Electronic Supplementary Material (ESI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2018

Oxidative Transformations involved during the Biosynthesis of Squalestatin S1

Karen E. Lebe and Russell J. Cox

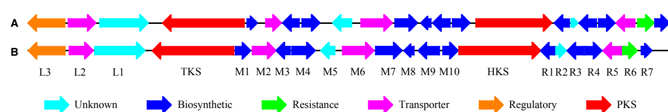
Institute for Organic Chemistry and BMWZ
Leibniz Universität Hannover
Schneiderberg 38
30167 Hannover, Germany

Electronic Supplementary Information

1.	Details of Cluster Annotation		2
2.	Phyre 2 Results Mfr1		3
3.	*In silico* Results Mfm1		3
4.	KO experiments		4
	4.1	Cloning Strategy	5
	4.2	LCMS results from KO strains	9
5.	Expression Experiments		10
	5.1	Cloning Strategy	10
	5.2	LCMS results of Expression Experiments	11
	5.3	PCR analysis of Expression Strains	1
6.	Fermentation in low Cu media		12
7.	Purification and Characterisation of Compounds		13
	7.1	Analytical and Preparative LCMS	13
	7.2	Methylation procedures	13
	7.3	Characterisation of Compounds	14
	7.4	Unknown Compound from Expression of Benzoyl CoA Cassette	45
8.	General Experimental procedures		49
	8.1	General Fermentation Procedures	49
	8.2	Fermentaion of *Phoma*	50
	8.3	Fermentation of *A. oryzae*	50
	8.4	KO procedure in *Phoma*	51
	8.5	Expression procedure in *A. oryzae*	51
	8.6	Extraction Procedure	52
	8.7	NMR/HRMS instrumentation	52
9.	References		53

1. Details of Cluster Annotation

The 69.7Kb gene cluster is deposited in Genbank with accession number KU946987.1



MF5453 Gene	C2932 Gene	Putative function	aa % id. / sim.
mfl4	*cl4*	NADP dependent dehydrogenase	87 / 93
mfl3	*cl3*	Transcription factor	78 / 85
mfl2	*cl2*	MFS transporter	89 / 91
mfl1	*cl1*	Unknown	88 / 91
mfpks1	*phpks1*	MF-SQTKS (tetraketide synthase)	87 / 93
mfm1	*cm1*	Copper dependant oxygenase	93 / 96
mfm2	*cm2*	MFS transporter	91 / 94
mfm3	*cm3*	NADP dependent dehydrogenase	80 / 85
mfm4	*cm4*	Acyltransferase (AT1)	89 / 94
mfm5	*cm5*	Unknown	82 / 86
mfm6	*cm6*	MFS transporter	81 / 85
mfm7	*cm7*	Phenylalanine ammonia-lyase	88 / 93
mfm8	*cm8*	Esterase/lipase	90 / 94
mfm9	*cm9*	Acyl-CoA synthetase/ligase	78 / 82
mfm10	*cm10*	Hydrolase	78 / 83
mfpks2	*cpks2*	MF-SQHKS (hexaketide synthase)	93 / 95
mfr1	*cr1*	Possible non-heme iron dependent	93 / 95
mfr2	*cr2*	Possible non-heme iron dependent	92 / 97
mfr3	*cr3*	Citrate synthase	82 / 85
mfr4	*cr4*	Acyltransferase-3 superfamily (AT2)	76 / 86
mfr5	*cr5*	MFS transporter	86 / 91
mfr6	*cr6*	Squalene synthase	74 / 80
mfr7	*cr7*	$\alpha\beta$ Hydrolase	91 / 95
mfr8	*cr8*	Peptidase	90 / 93
mfr9	*cr9*	NTPase	79 / 88

Table S1.1 Comparison and annotation of the MF5453 and C2923 SQS1 cluster genes. Boundaries of BGC estimated by qRT-PCR.[4]

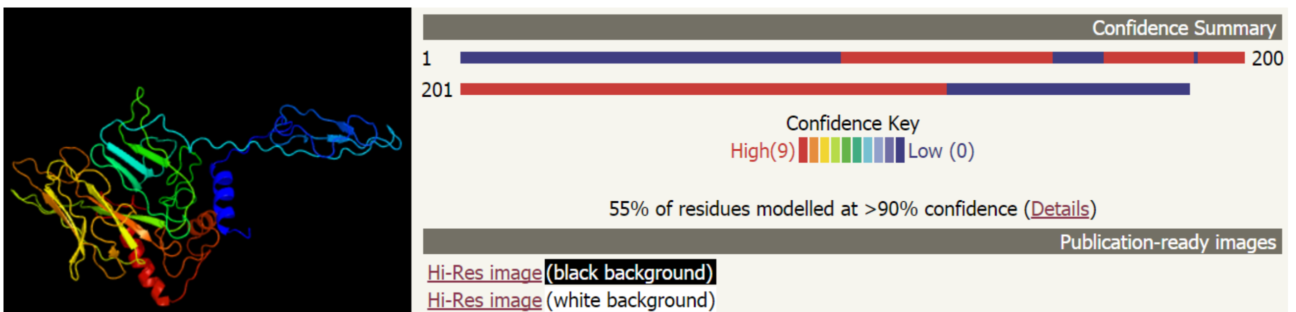
2. *In silico* results Mfr1

Mfr1 shows weak homology to the non-heme iron halogenase CytC3.

#	Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	c3gjbA_	Alignment		58.3	19	**PDB header:**biosynthetic protein **Chain:** A: **PDB Molecule:**cytc3; **PDBTitle:** cytc3 with fe(ii) and alpha-ketoglutarate Phyre2 Run Investigator

3. *In silico* results Mfm1

The central region of Mfm1 shows high likely structural homology to an oxidoreductase (1opm,) and has potential copper binding sites.



#	Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	c1opmA_	Alignment		90.9	15	**PDB header:**oxidoreductase **Chain:** A: **PDB Molecule:**protein (peptidylglycine alpha-hydroxylating **PDBTitle:** oxidized (cu2+) peptidylglycine alpha-hydroxylating monooxygenase2 (phm) with bound substrate ✓ View investigator results

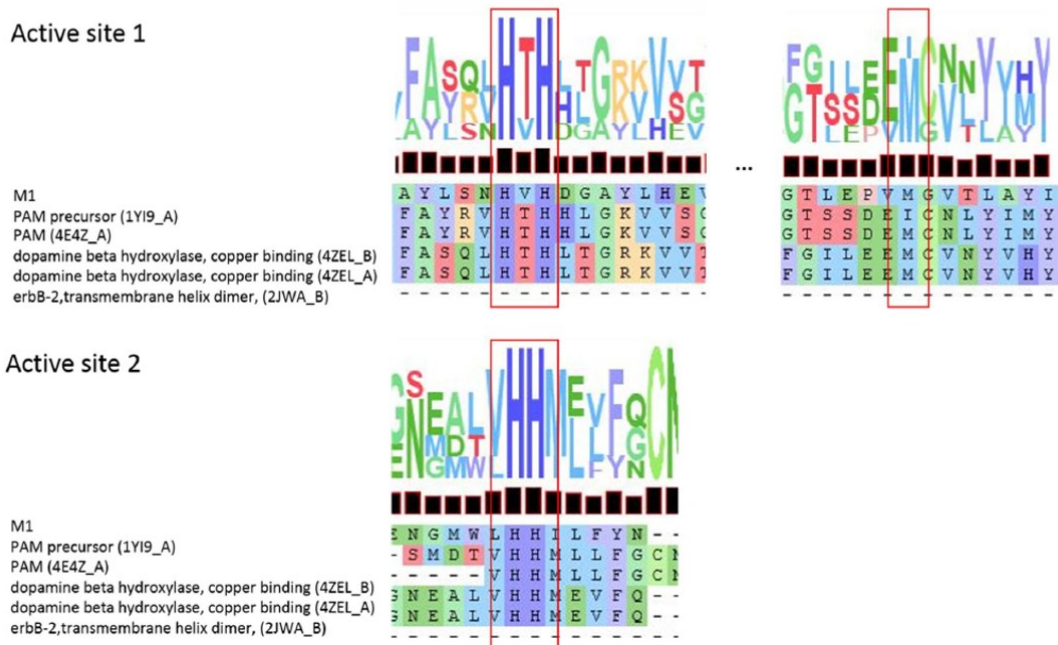


Figure S1: Multiple sequence alignment using HHpred web server[1] of Mfm1 with PDB database showing conserved amino acids in the active sites of the protein.

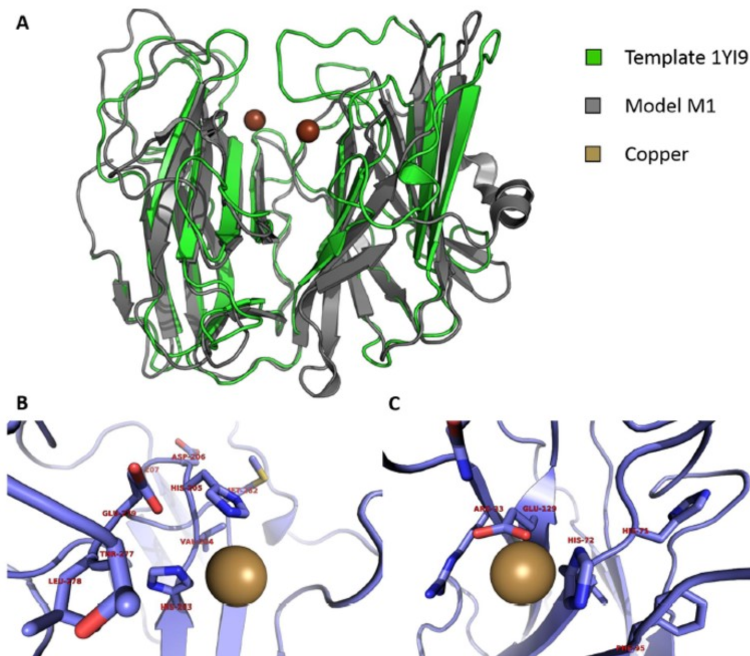


Figure S2: Active site model of the putative copper dependent monooxygenase Mfm1 using the peptidyl-glycine α-hydroxylating monooxygenase (PHM) from Rattus norvegicus (PDB.1YI9) as template (97.1 % probability and an E-value of 0.0063), which is known to require Cu(II) as a cofactor.[2] MODELLER program (version mod9.15) 114–117 with a loop refinement script was used. Active site 1 (B): Two histidines (His 203, His 205) and methionine (Met 282) in proximity to the copper are conserved to the CuM site of the PAM. Active site 2 (C) showed two histidines (His 71, His 72), an arginine (Arg 33) and a glutaminc acid (Glu 129) as ligands. The two histidines were shown to be conserved in the MSA.

4. Knockout results

4.1 Cloning strategy

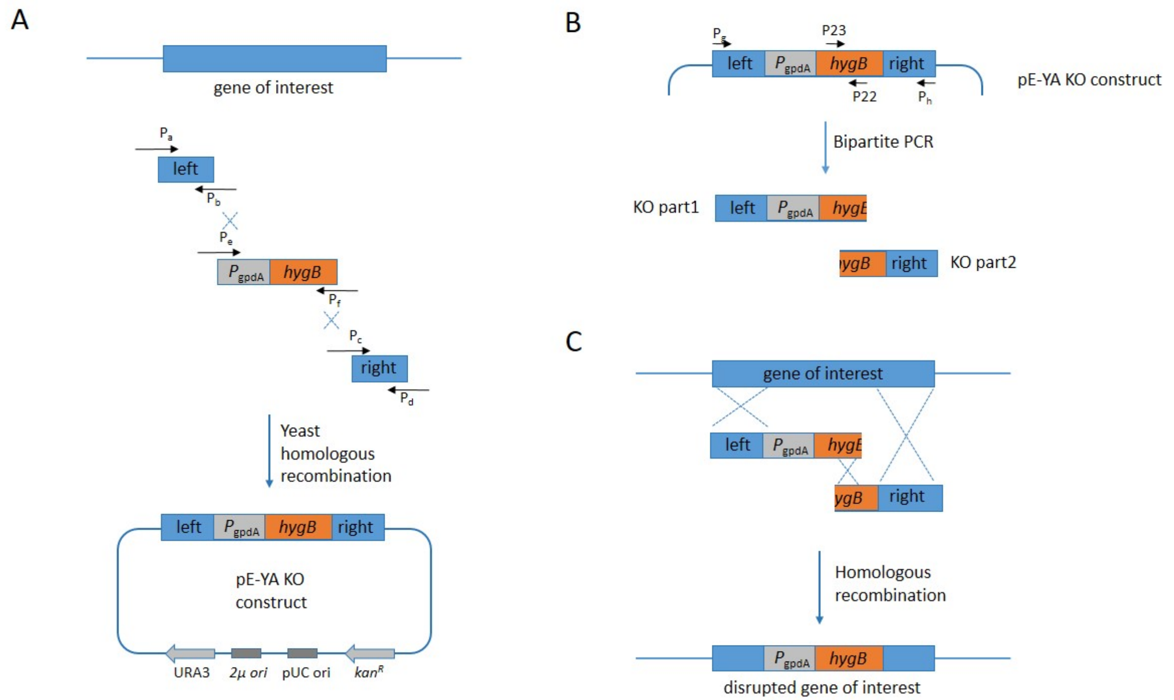


Figure S3: General KO method (A) Construction of KO plasmid using oligonucleotides to amplify flanking regions of the gene of interest and using yeast homologous recombination, (B) Bipartite PCR using forward primer of left flanking region in combination with P22 to give PCR fragment KO part1 and reverse primer of right flanking region in combination with P23 to give PCR fragment KO part2, (C) KO part1 & 2 were used to transform MF5453, three recombination events must take place for a successful KO.

Table S1A: Combination of oligonucleotides used for amplification of KO fragments

Construct ID	Target	Oligonucleotides for amplification of KO fragments	Oligonucleotides for construction in *S. cerevisiae*
KELI30	*mfr4*	P_g: P578 + P22 P23 + P_h: P579	P_{a+b}: P597+314 P_{c+d}: P598+315 P_{e+f}: P313+316
KELI27	*mfm1*	P_g: P584 + P22 P23 + P_h: P585	P_{a+b}: P589+567 P_{c+d}: P570+590 P_{e+f}: P568+569
KELI28	*mfr1+2*	P_g: P586 + P22 P23 + P_h: P587	P_{a+b}: P591+573 P_{c+d}: P576+592 P_{e+f}: P574+575
KELI38	*mfr1*	P_g: P767 + P22 P23 + P_h: P768	P_{a+b}: P591+573 P_{c+d}: P747+748 P_{e+f}: P574+746
KELI39	*mfr2*	P_g: P769 + P22 P23 + P_h: P770	P_{a+b}: P765+766 P_{c+d}: P753+754 P_{e+f}: P571+752
KELI29	*mfr3*	P_g: P580 + P22 P23 + P_h: P581	P_{a+b}: P593+461 P_{c+d}: P594+463 P_{e+f}: P462+464
KELI33	*mfm3*	P_g: P582 + P22 P23 + P_h: P583	P_{a+b}: P306+307 P_{c+d}: P605+606 P_{e+f}: P308+310

Table S1B: Oligonucleotide Sequences

Cox ID number	5'-3' sequence
22	CGTCAGGACATTGTTGGAG
23	CTGTCGAGAAGTTTCTGATCG
57	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGGACGTTTCCAAGGAAGA
60	TCCTCGCCATGCCTGTTATC
61	CCGCGGAATGGCTTGATAAC
62	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCTAAGAACCAGAGTTCTTCA
64	ATGCCAAGTTGCCAACAGGT
87	CTTCTTAAATATCGTTGTAACTGTTCCTGA
88	CGAAGTATATTGGGAGACTATAGCTACTAG
89	ATTCACCACTATTATTCCCACCCTATAATA
90	GAGACGAAACAGACTTTTTCATCGCTAAAA
91	CTTTTCTTTTCTCTTTCTTTTCCCATCTTC
92	TGACCTCCTAAAACCCCAGTG
306	TTTAAGAAGGAGCCCTTCACCAAGGGTGGGCCCGGGGAAACCTTGACTCACTGCGG
307	ACTAGAGGATCCCCATCATGGGCATATCTCGCACTTTGGT

308	ACCAAAGTGCGAGATATGCCCATGATGGGGATCCTCTAGT
310	AGGTCACGGTATAGCGCAGGCAGGTCGAGTGGAGATGTGG
313	TACGACAATCTCTCTCCACCCATGATGGGGATCCTCTAGT
314	ACTAGAGGATCCCCATCATGGGTGGAGAGAGATTGTCGTA
315	CCACATCTCCACTCGACCTGGGATCGTTGCGGAAATACAC
316	GTGTATTTCCGCAACGATCCCAGGTCGAGTGGAGATGTGG
461	ACTAGAGGATCCCCATCATGTGCATCGGGGCCGCTCCATG
462	CATGGAGCGGCCCCGATGCACATGATGGGGATCCTCTAGT
463	CCACATCTCCACTCGACCTGGGGATCGAGTCATCTCTTCG
464	CGAAGAGATGACTCGATCCCCAGGTCGAGTGGAGATGTGG
567	ACTAGAGGATCCCCATCATGCGACCACAGCTTGCAACATC
568	GATGTTGCAAGCTGTGGTCGCATGATGGGGATCCTCTAGT
569	AAGATGGCATTGTTCGTGGCCAGGTCGAGTGGAGATGTGG
570	CCACATCTCCACTCGACCTGGCCACGAACAATGCCATCTT
573	ACTAGAGGATCCCCATCATGCACTACCAAGAAGCTCGGCA
574	TGCCGAGCTTCTTGGTAGTGCATGATGGGGATCCTCTAGT
575	CCTACTTTGCCGTCTGTGGTCAGGTCGAGTGGAGATGTGG
576	CCACATCTCCACTCGACCTGACCACAGACGGCAAAGTAGG
578	CTATATGGGAGAGGGATCAG
579	TGCCATCTCATGATTATCAGAC
580	TTTATATCTCACGACGTGCA
581	ATTGGCTCTTGTATGTTCAC
582	GAAACCTTGACTCACTGCGG
583	TAAAAATTTGAACTTTCGCGAG
584	TGCATTTCTCACCATCGCAC
585	TCGTCTCGAGAGCCATTTCG
586	AGAACACCAAGGAGCGAACC
587	GGTCTACGAGCATCAGAGGC
589	GCCAACTTTGTACAAAAAAGCAGGCTCCGCTGCATTTCTCACCATCGCAC
590	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTCGTCTCGAGAGCCATTTCG
591	GCCAACTTTGTACAAAAAAGCAGGCTCCGCAGAACACCAAGGAGCGAACC
592	TGCCAACTTTGTACAAGAAAGCTGGGTCGGGGTCTACGAGCATCAGAGGC
593	GCCAACTTTGTACAAAAAAGCAGGCTCCGCTTTATATCTCACGACGTGCA
594	TGCCAACTTTGTACAAGAAAGCTGGGTCGGATTGGCTCTTGTATGTTCAC
597	GCCAACTTTGTACAAAAAAGCAGGCTCCGCCTATATGGGAGAGGGATCAG
598	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTGCCATCTCATGATTATCAG
605	CCACATCTCCACTCGACCTGAGATGCACCTCCTGTAACTA
606	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTCACATCCTTAGACCTTCTTC
609	GTCGACTGACCAATTCCGCAGCTCGTCAAAATGGCTATCGTCAACGGCGC
610	GGTTGGCTGGTAGACGTCATATAATCATACCTACAACTTGAACTTCTGCA
611	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGCTCAGAATGAAGAAGAT
612	TTCATTCTATGCGTTATGAACATGTTCCCTCTAGCGAGCTAAGAATGCTT
613	TAACAGCTACCCCGCTTGAGCAGACATCACCATGGCGACAGCTATCCTTCC
614	ACGACAATGTCCATATCATCAATCATGACCCTAAAAGAACACCAAGGAGC
615	GTCGACTGACCAATTCCGCAGCTCGTCAAAATGGCGACTGCAACAACAAC
616	GGTTGGCTGGTAGACGTCATATAATCATACCTAGTCAAAGAAAACCCAGA
746	GATGTCTACTTCGCGGAGTGCAGGTCGAGTGGAGATGTGG
747	CCACATCTCCACTCGACCTGCACTCCGCGAAGTAGACATC

748	AACTTTGTACAAGAAAGCTGGGTCGGCGCGCCCGGGCAAGAGCAGCAATTGACCAT
751	AAGCCGACATGCTATTGAGCCATGATGGGGATCCTCTAGT
752	CCACACATTGATCGATCCGTCAGGTCGAGTGGAGATGTGG
753	CCACATCTCCACTCGACCTGACGGATCGATCAATGTGTGG
754	AACTTTGTACAAGAAAGCTGGGTCGGCGCGCCCGGGTCCATTGTATGTCCATGCCT
765	GCCAACTTTGTACAAAAAAGCAGGCTCCGCGACTGCAACAACAACTCTAC
766	ACTAGAGGATCCCCATCATGTTGGGGTGCAGTGACATCGT
767	CAAGAGCAGCAATTGACCAT
768	AGAACACCAAGGAGCGAACC
769	GACTGCAACAACAACTCTAC
770	TCCATTGTATGTCCATGCCT
861	CCAATGTATACTGGTCTCACCCCATCCGCGAAACAATTGGCCACTGCCTC
862	GAGGCAGTGGCCAATTGTTTCGCGGATGGGGTGAGACCAGTATACATTGG
863	TGTAGAGGTCTGCCACGGTTGCCATATGATGAGAATGATAGCCCATTGGT
864	ACCAATGGGCTATCATTCTCATCATATGGCAACCGTGGCAGACCTCTACA
886	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGGACTTCCCCGGGAATTC
887	ACGACAATGTCCATATCATCAATCATGACCTCACCAATGTGTCGAAAGGA

4.2 LCMS results from KO strains

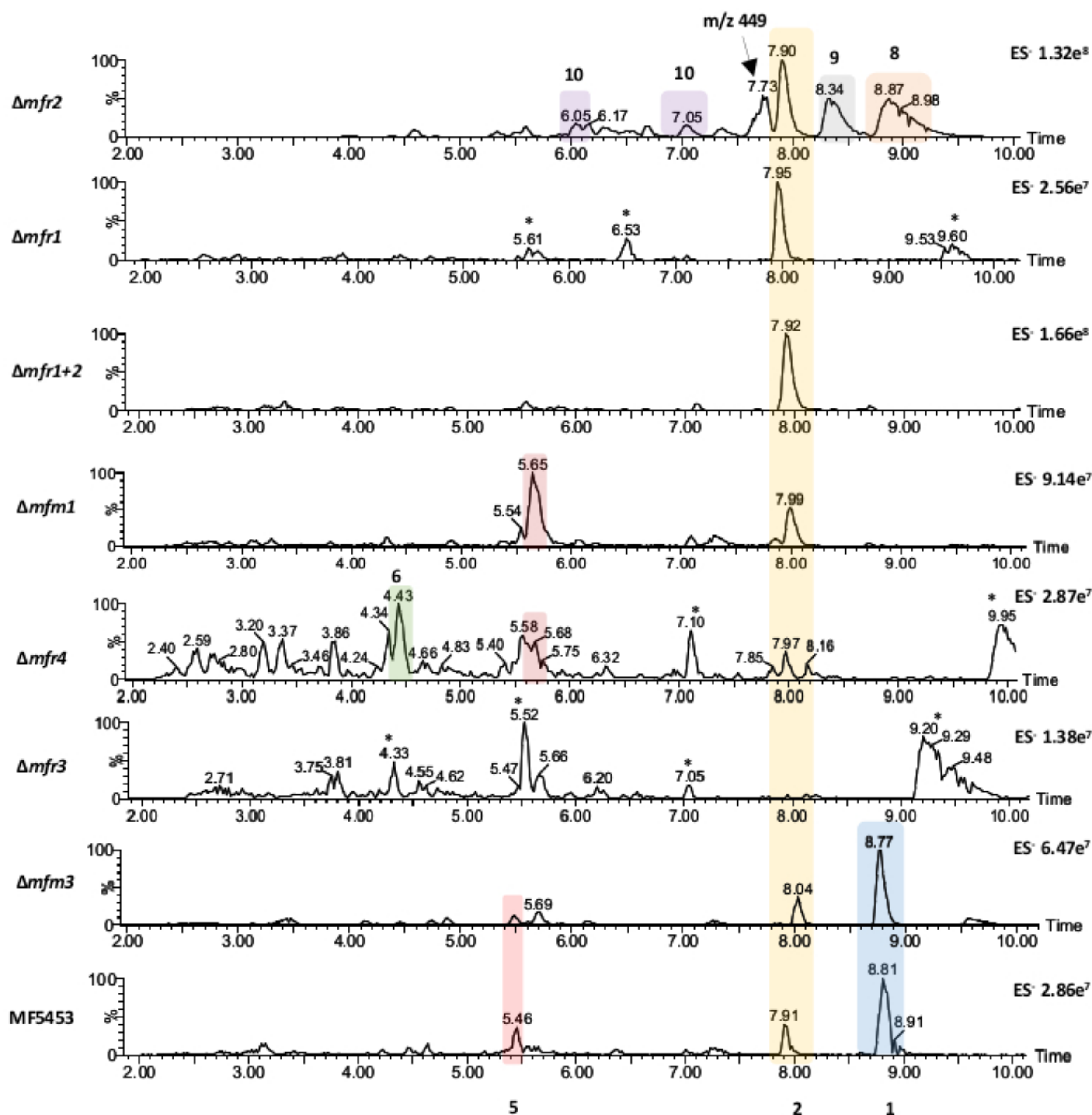


Figure S4: ES- traces of MF5453 wt and KO strains, in blue = compound **1**, in yellow = compound **2**, in red = compound **5**, further analytical data for produced compounds is described in section 7.3.

