

# Supporting Information

## Genomic Discovery and Structure-Activity Exploration of a Novel Family of Enzyme-Activated Covalent Cyclin-Dependent Kinase Inhibitors

Jack R. Davison<sup>†</sup>, Michalis Hadjithomas<sup>†</sup>, Stuart P. Romeril<sup>†</sup>, Yoon Jong Choi<sup>†</sup>, Keith W. Bentley<sup>†</sup>, John B. Biggins<sup>†</sup>, Nadia Chacko<sup>†</sup>, M. Paola Castaldi<sup>†</sup>, Lawrence K. Chan<sup>†</sup>, Jared N. Cumming<sup>†</sup>, Thomas D. Downes<sup>†</sup>, Eric L. Eisenhauer<sup>†</sup>, Fan Fei<sup>†</sup>, Benjamin M. Fontaine<sup>†</sup>, Venkatesh Endalur Gopinarayanan<sup>†</sup>, Srishti Gurnani<sup>†</sup>, Audrey Hecht<sup>†</sup>, Christopher J. Hosford<sup>†</sup>, Ashraf Ibrahim<sup>†</sup>, Annika Jagels<sup>†</sup>, Camil Joubran<sup>†</sup>, Ji-Nu Kim<sup>†</sup>, John P. Lisher<sup>†</sup>, Daniel D. Liu<sup>†</sup>, James T. Lyles<sup>†</sup>, Matteo N. Mannara<sup>†</sup>, Gordon J. Murray<sup>†</sup>, Emilia Musial<sup>†</sup>, Mengyao Niu<sup>†</sup>, Roberto Olivares-Amaya<sup>†</sup>, Marielle Percuoco<sup>†</sup>, Susanne Saalau<sup>†</sup>, Kristen Sharpe<sup>†</sup>, Anjali V. Sheahan<sup>†</sup>, Neroshan Thevakumaran<sup>†</sup>, James E. Thompson<sup>†</sup>, Dawn A. Thompson<sup>†</sup>, Aric Wiest<sup>†</sup>, Stephen A. Wyka<sup>†</sup>, Jason Yano<sup>†</sup>, Gregory L. Verdine<sup>\* †‡</sup>

<sup>†</sup>LifeMine Therapeutics, 30 Acorn Park Drive, Cambridge, MA, 02140, USA

<sup>‡</sup>Departments of Chemistry and Chemical Biology, and Stem Cell and Regenerative Biology, Harvard University and Harvard Medical School, 12 Oxford Street, Cambridge, MA, 02138, USA

<sup>\*</sup>Corresponding author: email - gregory\_verdine@harvard.edu

## Table of Contents

<b>Supplementary Figures .....</b>	<b>S5</b>
Figure S1. Activity of roseopurpurin C <b>2</b> on <i>Preussia flanaganii</i> Pho85 ETaG chimera. ....	S5
Figure S2. Taxonomic assignment of the <i>ros</i> production strain. ....	S6
Figure S3. Effect of <i>ros</i> -cluster single gene knockouts on BGC expression measured by RNAseq differential expression .....	S7
Figure S4. Comparison of LC-MS/MS data for reference standards of <b>3</b> and <b>5</b> to fungal extracts from gene inactivation mutants .....	S8
Figure S5. LC-MS data from heterologous expression of <i>ros</i> pathway genes .....	S9
Figure S6. The proposed relationship of known roseopurpurins and aculeatusquinones to the <i>ros</i> cluster products <b>6</b> , <b>1</b> and <b>2</b> .....	S10
Figure S7. Mass spectrometric data supporting covalent mechanism of action. ....	S11
Figure S8. Mechanism of covalent modification of CDK2 by <b>2</b> at Lys33. ....	S12
Figure S9. Target profile of <b>2</b> in human cell lysate using biotinylated analog <b>50</b> in a competitive chemoproteomics pulldown assay. ....	S13
Figure S10. Kinobead profiling of <b>2</b> in <i>Saccharomyces cerevisiae</i> lysate. ....	S14

**Supplementary Tables .....S15**

Table S1. Functional analysis of <i>ros</i> gene cluster.....	S15
Table S2. NMR assignments of <b>1</b> in THF- <i>d</i> <sub>8</sub> (assigned), MeOD (assigned) and DMSO- <i>d</i> <sub>6</sub> (listed).....	S16
Table S3. NMR assignments of <b>2</b> in THF- <i>d</i> <sub>8</sub> (assigned), MeOD (assigned) and DMSO- <i>d</i> <sub>6</sub> (listed) .....	S17
Table S4. Protein crystallography data collection and refinement statistics.....	S18
Table S5. List of nucleic acids used for fungal strain engineering .....	S21
Table S6. List of Uniprot accessions used for CDK phylogenetic tree.....	S22

**MS and NMR Spectra of fungal compounds isolated in this study .....S23**

MS/MS spectrum of compound <b>1</b> .....	S23
<sup>1</sup> H NMR of compound <b>1</b> (600 MHz, THF- <i>d</i> <sub>8</sub> ).....	S23
<sup>13</sup> C NMR of compound <b>1</b> (150 MHz, THF- <i>d</i> <sub>8</sub> ).....	S24
HSQC of compound <b>1</b> (600 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S24
HMBC of compound <b>1</b> (600 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S25
MS/MS spectrum of compound <b>2</b> .....	S25
<sup>1</sup> H NMR of compound <b>2</b> (600 MHz, THF- <i>d</i> <sub>8</sub> ).....	S26
<sup>13</sup> C NMR of compound <b>2</b> (150 MHz, THF- <i>d</i> <sub>8</sub> ).....	S26
HSQC of compound <b>2</b> (600 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S27
HMBC of compound <b>2</b> (600 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S27
MS/MS spectrum of compound <b>8</b> .....	S28
<sup>1</sup> H-NMR of compound <b>8</b> (400 MHz, MeOD).....	S28
<sup>13</sup> C-NMR of compound <b>8</b> (100 MHz, MeOD).....	S29
HSQC of compound <b>8</b> (400 MHz, MeOD).....	S29
HMBC of compound <b>8</b> (400 MHz, MeOD).....	S30
HMBC correlations of isolated compound <b>8</b> .....	S30

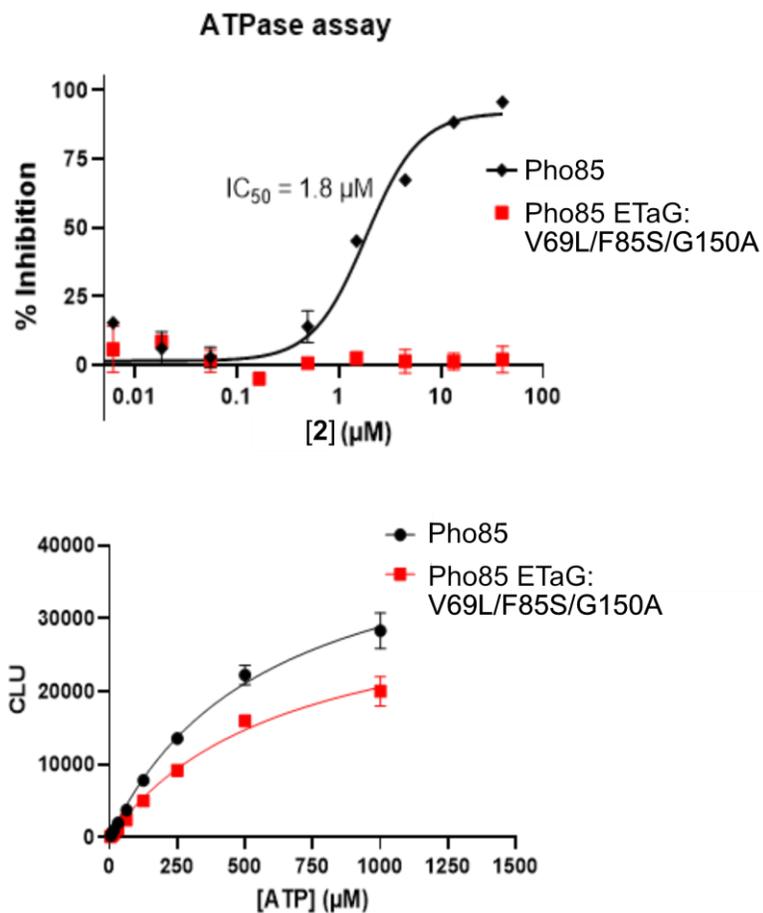
**MS and NMR spectra of compounds synthesized in this study.....S31**

<sup>1</sup> H NMR of compound ( <b>1</b> ) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S31
HPLC of compound ( <b>1</b> ) .....	S31
Chiral SFC of synthetic compound ( <b>1</b> ).....	S32
Chiral SFC of racemate ( <b>23</b> ). Reference for Chiral SFC of compounds <b>1</b> and <b>24</b> .....	S32
<sup>1</sup> H NMR of compound ( <b>2</b> ) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S34
LCMS of compound ( <b>2</b> ) .....	S34

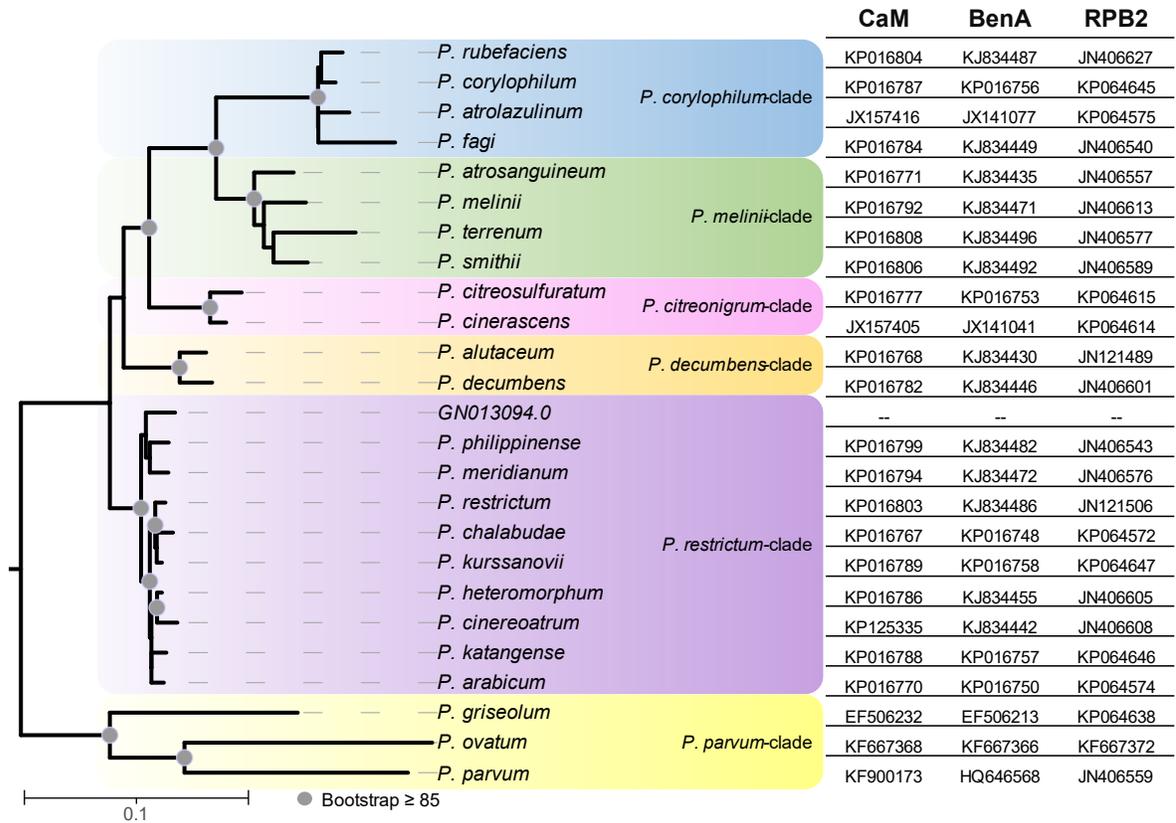
<sup>1</sup> H NMR of compound (24) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S35
HPLC of compound (24) .....	S35
Chiral SFC of compound (24) .....	S36
Chiral SFC of racemate (26). Reference for Chiral SFC of compounds 27 and 29.....	S37
<sup>1</sup> H NMR of compound (27) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S38
HPLC of compound (27) .....	S38
NOESY of compound (27) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S39
Chiral SFC of Compound (27) .....	S40
<sup>1</sup> H NMR of compound (28) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S41
LCMS of compound (28) .....	S41
NOESY of compound (28) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S42
<sup>1</sup> H NMR of compound (29) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S43
HPLC of compound (29) .....	S43
Chiral SFC of Compound (29) .....	S44
<sup>1</sup> H NMR of compound (38) (400 MHz, CDCl <sub>3</sub> ) .....	S45
LCMS of compound (38) .....	S45
<sup>1</sup> H NMR of compound (39) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S46
HPLC of compound (39) .....	S46
Chiral SFC of compound (39) .....	S47
Chiral SFC of racemic reference for Chiral SFC of compound 39. ....	S48
<sup>1</sup> H NMR of compound (40) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S49
LCMS of compound (40) .....	S49
<sup>1</sup> H NMR of compound (41) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S49
LCMS of compound (41) .....	S50
<sup>1</sup> H NMR of compound (42) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S51
LCMS of compound (42) .....	S51
<sup>1</sup> H NMR of compound (43) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S52
LCMS of compound (43) .....	S52
<sup>1</sup> H NMR of compound (44) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S53
LCMS of compound (44) .....	S53
<sup>1</sup> H NMR of compound (45) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S54
LCMS of compound (45) .....	S54
<sup>1</sup> H NMR of compound (47) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S55

LCMS of compound (47) .....	S55
<sup>1</sup> H NMR of compound (48) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S56
LCMS of compound (48) .....	S56
<sup>1</sup> H NMR of compound (49) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S57
LCMS of compound (49) .....	S57
<sup>1</sup> H NMR of compound (50) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S58
LCMS of compound (50) .....	S58
<sup>1</sup> H NMR of compound (51) (400 MHz, THF- <i>d</i> <sub>8</sub> ) .....	S59
LCMS of compound (51) .....	S59
<b>Small-molecule X-ray Crystallography Reports .....</b>	<b>S60</b>
X-ray crystallography report for compound 2 .....	S60
X-ray crystallography report for compound 43 (XC219).....	S73
<b>Supplementary Methods.....</b>	<b>S85</b>
Genomic search for CDK ETaGs .....	S85
Identification of putative resistance mutations in the ETaG .....	S85
Phylogenetic reconstruction .....	S86
Filamentous fungal engineering .....	S87
Yeast expression and assay of <i>rosA</i> .....	S89
<i>E. coli</i> expression and assay of <i>rosD</i> .....	S89
Expression of fungal Pho85 and human CDKs .....	S90
In vitro kinase functional assays .....	S93
Protein intact mass spectrometry.....	S94
Protein X-ray crystallography.....	S95
Fungal fermentation and GEM extraction .....	S95
RNA sequencing .....	S97
RNAseq expression analysis .....	S98
Human and fungal cell culture and lysis conditions for chemoproteomics experiments .....	S98
Preparation of Kinobead-Sepharose affinity matrix bead mix .....	S99
Kinobead chemoproteomics: Sample preparation, LC-MS analysis, and data processing.....	S100
Competitive chemoproteomics with biotin-probe: Sample preparation, LC-MS analysis, and data processing .....	S102
<b>References .....</b>	<b>S104</b>

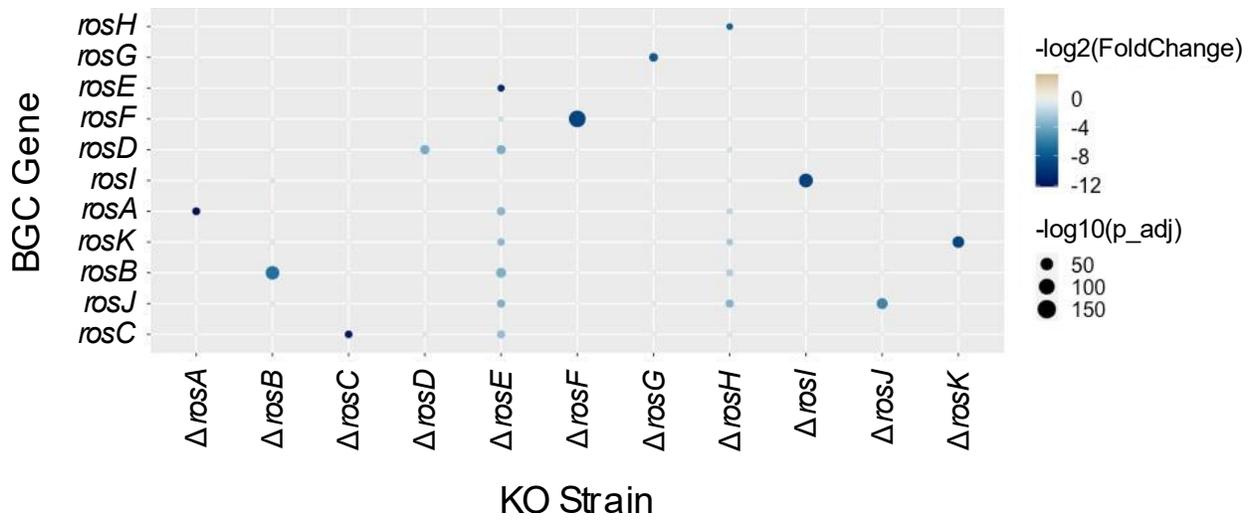
## Supplementary Figures



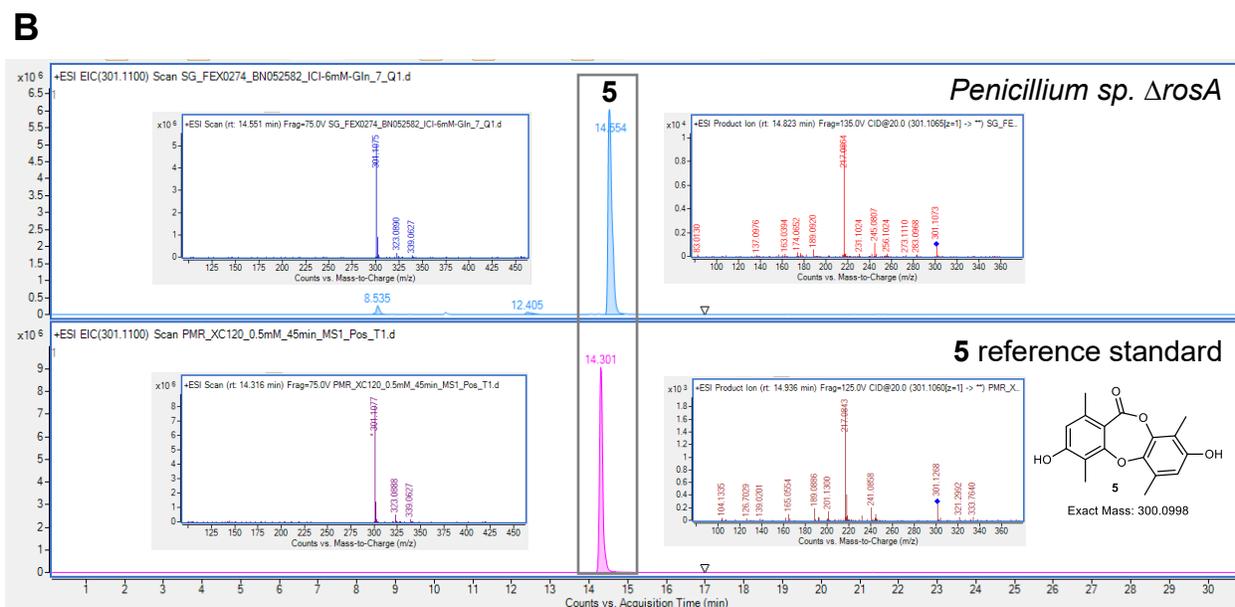
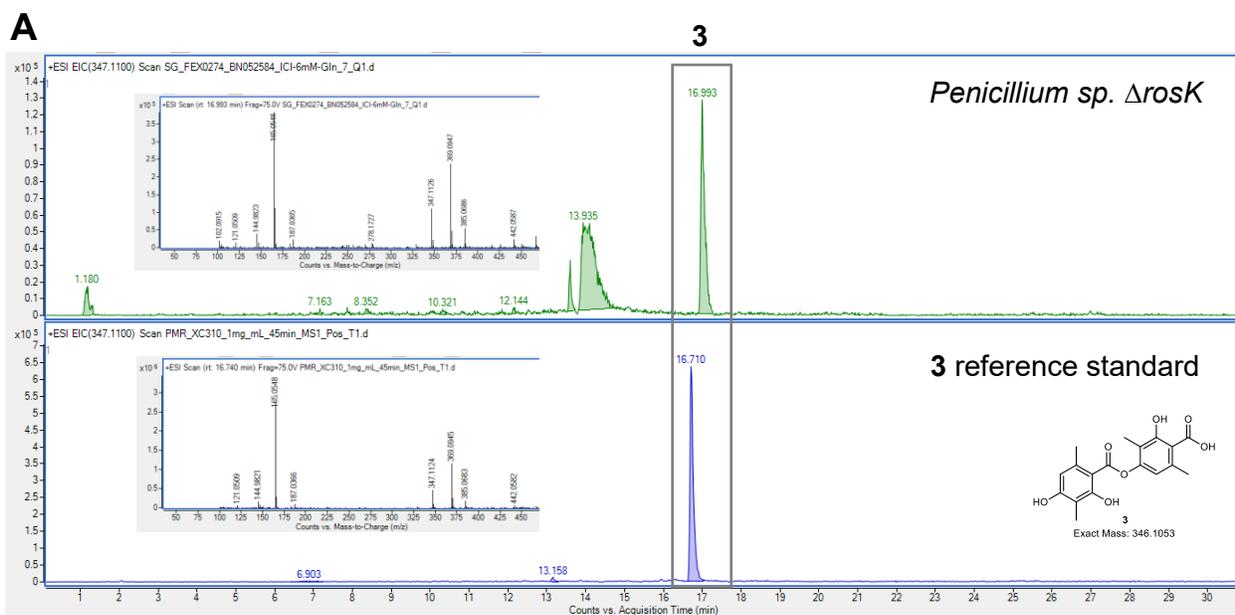
**Figure S1.** Activity of roseopurpurin C **2** on *Preussia flanaganii* Pho85:ETaG chimera. Displaying increased resistance to **2** in chimera vs WT Pho85, but comparable activity in the absence of inhibitor.



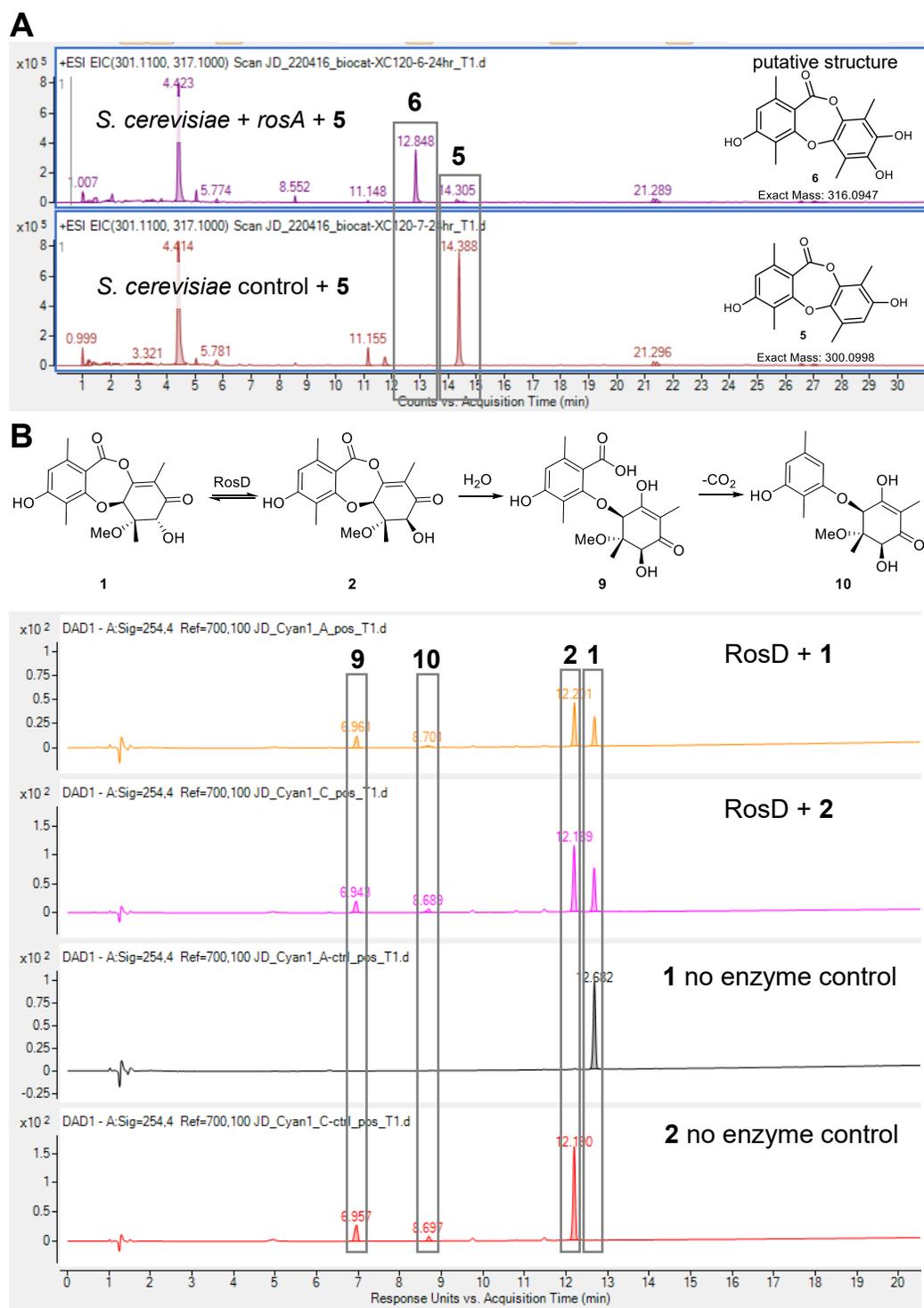
**Figure S2. Taxonomic assignment of the *ros* production strain.** Maximum likelihood phylogenetic analysis of concatenated marker genes placed the *ros* production strain *Penicillium* sp. (LifeBase accession GN013094.0) in the *P. restrictum* clade of *Penicillium* sect. Excilicaulis. NCBI accessions for the sequences used are listed in the inset table.



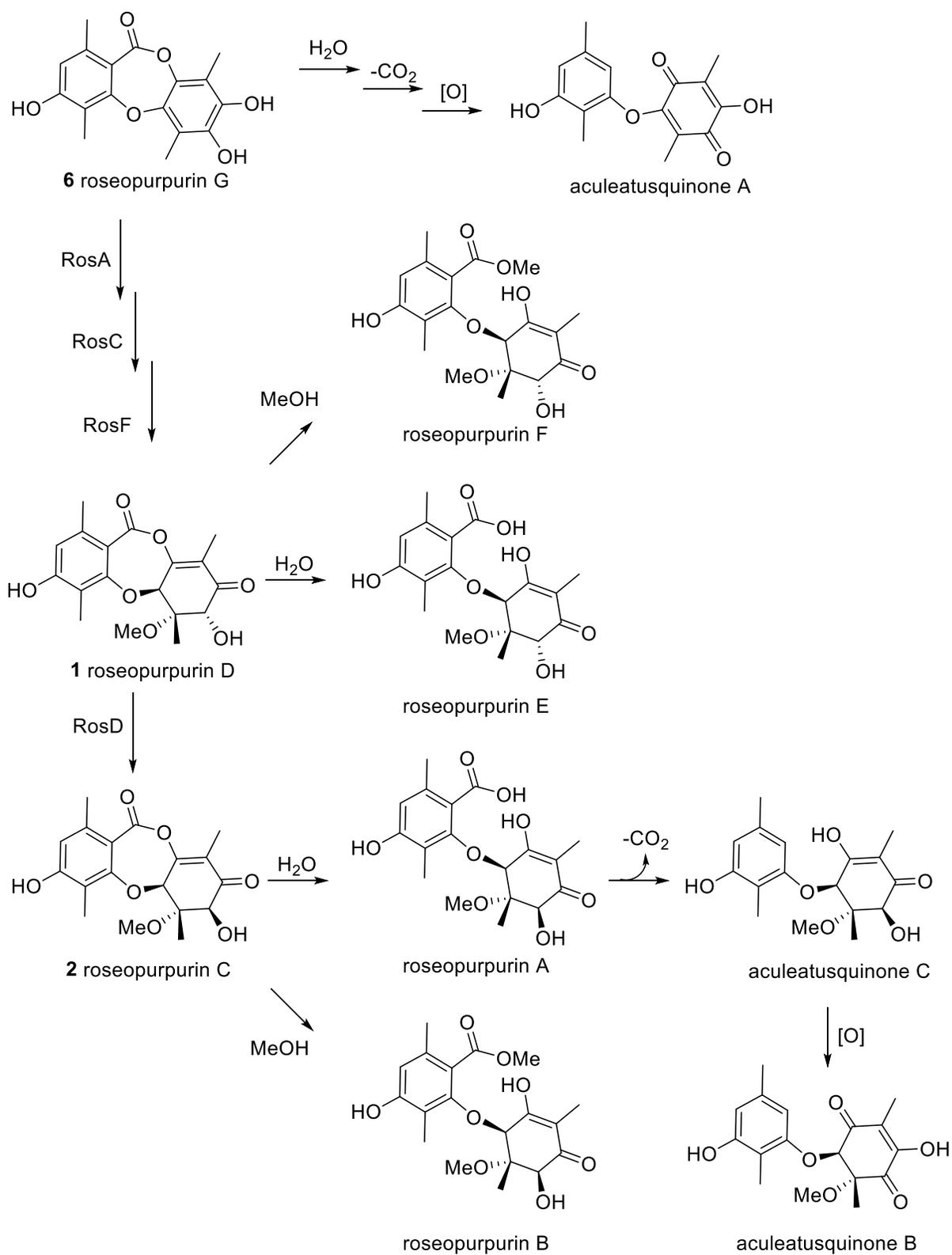
**Figure S3.** Effect of ros-cluster single gene knockouts on BGC expression measured by RNAseq differential expression. Fold change represented by color scale and significance ( $p_{adj}$ ) by marker size.



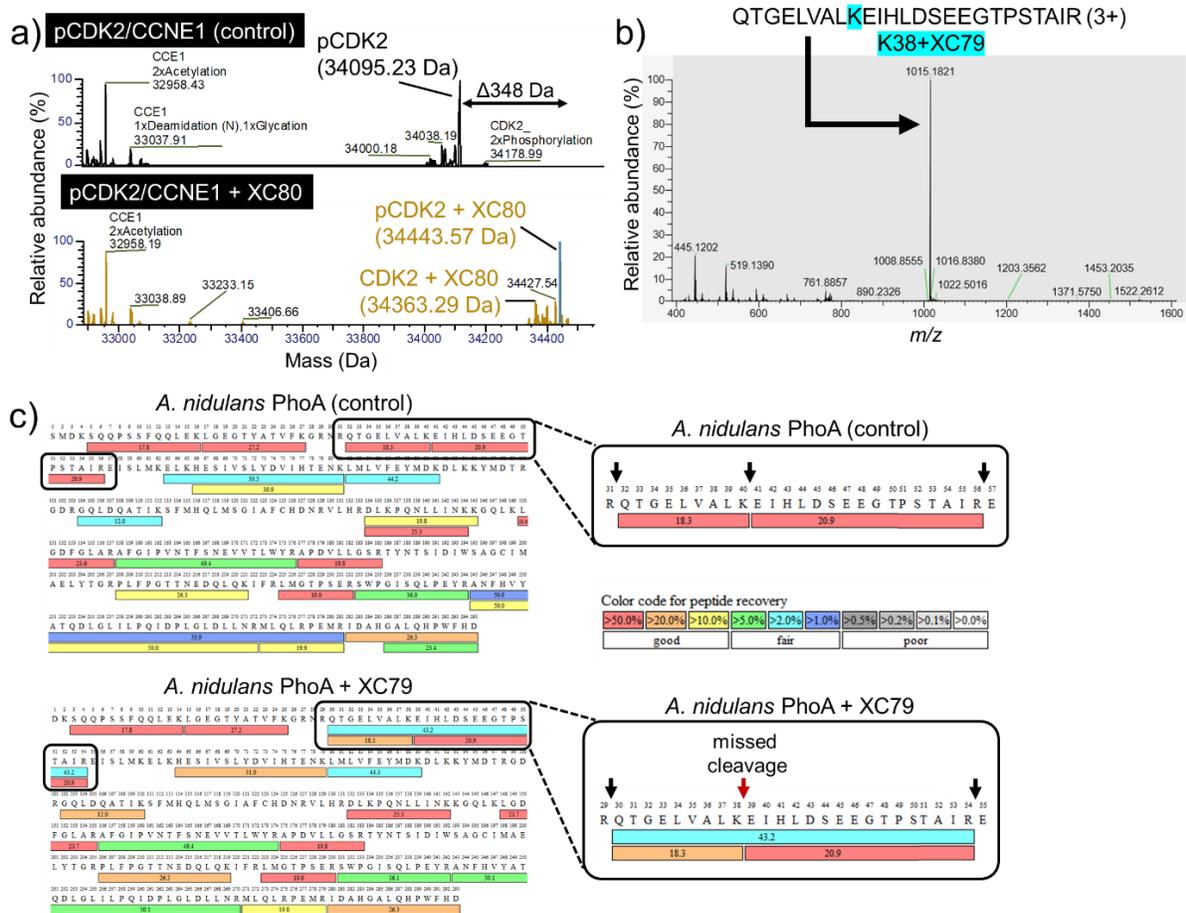
**Figure S4.** Comparison of LC-MS/MS data for reference standards of **3** and **5** to fungal extracts from gene inactivation mutants A)  $\Delta rosK$  and B)  $\Delta rosa$ .



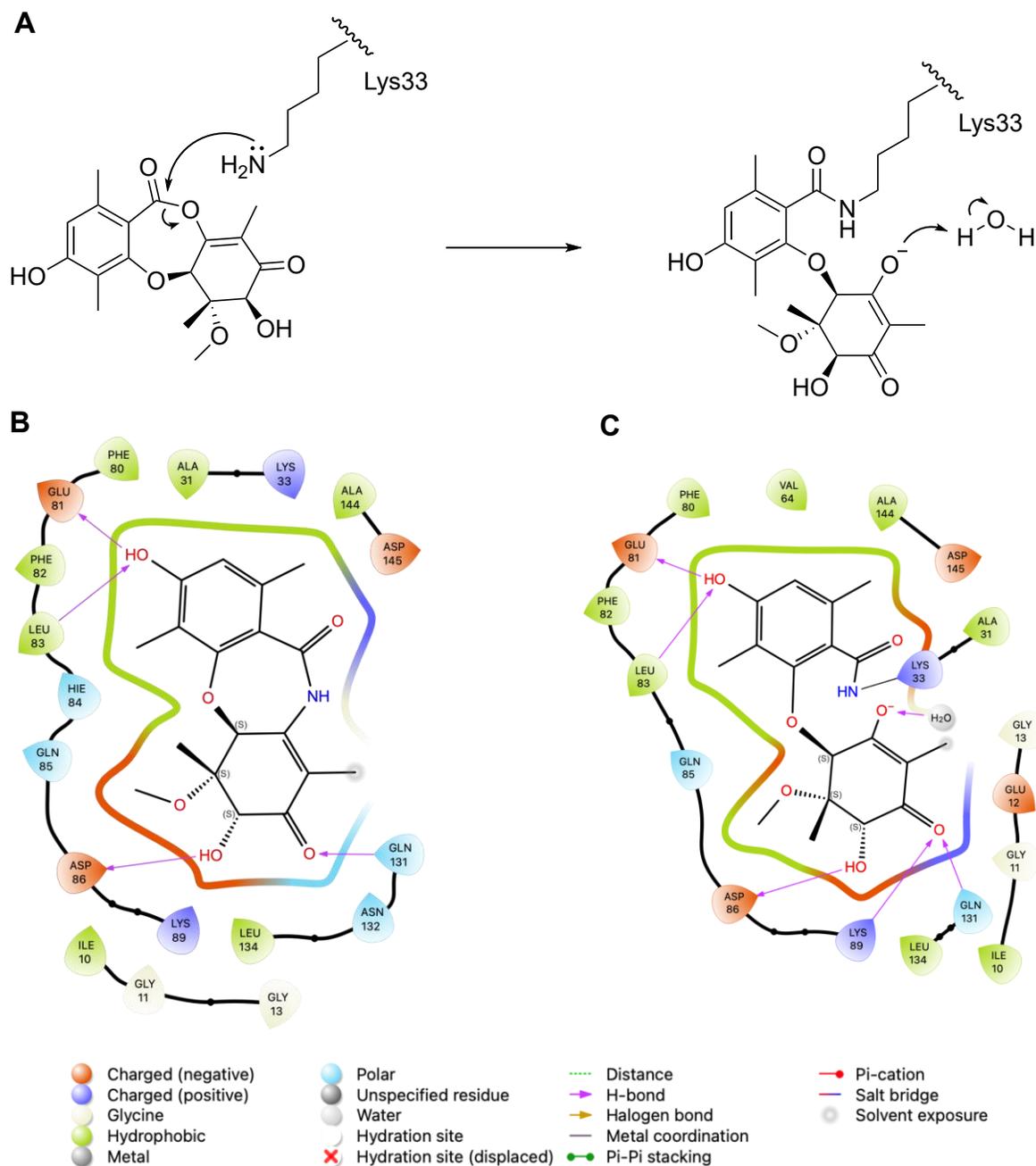
**Figure S5.** LC-MS data from heterologous expression of *ros* pathway genes. A) Whole-cell biotransformation by *rosA* expressed in *S. cerevisiae* showing conversion of **5** to **6**. B) In vitro assay of RosD from *E. coli*: epimerization of **1** and **2** was observed in the presence of enzyme; minor degradation to **9** and **10** was observed in reactions involving **2**, but not **1**, regardless of the presence or absence of enzyme.



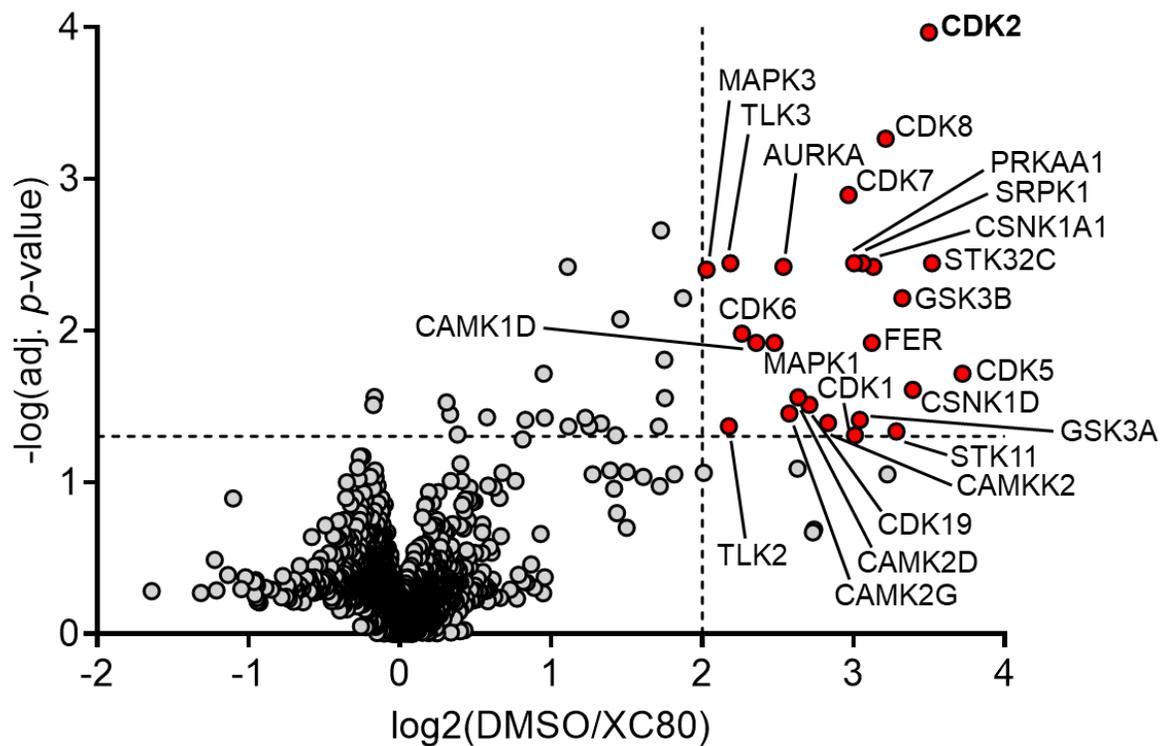
**Figure S6.** The proposed relationship of known roseopurpurins and aculeatusquinones to the *ros* cluster products **6**, **1** and **2**.



