2015

An NMR-guided approach to the isolation of secondary metabolites from NRPS and PKS gene clusters

Connor Craig
University of Richmond

An NMR-Guided Approach to the Isolation of Secondary Metabolites from NRPS and PKS Gene Clusters

by

Connor P. Craig

Honors Thesis

in

Program in Biochemistry and Molecular Biology
University of Richmond
Richmond, VA

Submitted April 24, 2015

Advisor: Dr. Jonathan Dattelbaum

This thesis has been accepted as part of the honors requirements in the Program in Biochemistry and Molecular Biology.

Janet Patterson — *(advisor signature)*

4/25/15 — *(date)*

Julie A Pollock — *(reader signature)*

04.24.15 — *(date)*

Abstract: Secondary metabolites continue to offer a promising source of novel chemical entities with therapeutic bioactivities. Our research aims to isolate secondary metabolites produced by non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) constructs. A metagenomic library was created from the marine bacterium *Pseudoalteromonas sp.*, which was isolated from the Chesapeake Bay sponge *Halichondria bowerbanki*. Using a functional screen for PPTase activity we were able to identify colonies that contained NRPS and PKS gene clusters. Cultures of this microbe have been produced and analyzed using a combination of chromatography and ^{1}H, ^{13}C, HMBC, HSQC, COSY, AND NOESY NMR spectra. We were able to isolate a potential partial product from one of these gene clusters. Further investigation will be needed to confirm the identity of this compound, and perform biological assays to screen for relevant therapeutic activity.

INTRODUCTION

Natural products have been used for medicinal purposes since 2600 B.C., when the ancient Mesopotamians documented the use of Cypress oils for the treatment of coughs, colds, and inflammation.[1] A proverbial 'Golden Age of Antibiotics' occurred beginning with the discovery of Penicillin in 1929 and continuing into the 1960s. During this time more than one half of the antibiotics used in the healthcare industry today were discovered.[2] This remarkable period of discovery began to dwindle as researchers encountered considerable difficulties due to replication (the isolation of the same compounds over and over). In an effort to combat replication and push natural product discovery, major pharmaceutical companies developed natural product discovery programs, which screened for bioactive compounds that could be developed into marketable pharmaceuticals.[3] Many pharmaceutical companies, however, ended their natural product discovery programs in the 1990s and early 2000s as the rise of high throughput screening led to a preference towards combinatorial chemistry for future drug

[1] Cragg, G.M.; Newman, D.J. Biodiversity: A Continuing Source of Novel Drug Leads. *Pure Appl. Chem*. 2005; 77: 7-24

[2] Davies, J. Where have all the antibiotics gone? *Can J Infect Dis Med Microbiol.* 2006; 17(5):287-290.

[3] Baker, D.D.; Chu, M.; Oza, U.; Rajgarhia, V. The value of natural products to future pharmaceutical discovery. *Nat Prod Rep.* 2007; 24(6):1225-44.

research.[4] Although natural product research by many pharmaceutical companies has declined in the last several decades, of the 1,184 new chemical entities studied from the period 1981-2006, 52% were natural products or derived from natural products, with only 32% being purely synthetic molecules.

Therefore, natural product isolation and characterization may still promise novel compounds with bioactive and therapeutic properties. These molecules are defined as secondary metabolites and differ from primary metabolites in that they are not a part of primary metabolism (i.e. the breakdown of proteins, fats, nucleic acids, and carbohydrates needed for survival).[5] Rather, natural products are essential to providing an organism with some sort of selective advantage in its natural environment and therefore often have important properties such as antibiotic, antitumor, or anti-inflammatory activity.[6] Indeed, some of the most famous drug discoveries of the last century have come from a variety of secondary metabolites such as: the antibiotics penicillin and vancomycin from the fungi *Penicillium notatum* and *Amycolatopsis orientalis* respectively; paclitaxel (Taxol®) from the bark of *Taxus brevifolia*; and the antiviral compound prostratin from *Homalanthus natuans,* which has shown promise in

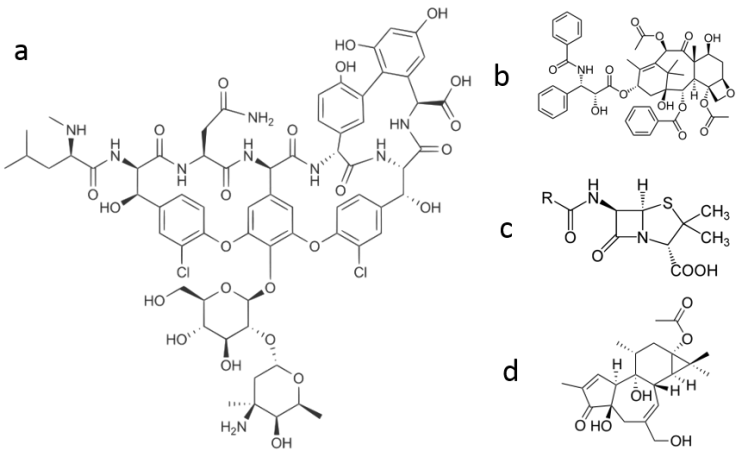


Figure 1. Notable secondary metabolites isolated and manufactured for pharmaceutical purposes. a) vancomycin b) paclitaxel c) penicillin d) prostratin.

[4] Nussbaum, F.V.; Brands, M.; Hinzen, B.; Weigand, S.; Habich, D. Antibacterial natural products in medicinal chemistry – exodus or revival? *Angew. Chem Int. Ed.* 2006; 45: 5072-5129

[5] Dweick, P.M. A Historical Overview of Natural Products in Drug Discovery. *Medicinal Natural Products: A Biosynthentic Approach*, 2nd ed.; John Wiley and Son: West Sussex, UK, 2002; p. 520

[6] Maplestone, R.A.; Stone, M.J.; Williams D.H. The evolutionary role of secondary metabolites—A review. *Gene* 1992, 115, 151-157

combatting HIV.[7] The structures in Figure 1 demonstrate the highly conjugated, aromatic, and stereochemically diverse nature of many natural products on the market today.

Approximately 99% of all microbes are not able to be cultured given their complex requirements for growth, which include specific pH, temperature, and nutrient conditions.[8] This makes them a generally underutilized source for isolating chemical entities. Of particular interest are microbes with non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) gene clusters. These large multifunctional enzyme systems are used to synthesize complex secondary metabolites using simple building blocks such as carboxylic acids and amino acids.[9] These enzyme complexes consist of multiple domains constructed in an assembly line-like fashion, where a substrate is shuttled from one domain to the next during synthesis and modification.

This multimodular construction makes possible the synthesis of highly complex and stereochemically diverse molecules. Each module contains a catalytic domain with one of several specified

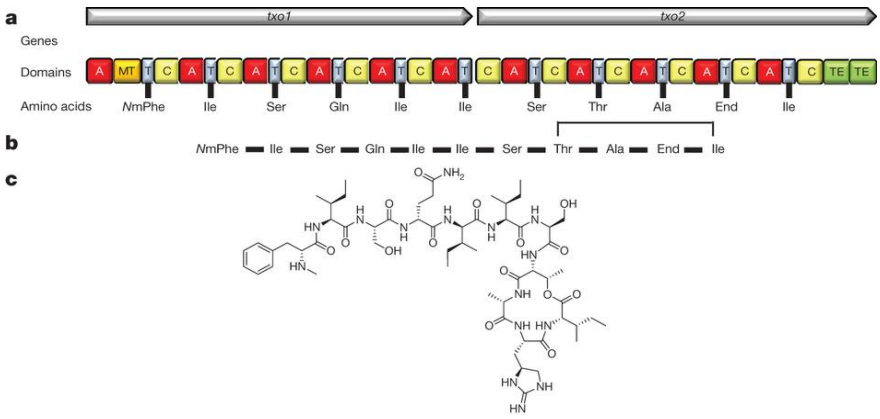


Figure 2. The predicted NRPS gene cluster responsible for the production of the newly discovered antibiotic teixobactin. a) the two genes, *txo1* and *txo2*, their associated enzymatic domains, and the amino acids incorporated by each. b) schematic structure of the teixobactin molecule. c) structure of teixobactin.

[7] Dias, D.A.; Urban, S.; Roessner, U. A Historical Overview of Natural Products in Drug Discovery. *Metabolites* 2012, 2, 303-336; doi: 10.3390/metabo2020303

[8] Lewis, K. Platforms for antibiotic discovery. *Nature Rev. Drug Discov.* 2013; 12(5):371-87. doi: 10.1038/nrd3975.

[9] Wenzel, S.C.; Muller, R. Formation of novel secondary metabolites by bacterial multimodular assembly lines: deviations from textbook biosynthetic logic. *Current Opinion in Chemical Biology* 2005, 9:447 458

functions. Figure 2[10] shows the NRPS construct responsible for the synthesis of the antibiotic teixobactin. In section 'a', one can observe the individual domains responsible for the chemical modifications required to produce the chemical entity of interest. These domains include: adenylation, methylation, thiolation, condensation, and thioesterase sites. The order of these domains and the identity of the substrate determine what the final product will be. Due to their highly complex and diverse nature, our research aims to isolate NRPS and PKS gene clusters from environmental samples using metagenomic library construction in conjunction with a novel screening technique.

In order to identify NRPS and PKS gene clusters we employed a novel functional screen of 4'-Phosphopantetheinyl transferase (PPTase) proteins. This class of enzymes is responsible for the conversion of 4-Phosphopantetheine (Ppant) carrier proteins from an inactive to an active state, [11,12] allowing them to covalently bond substrate intermediates such as reactive acyl thioesters, and shuttle them from one reactive subunit to the next in a multidomain enzymatic system.[13] Ppant carrier proteins are often vital components of NRPS-and PKS-associated synthesis complexes and are responsible for transferring the substrate of interest from one domain to the next, in a fashion much akin to an assembly line. For this to be possible, PPTases must also be actively transcribed and translated from their respective genes in NRPS and PKS gene clusters for these Ppant carrier proteins to be active. For these reasons we chose PPTase as the target molecule in our functional screen.

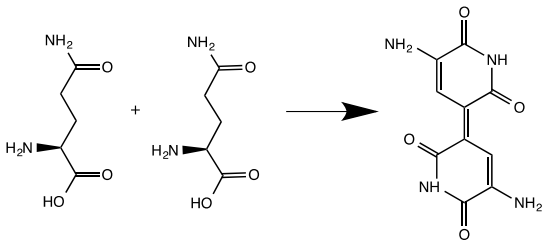


Figure 3. The condensation of 2 L-Glutamine molecules to the blue pigment indigoidine. Colonies producing indigoidine are positive for PPTase activity.

