTAAAGAGTGGCATGCGGAAAAGAAAGACAAACCAGTACCACCAGATGACGCAGAGGAGGAACGGTATACC
TTTTTAAATGCCCTTGCTCGAGAAGGAGTGGAGCCCAAGCAGATTCGGGACCAGATATTAAATATCTGTAAGTGCCATAATCGCTTCCTAGCAA
TTCGTGTAGAGCCGGCTGACAACTCGCCCTCAATTCTCAGTGGTTGCTGCCCGCGATACATCCGCCTGCCTCATGAGTGCGGCGGTATTTGAGC
TTGCCCGCAGACCCGAGTACCAAGTACAACTTCGACGGGAAATCGCCGAGAAATTAAGTGGACGCCAACCCACATTCGAAGATCTTAAAGATCT
CACCTTCCTGAATCATTTTGTCAAAGAAACTCTACGCATGTACCCCCCAGTGCCGCTCAACGCACGGGTAGCGAAGAATGACACAGTTCTACCC
CGTGGGGGTGGCTCGGATGGCATGGCCCCAATTTTTGTGCCCCGAGGGCAGCTCGTCGTGTATCAGGTTTATTCCATGCACAGGCGAGAAGATC
TGTGGGGTTCGGATGCGCACATATTCCGCCCTCAGCGATGGGAAACGACACGGCCTACTTTTGAGTATCTGCCATTCAATGCAGGACCGCGCAT
TTGCCCTGGTACGTTATGGTCATGTCTAACCTACTTTTTACAAAATTATTGTGTCATCTTCATGGTTCAAATACACAGCTAATAATTCTGTGAT
TCTGTTCTAGGCCAGCAATTCGCCTTGGTAGAGACAAGCTACGTCTTGGTCCGTTTGCTCCAGGAATATTCGAACATTGAGGCCCGTGGCAACA
GTGCACCGTGGAGGGAGCATCTGACCCTTACTTGCTCGGTTGGTCAAGGTGTGTGGGTGAGTCTGGTCAAATGAAGATCTTTGATGGTGAAATT
TGTTGCAAAGAGCGTACGTAAGCTCGCAGTTTTCTTTTTGTTTTCTAGATAACATACTATACTCCATAGGCCTGGAAATATTTCAGTTTTTCCA
TCATCGTCATCATCATTAGAAGCAGCATTTAA

Amino acid

LRSAPSKDPILGLDHFFRLGKAAHNSRYLEAFQEWFQAVGSTFGVNLMGDYVIFTNEPKNVQAVLVTKFKDFEIGQRRRDNSAELLGIGVFNAD
GQTWEHGRALVRPNFTRKQVADLGLFEKHVQRLFEALPKDGTAVDIQEWTLDTGTDVLFSESSDVLLPSATEVARKFAWAFNRGIDGIAQRIRL
GRFARFYYDPQYTAACKFVHDYVDEIVAKAVYRAKEWHAEKKDKPVPPDDAEEERYTFLNALAREGVEPKQIRDQILNILVAARDTSACLMSAA
VFELARRPEYQVQLRREIAEKLSGRQPTFEDLKDLTFLNHFVKETLRMYPPVPLNARVAKNDTVLPRGGGSDGMAPIFVPRGQLVVYQVYSMHR
REDLWGSDAHIFRPQRWETTRPTFEYLPFNAGPRICPGQQFALVETSYVLVRLLQEYSNIEARGNSAPWREHLTLTCSVGQGVWVSLVK

Table S4. List of primers used for cloning of *pspA*, *pspB*, and *pspC*

Primer name	DNA sequence 5' to 3'
pspA-IFpUKpnI-FW	CCGGAATTCGAGCTCGAATATGCTTGCACAAGATGTCG
*pspA*_IFpUKpnI-RV	ACTACAGATCCCCGGCGATATCTGTAGGGGACAGC
*pspA*_F1	GCAAGACCATCGCCACTGTGG
*pspA*_R1	CCACAGTGGCGATGGTCTTGC
pspB-IFpUKpnI-FW	CCGGAATTCGAGCTCGAAAATGGCCCCCATGGCCCTTC
pspB-IFpUKpnI-RV	ACTACAGATCCCCGGTCCATGACTCTGGCACAACC
*pspC*_IFpNotI-FW	TTTGAGCTAGCGGCCGAAATGGCTTATACTTTAACG
*pspC*_IFpUNotI-RV	GTCACTAGTGCGGCCCACCCCTGTCTGGCGAATGG

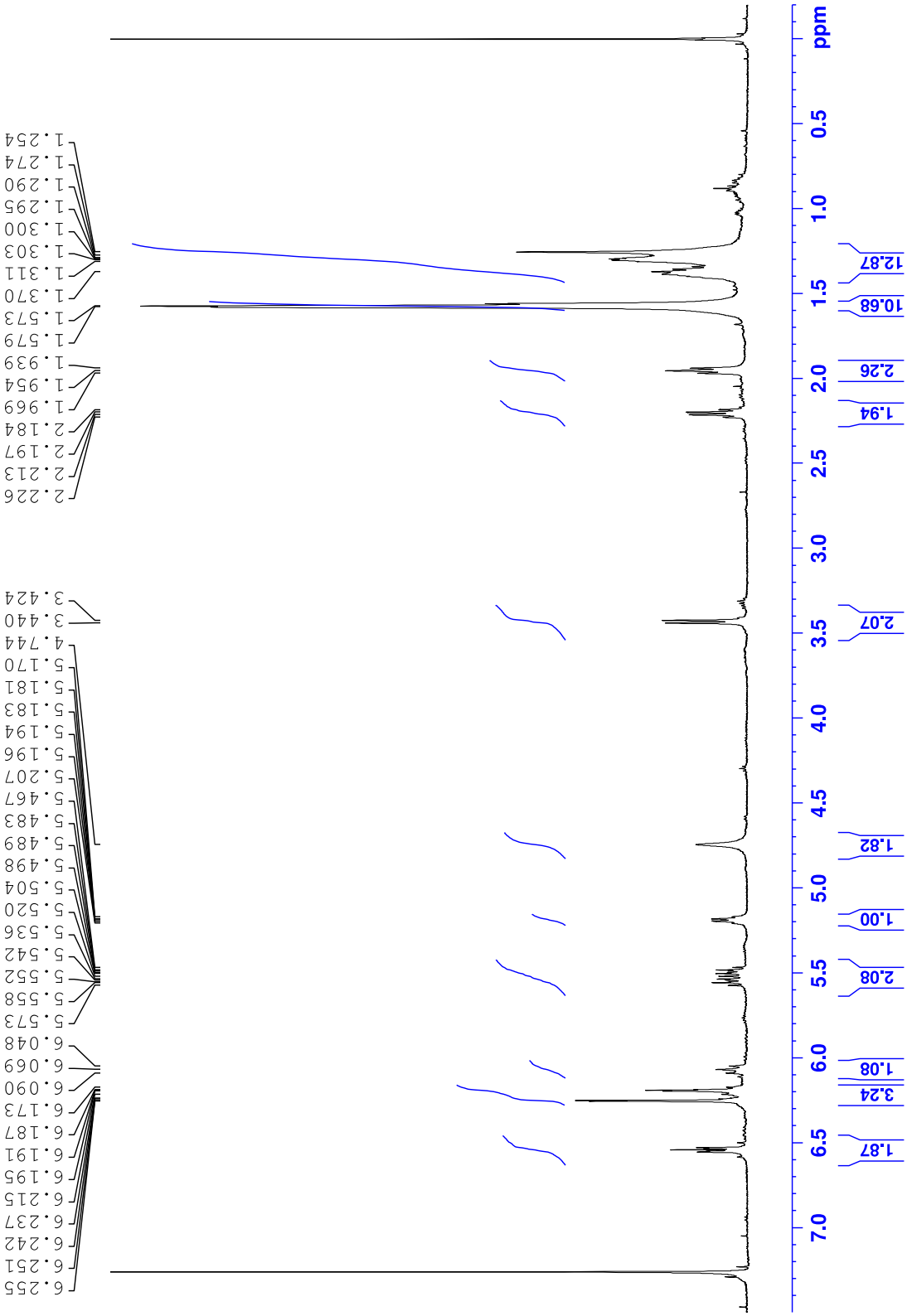
Table S5. Summary of construction of expression plasmids for *A. oryzae* transformation.