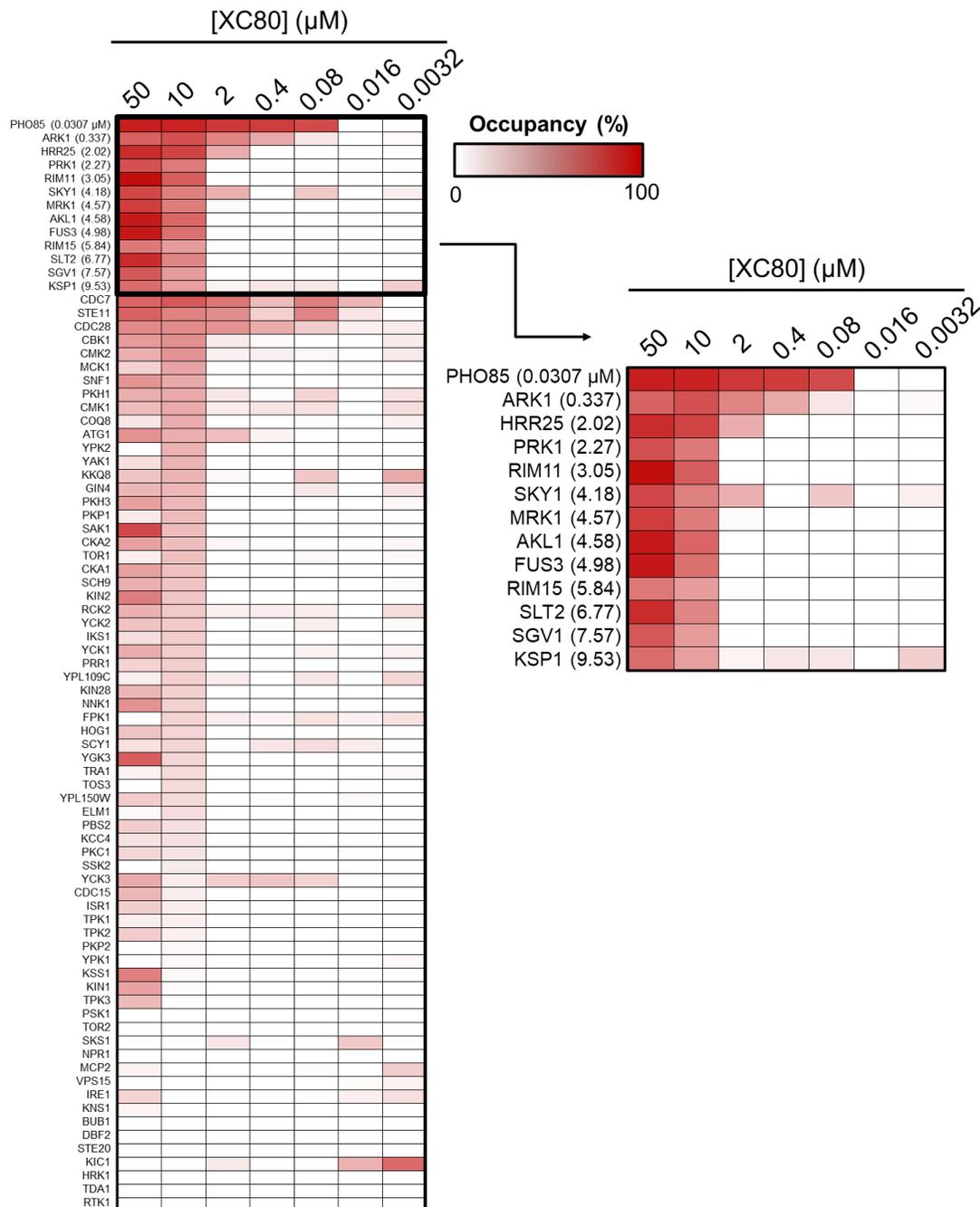
**Figure S7.** Mass spectrometric data supporting covalent mechanism of action. A) Intact MS spectrum of human CDK2/cyclin E1 (CCNE1) incubated with roseopurpurin **2** (XC80) showing the expected mass shift of 348 Da (bottom panel), relative to the CDK2/cyclin E1 spectrum in the absence of **2** (top panel). b) MS spectrum depicting the roseopurpurin D-modified peptide of *Aspergillus nidulans* PhoA following incubation of the intact protein with **1** (XC79) and subsequent tryptic digestion. c) Peptide maps depicting a missed tryptic cleavage site at Lys38 following incubation of **1** with intact *A. nidulans* PhoA and subsequent tryptic digestion (lower panel, red arrow and cyan bar). This missed cleavage event was not detected in the absence of **1** (upper panel), providing further evidence of covalent modification of Lys38 by **1**. No additional missed cleavages were detected in **1**-treated PhoA, consistent with site-specific modification of Lys38. Arrows denote tryptic cleavage sites, peptide map colors depict peptide recovery, and the number in the box denotes peptide retention time.



**Figure S8.** Mechanism of covalent modification of CDK2 by **2** at Lys33. A) Scheme showing nucleophilic attack of the amine group of Lys33 on the vinylogous anhydride of **2**. B) 2D interaction map of the active site of CDK2 in complex with **51**, showing the pre-reaction binding mode (generated from PDB:9BJC by Schrödinger Maestro). C) 2D interaction map of the active site of CDK2 in the post-reaction complex with **2** (generated from PDB:9BJB by Schrödinger Maestro).

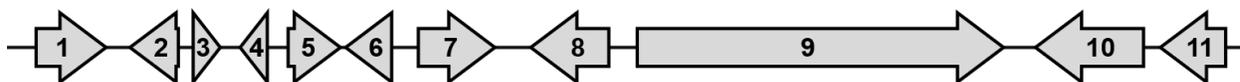


**Figure S9.** Target profile of **2** (XC80) in human cell lysate using biotinylated analog **50** in a competitive chemoproteomics pulldown assay. Statistically significant target proteins of **2** with >75% occupancy are colored in red.



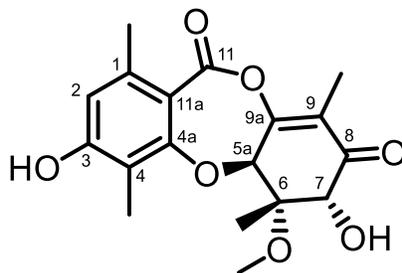
**Figure S10.** Kinobead profiling of roseopurpurin C 2 (XC80) in *S. cerevisiae* lysate. All detected kinases are shown in the heatmap, with EC<sub>50</sub> binding values for target kinases (in μM) listed in parentheses. Data for target kinases of **2** are shown in the inset. Data were filtered to identify kinases using the BioGRID *S. cerevisiae* kinome (<https://thebiogrid.org/project/2>).

## Supplementary Tables



Gene	Best hit (NCBI nr)	% id to nr hit (qcov)	Best hit (SwissProt)	% id to SP hit (qcov)	Domain hit (CDD)	Putative class
<b>1 <i>rosF</i></b>	XP_033540844.1	94 (89)	Q9UQY0.2	35 (86)	pfam00891	O-methyltransferase
<b>2 <i>rosG</i></b>	KAH8900317.1	72 (95)	Q6C7U8.1	69 (94)	cd07836	ST Kinase Pho85
<b>3 <i>rosE</i></b>	XP_031872704.1	44 (91)	-	-	-	hypothetical
<b>4 <i>rosD</i></b>	XP_033540866.1	81 (98)	-	-	cd16361	VOC superfamily
<b>5 <i>rosC</i></b>	XP_033540865.1	87 (97)	Q9UUN9.3	36 (90)	cd05227	extended-SDR dehydrogenase / reductase
<b>6 <i>rosB</i></b>	XP_033540864.1	89 (100)	G1XU02.1	40 (99)	pfam04909	amidohydrolase superfamily
<b>7 <i>rosA</i></b>	XP_033540863.1	84 (99)	A0A3Q9FEJ4.1	40 (95)	cd11058	cytochrome P450 CYP60B
<b>8 <i>rosK</i></b>	XP_033540862.1	86 (98)	A0A3Q9FEJ4.1	36 (99)	cd11058	cytochrome P450 CYP60B
<b>9 <i>rosJ</i></b>	XP_033540861.1	74 (99)	A0A8F4SKJ7.1	46 (98)	cd00833	polyketide synthase
<b>10 <i>rosI</i></b>	XP_033540860.1	81 (94)	F2SH39.2	52 (91)	cd17502	MFS transporter
<b>11 <i>rosH</i></b>	XP_033540859.1	54 (99)	-	-	smart00066	Zn2Cys6-domain transcription factor

**Table S1.** Functional analysis of the *ros* gene cluster. Top homologs identified (with > 30% identity / 85% query coverage) in NCBI nr database or the SwissProt database of curated and experimentally validated sequences, conserved domains and putative protein class identified by NCBI CD-Search.



Posn	This Study				HMBC	Literature <sup>1</sup>	
	$\delta^1\text{H}$ (mult., J (Hz)) in THF- <i>d</i> <sub>8</sub> , 600 MHz	$\delta^{13}\text{C}$ in THF- <i>d</i> <sub>8</sub> , 150 MHz	$\delta^1\text{H}$ (mult., J (Hz)) in MeOD, 500 MHz	$\delta^{13}\text{C}$ in MeOD, 125 MHz		$\delta^1\text{H}$ (mult., J (Hz)) in MeOD	$\delta^{13}\text{C}$ in MeOD
1	-	141.1	-	141.7	-	-	141.7
2	6.43 (s, 1H)	113.5	6.49 (s, 1H)	113.9	1-Me, 11a, 4, 3	6.49 (s, 1H)	113.9
3	-	160.3	-	161.0*	-	-	157.4
4	-	112.3	-	112.8	-	-	112.8
4a	-	157.1	-	157.4*	-	-	161.0
5a	4.90 (s, 1H)	78.8	4.98 (s, 1H)	78.7	6-Me, 7, 6, 9, 9a	4.98 (s, 1H)	78.7
6	-	81.0	-	81.5	-	-	81.5
7	4.37 (s, 1H)	77.2	4.39 (s, 1H)	77.3	6-Me, 6-OMe, 6, 8	4.39 (s, 1H)	77.3
8	-	197.3	-	198.3	-	-	198.3
9	-	117.9	-	119.3	-	-	119.3
9a	-	156.9	-	157.1	-	-	157.1
11	-	162.9	-	164.2	-	-	164.2
11a	-	109.9	-	109.8	-	-	109.8
1-Me	2.36 (s, 3H)	22.8	2.37 (s, 3H)	22.8	11a, 2, 1	2.37 (s, 3H)	22.7
3-OH	-	-	-	-	-	-	-
4-Me	2.08 (s, 3H)	9.0	2.08 (s, 3H)	8.9	4, 4a, 3	2.08 (s, 3H)	8.9
6-Me	1.66 (s, 3H)	16.3	1.67 (s, 3H)	16.4	7, 5a, 6	1.67 (s, 3H)	16.4
6-OMe	3.33 (s, 3H)	52.3	3.34 (s, 3H)	52.1	6	3.34 (s, 3H)	52.1
7-OH	-	-	-	-	-	-	-
9-Me	1.73 (s, 3H)	7.9	1.76 (s, 3H)	7.9	9, 9a, 8	1.76 (s, 3H)	7.9

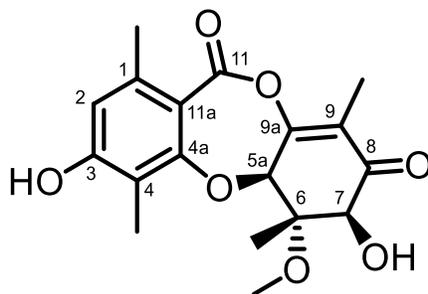
\* = Assignments of these two signals differ from literature reported assignments due to the HMBC correlation shown between H-2 (6.49 ppm, <sup>1</sup>H) and C-3 (161.0 ppm, <sup>13</sup>C) but not shown between H-2 (6.49 ppm, <sup>1</sup>H) and C-4a (157.4 ppm, <sup>13</sup>C).

NMR data for **1** (500MHz and 125 MHz) in DMSO-*d*<sub>6</sub><sup>2</sup>:

<sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  10.37 (s, 1H), 6.53 (s, 1H), 5.41 (d, J = 5.8 Hz, 1H), 5.12 (d, J = 0.9 Hz, 1H), 4.26 (d, J = 5.8 Hz, 1H), 3.23 (s, 3H), 2.30 (s, 3H), 1.99 (s, 3H), 1.65 (s, 3H), 1.52 (s, 3H).

<sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  196.5, 162.2, 159.4, 155.8, 155.1, 139.7, 117.5, 112.8, 111.0, 108.2, 80.0, 77.0, 75.4, 51.3, 22.3, 15.7, 8.6, 7.8.

**Table S2.** NMR Assignments of **1** in THF-*d*<sub>8</sub> (assigned), MeOD (assigned) and DMSO-*d*<sub>6</sub> (listed)



Posn	This Study				HMBC	Literature <sup>1</sup>	
	$\delta^1\text{H}$ (mult., J (Hz)) in THF- <i>d</i> <sub>8</sub> , 600 MHz	$\delta^{13}\text{C}$ in THF- <i>d</i> <sub>8</sub> , 150 MHz	$\delta^1\text{H}$ (mult., J (Hz)) in MeOD, 500 MHz	$\delta^{13}\text{C}$ in MeOD, 125 MHz		$\delta^1\text{H}$ (mult., J (Hz)) in MeOD	$\delta^{13}\text{C}$ in MeOD
1	-	143.8	-	144.5	-	-	144.5
2	6.52 (s, 1H)	115.3	6.58 (s, 1H)	115.7	4-Me, 1-Me, 11a, 4, 3	6.59 (s, 1H)	115.7
3	-	161.3	-	162.5	-	-	162.5
4	-	115.2	-	115.5	-	-	115.5
4a	-	161.7	-	161.2	-	-	161.2
5a	5.05 (q, J = 2Hz, 1H)	83.0	5.23 (q, J = 2.0 Hz, 1H)	82.0	6, 9, 4a, 9, 6-Me	5.24 (d, J = 2.0 Hz)	82.0
6	-	84.5	-	84.9	-	-	84.9
7	4.24 (s, 1H)	77.7	4.39 (s, 1H)	76.5	6-Me, 5a, 6, 8	4.40 (s, 3H)	76.4
8	-	197.9	-	198.9	-	-	198.9
9	-	117.3	-	118.6	-	-	118.7
9a	-	162.5	-	162.7	-	-	162.7
11	-	162.2	-	162.9	-	-	162.9
11a	-	112.5	-	112.5	-	-	112.5
1-Me	2.44 (s, 3H)	22.5	2.44 (s, 3H)	22.4	11a, 2, 1	2.45 (s, 3H)	22.3
3-OH	-	-	-	-	-	-	-
4-Me	2.18 (s, 3H)	8.7	2.18 (s, 3H)	8.6	4, 3, 4a	2.18 (s, 3H)	8.6
6-Me	1.15 (s, 3H)	12.4	1.14 (s, 3H)	13.3	7, 5a, 6	1.14 (s, 3H)	13.3
6-OMe	3.31 (s, 3H)	51.4	3.41 (s, 3H)	51.1	6	3.41 (s, 3H)	51.1
7-OH	-	-	-	-	-	-	-
9-Me	1.86 (d, J = 2Hz, 3H)	8.6	1.88 (d, J = 2.0 Hz, 3H)	8.6	9, 9a, 8	1.88 (d, J = 2.0)	8.6

NMR data for **2** (500MHz and 125 MHz) in DMSO-*d*<sub>6</sub>.<sup>2</sup>

<sup>1</sup>H NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  10.59 (s, 1H), 6.61 (s, 1H), 5.51 (d, J = 5.1 Hz, 1H), 5.34 (q, J = 1.9 Hz, 1H), 4.31 (d, J = 5.1 Hz, 1H), 3.27 (s, 3H), 2.37 (s, 3H), 2.07 (s, 3H), 1.76 (d, J = 2.0 Hz, 3H), 1.00 (s, 3H).

<sup>13</sup>C NMR (125 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  197.5, 161.1, 160.7, 160.1, 159.1, 142.0, 117.1, 114.5, 113.6, 110.9, 82.7, 80.3, 74.6, 50.1, 21.8, 13.3, 8.4, 8.3.

**Table S3.** NMR Assignments of **2** in THF-*d*<sub>8</sub> (assigned), MeOD (assigned) and DMSO-*d*<sub>6</sub> (listed)

	CDK2 / cyclin E1 + roseopurpurin C	CDK2 / cyclin E1 + XC208	CDK2 / cyclin E1 + XC219
<b>PDB ID</b>	9BJB	9BJC	9BJD
<b>Data collection</b>			
Space group	P 3 <sub>2</sub> 2 1	P 3 <sub>1</sub> 2 1	P 3 <sub>2</sub> 2 1
Cell dimensions			
a, b, c (Å)	90.00, 90.00, 120.00	90.00, 90.00, 120.00	90.00, 90.00, 120.00
$\alpha, \beta, \gamma$ (°)	76.10, 76.10, 255.65	76.32, 76.32, 258.85	76.04, 76.04, 258.65
Resolution (Å)	63.82 – 2.14 (2.20 – 2.14)	43.14 – 2.22 (2.28 – 2.22)	258.65 – 2.12 (2.18 – 2.12)
R <sub>merge</sub>	0.089 (0.628)	0.217 (2.775)	0.146 (2.269)
CC <sub>1/2</sub>	0.992 (0.813)	0.986 (0.414)	0.993 (0.478)
I / $\sigma$ I	11.2 (2.5)	6.8 (1.1)	7.4 (1.1)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Redundancy	7.3 (6.8)	11.0 (10.9)	10.5 (10.9)
<b>Refinement</b>			
Resolution (Å)	63.82 – 2.14	43.12 – 2.22	52.34 – 2.12
No. reflections	48,404	44,309	50,279
R <sub>work</sub> / R <sub>free</sub>	0.1939/0.2253	0.1940/0.2131	0.2022/0.2342
No. atoms			
Protein	4,914	4,860	4,629
Ligand/ion	25	25	52
Water	235	193	98
B-factors			
Protein	63.40	69.00	73.38
Ligand/ion	66.26	78.96	95.33
Water	48.93	52.69	58.90
R.m.s. deviations			
Bond lengths (Å)	0.007	0.008	0.004
Bond angles (°)	0.827	0.895	0.647

**Table S4.** Protein crystallography data collection and refinement statistics.





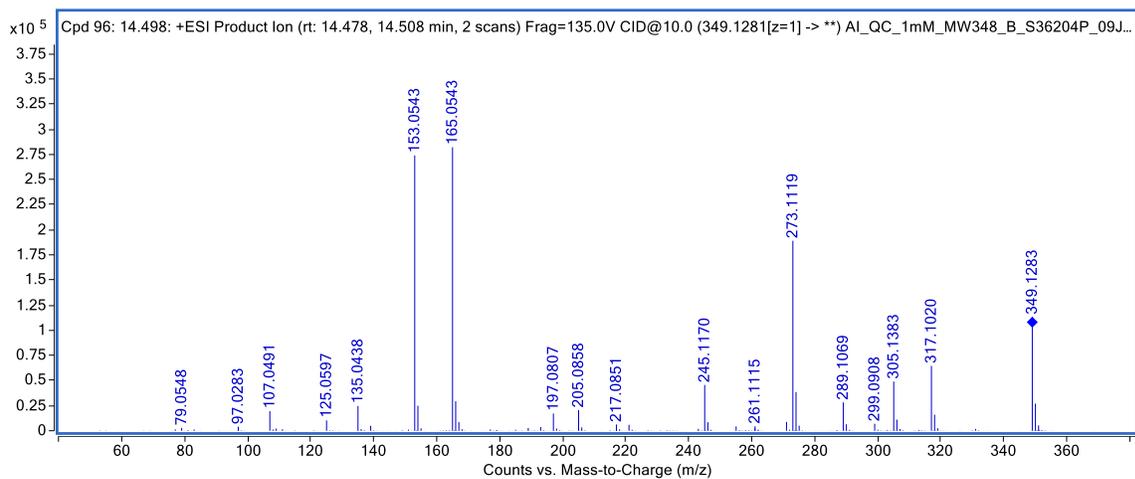
<b>PC00635</b>	CTCGATGTCCTCGACGGTCAG	Split marker Nat 5' R
<b>PC00632</b>	CAGCGAGAGCCTGACCTATTGC	Split marker Hyg 3' F
<b>PC00634</b>	TCTTGACGACACGGCTTACCG	Split marker Nat 3' F
<b>PC00631</b>	CGCCAAGCTCGAAATTAACCC	Split marker all 3' R
<b>diag1967</b>	gaacctctctgctctgggtc	Diagnostic PCR TF-OE 5'
<b>diag2454</b>	gccacataaaccactgg	Diagnostic PCR TF-OE 5'
<b>diag1968</b>	cttctcatcacactactcg	Diagnostic PCR TF-OE 3'
<b>diag2455</b>	ggcagcagggttcaaacc	Diagnostic PCR TF-OE 3'
<b>diag2742</b>	ggatcatgacaggaggacaagg	Diagnostic PCR PKS-KO 5'
<b>diag2743</b>	caggcgcaatcgaactatggg	Diagnostic PCR PKS-KO 5'
<b>diag2744</b>	catctgctctctctgggtgg	Diagnostic PCR PKS-KO 3'
<b>diag2745</b>	tgaggaaatcgctcttggctcc	Diagnostic PCR PKS-KO 3'
<b>sg1-g1</b>	GCACGTACAGTCCCACCAAG	$\Delta rosF$ sgRNA
<b>sg2-g1</b>	taactacttgaaccctgg	$\Delta rosF$ sgRNA
<b>sg1-g2</b>	TTCGCATGTCAAATGCTGCG	$\Delta rosG$ sgRNA
<b>sg2-g2</b>	TGTTCACTTATAATAAGACC	$\Delta rosG$ sgRNA
<b>sg1-g3</b>	CCTCATACTGGACTTGATTG	$\Delta rosE$ sgRNA
<b>sg2-g3</b>	TGCATCTAGGATGTCCAACA	$\Delta rosE$ sgRNA
<b>sg1-g4</b>	AGTGATGTTCGGGATCCGT	$\Delta rosD$ sgRNA
<b>sg2-g4</b>	CACCAGCTGCAATCACAGCT	$\Delta rosD$ sgRNA
<b>sg1-g5</b>	TTTCTCCGTGCGGTGCAAG	$\Delta rosC$ sgRNA
<b>sg2-g5</b>	GGCTCCCTTATCCTAATAAC	$\Delta rosC$ sgRNA
<b>sg1-g6</b>	AACCGATAATGTCTTTGCTG	$\Delta rosB$ sgRNA
<b>sg2-g6</b>	CTCAAGGAGTTCGCGAAACC	$\Delta rosB$ sgRNA
<b>sg1-g7</b>	TACTACAATGGCGCCACCT	$\Delta rosA$ sgRNA
<b>sg2-g7</b>	AATTTTCGTAGAATAGGACC	$\Delta rosA$ sgRNA
<b>sg1-g8</b>	ACTGGATCAACACTCACACC	$\Delta rosK$ sgRNA
<b>sg2-g8</b>	GCCTTCCAGCCATTTTCGAT	$\Delta rosK$ sgRNA
<b>sg1-g9</b>	TCTCCTATCAACGAAAAGCG	$\Delta rosJ$ sgRNA
<b>sg2-g9</b>	CGGACCTAAGCCATTCAATG	$\Delta rosJ$ sgRNA
<b>sg1-g10</b>	GCCCTCTACTACTGACACAG	$\Delta rosI$ sgRNA
<b>sg2-g10</b>	TCGTGCCGCAAATGTCTGT	$\Delta rosI$ sgRNA
<b>sg1-g11</b>	GCACGTACAGTCCCACCAAG	$\Delta rosH$ sgRNA
<b>sg2-g11</b>	taactacttgaaccctgg	$\Delta rosH$ sgRNA

**Table S5.** List of nucleic acids used for fungal strain engineering.

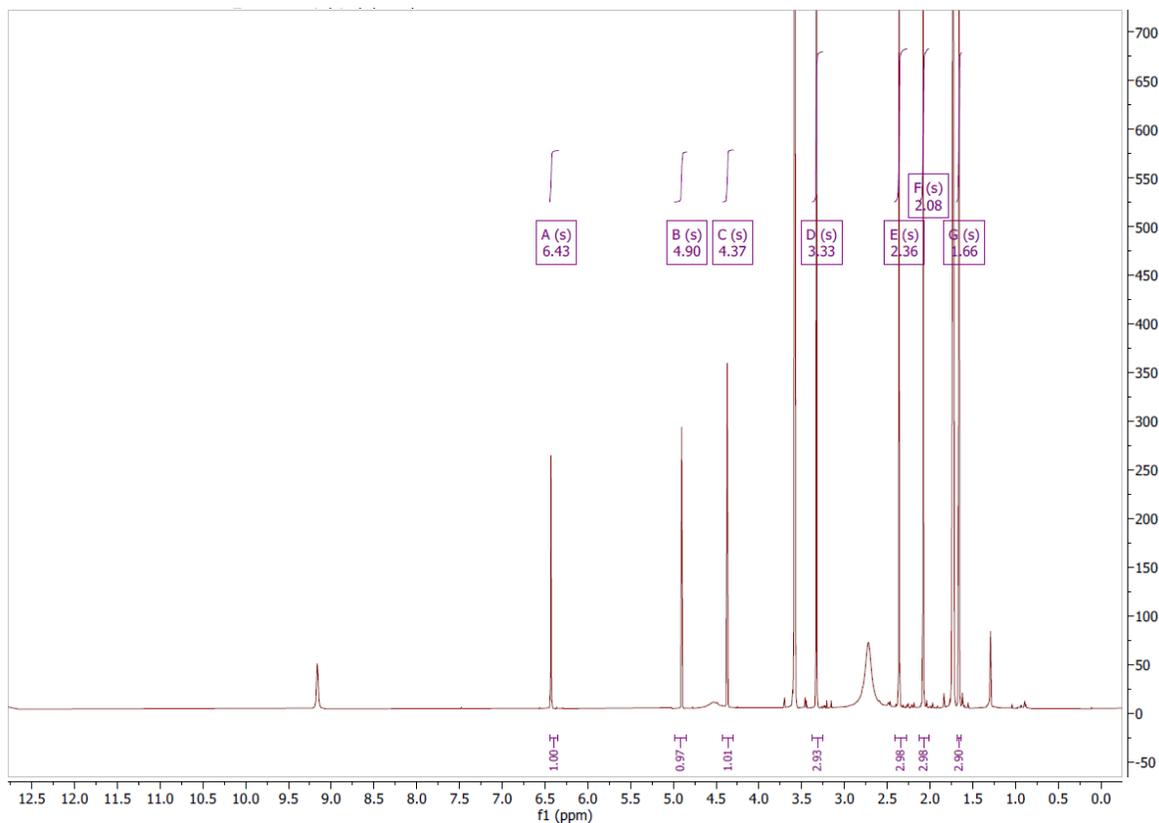
<b>Organism</b>	<b>Uniprot ID</b>	<b>Gene name</b>
<i>H. sapiens</i>	P06493	CDK1
<i>H. sapiens</i>	P24941	CDK2
<i>H. sapiens</i>	Q00526	CDK3
<i>H. sapiens</i>	P11802	CDK4
<i>H. sapiens</i>	Q00535	CDK5
<i>H. sapiens</i>	Q00534	CDK6
<i>H. sapiens</i>	P50613	CDK7
<i>H. sapiens</i>	P49336	CDK8
<i>H. sapiens</i>	P50750	CDK9
<i>H. sapiens</i>	Q15131	CDK10
<i>H. sapiens</i>	Q9UQ88	CDK11A
<i>H. sapiens</i>	P21127	CDK11B
<i>H. sapiens</i>	Q9NYV4	CDK12
<i>H. sapiens</i>	Q14004	CDK13
<i>H. sapiens</i>	O94921	CDK14
<i>H. sapiens</i>	Q96Q40	CDK15
<i>H. sapiens</i>	Q00536	CDK16
<i>H. sapiens</i>	Q00537	CDK17
<i>H. sapiens</i>	Q07002	CDK18
<i>H. sapiens</i>	Q9BWU1	CDK19
<i>H. sapiens</i>	Q8IZL9	CDK20
<i>S. cerevisiae</i>	P17157	PHO85
<i>S. cerevisiae</i>	P00546	CDC28
<i>S. cerevisiae</i>	Q03957	CTK1
<i>S. cerevisiae</i>	P23293	BUR1
<i>S. cerevisiae</i>	P39073	SRB10
<i>S. cerevisiae</i>	P06242	KIN28

**Table S6.** List of Uniprot accessions used for CDK phylogenetic tree.

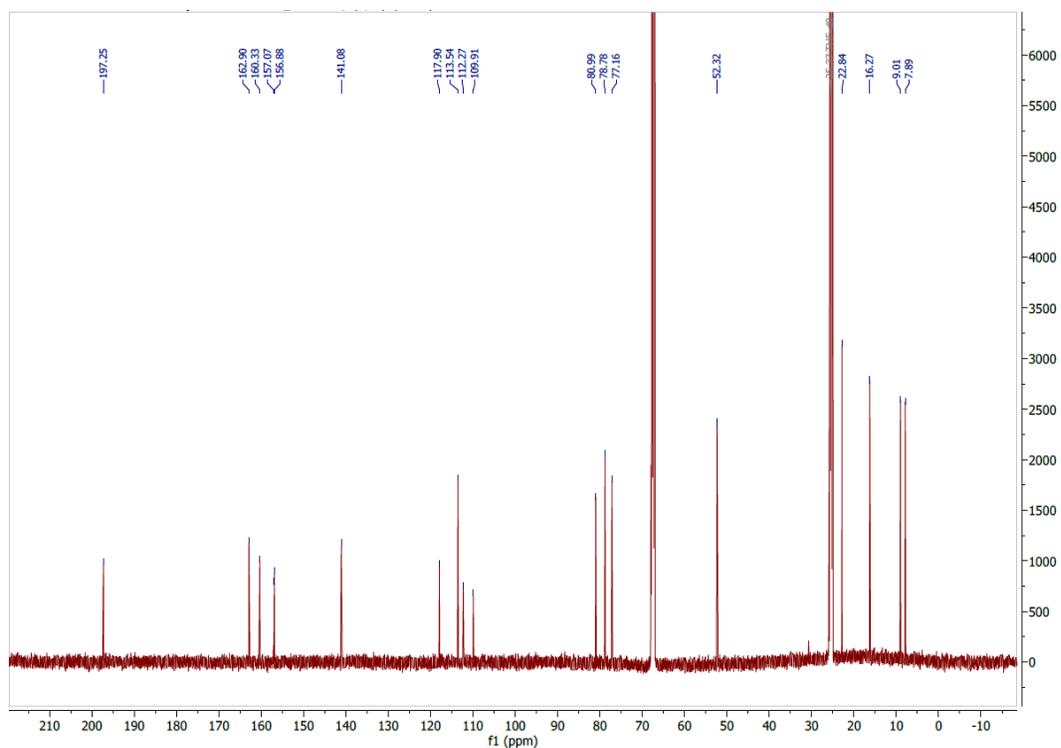
## MS and NMR Spectra of fungal compounds isolated in this study



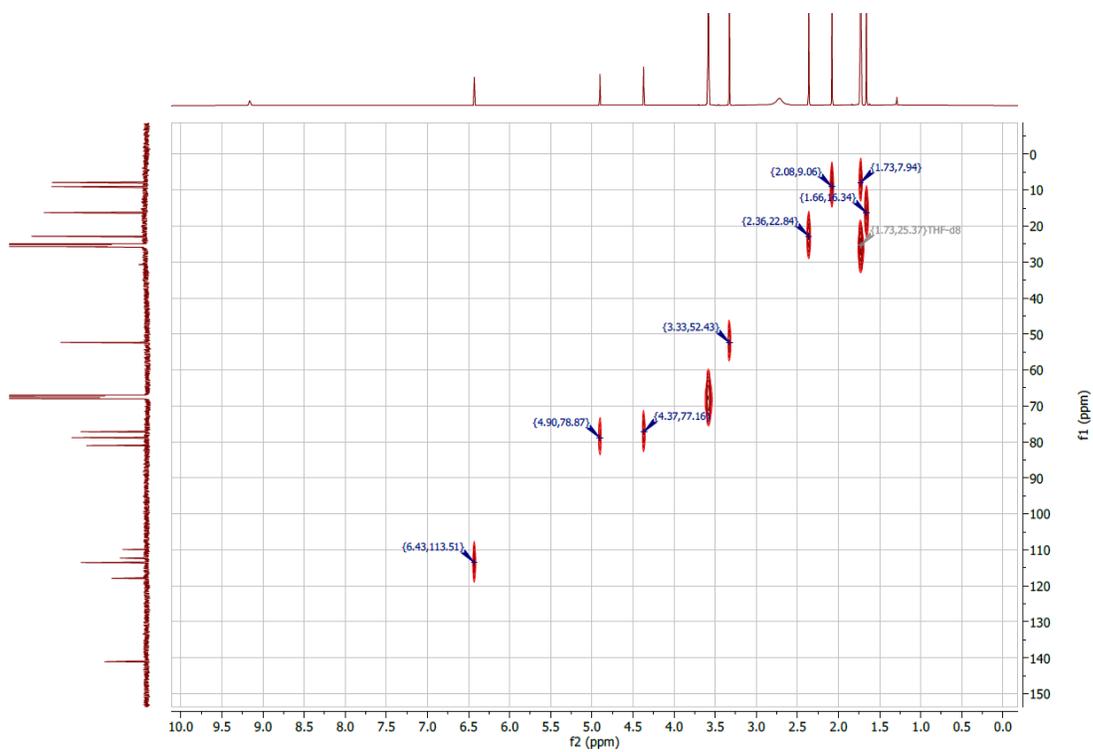
### MS/MS spectrum of compound 1



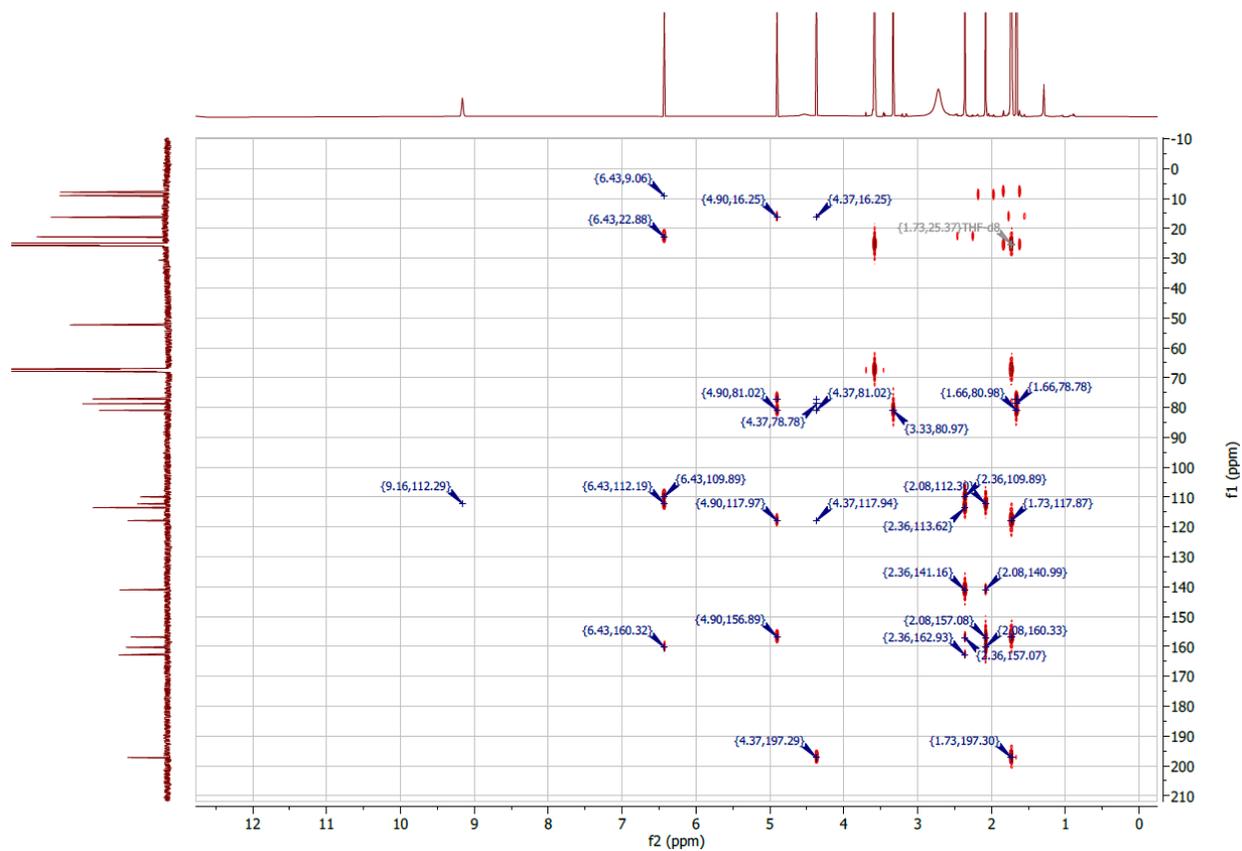
### <sup>1</sup>H NMR of compound 1 (600 MHz, THF-d<sub>8</sub>)



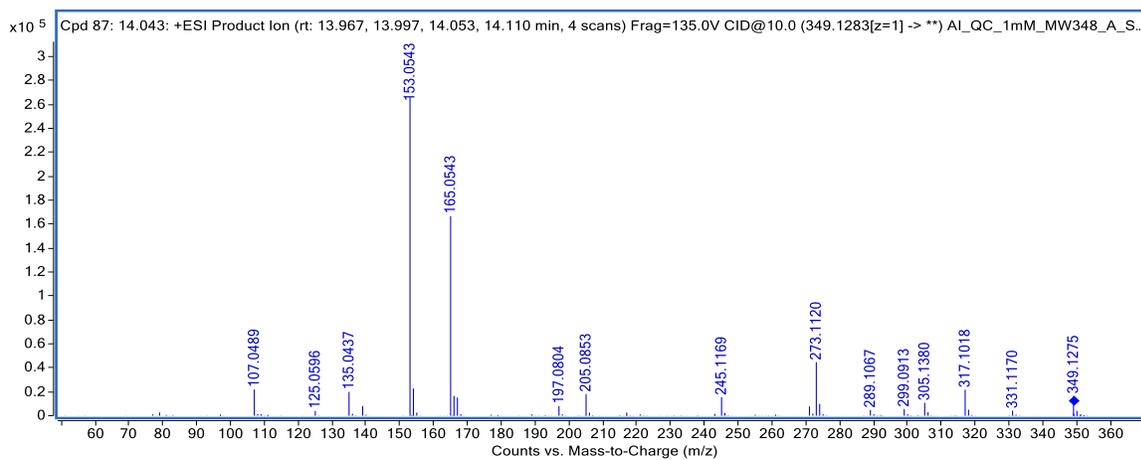
<sup>13</sup>C NMR of compound 1 (150 MHz, THF-d<sub>8</sub>)



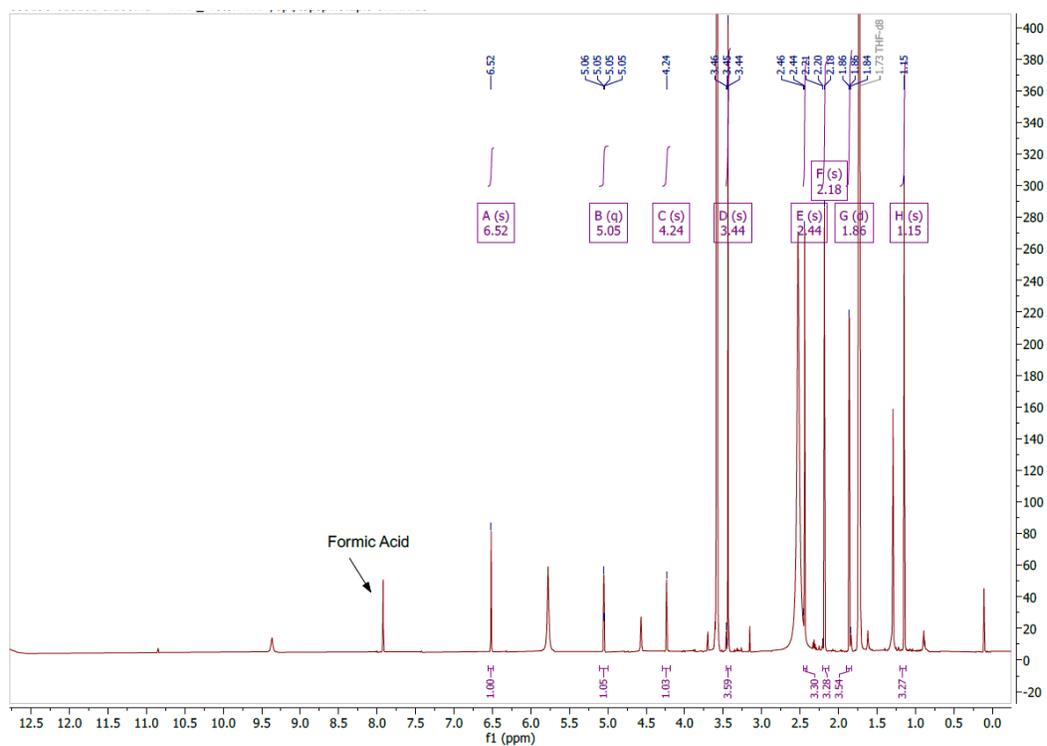
HSQC of compound 1 (600 MHz, THF-d<sub>8</sub>)



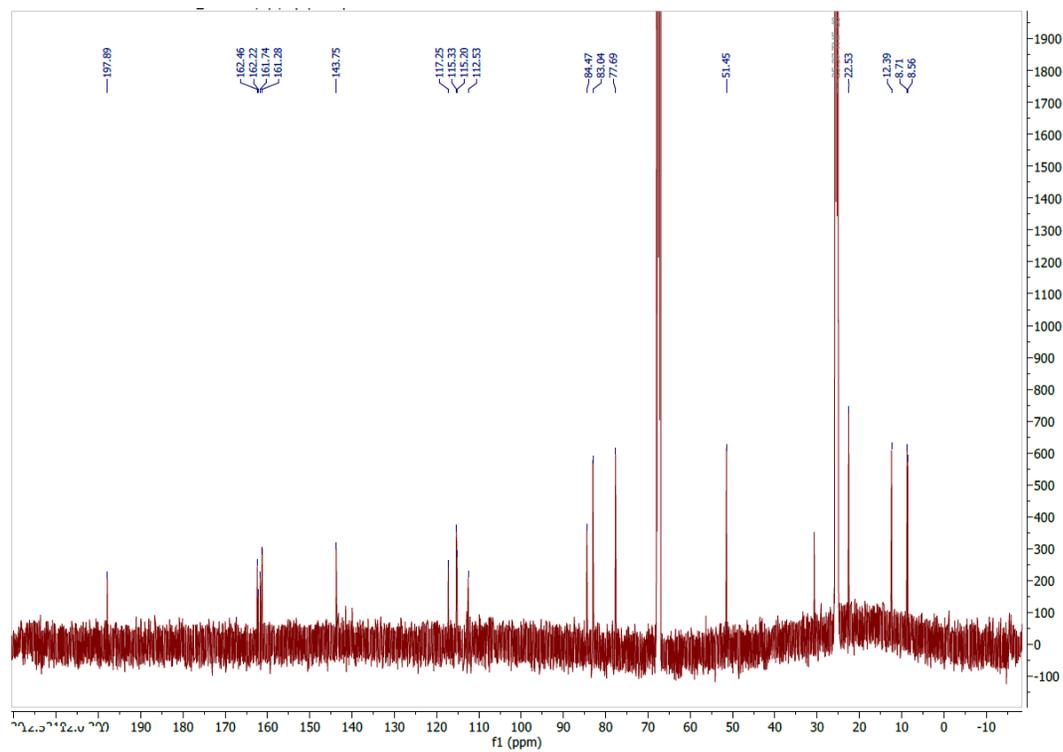
### MS/MS spectrum of compound 2



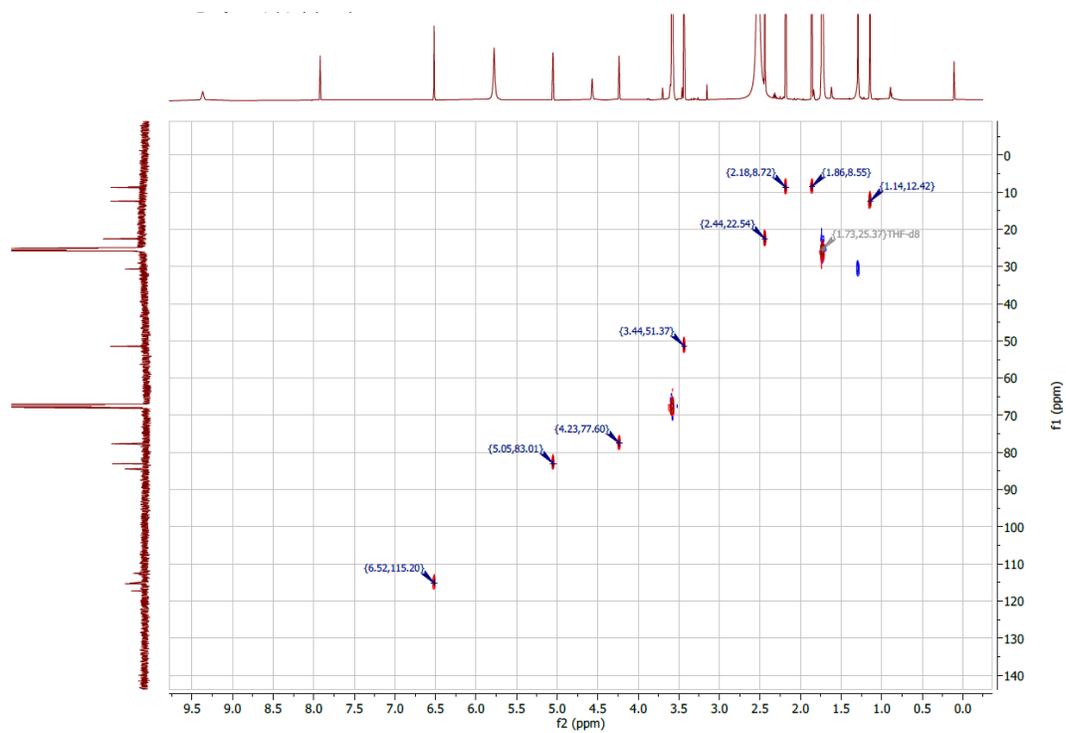
### HMBC of compound 1 (600 MHz, THF-*d*<sub>8</sub>)