[10] Losee, L.L. A new antibiotic kills pathogens without detectable resistance. *Nature* 2015; 517, 455–459

[11] Lambalot, R.H. et al. A new enzyme superfamily – the phosphopantetheinyl transferases. *Chem. Biol.* 3, 1996; 923-936

[12] Lambalot, R.H. and Walsh, C.T. Cloning, overproduction and characterization of the Escherichia coli Holo-acyl carrier protein synthase. *J. Biol. Chem.* 1995; 270, 24658-24661

[13] Walsh, C.T., Gehring, A.M., Weinreb, P.H., Quadri, L.E. and Flugel, R.S. Post-translational modification of polyketide and nonribosomal peptide synthases. *Curr. Opin. Chem. Biol.* 1997; 1, 309-315

Our functional screen employs the *Escherichia coli* cell line EcoBlue 1. EcoBlue 1 contains a reporter plasmid with a small NRPS taken from *Streptomyces lavendulae*. This NRPS is known as BpsA, and is able to condense two L-glutamine residues into a single blue pigment: indigoidine (Figure 3).[14] If a fosmid with a DNA insert containing a PPTase gene is transformed into an ecoBlue 1 cell, BpsA will be activated, resulting in the production of indigoidine and an easily identifiable blue colony (Figure 4).

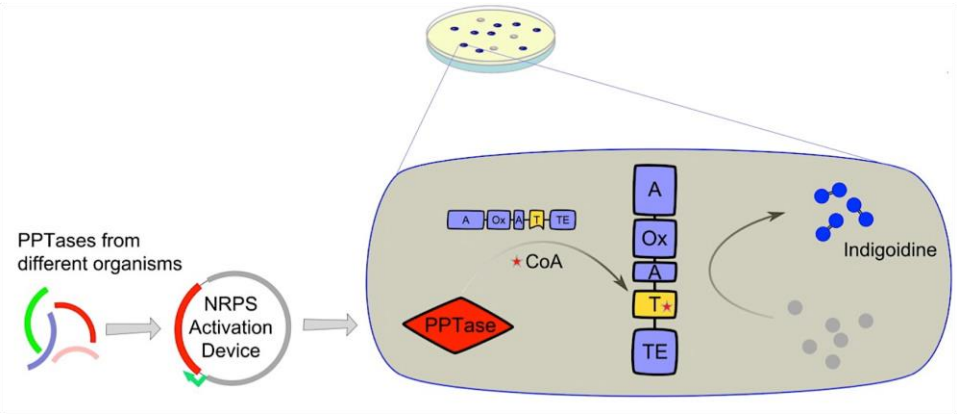


Figure 4. The BpsA reporter plasmid responsible for indigoidine production in ecoBlue 1 cells. Environmental DNA is sheered an inserted into a fosmid. The fosmid is transformed into the ecoBlue cell line. If a PPTase exists on a particular insert, BpsA will be activated and produce indigoidine, resulting in a blue colony.

Previous work in the Dattelbaum lab has generated two metagenomic fosmid libraries. One library was created from environmental DNA isolated from soil collected on the University of Richmond campus. The second was created from a cultivable strain of the marine bacterium *Pseudoalteromonas sp.*, arbitrarily named SW202. SW202 was isolated from the sponge *Halichondria bowerbanki* by Dr. Malcolm Hill of the University of Richmond, and generously provided for use in our lab. Screening of the SW202 fosmid library in our functional assay yielded five positive colonies: designated A, B, C, D, and E (a positive result being indicated by the production of blue pigment).

This paper will detail the work done to isolate novel secondary metabolites from positive colony A. The metabolites in question were isolated using an NMR-guided approach to

[14] Owen, J.G., Copp, J.N, and Ackerley, D.F. Rapid and flexible biochemical assays for evaluating 4'-phosphopantetheinyl transferase activity. *Biochem J.* 2011: 436: 709-717

natural product mining in conjunction with various chromatography techniques, including high-pressure liquid chromatography (HPLC). The results are promising, indicating that SW202A does in fact contain an as yet undetermined gene cluster responsible for the production of unique compounds not found in a control sample.

METHODS

ZYP-5052 auto-induction broth (2L) [10 g yeast extract, 20 g N-Z amines (combined in 1600 mL of ultrapure water and autoclaved), 13.16 g glutamine, 10 mL glycerol, 1.0 g D-glucose, 4 g α–lactose, 14.196 g sodium phosphate dibasic, 6.608 g ammonium sulfate dibasic, and 13.608 g potassium phosphate monobasic (combined in 400 mL ultrapure water and filter sterilized into the autoclaved solution described above)] was prepared to grow a culture from ecoBlue 1 cells of positive colony A (henceforth referred to as ecoSW202A) and non-transformed ecoBlue 1 cells. The two cultures were grown for six days at 16 °C and 200 RPM.

The cultures were centrifuged at 6750 xg to pellet the cells. The supernatant was collected and the cell pellets were stored at -20 °C for later use in methanol extractions. The supernatant underwent a first round of chromatography on a 30 mL Diaion® HP-20 open column. The column was prepared using successive washes of three column volumes of acetone, water, and methanol. The supernatant was then applied to the column and the flow-through discarded. Bound molecules were eluted from the column using washes of 25% and 100% acetone (three column volumes each). The eluted fractions were collected in 250 mL round bottom flasks and stored at 4 °C overnight.

The 100% acetone fractions from the ecoSW202A and ecoBlue 1 cultures were dried using rotary evaporation and transferred to 20 mL scintillation vials. NMR spectroscopy on a Bruker 500 MHz was performed on the two samples in CD_3OD.

Both 100% acetone fractions were then further separated on a 30 mL Diaion® HP-20 column into 50%, 75%, and 100% acetone fractions. ^1H NMR spectroscopy was

performed on the three fractions of each sample in conjunction with 2-D HMBC experiments. The resulting spectra for the corresponding ecoBlue 1 and ecoSW202A fractions were compared and unique peaks were identified to focus further chromatography on fractions with fosmid-associated compounds. The 50% acetone ecoSW202A fraction was believed to have novel peaks.

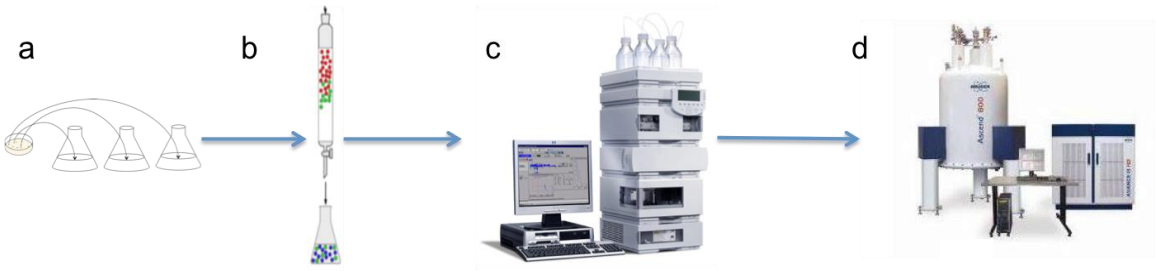


Figure 5. Protocol for isolating novel products from ecoSW202A cell cultures. a) 2 L cultures were grown, and the supernatant collected. b) HP20 open columns were used to fractionate metabolites by polarity. c) HPLC was used to further purify the existing fractions. d) NMR spectroscopy was used to characterize each fraction, and target NRPS and PKS compounds.

The 50% acetone ecoSW202A fraction was subjected to further purification using a C18 semi-prep column on an Agilent 1100 series HPLC system. The ecoSW202A 50% acetone fraction (224 mg) was dissolved in 10 mL of HPLC-grade methanol. The resulting solution was syringe-filtered to remove any insoluble debris. Injections (10 µL) of the solution were added for 20 minute runs with a flow rate of 4.0 mL/min. Runs began with 100% methanol, with consecutive runs increasing the water content by 10% until satisfactory separation was observed. The most effective solvent solution was a 1:1 water:methanol mixture. More

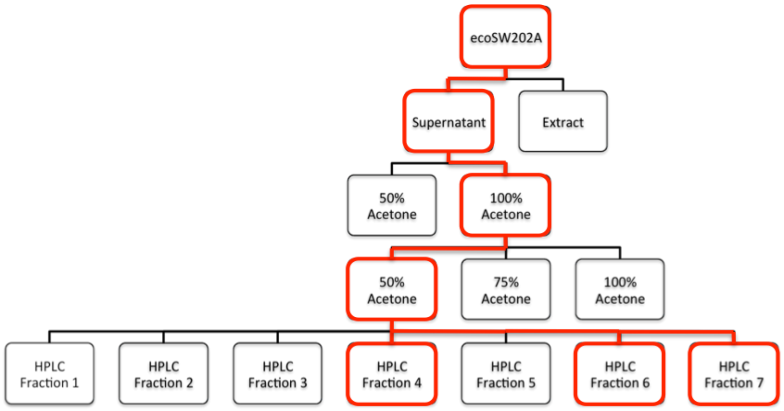


Figure 6. Overall fractionation scheme for a 2L culture of ecoSW202A. The red fractions indicate samples with unique proton peaks when compared to a spectra from the ecoBlue 1 control culture.

polar mixtures led to large changes in the column's properties, likely due to the very hydrophobic nature of the resin used.

Seven unique fractions were eluted for the ecoSW202A extract. The resulting fractions were back-loaded on a 3 mL Diaion® HP20SS column to remove water from the samples and reduce drying times. The column was prepared with successive 9 mL washes of acetone, water, and methanol. The fractions were then applied to the HP20SS column under vacuum, and the flow-through discarded. The bound product was eluted from the column in 9 mL of 100% acetone. The seven fractions were then dried via rotary evaporation.

Each fraction was further analyzed using ^{1}H NMR spectroscopy. The resulting spectra were compared to the original ecoBlue 1 control spectra to identify novel peaks. Fractions were subsequently subjected to further NMR analysis using ^{13}C, 2-D HMBC, HSQC, COSY, and NOESY experiments.

RESULTS AND DISCUSSION

FRACTION	MASS (mg)	RECOVERY (%)
HP-20 OPEN COLUMN (1)		
25% acetone	N/A*	
100 % acetone	320.9	
HP-20 OPEN COLUMN (2)		
50% acetone	224.0	
75% acetone	45.5	
100% acetone	2.2	
TOTAL	271.7	84.67
HPLC C18 SEMI-PREP COLUMN		
F1	9.5	
F2	12.0	
F3	9.6	
F4	5.2	
F5	4.9	
F6	5.0	
F7	7.3	
TOTAL	53.5	85.57

Table 1. Mass yields from chromatography. Open column HP-20 and C18 HPLC Semi-Prep chromatography were employed to purify a culture of ecoSW202A. The 100% acetone fraction from HP-20 open column (1) was fractionated on HP-20 open column (2). The 50% acetone fraction from HP-20 open column (2) was fractionated using HPLC. *The initial 25% acetone fraction was not processed, and thus no weight is available.

The mass for each chromatography fraction was recorded after thorough drying. The results are shown in Table 1. Initially, 320.9 mg were recovered from the first round of HP-20 chromatography in the 100% acetone fraction. After a second round of HP-20 chromatography 271.7 mg, or 84.67% of the original material, was recovered. Approximately 67.20 mg of the 224.0 mg 50% acetone fraction were subjected to HPLC purification. Of that 53.5 mg, or 85.57%, was recovered.

The purification progression of the ecoSW202A sample was tracked via ^{1}H NMR spectroscopy. After each round of chromatography, NMR spectra were obtained for the resulting fractions. Figure 7 illustrates the improvement seen in resolution as well as sample purity after an initial round of HPLC.

