Electronic Supplementary Information

MmfL catalyses formation of a phosphorylated butenolide intermediate in methylenomycin furan biosynthesis

Shanshan Zhou,[a] Nicolas R. Malet,[a] Lijiang Song,[a] Christophe Corre[a,b,c] and Gregory L. Challis[*a,c,d]

a. Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK.

b. School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK.

c. Warwick Integrative Synthetic Biology Centre, University of Warwick, Coventry CV4 7AL, UK.

d. Department of Biochemistry and Molecular Biology, Monash University, Clayton VIC 3800, Australia.

Materials and Methods

Overproduction and purification of MmfL

E. coli BL21star (DE3) cells were transformed with the previously described pET151-*mmfL* plasmid.[1] 15 µL of a glycerol stock of *E. coli* BL21star (DE3) cells carrying the pET151-*mmfL* plasmid was used to inoculate 15 mL of LB medium containing 100 µg/mL of ampicillin and the resulting culture was incubated for 16 hours (37 °C, 180 rpm). 10 mL of this preculture was used to inoculate 1 L of LB medium (1% v/v) containing 100 µg/mL of ampicillin and the resulting culture was grown for 3-5 hours (37 °C, 180 rpm) until the optical density at 600 nm had reached 0.6-0.8. Isopropyl-β-D-thiogalactopyranoside (ITPG) was then added to a final concentration of 0.5 mM and the culture was grown for a further 16 hours (15 °C, 180 rpm).

After incubation, cells were harvested by centrifugation at 5000 rpm for 20 min and resuspended in washing buffer (15 mL per litre of culture) containing 20 mM Tris-HCl pH 8, 100 mM NaCl, 20 mM imidazole and 10% v/v glycerol. Phenylmethanesuphonyl fluoride (PMSF) was also added to a final concentration of 1 mM. The resuspended cells were lysed using a Constant Systems E1061 cell disrupter and centrifuged at 17,000 rpm for 30 min to separate insoluble material. The supernatant was passed through a 0.2 µm filter and loaded onto a pre-equilibrated 1 mL HisTrap HP nickel affinity column (GE Healthcare). Unbound proteins were removed by washing with 10 mL of the above buffer and His_6-tagged MmfL was eluted using 3 mL each of elution buffers containing successively increasing amounts of imidazole (20 mM Tris-HCl, pH 8, 100 mM NaCl, 10% v/v glycerol, and 50 mM, 100 mM, 200 mM, or 300 mM imidazole). After analysis by SDS-PAGE, the fractions containing the target protein were collected and concentrated using a Vivaspin ultrafiltration column (Sartorius) with a 30 kDa molecular weight cutoff. The purified protein was exchanged into imidazole-free storage buffer (20 mM Tris-HCl, pH 8, 100 mM NaCl and 10% v/v glycerol) using a PD-10 column. The protein was aliquoted, flash frozen in liquid nitrogen and stored at -80 °C. It was used in subsequent experiments without further purification.

In vitro enzyme assays

The buffer used for storage of His_6-MmfL was used for the enzymatic reactions. 400 µM of *S*-(2-acetamidoethyl) 3-oxooctanethioate **1**, 400 µM of DHAP and 15 µM of His_6-MmfL in a total volume of 200 µL were incubated for 90 min at room temperature. One unit of shrimp alkaline phosphatase (New England Biolabs) was added and the resulting mixture was incubated for a further 45 min at room temperature. 200 µL of methanol was added and, after 5 min, the mixture was centrifuged at 14, 200 rpm for 10 min. The supernatant was passed through a 0.4 µm filter and analysed by LC-MS. For the negative control reaction employing denatured MmfL, protein solution was boiled for 10 min and allowed to cool, prior to use.

LC-MS

LC-MS analyses were performed on a Dionex Ultimate 3000 HPLC instrument equipped with a ZORBAX Eclipse Plus C18 column (2.1 × 100 mm, 1.8 µm) coupled to a Bruker MaXis Impact mass spectrometer [ESI in positive ion mode; full scan 50-2500 *m/z*; end plate offset, -500 V; capillary, -4500 V; nebulizer gas (N_2), 1.4 bar; dry gas (N_2), 8 L/min; dry temperature, 200 °C]. The solvents used for elution of the column were as follows: solvent A, 0.1% formic acid in water; solvent B, 0.1% formic acid in acetonitrile. The elution profile was as follows: 0-5 min, 5% B; 5-17 min, linear gradient from 5% B to 100% B; 17-22 min, 100% B; 22-25 min, linear gradient from 100% B to 5% B; 25-34 min, 5% B. The flow rate was 0.2 mL/min. The mass spectrometer was calibrated with 10 mM sodium formate at the beginning of each run.

