IDENTIFICATION OF SMALL MOLECULES THAT REGULATE NEMATODE – ENVIRONMENT INTERACTIONS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Maro John Kariya

December 2019

IDENTIFICATION OF SMALL MOLECULES THAT REGULATE NEMATODE‐ENVIRONMENT INTERACTIONS

Maro John Kariya, Ph.D.

Cornell University [2019]

ABSTRACT

One of the central channels of nematode-environment interactions is through small molecules (non-polymeric chemical entities with a molecular mass <1000 Daltons). Nematodes release an abundance of small molecules which affect their environment, such as the ascarosides, di-deoxy sugar lipids which control the development and behavior of nematodes and other members of their environment. Through the advancement of high-resolution mass spectrometry and analysis software, we now find a wealth of other small molecules produced by nematodes which play key roles in their environmental interactions.

Herein the author describes the integrated use of 2D NMR and high-resolution UHPLC-MS/MS to aid metabolomics of complex natural samples to characterize small molecules that regulate interactions between nematodes and their environment. Combining these analytical techniques with biological assay data, the author has

characterized small molecules from nematodes which elicit defensive responses in other nematodes and affect microbial and nematode growth. The library of small molecules presented in this dissertation provide a new layer of structural diversity in nematodes and define a path forward in discovering other molecules of their kind.

The novel small molecules shown in this thesis include - Branched chain sulfolipids (sufal/ac), molecules that cause defensive behavior in prey nematodes that appear to be linked to the development of the predator nematode's mouth dimorphism. Cyclic sulfates (cysul), a class of metabolites unprecedented in nature, with possible epoxide precursors that suggest a straightforward biosynthetic pathway. The cyclic sulfates, unlike the branched chain sulfolipids, do not affect defensive behaviors in other nematodes, but may play a yet undiscovered role in a nematode's interactions with its environment. Cyclopropyl lipids (cpfa), nematode derivatives of bacteriogenic metabolites which inhibit microbial growth in several species. The activity of these lipids was found to be stereochemically specific, suggesting that the effect on microbial growth may be through a signaling mechanism.

Through these studies the author hopes to convey connections between small molecules created by nematodes with their environment and draw parallels with their potential effect on other organisms.

BIOGRAPHICAL SKETCH

Maro J. Kariya grew up in Brooklyn, NY and attended Stony Brook University, where he found an interest in organic chemistry through the passion of his teachers. Through his organic chemistry recitation teaching assistant, Maro was introduced to Prof. Francis Johnson, a professor in the pharmacology department at Stony Brook University. Under faculty guidance, Maro helped characterize the structure of the first intermediate in the Pabon reaction, which is used to make curcumin. Through these studies, Maro found an interest in medicinal and organic chemistry. In 2012, he conducted research through a SURF fellowship at the Jet Propulsion Laboratory in Pasadena, CA, under the guidance of Dr. James Bradley Dalton III. There, he generated reference spectra for salts at temperatures and pressures resembling those on Europa's surface in order to model the composition of the moon.

Inspired through this internship, Maro decided to apply to graduate school. He applied to ten, was denied by eight, and went to one. Maro had explored various groups and, during a time of pondering at The Chapter House, chanced upon a conversation with Prof. Frank C. Schroeder. Maro soon noticed that he would be a good fit with the group. He initially studied fungal metabolites with the hope to link artificially activated gene clusters with metabolites and eventually was drawn to studying nematode metabolites. The rest of the thesis describes these ventures.

Dedicated to my family, friends, community, those who have helped me, those who have inspired me to get to this point, and those who are helping the world heal.

ACKNOWLEDGEMENTS

Special Committee Chair and Advisor: Prof. Frank C. Schroeder

Special Committee: Profs. Brian Crane, Jeremy Baskin, and Yimon Aye

Collaborating Principal Investigators: Profs. Ralf J. Sommer, Jagan Srinivasan, Sreekanth Chalasani, Dennis H. Kim, Erik J. Ragsdale.

Collaborating Researchers: Dr. Neelanjan Bose, Joshua J. Yim, Spencer Wong, Zheng Liu, Harini Sadeeshkumar, Jazmin Aguilar-Romero, Christopher D. Chute, Dr. Alex Artyukhin, Amy K. Pribadi, Joseph Biddle, Sarah G. Leinwand, Ada Tong, Kevin P. Curran.

Facility Managers: Dr. Ivan Keresztes, Anthony M. Condo, David Kiemle.

Past mentors: Prof. Francis Johnson and Dr. James Bradley Dalton III.

Organizations: Boyce Thompson Institute for Plant Research and Cornell University, Ithaca, NY.

Funding: NIH R01AT008764

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE
Defensive Behaviors from *C. elegans* to Sulfolipids from *Pristionchus pacificus*

1.1 ABSTRACT

Animals respond to predators by altering their behavior and physiological states, but the underlying signaling mechanisms are poorly understood. Using the interactions between *Caenorhabditis elegans* and its predator, *Pristionchus pacificus*, we show that perception by *C. elegans* of a predator-specific molecular signature induces instantaneous escape behavior and a prolonged reduction in oviposition. Chemical analysis revealed this predator-specific signature to consist of a class of sulfolipids, produced by a biochemical pathway required for developing predacious behavior and specifically induced by starvation.

Zheng Liu[1]*, Maro J. Kariya[2]*, Christopher D. Chute[3]*, Sarah G. Leinwand[1], Ada Tong[1], Amy K. Pribadi[1,4], Kevin P. Curran[1], Neelanjan Bose[2], Frank C. Schroeder[2], Jagan Srinivasan[3] and Sreekanth H. Chalasani[1,4,5].

[1] Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA.

[2] Boyce Thompson Institute and Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853.

[3] Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA 01605, USA.

[4] Division of Biological Sciences, University of California, San Diego, La Jolla, CA 92093, USA.

* These authors contributed equally to this work.

1.2 INTRODUCTION

Animal survival depends on the ability to sense predators and generate appropriate behavioral and physiological changes[1]. Such defensive behaviors[2], including the commonly observed 'flight or freezing' responses, are often hardwired into the genome of the prey–for example, mice reliably exhibit fear-like responses to cat odors despite not having encountered cats for hundreds of generations[3]. Despite this, the neuronal and signaling machinery that regulate defensive behaviors remains poorly understood. One approach to uncovering the nature of innate defensive responses is to identify the molecular signals between predators and prey and map the underlying neuronal and molecular machinery that drive defensive responses to these signals. Studies from both vertebrates and invertebrates indicate that signaling between predators and prey involves multiple sensory modalities including vision, audition, and most frequently olfaction[4-6]. Considerable progress has been made in identifying the sensory neurons that detect predator-released odors in several model systems. For example, in mice, the chemosensory neurons in the vomeronasal organ (VNO), Grueneberg ganglion, and main olfactory epithelium have been shown to facilitate defensive behaviors through detection of signals from cat urine and fox feces[3,7,8]. These neurons project to higher brain regions where predator odor information is processed to generate stereotyped defensive behaviors[9]. More generally, it is thought that circuit constancy typically accompanies behavioral stereotypy. While neural circuits that detect

odors vary between individuals[10], those sensing predator-released odors appear to be invariant between members of the same species[3]. However, the precise identities of the participating neurons, their connections, and the nature of the circuit computations driving these invariant defensive behaviors have remained elusive.

We approached these questions by analyzing the behavioral responses of the nematode *Caenorhabditis elegans*[11] to a predatory nematode *Pristionchus pacificus*[12]. These two nematodes likely shared a common ancestor around 350 million years ago[13]. Recent studies have shown that *P. pacificus* is a facultative predator. *P. pacificus* can bite and kill *C. elegans*, a process facilitated by the extensive re-wiring of the *P. pacificus* nervous system under crowded and/or starvation conditions[14,15]. *C. elegans*, with its fully mapped neural network comprising of just 302 neurons connected by identified synapses and powerful genetic tools, is ideally suited for a molecular and circuit-level analysis of complex behaviors[16,17]. We found that a novel class of sulfated small molecules excreted by *P. pacificus* trigger defensive responses in *C. elegans*.

1.3 RESULTS

A predator elicits defense responses in *C. elegans*.

C. elegans was originally isolated from compost heaps in the developmentally arrested dauer stage[18]. However, recent studies have isolated proliferating and feeding populations of *C. elegans* from rotting flowers and fruits[19], where they are often found to cohabit with other nematodes including the Diplogastrid *Pristionchus* (M-A. Felix, personal communication). Previous reports have shown that the terrestrial nematode, *P. pacificus* can kill and consume the smaller nematode *C. elegans*[20]. We hypothesized that the prey, *C. elegans*, detects the predator, *P. pacificus*, through chemical cues and thus tested *C. elegans* responses to *P. pacificus* excretions. We found that *C. elegans* showed immediate avoidance upon perceiving excretions of starved, but not well-fed predators (Figure 1a and Supplementary Figure S1). Excretions from *P. pacificus* collected after 21 h of starvation ('predator cue') were found to consistently repel genetically diverse *C. elegans* isolates (Figure 1b). Next, we tested whether volatile components could be responsible for the activity of the predator cue by analyzing prey responses using a chemotaxis assay optimized for volatiles. We found that predator cue had no significant effect on *C. elegans* taxis responses in this assay (Figure 1c, d), indicating that volatiles do not contribute to the activity of predator cue. Together, these results show that starving *P. pacificus* release potent non-volatile repellent(s) that induce rapid *C. elegans* avoidance.

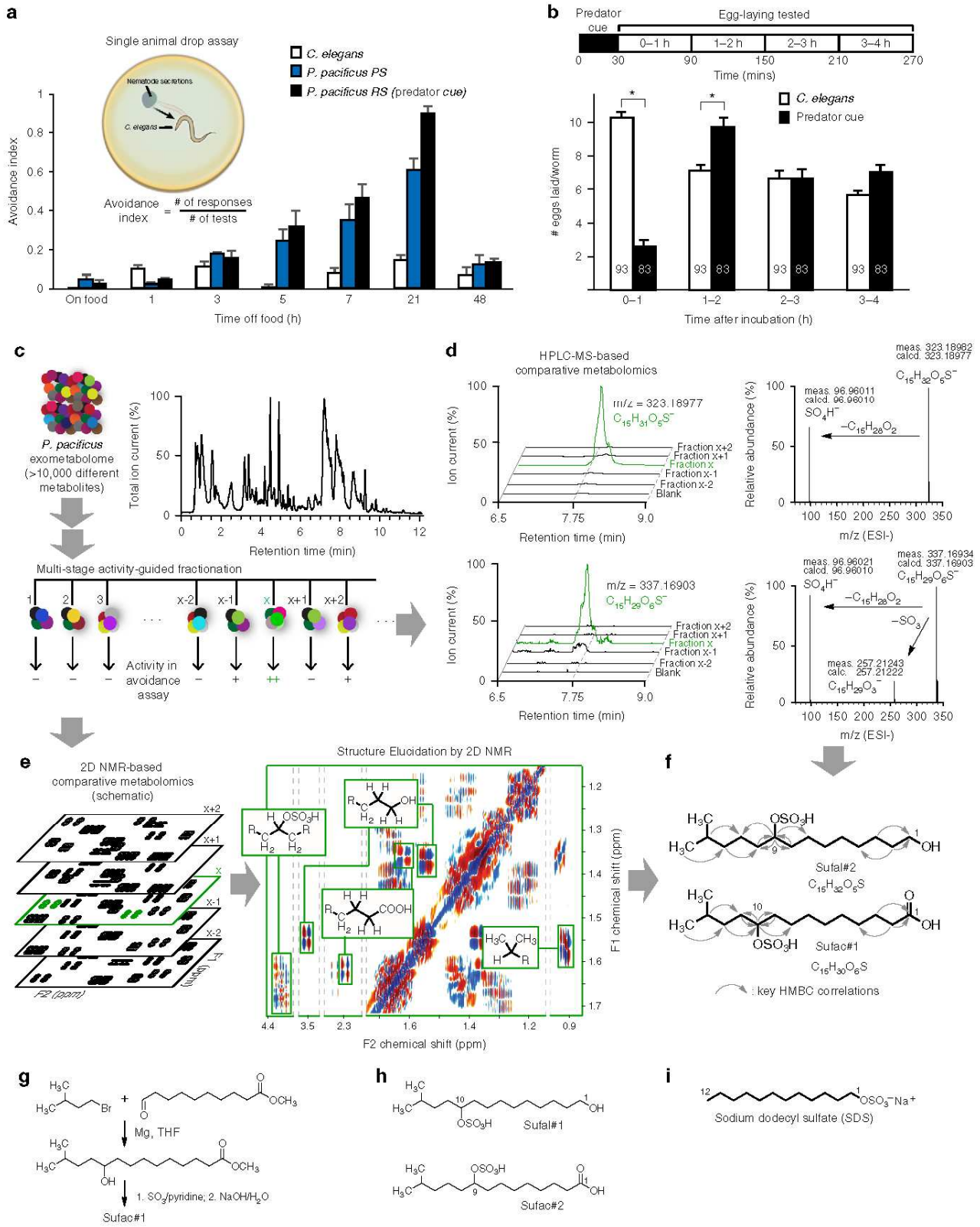

a
Single animal drop assay
Nematode secretions
C. elegans
Avoidance index = # of responses / # of tests
C. elegans
P. pacificus PS
P. pacificus RS (predator cue)
Avoidance index
Time off food (h)
b
Predator cue
Egg-laying tested
0–1 h
1–2 h
2–3 h
3–4 h
Time (mins)
C. elegans
Predator cue
eggs laid/worm
Time after incubation (h)
c
P. pacificus exometabolome (>10,000 different metabolites)
Total ion current (%)
Retention time (min)
Multi-stage activity-guided fractionation
Activity in avidance assay
d
HPLC-MS-based comparative metabolomics
m/z = 323.18977
Fraction x+2
Fraction x+1
Fraction x
Fraction x-1
Fraction x-2
Blank
Ion current (%)
Retention time (min)
meas. 323.18982
calcd. 323.18977
meas. 96.96011
calcd. 96.96010
Relative abundance (%)
m/z (ESI-)
m/z = 337.16903
meas. 337.16934
calcd. 337.16903
meas. 96.96021
meas. 257.21243
calc. 257.21222
e
2D NMR-based comparative metabolomics (schematic)
Structure Elucidation by 2D NMR
F1 chemical shift (ppm)
F2 chemical shift (ppm)
f
Sufal#2
Sufac#1
: key HMBC correlations
g
Mg, THF
1. SO3/pyridine; 2. NaOH/H2O
Sufac#1
h
Sufal#1
Sufac#2
i
Sodium dodecyl sulfate (SDS)

Figure 1 Predator-released sulfolipids drive *C. elegans* behaviors. a *C. elegans* avoid excretions from starving *P. pacificus* PS (PS312, domesticated) and RS (RS5725B a wild isolate, predator cue) strains. Inset shows a schematic of the avoidance assay. **b** Top shows schematic of the modified egg-laying assay and bottom, *C. elegans* lay fewer eggs after a 30 min exposure to concentrated predator cue, but recovers after 2 h. **c** UHPLC-HRMS analysis reveals a complex mixture of >10,000 metabolites, which was subjected to multistage activity-guided fractionation using reverse-phase chromatography. After four fractionation steps, most of the activity (++) was found in fraction x. Averages and s.e.m. are shown. n > 30 for each condition. **d** UHPLC-HRMS ion chromatograms (m/z value $\pm$ 5 ppm) of active fraction x and adjacent fractions for two sulfate-containing metabolites that were strongly enriched in the active fraction (left). MS‑MS analysis (right) confirms presence of sulfate moieties in both compounds. **e** Schematic representation of 2D NMR-based comparative metabolomics (left) of consecutive fractions (x − 2 to x+2) used to identify signals specific to fraction x. Cropped 2D NMR (dqfCOSY) spectrum (right) of active fraction highlighting signals that represent specific features of the identified metabolites (gray lines define edges of shown subsections). **f** Chemical structures of metabolites identified via comparative metabolomics from active fraction x, sufac#1 and sufal#2. Gray arrows indicate important correlations observed in heteronuclear 2D NMR (HMBC) spectra. **g** Synthesis of sufac#1; THF: tetrahydrofuran. **h** Homologous metabolites sufac#2 and sufal#1 were also detected by UHPLC-HRMS. **i** Chemical structure of sodium dodecyl sulfate (SDS). Averages and s.e.m. are shown and number of animals tested are indicated on each bar or condition. *p < 0.05 obtained by comparison with controls using Fisher's exact t-test with Bonferonni correction.

We further found that *C. elegans* exposed to predator cue did not lay eggs for many minutes following exposure, even when placed on food (bacterial lawn), suggesting that predator cue induced stress affects egg-laying behavior. Consistent with this idea, previous studies have shown that *C. elegans* retain eggs in the gonad when exposed to environmental stressors[21]. To test our hypothesis, we designed a behavioral assay wherein the prey was exposed to predator cue for 30 min, and egg-laying was monitored for many hours following cue removal. Animals exposed to predator cue laid significantly fewer eggs than controls during the initial 60 min following cue removal.

During the next hour (i.e., the 60–120 min post-cue time period), these animals laid more eggs than controls, suggesting that predator cue transiently modified egg-laying behavior, but not egg production (Figure 1b). Collectively, these results indicate that starving *P. pacificus* release a potent, non-volatile factor (predator cue) that elicits multiple prey responses, namely urgent escape behavior followed by up to one hour of reduced egg laying.

Predator-derived sulfolipids elicit defense responses. We aimed to identify the chemical structure(s) of the small molecule(s) excreted by *P. pacificus* that cause *C. elegans* avoidance behavior. As the *P. pacificus* exo-metabolome is highly complex, consisting of more than 20,000 distinct compounds detectable by UHPLC-HRMS (Figure 1c), we used a multistage activity-guided fractionation scheme (see Supplementary Methods). After three rounds of fractionation, chemical complexity of individual fractions had been reduced sufficiently to enable comparative metabolomics analysis of 2D NMR spectra[22,23] and high-resolution tandem mass spectrometry data of active and adjacent inactive fractions (Figure 1d, e). This analysis revealed several novel (ω-1)-branched-chain sulfolipids (sufac#1, and sufal#2) as major components of active, but not inactive fractions (Figure 1f). Further analysis revealed several additional sulfolipids with closely related structures, including sufac#2 and sufal#1, which differ in the position of attachment of the sulfate moiety (Figure 1h).

Next, we corroborated the structures of these compounds via total synthesis and tested their activity in the avoidance assay. The terminal alcohols sufal#1 and sufal#2 accounted for most of the isolated activity (Supplementary Figure S1, Supplementary methods). None of the identified *P. pacificus* sulfolipids could be detected in the metabolomes of Escherichia coli OP50 (used as food for nematodes) (Supplementary Figure S2), *C. elegans*, which were extracted and analyzed under identical conditions. Notably, the identified sulfolipids are structurally similar to sodium dodecyl sulfate (SDS, Figure 1i), which is a potent *C. elegans* avoidance cue[24].

1.4 DISCUSSION

We show that *P. pacificus* releases a mixture of sulfolipids that *C. elegans* perceives as a predator-specific molecular signature, or kairomone[25]. Perception of these sulfolipids initiates defensive responses including rapid avoidance. Sulfated fatty acids and related lipids have previously been described primarily from marine sources, including tunicates[26], sponges[27], crustaceans[28], and algae[29]. In addition, a family of sulfated fatty acids, the caeliferins, has been identified from grasshopper oral secretions[30,31]. In a striking parallel to the role of sulfated lipids in the nematode predator-prey system studied here, these herbivore-associated sulfolipids have been shown to elicit specific defense responses in plants[31]. Furthermore, the sulfolipids we identified from *P. pacificus* resemble sodium dodecyl sulfate (SDS), a known nematode repellent[24]. The sulfolipids we identified from *P. pacificus*, sufac#1 and sufal#2, and

several related compounds, appear to be derived from the monomethyl branched-chain fatty acid (mmBCFA), C15ISO, which is also produced by *C. elegans* and has been shown to be essential for *C. elegans* growth and development[32]. The biosynthesis of C15ISO in *C. elegans* requires the fatty acid elongase ELO-5, and several homologous elongases in *P. pacificus* exist that may be involved in the biosynthesis of the fatty acid precursors of sufac#1 and sufal#2. Additionally, the biosynthesis of sufac#1 and sufal#2 requires oxygenation at the (ω-5) or (ω-6) position in the fatty acid chain, respectively, followed by sulfation by sulfotransferase(s), a family of genes that has undergone major expansion in *P. pacificus*[33]. Notably, at least one sulfotransferase, EUD-1, functions as a central switch determining whether *P. pacificus* larvae will develop into a primarily bacterivorous, narrow-mouthed adult, or into a predacious, wide-mouthed adult that can feed on other nematodes[34]. It is intriguing that *C. elegans* has evolved the ability to detect a *Pristionchus*-specific trait (the extensive sulfation of small molecules) that is directly connected to the endocrine signaling pathway that controls development of the morphological features required for predation.

1.5 METHODS

P. pacificus predator cue

Nematodes (*C. elegans*, and *P. pacificus* PS312 or RS5275B) were cultured on 10 cm wide NGM plates seeded with lawns of OP50 *E. coli* (OD_{600} = 0.5). Animals from 40 such plates were harvested just before the bacteria were completely depleted (6 days), and washed 5 times with M9 buffer. Animals from these plates were pooled into microfuge tubes with approximately 100 μ l of worms and the secretions were collected at the starvation times indicated in about 100 μ l of M9 buffer. Predator Cue (RS5275B secretions) were tested and diluted in M9 buffer for further experiments. *C. elegans* secretions (N2 cue) were used in Figure 1. To further concentrate the predator cue for egg-laying assays, we passed the secretions through a Microcon 3 K Centrifugal Filter Column (Millipore), reducing the final volume to 1/10 of the initial volume. This concentrated predator cue was then used in behavioral assays.

Single animal avoidance assay

Adult *C. elegans* were exposed to a small volume of a test compound (0.05 μ l) near the head of the animal. Upon sensing a repellent, animals initiate a reversal followed by an omega bend. Positive responses to a test compound were scored only if they occurred within 4 s. Results are shown as the avoidance index, which is the ratio of

number of positive responses to the total number of trials. Each animal is tested thrice and data is presented as average of the animals tested on at least 3 days (Figure 1).

Nematode-derived modular metabolite (NDMM) nomenclature

Nematode metabolites are named using Small Molecule IDentifiers ("SMIDs"), representing searchable, gene-style identifiers that consist of four lower case non-italicized letters followed by a pound sign and a number. The SMID database (www.smid-db.org) is an electronic resource maintained by Profs. Frank C. Schroeder and Lukas Mueller at the Boyce Thompson Institute/Cornell University, in collaboration with Prof. Paul Sternberg at Caltech and WormBase (www.wormbase.org). This database catalogs newly identified nematode small molecules, assigns a unique four-letter SMID (a searchable gene-style identifier), and for each compound includes a list of other names and abbreviations used in the literature.

Instrumentation for chemical analyses

NMR spectra were recorded on a Bruker AVANCE III HD (800 MHz) and Varian INOVA-600 (600 MHz) instruments. UHPLC-high-resolution mass spectrometry (HRMS) was performed using a Thermo Scientific Dionex Ultimate 3000 UHPLC system equipped with an Agilent ZORBAX Eclipse XDB C18 column, connected to a Thermo Scientific Q

Exactive HF Hybrid Quadrupole-Orbitrap mass spectrometer. HPLC-MS and –MS/MS was performed using an Agilent 1100 Series HPLC system equipped with a diode array detector and an Agilent Eclipse XDB-C18 column (4.6 × 250 mm, 5 μ m particle diameter), connected to a Quattro II spectrometer (Micromass/Waters). Flash chromatography was performed using a Teledyne ISCO CombiFlash system. Preparative HPLC separation was performed using the Agilent 1100 Series HPLC system equipped with an Agilent Eclipse XDB-C18 or C8 column (9.4 × 250 mm, 5 μ m particle diameter) coupled to a Teledyne ISCO Foxy 200 fraction collector.

HPLC-MS/MS analyses

A 0.1% acetic acid water–acetonitrile solvent gradient was used at a flow rate of 1 ml/min, starting with an acetonitrile content of 5% for 5 min and increasing to 100% over a period of 40 min. Exo-metabolome fractions were analyzed by HPLC-ESI-MS in negative and positive ion modes using a capillary voltage of 3.5 kV and a cone voltage of −35 V and +20 V, respectively. The analytical HPLC protocol mentioned above was translated to a semi-preparative Agilent Eclipse XDB-C18 or C8 column (9.4 × 250 mm, 5 μm particle diameter) with a flow rate of 3.6 ml/min and used for MS-assisted enrichment of desired metabolites, as well as for synthetic sample purification. Data acquisition and processing for the HPLC-MS was controlled by Waters MassLynx software.

UHPLC-HRMS analyses

A 0.1% formic acid water −0.1% formic acid acetonitrile solvent gradient was used at a flow rate of 0.500 ml/min, starting with an acetonitrile content of 5% for 1.9 min, increasing to 100% over a period of 11 min, and then returning to 5% for 2 min. Data acquisition and processing for the UHPLC-HRMS was controlled by Thermo Scientific Xcalibur software.

P. pacificus strains and culture conditions

RS2333 was used for exo-metabolome preparation. Mixed stage worms from a populated 10 cm NGM agar plate seeded with *E. coli* OP50 were washed into 25 ml of S-complete medium and fed OP50 on days 1, 3 and 5 for a 7-day culture period, while shaking at 22 ° C, 220 r.p.m. The cultures were then centrifuged and worm pellets and supernatant frozen separately. For axenic cultures, *P. pacificus* (RS2333) gravid adults from ten 10 cm plates were washed with M9 buffer and treated with alkaline hypochlorite solution to isolate eggs. Isolated eggs were washed thoroughly with M9 buffer and allowed to hatch in fresh sterile M9 for 24 h. The M9 supernatant was prepared as described below.

Preparation of exo-metabolome extracts and fractionation

Protocol A

P. pacificus RS2333 liquid culture supernatant (3 l) was lyophilized to a fine powder and extracted with 750 ml of a 95:5 mixture of ethanol and water for 16 h (2 times). The exo-metabolome extract was then concentrated in vacuo, loaded onto 12 g of ethyl acetate-washed Celite and fractionated using a Teledyne ISCO CombiFlash system and a RediSep GOLD 30 g HP C18 reverse-phase column using a water-methanol solvent gradient, starting with 15 min of 98% water, followed by a linear increase of methanol content up to 100% at 60 min. The eluate was divided into 8 fractions, which were evaporated in vacuo and prepared for HPLC-MS/MS and NMR spectroscopic analyses. Subsequently a subset of fractions was further fractionated by HPLC. A 0.1% acetic acid water‑acetonitrile solvent gradient was used at a flow rate of 3.6 ml/min, starting with an acetonitrile content of 5% for 5 min and increasing to 100% over a period of 40 min.

Protocol B

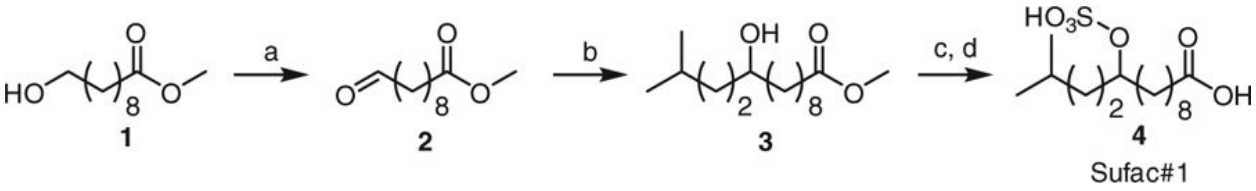
P. pacificus RS2333 liquid culture supernatant (3L) was lyophilized to a fine powder and extracted with 3 l of a 95:5 mixture of ethanol and water for 16 h with stirring. The exo-metabolome extract was then concentrated in vacuo, loaded onto 8 g of ethyl acetate-washed Celite and fractionated using a Teledyne ISCO CombiFlash

system over a RediSep and GOLD 100 g HP C18 reverse-phase column using a water (0.1% acetic acid)-acetonitrile solvent gradient, starting with 10 min of 100% water, followed by a linear increase of methanol content up to 100% at 92 min, which was maintained up to 100.7 min, thereby producing ~240 fractions, which were prepared for analysis by UHPLC-HRMS.

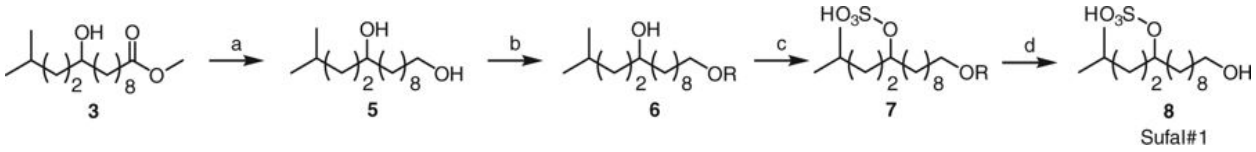
Chemical syntheses

General methods for chemical synthesis

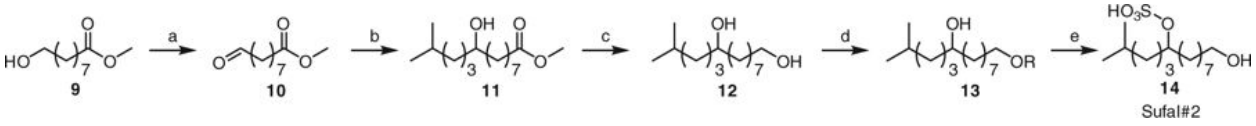
Thin-layer chromatography (TLC) was used to monitor progress of reactions unless stated otherwise, using J. T. Baker Silica Gel IB2-F. Unless stated otherwise, reagents were purchased from Sigma-Aldrich and used without further purification. *N*,*N*-dimethylformamide (DMF) and dichloromethane (DCM) were dried over 4 Å molecular sieves prior to use. Tetrahydrofuran (THF) was distilled over lithium aluminum hydride prior to use. Optical rotations were measured on a Perkin Elmer 341 polarimeter. Synthetic schemes are shown below.



Supplementary Scheme 1. Overview of synthesis of sufac#1. Reagents and conditions: **a** PCC, DCM, 0 ° C; **b** Mg, THF, argon; **c** SO_3/pyridine complex, pyridine; **d** NaOH, H_2O



Supplementary Scheme 2. Overview of synthesis of sufal#1. Reagents and conditions: **a** $LiAlH_4$, THF; **b** *t*-butyldiphenylsilyl chloride (TBDPS-Cl), imidazole, THF; **c** SO_3/pyridine complex, pyridine; **d** acetyl chloride, MeOH



Supplementary Scheme 3. Overview of synthesis of sufal#2. Reagents and conditions: **a** Pyridinium chlorochromate (PCC), DCM, 0 ° C; **b** Mg, THF, argon; **c** $LiAlH_4$, THF; **d** TBDPS-Cl, imidazole, THF; **e** SO_3/pyridine complex, pyridine

Preparation of **2** (methyl 10-oxo-decanoate)

1 (2.5 mL, 13.5 mmol) was added to a solution of pyridinium chlorochromate (2.91 g, 13.51 mmol) and DCM (50 ml) at 0 ° C. After stirring for 20 min, the reaction was quenched with ether (30 ml), filtered over Celite, and washed further with ether (20 ml). The filtrate was concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% ethyl acetate in hexanes afforded **2** at 52% purity (48% dimer) (226.12 mg, 1.13 mmol, 8%). ^{1}H NMR (400 MHz, chloroform-*d*): δ (p.p.m.) 9.76 (t, *J* = 2.0 Hz, 1 H), 3.56 (s, 3 H), 2.42 (td, *J* = 7.4, 1.6 Hz, 2 H), 2.30 (t, 7.6 Hz, 2 H), 1.55‑1.48 (m, 4 H), 1.25‑1.19 (m, 10 H).

Preparation of **3** (methyl 10-hydroxy-13-methyltetradecanoate)

Ethylene dibromide (2 drops, 0.58 mmol) was added to Mg turnings (729 mg, 30 mmol) in THF (5 ml) under argon and the mixture was heated briefly until reflux. After cooling to room temperature, a solution of 3-methylbromobutane (1.2 ml, 10 mmol) in THF (10 ml) was added dropwise to the reaction over 10 min with stirring. The reaction was stirred at reflux for 1 h, affording isopentylmagnesium bromide. The isopentylmagnesium bromide solution (100 µl, 0.9 mmol) was then added to a solution of **2** (175 mg, 874 µmol) in THF (1 ml) at −25 ° C. The reaction was allowed to return to room temperature while stirring, at which point the reaction was quenched with saturated NH_4Cl (10 ml) and extracted with hexanes (10 ml). Flash column

chromatography on silica using a gradient of 0‑100% ethyl acetate in hexanes afforded **3** at 70% purity (90.6 mg, 333 µmol, 38%). ^{1}H NMR (600 MHz, chloroform-*d*): δ (p.p.m.) 3.66 (s, 3 H), 3.55 (m, 1 H), 2.30 (t, 7.7 Hz, 2 H), 1.64‑1.32 (m, 9 H), 1.20 (m, 1 H), 1.30 (m, 10 H), 0.90 (d, 6.6 Hz, 3 H), 0.89 (d, 6.6 Hz, 3 H).

Preparation of **4** (13-methyl-10-(sulfooxy)tetradecanoic acid) (sufac#1)

Sulfur trioxide-pyridine complex (120 mg, 0.75 mmol) was added to a solution of **3** (70 mg, 257 µmol) and pyridine (1 ml). After stirring for 5 min, the solution was concentrated in vacuo and dissolved in NaOH (aq) (4 ml, 2.75 M). The reaction was then neutralized with acetic acid/TFA to a pH of 4 and concentrated in vacuo over Celite. Flash column chromatography on silica using a gradient of 0‑100% methanol in DCM afforded **4** (9.6 mg, 30 µmol, 12%). Spectroscopic data were in agreement with those for the natural product (see Supplementary Tables S2 and S3).

Preparation of **5** (13-methyltetradecane-1,10-diol)

3 (16.9 mg, 63.0 µmol) was added dropwise to a suspension of lithium aluminum hydride (20.14 mg, 530 µmol, 8.4 eq) in THF (20 mL) at 0° C. The reaction was warmed slowly to room temperature while stirring for 20 min. The reaction was then quenched by the Feiser method, filtered, and concentrated in vacuo over celite. Flash column chromatography on silica using a gradient of 0 -100% ethyl acetate in hexanes afforded 5 (13.8 mg, 60.0 µmol, 95%).

1H NMR (400 MHz, methanol-d4): δ (ppm) 3.54 (t, 6.50 Hz, 2H), 3.47 (m, 1H), 1.56-1.36 (m, 10H), 1.32 (m, 10H), 1.20 (m, 1H), 0.91 (d, 6.66 Hz, 3H), 0.90 (d, 6.59 Hz, 3H).

Preparation of **6** (14-((*t*-butyldiphenylsilyl)oxy)-2-methyltetradecan-5-ol)

TBDPSCl (30 µL, 0.11 mmol, ~3 eq) was added incrementally over the course of 24 h to a solution of **5** (10.2 mg, 40 µmol) and imidazole (3.51 mg, 0.13 mmol, ~3 eq) in THF (2 ml). The reaction was concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% ethyl acetate in hexanes afforded **6** at 54% purity (6.04 mg, 12.96 µmol, 32.4%). ^{1}H NMR (600 MHz, methanol-d_4): δ (p.p.m.) 7.72 (m, 4 H), 7.36 (m, 6 H), 3.66 (t, 6.4, 2 H), 3.48 (m, 1 H), 1.57‑1.25 (m, 20 H), 1.2 (m, 1 H), 1.03 (s, 9 H), 0.90 (d, 6.6 Hz, 3 H), 0.89 (d, 6.6 Hz, 3 H).

Preparation of **7** (14-((*t*-butyldiphenylsilyl)oxy)-2-methyltetradecan-5-yl hydrogen sulfate)

Sulfur trioxide-pyridine complex (excess) was added to a solution of **6**(6.04 mg, 12.96 µmol) in pyridine (1 ml) and stirred for 5 min at room temperature. The reaction was concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% methanol in DCM afforded **7** at 20% purity (3.3 mg, 6 µmol, 46% yield). ^{1}H NMR (600 MHz, methanol-d_4): δ (p.p.m.) 7.72 (m, 4 H), 7.36 (m, 6 H), 4.31 (quin, 6.08 Hz, 1 H), 3.66 (t, 6.4 Hz, 2 H), 1.70‑1.60 (m, 4 H), 1.6‑1.5 (m, 3 H), 1.47‑1.22 (m, 14 H), 1.03 (s, 9 H), 0.90 (d, 6.6 Hz, 3 H), 0.89 (d, 6.6 Hz, 3 H).

Preparation of **8** (14-hydroxy-2-methyltetradecan-5-yl hydrogen sulfate) (sufal#1)

Methanol (500 µl) with a catalytic amount of acetyl chloride (0.5 µl) was added to a solution of **7** (1.3 mg, 2.3 µmol). The solution was concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% methanol in DCM afforded **8** (648 µg, 2.0 µmol, 87% yield). Spectroscopic data were in agreement with those for the natural product: ^{1}H NMR (600 MHz, methanol-d_4): δ (p.p.m.) 4.31 (quin, 6.1 Hz, 1 H), 3.54 (t, 6.8 Hz, 2 H), 1.7‑1.58 (m, 5 H), 1.57‑1.49 (m, 3 H), 1.47‑1.23 (m, 13 H), 0.91 (d, 6.6 Hz, 3 H), 0.90 (d, 6.6 Hz, 3 H); ^{13}C NMR (800 MHz, methanol-d_4): δ (p.p.m.) 62.76, 33.40, 26.64, 30.28, 30.xx, 30.xx, 30.xx, 25.75, 35.00, 80.85, 32.91, 34.87, 28.98, 22.68, 22.68

Preparation of **10** (methyl-9-oxo-nonanoate)

9 (600 µl, 2.8 mmol) was added to a solution of PCC (900 mg, 4.2 mmol) in dry DCM (50 ml) under argon at −15 $^{\circ}$ C, and the mixture was stirred for 45 min. Reaction progress was monitored by TLC. The reaction mixture was allowed to return to room temperature and was stirred for an additional 40 min. The mixture was then filtered over Celite, and the residue was washed with ether. The filtrate was concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% ethyl acetate in hexanes afforded **10** at 70% purity (containing about 30% starting material) (305 mg, 1.64 mmol, 58% yield). ^{1}H NMR (600 MHz, chloroform-*d*): δ (p.p.m.) 9.76 (t,

1.79 Hz, 1 H), 3.66 (s, 3 H), 2.42 (td, 7.30, 1.77 Hz, 2 H), 2.30 (t, 7.60 Hz, 2 H), 1.62 (m, 4 H), 1.32 (m, 6 H).

Preparation of **11** (methyl 9-hydroxy-13-methyltetradecanoate)

Ethylene dibromide (6 drops, 1.74 mmol) were added to Mg turnings (500 mg, 20.6 mmol, ~2 eq) in THF (5 ml) under argon. And the mixture was heated to reflux for 20 min. After cooling to room temperature, 4-methylbromopentane (1.38 ml, 10 mmol) in THF (10 ml) was added quickly to the solution of activated Mg turnings. The mixture was then refluxed for 1 h. After cooling to room temperature, the Grignard reagent (1.36 ml, 0.9 mmol, 1 eq) was added to a solution of **10** (170 mg, 0.9 mmol) in THF (1 ml). The reaction was monitored by TLC using 1:4 acetone:hexanes. The reaction was quenched with saturated NH_4Cl (10 ml) and extracted with hexanes. Flash column chromatography on silica using a gradient of 0–100% ethyl acetate in hexanes afforded **11** at 60% purity (81.7 mg, 0.3 mmol, 32%). ^{1}H NMR (400 MHz, chloroform-*d*): δ (p.p.m.) 3.66 (s, 3 H), 3.58 (m, 1 H), 2.30 (t, 7.58 Hz, 2 H), 1.65–1.49 (m, 4 H), 1.47–1.36 (m, 4 H), 1.31 (m, 9 H), 1.16 (m, 1 H), 0.88 (d, 6.64 Hz, 3 H), 0.88 (d, 6.64 Hz, 3 H).

Preparation of **12** (13-methyltetradecane-1,9-diol)

11 (81.7 mg, 300 µmol) was added dropwise to a suspension of lithium aluminum hydride (102.3 mg, 2.69 mmol) in THF (20 ml) at 0 ° C. The solution was warmed

slowly to room temperature while stirring for 20 min. The reaction was quenched by the Fieser method as described above, filtered over Celite, and concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% ethyl acetate in hexanes afforded **12** at 90% purity (57.0 mg, 234 µmol, 77.4%). ^{1}H NMR (400 MHz, methanol-d_4): δ (p.p.m.) 3.54 (t, 6.53 Hz, 2 H), 3.50 (m, 1 H), 1.61‑1.13 (m, 21 H), 0.89 (d, 6.6 Hz, 3 H), 0.89 (d, 6.6 Hz, 3 H).

Preparation of **13** (14-((*tert*-butyldiphenylsilyl)oxy)-2-methyltetradecan-6-ol)

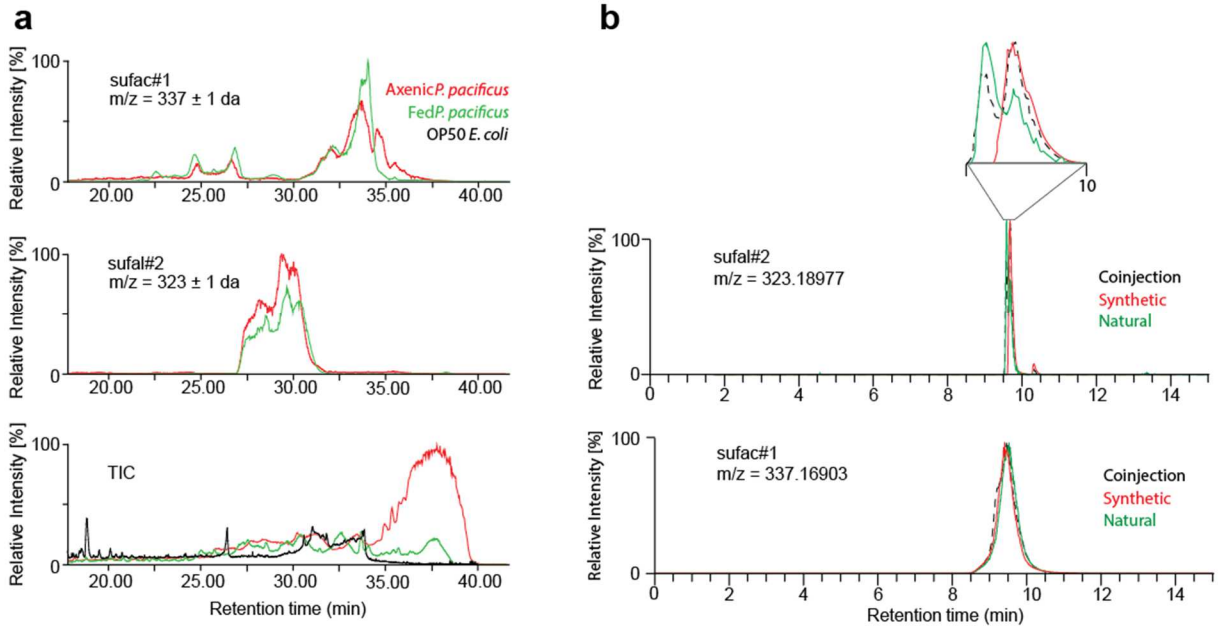
TBDPS-Cl (97.6 µl, 370 µmol) was added incrementally over the course of 3 h to a solution of **12** (57.0 mg, 234 µmol) in THF (1 ml) and DMF (0.5 ml). The reaction was concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% ethyl acetate in hexanes afforded **13** at 30% purity (13.1 mg, 23.4 µmol, 10%). ^{1}H NMR (400 MHz, methanol-d_4): δ (p.p.m.) 7.67‑7.64 (m, 4 H), 7.41‑7.35 (m, 6 H), 3.66 (t, 6.1 Hz, 2 H), 3.50 (m, 1 H), 1.61‑1.13 (m, 21 H), 1.03 (s, 9 H), 0.89 (d, 6.6 Hz, 3 H), 0.89 (d, 6.6 Hz, 3 H).

Preparation of **14** (14-hydroxy-2-methyltetradecan-6-yl hydrogen sulfate) (sufal#2)

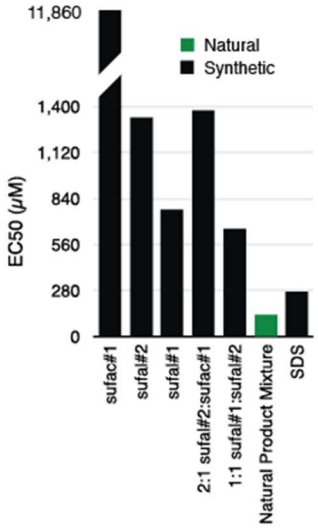
Sulfur trioxide-pyridine complex (50 mg, excess) was added to a solution of **13** (20 µmol) in acetonitrile (2 ml). The solution was concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% methanol in DCM

afforded **14** (0.65 mg, 2 µmol, 10%). Spectroscopic data were in agreement with those for the natural product (see Supplementary Tables S1 and S3).

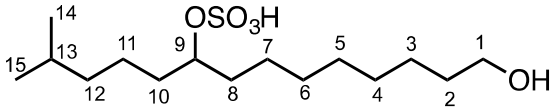
1.6 SUPPLEMENTARY FIGURES



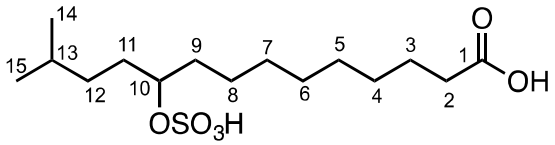
Supplementary Figure S1. Properties of *P. pacificus* sulfates. a. HPLC-MS analysis of exo- metabolome extracts from *P. pacificus* under fed (green) and axenic (red) conditions and an extract from OP50 *E. coli* (black). Traces for sufac#1 and sufal#2 are shown in the upper panels and the total ion chromatogram (TIC) in the lower panel (all intensities are normalized to pasc#9 intensity). Both fed and axenic cultures produced sufac#1 and sufal#2. **b**. UHPLC-HRMS analysis of natural (green), synthetic (red), and natural/synthetic-coinjected sufac#1 and sufal#2 (black-dotted).



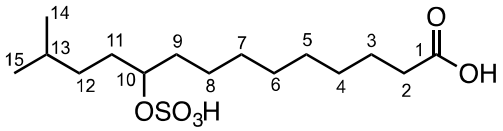
Supplementary Figure S2: *C. elegans* responses to sulfates. EC50 concentrations for *C. elegans* avoidance responses to synthetic sulfates and sulfate mixtures, compared to avoidance responses to a mixture of sulfates isolated from *P. pacificus* exo-metabolome.