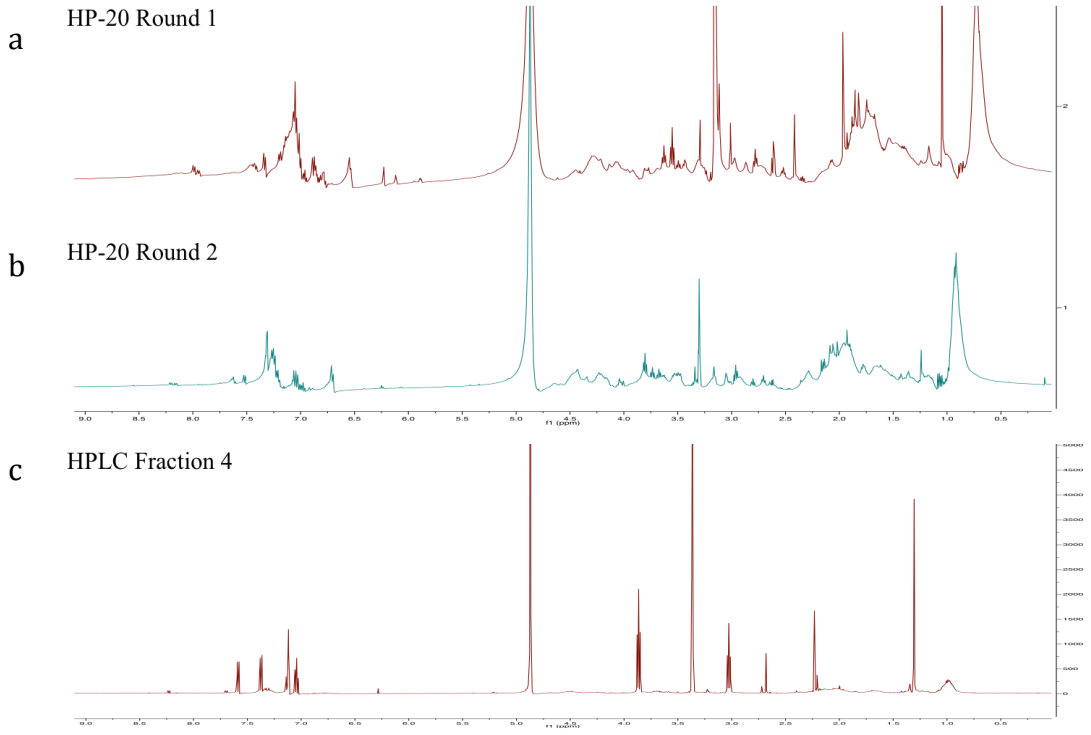


Figure 7. Results of chromatography purification. a) ecoSW202A 100% acetone fraction [from HP-20 open column (1)]. b) ecoSW202A 50% acetone fraction [from HP-20 open column (2)]. c) ecoSW202A F4 (from HPLC C18 Semi-Prep column).

When we transitioned to HPLC it was important to determine which experimental

conditions would yield the best separation of our sample. To do this small injection volumes were run with varying solvent mixtures to determine the most favorable. As discussed above, a 1:1 methanol:water mixture provided the greatest separation without altering the properties of the resin. Unfortunately, much of the resolution that was obtained using the 10 µL HPLC injections was lost when scaling up to 100 µL injections (Figure 8a-b). By observing peak intensity and separation at A_{254}, we were able to identify seven chromatography fractions. Fractions 1-7 were collected at 1.7, 2.7, 4.0, 4.7, 5.5, and 9.0 min, respectively (Figure 8b), and characterized via ^{1}H NMR.

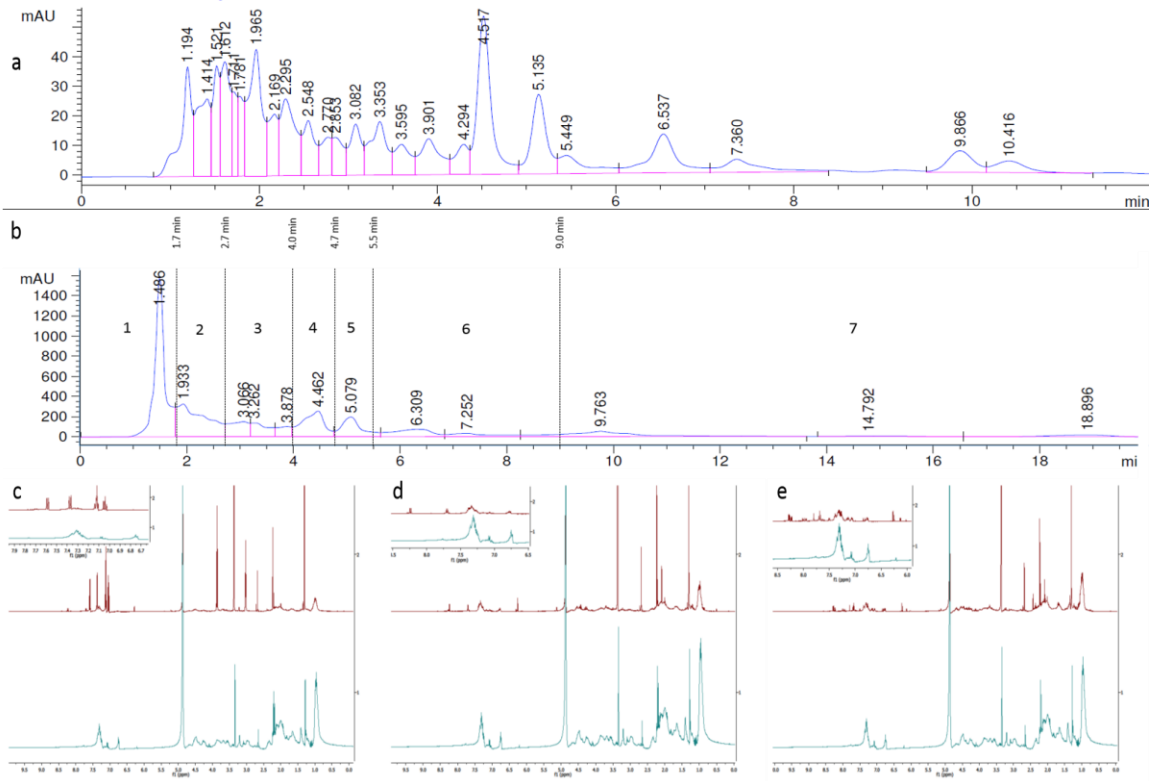


Figure 8. Using HPLC and ^{1}H NMR spectroscopy to identify novel proton peaks. a) UV-Vis readings at 254 nm for original 10 µL injection of sample on C18 Semi-Prep column. b) UV-Vis readings for 100 µL sample injection on C18 Semi-Prep column. Dotted lines indicate the switch to a new fraction. c) Top (red): ^{1}H NMR from HPLC faction 4; bottom (blue): ^{1}H NMR from ecoBlue 1 control. d) Top (red): ^{1}H NMR from HPLC faction 6; bottom (blue): ^{1}H NMR from ecoBlue 1 control. e) Top (red): ^{1}H NMR from HPLC faction 7; bottom (blue): ^{1}H NMR from ecoBlue 1 control. (For expanded NMR spectra see Appendix II).

The NMR spectra from Fractions 1-7 were compared to the ecoBlue 1 control spectrum

to identify unique peaks (Figure 8c-e). As demonstrated by the natural product examples presented in the introduction, many natural products have aromatic properties. Thus particular attention was paid to comparing the 6.0-8.0 ppm regions of the control and experimental spectra to target aromatic compounds. F4, F6, and F7 showed potentially novel peaks in the aromatic region of their spectra when compared with the control spectrum. F4 was ultimately selected for further analysis. To further analyze F4, ^{13}C, HMBC, HSQC, COSY and NOESY experiments were performed (Figure 9) in order to propose a structure for the isolated compound. The proposed structure and related NMR data are presented in Figure 10.

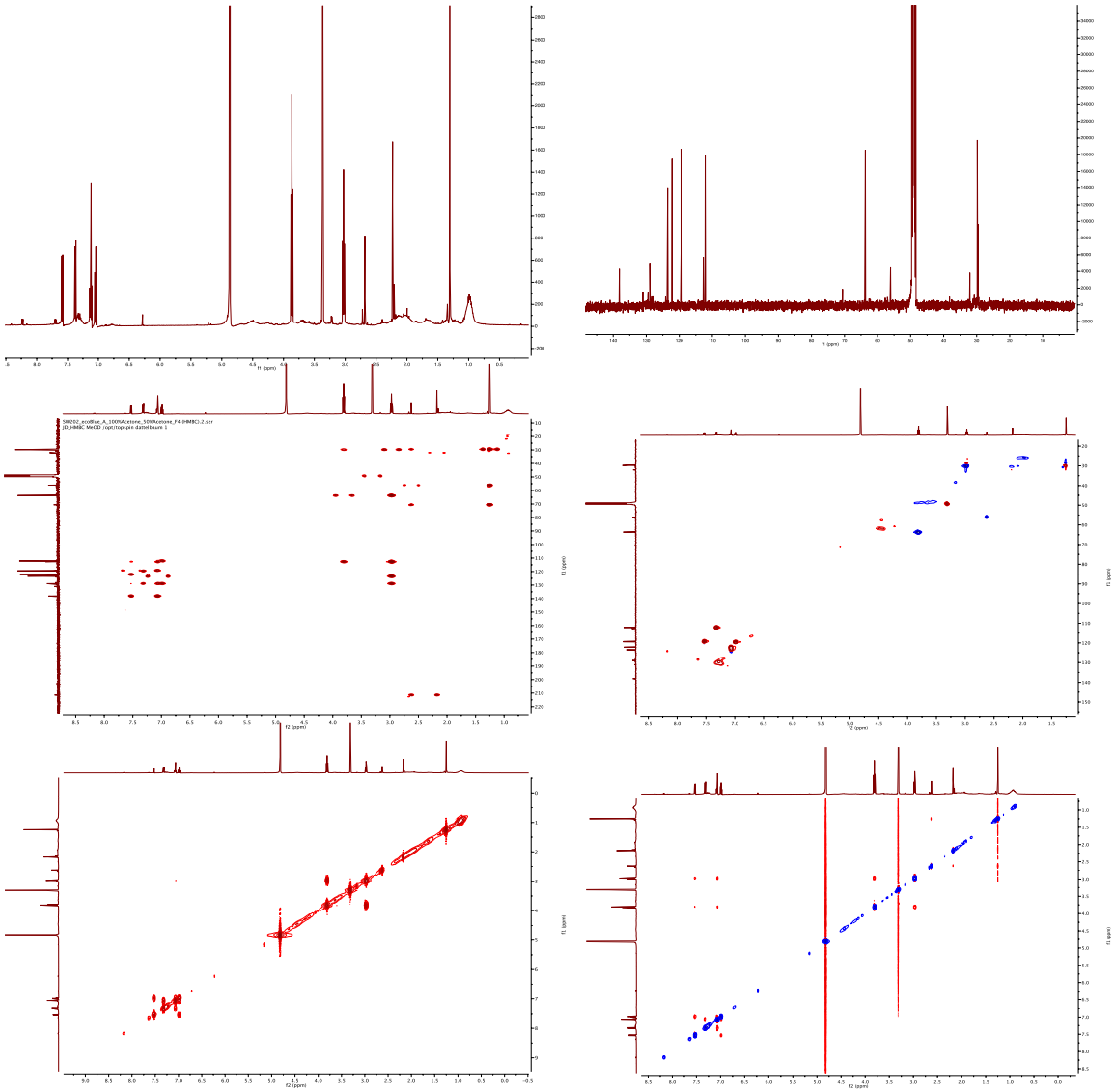


Figure 9. Fraction 4 NMR spectra. a) ^{1}H NMR spectrum b) ^{13}C NMR spectrum c) HMBC NMR spectrum d) HSQC NMR spectrum