For comparison of the MS/MS fragmentation pattern of deprotected **4** and the dephosphorylated product of the MmfL-catalysed reaction, LC-MS/MS analyses were performed on an Agilent 1260 Infinity HPLC equipped with

an Agilent Eclipse XDB-C18 column (4.6 × 150 mm, 5 µm) coupled to a Bruker AmaZon X ion trap mass spectrometer [ESI in positive ion mode; full scan 100-3000 *m/z*; end plate offset, -500 V; capillary, -4500 V; nebulizer gas (N_2), 40 p.s.i.; dry gas (N_2), 10 L/min; dry temperature, 200 °C]. The solvents used for elution of the column were as follows: solvent A, 0.1% formic acid in water; solvent B, 0.1% formic acid in acetonitrile. The elution profile was as follows: 0-5 min, 5% B; 5-30 min, linear gradient from 5% B to 100% B; 30-35 min, 100% B; 35-40 min, linear gradient from 100% B to 5% B. The flow rate was 1.0 mL/min.

Organic synthesis

General information

All chemicals were purchased from commercial sources and were used as received. ^{1}H and ^{13}C NMR spectra were measured at 300 or 400 MHz and 75 or 100 MHz, respectively. All chemical shifts are given as δ values in ppm with reference to the residual protiated solvent peaks at δ_H 7.26 and δ_C 77.00 for $CDCl_3$. Data for ^{1}H NMR spectra are reported as follows: chemical shift (δ, ppm), multiplicity (s = singlet, t = triplet, dd = doublet of doublets, td = triplet of doublets, m = multiplet, br = broad), integration and coupling constant (Hz). HRMS spectra were measured on a Bruker MaXis impact mass spectrometer. NMR spectra are shown in **Fig. S3-S10**. High resolution mass spectra of compounds **3** and **4** are shown in **Fig. S11**.

1-((tert-Butyldimethylsilyl)oxy)-3-hydroxypropan-2-one.

To a mixture of dihydroxyacetone (5.00 g, 55.6 mmol) and imidazole (1.51 g, 22.2 mmol) in 50 mL DMF at 0 °C was added a solution of *tert*-butyldimethylsilyl chloride (2.67 g, 17.8 mmol) in 10 mL DMF. The resulting mixture was stirred at room temperature (RT) for 17 h, then 50 mL of water was added and the mixture was extracted with Et_2O. The organic phase was washed with brine, dried over $MgSO_4$ and concentrated under vacuum. The residue was purified by silica gel flash column chromatography (pet ether/EtOAc = 5:1 v/v) to give the product (1.82 g, 50%) as a colourless oil. The spectroscopic data are consistent with those reported previously.[2] ^{1}H NMR (300 MHz, $CDCl_3$) δ: 0.12 (s, 6H), 0.90 (s, 9H), 2.99 (br s, 1H), 4.31 (s, 2H), 4.50 (s, 2H), further signals at 3.43 - 4.24 are due to reversible formation of various ketals. ^{13}C NMR (75 MHz, $CDCl_3$) δ: 210.88, 25.74, 18.15, -5.56, further signals at 63.07 - 111.38 are due reversible formation of various ketals. HR-MS: *m/z* calculated for $C_9H_{20}NaO_3Si$ $[M+Na]^+$: 227.1074; found: 227.1076.

5-(1-Hydroxyhexylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione (acyl Meldrum's acid).

To a solution of hexanoic acid (1.20 g, 10.3 mmol) in DCM was added EDC·HCl (2.96 g, 15.5 mmol) and the mixture was cooled to 0 °C. DMAP (3.78 g, 30.9 mmol) dissolved in DCM was added, followed by Meldrum's acid (1.48 g, 10.3 mmol) dissolved in DCM. After stirring at room temperature for 18 h, the reaction mixture was diluted with DCM, and washed with 1 M HCl, water and brine. The organic phase was dried over $MgSO_4$ and concentrated under vacuum. The residue was purified by silica gel flash column chromatography to yield the product (1.67 g, 67%) as a yellow oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.78 (t, 3H, J = 7.0 Hz), 1.17 - 1.31 (m, 4H), 1.54 - 1.58 (m, 2H), 1.61(s, 6H), 2.93 (t, 2H, J = 7.5 Hz), 15.18 (br s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$) δ: 170.23, 159.77, 104.37, 90.96, 35.28, 31.14, 26.38, 25.46, 21.93, 13.48. LR-MS: *m/z* calculated for $C_{12}H_{17}O_5$ $[M-H]^-$: 241.1; found: 241.1.