Plasmid name	Original vector	Gene 1 (KpnI site)	Gene 1 (NotI site)
pUARA2-*pspA*	pUARA2	*pspA*	
pUAdeA2-*pspB*	pUAdeA2	*pspB*	
pUARA2-*pspABC*	pUARA2	*pspA*	*pspB, pspC*

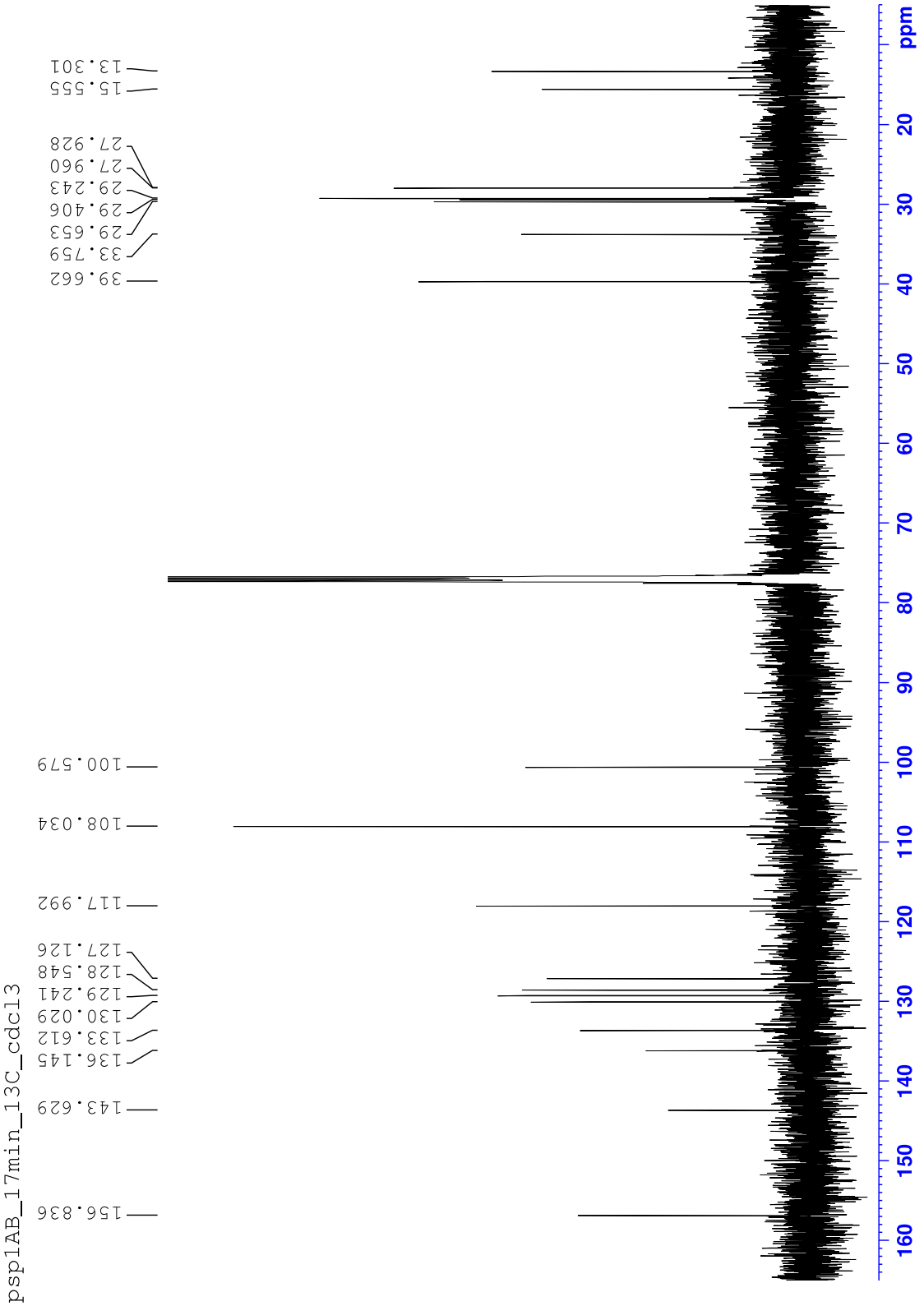
Table S6. Summary of *A. oryzae* transformant in this study.

Transformant name	Plasmid 1	Plasmid 2
AO-*pspA*	pUARA2-*pspA*	
AO-*pspB*	pUAdeA2-*pspB*	
AO-*pspAB*	pUARA2-*pspA*	pUAdeA2-*pspB*
AO-*pspABC*	pUARA2-*pspABC*	

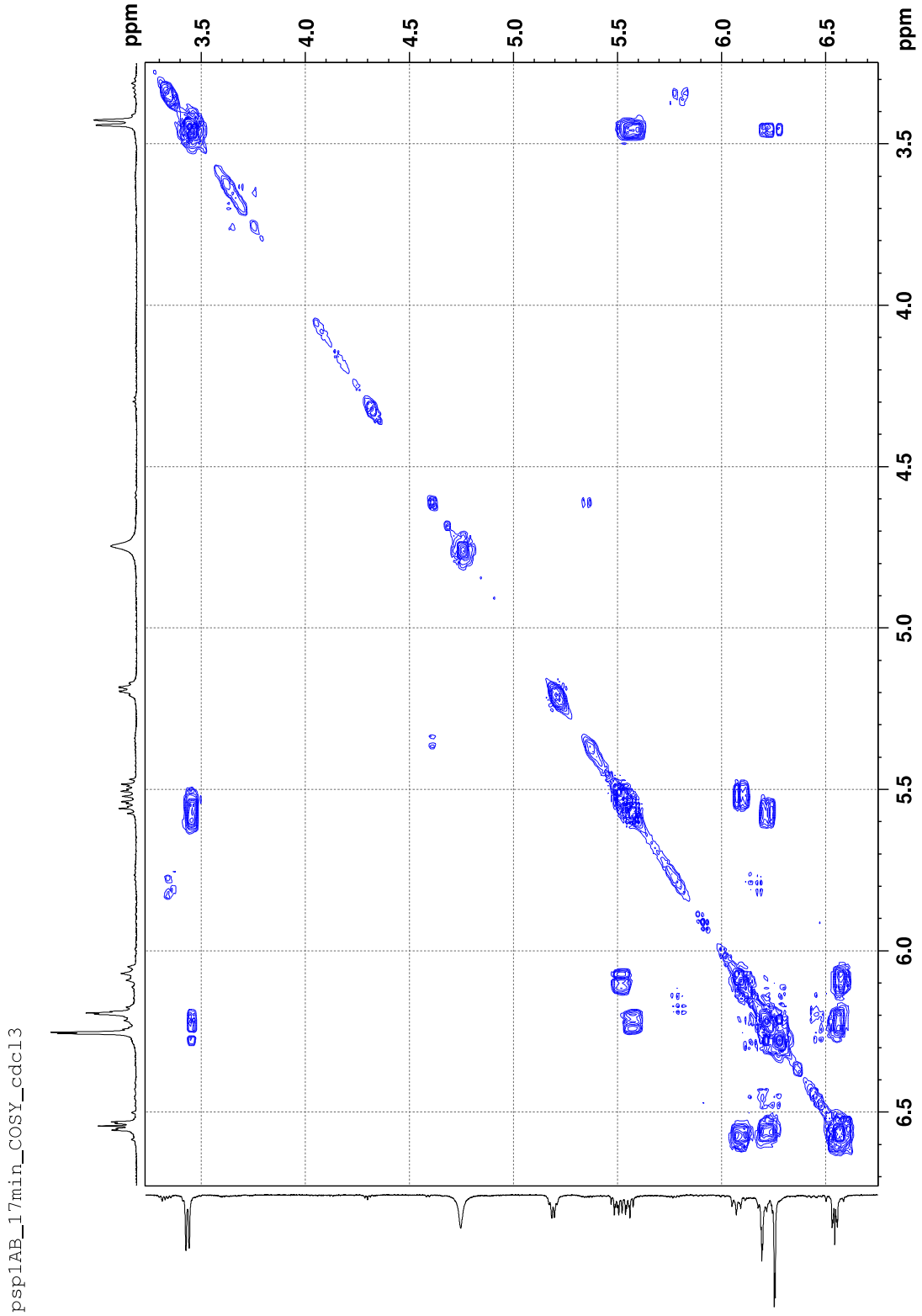
^{1}H NMR spectrum of soppiline B (2)