5. Expression Experiments

5.1 Cloning Strategy

Table S2: Construction details for fungal transformation plasmids used in this study

Construct ID	Vector backbone	Template	Oligonucleotides for construction in *S. cerevisiae*
KELI47	pE-YA	BB_*sqhks*_old	*sqhks* part1: P57+861; *sqhks* part2: P862+863; *sqhks* part3: P864+60; *sqhks* part4: P61+62;
KELI50	pTYGSarg	cDNA	*mfr3*: P609+610; *mfm8*: P886+P887
KELI50LR	pTYGSarg	-	LR with KELI47
KELI52A	pTYGSmet	cDNA	*mfr1*: P613+614; Patch_PadH: P89+90; Patch_Peno: P87+88
KELI52B	pTYGSmet	cDNA	*mfr2*: P615+616; Patch_PadH: P89+90; Patch_PgpdA: P91+92
KELI52C	pTYGSmet	cDNA	*mfr1*: P613+614; *mfr2*: P615+616; Patch_PadH: P89+90
KELI52D	pTYGSmet	cDNA	*mfr1*: P613+614; *mfr2*: P615+616; *mfm1*: P611+612

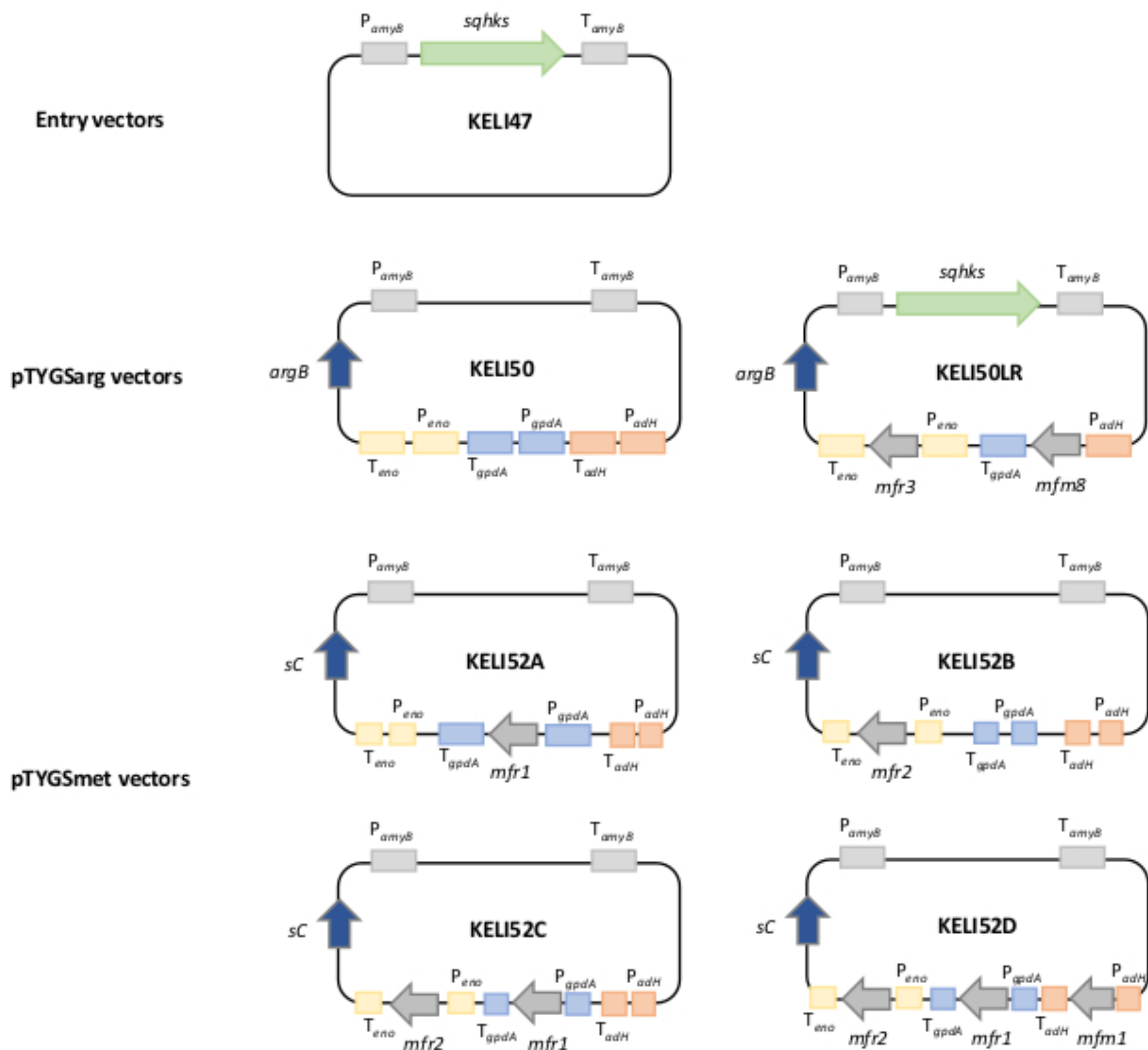


Figure S5: Map of plasmids constructed in this study.

5.2 LCMS results of Expression Experiments

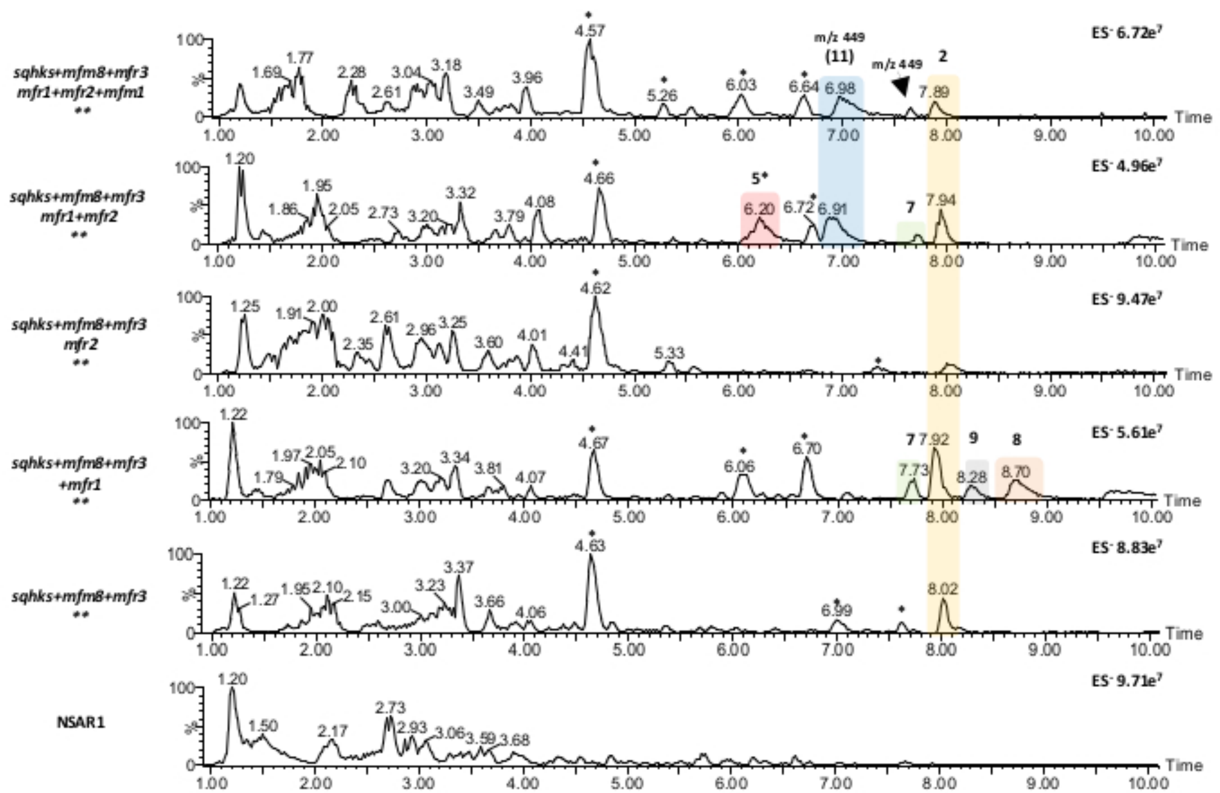


Figure S6: ES- traces of *A. oryzae* NSAR1 and constructed expression strains, in red = compound **5**, in blue = compound **11** and m/z 449, in yellow = compound **2**, in green = compound **7**, in orange = compound **8**, in purple = compound **9**, further analytical data for produced compounds is described in section 7.3 and peaks marked with "*" are unrelated compounds.

5.3 PCR analysis of Expression Strains

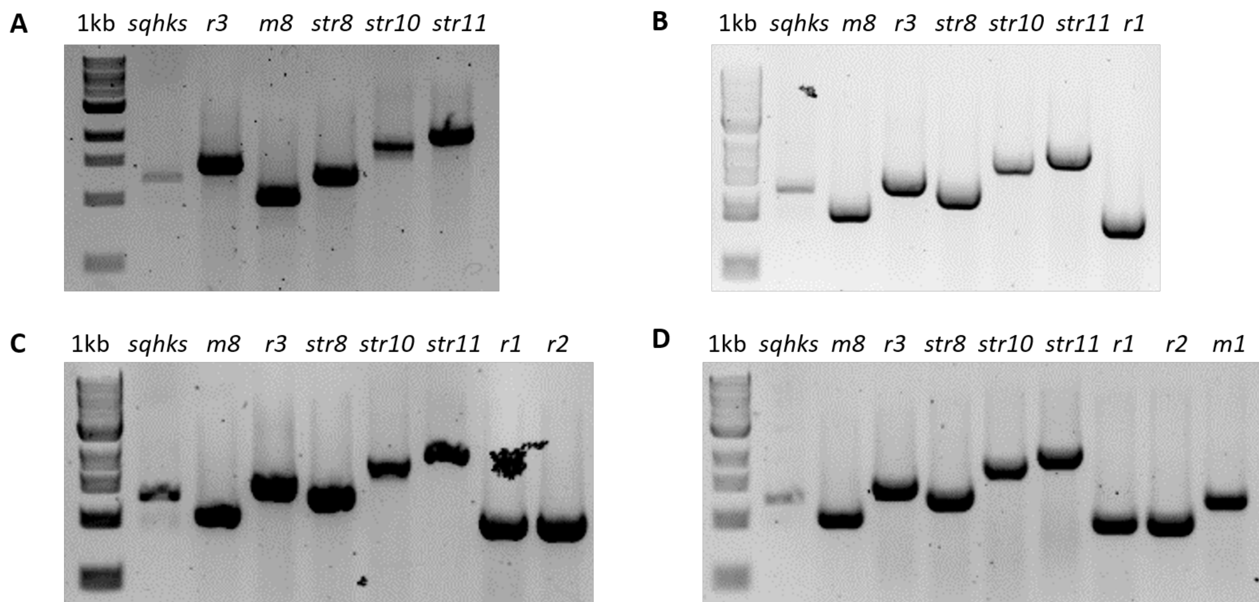


Figure S7: Genetic analysis of expression transformants KELI52 A to D using oligonucleotides described in table S2. All expressed genes were integrated into the genome of *A. oryzae* NSAR1.

6. Fermentation in low Cu media results

According to *in silico* analysis (see section 3) Mfm1 contains conserved copper binding sites. Fermentation experiments underlined the hypothesis of Mfm1 being a copper dependent monoxygenase as it was shown that the MF5453 wildtype strain produced almost exclusively SQS1 **1** by growing in YMG medium supplemented with trace element solution (containing copper amongst others, see section 8.1). In an experiment in which the wildtype strain was grown in YMG medium without trace elements, just a small amount of SQS1 was observed compared to compound **5** (Fig. S8). The *mfm1* KO strain showed almost the same chromatogram apart from production of **1**.

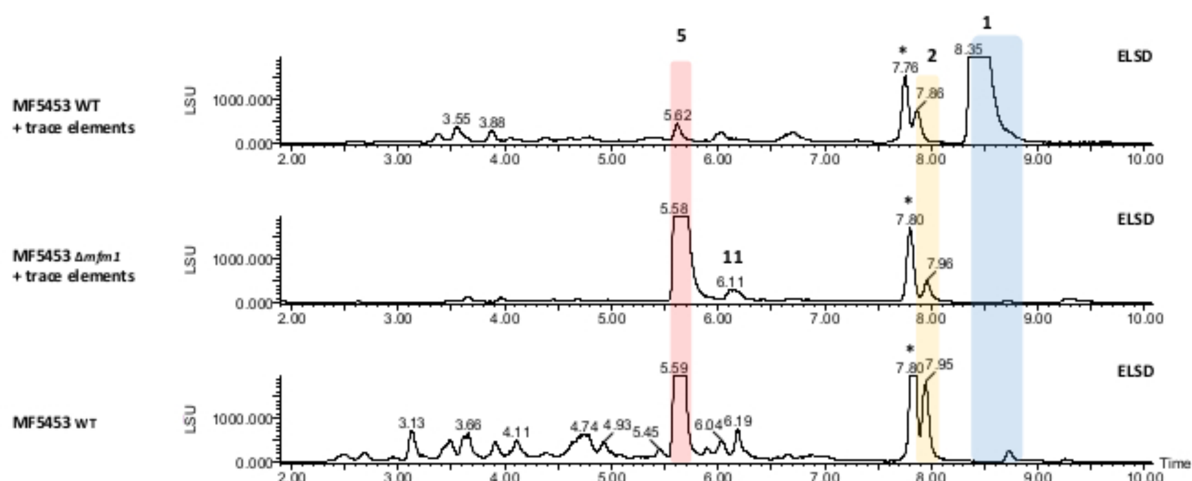


Figure S8: ELSD chromatograms of MF5453 wt strain growing in YMG medium + trace element solution (TES) containing copper or without TES and *mfm1* KO strain growing in YMG medium + TES, peaks marked with an * were shown to be unrelated, in blue = compound **1**, in yellow = compound **2**, in red = compound **5.**

7. Purification and Characterisation of Compounds

7.1 Analytical and Preparative LCMS

Analytical LCMS data were obtained with a Waters 2767 sample manager connected to waters 2545 pumps and SFO, a Phenomenex Kinetex column (2.6 μ, C_{18}, 100 Å, 4.6 × 100 mm) equipped with a Phenomenex Security Guard precolumn (Luna C_5 300 Å) eluted at 1.0 ml·min^{-1}, with a waters 2998 Diode Array detector (200-600 nm) and Waters 2424 ELSD and Waters SQD-2 mass detector operating simultaneously in ES^+ and ES^- modes between 100 *m/z* and 650 *m/z*. Solvents were: **A**, HPLC grade H_2O containing 0.05% formic acid; **B**, HPLC grade CH_3CN containing 0.045% formic acid. The Gradient was as follows: 0 min, 10% **B**; 10 min, 90% **B**; 12 min, 90% **B**; 13 min, 10% **B**; 15 min, 10% **B**. Purification of all compounds was generally achieved using a Waters mass-directed autopurification system comprising of a Waters 2767 autosampler, Waters 2545 pump system and SFO, a Phenomenex Kinetex Axia column (5μ, C_{18}, 100 Å, 21.2 × 250 mm) equipped with a Phenomenex Security Guard precolumn (Luna C_5 300 Å) eluted at 20 ml·min^{-1} at ambient temperature. Solvents as above. The post-column flow was split (100:1) and the minority flow was made up with HPLC grade MeOH + 0.045% formic acid to 1 ml·min^{-1} for simultaneous analysis by diode array (Waters 2998), evaporative light scattering (Waters 2424) and ESI mass spectrometry in positive and negative modes (Waters SQD-2). Detected peaks were collected into glass test tubes. Combined fractions were evaporated (vacuum centrifuge), weighed, and residues dissolved directly in deuterated solvent for NMR.

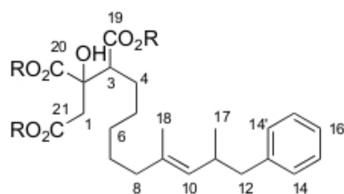
7.2 Methylation procedure

Isolated compounds (4-6 mg) were dissolved in 1 ml HPLC grade MeOH and (Trimethylsilyl)diazomethane solution (2.0 M in diethyl ether) (45 μl, 0.09 mmol) was added. The mixture was stirred overnight at room temperature and then evaporated under reduced pressure. The residual oil was purified by preparative LCMS, using a gradient from 70 % to 90 % acetonitrile.

7.3 Characterisation of Compounds

Compound 2/2A

Liquid culture of *mfr1+mfr2* KO strain (1L) was grown and extracted. Impure **2** (5.8 mg) was isolated from the resulting raw extract using preparative LCMS. The compound was treated with $TMSCHN_2$ to give **2A** (1.0 mg) which was repurified and fully characterised.



2 R=H

2A R=CH_3

HRMS of **2** (ESI-) *m/z* calc. for $C_{23}H_{31}O_7$ [M-H] 419.2070, found 419.2070.

HRMS of 2A (ESI+) *m/z* calc. for $C_{26}H_{38}O_7Na$ [M+Na]+ 485.2515, found 485.2513.

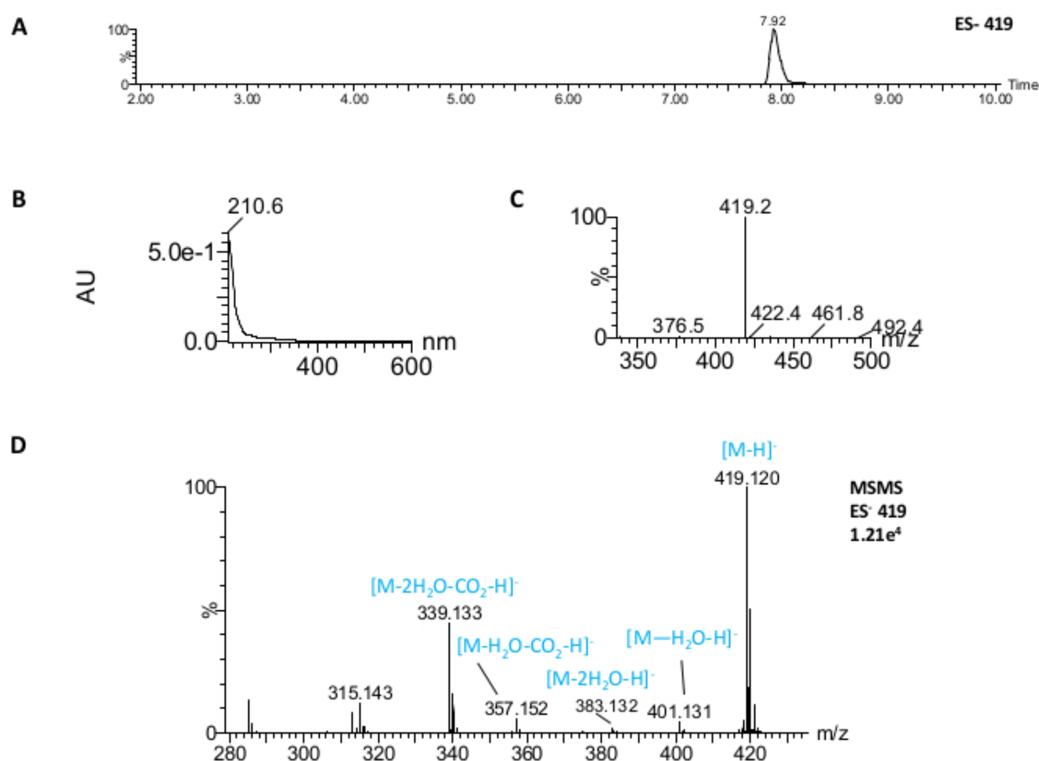


Figure S9: LCMS data for compound **2**, (A) Extracted ion chromatogram (ES- 419), (B) UV chromatogram, (C) mass spectrum (ES-), (D) MSMS fragmentation pattern of compound **2.**

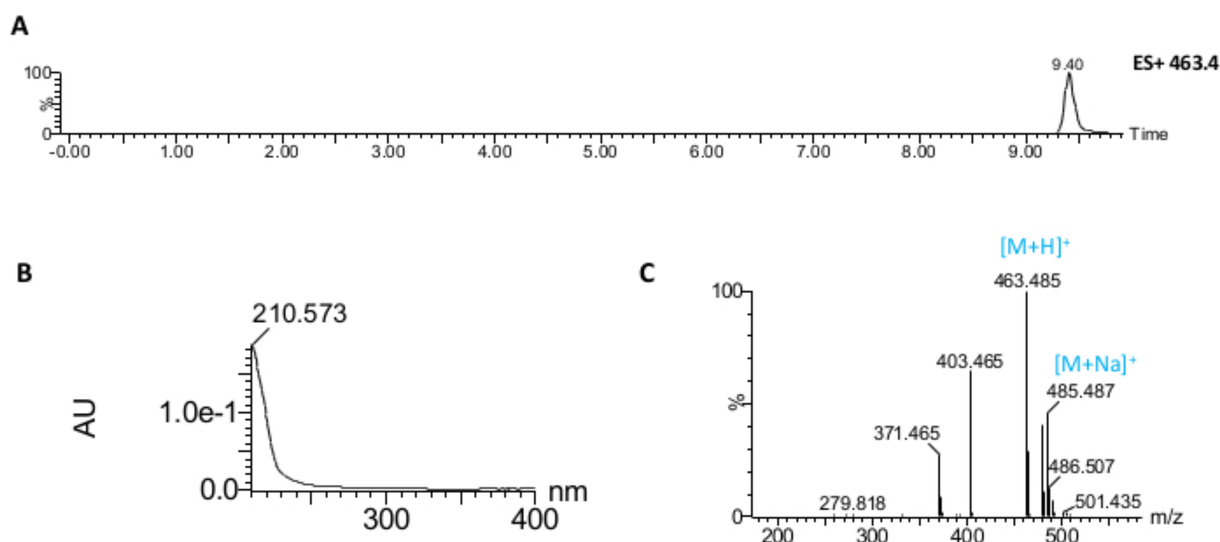


Figure S10: LCMS data for compound **2A**, (A) Extracted ion chromatogram (ES+ 463.4), (B) UV chromatogram, (C) mass spectrum (ES+).

Table S3: NMR assignment of compound **2A** in comparison to literature data of compound **2**

	Compound **2** Literature[3] CD_3CN		Compound **2A** $CDCl_3$				
Pos	**δ_C / ppm (125 MHz)**	**δ_H / ppm (500 MHz)**	**δ_C / ppm (125 MHz)**	**δ_H / ppm (500 MHz)**	***J* / Hz**	**^{1}H-^{1}H COSY**	**HMBC**
1	41.8	2.65 (d, 16.5) & 3.02 (d, 16.5)	40.9	2.72 & 3.10	d, 16.4 & d, 16.4	-OH	2, 20, 21
2	76.5		76.1				1, 3, OH
3	53.9	2.60 (m)	53.9	2.69	dd, 11.8, 3.1		2, 19, 20
4	27.9	1.40 (m) & 1.66 (m)					
5	28.1	1.22 (m)	27.7/ 27.8/ 29.8/29.9	1.18-1.34	m, 8H		
6	29.3	1.22 (m)					
7	28.2	1.22 (m)					
8	40	1.86 (brt, 7)	39.6	1.88	m	1.29 ppm, 10	27.7 ppm, 9, 10, 18
9	134.9		134.1				8,18
10	131.3	4.94 (dq, 1.2, 9.0)	130.5	4.93	dd, 1.5, 9.3	8, 11, 18	8
11	35.3	2.64 (m)	34.6	2.60	m	10, 12, 17	
12	44.5	2.46 (dd, 8.1, 12.8) & 2.57 dd 6.2, 12.3)	44.2	2.51 & 2.53	m	11	11, 13, 14, 17
13	142.3		141.3				12, 15
14	130.2		129.4	7.13	m		12, 15, 16
15	126.5		128.1	7.24	m		13, 14
16	128.9		125.7	7.13	m		14
17	21.3	0.92 (d, 6.4)	21.0	0.93	d, 6.5	11, 12	10, 11, 12
18	16	1.36 (d, 1.9)	16.0	1.39	d, 1.3	10	8, 9, 10
19	174.1		173.0				22
20	174.6		174.2				3, 23
21	172.1		171.2				1, 3, 4, 24
22	-		52.1	3.71	s		21
23	-		53.3	3.82	s		3
24	-		52.1	3.67	s		19

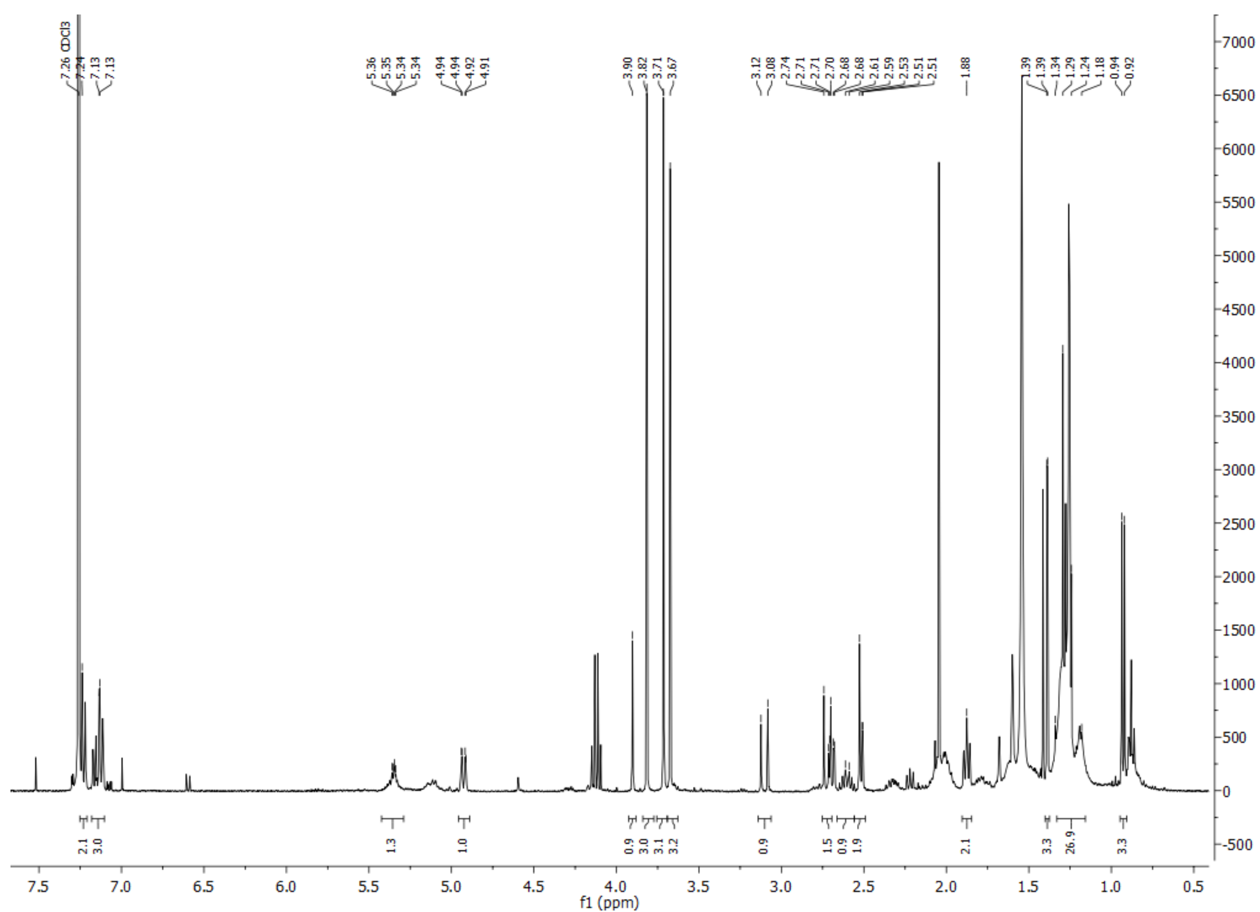


Figure S11: ^{1}H NMR of compound **2A** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. See Table S3 for assignment.