*S-(2-Acetamidoethyl) 3-oxooctanethioate **1**.*

To a 0.5 M solution of 5-(1-hydroxyhexylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione (0.63 g, 2.6 mmol) in 1,2-dichloroethane under argon was added *N*-acetylcysteamine (0.20 g, 1.7 mmol) and the reaction mixture was refluxed for 3 h. After removal of the solvent under vacuum, the residue was purified by silica gel flash column chromatography with a 10 cm plug of $CuSO_4$-impregnated silica gel on top to afford **1** (0.35 g, 80%) as a white solid,

mp 89-90 °C. The spectroscopic data are consistent with those reported previously.[3] ^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.86 - 0.90 (m, 3H), 1.24 - 1.31 (m, 4H), 1.58 (quintet, 2H, *J* = 7.0 Hz), 1.96 (s, 3H), 2.16 (t, 2H, *J* = 7.5 Hz), 2.51 (t, 2H, *J* = 7.5 Hz), 3.05 - 3.09 (m, 2H), 3.42 - 3.48 (m, 2H), 3.68 (s, 2H), 5.45 (s, 1H), 6.02 (br s, 1H), 12.59 (br s, 1H). The peaks underlined are due to the enol form of the product (approximately 25%). ^{13}C NMR (100 MHz, $CDCl_3$) δ: 194.24, 192.34, 177.63, 170.41, 99.09, 57.14, 43.37, 39.86, 39.14, 34.82, 31.19, 31.07, 29.17, 27.78, 25.86, 23.11, 23.05, 22.32, 22.28, 13.82. HR-MS: *m/z* calculated for $C_{12}H_{21}NNaO_3S$ $[M+Na]^+$: 282.1134; found: 282.1147.

*4-(((tert-Butyldimethylsilyl)oxy)methyl)-3-hexanoylfuran-2(5H)-one **4**.*

A solution of 1-((tert-butyldimethylsilyl)oxy)-3-hydroxypropan-2-one (0.35 g, 1.7 mmol) and acyl Meldrum's acid (0.48 g, 2.0 mmol) in 2 mL of toluene was refluxed for 3 h, then additional acyl Meldrum's acid (0.22 g, 0.9 mmol) was added. After stirring for an additional 5 h, the reaction mixture was allowed to stand at -20 °C for 19 h and purified directly without removal of the toluene using silica gel flash column chromatography (hexanes/acetone = 60:1 v/v) to yield **26** (0.17 g, 30%) as a yellowish oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ: 0.08 (s, 6H), 0.81 - 0.92 (m, 12H), 1.29 - 1.32 (m, 4H), 1.59 (quintet, 2H, *J* = 7.0 Hz), 2.95 (t, 2H, *J* = 7.5 Hz), 4.97 (s, 2H), 5.05 (s, 2H). ^{13}C NMR (100 MHz, $CDCl_3$) δ: 197.08, 181.16, 170.61, 122.43, 70.17, 61.81, 41.52, 31.19, 30.80, 25.65, 22.80, 22.44, 13.85, -5.70. HR-MS: *m/z* calculated for $C_{17}H_{30}NaO_4Si$ $[M+Na]^+$: 349.1806; found: 349.1810.

3-Hexanoyl-4-(hydroxymethyl)furan-2(5H)-one **3**.

To a solution of **4** (2 mg) in 200 µL of THF at 0 °C was added 2 drops of hydrogen fluoride-pyridine and the resulting mixture was stirred for 40 min at 0 °C. The reaction was quenched with 500 µL saturated $NaHCO_3$ solution, concentrated under vacuum to remove the THF and extracted with EtOAc. The organic layers were combined, concentrated and re-dissolved in 700 µL of 1:1 v/v methanol/water. The sample was passed through a 0.4 µm filter and analysed by LC-MS. HR-MS: *m/z* calculated for $C_{11}H_{16}NaO_4$ $[M+Na]^+$: 235.0941; found: 235.0941.