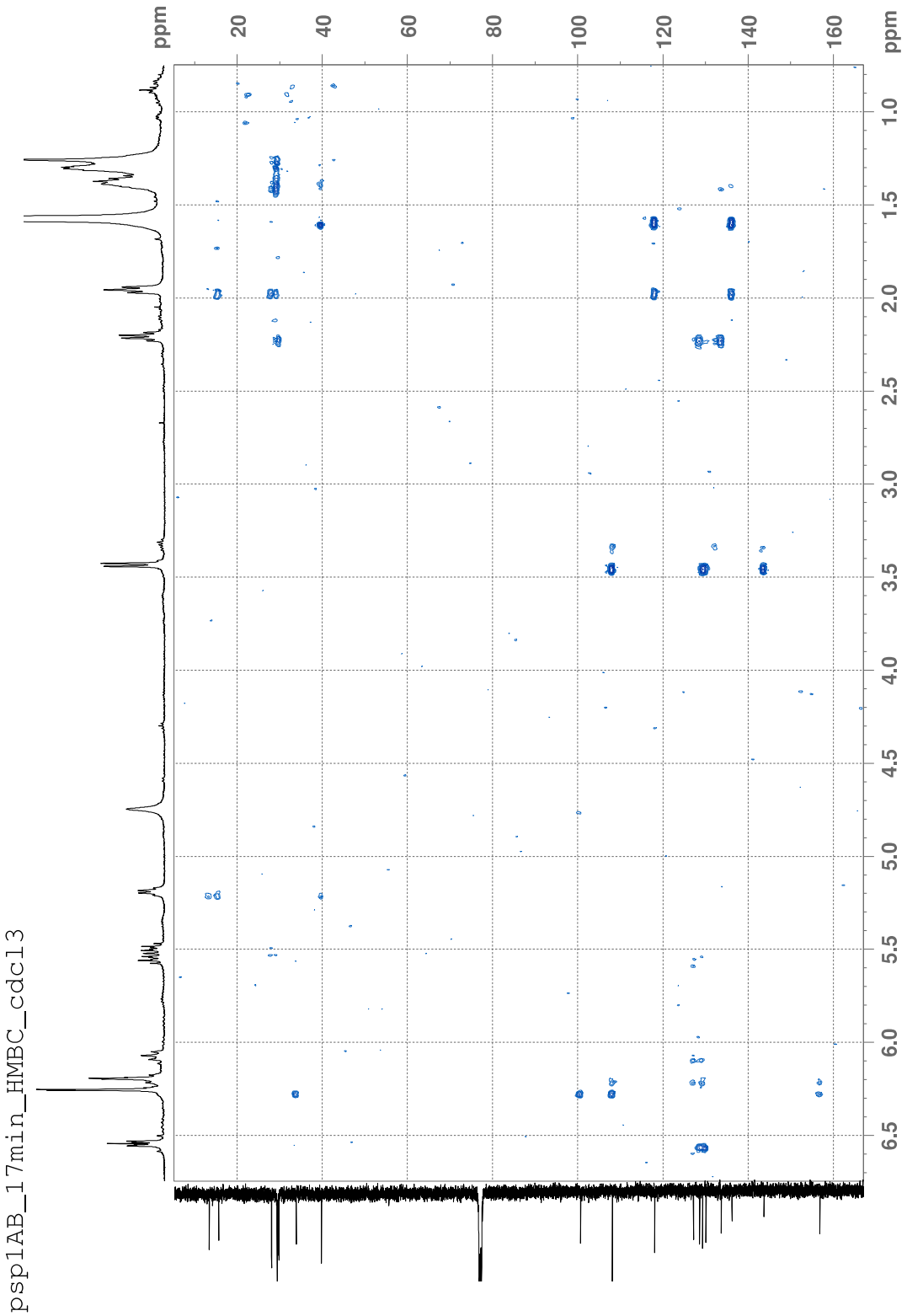
^{13}C NMR spectrum of soppiline B (2)



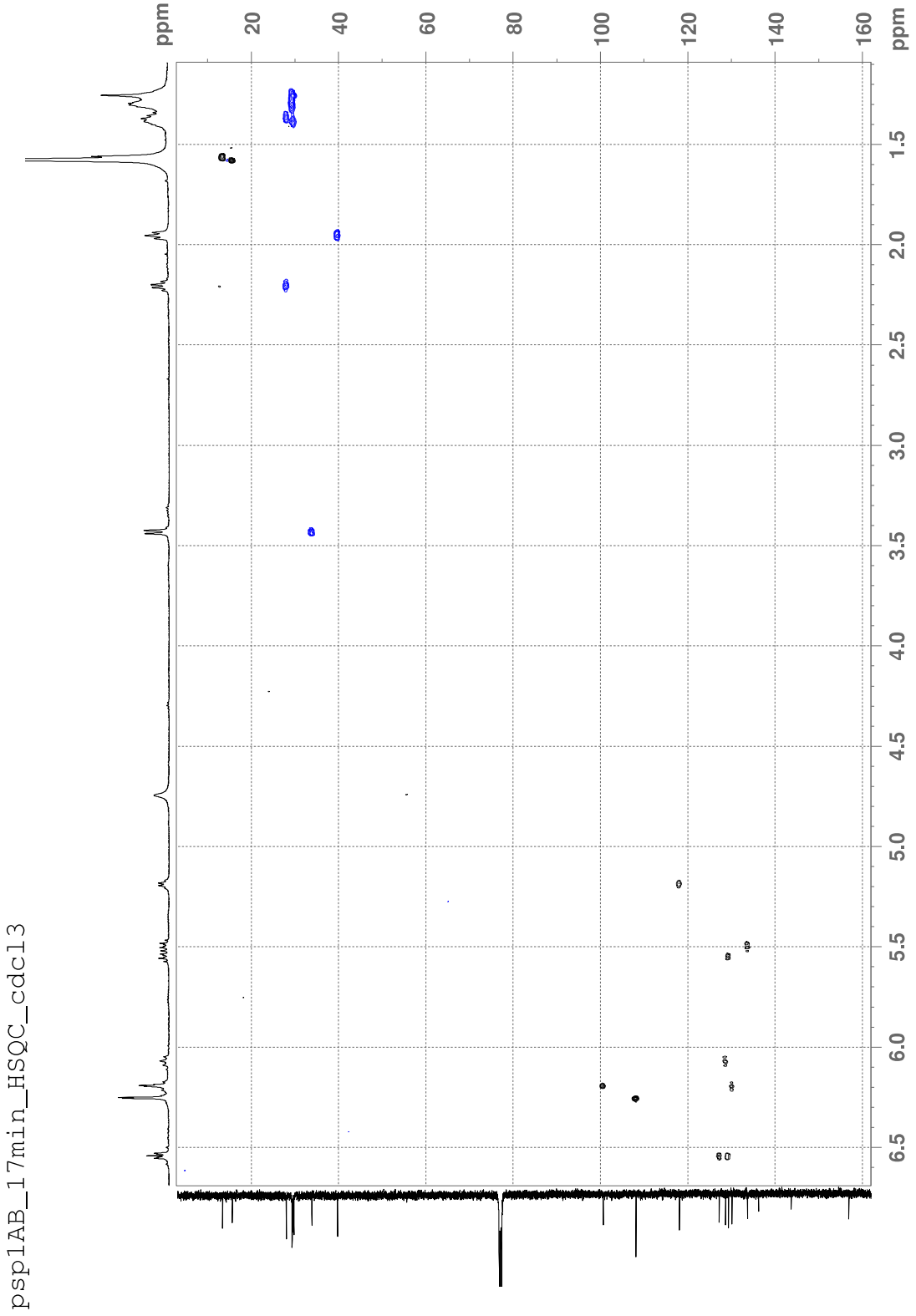
^{1}H-^{1}H COSY spectrum of soppiline B (2)



