Supplementary Information

Ketoreductase domain mutagenesis reprogrammes chain-length control in alternapyrone biosynthesis

Inthira Tapeng,[a] Jaiyfungkhong Phakeovilay[b] and Pakorn Wattana-Amorn*[a,b]

[a]Department of Chemistry and Center of Excellence for Innovation in Chemistry, Faculty of Science, Kasetsart University, Bangkok, 10900, Thailand

[b]Interdisciplinary Program in Genetic Engineering and Bioinformatics, Graduate School, Kasetsart University, Bangkok, 10900, Thailand

*email: fscipwa@ku.ac.th

Table of contents **Page**

Experimental Procedures

1. Plasmid construction

The intronless alternapyrone polyketide synthase gene from *M. theobromae* BCC 4162 (*mtaltA*) was constructed from the intron-containing *mtaltA* by amplification into three fragments, followed by site-directed mutagenesis (SDM) to introduce a substitution at the catalytic residue Y2384 in the ketoreductase (KR) domain. The primer pairs are listed in Table S1. Polymerase chain reactions (PCR) were performed using KOD-Plus-Neo DNA polymerase (TOYOBO, Japan). The thermocycling conditions for amplifying each fragment of intronless *mtaltA* were as follows: initial denaturation at 94 ºC for 2 min, followed by 35 cycles of denaturation at 98 ºC for 10 sec, annealing at 61 ºC (for fragments 1 and 2) and 60 ºC (for fragments 3) for 30 sec, and extension at 68 ºC for 1 min/kb. A final extension was performed at 68 °C for 5 min. The PCR products were then cloned into the *E. coli*-yeast shuttle vector, pEYA2eGFP, and the expression vector pTYGSarg using a previously described method.[1] The resulting plasmids were named pEYA2-MtAltA*eGFP and pTYarg-MtAltA*eGFP, respectively.

For site-directed mutagenesis, pEYA2-MtAltA*eGFP was used as a template. The same thermocycling conditions were used as above, except the reaction was performed for 45 cycles, with an annealing temperature of 66.8 °C for 30 sec. After amplification, PCR products were digested with *Dpn*I (Thermo Scientific, USA). The reaction mixture was incubated at 37 °C for 4 hrs, followed by transformation into *E. coli* TOP10 via electroporation. The resulting plasmid was subsequently introduced into the expression vector pTYGSarg using the Gateway® cloning system with LR Clonase II (Invitrogen, USA). The nucleotide sequences of all resulting expression vectors were verified by Sanger sequencing and Barcode Taq sequencing (Celemics, Inc., South Korea).

2. Protoplast-mediated transformation of *Aspergillus oryzae* NSAR1

The expression vector was introduced into *A. oryzae* NSAR1 through protoplast-mediated transformation.[2] For protoplast preparation, spores of *A. oryzae* NSAR1 were cultured on malt extract agar (MEA), containing 3% (w/w) malt extract broth, 1% (w/w) bacteriological peptone, 0.04% (w/w) adenine, and 1.5% (w/w) agar, at 28 °C for 7 days. The spores were then inoculated into 50 mL of liquid GN medium (2% (w/w) nutrient broth, 1% (w/w) glucose) and cultured at 28 °C, 200 rpm for 16-18 hrs. Subsequently, the culture was centrifuged at 8000 x g for 5 min to collect the cells. Ten millilitres of sterile protoplasting solution, containing 10 mg/mL of filter-sterilised Vinotaste lysing enzyme in 0.8 M NaCl, were added to the cell pellet. The mixture was shaken at 30 °C and 70 rpm for 2 hrs. The solution was then filtered through sterile Miracloth to isolate protoplasts. The filtrate was centrifuged at 3000 x g for 5 min. The supernatant was discarded, leaving only the protoplasts. The protoplasts were chilled on ice. Then 100 µL of solution I (0.8 M NaCl, 10 mM $CaCl_2$, and 50 mM Tris-HCl, pH 7.5) was added to the protoplasts. Ten microliters of the expression vectors (700-1000 ng/µL) were added to each reaction and the mixtures were chilled on ice for 2 min. For the control, empty

expression vectors were added to the protoplasts. After chilling, 1 mL of solution II (60% (w/w) PEG 3350, 0.8 M NaCl, 10 mM $CaCl_2$, and 50 mM Tris-HCl, pH 7.5) was added to the mixture, and the entire mixture was incubated at room temperature for 20 min. Five milliliters of warm (approximately 50°C) molten Czapek-Dox/Sorbitol (CZD/S) top-up medium lacking arginine (3.5% (w/w) Czapek-Dox,18.22% (w/) sorbitol, 0.05% (w/w) adenine, 0.1% (w/w) $(NH_4)_2SO_4$, 0.15% (w/w) methionine and 0.8% (w/w) agar) was then added to the transformation mixture and poured onto CZD/S agar plates lacking arginine (containing 1.5% (w/w) agar). The plates were incubated at 28 °C for 7 days. Transformants were then serially subcultured on CZD/S agar plates lacking arginine agar until genetically pure.

3. Fermentation and extraction of transformant cultures

The spore suspensions of *A. oryzae* NSAR1 transformants were cultured in 50 mL of CMP medium (3.5% (w/w) Czapek-Dox, 2% (w/w) maltose, and 1% (w/w) bacteriological peptone) for 7 days. The cultures were filtered under vacuum to separate the mycelia from the liquid phase. The latter was extracted with a solvent mixture of ethyl acetate, methanol, and acetic acid (89.5:10.0:0.5) as previously reported.[3] The mycelia were rinsed with water to remove residual culture broth before soaking in acetone. Crude extracts from both the culture broth and the mycelia were then evaporated to dryness prior to metabolite screening.

4. Fluorescence analysis

The expression of the KR mutant was analysed by fluorescence microscopy using a Leica Thunder Imager 3D microscope system. The excitation wavelength for eGFP was 450-490 mm.

5. Metabolite screening by High-Performance Liquid Chromatography (HPLC) and Liquid Chromatography – Mass Spectrometer (LC-MS)

Crude extracts were dissolved in HPLC-grade methanol to a final concentration of 50 mg/mL and analysed using an HPLC-DAD system (Agilent Technologies 1200) equipped with a C-18 Kaseisorb LC ODS 2000 column (4.6 mm × 150 mm) (Tokyo Chemical Industry). Elution was performed using an acetonitrile-water gradient (containing 0.05% formic acid) as follows: 5-25% over 8 min, 25-27% over 22 min and 27-95% over 4 min, at a flow rate of 1 mL/min. UV absorbance was measured at 215 and 280 nm. Metabolite analysis was also performed by LC-MS using a 6420 Triple Quadrupole LC/MS system (Agilent 1260 Infinity Series) with the same conditions as described for HPLC-DAD.

6. Isolation and structure elucidation of compounds (1) and (2)

The crude extract containing compounds **1** and **2** was purified using a semi-preparative C18 Kaseisorb LC ODS 2000 column (10.0 mm × 250 mm) (Tokyo Chemical Industry), connected to an HPLC-DAD system (Agilent Technologies 1200). An acetonitrile-water gradient (with 0.05% formic acid) was used for elution as follows: 5-25% over 18 min, 25-27% over 42 min and 27-30% over 12 min, at a flow rate of 3 mL/min. UV absorbance was measured at 215 and 280 nm, yielding 4-hydroxy-3,5,6-trimethyl-2-pyrone (**1**) and 6-ethyl-4-hydroxy-3,5-dimethyl-2-pyrone (**2**). The structures were elucidated by NMR and HRMS. All NMR data (^{1}H, ^{13}C, COSY, HSQC and HMBC) were recorded on a Bruker AVANCE NEO 500 MHz NMR spectrometer equipped with a cryoprobe, using CD_3OD as the solvent.

7. Feeding experiment

Equal numbers of spores from suspensions of *A. oryzae* NSAR1 carrying KR-mutated *mtaltA* were cultured in 25 mL of CMP medium at pH 7. Fatty acids with chain lengths from C4 to C8 (5 mM) were used as precursors and added to the cultures after 72 hrs of initial growth. Each feeding experiment was performed in triplicate. Following precursor addition, the cultures were further incubated for a total of 7 days. The culture broth was separated from the mycelia by vacuum filtration, followed by extraction as described above for *A. oryzae* NSAR1 transformant cultures prior to analysis by UHPLC-QTOF-MS. Feeding cultures of *A. oryzae* NSAR1 transformants expressing *mtaltA* with fatty acids were also performed as described for KR-mutated *mtaltA*.

8. Metabolite analysis by Ultra-High-Performance Liquid Chromatography coupled with Quadrupole Time-of-Flight Mass Spectrometry (UHPLC-QTOF-MS)

High-resolution mass spectrometry (HRMS) was performed using a UHPLC system (ExionLC™AD) coupled to a quadrupole time-of-flight mass spectrometer (X500R QTOF system), equipped with an Avantor ACE SuperC18™ column (2.1 mm × 100 mm). UHPLC analysis was performed with an injection volume of 1 µL of crude extracts at 25 mg/mL. Gradient elution was applied using water containing 0.05% (v/v) formic acid (A) and methanol containing 0.05% (v/v) formic acid (B) with a flow rate of 0.4 mL/min. The gradient was as follows: 5% B for 2 min, increased linearly to 95% B over 5 min, and held at 95% B for 2 min. For MS/MS data collection, SWATH acquisition was employed, and the parameters for TOF-MS and MS/MS were as described in previous work.[1] All MS data were analysed using the Explorer and Analytics workspaces on the SCIEX OS software.

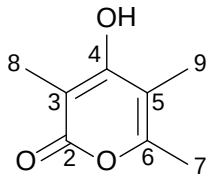
9. Protein structure prediction and sequence analysis

The structure of the MtAltA KR domain was predicted using AlphaFold 3.[4] The resulting model was visualised in UCSF ChimeraX.[5] Structure-based sequence alignment was performed using Expresso[6] and rendered with ESPript 3.0.[7]

Table S1 Primers used in this study.

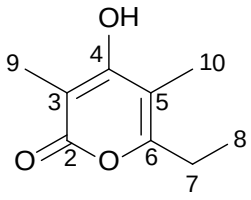
Primer	**Sequence (5′ to 3′)**
For constructing intronless *mtaltA* (Underlined nucleotides are regions of overlap with pEYA2eGFP)	
Fragment 1	
mtaltA-F1-fwd	TAATGCCAACTTTGTACAAAAAAGCAGGCTCCATGGCCCAACATACAGCCCG
mtaltA-F1-rev	CTGTCCGGGTGCCAGCAATCCGATTCAAGCCCAGTCCTGC
Fragment 2	
mtaltA-F2-fwd	GCAGGACTGGGCTTGAATCGGATTGCTGGCACCCGGACAG
mtaltA-F2-rev	GAAGCGGGCCCATGATGAAGCCGCTGAAGATACGCTTGATG
Fragment 3	
mtaltA-F3-fwd	CATCAAGCGTATCTTCAGCGGCTTCATCATGGGCCCGCTTC
mtaltA-F3-rev	GGTGAACAGCTCCTCGCCCTTGCTCACCATTTCGCCAGCCTTGGCAGCCT
For site-directed mutagenesis (Mutated nucleotides are underlined)	
Y2384F-fwd	GACAGGGCAACTTCGGCATCGGCAACATC
Y2384F-rev	GATGTTGCCGATGCCGAAGTTGCCCTGTC

Table S2 ^{1}H NMR (500 MHz) and ^{13}C NMR (125 MHz) data of 4-hydroxy-3,5,6-trimethyl-2-pyrone (**1**) in CD_3OD.