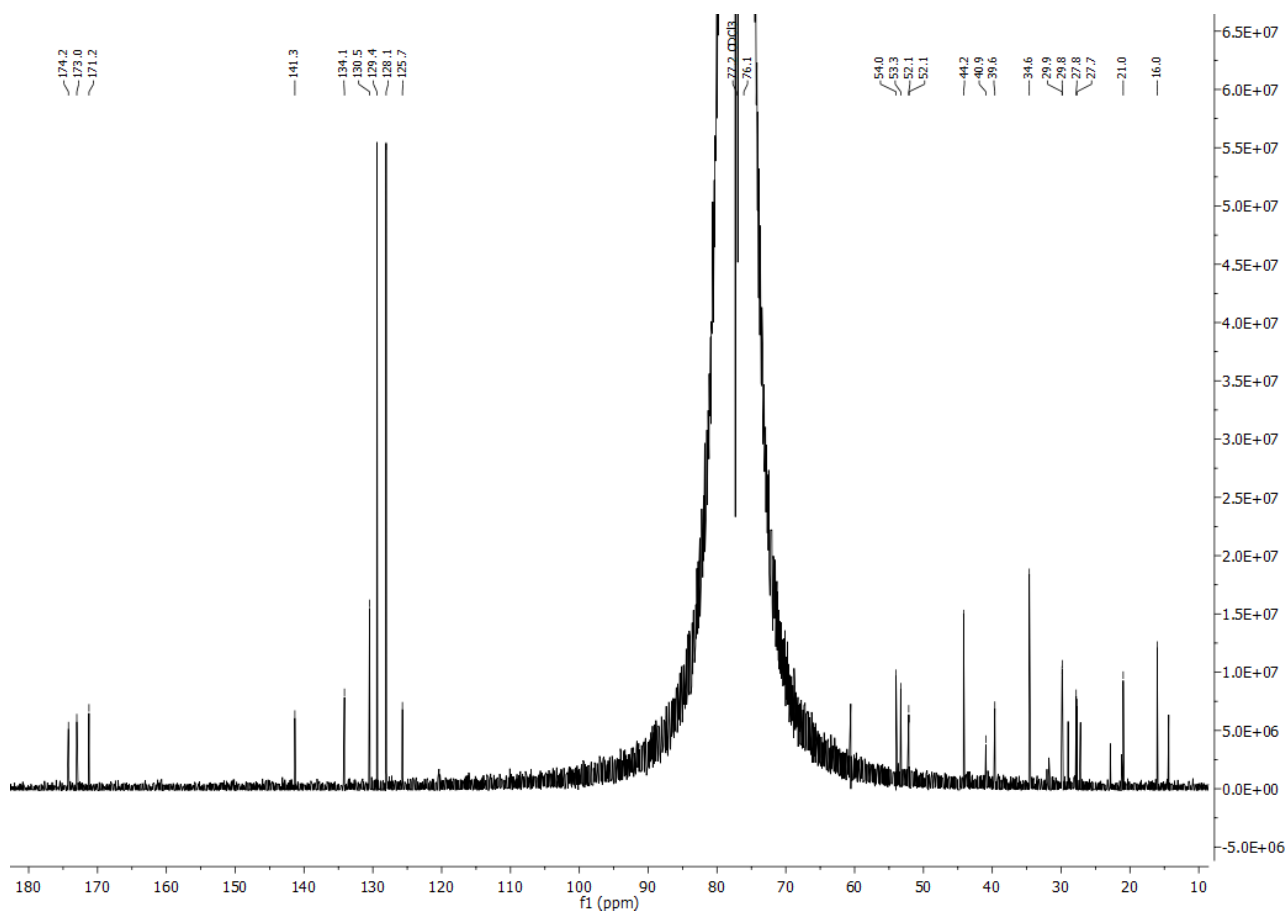


Figure S12: ^{13}C NMR of compound **2A** in $CDCl_3$ (125 MHz) referenced to $CDCl_3$. See Table S3 for assignment.

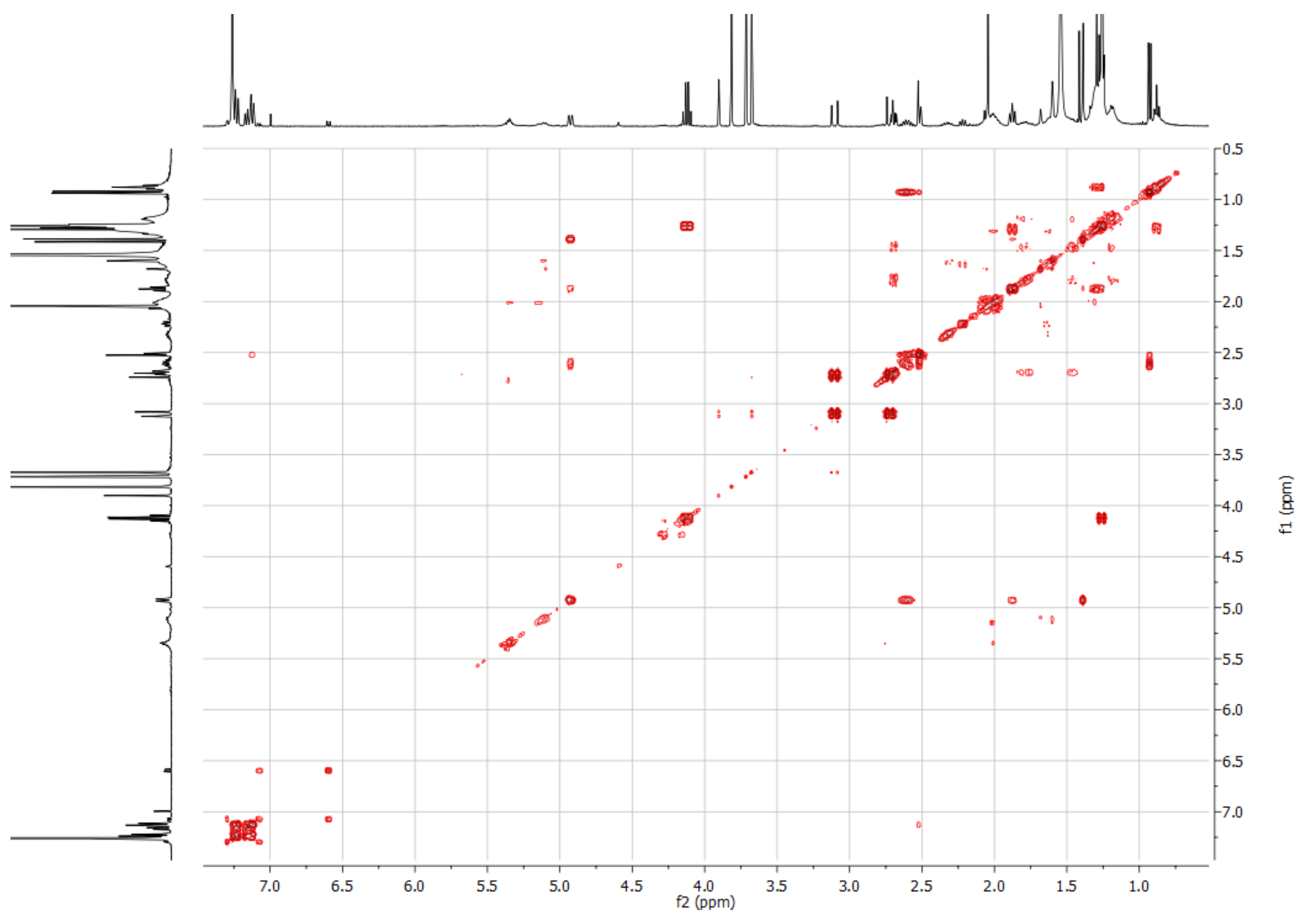


Figure S13: ¹H, ¹H COSY of compound **2A**.

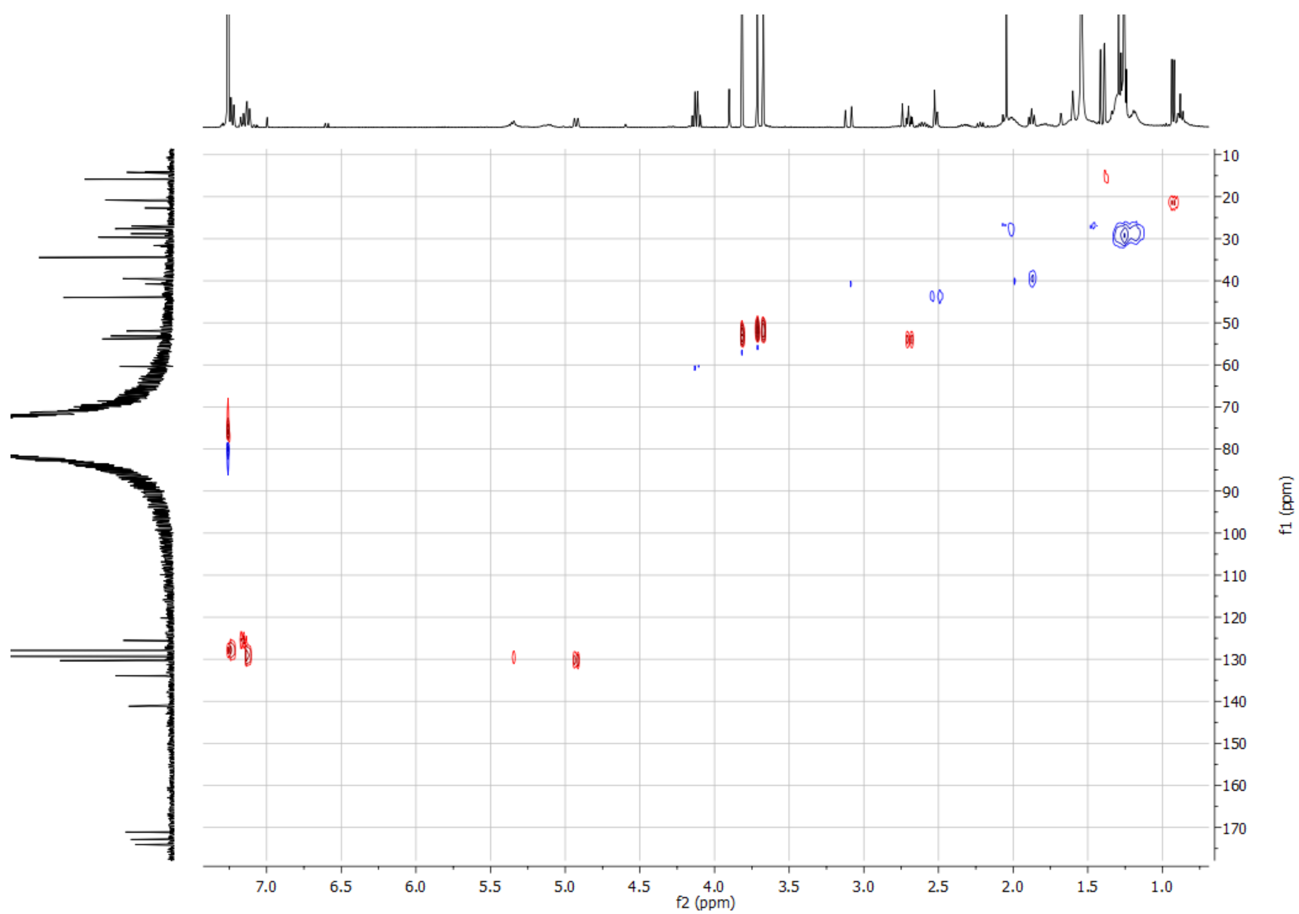


Figure S14: HSQC of compound **2A**.

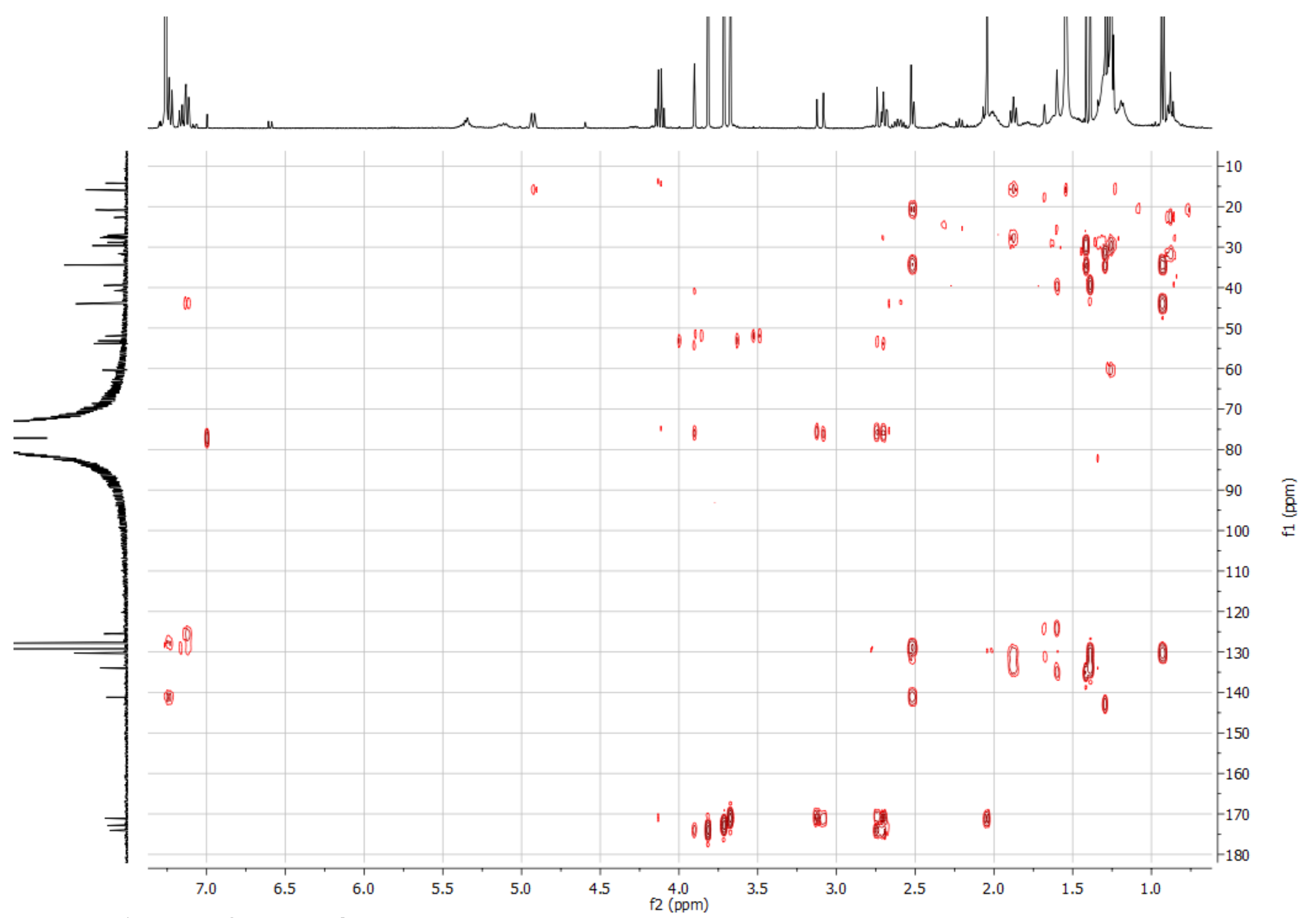
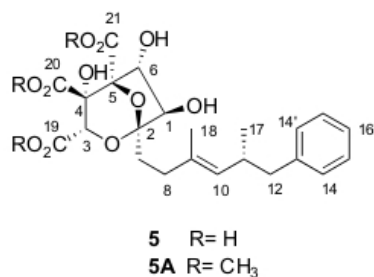


Figure S15: HMBC of compound **2A**.

Compound 5

Isolation of **5** from raw extract of *mfm1* KO strain (1L) yielded 4.8 mg product. This was treated with $TMSCHN_2$ and repurified to yield **5A**



HRMS 5 (ESI-) *m/z* calc. for $C_{23}H_{27}O_{11}$ $[M-H]^-$ 479.1553, found 479.1552.

HRMS 5A (ESI+) *m/z* calc. for $C_{26}H_{34}O_{11}Na$ $[M+Na]^+$ 522.1999, found 522.1998.

Figure S16: LCMS data for compound **5**, (A) Extracted ion chromatogram (ES- 479), (B) UV chromatogram, (C) mass spectrum (ES-), (D) MSMS fragmentation pattern of compound **5.**

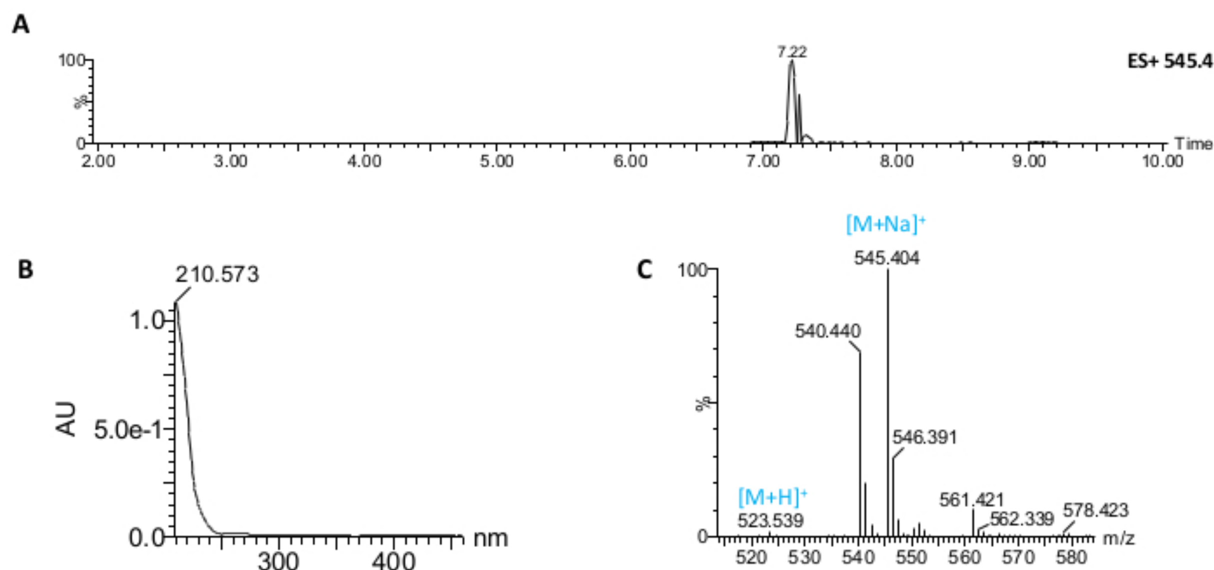


Figure S17: LCMS data for compound **5A**, (A) Extracted ion chromatogram (ES+ 545.4), (B) UV chromatogram, (C) mass spectrum (ES+).

Table S4: NMR assignment of compound **5**

Compound 5 (CD_3OD)					
Pos.	**δ_C / ppm (125 MHz)**	**δ_H / ppm (500 MHz)**	***J* / Hz**	**^{1}H-^{1}H COSY**	**HMBC**
1	82.4	4.07	d, 2.2	6	6, 7
2	105.3	-	-	-	7
3	75.3	5.14	s	-	19
4	74.6	-	-	-	6
5	91.8	-	-	-	-
6	78.0	5.17	d, 2.2	1	1
7	34.2	1.94	m	8	1, 2, 8
8	32.2	2.17 - 2.26 & 2.27 - 2.35	m	7	7, 9, 10, 18
9	133.9	-	-	-	18
10	129.8	5.06	dd, 9.3, 1.3	11, 18	8, 11, 17, 18
11	34.5	2.65	ddt, 9.2, 7.9, 6.4	10, 12, 17	9, 10, 12, 17
12	43.7	2.53	m	11	10, 11, 13, 14, 17
13	141.0	-	-	-	12, 15
14	128.9	7.13	d, 7.2	15,	12, 15, 16
15	127.6	7.22	dd, 8.1, 6.9	14	13, 14
16	125.2	7.13	d, 7.2		14,
17	19.9	0.97	d, 6.6	11	10, 11, 12
18	14.9	1.43	d, 1.3	10	8, 9, 10
19	169.7	-	-	-	3
20	172.1	-	-	-	-
21	168.7	-	-	-	-

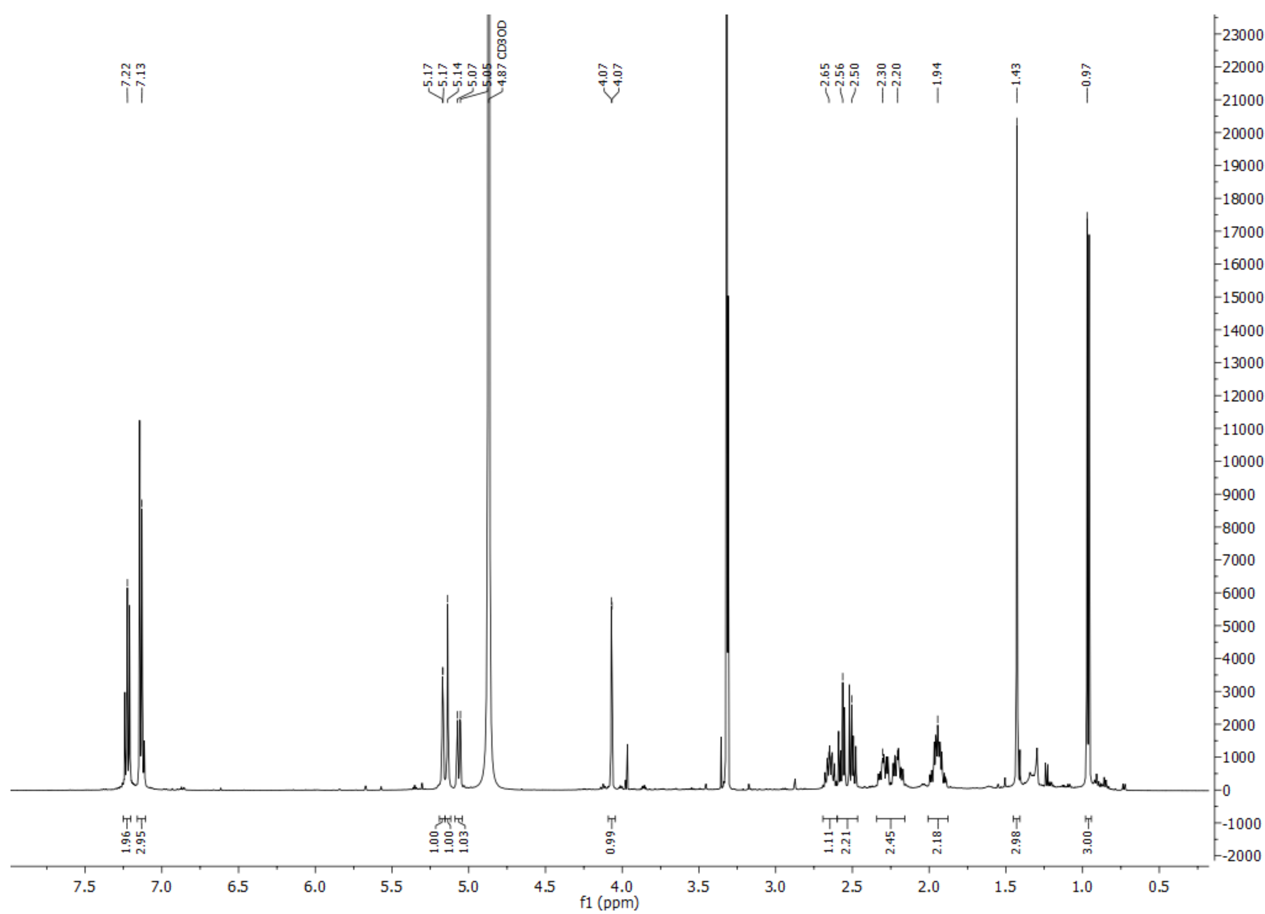


Figure S18: ^{1}H NMR of compound **5** in CD_3OD (500 MHz) referenced to CD_3OD. See Table S4 for assignment.

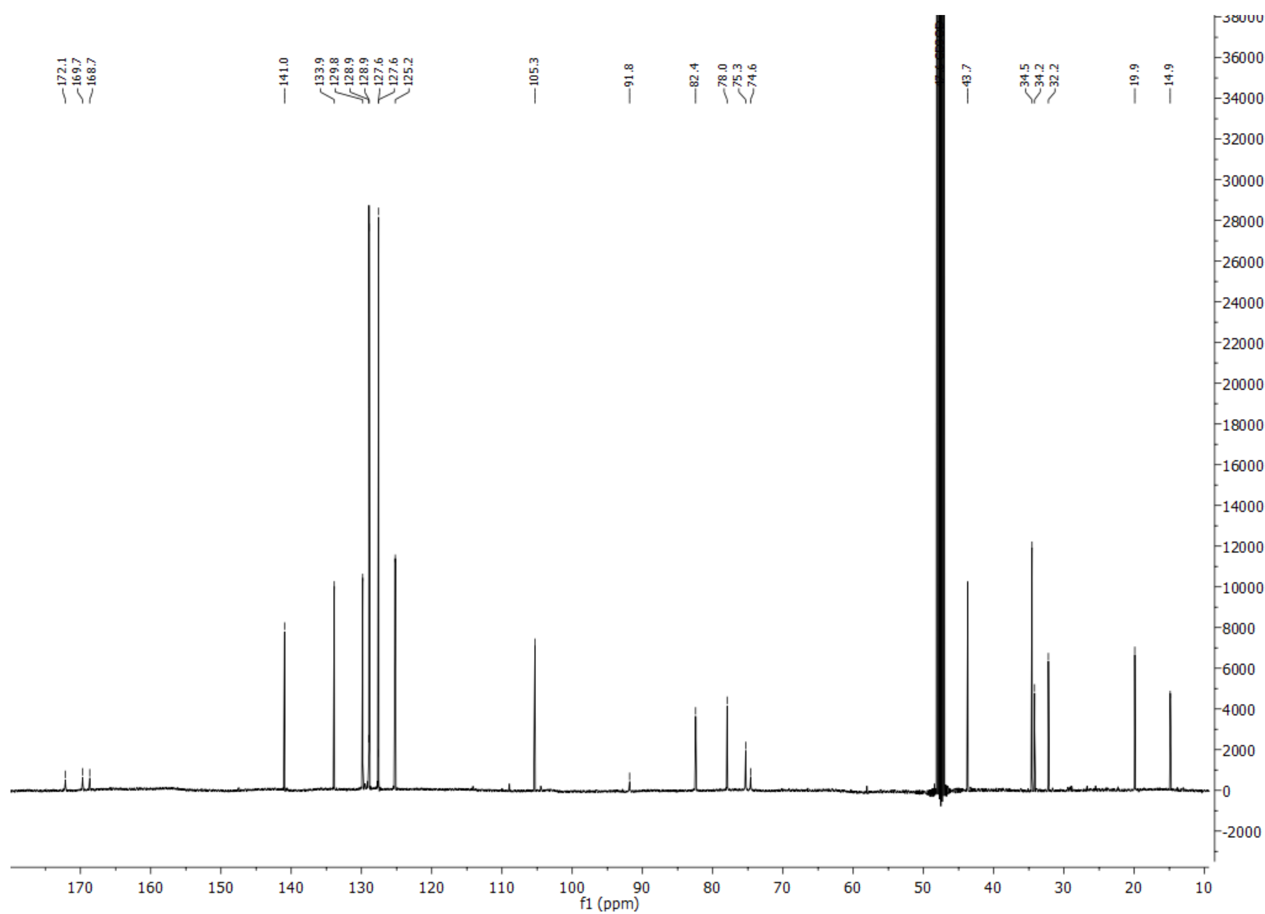


Figure S19: ^{13}C NMR of compound **5** in CD_3OD (500 MHz) referenced to CD_3OD. See Table S4 for assignment.