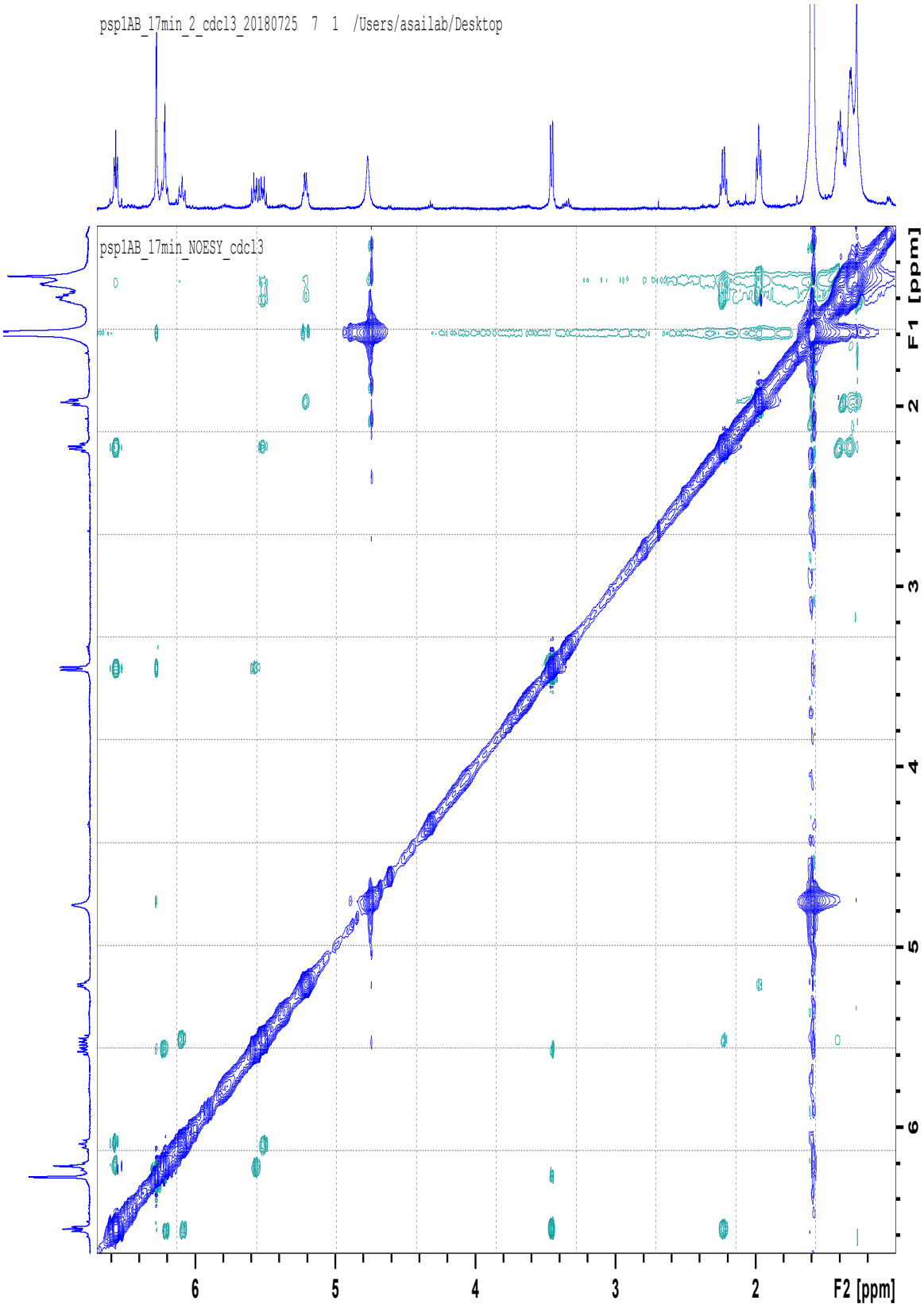
HMBC spectrum of soppiline B (2)



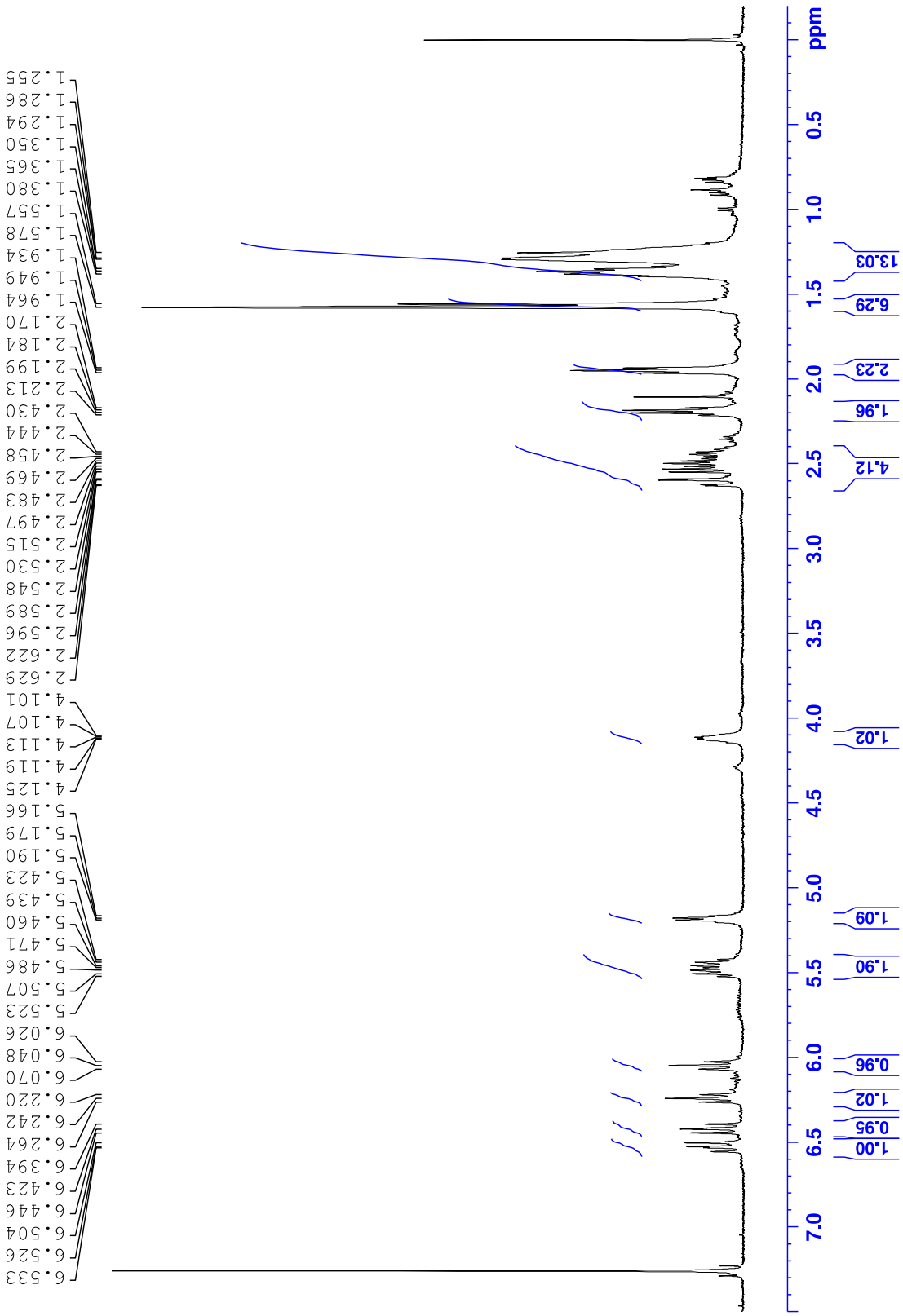
HSQC spectrum of soppiline B (2)