1-(2-nitrosoethyl)-2-benzofuran

Position	^{13}C	^{1}H	Mult.	J (Hz)
1	63.7	3.86 (2H)	t	7.3
2	29.8	3.03 (2H)	td	7.3, 0.9
3	112.7			
4	128.9			
5	112.2	7.37 (1H)	dt	8.1, 1.0
6	119.5	7.04 (1H)	ddd	8.0, 7.0, 1.1
7	122.3	7.12 (1H)	ddd	8.0, 6.9, 1.1
8	119.2	7.58 (1H)	dt	7.9, 1.0
9	138.1			
10	123.5	7.12 (1H)	s	

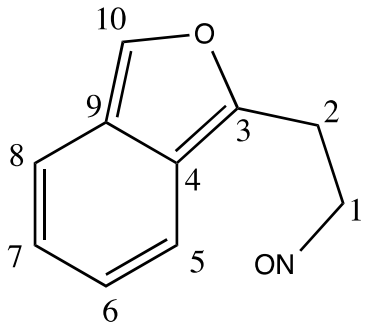


Figure 10. H^1 and C^{13} chemical shift assignments and proposed structure.

^{1}H NMR (500 MHz, Methanol-d_4) δ 7.58 (dt, J = 7.9, 1.0 Hz, 1H), 7.37 (dt, J = 8.1, 1.0 Hz, 1H), 7.12 (ddd, J = 8.0, 6.9, 1.1 Hz, 2H), 7.12 (s, 1H), 7.04 (ddd, J = 8.0, 7.0, 1.1 Hz, 1H), 3.86 (t, J = 7.3 Hz, 2H), 3.03 (td, J = 7.3, 0.9 Hz, 2H).

^{13}C NMR (126 MHz, Methanol-d_4) δ 138.13, 128.92, 123.51, 122.18, 119.47, 119.23, 112.73, 112.15, 63.68, 29.81.

^{1}H NMR showed five peaks in the aromatic region (7.04, 7.12, 7.12, 7.37, and 7.58 ppm) with equal integration values. There was an overlap of peaks occurring at 7.12 ppm of a singlet and a doublet of doublet of doublets. Four of the peaks (7.04, ddd; 7.12, ddd; 7.37, dt; 7.58, dt) were indicative of a 1,2 di-substituted benzene ring, with the downfield proton peaks located *ortho* relative to the two substituents in positions 9 and 4 above, and the upfield proton peaks *meta* relative to the two substituents in positions 6 and 7 above. This was supported by the splitting patterns and coupling constants: the ddd splitting with two coupling constants of 7.0 Hz and 8.0 Hz (located at positions 6 and 7) is likely due to *ortho* coupling, as protons at the 6 and 7 position on the benzene ring would each be located on adjacent carbons, with each having two non-equivalent protons in the *ortho* positions (Figure 11). The carbon-proton pairs were discerned using the HSQC data. HMBC was then used in conjunction with the proton splitting patterns discussed above to determine the location of each proton-carbon pair within the aromatic ring. COSY and NOESY spectra supported the relative position of protons located on the isobenzofuran ring system, as well as on the ethyl-based substituent. The final proposed assignments can be seen in Figure 10.

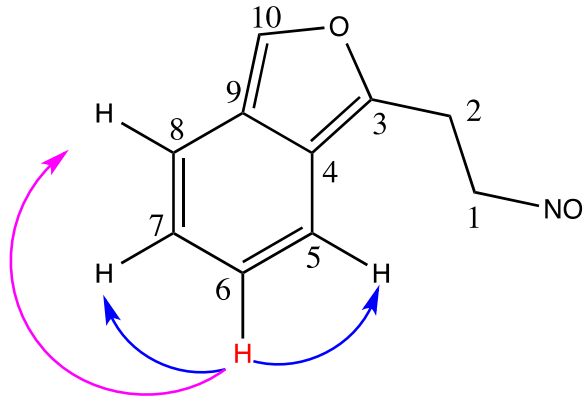


Figure 11. Proton coupling resulting in a ddd splitting pattern. The proton at position six shows *ortho* coupling to two non-equivalent protons on adjacent carbons (blue arrows). The same proton also shows *meta* coupling to the proton at position 8 (purple arrow). The same pattern is observed for the proton at position 7.

The dt peaks at 7.37 and 7.58 ppm were more difficult to rationalize. The coupling constant associated with the doublet splitting pattern (close to 8.0 Hz for both dt protons) indicates *ortho* coupling with each proton's respective neighbor (at positions 5 and 6). It seems likely that the triplet splitting pattern is the result of fine structure. The low coupling constants of the triplets (< 1.0 Hz in most cases), would seem to suggest *para* coupling across the ring. The combination of *ortho*, *meta*, and *para* coupling at the 3 and

6 positions along the ring may suggest that in fact the dt is a ddd, with the fine structure resulting from the para coupling overlapping to some extent.

The ^{1}H spectrum shows a very clear triplet at 3.86 ppm and a triplet of doublets at 3.04 ppm, both integrating to two protons each. HSQC data showed that these two peaks coupled with carbons at 63.68 and 29.81 ppm, respectively. HMBC confirmed that the two methylenes were likely neighbors, as there was strong coupling observed between both of the groups. This indicated the presence of an ethyl-based substituent group. The protons located on the methylenes did not, however, couple into the benzene ring system according to the HMBC data, making it less likely that this ethyl group was a substituent located at position 4 and 9 on di-substituted benzene ring.

By comparing HMBC and HSQC results, four carbons were identified that were likely quaternary carbons. These suspected quaternary carbons were ^{13}C peaks located at 138.1, 128.9, 123.5, and 112.7 ppm. The downfield chemical shifts of these peaks indicated that they were likely aromatic carbons. HMBC showed that the peaks at 138.1 and 128.9 ppm both coupled strongly with the methines identified as part of the 1,2 disubstituted benzene ring. It was determined that the peak at 138.13 ppm was located at the 4 position, and the peak at 128.92 ppm at the 9 position using the HMBC data provided. The carbon peak at 138.1 ppm in turn coupled to the singlet peak at 7.12 ppm. The carbon peak at 128.9 coupled to the quaternary carbon located at 112.7 ppm. This quaternary carbon coupled strongly into the ethyl group discussed above, indicating it was likely the position where the ethyl substituent attached to the ring system (position 9). Of the two methylene carbons on the ethyl substituent, the carbon at 29.8 ppm appeared to couple fairly strongly into the ring system, indicating it was the carbon attached to the ring. The second ethyl carbon at 63.7 ppm appears to be split solely by the two protons located on the methylene at 29.8 ppm. The downfield shift to 63.7 ppm would appear to be associated with a nitro-like group.

Ultimately a nitroso group was selected based upon the chemical shifts predicted by ChemDraw ^{1}H NMR prediction software. Another likely candidate for this position is a

halogen species. Placing a bromine atom on the ethyl substituent yielded a similar downfield shift as the nitroso group according to the ChemDraw prediction software. If this is the case, mass spectrometry data should show two peaks of approximately equal intensity with a difference of two m/z units between them, each corresponding to the molecular ion but one with the ^{79}Br isotope and the other containing the ^{81}Br isotope. An oxygen was selected to complete the five-membered ring, however this too cannot be confirmed without a molecular weight provided by mass spectrometry data. A fused benzene-furan with a 2-nitroso ethyl substituent seems a likely potential candidate for the isolated molecule.

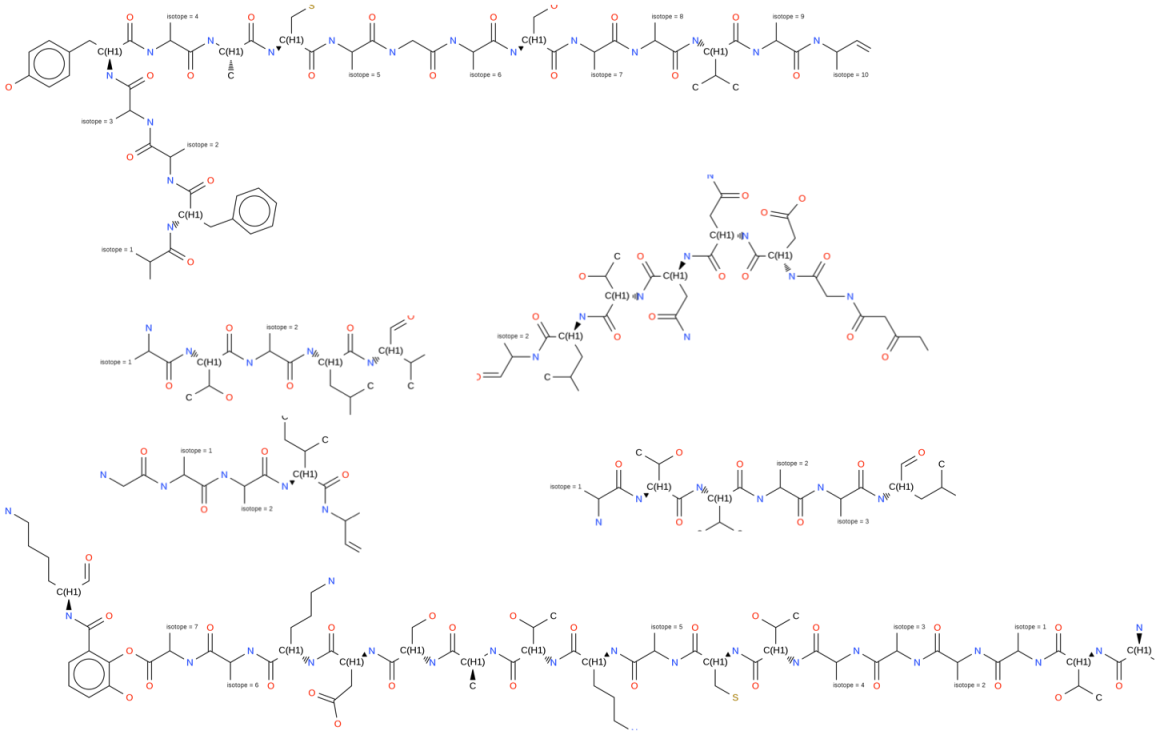


Figure 12. Predicted synthesis products based on antiSMASH analysis. Products of six sequenced gene clusters are shown above. The predicted structures do not include tailoring reactions that would occur to create the products' final form. All reported products contained greater than 10 carbons.

antiSMASH analysis[15] of the SW202 genome done by our lab provided predicted structures for the products of the NRPS and PKS gene clusters present in the SW202 microbe (Figure 12). All of these proposed structures had greater than 10 carbons. This may indicate that the proposed 1-(2-nitrosoethyl)-benzofuran isolate is a partial product of an NRPS and PKS cluster. In order to confirm this the fosmid insert from colony A will have to be sequenced.