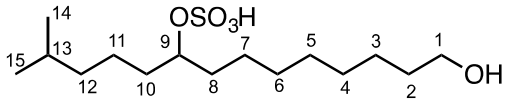
Position	^{13}C [ppm]	^{1}H [ppm]	^{1}H-^{1}H coupling constants (Hz)	HMBC correlations
1	62.76	1a = 3.53 1b = 3.53	$J_{1,2}$ = 6.80	2, 3
2	33.40	2a =1.53 2b = 1.53		1, 3, 4
3	26.64	3a = 1.34 3b = 1.34		1, 2, 4, 5
4	30.28	4a = 1.32 4b = 1.32		2, 3, 5, 6
5	30.32/ 30.29	5a =1.32-1.35 5b =1.32-1.35		-
6	30.29/ 30.32	6a =1.32-1.35 6b =1.32-1.35		-
7	25.75	7a = 1.41 7b = 1.41		8, 9
8	35.02	8a = 1.62 8b = 1.65		7, 9, 10
9	80.70	4.33	$J_{9,\ 8a\ 8b\ 10a\ 10b}$ = 6.13	7, 8, 10, 11
10	35.27	10a = 1.59 10b = 1.64		8, 9, 11, 12
11	23.59	11a = 1.39 11b = 1.43		9, 10, 12, 13,
12	39.94	12a = 1.18 12b = 1.22		10, 11, 13, 14
13	28.82	1.55		11, 12, 14, 15
14	22.74	0.89	$J_{14,13}$ = 6.36	12, 13, 15
15	22.74	0.89	$J_{15,13}$ = 6.36	12, 13, 14



Position	^{13}C [ppm]	^{1}H [ppm]	^{1}H-^{1}H coupling constants (Hz)	HMBC correlations
1	177.40	---		-
2	34.70	2a = 2.28 2b = 2.28	$J_{2,3}$ = 7.55	1, 3, 4
3	25.83	3a = 1.60 3b = 1.60		1, 2, 4
4	29.92	4a = 1.34 4b = 1.34		-
5	30.09	5a = 1.35 5b = 1.35		3, 4
6	30.22/30.37	6a = 1.32-1.35 6b = 1.32-1.35		-
7	30.37/30.22	7a = 1.32-1.35 7b = 1.32-1.35		-
8	25.75	8a = 1.39 8b = 1.39		-
9	35	9a = 1.63 9b = 1.65		8, 10, 11
10	80.85	4.31	$J_{10,\ 9a\ 9b\ 11a\ 11b}$ = 6.15	8, 9, 11, 12
11	32.91	11a = 1.63 11b = 1.67		9, 10, 12, 13
12	34.87	12a = 1.26 12b = 1.31		10, 11, 13, 14, 15
13	28.98	1.54		11, 12, 14, 15
14	22.68	0.90	$J_{14,13}$ = 6.59	12, 13, 15
15	22.68	0.90	$J_{15,13}$ = 6.59	12, 13, 14



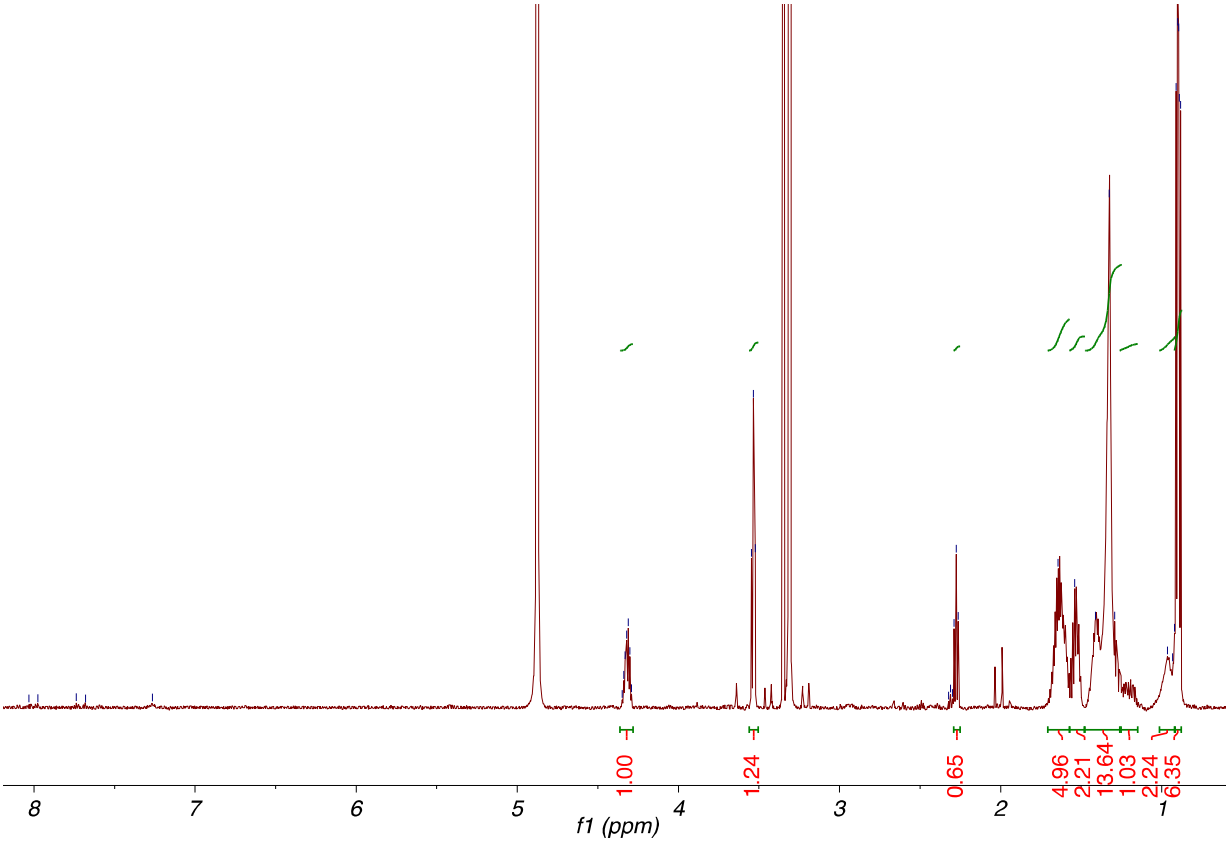
sufac#1			
Position	^{13}C Natural [ppm]	^{13}C Synthetic [ppm]	Δppm
1	177.40	178.46	-1.06
2	34.70	35.10	-0.40
3	25.83	26.31	-0.48
4	29.92	30.54	-0.62
5	30.09	30.73	-0.64
6	n/a	30.39	n/a
7	n/a	30.29	n/a
8	25.75	25.99	-0.24
9	35.00	35.56	-0.56
10	80.85	81.13	-0.28
11	32.91	33.09	-0.18
12	34.87	35.23	-0.36
13	28.98	29.26	-0.28
14	22.68	23.04	-0.36
15	22.68	22.96	-0.28



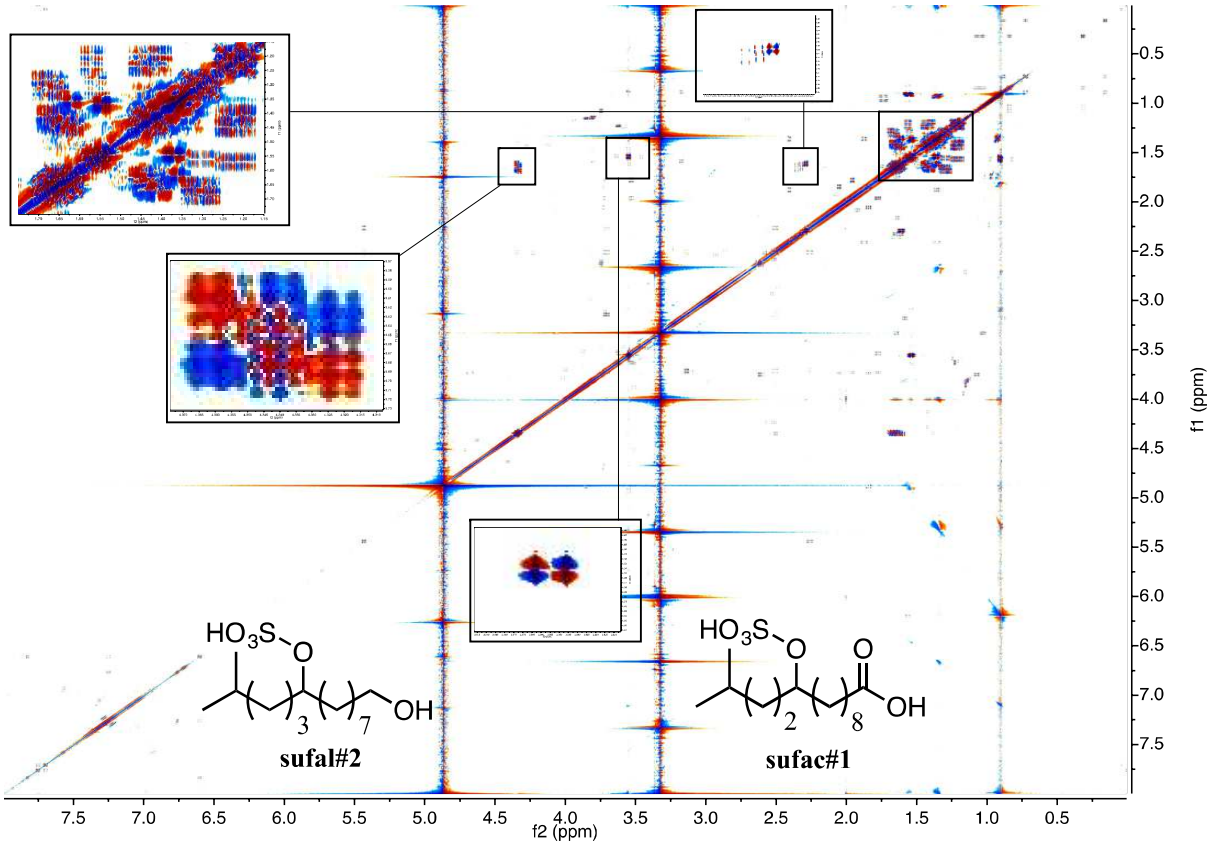
sufal#2			
Position	^{13}C Natural [ppm]	^{13}C Synthetic [ppm]	Δppm
1	62.76	63.03	-0.27
2	33.40	33.69	-0.29
3	26.64	26.93	-0.29
4	30.28	30.54	-0.26
5	n/a	30.74	n/a
6	n/a	30.67	n/a
7	25.75	26.03	-0.28
8	35.02	35.30	-0.28
9	80.70	81.04	-0.34
10	35.27	35.55	-0.28
11	23.59	23.87	-0.28
12	39.94	40.23	-0.29
13	28.82	29.11	-0.29
14	22.74	23.03	-0.29
15	22.74	23.03	-0.29

1.8 SUPPLEMENTARY NMR SPECTRA

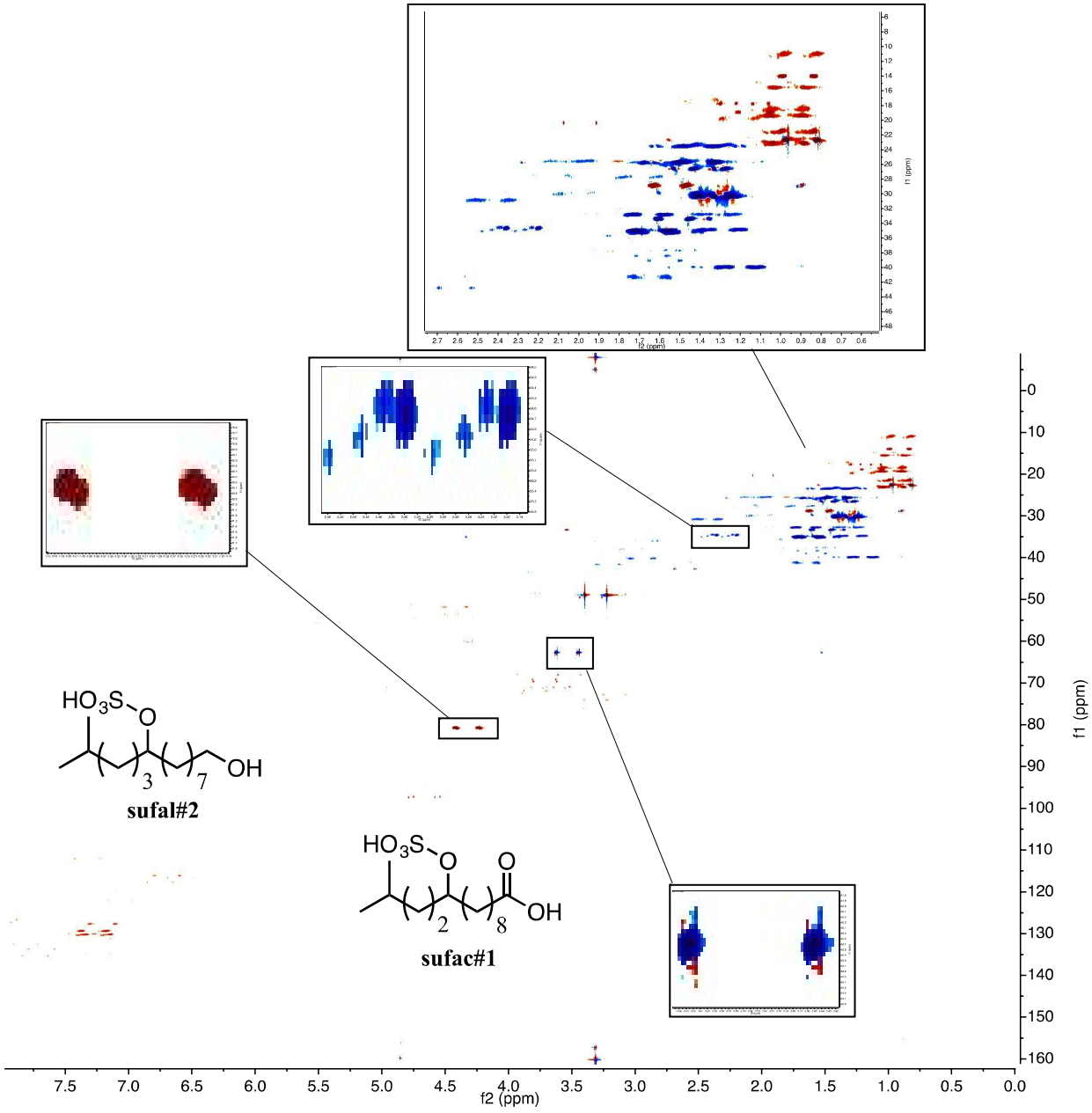
^{1}H NMR spectrum of active metabolome fraction containing sufac#1 and sufal#2 (600 MHz, methanol-d_4).



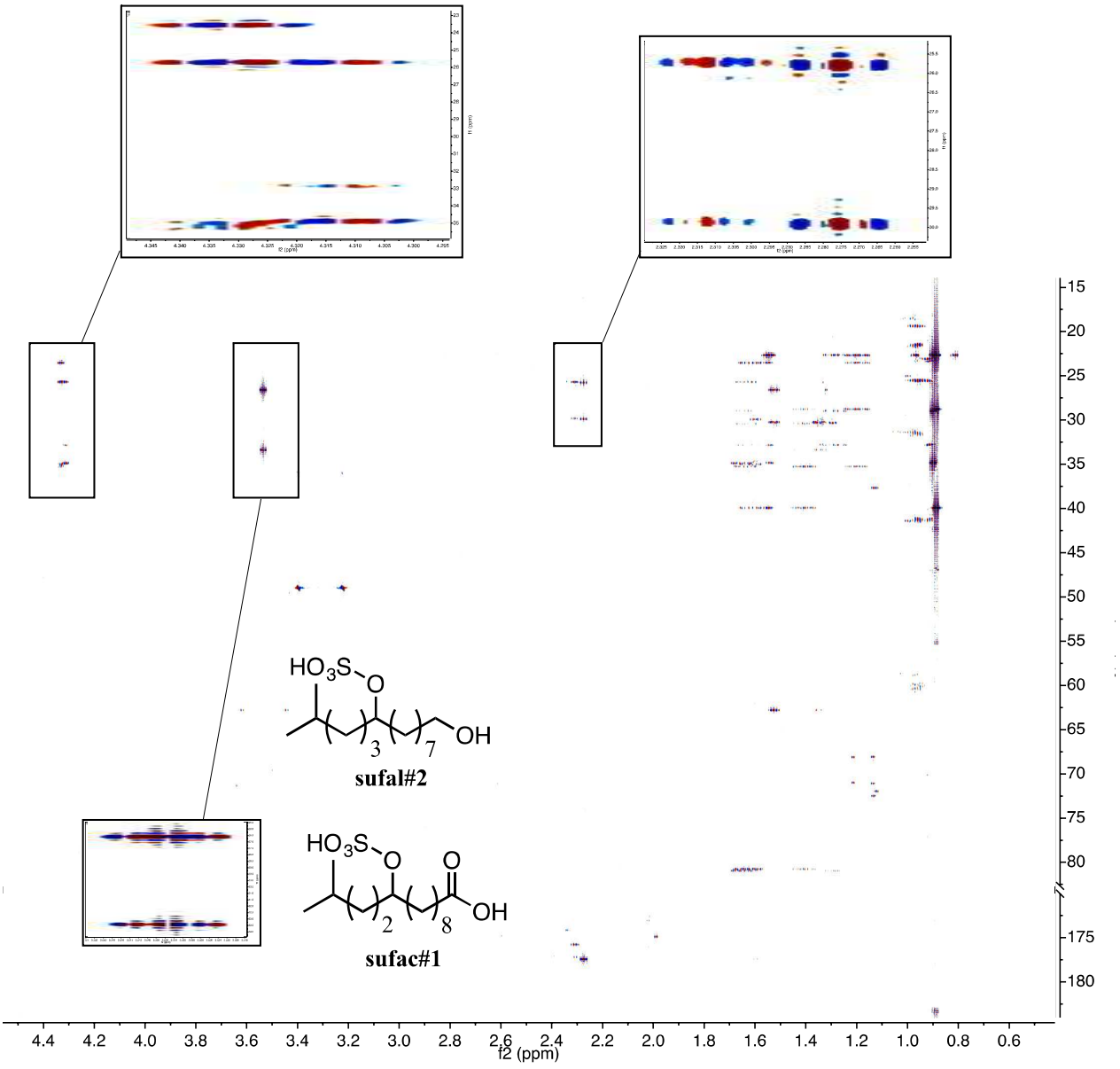
dqfCOSY spectrum of active metabolome fraction containing sufac#1 and sufal#2 (600 MHz, methanol-d_4).



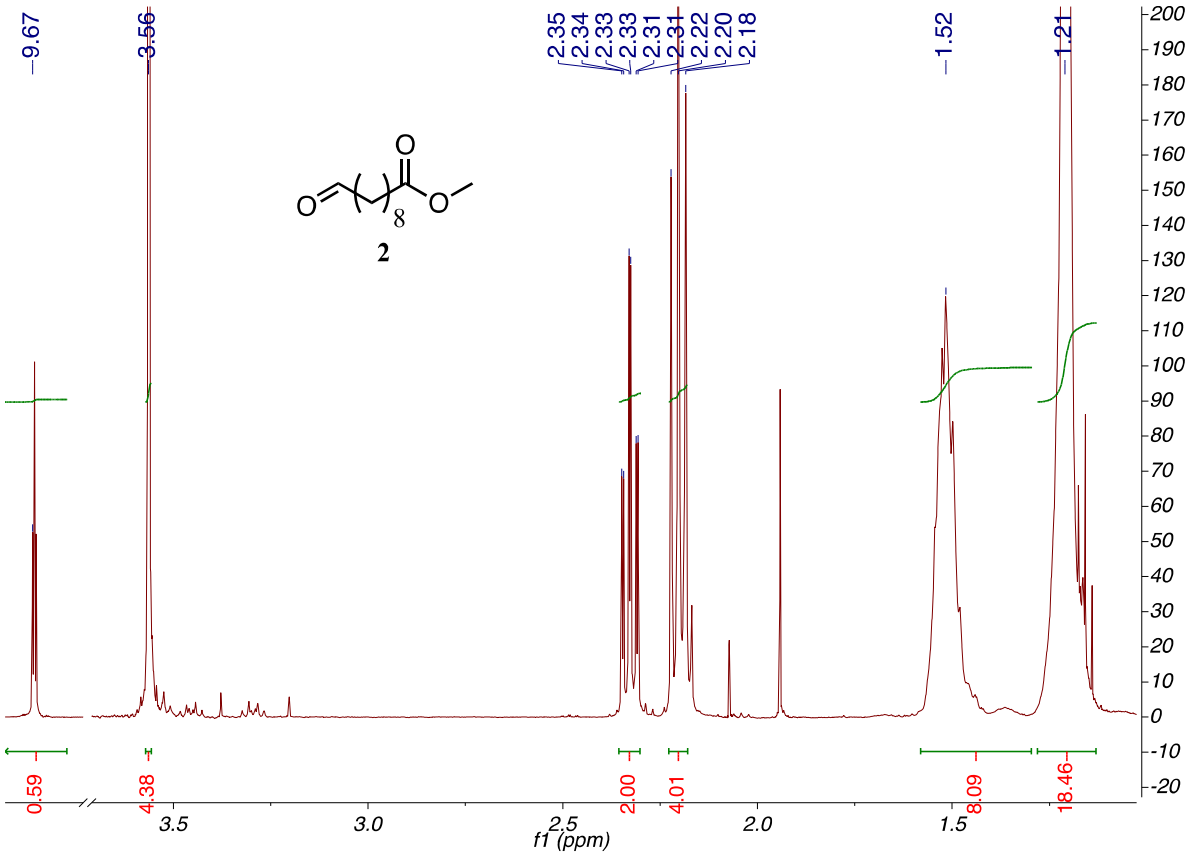
Coupled HSQC spectrum of active metabolome fraction containing sufac#1 and sufal#2 (600 MHz, methanol-d_4).



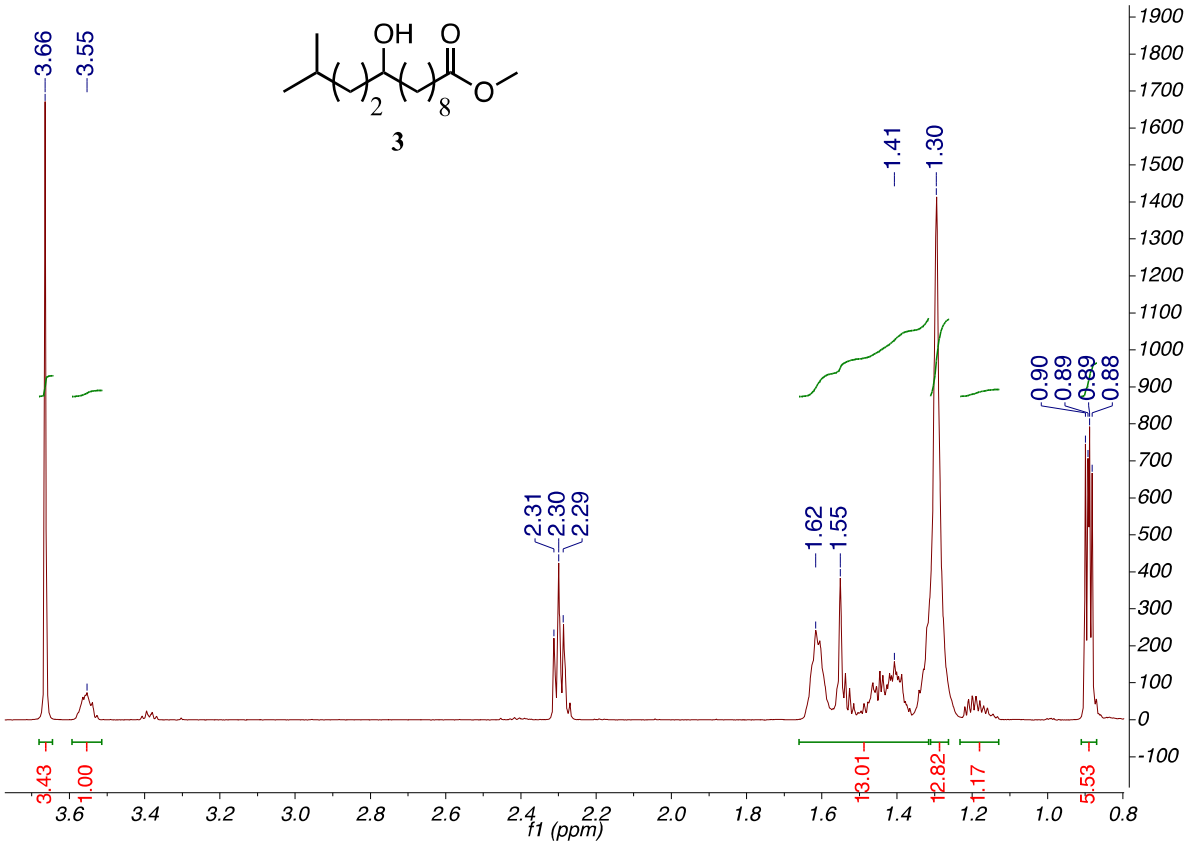
Constant-time HMBC spectrum of active metabolome fraction containing sufac#1 and sufal#2 (600 MHz, methanol-d_4).



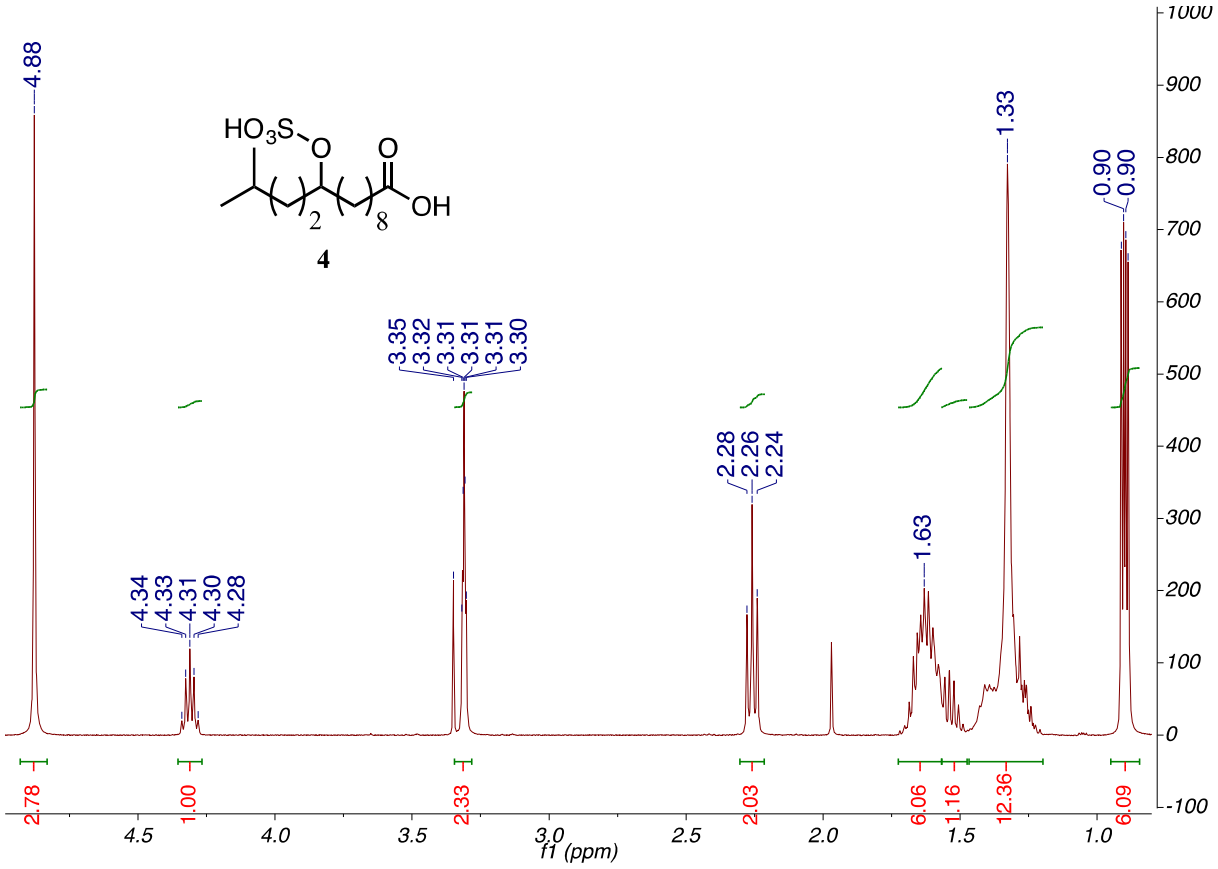
^{1}H NMR Spectrum (400 MHz, chloroform-*d*) of 2



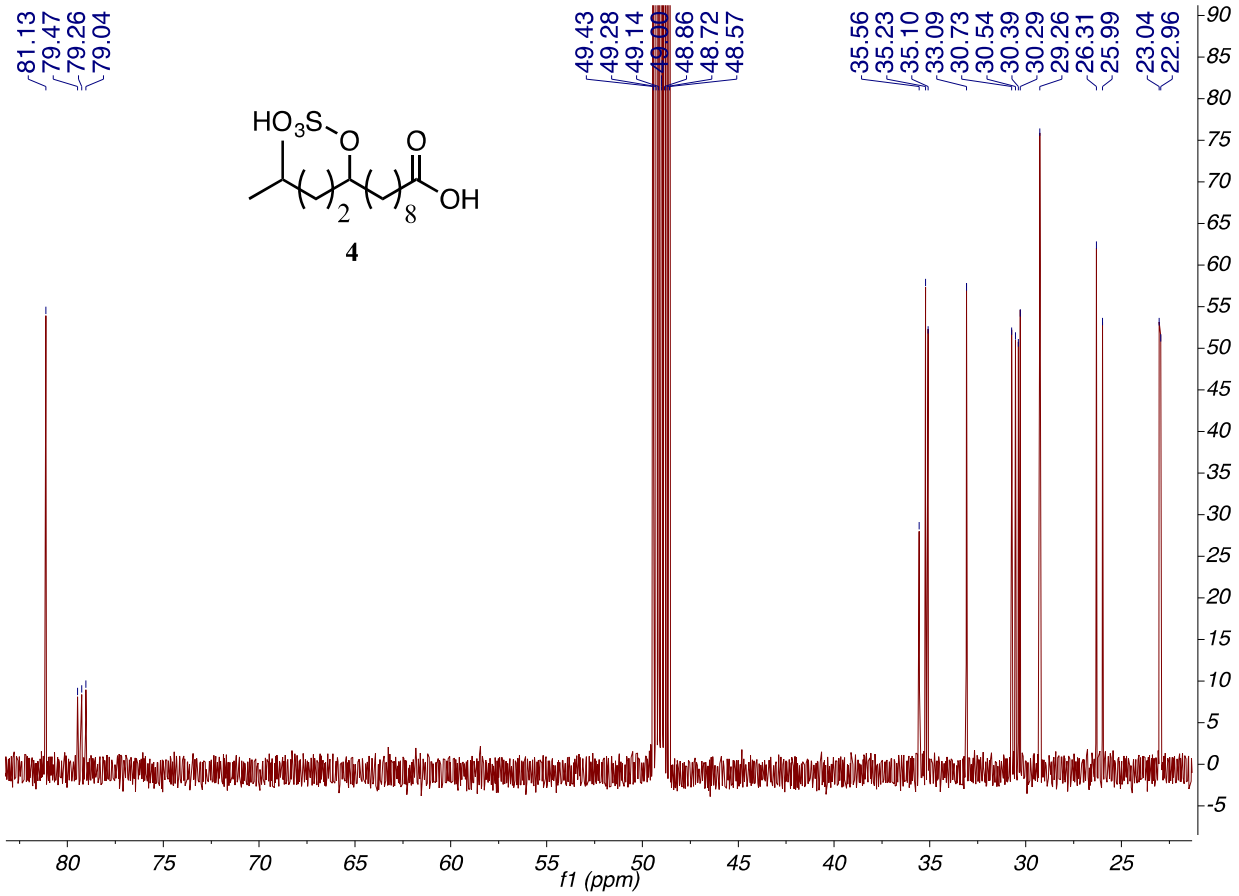
^{1}H NMR Spectrum (600 MHz, chloroform-*d*) of 3



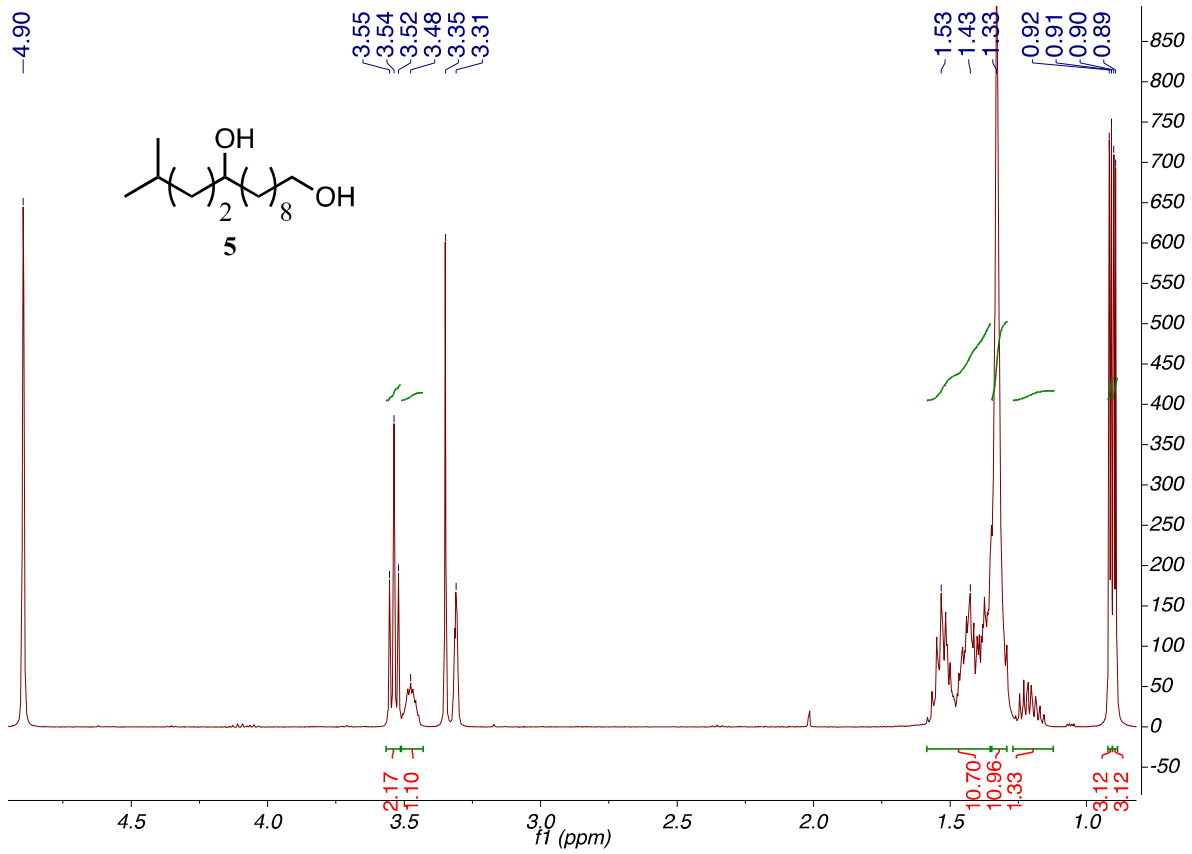
^{1}H NMR Spectrum (600 MHz, methanol-d_4) of 4



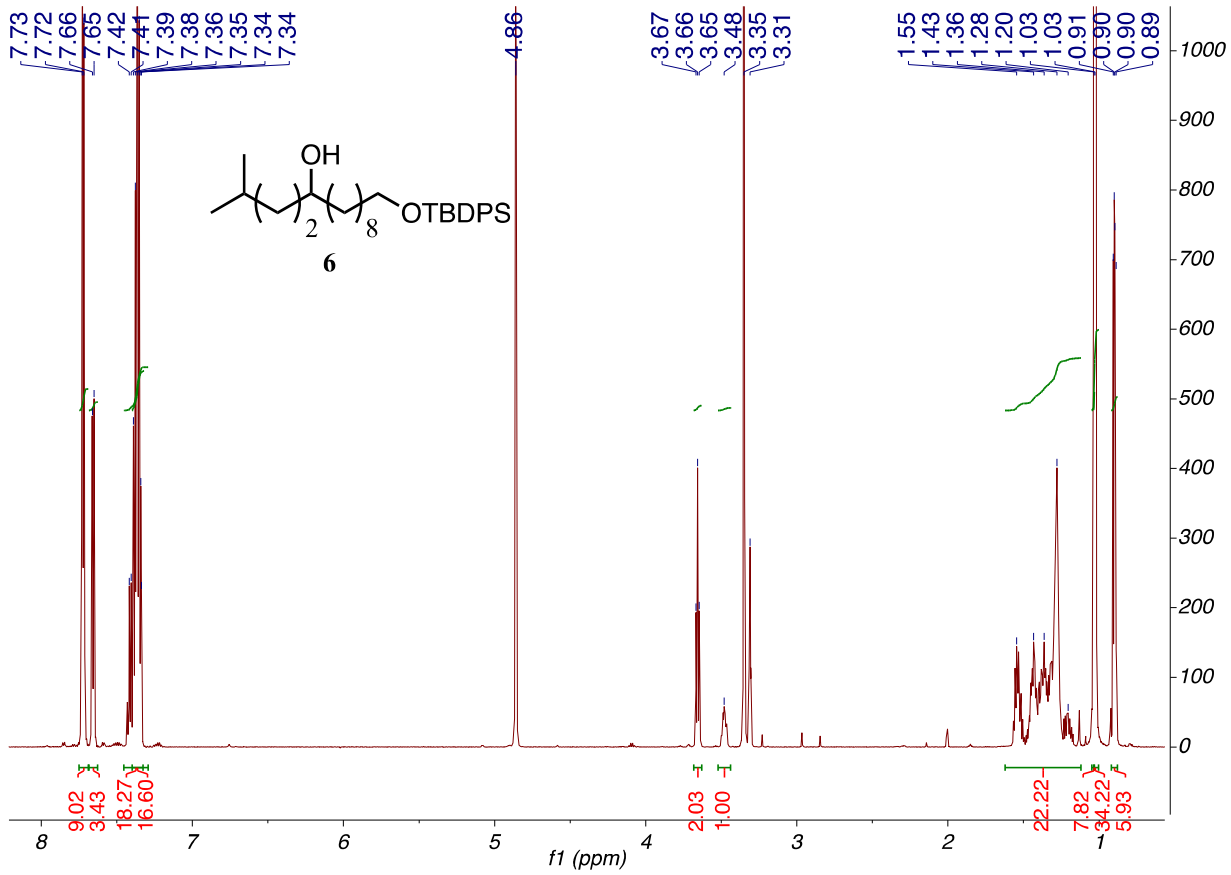
^{13}C NMR Spectrum (600 MHz, methanol-d_4) of 4



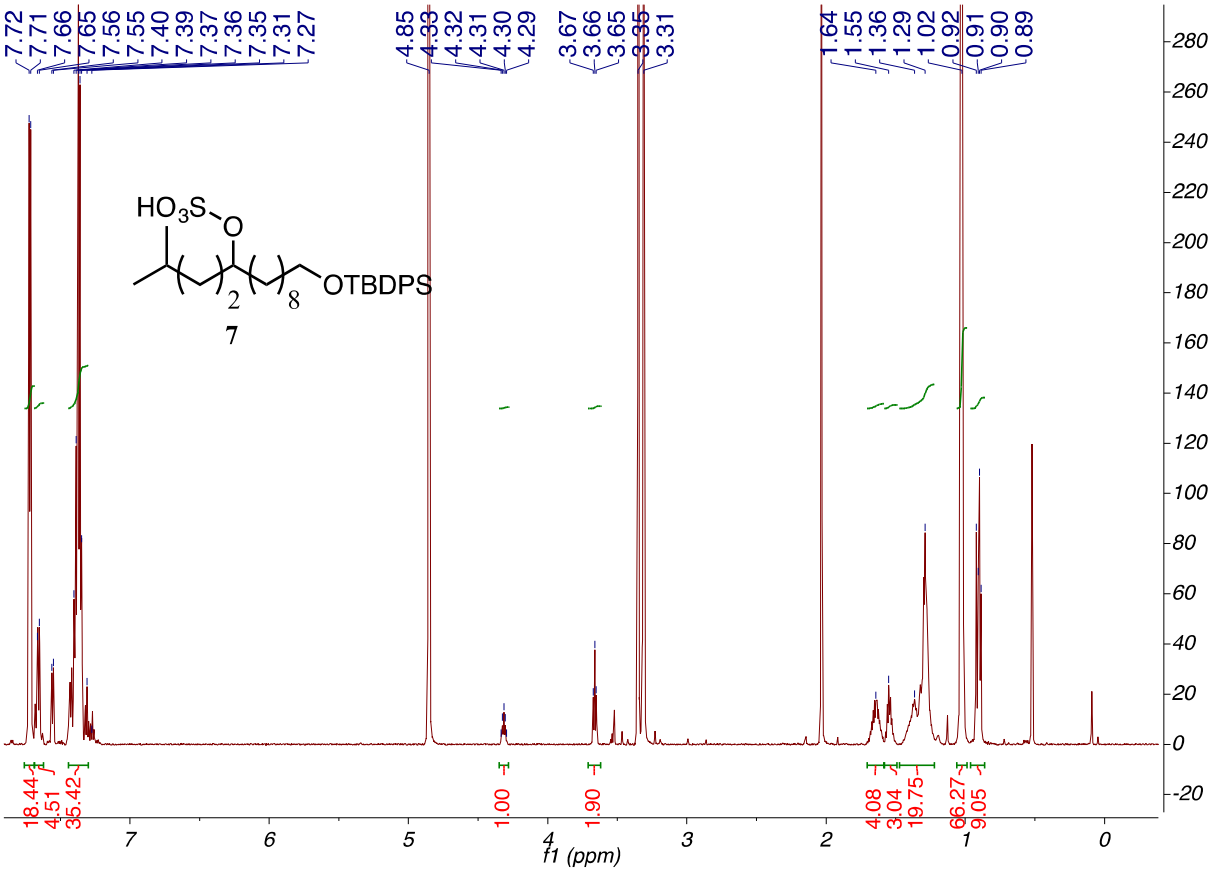
^{1}H NMR Spectrum (400 MHz, methanol-d_4) of 5



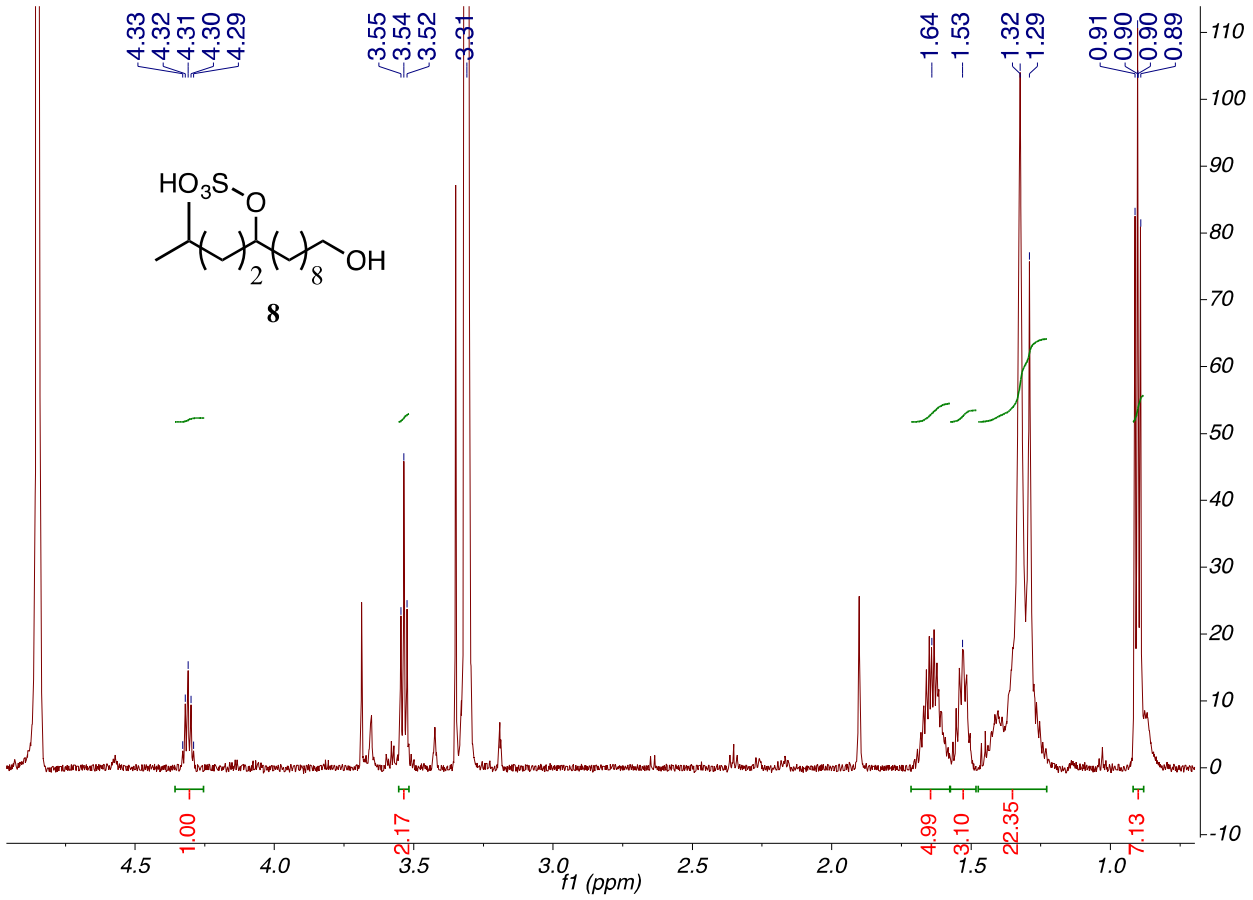
^{1}H NMR Spectrum (600 MHz, methanol-d_4) of 6



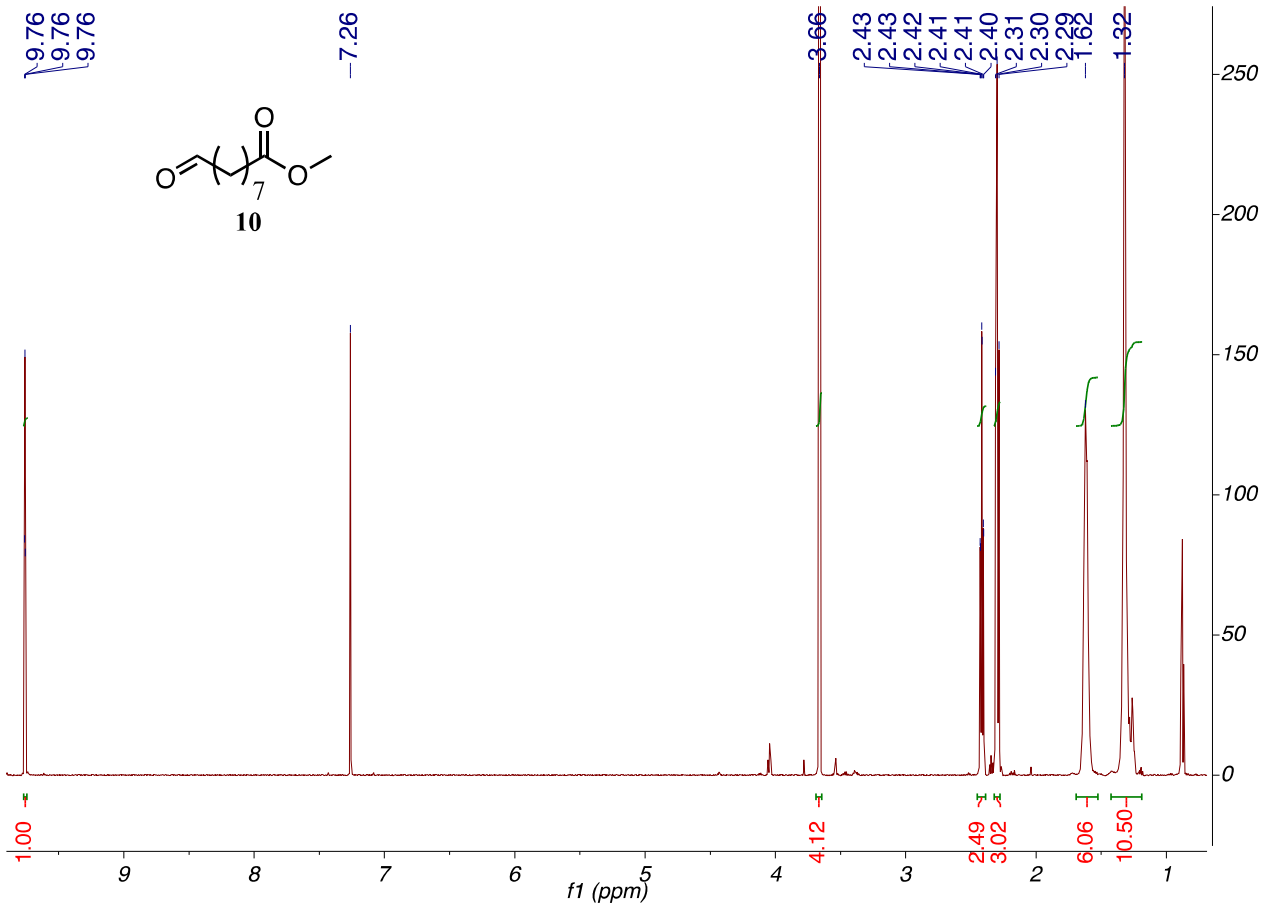
^{1}H NMR Spectrum (600 MHz, methanol-d_4) of 7



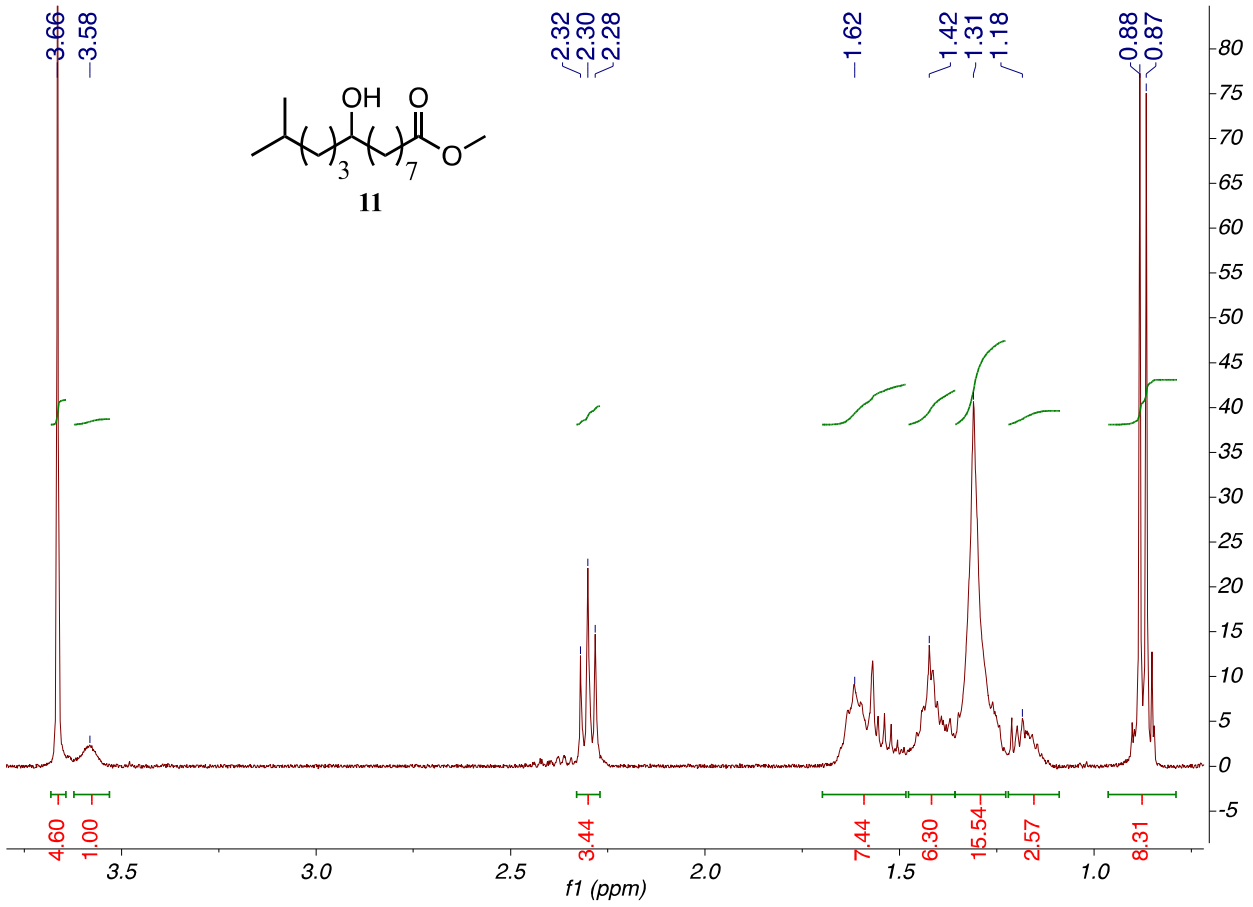
^{1}H NMR Spectrum (600 MHz, methanol-d_4) of 8



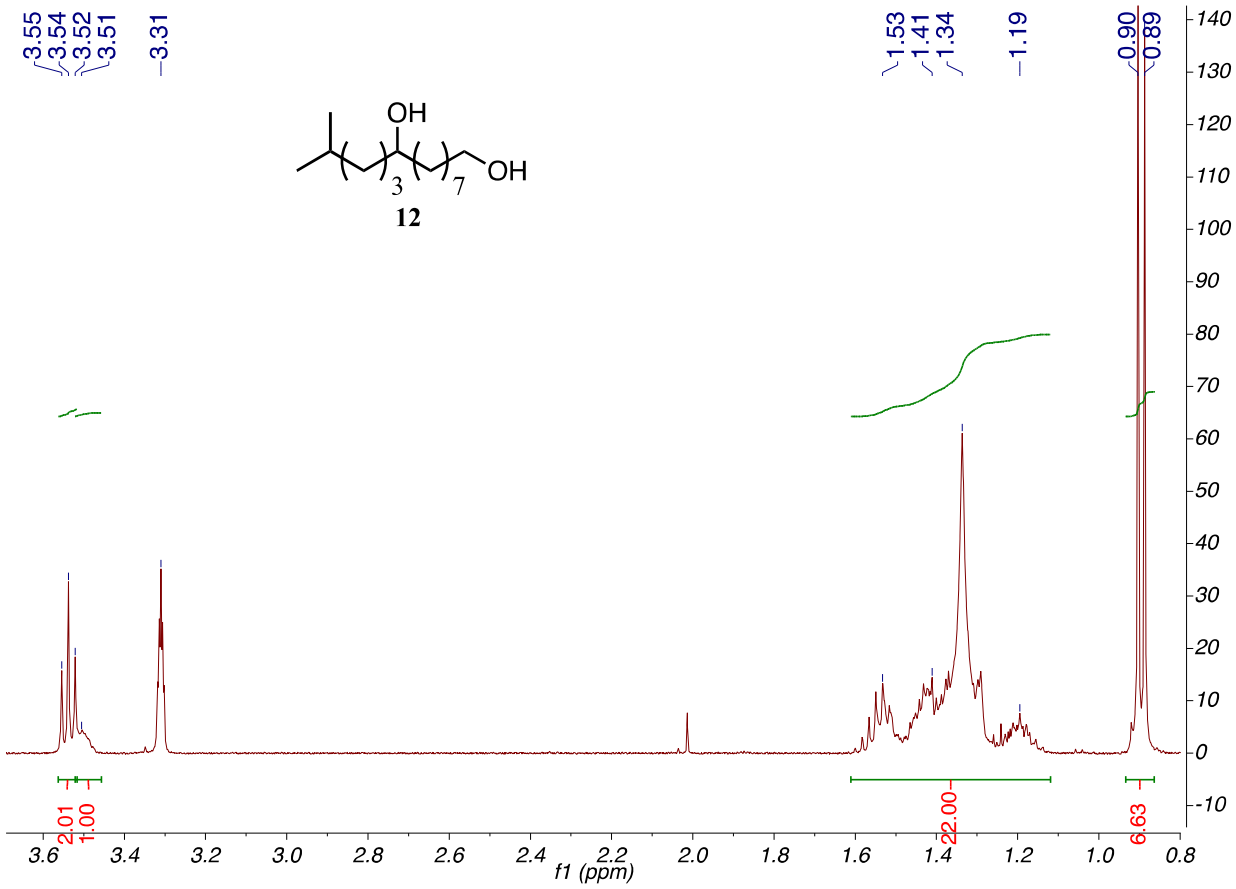
^{1}H NMR Spectrum of (600 MHz, chloroform-*d*) of 10