Incorporation of $^{18}O_2$ into MMF1

Streptomyces coelicolor W89[4] was grown on a modified SMMS agar medium,[5] containing 5mM each of NaH_2PO_4 and K_2HPO_4 under an atmosphere of 80% N_2 and 20% $^{18}O_2$. After incubation at 30 °C for 72 h, the plates were frozen overnight, then defrosted, and the agar and mycelia were removed by filtration through cotton wool by centrifugation for 5 mins at 2000 rpm. The resulting culture supernatant was analysed by positive ion LC-MS and negative ion LC-MS/MS as described previously.[1]

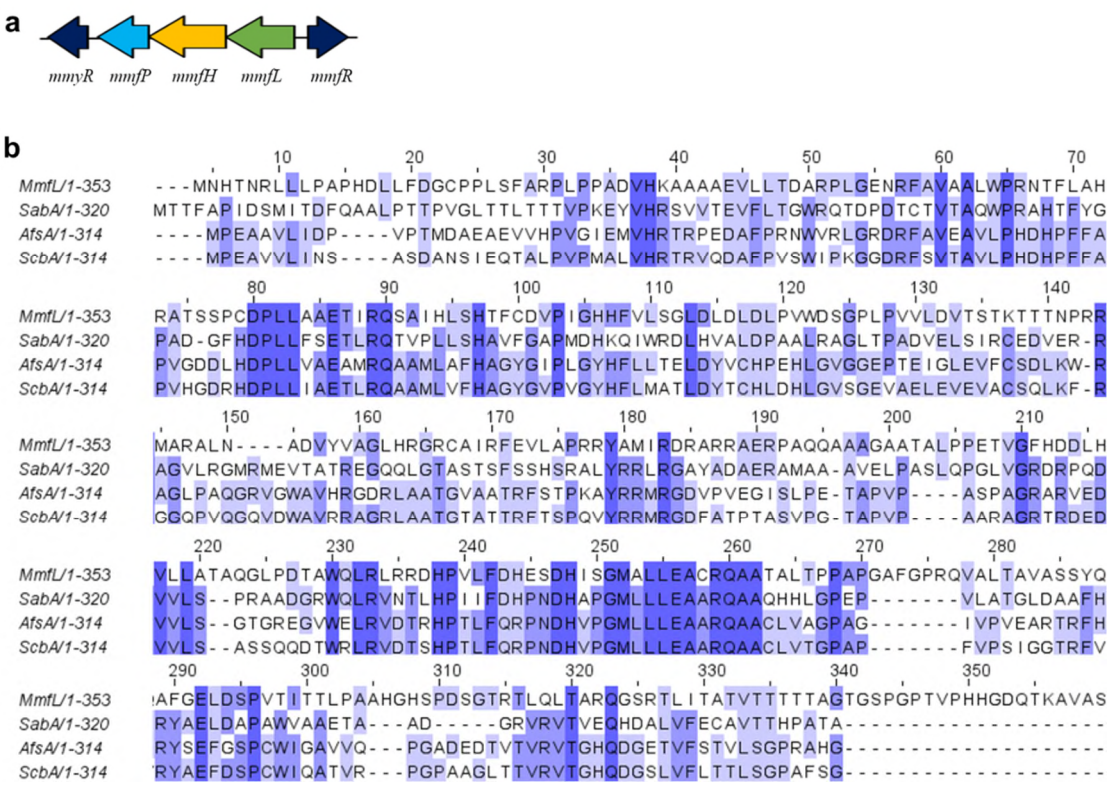

Fig. S1 (a) Gene organization in the neighbourhood of *mmfL*. MmfP: putative phosphatase; MmfH: putative flavin-dependent oxidoreductase; MmfR and MmyR: putative DNA-binding receptor proteins. **(b)** Multiple sequence alignment of MmfL (UniProt ID Q7APH6), AfsA (UniProt ID B1VN93), ScbA (UniProt ID Q7AKF0) and SabA[6].