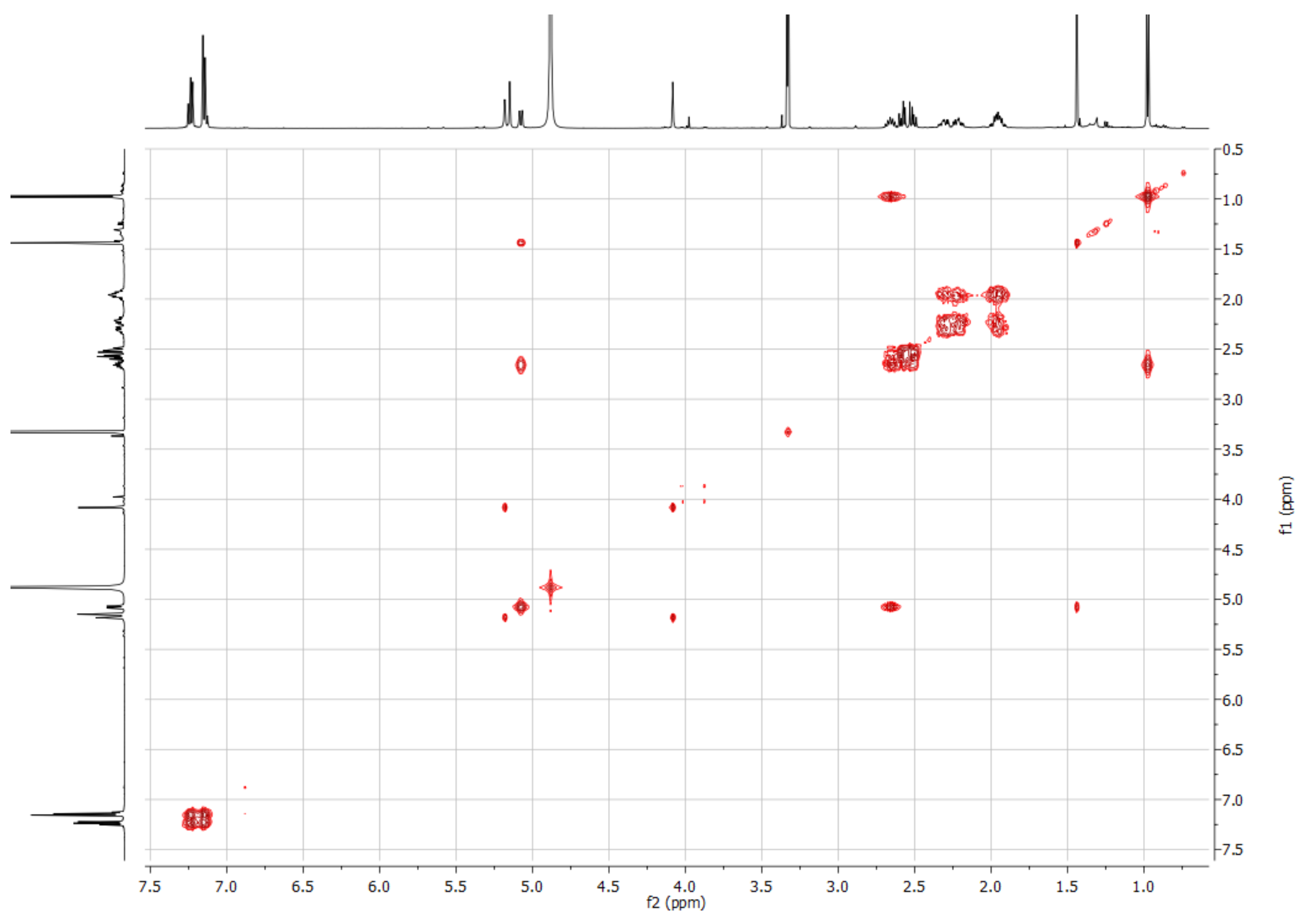


Figure S20: ^{1}H, ^{1}H COSY of compound **5**.

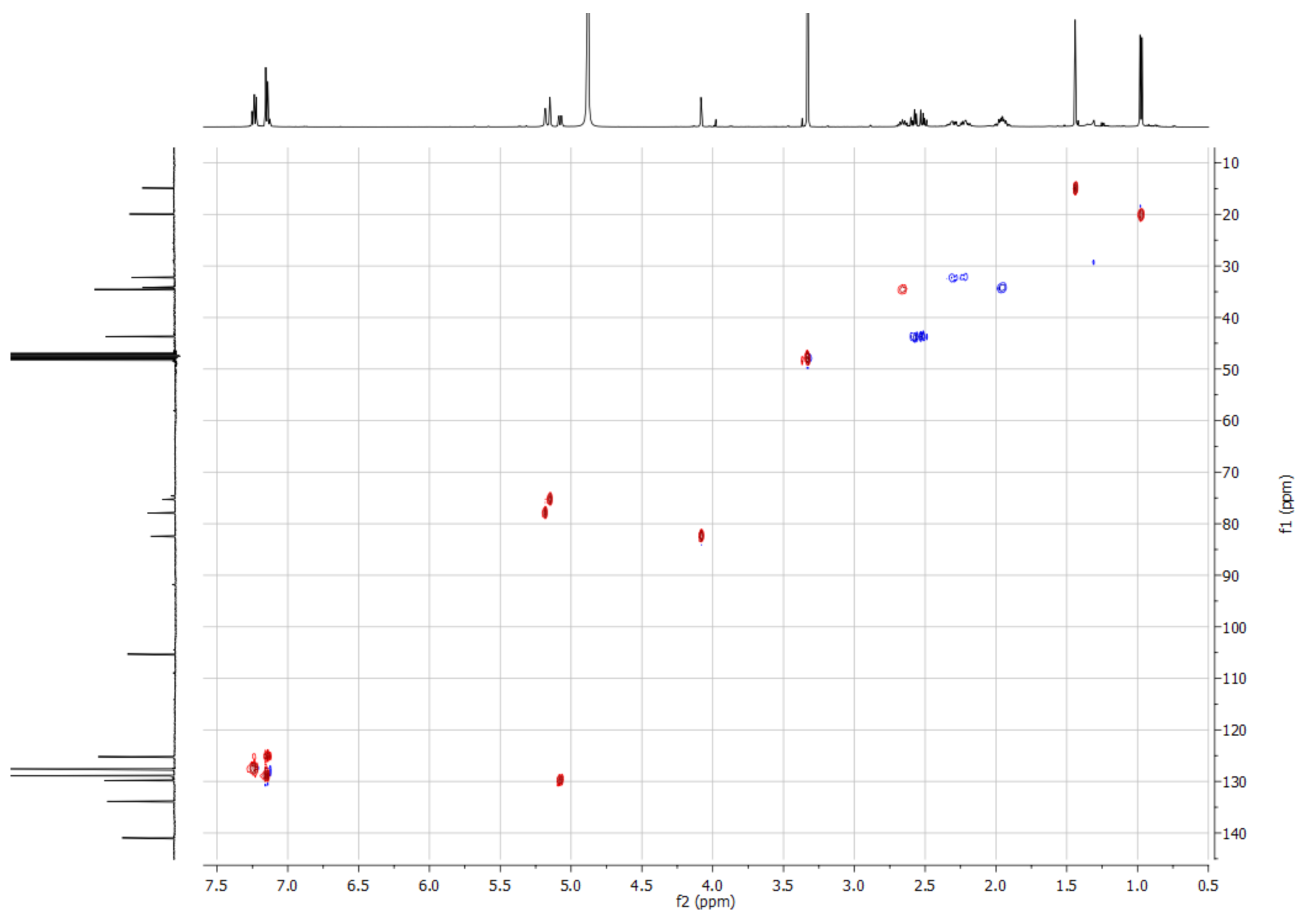


Figure S21: HSQC of compound **5**.

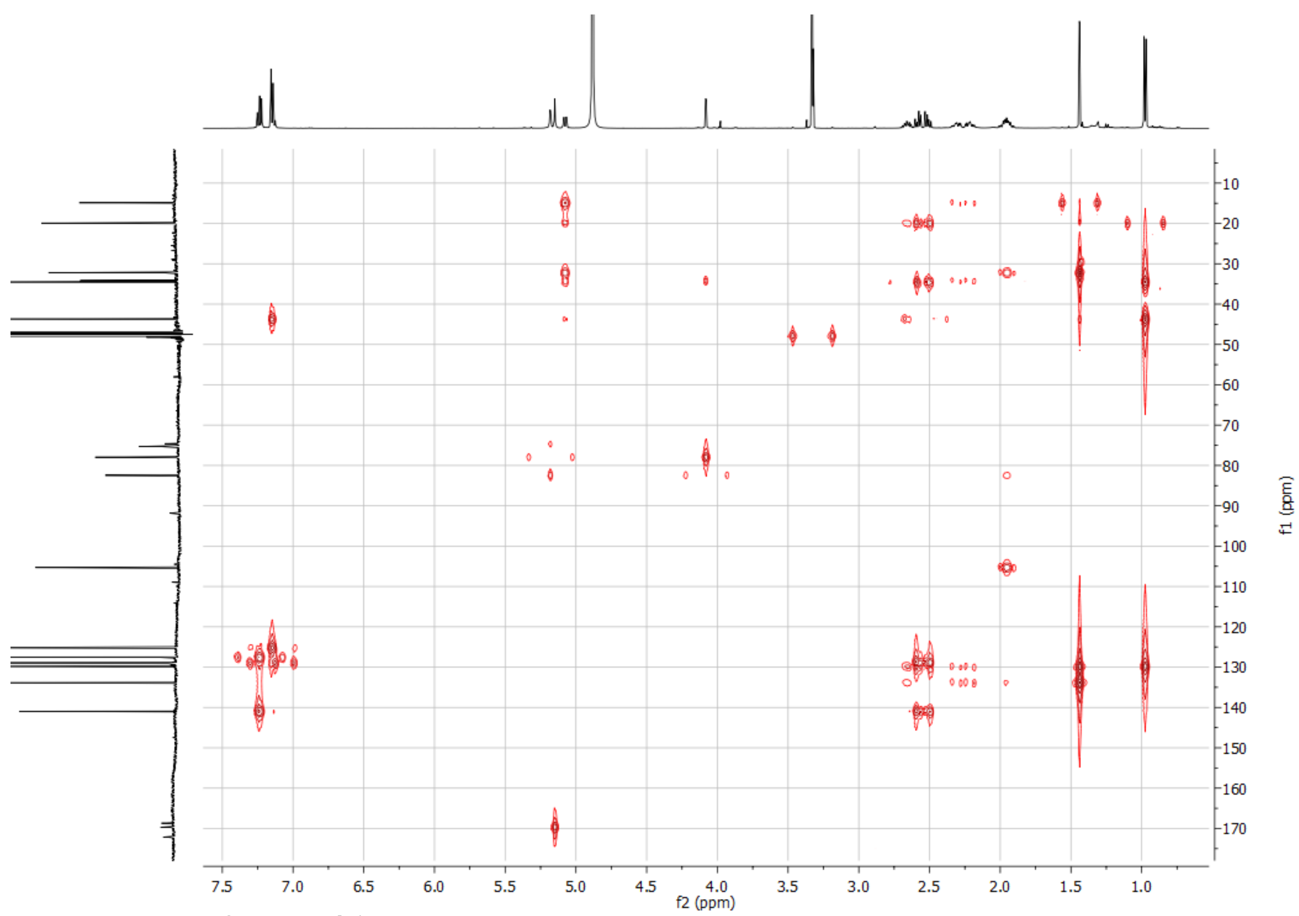


Figure S22: HMBC of compound **5**.

Table S4: NMR assignment of compound **5A**

Compound 5A ($CDCl_3$)					
Pos.	**δ_C / ppm (125 MHz)**	**δ_H / ppm (500 MHz)**	***J* / Hz**	**^1H-^1H COSY**	**HMBC**
1	82.5	4.15	d, 2.2	6	6
2	106.3	-	-	-	7, 8
3	75.6	5.16	s	22	19
4	74.8	-	-	-	6, OH
5	91.7	-	-	-	OH
6	78.6	5.14	d, 2.1	1	1
7	33.9	2.06	m	8	1, 6, 8
8	32.7	2.28	m	7	7, 9, 10, 11 18
9	133.6	-	-	-	8, 18
10	131.1	5.07	dq, 9.2, 1.3	11, 18	8, 12, 18
11	34.6	2.64	m	10, 12, 17	9, 10, 12, 17
12	44.0	2.53	d, 7.2	11	11, 13, 14, 17
13	141.1	-	-	-	12
14	129.4	7.13	d, 7.2	12, 15	12, 13, 16
15	128.2	7.22	dd, 8.1, 6.9	14	14, 16
16	125.8	7.13	d, 7.2	-	15
17	20.8	0.93	d, 6.6	11	10, 11, 12
18	16.3	1.49	d, 1.3	10	8, 9, 10
19	167.2	-	-	-	3, 22
20	169.8	-	-	-	23
21	167.0	-	-	-	24
22	52.8	3.76	s	-	19
23	53.8	3.91	s	-	20
24	53.2	3.81	s	-	21

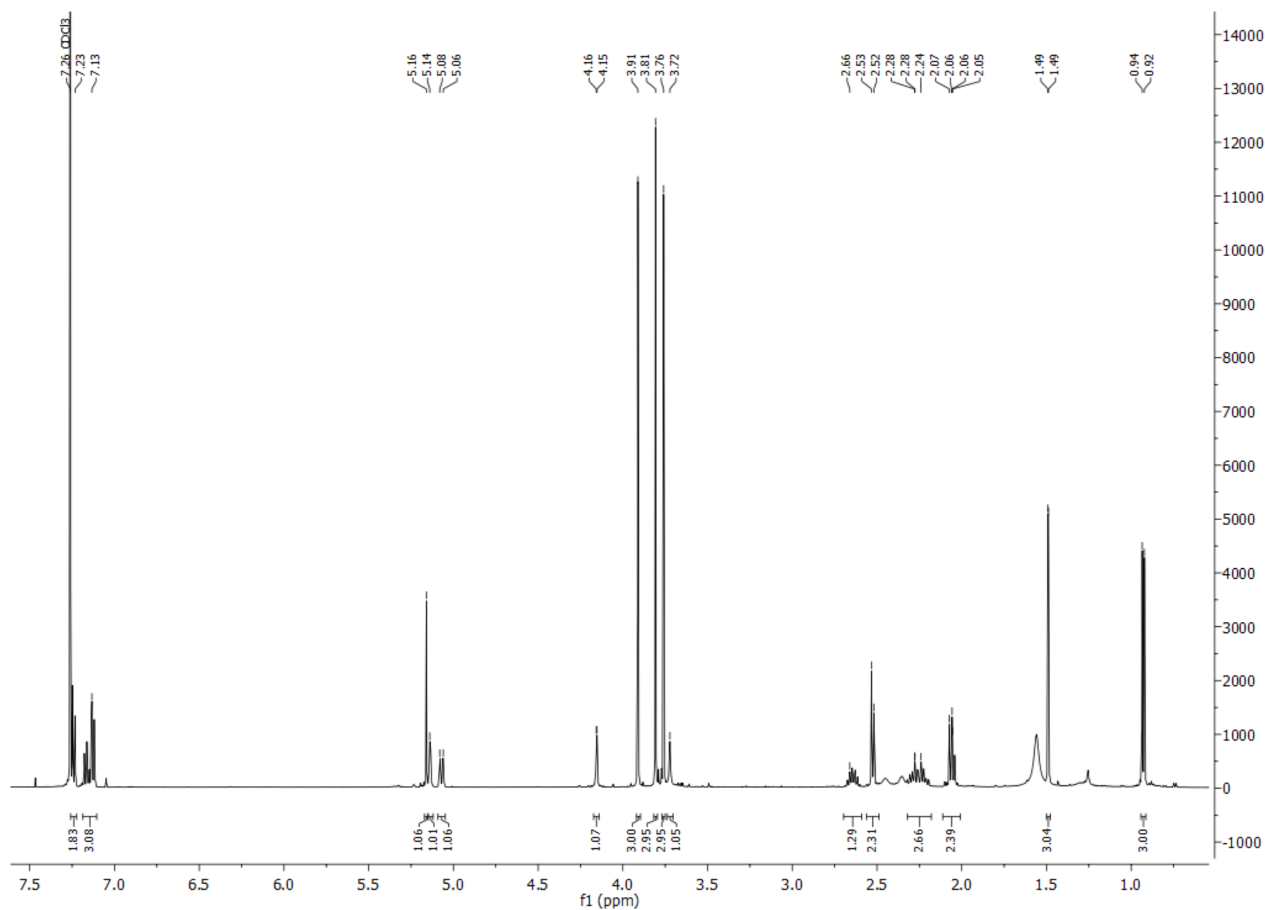


Figure S23: ^{1}H NMR of compound **5A** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. See Table S5 for assignment.

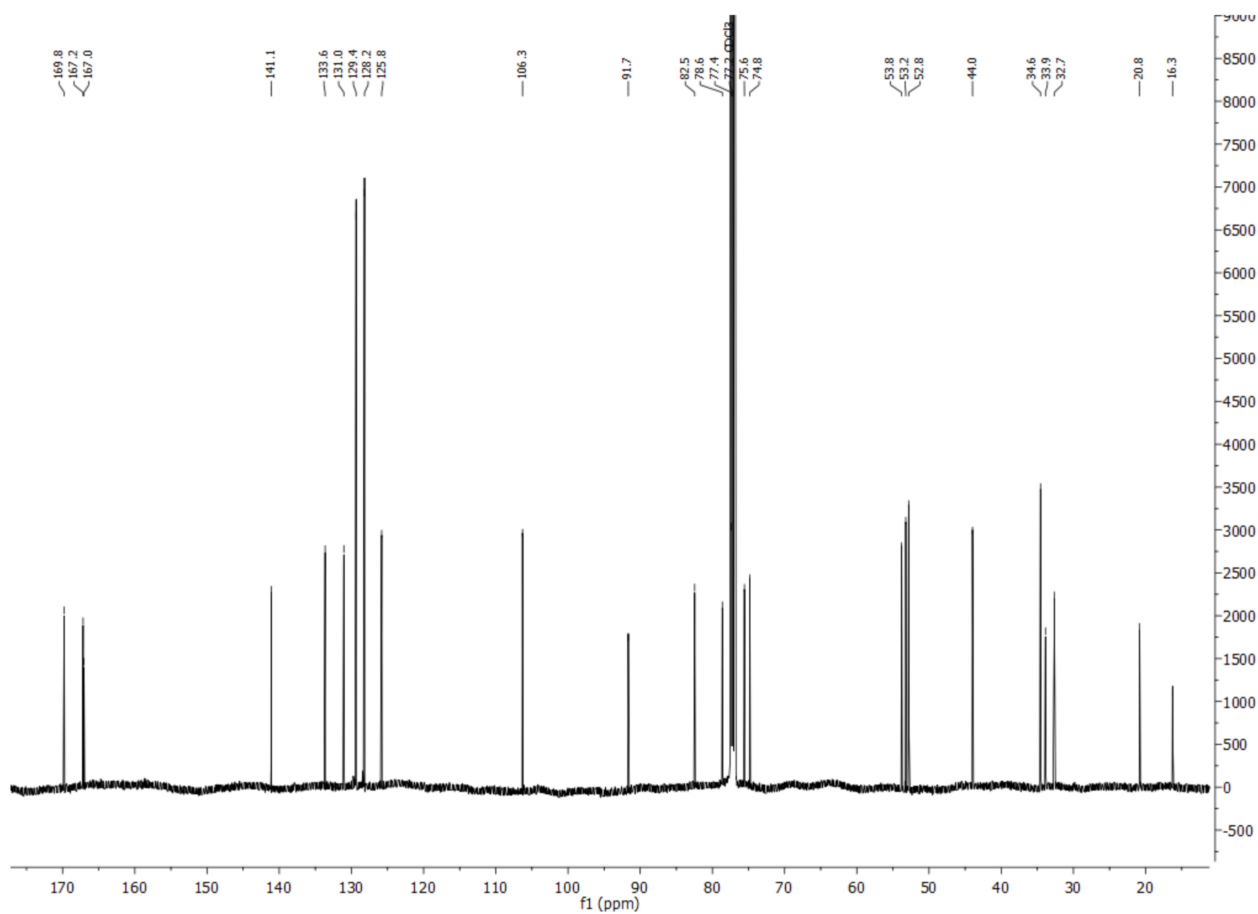


Figure S24: ^{13}C NMR of compound **5A** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. See Table S5 for assignment.

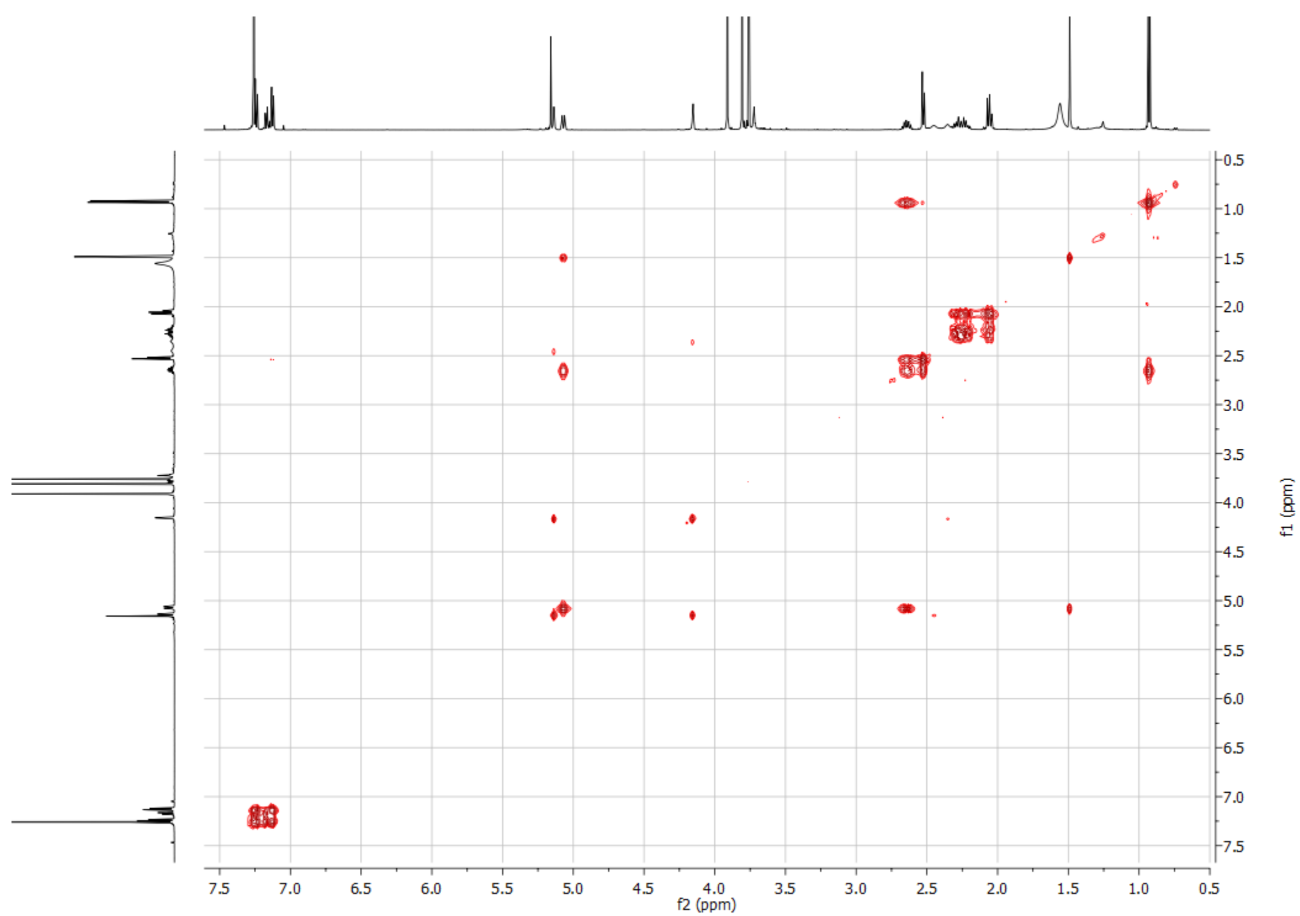


Figure S25: ^{1}H, ^{1}H COSY of compound **5A**.