No.	δ_H (multi)	δ_C
2		168.7
3		98.5
4		168.9
5		109.6
6		156.7
7	2.22 (s)	17.1
8	1.90 (s)	8.8
9	1.94 (s)	10.2

Table S3 ^{1}H NMR (500 MHz) and ^{13}C NMR (125 MHz) data of 6-ethyl-4-hydroxy-3,5-dimethyl-2-pyrone (**2**) in CD_3OD.



No.	δ_H (multi, *J* in Hz)	δ_C
2		168.9
3		98.4
4		169.7
5		109.2
6		161.0
7	2.57 (q, *J* = 7.6 Hz)	25.1
8	1.19 (t, *J* = 7.6 Hz)	12.0
9	1.90 (s)	9.9
10	1.95 (s)	8.9

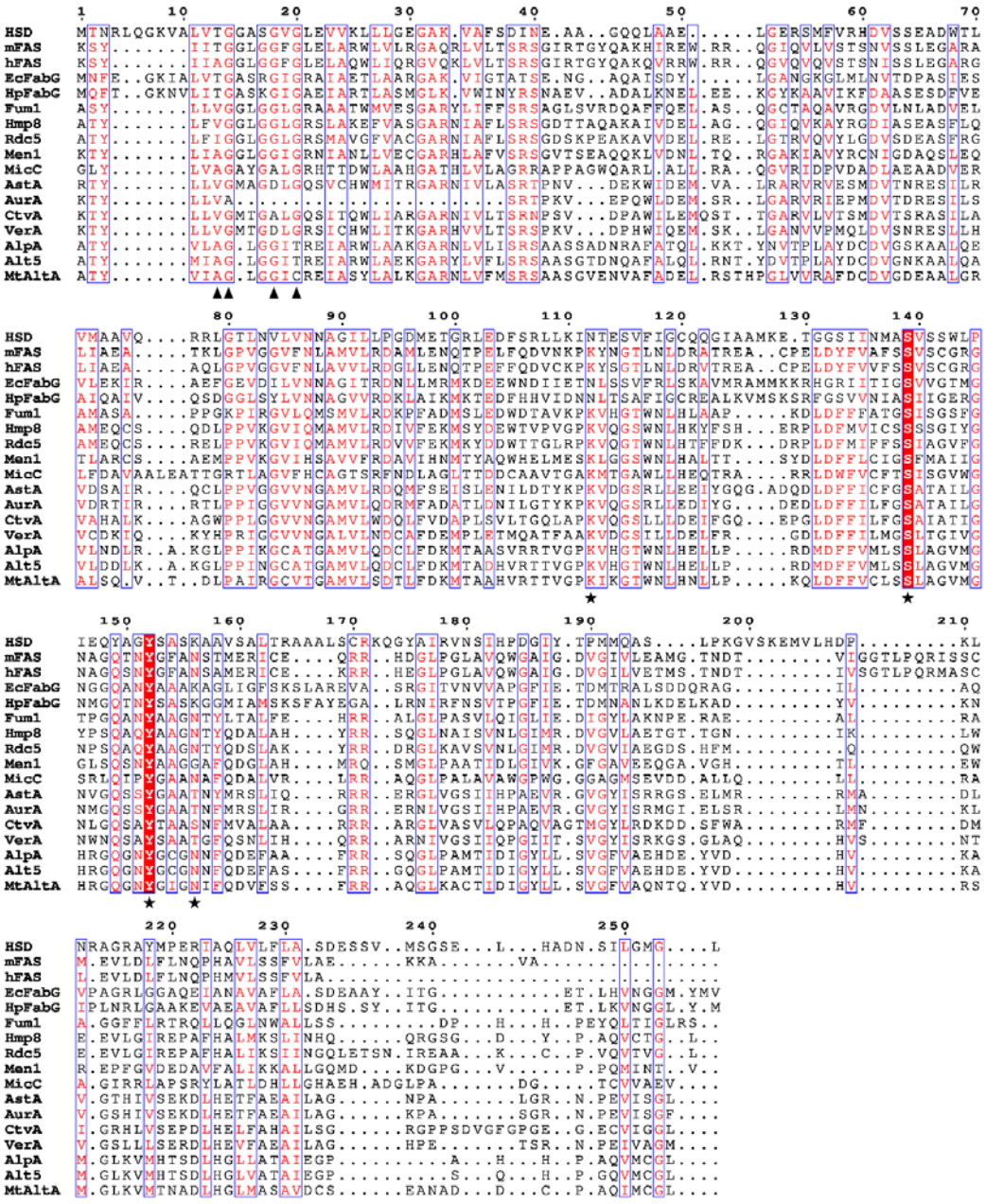


Fig. S1 Structure-based sequence alignment of KRs of type I (mFAS (NP_032014), hFAS (NP_004095)) and type II (EcFabG (WP_001499617), HpFabG (WP_078215644)) fatty acid synthases (FASs), HR-PKSs (Fum1 (AAD43562), Hmp8 (B3FWT3), Rdc5 (B3FWU0), Men1(A0A6F9DXA0)), and a hybrid partially reducing PKS-NRPS (MicC(CAD15508)), including HR-PKSs involved in the biosynthesis of methylated polyene α-pyrones (AstA(UOK93122), AurA(A0A0M4L8I7), CtvA(Q0C9L7), VerA (A0A1V6NYI6), AlpA(XP_001796187), Alt5(Q5KTM9), MtAltA (URN71259)), compared with bacterial 3β/17β-hydroxysteroid dehydrogenase, HSD (WP_003080542). The NADPH-binding region (TGXXXGXG) and the catalytic tetrad are marked with black triangles and stars, respectively. GenPept accession numbers are given in brackets. The amino acid positions of HSD are shown at the top of the alignment.

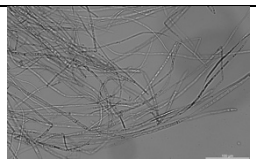
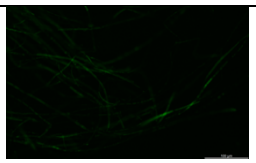
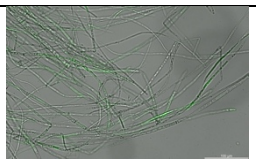
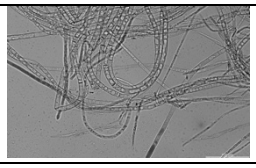
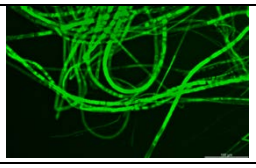
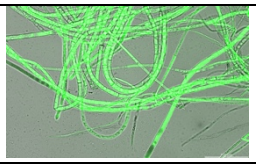
Transformant	Bright field	eGFP	Merged
Negative control			
KR-mutated *mtaltA*			

Fig. S2 Fluorescence imaging of *A. oryzae* NSAR1 transformants carrying negative control (empty expression vector) and KR-mutated *mtaltA.*

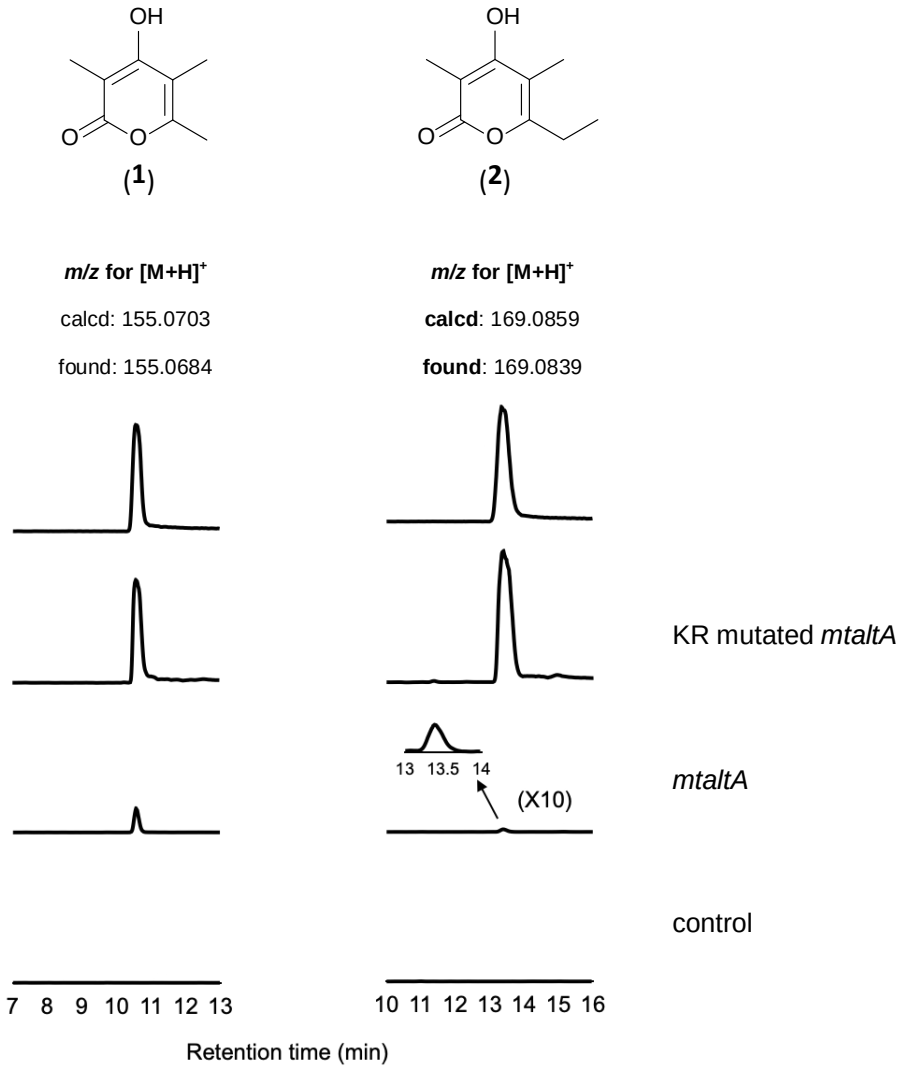


Fig. S3 Extracted ion chromatograms of crude extracts from the culture broth of *A. oryzae* NSAR1 transformants carrying *mtaltA* or KR-mutated *mtaltA*, compared with the control transformant. Structures of compounds **1** and **2** are shown at the top of chromatograms, along with calculated and observed *m*/*z* for $[M+H]^+$.