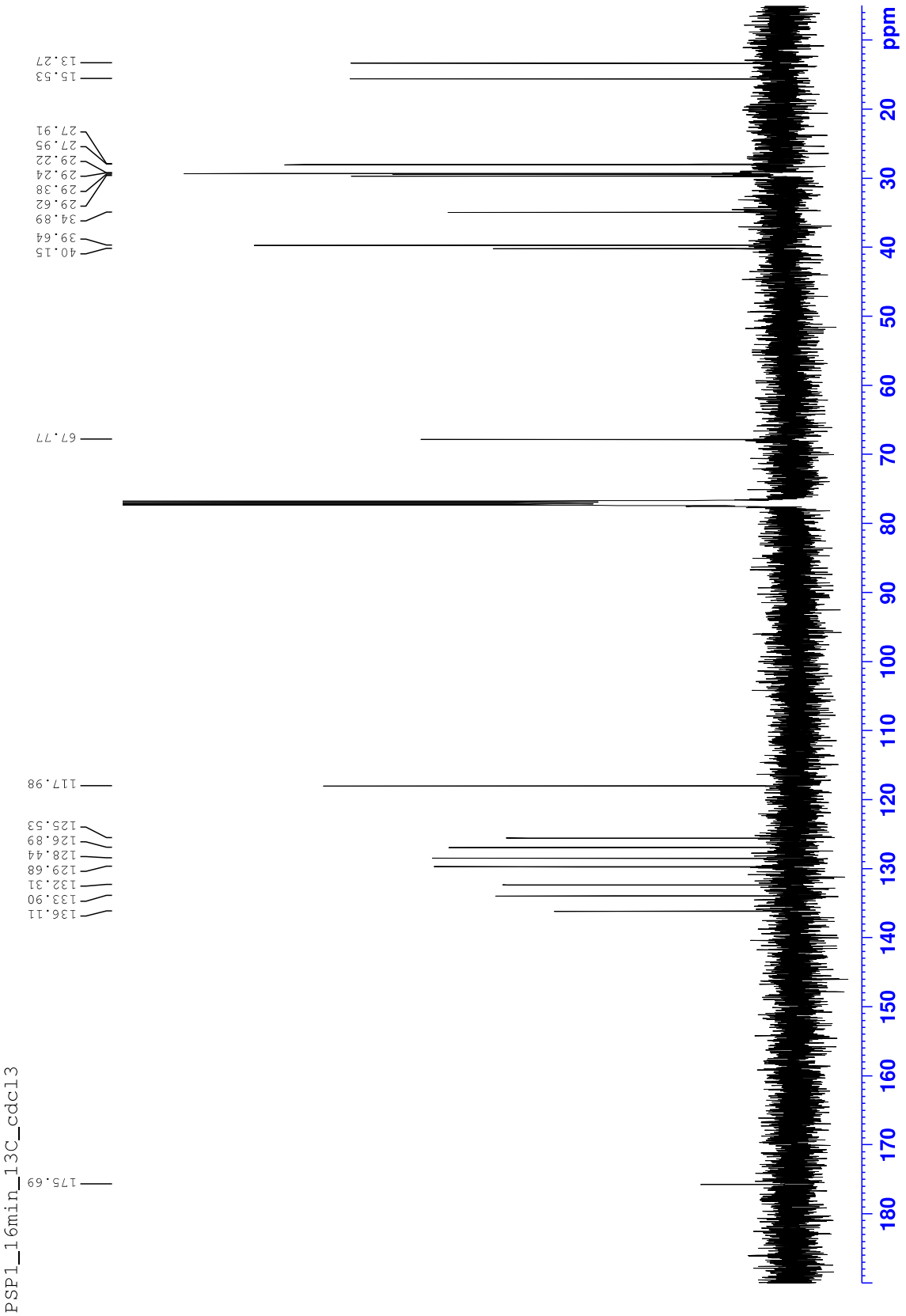
NOESY spectrum of soppiline B (2)



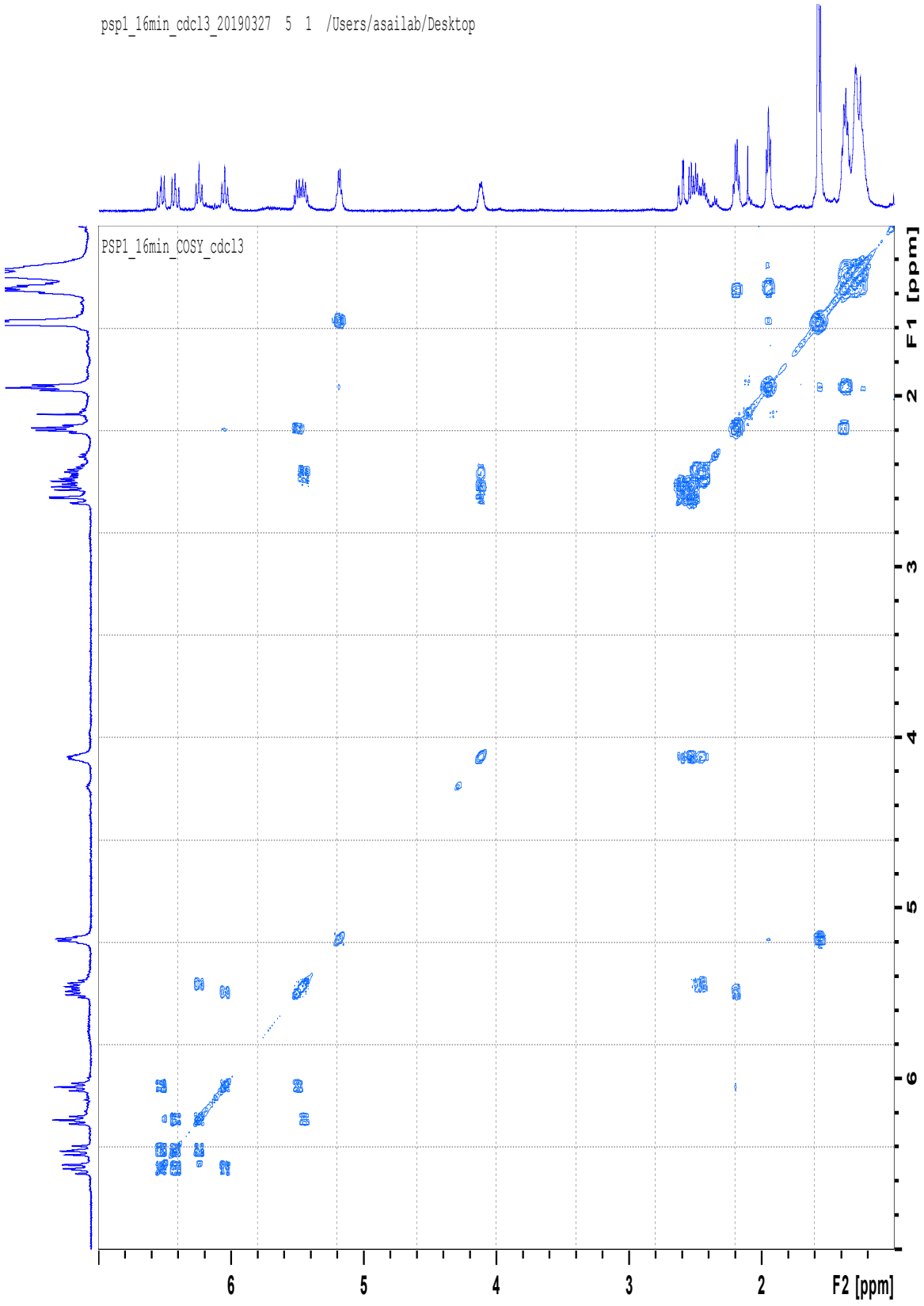
^{1}H NMR spectrum of soppiline A (1)