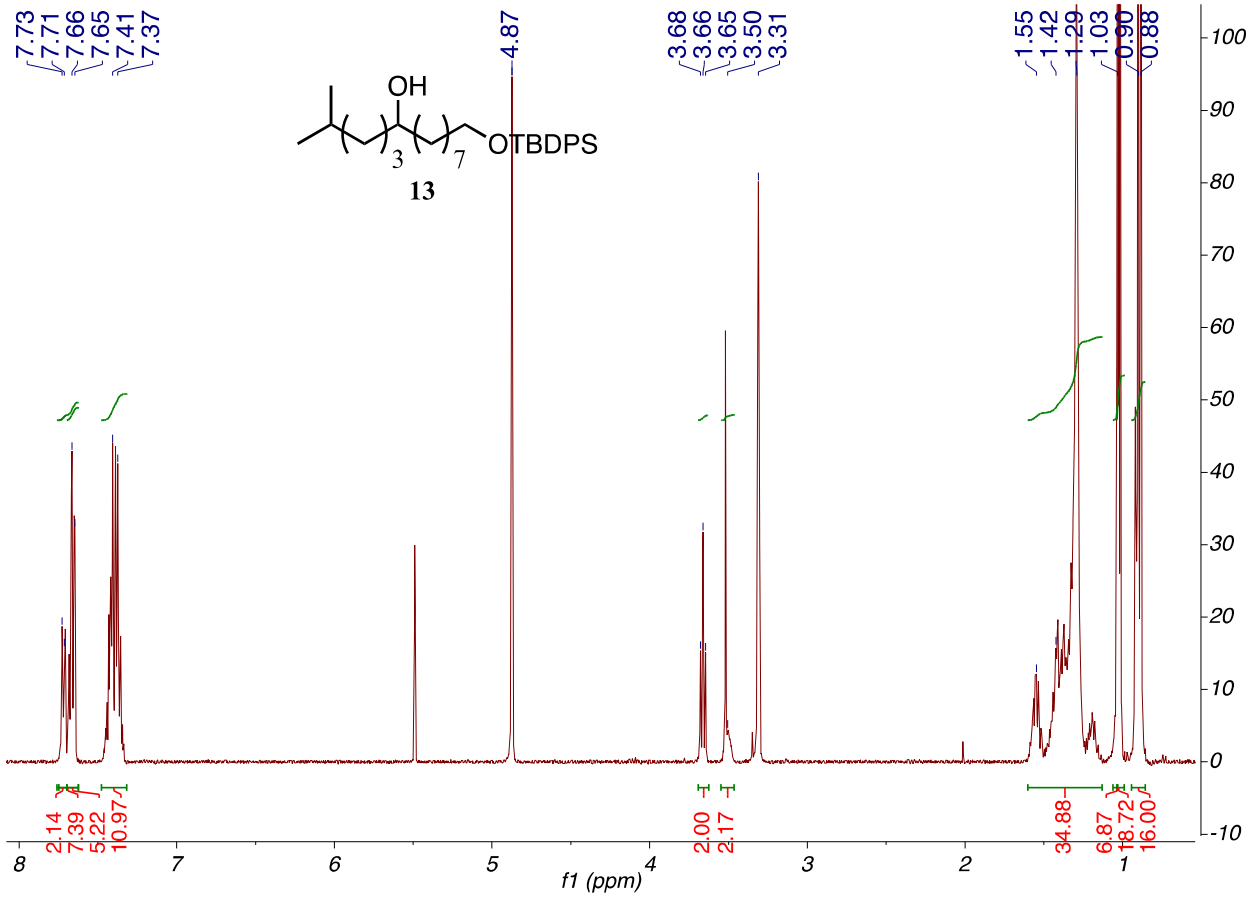
^{1}H NMR Spectrum (400 MHz, chloroform-*d*) of 11



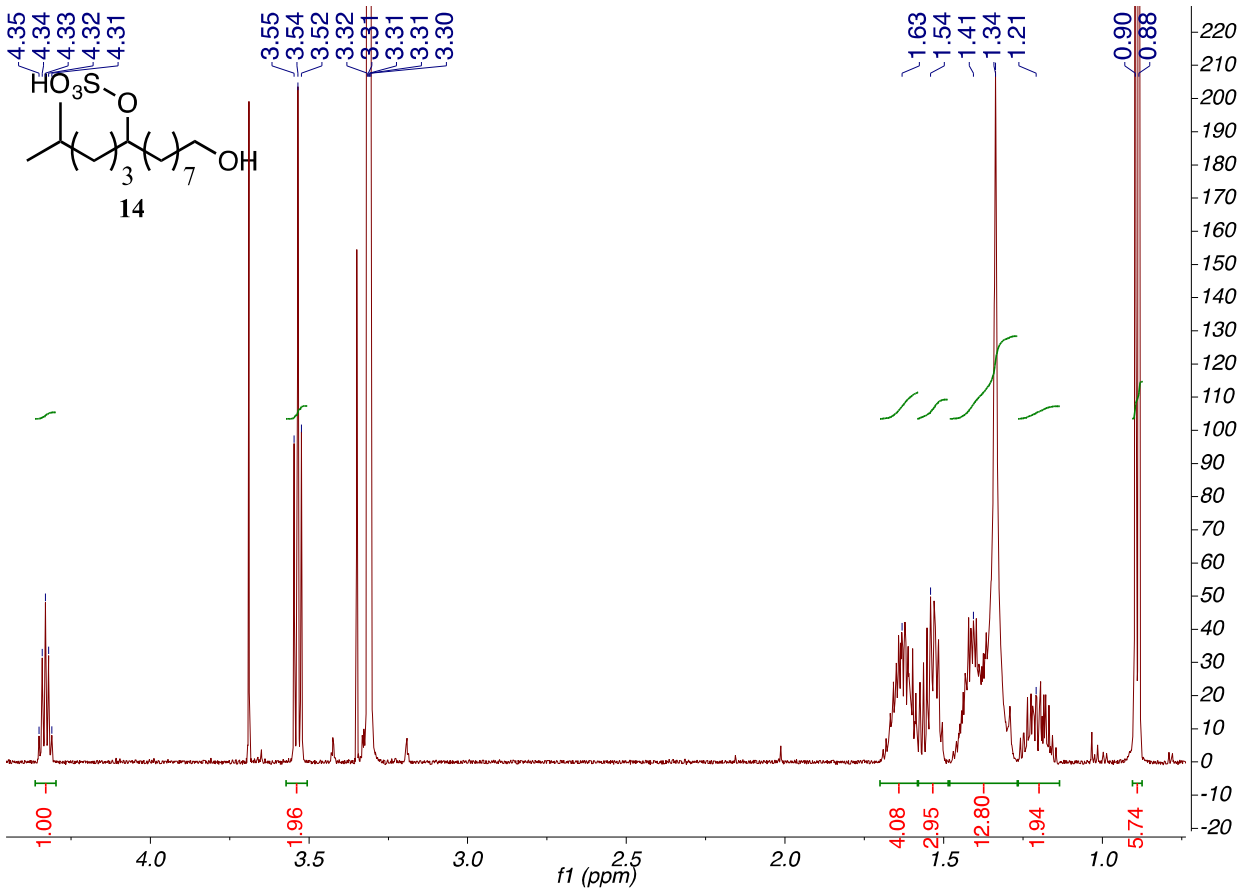
^{1}H NMR Spectrum (400 MHz, methanol-d_4) of 12



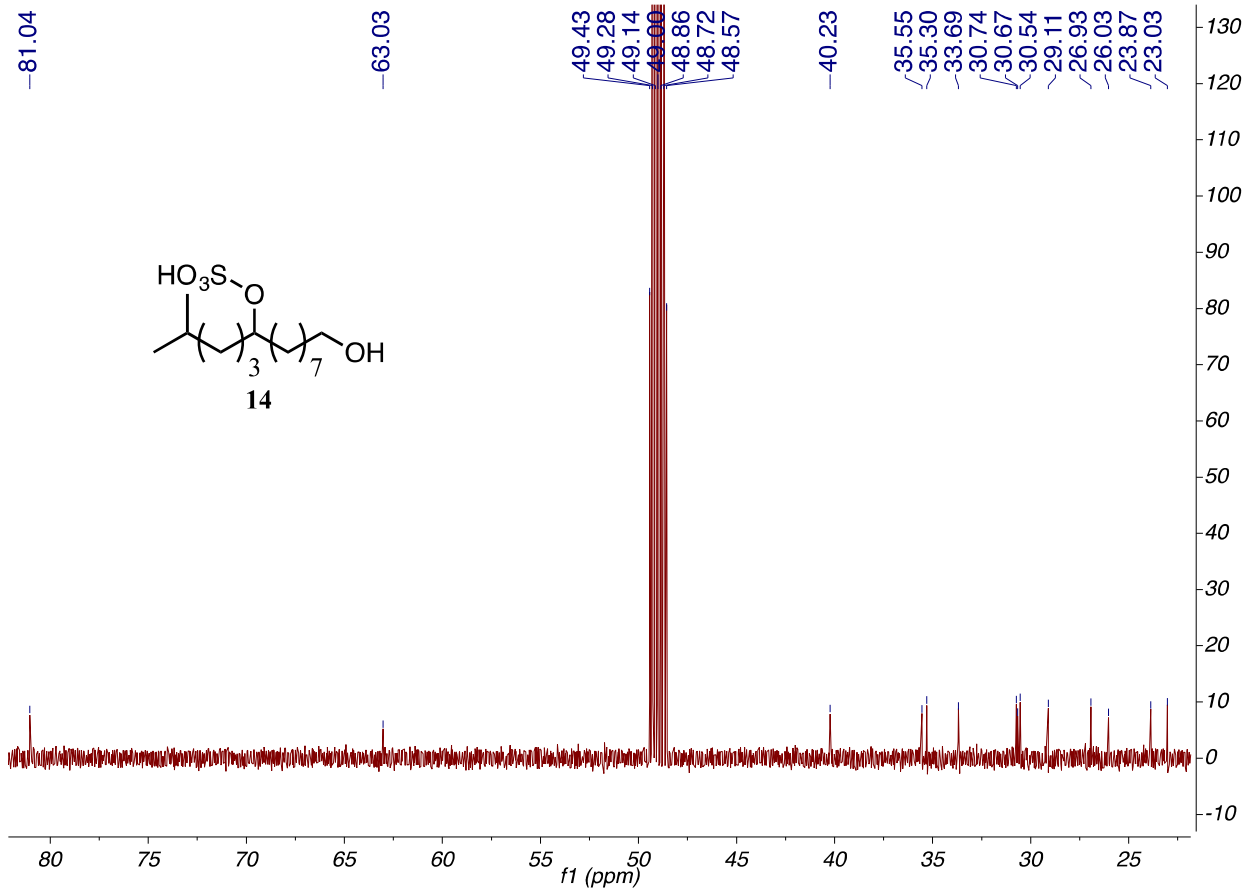
^{1}H NMR Spectrum (400 MHz, methanol-d_4) of 13



^{1}H NMR Spectrum (600 MHz, methanol-d_4) of 14



^{13}C NMR Spectrum (600 MHz, methanol-d_4) of 14



1.9 AUTHOR'S NOTE

All avoidance assays were conducted by Zheng Liu and Amy Pribadi under the instruction of Prof. Sreekanth Chalasani. Initial metabolomic samples were provided by Dr. Neelanjan Bose and Joshua Yim. Z.L., M.J.K., C.D.C., A.K.P., S.G.L., A.T., K.P.C., and N.B. performed experiments, developed experimental methods and reagents. F.C.S., J.S., and S.H.C. designed and interpreted the experiments and wrote the paper.

REFERENCES

1. Plutchik R. A.; general psychoevolutionary theory of emotion. *Theor. Emot.* 1, 3–31 (1980).
2. Apfelbac, R.; Blanchard C. D.; Blanchard R. J.; Hayes, R. A. & McGregor; I.S. The effects of predator odors in mammalian prey species: a review of field and laboratory studies. *Neurosci. Biobehav. Rev.* 29, 1123–1144, S0149-7634 (05)00098-9 [pii] (2005).
3. Papes F.; Logan D. W. & Stowers L.; The vomeronasal organ mediates interspecies defensive behaviors through detection of protein pheromone homologs. *Cell* 141, 692–703, S0092-8674(10)00355-7 [pii] (2010).
4. Yager D. D.; Predator detection and evasion by flying insects. *Curr. Opin. Neurobiol.* 22, 201–207, S0959-4388(11)00229-7 [pii] (2012).
5. Takahashi L. K.; Olfactory systems and neural circuits that modulate predator odor fear. *Front Behav. Neurosci.* 8, 72 (2014).
6. Stevens M. & Ruxton G. D.; Linking the evolution and form of warning coloration in nature. *Proc Biol Sci* 279, 417–426, rspb.2011.1932 [pii] (2012).
7. Kobayakawa K. et al.; Innate versus learned odour processing in the mouse olfactory bulb. *Nature* 450, 503–508, nature06281 [pii] (2007).
8. Brechbuhl J.; Klaey, M. & Broillet M. C.; Grueneberg ganglion cells mediate alarm pheromone detection in mice. *Science* 321, 1092–1095 (2008).
9. Tirindelli R.; Dibattista M.; Pifferi S. & Menini A.; From pheromones to behavior. *Physiol. Rev.* 89, 921–956, 89/3/921 [pii] (2009).
10. Stettler D. D. & Axel R.; Representations of odor in the piriform cortex. *Neuron* 63, 854–864 (2009).
11. Brenner S.; The genetics of Caenorhabditis elegans. *Genetics* 77, 71–94 (1974).
12. Sommer R. J.; Pristionchus pacificus. *WormBook*, 1–8, https://doi.org/10.1895/wormbook.1.102.1 (2006).
13. Hong R. L. & Sommer R. J.; Pristionchus pacificus: a well-rounded nematode. *Bioessays* 28, 651–659 (2006).
14. Bumbarger D. J.; Riebesell M.; Rodelsperger C. & Sommer R. J.; System-wide Rewiring Underlies Behavioral Differences in Predatory and Bacterial-Feeding Nematodes. *Cell* 152, 109–119, S0092-8674(12)01500-0 [pii] (2013).
15. Serobyan V.; Ragsdale E. J.; Muller M. R. & Sommer R. J.; Feeding plasticity in the nematode Pristionchus pacificus is influenced by sex and social context and is linked to developmental speed. *Evol. Dev.* 15, 161–170 (2013).
16. White J. G.; Southgate E.; Thomson J. N. & Brenner S.; The structure of the nervous system of the nematode Caenorhabditis elegans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 314, 1–340 (1986).
17. de Bono M. & Maricq A. V.; Neuronal substrates of complex behaviors in C. elegans. Annu Rev. *Neurosci.* 28, 451–501 (2005).

18. Barriere A. & Felix M. A.; Isolation of C. elegans and related nematodes. *WormBook*, 1‑19, https://doi.org/10.1895/wormbook.1.115.2 (2014).
19. Felix M. A. et al.; Species richness, distribution and genetic diversity of Caenorhabditis nematodes in a remote tropical rainforest. *BMC Evol. Biol*. 13, 10 (2013).
20. Serobyan V.; Ragsdale E. J. & Sommer R. J.; Adaptive value of a predatory mouth-form in a dimorphic nematode. *Proc. Biol. Sci*. 281, 20141334 (2014).
21. Schafer W. R.; Egg-laying. *WormBook*, 1‑7, https://doi.org/10.1895/wormbook.1.38.1 (2005).
22. Forseth R. R. & Schroeder F. C.; NMR-spectroscopic analysis of mixtures: from structure to function. *Curr. Opin. Chem. Biol*. 15, 38‑47 (2011).
23. Pungaliya C. et al.; A shortcut to identifying small molecule signals that regulate behavior and development in Caenorhabditis elegans. *Proc. Natl Acad. Sci.* USA 106, 7708‑7713, 0811918106 [pii] (2009).
24. Hilliard M. A.; Bergamasco C.; Arbucci S.; Plasterk R. H. & Bazzicalupo; P. Worms taste bitter: ASH neurons, QUI-1, GPA-3 and ODR-3 mediate quinine avoidance in Caenorhabditis elegans. *Embo J.* 23, 1101‑1111 (2004).
25. Stowe M. K.; Turlings T. C.; Loughrin J. H.; Lewis W. J. & Tumlinson J. H.; The chemistry of eavesdropping, alarm, and deceit. *Proc. Natl Acad. Sci.* USA 92, 23‑28 (1995).
26. Fujita M.; Nakao Y.; Matsunaga S.; Nishikawa T. & Fusetani; N. Sodium 1-(12-hydroxy)octadecanyl sulfate, an MMP2 inhibitor, isolated from a tunicate of the family Polyclinidae*. J. Nat. Prod.* 65, 1936‑1938 (2002).
27. Nakao Y.; Matsunaga S. & Fusetani N.; Toxadocials B, C and toxadocic acid A: thrombin-inhibitory aliphatic tetrasulfates from the marine sponge, Toxadocia cylindrica. *Tetrahedron* 49, 11183‑11188 (1993).
28. Yasumoto K. et al. ;Aliphatic sulfates released from Daphnia induce morphological defense of phytoplankton: isolation and synthesis of kairomones. *Tetrahedron Lett.* 46, 4765‑4767 (2005).
29. White A. R.; Duggan B. M.; Tsai S. C. & Vanderwal C. D.; The Alga Ochromonas danica Produces Bromosulfolipids. *Org. Lett.* 18, 1124‑1127 (2016).
30. O'Doherty I.; Yim J. J.; Schmelz E. A. & Schroeder F. C.; Synthesis of caeliferins, elicitors of plant immune responses: accessing lipophilic natural products via cross metathesis. *Org. Lett.* 13, 5900‑5903 (2011).
31. Alborn H. T. et al.; Disulfooxy fatty acids from the American bird grasshopper Schistocerca americana, elicitors of plant volatiles. *Proc. Natl Acad. Sci.* USA 104, 12976‑12981 (2007).
32. Kniazeva M.; Crawford Q. T.; Seiber M.; Wang C. Y. & Han M.; Monomethyl branched-chain fatty acids play an essential role in Caenorhabditis elegans development. *PLoS Biol.* 2, E257 (2004).

33. Ragsdale E. J. & Ivers; N. A. Specialization of a polyphenism switch gene following serial duplications in Pristionchus nematodes. *Evolution* 70, 2155‐2166 (2016).
34. Ragsdale E. J.; Muller M. R.; Rodelsperger C. & Sommer R. J.; A developmental switch coupled to the evolution of plasticity acts through a sulfatase. *Cell* 155, 922‐933 (2013).

CHAPTER TWO
The Sulfatome of *P. pacificus*

Authors: Kariya M.J.; Bose N.; Wong S.; Biddle J.; Yim J.J.; Aguilar-Romero J.; Ragsdale E.J.; Schroeder F.C.

2.1 ABSTRACT

From tRNA-derived ascarosides to xylose-based nucleosides, *Pristionchus pacificus* has been shown to make a large diversity of molecules. Using an activity-guided search to find the structure of an alert pheromone from *P. pacificus*, we found a collection of novel sulfolipids from its metabolome. Based on key characteristics of the sulfolipids, we used targeted fragment traces in HPLC-MSMS to uncover a library of sulfolipids from the *P. pacificus* metabolome. Through these studies we came across over 30 unique compounds which contain sulfate, two of which are unprecedented cyclic sulfates. We propose the biosynthesis and function of the sulfolipids based on their effects on microbes and nematodes.

2.2 INTRODUCTION

Sulfation has been well described in the literature as a deactivation mechanism of otherwise toxic xenobiotic compounds[35]. Sulfation inactivates some xenobiotic compounds and allows for their removal from the organism through physiological fluids[45]. Outside the context of detoxification, we find sulfation in signaling events, such as with heparan sulfate and its binding of fibroblast growth factors[36] and the regulation

of steroid action through sulfation[37], in addition sulfated nucleosides have been found in spider venom[38].

In the organism *P. pacificus*, we find an unusually large collection of sulfates that may be important in the life history and interactions of the organism. Through a search for a *P. pacificus*-produced molecule that elicits fear-like behaviors in*. elegans*, we uncovered sulfated branched chain lipids with varying chain lengths[39] that resemble the known *C. elegans* deterrent, SDS. The identification of these sulfolipids, sufac and sufal, motivated an in-depth analysis of the *P. pacificus* metabolome using tandem mass spectrometry and subsequently revealed a large collection of structurally diverse sulfated lipids.

The sulfated lipids we found include cyclic sulfate lipids, unprecedented in nature, sulfated epoxide lipids, trihydroxyl sulfolipids, and sulfated branched chain lipids of varying chain lengths. The compounds appear to be derived from fatty acid biosynthesis and are produced *de novo* by *P. pacificus*.

2.3 RESULTS

Finding Sulfated Molecules in the *P. pacificus* Metabolome

Through our search for an alert pheromone for *C. elegans* described in the previous chapter, we became aware of a class of sulfolipids from *P. pacificus* which are related to both the known nematode deterrent SDS and a molecule critical for *C. elegans* development, 13-methyl myristic acid (C15ISO). We subsequently found that

these comprise only a small portion of many sulfates within the *P. pacificus* metabolome, which we define as a "sulfatome".

The sulfatome of *P. pacificus* was uncovered through investigating the characteristic fragmentation of sufal and sufac by mass spectrometry. Each one of these sulfates cleaves a reproducible sulfate fragment, HO_4S^-, in tandem mass spectrometry (Figure 2a). We expanded our search to find parents of the sulfate fragment, HO_4S^-, in the *P. pacificus* metabolome using both manual chromatogram searches and using the Global Natural Products Network (GNPN) software. Through this method we found approximately 82 ions that clustered due to a common sulfate fragment (Figure 2b), about 40 of which were unique. Although the molecules clustered based on their similarity of having a sulfate fragment, we did not notice any other defining relations amongst the ions in the cluster. We manually observed each ion and picked out the most abundant ones amongst the TIC of *P. pacificus* exometabolome (Figure 2c, Supplementary Table S4).

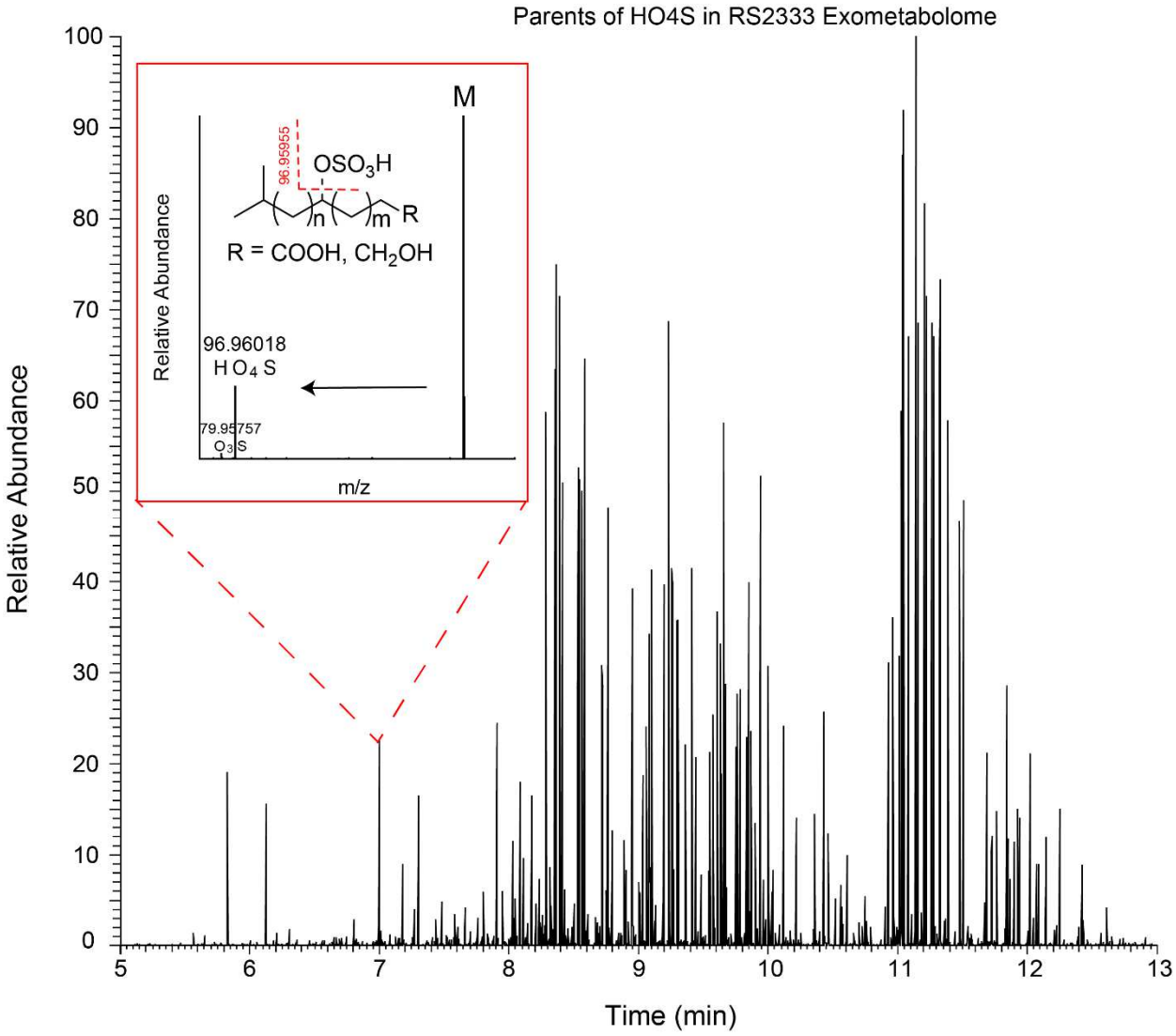


Figure 2a. Search for sulfates in *P. pacificus* metabolome**,** MSMS EIC of HO_4S from RS2333 (*P. pacificus* wt) supernatant methanol extract in ESI- mode.

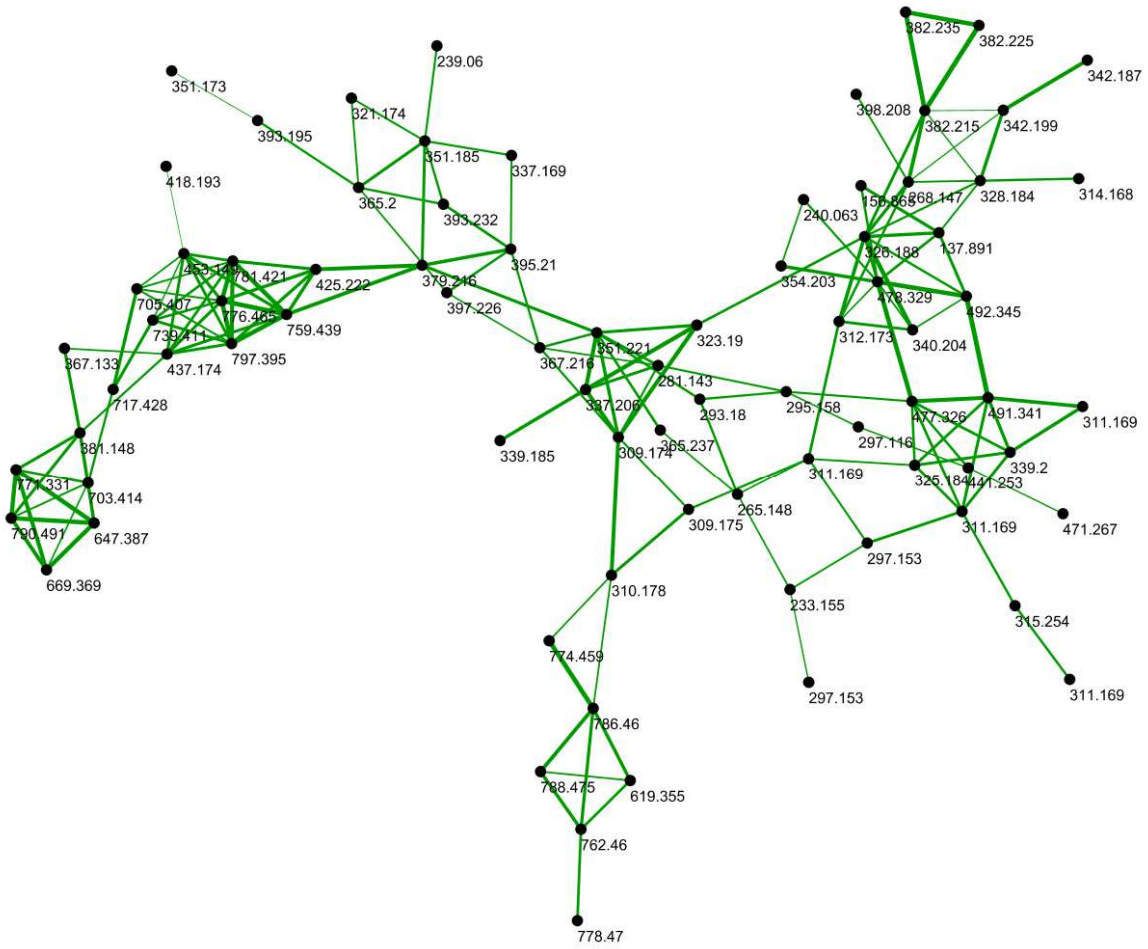


Figure 2b. MSMS network cluster of compounds in RS2333 exometabolome which contain a sulfate fragment, HO4S.

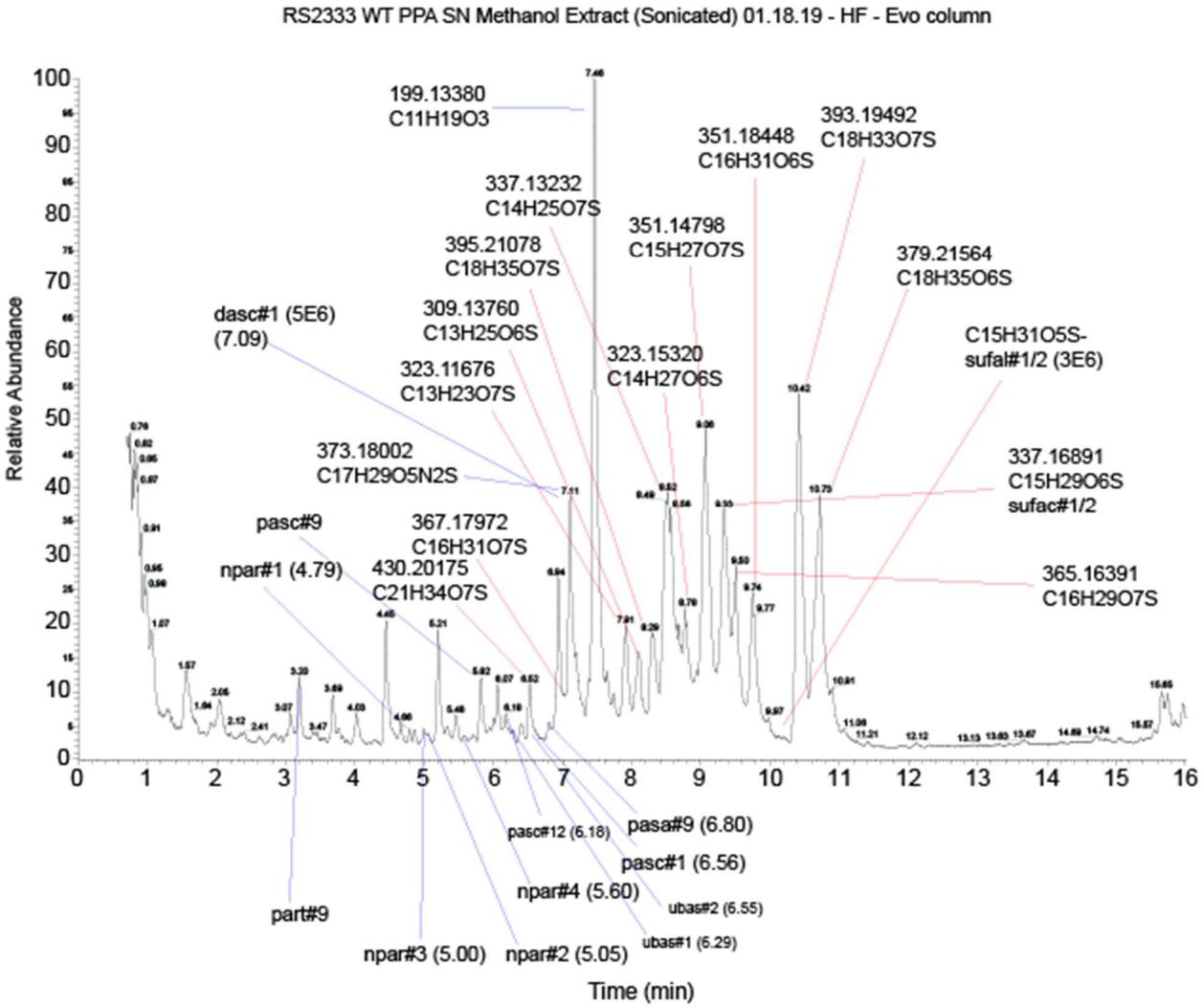


Figure 2c. TIC of RS2333 (*P. pacificus* wt) supernatant methanol extract in ESI- mode. Peaks indicated with red are sulfated ions that cleave a characteristic HSO4- fragment in MSMS mode, whereas peaks indicated in blue are known compounds in *P. pacificus*.

Identification of Cyclic Sulfates and Sufal#3

After the molecular formulae of these sulfolipids were determined, we isolated them using column chromatography and characterized several of them using a combination analysis of 2D NMR spectra and high-resolution tandem mass spectrometry. We started with m/z 309.17412, a sulfolipid alcohol structurally similar to

sufal#1/2 with one less methylene, which we named sufal#3. After isolation and characterization (Figure 3a), we synthesized sufal#3 in a similar method to sufal#1/2 (see methods).

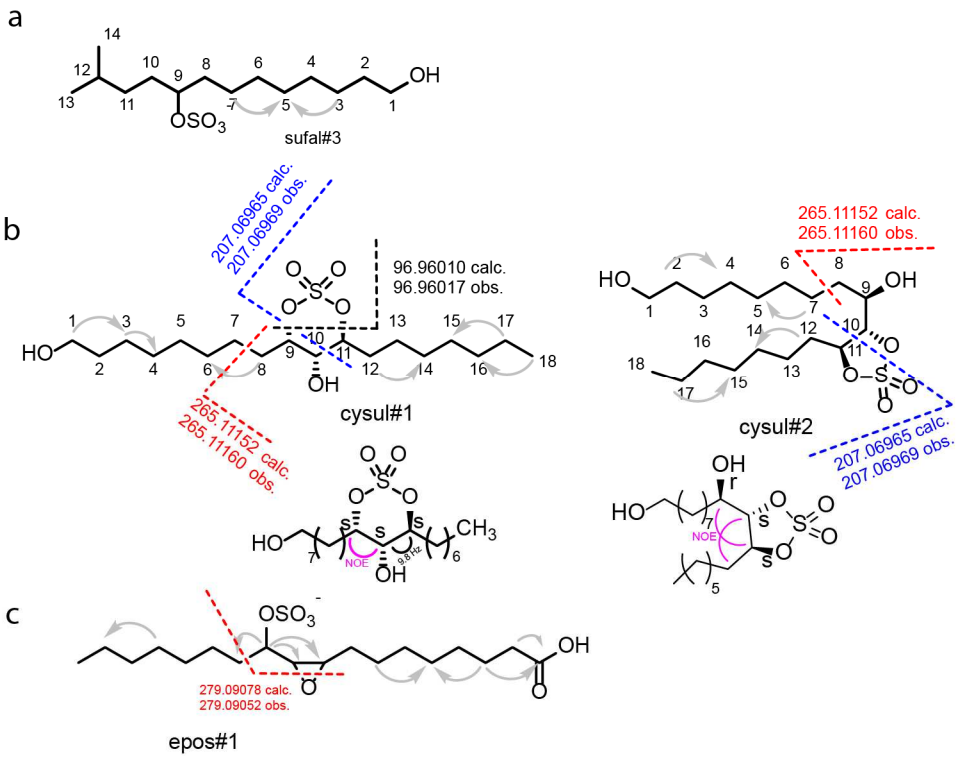


Figure 3 Elucidation of cyclic sulfates and other sulfolipids **a** Structure of sufal#3 with key HMBC correlations. **b** Structures of cysul#1 and cysul#2, with key HMBC correlations, MSMS fragments, and NOE correlations. **c.** Structure of epos#1 with key HMBC correlations and MSMS fragments.

We then moved on to two features with m/z = 379.21564, that are significantly less polar than sufal#3 based on retention time on C18. We isolated these compounds through HPLC, and using 2D NMR and UHPLC-MSMS analysis, we characterized these features as two cyclic sulfates (Figure 3b), one six-membered, cysul#1, and one five-

membered, cysul#2, the first of their kind in nature. Both compounds are formally derived from a trihydroxylated lipid that has two six-CH2 long chains, one of which bears a hydroxy at the end. Although we were unable to determine the orientation of the side chains to cysul through HMBC because of the stark similarity of carbon chemical shift values of methylene groups within the side chain, we were able to determine it through MSMS fragmentation data. Both cysul#1 and cysul#2 fragment through their sulfate ring at the 10^{th} and 11^{th} carbons to give a fragment of approximately m/z = 207. This fragment of 207 between the 10^{th} and 11^{th} carbon could only occur in the 5-membered ring of cysul#2 if the sulfate is attached to the 11^{th} carbon. Through a ROESY spectrum of cysul#2 we detected NOE correlations between protons 9 and 11, as well as 8 and 10, suggesting that the ring is oriented in *trans* and has a stereochemistry configuration of 9*R*, 10*S*, 11*S*, or the enantiomer (Supplementary Table S7). For cysul#1, we see a NOE correlation between protons 9 and 10 and a diaxial coupling constant between 10 and 11 (9.88 Hz), implying a stereochemical configuration of 9*S*, 10*S*, 11*R* or the enantiomer.

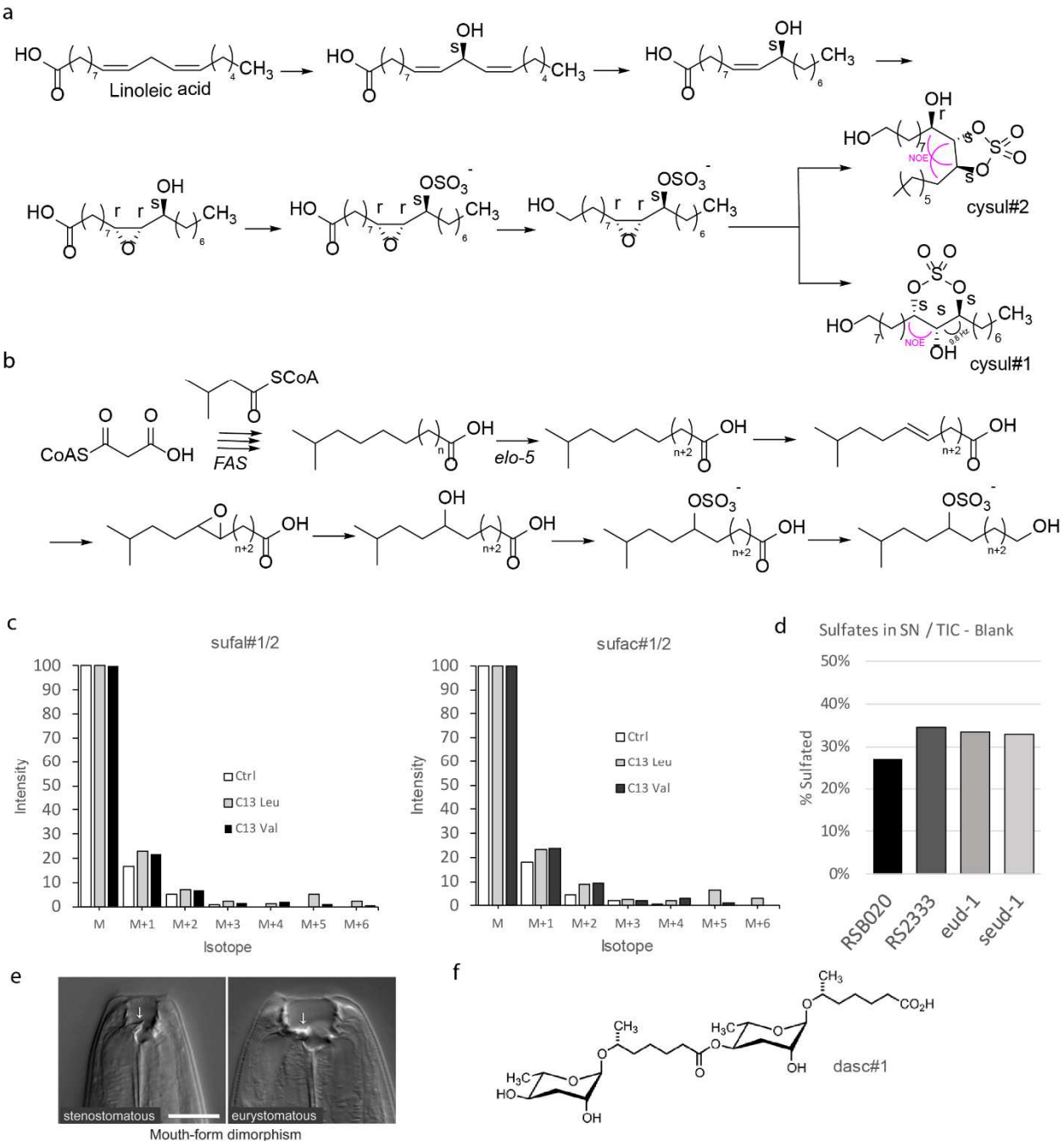


Figure 4 Proposed biosynthesis of sulfolipids **a**. cysul **b**. sufal **c**. abundance of isotopes of sulfolipids sufal #1/2 and sufac #1/2 after being fed water, C13 Leucine, and C13 Valine **d**. Percentage of known sulfated compounds compared to total TIC normalized to blank in various *P. pacificus* supernatant extracts **e.** mouth dimorphism of *P. pacificus*[54] **f.** structure of known mouth form effector dasc#1[54].

Investigating the biosynthesis of sulfolipids from *P. pacificus*

Next, we investigated the biosynthesis of the sulfolipids. Both cysul#1 and cysul#2 are formally derived from a 9,10,11-trihydroxylated 18-carbon chain, so we posited that they are derived from the same precursor metabolite. Through examining the metabolome of *P. pacificus* further we had characterized an epoxidated sulfate, epos#1, that could represent the precursor to both cysul#1 and cysul#2 (Figure 3c, Supplementary Figure S6). In addition, we also find evidence via MSMS for a non-sulfated epoxide acid (Supplementary Figure S7). We believe that the cyclic moiety is generated from a nucleophilic addition into an epoxide of a neighboring sulfate to generate a five membered or six membered ring (Figure 4a). Based on the predicted stereochemistry of cysul#2, we propose a stereochemistry of 9*R*, 10*R*, 11*S* for epos#1 (or the enantiomer). This stereochemistry also aligns with the predicted stereochemistry of cysul#1 after inversion of stereochemistry at the 9th carbon through nucleophilic addition. These compounds may ultimately be derived from oleic or linoleic acid.

The regularity to the branched chain sulfolipids, sufal, and sufac hints that the compounds are developed from branched chain fatty acid biosynthesis. We proposed that the biosynthesis is initiated with a coupling of leucine and malonyl-coa to generate a branched chain lipid structure of varying chain length (Figure 4b). We then propose that this scaffold would then be oxidized, epoxidated, hydrolyzed at either epoxide carbon, then sulfated. To investigate the first steps of this proposed biosynthesis, we fed

P. pacificus C13-labeled leucine and valine and checked the extracted metabolome using HPLC-HRMS for the presence of labeled sulfolipids. We found an upregulated M+5 for both sufal#1/2, sufac#1/2, and other various chain length sulfolipids (Supplementary Table S4), implying that a fully-labeled, leucine-derived, isovaleryl-starter unit was incorporated into these branched chain sulfolipids (Figure 4c). Importantly, the M+5 peak in the mass spectra of the leucine-fed samples was much larger than the M+4 peak, indicating that a complete 5-carbon unit was incorporated, as opposed to percolation of label through breakdown of the fed leucine into acetate units. However, we do find some label incorporation into cysul#1/2 with upregulated M+1 through M+6, suggesting that part of the fed labeled amino acids are fully broken and fed back into fatty acid biosynthesis through acetyl-CoA.

We then considered the final steps of sulfolipid biosynthesis in *P pacificus*, specifically how the lipid alcohols are sulfated. Unlike *C. elegans*, *P. pacificus* has an expanded number of predicted sulfotransferases that could play a role in the biosynthesis of the sulfolipids[34, 43]. We had tested two known sulfatase and sulfotransferase knockout mutants, *eud-1* (tu445) and *seud-1* (iub7)[43], respectively. Each sulfotransferase mutant had somewhat variable levels of sulfolipids (Supplementary Figure S4), however, we did not detect complete abolishment of any specific sulfolipid. This may imply that there are multiple sulfotransferases with redundant activities that are responsible for the sulfation of a single precursor alcohol. Given that axenic worms

(worms grown on an artificial diet, without bacteria) also produce sulfolipids (Supplementary Figure S1), the sulfation does appear to be conducted by *P. pacificus*.

Searching for specific metabolites made by sulfotransferases in *P. pacificus*.

Ragsdale et al. showed that the putative sulfotransferases *eud-1* and *seud-1* have an important role in the development and life history of *P. pacificus*[34, 43]. A knockout of the sulfatase *eud-1* causes *P. pacificus* to develop exclusively into a closed mouth form (stenostomatous), whereas a knockout in the sulfotransferase *seud-1* causes a constitutively open mouth form (eurystomatous)[34, 43] (Figure 4e). Through this screen of sulfotransferase mutants, we aimed to find a small molecule substrate for either *eud-1* or *seud-1* which may be inducing this phenotypic dimorphism. Through preliminary comparative metabolomics we detected several candidate ions that are differentially regulated by *seud-1* and *eud-1* (Supplementary Figure S13). We have yet to determine the identity and assay for any effects of these putative *eud-1* or *seud-1* products or substrates on the phenotypic dimorphism, however, these data suggest that there are effectors of the mouth dimorphism from the *P. pacificus* metabolome in addition to the previously identified ascaroside dimer, dasc#1 (Figure 4f, Supplementary Figure S8).

Given the wide diversity and abundance of sulfated lipids *in P. pacificus*, it appeared that sulfolipids account for a large percentage of the *P. pacificus* metabolome. To get a rough estimate of the contribution of sulfolipids to the detected metabolome,

we calculated the sum of integration of traces for the list of sulfated lipids and compared that to the total TIC. This showed that the total content of ionizable sulfates in *P. pacificus* exometabolome comprises roughly 25% of the total exometabolome TIC (Figure 4d, Supplementary Figure S3). Sulfates are excellent charge carriers in ESI mass spectrometry, so this implies that less than a quarter of *P. pacificus*' metabolome is sulfated. It is interesting to note that the *C. elegans* metabolome contains no detectable sulfolipids (Supplementary Figure S9).

We further found that this abundance of sulfolipids is not specific to the *P. pacificus* laboratory strain RS2333 alone and is shared amongst all studied strains (Supplementary Figure S10). This implies that the biosynthetic machinery needed to create these sulfolipids has not yet been significantly affected through propagation in a laboratory setting, in contrast to some ascaroside-related traits in *C. elegans*[55].

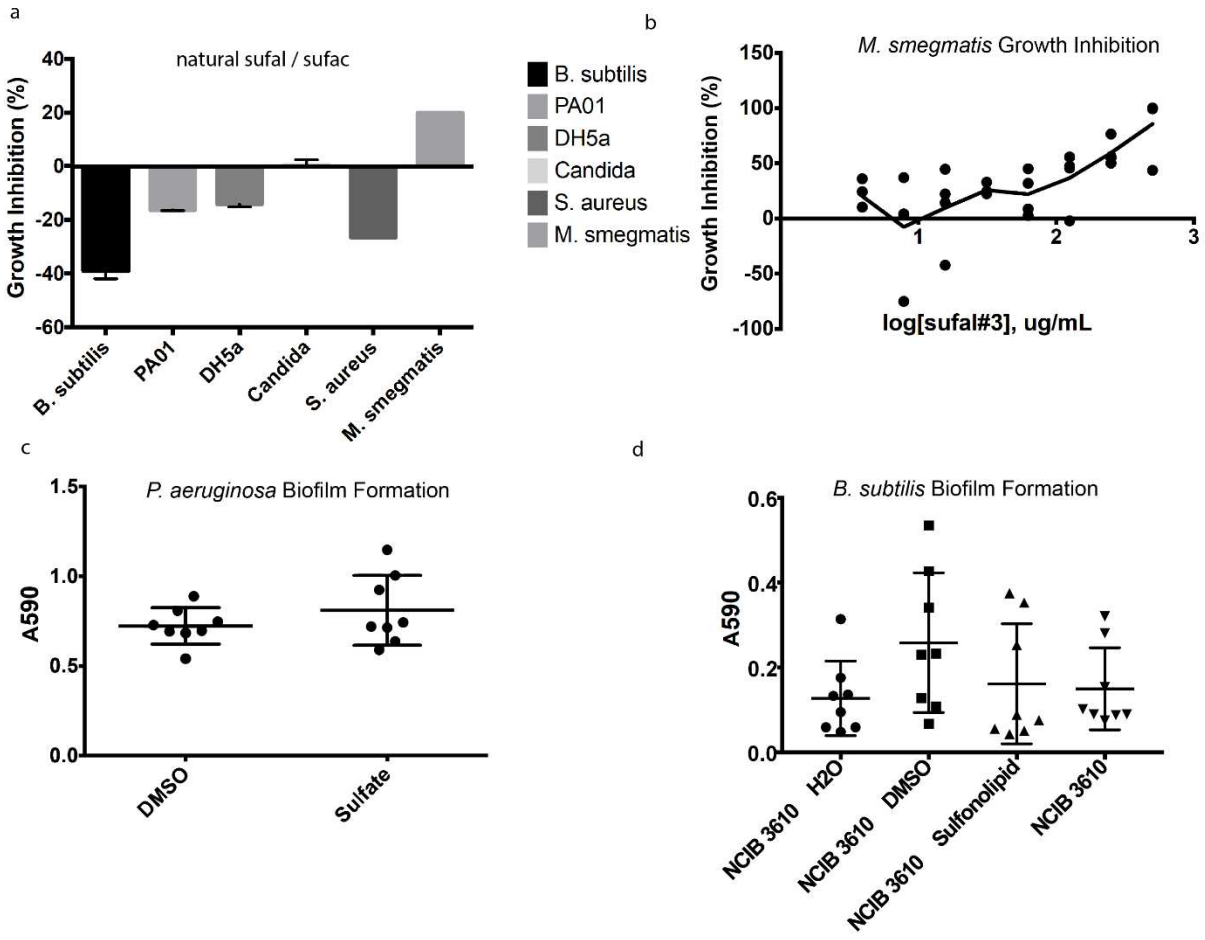


Figure 5 Sulfolipid effect on microbial growth **a**. effect of natural sufal/sufac mixture on microbial growth **b**. effect of synthetic sufal#3 on *M. smegmatis* growth with MIC between 0.8 – 1.6 mM **c**. natural sufal/sufac mixture on biofilm formation of *P. aeruginosa* (PA14) and **d**. *B. subtilis* (NCIB 3610) at 0.25 mg/mL

We probed why *P. pacificus* makes such an abundant amount of these sulfolipids (estimated to be on the order of 60 mg of sulfolipids per 30 million worms). We believe that *P. pacificus* uses these sulfolipids as a means for digestion of bacteria or for selective growth of bacteria in place of a bacterial grinder. We tested a mixture of sulfolipids on various microbial organisms (Figure 5a) to see if there were any effects on growth inhibition. Treatment with sulfolipids at 0.25 mg/mL led to enhanced growth of

many of the microbes that we had tested, including a known symbiont *Bacillus subtilis*, and inhibition of others, such as *Mycobacterium smegmatis,* a known human pathogen (Figure 5b). This implies that these sulfolipids can affect microbial growth, selectively. It is known that SDS can affect biofilm formation in microbes[49]*,* so we tested sufal#3 for biofilm formation using a crystal violet assay (Figure 5c, d), yet found no significant effect. Unlike sufal and sufac, which have drastic effects on microbial growth, cysul does not (Supplementary Files 2, E11).

Looking further into sufal#3's effect in its environment, we tested it in a population dependent developmental acceleration (PDDA) assay[47] against *C. elegans*. Both SDS and sufal#3 had a marked effect on egg laying time, significantly shortening the time until first egg laid, implying a decrease in life span of worms treated with either sulfolipids (Supplementary Figure S7).

2.4 DISCUSSION

Bioactivity

Amongst the sulfated lipid metabolites we identified in *P. pacificus*, we came across two unprecedented cyclic sulfates. Cyclic sulfates have been used as a synthetic tool for decades[40, 41]. They function similar to epoxides[40], yet with greater reactivity towards certain nucleophiles[41]. In addition, their products are easier to separate due to the anionic nature of an open sulfate, and they have the ability to impart two contiguous stereocenters in a scaffold. Due to this versatility, a cyclic sulfate synthon was used as an intermediate in the synthesis of ascarosides[42], one important family of signaling molecules found in nematodes[42].

Although the function of the cyclic sulfates in *P. pacificus* is yet to be determined, we show that they seem to be formed from epoxide lipids, some of which are known to be biologically active, such as epoxide derivatives of unsaturated fatty acids[50, 51] that lose their activity upon hydrolysis. Perhaps these cyclic sulfates are a result of intramolecular hydrolysis of an otherwise active precursor. On the other hand, synthetic cyclic sulfates have been shown to act as signaling molecules similarly to an epoxide as an irreversible glycosidase inhibitor[46]. Given that both cyclic sulfates and epoxides have been shown to be active, and that the two cyclic sulfates found appear to be derived from the same precursor, it seems that cyclic sulfates are intentionally produced metabolites, rather than products of detoxification.

Biosynthesis

We showed that the branched chain sulfolipids affect nematode and microbial growth, including a known symbiont. We believe that these sulfates serve to help *P. pacificus* digest and selectively grow food. This idea resonates with two key differences between *P. pacificus* and *C. elegans*: whereas *C. elegans* has a grinder for bacteria and no sulfolipids (Supplementary Figure S9), *P. pacificus* does not have a grinder and has many sulfolipids. In addition, the well-studied sulfolipid SDS has been shown to cause cellular rearrangement and membrane perforation in *E. coli* at 1.7 mM[48], and we estimate that the local concentrations of sulfolipids in a single worm of *P. pacificus* to be on the order of 0.3-0.5 mM (we estimate to collect approximately 60 mg sulfolipid from 30 million worms). These sulfates may be an indicator of divergent evolution from mechanical digestion (via grinder, which *Pristionchus* lacks) to chemical digestion.