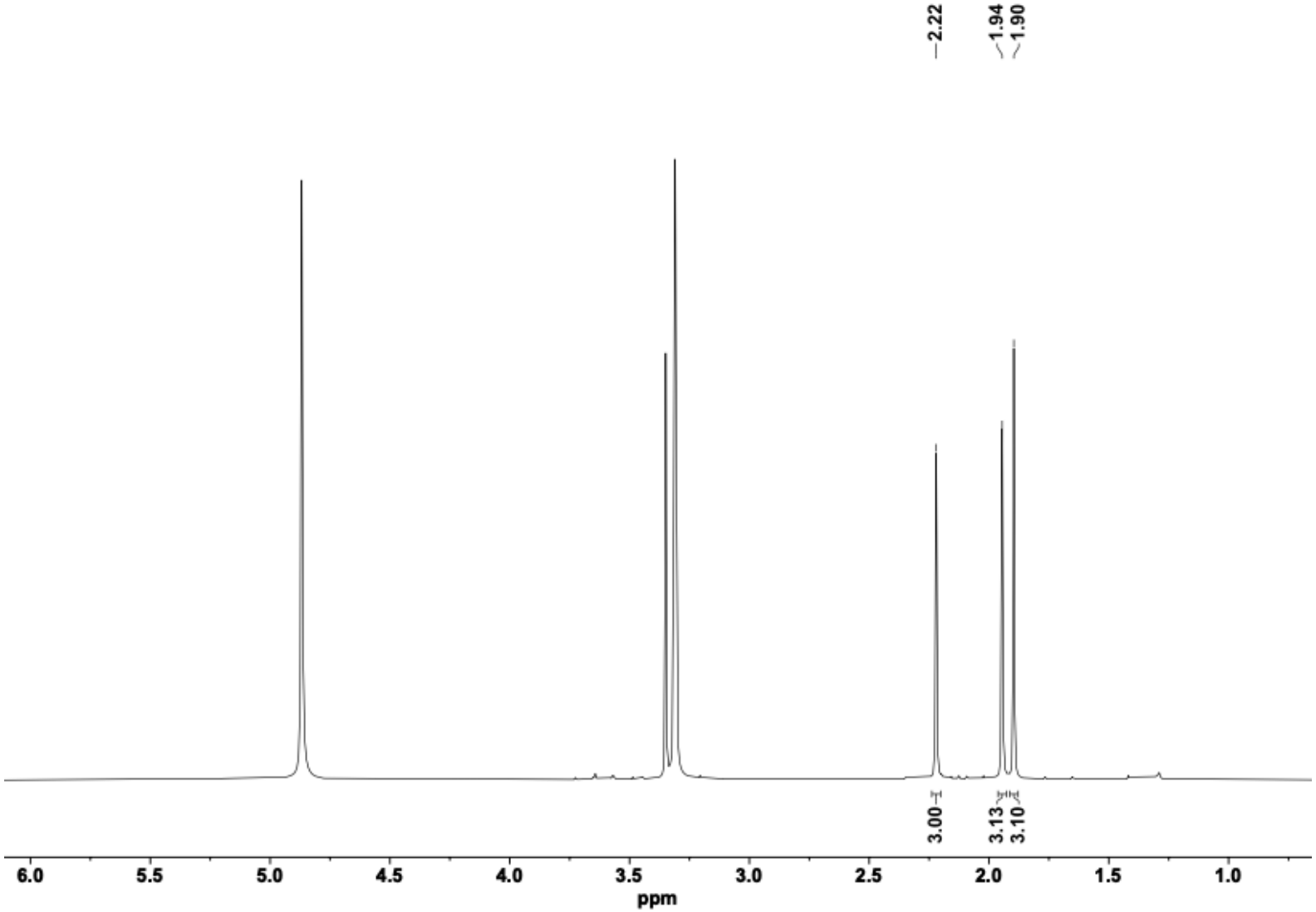


Fig. S4 ^{1}H NMR (500 MHz) with water suppression of compound **1** in CD_3OD. Residual solvent and water signals appear at 3.31, 3.35 and 4.87 ppm, respectively.

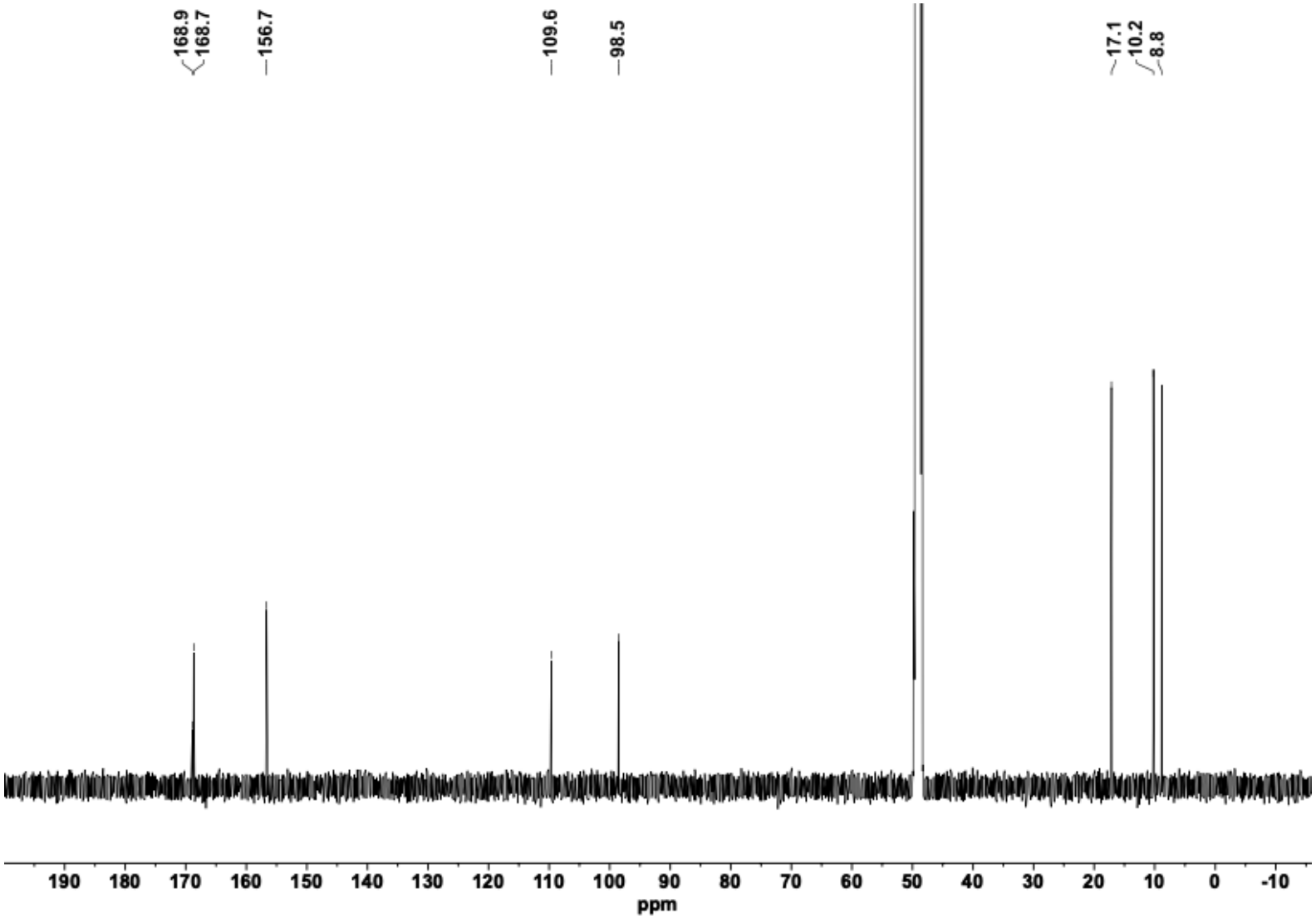


Fig. S5 ^{13}C NMR (125 MHz) of compound **1** in CD_3OD.

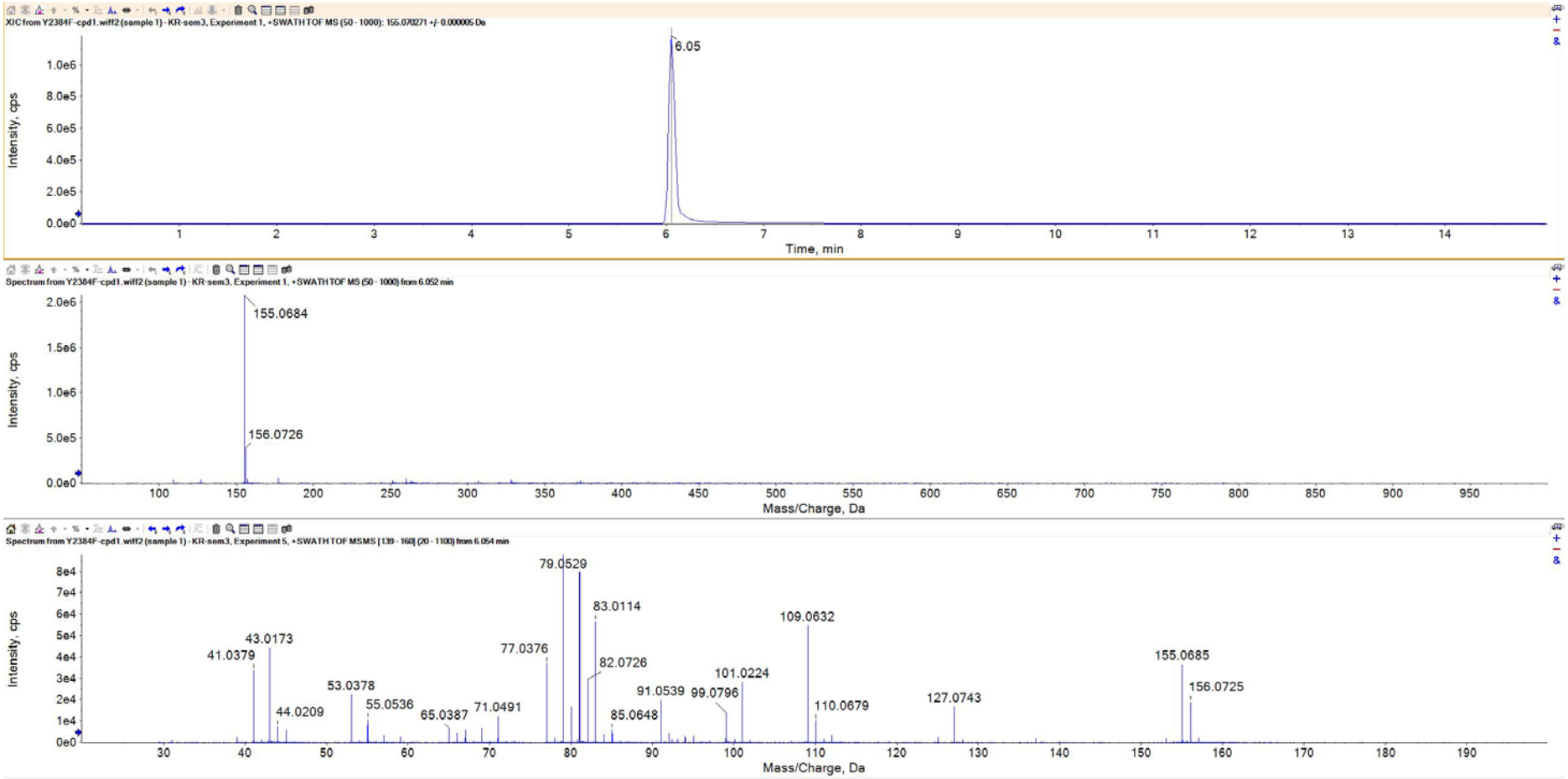


Fig. S6 Extracted ion chromatogram (top), MS (middle), and MS/MS (bottom) spectra of compound **1**. HRMS (ESI-TOF) (*m*/*z*): $[M+H]^+$ calculated for $C_8H_{11}O_3$ 155.0703; found 155.0684.