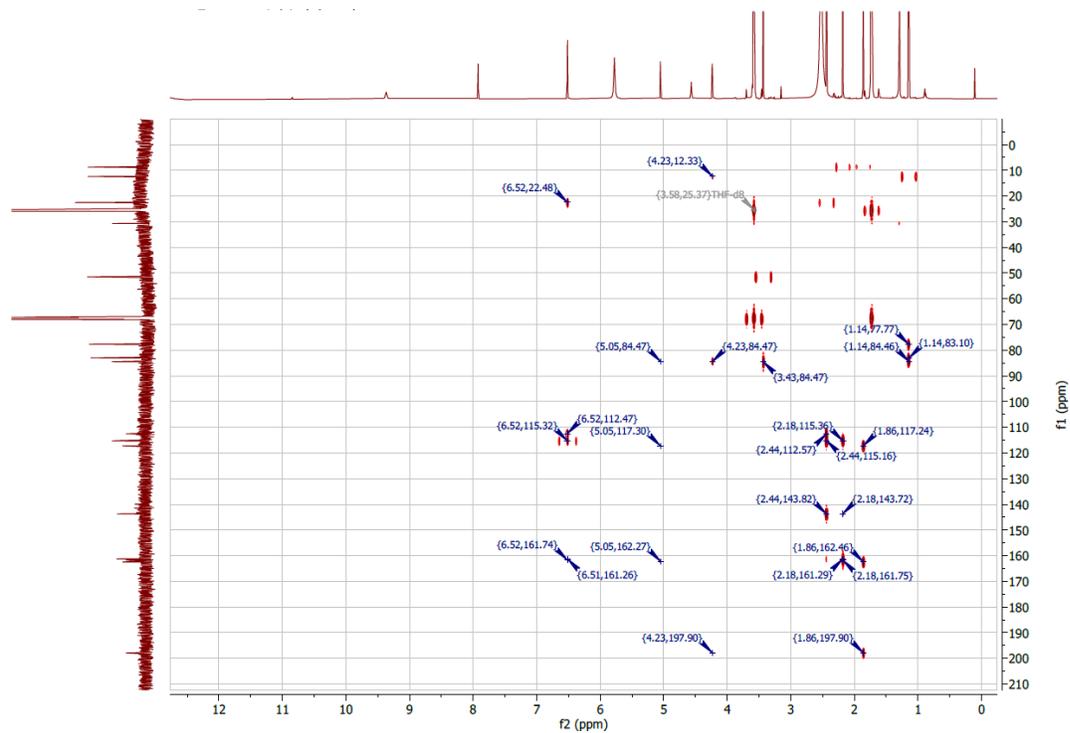
**<sup>1</sup>H NMR of compound 2 (600 MHz, THF-*d*<sub>8</sub>)**



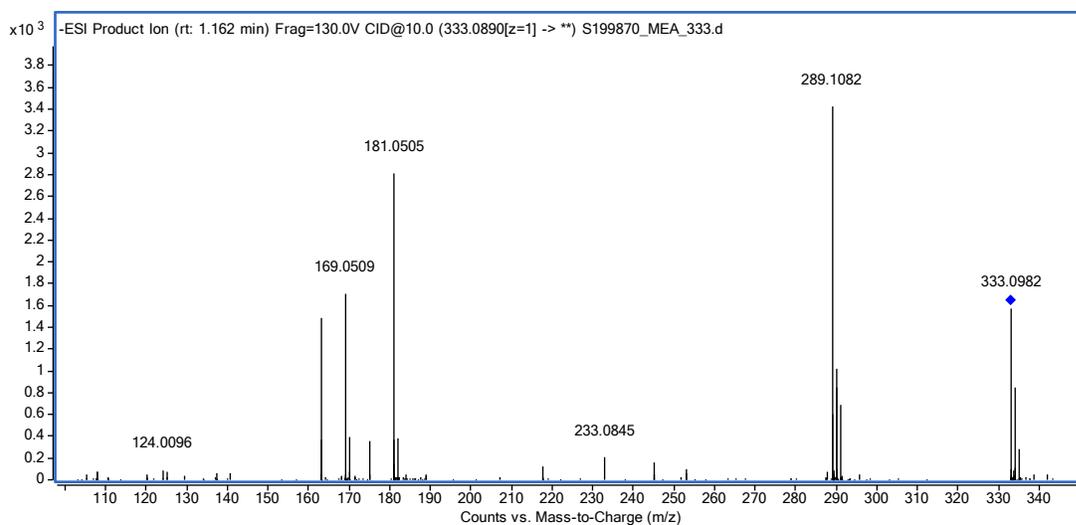
**<sup>13</sup>C NMR of compound 2 (150 MHz, THF-*d*<sub>8</sub>)**



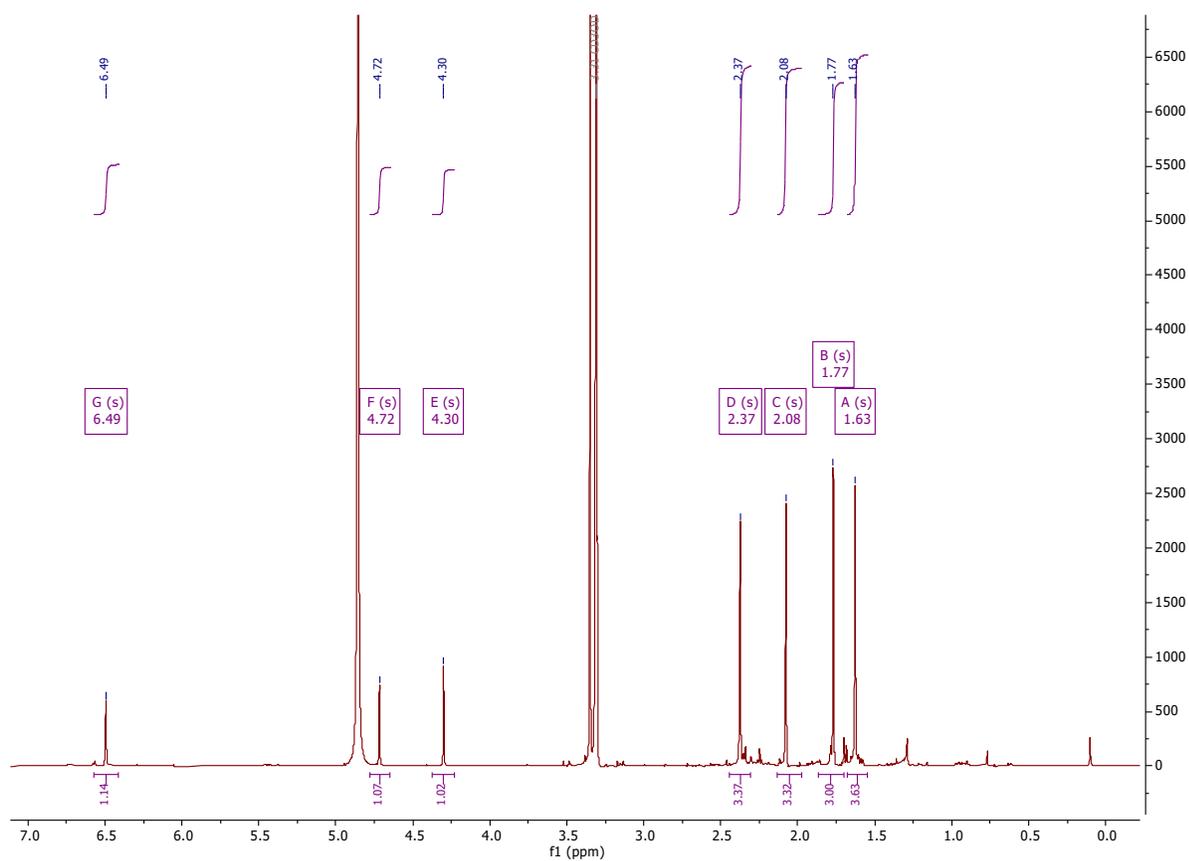
HSQC of compound 2 (600 MHz, THF-*d*<sub>8</sub>)



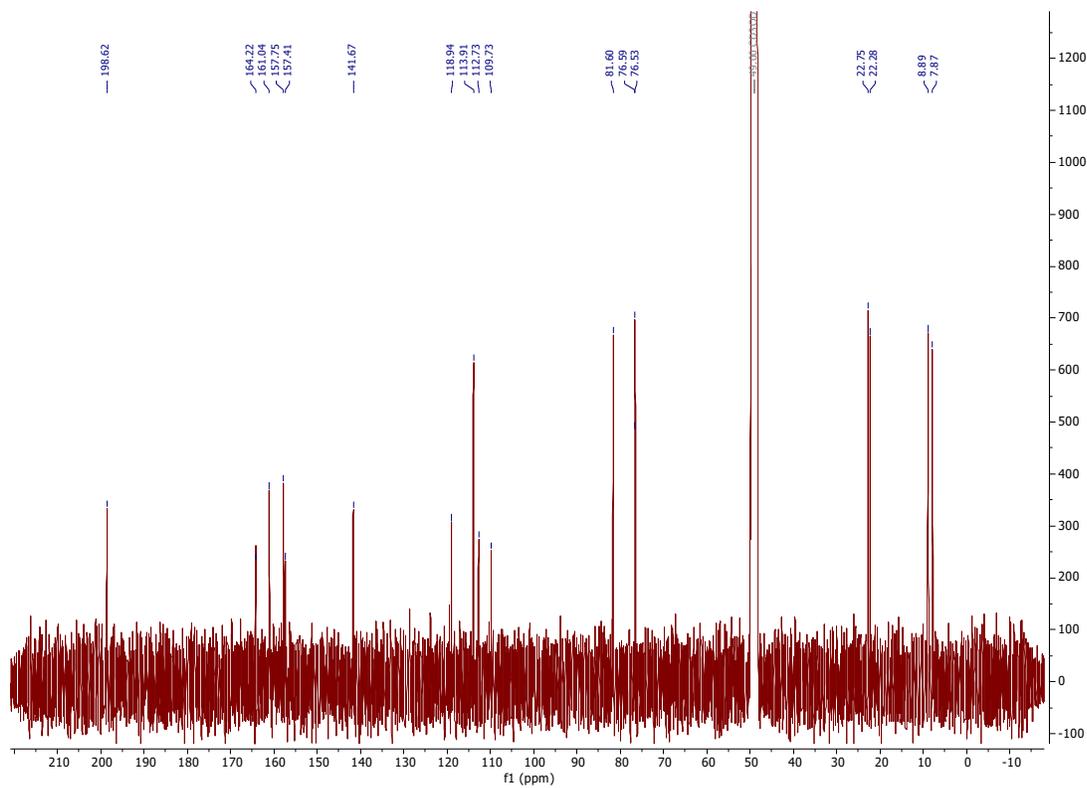
HMBC of compound 2 (600 MHz, THF-*d*<sub>8</sub>)



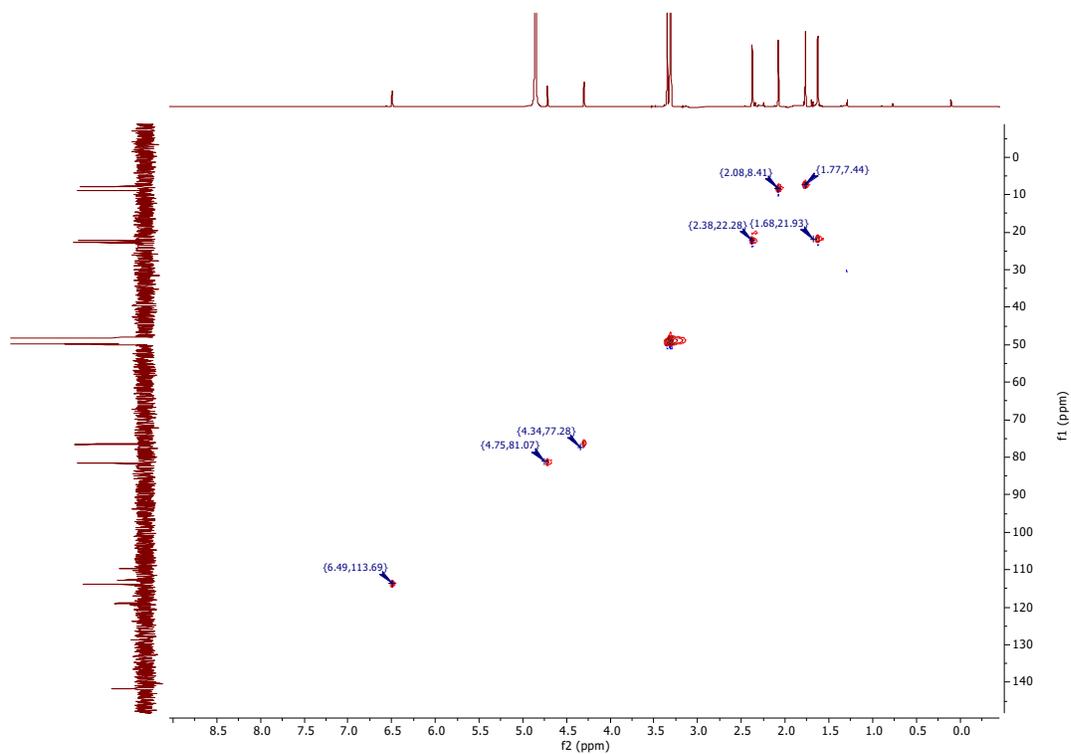
MS/MS spectrum of compound **8**



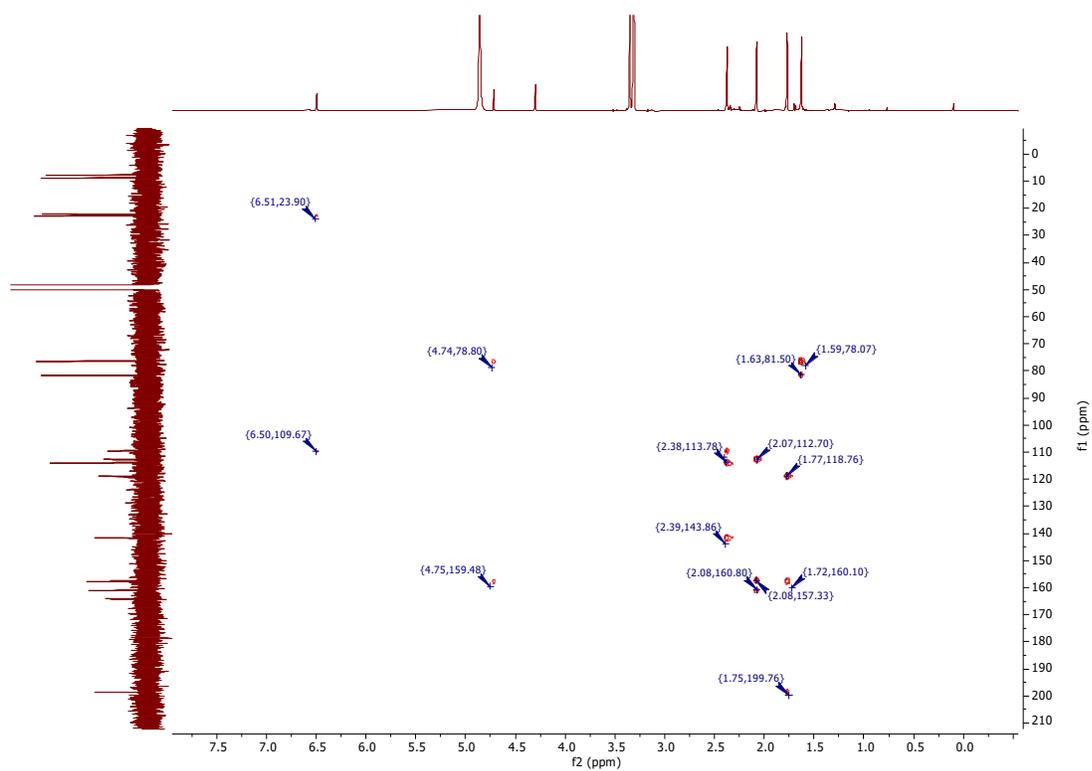
<sup>1</sup>H-NMR of compound **8** (400 MHz, MeOD)



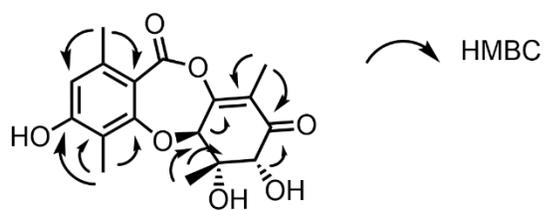
$^{13}\text{C-NMR}$  of compound **8** (100 MHz, MeOD).



HSQC of compound **8** (400 MHz, MeOD).

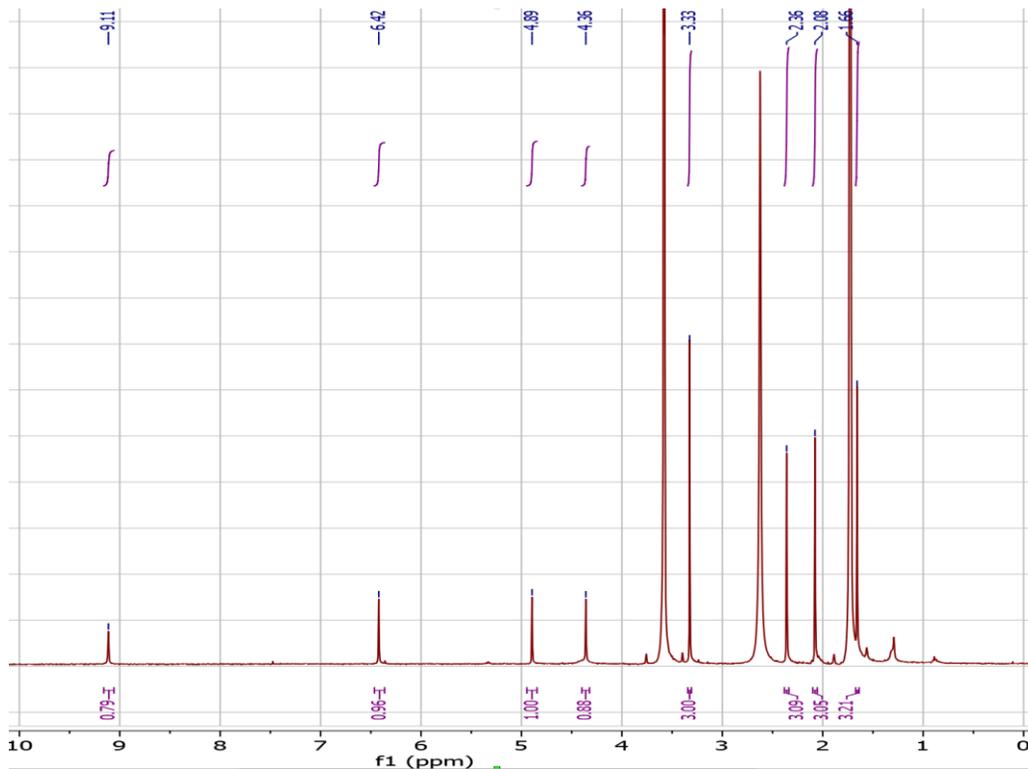


HMBC of compound **8** (400 MHz, MeOD).

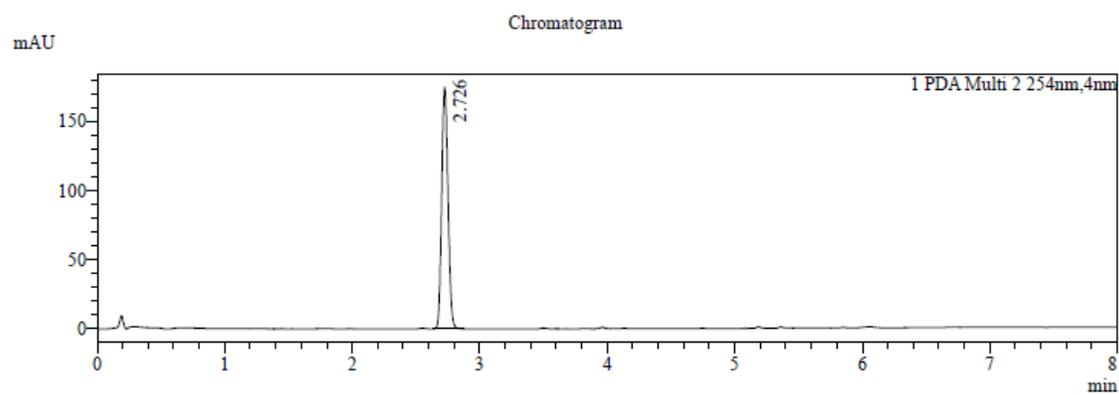


HMBC correlations of isolated compound **8**.

## MS and NMR spectra of compounds synthesized in this study



<sup>1</sup>H NMR of compound (1) (400 MHz, THF-d<sub>8</sub>)



1 PDA Multi 2 / 254nm, 4nm

### Integration Result

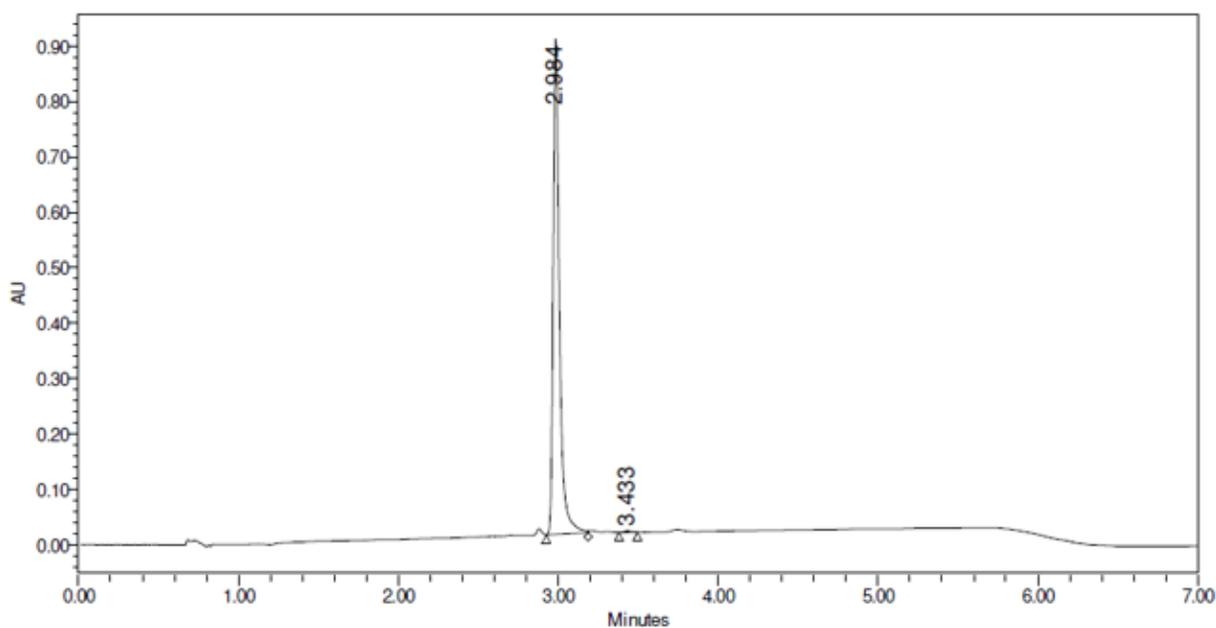
Peak Table						
Peak#	Ret. Time	Height	Height%	USP Width	Area	Area%
1	2.726	173420	100.000	0.092	601733	100.000

HPLC of compound (1)

## CHIRAL SFC REPORT

### SAMPLE INFORMATION

Sample ID:	LMT-0011-146-P1F1_L1	Acquired By:	System
Compound ID:	N/A		
Acq Method:	AS_3_EtOH_DEA_5_40_25ML_7MIN	Vial:	1:F,2
Channel Name:	PDA Ch1 220nm@4.8nm -Compens.	Injection Volume:	6.00 ul
Proc. Chnl. Descr.:	PDA Ch1 220nm@4.8nm -Compens.		
Date Acquired:	1/20/2022 1:48:47 AM CST		
Date Processed:	1/20/2022 10:20:52 AM CST		
Instrument:	CAS-SH-ANA-SFC-L (Waters UPCC with PDA Detector)		



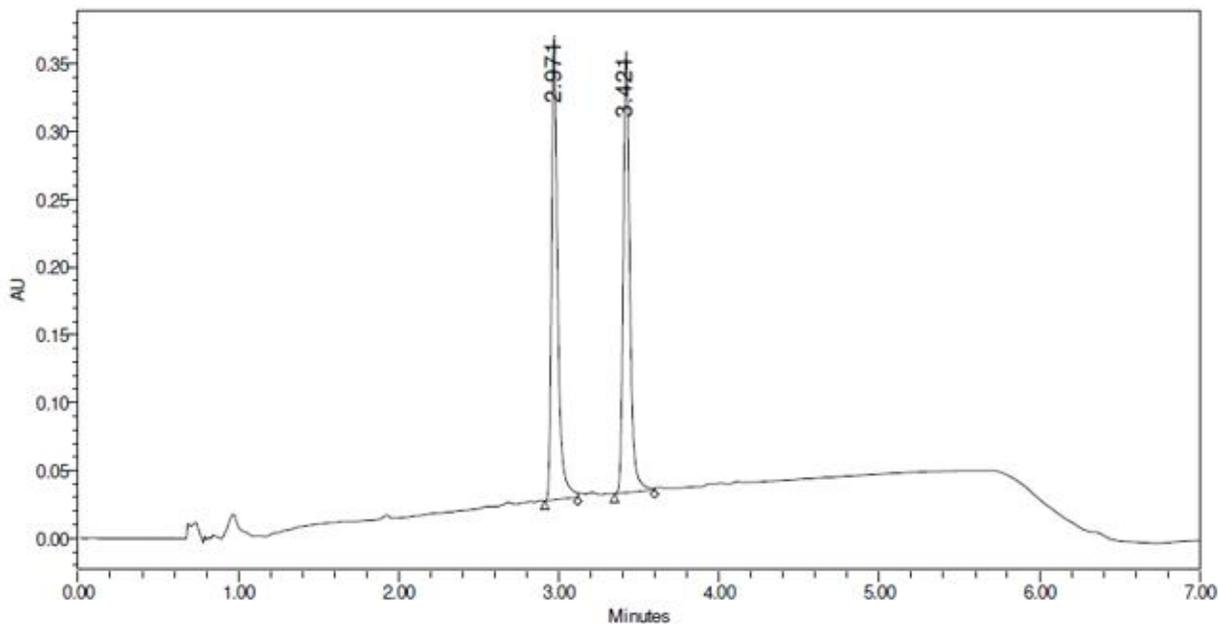
	RT	Area	% Area
1	2.984	2428105	99.67
2	3.433	7976	0.33

Chiral SFC of synthetic compound (1).

## CHIRAL SFC REPORT

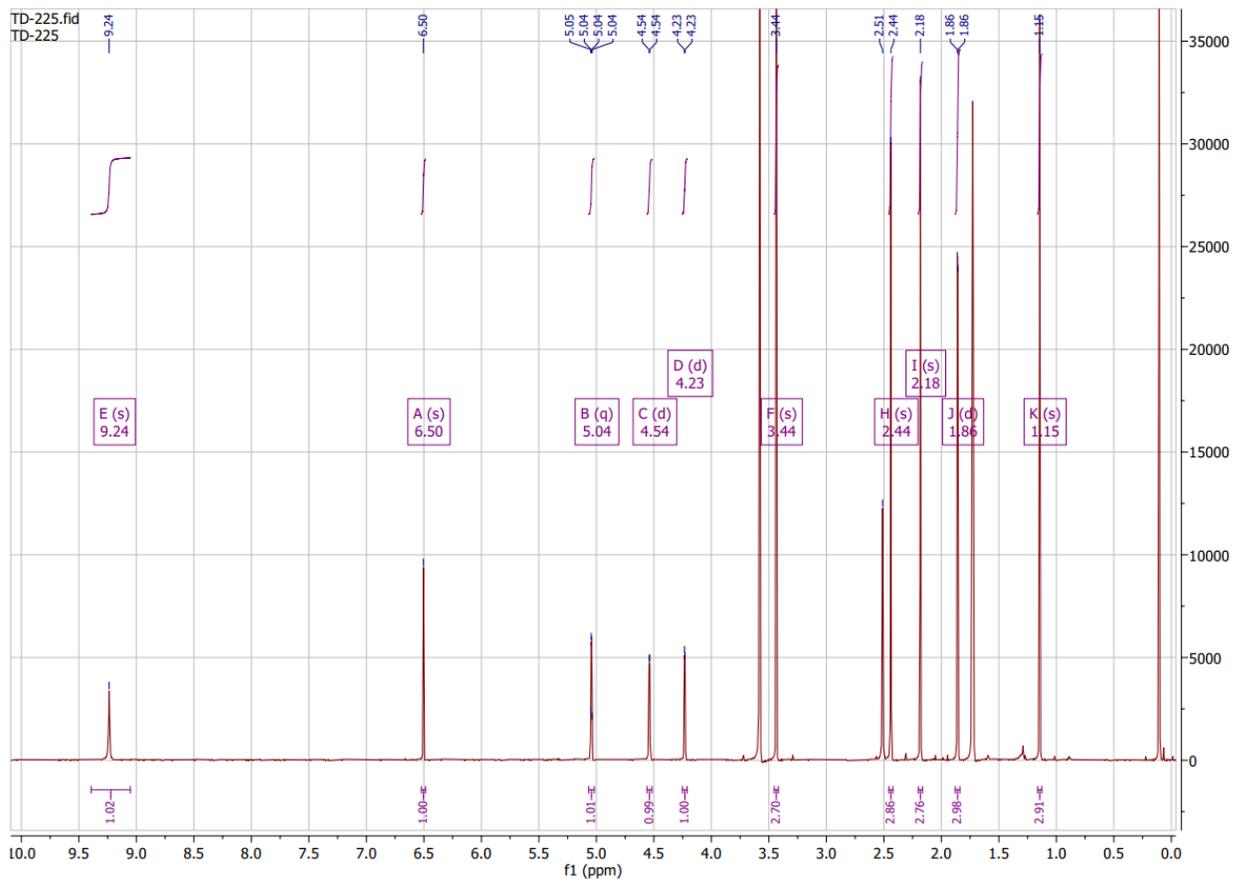
### SAMPLE INFORMATION

Sample ID:	ES17794-383-P1F1_L2	Acquired By:	System
Compound ID:	N/A		
Acq Method:	AS_3_EtOH_DEA_5_40_25ML_7MIN	Vial:	2:E,4
Channel Name:	PDA Ch1 220nm@4.8nm -Compens.	Injection Volume:	5.00 ul
Proc. Chnl. Descr.:	PDA Ch1 220nm@4.8nm -Compens.		
Date Acquired:	10/27/2021 2:11:08 AM CST		
Date Processed:	10/27/2021 10:15:01 AM CST		
Instrument:	CAS-SH-ANA-SFC-L (Waters UPCC with PDA Detector)		

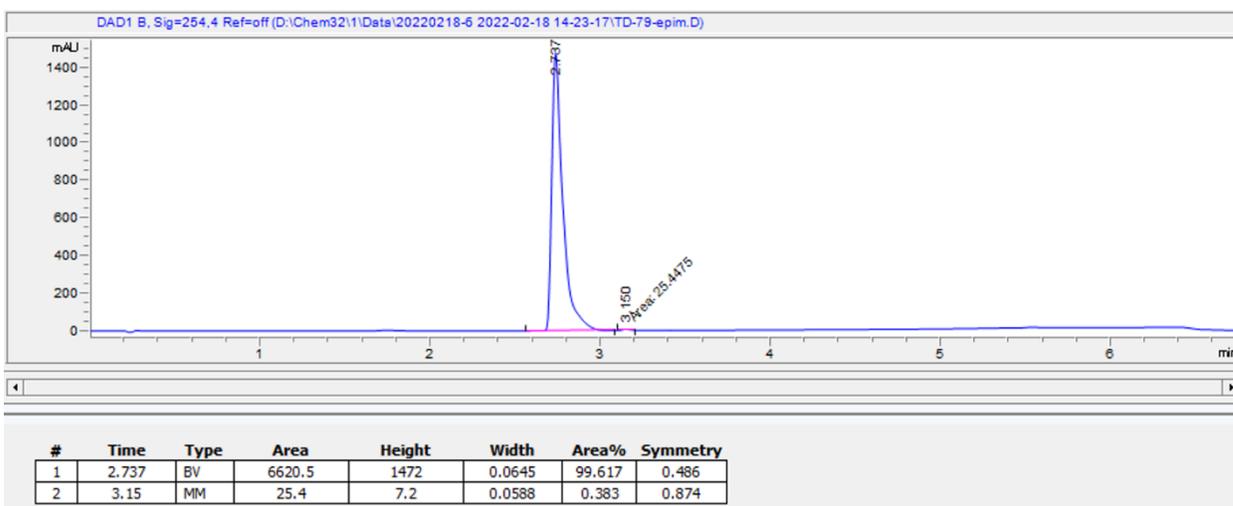


	RT	Area	% Area
1	2.971	909320	49.24
2	3.421	937282	50.76

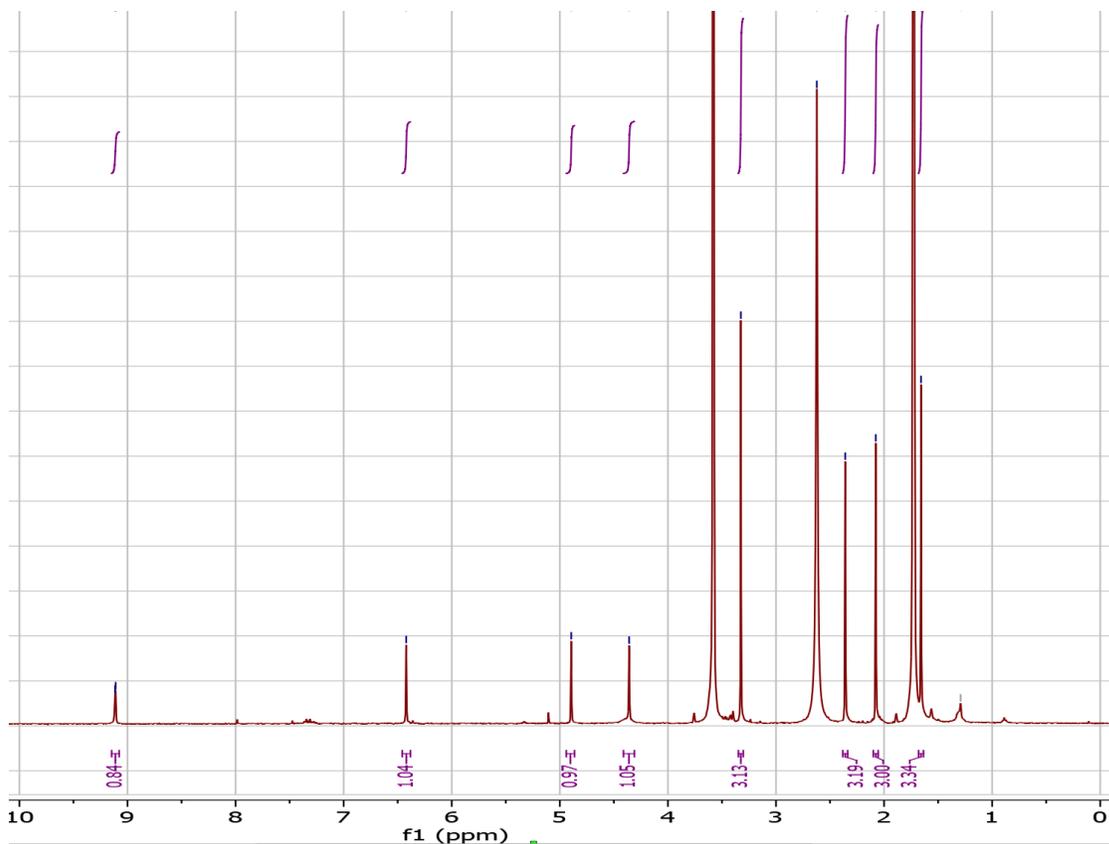
Chiral SFC of racemate (**23**). Reference for Chiral SFC of compounds **1** and **24**.