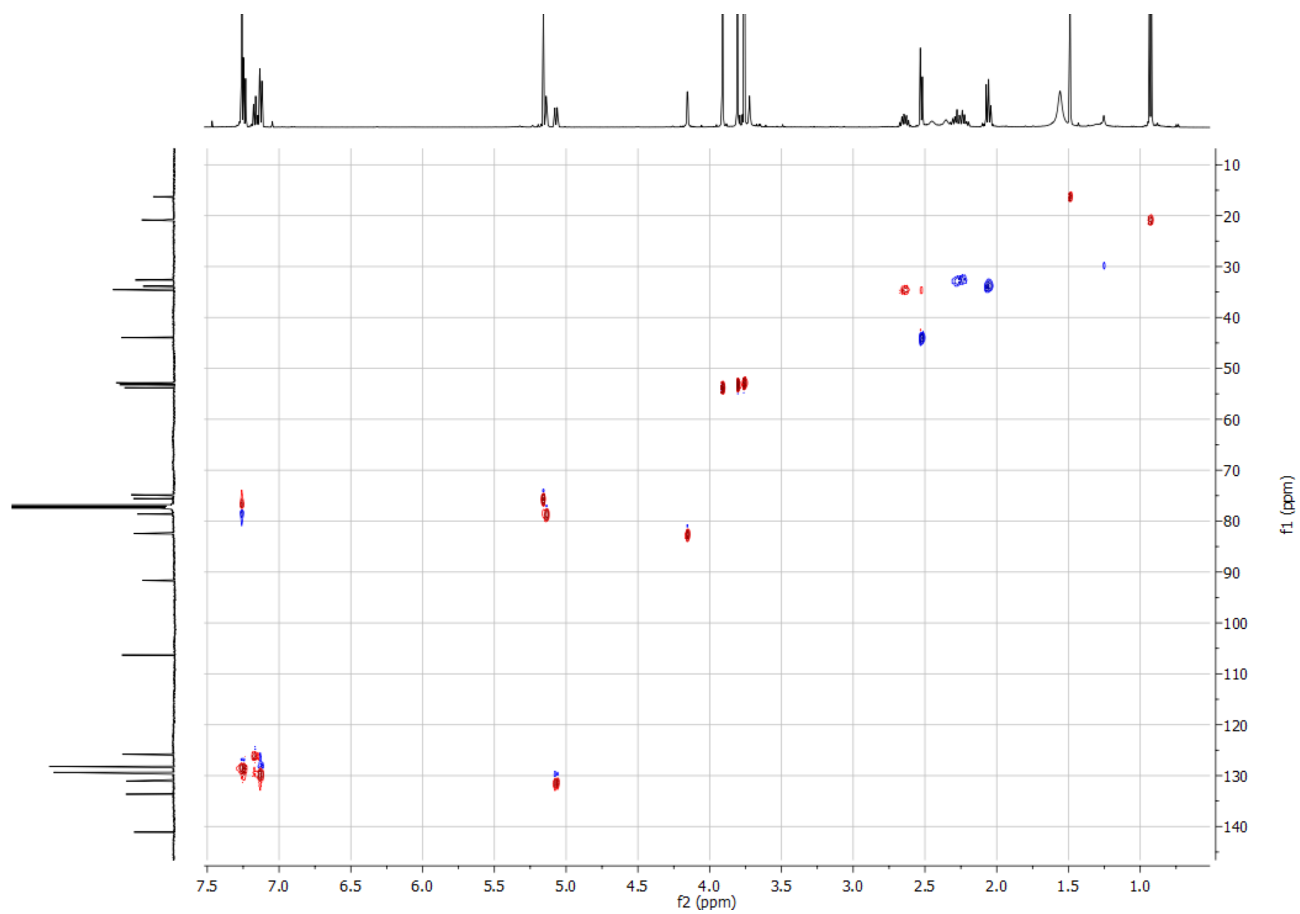


Figure S26: HSQC of compound **5A**.

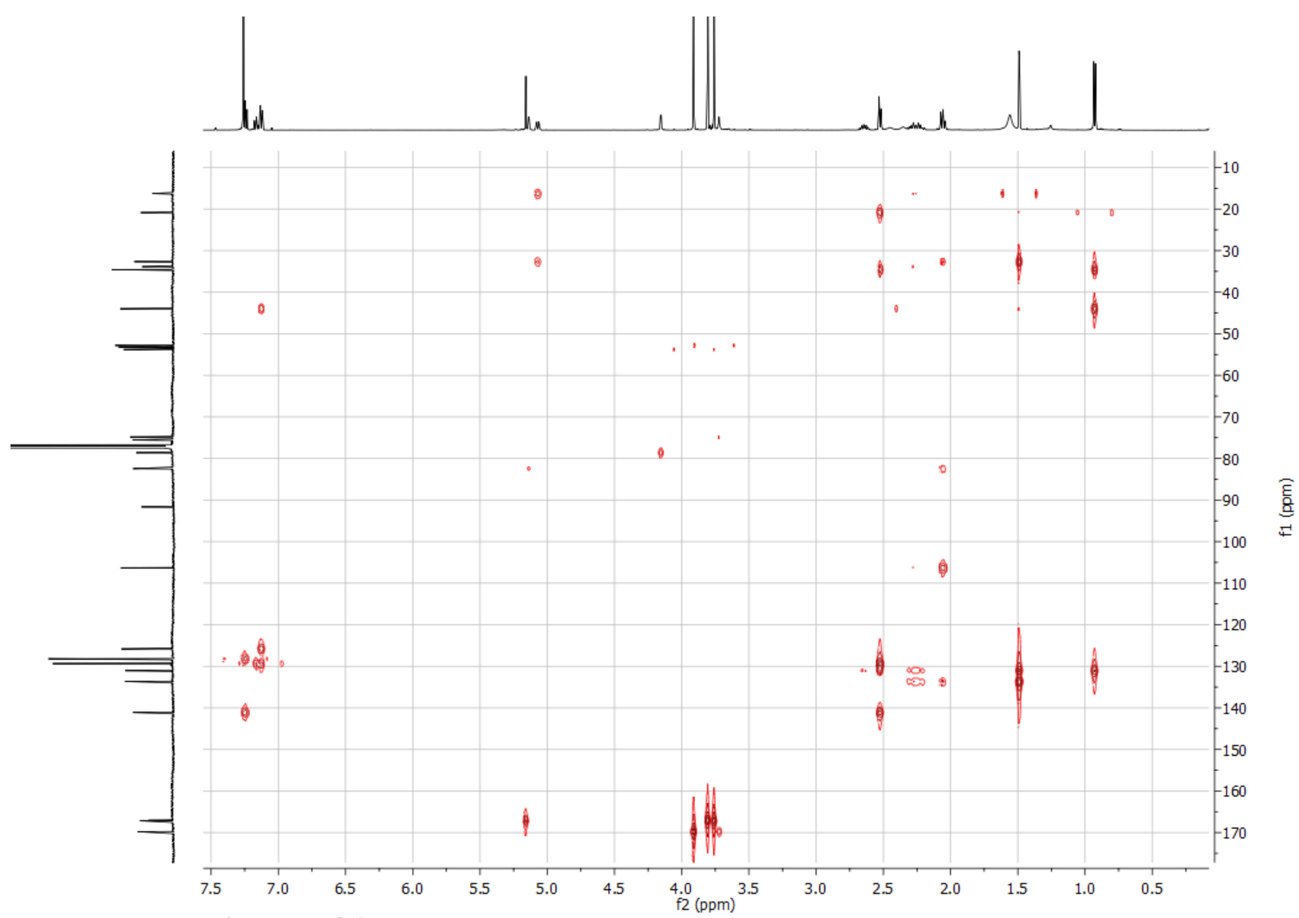
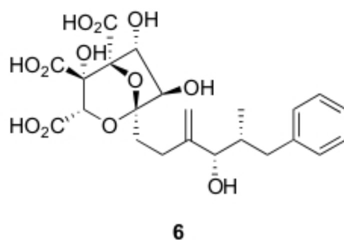


Figure S27: HMBC of compound **5A**.

Compound 6[4,5]

Isolation of **6** from raw extract of *mfr4* KO strain yielded in 0.5 mg product. Comparison to a standard of **6**[4] by HRMS and MSMS measurements confirmed the structure.



HRMS (ESI-) *m/z* calc. for $C_{23}H_{27}O_{12}$ [M-H]$^-$ 495.1503, found 495.1507.

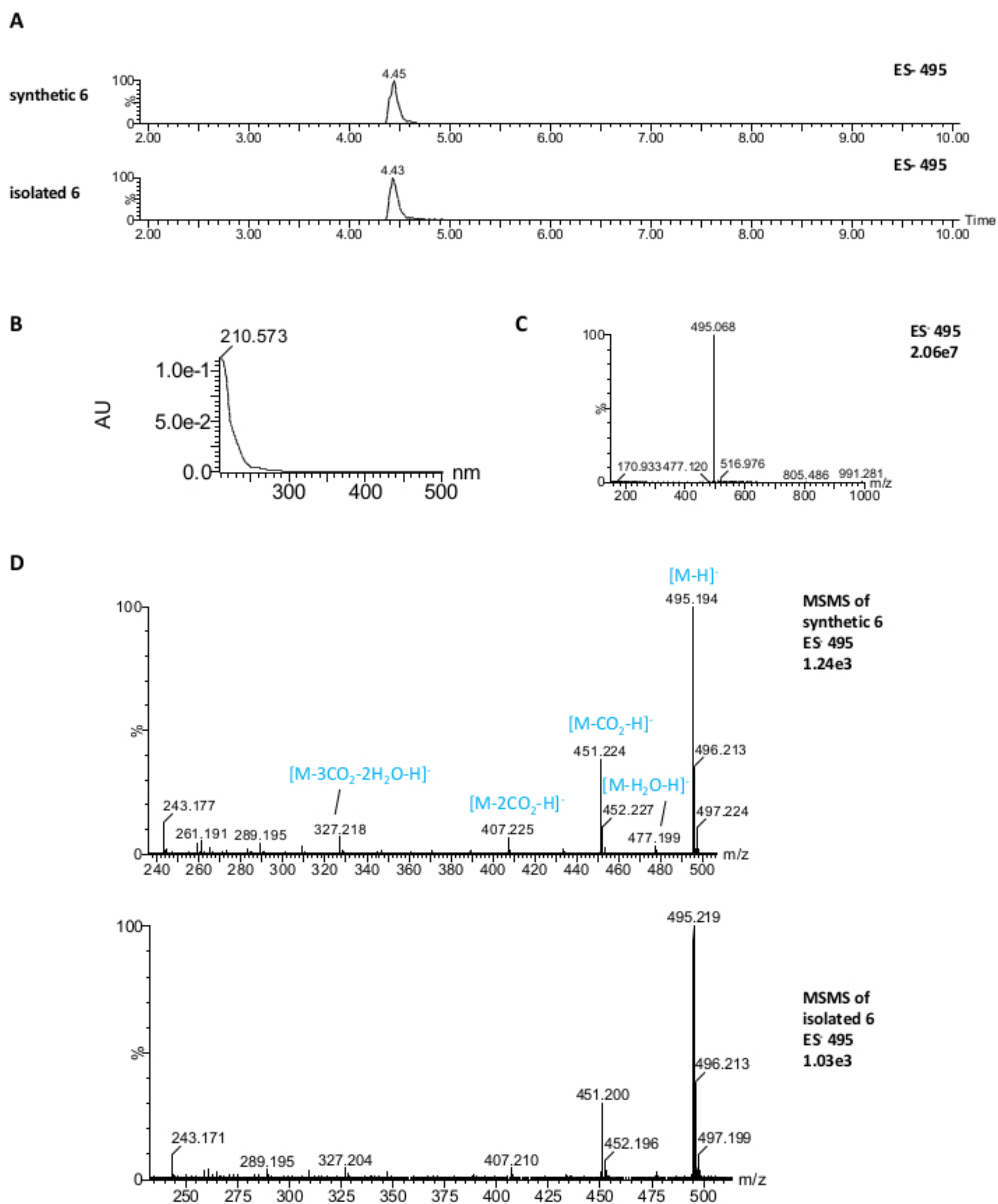
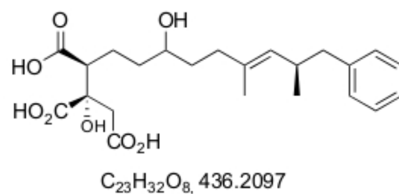


Figure S28: LCMS data for compound **6** in comparison to a synthetic standard of **6**, (A) Extracted ion chromatograms (ES-495), (B) UV chromatogram, (C) mass spectrum (ES-), (D) MSMS fragmentation pattern of compound **6** in comparison to a synthetic standard of **6**.

Compound 7



HRMS (ESI-) *m/z* calc. for $C_{23}H_{32}O_8$ $[M-H]^-$ 435.2019, found 435.2020

A

ES- 435
1.36e7

7.69

100
%
0
2.00 3.00 4.00 5.00 6.00 7.00 8.00 9.00 10.00 Time

B

210.573

AU
4.0e-1
2.0e-1
0.0
300 400 500 600 nm

C

[M-H]-
435.199
433.219
449.241
273.168
417.448
451.330
517.401

100
%
0
300 400 500 600 m/z

D

[M-H]-
435.130

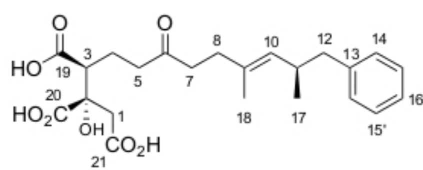
MSMS
ES- 435

[M-H2O-H]-
417.128
434.107
418.123
[M-2CO2-H2O-H]-
329.159
[M-CO2-2H2O-H]-
355.119
301.156

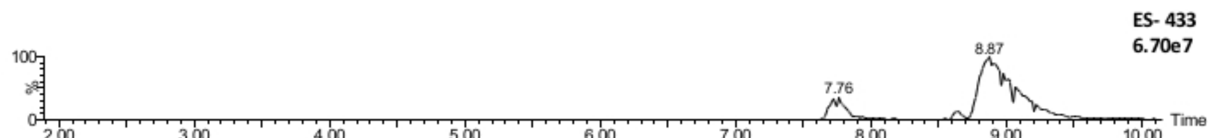
100
%
0
300 320 340 360 380 400 420 m/z

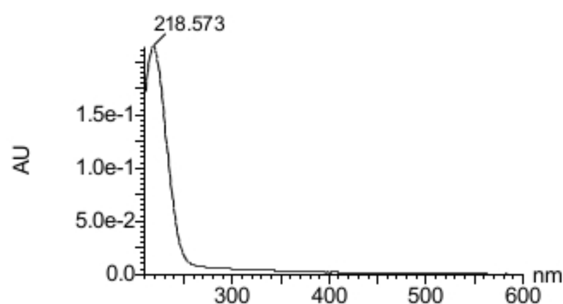
Figure S29: LCMS data for compound **7**, (A) Extracted ion chromatogram (ES- 435), (B) UV chromatogram (C) mass spectra (ES-), (D) MSMS fragmentation pattern of compound **7**.

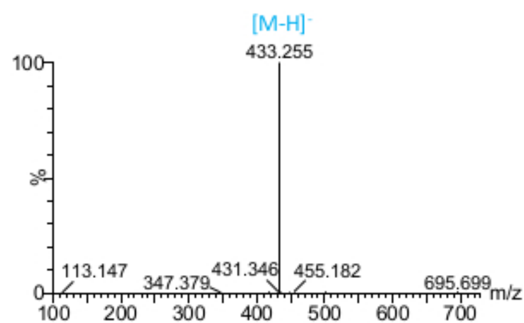
Compound 8

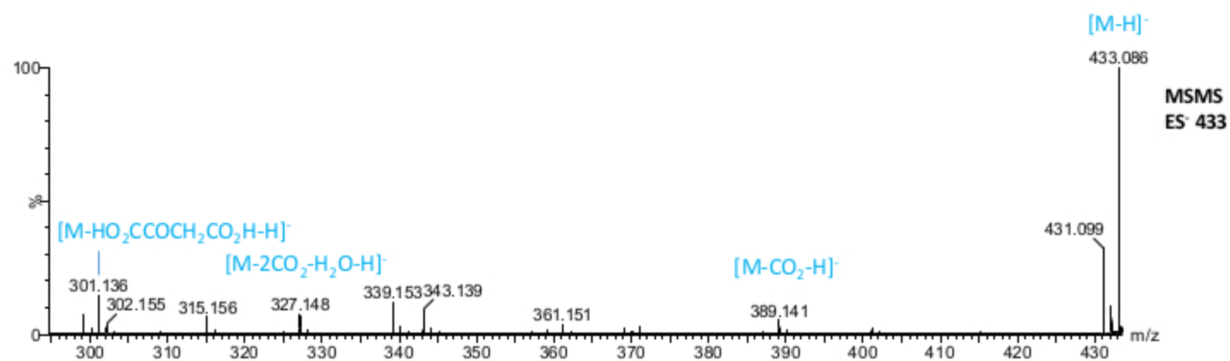


HRMS (ESI-) *m/z* calc. for $C_{23}H_{30}O_8$ $[M-H]^-$ 433.1862, found 433.1861









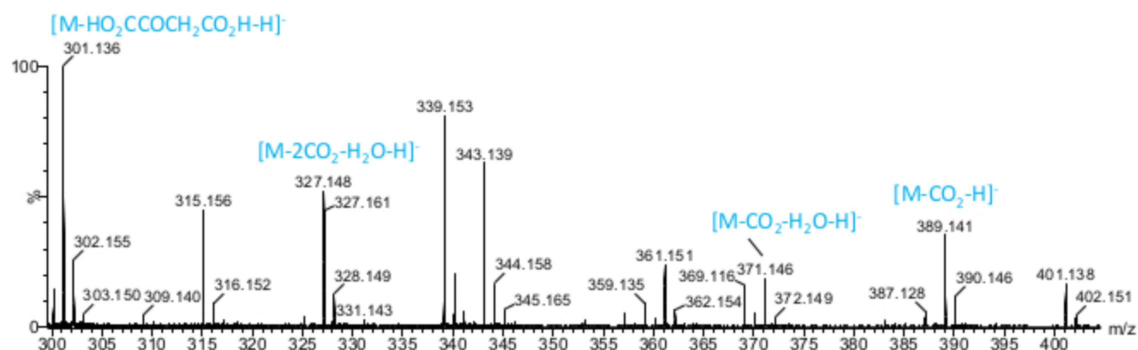


Figure S30: LCMS data for compound **8**, (A) Extracted ion chromatogram (ES- 433), (B) UV chromatogram (C) mass spectra (ES-), (D) MSMS fragmentation pattern of compound **8**.

Table S5: Partial NMR assignment of compound 8 – C1-9, C11 and C19 to 21 could not be assigned

Pos.	δ_C / ppm (125 MHz)	δ_H / ppm (500 MHz)	*J* / Hz
1			
...			
9			
10	129.5	5.36	m
11			m
12	43.7	2.47-2.54	m
13	141.0	q	
14	129.2	7.17	m
15	127.7	7.28	m
16	125.8	7.17	m
17	20.6	0.92	m
18	15.4	1.37	m
19			
...			
21			

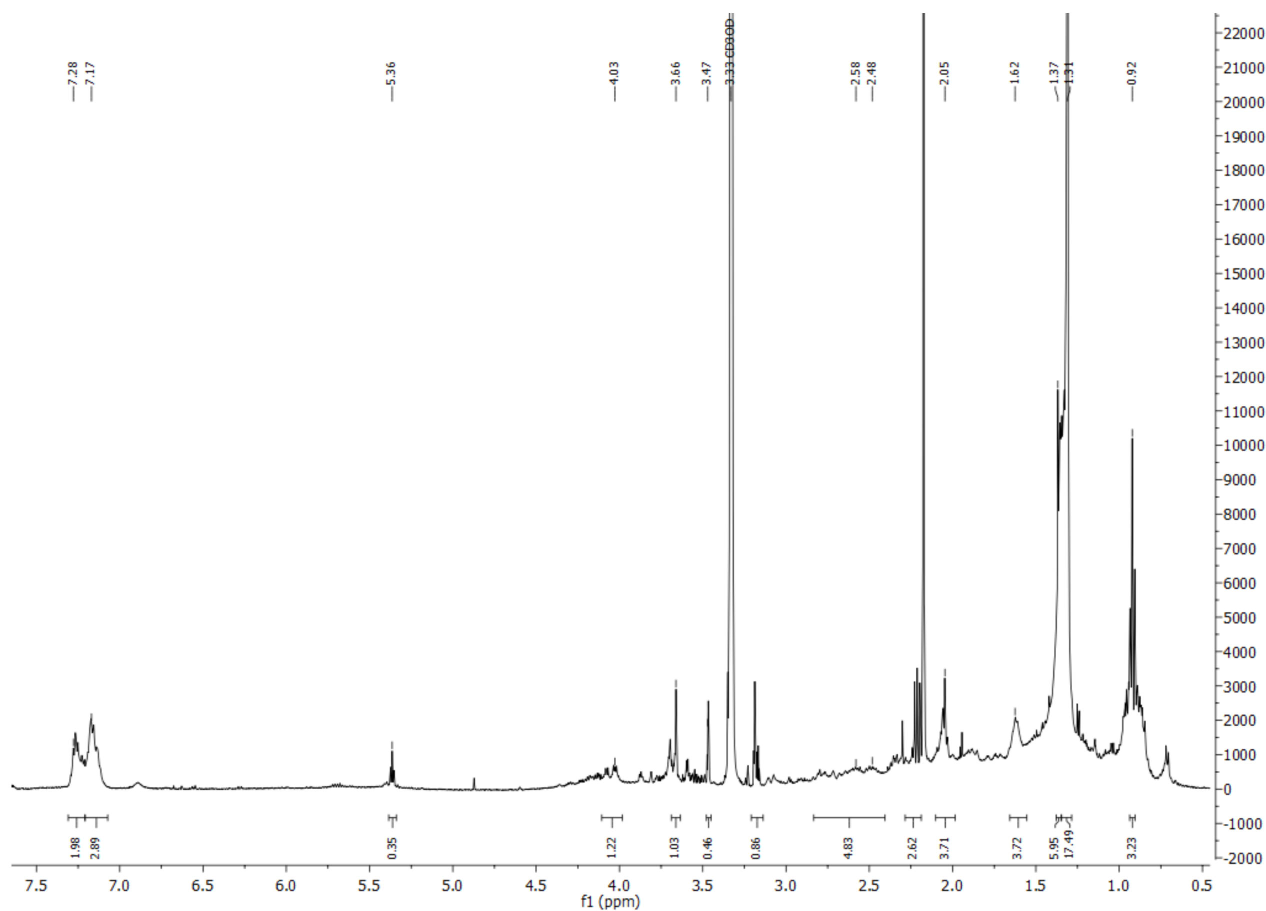


Figure S31: ^{1}H NMR of compound **8** in CD_3OD (500 MHz) referenced to CD_3OD. See Table S6 for partial assignment.

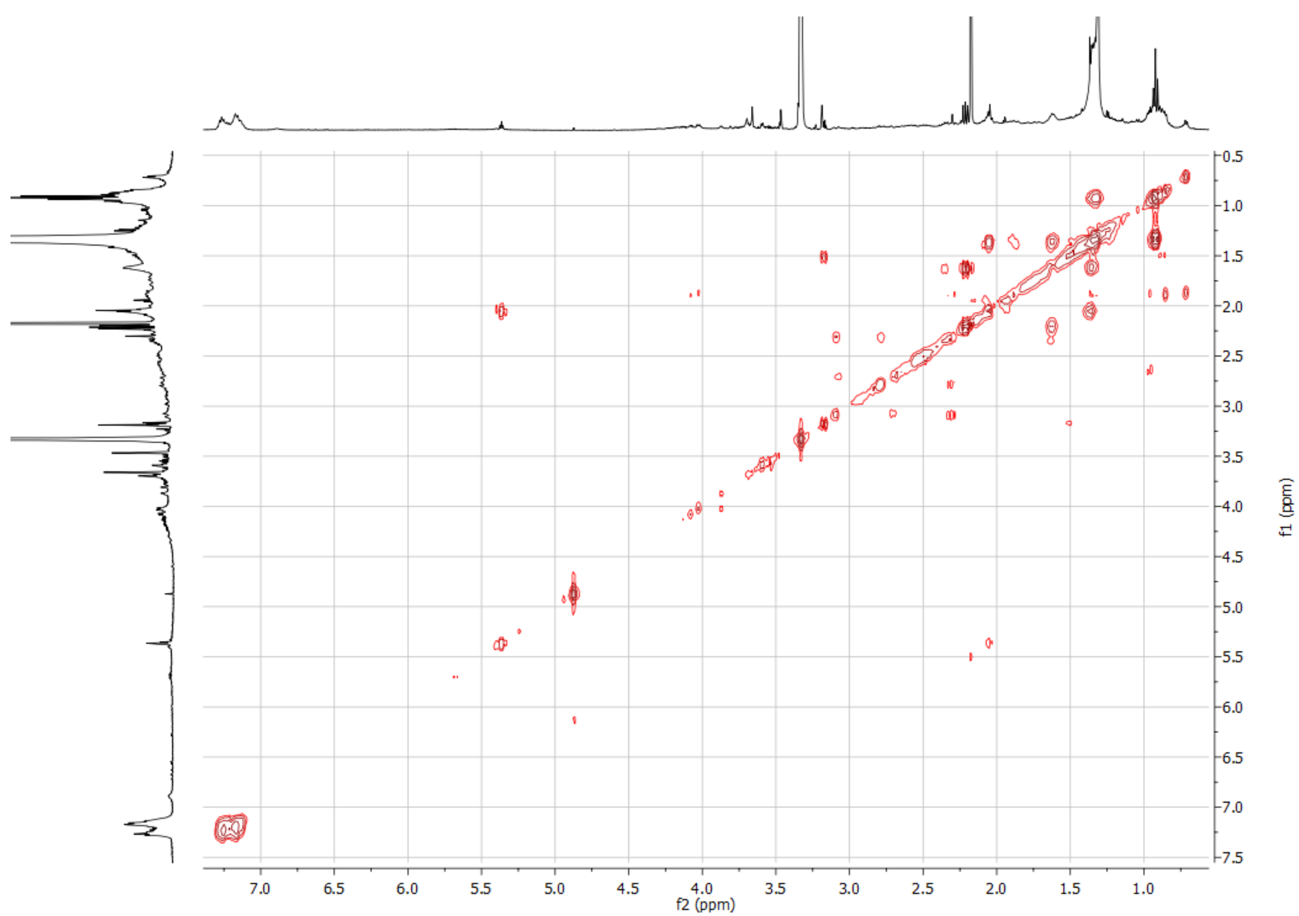


Figure S32: COSY of compound **8**.