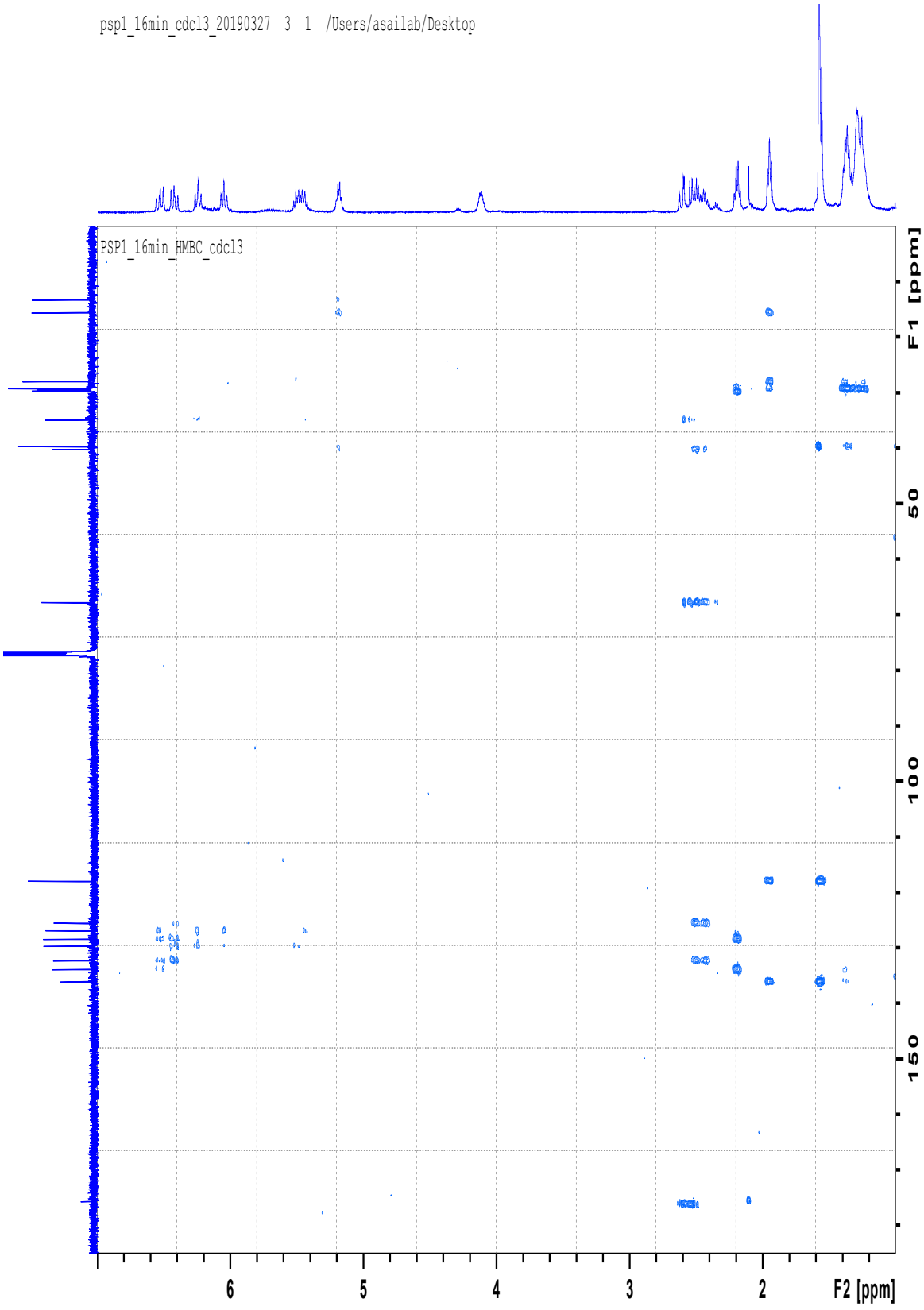
^{13}C NMR spectrum of soppiline A (1)



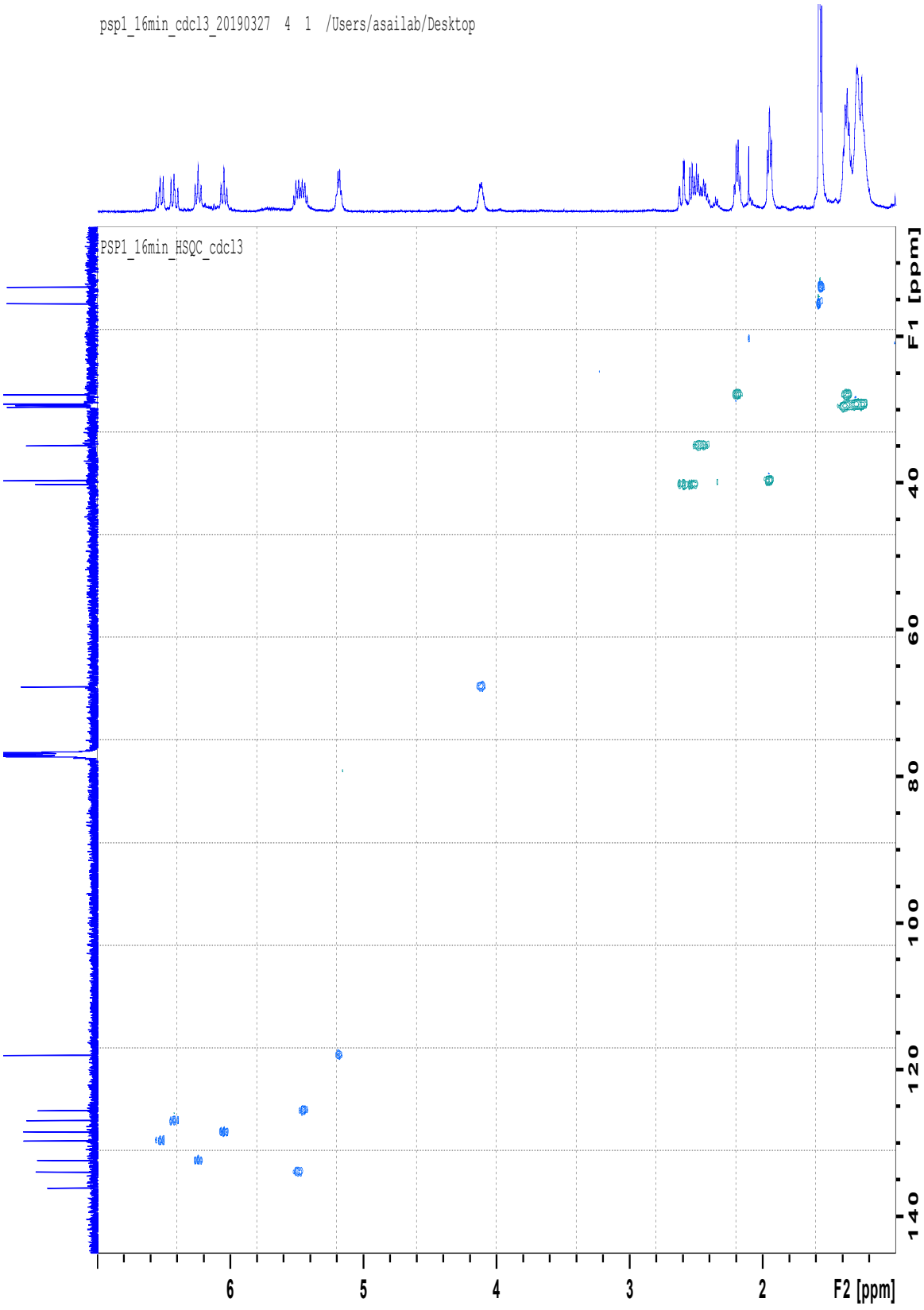
^{1}H-^{1}H COSY spectrum of soppiline A (1)