CONCLUSION

According to the NMR data collected, we were able to successfully isolate a chemical entity not present in a control culture of ecoBlue 1 cells. The proposed identity of this chemical entity is a 1-(2-nitrosoethyl)-benzofuran, however further analysis using mass spectrometry will be required to determine the presence of any nitrogen and halogen species. Further research will also be required to determine whether or not the isolate in question is the product of an NRPS or PKS gene cluster. While our functional screen has determined the presence of PPTase in the ecoSW202A sample, it is unclear if a full gene cluster was isolated, and if it was, whether or not this is a direct product.

This compound, as well as the various collected fractions, will need to be screened for bioactivity. If a promising isolate is characterized, we will apply for the NCI-60 screening program operated by the National Cancer Institute. This program screens novel, primarily heterocyclic aromatic compounds, for activity against 60 different cancer cell lines. In addition, the lab will explore chemical genomic screens on yeast strains to determine the biological pathways involved in the mechanism of action for promising isolates. Disk diffusion assays against Gram-negative and Gram-positive bacterial strains will also be performed.

[15]Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Breitling, R.; Takano, E. antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters. *Nucleic Acids Research* 2011; doi: 10.1093/nar/gkr466.

The current data collected for ecoSW202A are promising. With luck, further work will be able to determine if our isolate has any promising bioactive, and successfully link it to one of the NRPS or PKS gene clusters that we have sequenced from the SW202 genome.

APPENDIX I: COLUMN METHODS

Diaion ® HP-20 Column

The Diaion ® HP-20 column was used to fractionate 2 L cell cultures. For 2L cultures, a 30 mL column was packed. Cultures were centrifuged at 6,000 xg. The sample was then decanted, and the supernatant collected. The cell pellet was stored at -20 °C for later use.

Supernatant

The column was prepared with successive washes of 100% acetone, water, and methanol (three column volumes each). Once prepped, the supernatant was applied to the column and the flow-through discarded. The bound molecules were eluted from the column in 50%, 75%, and 100% acetone solutions (three column volumes each). The fractions were dried via rotary evaporation (when about 3 mL of solution was left in the round-bottom flask, the solution was transferred to a 20 mL scintillation vial) and placed on the freeze-drier overnight. The dried fractions were stored at -20 °C for later use.

Cell Pellet

Methanol (20 mL) was added to the cell pellet and vortexed. The extraction mixture was allowed to sit overnight. The mixture was then centrifuged at 6000 xg, and the supernatant was collected. The remaining cell pellet underwent a second extraction in methanol overnight. The mixture was once again centrifuged as described above, and the supernatants combined. The column was then prepared with successive washes of 100% acetone, water, and methanol (three column volumes each). The methanol extract was then applied to the column and the flow-through collected and diluted 1:1 with water. The solution was then reapplied to the column. This process was repeated a two additional times (additional dilutions may be necessary if the flow-through is not clear). The final flow-through was discarded and the bound molecules eluted in three column volumes of 50%, 75%, and 100% acetone solutions. The fractions were dried using the method described in the above section. The dried fractions were stored at -20 °C.

Diaion® HP20SS Backloading

Diaion® HP20SS columns were used to remove water from samples after HPLC chromatography to remove water and thus reduce drying times. The procedure is as follows.

Fractions were diluted until approximately 25% organic solvent by volume. A 3 mL HP20SS column was packed and prepped using successive washes of 100% acetone, water and methanol (three column volumes each). The fraction was then applied to the column and the flow-through discarded. The bound molecules were then eluted in three column volumes of 50%, 75%, and 100% acetone. The sample was then dried via rotary evaporation and placed on the freeze drier overnight.

Diol Columns

Diol columns were used to fractionate samples after going through one round of HP-20 column chromatography, and before HPLC. Diol has a binding capacity of 6 mg/mL.

A Diol column was packed and prepped with three column volumes of the most nonpolar solvent to be used. For the majority of the samples run dichloromethane (DCM), ethyl acetate (EtAc), and methanol (MeOH) were the desired solvents. Mixtures of these solvents were prepared prior to running the column. The fractions selected were mixed in solutions of increasing polarity (i.e. in order: DCM, 3:1 DCM:EtAc, 1:1 DCM:EtAc, 1:3 DCM:EtAc, EtAc, 3:1 EtAc:MeOH etc.). The solution of interest was applied to the prepped column, and the supernatant discarded. Starting with the most non-polar solvent solution and progressing to the most polar, the bound molecules were eluted from the column. The fractions were then dried via rotary evaporation and then placed on the freeze drier overnight. The dried fractions were stored at -20 °C.

HPLC

A C-18 Semi-Prep column was used for all HPLC described in this paper. The column was equilibrated incrementally, beginning with 100% methanol at 4 mL/min. The water content of the solvent solution was increased by 10% every ten minutes until a 1:1

methanol:water solution was reached. For the Agilent 1100 series the maximum injection volume is 100 μL, and the maximum solute mass is 10 μg per run. The 224.0 mg sample discussed in this paper was dissolved in 10 mL of methanol. Ultimately, 100 μL injections were used with a 1:1 methanol:water solvent solution at 4 mL/min with 20 min runs. The UV-vis was set to detect at 254 nm.

APPENDIX II: NMR TECHNIQUES AND RESULTS

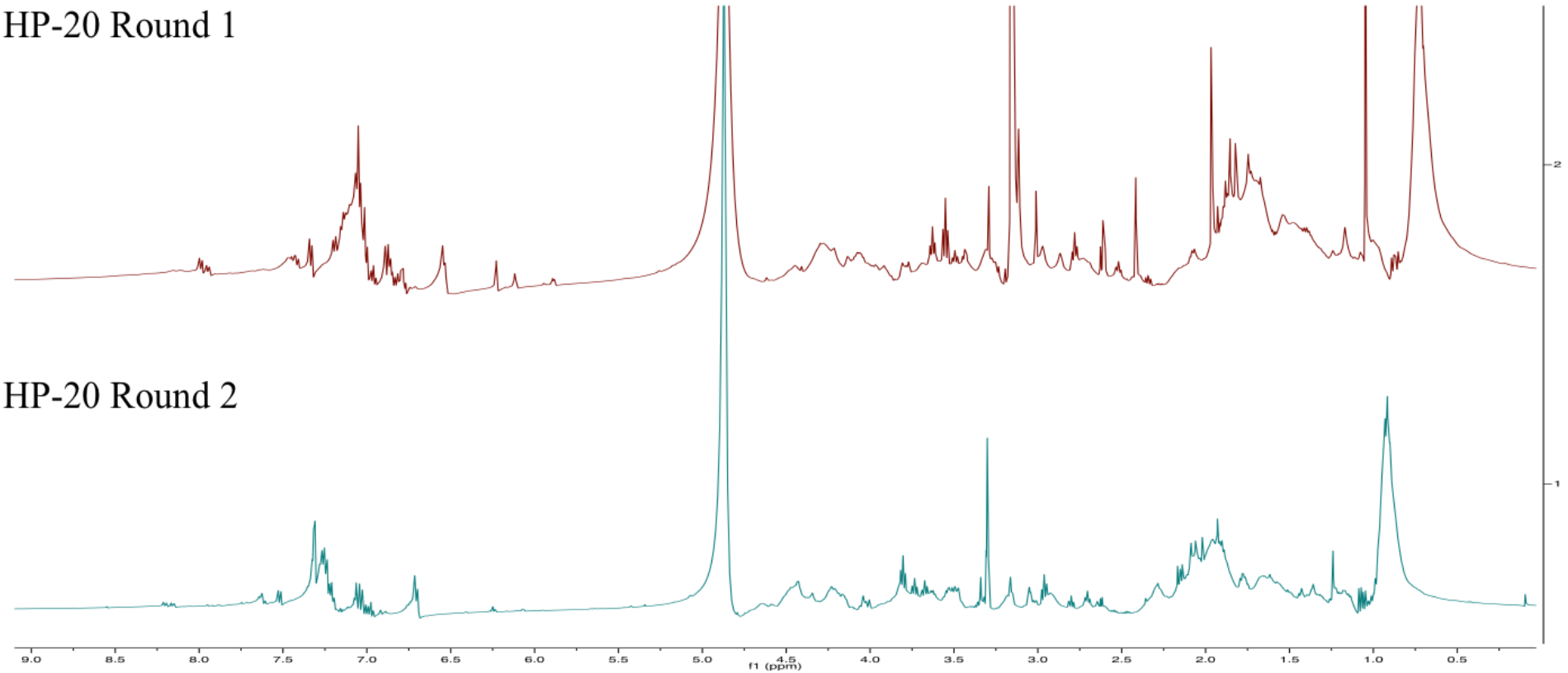
NMR Programs for Bruker 500

EXPERIMENT	PROGRAM	RUN DURATION
^{1}H	PROTON_UR	10+ min
^{13}C	CARBON_UR	overnight/weekend
HMBC	JD_HMBC	overnight
HSQC	JD_HSQCED	overnight
NOESY	COSY_UR	overnight
COSY	NOESY_UR	overnight

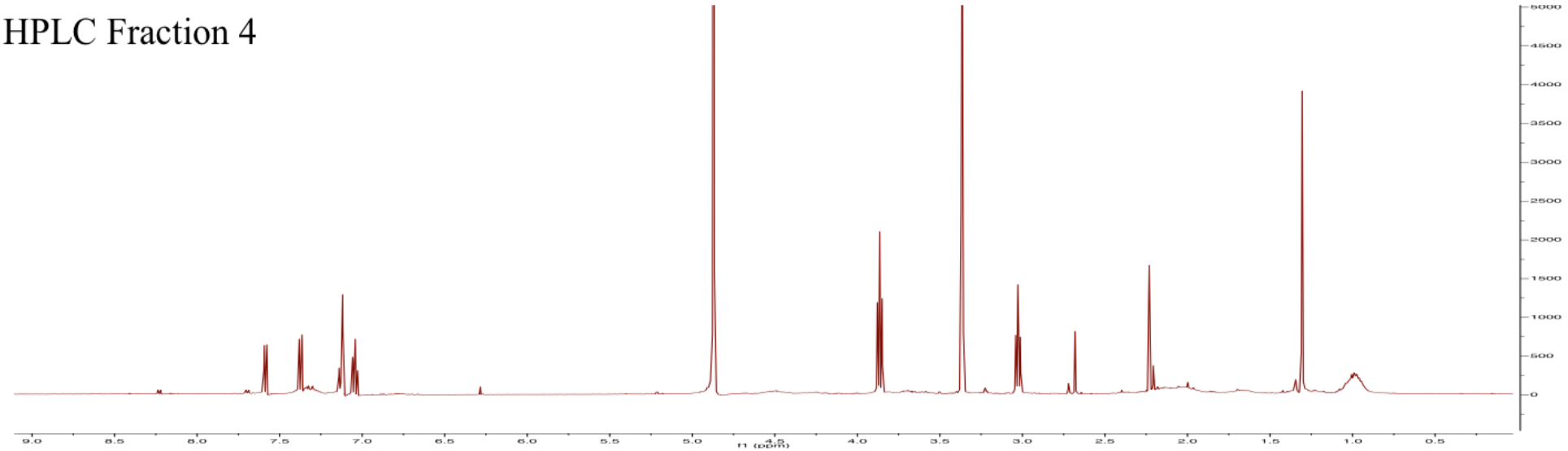
Table 2. NMR Experiments with associated University of Richmond Bruker Topspin programs and typical run times.