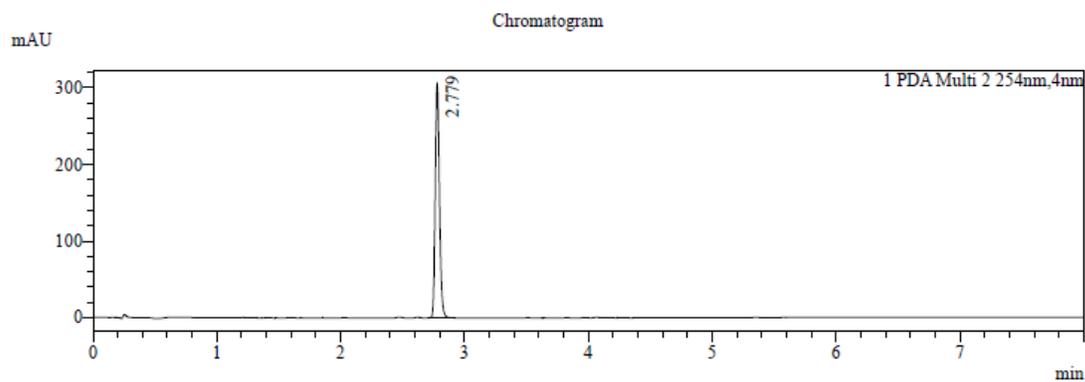
$^1\text{H}$  NMR of compound (**2**) (400 MHz,  $\text{THF-d}_8$ )



LCMS of compound (**2**)



$^1\text{H}$  NMR of compound (**24**) (400 MHz,  $\text{THF-}d_8$ )



1 PDA Multi 2 / 254nm,4nm

=====  
 Integration Result  
 =====

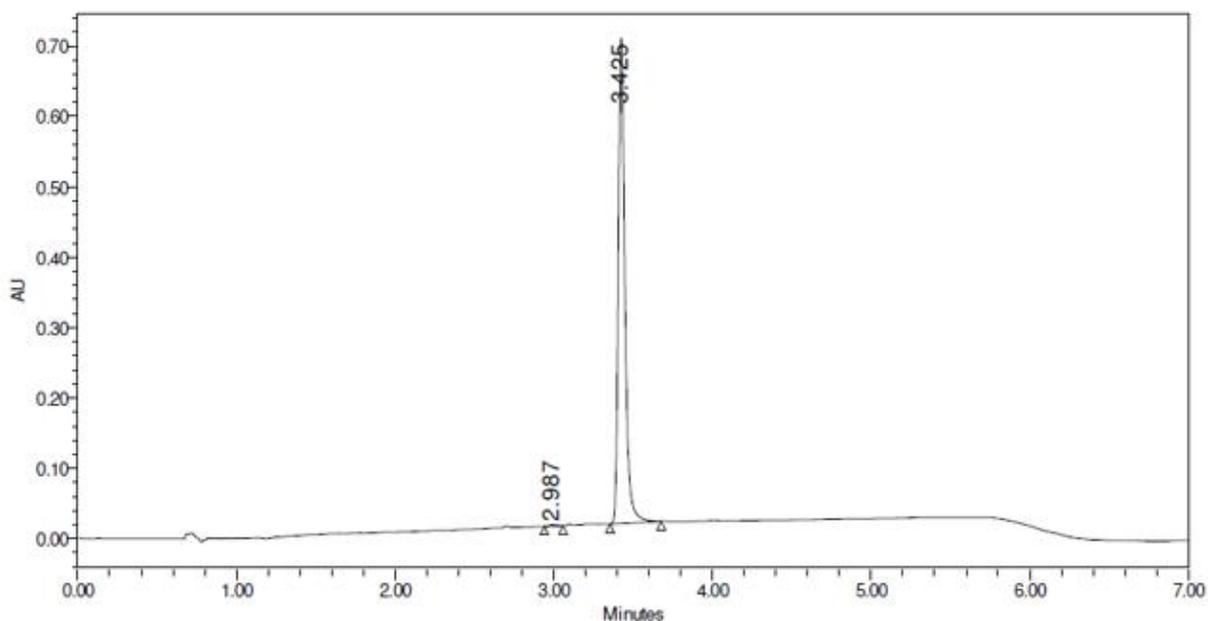
PDA Ch2 254nm		Peak Table					
Peak#	Ret. Time	Height	Height%	USP Width	Area	Area%	
1	2.779	300879	100.000	0.062	721953	100.000	

HPLC of compound (**24**)

# CHIRAL SFC REPORT

## SAMPLE INFORMATION

Sample ID:	LMT-0011-146-P2F1_L1	Acquired By:	System
Compound ID:	N/A		
Acq Method:	AS_3_EtOH_DEA_5_40_25ML_7MIN	Vial:	1:F,1
Channel Name:	PDA Ch1 220nm@4.8nm -Compens.	Injection Volume:	4.00 ul
Proc. Chnl. Descr.:	PDA Ch1 220nm@4.8nm -Compens.		
Date Acquired:	1/20/2022 1:40:46 AM CST		
Date Processed:	1/20/2022 10:20:40 AM CST		
Instrument:	CAS-SH-ANA-SFC-L (Waters UPCC with PDA Detector)		

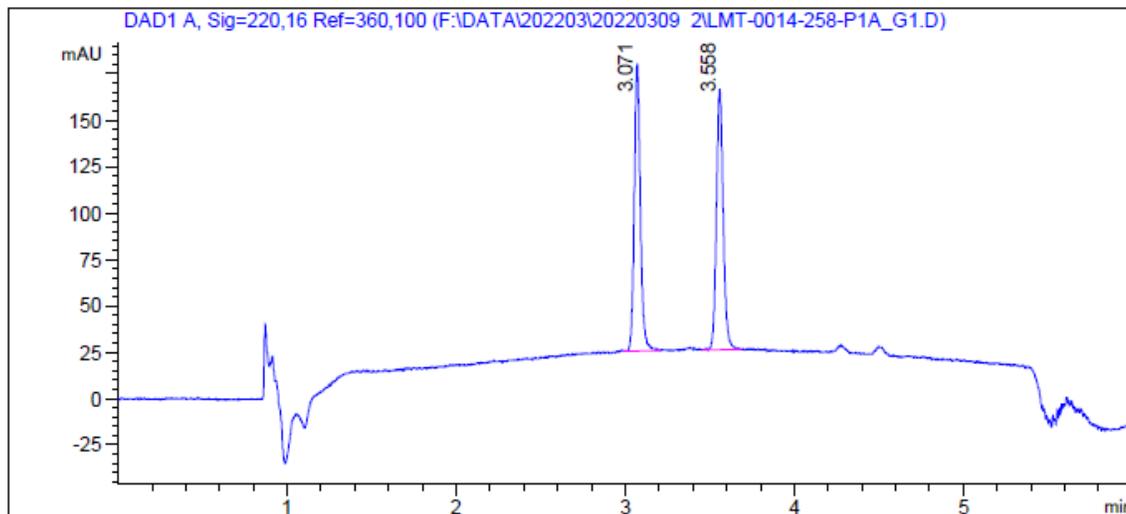


	RT	Area	% Area
1	2.987	5932	0.30
2	3.425	1993507	99.70

Chiral SFC of compound (**24**)

CHIRAL SFC REPORT

Compound ID : LMT-0014-258-P1A  
 Sample ID : LMT-0014-258-P1A\_G1  
 Injection Date : 3/10/2022  
 Acq Method : F:\Data\202203\20220309 2\OJ-H\_EtOH(DEA)\_5\_40\_25ML\_ ->  
 Raw Data : F:\DATA\202203\20220309 2\LMT-0014-258-P1A\_G1.D  
 Instrument : CAS-SH-ANA-SFC-G (Agilent 1260 with DAD detector)  
 Method Comments : Column: ChiralCel OJ-H 150x4.6mm I.D., 5um  
 Mobile phase: A: CO2 B:Ethanol (0.05% DEA)  
 Gradient: from 5% to 40% of B in 4.5min , then 5% of B  
 for 1.5 min  
 Flow rate: 2.5mL/min Column temp.:40°C  
 Back pressure: 100 bar



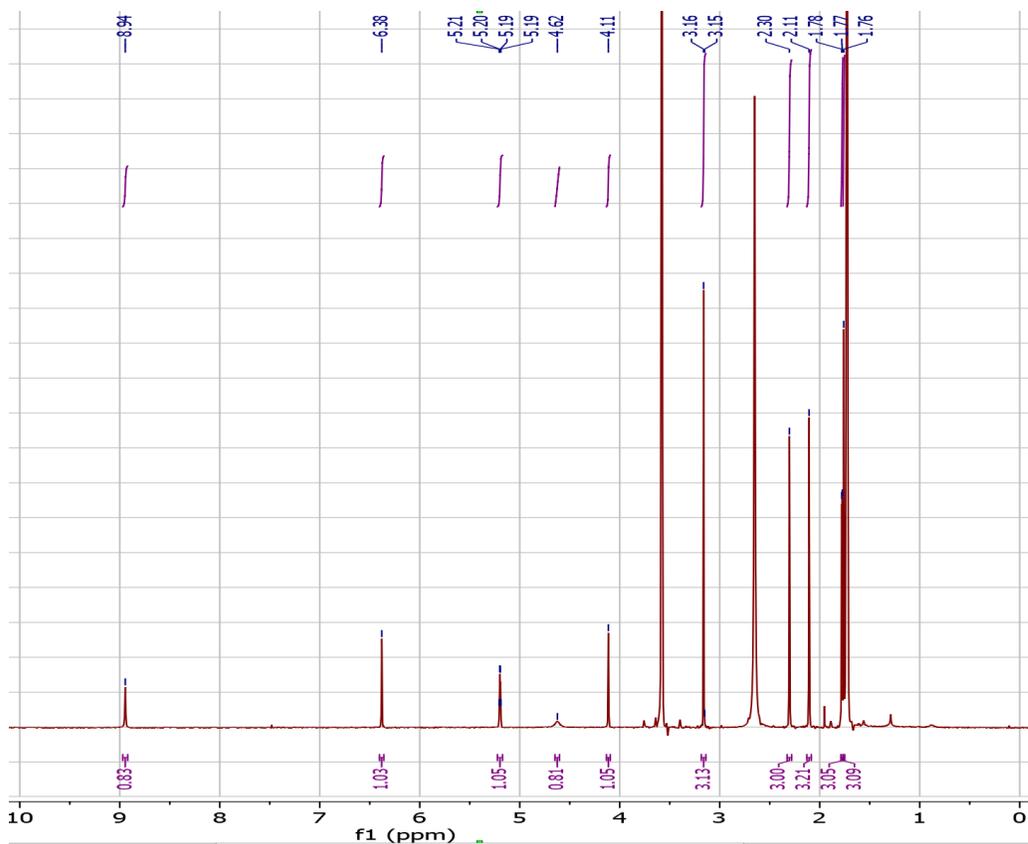
=====

DAD1 A, Sig=220,16 Ref=360,100

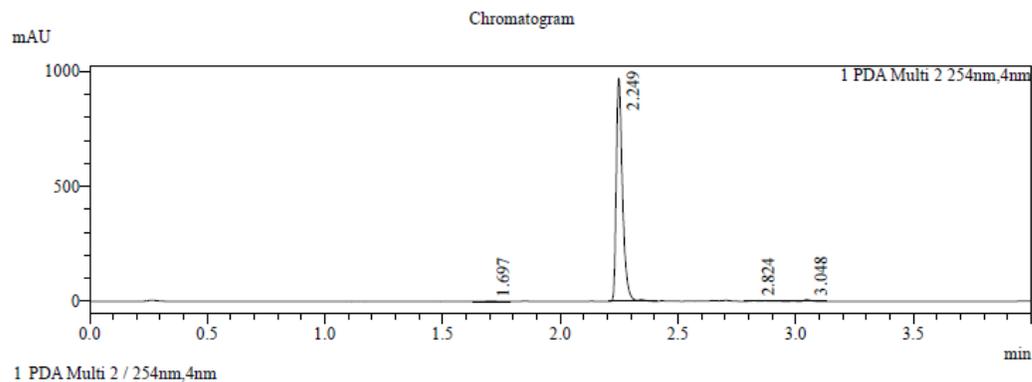
#	Meas.	Ret. Time	Height	Height %	Width	Area	Area %
1		3.071	155.272	52.501	0.038	392.083	49.962
2		3.558	140.480	47.499	0.043	392.686	50.038

-----

Chiral SFC of racemate (**26**). Reference for Chiral SFC of compounds **27** and **29**.



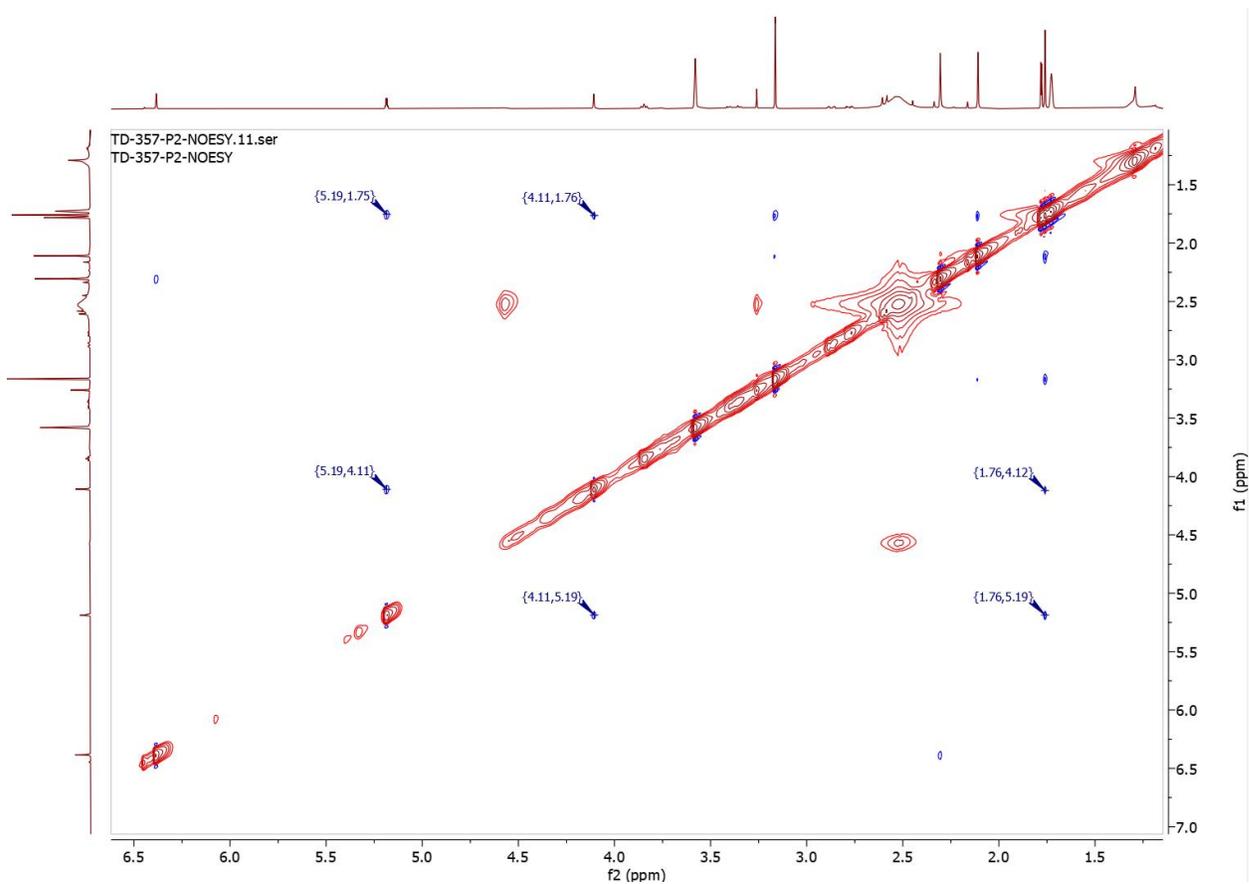
$^1\text{H}$  NMR of compound (**27**) (400 MHz,  $\text{THF-}d_8$ )



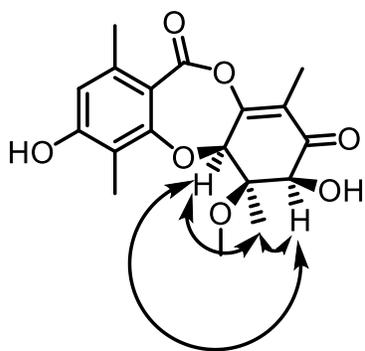
Integration Result

Peak Table						
Peak#	Ret. Time	Height	Height%	USP Width	Area	Area%
1	1.697	1522	0.167	0.093	6012	0.331
2	2.249	901487	98.976	0.048	1791851	98.545
3	2.824	1787	0.196	0.056	9148	0.503
4	3.048	6019	0.661	0.048	11302	0.622

HPLC of compound (**27**)



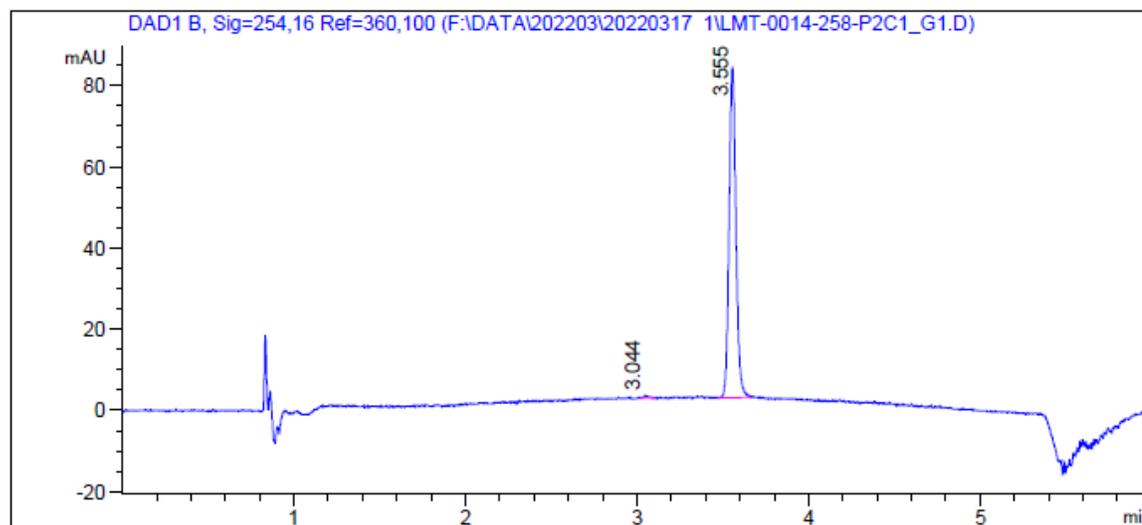
NOESY of compound (**27**) (400 MHz, THF- $d_8$ )



Correlations between C5a and C-6Me, C7 and C6-Me and C5a and C7

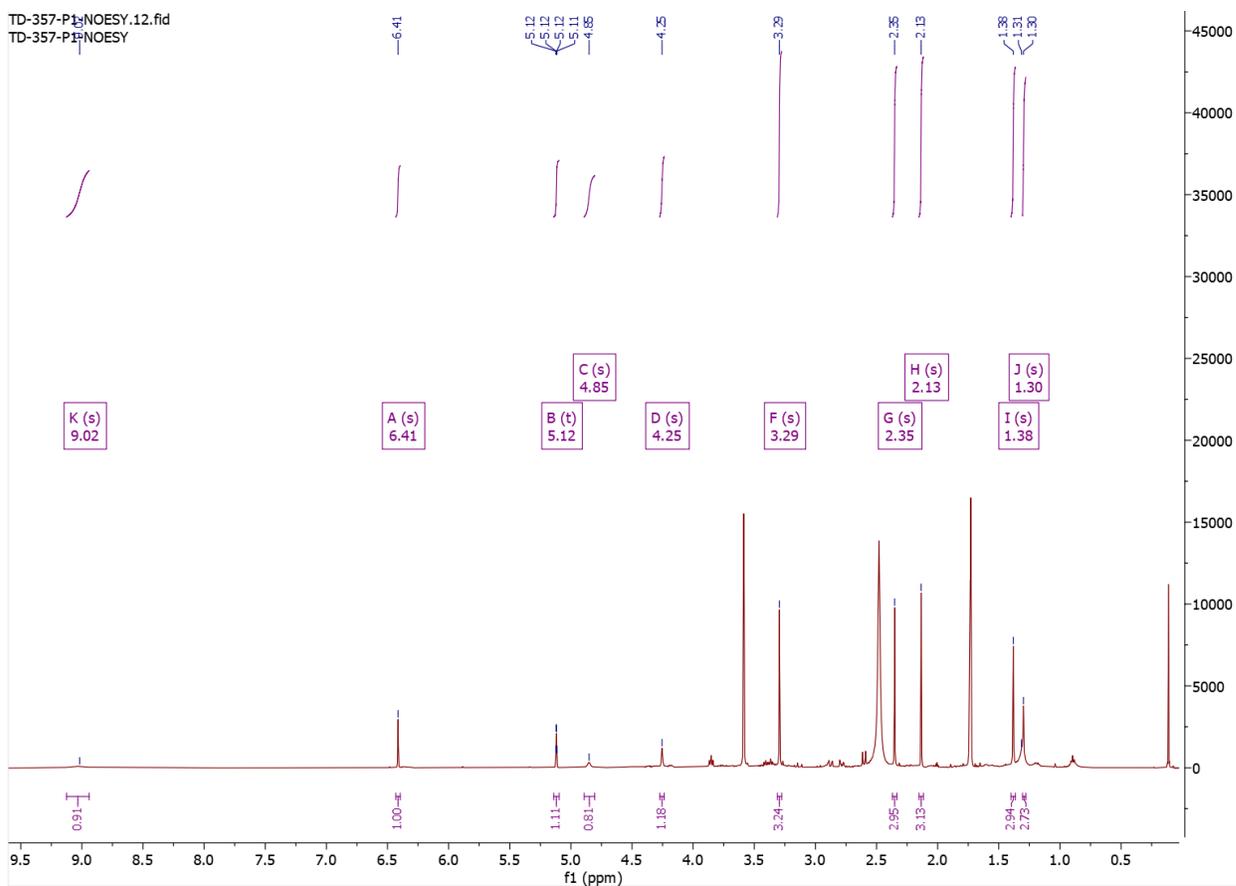
CHIRAL SFC REPORT

Compound ID : LMT-0014-258-P2C1  
 Sample ID : LMT-0014-258-P2C1\_G1  
 Injection Date : 3/17/2022  
 Acq Method : F:\Data\202203\20220317 1\OJ-H\_EtOH(DEA)\_5\_40\_25ML\_ ->  
 Raw Data : F:\DATA\202203\20220317 1\LMT-0014-258-P2C1\_G1.D  
 Instrument : CAS-SH-ANA-SFC-G (Agilent 1260 with DAD detector)  
 Method Comments : Column: ChiralCel OJ-H 150x4.6mm I.D., 5um  
 Mobile phase: A: CO2 B:Ethanol (0.05% DEA)  
 Gradient: from 5% to 40% of B in 4.5min , then 5% of B  
 for 1.5 min  
 Flow rate: 2.5mL/min Column temp.:40°C  
 Back pressure: 100 bar

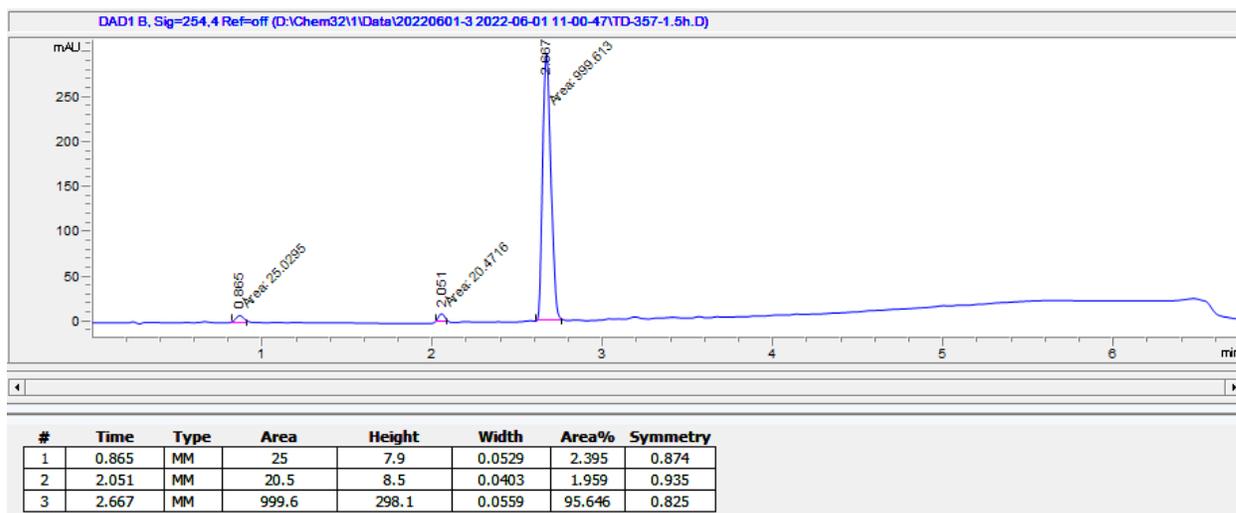


DAD1 B, Sig=254,16 Ref=360,100							
#	Meas.	Ret. Time	Height	Height %	Width	Area	Area %
1		3.044	0.602	0.733	0.041	1.490	0.655
2		3.555	81.541	99.267	0.042	226.188	99.345

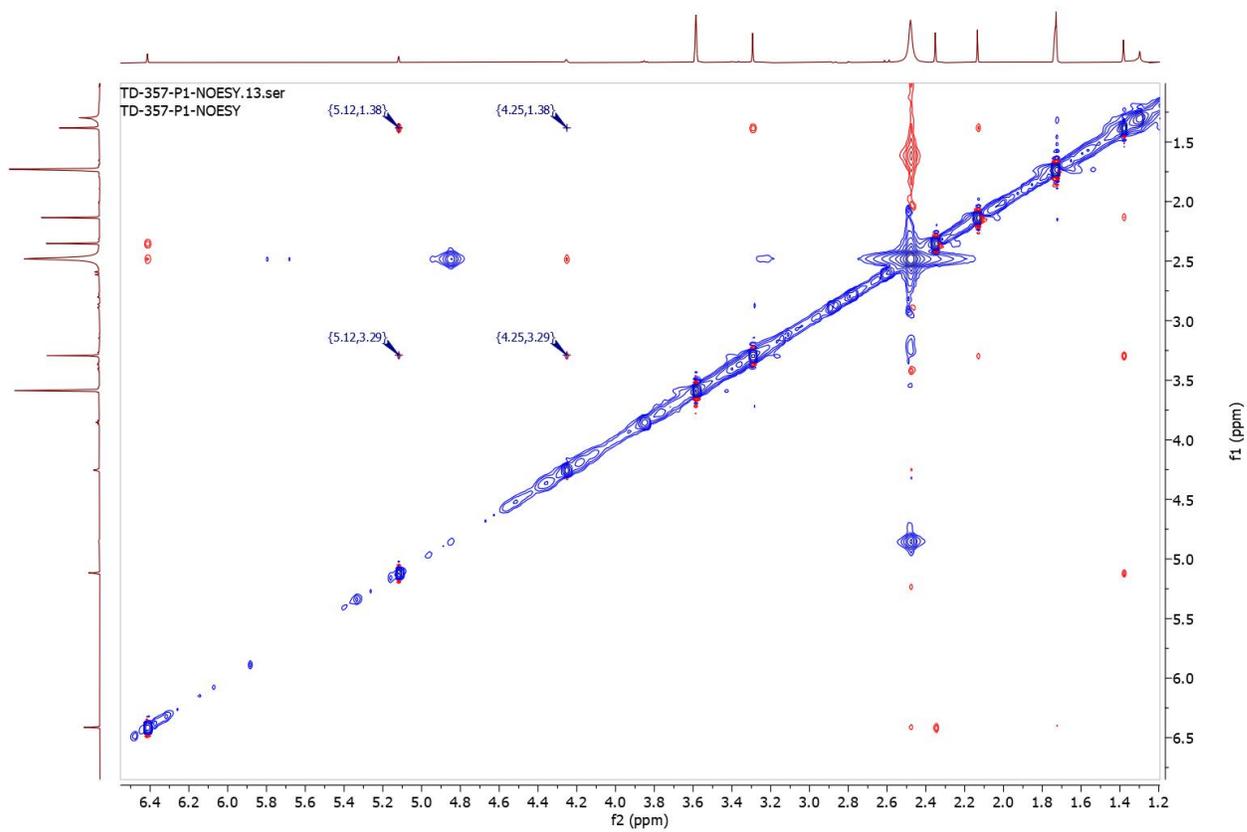
Chiral SFC of Compound (27)



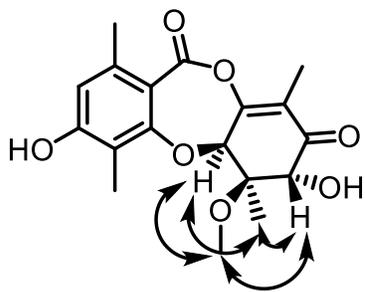
<sup>1</sup>H NMR of compound (28) (400 MHz, THF-d<sub>8</sub>)



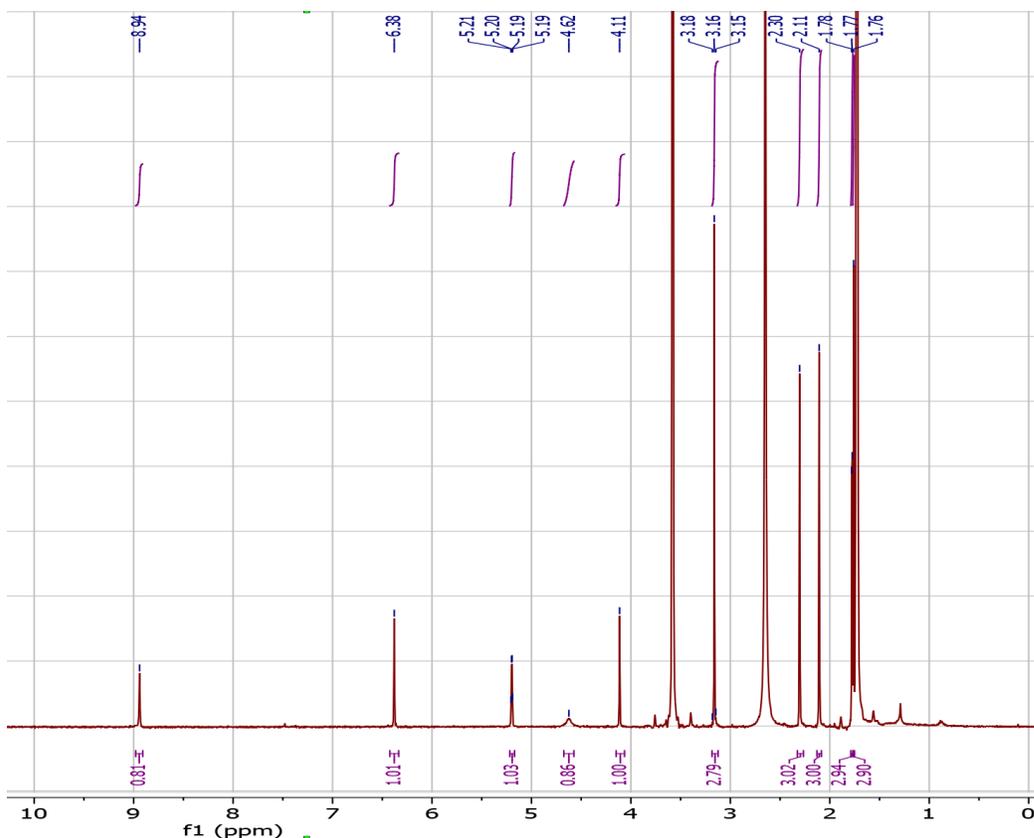
LCMS of compound (28)



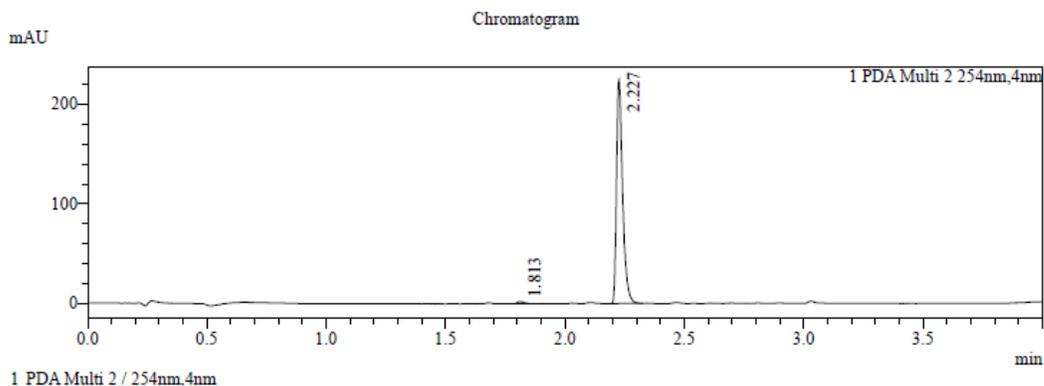
NOESY of compound (**28**) (400 MHz, THF- $d_8$ )



Correlations between C5a and C6-Me, C5a and C6-OMe, C-7 and C6-Me, C-7 and C-6OMe. No correlation between C5a and C7.



$^1\text{H}$  NMR of compound (**29**) (400 MHz,  $\text{THF-}d_8$ )



=====  
 Integration Result  
 =====

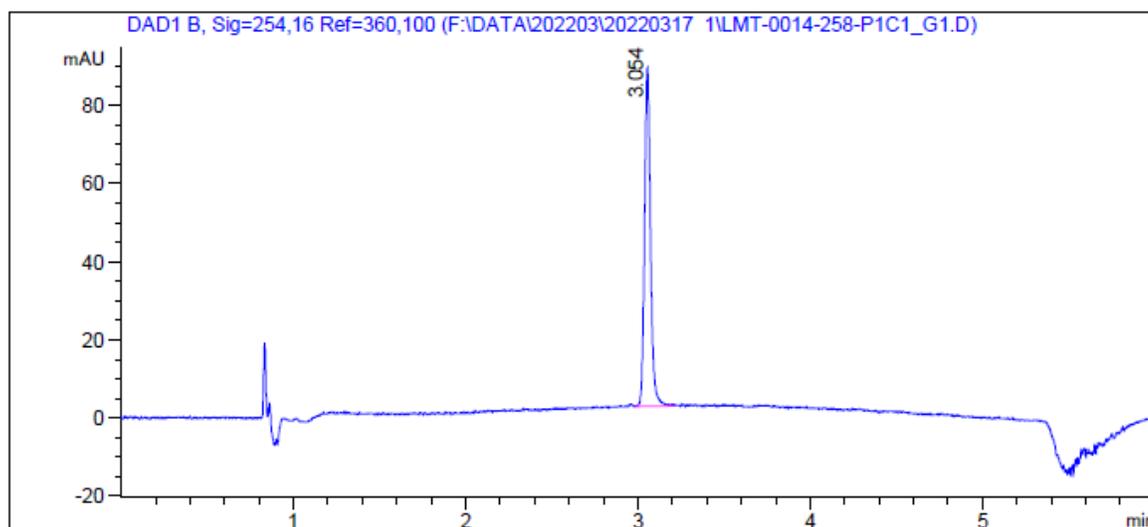
Peak Table

Peak#	Ret. Time	Height	Height%	USP Width	Area	Area%
1	1.813	2381	1.090	0.051	5298	1.310
2	2.227	216015	98.910	0.046	399228	98.690

HPLC of compound (**29**)

CHIRAL SFC REPORT

Compound ID : LMT-0014-258-P1C1  
 Sample ID : LMT-0014-258-P1C1\_G1  
 Injection Date : 3/17/2022  
 Acq Method : F:\Data\202203\20220317 1\OJ-H\_EtOH(DEA)\_5\_40\_25ML\_ ->  
 Raw Data : F:\DATA\202203\20220317 1\LMT-0014-258-P1C1\_G1.D  
 Instrument : CAS-SH-ANA-SFC-G (Agilent 1260 with DAD detector)  
 Method Comments : Column: ChiralCel OJ-H 150x4.6mm I.D., 5um  
 Mobile phase: A: CO2 B:Ethanol (0.05% DEA)  
 Gradient: from 5% to 40% of B in 4.5min , then 5% of B  
 for 1.5 min  
 Flow rate: 2.5mL/min Column temp.:40°C  
 Back pressure: 100 bar



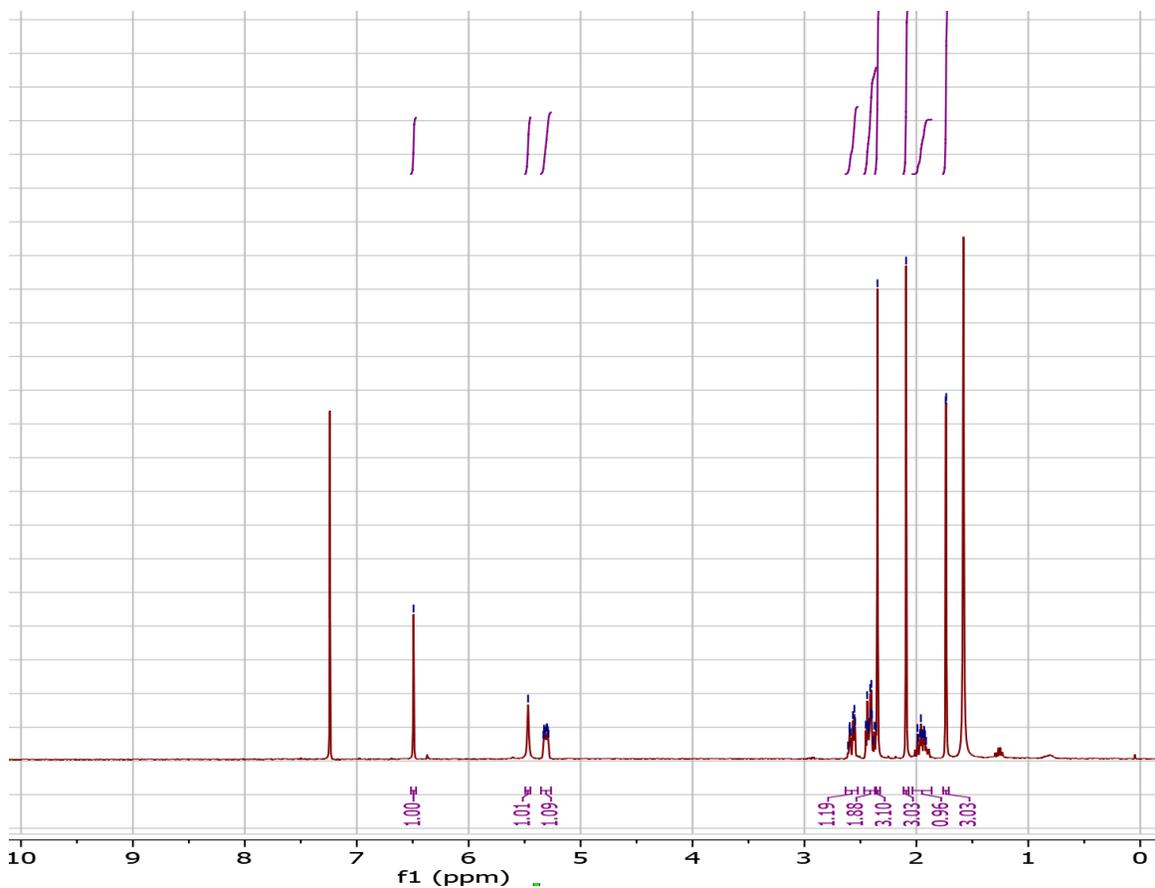
=====

DAD1 B, Sig=254,16 Ref=360,100

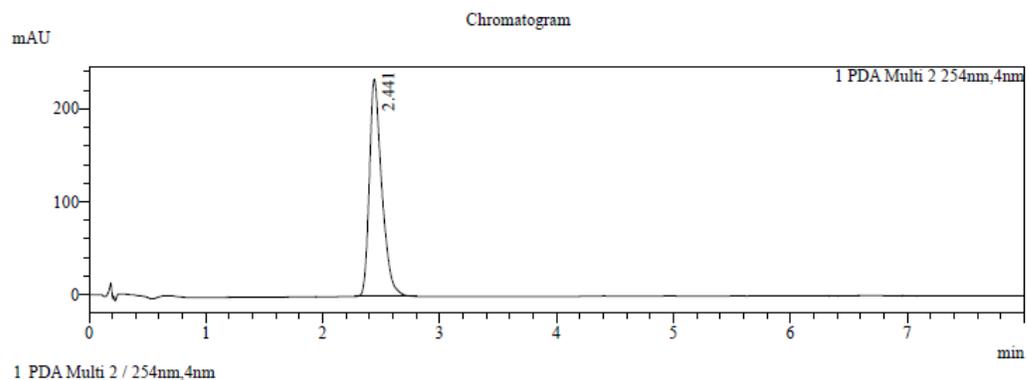
#	Meas.	Ret. Time	Height	Height %	Width	Area	Area %
1		3.054	86.611	100.000	0.037	208.779	100.000

=====

Chiral SFC of Compound (29)



<sup>1</sup>H NMR of compound (**38**) (400 MHz, CDCl<sub>3</sub>)



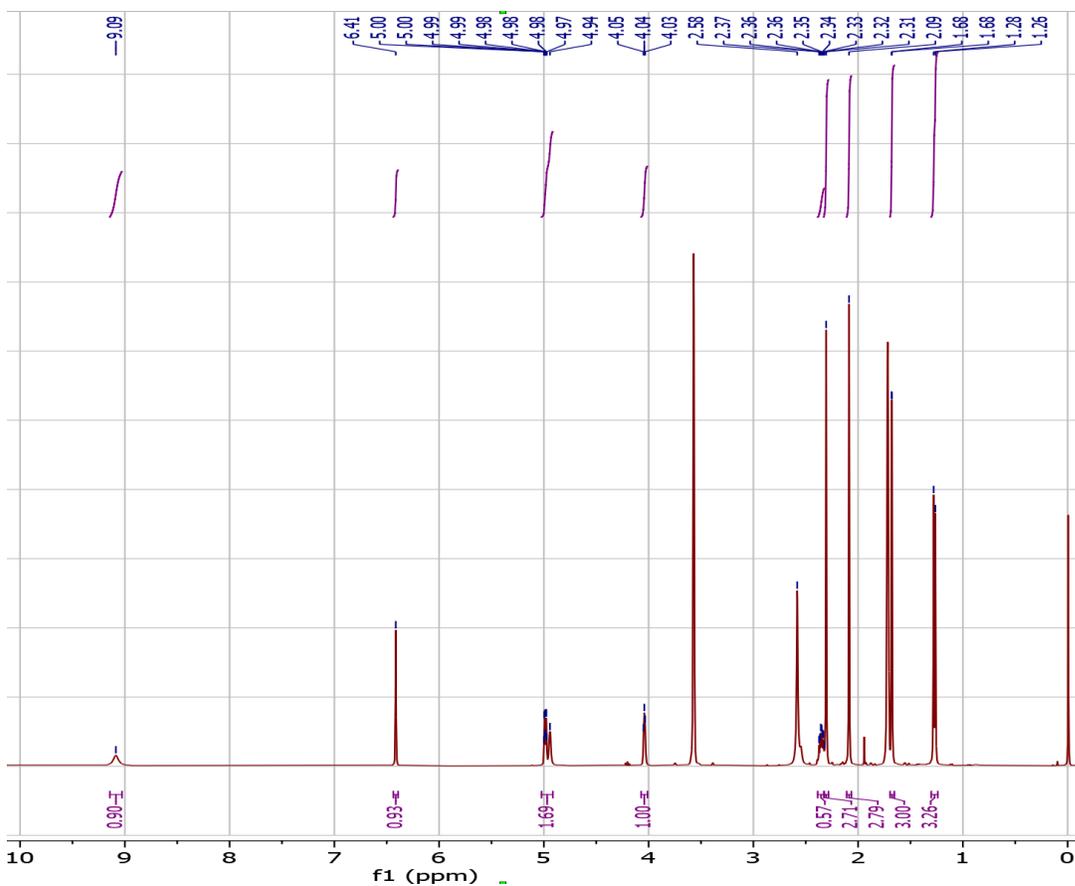
1 PDA Multi 2 / 254nm, 4nm

=====  
 Integration Result  
 =====

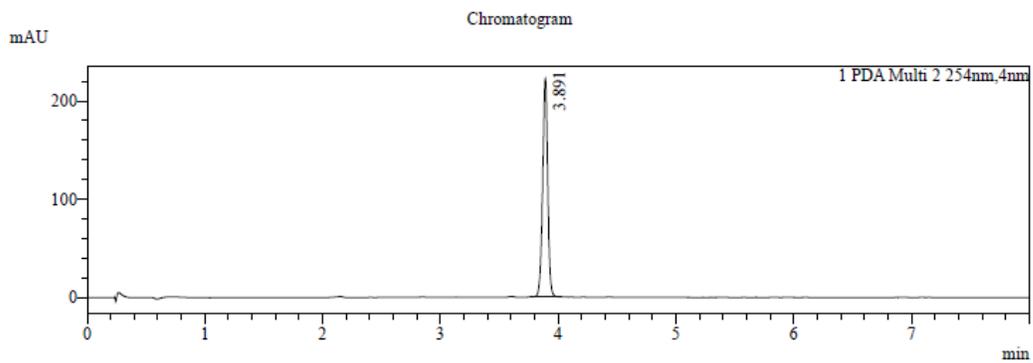
Peak Table

Peak#	Ret. Time	Height	Height%	USP Width	Area	Area%
1	2.441	232725	100.000	0.189	1731551	100.000

LCMS of compound (**38**)



$^1\text{H}$  NMR of compound (**39**) (400 MHz,  $\text{THF-}d_8$ )



1 PDA Multi 2 / 254nm,4nm

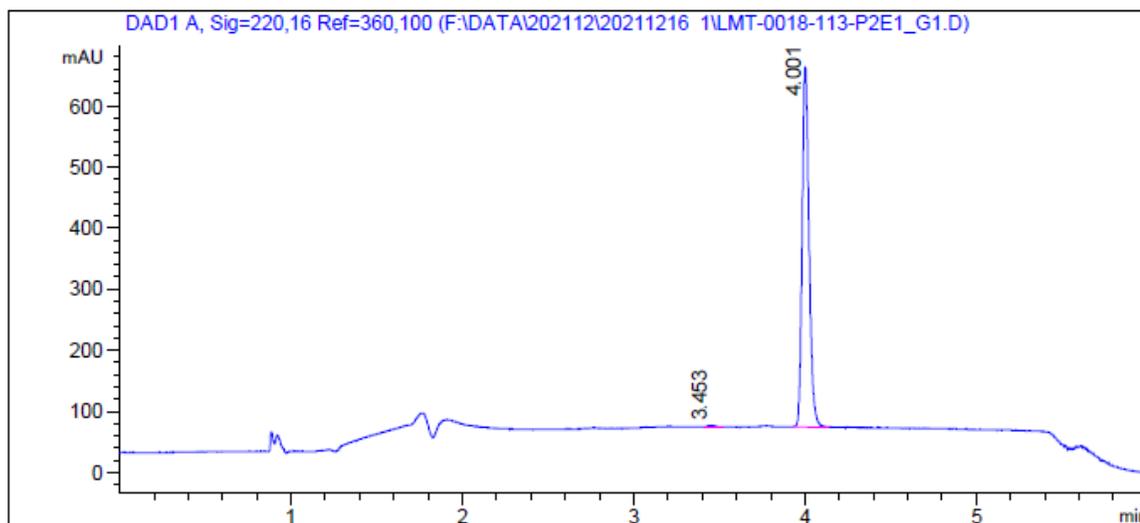
=====  
 Integration Result  
 =====

Peak Table						
Peak#	Ret. Time	Height	Height%	USP Width	Area	Area%
1	3.891	219675	100.000	0.079	668074	100.000