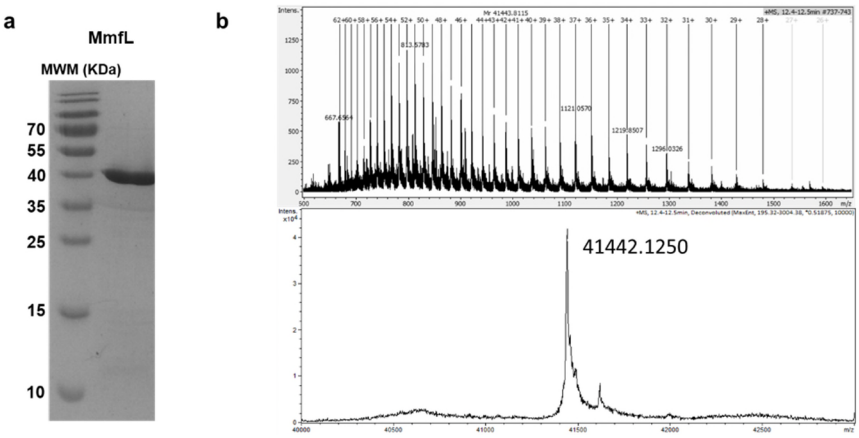

Fig. S2 Analysis of purified recombinant His_6-MmfL. a, SDS-PAGE. **b**, Mass spectrometric analysis: measured (top) and deconvoluted (bottom) mass spectra of His_6-MmfL (calculated mass = 41442.0 Da).

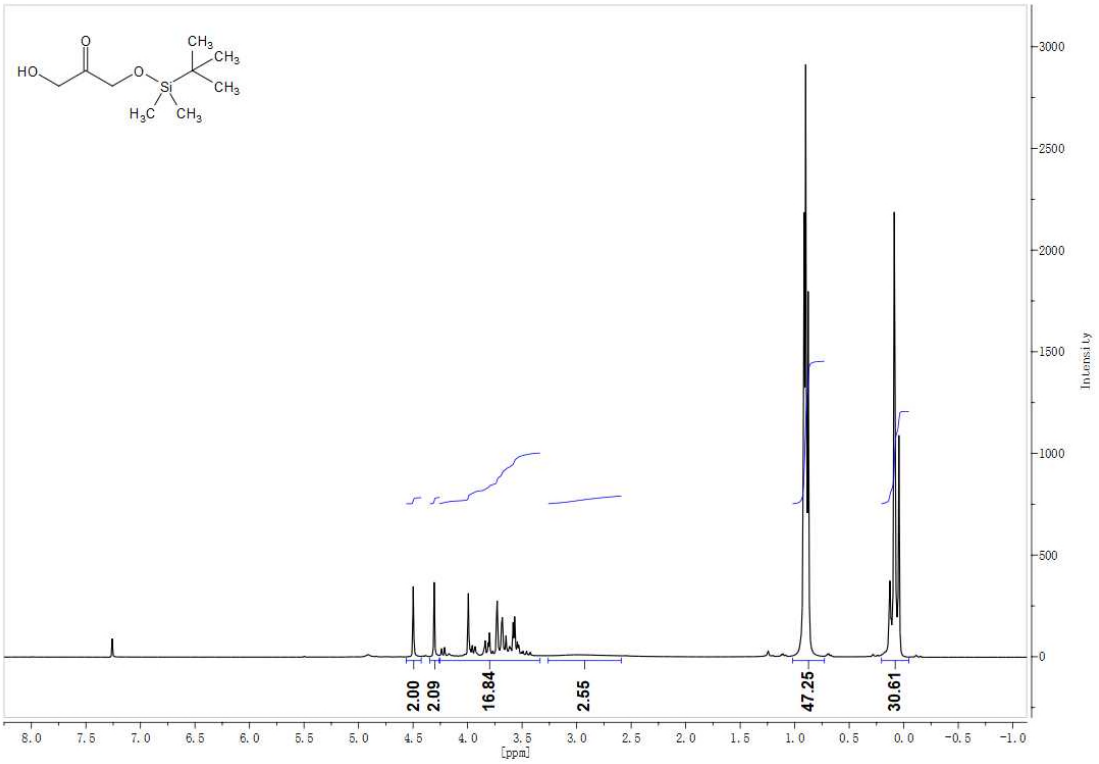

Fig. S3 ^{1}H NMR spectrum (300 MHz, $CDCl_3$) of 1-((*tert*-butyldimethylsilyl)oxy)-3-hydroxypropan-2-one.