HP-20 Round 1

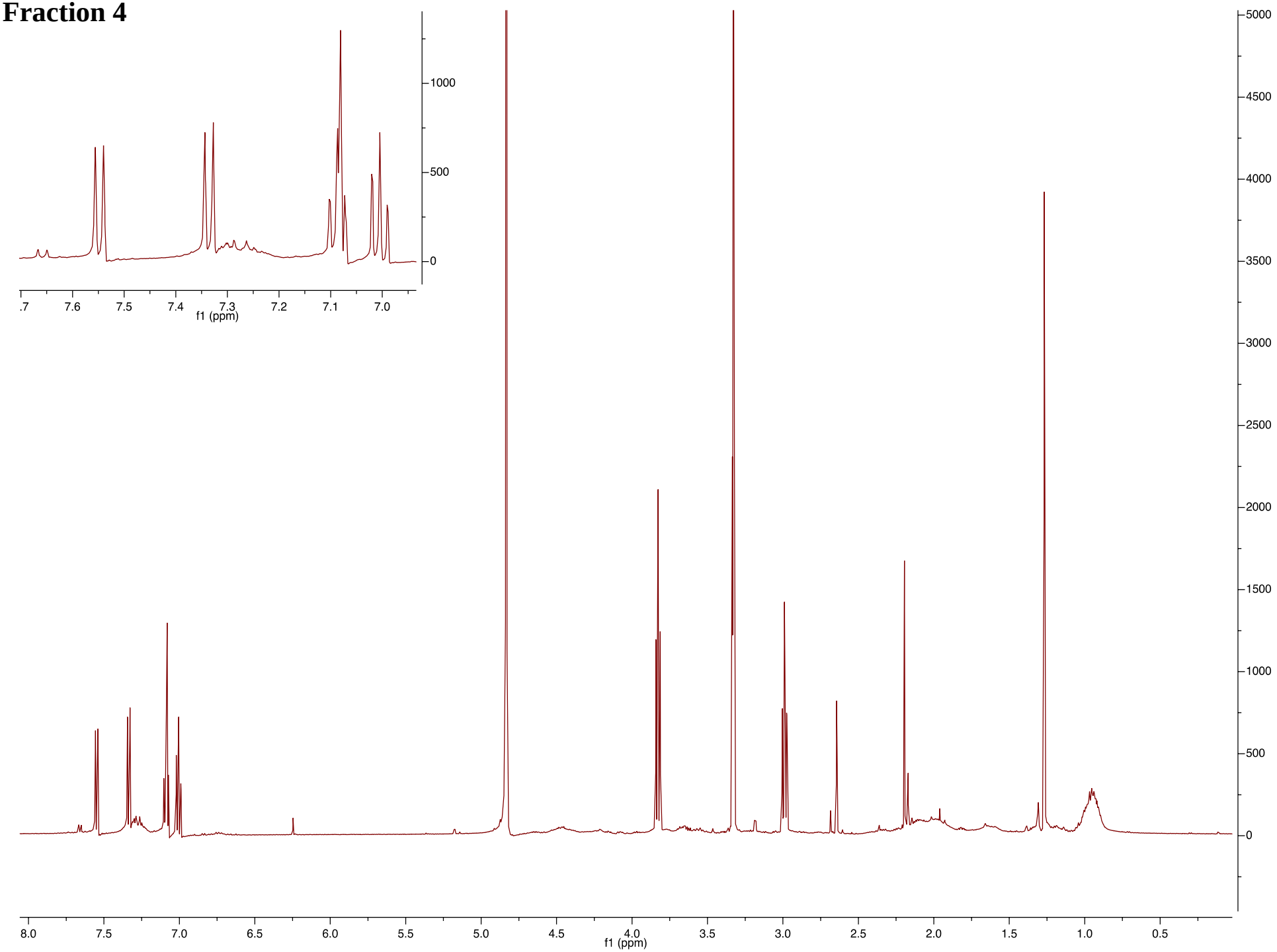
HP-20 Round 2



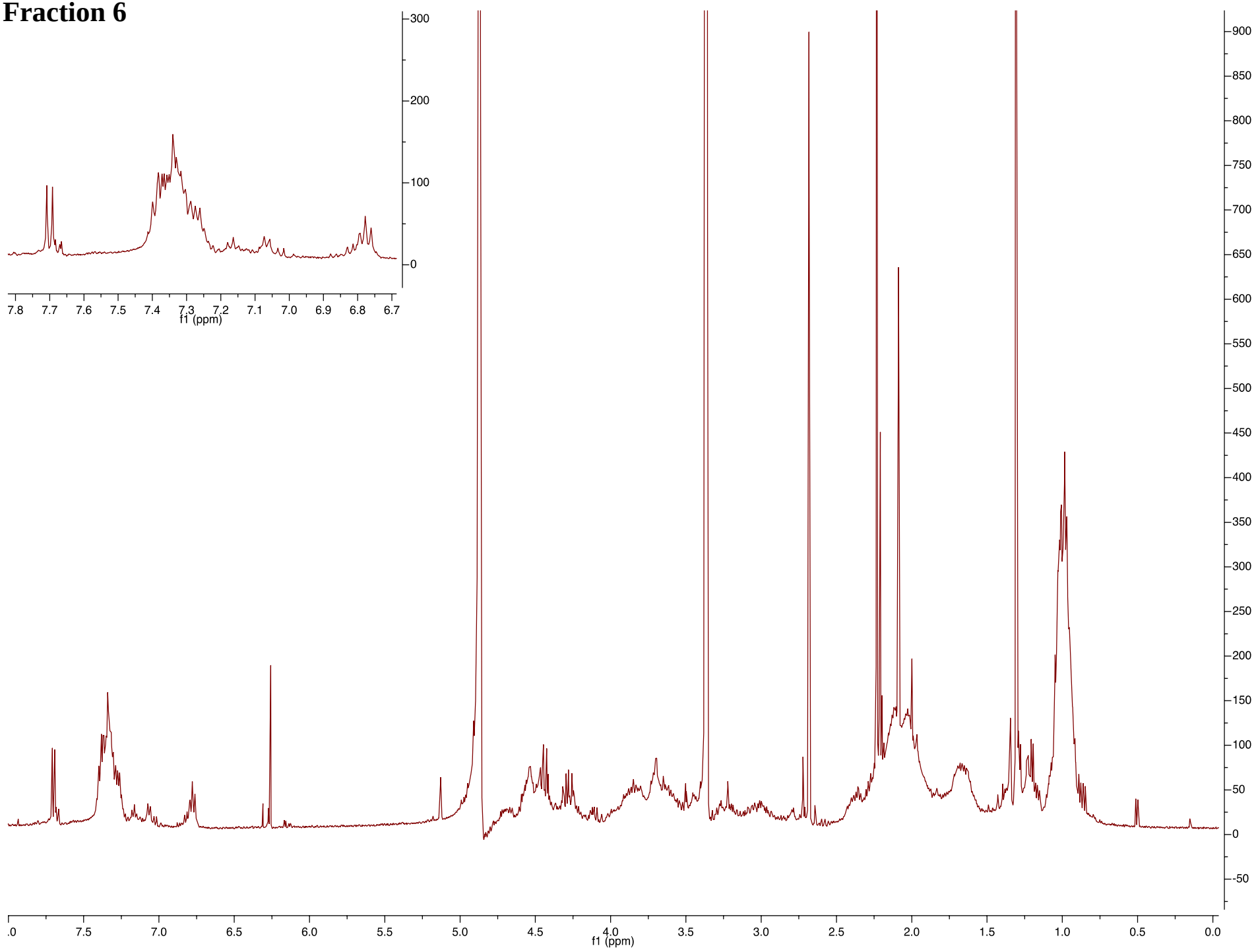
HPLC Fraction 4



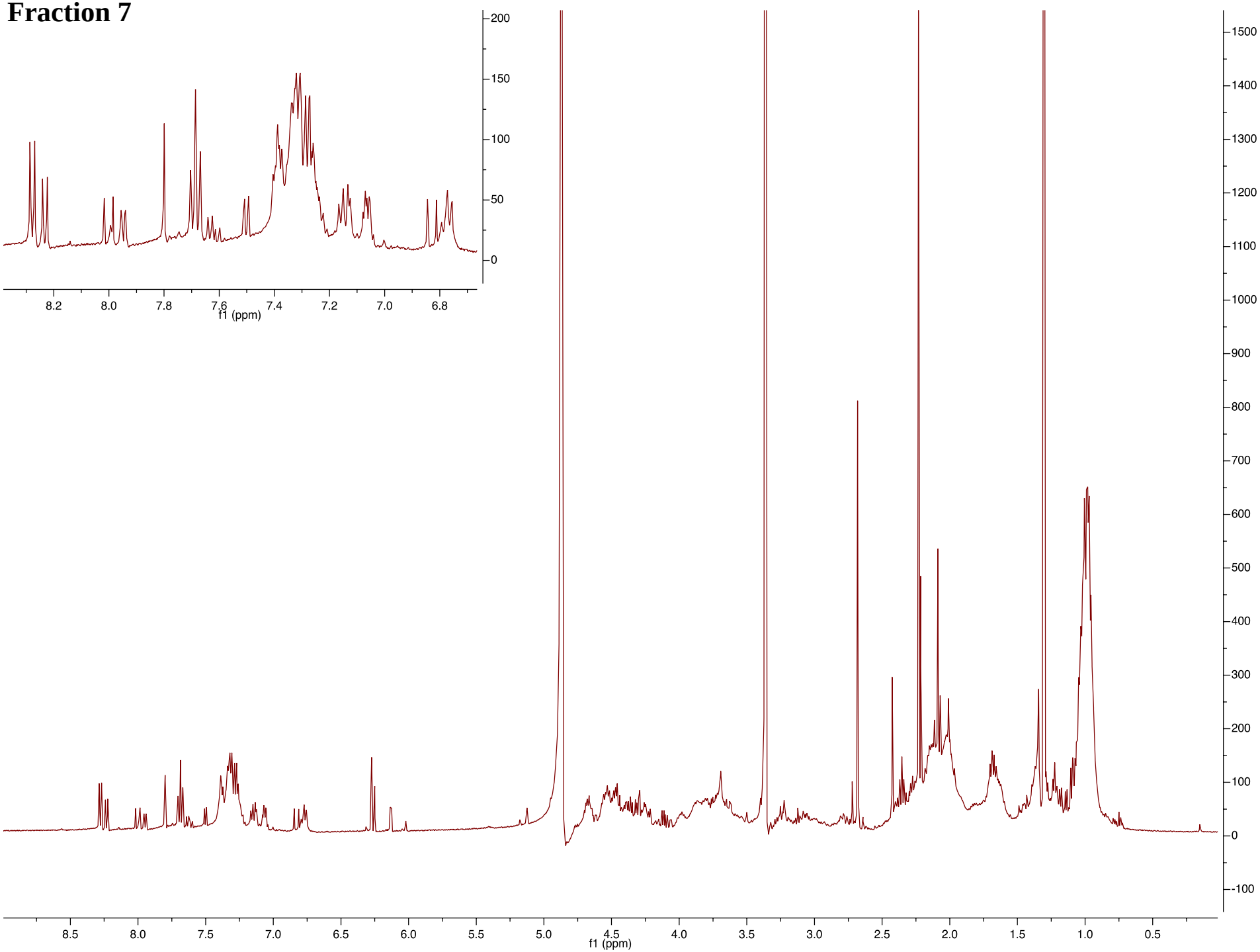
Fraction 4


1000
500
0
.7
7.6
7.5
7.4
7.3
7.2
7.1
7.0
f1 (ppm)
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
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
1.0
0.5
f1 (ppm)