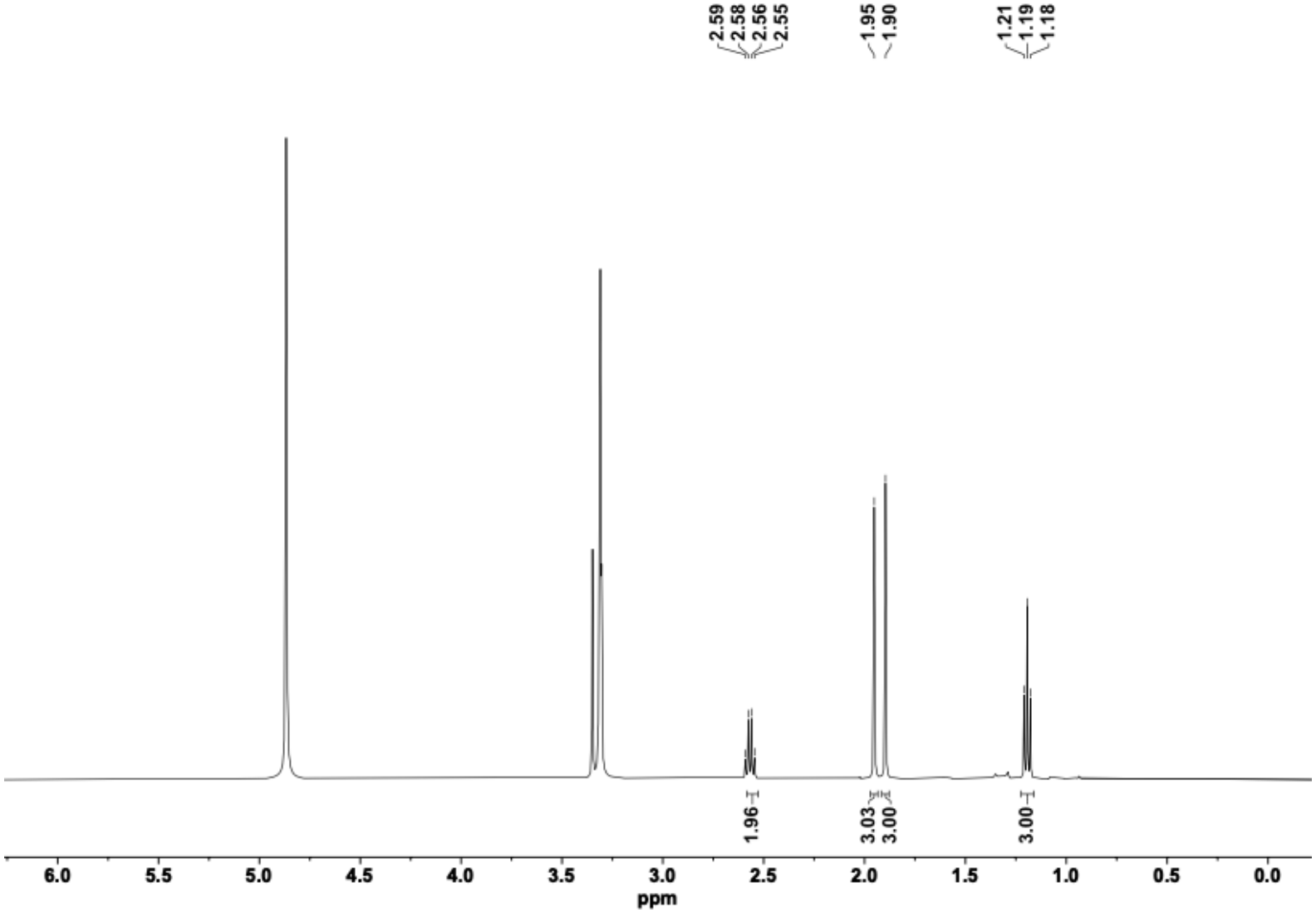


Fig. S7 ^{1}H NMR (500 MHz) with water suppression of compound **2** in CD_3OD. Residual solvent and water signals appear at 3.31, 3.35 and 4.87 ppm, respectively.

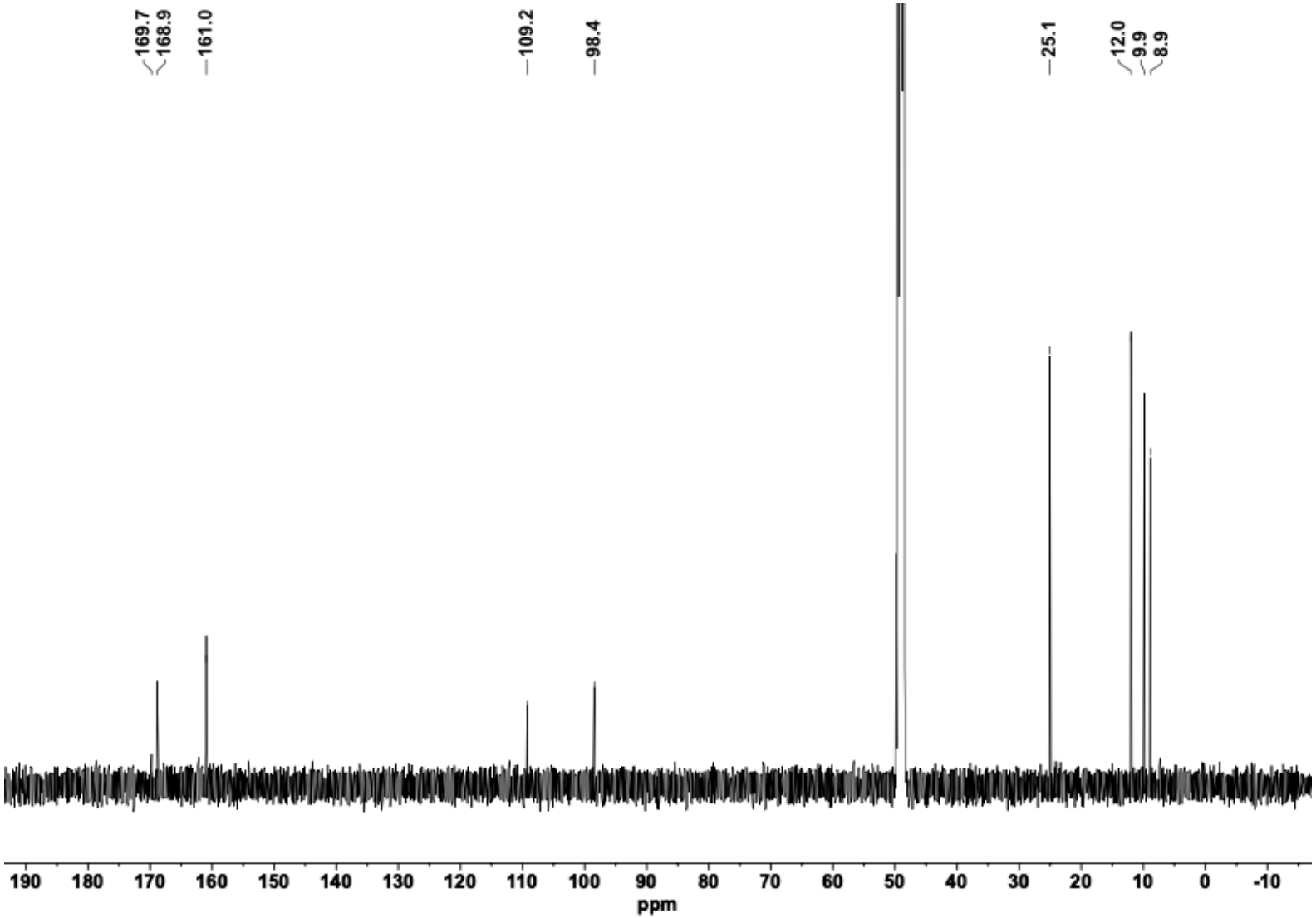


Fig. S8 ^{13}C NMR (125 MHz) of compound **2** in CD_3OD.