Through isotopically labelled leucine and valine feeding experiments, we showed that the branched chain sulfolipids from *P. pacificus* are derived from branched chain fatty acid biosynthesis. We also provide evidence that the sulfolipids are produced by *P. pacificus de novo.* This is through the discovery of different sulfolipids across the metabolomes of various strains of *P. pacificus*, the presence of sulfolipids within starved J2 *P. pacificus*, and the exhibition of regulation of sulfolipid abundance by sulfotransferase mutants. Although we have not proven the source of sulfate, it has

been shown that 3'-phosphoadenosine 5'-phosphosulfate (PAPS) is a universal donor of sulfate in sulfation reactions and has been shown to be produced by inorganic sulfate[52].

Metabolomics

We believe that the substrate affected by *eud-1* and *seud-1* may be involved in the mouth formation event. Although characterization of the metabolomic changes within these samples has remained challenging, going forward we plan to systematically identify molecules that are differential using analytical tools such as Metaboseek (metaboseek.com). Although initially starting out untargeted, we will prioritize potential candidates that share characteristics with known sulfolipid signaling molecules such as: the sulfonolipid RIF-1 produced by a prey bacterium regulates morphogenesis in its predator[53], sulfatides, sulfonosteroids, or sulfated N-acetylglycosaminoglycans (their enzymatic machinery has been found to colocalize with eud-1[44]). Through preliminary results we have found differentially regulated features (Supplementary Figure S12) that are to be confirmed.

2.5 METHODS

Mouth Form Assay

Extracts are suspended in 100% ethanol (150 uL). These extracts are subsequently diluted in 1:1 H_2O:EtOH and plated on 35 mm NGM plates in a sterile hood. After sitting for ~4 hr OP50 (60 uL) is placed on the plates. After the bacteria are allowed to grow

overnight two worms are put on each plate and then monitored until the F1 are adults. The adult F1 population are then screened for their mouth form.

Metabolomic analysis

The GNPN database and Metaboseek were used as analytical tools to find connections in MSMS spectra. UHPLC-MS protocol is described in the previous chapter.

P. pacificus strains and culture conditions

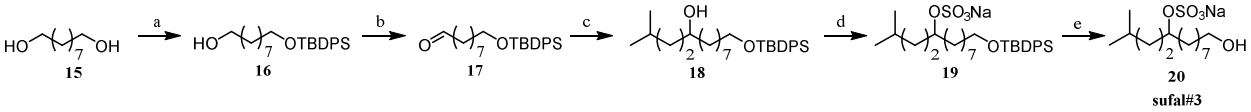
RS2333 exometabolome extracts were used to identify the cyclic sulfates and other sulfolipids. eud-1 (tu445), seud-1 (iub7), RSB020, and RS2333 metabolomic extracts were used to find sulfatase-specific metabolites. Mixed stage worms from a populated 10 cm NGM agar plate seeded with *E. coli* OP50 were washed into 25 ml of S-complete medium and fed OP50 on days 1, 3 and 5 for a 7-day culture period, while shaking at 22C, 220 r.p.m. The cultures were then centrifuged and worm pellets and supernatant frozen separately.

Population Dependent Developmental Acceleration Assay

Developmental acceleration assays, measuring the time point of the first egg lay[4], were performed using sufal#3 or SDS. The compound or pure methanol (as a control) was then spread on 3.5-cm NGM plates, and plates were dried for 30 min, seeded with 35 µl OP50 *E. coli* bacteria and incubated at 20 °C overnight. These plates were then used in

acceleration assays, measuring the time point of first egg laying, as described[47]. Developmental acceleration (in hours) was calculated as the time difference between the time of the first egg lay on untreated ISO plates (plates with a single worm) minus the time of the first egg lay on conditioned ISO (or high-density (HD)) plates. Assays were repeated at least three times, with 10–15 ISO plates used for assaying each metabolome sample.

Chemical syntheses



Supplementary Scheme 4. Overview of synthesis of sufal#3. Reagents and conditions: (a) TBDPS-Cl, imidazole, THF; (b) PCC; (c) 1-bromo-3-methyl-butane, Mg, THF (d) SO3/Pyr, Pyridine; (e) acetyl chloride, MeOH.

Preparation of **16**

15 (4.21 g, 26.3 mmol) was stirred with imidazole (2.15 g, 31.5 mmol, 1.2 eq) in dichloromethane for 35 minutes. Subsequently TBDPS-Cl (4.82 mL, 18.4 mmol, 0.7 eq) was added to the mix. After stirring for 10 min, the reaction was quenched with saturated sodium bicarbonate and extracted with dichloromethane (3 x 15 mL) and dried over magnesium sulfate. The filtrate was concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% methanol in dichloromethane afforded 16 (4.68 g, 11.74 mmol, 45%). ^{1}H NMR (500 MHz, chloroform-*d*): δ (p.p.m.)

7.66 (m, 4 H), 7.39 (m, 6 H), 3.64-3.63 (m, 4 H), 1.59-1.51 (m, 4 H), 1.37 – 1.23 (m, 10 H), 1.04 (s, 9 H).

Preparation of **17**

Oxalyl chloride (3.58 g, 28.2 mmol, 3.6 eq) was cooled to -78C in dichloromethane in a flame-dried flask. DMSO (2.752 mL, 35.22 mmol, 2.75 eq) was added and the reaction was stirred for 20 min. **16** (4.68 g, 11.74 mmol, 1 eq) in dichloromethane was added to the reaction and stirred for 20 min. Triethylamine (8.32 g, 82.18 mmol, 7 eq) was added to the reaction mixture and then the mixture was warmed to room temperature. After stirring for 1 hr, the reaction was quenched with saturated sodium bicarbonate and extracted with dichloromethane (3 x 40 mL), the organic layer was washed with brine, and dried with magnesium sulfate. Through crude NMR it was found that the methoxy acetal was formed, so the reaction was treated with HCl (1M, 10 mL) in acetonitrile overnight. The reaction was quenched with saturated sodium bicarbonate, extracted in dichloromethane, and concentrated in vacuo. Flash column chromatography on silica using a gradient of 0-100% methanol in dichloromethane afforded 17 (1.55 g, 3.95 mmol, 34%). ^{1}H NMR (400 MHz, chloroform-*d*): δ (p.p.m.) 9.76 (t, *J* = 2.0 Hz, 1 H), 7.66 (m, 4 H), 7.37 (m, 6 H), 3.64 (t, *J* = 6.40 Hz, 2 H), 2.41 (td, J =7.29 Hz, 2.0 Hz, 2 H), 1.61 (p, J = 7.08 Hz, 2 H), 1.55 (p, J = 7.08 Hz, 2H), 1.36 – 1.24 (m, 8 H), 1.20 – 1.15 (m, 2 H), 1.03 (s, 9 H).

Preparation of **18**

1,2 dibromoethane (4 drops) were added to Mg turnings (365 mg, 15 mmol, 1.5 eq) in THF (5 mL) under argon. The mixture was heated to reflux and stirred for 5 min. 3-methylbromobutane (1.25 ml, 10 mmol) in THF (15 ml) was added dropwise to the solution of activated Mg turnings. The mixture was then refluxed at 110C for 1 h and turned amber, then dark green. After cooling to room temperature, the mixture was added to a solution of **17** (1.55 g, 3.95 mmol) in THF at -20C. After stirring at -20C for 30 min, the reaction was warmed to room temperature and subsequently stirred for 20 min before being quenched with saturated NH_4Cl (10 ml), extracted with hexanes, dried on magnesium sulfate, and concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% ethyl acetate in hexanes afforded **18** (1.31 g, 2.8 mmol, 70%). ^{1}H NMR (500 MHz, chloroform-d): δ (p.p.m.) 7.66 (m, 4 H), 7.37 (m, 6 H), 3.64 (t, *J* = 6.40 Hz, 2 H), 3.56 (m, 1 H), 1.55 (m, 4 H), 1.47 ‑ 1.12 (m, 15 H), 1.04 (s, 9 H), 0.91 ‑ 0.84 (m, 6H).

Preparation of **19**

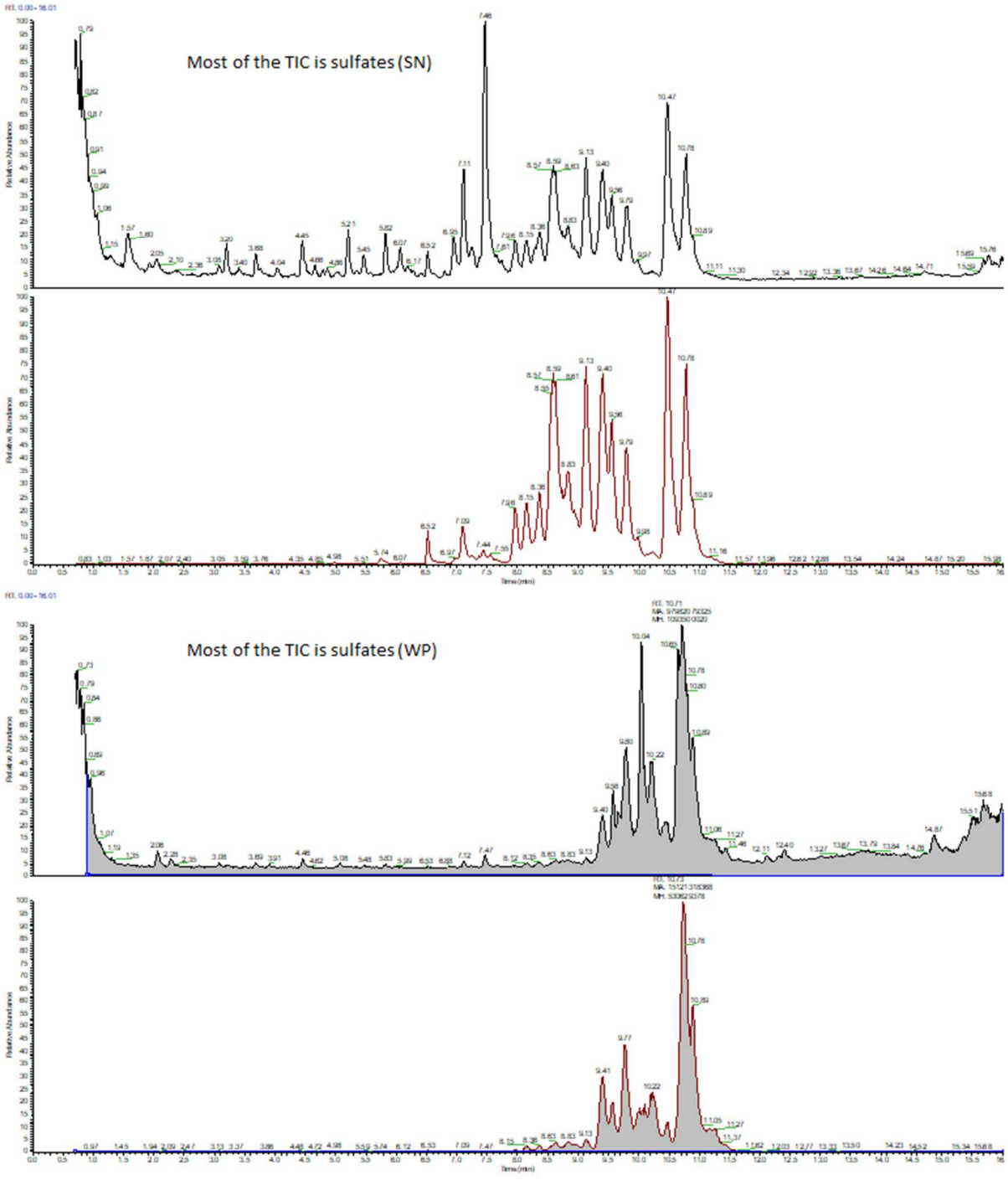
Sulfur trioxide-pyridine complex (1.00 g, 6.28 mmol, 2.2 eq) was added to a solution of **18** (1.31 g, 2.8 mmol) in dry dimethyl formamide (10 ml) at 0C. The reaction was warmed to room temperature and stirred for 20 min. The reaction was then cooled to 0C and neutralized with saturated sodium bicarbonate solution (aq). The reaction was

concentrated in vacuo and extracted with hexanes. The organic layer was dried with magnesium sulfate, filtered, and concentrated in vacuo. The concentrate was resuspended in dichloromethane and flash column chromatography on silica using a gradient of 0‑100% methanol in DCM afforded **19** at 70% purity (0.44 g, 0.8 mmol, 32% yield). ^{1}H NMR (500 MHz, chloroform-d): δ (p.p.m.) 7.64 (m, 4 H), 7.35 (m, 6 H), 4.35 (p, J = 6.05 Hz, 1 H), 3.62 (t, J = 6.60 Hz, 2 H), 1.64 ‑ 1.57 (m, 4 H), 1.56 ‑ 1.46 (m, 4H), 1.34 ‑ 1.15 (11 H), 1.02 (s, 9 H), 0.85 (dd, J = 6.55 Hz, 3.14 Hz, 6 H).

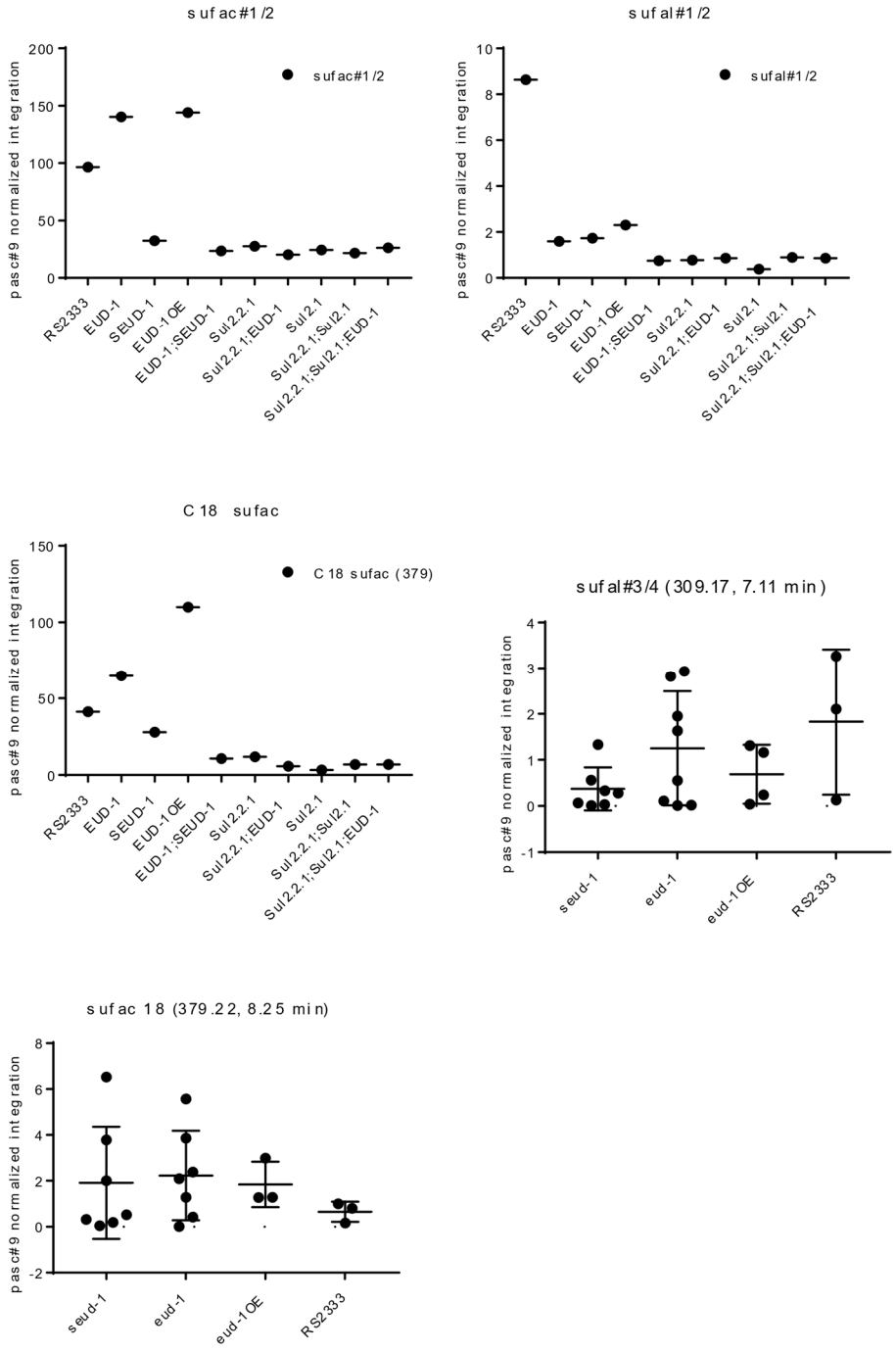
Preparation of **20**, **sufal#3**

A catalytic amount of acetyl chloride (8.56 µl, 0.12 mmol, 0.15 eq) was added to a solution of **7** (0.41 g, 0.77 mmol) in dry methanol (10 mL). The mixture was stirred for 5 hr and the reaction was quenched with saturated sodium bicarbonate solution (aq). The solution was concentrated in vacuo. Flash column chromatography on silica using a gradient of 0‑100% methanol in DCM afforded **20, sufal#3** (116.9 mg, 0.54 mmol, 67% yield). Spectroscopic data were in agreement with those for the natural product: ^{1}H NMR (500 MHz, methanol-d_4): δ (p.p.m.) 4.32 (quin, 5.82 Hz, 1 H), 3.53 (t, 6.86 Hz, 2 H), 1.69‑1.59 (m, 4 H), 1.57‑1.49 (m, 3 H), 1.43‑1.23 (m, 12 H), 0.91 (d, 6.64 Hz, 3 H), 0.89 (d, 6.48 Hz, 3 H); ^{13}C NMR (800 MHz, methanol-d_4): δ (p.p.m.) 81.48, 63.03, 35.25, 35.15, 33.68, 33.13, 30.74, 30.67, 30.54, 29.29, 26.92, 26.02, 23.04, 22.95.

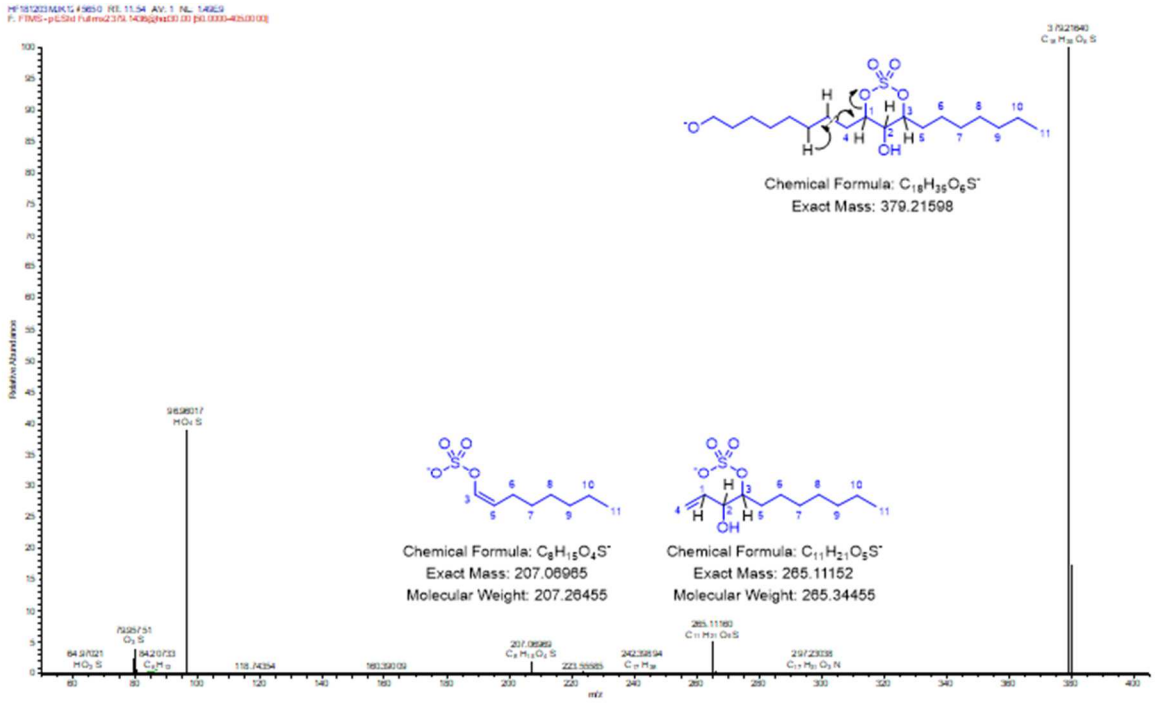
2.6 SUPPLEMENTARY FIGURES



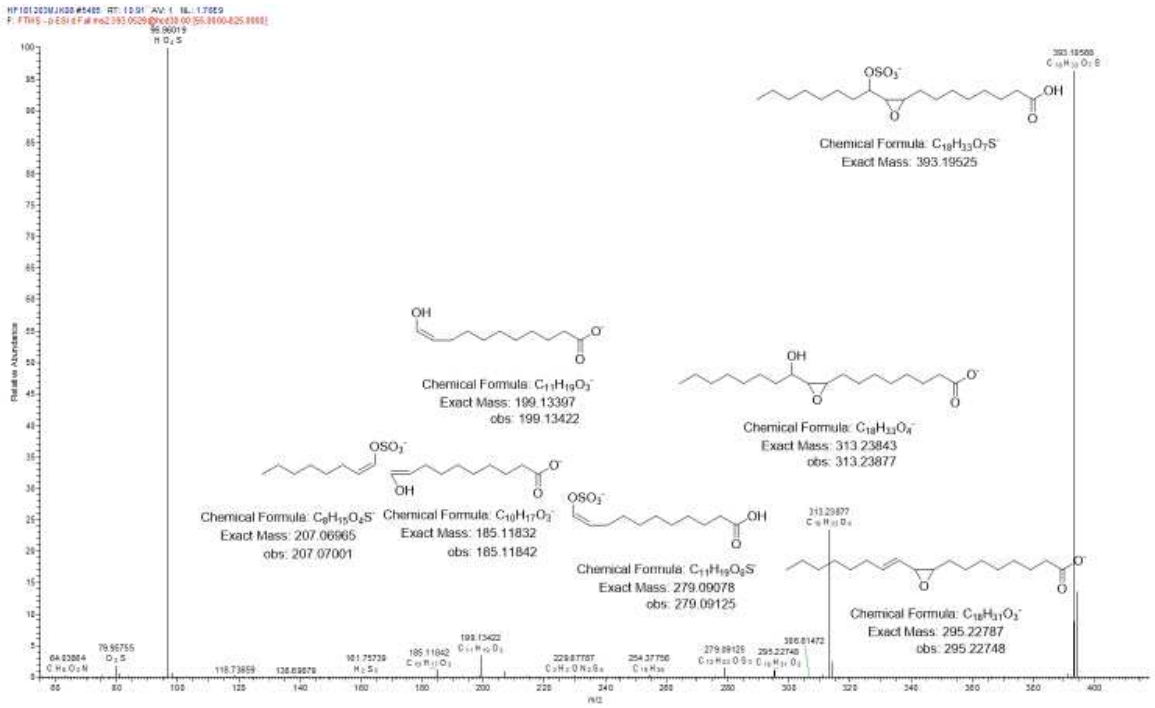
Supplementary Figure S3. TIC and EIC for all known sulfolipids in **a.** exometabolome and **b.** endometabolome.



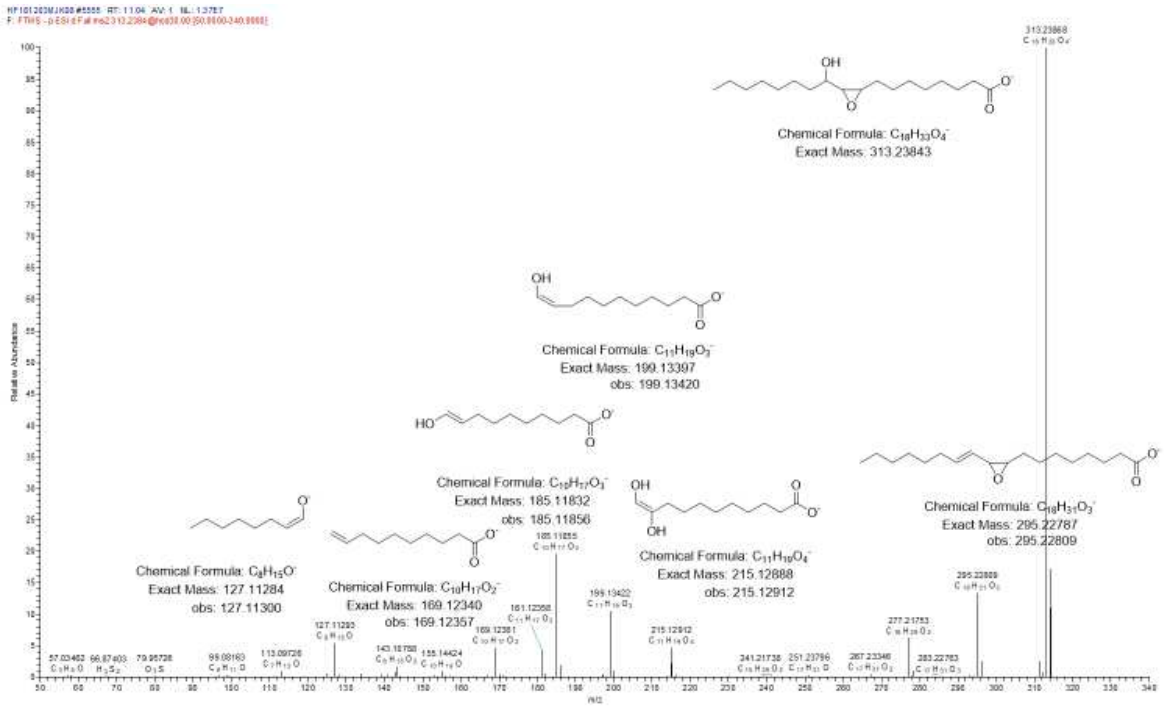
Supplementary Figure S4. Pasc#9 normalized abundance of sulfolipids from sulfatase mutants eud-1 and seud-1.



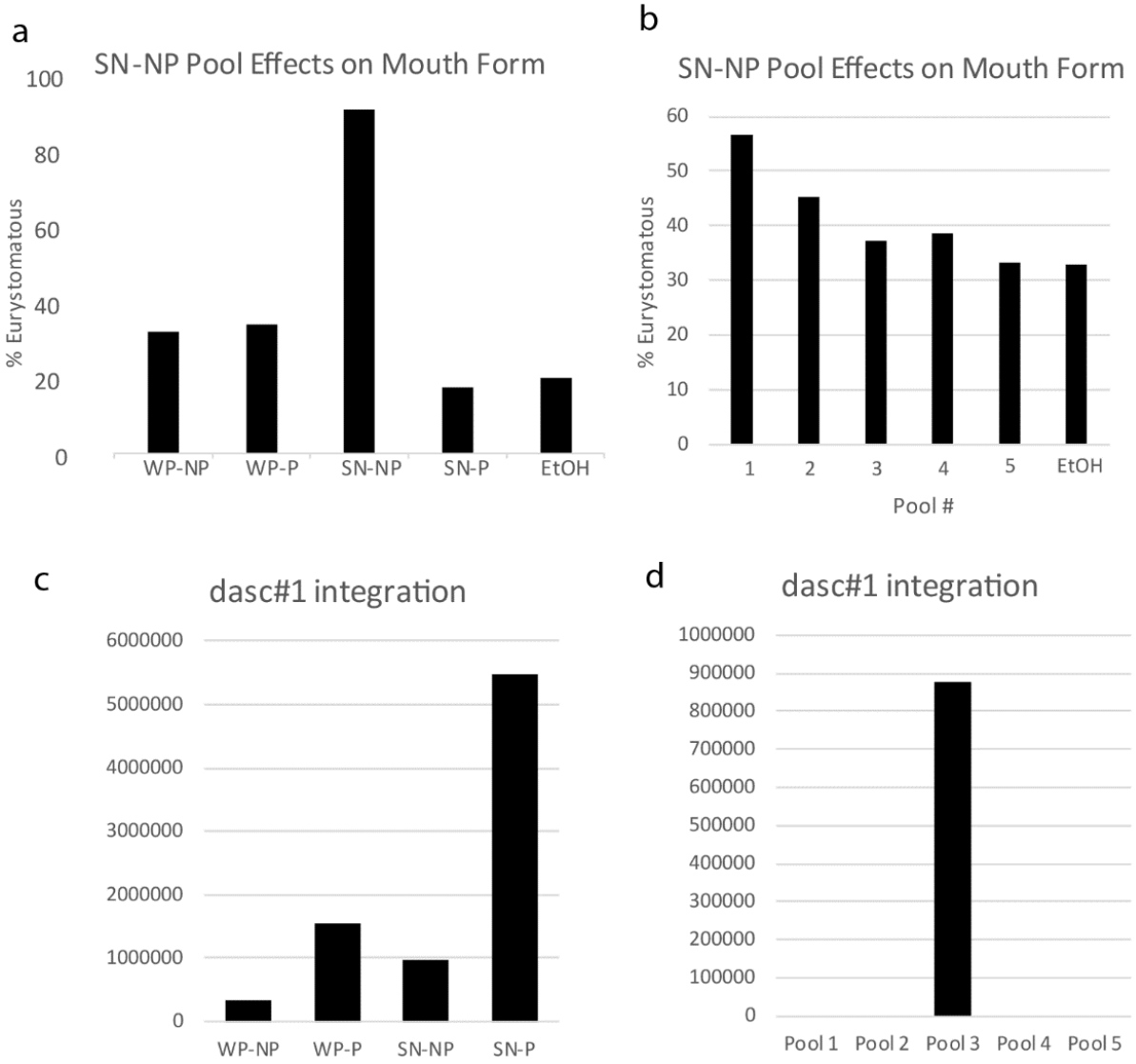
Supplementary Figure S5. MSMS spectrum for cysul#1.



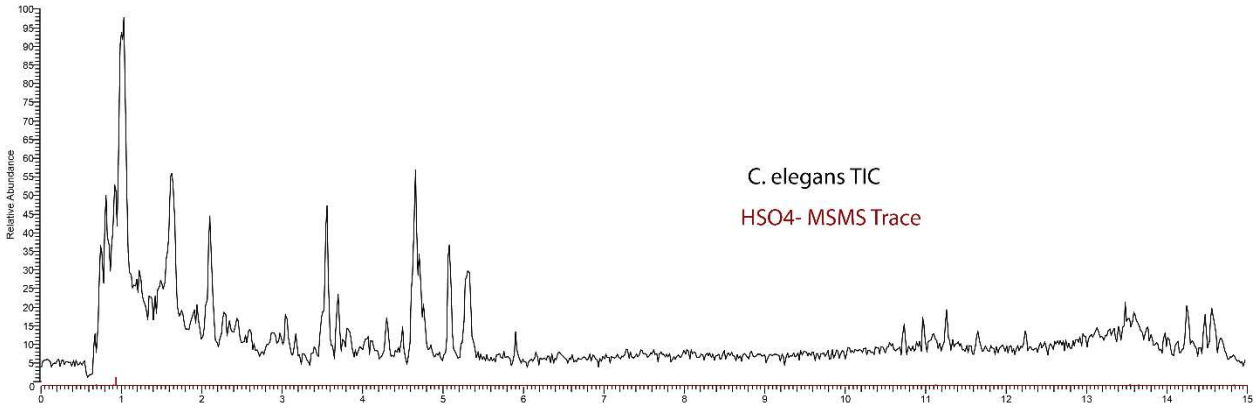
Supplementary Figure S6. MSMS spectrum for epos#1.



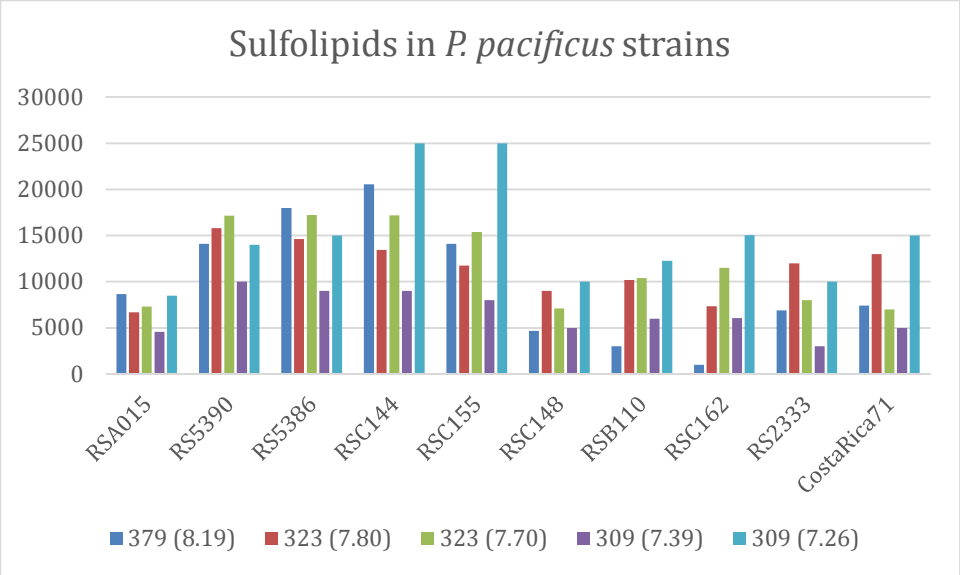
Supplementary Figure S7. MSMS spectrum for epoxide intermediate.



Supplementary Figure S8. Mouth form assays with **a.** nonpolar (ethyl acetate) and polar (ethanol) supernatant (sn) and pellet (wp) extracts from *P. pacificus* metabolome and **b.** fraction pools from nonpolar supernatant extract **c.** integration of dasc#1 in metabolome extracts and **d**. pools from nonpolar supernatant extract.

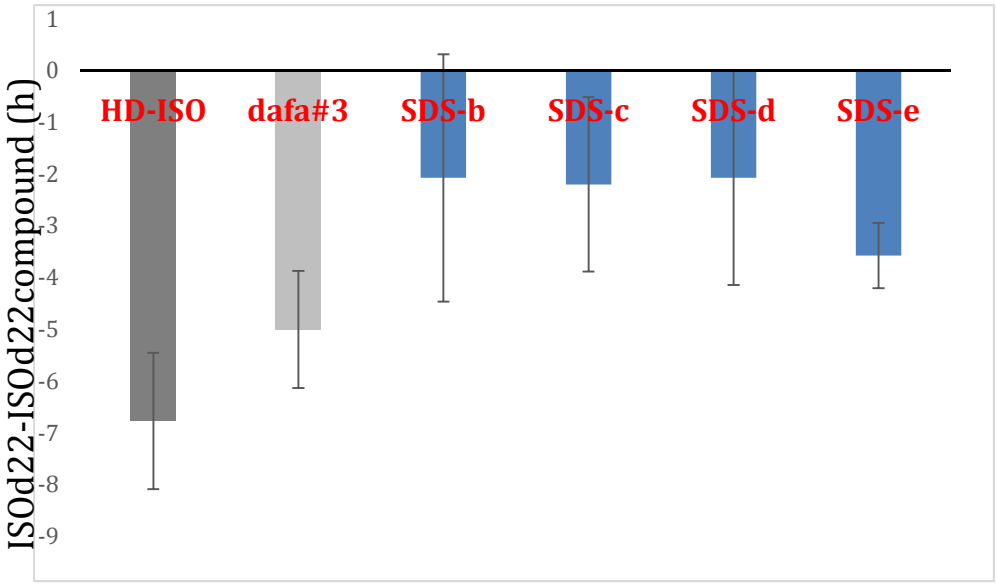


Supplementary Figure S9. TIC of *C. elegans* exometabolome with MSMS EIC of HSO_4^-

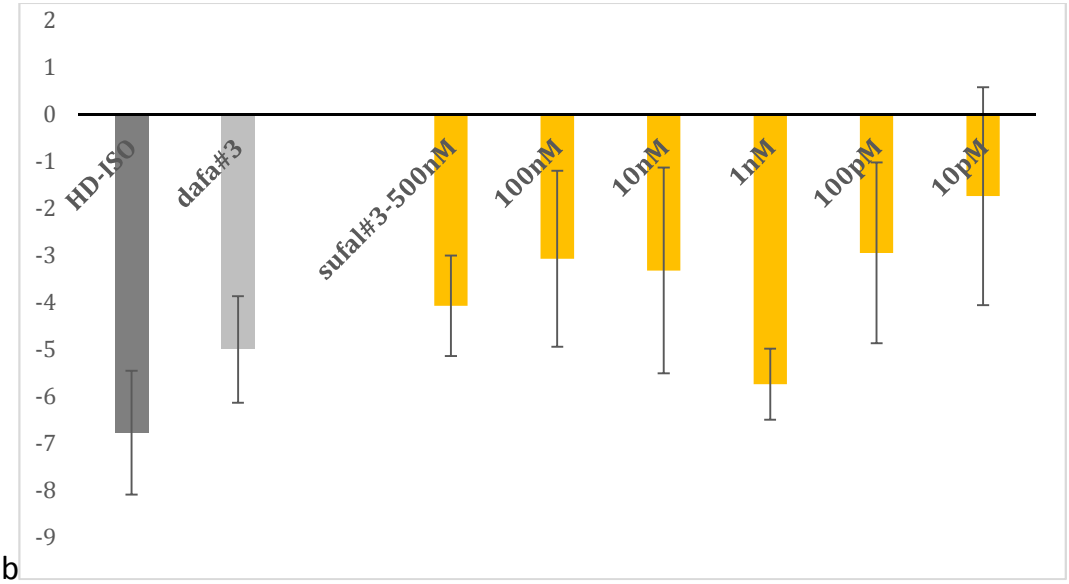


Supplementary Figure S10. Abundance of sulfolipids in various *P. pacificus* strains

a

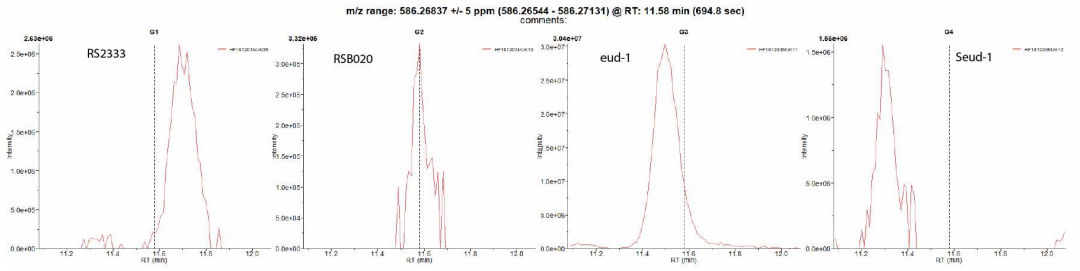


b

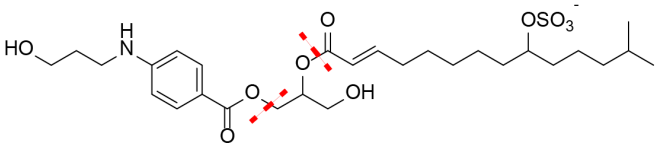


Supplementary Figure S11. Population density dependent development acceleration assay in *daf-22* (ok693) using a. SDS (b = 0.1%, c = 0.01%, d = 0.001%, e = 0.0001%) and b. sufal#3 with known developmental accelerator dafa#3 (100 nM).

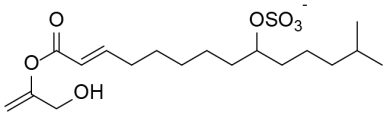
a



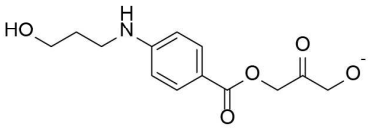
Hypothetical Structure based on MSMS Fragmentation



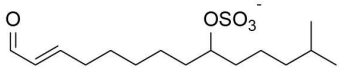
Chemical Formula: $C_{28}H_{44}NO_{10}S^-$
Exact Mass: 586.26914



Chemical Formula: $C_{18}H_{31}O_7S^-$
Exact Mass: 391.17960

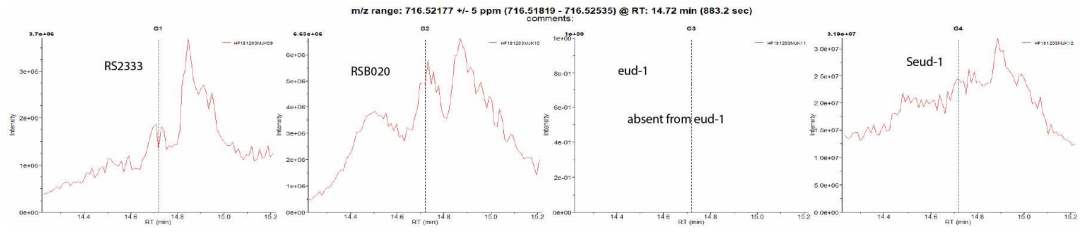


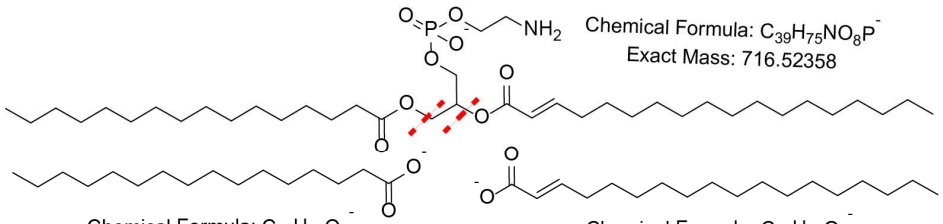
Chemical Formula: $C_{13}H_{16}NO_5^-$
Exact Mass: 266.10340



Chemical Formula: $C_{15}H_{27}O_5S^-$
Exact Mass: 319.15847

b

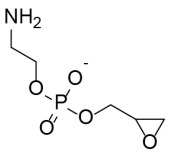




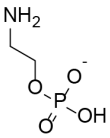
Chemical Formula: $C_{39}H_{75}NO_8P^-$
Exact Mass: 716.52358

Chemical Formula: $C_{16}H_{31}O_2^-$
Exact Mass: 255.23295

Chemical Formula: $C_{18}H_{33}O_2^-$
Exact Mass: 281.24860



Chemical Formula: $C_5H_{11}NO_5P^-$
Exact Mass: 196.03803



Chemical Formula: $C_2H_7NO_4P^-$
Exact Mass: 140.01182

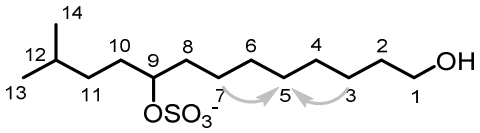
Supplementary Figure S12. Differential features found in sulfotransferase mutant strains and their proposed structures based on MSMS fragmentation.

2.7 SUPPLEMENTARY TABLES

Supplementary Table S4. List of predicted formulae for ions in *P. pacificus* metabolome that cleave a sulfate fragment.

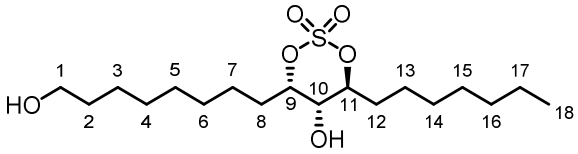
m/z	Predicted formula	SMID	m/z	Predicted formula	SMID
239.05972	C8H15O6S-		351.22141	C17H35O5S-	
253.07516	C9H17O6S-		353.12784	C14H25O8S-	
267.09109	C10H19O6S-		353.16428	C15H29O7S-	
281.10651	C11H21O6S-		353.20032	C16H33O6S-	
281.14301	C12H25O5S-		355.1799	C15H31O7S-	
295.15869	C13H27O5S-		363.22162	C18H35O5S-	
307.12247	C13H23O6S-		365.23703	C18H37O5S-	
309.13812	C13H25O6S-		367.17975	C16H31O7S-	
309.17447	C14H29O5S-	sufal#3/4	379.21634	C18H35O6S-	cysul#1/2
311.15326	C13H27O6S-		395.21063	C18H35O7S-	osul#1
323.19009	C15H31O5S-	sufal#1/2	409.19019	C17H33O8S-	
325.13297	C13H25O7S-		411.20621	C18H35O8S-	
325.1694	C14H29O6S-		435.31519	C23H47O5S-	
335.15378	C15H27O6S-		471.22736	C20H39O10S-	
337.16903	C15H29O6S-	sufac#1/2	485.2431	C21H41O10S-	
337.20596	C16H33O5S-		493.32077	C25H49O7S-	
339.18484	C15H31O6S-		505.32117	C26H49O7S-	
341.16418	C14H29O7S-		533.35211	C28H53O7S-	
349.20575	C17H33O5S-		547.36774	C29H55O7S-	
351.18491	C16H31O6S-		561.38434	C30H57O7S-	

Supplementary Table S5. ^{1}H and ^{13}C NMR spectroscopic data for **sufal#3** in methanol-d_4. Chemical shifts were referenced to $\delta(CD_2\underline{H}OD) = 3.31$ and $\delta(^{13}\underline{C}D_3OD) = 49.0$. ^{13}C chemical shifts were determined via HMBC and HSQC spectra. Natural product and synthetic spectra were acquired using the Bruker Avance 800 and 500 spectrometer, respectively. ^{1}H, ^{1}H-J-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



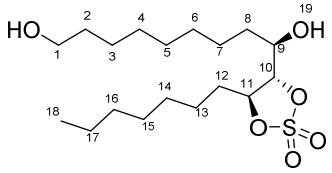
Position	^{13}C [ppm] - Synthetic	^{1}H [ppm] (Synthetic)	^{1}H-^{1}H coupling constants (Hz)	^{13}C [ppm] - Natural	^{1}H [ppm] (Natural)	^{1}H-^{1}H coupling constants (Hz)	HMBC Correlations
1	63.03	1a = 3.54 1b = 3.54	$J_{1,2}$ = 6.72	62.73	1a = 3.53 1b = 3.53	$J_{1,2}$ = 6.72	2, 3
2	33.68	2a = 1.53-1.54 2b = 1.53-1.54	n/a	33.39	2a = 1.53-1.54 2b = 1.53-1.54	n/a	1, 3, 4
3	26.92	3a = 1.32-1.36 3b = 1.32-1.36	n/a	26.65	3a = 1.32-1.36 3b = 1.32-1.36	n/a	1, 2, 4, 5
4	30.74	4a = 1.31-1.35 4b = 1.31-1.35	n/a	30.38	4a = 1.31-1.35 4b = 1.31-1.35	n/a	
5	30.54	5a = 1.32-1.35	n/a	30.5-30.7	5a = 1.32-1.35	n/a	
6	30.68	5b = 1.32-1.35	n/a	30.5-30.7	5b = 1.32-1.35	n/a	
7	26.02	7a = 1.40-1.43 7b = 1.40-1.43	n/a	25.70	7a = 1.40-1.43 7b = 1.40-1.43	n/a	9, 8, 6, 5
8	35.15	8a = 1.60-1.70 8b = 1.60-1.70	n/a	35.15	8a = 1.60-1.70 8b = 1.60-1.70	n/a	10, 9,
9	81.48	4.32	$J_{9,10\ \&\ 9,8}$ = 5.92	80.76	4.32	$J_{9,10\ \&\ 9,8}$ = 6.04	*7, 8, 10*
10	35.25	10a = 1.60-1.70 10b = 1.60-1.70	n/a	35.15	10a = 1.60-1.70 10b = 1.60-1.70	n/a	11, 9, 8,
11	33.14	11a = 1.22-1.30 11b = 1.22-1.30	n/a	33.14	11a = 1.22-1.30 11b = 1.22-1.30	n/a	
12	29.30	1.54	n/a	29.03	1.54	n/a	14, 13, 11, 10
13	22.95	0.91	$J_{13,12}$ = 6.56	22.70	0.91	$J_{13,12}$ = 6.81	
14	23.04	0.90	$J_{14,12}$ = 6.46	22.75	0.89	$J_{14,12}$ = 6.68	13, 12, 10

Supplementary Table S6. ^{1}H and ^{13}C NMR spectroscopic data for **cysul#1** in methanol-d_4. Chemical shifts were referenced to δ ($CD_2\underline{H}OD$) = 3.31 and δ ($^{13}\underline{C}D_3OD$) = 49.0. ^{13}C chemical shifts were determined via HMBC and HSQC spectra. Spectra were acquired using the Bruker Avance 800 spectrometer. ^{1}H, ^{1}H-J-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom. ROESY in ACN.



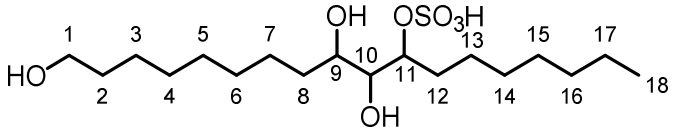
Position	^{13}C[ppm]	^{1}H [ppm]	1H-1H coupling constants (Hz)	HMBC Correlations	ROESY Correlations
1	62.72	1a = 3.54 1a = 3.54	$J_{1,2}$ = 6.69	2, 3	
2	33.40	2a = 1.53 2b = 1.53	$J_{2,3}$ = 5.93	1, 3, 4	
3	26.64	3a = 1.36 3b = 1.36	n/a	4	
4	30.20	4a = 1.35 4b = 1.35	n/a		
5	28.80-30.20	5a = 1.30-1.40 5b = 1.30-1.40	n/a		
6	29.68	6a = 1.38 6b = 1.35-1.40	n/a		
7	25.49	7a = 1.42 7b = 1.49	n/a		
8	30.11	8a = 1.93 8b = 1.74	n/a	6, 7, 9, 10	
9	87.00	4.82	m	7, 8, 10	10
10	68.76	3.72	J = 3.79 J = 2.48	8, 9	9, 8 (weak)
11	91.92	4.55	J = 9.88 J = 4.17	9, 10, 12, 13	
12	32.14	12a = 2.05 12b = 1.79	m	10, 11, 13, 14	
13	26.06	13a = 1.56 13b = 1.44	n/a	--	
14	29.78	14a = 1.39 14b = 1.30-1.40	n/a	--	
15	32.32	15a = 1.30-1.40 15b = 1.30-1.40	n/a	--	
16	32.57	16a = 1.33 16b = 1.34	n/a	--	
17	23.37	17a = 1.34 17b = 1.34	n/a	18, 16, 15	
18	14.07	0.91	$J_{17,18}$ =	17, 16	
19	OH	3.79			

Supplementary Table S7. ^{1}H and ^{13}C NMR spectroscopic data for **cysul#2** in methanol-d_4. Chemical shifts were referenced to δ ($CD_2\underline{H}OD$) = 3.31 and δ ($^{13}\underline{C}D_3OD$) = 49.0. ^{13}C chemical shifts were determined via HMBC and HSQC spectra. Spectra were acquired using the Bruker Avance 800 spectrometer. ^{1}H, ^{1}H-J-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom. ROESY in ACN.