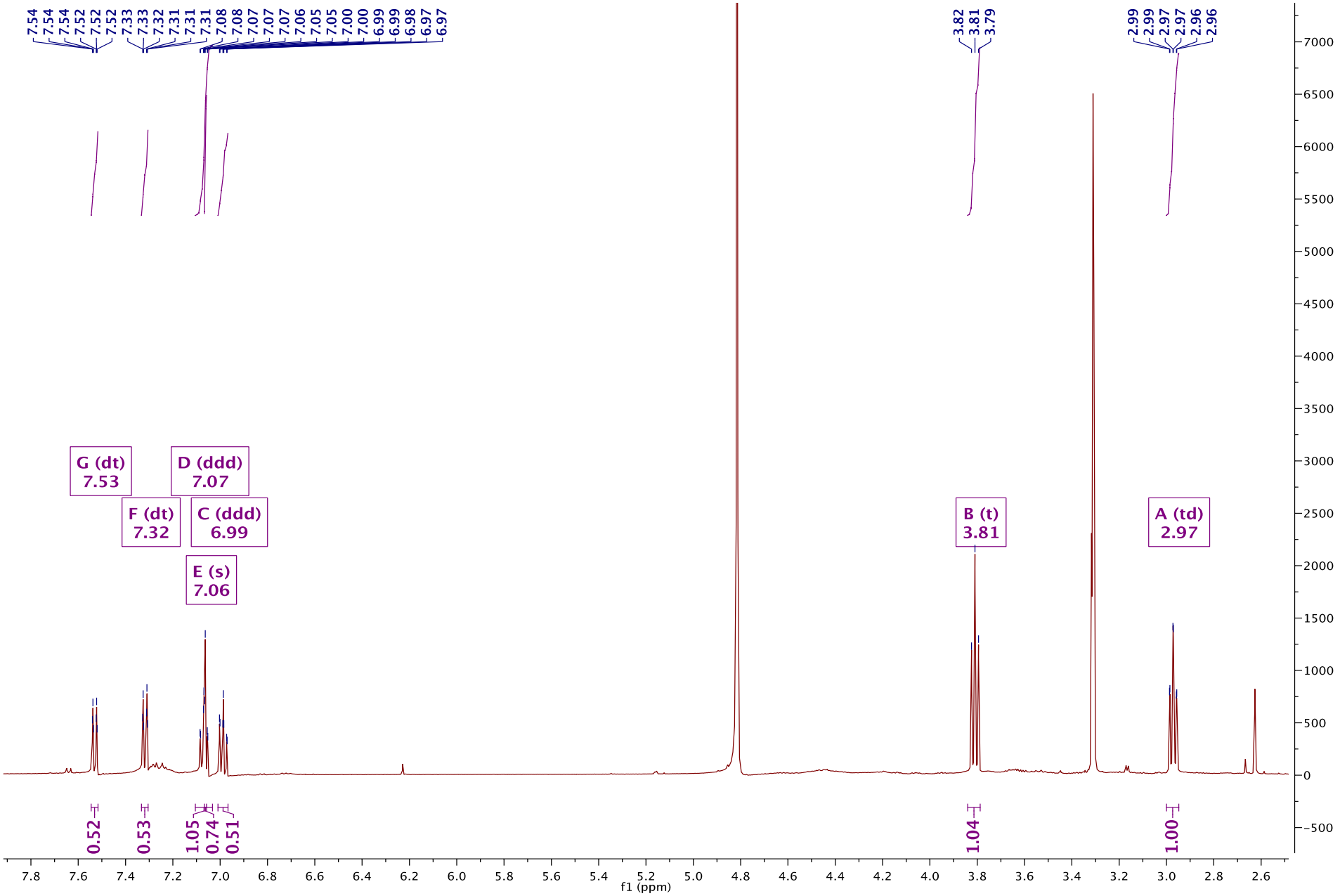
Fraction 6


300
200
100
0
7.8 7.7 7.6 7.5 7.4 7.3 7.2 7.1 7.0 6.9 6.8 6.7
f1 (ppm)
900
850
800
750
700
650
600
550
500
450
400
350
300
250
200
150
100
50
0
-50
.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 1.0 0.5 0.0
f1 (ppm)

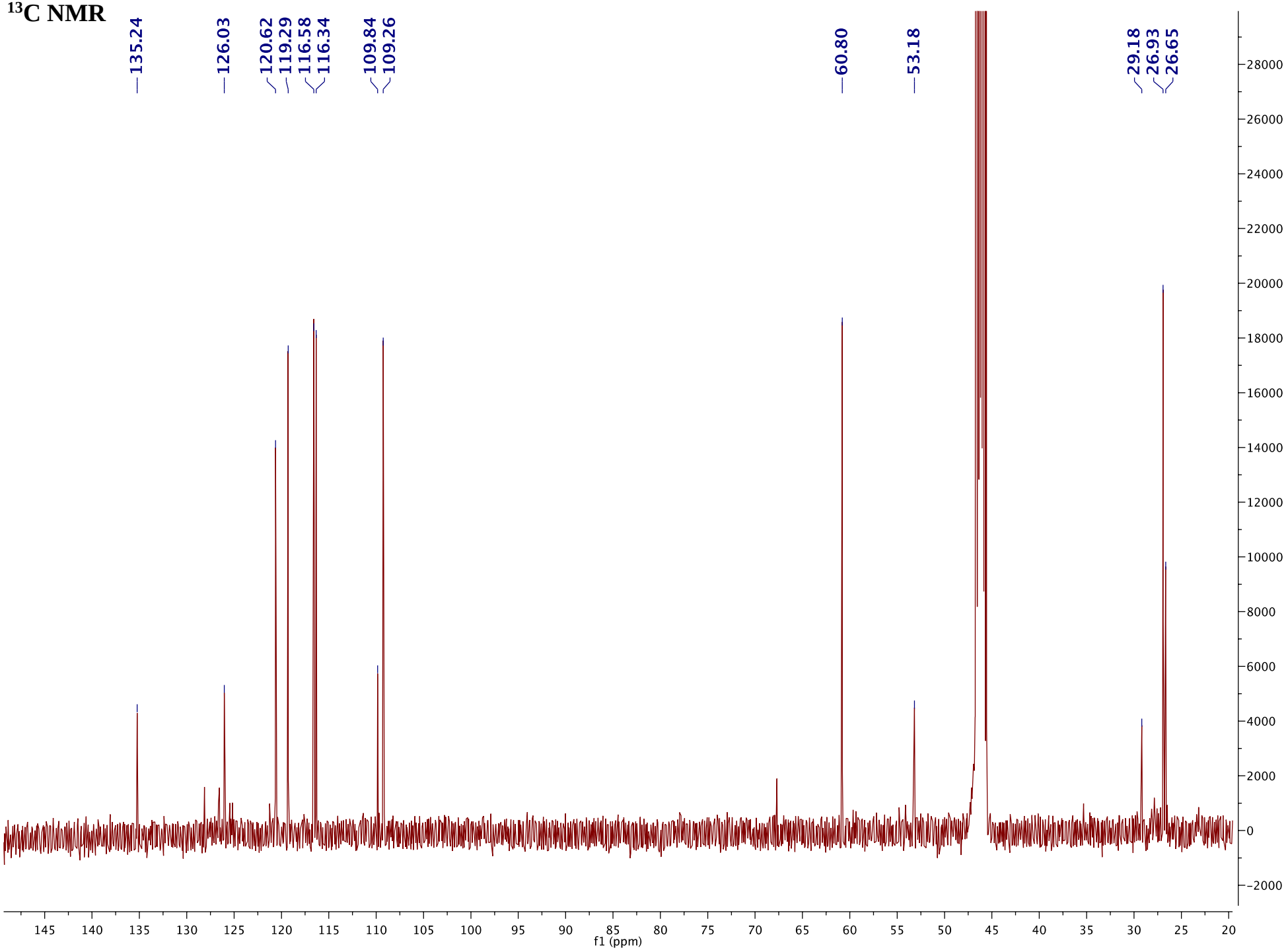
Fraction 7


200
150
100
50
0
8.2
8.0
7.8
7.6
7.4
7.2
7.0
6.8
f1 (ppm)
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
-100
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
1.0
0.5
f1 (ppm)