HPLC of compound (**39**)

CHIRAL SFC REPORT

Compound ID : 1  
 Sample ID : LMT-0018-113-P2E1\_G1  
 Injection Date : 12/16/2021  
 Acq Method : F:\Data\202112\20211216 1\AD-3\_EtOH(DEA)\_5\_40\_25ML\_ ->  
 Raw Data : F:\DATA\202112\20211216 1\LMT-0018-113-P2E1\_G1.D  
 Instrument : CAS-SH-ANA-SFC-G (Agilent 1260 with DAD detector)  
 Method Comments : Column: ChiralPak AD-3 150x4.6mm I.D., 3um  
 Mobile phase: A: CO2 B:Ethanol (0.05% DEA)  
 Gradient: from 5% to 40% of B in 4.5min , then 5% of B  
 for 1.5 min  
 Flow rate: 2.5mL/min Column temp.:40°C  
 Back pressure: 100 bar



=====

DAD1 A, Sig=220,16 Ref=360,100

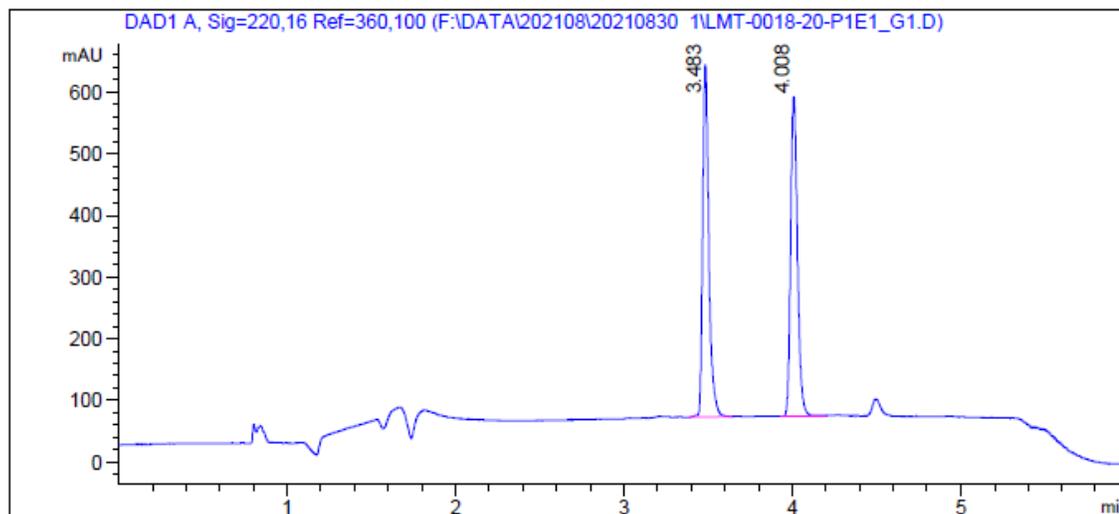
#	Meas. Ret. Time	Height	Height %	Width	Area	Area %
1	3.453	2.645	0.446	0.032	6.739	0.405
2	4.001	590.263	99.554	0.042	1658.531	99.595

-----

Chiral SFC of compound (39)

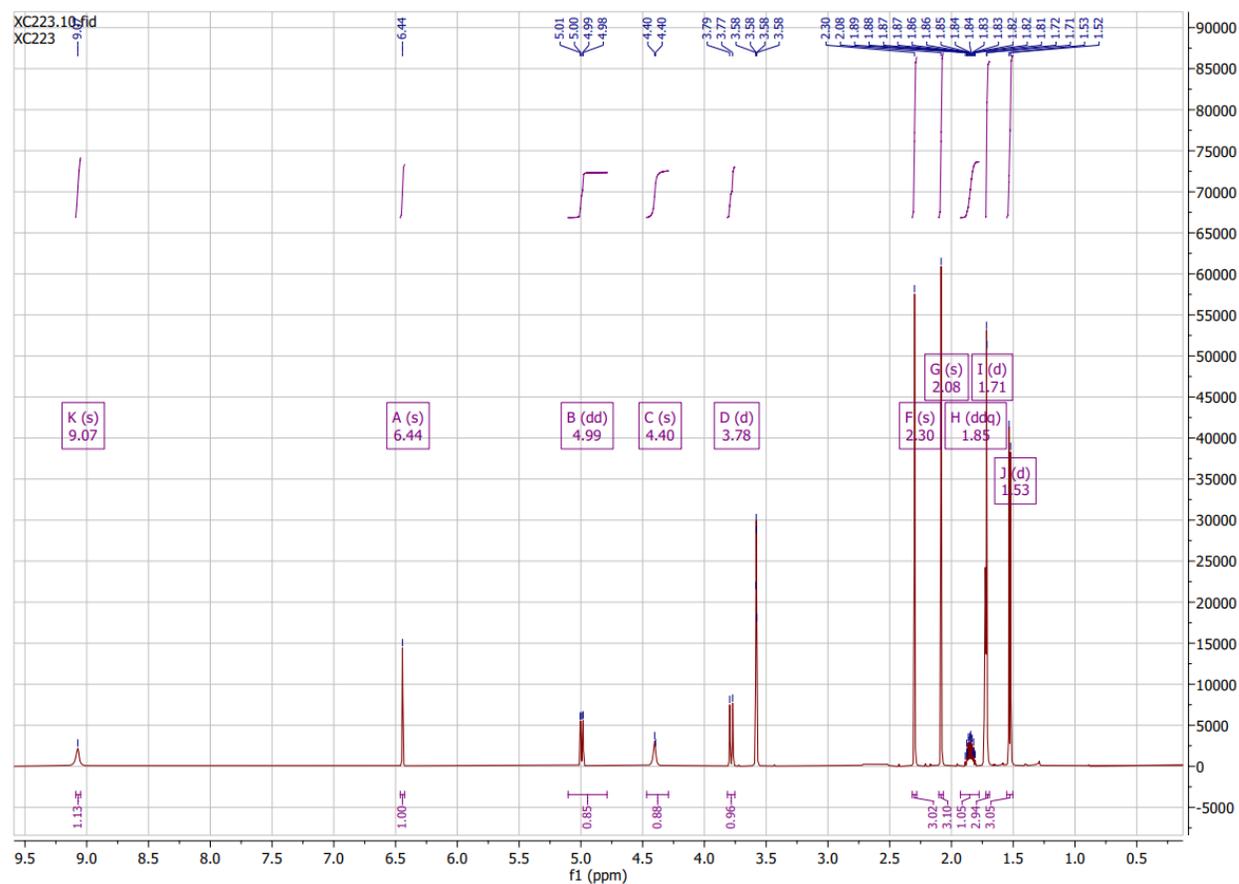
CHIRAL SFC REPORT

Compound ID : 1  
 Sample ID : LMT-0018-20-P1E1\_G1  
 Injection Date : 8/30/2021  
 Acq Method : F:\Data\202108\20210830 1\AD-3\_EtOH(DEA)\_5\_40\_25ML\_ ->  
 Raw Data : F:\DATA\202108\20210830 1\LMT-0018-20-P1E1\_G1.D  
 Instrument : CAS-SH-ANA-SFC-G (Agilent 1260 with DAD detector)  
 Method Comments : Column: ChiralPak AD-3 150x4.6mm I.D., 3um  
 Mobile phase: A: CO2 B:Ethanol (0.05% DEA)  
 Gradient: from 5% to 40% of B in 4.5min , then 5% of B  
 for 1.5 min  
 Flow rate: 2.5mL/min Column temp.:40°C  
 Back pressure: 100 bar

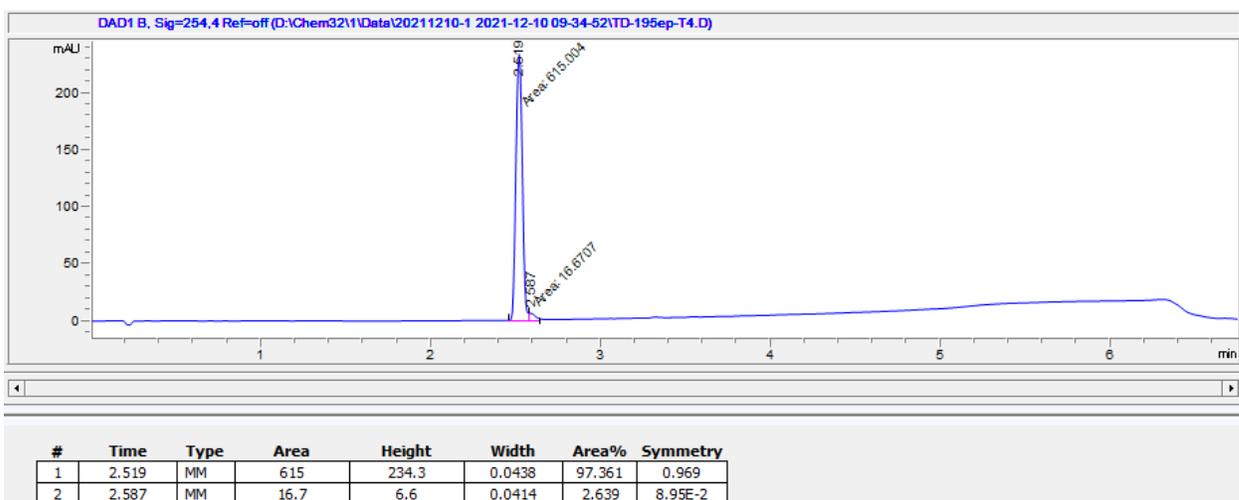


DAD1 A, Sig=220,16 Ref=360,100							
#	Meas.	Ret. Time	Height	Height %	Width	Area	Area %
1		3.483	572.260	52.433	0.039	1446.360	49.850
2		4.008	519.149	47.567	0.044	1455.092	50.150

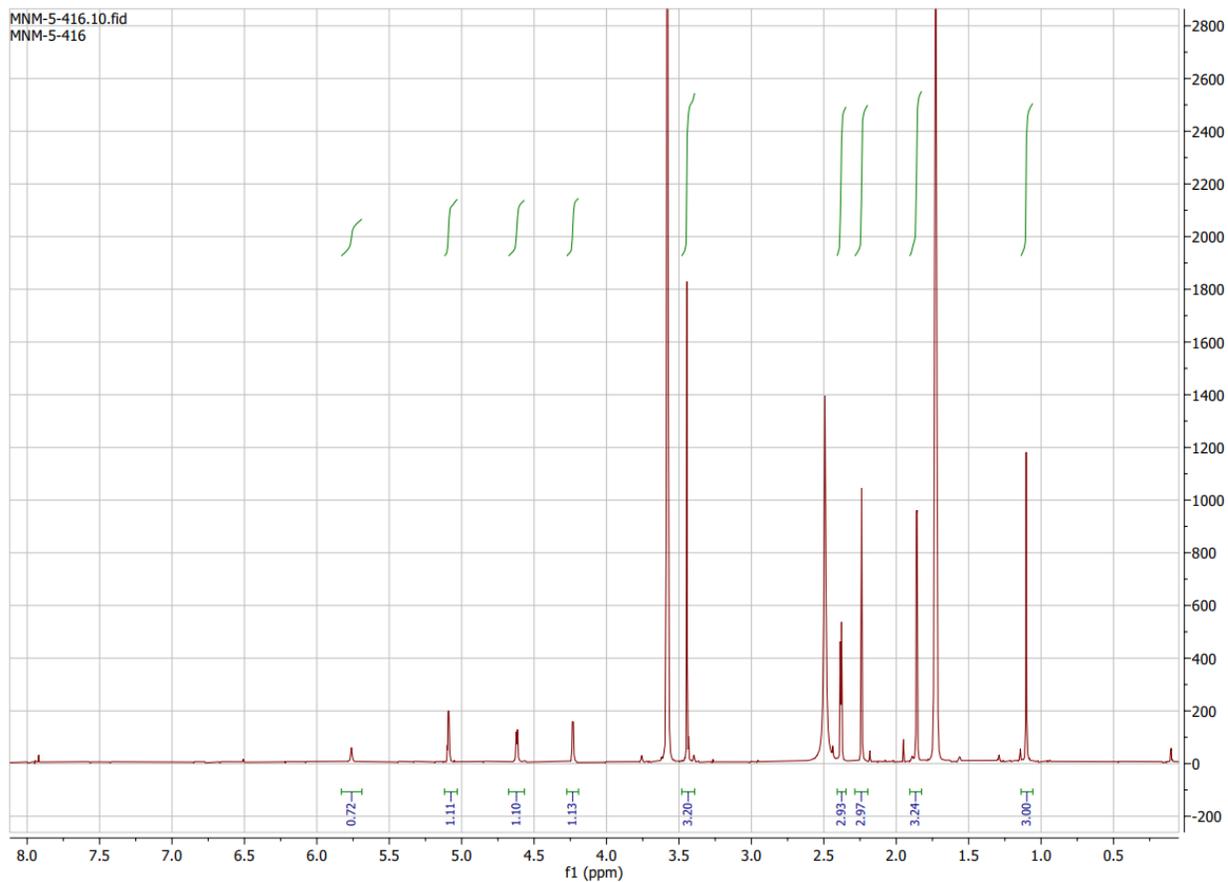
Chiral SFC of racemic reference for Chiral SFC of compound **39**.



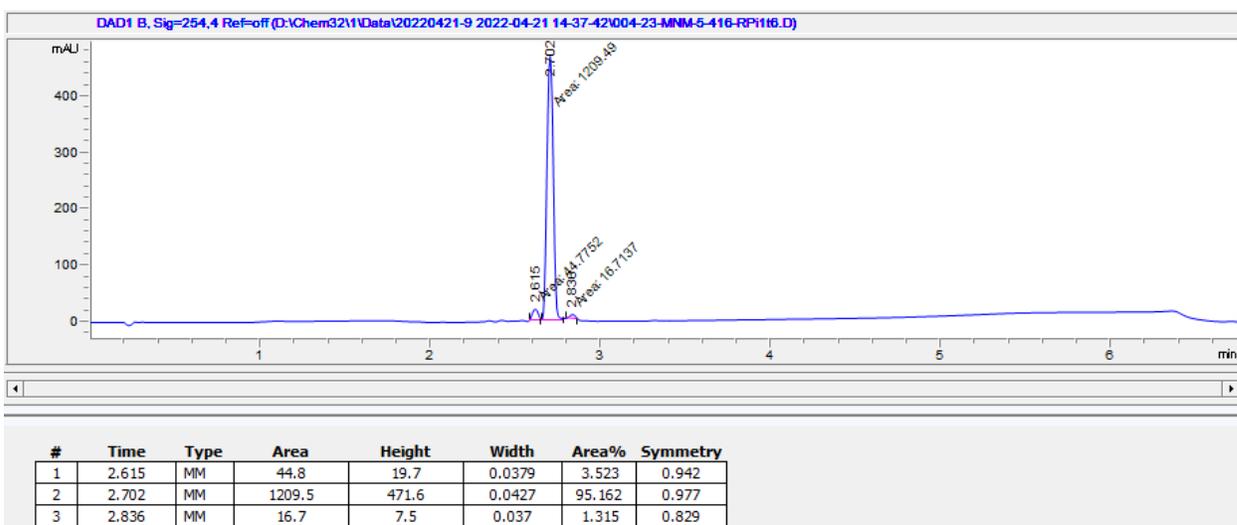
<sup>1</sup>H NMR of compound (40) (400 MHz, THF-d<sub>8</sub>)



LCMS of compound (40)

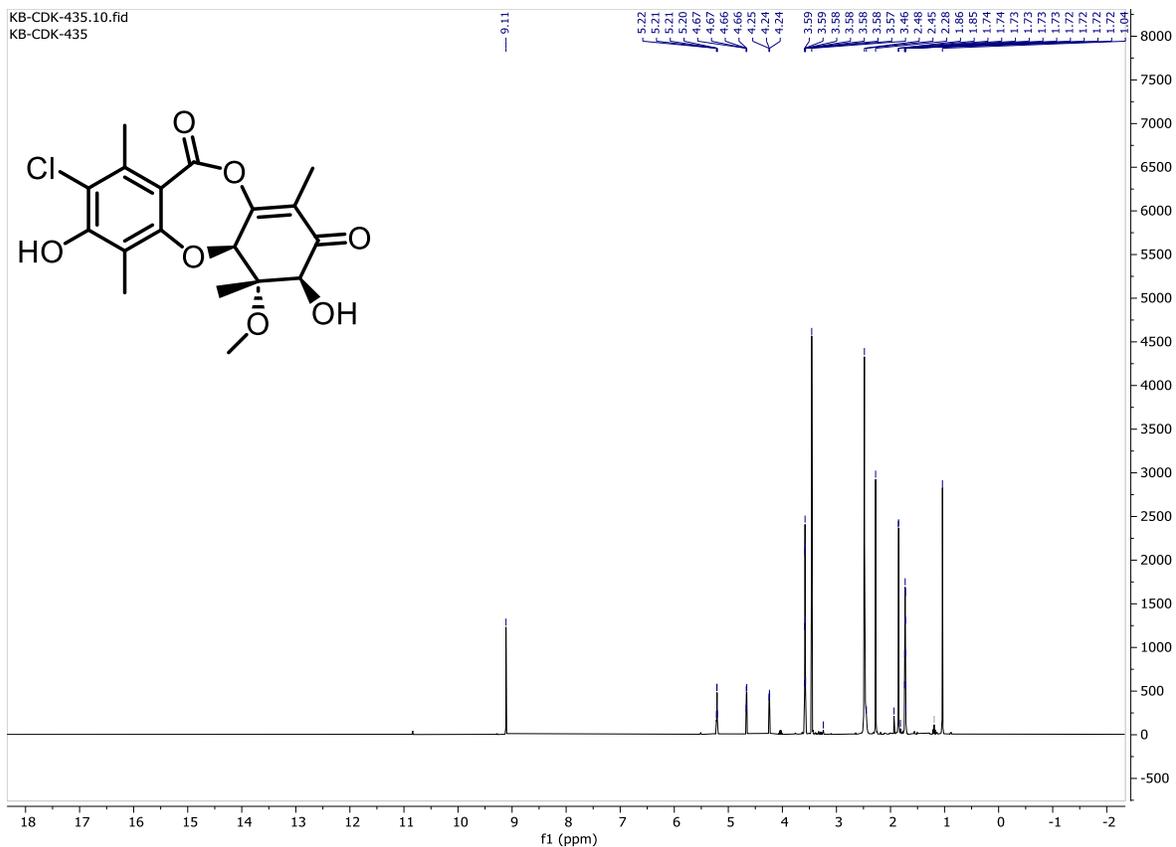


$^1\text{H}$  NMR of compound (**41**) (400 MHz,  $\text{THF-d}_8$ )

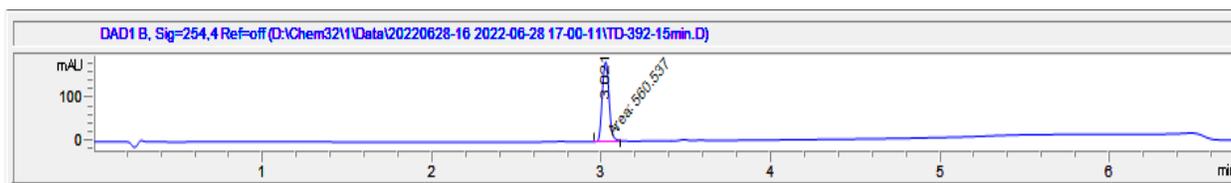


LCMS of compound (**41**)





$^1\text{H}$  NMR of compound (**43**) (400 MHz,  $\text{THF-d}_8$ )



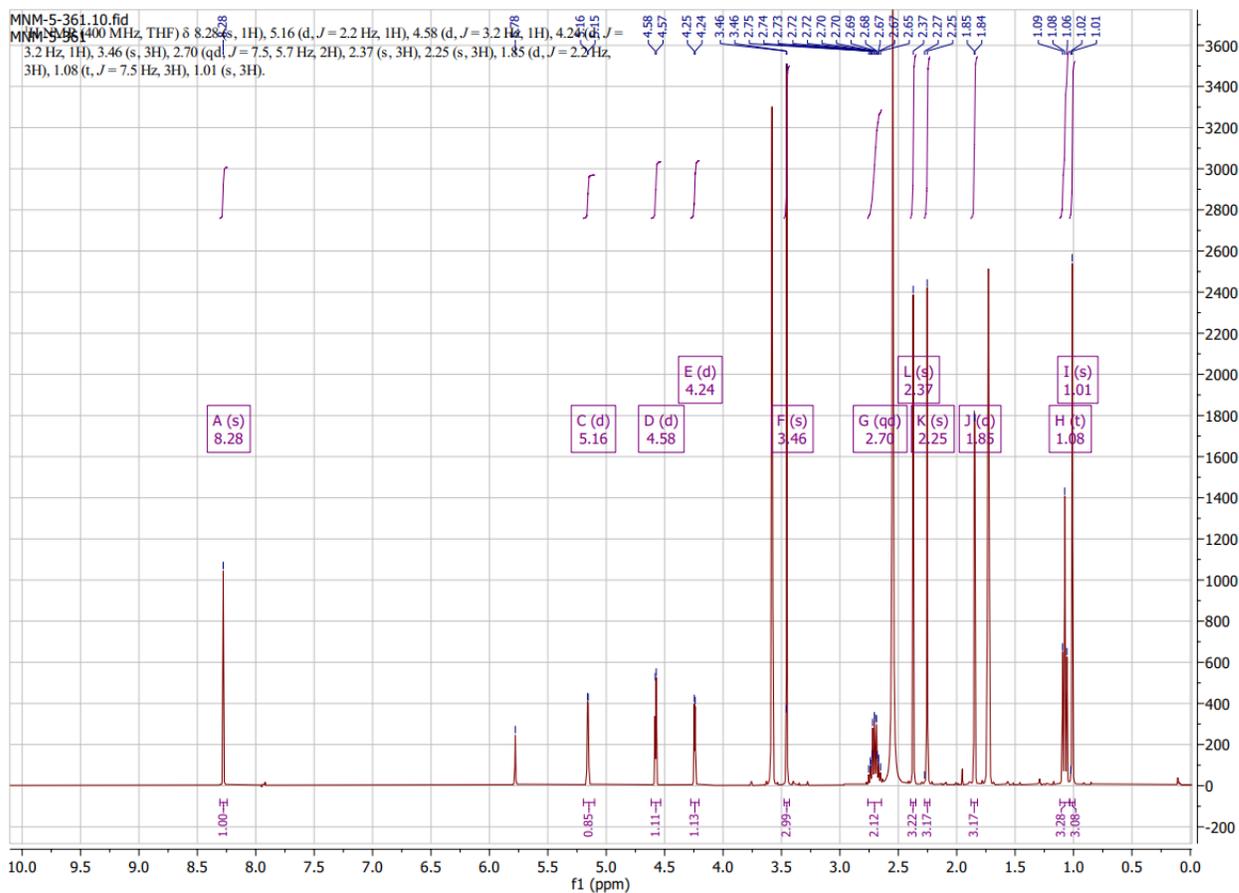
Signal 2: DAD1 B, Sig=254,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	3.021	MM	0.0483	560.53668	193.29146	100.0000

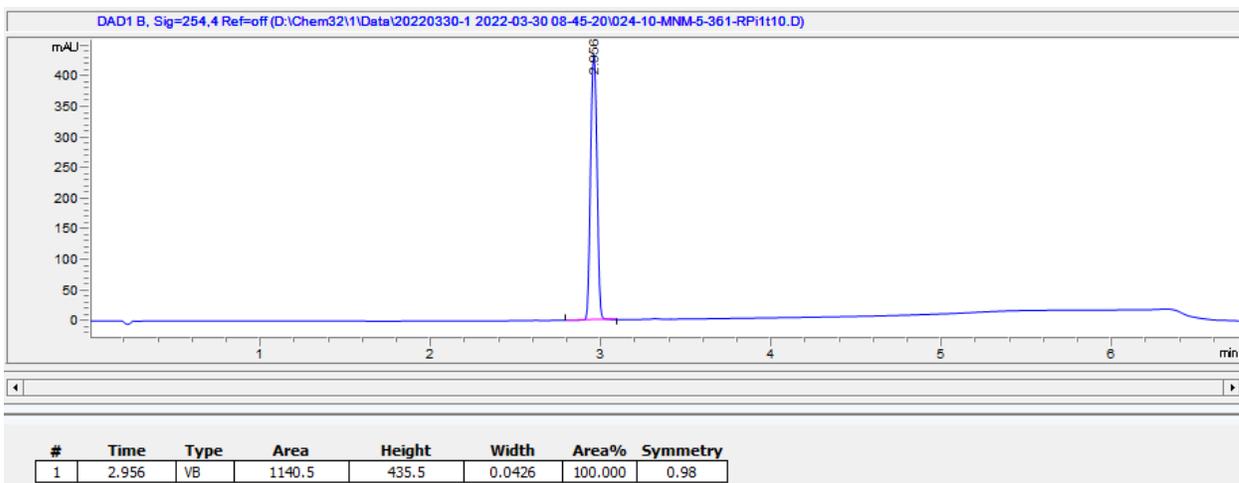
Totals : 560.53668 193.29146

LCMS of compound (**43**)



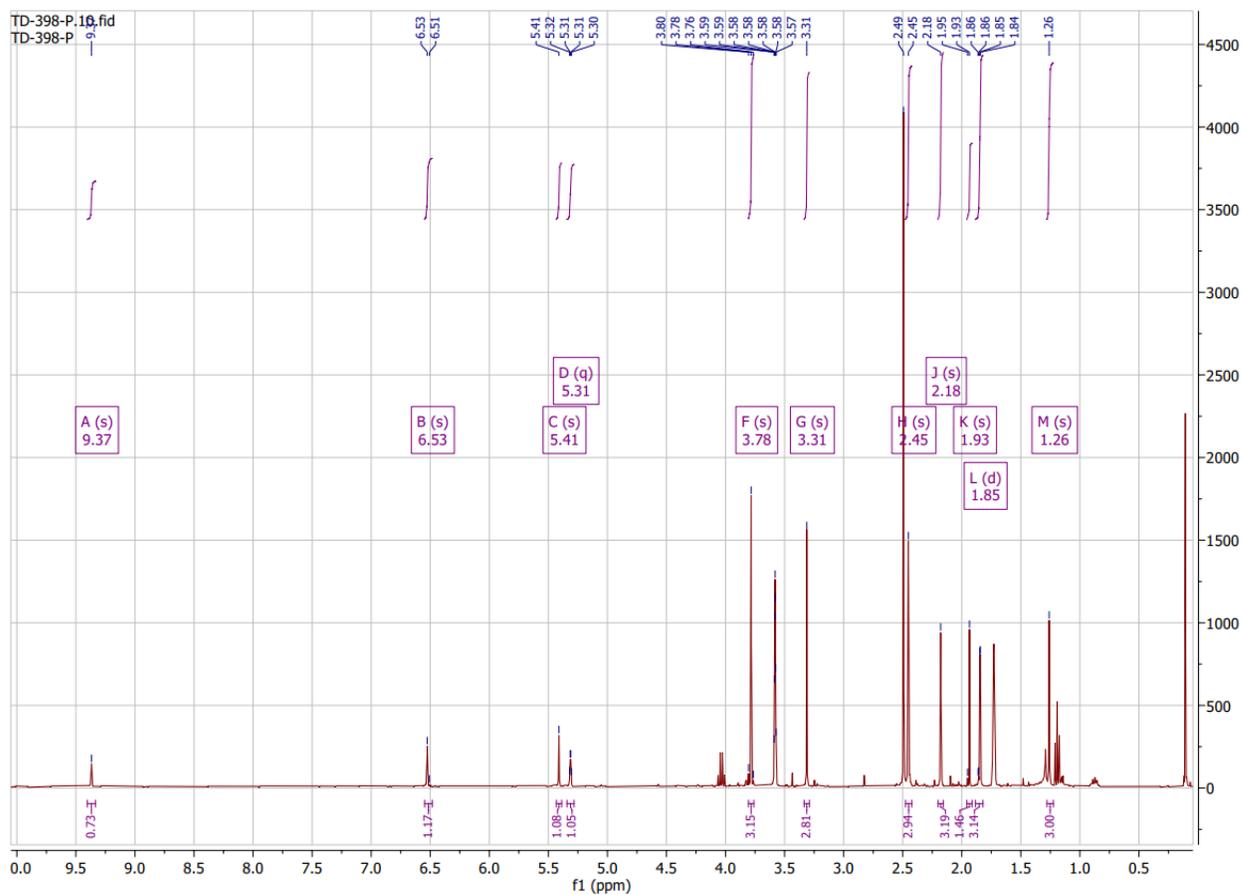


$^1\text{H}$  NMR of compound (**45**) (400 MHz, THF- $d_8$ )

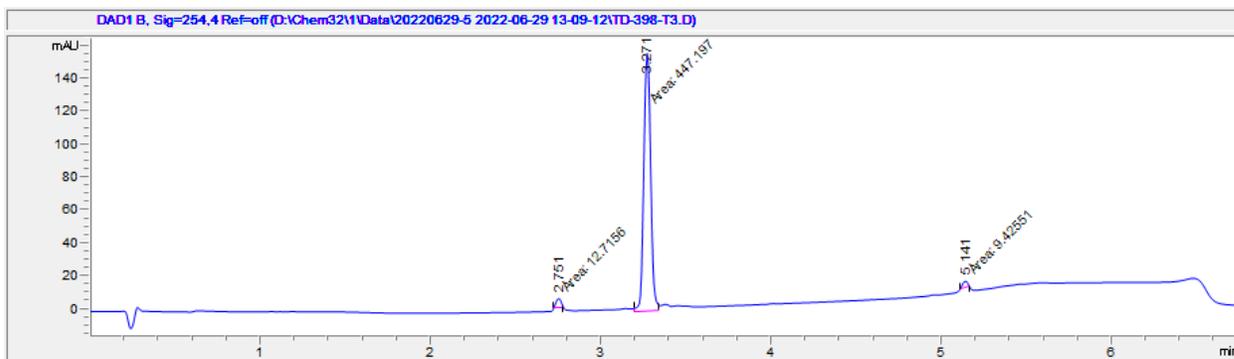


LCMS of compound (**45**)



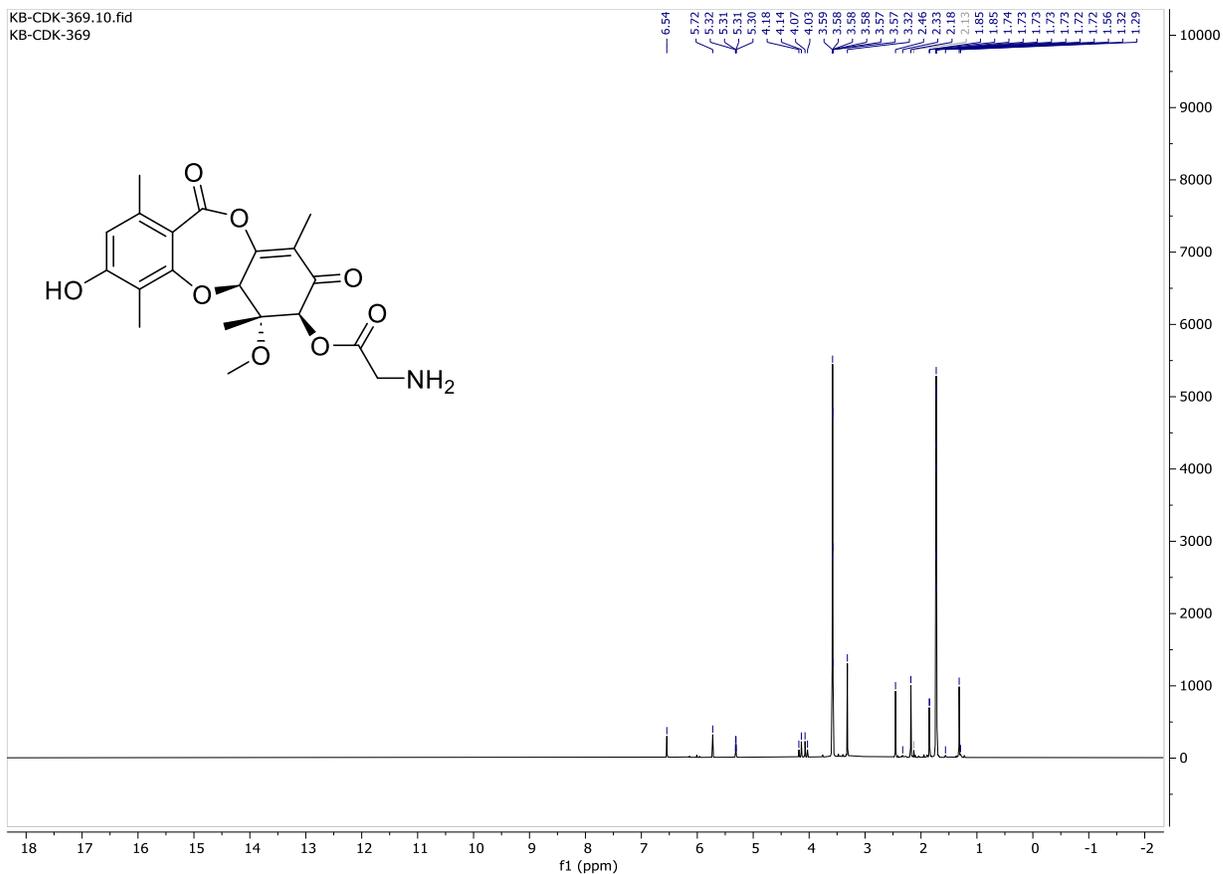


$^1\text{H}$  NMR of compound (**48**) (400 MHz,  $\text{THF-d}_8$ )

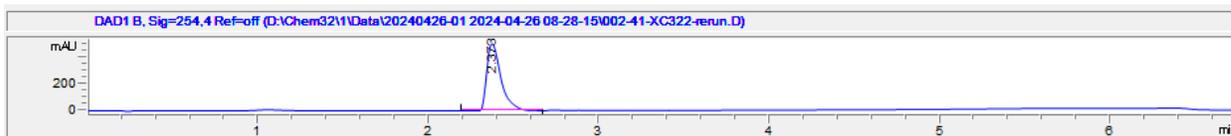


#	Time	Type	Area	Height	Width	Area%	Symmetry
1	2.751	MM	12.7	6	0.035	2.709	1.041
2	3.271	MM	447.2	158.5	0.047	95.282	0.98
3	5.141	MM	9.4	4.5	0.0348	2.008	1.28

LCMS of compound (**48**)



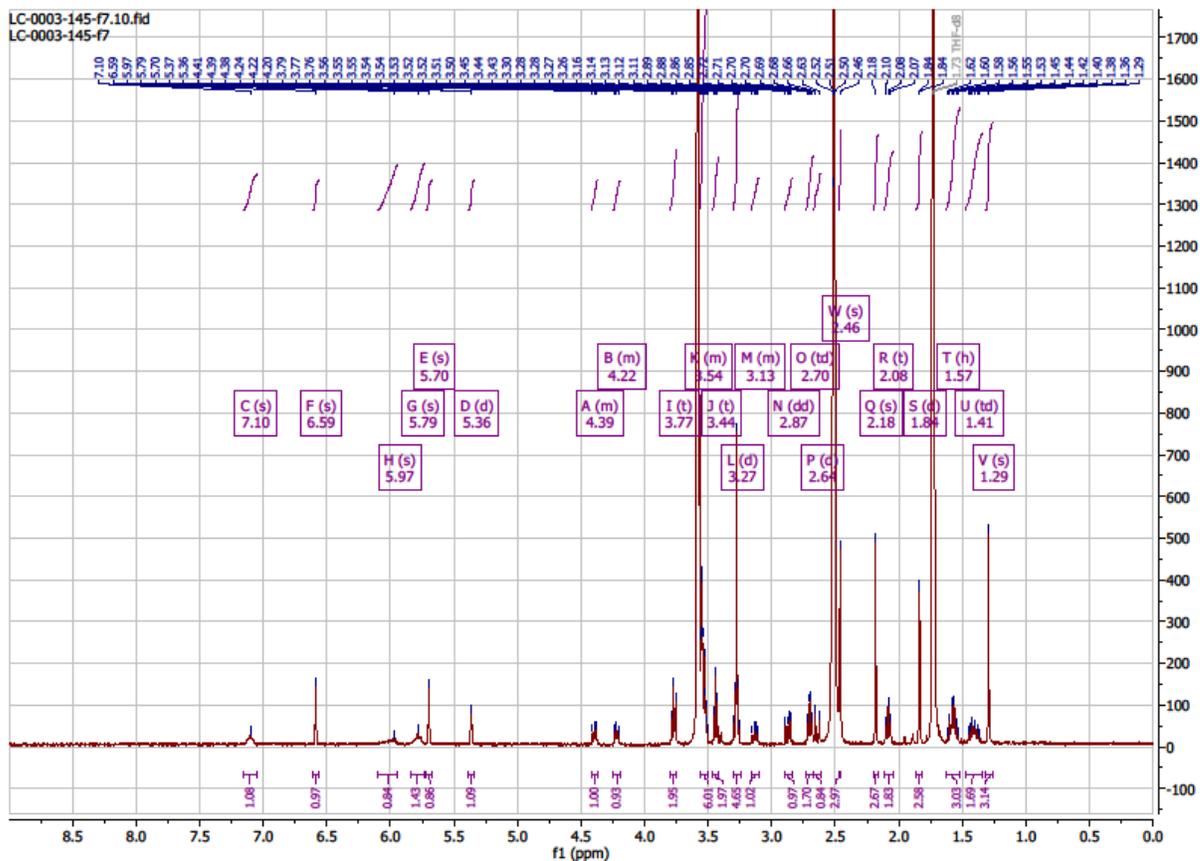
<sup>1</sup>H NMR of compound (49) (400 MHz, THF-*d*<sub>8</sub>)



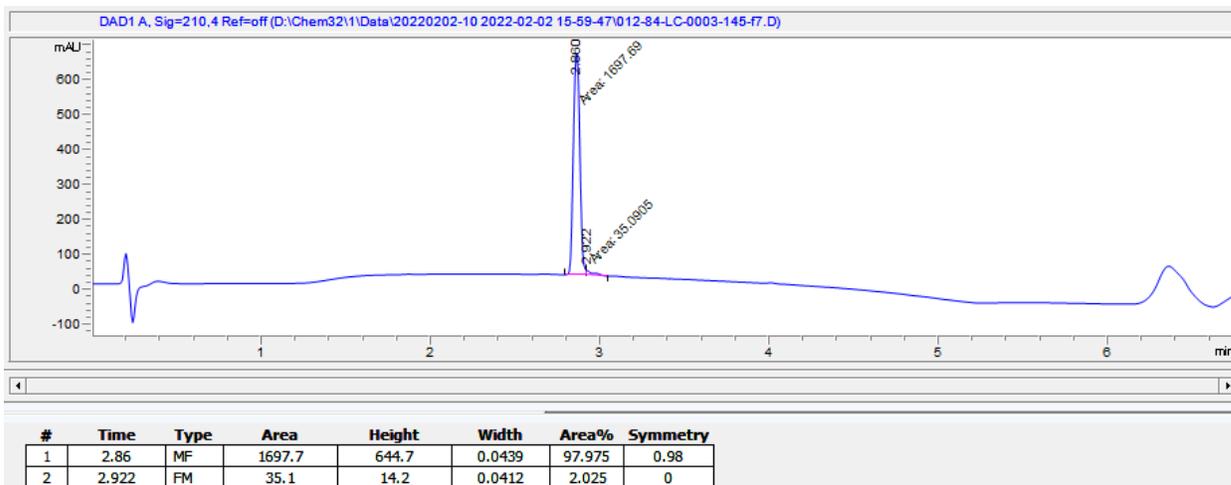
Signal 2: DAD1 B, Sig=254,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.373	BV	0.0873	2933.30444	508.43777	100.0000

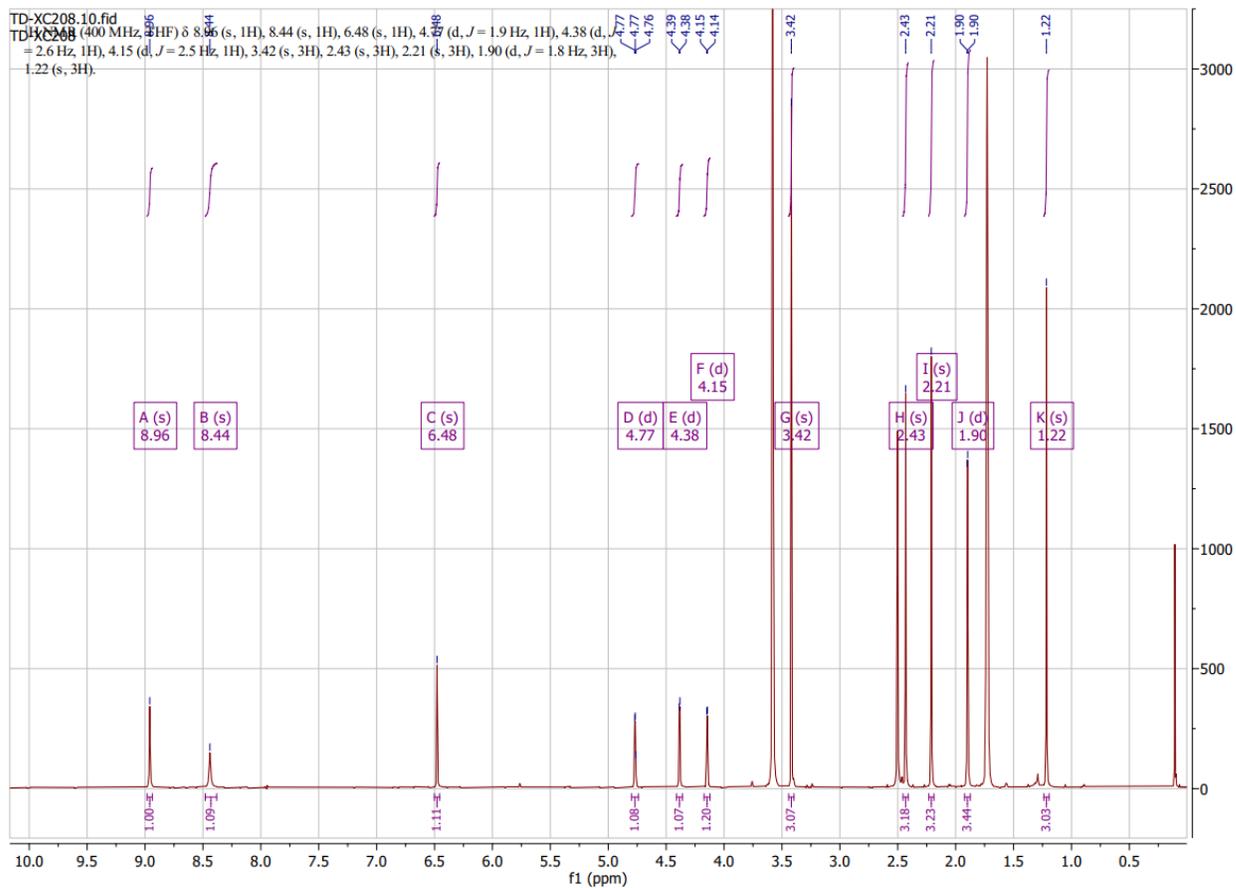
LCMS of compound (49)



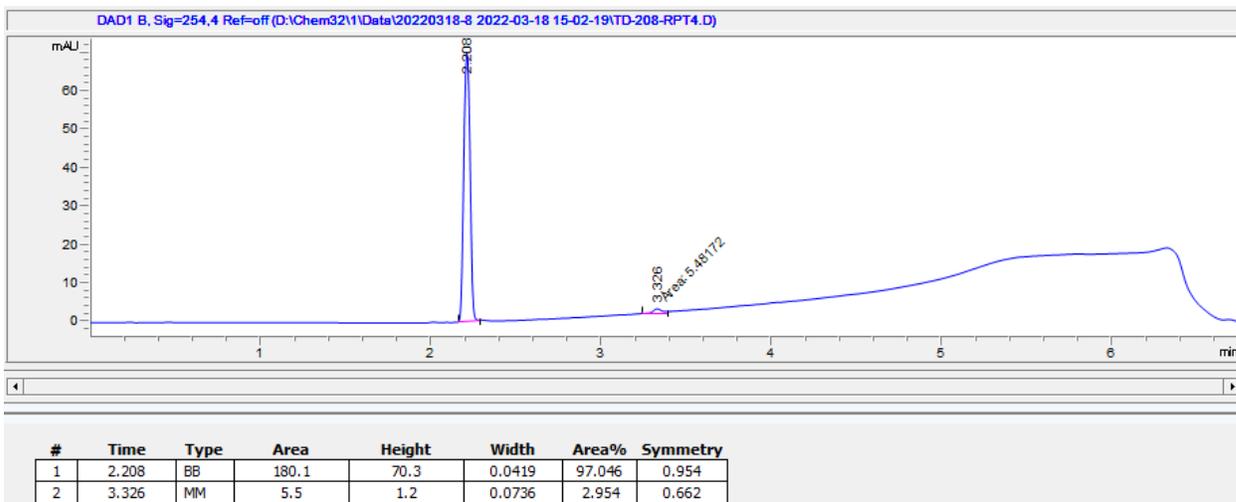
$^1\text{H}$  NMR of compound (**50**) (400 MHz,  $\text{THF-}d_8$ )



LCMS of compound (**50**)



<sup>1</sup>H NMR of compound (**51**) (400 MHz, THF-d<sub>8</sub>)



LCMS of compound (**51**)

# Small-molecule X-ray Crystallography Reports

X-ray crystallography report for compound **2**

## Compound 2

Submitted by: **Stuart Romeril**

Solved by: **John Bacsá**

**$R_1=2.1\%$**

## Crystal Data and Experimental

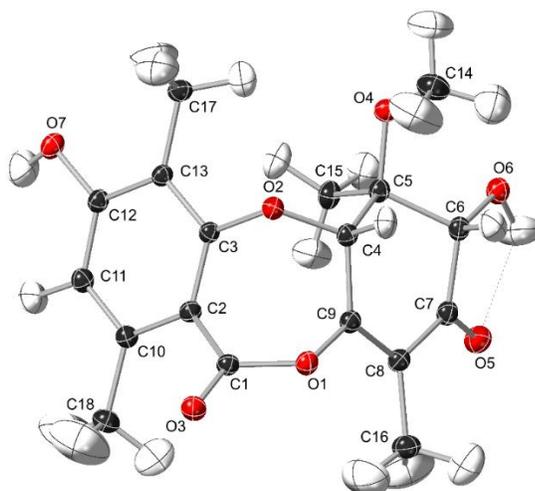


Figure S11. X-ray crystal structure of **2**

**Experimental.** Excellent quality, single, colorless plate/prism-shaped crystals of compound **2** were chosen from the sample as supplied. A suitable crystal with dimensions  $0.21 \times 0.18 \times 0.15 \text{ mm}^3$  was selected and mounted on a loop with paratone on a Rigaku XtaLAB Synergy S diffractometer. The crystal was kept at a constant  $T = 100.00(10) \text{ K}$  during data collection. The structure was solved with the ShelXT 2018/2 (Sheldrick, 2018) solution program using dual methods and by using Olex2 1.5-alpha (Dolomanov et al., 2009) as the graphical interface. The model was refined with olex2.refine 1.5-alpha (Bourhis et al., 2015) using full matrix least squares minimisation on  $F^2$ .