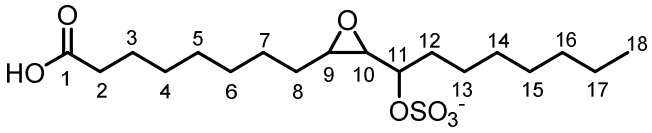
Position	^{13}C [ppm]	^{1}H [ppm]	^{1}H-^{1}H coupling constants (Hz)	HMBC Correlations	ROESY correlations
1	62.72	1a = 3.54 1b = 3.54	$J_{1,2}$ = 6.69	2, 3	
2	33.40	2a = 1.53 2b = 1.53	$J_{2,3}$ = 5.93	1, 3, 4	
3	26.64	3a = 1.36 3b = 1.36	n/a	1, 2	
4	30.23	4a = 1.35 4b = 1.35	n/a	--	
5	29.80-30.20	5a = 1.30-1.40 5b = 1.30-1.40	n/a	--	
6	29.97	6a = 1.35 6b = 1.30-1.40	n/a	--	
7	26.49	7a = 1.54 7b = 1.39	n/a	--	
8	33.92	8a = 1.58 8b = 1.60	n/a	6, 7, 9, 10	9
9	69.08	3.69	m	7, 8	7, 10, 11, 19
10	90.30	4.59	$J_{9,10}$ = 8.28 $J_{10,11}$ = 2.34	8, 9, 11, 12	7, 9, 12, 13
11	85.74	5.01	$J_{9,8}$ = 3.91 $J_{9,10}$ = 8.28	9, 10, 13	8, 12
12	32.82	12a = 1.81 12b = 1.87	M	10, 11, 13, 14	9, 10, 11, 13
13	26.05	13a = 1.48 13b = 1.47	n/a	11, 12, 14	
14	29.87	14a = 1.34 14b = 1.34	n/a	--	
15	32.32	15a = 1.30-1.40 15b = 1.30-1.40	n/a	--	
16	32.57	16a = 1.33 16b = 1.34	n/a	17	
17	23.37	17a = 1.34 17b = 1.34	n/a	15, 16, 18	
18	14.08	0.91	$J_{17,18}$ =	16, 17	
19		3.30			7, 8, 9, 11, 13

Supplementary Table S8. 1H and 13C NMR spectroscopic data for **osul#1** in methanol-d_4. Chemical shifts were referenced to δ ($CD_2\underline{H}OD$) = 3.31 and δ ($^{13}\underline{C}D_3OD$) = 49.0. ^{13}C chemical shifts were determined via HMBC and HSQC spectra. Spectra were acquired using the Bruker Avance 800 spectrometer. ^1H, ^1H-J-coupling constants were determined from the acquired ^1H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



Position	^{13}C [ppm]	^1H [ppm]	^1H-^1H coupling constants (Hz)	HMBC Correlations
1	62.72	1a = 3.54 1b = 3.54	$J_{1,2}$ = 6.73	2, 3
2	33.39	2a = 1.53 2b = 1.53		1, 3, 4
3	26.64	3a = 1.36 3b = 1.36		1, 2
4	30.20	4a = 1.35 4b = 1.35		--
5		5a = 1.30-1.40 5b = 1.30-1.40		--
6	30.24	6a = 1.34 6b = 1.34		--
7	26.72	7a = 1.35-1.50 7b = 1.35-1.50		9, 8, 6
8	30.38	8a = 1.57 8b = 1.70		6, 7, 9, 10
9	81.93	3.35		7, 10, 11
10	74.14	3.62		12, 11, 9, 8
11	81.22	4.42	$J_{11,12}$ = 4.43 $J_{11,10}$ = 10.82	12, 13
12	31.58	12a = 1.82 12b = 1.69		10, 11, 13, 14
13	25.92	13a = 1.40-1.55 13b = 1.40-1.55		--
14	30.10	14a = 1.30-1.40 14b = 1.30-1.40		--
15		15a = 1.30-1.40 15b = 1.30-1.40		--
16	32.83	16a = 1.31 16b = 1.31		17
17	23.60	17a = 1.34 17b = 1.34		18, 16, 15
18	14.14	0.91		17, 16

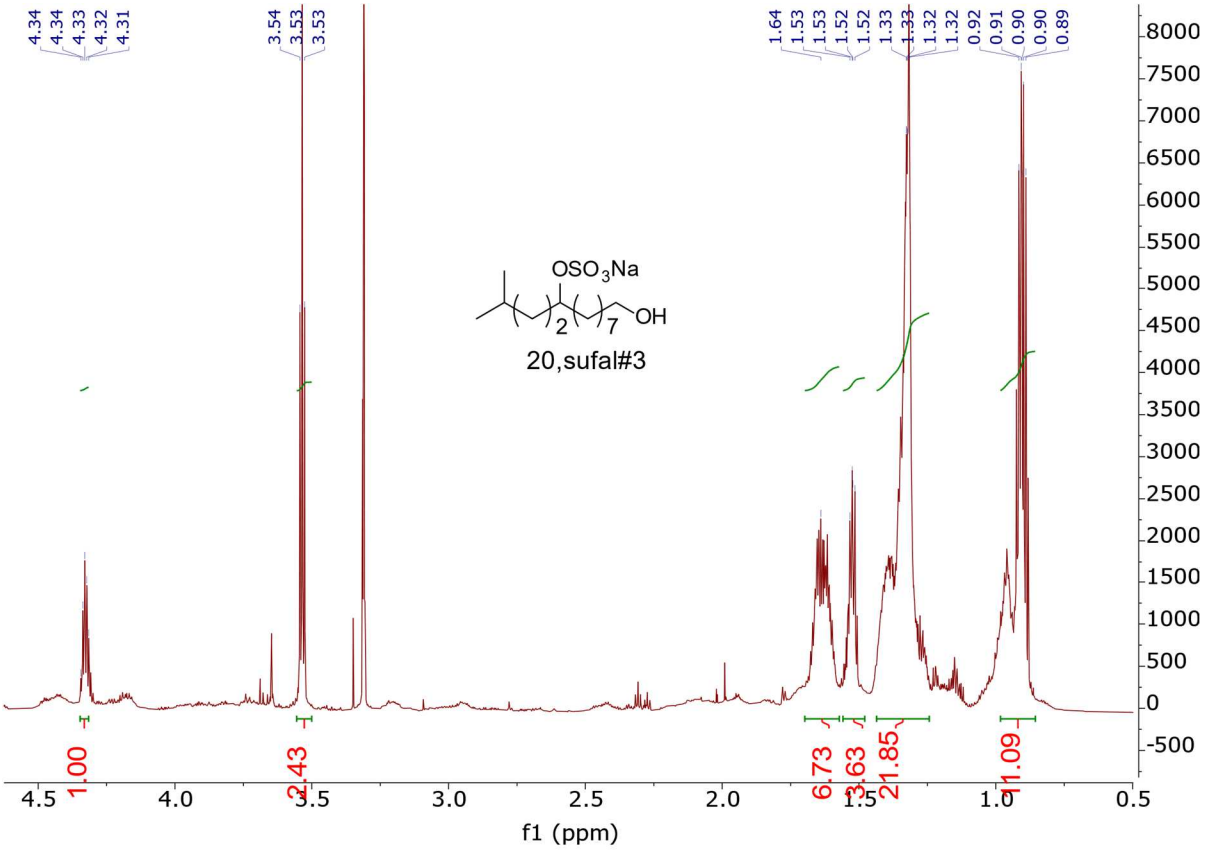
Supplementary Table S9. 1H and 13C NMR spectroscopic data for **epos#1** in methanol-d4. Chemical shifts were referenced to δ (CD2HOD) = 3.31 and δ (13CD3OD) = 49.0. 13C chemical shifts were determined via HMBC and HSQC spectra. Spectra were acquired using the Bruker Avance 800 spectrometer. 1H, 1H-J-coupling constants were determined from the acquired 1H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated 13C atom.



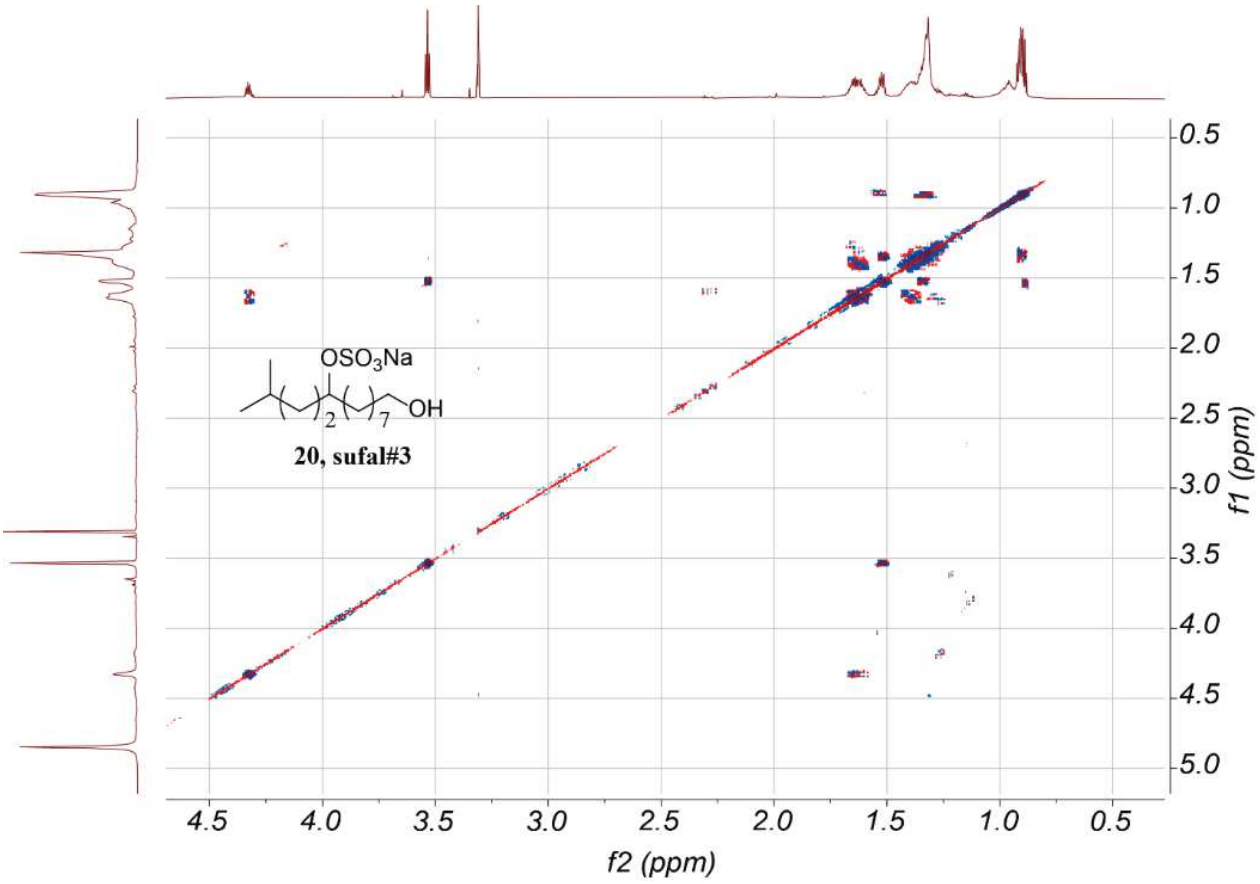
Position	^{13}C [ppm]	^{1}H [ppm] Corrected	^{1}H-^{1}H coupling constants (Hz)	HMBC Correlations
1	177.17			
2	34.29	2a = 2.25 2b = 2.25		1, 3, 4
3	25.05	3a = 1.60 3b = 1.60		1, 2, 4
4	29.02	4a = 1.34 4b = 1.34		2, 3, 5, 6
5	29.41	5a = 1.33-1.38 5b = 1.33-1.38		
6	31.64	6a = 1.28-1.38 6b = 1.28-1.38		
7	26.52	7a = 1.53 7b = 1.48		
8	28.09	8a = 1.91 8b = 1.36		7, 6
9	58.51	3.00		7, 8, 10
10	58.11	3.00		8, 9, 11
11	74.50	4.14		9, 10, 12
12	32.85	12a = 1.75 12b = 1.75		6, 13, 11
13	24.10	13a = 1.53 13b = 1.53		
14	29.35	14a = 1.33-1.38 14b = 1.33-1.38		
15	25.51	15a = 1.28-1.38 15b = 1.28-1.38		
16	31.65	16a = 1.40 16b = 1.40		
17	22.31	17a = 1.33 17b = 1.33		
18	13.05	0.91		

2.8 SUPPLEMENTARY NMR SPECTRA

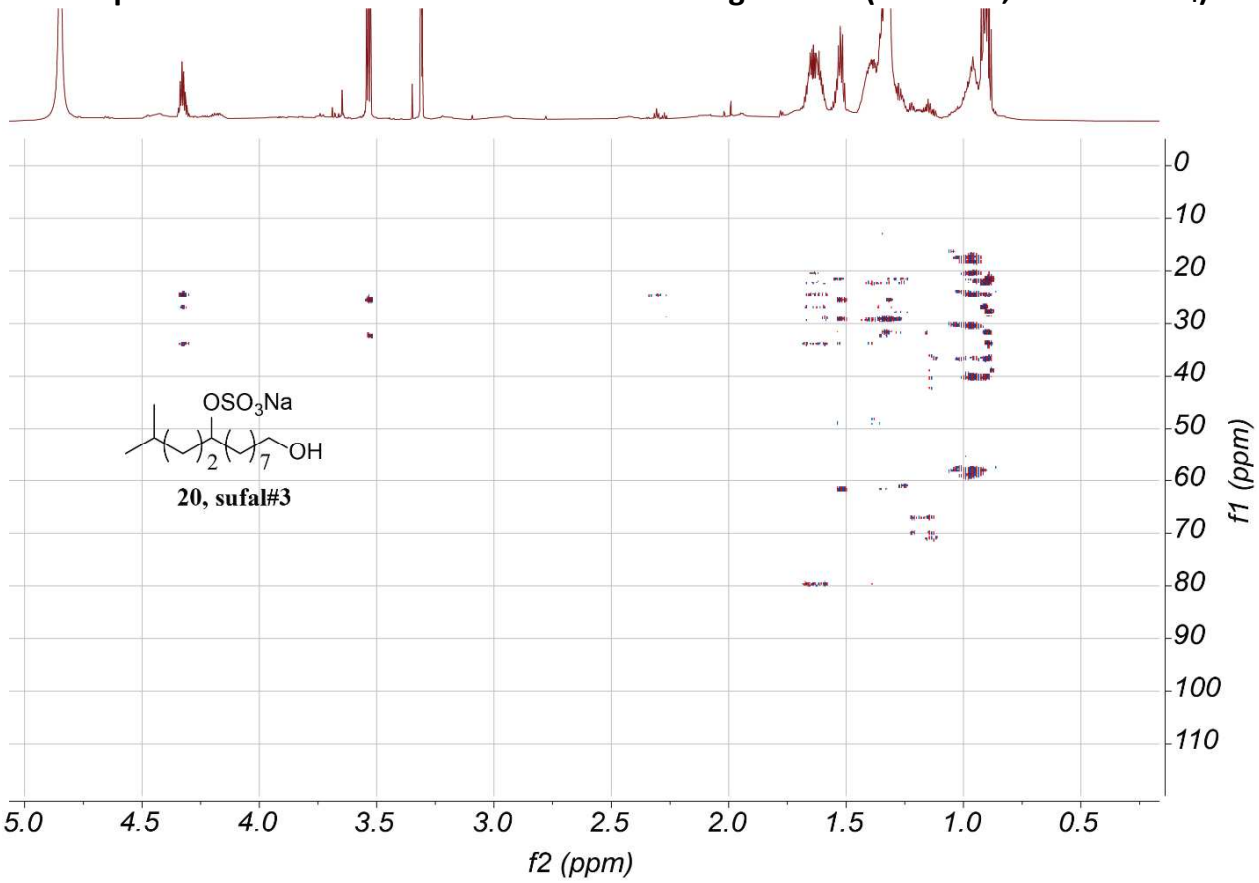
^{1}H NMR spectra of metabolome fraction containing sufal#3 (800 MHz, methanol d_4)



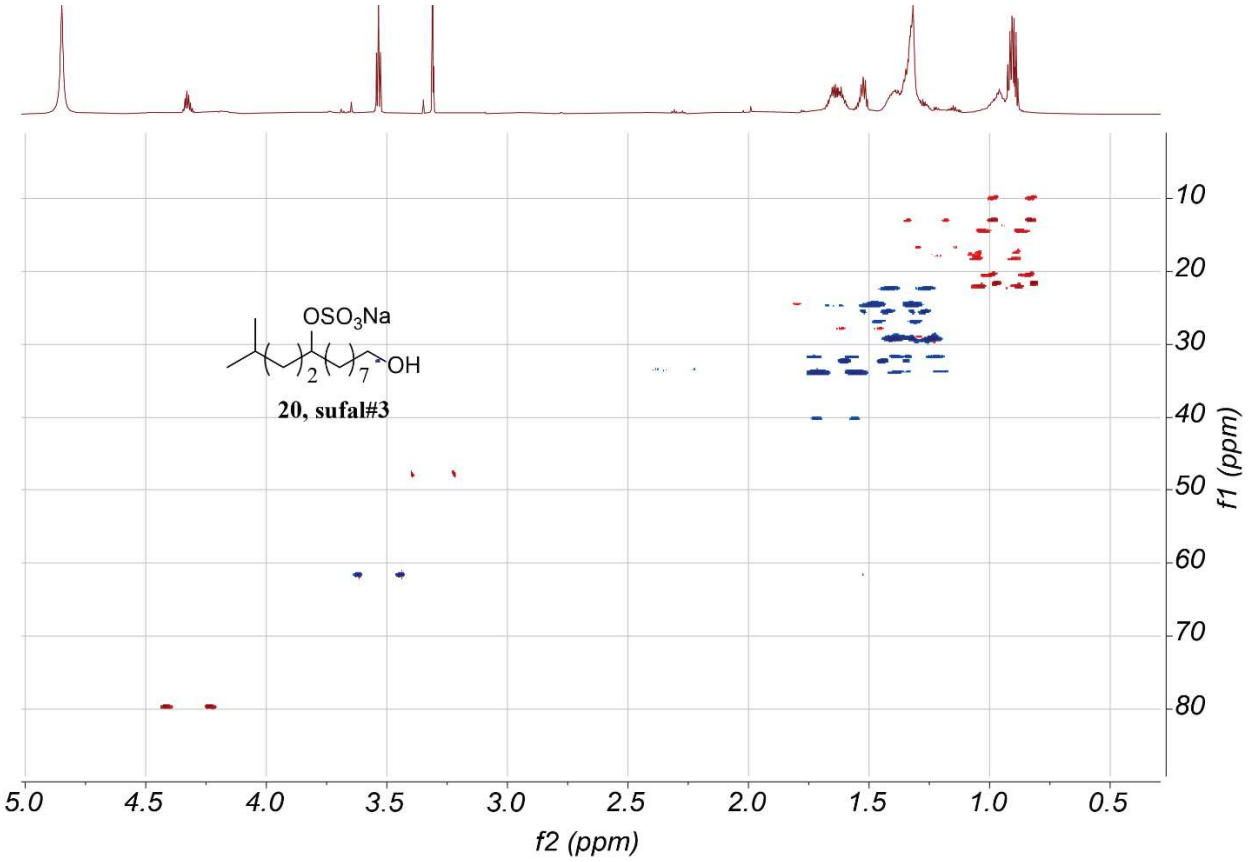
dqfCOSY spectrum of metabolome fraction containing sufal#3 (800 MHz, methanol-d_4)



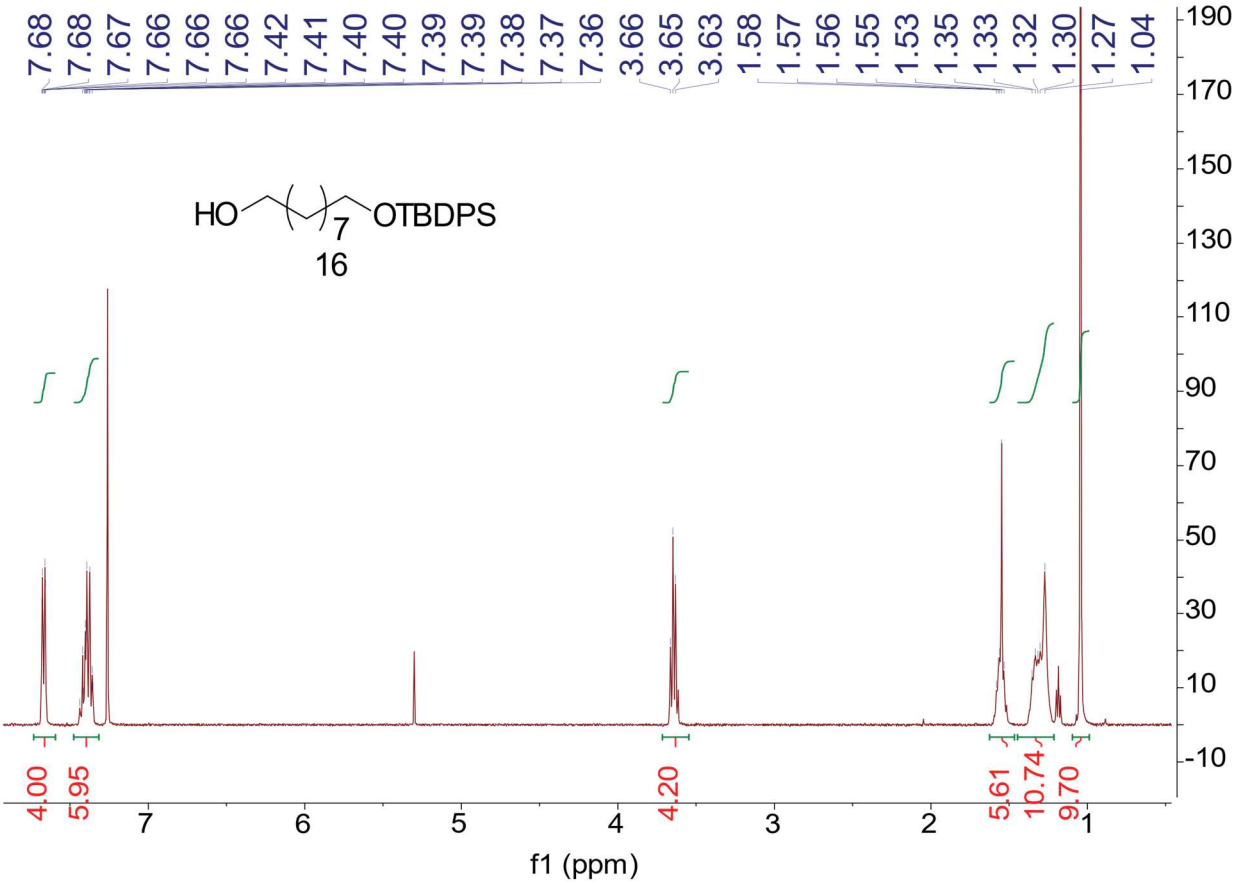
HMBC spectrum of metabolome fraction containing sufal#3 (800 MHz, methanol-d_4)



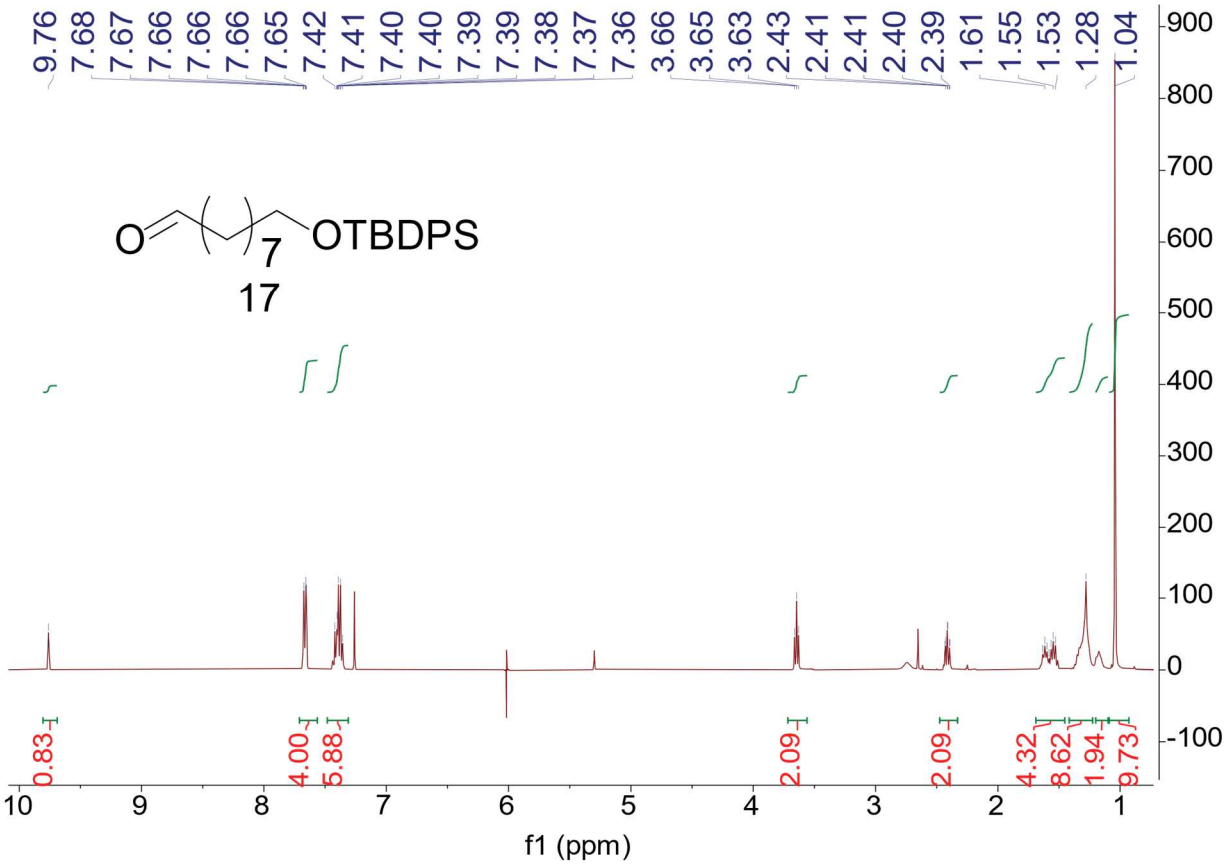
HSQC spectrum of metabolome fraction containing sufal#3 (800 MHz, methanol-d_4)



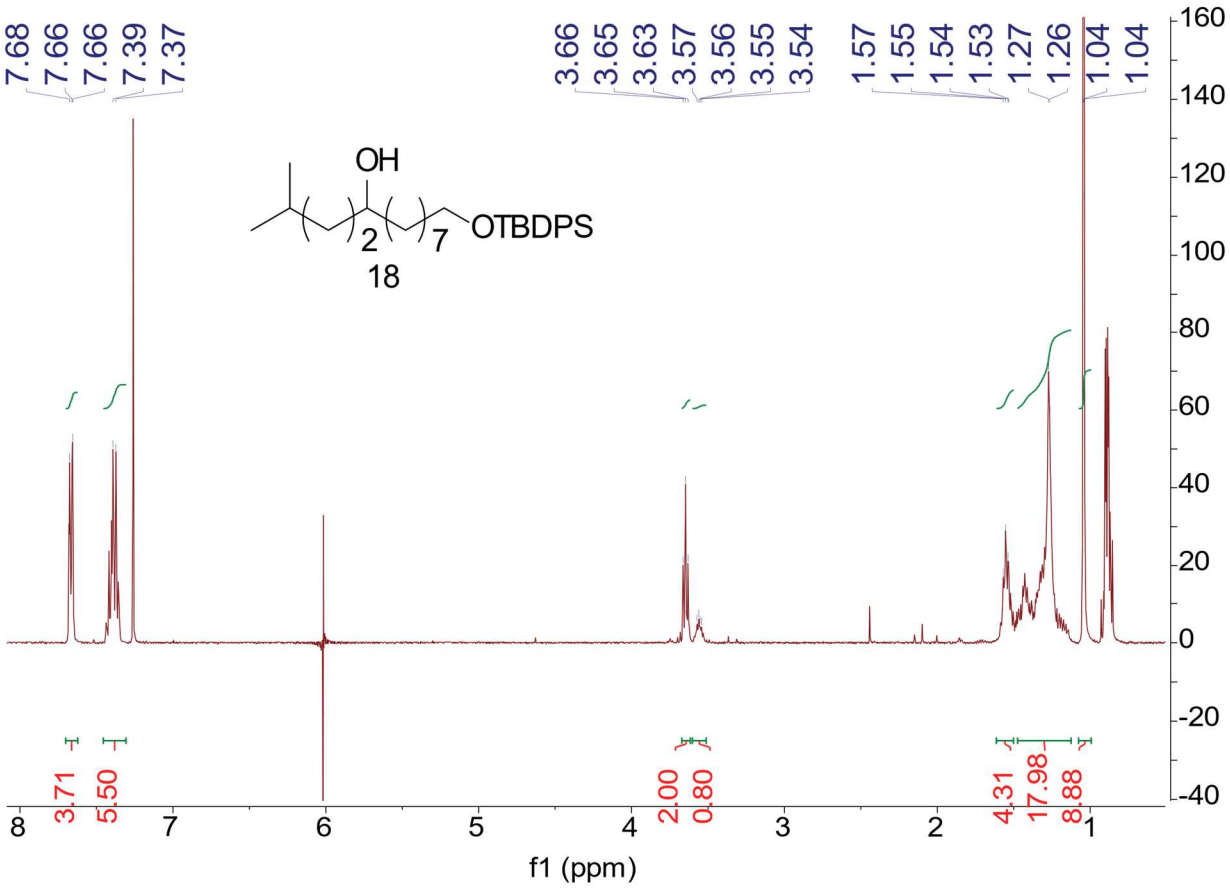
^{1}H NMR Spectrum (600 MHz, Chloroform-*d*) of 16



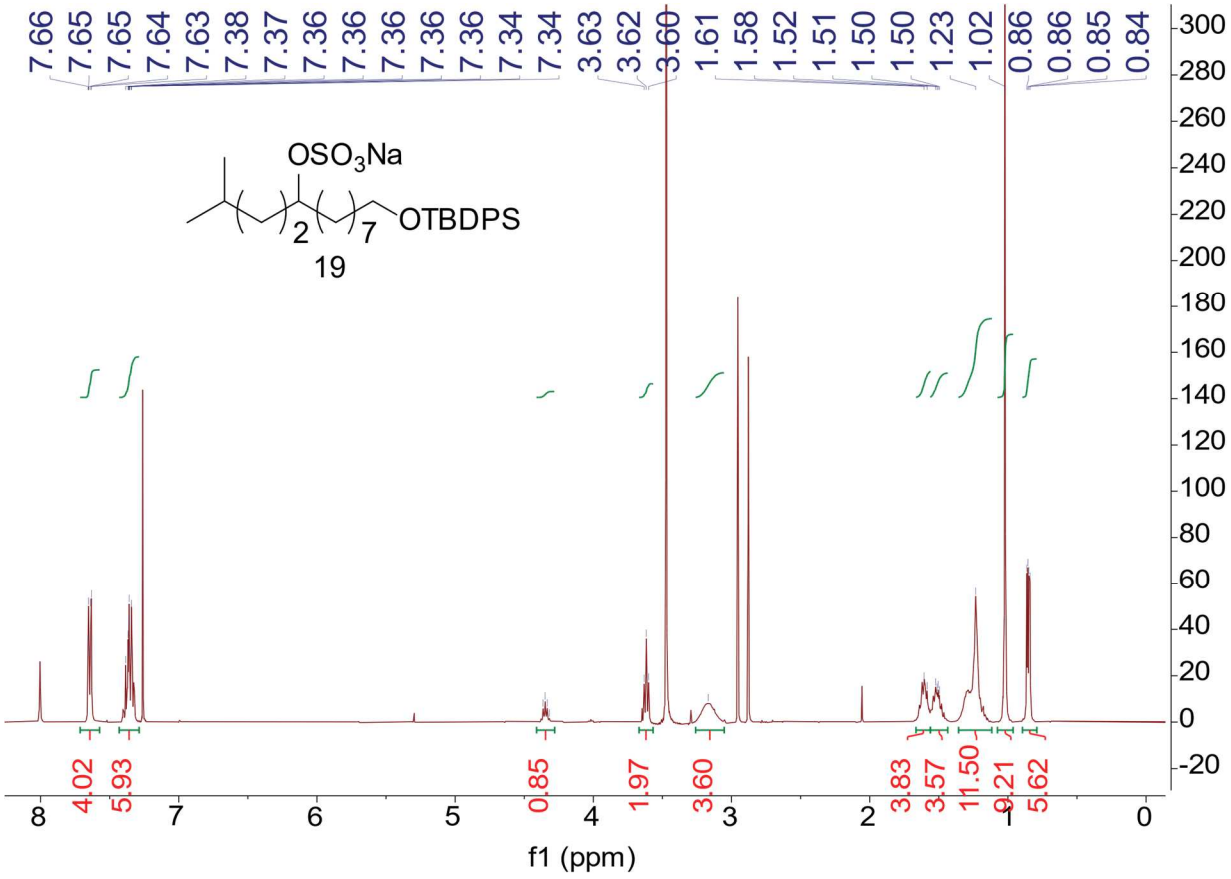
^{1}H NMR Spectrum (600 MHz, methanol-d_4) of 17



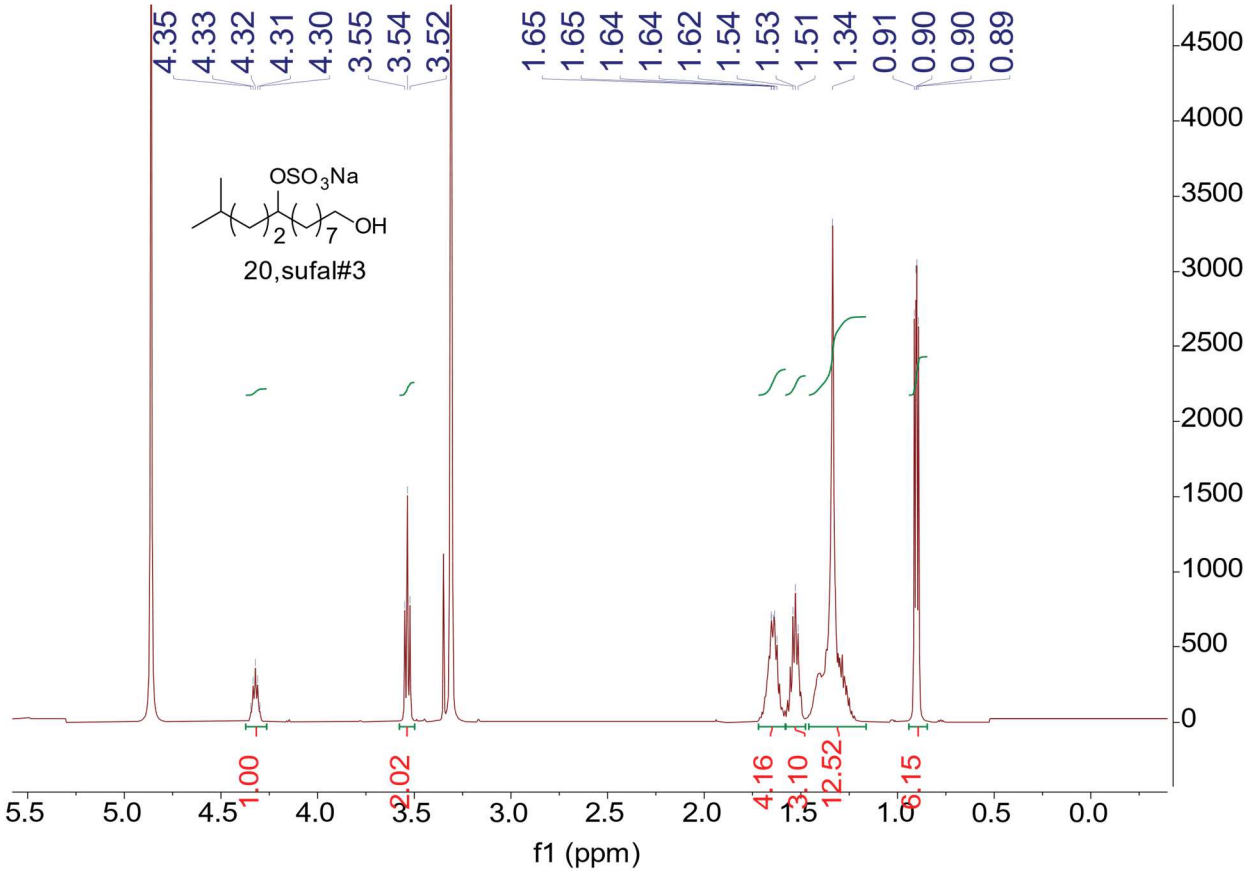
^{1}H NMR Spectrum (600 MHz, methanol-d_4) of 18



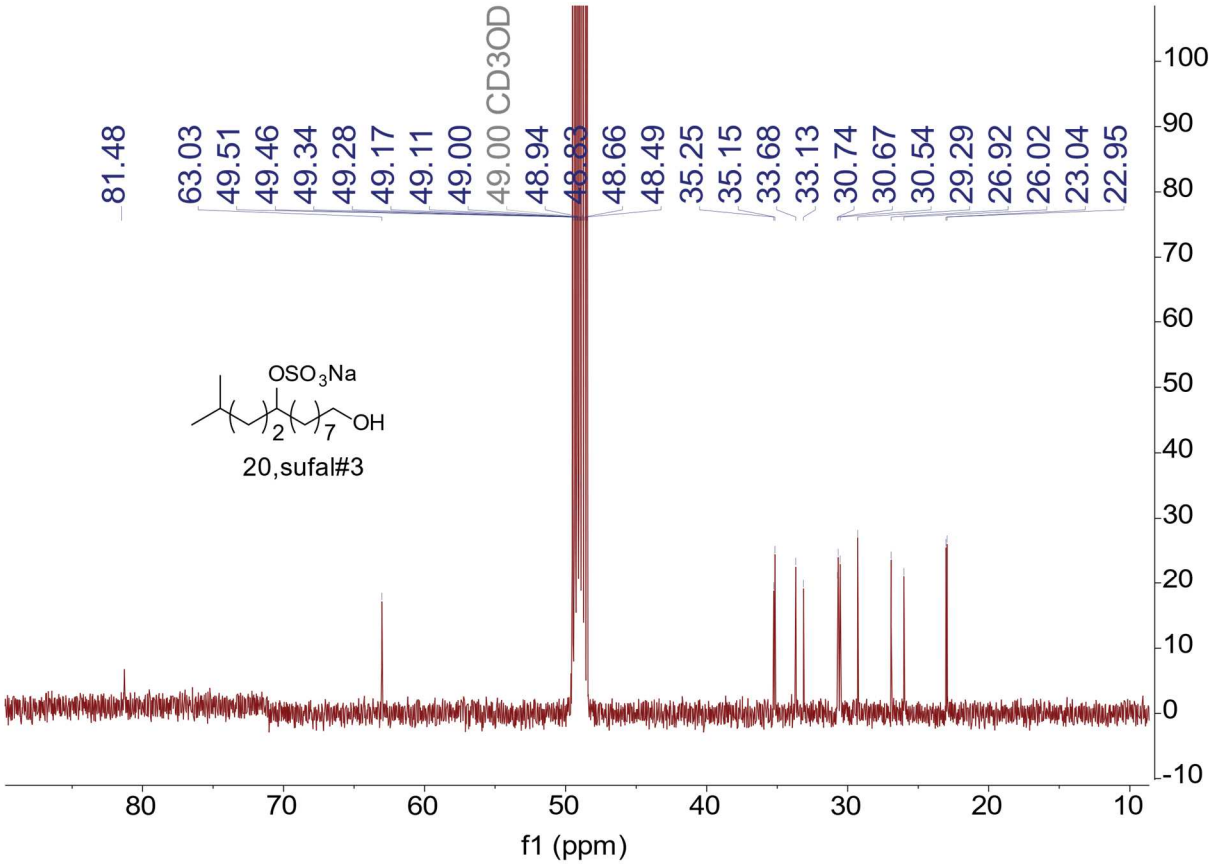
^{1}H NMR Spectrum (600 MHz, methanol-d_4) of 19



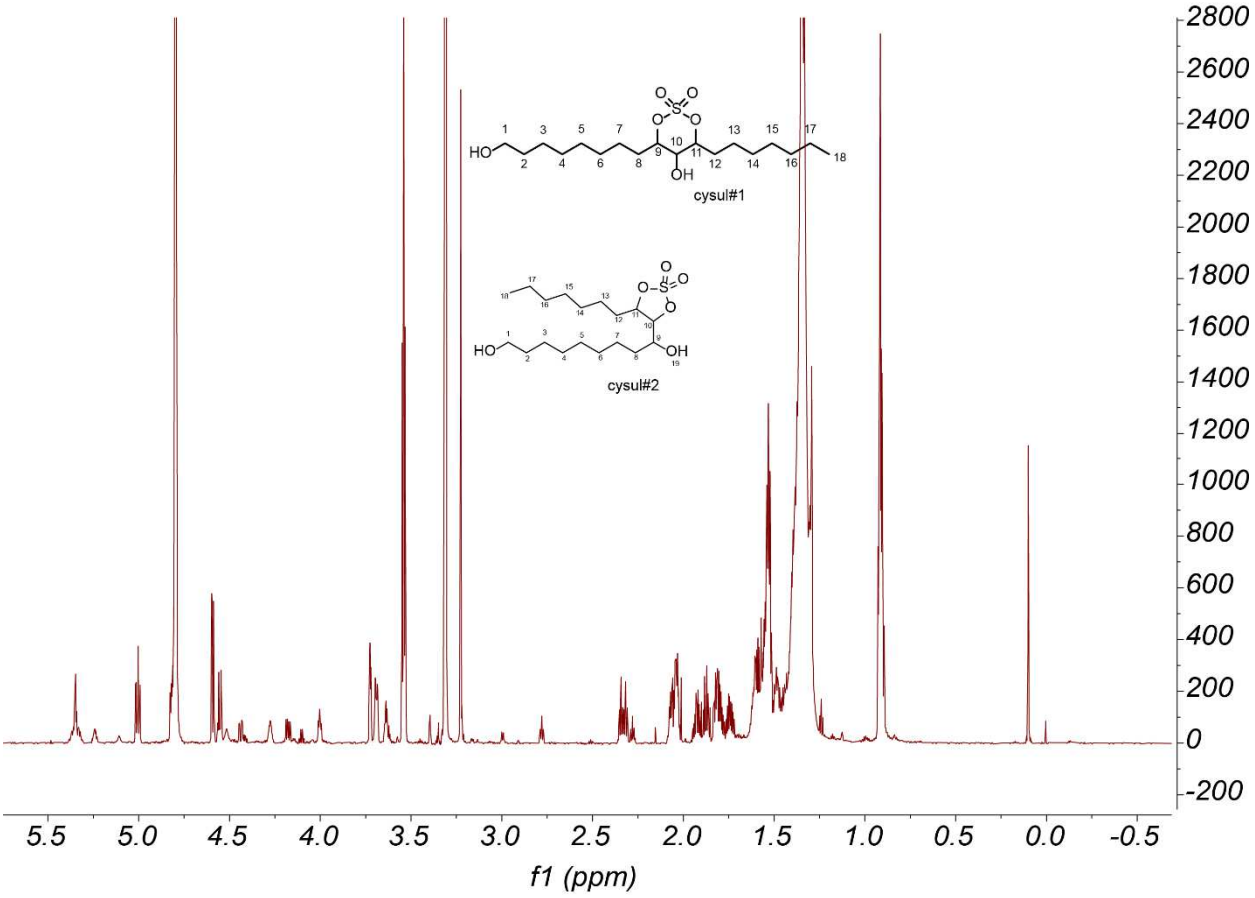
^{1}H NMR Spectrum (600 MHz, methanol-d_4) of 20



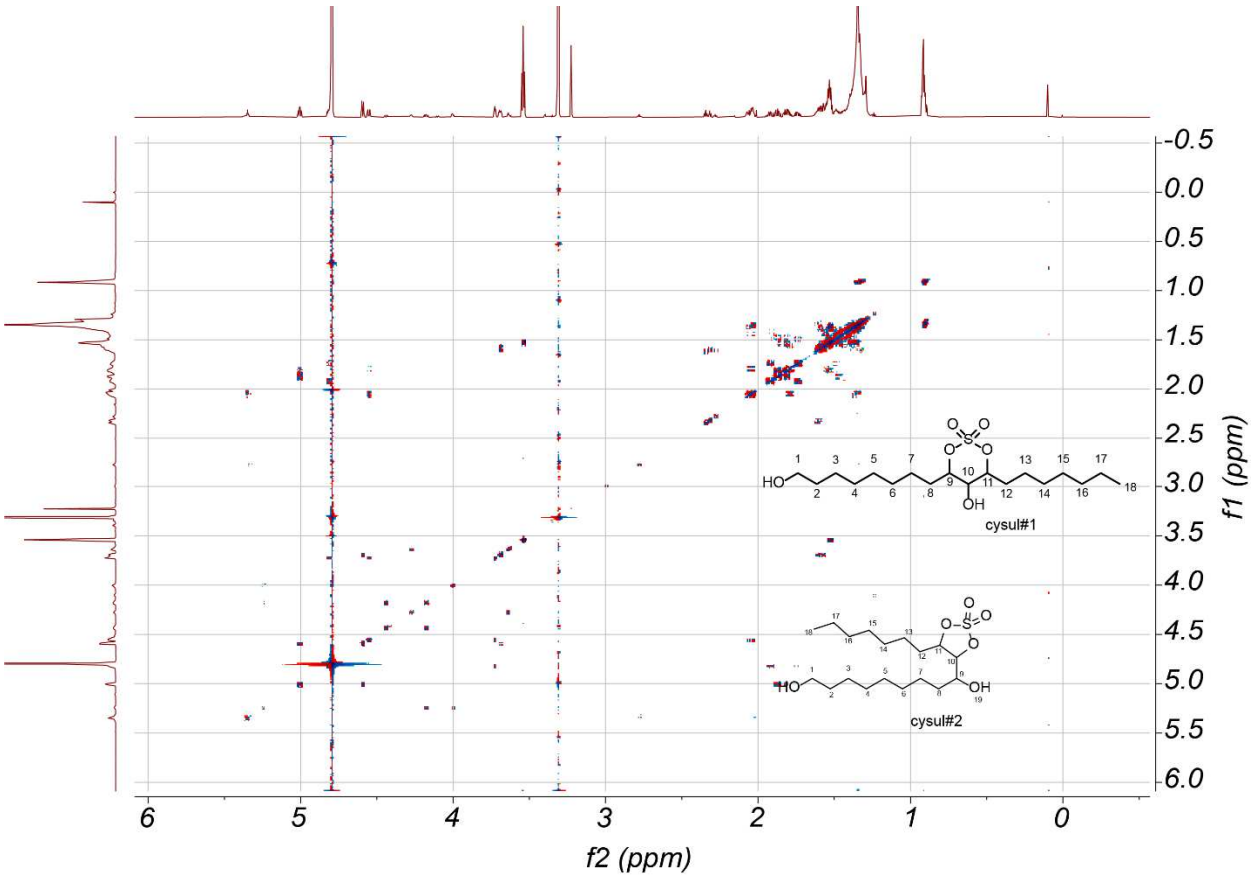
^{13}C NMR Spectrum (600 MHz, methanol-d_4) of 20



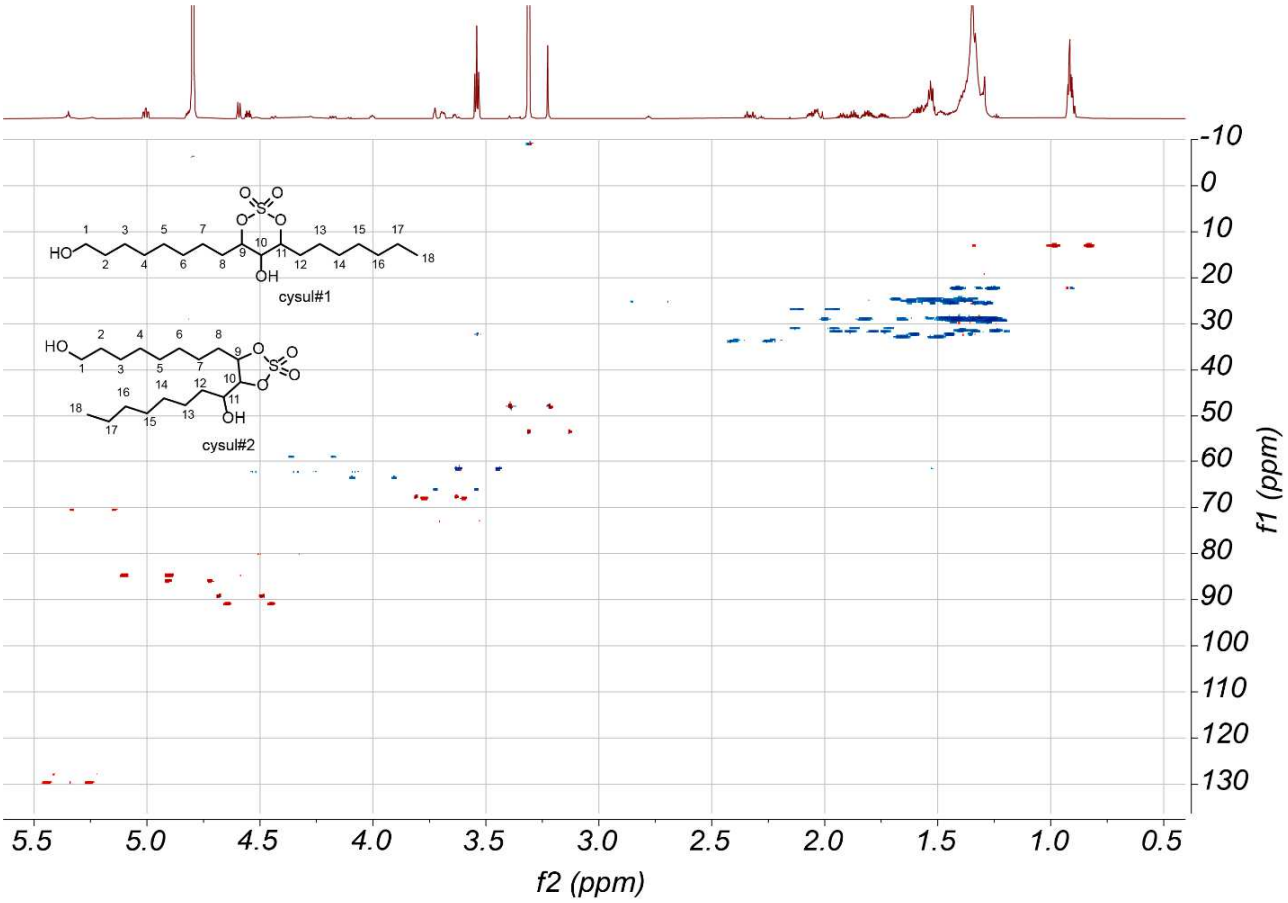
^{1}H NMR spectra of metabolome fraction containing 21 / 22, cysul#1 and cysul#2 (800 MHz, methanol-d_4)



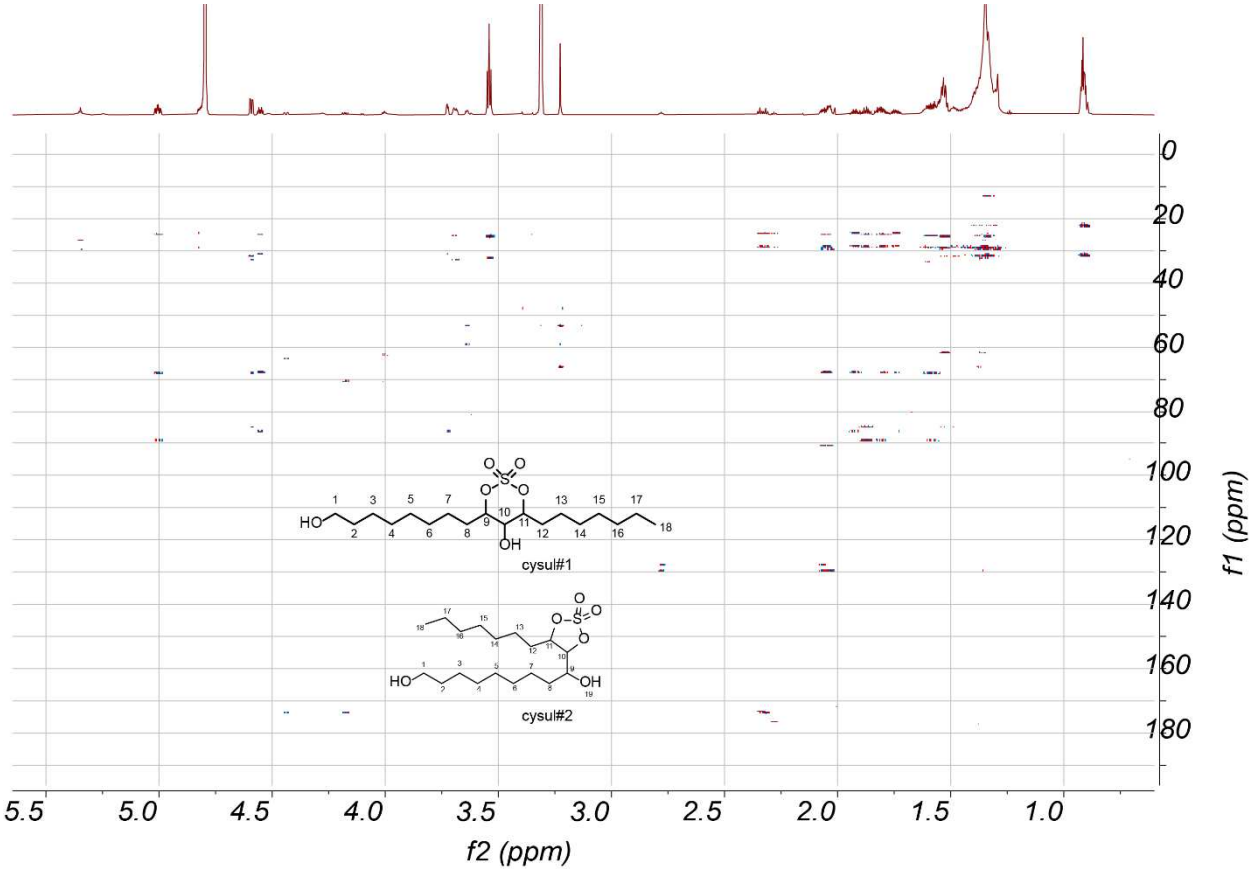
dqfCOSY spectrum of metabolome fraction containing cysul#1, 2 (800 MHz, methanol-d_4)