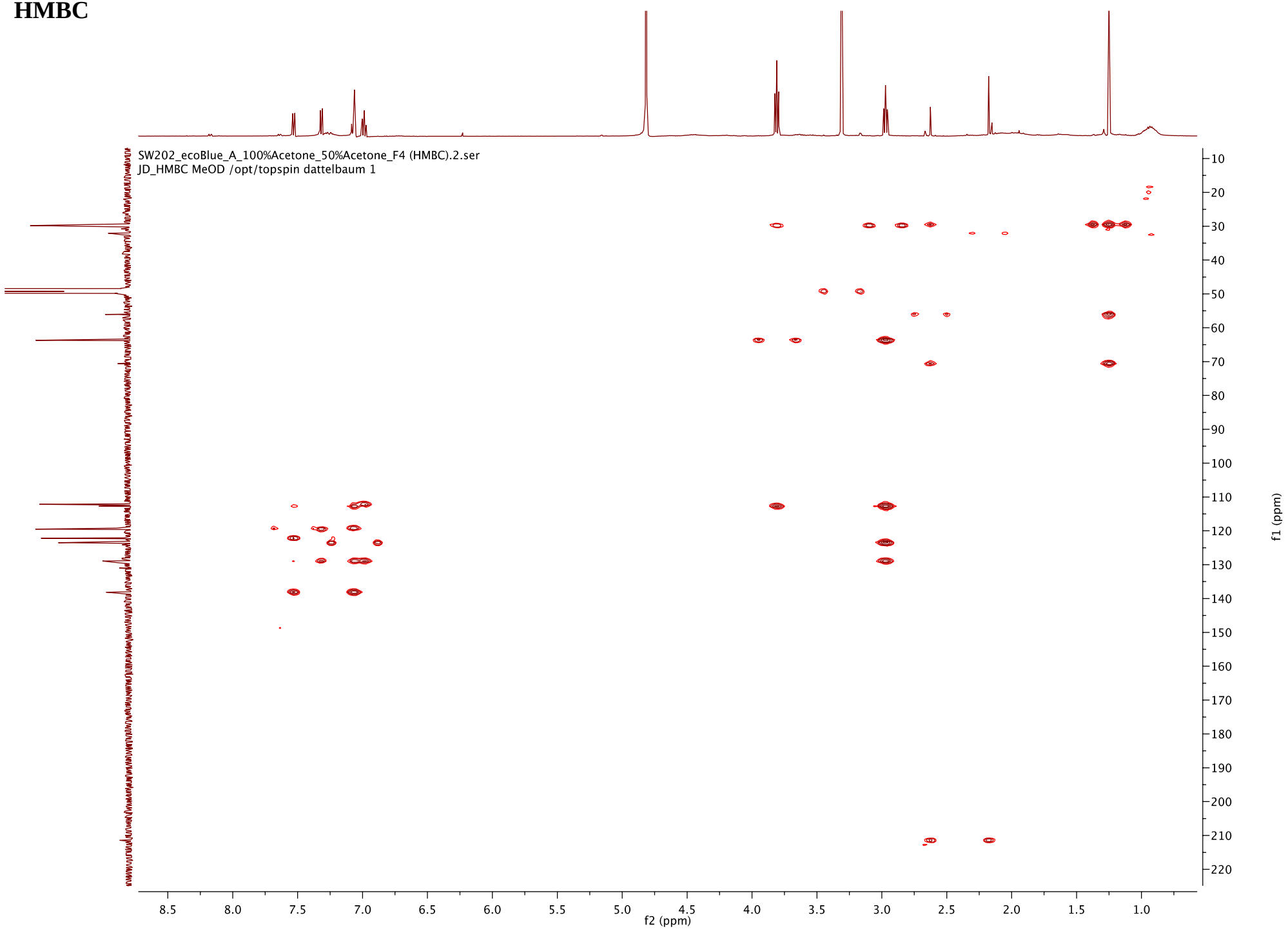
^{1}H NMR


7.54 7.54 7.54 7.52 7.52 7.52 7.33 7.33 7.32 7.31 7.31 7.31 7.08 7.08 7.07 7.07 7.07 7.06 7.05 7.05 7.00 7.00 6.99 6.99 6.98 6.97 6.97
3.82 3.81 3.79
2.99 2.99 2.97 2.97 2.96 2.96
G (dt) 7.53
F (dt) 7.32
D (ddd) 7.07
C (ddd) 6.99
E (s) 7.06
B (t) 3.81
A (td) 2.97
0.52 0.53 1.05 0.74 0.51 1.04 1.00
f1 (ppm)

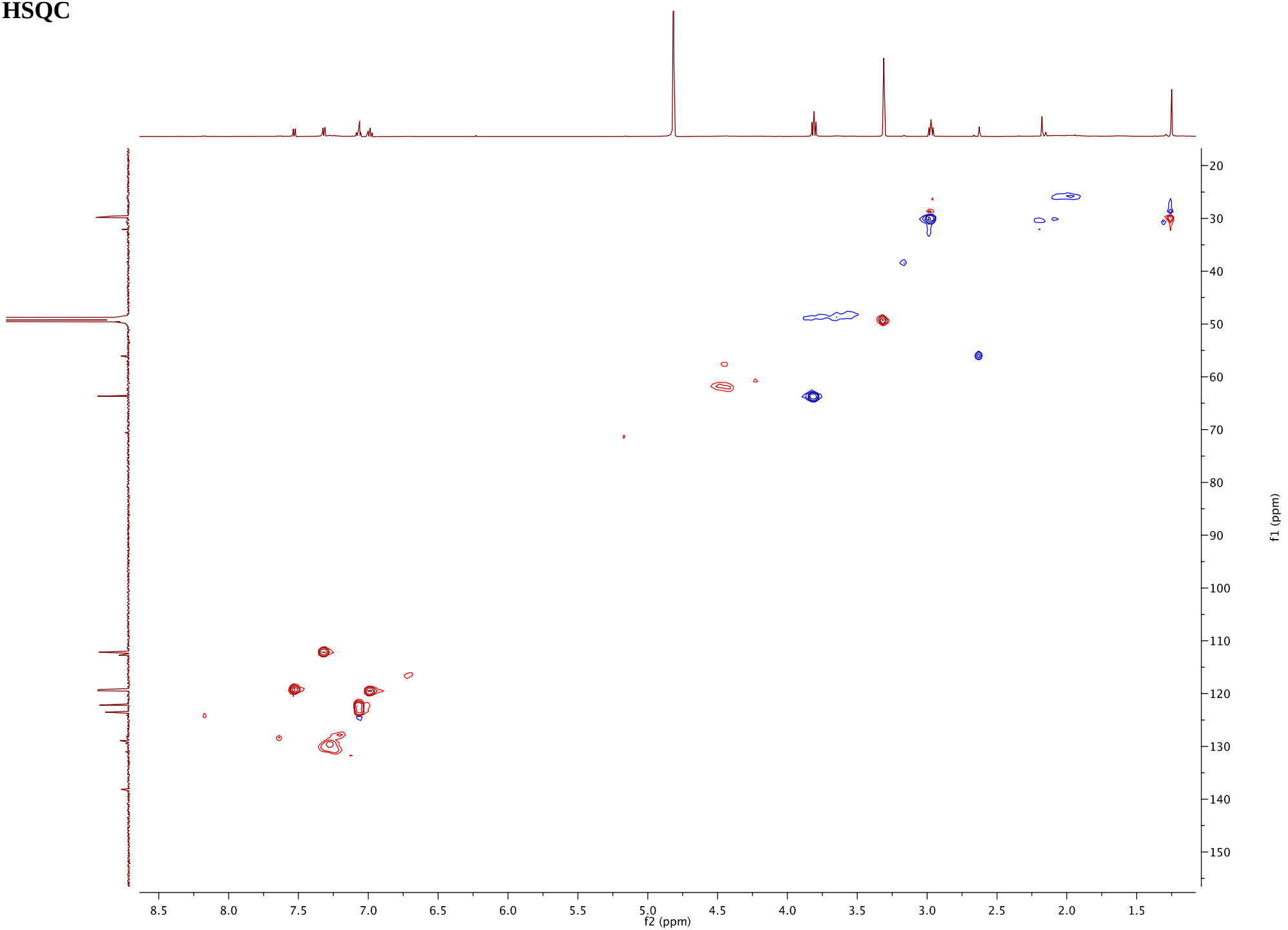
^{13}C NMR


135.24
126.03
120.62
119.29
116.58
116.34
109.84
109.26
60.80
53.18
29.18
26.93
26.65
28000
26000
24000
22000
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
-2000
145
140
135
130
125
120
115
110
105
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
f1 (ppm)

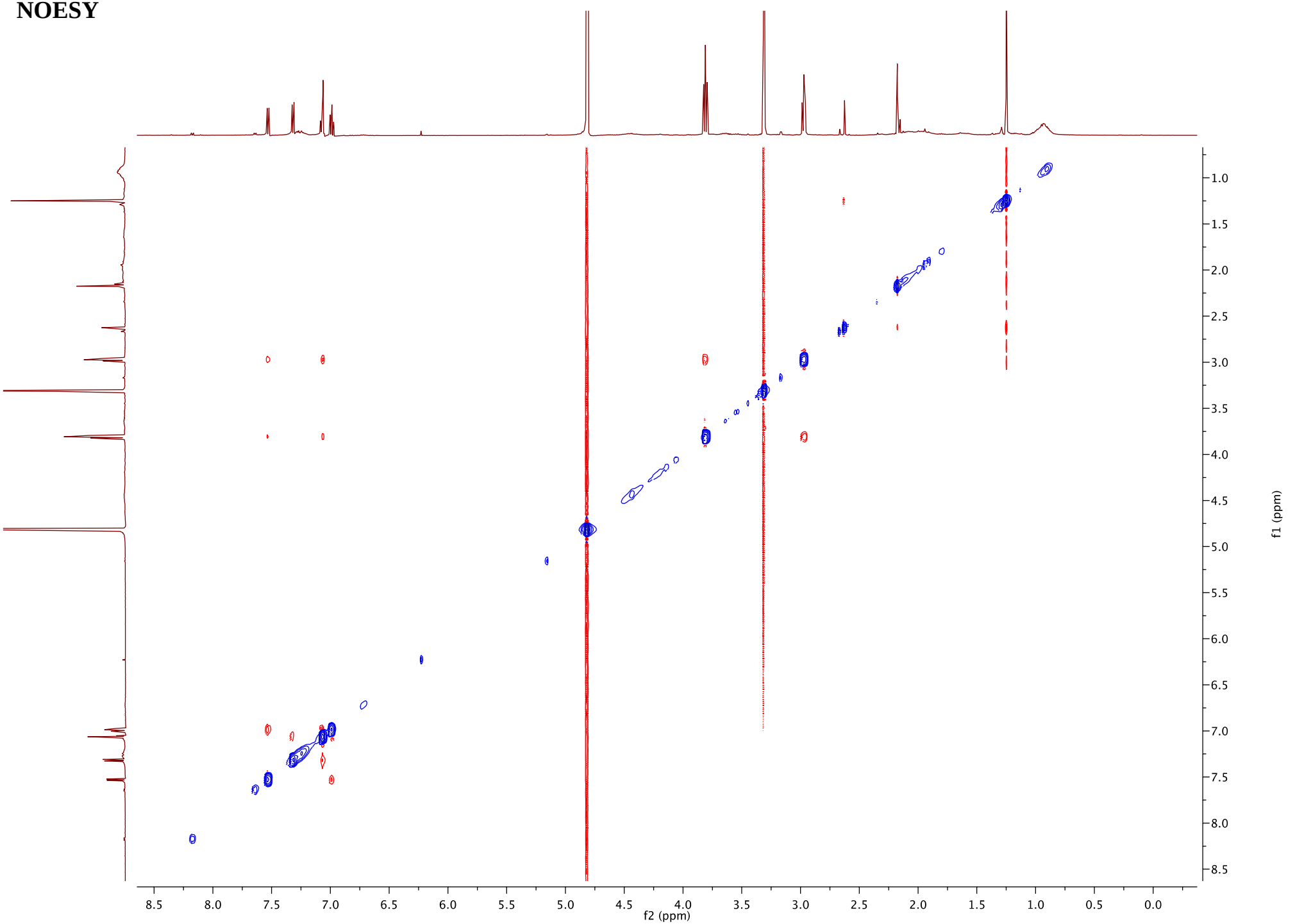
HMBC


SW202_ecoBlue_A_100%Acetone_50%Acetone_F4 (HMBC).2.ser
JD_HMBC MeOD /opt/topspin dattelbaum 1
f2 (ppm)
f1 (ppm)

HSQC


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)
20
30
40
50
60
70
80
90
100
110
120
130
140
150
f1 (ppm)

NOESY


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
1.0
0.5
0.0
f2 (ppm)
1.0
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
f1 (ppm)

COSY