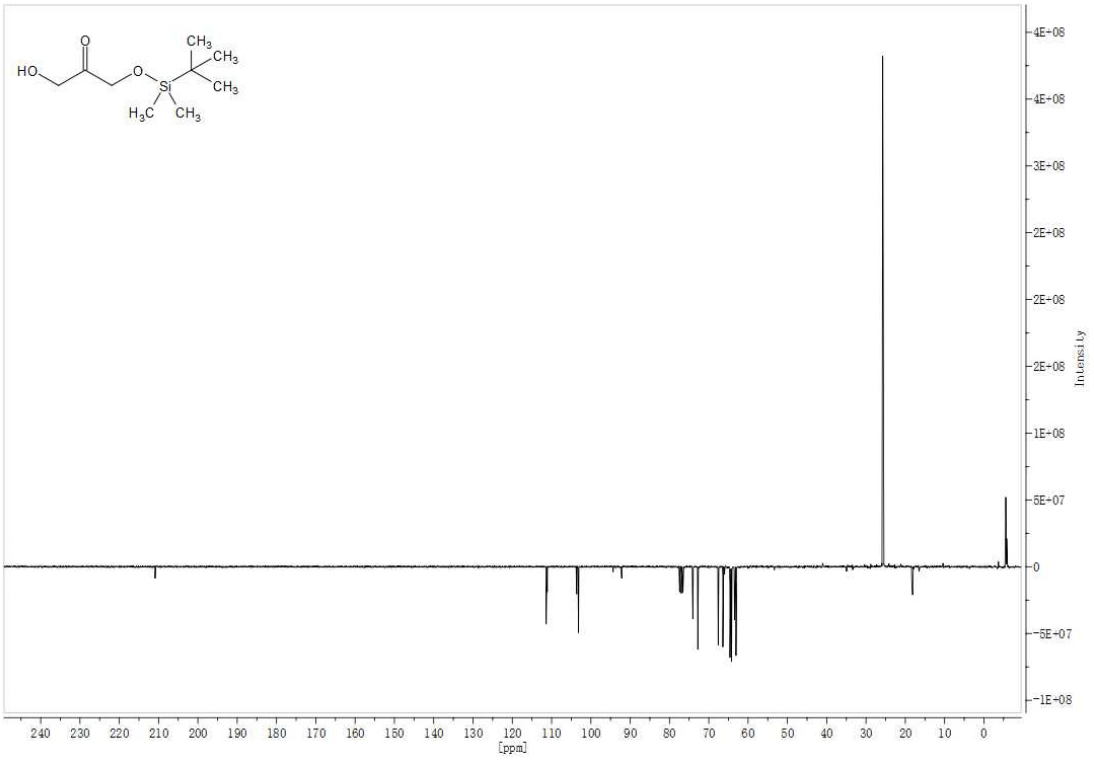

Fig. S4 ^{13}C NMR spectrum (75 MHz, $CDCl_3$) of 1-((*tert*-butyldimethylsilyl)oxy)-3-hydroxypropan-2-one.