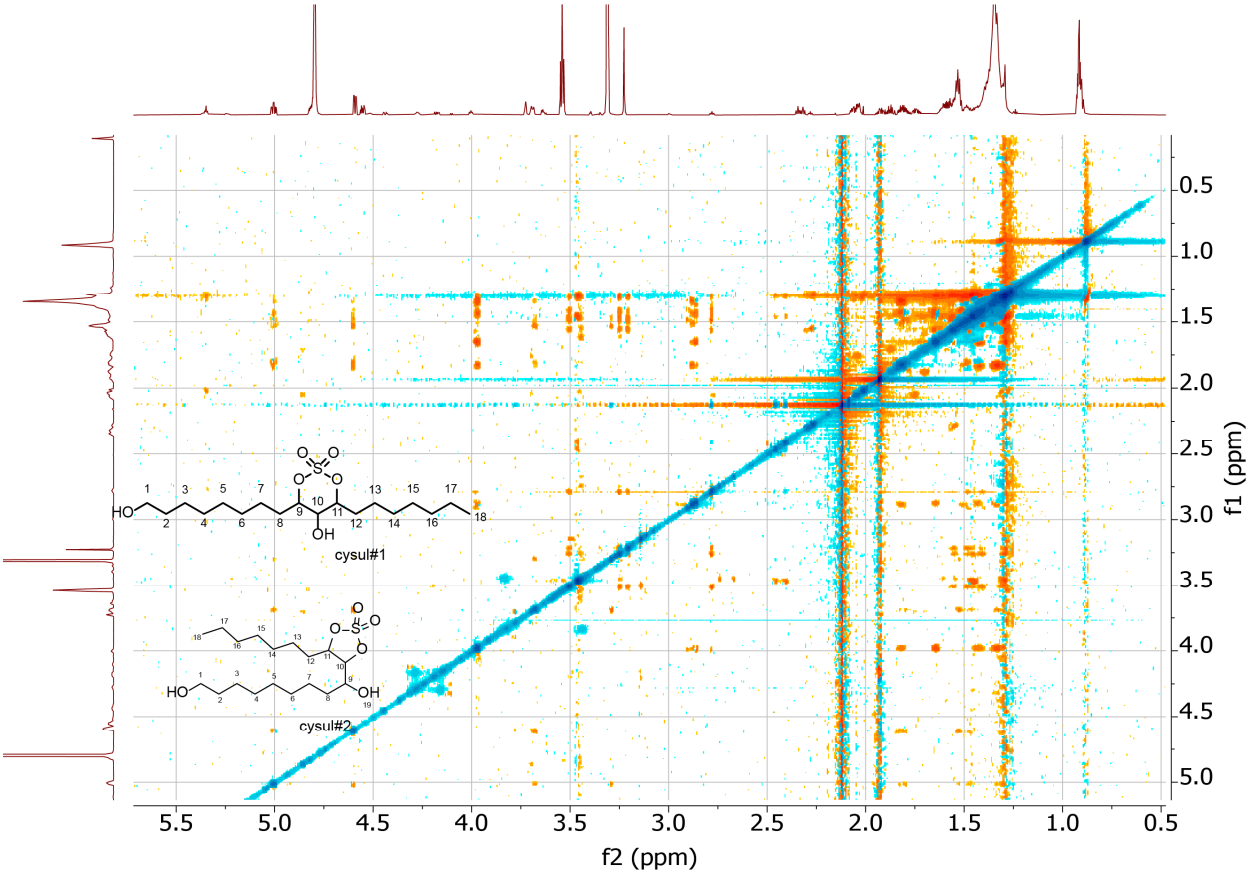
HSQC spectrum of metabolome fraction containing cysul#1, 2 (800 MHz, methanol-d_4)



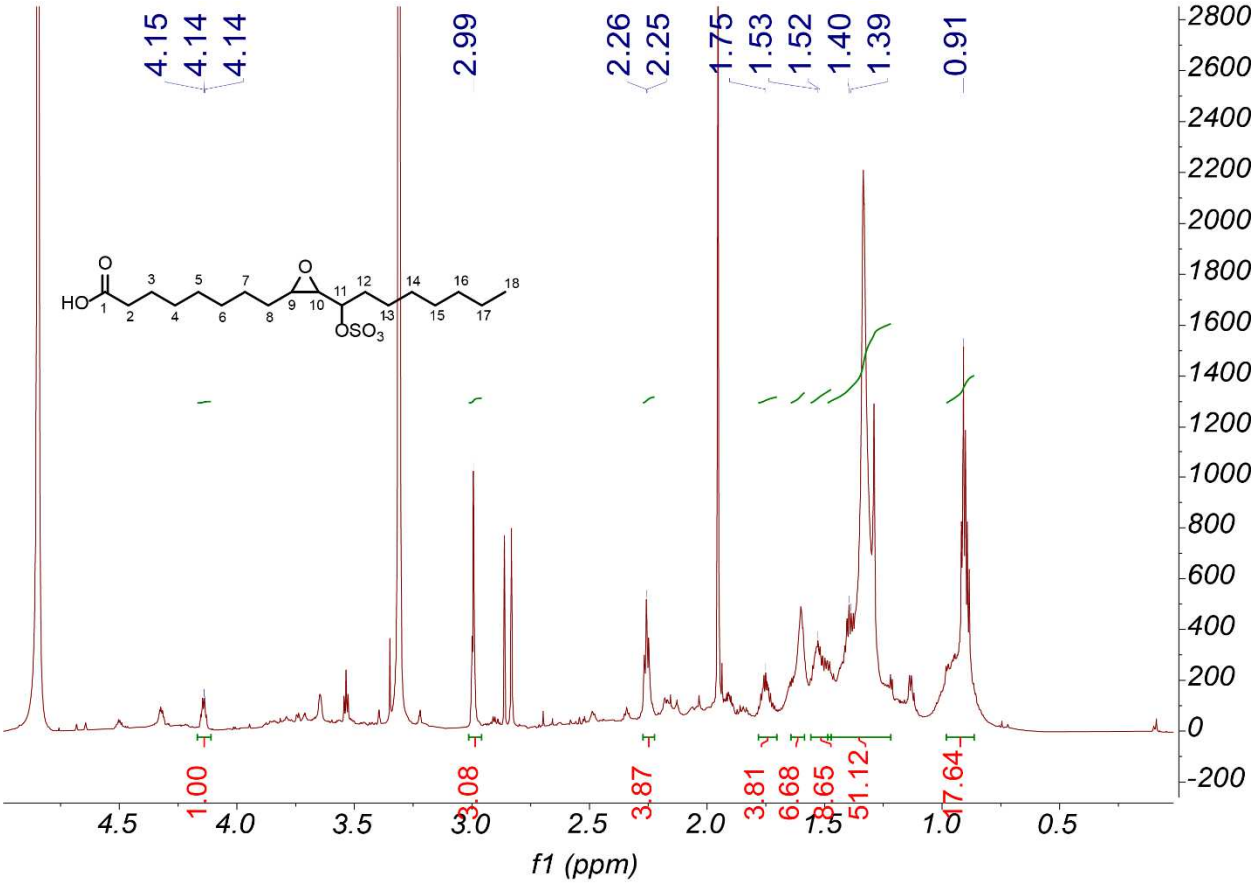
HMBC spectrum of metabolome fraction containing cysul#1, 2 (800 MHz, methanol-d_4)



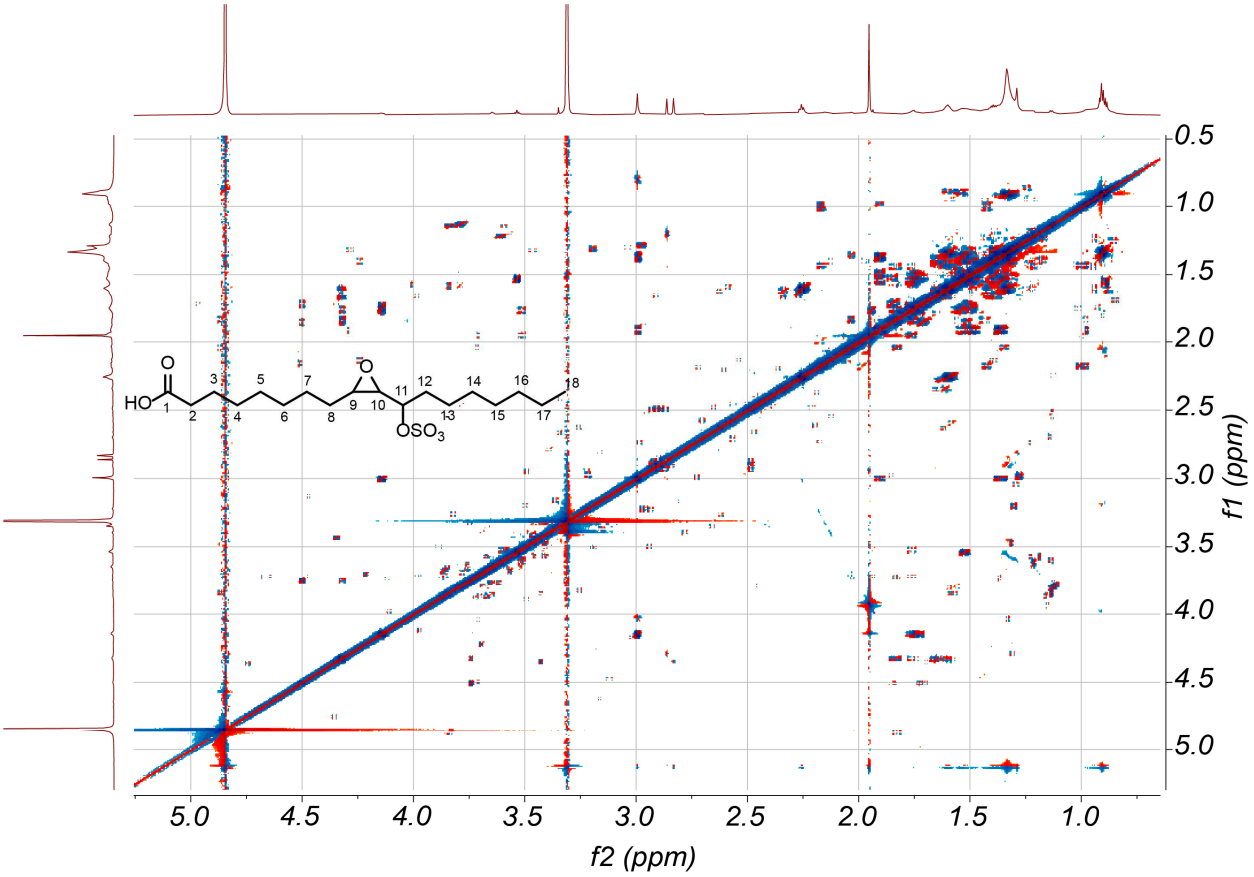
ROESY spectrum of metabolome fraction containing cysul#1, 2 (800 MHz, methanol-d_4)



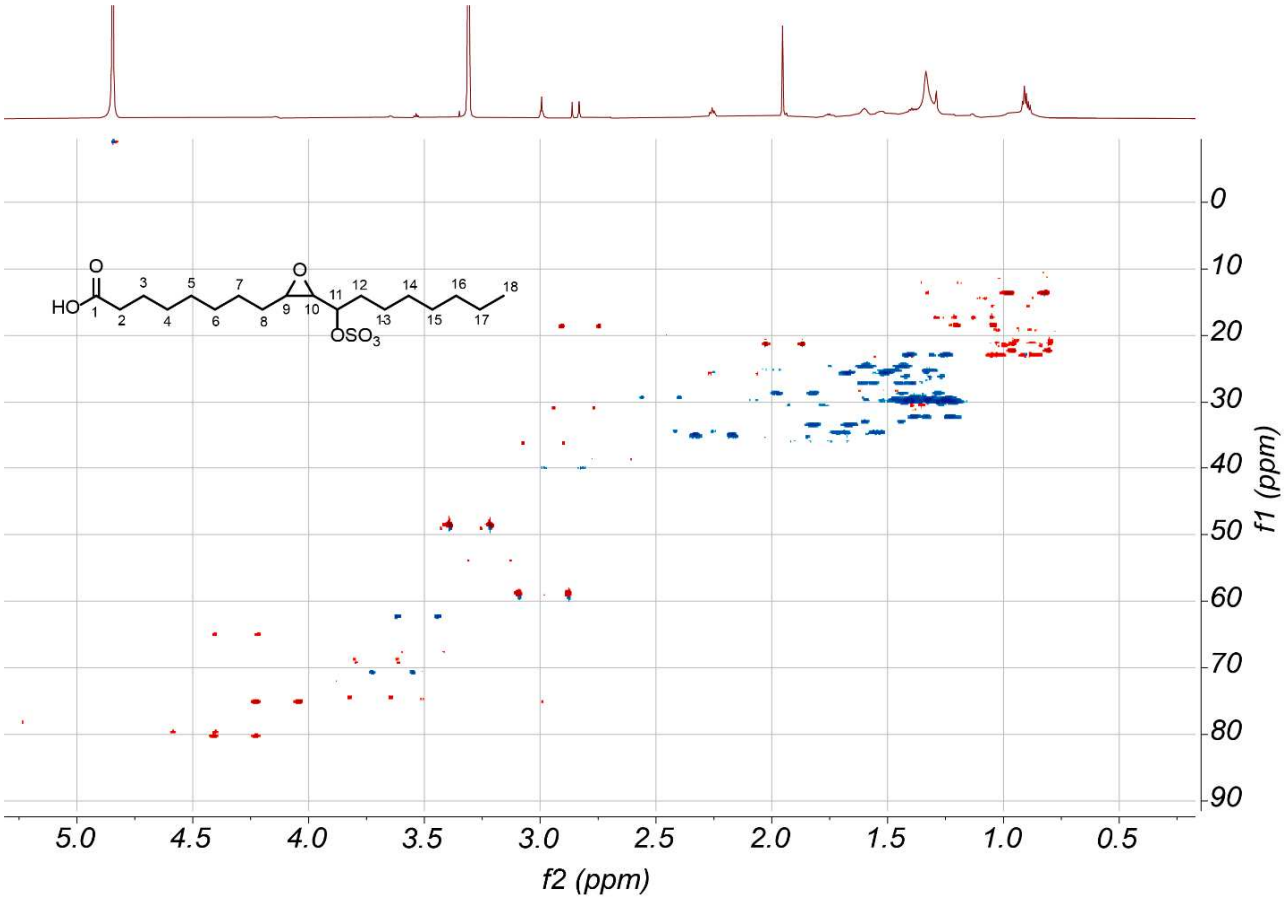
^{1}H NMR spectra of metabolome fraction containing epos#1, 30. (800 MHz, methanol-



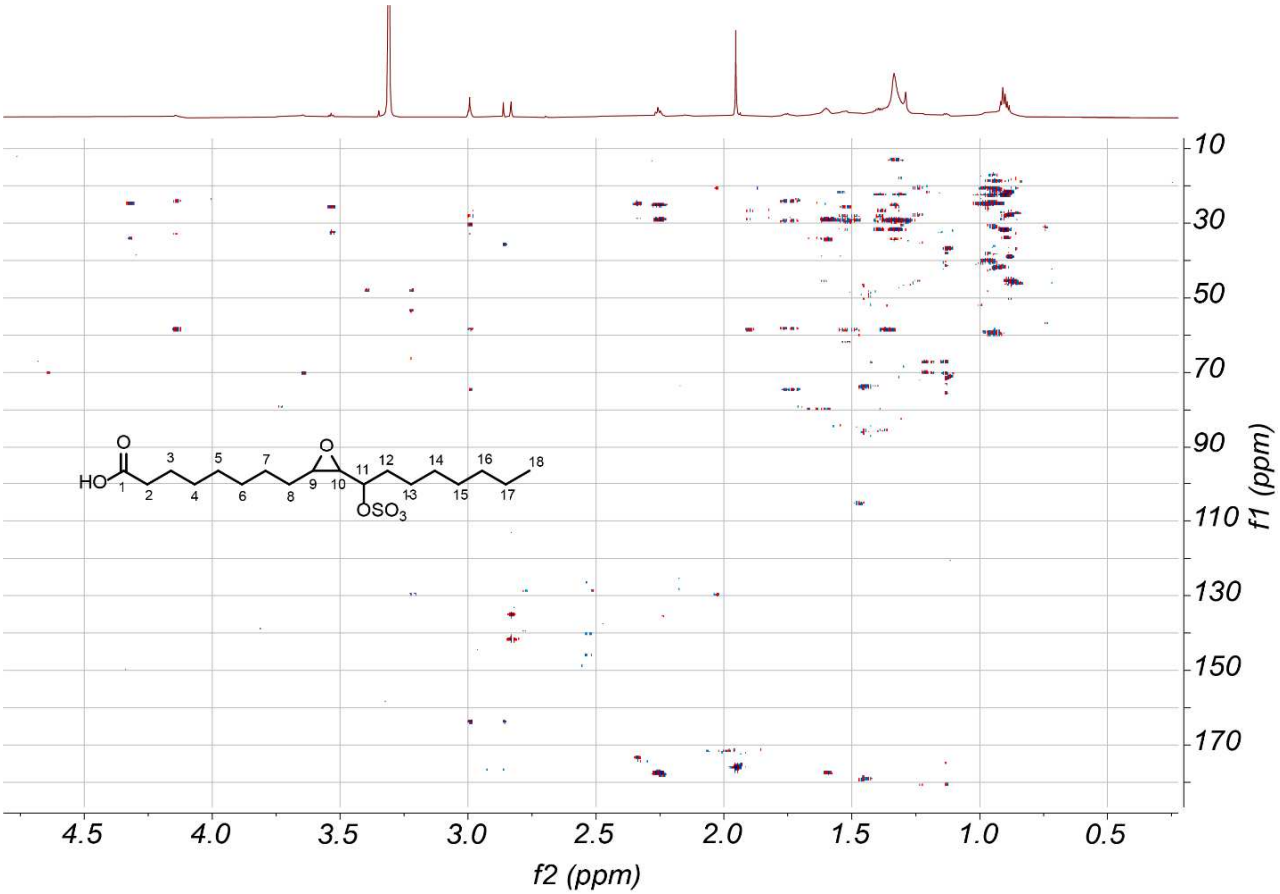
dqfCOSY spectrum of metabolome fraction containing epos#1 (800 MHz, methanol-d_4)



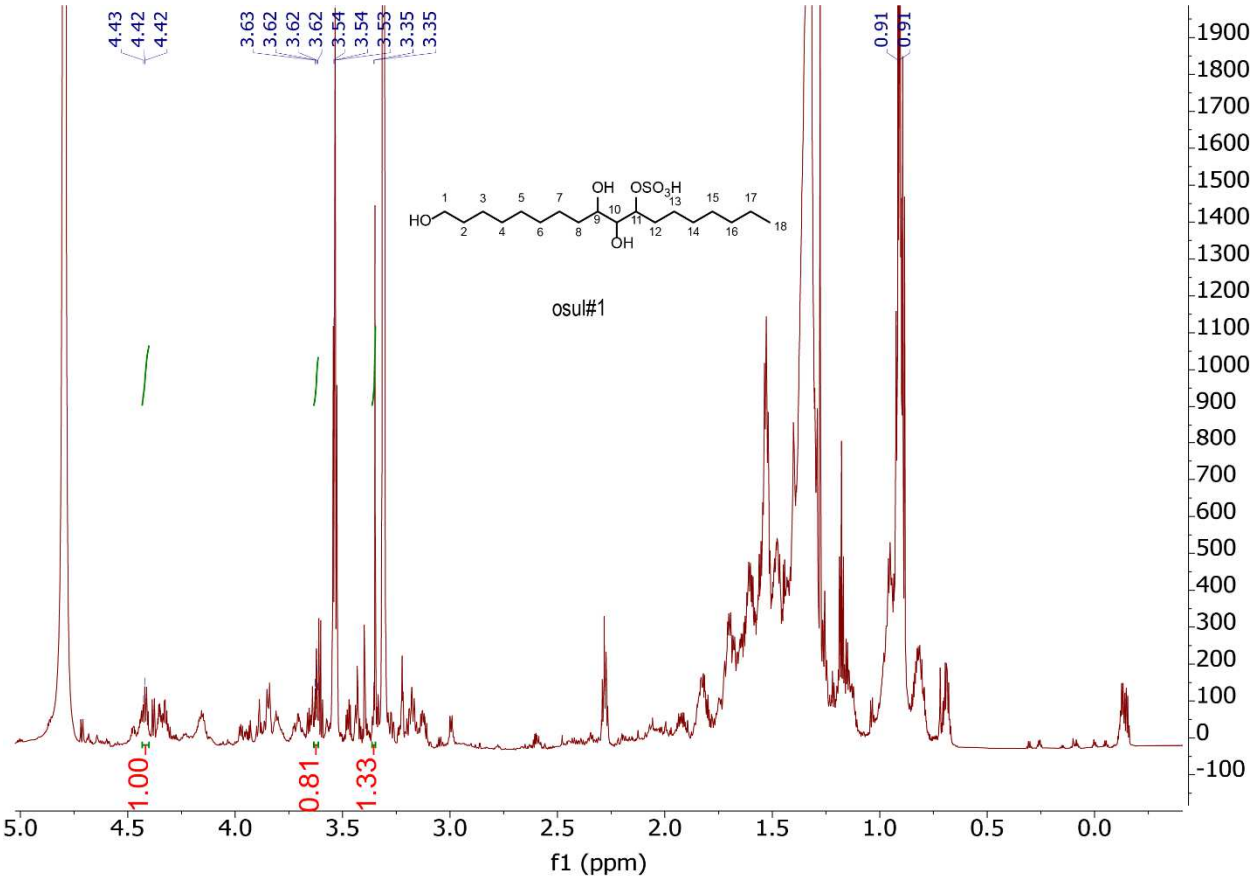
HSQC spectrum of metabolome fraction containing epos#1 (800 MHz, methanol-d_4)



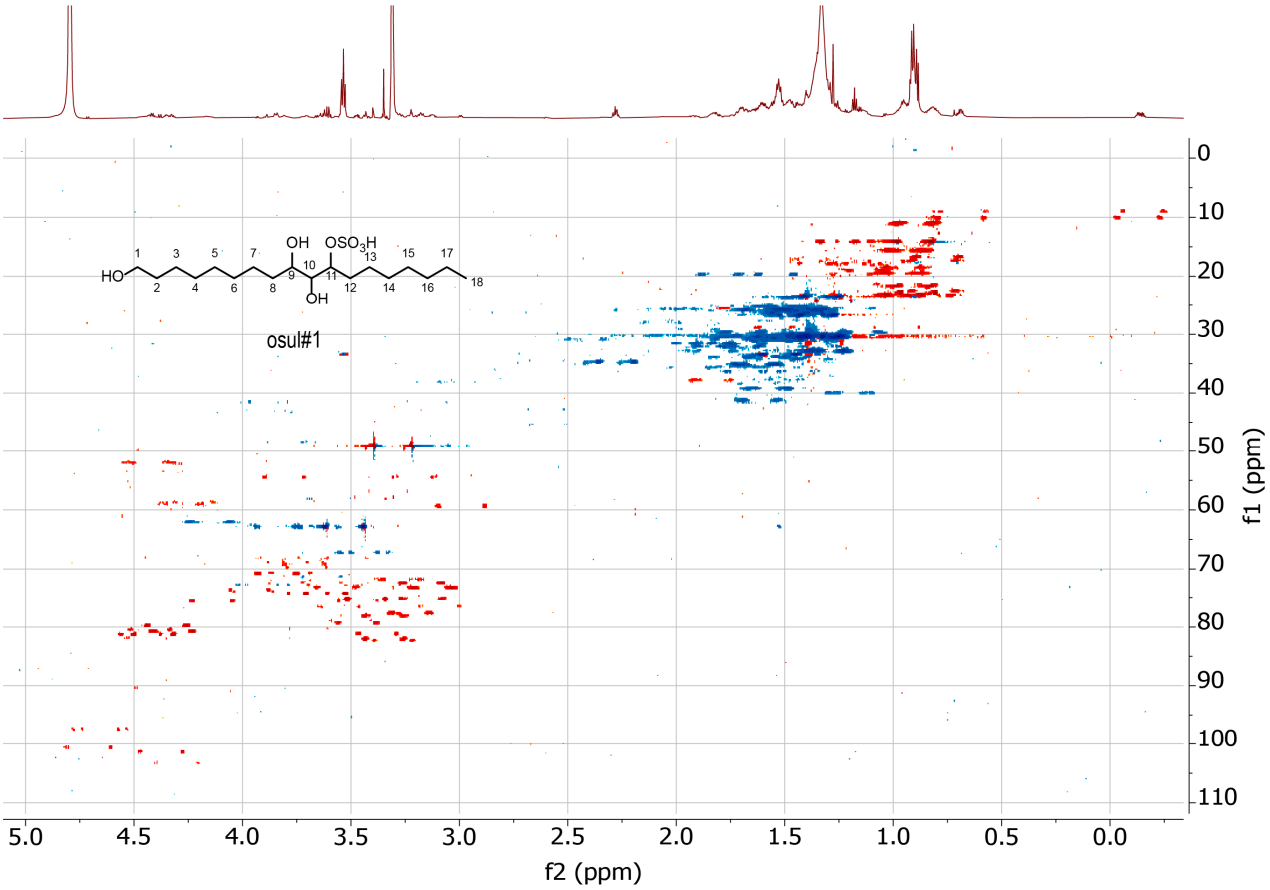
HMBC spectrum of metabolome fraction containing epos#1 (800 MHz, methanol-d_4)



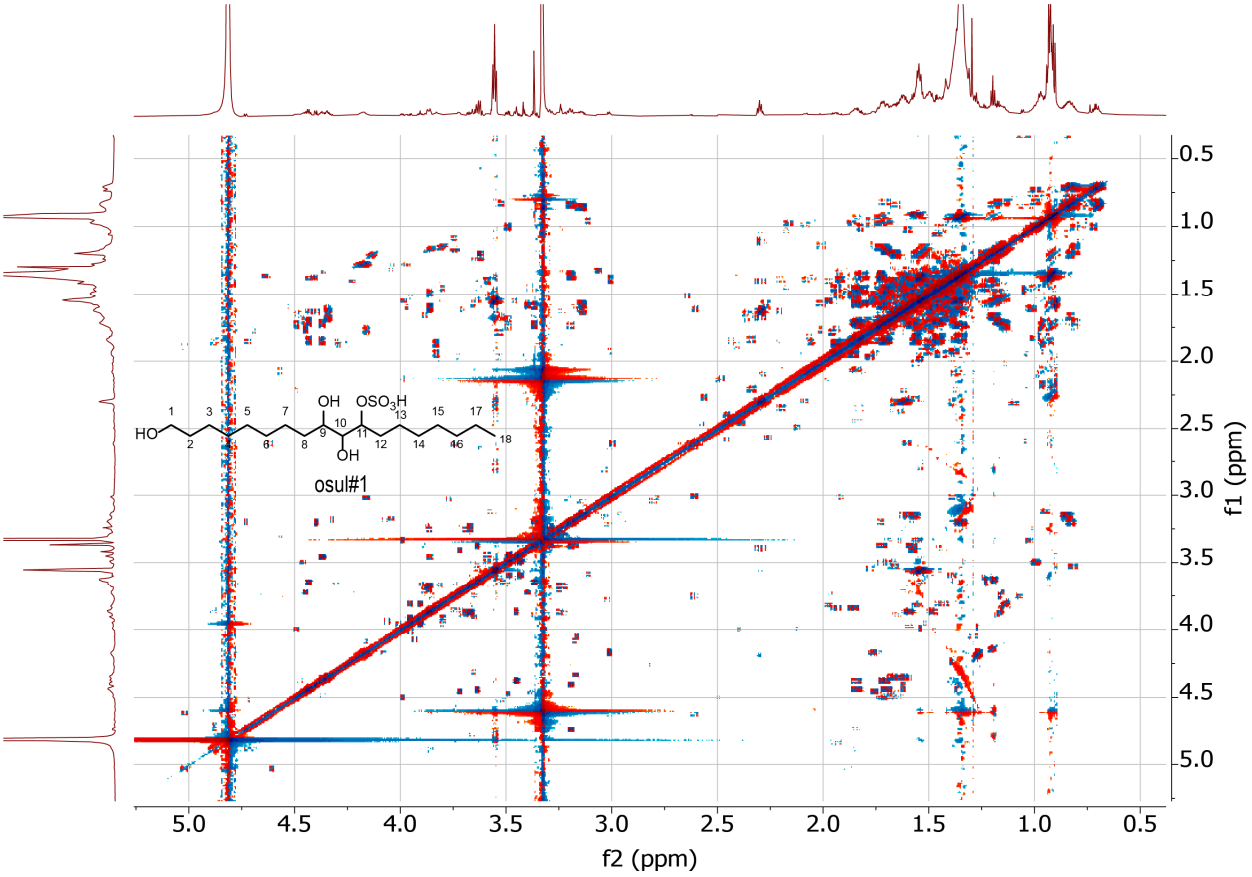
1D spectrum of metabolome fraction containing osul#1 (800 MHz, methanol-d_4)



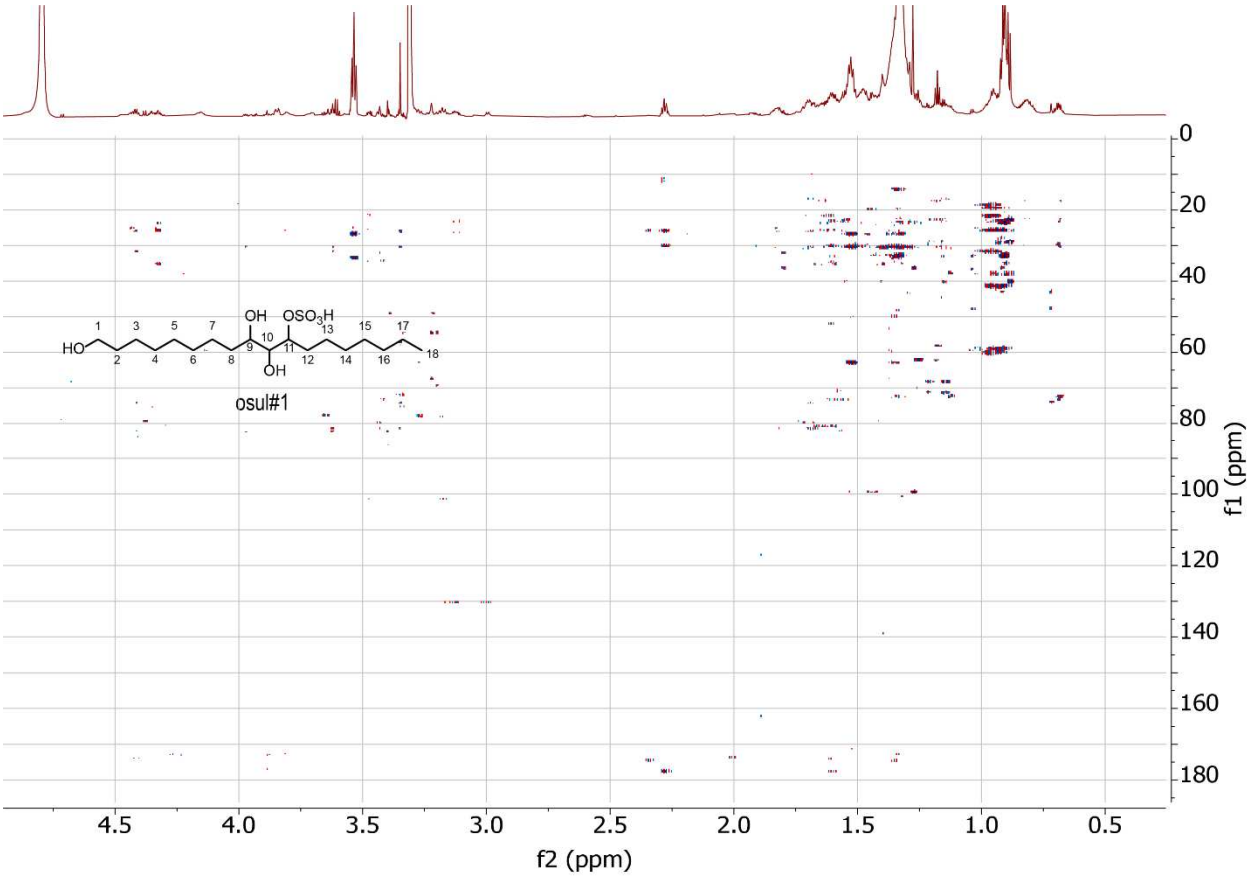
HSQC spectrum of metabolome fraction containing osul#1 (800 MHz, methanol-d_4)



dqfCOSY spectrum of metabolome fraction containing osul#1 (800 MHz, methanol-d_4)



HMBC spectrum of metabolome fraction containing osul#1 (800 MHz, methanol-d_4)



2.9 AUTHOR'S NOTE

Initial discovery of the cyclic sulfates was through the attention and samples from Dr. Neelanjan Bose and Prof. Frank C. Schroeder. The author had completed the characterization through subsequent analysis. PDDA assays were conducted by Dr. Hanno Andreas Ludewig, microbial inhibition assays were conducted by Spencer Wong from Prof. Dennis Kim's lab at MIT, sulfotransferase mutants eud-1 and seud-1 were provided by Prof. Erik J. Ragsdale. Mouth form assays were conducted by Joseph Biddle and Prof. Erik J. Ragsdale. The natural variants of *P. pacificus* were collected and grown by Prof. Roddelsburger's lab and the strains were run on the HPLC-MS by Dr. Joshua Baccile. Sulfatase strains were grown and maintained by Jazmin Aguilar-Romero.

REFERENCES

35. Mulder G.J. and Jakoby W.B.; Sulphation. In Mulder G.J. (ed.) Conjugation Reactions in Drug Metabolism. Taylor & Francis, London, pp. 107-161 (1990).

36. Deepa S.S.; Yamada S.; Zako M.; Goldberger O.; Sugahara K.; Chondroitin sulfate chains on syndecan-1 and syndecan-4 from normal murine mammary gland epithelial cells are structurally and functionally distinct and cooperate with heparan sulfate chains to bind growth factors. A novel function to control binding of midkine, pleiotrophin, and basic fibroblast growth factor. *J. Biol. Chem.,* 279, pp. 37368-37376 (2004).

37. Mueller J.W.; Gilligan L.C.; Idkowiak J.; Arlt W.; Foster P.A.; The Regulation of Steroid Action by Sulfation and Desulfation. *Endocr. Rev*. Oct;36(5):526-63 (2015).

38. Schroeder F.C.; Taggi A.E.; Gronquist M.; Malik R.U.; Grant J.B.; Eisner T.; Meinwald J.; NMR-spectroscopic screening of spider venom reveals sulfated nucleosides as major components for the brown recluse and related species. *PNAS*. 105(38), 14283-7 (2008).

39. Zheng L.; Kariya M.J.; Chute C.D.; Leinwand S.G.; Tong A.; Pribadi A.K.; Curran K.P.; Bose N.; Schroeder F.C.; Srinivasan J.; Chalasani S.H.; Predator-secreted sulfolipids induce defensive responses in *C. elegans*. *Nat. Comm.* 9 article 1128 (2018).

40. Gao Y;, Sharpless B.; Vicinal diol cyclic sulfates. Like epoxides only more reactive *J. Am. Chem. Soc.* (1988)

41. Megia-Fernandez A.; Morales-Sanfrutos J.; Hernandez-Mateo F.; Santoyo-Gonzalez F.; Synthetic Applications of Cyclic Sulfites, Sulfates and Sulfamidates in Carbohydrate Chemistry. *Curr. Org. Chem.*, 15, 401−432 (2011).

42. Zhang Y.K; Sanchez-Ayala M.A.; Sternberg P.W.; Srinivasan J.; Schroeder F.C.; Improved Synthesis for Modular Ascarosides Uncovers Biological Activity. *Org. Lett.* 19112837-2840 (2017).

43. Bui L.T.; Ivers N.A.; Ragsdale E.J. A sulfotransferase dosage-dependently regulates mouthpart polyphenism in the nematode Pristionchus pacificus. *Nat. Comm.* (2018)

44. Sieriebriennikov B.; Prabh N.; Dardiry M.; Witte H.; Röseler W.; Kieninger M.R.; Rödelsperger C.; Sommer R.J.; A Developmental Switch Generating Phenotypic Plasticity Is Part of a Conserved Multi-gene Locus. Cell Reports, 23(1) pp. 2835-2843 (2018).

45. Gregus Z.; Kim H. J.; Madhu C.; Liu Y.; Rozman P.; Klassen C. D.; Sulfation of acetaminophen and acetaminophen-induced alterations in sulfate and 3'-phosphoadenosine 5'-phosphosulfate homeostasis in rats with deficient dietary intake of sulfur. Drug Metabolism and Disposition, Vol. 22, No. 5 (1994).

46. Artola M.; Wu, L.; Ferraz, M. J.; Kuo C. L.; Raich L.; Breen I. Z.; Offen W.A.; Codee J.D.C.; van der Marel G.A.; Rovira C.; Aerts J.M.F.G; Davies G.J.; Overkleeft H. S.; 1,6-Cyclophellitol Cyclosulfates: A New Class of Irreversible Glycosidase Inhibitor. *ACS central science*, *3*(7), 784‑793 (2017).

47. Ludewig A.H.; Gimond C.; Judkins J.C.; Thornton S.; Pulido D.C.; Micikas R.J.; Doring F.; Antebi A.; Braendle C.; Schroeder F.C.; Larval crowding accelerates C. elegans development and reduces lifespan. *PLOS*, (2017).
48. Woldrigh C.L.; Iterson W.V.; Effects of Treatment with Sodium Dodecyl Sulfate on the Ultrastructure of *Escherichia coli. Journal of Bacteriology*, vol. 111 (3), p801-813 (1972).
49. Simoes M.; Simoes L. C.; Pereira M. O.; Vieira M. J.; Sodium dodecyl sulfate allows the persistence and recovery of biofilms of Pseudomonas fluorescens formed under different hydrodynamic conditions. *Biofouling*, 24(1):35-44 (2008).
50. Morisseau C.; Inceoglu B.; Schmelzer K.; Tsai H.J.; Jinks S.L.; Hegedus C.M.; Hammock B.D.; Naturally occurring monoepoxides of eicosapentaenoic acid and docosahexaenoic acid are bioactive antihyperalgesic lipids. *J. Lipid Res.* 51 (2010).
51. Ulu A.; Harris T.R.; Morisseau C.; Miyabe C.; Inoue H.; Schuster G.; Dong H.; Iosif A.M.; Liu J.Y.; Weiss R.H.; Chiamvimonvat N.; Imig J.D.; Hammock B.D.; J. Anti-inflammatory effects of ω-3 polyunsaturated fatty acids and soluble epoxide hydrolase inhibitors in angiotensin-II-dependent hypertension. *Cardiovasc. Pharmacol.* 62 (2013).
52. Klaassen C. D.; Boles J. W.; Sulfation and sulfotransferases 5: the importance of 3'-phosphoadenosine 5'-phosphosulfate (PAPS) in the regulation of sulfation. *The FASEB Journal*, 11 6, 404 418, 0892-6638 (1997).
53. Alegado R.A.; Brown L.W.; Cao S.; Dermenjian R.K.; Zuzow R.; Fairclough S.R.; Clardy J.; King N.; A bacterial sulfonolipid triggers multicellular development in the closest living relatives of animals. *Elife*. (2012).
54. Bose N.; Ogawa A.; von Reuss S.H.; Yim J.J.; Ragsdale E.J.; Sommer R.J.; Schroeder F.C.; Complex small-molecule architectures regulate phenotypic plasticity in a nematode. *Angewandte Comm.* (2012)
55. McGrath P.T.; Xu Y.; Ailion M.; Garrison J.L.; Butcher R.A. & Bargmann C.I.; Parallel evolution of domesticated Caenorhabditis species targets pheromone receptor genes. *Nature* (2011).

CHAPTER THREE
Antibiotics from *P. pacificus*

Authors: Kariya M.J.; Wong S.; Sadeeshkumar H.; Aguilar-Romero J.; Bose N.; Yim J.J.; Kim D.H. & Schroeder F.C.

3.1 ABSTRACT

Antibiotic resistant microbes are on the rise, and there is a dire need to identify new classes of antibacterial and antifungal scaffolds to add to our arsenal. To find new compounds with antimicrobial activity, we look to animals that thrive in microbe-rich environments, such as *P. pacificus*. We used a growth inhibition assay on fractions from the *P. pacificus* metabolome and focused on a group of fractions that inhibited growth of *C. albicans*. We elucidated a cyclopropyl containing lipid, cpfa#1, that was responsible for the activity. A 2:1 mixture of cpfa#1 and its isomer cpfa#1.1 against *C. albicans* has an MIC of 0.515 mM, and this concentration is toxic for both HEK cells and *C. elegans* and *P. pacificus*, yet *P. pacificus* shows higher resistance. We further show that the lipid is produced *de novo* by *P. pacificus*, using a cyclopropyl fatty acid from the bacteria that it feeds on. Additionally, we identify another example of a cyclopropyl lipid from *P. pacificus* of shorter chain length that retains a different activity profile.

3.2 INTRODUCTION

With the rise in antibiotic-resistant bacteria and fungi, we are desperate to find alternative sources of antibiotics. For the past 70 years we have seen a paucity in the discovery of new classes of antibiotics coupled with a stark increase in antibiotic-resistant microbes[62]. To find new classes of antibiotics, it may be necessary to look in the metabolomes of animals that thrive in microbe-rich environments because these organisms may produce chemical defense responses. Antibiotics have been found from some marine animals, like sponges[56], giving some hope to this pathway. We were interested in assaying the metabolome of animals that both thrive in environments rich in microbial life and are easy to cultivate in a laboratory setting, such as nematodes. We prioritized the nematode *P. pacificus* because unlike *C. elegans*, which uses a grinder to digest bacteria, *P. pacificus* leaves microbes intact as they pass through the worm. Thus, we believe that *P. pacificus* must somehow control the proliferation of bacteria and fungi in order for this to occur, perhaps through the generation of antibiotics.

We collected the metabolome of *P. pacificus* and performed antibiotic assays on metabolomic fractions. We found a wide activity response from these fractions to several microbes and followed through with one of the many hits found. We characterized the active component of one of the active fractions to be a hydroxylated cyclopropyl fatty alcohol. Through total synthesis we confirmed the activity to be attributed to the anti-hydroxy-cis-cyclopropyl fatty acid, cpfa#1.

3.3 RESULTS

We set about testing individual fractions using the fractionation protocol in previous chapters. We tested each fraction against a set of various microbes (see Supplementary File 1 for full data) and found a wide range of activity from various fractions. We focused on fractions D11 to D15, of which D11, D13, and D14 had consistent inhibitory effects against *Candida albicans* across the four technical replicates (Figure 6a). Fraction D13 also showed a degree of selectivity in efficacy against *C. albicans* over the other microbes tested (Figure 6b).

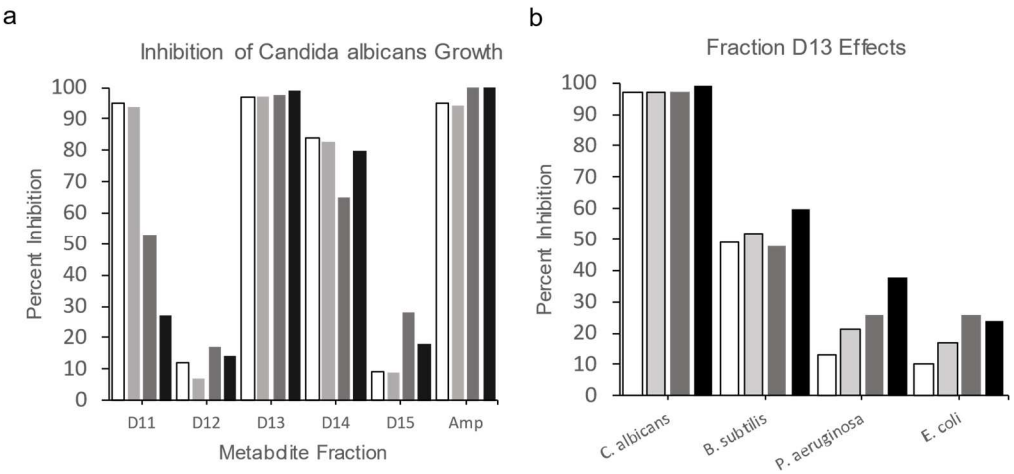


Figure 6 Antibiotic effects of *P. pacificus* metabolome fractions a. effects on *C. albicans* growth **b.** effects of fraction D13 on various microbial species.

We compared NMR and HPLC-MS spectra of fractions D11-D15 to pick out unique features in D13 (Supplementary Figure S13, S14). From these unique features, we used 2D NMR spectroscopy and high resolution HPLC-MS to determine the structure of the

active compound in fraction D13 (Supplementary Figure S14), an alpha hydroxylated cyclopropyl containing fatty alcohol (cpfa#1, Figure 7).

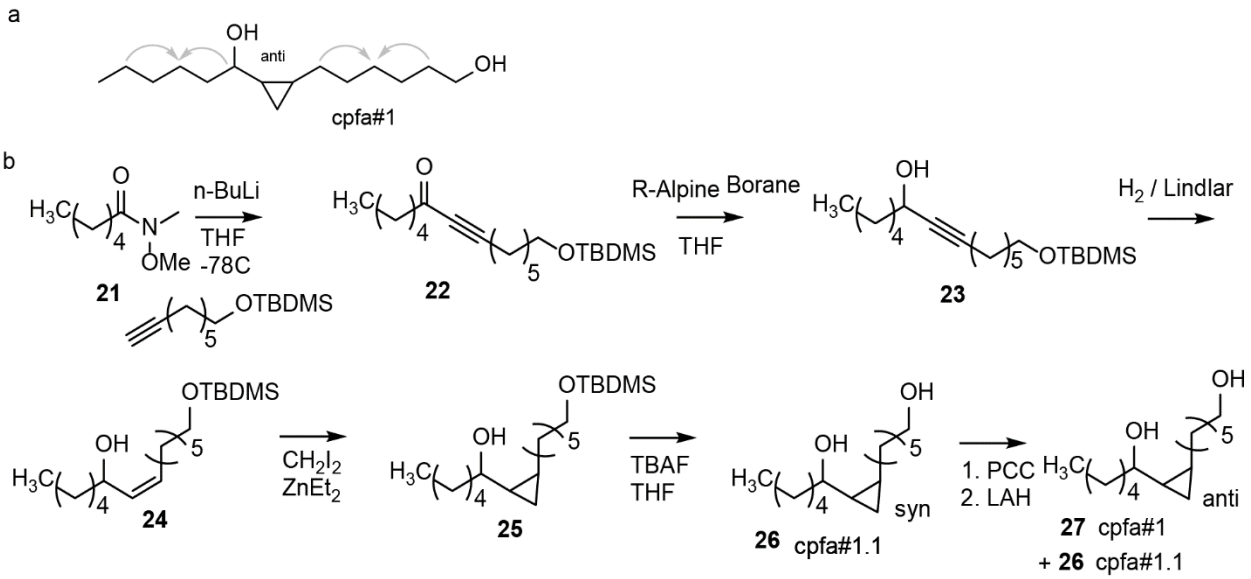


Figure 7 a. Structure of cpfa#1 with key HMBC correlations. **b**. total synthesis of cpfa#1 and its isomer cpfa#1.1.

Comparing the chemical shift values and coupling constants of the cyclopropyl functional group and neighboring protons to known compounds[58], we deduced that the cyclopropyl group was cis. We confirmed the relative stereochemistry of the alcohol to the cyclopropyl ring to be anti via total synthesis. We initiated the synthesis via coupling of Weinreb amide **21** with a protected octynol to afford the ynone **22**. We reduced the ketone through a midland reduction to ynol **23**, which was subsequently reduced to the enol **24**. Cyclopropanation through the Simmons-Smith method and deprotection afforded a the syn alcohol, **26,** cpfa#1.1, which showed markedly different chemical shift values at the cyclopropyl ring and eluted at a different time than the natural product

(Supplementary Figure S14). We oxidized **26** using pyridinium chlorochromate and reduced the compound using lithium aluminum hydride which afforded the anti alcohol, **27**, cpfa#1, in a 2:1 ratio of anti:syn. The synthesized product **27** shares the same chemical shift values as that of the natural product and co-elutes with the natural product on the HPLC-MS (Supplementary Figure S15, Supplementary Table S10).

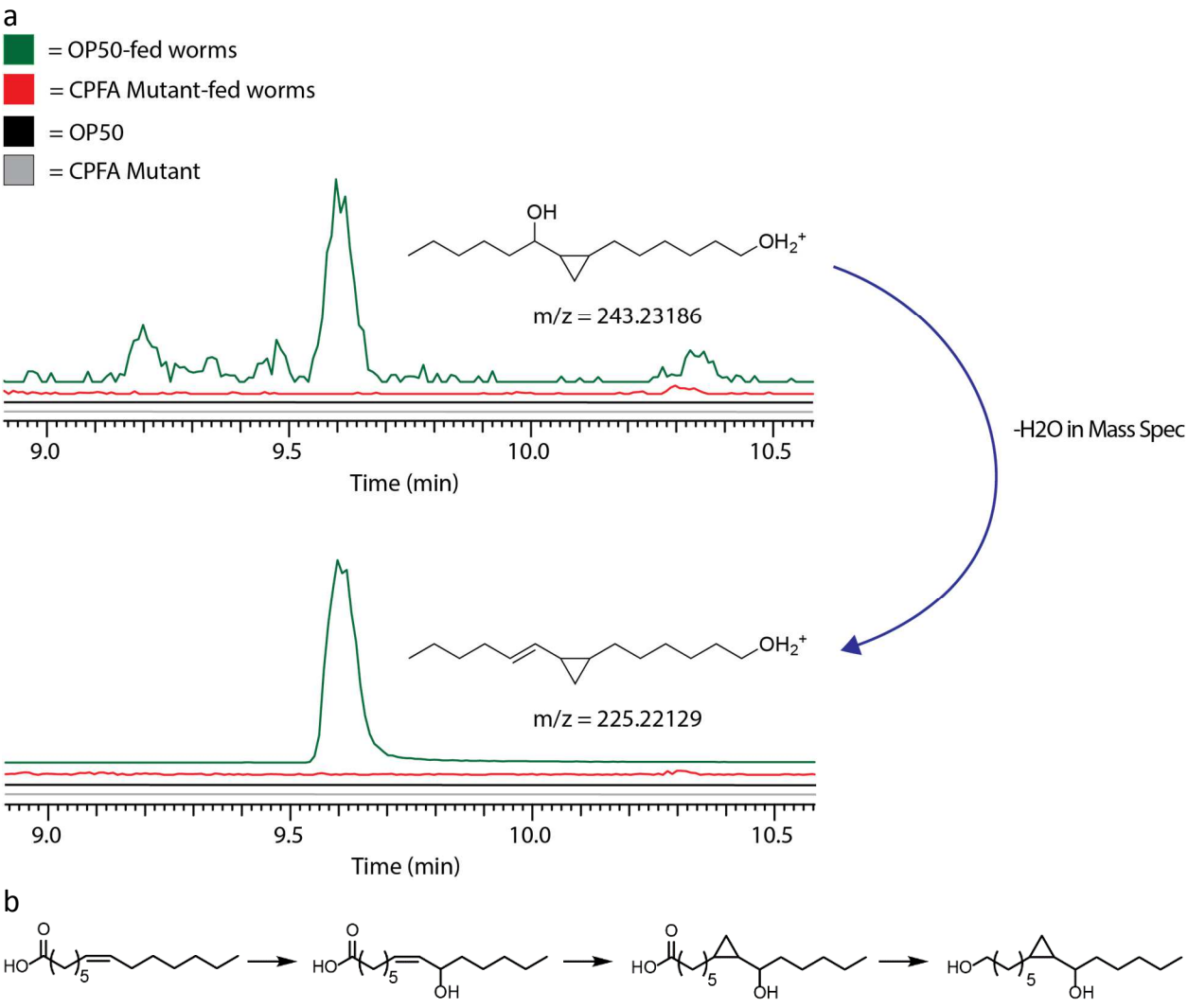


Figure 8 Biosynthetic origins of cpfa#1 a. EICs for cpfa#1 in OP50-fed worms, CPFA mutant *E. coli*-fed worms, OP50, and CPFA Mutant *E. coli*. b. proposed biosynthesis of cpfa#1.