**Crystal Data.**  $\text{C}_{18}\text{H}_{20}\text{O}_7$ ,  $M_r = 348.355$ , orthorhombic,  $P2_12_12_1$  (No. 19),  $a = 7.9775(1) \text{ \AA}$ ,  $b = 11.7069(2) \text{ \AA}$ ,  $c = 17.2946(2) \text{ \AA}$ ,  $\alpha = \beta = \gamma = 90^\circ$ ,  $V = 1615.17(4) \text{ \AA}^3$ ,  $T = 100.00(10) \text{ K}$ ,  $Z = 4$ ,  $Z' = 1$ ,  $\mu(\text{Cu K}\alpha) = 0.931$ , 22253 reflections measured, 3380 unique ( $R_{\text{int}} = 0.0607$ ) which were used in all calculations. The final  $wR_2$  was 0.0510 (all data) and  $R_1$  was 0.0210 ( $I \geq 2 \sigma(I)$ ).

<b>Compound</b>	<b>2</b>
Formula	C <sub>18</sub> H <sub>20</sub> O <sub>7</sub>
<i>D</i> <sub>calc.</sub> / g cm <sup>-3</sup>	1.433
<i>μ</i> /mm <sup>-1</sup>	0.931
Formula Weight	348.355
Color	colorless
Shape	prism-shaped
Size/mm <sup>3</sup>	0.21×0.18×0.15
<i>T</i> /K	100.00(10)
Crystal System	orthorhombic
Flack Parameter	0.05(6)
Hooft Parameter	0.05(6)
Space Group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
<i>a</i> /Å	7.9775(1)
<i>b</i> /Å	11.7069(2)
<i>c</i> /Å	17.2946(2)
<i>α</i> /°	90
<i>β</i> /°	90
<i>γ</i> /°	90
<i>V</i> /Å <sup>3</sup>	1615.17(4)
<i>Z</i>	4
<i>Z</i> '	1
Wavelength/Å	1.54184
Radiation type	Cu K <sub>α</sub>
2θ range/°	10.22 to 159.92
	(0.78 Å)
Index ranges	-10 ≤ <i>h</i> ≤ 9 -14 ≤ <i>k</i> ≤ 14 -21 ≤ <i>l</i> ≤ 22
Reflections collected	22253
Independent reflections	3380
	<i>R</i> <sub>int</sub> = 0.0607
	<i>R</i> <sub>sigma</sub> = 0.0330
Refl's I ≥ 2 σ( <i>I</i> )	3302
Parameters	406
Restraints	390
Largest peak/hole/eÅ <sup>-3</sup>	0.11/-0.10
Goof	1.1022
<i>wR</i> <sub>2</sub> (all data)	0.0510
<i>wR</i> <sub>2</sub>	0.0506
<i>R</i> <sub>1</sub> (all data)	0.0217
<i>R</i> <sub>1</sub>	0.0210

**Table S7.** Crystallography statistics for compound **2**

## Structure Quality Indicators

**Reflections:**  $d \min (\text{Cu}\backslash\text{a}) [2\theta=159.9^\circ] = 0.78$ ;  $I/\sigma(I) = 30.3$ ;  $R_{\text{int}} = 6.07\%$ ; Full  $135.4^\circ$  (96% to  $159.9^\circ$ ) = 99.7

**Refinement:** Shift = -0.001; Max Peak = 0.1; Min Peak = -0.1;  $GooF = 1.102$ ;  $Hooft = 0.05(6)$

A colourless plate-shaped crystal with dimensions  $0.21 \times 0.18 \times 0.15 \text{ mm}^3$  was mounted on a loop with paratone. Data were collected using a XtaLAB Synergy, Dualflex, HyPix diffractometer equipped with an Oxford Cryosystems low-temperature device operating at  $T = 100.00(10) \text{ K}$ .

Data were measured using  $\omega$  scans with  $\text{Cu K}\alpha$  radiation. The diffraction pattern was indexed and the total number of runs and images was based on the strategy calculation from the program CrysAlisPro 1.171.41.98a (Rigaku Oxford Diffraction, 2021). The maximum resolution that was achieved was  $\Theta = 79.96^\circ$  ( $0.78 \text{ \AA}$ ).

The unit cell was refined using CrysAlisPro 1.171.41.98a (Rigaku OD, 2021) on 4689 reflections, 21% of the observed reflections.

Data reduction, scaling and absorption corrections were performed using CrysAlisPro 1.171.41.98a (Rigaku OD, 2021). The final completeness is 99.88 % out to  $79.96^\circ$  in  $\Theta$ . A numerical absorption correction based on gaussian integration over a multifaceted crystal model was performed using CrysAlisPro 1.171.41.98a (Rigaku OD, 2021). An empirical absorption correction using spherical harmonics, implemented in SCALE3 ABSPACK scaling algorithm was also performed. The absorption coefficient  $\mu$  of this material is  $0.931 \text{ mm}^{-1}$  at this wavelength ( $\lambda = 1.54184 \text{ \AA}$ ) and the minimum and maximum transmissions are 0.567 and 1.000.

The structure was solved and the space group  $P2_12_12_1$  (# 19) determined by the ShelXT 2018/2 structure solution program<sup>3</sup> using dual methods and refined by full matrix least squares minimisation on  $F^2$  using version of olex2.refine 1.5-alpha.<sup>4</sup> All atoms, including hydrogen atoms, were refined anisotropically. Hydrogen atom positions were located from the electron densities and freely refined using Hirshfeld scattering factors. Refinement was by using NoSpherA2, an implementation of non-spherical atom-form-factors.<sup>5</sup> NoSpherA2 implementation of HAR makes use of tailor-made aspherical atomic form factors calculated from a Hirshfeld-partitioned electron density (ED) not from spherical-atom form factors. The ED was calculated from a Gaussian basis set single determinant SCF wavefunction from DFT using selected functionals for a fragment of this crystal. SOFTWARE: ORCA 5.0 PARTITIONING: NoSpherA2 INT ACCURACY: Normal METHOD: PBE0 BASIS SET: def2-TZVP CHARGE: 0 MULTIPLICITY: 1 SOLVATION: Ethanol DATE: 2022-08-23\_20-00-57

There is a single formula unit in the asymmetric unit, which is represented by the reported sum formula. In other words: Z is 4 and Z' is 1. The moiety formula is  $\text{C}_{18} \text{H}_{20} \text{O}_7$ .

The Flack parameter was refined to 0.05(6). Determination of absolute structure using Bayesian statistics on Bijvoet differences using the Olex2 results in 0.05(6). The chiral atoms in this structure are: C4(R), C5(S), C6(S). Note: The Flack parameter is used to determine chirality of the crystal studied, the value should be near 0, a value of 1 means that the stereochemistry is wrong and the model should be inverted. A value of 0.5 means that the crystal consists of a racemic mixture of the two enantiomers.

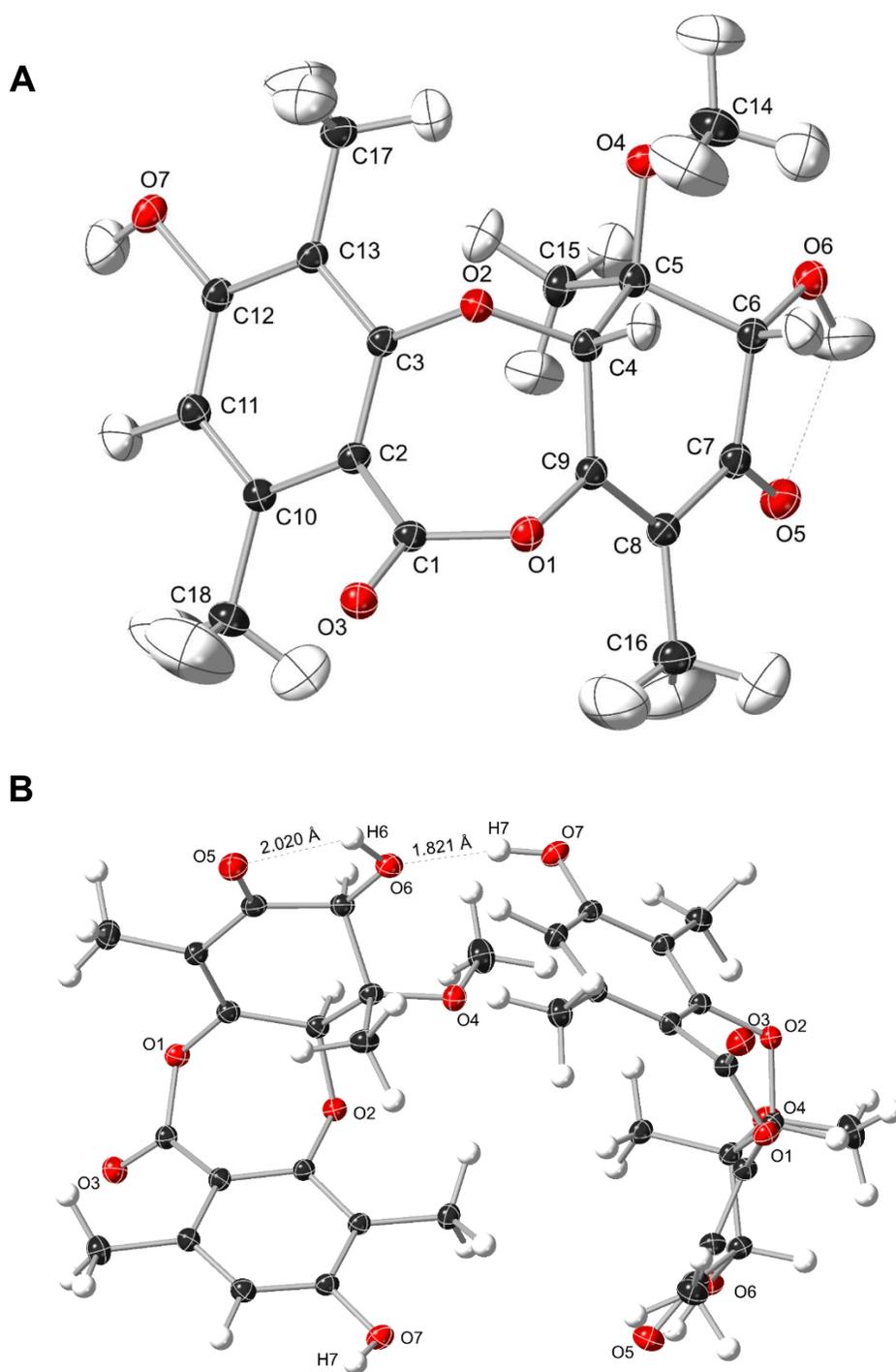


Figure S12. A) The molecular structure of compound **2**. There are three chiral centers, labelled C4, C5 and C6 and they have *S*, *S* and *S* configuration respectively. There is a single molecule in the asymmetric unit, which is represented by the reported sum formula. In other words: *Z* is 4 and *Z'* is 1. B) The intra and intermolecular hydrogen bonds.



Total reflections (after filtering)	22253	Unique reflections	3400
Completeness	0.965	Mean $I/\sigma$	24.26
$hkl_{\max}$ collected	(9, 14, 22)	$hkl_{\min}$ collected	(-10, -14, -21)
$hkl_{\max}$ used	(10, 14, 22)	$hkl_{\min}$ used	(-10, 0, 0)
Lim $d_{\max}$ collected	100.0	Lim $d_{\min}$ collected	0.77
$d_{\max}$ used	9.69	$d_{\min}$ used	0.78
Friedel pairs	2331	Friedel pairs merged	0
Inconsistent equivalents	3	$R_{\text{int}}$	0.0607
$R_{\text{sigma}}$	0.033	Intensity transformed	0
Omitted reflections	0	Omitted by user (OMIT)	62
		hkl)	
Multiplicity	(2935, 2216, 1610, 1008, 561, 277, 138, 45, 22, 0, 3)	Maximum multiplicity	15
Removed systematic absences	0	Filtered off (Shel/OMIT)	0

**Table S8.** Reflection statistics for crystal structure of **2**

Atom	x	y	z	$U_{eq}$
O2	0.85727(7)	0.53243(5)	0.44568(3)	0.01569(12)
O1	0.64322(8)	0.70229(5)	0.49422(3)	0.01803(12)
O4	1.00990(8)	0.38519(5)	0.56405(4)	0.02019(13)
O3	0.52665(8)	0.75078(5)	0.38403(4)	0.02097(13)
O6	0.84611(9)	0.38323(5)	0.71215(3)	0.02167(13)
O7	0.63254(9)	0.24702(5)	0.28432(3)	0.02151(13)
O5	0.56197(8)	0.49171(6)	0.72232(4)	0.02436(14)
C12	0.60859(11)	0.34448(7)	0.32427(5)	0.01600(16)
C13	0.74714(10)	0.38779(6)	0.36445(5)	0.01509(15)
C1	0.57545(10)	0.67334(6)	0.42387(5)	0.01571(15)
C3	0.72280(10)	0.48932(6)	0.40520(4)	0.01395(15)
C2	0.57262(11)	0.55157(7)	0.40090(4)	0.01505(15)
C10	0.43375(11)	0.50589(7)	0.36132(4)	0.01577(15)
C7	0.65127(12)	0.52474(7)	0.66971(4)	0.01816(16)
C11	0.45318(11)	0.40055(7)	0.32507(5)	0.01763(17)
C9	0.68252(10)	0.61705(6)	0.54666(5)	0.01627(16)
C17	0.91242(11)	0.32667(7)	0.36500(5)	0.01856(17)
C6	0.82586(11)	0.47534(7)	0.66049(5)	0.01745(17)
C4	0.83701(10)	0.54913(6)	0.52718(4)	0.01559(16)
C8	0.59799(11)	0.61265(7)	0.61391(5)	0.01781(17)
C5	0.85175(11)	0.43774(7)	0.57528(5)	0.01592(15)
C15	0.72444(12)	0.34540(7)	0.55501(5)	0.01953(17)
C18	0.26739(11)	0.56622(7)	0.35802(5)	0.02081(17)
C16	0.45457(12)	0.68901(8)	0.63493(6)	0.02416(19)
C14	1.15577(12)	0.44883(8)	0.58169(6)	0.02668(19)
H11	0.3489(16)	0.3631(10)	0.2938(7)	0.034(3)
H17a	0.996(2)	0.3615(12)	0.3216(9)	0.045(3)
H4	0.9419(14)	0.6039(9)	0.5437(6)	0.023(2)
H14a	1.2624(18)	0.3953(12)	0.5687(9)	0.051(3)
H14b	1.1654(19)	0.5321(13)	0.5495(10)	0.059(4)
H15a	0.7515(17)	0.3064(10)	0.4996(7)	0.033(3)
H6a	0.9178(15)	0.5435(10)	0.6722(7)	0.030(3)
H15b	0.7320(19)	0.2789(10)	0.6011(7)	0.037(3)
H15c	0.5969(16)	0.3774(9)	0.5529(8)	0.033(3)
H16a	0.352(2)	0.6462(14)	0.6557(11)	0.070(4)
H14c	1.1626(18)	0.4740(14)	0.6410(8)	0.050(3)
H17b	0.9802(18)	0.3369(12)	0.4200(7)	0.043(3)
H16b	0.486(2)	0.7396(15)	0.6875(9)	0.062(4)
H18a	0.273(2)	0.6371(15)	0.3185(12)	0.077(5)
H17c	0.8978(18)	0.2351(10)	0.3520(10)	0.048(4)
H7	0.525(2)	0.2116(13)	0.2752(8)	0.038(3)
H16c	0.415(2)	0.7443(15)	0.5900(9)	0.072(5)
H18b	0.229(2)	0.6010(17)	0.4118(9)	0.071(5)
H18c	0.166(2)	0.5141(14)	0.3404(12)	0.080(5)
H6	0.753(2)	0.3853(14)	0.7486(9)	0.050(5)

**Table S9.** Fractional Atomic Coordinates ( $\times 10^4$ ) and Equivalent Isotropic Displacement Parameters ( $\text{\AA}^2$ ) for compound **2**.  $U_{eq}$  is defined as  $1/3$  of the trace of the orthogonalised  $U_{ij}$ .

Atom	$U_{11}$	$U_{22}$	$U_{33}$	$U_{23}$	$U_{13}$	$U_{12}$
O2	15.6(3)	15.6(3)	15.9(2)	-1.6(2)	0.5(2)	0.0(2)
O1	24.0(3)	12.7(2)	17.4(3)	1.2(2)	-2.1(2)	-1.2(2)
O4	22.1(3)	18.1(3)	20.3(3)	4.0(2)	-0.3(2)	-0.5(2)
O3	27.6(3)	13.9(3)	21.5(3)	3.7(2)	-3.6(2)	1.2(2)
O6	29.0(3)	20.3(3)	15.7(3)	3.4(2)	-1.2(3)	1.8(2)
O7	26.9(3)	16.7(3)	20.8(3)	0.5(2)	-1.7(3)	-5.7(2)
O5	26.8(3)	27.2(3)	19.1(3)	1.8(3)	4.7(2)	2.3(2)
C12	19.7(4)	13.5(3)	14.8(3)	-0.3(3)	-0.6(3)	-1.2(3)
C13	17.4(4)	13.0(3)	14.9(4)	0.5(3)	-0.3(3)	-1.4(3)
C1	18.3(4)	12.4(3)	16.5(3)	1.8(3)	-0.1(3)	0.5(3)
C3	16.1(4)	12.3(3)	13.5(3)	0.3(3)	0.4(3)	-0.0(3)
C2	16.3(4)	12.5(3)	16.4(4)	0.1(3)	0.3(3)	-0.3(3)
C10	16.5(4)	14.6(3)	16.2(3)	-0.7(3)	-0.3(3)	0.7(3)
C7	21.0(4)	19.3(3)	14.2(3)	0.5(3)	0.2(3)	-0.5(3)
C11	19.1(4)	15.5(4)	18.3(4)	-1.5(3)	-2.4(3)	-0.5(3)
C9	19.2(4)	14.8(3)	14.8(4)	0.7(3)	-0.6(3)	-0.0(3)
C17	19.7(4)	16.4(4)	19.5(4)	3.2(3)	1.0(3)	-0.6(3)
C6	20.5(4)	16.9(3)	15.0(4)	0.5(3)	-1.5(3)	-0.8(3)
C4	17.0(4)	13.6(3)	16.2(3)	-1.1(3)	-0.9(3)	-1.2(3)
C8	19.4(4)	17.9(4)	16.2(4)	1.6(3)	0.3(3)	-1.2(3)
C5	18.8(4)	14.1(3)	14.9(3)	-1.1(3)	-0.9(3)	-0.8(3)
C15	25.9(5)	15.8(4)	16.9(4)	-4.9(3)	-1.1(3)	-0.2(3)
C18	16.2(4)	20.7(4)	25.5(4)	0.6(3)	-0.9(3)	2.1(3)
C16	22.6(5)	25.8(4)	24.1(4)	5.7(4)	3.3(3)	0.4(3)
C14	19.6(4)	25.8(4)	34.6(5)	0.4(4)	1.6(4)	5.8(4)
H11	27(5)	32(6)	43(7)	-4(3)	-13(3)	-12(4)
H17a	39(7)	48(7)	46(6)	3(4)	17(3)	12(4)
H4	24(5)	26(5)	20(5)	-9(2)	-3(3)	-2(3)
H14a	31(5)	53(7)	70(9)	15(3)	4(4)	-1(5)
H14b	40(8)	46(5)	89(9)	-5(4)	0(6)	31(3)
H15a	40(7)	34(6)	26(4)	-9(4)	3(3)	-9(2)
H6a	31(5)	23(5)	35(6)	-4(2)	-10(3)	-3(3)
H15b	50(8)	29(5)	32(5)	-5(4)	0(4)	9(3)
H15c	26(4)	14(5)	59(8)	-6(2)	-3(2)	-1(4)
H16a	53(6)	52(7)	104(11)	-9(3)	40(4)	-10(5)
H14c	35(8)	69(8)	47(4)	-6(5)	2(3)	-7(2)
H17b	39(7)	60(8)	31(4)	6(5)	-10(3)	-9(3)
H16b	69(10)	66(9)	52(6)	8(5)	-4(4)	-26(3)
H18a	61(10)	70(7)	100(9)	8(5)	5(6)	53(4)
H17c	35(8)	23(4)	87(10)	4(2)	4(5)	-17(2)
H7	33(5)	40(7)	42(9)	-11(3)	-4(3)	-10(6)
H16c	81(10)	85(9)	50(6)	41(5)	9(4)	27(4)
H18b	68(10)	98(11)	47(5)	30(6)	5(3)	-18(3)
H18c	36(6)	53(7)	151(14)	-6(3)	-25(4)	-21(5)
H6	55(7)	48(9)	46(8)	24(5)	27(4)	25(6)

**Table S10.** Anisotropic Displacement Parameters ( $\times 10^4$ ) for compound **2**. The anisotropic displacement factor exponent takes the form:  $-2\pi^2[h^2a^{*2} \times U_{11} + \dots + 2hka^* \times b^* \times U_{12}]$

Atom	Atom	Length/Å
O2	C3	1.3768(10)
O2	C4	1.4321(9)
O1	C1	1.3738(10)
O1	C9	1.3843(9)
O4	C5	1.4170(10)
O4	C14	1.4150(12)
O3	C1	1.2035(10)
O6	C6	1.4097(9)
O7	C12	1.3475(10)
O5	C7	1.2184(11)
C12	C13	1.4006(12)
C12	C11	1.4029(12)
C13	C3	1.3954(10)
C13	C17	1.5001(11)
C1	C2	1.4800(10)
C3	C2	1.4043(11)
C2	C10	1.4078(12)
C10	C11	1.3920(11)
C10	C18	1.5045(11)
C7	C6	1.5165(12)
C7	C8	1.4736(11)
C9	C4	1.5049(11)
C9	C8	1.3454(12)
C6	C5	1.5518(10)
C4	C5	1.5513(10)
C8	C16	1.4968(12)
C5	C15	1.5241(11)

**Table S11.** Bond Lengths in Å for compound **2**.

Atom	Atom	Atom	Angle/°
C4	O2	C3	117.55(6)
C9	O1	C1	119.43(6)
C14	O4	C5	118.30(6)
C13	C12	O7	116.68(7)
C11	C12	O7	121.77(7)
C11	C12	C13	121.54(7)
C3	C13	C12	116.69(7)
C17	C13	C12	121.61(7)
C17	C13	C3	121.69(7)
O3	C1	O1	116.64(7)
C2	C1	O1	118.78(6)
C2	C1	O3	124.54(7)
C13	C3	O2	117.43(7)
C2	C3	O2	120.09(6)
C2	C3	C13	122.28(7)
C3	C2	C1	118.21(7)
C10	C2	C1	120.56(7)
C10	C2	C3	120.00(7)
C11	C10	C2	117.90(7)
C18	C10	C2	122.29(7)
C18	C10	C11	119.80(7)
C6	C7	O5	119.64(7)
C8	C7	O5	122.82(8)
C8	C7	C6	117.54(7)
C10	C11	C12	121.15(8)
C4	C9	O1	114.82(7)
C8	C9	O1	118.71(7)
C8	C9	C4	125.72(7)
C7	C6	O6	109.28(7)
C5	C6	O6	111.69(6)
C5	C6	C7	109.29(6)
C9	C4	O2	112.64(7)
C5	C4	O2	113.85(6)
C5	C4	C9	112.72(6)
C9	C8	C7	116.64(7)
C16	C8	C7	118.59(8)
C16	C8	C9	124.77(8)
C6	C5	O4	111.82(6)
C4	C5	O4	111.04(6)
C4	C5	C6	105.11(6)
C15	C5	O4	104.71(6)
C15	C5	C6	109.32(7)
C15	C5	C4	114.99(7)

**Table S12.** Bond Angles in ° for compound **2**.

<b>Atom-Atom</b>	<b>Length [Å]</b>
O2-C3	1.3768(10)
O2-C4	1.4321(9)
O1-C1	1.3738(10)
O1-C9	1.3843(9)
O4-C5	1.4170(10)
O4-C14	1.4150(12)
O3-C1	1.2035(10)
O6-C6	1.4097(9)
O6-H6	0.973(17)
O7-C12	1.3475(10)
O7-H7	0.969(15)
O5-C7	1.2184(11)
C12-C13	1.4006(12)
C12-C11	1.4029(12)
C13-C3	1.3954(10)
C13-C17	1.5001(11)
C1-C2	1.4800(10)
C3-C2	1.4043(11)
C2-C10	1.4078(12)
C10-C11	1.3920(11)
C10-C18	1.5045(11)
C7-C6	1.5165(12)
C7-C8	1.4736(11)
C11-H11	1.085(12)
C9-C4	1.5049(11)
C9-C8	1.3454(12)
C17-H17a	1.082(14)
C17-H17b	1.101(13)
C17-H17c	1.101(12)
C6-C5	1.5518(10)
C6-H6a	1.102(11)
C4-C5	1.5513(10)
C4-H4	1.092(11)
C8-C16	1.4968(12)
C5-C15	1.5241(11)
C15-H15a	1.083(11)
C15-H15b	1.115(12)
C15-H15c	1.084(13)
C18-H18a	1.075(15)
C18-H18b	1.060(15)
C18-H18c	1.059(15)
C16-H16a	1.025(16)
C16-H16b	1.114(15)
C16-H16c	1.060(15)
C14-H14a	1.080(14)

**Table S13.** Bond Lengths including those for hydrogen atoms in Å for compound **2**.

Atom	Atom	Atom	Atom	Angle/ °
O2	C3	C13	C12	179.34(7)
O2	C3	C13	C17	0.60(8)
O2	C3	C2	C1	14.41(8)
O2	C3	C2	C10	-178.21(7)
O2	C4	C9	O1	-35.47(7)
O2	C4	C9	C8	154.69(6)
O2	C4	C5	O4	58.99(7)
O2	C4	C5	C6	-179.92(7)
O2	C4	C5	C15	-59.65(8)
O1	C1	C2	C3	-51.37(8)
O1	C1	C2	C10	141.32(7)
O1	C9	C4	C5	-165.99(7)
O1	C9	C8	C7	-176.43(7)
O1	C9	C8	C16	4.32(10)
O4	C5	C6	O6	-55.81(7)
O4	C5	C6	C7	-176.86(6)
O4	C5	C4	C9	-171.12(6)
O3	C1	C2	C3	126.26(9)
O3	C1	C2	C10	-41.04(10)
O6	C6	C7	O5	8.07(8)
O6	C6	C7	C8	-172.48(6)
O6	C6	C5	C4	-176.39(7)
O6	C6	C5	C15	59.67(8)
O7	C12	C13	C3	179.39(7)
O7	C12	C13	C17	-1.87(9)
O7	C12	C11	C10	-174.70(7)
O5	C7	C6	C5	130.57(8)
O5	C7	C8	C9	-160.40(8)
O5	C7	C8	C16	18.90(10)
C12	C13	C3	C2	-5.80(9)
C12	C11	C10	C2	-3.46(9)
C12	C11	C10	C18	177.23(7)
C13	C3	C2	C1	-160.31(7)
C13	C3	C2	C10	7.07(9)
C1	C2	C10	C11	164.85(7)
C1	C2	C10	C18	-15.86(9)
C3	C2	C10	C11	-2.23(8)
C3	C2	C10	C18	177.07(7)
C7	C6	C5	C4	62.57(7)
C7	C6	C5	C15	-61.38(7)
C7	C8	C9	C4	-6.96(9)
C9	C4	C5	C6	-50.02(7)
C9	C4	C5	C15	70.24(8)

**Table S14.** Torsion Angles in ° for compound **2**.

Atom	x	y	z	$U_{eq}$
H11	3489(16)	3631(10)	2938(7)	34(3)
H17a	9960(20)	3615(12)	3216(9)	45(3)
H4	9419(14)	6039(9)	5437(6)	23(2)
H14a	12624(18)	3953(12)	5687(9)	51(3)
H14b	11654(19)	5321(13)	5495(10)	59(4)
H15a	7515(17)	3064(10)	4996(7)	33(3)
H6a	9178(15)	5435(10)	6722(7)	30(3)
H15b	7320(19)	2789(10)	6011(7)	37(3)
H15c	5969(16)	3774(9)	5529(8)	33(3)
H16a	3520(20)	6462(14)	6557(11)	70(4)
H14c	11626(18)	4740(14)	6410(8)	50(3)
H17b	9802(18)	3369(12)	4200(7)	43(3)
H16b	4860(20)	7396(15)	6875(9)	62(4)
H18a	2730(20)	6371(15)	3185(12)	77(5)
H17c	8978(18)	2351(10)	3520(10)	48(4)
H7	5250(20)	2116(13)	2752(8)	38(3)
H16c	4150(20)	7443(15)	5900(9)	72(5)
H18b	2290(20)	6010(17)	4118(9)	71(5)
H18c	1660(20)	5141(14)	3404(12)	80(5)
H6	7530(20)	3853(14)	7486(9)	50(5)

**Table S15.** Hydrogen Fractional Atomic Coordinates ( $\times 10^4$ ) and Equivalent Isotropic Displacement Parameters ( $\text{\AA}^2 \times 10^3$ ) for compound **2**.  $U_{eq}$  is defined as 1/3 of the trace of the orthogonalised  $U_{ij}$ .

D	H	A	d(D-H)/\AA	d(H-A)/\AA	d(D-A)/\AA	D-H-A/deg
O7	H7	O6 <sup>1</sup>	0.969(15)	1.819(15)	2.7477(10)	159.7(14)
-----						
<sup>1</sup> -1/2+x,1/2-y,1-z						

**Table S16.** Hydrogen Bond information for compound **2**.

## X-ray crystallography report for compound 43 (XC219)

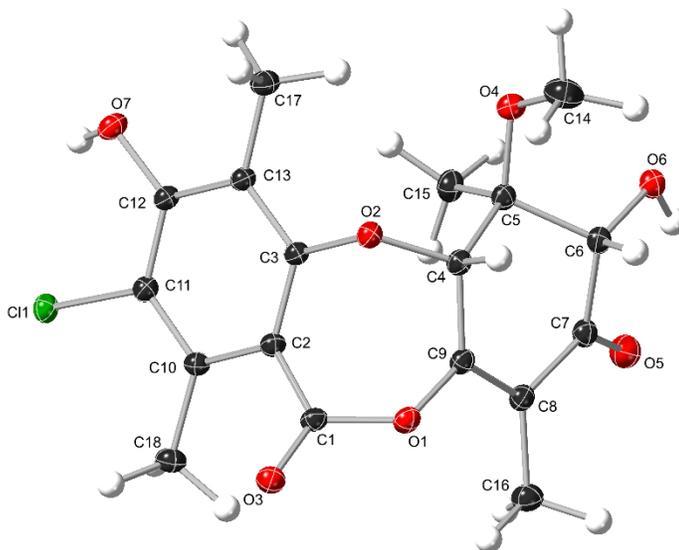
# Compound 43 (XC219)

Submitted by: **Stuart Romeril**

Solved by: **John Bacsá**

**$R_1=1.42\%$**

## Crystal Data and Experimental



**Figure S14.** X-ray crystal structure of **43**

**Experimental.** Single colorless plate-shaped crystals of **XC219 (43)** were used as supplied. A suitable crystal with dimensions  $0.28 \times 0.20 \times 0.11 \text{ mm}^3$  was selected and mounted on a Rigaku XtaLAB Synergy S diffractometer. The crystal was kept at a steady  $T = 100.00(10) \text{ K}$  during data collection. The structure was solved with the ShelXT 2018/2 solution program<sup>3</sup> using dual methods and by using Olex2 1.5-alpha as the graphical interface.<sup>6</sup> The model was refined with olex2.refine 1.5-alpha<sup>4</sup> using full matrix least squares minimisation on  $F^2$ .

**Crystal Data.**  $\text{C}_{18}\text{H}_{19}\text{ClO}_7$ ,  $M_r = 382.800$ , orthorhombic,  $P2_12_12_1$  (No. 19),  $a = 8.0816(1) \text{ \AA}$ ,  $b = 11.9289(1) \text{ \AA}$ ,  $c = 17.8225(2) \text{ \AA}$ ,  $\alpha = \beta = \gamma = 90^\circ$ ,  $V = 1718.17(3) \text{ \AA}^3$ ,  $T = 100.00(10) \text{ K}$ ,  $Z = 4$ ,  $Z' = 1$ ,  $\mu(\text{Cu K}\alpha) = 2.328$ , 27395 reflections measured, 3588 unique ( $R_{\text{int}} = 0.0433$ ) which were used in all calculations. The final  $wR_2$  was 0.0328 (all data) and  $R_1$  was 0.0142 ( $I \geq 2 \sigma(I)$ ).

<b>Compound</b>	<b>XC219 (43)</b>
Formula	C <sub>18</sub> H <sub>19</sub> ClO <sub>7</sub>
<i>D</i> <sub>calc.</sub> / g cm <sup>-3</sup>	1.480
$\mu$ /mm <sup>-1</sup>	2.328
Formula Weight	382.800
Color	colorless
Shape	plate-shaped
Size/mm <sup>3</sup>	0.28×0.20×0.11
<i>T</i> /K	100.00(10)
Crystal System	orthorhombic
Flack Parameter	0.002(4)
Hooft Parameter	0.002(4)
Space Group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
<i>a</i> /Å	8.0816(1)
<i>b</i> /Å	11.9289(1)
<i>c</i> /Å	17.8225(2)
$\alpha$ /°	90
$\beta$ /°	90
$\gamma$ /°	90
<i>V</i> /Å <sup>3</sup>	1718.17(3)
<i>Z</i>	4
<i>Z'</i>	1
Wavelength/Å	1.54184
Radiation type	Cu K $\alpha$
$\theta$ <sub>min</sub> /°	4.96
$\theta$ <sub>max</sub> /°	79.64
Measured Refl's.	27395
Indep't Refl's	3588
Refl's I $\geq$ 2 $\sigma$ (I)	3553
<i>R</i> <sub>int</sub>	0.0433
Parameters	431
Restraints	390
Largest Peak	0.0976
Deepest Hole	-0.0947
GooF	1.0869
<i>wR</i> <sub>2</sub> (all data)	0.0328
<i>wR</i> <sub>2</sub>	0.0327
<i>R</i> <sub>1</sub> (all data)	0.0143
<i>R</i> <sub>1</sub>	0.0142

**Table S17.** Crystallography statistics for compound **43**

## Structure Quality Indicators

**Reflections:**  $d \min (\text{Cu}\alpha) [2\theta=159.9^\circ] = 0.78$ ;  $I/\sigma(I) = 44.5$ ;  $R_{\text{int}} = 4.33\%$ ; Full  $135.4^\circ$  (96% to  $159.9^\circ$ ) = 99.9

**Refinement:** Shift = -0.000; Max Peak = 0.1; Min Peak = -0.1; GooF = 1.087; Hooft = 0.02(4)

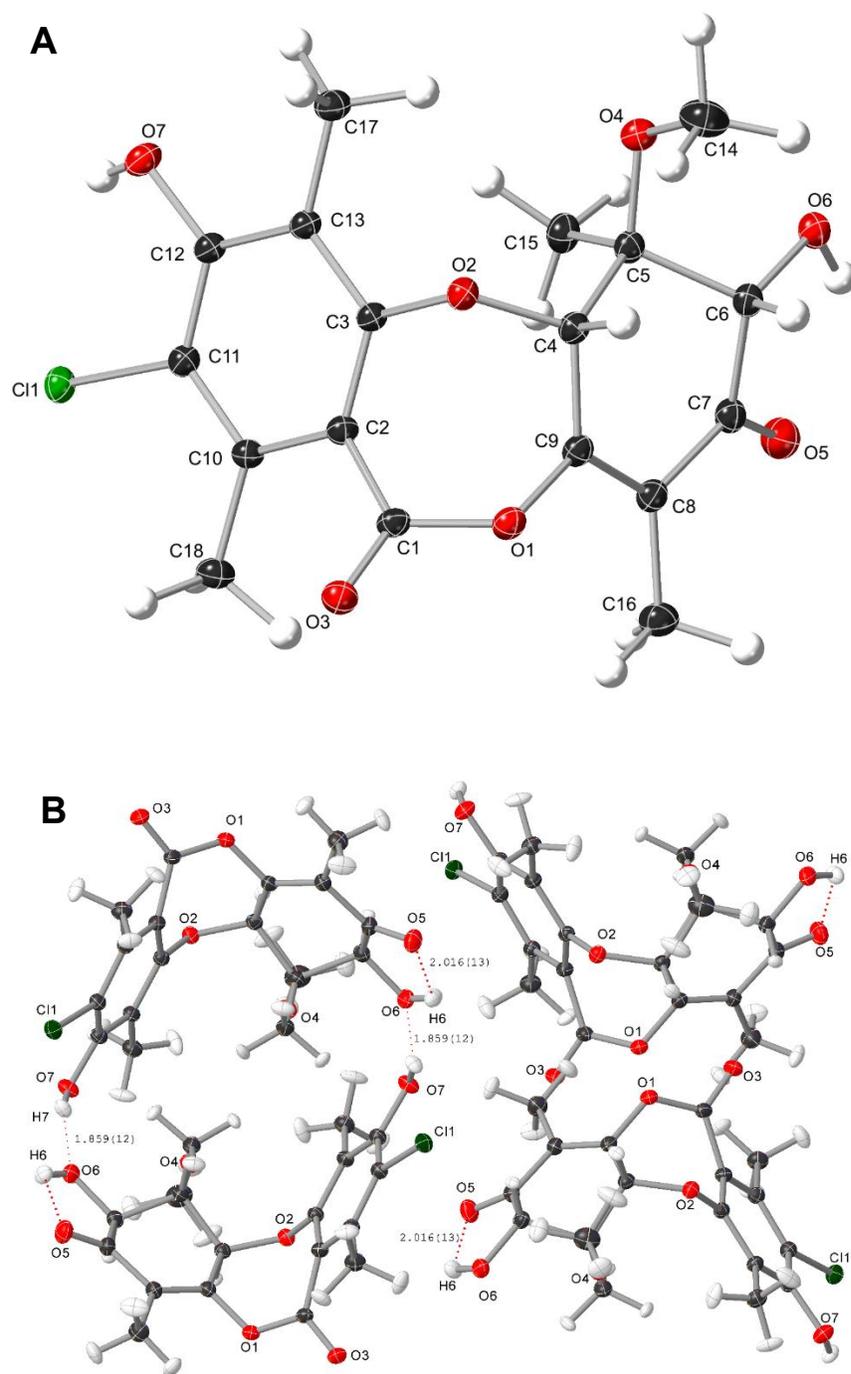
A colorless plate-shaped crystal with dimensions  $0.28 \times 0.20 \times 0.11 \text{ mm}^3$  was mounted. Data were collected using a XtaLAB Synergy, Dualflex, HyPix diffractometer equipped with an Oxford Cryosystems low-temperature device operating at  $T = 100.00(10) \text{ K}$ .