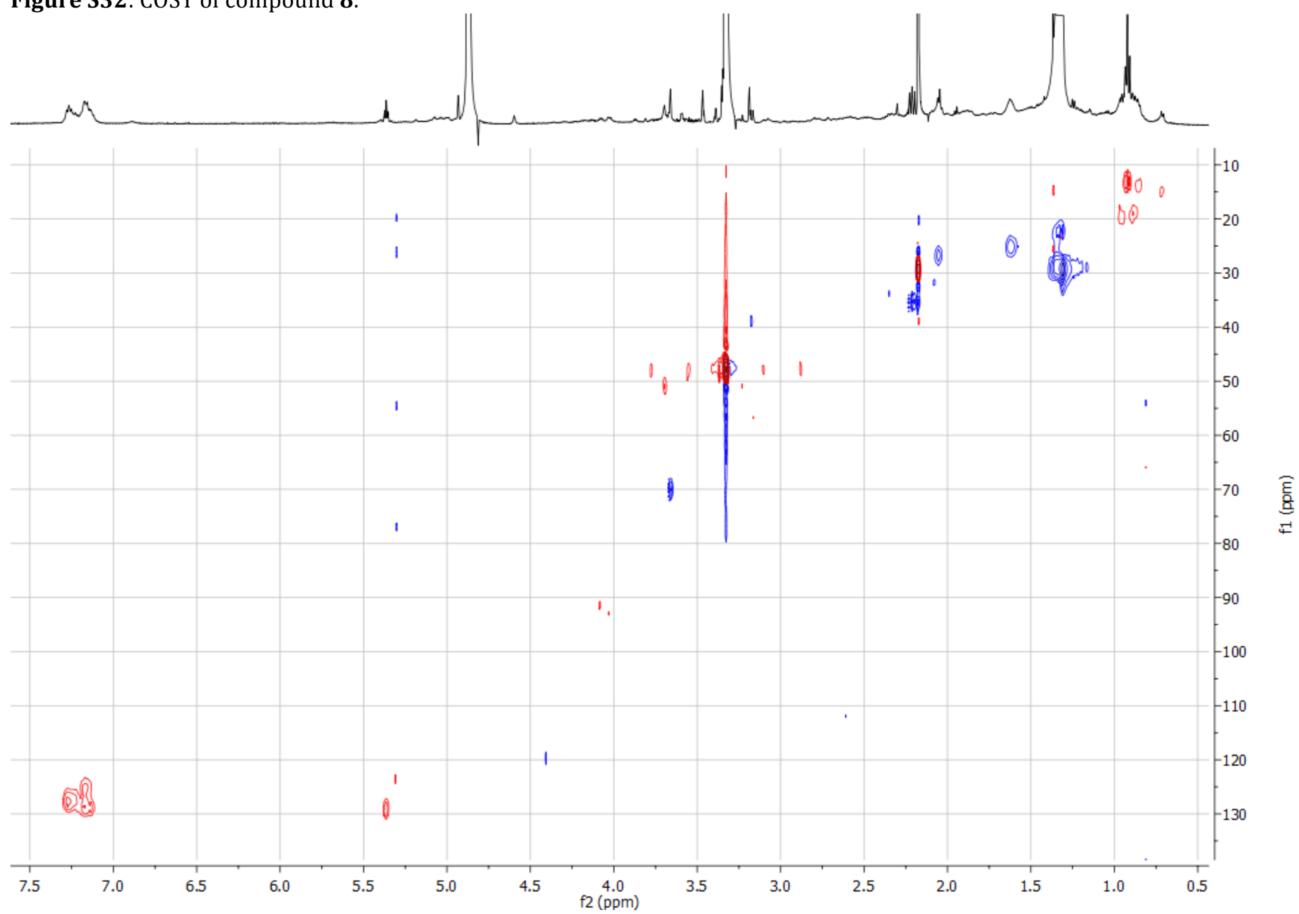


Figure S33: HSQC of compound **8**.

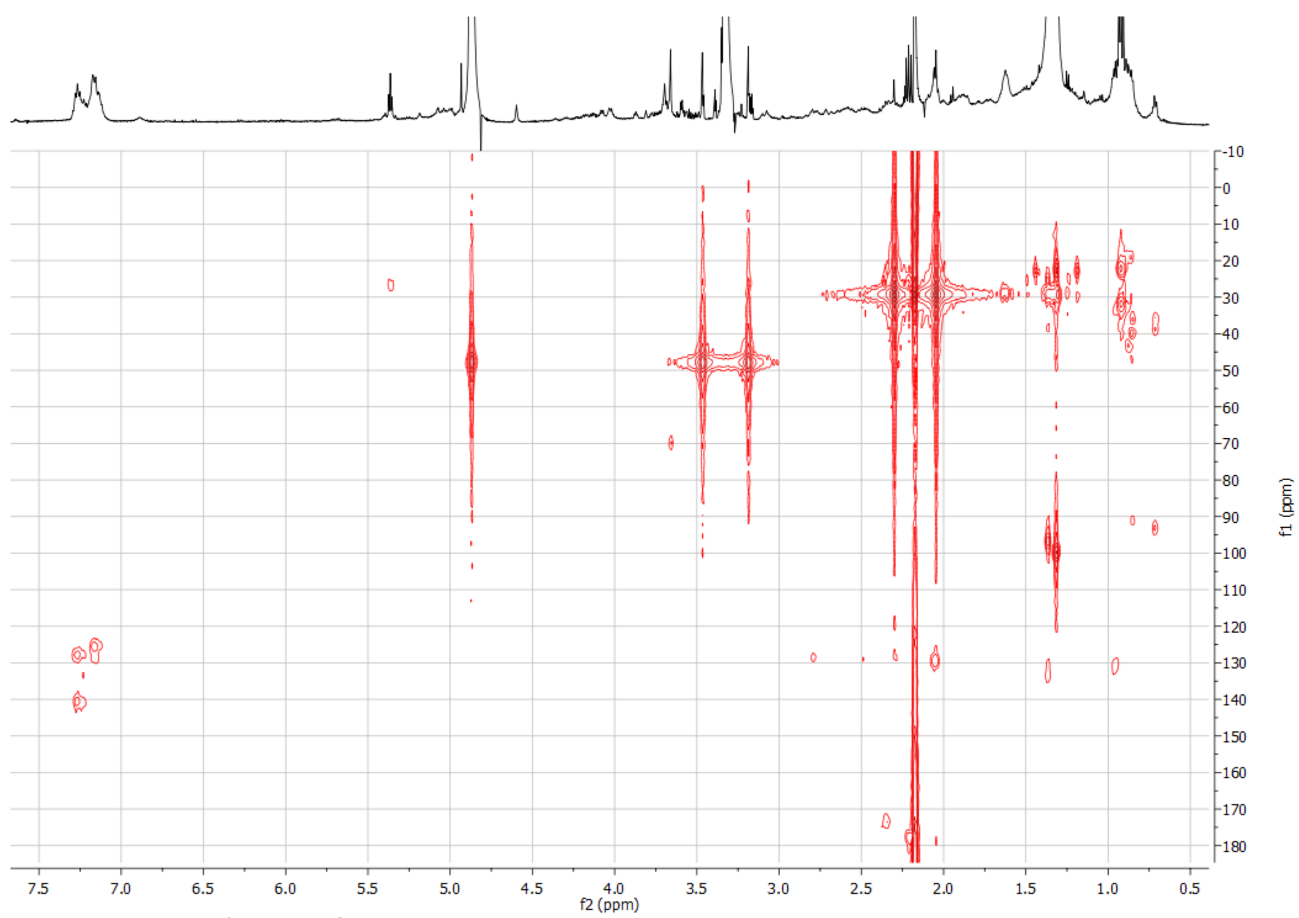
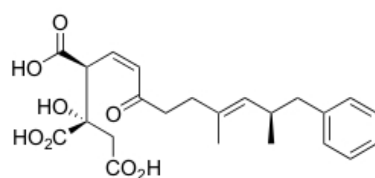


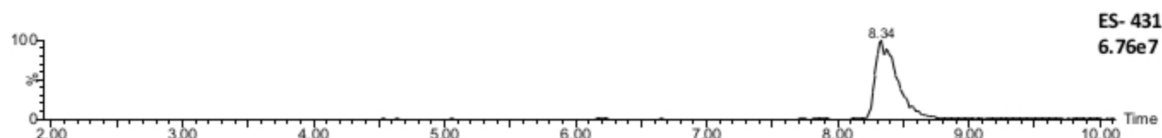
Figure S34: HMBC of compound **8.**

Compound 9

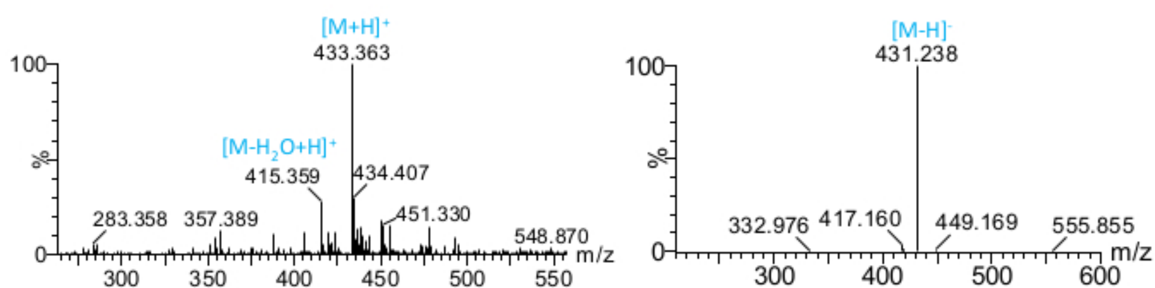


HRMS (ESI-) *m/z* calc. for $C_{23}H_{28}O_8$ $[M-H]^-$ 431.1706, found 431.1702

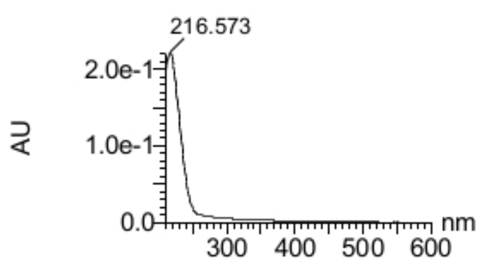
A



B



C



D

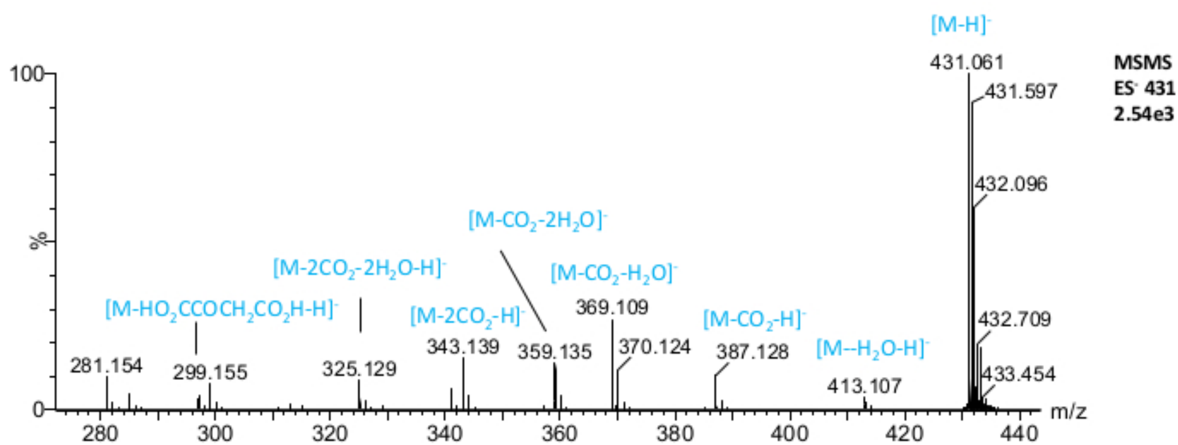
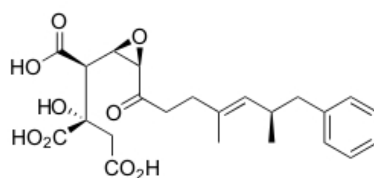


Figure S35: LCMS data for compound **9,** (A) Extracted ion chromatogram (ES- 431), (B) mass spectra (ES- and ES+), (C) UV chromatogram, (D) MSMS fragmentation pattern of compound **9**.

Compound 10



HRMS (ESI-) *m/z* calc. for $C_{23}H_{28}O_9$ $[M-H]^-$ 477.1655, found 447.1655

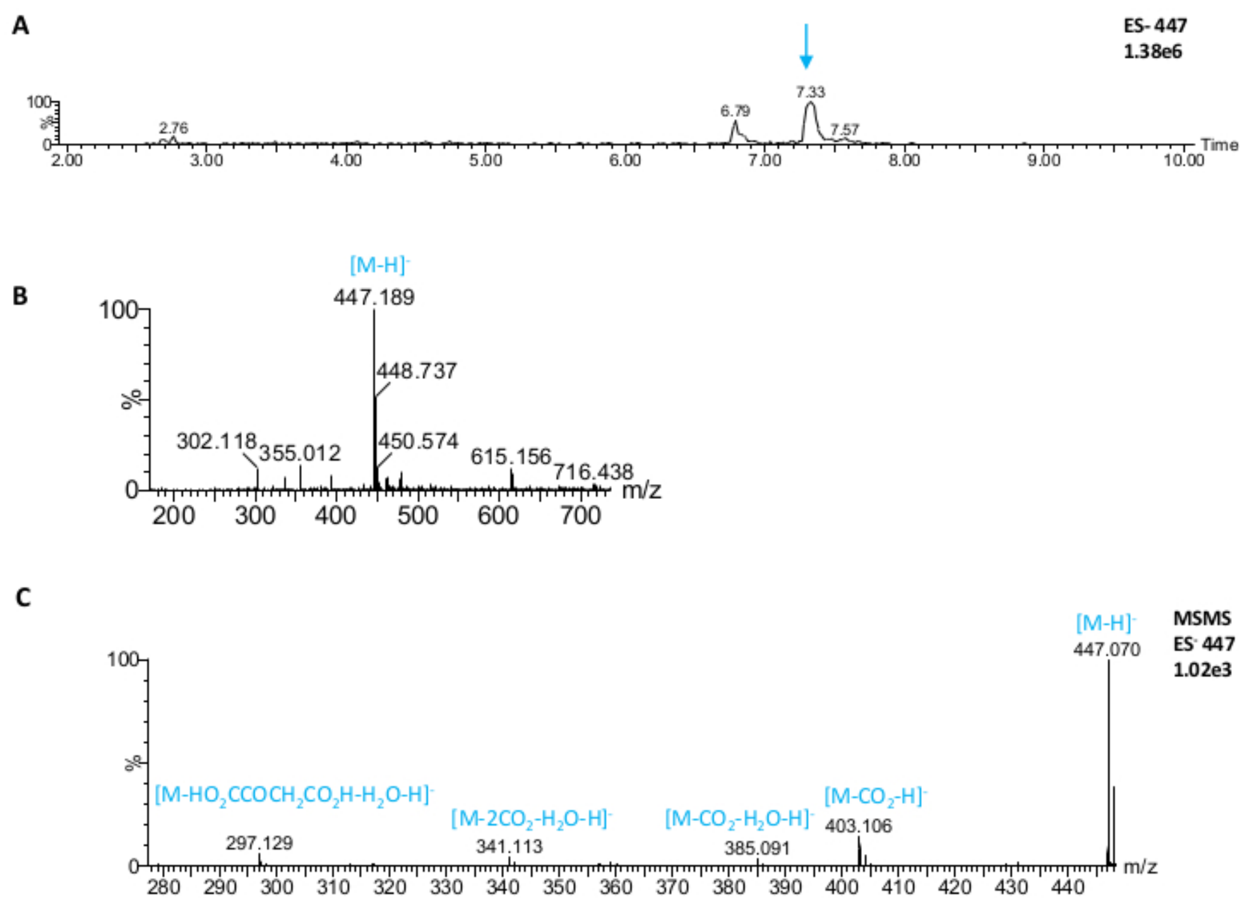
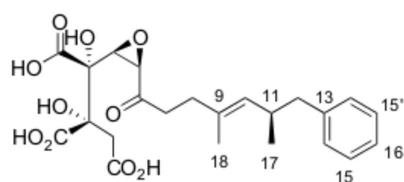


Figure S36: LCMS data for compound **10,** (A) Extracted ion chromatogram (ES- 448), (B) mass spectra (ES-), (C) MSMS fragmentation pattern of compound **10**.

Compound 11



HRMS (ESI-) *m/z* calc. for $C_{23}H_{27}O_{10}$ $[M-H]^-$ 463.1604, found 463.1605

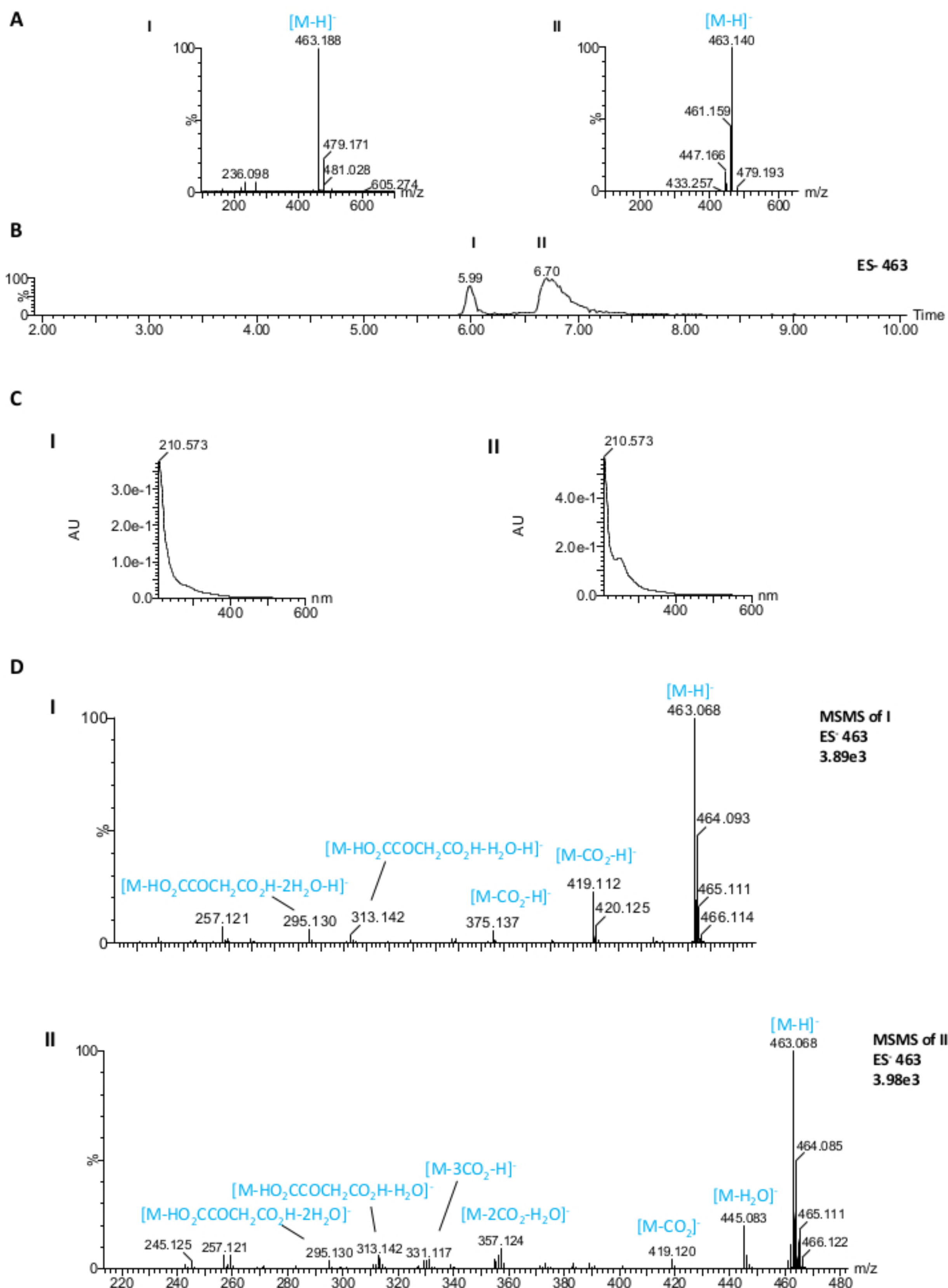


Figure S37: LCMS data for compound **11**, (A) Mass spectra of I and II (ES-), (B) Extracted ion chromatogram (ES- 463) showing two compounds (I and II), (C) UV chromatograms of I and II, (D) MSMS fragmentation pattern of both compounds with the mass of 464 (see chromatogram A) **11_I** and **11_II**.

Table S6: Partial NMR assignment of compound **11_II –** C1-8 and C19 to 21 could not be assigned

		CD$_3$OD			
Pos.	**δ_C / ppm (125 MHz)**	**δ_H / ppm (500 MHz)**	***J* / Hz**	**^{1}H-^{1}H COSY**	**HMBC**
1					
2					
3					
4					
5					
6					
7					
8					
9	135.2	q			18
10	131.3	5.02-5.06	m	11, 18	8, 11, 17, 18
11	36	2.66	m	10, 17	10, 12, 17
12	45.1	2.46-2.58	m	11	11, 13, 14, 17
13	142.3	q			12, 15
14	130.2	7.14	m	15	15, 16
15	128.8	7.24	m	14	14, 13
16	126.4	7.14	m		14, 15
17	21.6	0.97	m	11	10, 11, 12
18	16.3	1.43	d		8, 9, 10
19					
20					
21					

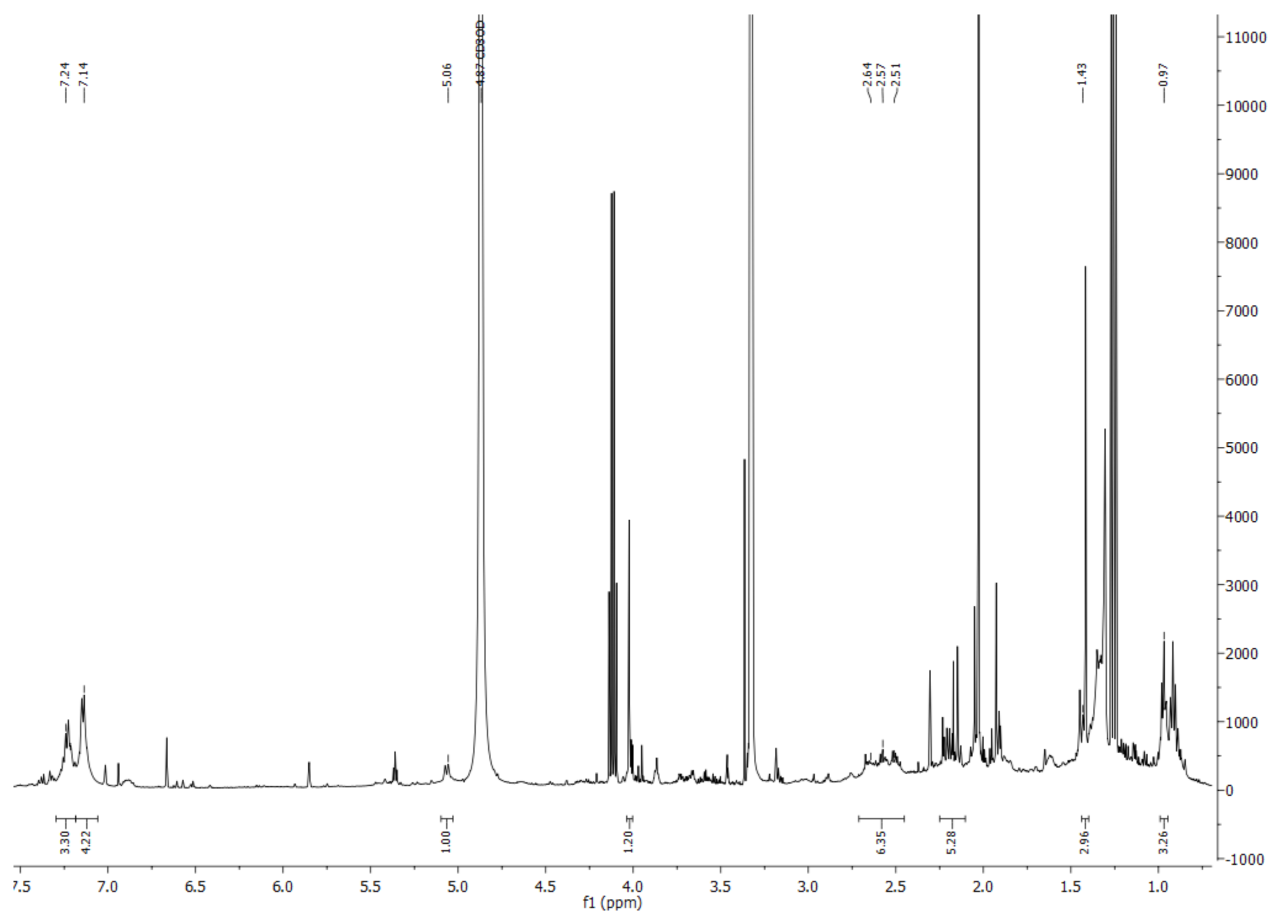


Figure S38: ^{1}H NMR of compound **11_II** in CD_3OD (500 MHz) referenced to CD_3OD. See Table S7 for partial assignment.