Through HRMS we do not see an abundant amount of the parent ion for cpfa#1. Instead, we see the dehydrated ion (Figure 8a), indicating a high tendency to eliminate water from the molecular ion. We were curious as to the origin of cpfa#1 and if it was a compound that came directly from *P. pacificus* or of it was produced by the OP50 *E. coli* that *P. pacificus* feeds on in laboratory settings. Although we found that *E. coli* does not

produce cpfa#1 (Figure 8a), *P. pacificus* fed with cyclopropyl-lacking bacteria do not produce cpfa#1 (Figure 8a). This implies that cpfa#1 is produced synergistically between bacteria and *P. pacificus*. The cyclopropyl group, or more likely the entire fatty alcohol scaffold may be contributed by the bacteria, which is then hydroxylated by *P. pacificus*.

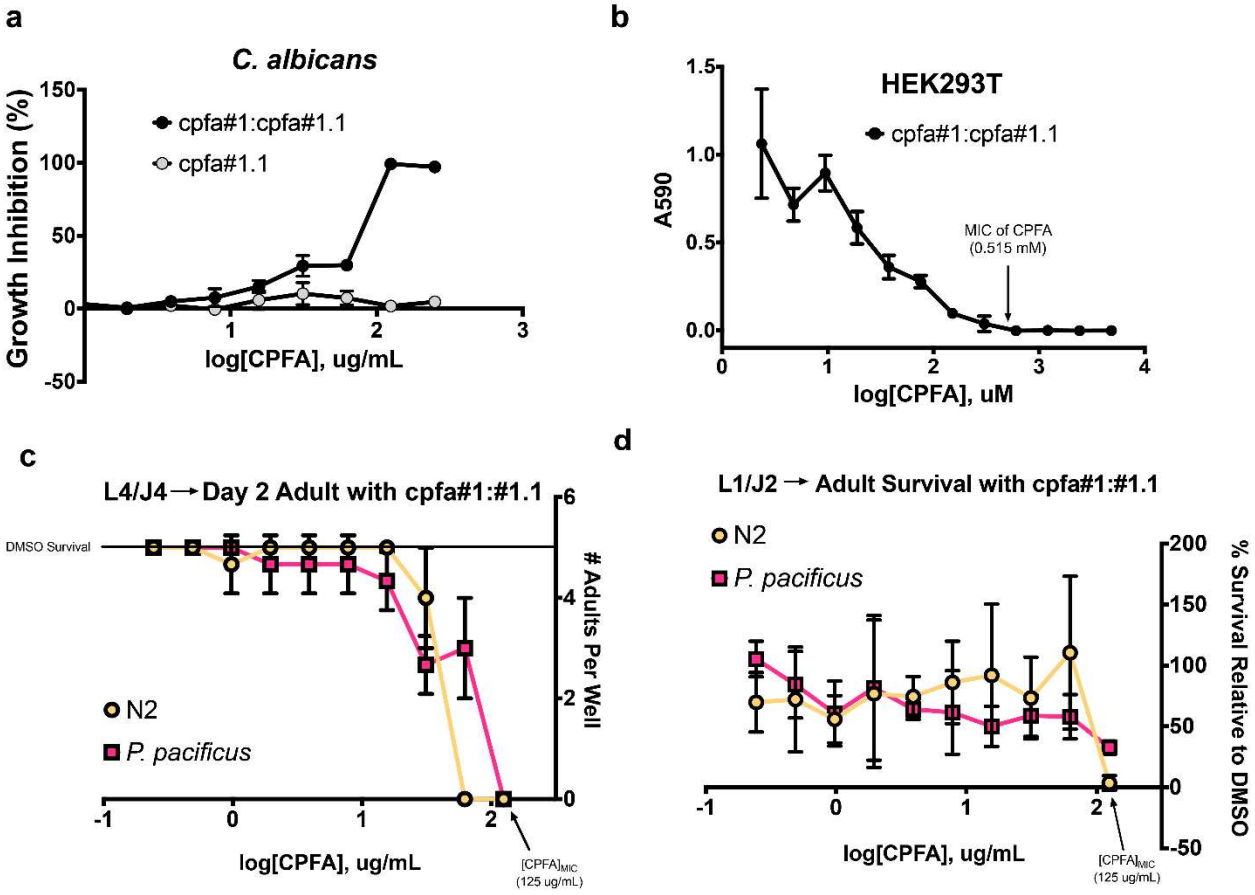


Figure 9 Bioactivity of cpfa#1. a. Growth inhibition against *C. albicans* of synthetic cpfa#1.1 compared to mixture of cpfa#1 and cpfa#1.1 **b.** Cytotoxicity test against HEK293T cells with cpfa#1:cpfa#1.1 mixture **c.** Toxicity to nematodes of a cpfa#1:cpfa#1.1 mixture when treated at L4/J4 stage and **d.** L1/J2 stage.

Only the anti-version (cpfa#1) retained any antimicrobial activity, indicating structural specificity for activity (Figure 9a) with a MIC of about 0.515 mM (125 ug/mL)

for *C. albicans*. The necessity for a particular stereochemistry for activity may imply that the antimicrobial effect is induced through signaling. At the MIC of 0.515 mM for *C. albicans*, cpfa#1 appears to be toxic to mammalian cells (Figure 9b). Additionally, cpfa#1 appears toxic to L4s/J4s (Figure 9c), being lethal at 125 ug/mL and tolerable at 62.5 ug/mL to *P. pacificus*. However, it appears that *P. pacificus* grown with cpfa#1 since J2 (Figure 9d) will survive to adulthood at around a 30% rate, even at the MIC of 125 ug/mL, whereas same concentration is fatal for *C. elegans*. Thus, *C. elegans* appear to be slightly more sensitive to cpfa#1 than *P. pacificus*. Together, this may imply that *P. pacificus* has a higher tolerance to cpfa#1 than *C. elegans*, but both organisms are susceptible to toxicity at the MIC for *C. albicans*.

3.4 DISCUSSION

Bioactivity

We show that the *P. pacificus* exometabolome contains molecules which can affect microbial proliferation. Of these many fractions, we have elucidated the structure of cpfa#1, and verified its activity with total synthesis. Of the microbes that this compound showed efficacy against, *Candida albicans* proved the most susceptible. Candida is an invasive fungus that is deadly in immunocompromised individuals and there is a dire need for new classes of antibiotics to battle against antibiotic resistant candida. cpfa#1 somewhat resembles mycolic acids from the membrane of

Mycobacterium tuberculosis, which requires the cyclopropyl group to maintain its virulence to human cells[57]. Similar to the mycolic acids, it's clear that the cyclopropyl group in cpfa#1 plays a key role in its activity because the anti-configuration (cpfa#1) contains activity, whereas the syn product (cpfa#1.1) does not.

Although cpfa#1 does not seem promising as an effective antibiotic because of its high MIC value and toxicity against mammalian cells and nematodes, the method for identification has proven effective in discovering new antibiotics. In a further search for antibiotic compounds in *P. pacificus* we came across several fractions with compounds containing cyclopropyl groups that have activity in the growth inhibition assay. One such a compound is cpfa#2, which seems to show antimicrobial effect against *Mycobacterium smegmatis* (Supplementary Table S11, Supplementary File - F11).

Although the cause of cpfa#1 toxicity is unknown, a number of analogous free fatty acids have been shown to have antimicrobial activity, and their mechanism of action has been determined[59]. For example, oleic acid causes an increase in membrane fluidity of *Staphylococcus aureus* and eventually cell death[60]. Oleic acid is also a potent glucosyltransferase inhibitor[61], implying that the antimicrobial activity of fatty acids is partially due to inhibition of membrane forming enzymes[59].

Biosynthesis and overview

Through investigating the biosynthesis of cpfa#1 we found that it is produced through a combination of bacterial and nematode metabolism, which may mean that the

cyclopropyl lipid backbone may be entirely from *E. coli*. This cyclopropyl lipid is similar in structure to the known cyclopropyl lipid methyl lactobacillate[58], of which the absolute stereochemistry was determined. Methyl lactobacillate has a 9*R*, 10*S* configuration for the cyclopropyl group, which may indicate that cpfa#1 shares a similar stereochemistry of 7*R*, 8*S*, and if the hydroxy is anti, 9*R*.

Although cpfa#1 may not be effective as an antibiotic because of its toxicity to human cells, we have proven that through this method we can potentially find new antibiotics. There are still many other metabolomic fractions from this set that have the potential for new discovery.

3.5 METHODS

Growth Inhibition Assay

Metabolomic fractions were reconstituted in DMSO and arranged in a 96-well plate. The concentration was brought to 6.25 mg/mL depending on starting mass for the metabolomic fraction. 100 uL of culture was added to each well in a plate at a concentration of $5x10^5$ CFU/mL. 1uL of each fraction was added to its corresponding well. 2 ug/mL of Amphotericin or 5 ug/mL of Cipro was added to a positive control well. The plate was allowed to incubate for 16-18 hours at 37C in a humidity chamber. The O.D.s were then read in a plate reader. The percent inhibition was determined through the following formula:

((O.D negative control - O.D. fraction)/ O.D. negative control)*100%

Biofilm Assay

A 2 hr subculture is developed from an overnight culture. The culture is then diluted to OD600 0.0025. 100 uL of cell culture is grown in 8x replicate with 1 uL DMSO/metabolomic fraction for 25 hr at 30C. The wells are then washed to rid of motile bacteria and then stained with Crystal Violet. A590 is then read.

Metabolomic fraction collection

See previous chapter methods

Strains and culture conditions

The Keio K12 methyl transferase mutant was used for feeding experiments to determine the source of cyclopropyl in cpfa#1. *See previous chapter methods* for additional info.

HEK Cell Toxicity Assay

In 96 well plates, starting with 3000 cells/well, cells were allowed to grow adherent overnight prior to replacement of DMEM media with the addition of cpfa at the stated concentrations. Cells were incubated for 3 days at 37C and incubated with MTT on the final day for 3 hours. Processed MTT metabolites (signifying actively metabolizing cells) were then solubilized and absorbance was recorded with a plate reader. A high A590 value means that the cells are living. To give some relative values, the MIC is around 0.5

mM, which corresponds to somewhere between the 4th and 5th from the right concentration. 4.85 mM (the maximum possible concentration) was started with and serial dilutions were conducted 2-fold per each well. The DMSO control had an A590 of roughly 0.716 and taxol (positive control for killing) was 0.082.

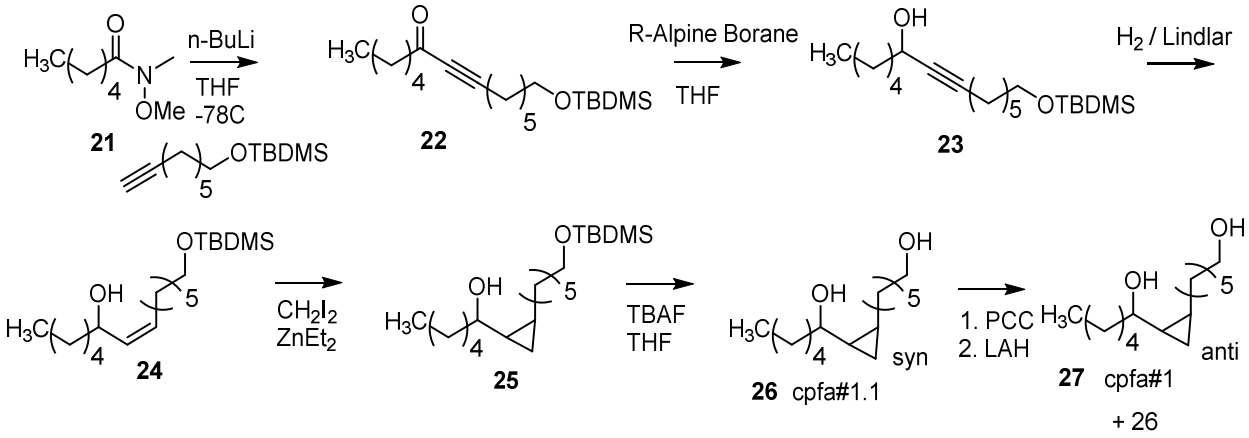
L1/J2 to day 1 adult developmental assay:

Using egg prepped N2 or RS2333, arrested L1s / J2s, were diluted in S-Complete suspension until there was 1 worm / uL. 30 worms were then added to each 96 well plate's well (on average based on the dilution calculation), which had a total of 100 uL of S complete with concentrated OP50 and 1 uL of CPFA. The CPFA was serially diluted across wells to different concentrations. The plate was allowed to incubate in a humidity chamber at 20C for 2 days until L4s and counted if there was movement or not. The next day, the wells were washed to get a more accurate count of the adults.

L4/J4 to day 2 adult assay,

Egg prepped N2 or RS2333 worms were placed in OP50 for 2 days at 20C. 5 L4/J4 worms were then picked into each 96 well. The CPFA was serially diluted as before and each well had OP50. The worms were grown at 20C overnight in a humidity chamber and

counted how many were alive in each well. The worms were allowed to grow for another day at 20C and survival score was counted.



Synthetic Scheme 4 Overview of synthesis of a. cpfa#1.1, and b. cpfa#1 Reagents and conditions labeled in figure.

Preparation of 22

n-BuLi (2.8 mL of 2.5 M soln. in THF, 7 mmol) was added dropwise to o-TBDMS-octynol (1.472 g, 6.128 mmol) at -78C. After stirring for 1 hour, **21** (1.114 g, 7 mmol) was added and the mixture was warmed to 0C. After stirring for 2 hours the mixture was quenched with NH_4Cl (sat, aq) and extracted with diethyl ether (2 x 15 mL). The organic layer was washed with brine, dried with magnesium sulfate, filtered, then concentrated in vacuo. Flash column chromatography on silica using a gradient of 0–100% ethyl acetate in hexanes afforded **22** (1.21 mmol, 17%).

^{1}H NMR (500 MHz, chloroform-*d*): δ (p.p.m.) 3.60 (t, J = 6.45 Hz, 2 H), 2.51 (t, J = 7.38 Hz, 2 H), 2.36 (t, J = 7.10 Hz, 2 H), 1.66 (quin, J = 7.77 Hz, 2 H), 1.57 (quin, J = 7.77 Hz , 2 H), 1.56 – 1.49 (m, 4 H), 1.45 – 1.26 (m, 7 H), 0.91 – 0.89 (s, 12 H), 0.04 (s, 6 H).

Preparation of 23

22 (410.3 mg, 1.21 mmol) was cooled to 0C and S-Alpine Borane (0.5 M in THF, 3 mmol) was added. The reaction was then concentrated and put under argon. After stirring for 3 days, propionaldehyde (75 uL) was added, and the mixture was concentrated under vacuum at 40C. The mixture was diluted with THF (1.5 mL), cooled to 0C, then NaOH (3M, 1.5 mL) and H_2O_2 (1.5 mL, 30%) were added. The mixture was warmed to 40C and stirred for 4 hours. The reaction was extracted with ether, dried with magnesium sulfate, and concentrated in vacuo. Flash column chromatography on silica using a gradient of 0–100% methanol in DCM afforded **23** (0.56 mmol, 46%). ^{1}H NMR (500 MHz, chloroform-*d*): δ (p.p.m.) 4.34 (q, J = 8.82 Hz, 1 H), 3.60 (t, J = 6.62 Hz, 2 H), 2.40 (m, 1 H), 2.20 (td, 2.11 Hz, 7.47 Hz, 2 H), 1.88 (m, 1 H), 1.79 – 1.60 (m, 3 H), 1.57 – 1.25 (m, 12 H), 0.91 – 0.89 (12 H), 0.05 (s, 6 H).

Preparation of 24

Quinoline (31.2 uL, 0.27 mmol) and Lindlar's catalyst (8.3 mg) were added to **23** (189.3 mg, 0.56 mmol) in hexanes. The reaction vial was flooded with argon, then H_2. After stirring overnight, flash column chromatography on silica using a gradient of 0–100% methanol in DCM afforded **24** (0.47 mmol, 84%).

^{1}H NMR (500 MHz, chloroform-*d*): δ (p.p.m.) 5.47 (m, 1 H), 5.36 (m, 1 H), 4.42 (m, 1 H), 3.60 (t, J = 6.66 Hz, 2 H), 2.13 – 2.03 (m, 2 H), 1.59 (m, 1 H), 1.50 (quin, J = 6.67 Hz, 2 H), 1.45 – 1.25 (15 H), 0.92 – 0.89 (m, 12 H), 0.04 (s, 6 H).

Preparation of 25

24 (0.47 mmol) was cooled to -10C in dry dichloromethane and then $ZnEt_2$ (4.24 mL, 4.24 mmol, 1M in THF, 5 eq) was added dropwise. In the dark, CH2I2 (4.24 mmol, 5 eq) was added dropwise. After stirring for one day flash column chromatography on silica using a gradient of 0–100% methanol in DCM afforded **25** (0.316 mmol, 67%).

^{1}H NMR (500 MHz, chloroform-*d*): δ (p.p.m.) 3.64 (t, J = 6.74 Hz, 2 H), 3.12 (m, 1 H), 1.59 – 1.54 (m, 2 H), 1.55 – 1.48 (m, 4 H), 1.47 – 1.27 (m, 12 H), 0.93 – 0.88 (m, 12 H), 0.82 (m, 2 H), 0.67 (m, 1 H), 0.07 (q, J = 5.00 Hz, 1 H), 0.06 (s, 6 H).

Preparation of 26

TBAF (0.6 mmol, 2 eq) was mixed with **25** (109.5 mg, 0.3 mmol) in THF under argon. After stirring for 3 hours, the reaction was quenched with water (1 mL), then flash column chromatography on silica using a gradient of 0–100% methanol in DCM afforded **26**, cpfa#1.1 (0.27 mmol, 90% yield).

^{1}H NMR (500 MHz, methanol-d_4): δ(p.p.m.) 3.54 (t, J = 6.46 Hz, 2 H), 3.12 (m, 1 H), 1.58 – 1.48 (m, 7 H), 1.46 – 1.26 (m, 10 H), 1.07 (m, 1 H), 0.92 (t, J = 7.25 Hz, 3 H), 0.82 (m, 2 H), 0.70 (m, 1 H), 0.08 (q, J = 5 Hz, 1 H). ^{13}C NMR (500 MHz, methanol-d_4): δ(p.p.m.) 73.27, 62.99, 39.32, 33.66, 33.12, 31.30, 30.51, 30.30, 26.98, 26.46, 23.91, 23.73, 17.50, 14.42, 10.63.

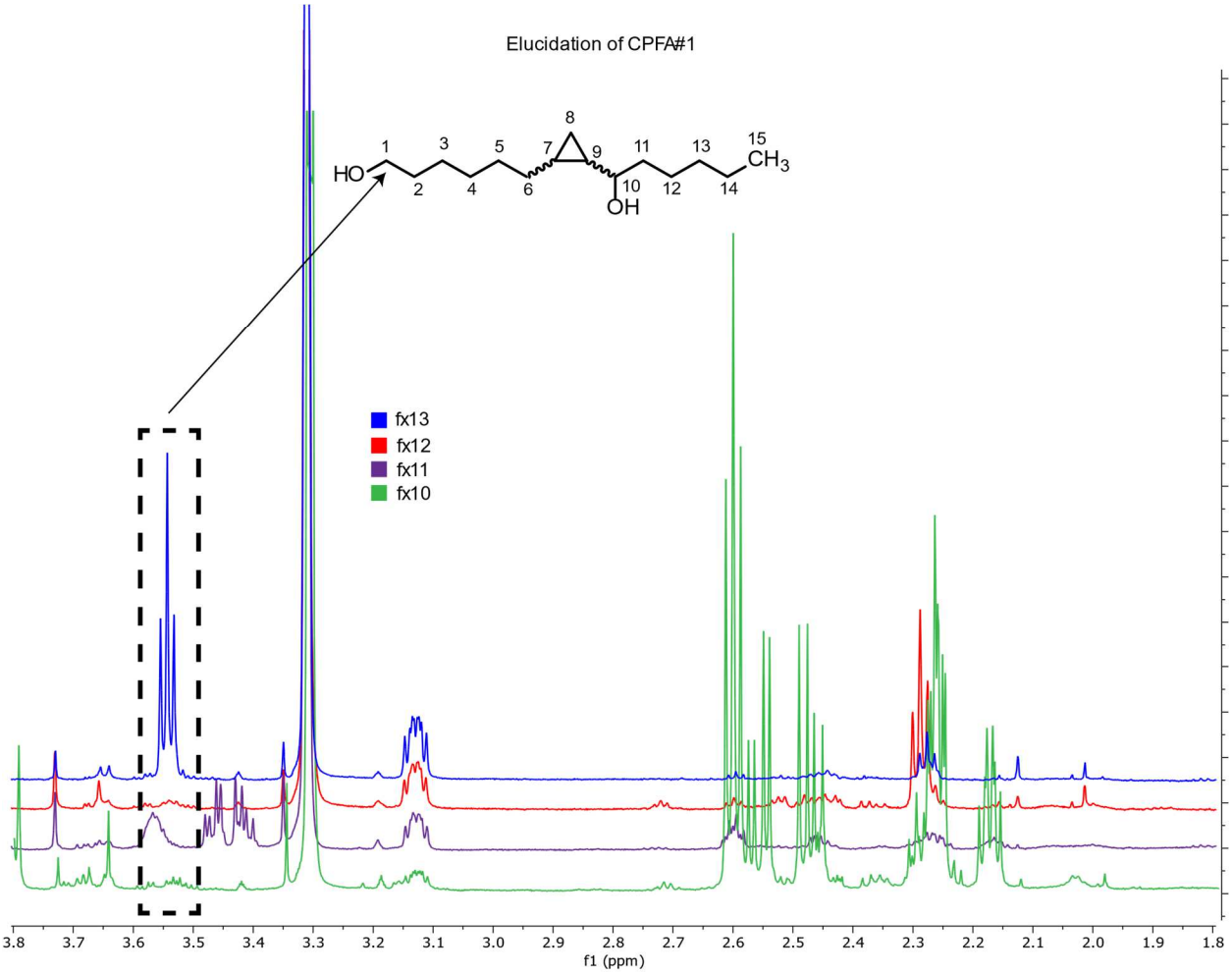
Preparation of 27

26 (25.5 mg, 0.1 mmol) was mixed with pyridinium chlorochromate (43.1 mg, 0.2 mmol). After one day of stirring, flash column chromatography on silica using a gradient of 0–100% ethyl acetate in hexanes afforded the ketone product (19.4 mg, 0.082 mmol, 81.5%). Lithium aluminum hydride (one pipette-tip full) was then added to the ketone in THF. After 1 hr of stirring, flash column chromatography on silica using a gradient of 0–100% ethyl acetate in hexanes afforded **27**, cpfa#1.

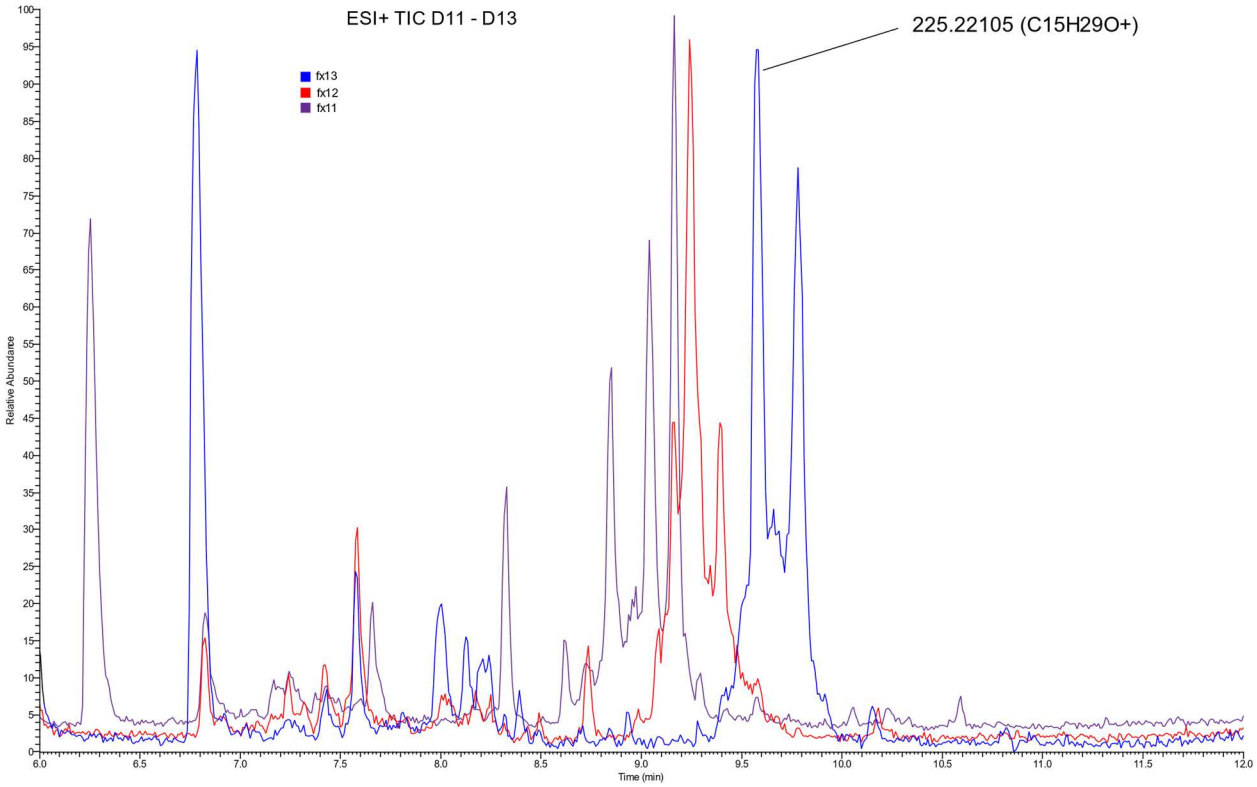
^{1}H NMR (500 MHz, methanol-d_4): δ (p.p.m.) 3.54 (6.53 Hz, 2 H), 3.13 (m, 1 H), 1.70 (m, 1 H), 1.60 – 1.43 (m, 7 H), 1.40 – 1.28 (m, 9 H), 1.14 (m, 1 H), 0.92 (m, 3 H), 0.84 – 0.79 (m, 2 H), 0.69 (m, 1 H), -0.13 (q, J = 5 Hz, 1 H). ^{13}C NMR (500 MHz, methanol-d_4): δ (p.p.m.) 73.39, 63.03, 39.50, 33.69, 33.25, 31.15, 30.60, 29.84, 26.99, 26.59, 23.75, 23.51, 17.10, 14.41, 10.32.

3.6 SUPPLEMENTARY FIGURES

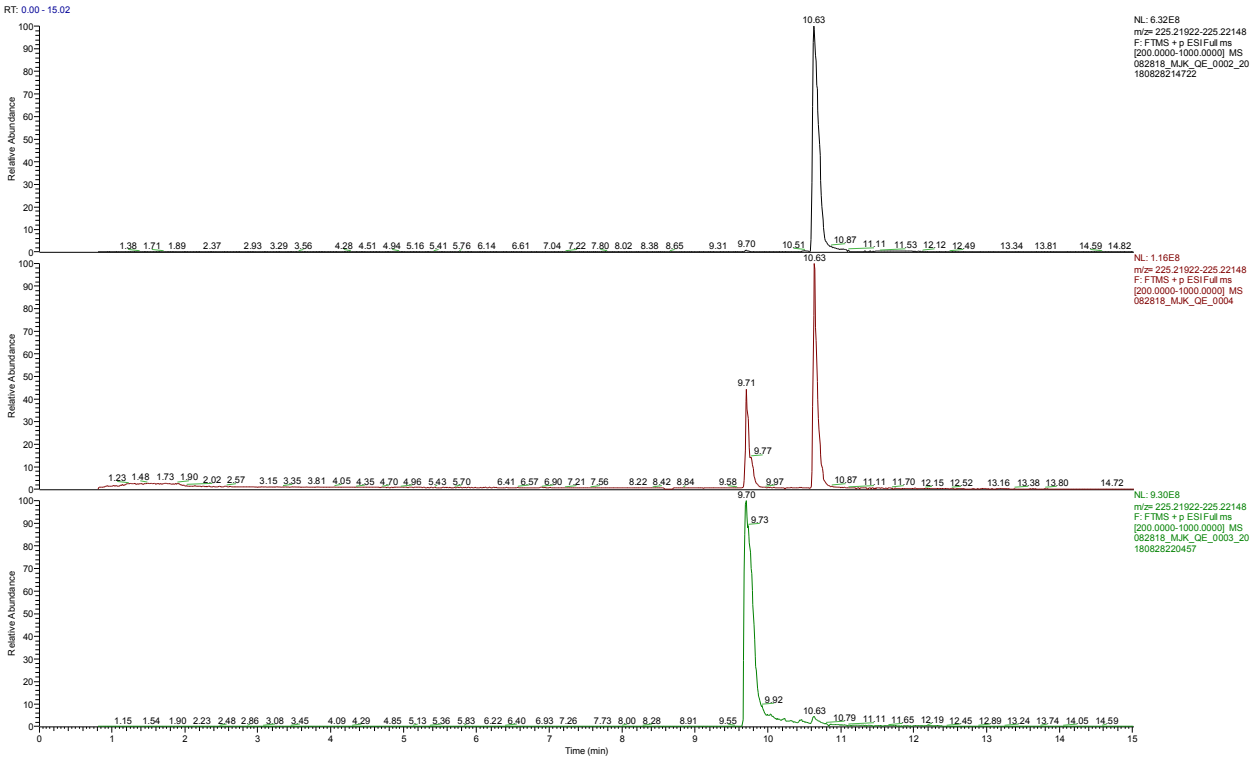
Results from antibiotic assays conducted are located in Supplementary File 1



Supplementary Figure S13. Differential analysis of NMR spectra from fractions D10-13. Highlighted in black are the CH2-group spin system which was the first identified differential feature.



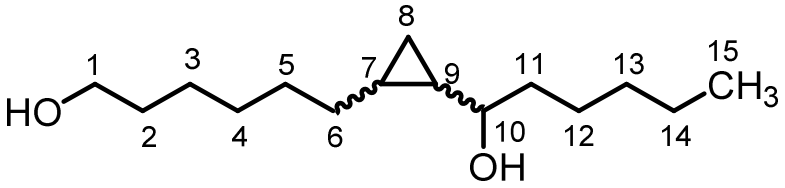
Supplementary Figure S14. Differential analysis of MS spectra from fractions D11-13. The feature belonging to CPFA#1 is indicated.



Supplementary Figure S15. Elution profiles of natural product cpfa#1 (top), cpfa#1.1 + cpfa#1 mixture (middle), and cpfa#1.1 (bottom).

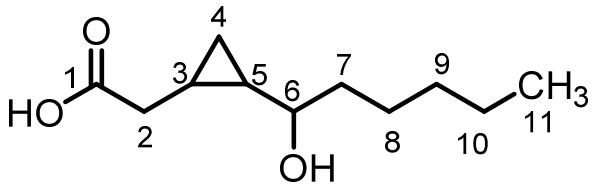
3.7 SUPPLEMENTARY TABLES

Supplementary Table S10. ^{1}H and ^{13}C NMR spectroscopic data for cpfa#1 in methanol-d_4. Chemical shifts were referenced to δ ($CD_2\underline{H}OD$) = 3.31 and δ ($^{13}\underline{C}D_3OD$) = 49.0. ^{13}C chemical shifts were determined via HMBC and HSQC spectra. Spectra were acquired using the Bruker Avance 800 spectrometer. ^{1}H, ^{1}H-J-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom. ROESY in ACN.



Position	^{13}C[ppm]	^{1}H [ppm]	1H-1H coupling constants (Hz)	HMBC Correlations	ROESY Correlations	^{13}C [ppm] synthetic	^{1}H [ppm] synthetic
1	61.63	1a = 3.53 1a = 3.53	$J_{1,2}$ = 6.74	2, 3		63.03	3.54
2	33.36	2a = 1.52 2b = 1.53	n/a	1, 3, 4		33.69	n/a
3	26.26	3a = 1.47 3b = 1.39	n/a	--		26.99	n/a
4	29.16	4a = 1.37 4b = 1.29	n/a	3		30.60	n/a
5	30.73	5a = 1.47 5b = 1.50	n/a	3, 6		31.15	n/a
6	29.54	6a = 1.69 6b = 1.13	n/a	4, 5, 9	6a = 6b = 8b	29.84	n/a
7	16.73	0.79	n/a	5, 8, 9, 10	8b	17.10	0.78-0.81
8	10.10	8a = 0.68 8b = -0.15	m	6, 7, 9, 10	8a = 8b = 7, 6b	10.32	8a = 0.69 8b = -0.13
9	23.17	0.83		6, 7, 8, 10, 11		23.51	0.81-0.83
10	73.08	3.12				73.38	3.13
11	39.20	11a = 1.54 11b = 1.58	m	10, 12, 13		39.50	n/a
12	26.56	12a = 1.47 12b = 1.39	n/a	10, 13, 14		26.59	n/a
13	33.01	13a = 1.30 13b = 1.30	n/a	--		33.25	n/a
14	23.48	14a = 1.33 14b = 1.34	n/a	12, 13, 15		23.75	n/a
15	14.05	0.89	n/a	13, 14		14.41	0.91

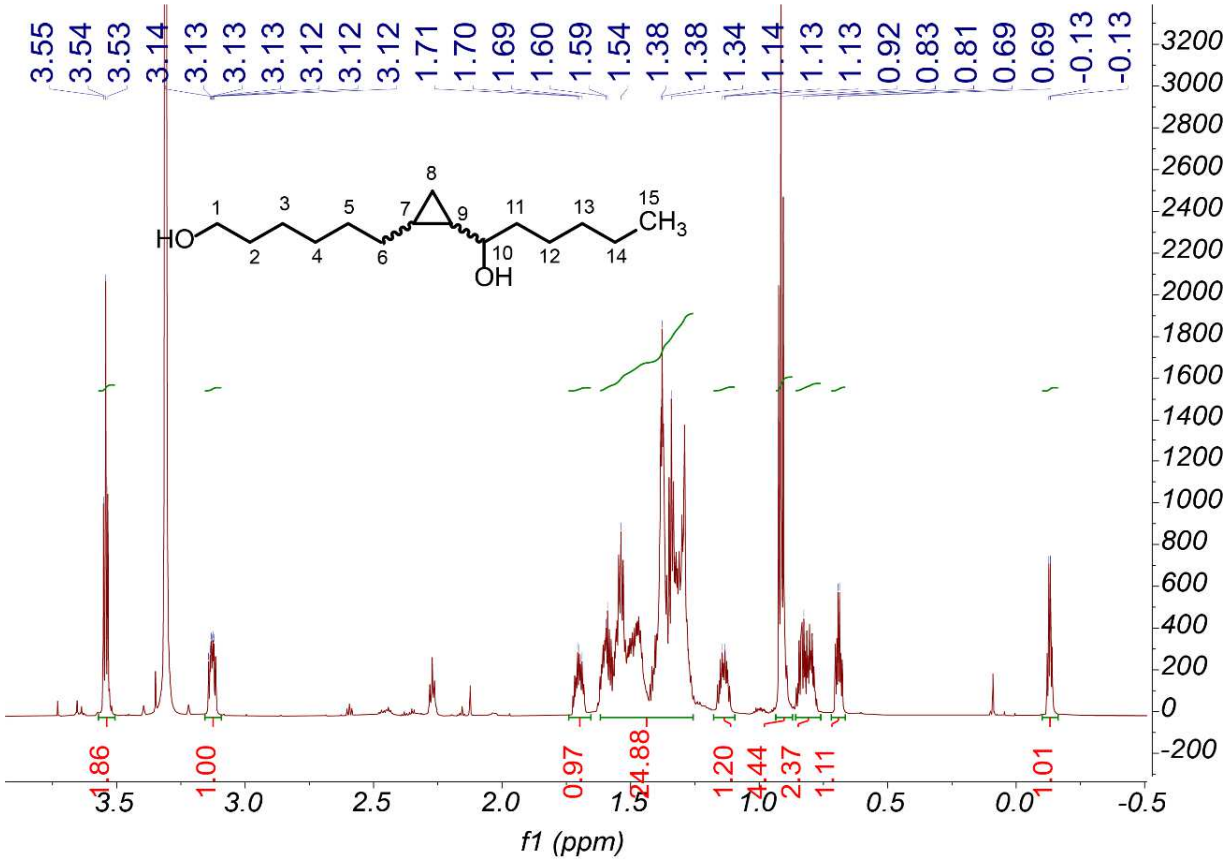
Supplementary Table S11. ^{1}H and ^{13}C NMR spectroscopic data for cpfa#2 in methanol-d_4. Chemical shifts were referenced to δ ($CD_2\underline{H}OD$) = 3.31 and δ ($^{13}\underline{C}D_3OD$) = 49.0. ^{13}C chemical shifts were determined via HMBC and HSQC spectra. Spectra were acquired using the Bruker Avance 800 spectrometer. ^{1}H, ^{1}H-J-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom. ROESY in CD_3OD.



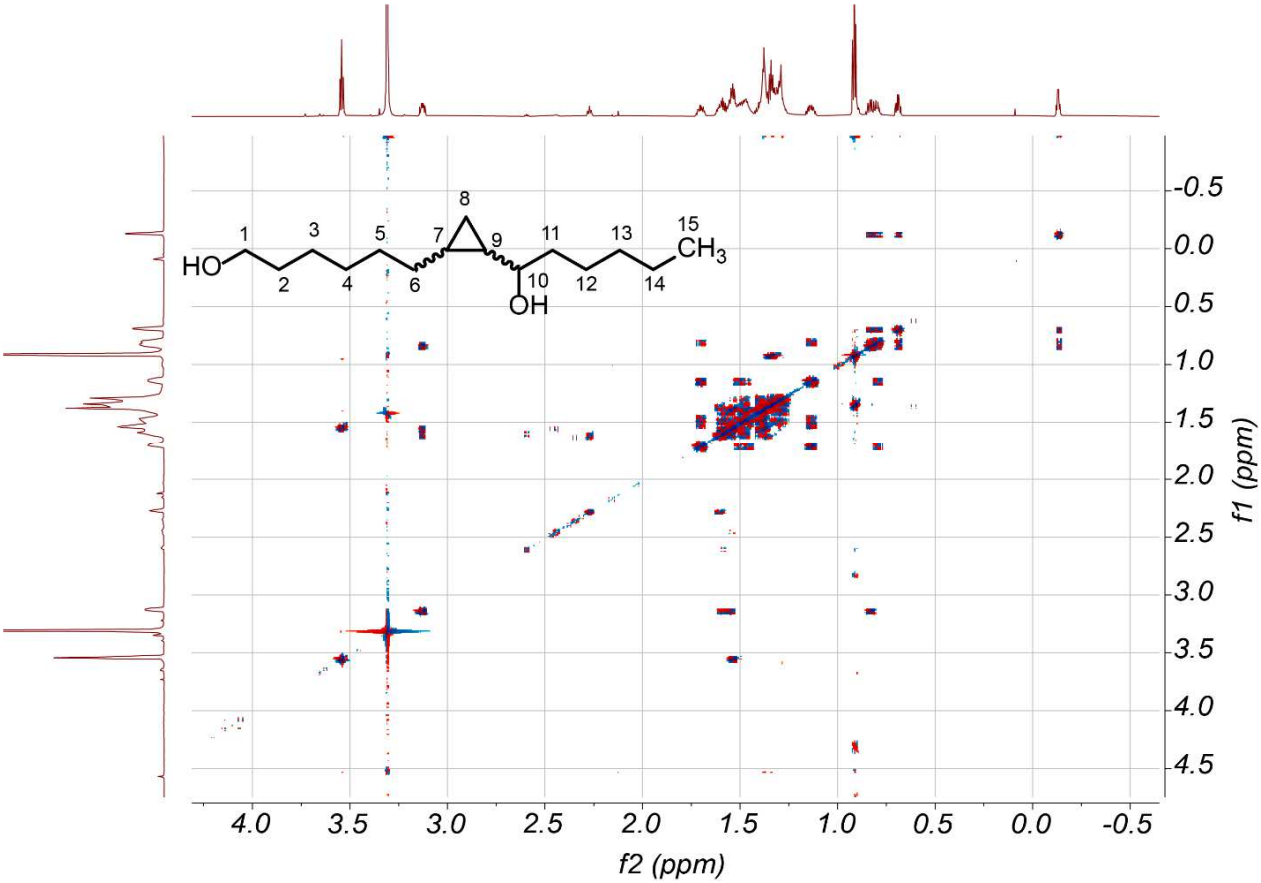
Position	^{13}C[ppm]	^{1}H [ppm]	1H-1H coupling constants (Hz)	HMBC Correlations	ROESY Correlations
1	189.45				
2	36.08	2a = 2.30 2b = 2.48		1, 3, 4, 5	
3	13.34	1.12	n/a	--	
4	9.40	4a = 0.05 4b = 0.78	$J_{4a,3,5}$ = 5.25 $J_{4b,\ 4a}$ = 8.43 $J_{4b,3,5}$ = 4.84	2, 3, 5, 6	4a = 4b 4a = 2a, 2b
5	23.37	0.98	n/a	2, 6, 7,	
6	73.21	3.20	n/a	3, 5, 7, 8	6 = 2a, 7b
7	38.29	7a = 1.60 7b = 1.60	n/a	5, 6, 8, 9	
8	26.28	8a = 1.41 8b = 1.48	n/a	6, 7, 9, 10	
9	32.94	9a = 1.31 9b = 1.31	n/a	7, 8, 10, 11	
10	23.37	10a = 1.35 10b = 1.35	n/a	8, 9, 11	
11	15.10	0.91	$J_{11,10}$ = 7.19	9, 10	

3.8 SUPPLEMENTARY NMR SPECTRA

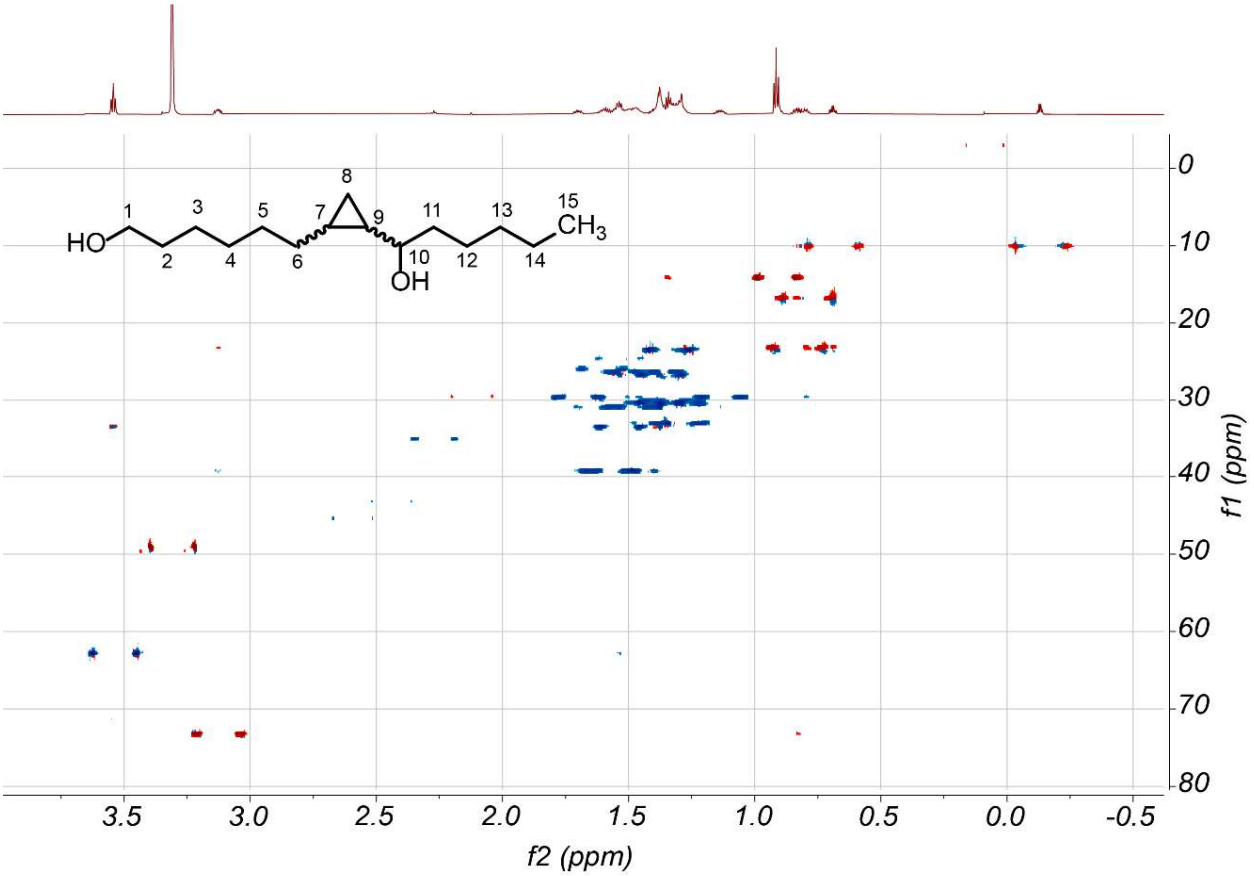
^{1}H NMR spectrum of active metabolome fraction containing cpfa#1, 31 (800 MHz, methanol- d_4).



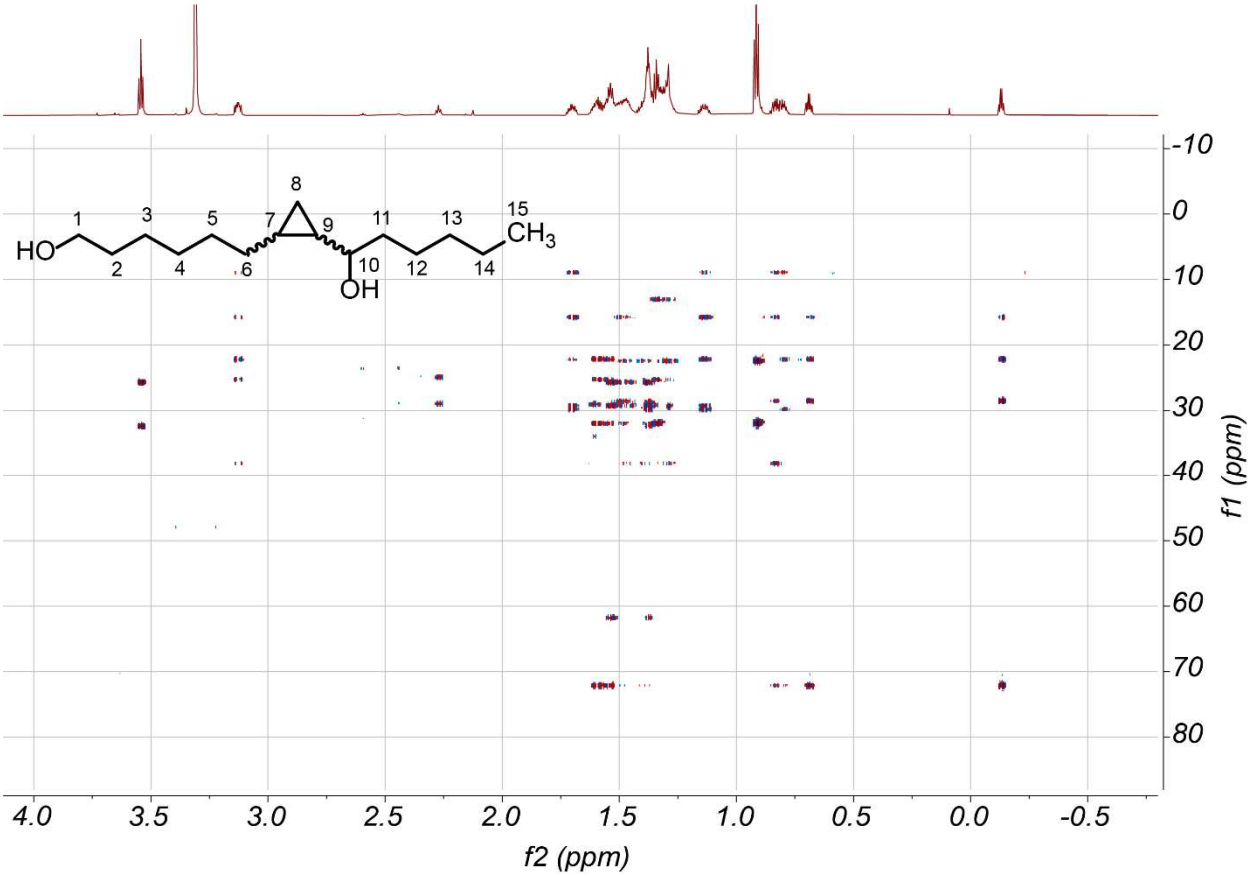
dqfCOSY spectrum of active metabolome fraction containing cpfa#1 (800 MHz, methanol-d_4).



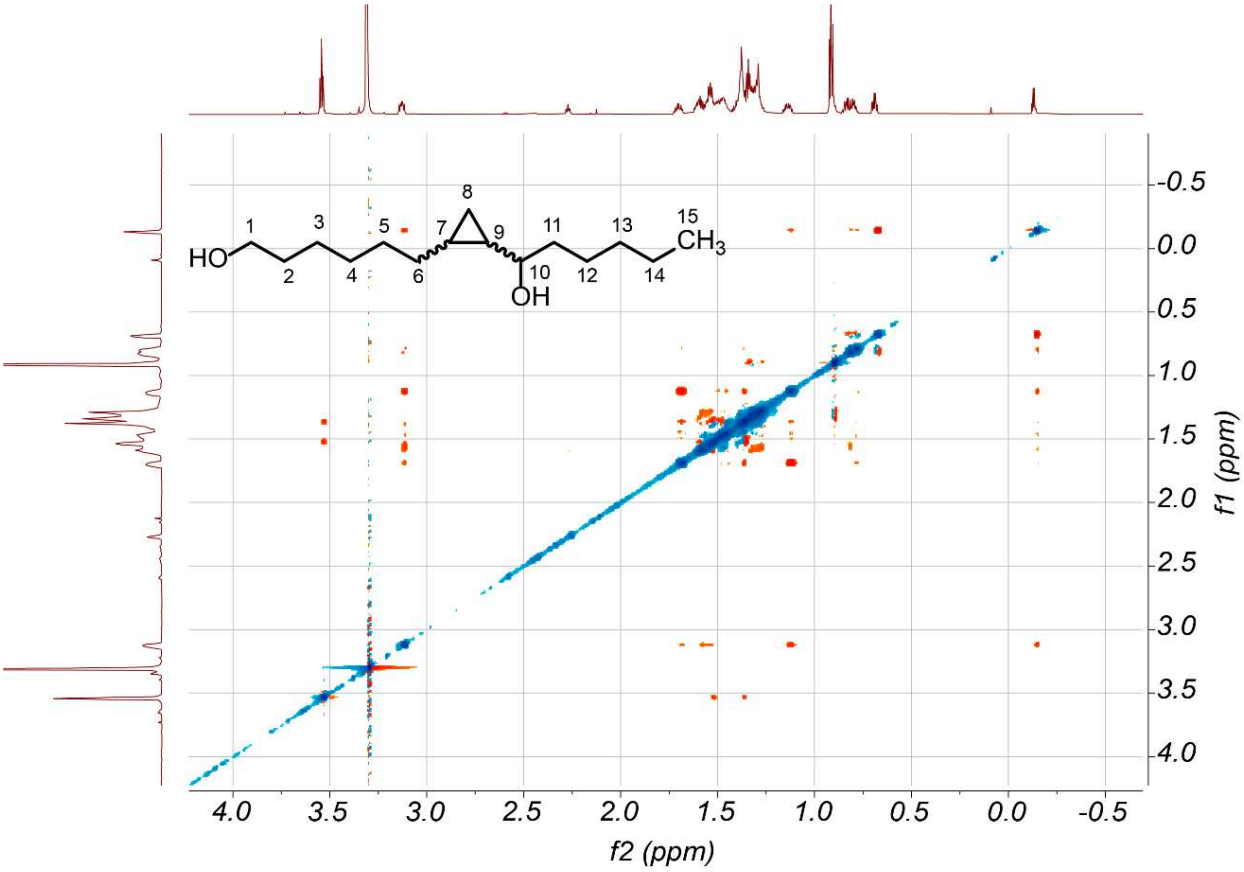
HSQC spectrum of active metabolome fraction containing cpfa#1 (800 MHz, methanol-d_4).