Data were measured using  $\omega$  scans with a narrow frame width with Cu  $K_\alpha$  radiation. The diffraction pattern was indexed and the total number of runs and images was based on the strategy calculation from the program CrysAlisPro 1.171.41.98a (Rigaku OD, 2021). The maximum resolution that was achieved was  $\theta = 79.64^\circ$  (0.78 Å).

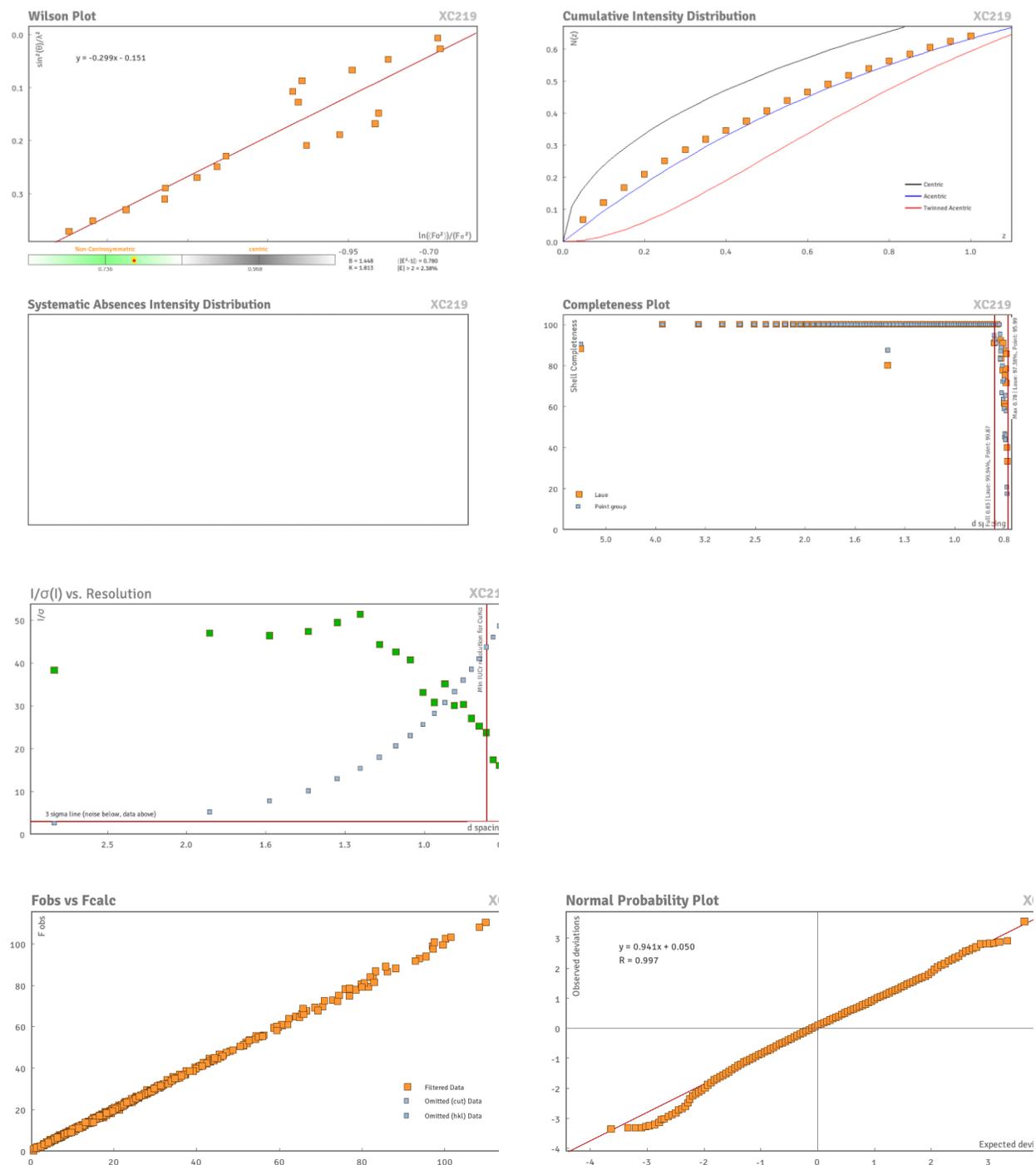
The unit cell was refined using CrysAlisPro 1.171.41.98a (Rigaku OD, 2021) on 6861 reflections, 25% of the observed reflections.

Data reduction, scaling and absorption corrections were performed using CrysAlisPro 1.171.41.98a (Rigaku OD, 2021). The final completeness is 99.78 % out to  $79.64^\circ$  in  $\theta$ . A numerical absorption correction based on gaussian integration over a multifaceted crystal model was performed using CrysAlisPro 1.171.41.98a (Rigaku OD, 2021). An empirical absorption correction using spherical harmonics, implemented in SCALE3 ABSPACK scaling algorithm was also performed. The absorption coefficient  $\mu$  of this material is  $2.328 \text{ mm}^{-1}$  at this wavelength ( $\lambda = 1.54184\text{Å}$ ) and the minimum and maximum transmissions are 0.451 and 1.000.

The structure was solved and the space group  $P2_12_12_1$  (# 19) determined by the ShelXT 2018/2 structure solution program<sup>3</sup> using dual methods and refined by full matrix least squares minimisation on  $F^2$  using version of olex2.refine 1.5-alpha.<sup>4</sup> All non-hydrogen atoms were refined anisotropically. Hydrogen atom positions were calculated geometrically and refined using the riding model. Hydrogen atom positions were located from the electron densities and freely refined using Hirshfeld scattering factors. Refinement was by using NoSpherA2, an implementation of non-spherical atom-form-factors.<sup>5</sup> NoSpherA2 implementation of HAR makes use of tailor-made aspherical atomic form factors calculated from a Hirshfeld-partitioned electron density (ED) not from spherical-atom form factors. The ED was calculated from a Gaussian basis set single determinant SCF wavefunction from DFT using selected functionals for a fragment of this crystal. SOFTWARE: ORCA 5.0 PARTITIONING: NoSpherA2 INT ACCURACY: Normal METHOD: B3LYP BASIS SET: def2-TZVP CHARGE: 0 MULTIPLICITY: 1 SOLVATION: Ethanol DATE: 2022-08-24\_15-24-24



**Figure S15.** A) The crystal structure of sample **XC219 (43)**. Structurally, the crystal structure is remarkably similar (almost identical) to compound **2**. Replacing the H atom on C11 with a Cl atom has a minimal effect on the symmetry and geometry of the molecule. There are three chiral centers, labelled C4, C5 and C6 and they have S, S and S configuration respectively. There is a single molecule in the asymmetric unit, which is represented by the reported sum formula. In other words: Z is 4 and Z' is 1. B) The intra and intermolecular hydrogen bonds. The Cl does not participate in strong hydrogen bonding.



**Figure S16.** Crystallography diffraction and refinement data for **43**.

Total reflections (after filtering)	27395	Unique reflections	3593
Completeness	0.96	Mean I/ $\sigma$	35.36
hkl <sub>max</sub> collected	(9, 14, 21)	hkl <sub>min</sub> collected	(-10, -13, -21)
hkl <sub>max</sub> used	(10, 14, 21)	hkl <sub>min</sub> used	(-10, 0, 0)
Lim d <sub>max</sub> collected	100.0	Lim d <sub>min</sub> collected	0.77
d <sub>max</sub> used	9.91	d <sub>min</sub> used	0.78
Friedel pairs	2994	Friedel pairs merged	0
Inconsistent equivalents	0	R <sub>int</sub>	0.0433
R <sub>sigma</sub>	0.0225	Intensity transformed	0
Omitted reflections	0	Omitted by user (OMIT)	45
Multiplicity	(2838, 2217, 1559, 1145, 773, 494, 313, 127, 46, 24, 16)	Maximum multiplicity	18
Removed systematic absences	0	Filtered off (Shel/OMIT)	0

**Table S18.** Reflection statistics for crystal structure of **43**

Atom	x	y	z	$U_{eq}$
C11	6912.6(6)	6339.3(4)	2743.6(3)	16.95(15)
O1	3254.4(7)	2965.7(4)	4950.4(3)	16.96(11)
O2	1259.2(6)	4692.3(4)	4476.5(3)	15.16(11)
O4	-280.8(7)	6116.6(4)	5627.0(3)	20.63(12)
O7	3635.4(7)	7426.3(4)	2905.6(3)	19.31(12)
O3	4490.0(7)	2484.6(4)	3901.4(3)	20.04(12)
O6	1139.0(7)	6000.2(4)	7096.7(3)	20.63(12)
O5	3962.8(7)	4939.3(5)	7211.5(3)	24.03(12)
C12	3825.9(9)	6481.7(6)	3306.8(4)	14.64(14)
C3	2631.1(8)	5093.2(6)	4105.3(4)	13.37(14)
C2	4078.6(9)	4441.2(6)	4067.0(4)	14.16(14)
C7	3114.4(10)	4641.0(6)	6681.7(4)	16.88(14)
C1	3991.8(9)	3244.6(6)	4284.1(4)	15.26(14)
C9	2869.7(9)	3793.4(6)	5465.3(4)	14.87(14)
C13	2460.7(8)	6097.0(6)	3714.9(4)	13.94(14)
C10	5471.3(9)	4839.0(6)	3672.5(4)	14.37(14)
C8	3671.8(9)	3807.1(6)	6127.1(4)	15.88(14)
C17	858.9(9)	6732.7(7)	3710.4(5)	17.93(15)
C6	1400.5(9)	5136.6(6)	6574.4(4)	16.86(15)
C4	1381.7(9)	4491.4(6)	5265.5(4)	14.39(14)
C18	7075.5(9)	4197.9(6)	3639.6(5)	18.74(15)
C15	2561.6(10)	6449.4(7)	5600.6(4)	19.36(16)
C5	1241.7(9)	5559.7(6)	5752.9(4)	15.68(14)
C11	5305.0(9)	5860.1(6)	3302.3(4)	14.49(14)
C16	5088.3(10)	3055.7(7)	6329.8(5)	20.92(16)
C14	-1765.7(10)	5524.2(7)	5777.1(6)	27.81(19)

**Table S19.** Fractional Atomic Coordinates ( $\times 10^4$ ) and Equivalent Isotropic Displacement Parameters ( $\text{\AA}^2 \times 10^3$ ) for **XC219 (43)**.  $U_{eq}$  is defined as 1/3 of the trace of the orthogonalised  $U_{ij}$ .

Atom	$U_{11}$	$U_{22}$	$U_{33}$	$U_{23}$	$U_{13}$	$U_{12}$
C11	14.3(3)	20.0(3)	16.6(3)	-2.7(2)	1.6(2)	2.4(2)
O1	22.6(3)	11.7(2)	16.6(3)	-0.12(19)	-0.0(2)	1.03(19)
O2	13.9(2)	16.2(2)	15.4(2)	-1.81(19)	-0.78(19)	1.40(19)
O4	22.1(3)	18.9(3)	20.9(3)	4.5(2)	1.4(2)	2.9(2)
O7	18.4(3)	17.4(2)	22.0(3)	0.5(2)	-1.2(2)	6.7(2)
O3	27.5(3)	13.5(2)	19.2(3)	3.1(2)	1.9(2)	-2.0(2)
O6	26.9(3)	19.3(3)	15.7(3)	3.3(2)	3.0(2)	-0.4(2)
O5	26.4(3)	28.1(3)	17.6(3)	2.1(2)	-4.3(2)	-3.6(2)
C12	14.5(3)	14.1(3)	15.4(3)	0.2(3)	-1.5(3)	1.7(3)
C3	13.2(3)	12.9(3)	14.0(3)	0.0(2)	-1.1(2)	0.8(3)
C2	14.8(3)	12.2(3)	15.4(3)	0.5(2)	-0.2(3)	0.3(3)
C7	19.3(3)	17.7(3)	13.6(3)	0.1(3)	0.1(3)	-0.3(3)
C1	18.1(3)	11.9(3)	15.8(3)	1.6(3)	-1.6(3)	-0.1(3)
C9	17.2(3)	13.1(3)	14.4(3)	-0.8(3)	-0.4(3)	1.5(2)
C13	13.2(3)	13.3(3)	15.3(3)	0.2(2)	-0.7(2)	1.6(3)
C10	13.5(3)	14.1(3)	15.5(3)	1.1(2)	-0.7(3)	-0.1(3)
C8	17.2(3)	16.1(3)	14.3(3)	0.3(3)	-0.5(3)	1.2(3)
C17	15.2(3)	17.2(4)	21.4(4)	1.6(3)	-0.9(3)	2.8(3)
C6	19.2(4)	16.4(3)	14.9(4)	0.7(3)	2.2(3)	1.6(3)
C4	14.7(3)	13.4(3)	15.1(3)	-1.8(2)	-0.1(3)	1.2(3)
C18	15.1(4)	17.7(4)	23.4(4)	3.0(3)	-0.2(3)	0.1(3)
C15	25.0(4)	15.8(4)	17.3(4)	-4.4(3)	1.9(3)	-0.0(3)
C5	18.4(3)	13.8(3)	14.8(3)	-0.3(3)	1.5(3)	1.0(3)
C11	13.6(3)	14.8(3)	15.1(3)	-1.0(3)	-1.0(3)	1.0(3)
C16	18.7(4)	22.0(4)	22.0(4)	3.1(3)	-3.1(3)	0.4(3)
C14	19.1(4)	27.0(4)	37.3(5)	3.2(3)	-1.4(4)	-2.3(4)
H4	19(4)	27(5)	32(5)	-7(2)	7(3)	-4(3)
H16a	23(4)	46(6)	98(10)	-7(3)	-15(4)	9(5)
H17a	38(6)	24(4)	88(9)	6(2)	9(5)	17(2)
H15a	55(7)	30(5)	32(5)	-14(4)	10(4)	-13(3)
H18a	47(7)	66(8)	38(4)	20(5)	-2(3)	-20(3)
H18b	25(4)	30(5)	74(8)	-4(2)	-4(4)	-3(4)
H14a	52(8)	56(5)	106(9)	-2(4)	-5(6)	-40(4)
H6a	23(4)	21(4)	28(5)	-4(2)	7(3)	-1(3)
H17b	34(5)	56(7)	70(7)	15(4)	-26(3)	-27(4)
H18c	45(6)	34(5)	49(6)	12(4)	1(4)	17(3)
H15b	26(4)	29(6)	68(8)	-1(2)	5(3)	8(5)
H16b	58(7)	44(5)	45(6)	14(4)	-8(4)	-18(3)
H14b	29(5)	53(6)	71(8)	17(3)	0(4)	9(4)
H17c	40(6)	82(9)	36(4)	20(5)	15(3)	18(3)
H14c	45(7)	76(8)	50(4)	-13(6)	-2(2)	12(2)
H7	25(4)	41(6)	37(7)	-10(2)	-1(3)	11(5)
H16c	53(7)	63(8)	32(4)	11(5)	1(3)	20(3)
H6	40(5)	41(7)	41(6)	6(4)	-14(3)	-6(5)
H15c	59(7)	29(5)	24(4)	-13(4)	-8(3)	9(2)

**Table S20.** Anisotropic Displacement Parameters ( $\times 10^4$ ) for **XC219 (43)**. The anisotropic displacement factor exponent takes the form:  $-2\pi^2[h^2a^{*2} \times U_{11} + \dots + 2hka^* \times b^* \times U_{12}]$

Atom	Atom	Length/Å
C11	C11	1.7338(8)
O1	C1	1.3697(9)
O1	C9	1.3834(9)
O2	C3	1.3768(8)
O2	C4	1.4298(9)
O4	C5	1.4162(9)
O4	C14	1.4181(10)
O7	C12	1.3435(9)
O3	C1	1.2039(9)
O6	C6	1.4045(9)
O5	C7	1.2200(10)
C12	C13	1.3988(10)
C12	C11	1.4067(10)
C3	C2	1.4064(9)
C3	C13	1.3918(10)
C2	C1	1.4805(10)
C2	C10	1.4094(10)
C7	C8	1.4729(10)
C7	C6	1.5181(10)
C9	C8	1.3460(10)
C9	C4	1.5054(10)
C13	C17	1.5003(9)
C10	C18	1.5064(10)
C10	C11	1.3918(10)
C8	C16	1.4981(10)
C6	C5	1.5540(10)
C4	C5	1.5463(10)
C15	C5	1.5291(10)

**Table S21.** Bond Lengths in Å for **XC219 (43)**.

Atom	Atom	Atom	Angle/°
C9	O1	C1	119.97(5)
C4	O2	C3	118.35(5)
C14	O4	C5	118.14(6)
C13	C12	O7	117.49(6)
C11	C12	O7	122.44(6)
C11	C12	C13	120.02(6)
C2	C3	O2	120.07(6)
C13	C3	O2	117.35(6)
C13	C3	C2	122.26(6)
C1	C2	C3	118.75(6)
C10	C2	C3	120.15(6)
C10	C2	C1	119.53(6)
C8	C7	O5	122.99(7)
C6	C7	O5	119.78(7)
C6	C7	C8	117.23(6)
O3	C1	O1	116.98(6)
C2	C1	O1	118.78(6)
C2	C1	O3	124.20(7)
C8	C9	O1	118.80(6)
C4	C9	O1	114.67(6)
C4	C9	C8	125.83(6)
C3	C13	C12	117.66(6)
C17	C13	C12	120.80(6)
C17	C13	C3	121.53(6)
C18	C10	C2	122.39(6)
C11	C10	C2	117.01(6)
C11	C10	C18	120.60(7)
C9	C8	C7	116.67(6)
C16	C8	C7	118.43(6)
C16	C8	C9	124.89(7)
C7	C6	O6	109.84(6)
C5	C6	O6	111.95(6)
C5	C6	C7	108.69(6)
C9	C4	O2	112.38(6)
C5	C4	O2	114.16(6)
C5	C4	C9	112.42(6)
C6	C5	O4	111.92(6)
C4	C5	O4	111.17(6)
C4	C5	C6	104.81(6)
C15	C5	O4	104.61(6)
C15	C5	C6	109.57(6)
C15	C5	C4	114.92(6)
C12	C11	Cl1	117.79(5)
C10	C11	Cl1	119.24(6)
C10	C11	C12	122.73(7)

**Table S22.** Bond Angles in ° for **XC219 (43)**.

Atom	Atom	Atom	Atom	Angle/ °
C11	C11	C12	O7	-0.89(7)
C11	C11	C12	C13	176.41(6)
C11	C11	C10	C2	-174.69(6)
C11	C11	C10	C18	4.80(7)
O1	C1	C2	C3	-49.28(7)
O1	C1	C2	C10	144.97(6)
O1	C9	C8	C7	-174.69(6)
O1	C9	C8	C16	6.33(8)
O1	C9	C4	O2	-36.43(6)
O1	C9	C4	C5	-166.87(6)
O2	C3	C2	C1	12.87(8)
O2	C3	C2	C10	178.53(6)
O2	C3	C13	C12	-176.99(6)
O2	C3	C13	C17	1.28(8)
O2	C4	C9	C8	153.38(6)
O2	C4	C5	O4	58.67(6)
O2	C4	C5	C6	179.77(6)
O2	C4	C5	C15	-59.90(7)
O4	C5	C6	O6	-53.70(6)
O4	C5	C6	C7	-175.21(6)
O4	C5	C4	C9	-171.81(6)
O7	C12	C13	C3	177.30(6)
O7	C12	C13	C17	-0.98(8)
O7	C12	C11	C10	-175.18(7)
O3	C1	C2	C3	128.54(8)
O3	C1	C2	C10	-37.21(8)
O6	C6	C7	O5	6.60(7)
O6	C6	C7	C8	-173.43(5)
O6	C6	C5	C4	-174.31(6)
O6	C6	C5	C15	61.88(7)
O5	C7	C8	C9	-160.85(7)
O5	C7	C8	C16	18.20(9)
O5	C7	C6	C5	129.39(7)
C12	C13	C3	C2	-3.48(8)
C12	C11	C10	C2	-0.48(8)
C12	C11	C10	C18	179.01(7)
C3	C2	C10	C18	177.45(7)
C3	C2	C10	C11	-3.07(8)
C7	C8	C9	C4	-4.86(8)
C7	C6	C5	C4	64.18(6)
C7	C6	C5	C15	-59.63(7)
C9	C4	C5	C6	-50.71(6)
C9	C4	C5	C15	69.62(7)

**Table S23.** Torsion Angles in ° for **XC219 (43)**.

Atom	x	y	z	$U_{eq}$
H4	292(11)	3973(8)	5376(5)	26(2)
H16a	6226(14)	3505(10)	6370(8)	56(3)
H17a	1039(14)	7590(9)	3570(7)	50(3)
H15a	2511(14)	7068(9)	6047(6)	39(3)
H18a	7230(14)	3790(11)	3091(7)	50(3)
H18b	8105(14)	4753(8)	3727(7)	43(3)
H14a	-1899(17)	4774(12)	5444(9)	71(4)
H6a	481(11)	4449(8)	6663(5)	24(2)
H17b	31(14)	6370(11)	3308(7)	54(3)
H18c	7139(14)	3562(9)	4073(7)	43(3)
H15b	3789(13)	6102(9)	5583(7)	41(3)
H16b	5221(16)	2385(10)	5936(7)	49(3)
H14b	-2782(13)	6068(10)	5626(8)	51(3)
H17c	215(14)	6655(11)	4232(6)	53(3)
H14c	-1911(16)	5333(11)	6363(7)	57(3)
H7	4690(15)	7795(10)	2830(7)	34(3)
H16c	4898(15)	2696(11)	6886(7)	50(3)
H6	2086(18)	6001(11)	7436(7)	40(3)
H15c	2325(14)	6869(8)	5071(6)	37(3)

**Table S24.** Hydrogen Fractional Atomic Coordinates ( $\times 10^4$ ) and Equivalent Isotropic Displacement Parameters ( $\text{\AA}^2 \times 10^3$ ) for XC219 (**43**).  $U_{eq}$  is defined as 1/3 of the trace of the orthogonalised  $U_{ij}$ . ( $^{1/2}+x, 3/2-y, 1-z$ )

D	H	A	d(D-H)/\AA	d(H-A)/\AA	d(D-A)/\AA	D-H-A/deg
O7	H7	O6 <sup>1</sup>	0.968(12)	1.858(12)	2.7599(8)	153.6(11)

**Table S25.** Hydrogen Bond information for XC219 (**43**).

Atom	Atom	Length/\AA
O7	H7	0.968(12)
O6	H6	0.975(14)
C17	H17a	1.063(11)
C17	H17b	1.072(11)
C17	H17c	1.070(11)
C6	H6a	1.118(9)
C4	H4	1.094(9)
C18	H18a	1.100(12)
C18	H18b	1.075(11)
C18	H18c	1.084(11)
C15	H15a	1.085(10)
C15	H15b	1.075(10)
C15	H15c	1.086(10)
C16	H16a	1.067(11)
C16	H16b	1.070(11)
C16	H16c	1.091(11)
C14	H14a	1.080(13)
C14	H14b	1.081(11)
C14	H14c	1.076(13)

**Table S26.** Selected Bond Lengths in \AA for XC219 (**43**).

Atom	Atom	Atom	Angle/°
H7	O7	C12	110.7(7)
H6	O6	C6	107.1(7)
H17a	C17	C13	111.7(6)
H17b	C17	C13	109.8(6)
H17b	C17	H17a	108.5(10)
H17c	C17	C13	111.8(6)
H17c	C17	H17a	110.8(10)
H17c	C17	H17b	104.0(10)
H6a	C6	O6	110.1(5)
H6a	C6	C7	107.6(5)
H6a	C6	C5	108.5(5)
H4	C4	O2	102.5(5)
H4	C4	C9	106.7(5)
H4	C4	C5	107.8(5)
H18a	C18	C10	110.9(6)
H18b	C18	C10	110.4(6)
H18b	C18	H18a	108.3(9)
H18c	C18	C10	111.6(6)
H18c	C18	H18a	108.6(10)
H18c	C18	H18b	106.9(9)
H15a	C15	C5	108.4(6)
H15b	C15	C5	112.4(6)
H15b	C15	H15a	108.5(9)
H15c	C15	C5	110.6(6)
H15c	C15	H15a	108.5(8)
H15c	C15	H15b	108.4(9)
H16a	C16	C8	112.0(6)
H16b	C16	C8	111.5(6)
H16b	C16	H16a	109.5(10)
H16c	C16	C8	110.3(6)
H16c	C16	H16a	105.0(10)
H16c	C16	H16b	108.4(10)
H14a	C14	O4	113.2(8)
H14b	C14	O4	107.3(6)
H14b	C14	H14a	106.5(10)
H14c	C14	O4	112.4(7)
H14c	C14	H14a	110.4(11)
H14c	C14	H14b	106.6(10)

**Table S27.** Selected Bond Angles in ° for **XC219 (43)**.

## Supplementary Methods

### Genomic search for CDK ETaGs

We searched our proprietary database of fully annotated genomes (LifeBase) for biosynthetic regions (annotated by antiSMASH v5<sup>7</sup>) that contained a homolog of human CDK2 (Uniprot ID P24941) with a minimum protein sequence identity of 40% using DIAMOND BlastP v.2.0.9.<sup>8</sup> To establish whether the identified homologs represented paralogs of housekeeping CDK proteins, we identified all CDK homologs in the candidate genome and in the reference genomes of *Saccharomyces cerevisiae* (UP000002311), *Aspergillus nidulans* (UP000000560), *Neurospora crassa* (UP000001805), and *Homo sapiens* (UP000005640) from Uniprot. We aligned sequences with MAFFT v7.508,<sup>9</sup> removed gaps with TrimAl v.1.4 (--automated1),<sup>10</sup> and built phylogenetic trees with IQ-TREE 2 v.2.2.0.3,<sup>11</sup> using default parameters unless otherwise noted.

To establish the boundaries of the Biosynthetic Gene Cluster (BGC), we identified bidirectional best hits to the core-synthase of the candidate region and perform gene cluster comparison using comparative genomic heatmaps<sup>12</sup> and with Clinker.<sup>13</sup> Once we established the boundaries of the BGC, we identified homologous BGCs in LifeBase by searching for loci that contained a sequence homolog to the core-synthase (*rosJ*) in addition to a majority of the genes predicted to encode biosynthetic enzymes (6 out of 7 in the case of *ros*).

### Identification of putative resistance mutations in the ETaG

We predicted putative resistance mutations in the ETaG by comparing the conservation of residues in the ETaG with that in housekeeping Pho85. We identified 38 diverse ETaGs by dereplicating the *ros* ETaG protein dataset at 90% sequence identity using all versus all sequence comparison (DIAMOND BlastP v.2.0.9) and mcl v22-282.<sup>14</sup> To construct a reference set with housekeeping Pho85 sequences, we identified 100 genomes in LifeBase that 1) had a single copy of Pho85, 2) did not contain the *ros* BGC and 3) maximized species diversity. We aligned the sequences in the ETaG and reference datasets using MAFFT v7.490 and then used the Two Sample Logo software<sup>15</sup> to identify

positions with statistically significant differences in their amino acid composition between the two datasets.

### Phylogenetic reconstruction

To create a CDK tree (Figure 2), CDK protein sequences from humans (CDK1 – CDK20) and *S. cerevisiae* were imported from UniProt (Table S6). All sequences were used to identify CDK homologs in *P. flanaganii* using DIAMOND BlastP v.2.0.9 on default settings with an E-value cutoff of  $1e^{-5}$ . Sequences were then aligned using the L-INS-I parameter in MAFFT v7.508 and subsequently trimmed using TrimAl v.1.4 with the --automated1 parameter. Maximum likelihood (ML) phylogenetic reconstruction was performed using IQ-TREE 2 v.2.2.0.3 (--seed 12345, 1000 bootstraps replicates).<sup>16</sup> Sequence homologs in *P. flanaganii* that we placed outside of the human CDK clade were removed and the resulting tree was rooted at the midpoint.

To create a fungal species tree (Figure 4), single copy protein markers were identified in our genomes using BUSCO v.2<sup>17</sup> and the Dikarya database. We selected a set of 100 complete protein marker sequences that were shared among all fungal genomes of interest and were used to create a multilocus phylogenetic species tree. Each marker protein set was individually aligned using MAFFT v7.508 (--maxiterate 1000) and the resulting alignment was trimmed with TrimAl v.1.4 (--automated1). The trimmed alignments were then concatenated to create a super-alignment, which was used for maximum likelihood (ML) phylogenetic reconstruction with FastTreeMP v2.1.11<sup>18</sup> (--gamma parameter, 1000 bootstrap replicates).

To create a *Penicillium* species tree (Figure S2), we imported the CaM, BenA, and RPB2 gene sequences for 24 *Penicillium* species in the sect. *Excilicaulis* (with a focus on the *P. restrictum* clade) from public sources and from our proprietary *P. flanaganii* genome. Datasets containing sequences from each gene were individually aligned using the L-INS-I parameter in MAFFT v7.508 and subsequently trimmed using TrimAl v.1.4 (--automated1). The resulting trimmed alignments were then concatenated and used for maximum likelihood (ML) phylogenetic reconstruction using IQ-TREE 2 v.2.2.0.3 (K2P+G model, --seed 12345, 1000 bootstraps replicates).

## Filamentous fungal engineering

Engineering constructs for promoter replacement of *rosH* and KO of *rosJ* in *Mycophilomyces sp.*, as well as KO of *rosA—rosK* in *Penicillium sp.*, were built using a standard HiFi Assembly protocol (NEB) to assemble overlapping fragments with a pBIIKS vector backbone (linearized by restriction digest using XbaI, KpnI (NEB)). For the TF-OE promoter replacement construct, fragments consisted of *i*) 5' target homology, PCR from gDNA (VEG2426, VEG2427) *ii*) *Mycophilomyces sp. gpdA* promoter, Gene Fragment (gVEG30, GeneWiz) *iii*) *hygR* hygromycin resistance cassette, PCR from preassembled vector: [pBIIKS backbone; *E. coli hph* (UniProt:P00557, GeneWiz); in-frame CDS fusion with eGFP (UniProt:C5MKY7), Gene Fragment (GeneWiz); *Aspergillus nidulans A4 trpC* terminator Gene Fragment (trpCter, GeneWiz); *iv*) *Mycophilomyces sp. actin* promoter, Gene Fragment (gVEG32, GeneWiz); *v*) 3' target homology, PCR from gDNA (VEG2430, VEG2431). For the PKS-KO knockout construct, fragments consisted of *i*) 5' target homology, PCR from gDNA (VEG2395, VEG2396); *ii*) *Mycophilomyces sp. tef1* promoter, Gene Fragment (gVEG28, GeneWiz); *iii*) *nrsR* nourseothricin resistance cassette, PCR from preassembled vector: [pBIIKS backbone; *Streptomyces noursei nat1* (UniProt:Q08414), Gene Fragment (GeneWiz); *Aspergillus nidulans A4 trpC* terminator, Gene Fragment (trpCter, GeneWiz)]; *iv*) 3' target homology, PCR from gDNA (VEG2397, VEG2398). For individual *Penicillium sp. ros* gene KO constructs, fragments consisted of *i*) 5' target homology, PCR from gDNA (PC00379—PC00422); *ii*) GFP/*hygR* hygromycin resistance cassette, PCR from preassembled vector [pBIIKS backbone; mGFP5, Gene Fragment (GeneWiz); *Aspergillus nidulans A4 tubulin* terminator, Gene Fragment (tubter, GeneWiz); *Aspergillus nidulans A4 trpC* terminator, Gene Fragment (trpCter, GeneWiz); *E. coli hph* (UniProt:P00557), Gene Fragment (GeneWiz)]; *iii*) *Penicillium sp. tef1* promoter, Gene Fragment (g001019, Twist); 3' target homology, PCR from gDNA (PC00379—PC00422). Correct assembly of fragments without mutations was confirmed by NGS (MiSeq, Illumina). Before transformation, constructs were amplified by PCR in two parts (PC00628—PC00635), with the coding sequence of the resistance marker split between each part and 325 bp (*nrsR*) or 762 bp (*hygR*) of overlap between parts. The resistance marker is functionally restored upon homologous recombination at the target site. Targeted integration was supported by introduction of Cas9 as a ribonucleoprotein

(RNP) complex with two synthetic guide RNAs (sg1/2-g1:g11, IDT) binding the target site during transformation.

Transformation of *Penicillium sp.* (*Psp*) and *Mycophilomyces sp.* (*Msp*) was achieved by inoculating potato dextrose broth (50 mL, Difco BD Biosciences) with a suspension of  $10^8$  conidia followed by 18 hr incubation at 220 rpm, 25 °C. Germinated conidia were collected by centrifugation (*Psp*, 4000 g, 10 min) or by filtration through a double layer of Miracloth (*Msp*, 22-25  $\mu$ M average pore size), washed twice with a solution of MgSO<sub>4</sub> (0.6 M), and resuspended in a protoplasting solution consisting of Lysing Enzymes from *Trichoderma harzanium* (10 mg/mL, Sigma), Yatalase (2 mg/mL [*Psp*] or 1 mg/mL [*Msp*], Takara) in OM buffer (MgSO<sub>4</sub> 1.2 M, phosphate buffer 10 mM pH 5.8, BSA 1.2 mg/mL) followed by incubation at 32 °C, 100 rpm, 1 hr (*Psp*) or 30 °C, 100 rpm, 3 hr (*Msp*). Protoplasts were collected by centrifugation (3500 g, 15 min [*Psp*] or 2750 g, 20 min [*Msp*]) and resuspended in 15 mL ST buffer (sorbitol 1M, tris 10 mM pH 7.5), followed by two iterations of centrifugation and resuspension in 35 mL STC buffer (sorbitol 1.2 M, tris 10 mM pH 7.5, CaCl<sub>2</sub> 10 mM), then a final centrifugation and resuspension in 1 mL STC buffer. Cas9 enzyme (1.5  $\mu$ L, EnGen Spy, NEB) was incubated with two sgRNA samples (0.75  $\mu$ L each, 100  $\mu$ M) for 5 minutes at room temperature (RT). Each of the two parts of the engineering construct prepared by PCR was added to the Cas9 rnp (3.5  $\mu$ L, 0.5  $\mu$ g) and incubated for 10 minutes at RT, followed by addition of STC buffer (10  $\mu$ L) and a suspension of protoplasts (100  $\mu$ L,  $10E8$ /mL). The mixture was incubated for 20 minutes at RT, followed by addition of PTC buffer (1 mL, PEG 4000 60% v/v, tris 10 mM pH 7.5, CaCl<sub>2</sub> 10 mM) and incubation for 20 minutes at RT. Protoplasts were transferred to a mixture of STC buffer (1 mL) and recovery medium (2 mL, Czapek-Dox broth (Difco, BD Biosciences), sorbitol 1M), before incubation at 37 °C, 80 rpm for 2 hr (*Psp*) or 30 °C, 100 rpm overnight (*Msp*). Molten recovery agar (15 mL, Czapek-Dox agar (Difco, BD Biosciences), sorbitol 1M) equilibrated to 55 °C in a water bath was added to the protoplasts, followed by selection antibiotic (hygromycin 20  $\mu$ g/mL or nourseothricin 500  $\mu$ g/mL) and the mixture was plated and incubated for 3-5 days at 30 °C until transformants were visible. Transformants were selected and transferred to a secondary plate (Czapek-Dox agar with antibiotic) before incubation for 3 days at 30 °C. Colonies growing on secondary plates were streaked on isolation plates (Czapek-Dox agar with antibiotic) for

single colony selection, and single colonies were screened by WGS (MiSeq, Illumina) for expected integration at the target site and single copy integration of the construct (*Psp*) or by diagnostic PCR (*Msp*) using primers spanning each homology targeting region (diag1967—diag2745).

### **Yeast expression and assay of *rosA***

The gene *rosA* and its partner cytochrome P450: reductase (*cprA*) from *Penicillium sp.* were codon-optimized for *S. cerevisiae* and ordered as synthetic genes cloned in pD1204 and pD1201 backbones respectively (ATUM).

Vectors containing *rosA* and *cprA* were cotransformed into *S. cerevisiae* BY4742 (ATCC) using the Frozen-EZ Yeast Transformation II kit (Zymo) and transformants were selected on SC-URA/LEU (Sunrise Science). The GAL1 promoter controlling expression of *rosA* and *cprA* was induced in buffered YP + 2% galactose broth using the protocol specified for the pD1204/pD1201 vectors (ATUM) scaled up to 100 mL volume. After induction, cultures were incubated for 24 hr, and cell pellets were collected by centrifugation and lyophilized. Biotransformation assays were set up in 2mL vials with 10 mg lyophilized cells resuspended in 180  $\mu$ L tris pH 8 buffer + glycerol (10%), followed by addition of 10  $\mu$ L compound **5** (5 mM in DMSO) and 10  $\mu$ L NADPH (20 mM in water). Reactions were incubated at 30°C for 3 hrs with shaking at 750 rpm (ThermoMixer C, Eppendorf). Samples were quenched with addition of 1 vol acetonitrile, centrifugation and filtration of the supernatant (0.22  $\mu$ m) before analysis by LC-MS.

### ***E. coli* expression and assay of *rosD*.**

The gene *rosD* was codon optimized for *E. coli* expression (Geneious Prime) and ordered as a Gene Fragment with terminal Bsal digestion sites and AACC/GTCC overhangs (Twist). The *rosD* expression plasmid was assembled by Golden Gate cloning of the Gene Fragment into a pET21a backbone (amplified by PCR with compatible overhangs) with introduction of an in-frame His6-TEV N-terminal fragment (Twist). The plasmid was used to transform *E. coli* BL21 (DE3) and expressed with autoinduction in ZYM-5052 medium (N-Z amine 1% w/v, yeast extract 0.5% w/v, Na<sub>2</sub>HPO<sub>4</sub> 25 mM, KH<sub>2</sub>PO<sub>4</sub> 25 mM, NH<sub>4</sub>Cl 50 mM, Na<sub>2</sub>SO<sub>4</sub> 5 mM, MgSO<sub>4</sub> 2 mM, trace metals 0.2x, glycerol 0.5% w/v, glucose 0.05% w/v, lactose 0.2% w/v) at 25 °C (24 hrs). After collection of the cell pellet by

centrifugation (4000 *g*, 5 min) and lysis using BugBuster Protein Extraction Reagent (EMD Millipore), the protein was purified on a HisTrap FF (Cytiva) column using the vendor protocol, and buffer exchanged into 2X phosphate buffered saline (PBS, Gibco) on a PD-10 column (Cytiva). Functional assays were set up in 100  $\mu$ L total volume with sodium phosphate (50 mM, pH 7.5) buffer, RosD enzyme (7  $\mu$ M), substrate **1** (350  $\mu$ M), DMSO (1% v/v); incubation for 1 hr at 30 °C with shaking at 500 rpm (ThermoMixer C, Eppendorf) before quenching with 100  $\mu$ L acetonitrile, centrifugation at 16000 *g* for 5 min, filtration (0.22  $\mu$ M) and analysis by LC-MS [3  $\mu$ L injection, Kinetex 2.6  $\mu$ m polar C18 100 Å 150x3.0 mm column, 45 min gradient (A: water + 0.1% formic acid, B: ACN + 0.1 % formic acid; 0 min 5% B, 30 min 95% B, 32 – 36 min 100% B, 37 – 45 min 5% B, 0.8 mL/min)].

### **Expression of fungal Pho85 and human CDKs**

The housekeeping *pho85* sequence (encoding residues 63–355) from a LifeBase strain *Aspergillus sp.* (Sect. *Nigri*, *A. aculeatus* clade) containing the *ros* BGC, as well the WT and a mutant sequence (V69L, F85S, G150A) of *pho85* from *Mycophilomyces sp.* (residues 1–294) were codon optimized for *E. coli* expression and cloned into pET28a with a synthetic N-terminal His6-Avi-TEV fusion (GenScript). Expression plasmids were used to transform *E. coli* BL21 (DE3) RIL (Agilent), and induced in TB medium (3L, Fisher Bioreagents) after incubation at 37 °C, 180 rpm to OD<sub>600</sub> 1.0 by addition of IPTG (350 mM) followed by incubation for 18 hr at 18 °C. Cell pellets were collected by centrifugation (10000 *g*, 10 min), frozen at -20 °C, thawed at room temperature and resuspended to homogeneity in lysis buffer 1 (50 mL, 50 mM HEPES pH 8.0, 150 mM NaCl, 1 mM MgCl<sub>2</sub>, 5 mM  $\beta$ -mercaptoethanol, 1X ProBlock Gold protease inhibitor cocktail (GoldBio), 10  $\mu$ g/mL DNase I (Roche), 1 mg/mL lysozyme). Cell lysis was achieved by sonication (Branson SFX550) in 20 cycles (30 s on, 60 s off), and suspensions were cleared by centrifugation at 25000 *g* for 30 min at 4 °C. Cleared lysates were loaded onto Ni-NTA resin (3 mL, Qiagen) in a gravity column, washed with 60 mL purification buffer 1 (50 mM HEPES pH 8.0, 150 mM NaCl, 1 mM MgCl<sub>2</sub>, 5 mM  $\beta$ -mercaptoethanol) + 15 mM imidazole, followed by 60 mL purification buffer 1 + 30 mM imidazole. His-tagged proteins were eluted by adding purification buffer 1 + 250 mM imidazole (5 mL) to the resin, incubating for 2 minutes, then eluting; elution was repeated four times. Pooled eluant

fractions were concentrated to 1 mL in a centrifugal filter unit (Amicon Ultra 3k MWCO, Millipore Sigma), and diluted with purification buffer 1 (18.5 mL). TEV protease (NEB) was added in a 1:50 ratio (w/w) with total protein measured by NanoDrop (Thermo), followed by incubation at 4 °C, 18 hrs. After tag cleavage, the protein solution was added to 2 mL Ni-NTA resin in a gravity column and incubated for 20 min at 4 °C. Untagged proteins were collected in the flowthrough, followed by washing twice with purification buffer 1 + 15 mM imidazole (5 mL). Combined protein fractions were concentrated to 10 mL in a centrifugal filter unit (3k MWCO) and loaded onto a HiLoad 26/600 Superdex 75 size exclusion column (Cytiva) equilibrated in SEC buffer (50 mM HEPES pH 8.0, 150 mM NaCl, 5 mM DTT, 10% glycerol) for final purification of the intended protein. Protein was concentrated to 2.5 mg/mL and stored at -80 °C before assay.