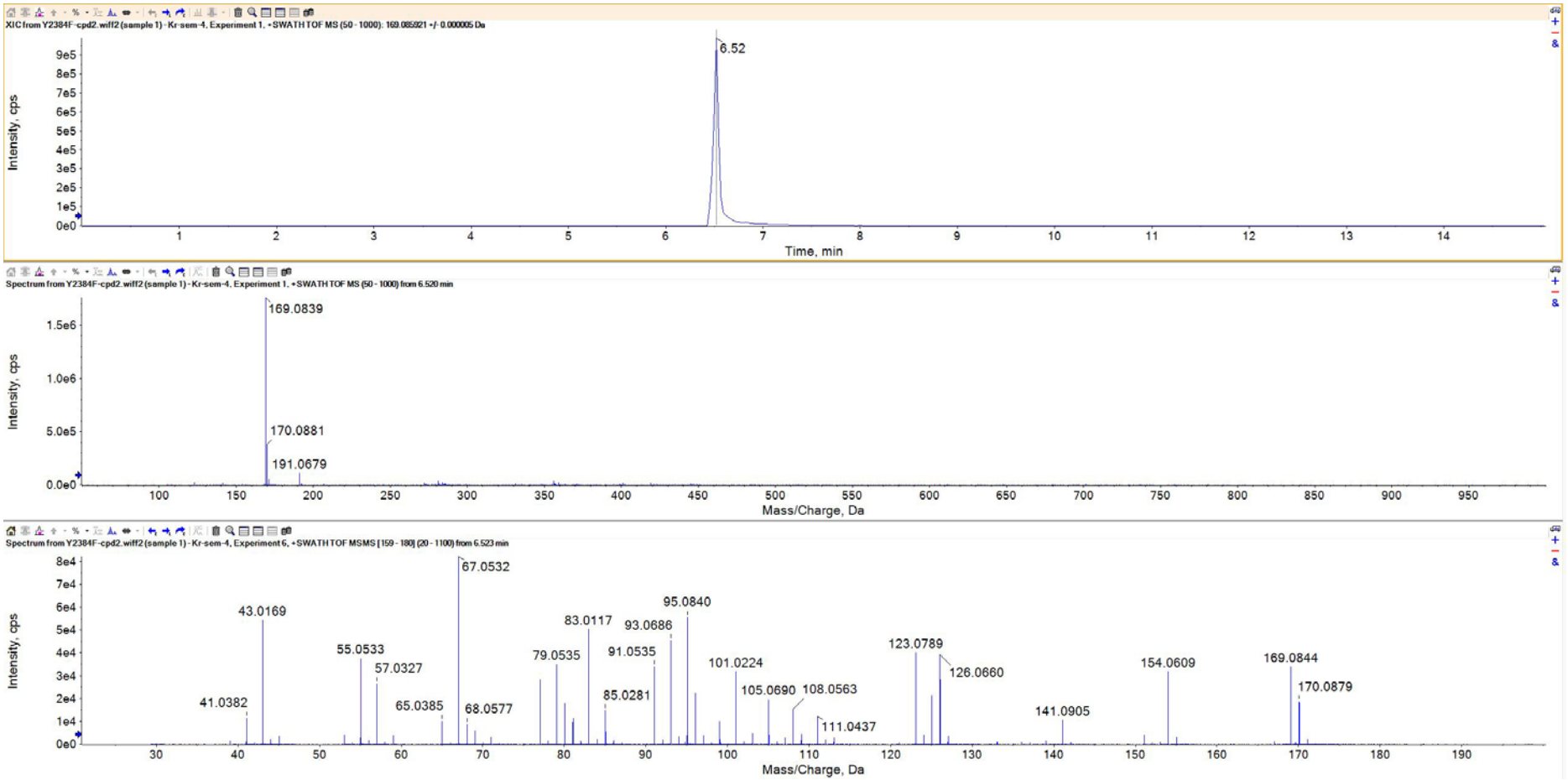


Fig. S9 Extracted ion chromatogram (top), MS (middle), and MS/MS (bottom) spectra of compound **2**. HRMS (ESI-TOF) (*m*/*z*): $[M+H]^+$ calculated for $C_9H_{13}O_3$ 169.0859; found 169.0839, $[M+Na]^+$ calculated for $C_9H_{12}O_3Na$ 191.0679; found 191.0679.

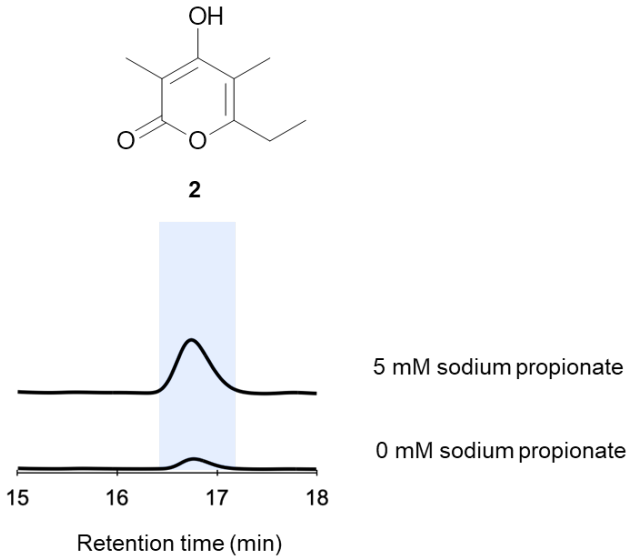


Fig. S10 HPLC chromatograms of crude extracts from the culture broth of *A. oryzae* NSAR1 transformants carrying KR-mutated *mtaltA* supplemented with 5 mM of sodium propionate.

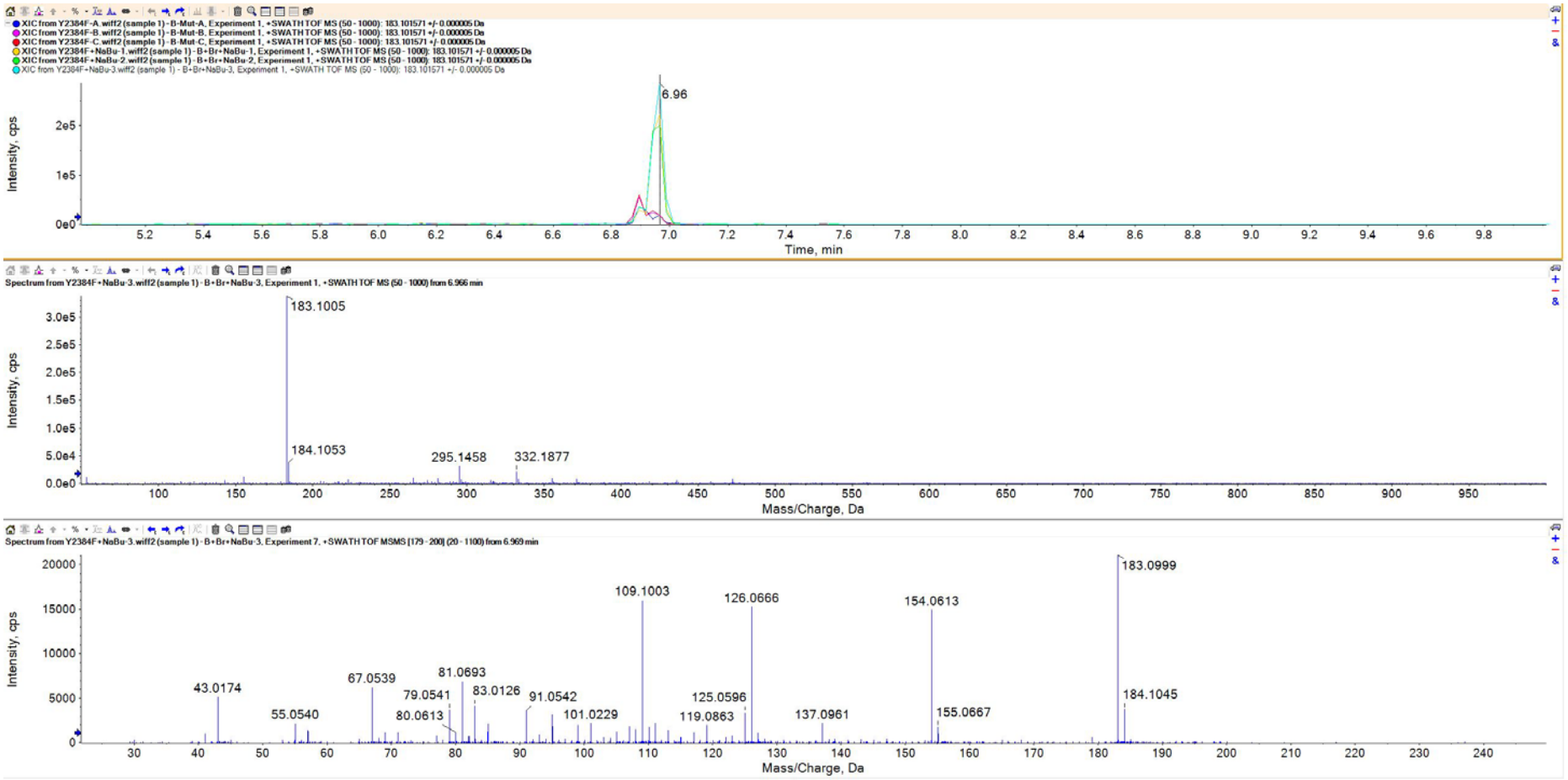


Fig. S11 Extracted ion chromatograms (top) of compound **3** from crude culture broth extracts of transformants carrying KR-mutated *mtaltA* fed with sodium butanoate (coloured in yellow, green and cyan), compared with those without feeding (coloured in blue, pink and red), along with MS (middle), and MS/MS (bottom) spectra of compound **3** from one of the sodium butanoate-fed transformant cultures. HRMS (ESI-TOF) (*m*/*z*): $[M+H]^+$ calculated for $C_{10}H_{15}O_3$ 183.1016; found 183.1005.

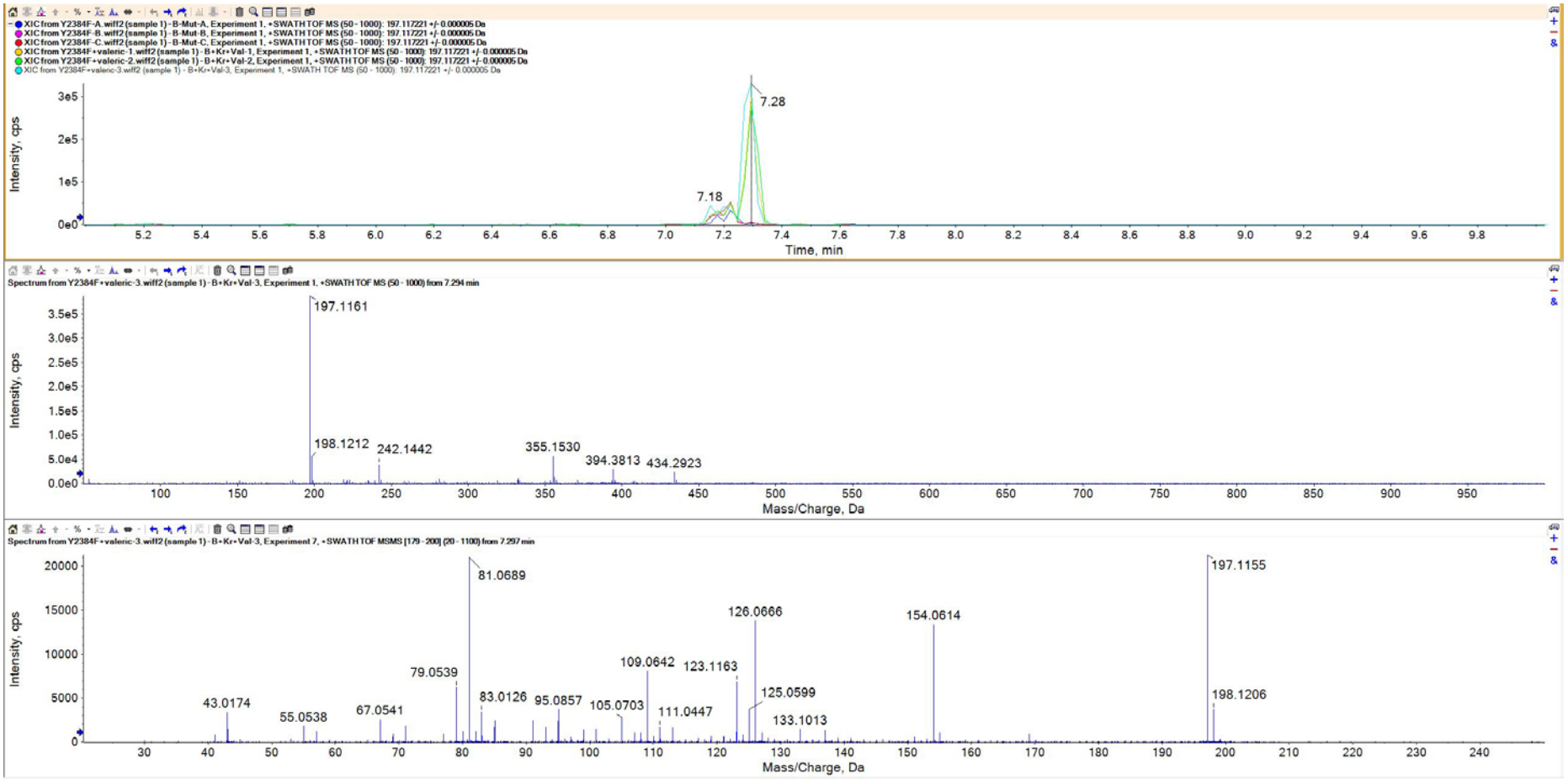


Fig. S12 Extracted ion chromatograms (top) of compound **4** from crude culture broth extracts of transformants carrying KR-mutated *mtaltA* fed with pentanoic acid (coloured in yellow, green and cyan), compared with those without feeding (coloured in blue, pink and red), along with MS (middle), and MS/MS (bottom) spectra of compound **4** from one of the pentanoic acid-fed transformant cultures. HRMS (ESI-TOF) (*m/z*): $[M+H]^+$ calculated for $C_{11}H_{17}O_3$ 197.1172; found 197.1161.