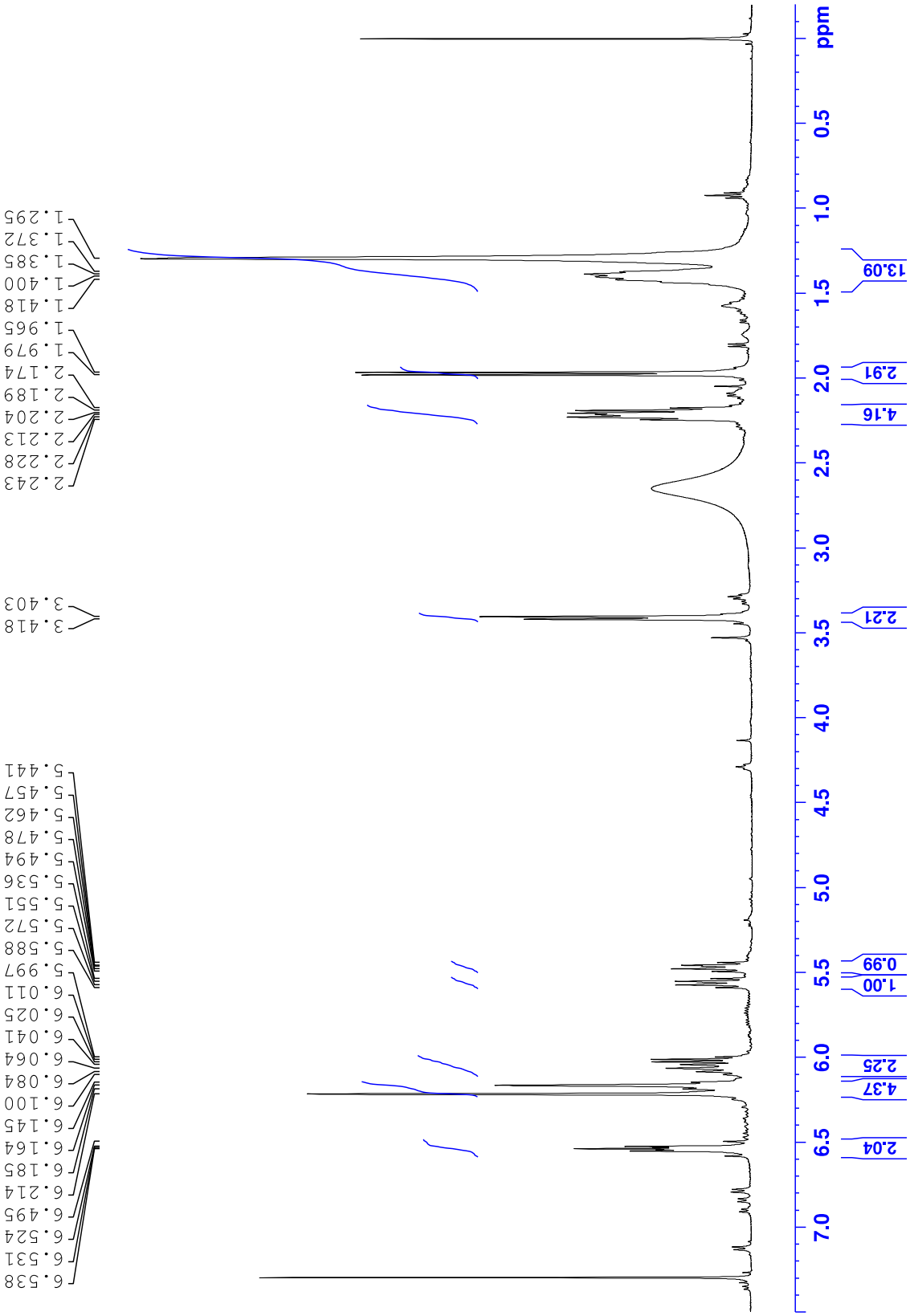
HMBC spectrum of soppiline A (1)



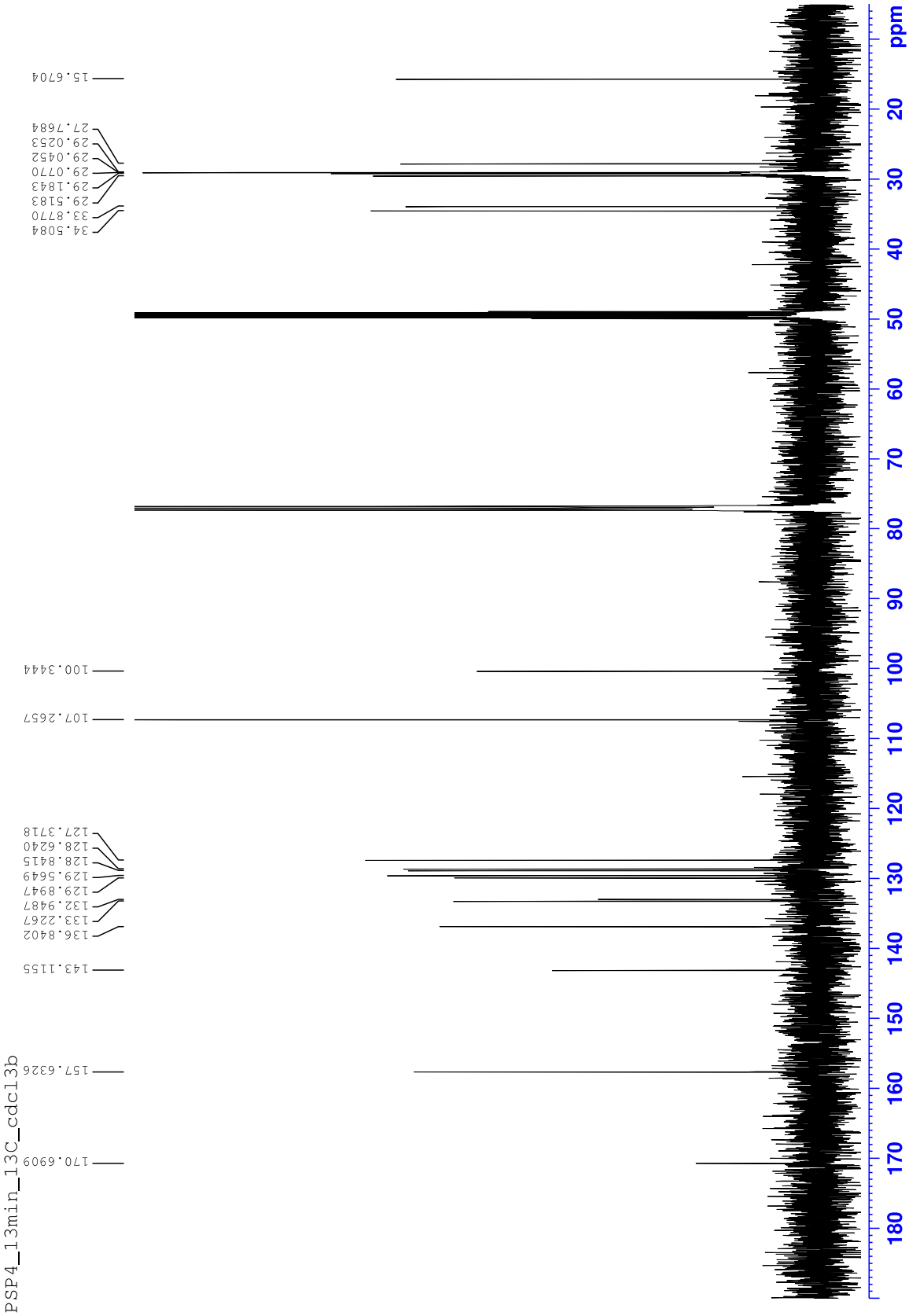
HSQC spectrum of soppiline A (1)