Yeast CAK1 (P43568 residues 1-368), human CDK2 (P24941 residues 1-298), and CCNA2 (P24864 residues 172-432) sequences were codon optimized for expression in *E. coli*; CDK2 with an N-terminal His6-GST fusion / TEV cleavage site and CAK1 with an N-terminal GST fusion were synthesized and cloned together into pETDuet1 (GenScript), while CCNA2 with an N-terminal His6 fusion / TEV cleavage site was synthesized and cloned into pET28a (GenScript). Each construct was used to transform *E. coli* BL21 (DE3), and protein production was induced in expression medium (10 L TB [CDK2/CAK1] or 3 L LB [CCNA2], Fisher Bioreagents) after incubation at 37 °C, 180 rpm to OD<sub>600</sub> 1.0 by addition of IPTG (80 µM) followed by incubation for 20 hr at 18 °C. Cell pellets were collected by centrifugation (10000 g, 10 min) and frozen at -20 °C. Pellets were resuspended to homogeneity in lysis buffer 2 (50 mL, 50 mM HEPES pH 8.0, 250 mM NaCl, 1 mM MgCl<sub>2</sub>, 5 mM β-mercaptoethanol, 5% glycerol, 1X ProBlock Gold protease inhibitor cocktail (GoldBio), 10 µg/mL DNase I (Roche), 1 mg/mL lysozyme). Cell lysis was achieved by sonication (Branson SFX550) in 10 cycles (30 s on, 60 s off), and suspensions were cleared by centrifugation at 35000 g for 30 min at 4 °C followed by filtration (0.45 µm). Cyclin A2 clarified lysate was loaded onto a HisTrap FF 5 mL column (Cytiva), washed with 25 mL wash buffer (50 mM HEPES pH 8.0, 250 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol), 10 mL wash buffer + 3% elution buffer (50 mM HEPES pH 8.0, 250 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol, 500 mM imidazole), 10 mL wash buffer + 6% elution buffer, followed by elution in a linear gradient

from 6% to 100% elution buffer over 50 mL. Fractions containing cyclin A2 were combined and loaded on a HiPrep 26/10 Desalt column equilibrated with wash buffer then eluted with 1.5 column volumes of wash buffer. Fractions containing cyclin A2 were combined and the N-terminal tag was cleaved by addition of TEV protease (1:50 ratio w/w with total protein) and incubation at 4 °C for 16 hr. The protein solution was loaded onto a HisTrap FF 5 mL column equilibrated in wash buffer and eluted with 5 mL wash buffer. Fractions containing cyclin A2 were pooled and loaded onto a HiPrep 26/600 75 µg size exclusion column equilibrated in SEC buffer 2 (50 mM HEPES pH 8.0, 150 mM NaCl, 5 % glycerol, 1 mM TCEP) and eluted in 1 column volume SEC buffer 2. Fractions containing cyclin A2 were pooled and concentrated to 2.5 mg/mL in a centrifugal filter unit (Amicon Ultra 3k MWCO, Millipore Sigma). CDK2 / CAK1 clarified lysate was mixed with Glutathione Agarose Resin (4 mL, GoldBio) with stirring for 1 hr before transfer to a gravity column. The resin was washed three times with 50 mL lysis buffer 2 and bound protein eluted with 20 mL lysis buffer 2 + glutathione (5 mM). To complete phosphorylation, MgCl<sub>2</sub> (5 mM) and ATP (2 mM) were added to the eluant and incubated for 2 hr. Fractions containing CDK2 / CAK1 were combined and loaded on a HiPrep 26/10 Desalt column equilibrated with wash buffer then eluted with 1.5 column volumes of wash buffer. Glutathione Agarose Resin (4 mL, GoldBio) was added to the protein solution and stirred at 4 °C, 1 hr before transfer to a gravity column and flow-through of the solution. Purified cyclin A2 solution was added to the resin and incubated for 10 minutes to allow CDK2 / cyclin A2 complex formation, before flow-through of the unbound protein; this step was repeated once, then the resin was washed twice with 50 mL lysis buffer 2. Bound proteins were eluted from the resin with 20 mL lysis buffer 2 + glutathione (5 mM). TEV protease was added to the eluant (1:50 w/w ratio with total protein) and incubated at 4 °C, 16 hr. The protein solution was mixed with Glutathione Agarose Resin (4 mL, GoldBio) with stirring for 1 hr, to bind residual CAK1. The mixture was transferred to a gravity column and the flow-through collected. The flow-through was loaded onto a HiPrep 26/600 75 µg (Cytiva) size exclusion column equilibrated in SEC buffer 2, and eluted with 1 column volume SEC buffer 2. The purified CDK2 / cyclin A2 complex was concentrated to 2.5 mg/mL on a centrifugal filter unit (Amicon Ultra 3k MWCO, Millipore Sigma) and stored at -80 °C.

CAK1 (residues 1-368), CDK2 (residues 1–298) and CCNE1 (residues 96–378) sequences were codon optimized for expression in an insect cell system (GenScript); CDK2 with an N-terminal His6-GST fusion / TEV cleavage site and CCNE1 with an N-terminal GST fusion / TEV cleavage site were synthesized and cloned together into pFastBacDual (GenScript); CAK1 was synthesized and cloned into pFastBac HTb (GenScript). The vector was used to transfect Sf9 insect cells from the Bac-to-Bac Baculovirus Expression System (Gibco) using standard manufacturer protocols. High Five cells (Gibco,  $2 \times 10^6$  / mL) were infected with both the CDK2/CCNE1 and CAK1 constructs via P2 virus at a ratio of 1:100 (v/v) and 1:500 (v/v), incubated at 27 °C for 48 hr, then collected by centrifugation. The pellet was resuspended to homogeneity in lysis buffer 3 (50 mL, 50 mM HEPES pH 8.0, 250 mM NaCl, 5% glycerol, 0.5 mM TCEP 1X ProBlock Gold protease inhibitor cocktail (GoldBio), 10 µg/mL DNase I (Roche), 1 mg/mL lysozyme). Cell lysis was achieved by sonication (Branson SFX550) in 20 cycles (30 s on, 60 s off), and suspensions were cleared by centrifugation at 40000 g for 50 min at 4 °C followed by filtration (0.45 µm). Clarified lysate was mixed with Glutathione Agarose Resin (14 mL, GoldBio) with stirring at 4 °C, 1 hr before transfer to a gravity column. The resin was washed four times with 100 mL lysis buffer 2 followed by addition of 15 mL lysis buffer 2 + TEV protease (11 mg) to the column, mixing, and incubation for 24 hr at 4 °C. Untagged protein was eluted from the column with 10 mL lysis buffer 2. Eluted proteins were loaded onto a HisTrap HP column (5 mL, Cytiva), washed with 5 mL wash buffer 2 (50 mM HEPES pH 8.0, 250 mM NaCl, 5% glycerol, 0.5 mM TCEP) and eluted with a linear gradient of 0–100% elution buffer 2 (50 mM HEPES pH 8.0, 250 mM NaCl, 5% glycerol, 0.5 mM β-mercaptoethanol, 500 mM imidazole) over 75 mL. Fractions containing CDK2 / cyclin E1 were concentrated on a centrifugal filter unit (Amicon Ultra 10k MWCO) and loaded onto a HiLoad 16/600 200 µg size exclusion column (Cytiva) equilibrated in SEC buffer 2, followed by elution in 1.5 column volumes SEC buffer 2. Purified protein was concentrated on a centrifugal filter unit (Amicon Ultra 10k MWCO) to 8.7 mg/mL and stored at -80 °C.

### **In vitro kinase functional assays**

Pho85 inhibition assays were carried out on fractionated fungal extracts and purified GEMs via chemiluminescent assay of ATPase activity. Compounds (1 µL) were added to

ATP (12.5  $\mu$ L, 0.6 mM) and purified Pho85 (12.5  $\mu$ L, 2  $\mu$ M) in a black 96-well assay plate, and incubated for 3 hr. Samples (5  $\mu$ L) were transferred to a white 96-well detection plate alongside ATP-Glo (5  $\mu$ L, Promega) and incubated for 40 min. Kinase detection reagent (10  $\mu$ L, Promega) was added, incubated for 30 min, and chemiluminescence was measured at gain 2500 on a Clariostar plate reader (BMG Labtech).

Kinetic assays were performed by monitoring phosphorylation of a Sox peptide (AQT0297, AssayQuant) by CDK2 / cyclin E1 (Carna Biosciences) using a stopped flow fluorescence spectrometer (SX-20, Applied Photophysics). Solution A: ATP (50  $\mu$ M) and AQT0297 (10  $\mu$ M) with varied concentrations of inhibitor in assay buffer 1 (50 mM HEPES pH 7.5, 10 mM MgCl<sub>2</sub>, 0.01% n-octyl  $\beta$ -D glucopyranoside, 50  $\mu$ M TCEP, 4% DMSO), and solution B: CDK2 / cyclin E1 (35 nM) in assay buffer 1 were prepared; each solution was used to fill one syringe of the stopped flow system. Each analyte concentration was flushed five times using the drive function followed by three acquisitions with 0.5 s intervals (excitation 360 nm, PMT filter 435 nm, voltage 520 V).

### **Protein intact mass spectrometry**

Intact MS analysis to identify covalent binding of inhibitors was achieved by dilution of compound **2** (1  $\mu$ L, 10 mM in DMSO) in assay buffer 2 (24  $\mu$ L, 50 mM HEPES pH 7.5, 25 mM NaCl, 10 mM MgCl<sub>2</sub>, 1 mM TCEP, 0.01% n-octyl  $\beta$ -D glucopyranoside). Compound **1** or **2** (20  $\mu$ L) was mixed with protein samples (20  $\mu$ L, 10  $\mu$ M in assay buffer 2, *A. nidulans* PhoA or CDK2 / cyclin E1), centrifuged at 1000 g, 5 min, and incubated for 1 hr at room temperature. Reactions (5  $\mu$ L) were analyzed by LC-MS (2.1x50 mm C4, Phenomenex; Q-Exactive, Thermo). For peptide mapping analysis, samples were proteolyzed by buffer exchange into NH<sub>4</sub>HCO<sub>3</sub> (50 mM pH 7.8) followed by addition of trypsin (1:50 w/w ratio with total protein) and incubation at 37 °C for 18 hr. The reaction was quenched by addition of TFA (0.5%) and desalted on a SepPak C18 column (50 mg, 1cc, Waters) conditioned in water, washed twice with trifluoroacetic acid (0.1%) in water (500  $\mu$ L) and eluted with two applications of 80% acetonitrile / 20% formic acid (0.1%) in water (250  $\mu$ L). The eluant was dried in a vacuum centrifuge for 6 hr at room temperature, reconstituted in 3% acetonitrile / 97% formic acid (0.1%) in water before analysis by LC-MS (EASY-Spray Nano, Thermo; Eclipse Tribrid, Thermo).

### **Protein X-ray crystallography**

The CDK2 / cyclin E1 protein complex was diluted to 5.0 mg/mL with dilution buffer (20 mM Tris-HCl, 150 mM NaCl, pH 8.0). For co-crystallization with ligand **2**, the protein solution was supplemented with **2** to 0.256 mM using a 100mM stock in DMSO. Crystals were obtained by mixing 210 nL of CDK2 / cyclin E1 / ligand **2** to 70 nL of precipitant containing 100 mM Tris-HCl, 2.8 M NaCl, pH 8.0 and equilibrated against 80  $\mu$ L of reservoir. For soaking with XC208, apo crystals were obtained by mixing 400 nL CDK2 / cyclin E1 to 200 nL precipitant containing 100 mM Tris-HCl, 3.6 M NaCl, pH 7.5 and equilibrated against 80  $\mu$ L of reservoir. Well solution was used to dilute XC208 to 2.5 mM using a 100 mM stock in DMSO; apo crystals were moved to 2.5  $\mu$ L XC208 solution and soaked overnight. For soaking with XC219, apo crystals were obtained by mixing 400 nL CDK2 / cyclin E1 to 200 nL precipitant containing 100 mM Tris-HCl, 3.76 M NaCl, pH 7.0 and equilibrated against 80  $\mu$ L of reservoir. Well solution was used to dilute XC208 to 2.5 mM using a 50 mM stock in DMSO; apo crystals were moved to 2.5  $\mu$ L XC219 solution and soaked for 3 hours. Crystals typically appeared within 1–3 days at 20 °C and were cryo-protected with 4.0M Lithium formate and frozen in liquid nitrogen. Single-wavelength data were collected remotely at the EMBL-ESRF ID30A-1 beamline at 12.841 keV (0.96550 Å) for ligand **2**, BNL-NLSL-II 17-ID-2 beamline at 12.659 keV (0.97934 Å) for XC208, and EMBL-ESRF ID30B beamline at 12.668 keV (0.97872 Å) for XC219. Data were integrated and scaled using the Xia2<sup>19</sup> automated pipeline which utilizes CCP4,<sup>20</sup> DIALS,<sup>21</sup> and AIMLESS<sup>22</sup> and phased using molecular replacement in PHASER.<sup>23</sup> Iterative rounds of model building were performed in COOT<sup>24</sup> with refinement in REFMAC<sup>25</sup> for ligand **2** and XC208 or PHENIX<sup>26</sup> for XC219 respectively.

### **Fungal fermentation and GEM extraction**

Seed cultures were prepared by inoculation of KF medium (55 mL, tomato paste 40 g/L, corn steep solids 5 g/L, oat flour 10 g/L, glucose 10 g/L, KF trace elements 10 g/L [FeSO<sub>4</sub>.7H<sub>2</sub>O 1 g/L, MnSO<sub>4</sub>.4H<sub>2</sub>O 1 g/L, CuCl<sub>2</sub>.2H<sub>2</sub>O 0.025 g/L, CaCl<sub>2</sub> 0.1 g/L, H<sub>3</sub>BO<sub>3</sub> 0.056 g/L, (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>.4H<sub>2</sub>O 0.019 g/L, ZnSO<sub>4</sub>.7H<sub>2</sub>O 0.2 g/L], pH 6.8) with frozen vegetative mycelial stocks of fungal strains, incubation for 3 days, 25 °C, 220 rpm. For analytical-scale experiments, seed culture was used to inoculate three points (0.24% v/v) on agar plates (25 mL, 100 mm), or was used to inoculate (5% v/v) agar flasks (45 mL

agar in 250 mL erlenmeyer flask) consisting of a standard panel of seven media (CM [glucose 10 g/L, peptone 2 g/L, yeast extract 1 g/L, casein hydrolysate 1 g/L, sodium nitrate 6 g/L, potassium chloride 0.52 g/L, magnesium sulfate heptahydrate 0.52 g/L, potassium phosphate monobasic 1.52 g/L, Kao and Michayluk Vitamin Solution 1X, Hutner's Trace Elements 1X, agar 20 g/L, pH 6.5], PDA (Difco, BD Biosciences), MEA (Difco, BD Biosciences), ICI [glucose 80 g/L, potassium phosphate monobasic 5 g/L, magnesium sulfate heptahydrate 1 g/L, ICI trace elements 2 g/L [FeSO<sub>4</sub>·7H<sub>2</sub>O 1 g/L, CuSO<sub>4</sub>·5H<sub>2</sub>O 0.15 g/L, ZnSO<sub>4</sub>·7H<sub>2</sub>O 1.61 g/L, MnSO<sub>4</sub>·4H<sub>2</sub>O 0.08 g/L, (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>·4H<sub>2</sub>O 0.1 g/L], agar 15 g/L, pH 5.5] + 120 mM NaNO<sub>3</sub>, ICI + 6 mM NaNO<sub>3</sub>, ICI + 60 mM glutamine, ICI + 6 mM glutamine) followed by incubation at 25 °C for 7 or 14 days. For preparative-scale fermentation, seed culture was used to inoculate (5% v/v) 450 mL ICI-variant medium [glucose 20 g/L, potassium phosphate monobasic 5 g/L, magnesium sulfate heptahydrate 1 g/L, ICI trace elements 2 g/L, pH 5.5] + 6 mM glutamine in Fernbach flasks (2.8 L), followed by static incubation for 7 days.

For chemical extraction at analytical scale, colonies were removed with attached agar from microplates using a knife, and DCM was added (60 mL), followed by homogenization and shaking at room temperature (200 rpm, 18 hr). Samples were filtered (20 µm) and the filtrate evaporated under vacuum. Extracts were dissolved in acetonitrile (100 µL), sonicated in a water bath for 15 min, followed by addition of water (300 µL). An aliquot (200 µL) of each extract was filtered (0.45 µm) and analyzed by LC-MS (Kinetex 2.6 µm Polar C18 100 Å 50 x 3.0 mm [Phenomenex]; solvent A: water + 0.1% formic acid, solvent B: acetonitrile + 0.1% formic acid; isocratic 50% B; 0.8 mL/min; targeted MS detection of roseopurpurins using QTRAP 6500+ [SciEx] with MRM [255.000->209.000, 255.000->105.000, 379.000->253.000, 349.000->165.000, 367.000->165.000]). For preparative isolation of roseopurpurin C, batches of up to 100 Fernbach flasks (45 L total volume) were filtered on two layers of Miracloth (Calbiochem), collecting both the filtrate in 10 L carboys (8 L / carboy) and the retained mycelium. For each batch of 8 L filtrate, a mixture of Amberlite XAD7HP (600 g, Supelco) and XAD16N (150 g, Supelco) resins was prepared by washing three times with water (5 volumes), three times with hexane (5 volumes), three times with isopropyl alcohol, then three times with water, stirring for 15 min between each wash, before addition to the culture filtrate and incubation at room

temperature with stirring (16 hr, 300 rpm). The resin was collected by vacuum filtration and extracted three times with ethyl acetate (1.2 L) and once with isopropyl alcohol (1 L). The residual aqueous fraction in the extracts was collected by separatory funnel and extracted two times with ethyl acetate (1 L). The filtered mycelium was homogenized with 1:1 acetonitrile / acetone (2L) in a blender, sonicated for 30 min and the organic extract was collected after vacuum filtration. All organic extracts were combined and evaporated under vacuum. Roseopurpurins were purified from crude culture extracts in two steps by flash column chromatography (RediSep Silica 300 g Gold [Teledyne ISCO]; solvent A: hexane, solvent B: ethyl acetate; %B: 0 min 0%, 4 min 0%, 6 min 35%, 31 min 35%, 34 min 100%, 44 min 100%; 200 mL/min) followed by preparative HPLC (Luna C18(2) 100 Å 30 x 100 mm [Phenomenex]; solvent A: water + 0.1% formic acid, solvent B: acetonitrile + 0.1% formic acid; %B: 0 min 35%, 11.5 min 35%, 11.75 min 100%, 14.75 min 100%, 15 min 35%, 17 min 100%; 40 mL/min).

For RNA extraction, plugs were taken from the outer edge and center of each colony using a Transfertube (1.6 mm ID, Spectrum) and combined in a collection tube prefilled with lysing beads (0.5 & 1.0 mm ZR BashingBead Lysis Matrix, Zymo) followed by addition of DNA/RNA Shield reagent (500 µL, Zymo) before freezing in liquid nitrogen and storage at -80 °C. Samples were thawed (30 min, RT) and lysed by bead beating (150 s, Mini-BeadBeater-96 [BioSpec]), then centrifuged (4000 g, 4 °C, 5 min). RNA was purified from the supernatant (300 µL) using the Quick-RNA 96 kit (Zymo) after addition of lysis buffer (300 µL), by following the manufacturers protocol, eluting from the kit columns in RNase-free water (60 µL). Aliquots of each RNA sample (44 µL) were added to Turbo DNase master mix (1X, 6 µL, Thermo) and incubated (30 min, 37 °C) followed by addition of EDTA (1.5 µL, 0.5 M) and enzyme inactivation (10 min, 75 °C). Samples were cleaned using the RNA Clean-Up and Concentration 96-Well Kit (Norgen) following the manufacturers protocol, eluting in 75 µL elution buffer. RNA samples were stored at -80 °C.

### **RNA sequencing**

RNA samples were quantitated using the Quant-iT RNA BR Assay Kit (Invitrogen) and integrity (RIN) was analyzed using a Fragment Analyzer (Agilent). Strand-specific RNA

sequencing libraries were constructed using the RNA Hyperprep Kit (KAPA), with rRNA depletion using the QIAseq FastSelect rRNA Yeast Kit (Qiagen) and addition of ERCC Spike-In Mix 1 (Invitrogen), following all manufacturers protocols. Fragmentation was performed dependent on RIN: samples with RIN < 5 were fragmented with cycling program A (75 °C, 2 min; 70 °C, 2 min; 65 °C, 2 min; 60 °C, 2 min; 55 °C, 2 min; 37 °C, 5 min; 25 °C, 5 min), while those with RIN > 5 were fragmented with cycling program B (85 °C, 2 min before proceeding with cycling program A). Libraries were quantitated with the Quant-iT™ 1X dsDNA Kit (Invitrogen) and size distribution was measured on a Fragment Analyzer. Samples were pooled, normalized to 2 nM, and underwent a final cleanup by addition of AMPure XP beads (1:1 ratio, Beckman), incubation for 5 min, removal of supernatant, washing twice with 80% EtOH (200 µL), and elution with elution buffer (1 vol wrt input). Final libraries were sequenced on a NovaSeq 6000 (Illumina) using an S2 flowcell with 50 bp paired end reads.

### **RNAseq expression analysis**

RNAseq FASTQ were subject to quality control using fastqc v0.11.9 (Andrews 2010) and trimmed using Trimmomatic v39-2<sup>27</sup> in paired-end mode with the following parameters ILLUMINACLIP:TrueSeq3-PE.fa:2:30:7:2:TRUE LEADING:10 TRAILING:10 SLIDINGWINDOW:4:12 MINLEN:20. Ribosomal and mitochondrial RNA was computationally depleted using Bowtie2 v2.4.4<sup>28</sup> with the -fast-local parameter. Reads were then mapped to respective genomes using STAR v.2.7.9.a.<sup>29</sup> STAR was first index using the genomeGenerate run mode and --genomeSAindexNbases 10. Subsequent alignment was performed with the following parameters --outFilterMultimapNmax 1 --outFilterMismatchNoverLmax .05 --alignIntronMax 3000. Feature quantification was performed using featureCounts v2.0.3<sup>30</sup> and normalized to transcripts per million (TPM). Differential expression analysis was performed using DESeq2 v1.42.1 with standard parameters.<sup>31</sup>

### **Human and fungal cell culture and lysis conditions for chemoproteomics experiments**

All cell lines (COLO 205, K-562, SK-N-BE(2), and MV-4-11) were maintained in a humidified incubator at 37 °C and were cultured in the ATCC-recommended media

supplemented with 10% (v/v) fetal bovine serum and 2 mM L-glutamine. Suspension cell lines (K-562 and MV-4-11) were harvested by centrifugation at 400 g for 5 min at room temperature (RT), followed by washing with PBS, flash-freezing in liquid nitrogen, and storage at -80 °C. Mixed adherent-suspension cell lines (COLO 205 and SK-N-BE(2)) were harvested as described above to collect the suspended portion of the cells, and the adherent portion was collected by trypsinization, followed by washing and storage, as described above. Cell lysate was prepared by resuspending frozen cell pellets in cell lysis buffer (50 mM Tris-HCl pH 7.5, 5% glycerol, 1.5 mM MgCl<sub>2</sub>, 150 mM NaCl, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 1 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl), 25 mM NaF, 0.4% NP-40, and 1x Halt EDTA-free protease/phosphatase inhibitor cocktail (Thermo Scientific 78440)), followed by probe-sonication on ice. The resulting lysates from each cell line were centrifuged at 15,000 g for 15 min at 4 °C and the supernatant was collected, normalized to a protein concentration of 5 mg/mL by BCA assay (Thermo Scientific 23225), and combined in equal portions to provide the final mixed cell lysate.

### ***Saccharomyces cerevisiae* culture and lysis conditions for chemoproteomics experiments**

*Saccharomyces cerevisiae* cultures were initiated from a frozen stock inoculated into potato dextrose broth (PDB). After overnight incubation at 25 °C with 220 rpm shaking, these starter cultures were used to inoculate 50-mL of PDB and incubated under same conditions for 48 h. Cells subsequently were harvested by centrifugation and lysed into the same cell lysis buffer indicated above by bead-beating using a mixture of 0.1- and 0.5-mm beads (Zymo Research S6002-96-3) with five 1-min bead-beating cycles (resting samples on ice between each cycle). The resulting lysate was centrifuged at 15,000 g at 4 °C for 20 min to remove cell debris.

### **Preparation of Kinobead-Sepharose affinity matrix bead mix**

The Kinobead-Sepharose affinity matrices consisted of a combination of five promiscuous kinase inhibitors (probe set KBy) that has been optimized previously for maximal kinome coverage, and the probe molecules were prepared according to published procedure<sup>32</sup>. Briefly, the coupling reactions between compounds and beads were performed overnight in dimethyl sulfoxide (DMSO) using a 1:1 bead slurry, 1 mM kinase probe, and 15 mM

triethylamine, followed by quenching with 50 mM ethanolamine. After the coupling reaction, the five individual Kinobead-Sepharose affinity matrices were combined in equal portions to provide the final Kinobead-Sepharose affinity matrix bead mix (stored as a 1:1 slurry in isopropanol at 4 °C).

### **Kinobead chemoproteomics: Sample preparation, LC-MS analysis, and data processing**

Mixed human cell lysate or *Saccharomyces cerevisiae* lysate was distributed into a deep 96-well plate (1.5 mg protein per well), treated with either DMSO or test compound in duplicate, and incubated at 4 °C for 45 min on a rocker. The Kinobead-Sepharose affinity matrix slurry was distributed into a 96-well fritted filter plate (30 µL per well) and pre-washed three times with cell lysis buffer, followed by addition of compounds in cell lysate. Samples were incubated at 4 °C for 30 min on a rocker, washed twice with ice-cold cell lysis buffer, twice with ice-cold cell lysis buffer lacking NP-40, and twice with ice-cold phosphate-buffered saline (PBS) using a vacuum manifold. Beads were resuspended in 125 µL of 50 mM triethylammonium bicarbonate (TEAB) and treated with 5 mM TCEP for 30 min at 37 °C on a shaker. Reduced cysteines were alkylated with 10 mM iodoacetamide (IA) for 30 min at RT on a shaker, followed by digestion using a mixture of Lys-C and trypsin proteases (1 µg) overnight at 37 °C on a shaker. The filter plate was chilled at 4 °C for 10 min and the peptides were eluted into a deep 96-well plate using a vacuum manifold. Beads were washed twice with 200 µL of 50% water / acetonitrile (MeCN) with 0.1% formic acid (FA) for further peptide elution, and the combined eluent was concentrated to dryness by vacuum centrifugation. The samples were desalted using 96-well Sep-Pak C18 plates (Waters 186003966) and redissolved in 30 µL of water containing 3% MeCN and 1% FA prior to LC-MS analysis.

LC-MS data were acquired in data-independent analysis (DIA) mode using either a QExactive Classic (Thermo Scientific) or an Orbitrap Eclipse Tribrid (Thermo Scientific) mass spectrometer coupled to an EASY nano-LC 1200 instrument (Thermo Scientific) equipped with an EASY-Spray C18 analytical column (75 µm x 50 cm, 2 µm particle size, Thermo Scientific ES903) installed downstream of an Acclaim PepMap 100 C18 trap column (75 µm x 2 cm, 3 µm particle size, Thermo Scientific 164946). Peptides (10-µL

injections) were separated with a 115-min gradient from 3-26% MeCN in water containing 0.1% FA. The DIA-MS method on the Eclipse utilized a 120k-resolution MS1 scan in the Orbitrap ( $m/z$  350-1650) with an automatic gain control (AGC) target of 750% and a maximum injection time (MIT) of 20 ms. Peptide precursor ions were subjected to quadrupole isolation in 36 variable-width  $m/z$  windows spanning the range of 350-1650 Th. Following isolation, precursors were fragmented by higher energy collisional dissociation (HCD) using stepped activation energies of 27, 30, and 33%, and MS2 spectra were recorded in the Orbitrap at 30k resolution using an AGC target of 1000% and MIT of 54 ms. An analogous DIA-MS method was utilized on the QExactive, which deployed an initial 70k-resolution MS1 scan ( $m/z$  400-1200) (AGC target of 5e6 and MIT of 20 ms), followed by quadrupole isolation in 17 50-Th windows spanning the range of 400-1200 Th. Precursors were fragmented by HCD using activation energy of 27, and MS spectra were acquired at 35k resolution (AGC target of 3e6 and MIT set to 'auto').

Raw LC-MS data files were processed in Spectronaut 16 (Biognosys Inc) using the directDIA database search algorithm employing default settings for peptide/protein identification and quantification. Briefly, raw data files were searched against either the Uniprot database of human protein sequences (20,396 entries) or the Uniprot database of *Saccharomyces cerevisiae* (ATCC 204508) sequences (6,727 entries), with cysteine carbamidomethylation as a static modification, along with methionine oxidation, *N*-terminal acetylation, and *N*-terminal methionine loss as variable modifications, and up to two missed cleavages were allowed. Identifications of peptide-spectrum matches (PSMs), peptides, and proteins were filtered to a false-discovery rate (FDR) threshold of 0.01, and quantification was performed based on the MS2 peak area. To enable direct comparison between the human kinome selectivity profiles of the test compounds, the datasets were filtered for protein kinases that were quantified across all experiments. The resulting protein abundance ratios of compound-treated:DMSO groups were plotted and visualized using the online tool Coral<sup>33</sup> (<http://phanstiel-lab.med.unc.edu/CORAL/>).

## **Competitive chemoproteomics with biotin-probe: Sample preparation, LC-MS analysis, and data processing**

Mixed human cell lysate (2 mg/mL total protein) was used for competitive chemoproteomic profiling studies with a biotinylated analog of **2**, compound **50**. Cell lysate (400 µg protein) was incubated with DMSO or 25 µM **2** in triplicate for 1 h at RT on a shaker, followed by incubation with **50** (10 µM) for 1 h at RT. All samples were treated with 1 mL of cold acetone (pre-chilled at -20 °C), centrifuged at 12,000 g at 4 °C for 10 min, and the acetone was discarded. The protein pellet was resuspended in 800 µL of cold methanol (pre-chilled at -20 °C) by probe-sonication on ice and the centrifugation step was repeated to wash the pellet. The pellet then was resuspended in 200 µL of PBS containing 1.2% SDS by probe-sonication at RT, followed by heating at 90 °C for 5 min. Samples were diluted with 1 mL of PBS and incubated with pre-washed streptavidin-agarose beads (100 µL of a 1:1 slurry in PBS, Thermo Scientific 20353) overnight at 4 °C with end-over-end rotation. All samples were transferred into a 96-well fritted filter plate and the beads were washed once with 1 mL of PBS containing 0.2% SDS, followed by three additional washes with 1 mL of PBS, and three final washes with 1 mL of water using a vacuum manifold. The beads were resuspended in 250 µL of 50 mM TEAB, incubated with 10 mM dithiothreitol (DTT) for 30 min at 37 °C on a shaker, followed by incubation with 20 mM IA for 30 min at RT. The beads containing the reduced and alkylated proteins were washed once with 1 mL of PBS, resuspended in 125 µL of 50 mM TEAB containing 1 µg of trypsin, and incubated overnight at 37 °C on a shaker. Following digestion, samples were chilled at 4 °C for 10 min and the peptides were eluted into deep 96-well plate using a vacuum manifold. Beads were washed twice with 125 µL of 50 mM TEAB for further peptide elution. The combined eluent for each sample was subjected to Tandem Mass Tag (TMT) labeling using isobaric 6-plex TMT reagents (0.8 mg per sample, Thermo Scientific 90061) according to the manufacturer's instructions. After confirmation of TMT labeling efficiency >97% by LC-MS analysis, all samples were incubated with hydroxylamine (0.4% final concentration) for 30 min at RT on a shaker. All six samples were pooled to provide a single multiplexed sample, which was concentrated to dryness by vacuum centrifugation, desalted using Peptide Desalting Spin Columns (Thermo

Scientific 89851), and concentrated to dryness again. The final sample was redissolved in 50  $\mu$ L of water containing 3% MeCN and 1% FA prior to LC-MS analysis.

LC-MS data were acquired on an Orbitrap Eclipse Tribrid mass spectrometer (Thermo Scientific) using synchronous precursor selection (SPS) and real-time search (RTS) technologies in tandem with MS3-based TMT reporter ion quantification. The LC setup was identical to that described for the Kinobead assay, and peptides (5- $\mu$ L injection) were separated with a 105-min gradient from 3-26% MeCN in water with 0.1% FA. MS analysis consisted of a 120k-resolution MS1 scan ( $m/z$  400-1600) in the Orbitrap with the AGC target set to 'standard' and MIT of 50 ms. Peptide precursor ions (charge states 2-6) were isolated for data-dependent analysis (DDA) using the quadrupole with an isolation width of 0.7 Th. Ions were subsequently fragmented by collision-induced dissociation (CID) with an energy of 35%, followed by MS2 acquisition in the ion trap using the 'turbo' scan rate, AGC target of 200%, and MIT of 35 ms. During data acquisition, MS2 spectra were subjected to a RTS of the SwissProt canonical human protein database with carbamidomethylation (cysteine) and TMT-6plex (lysine and peptide *N*-termini) set as static modifications, along with oxidation (methionine) as a variable modification. The results of the RTS step were utilized to inform the selection of ten MS2 peptide fragment ions via SPS in the ion trap, which were subsequently fragmented by HCD using an energy of 50%. MS3 spectra of the resulting TMT reporter ions then were acquired in the Orbitrap at 50k resolution (AGC target of 200% and MIT of 86 ms). A cycle time of 3 sec was specified for each complete DDA scan.

Raw LC-MS data files were processed in Proteome Discoverer 2.5 (PD, Thermo Scientific) via a processing workflow consisting of parallel searches of the SwissProt human protein database (version 2022-08-03) using MS Amanda 2.0 and Sequest HT algorithms. Oxidation (methionine) was set as a variable modification, along with carbamidomethylation (cysteine) and TMT-6plex (lysine and peptide *N*-termini) set as static modifications, and up to two missed cleavages were allowed. Identifications were filtered to a FDR of 0.01 (strict) or 0.05 (relaxed) using the Percolator node within PD. TMT reporter ion MS3 intensities were quantified in the consensus workflow using a co-isolation interference threshold of 50%, a reporter ion signal:noise threshold of 10%, and

a minimum threshold of 65% for SPS matches. Data were normalized within PD based on total peptide amount, and Benjamini-Hochberg-adjusted *p*-values were calculated using the ANOVA function within PD.

## References

- (1) Shang, Z.; Khalil, Z.; Li, L.; Salim, A. A.; Quezada, M.; Kalansuriya, P.; Capon, R. J. Roseopurpurins: Chemical Diversity Enhanced by Convergent Biosynthesis and Forward and Reverse Michael Additions. *Org Lett* **2016**, *18* (17), 4340–4343. <https://doi.org/10.1021/acs.orglett.6b02099>.
- (2) Hooda, D. U.; Haruyoshi, S.; Yuki, T.; Michael, W. P.; Nami, Y. Dibenzo-Oxazepine and -Dioxepine Derivatives and Their Use as Anti-Tumor Agents. WO1997047611A1, 1997.
- (3) Sheldrick, G. M. SHELXT– Integrated Space-group and Crystal-structure Determination. *Acta Crystallogr. Sect. A* **2015**, *71* (1), 3–8. <https://doi.org/10.1107/s2053273314026370>.
- (4) Bourhis, L. J.; Dolomanov, O. V.; Gildea, R. J.; Howard, J. A. K.; Puschmann, H. The Anatomy of a Comprehensive Constrained, Restrained Refinement Program for the Modern Computing Environment – Olex2 Dissected. *Acta Crystallogr. Sect. A: Found. Adv.* **2015**, *71* (1), 59–75. <https://doi.org/10.1107/s2053273314022207>.
- (5) Kleemiss, F.; Dolomanov, O. V.; Bodensteiner, M.; Peyerimhoff, N.; Midgley, L.; Bourhis, L. J.; Genoni, A.; Malaspina, L. A.; Jayatilaka, D.; Spencer, J. L.; White, F.; Grundkötter-Stock, B.; Steinhauer, S.; Lentz, D.; Puschmann, H.; Grabowsky, S. Accurate Crystal Structures and Chemical Properties from NoSpherA2. *Chem. Sci.* **2020**, *12* (5), 1675–1692. <https://doi.org/10.1039/d0sc05526c>.
- (6) Dolomanov, O. V.; Bourhis, L. J.; Gildea, R. J.; Howard, J. A. K.; Puschmann, H. OLEX2: A Complete Structure Solution, Refinement and Analysis Program. *J. Appl. Crystallogr.* **2009**, *42* (2), 339–341. <https://doi.org/10.1107/s0021889808042726>.
- (7) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Res.* **2019**, *47* (W1), W81–W87. <https://doi.org/10.1093/nar/gkz310>.
- (8) Buchfink, B.; Reuter, K.; Drost, H.-G. Sensitive Protein Alignments at Tree-of-Life Scale Using DIAMOND. *Nat. Methods* **2021**, *18* (4), 366–368. <https://doi.org/10.1038/s41592-021-01101-x>.

- (9) Katoh, K.; Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30* (4), 772–780. <https://doi.org/10.1093/molbev/mst010>.
- (10) Capella-Gutiérrez, S.; Silla-Martínez, J. M.; Gabaldón, T. TrimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses. *Bioinformatics* **2009**, *25* (15), 1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
- (11) Minh, B. Q.; Schmidt, H. A.; Chernomor, O.; Schrempf, D.; Woodhams, M. D.; Haeseler, A. von; Lanfear, R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **2020**, *37* (5), 1530–1534. <https://doi.org/10.1093/molbev/msaa015>.
- (12) Hadjithomas, M.; Chen, I.-M. A.; Chu, K.; Huang, J.; Ratner, A.; Palaniappan, K.; Andersen, E.; Markowitz, V.; Kyrpides, N. C.; Ivanova, N. N. IMG-ABC: New Features for Bacterial Secondary Metabolism Analysis and Targeted Biosynthetic Gene Cluster Discovery in Thousands of Microbial Genomes. *Nucleic Acids Res.* **2017**, *45* (D1), D560–D565. <https://doi.org/10.1093/nar/gkw1103>.
- (13) Gilchrist, C. L. M.; Chooi, Y.-H. Clinker & Clustermap.js: Automatic Generation of Gene Cluster Comparison Figures. *Bioinformatics* **2021**, *37* (16), 2473–2475. <https://doi.org/10.1093/bioinformatics/btab007>.
- (14) Dongen, S. V. Graph Clustering Via a Discrete Uncoupling Process. *SIAM J. Matrix Anal. Appl.* **2008**, *30* (1), 121–141. <https://doi.org/10.1137/040608635>.
- (15) Vacic, V.; Iakoucheva, L. M.; Radivojac, P. Two Sample Logo: A Graphical Representation of the Differences between Two Sets of Sequence Alignments. *Bioinformatics* **2006**, *22* (12), 1536–1537. <https://doi.org/10.1093/bioinformatics/btl151>.
- (16) Hoang, D. T.; Chernomor, O.; Haeseler, A. von; Minh, B. Q.; Vinh, L. S. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **2018**, *35* (2), 518–522. <https://doi.org/10.1093/molbev/msx281>.
- (17) Waterhouse, R. M.; Seppey, M.; Simão, F. A.; Manni, M.; Ioannidis, P.; Klioutchnikov, G.; Kriventseva, E. V.; Zdobnov, E. M. BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Mol. Biol. Evol.* **2017**, *35* (3), 543–548. <https://doi.org/10.1093/molbev/msx319>.
- (18) Price, M. N.; Dehal, P. S.; Arkin, A. P. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **2010**, *5* (3), e9490. <https://doi.org/10.1371/journal.pone.0009490>.
- (19) Winter, G. Xia2: An Expert System for Macromolecular Crystallography Data Reduction. *J. Appl. Crystallogr.* **2010**, *43* (1), 186–190. <https://doi.org/10.1107/s0021889809045701>.

- (20) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr. Sect. D* **2011**, *67* (4), 235–242. <https://doi.org/10.1107/s0907444910045749>.
- (21) Beilsten-Edmands, J.; Winter, G.; Gildea, R.; Parkhurst, J.; Waterman, D.; Evans, G. Scaling Diffraction Data in the DIALS Software Package: Algorithms and New Approaches for Multi-crystal Scaling. *Acta Crystallogr. Sect. D* **2020**, *76* (4), 385–399. <https://doi.org/10.1107/s2059798320003198>.
- (22) Evans, P. Scaling and Assessment of Data Quality. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2006**, *62* (1), 72–82. <https://doi.org/10.1107/s0907444905036693>.
- (23) Bunkóczi, G.; Echols, N.; McCoy, A. J.; Oeffner, R. D.; Adams, P. D.; Read, R. J. Phaser.MRage: Automated Molecular Replacement. *Acta Crystallogr. Sect. D* **2013**, *69* (11), 2276–2286. <https://doi.org/10.1107/s0907444913022750>.
- (24) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2010**, *66* (4), 486–501. <https://doi.org/10.1107/s0907444910007493>.
- (25) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. Sect. D* **2011**, *67* (4), 355–367. <https://doi.org/10.1107/s0907444911001314>.
- (26) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkóczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix. *Acta Crystallogr. Sect. D* **2019**, *75* (Pt 10), 861–877. <https://doi.org/10.1107/s2059798319011471>.
- (27) Bolger, A. M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30* (15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- (28) Langmead, B.; Salzberg, S. L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9* (4), 357–359. <https://doi.org/10.1038/nmeth.1923>.
- (29) Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29* (1), 15–21. <https://doi.org/10.1093/bioinformatics/bts635>.

(30) Liao, Y.; Smyth, G. K.; Shi, W. FeatureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features. *Bioinformatics* **2014**, *30* (7), 923–930. <https://doi.org/10.1093/bioinformatics/btt656>.

(31) Love, M. I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15* (12), 550. <https://doi.org/10.1186/s13059-014-0550-8>.

(32) Médard, G.; Pahl, F.; Ruprecht, B.; Klaeger, S.; Heinzlmeir, S.; Helm, D.; Qiao, H.; Ku, X.; Wilhelm, M.; Kuehne, T.; Wu, Z.; Dittmann, A.; Hopf, C.; Kramer, K.; Kuster, B. Optimized Chemical Proteomics Assay for Kinase Inhibitor Profiling. *J. Proteome Res.* **2015**, *14* (3), 1574–1586. <https://doi.org/10.1021/pr5012608>.

(33) Metz, K. S.; Deoudes, E. M.; Berginski, M. E.; Jimenez-Ruiz, I.; Aksoy, B. A.; Hammerbacher, J.; Gomez, S. M.; Phanstiel, D. H. Coral: Clear and Customizable Visualization of Human Kinome Data. *Cell Syst.* **2018**, *7* (3), 347-350.e1. <https://doi.org/10.1016/j.cels.2018.07.001>.