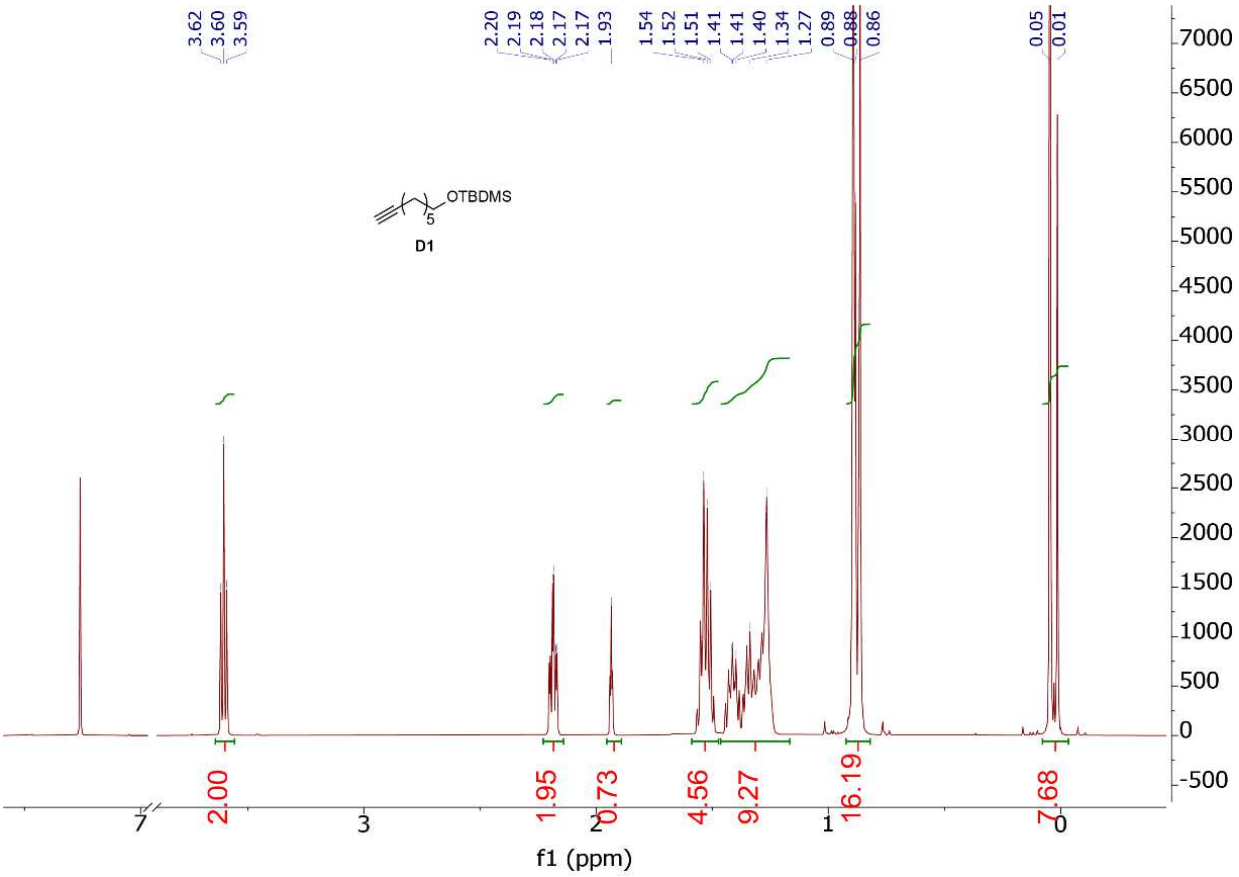
HMBC spectrum of active metabolome fraction containing cpfa#1 (800 MHz, methanol-d_4).



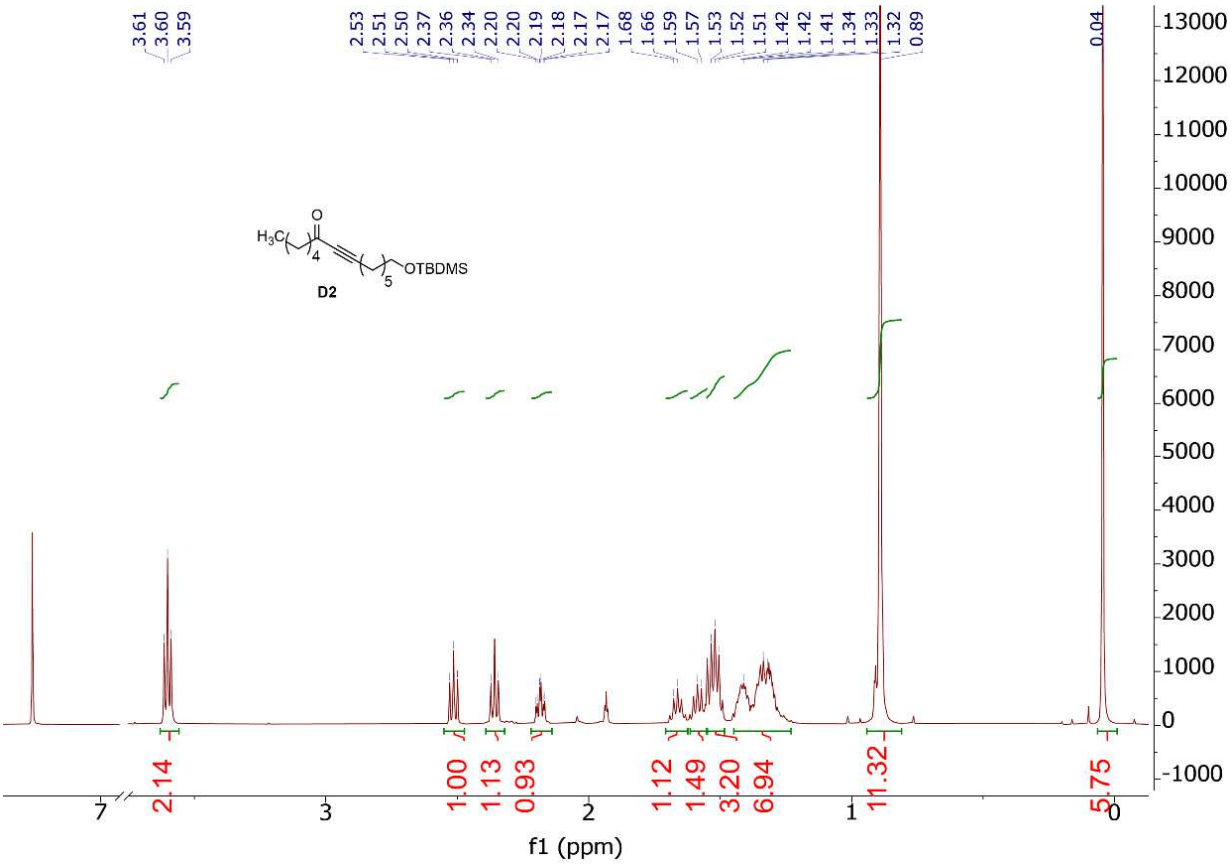
ROESY spectrum of active metabolome fraction containing cpfa#1 (800 MHz, methanol-d_4).



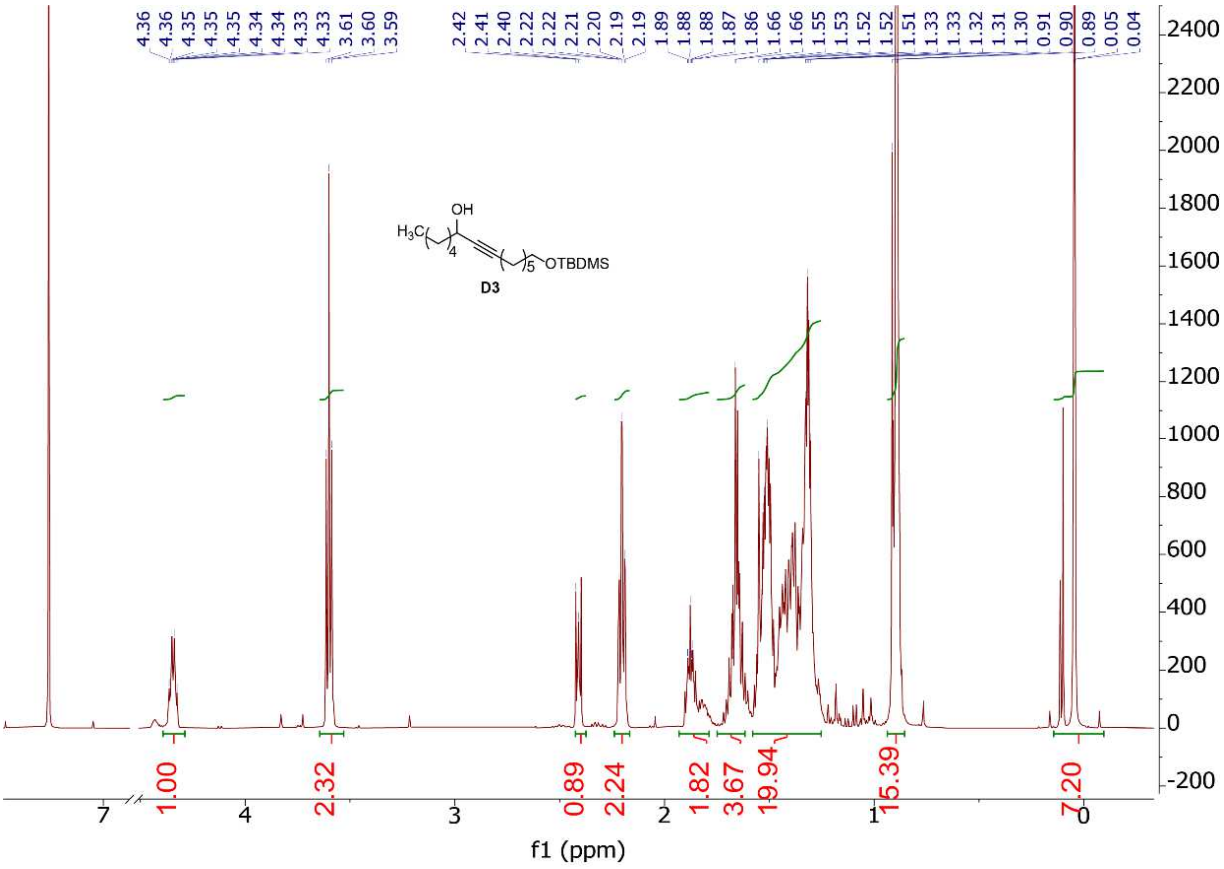
^{1}H NMR Spectrum (600 MHz, chloroform-d) of 21



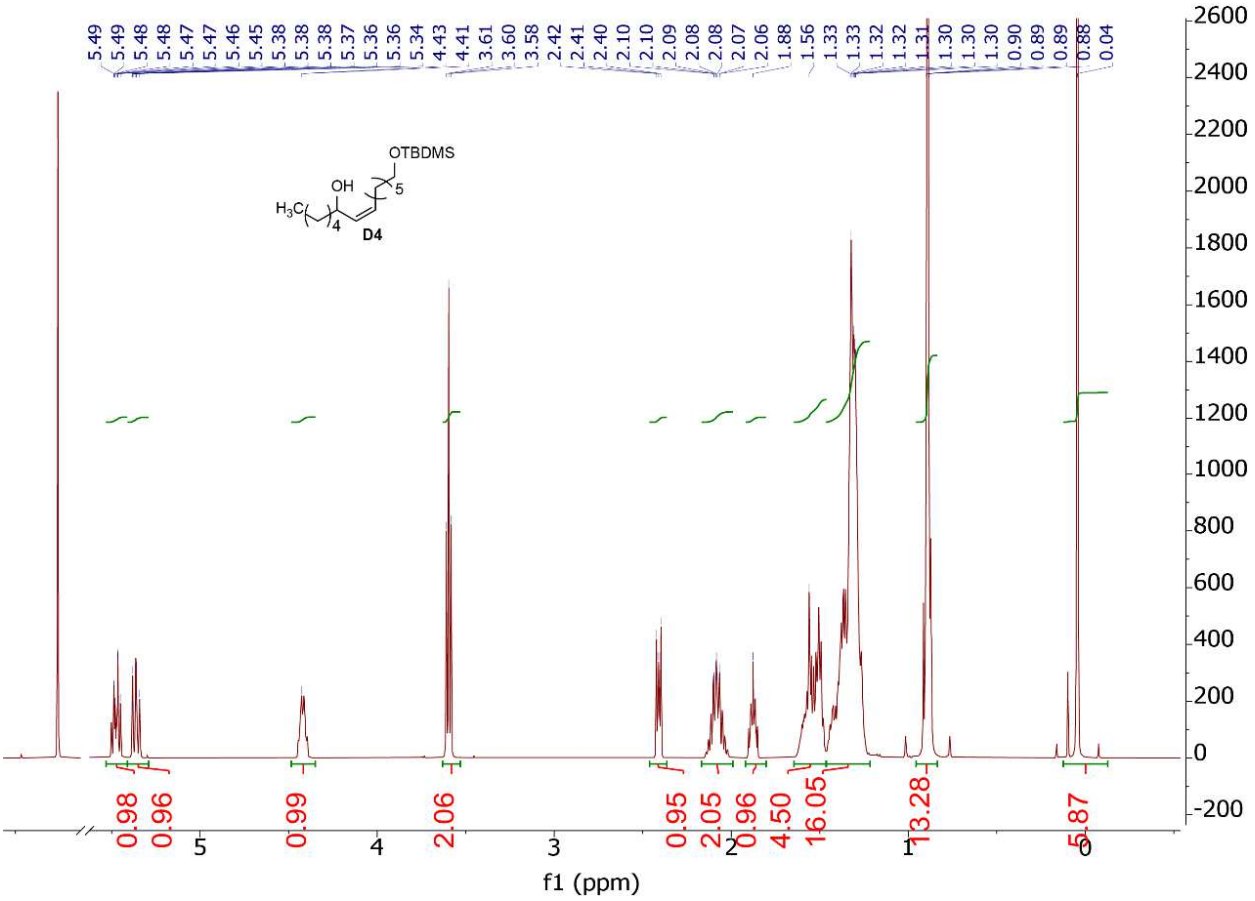
^{1}H NMR Spectrum (600 MHz, chloroform-d) of 22



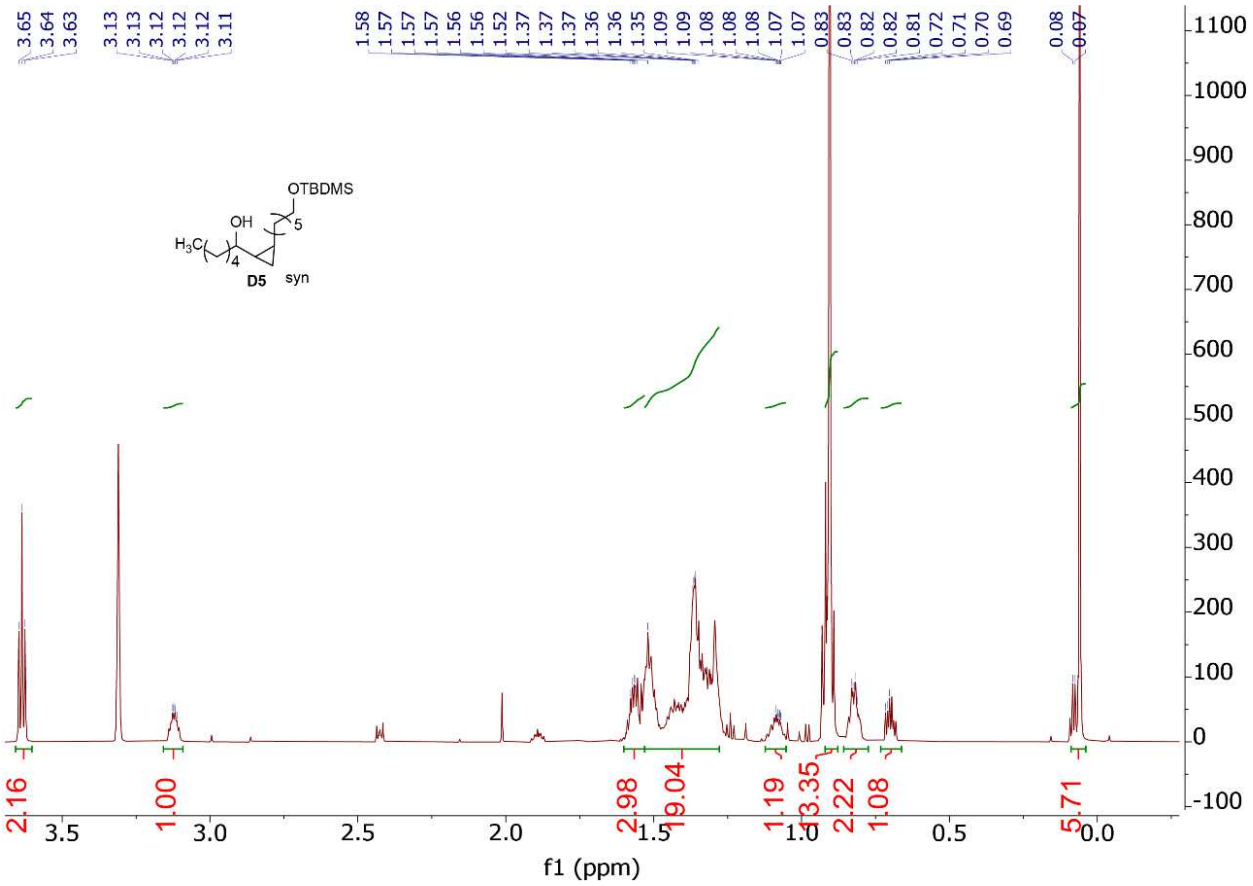
^{1}H NMR Spectrum (600 MHz, chloroform-d) of 23



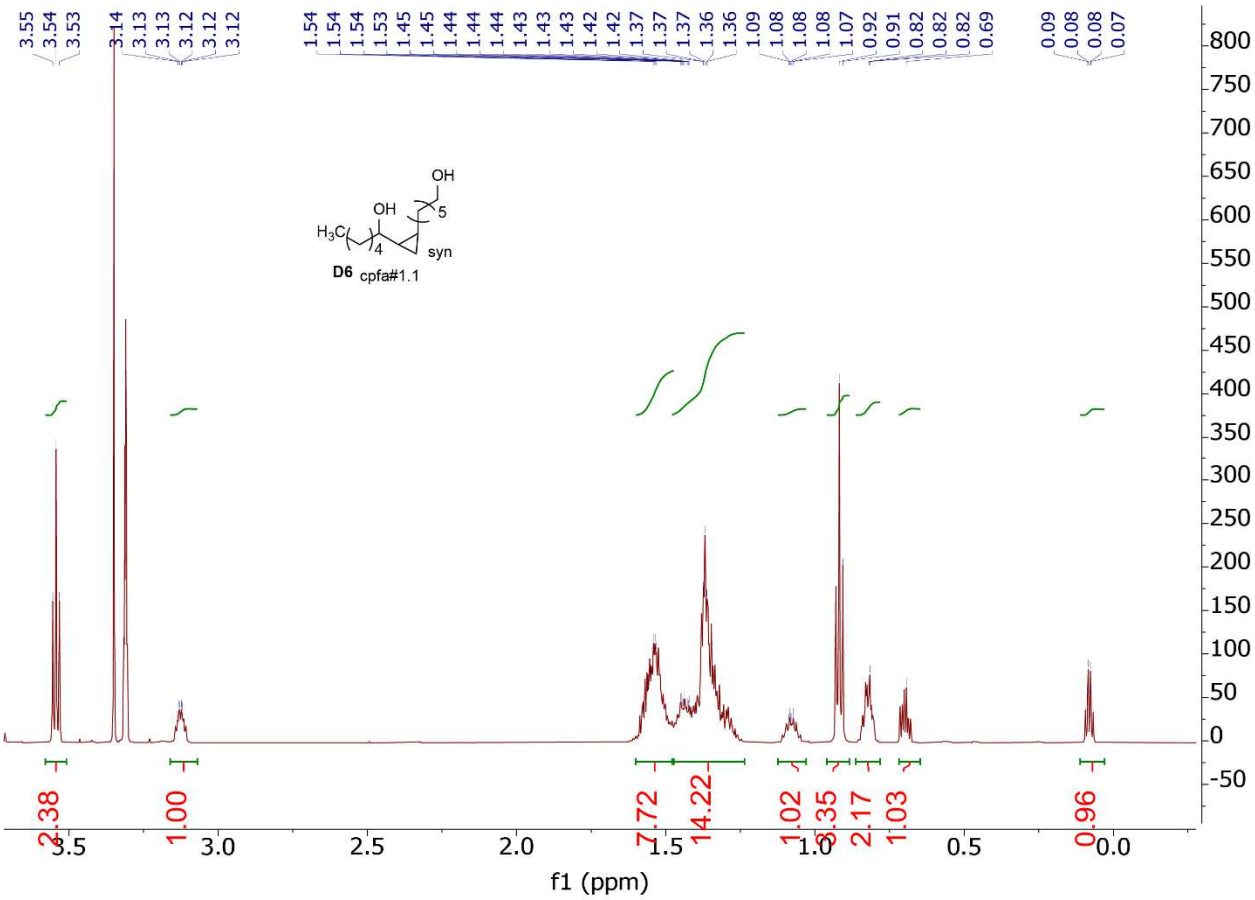
^{1}H NMR Spectrum (600 MHz, chloroform-d) of 24



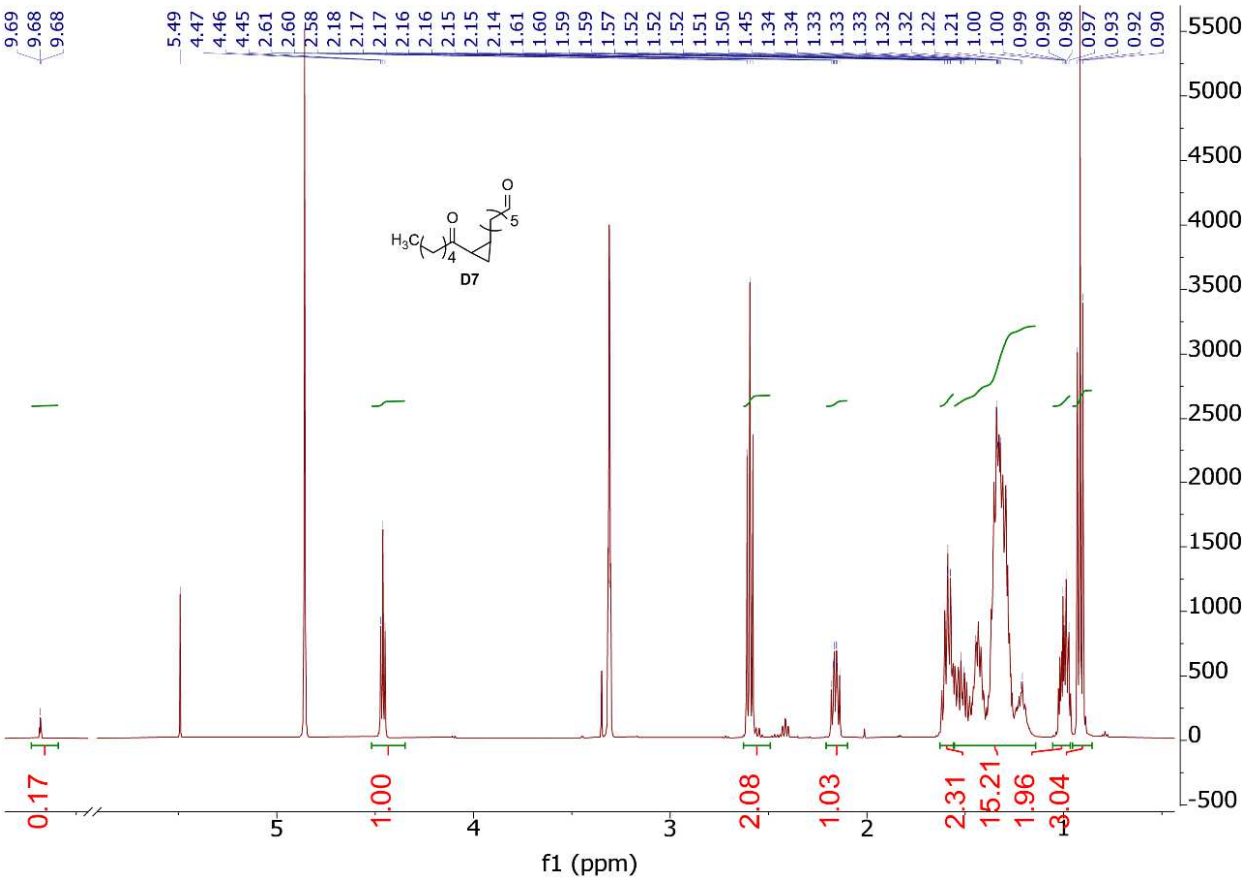
^{1}H NMR Spectrum (600 MHz, chloroform-d) of 25



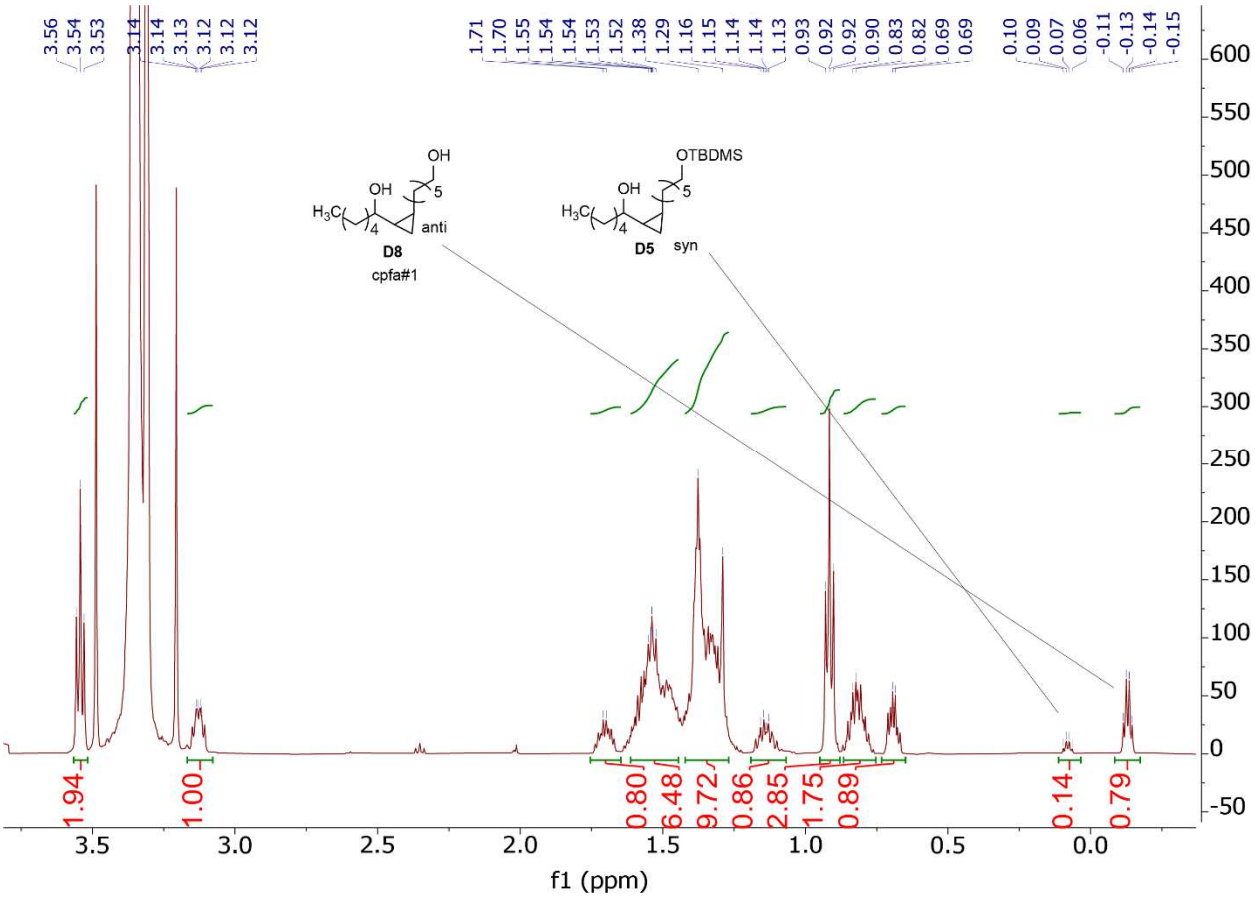
^{1}H NMR Spectrum (600 MHz, chloroform-d) of 26, cpfa#1.1



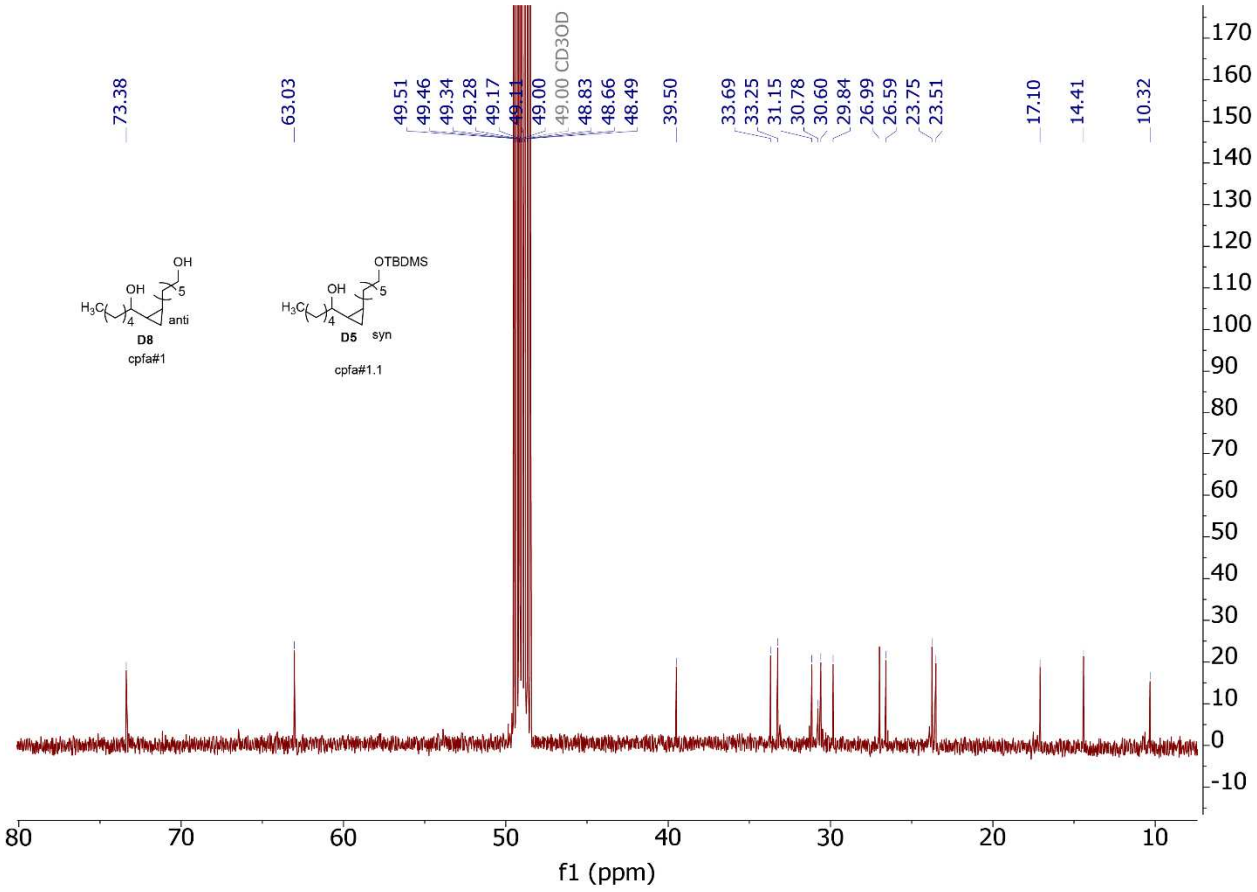
^{1}H NMR Spectrum (600 MHz, chloroform-d) of oxidized 26



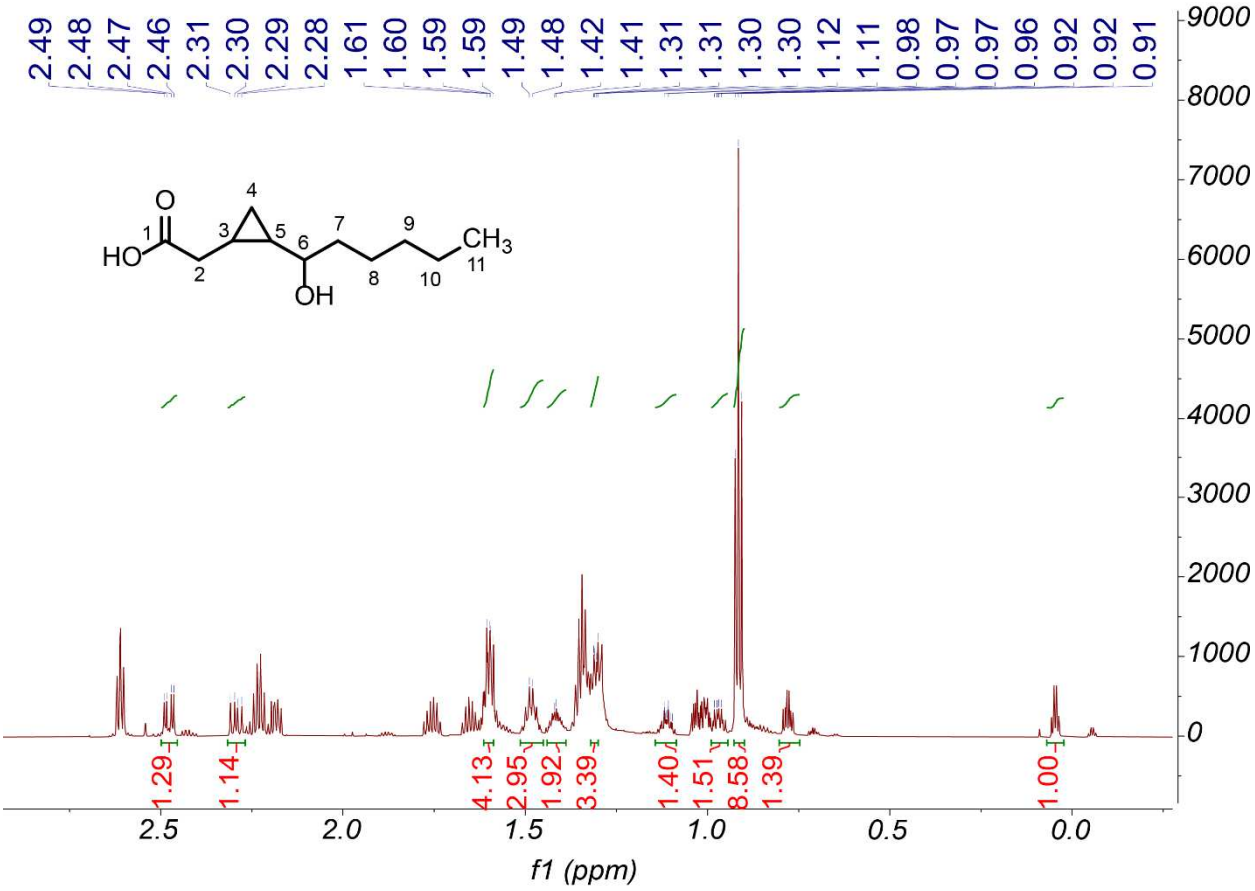
^{1}H NMR Spectrum (600 MHz, methanol-d_4) of cpfa#1.1, 26, and cpfa#1, 27



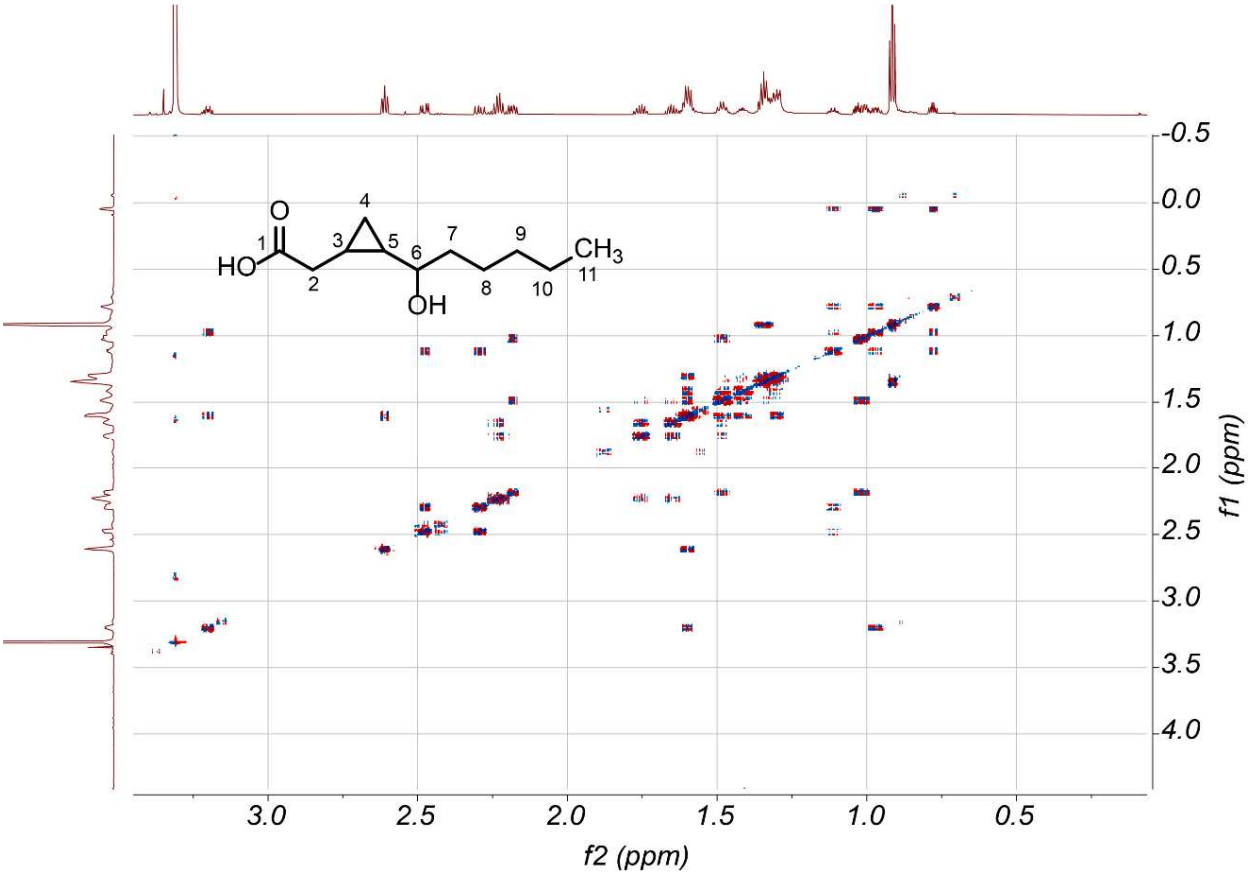
^{13}C NMR Spectrum (600 MHz, methanol-d_4) of cpfa#1.1, 26, and cpfa#1, 27



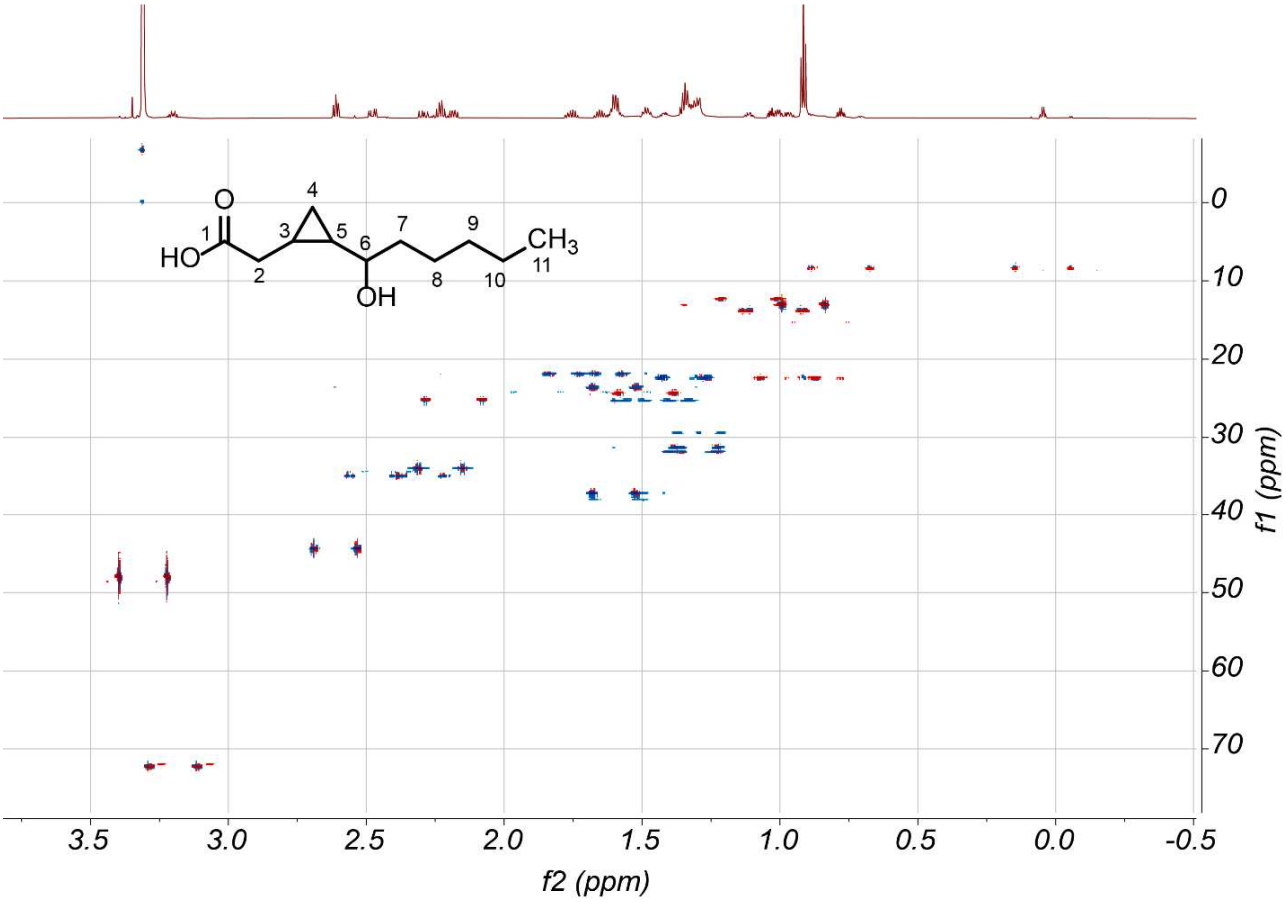
^{1}H NMR spectrum of active metabolome fraction containing cpfa#2, 40 (800 MHz, methanol-d_4)



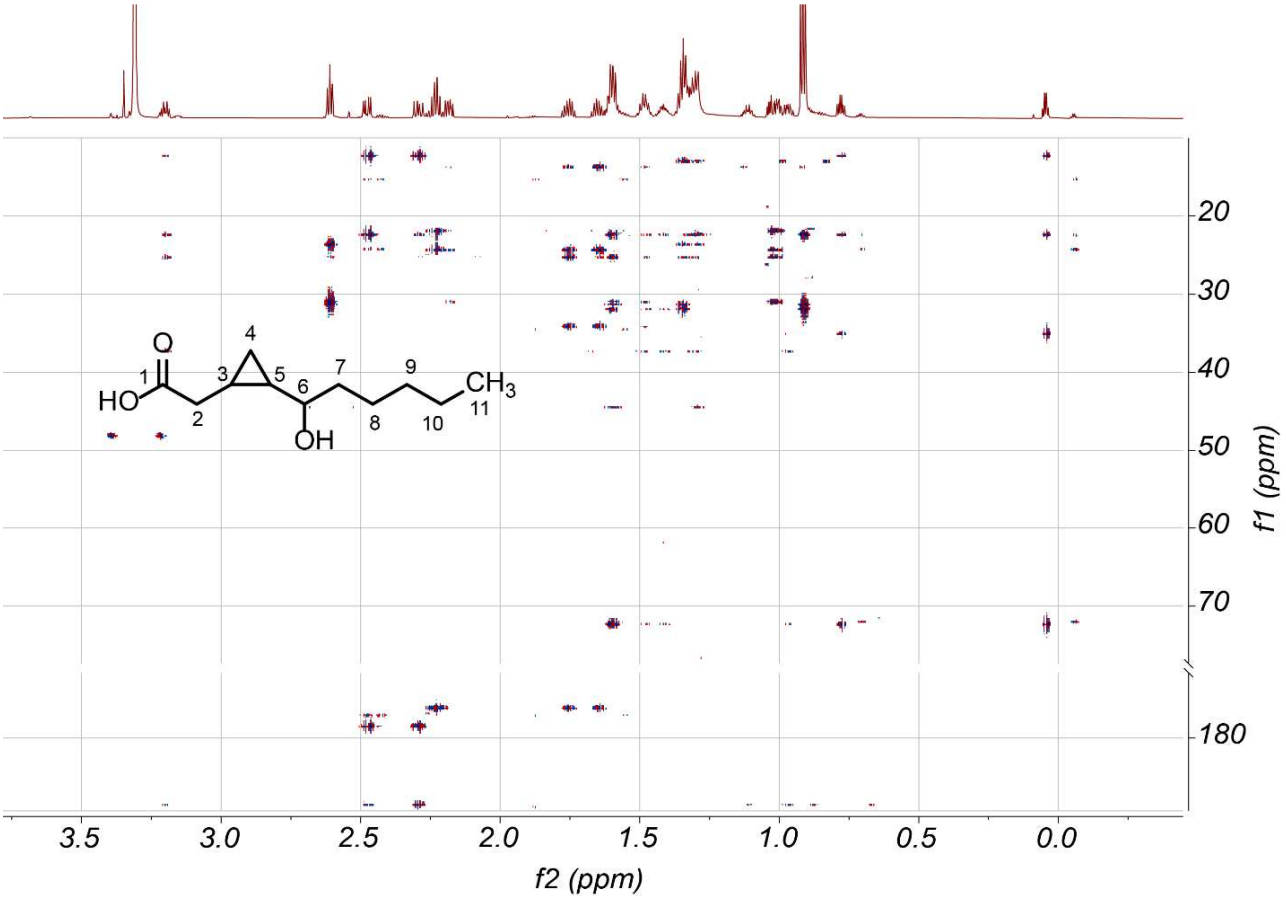
dqfCOSY spectrum of active metabolome fraction containing cpfa#2, 40 (800 MHz, methanol-d_4)



HSQC spectrum of active metabolome fraction containing cpfa#2, 40 (800 MHz, methanol-d_4)



HMBC spectrum of active metabolome fraction containing cpfa#2, 40 (800 MHz, methanol-d_4)



3.9 AUTHOR'S NOTE

All antibiotic assays were conducted by Harini Sadeeshkumar and Spencer Wong from Prof. Dennis Kim's lab. The initial samples for antibiotic assays were grown by Dr. Neelanjan Bose and Joshua Yim. Synthesis of cpfa#1 was partially conducted by Jazmin Aguilar-Romero.

REFERENCES

56. Beesoo R.; Bhagooli R.; Neergheen-Bhujun V.S.; Li W.W.; Kagansky A.; Bahorun T.*;* Antibacterial and antibiotic potentiating activities of tropical marine sponge extracts. *Comp. Biochem* (2017)

57. Glickman M.S.; Cox J.S.; Jacobs W.R.; A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of Mycobacterium tuberculosis. *Molecular Cell* (2000)

58. Stuart L.J.; Buck J.P.; Tremblay A.E.; Buist P.H.; Configurational analysis of cyclopropyl fatty acids isolated from Escherichia coli. *Org. Lett.*, Vol. 8 (1), 79-81 (2006)

59. Yoon B.K.; Jackman J.A.; Valle-Gonzalez E.R.; Cho N.J.; Antibacterial Free Fatty Acids and Monoglycerides: Biological Activities, Experimental Testing, and Therapeutic Applications. *Int. J. Mol. Sci* (2018).

60. Chamberlain N.R.; Mehrtens B.G.; Xiong Z.; Kapral F.A.; Boardman J.L.; Rearick J.I.; Correlation of carotenoid production, decreased membrane fluidity, and resistance to oleic acid killing in Staphylococcus aureus 18Z. *Infect Immun.* (1991).

61. Won S.R.; Hong M.J.; Kim Y.M.; Li C.Y.; Kim J.W.; Rhee H.I.; Oleic acid: an efficient inhibitor of glucosyltransferase. *FEBS Lett.* (2007)

62. Ventola C.L.; The antibiotic resistance crisis: part 1: causes and threats. *P T*.;40(4):277–283 (2015).

OUTLOOK

In this dissertation the author identifies small molecule metabolites produced by the nematode *P. pacificus* that regulate interactions with its environment. Namely, small molecules that can cause defensive behaviors in other nematodes, modulate bacterial growth, and affect nematode development. In chapter one, a combination of defensive behavior assays and UHPLC-MSMS/2D NMR analysis of metabolomic fractions were used to identify a class of sulfolipids from *P. pacificus* that cause defensive behaviors in *C. elegans*.

The small molecule analysis conducted in chapter one revealed a key mass spectral fragment of the sulfolipids that enabled us to uncover the 'sulfatome' of *P. pacificus* in chapter two. Amongst the dozens of lipids that comprise the sulfatome we found two unprecedented cyclic sulfates. We showed that the sulfolipids are produced *de novo* by *P. pacificus* and that the cyclic sulfates are likely derived from an epoxide sulfolipid. In addition to the defensive behaviors they elicit in *C. elegans*, the sulfolipids also modulate growth in microbes and development in *C. elegans*.

In chapter three, we used the same analytical techniques from chapter one combined with a growth inhibition assay, and identified a cyclopropyl fatty acid, cpfa#1, from the *P. pacificus* exometabolome that inhibits growth of *C. albicans* at an MIC of 0.515 mM. We found that the cyclopropyl group and its orientation is necessary for activity, and that the compound is toxic for both mammalian cells and nematodes at the MIC. Further,

we found that *P. pacificus* biosynthesizes this lipid *de novo* from cyclopropyl fatty acids from its *E. coli* diet. Due to the high MIC, this compound may not be suitable as an effective antibiotic, yet in combination with other secreted compounds it may play a role in active modulation of microbiome composition by the nematodes. Still, there exists potential for more discovery of new antibiotics from the *P. pacificus* exometabolome alone; based on the activity profile of our initial studies there are at least 13 other possible antibiotic molecules that remain to be discovered within this set.

A significant portion of sulfolipids have been identified and characterized in this dissertation, yet there are still many more detected peaks that produce sulfate-ions in MSMS analyses that await to be characterized, including other putative cyclic sulfates. Due to their unusually high abundance (estimated 60 mg/30 million worms) and similarity to critical metabolites from *C. elegans* (C15ISO), we anticipate that the sulfolipids play a critical role in the worm. We hypothesize that, in place of a grinder, *P. pacificus* uses these sulfolipids as a means of chemical digestive or fertilizer. Digestive chemicals that also function as kairomones have been found in marine life, such as the glutamine lipids from *Chaoborus* that induce morphological defenses (neckteeth) in *Daphnia*[i] In parallel, *Daphnia* also releases sulfolipid kairomones[28] that resemble those from *P. pacificus* and cause colony formation in phytoplankton. In an interpretation of the *Daphnia* kairomones by Van Donk[ii], the defensive behavior of colony formation in phytoplankton may prevent grazing from certain predators but may offer a better 'bite-size' for larger predators.

Perhaps some of the sulfolipids produced by *P. pacificus* function in a similar way to improve the quality of their prey.

Of the many applications of this research, the most relevant seem to be related to agriculture and human health. Through preliminary studies, the identified sulfolipids deter *C. elegans*, yet they do not appear to deter agricultural pest nematodes such as root-knot nematode (unpublished). The knowledge we gained from how *C. elegans* perceive these molecules, and the methodology we used to discover the sulfolipids, however, can be used to find effective pest deterrents. A targeted study using metabolomic fractions from a known predator of agricultural pests may be a more effective route to study this. It's yet to be determined if these sulfolipids are capable of treating parasitic nematode infections in animals.

The observation that these surfactants modulate microbial and nematode growth raises concern at how these surfactants affect humans long term. Every day many of us are exposed to analogues of these sulfolipids, whether it be through using detergents containing SDS or SDS analogues, through the soap we use on our bodies, or even in the food we eat. More studies should be conducted on the long-term effects of exposure to sulfated surfactants.

Through this dissertation, the author conveyed that small molecules generated from nematodes can affect their environment, and that we can use the information from these interactions to solve real world problems. The author hopes that reading through

the methodology outlined in this dissertation, one can conduct similar approaches to better understand interactions between organisms and their environments.

REFERENCES

i. Weiss L.C. et al; Identification of *Chaoborus* kairomone chemicals that induce defences in *Daphnia*. *Nat. Chem Bio* (2018).
ii. Van Donk E.; Induced defences in marine and freshwater phytoplankton: a review *Hydrobiologia* (2011).