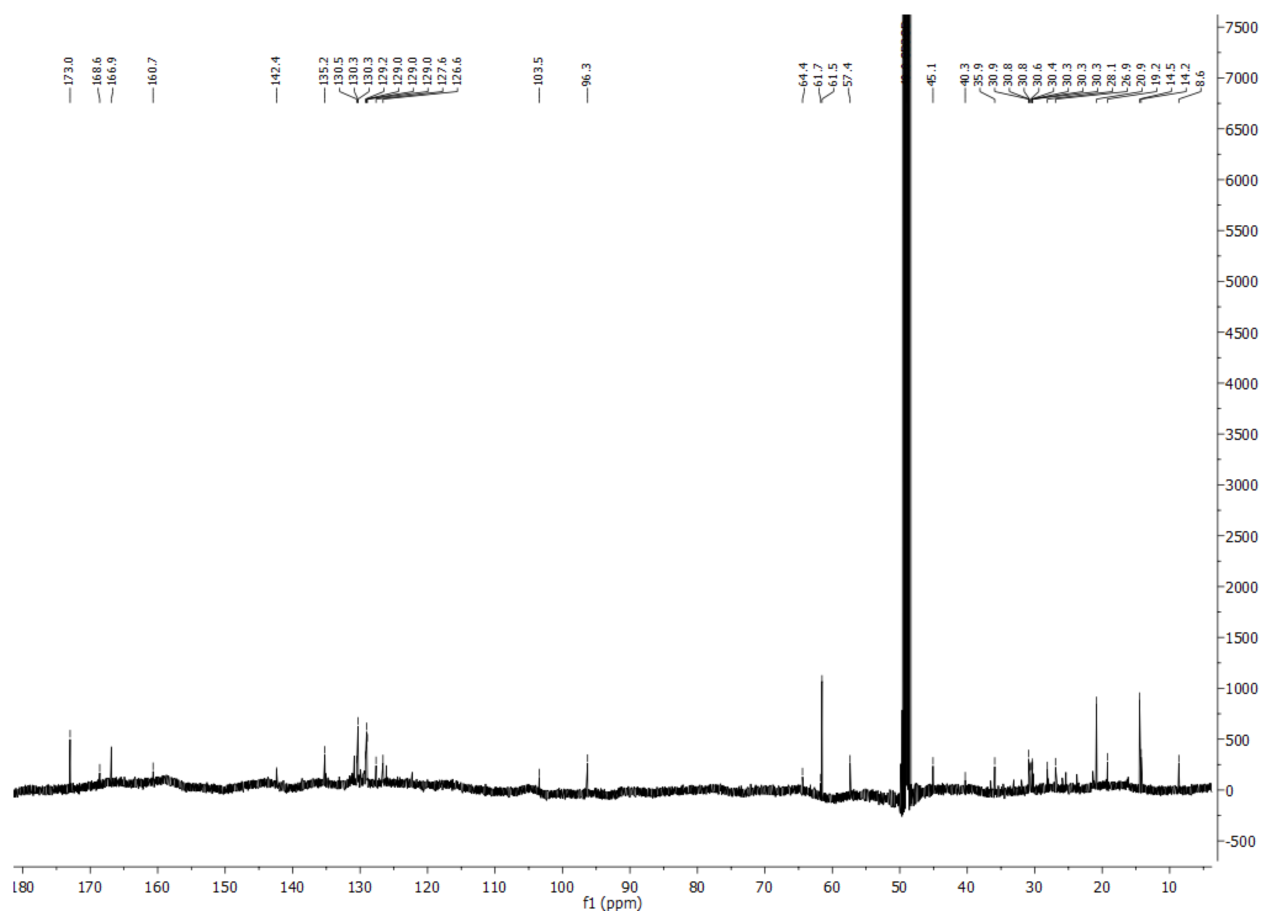


Figure S39: ^{13}C NMR of compound **11_II** in CD_3OD (500 MHz) referenced to CD_3OD. See Table S7 for partial for assignment.

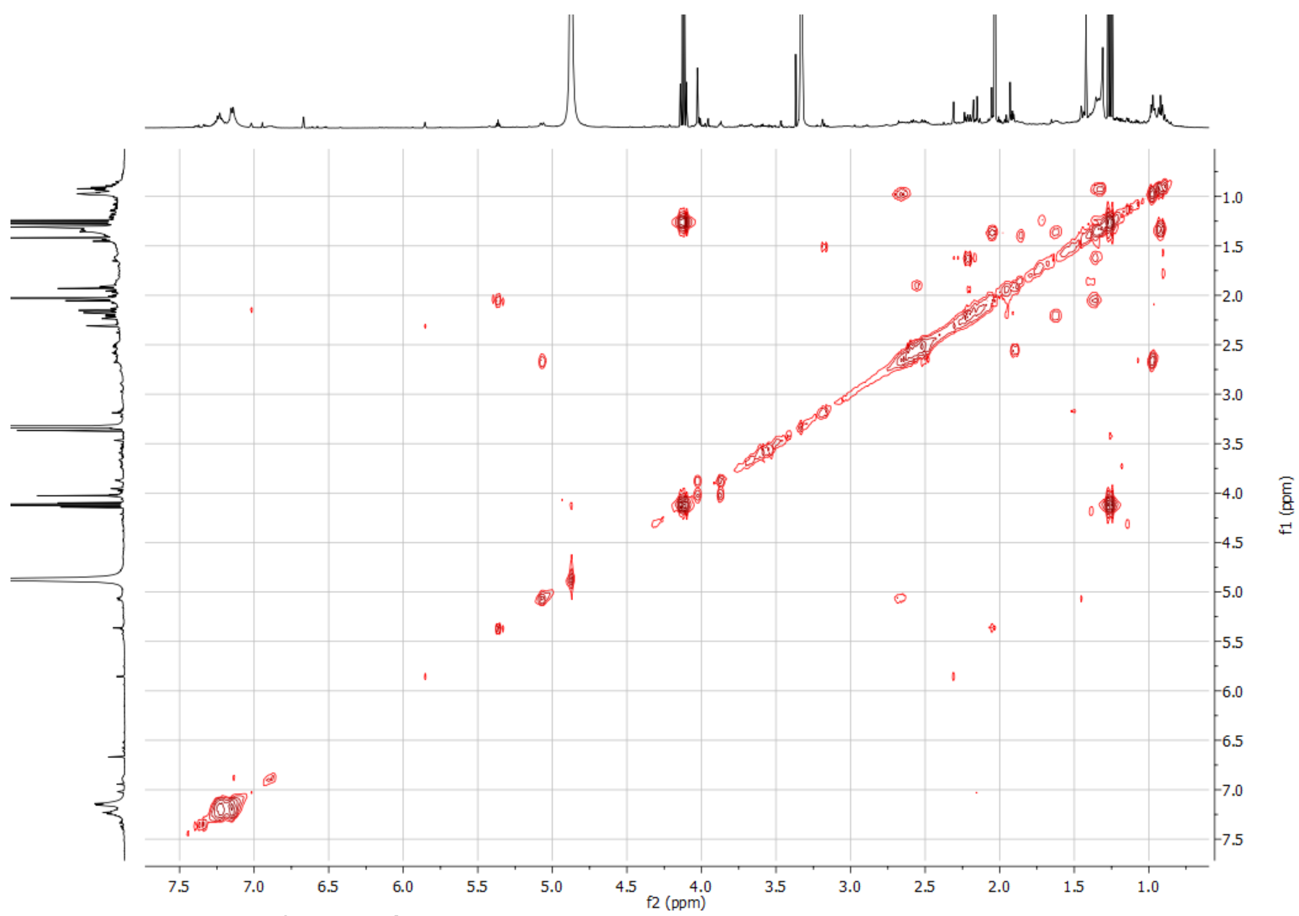


Figure S40: ^{1}H, ^{1}H COSY of compound **11_II**.

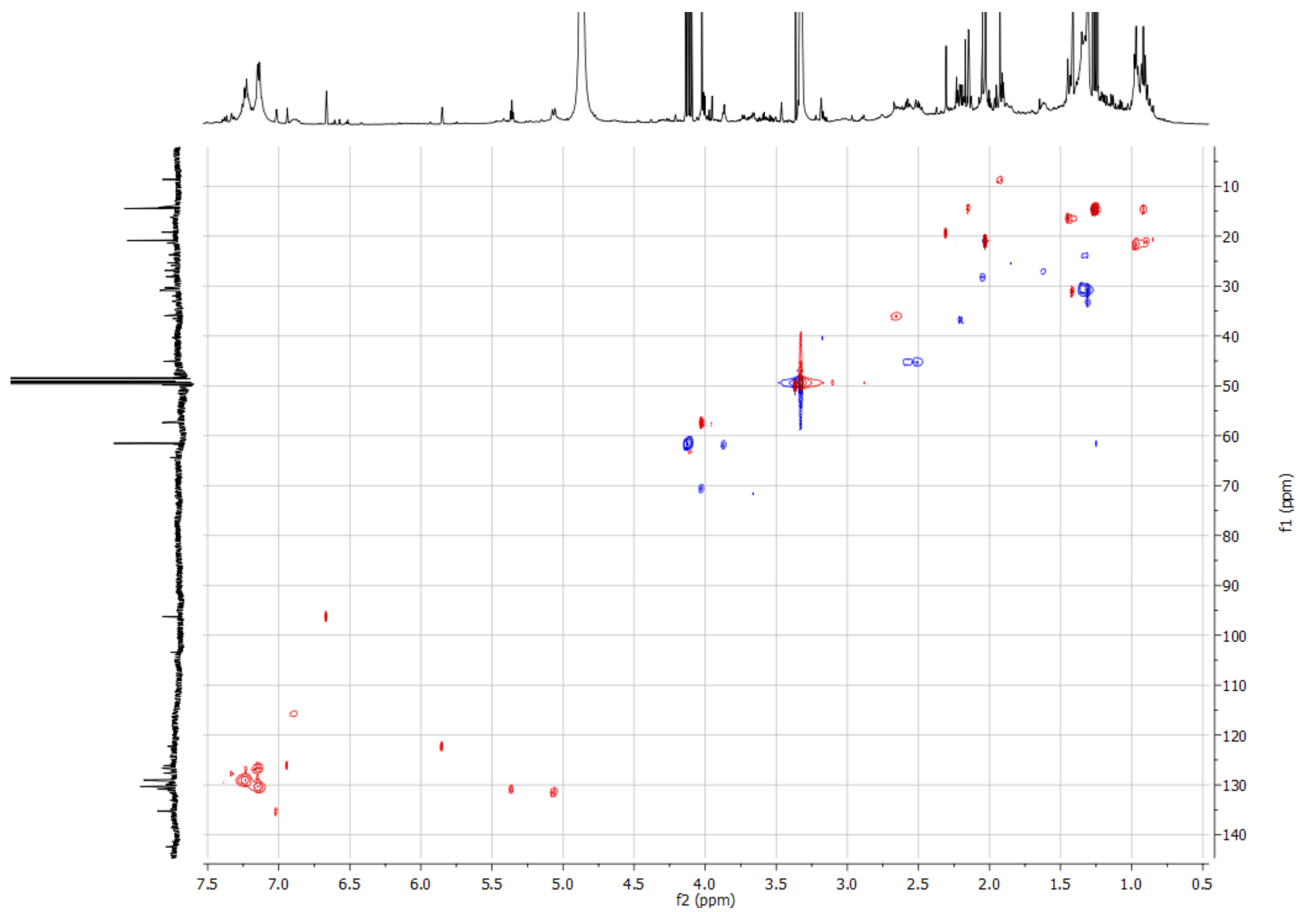


Figure S41: HSQC of compound **11_II**.

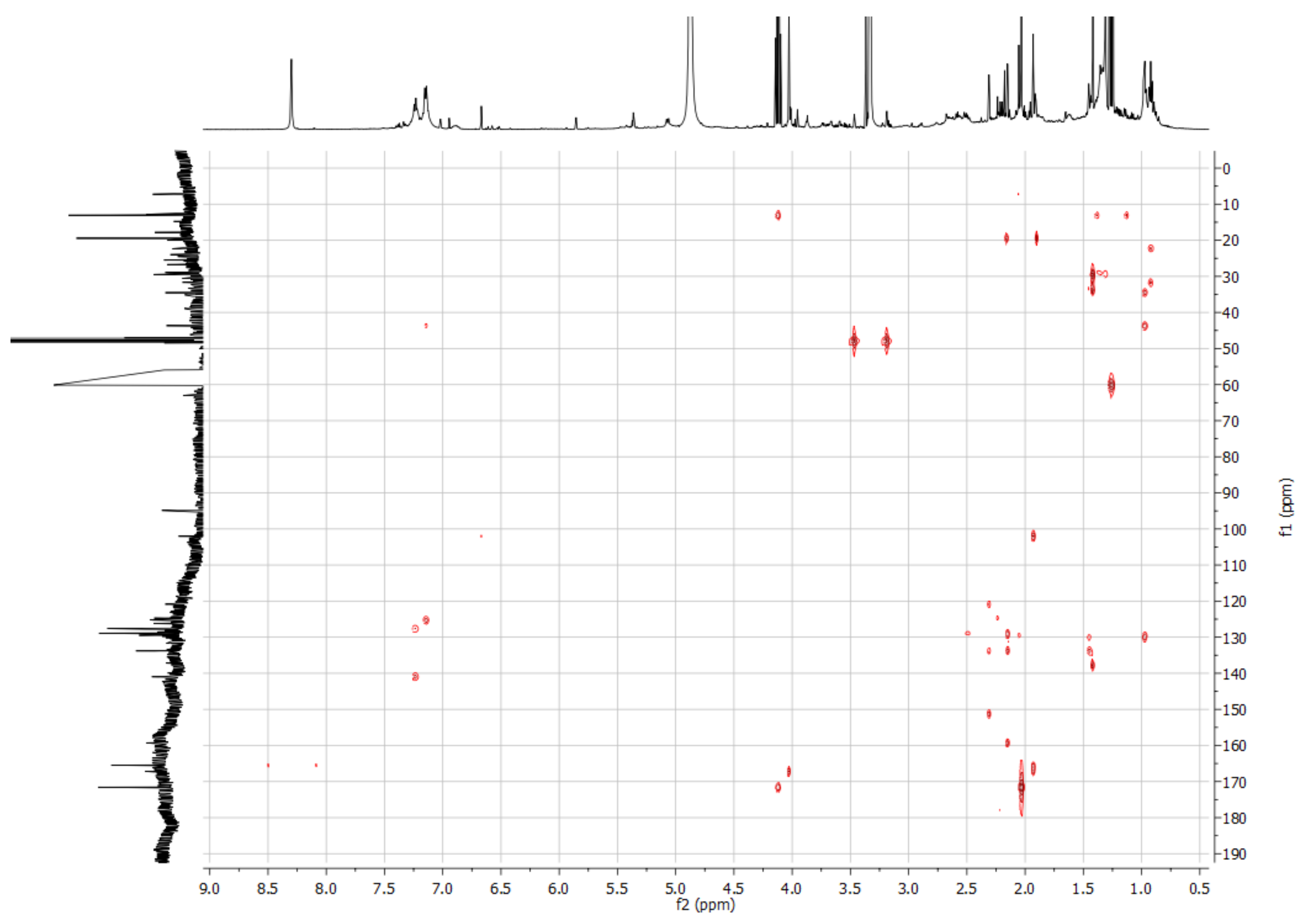
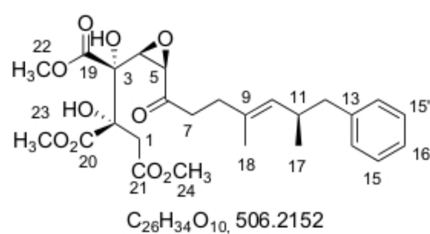


Figure S42: HMBC of compound **11_II**.

Compound 11A

Isolation of **11** from raw extract of M1 KO strain (1L) yielded 1.5 mg product. This was treated with $TMSCHN_2$ and repurified to yield 0.4 mg of **11A.**



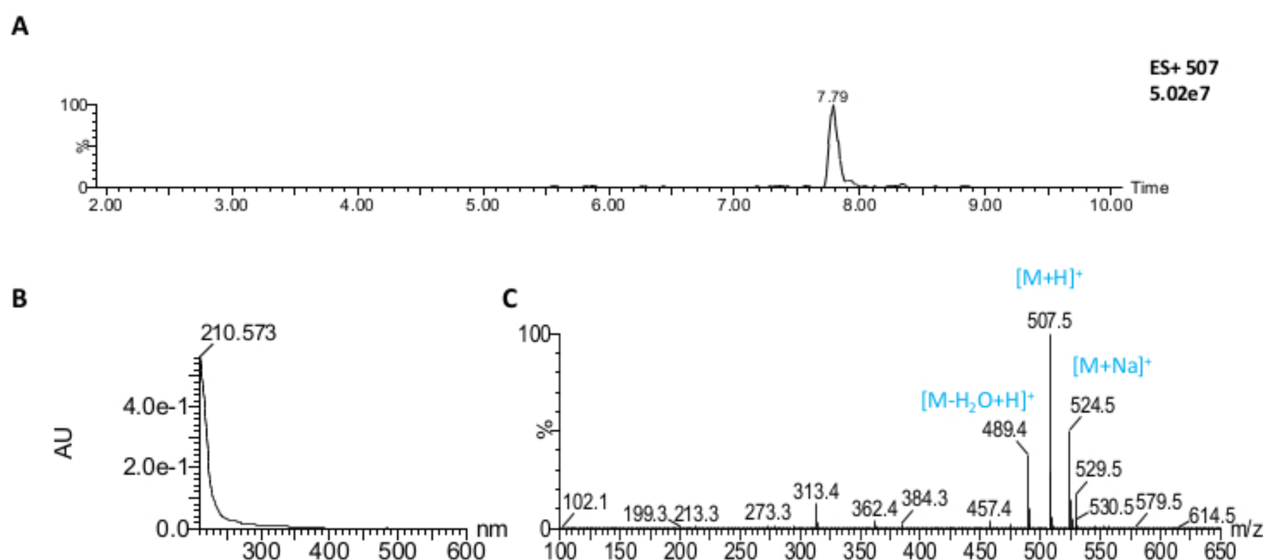


Figure S43: LCMS data for compound **11A**, (A) Mass spectra of compound **11A** (ES+), (B) UV chromatograms (C) Extracted ion chromatogram (ES+ 507).

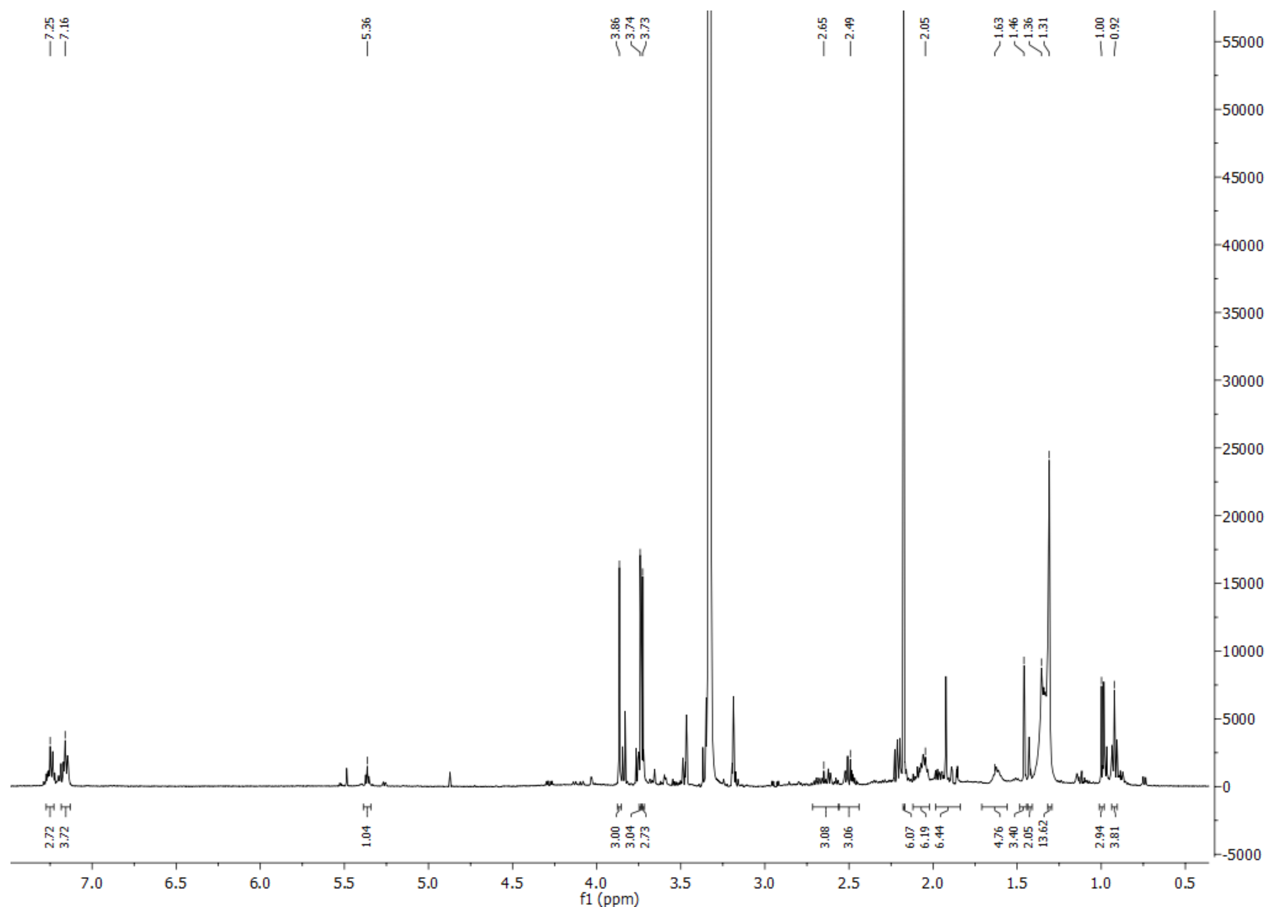


Figure S44: ^{1}H NMR of compound **11A** in CD_3OD (500 MHz) referenced to CD_3OD. See Table S8 for partial assignment.

Table 7: Partial NMR assignment of compound **11A –** C1-9 and C-19 to C-21 could not be assigned

Pos.	**δ_H / ppm (500 MHz)**	***J* / Hz**
1		
...		
8		
9		
10	5.36	m
11	2.65	m
12	2.49	m
13	q	
14	7.16	m
15	7.25	m
16	7.16	m
17	0.92	m
18	1.46	d, 1.3
19		
...		
22-24	3.96, 3.74, 3.73	s

7.4. Unknown Compound from Expression of Benzoyl CoA Cassette

A. oryzae was transformed with a plasmid containing the genes: *str11* (encoding a phenylalanine ammonia lysase); *str8* (encoding an oxidase); and *str10* (encoding a CoA ligase). In the absence of other genes a new (unidentified) compound was produced eluting at 4.75 min. The same compound was produced in the presence of SQHKS and other squalestatin genes (Figure 1G, 1F, 1E, main manuscript, Figure S45). The peak was isolated and preliminary NMR (Figures S46-S50) shows it is unrelated to the squalestatin compounds.

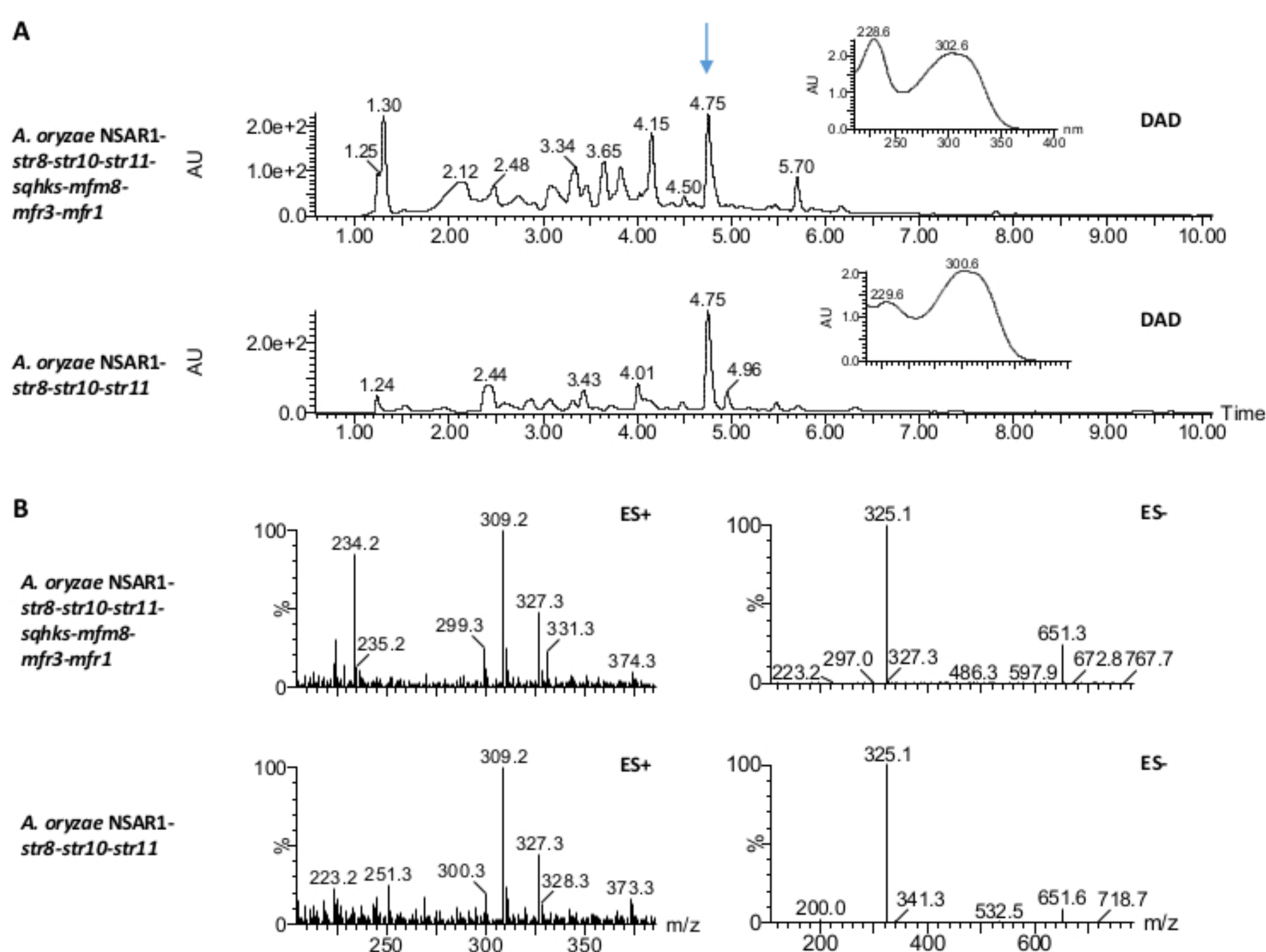


Figure S45. **A**, comparison of LCMS data for extraction of *A. oryzae* with benzoyl CoA genes only (lower) with LCMS data for extraction of *A. oryzae* with benzoyl CoA and squalestatin genes. Insets show uv data for 4.75 min peak; **B**, MS data for the 4.75 min peak.