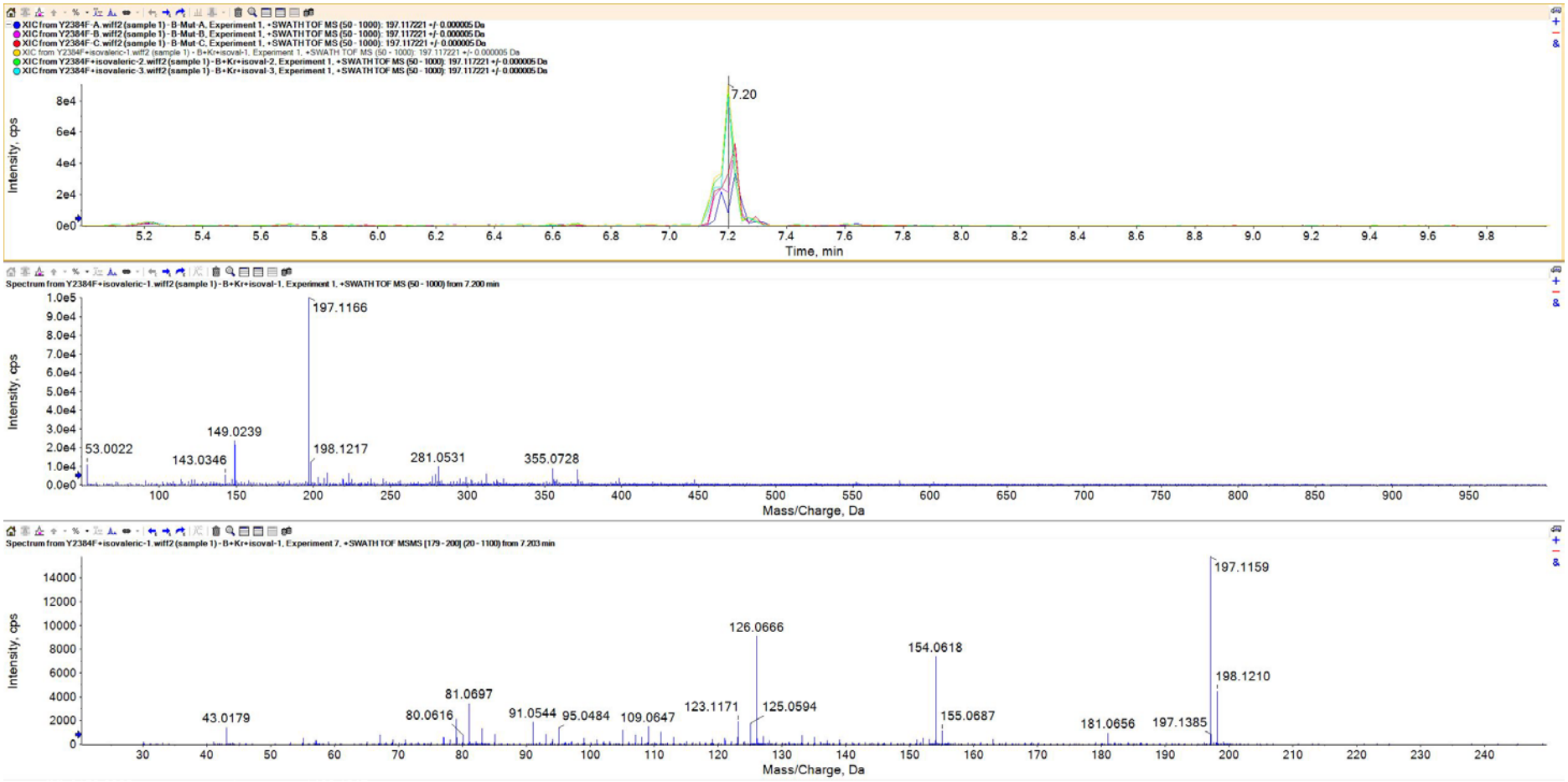


Fig. S13 Extracted ion chromatograms (top) of compound **5** from crude culture broth extracts of transformants carrying KR-mutated *mtaltA* fed with 3-methylbutanoic acid (coloured in yellow, green and cyan), compared with those without feeding (coloured in blue, pink and red), along with MS (middle), and MS/MS (bottom) spectra of compound **5** from one of the 3-methylbutanoic acid-fed transformant cultures. HRMS (ESI-TOF) (*m/z*): $[M+H]^+$ calculated for $C_{11}H_{17}O_3$ 197.1172; found 197.1166.

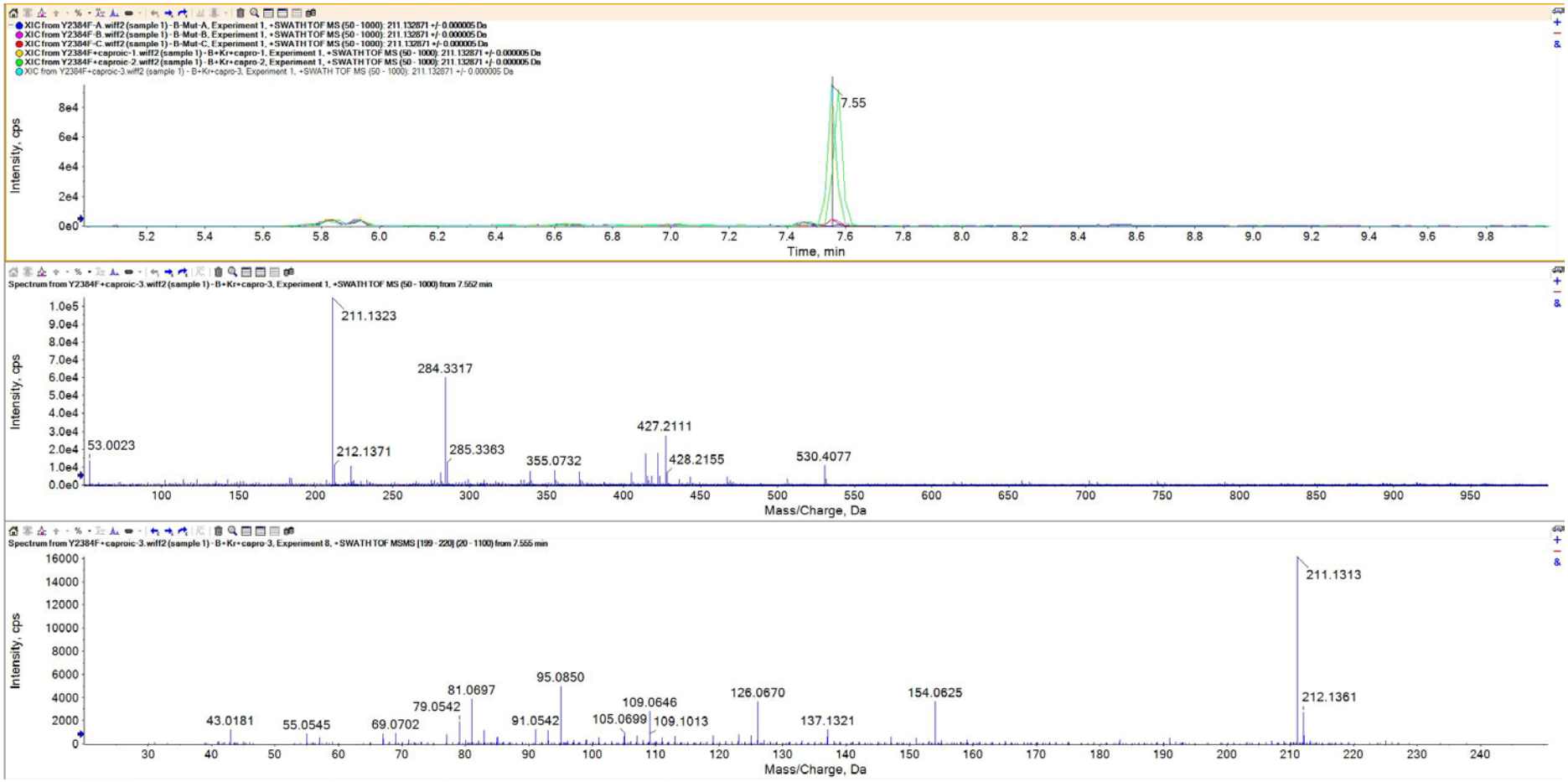


Fig. S14 Extracted ion chromatograms (top) of compound **6** from crude culture broth extracts of transformants carrying KR-mutated *mtaltA* fed with hexanoic acid (coloured in yellow, green and cyan), compared with those without feeding (coloured in blue, pink and red), along with MS (middle), and MS/MS (bottom) spectra of compound **6** from one of the hexanoic acid-fed transformant cultures. HRMS (ESI-TOF) (*m/z*): $[M+H]^+$ calculated for $C_{12}H_{19}O_3$ 211.1329; found 211.1323.