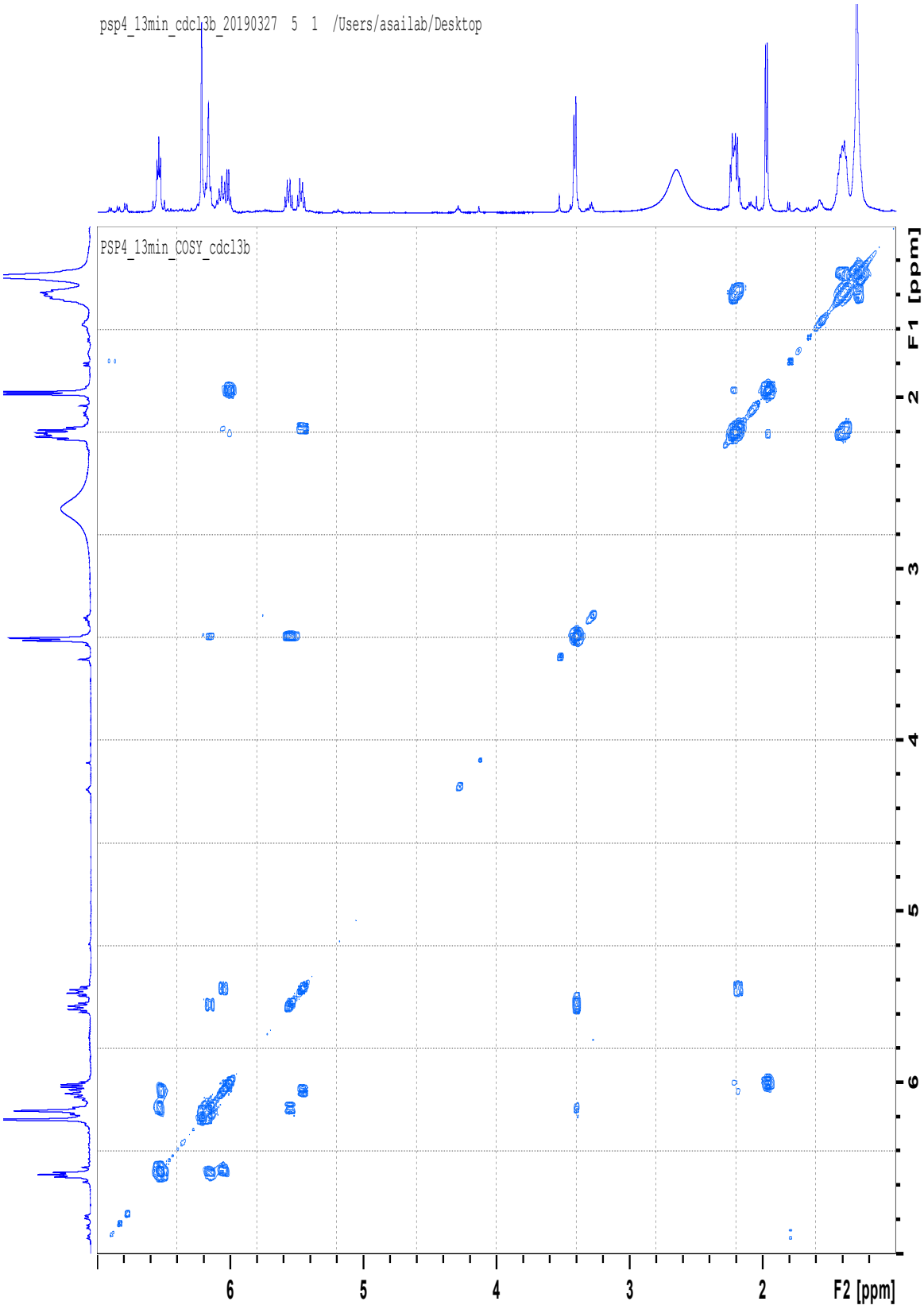
^{1}H NMR spectrum of soppiline C (3)



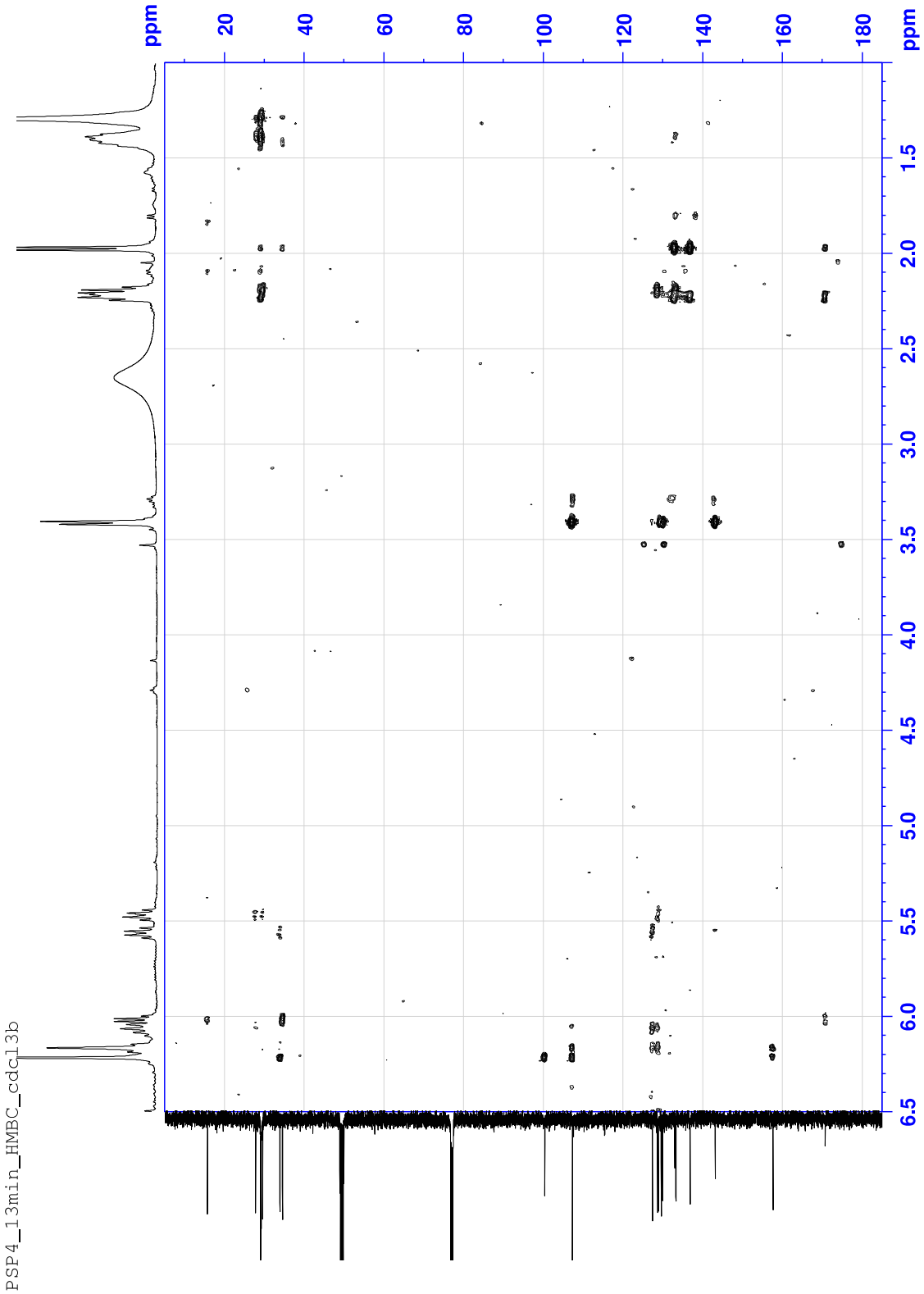
^{13}C NMR spectrum of soppiline C (3)



^{1}H-^{1}H COSY spectrum of soppiline C (3)



HMBC spectrum of soppiline C (3)



HSQC spectrum of soppiline C (3)