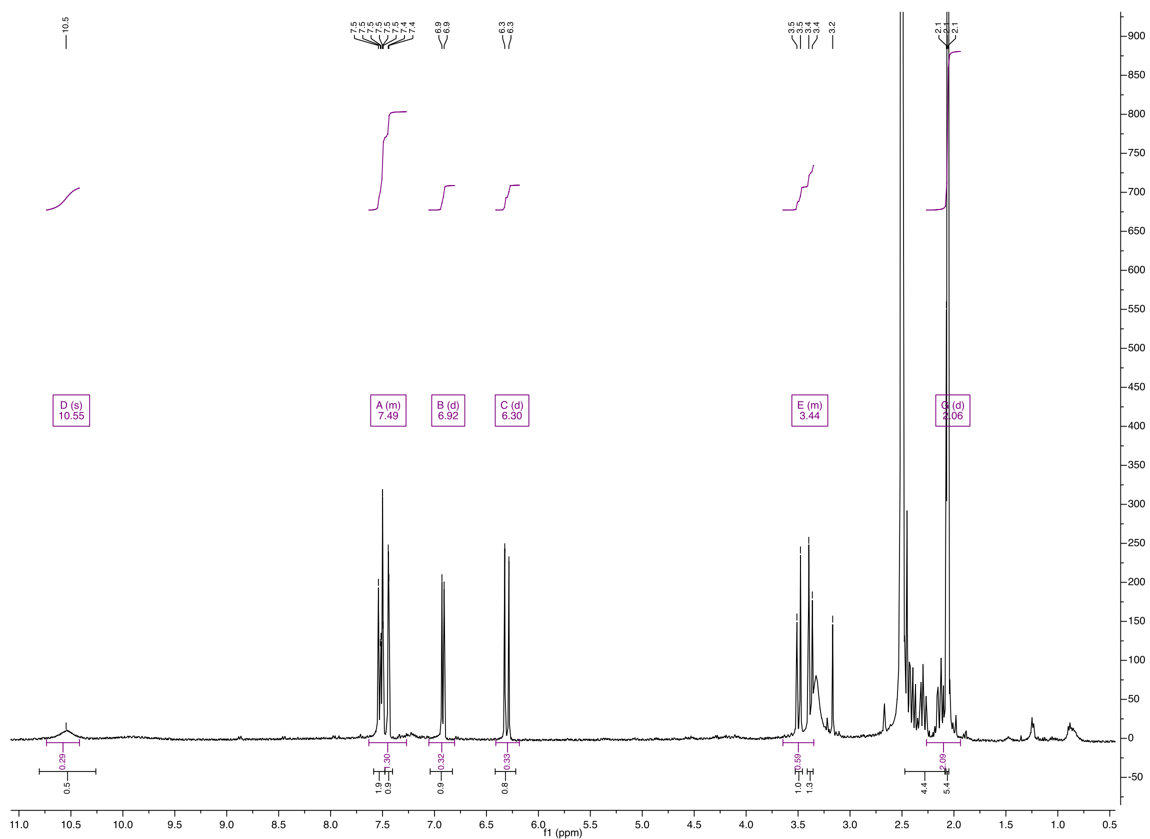


Figure S46. ^{1}H NMR of unknown compound, 400 MHz, DMSO-d6.

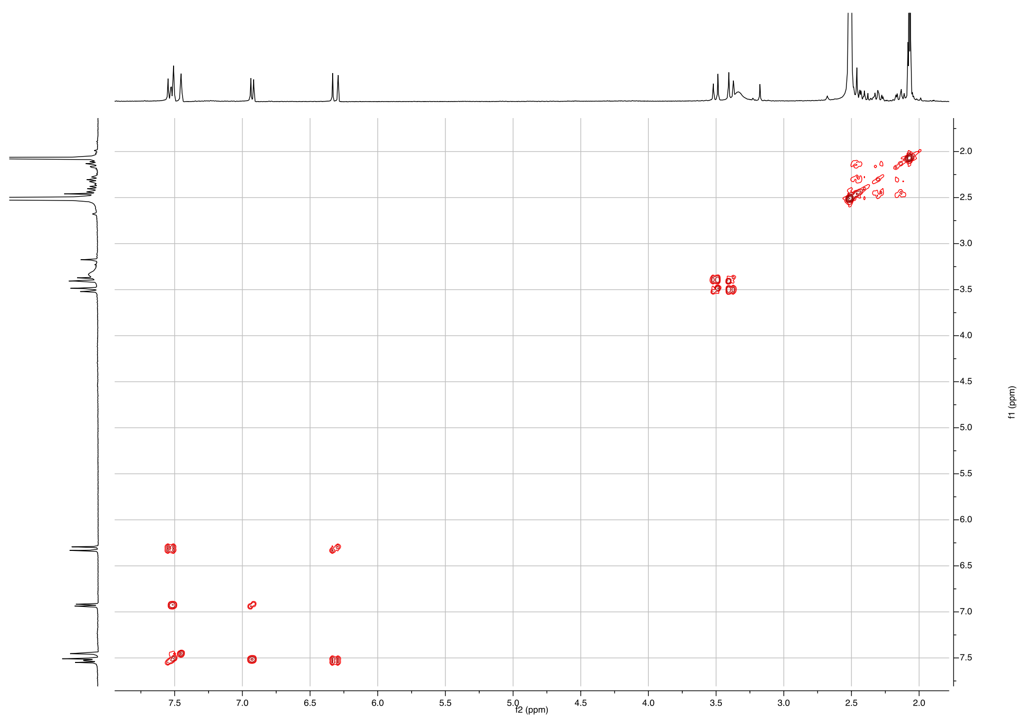


Figure S47. COSY spectrum of unknown compound.

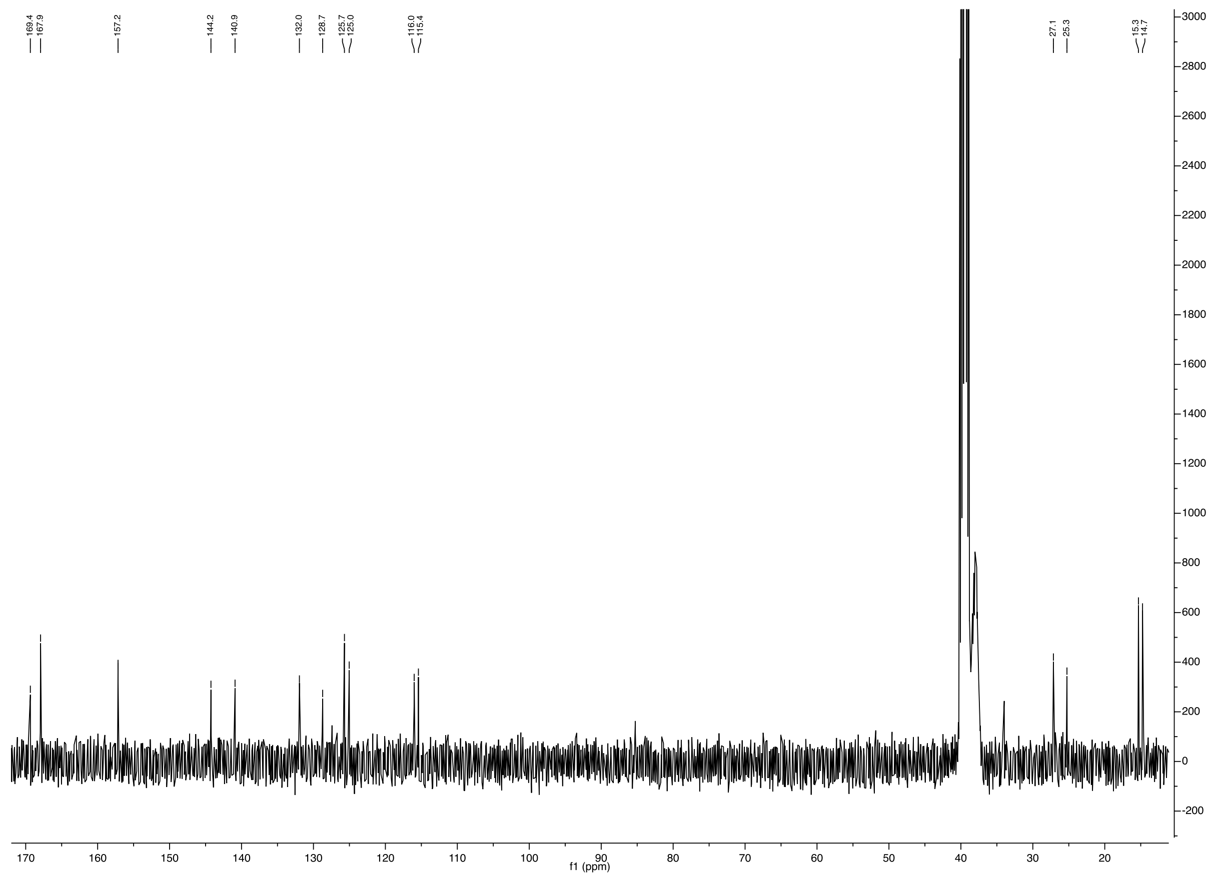


Figure S48. ^{13}C NMR spectrum of unknown compound, 100 MHz, DMSO-d6.

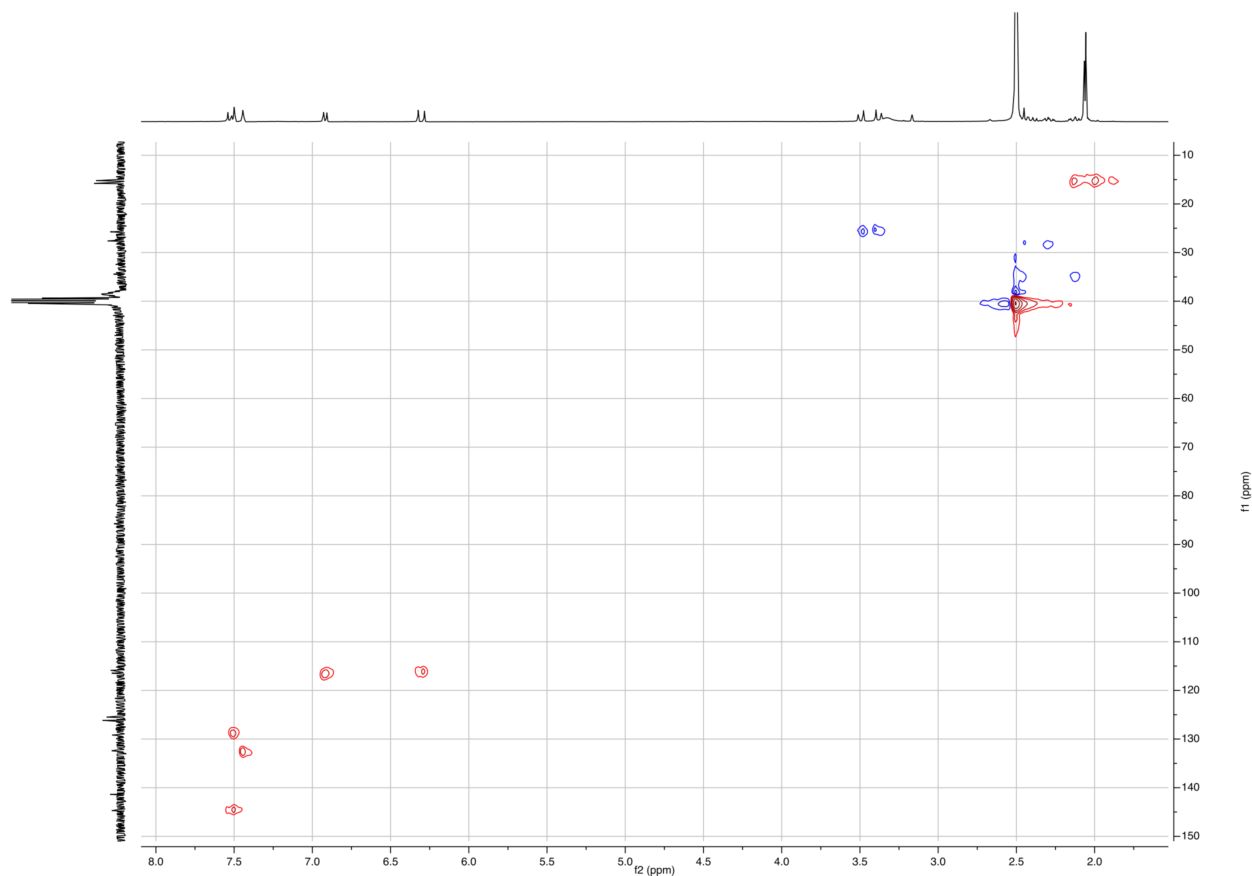


Figure S49. HSQC spectrum of unknown compound, DMSO-d6.

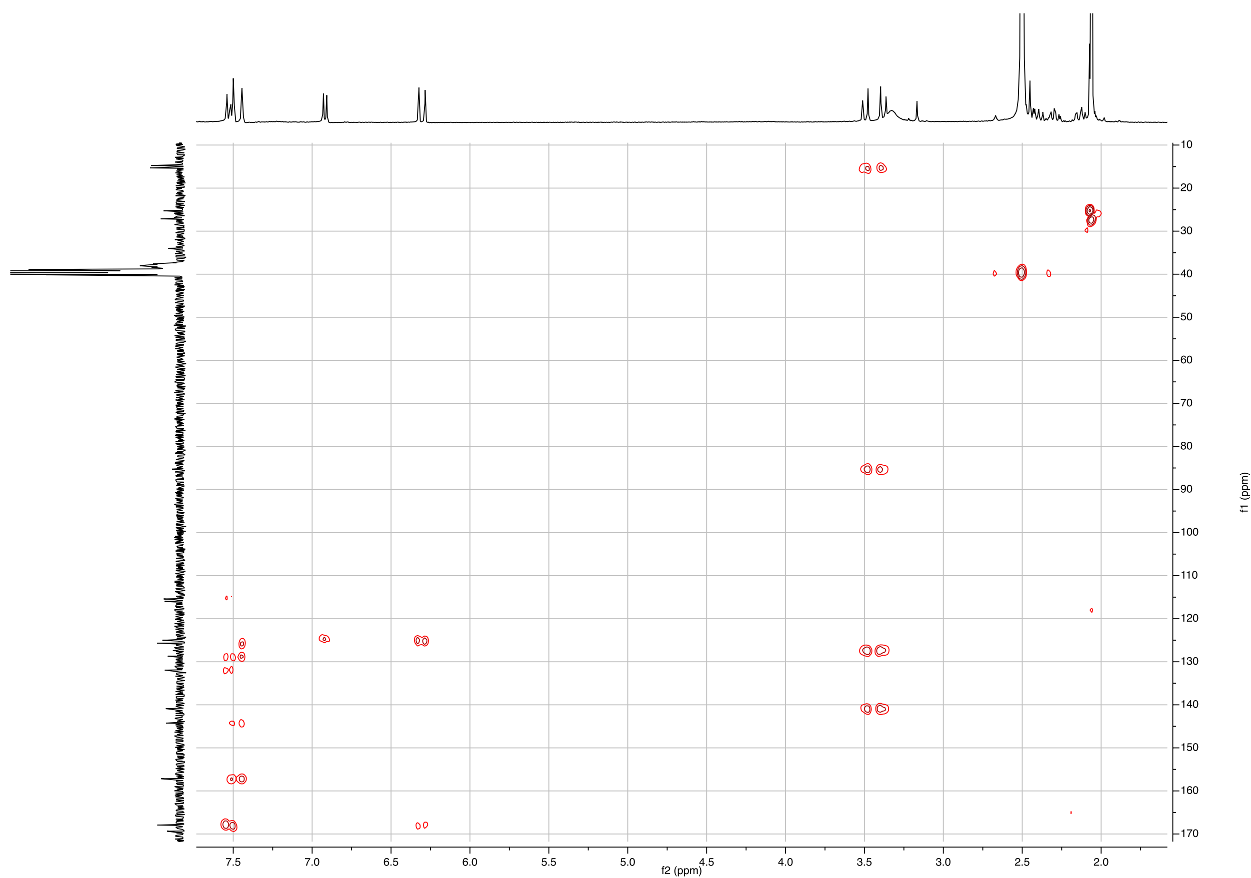


Figure S50. HMBC spectrum of unknown compound, DMSO-d6.

8. General Experimental Procedures

8.1 General Fermentation Procedures and Media

All liquid cultures were grown at 25 ˚C (for MF5453)/ 28 °C for *A. oryzae* NSAR1 and 200 rpm in a shaking incubator. The cultivation was carried out in 500 mL Erlenmeyer flasks filled with 100 mL of sterile medium.

ME Agar (growing MF5453): 2.00 % (*w/v*) Malt extract (Roth), 0.10 % (*w/v*) Peptone ex Soya (Roth), 2.00 % (*w/v*) D(+)-Glucose Monohydrate (Roth)

DPY Agar (growing NSAR1): 2.00 % (*w/v*) Dextrin from potato starch (Sigma Aldrich), 1.00 % (*w/v*) Polypeptone (Roth), 0.50 % (*w/v*) Yeast extract (Duchefa Biochemie, 0.50 % (*w/v*) Monopotassium phosphate (Roth), 0.05 % (*w/v*) Magnesium sulfate x H_2O (Roth), 2.50 % (*w/v*) Agar (Roth)

CM2 (producing medium): 2.00 % (*w/v*) Cottonseed flour (Sigma Aldrich), 10.0 % (*w/v*) Lactose Monohydrate (Roth), 2.00 % (*v/v*) Trace Element Solution, adjusted to pH 6.8

YMG Medium (producing medium): 0.4 % (*w/v*) D(+)-Glucose Monohydrate (Roth), 0.4 % (*w/v*) Yeast extract (Duchefa Biochemie), 1.0 % (*w/v*) Malt extract (Roth).

Trace Element Solution (TES): 0.10 % (*w/v*) Iron-(II)-sulfate heptahydrate (Sigma Aldrich), 0.0758 % (*w/v*) Manganese-II-sulfate tetrahydrate (Roth), 0.0025 % (*w/v*) Copper-II-chloride dihydrate (Roth), 0.01 % (*w/v*) Calcium chloride dihydrate (Roth), 0.0056 % (*w/v*) Boric acid (Roth), 0.0019 % (*w/v*) Ammonium molybdate tetrahydrate (Roth), 0.02 % (*w/v*) Zinc sulfate heptahydrate (Acros, dissolved in 0.6 M HCl)

GN Medium (overnight culture medium NSAR1 for transformation): 2.00 % (*w/v*) D(+)-Glucose Monohydrate (Roth), 1.00 % (*w/v*) Nutrient broth Nr. 2 from Oxoid (Fisher Scientific)

DPY Medium: 2.00 % (*w/v*) Dextrin from potato starch (Sigma Aldrich), 1.00 % (*w/v*) Polypeptone (Roth), 0.50 % (*w/v*) Yeast extract (Duchefa Biochemie, 0.50 % (*w/v*) Monopotassium phosphate (Roth), 0.05 % (*w/v*) Magnesium sulfate x H_2O (Roth)

CZD/S Medium (transformation medium NSAR1, arginine): 3.50 % Czapek Dox broth, 18.22 % D-Sorbitol (1 M), 0.10 % Ammonium sulphate, 0.05 % Adenine, 0.15 % L-Methionine

CZD/S1 Medium (transformation medium NSAR1, arginine): 3.50 % Czapek Dox broth, 18.22 % D-Sorbitol (1 M), 0.10 % Ammonium sulphate, 0.15 % L-Methionine

CZD/S1 Medium w/o methionine (transformation medium NSAR1, arginine): 3.50 % Czapek Dox broth, 18.22 % D-Sorbitol (1 M), 0.10 % Ammonium sulphate

CD1 Medium (NSAR1 selection medium): 3.50 % Czapek Dox broth, 4.68 % NaCl (0.8 M), 0.10 % Ammonium sulphate, 0.05 % Adenine, 0.15 % L-Methionine

CD2 Medium (NSAR1 selection medium): 3.50 % Czapek Dox broth, 4.68 % NaCl (0.8 M), 0.10 % Ammonium sulphate, 0.15 % *L*-Methionine

CD2 Medium w/o Methionine (NSAR1 selection medium): 3.50 % Czapek Dox broth, 4.68 % NaCl (0.8 M), 0.10 % Ammonium sulphate

LB (non producing): 0.5 % (*w/v*) Yeast extract (Duchefa Biochemie), 1.0 % (*w/v*) Tryptone (Duchefa Biochemie), 0.5 % (*w/v*) Sodium chloride (Honeywell).

8.2 Fermentaion of *Phoma*

Unidentified fungus MF5453 was grown on solid ME agar or in different liquid media (CM2, YMG, YMG+TES) at 25 °C for production of **1** or KO intermediates (200 rpm, one week, if liquid culture). For long term storage MF5453 mycelium of liquid culture was mixed with glycerol (50% glycerol final conc.).

8.3 Fermentation of *A. oryzae*

A. oryzae NSAR1[6] was grown on DPY agar plates (28 °C). For long term storage spores were collected from plates and mixed with glycerol (25% glycerol final conc.) and stored at −80 °C. For production of intermediates *A. oryzae* NSARI transformants were grown at 28 °C, 110 rpm for 5-7 days.

8.4 KO procedure in *Phoma*

Phoma MF5453 was grown in CM2 medium for three days at 25 °C and 250 rpm. Larger particles of cotton seed flour were removed by filtration through a coarse sieve. The fine mycelia was collected on a sterile filter paper under vacuum and washed with a 0.8 M aqueous solution of sodium chloride. 10 ml of filter-sterilized protoplasting solution (20 mg/ml lysing enzyme, Sigma L1412, 10 mg/ml of driselase in 0.8 M sodium chloride/10 mM sodium phosphate buffer, pH 6) was used for 1 g of cells. The digestion mixture was incubated at 30 °C for 3 hours with gentle shaking. The mixture was filtered through a sterile syringe with glass wool inside, followed by centrifugation at 4000g for 10 min. The pelleted protoplasts were washed with 0.8 M sodium chloride solution twice and subsequently with solution 1 (0.8 M sodium chloride solution, 10 mM $CaCl_2$, 50 mM TrisHCl pH 7.5). The cells were suspended in solution 1 and 1/5 of volume of solution 2 (PEG 4000 (60% *w/v*) 0.8 M sodium chloride solution, 10 mM $CaCl_2$, 50 mM TrisHCl pH 7.5, 50 mM $CaCl_2$). 0.2 ml for each transformation were mixed with both knock-out fragments (each 10 µl with 1 µg/µL) and placed on ice for 30 min. 1 ml of solution 2 were added at room temperature and incubated for 20 min. 10 ml soft agar (45 °C, 1 % agar containing 5 % NaCl) were poured in each tube, mixed gently and distributed on to two thin MEA pates with 50 µg/mL hygromycin B (37 °C pre-warmed). After incubation at 25 ˚C for 1 day the plates were overlaid with 10 mL MEA agar containing 150 µg/mL hygromycin B. After incubation at 25 ˚C for 5-10 days, 12 growing colonies were selected and cultivated on secondary plates with 100 µg/mL hygromycin B. In total 10 colonies were picked and transferred to liquid medium for subsequent fermentation, extraction and analysis by LC-MS.

8.5 Expression procedure in *A. oryzae*

A. oryzae NSAR1 conidia from a sporulating plates were inoculated into 100 ml GN medium in a 250 ml flask, which was incubated overnight for 28 °C, 180 rpm. The grown mycelia was collected by filtration through sterile Miracloth and was incubated in 10 ml protoplasting solution containing 20 mg/mL lysing enzyme, which was sterilized through a 0.45µm syringe filter, to bild protoplasts (lying falcon, 30 °C, 70 rpm, 3 h). The protoplasts were released from hyphal strands by gentle pipetting with wide–bore pipette. Afterwards the mycelia was collected by filtration through sterile miracloth and washed with NaCl (0.8 M). The supernatant was centrifuged (5 min, 3000 x g) for collecting protoplasts. Resulting supernatant was discarded and collected protoplasts were resuspended in solution 1 (100 µl per transformation) and 10 µl of purified plasmid DNA was added. The transformation

mixture was incubated on ice for 2 min before adding 1 ml solution 2 and incubating again for 20 min at room temperature. Afterwards pre–warmed 5 ml CZD/S (or CZD/S1, CZDS1-met) softagar (50 °C) was added to the mixture and overlaid onto prepared CZD/S (or CZD/S1, CZDS1-met) plates. Plates were incubated at 28°C for 3-5 days until colonies appeared. For selection on arginine (or adenine/methionine)–deficient media, the selection has to be repeated twice in order to prevent false positives due to consuming arginine (or adenine/methionine) from dead cell material. A few spores of the grown colony were picked with a tooth pick and transferred to a CD1 (or CD2/ CD2 w/o methionine) plate, without sorbitol. The plates were incubated at 28 °C for 3–4 days and the step was repeated by streaking out single colonies. In preparation for fermentation, a single colony was grown on DPY agar for one week.

8.6 Extraction Procedure

MF5453 cultures were clarified by Büchner filtration, supernatant acidified to pH 2 with concentrated H_2SO_4 and extracted twice with ethyl acetate (1:1). Combined organic layers were dried over $MgSO_4$ and solvent removed under vacuum. Extracts were dissolved in acetonitrile/H_2O (1:1) to a concentration of 10 mg/ml, filtered over glass wool and analysed or purified byLC/MS.

A. oryzae cultures were homogenized using a blender and stirred with an equal volume of ethyl acetate for 15 min. Homogenised cultures were clarified by Büchner filtration, supernatant was acidified to pH 2 with 2 M HCl and extracted twice with ethyl acetate (1:1). Combined organic layers were dried over $MgSO_4$ and solvent removed under vacuum. Extracts were dissolved in acetonitrile/H_2O (1:1) to a concentration of 10 mg/ml, filtered over glass wool and analysed by LC/MS.

8.7 NMR/HRMS instrumentation

NMR was obtained using a Bruker Avance 500 instrument equipped with a cryo-cooled probe at 500 MHz (^1H) and 125 MHz (^{13}C). 2D spectra (COSY, HSQC and HMBC) were obtained using standard parameters. Samples were dissolved in the indicated solvents. ^1H and ^{13}C spectra are referenecd relative to Me_4Si. All δ values are quoted in ppm and all J values in Hz. HRMS was obtained using a UPLC system (Acquity Ultraperformance, *Waters*) connected to a Q-TOF Premier massspectrometer.

9. References

1. A. Hildebrand, M. Remmert, A. Biegert and J. Söding, Proteins, 2009, 77 Suppl 9, 128–132.
2. X. Siebert, B. A. Eipper, R. E. Mains, S. T. Prigge, N. J. Blackburn and L. M. Amzel, *Biophys. J.*, 2005, 89, 3312–3319.
3. G. Harris, C. Dufresne, H. Joshua, L. Koch, D. Zink, P. Salmon, K. Göklen, M. Kurtz, D. Rew, J. Bergstrom, and K. Wilson, *Bioorg. Med. Chem. Lett.*, 1995, **5**, 2403–2408.
4. B. Bonsch, V. Belt, C. Bartel, N. Duensing, M. Koziol, C. M. Lazarus, A. M. Bailey, T. J. Simpson, R. J. Cox, *Chem. Commun.,* **2016**, *52*, 6777–6780.
5. P. A. Procopiou, E. J. Bailey, M. J. Bamford, A. P. Craven, B. W. Dymock, J. G. Houston, J. L. Hutson, B. E. Kirk, A. D. McCarthy, M. Sareen, J. J. Scicinski, P. J. Sharratt, M. A. Snowden, N. S. Watson and R. J. Williams, *J. Med. Chem*. **1994**, *37*, 3274 - 3281.
6. K. A. K. Pahirulzaman, K. Williams and C. M. Lazarus, *Meth. Enzymol.*, 2012, **517**, 241–260.