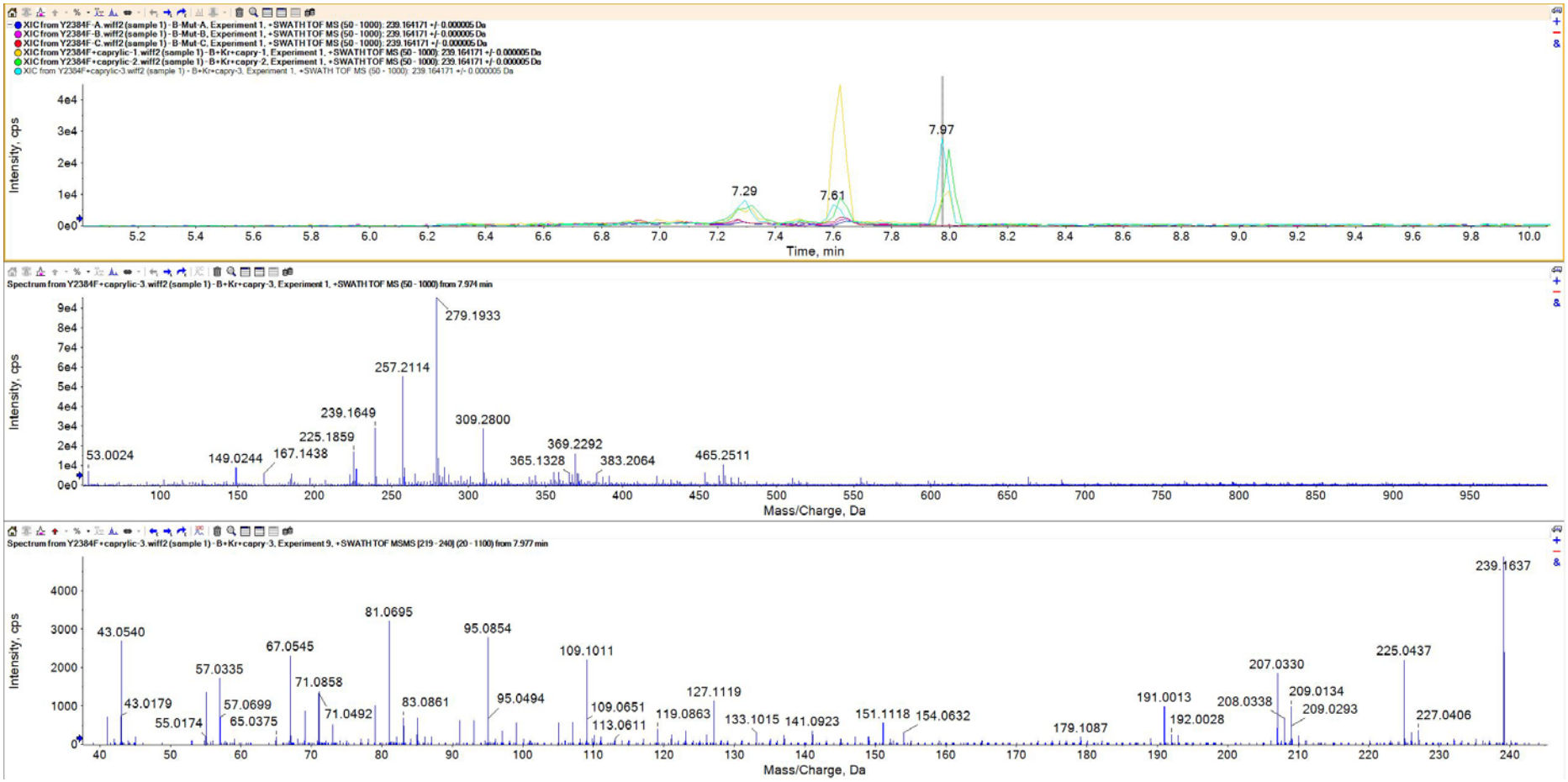


Fig. S15 Extracted ion chromatograms (top) of compound **7** from crude culture broth extracts of transformants carrying KR-mutated *mtaltA* fed with octanoic acid (coloured in yellow, green and cyan), compared with those without feeding (coloured in blue, pink and red), along with MS (middle), and MS/MS (bottom) spectra of compound **7** from one of the octanoic acid-fed transformant cultures. HRMS (ESI-TOF) (*m/z*): $[M+H]^+$ calculated for $C_{14}H_{23}O_3$ 239.1642; found 239.1649.

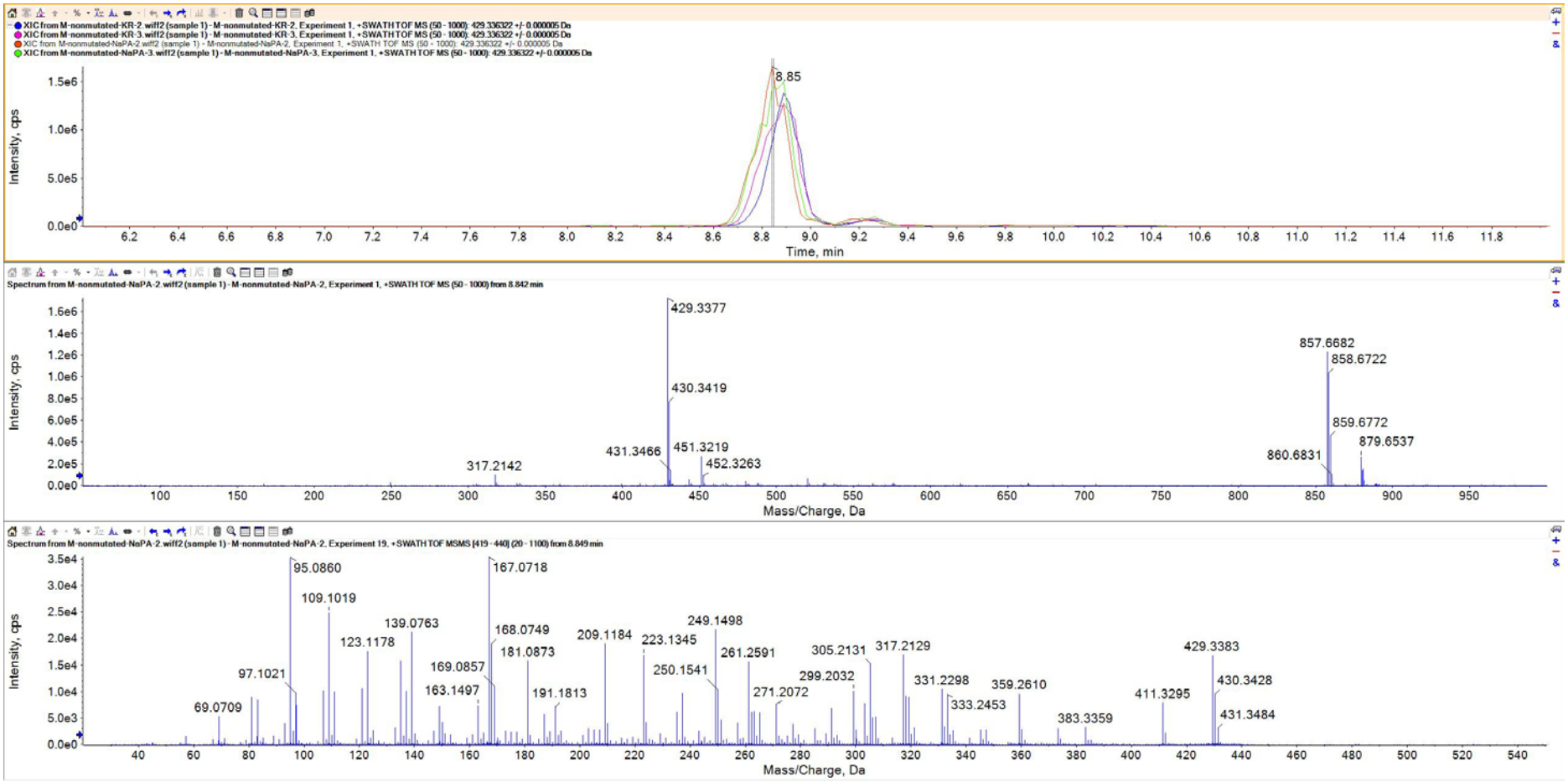


Fig. S16 Extracted ion chromatograms (top) of alternapyrone from crude mycelial extracts of transformants carrying *mtaltA* fed with sodium propionate (coloured in red and green), compared with those without feeding (coloured in blue and pink), along with MS (middle), and MS/MS (bottom) spectra of alternapyrone from one of sodium propionate-fed transformant cultures. HRMS (ESI-TOF) (m/z): $[M+H]^+$ calculated for $C_{28}H_{45}O_3$ 429.3363; found 429.3377, $[M+Na]^+$ calculated for $C_{28}H_{44}O_3Na$ 451.3183; found 451.3219, $[2M+H]^+$ calculated for $C_{56}H_{89}O_6$ 857.6654; found 857.6682, $[2M+Na]^+$ calculated for $C_{56}H_{88}O_6Na$ 879.6473; found 879.6537.

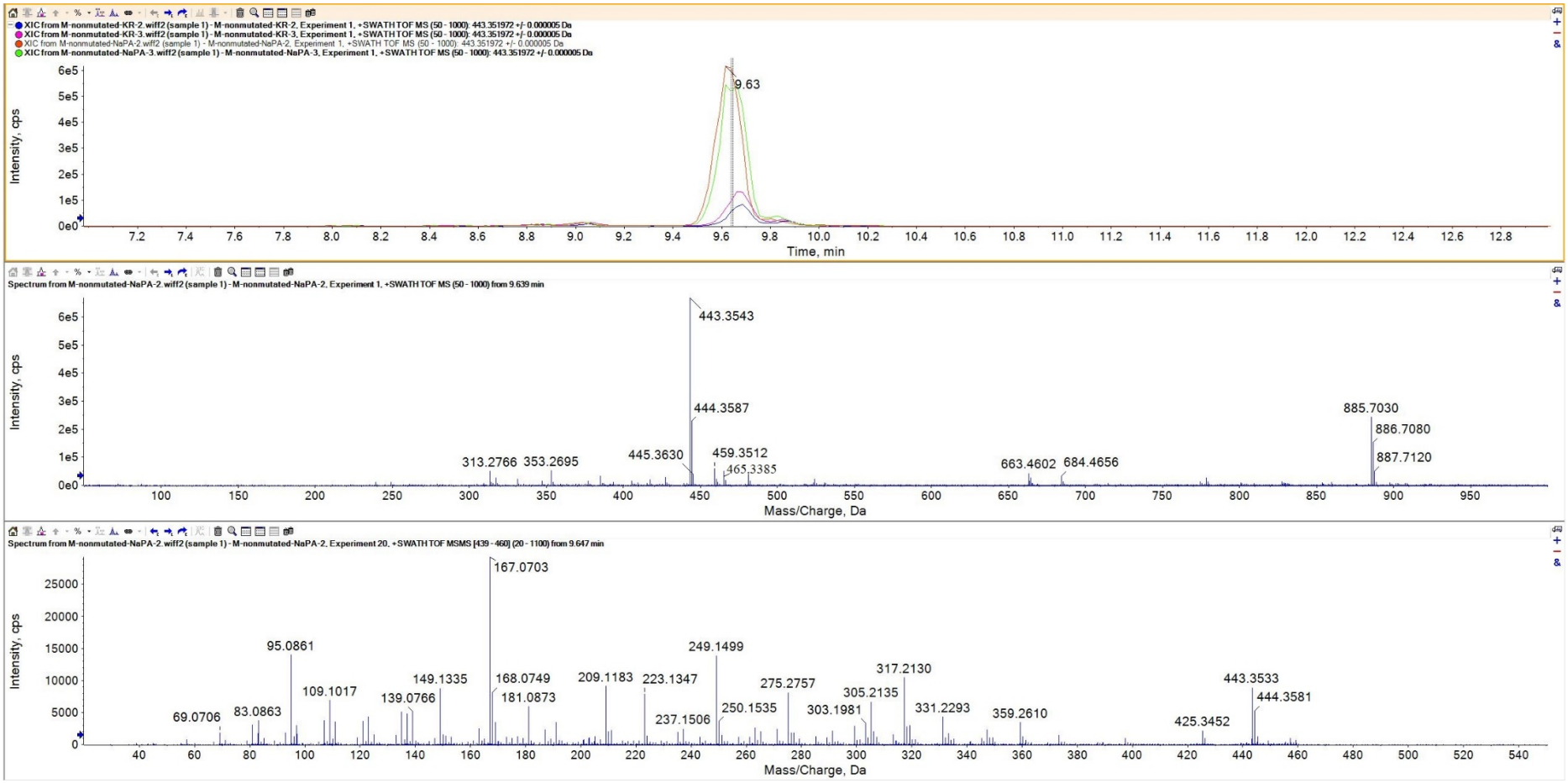


Fig. S17 Extracted ion chromatograms (top) of a propionate-derived alternapyrone derivative from crude mycelial extracts of transformants carrying *mtaltA* fed with sodium propionate (coloured in red and green), compared with those without feeding (coloured in blue and pink), along with MS (middle), and MS/MS (bottom) spectra of the propionate-derived alternapyrone derivative from one of sodium propionate-fed transformant cultures. HRMS (ESI-TOF) (m/z): $[M+H]^+$ calculated for $C_{29}H_{47}O_3$ 443.3520; found 443.3543, $[M+Na]^+$ calculated for $C_{29}H_{46}O_3Na$ 465.3339; found 465.3385, $[2M+H]^+$ calculated for $C_{58}H_{93}O_6$ 885.6967; found 885.7030.

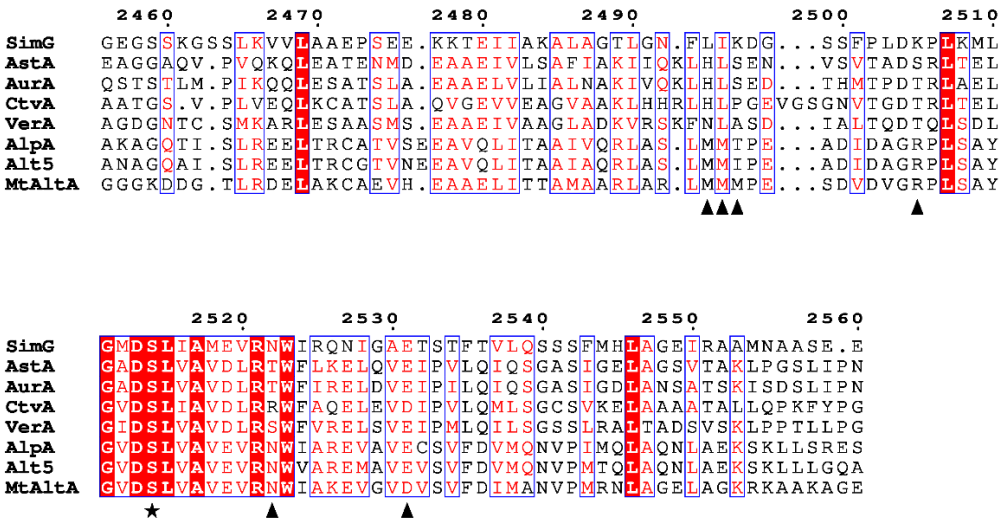


Fig. S18 Structure-based sequence alignment of ACP domains of HR-PKSs involved in methylated polyene α-pyrone biosynthetic pathways. Residues contributing significantly to AT-domain binding in the SimG ACP domain are indicated by black triangles. The phosphopantetheinyl attachment-site serine is marked with a star. The amino acid positions of SimG ACP domain are shown at the top of the alignment.

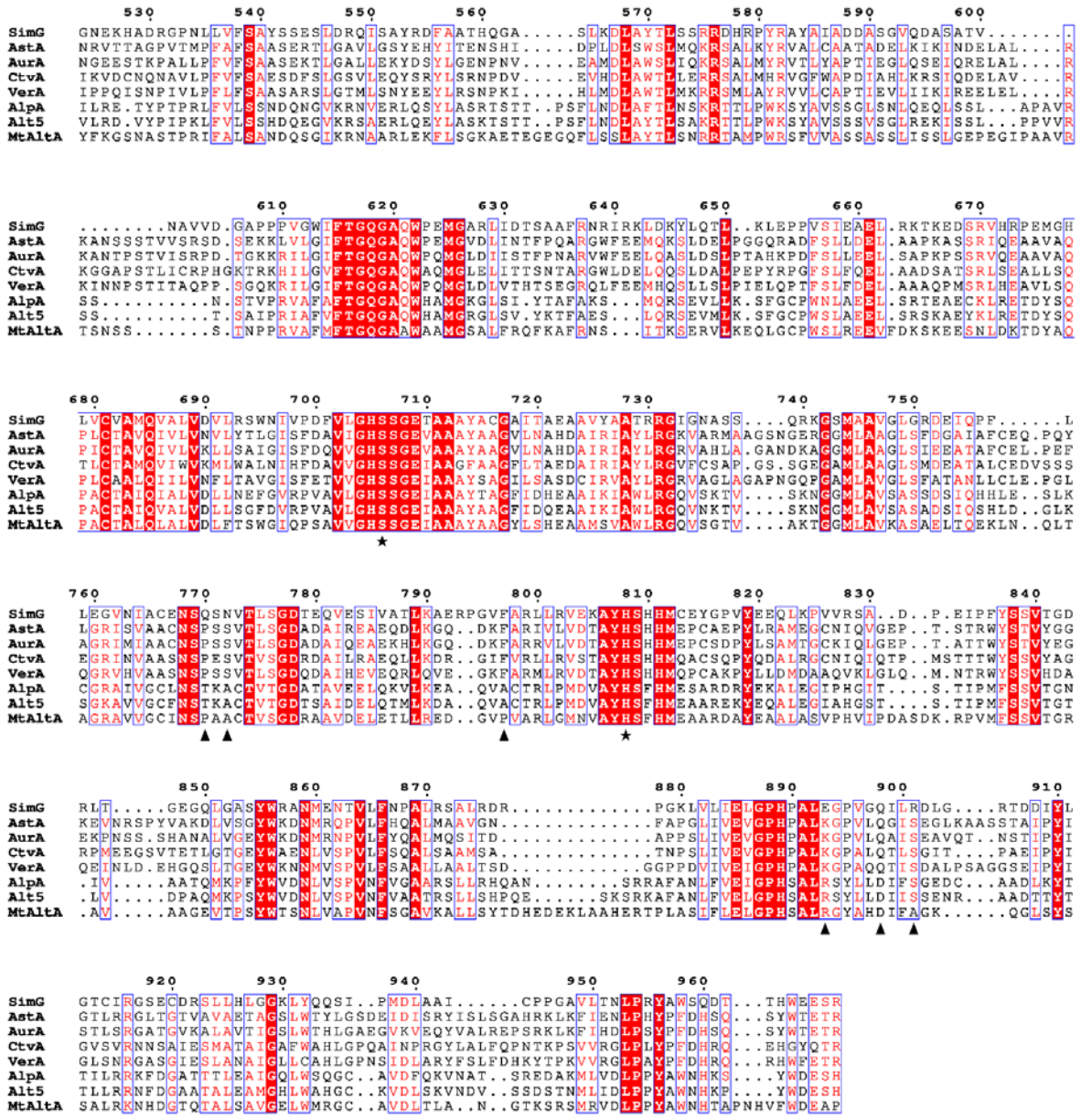


Fig. S19 Structure-based sequence alignment of AT domains of HR-PKSs associated with methylated polyene α-pyrone biosynthesis. Residues of the SimG AT domain surrounding the phosphopantetheinyl group of the ACP domain and in the catalytic region that contribute to ACP-domain binding are indicated by black triangles. The catalytic dyad is marked with black stars. The amino acid positions of SimG AT domain are shown at the top of the alignment.

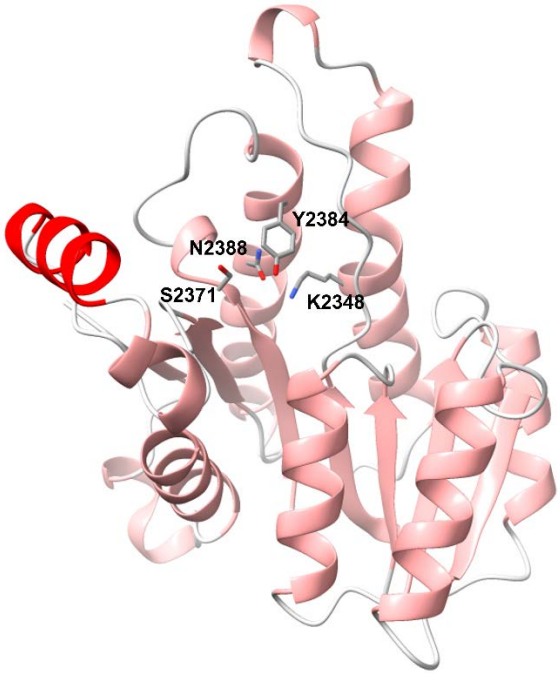


Fig. S20 Predicted structure of the MtAltA KR domain by AlphaFold 3 with a substrate-binding helix coloured in red.

References

1. J. Phakeovilay, W. Imaram, S. Vuttipongchaikij, W. Bunnak, C. M. Lazarus and P. Wattana-Amorn, *Org. Biomol. Chem.*, 2022, **20**, 5050-5054.
2. M. N. Heneghan, A. A. Yakasai, L. M. Halo, Z. Song, A. M. Bailey, T. J. Simpson, R. J. Cox and C. M. Lazarus, *Chembiochem*, 2010, **11**, 1508-1512.
3. H. Li, J. Hu, H. Wei, P. S. Solomon, D. Vuong, E. Lacey, K. A. Stubbs, A. M. Piggott and Y.-H. Chooi, *Org. Lett.*, 2018, **20**, 6148-6152.
4. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C. C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Zemgulyte, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Zidek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis and J. M. Jumper, *Nature*, 2024, **630**, 493-500.
5. E. C. Meng, T. D. Goddard, E. F. Pettersen, G. S. Couch, Z. J. Pearson, J. H. Morris and T. E. Ferrin, *Protein Sci.*, 2023, **32**, e4792.
6. F. Armougom, S. Moretti, O. Poirot, S. Audic, P. Dumas, B. Schaeli, V. Keduas and C. Notredame, *Nucleic Acids Res.*, 2006, **34**, W604-608.
7. X. Robert and P. Gouet, *Nucleic Acids Res.*, 2014, **42**, W320-324.