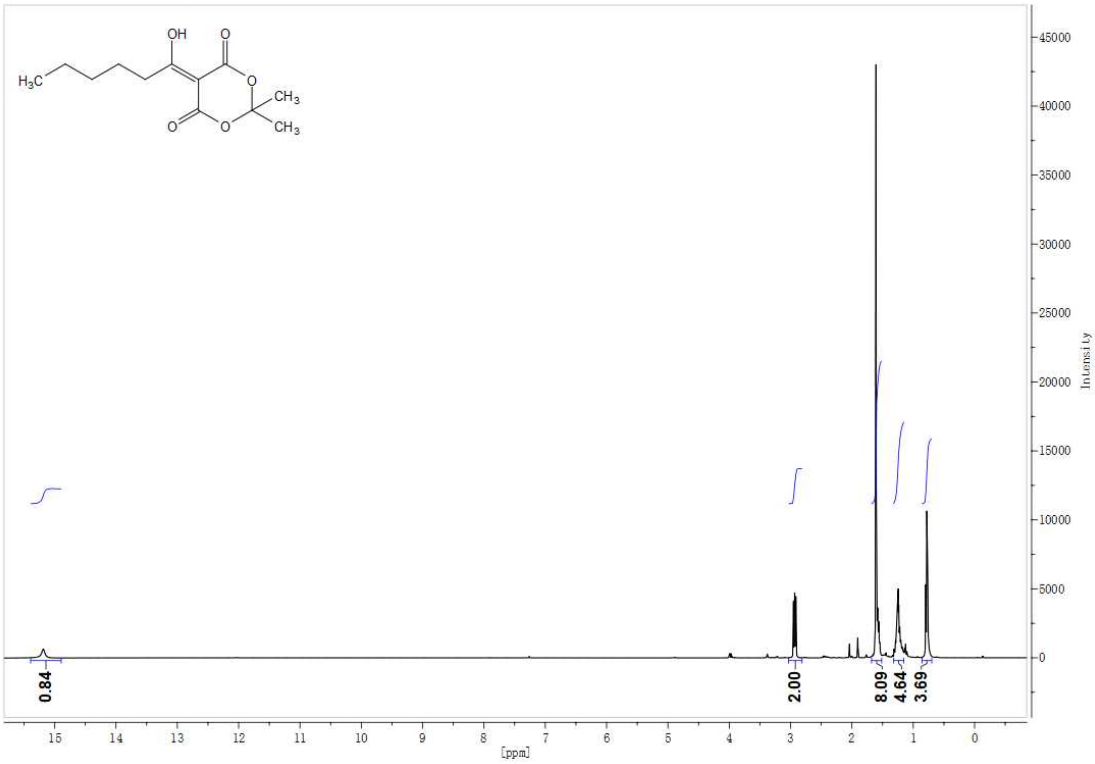

Fig. S5 ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of 5-(1-hydroxyhexylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione.

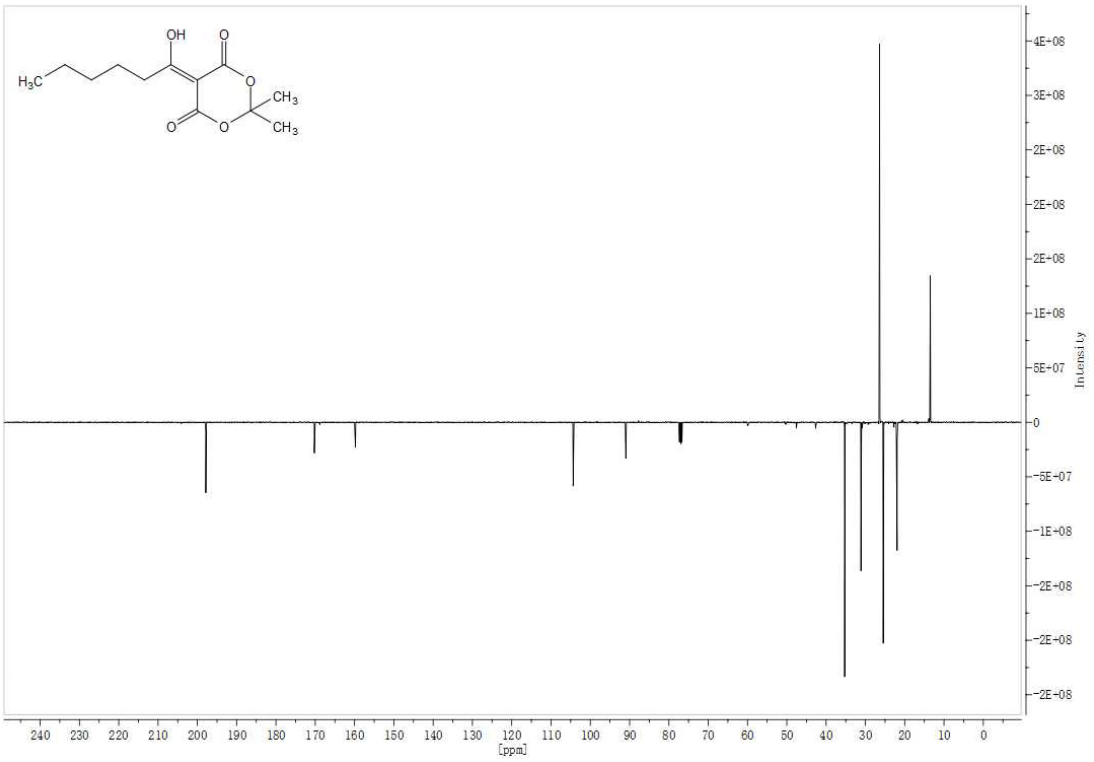

Fig. S6 ^{13}C NMR spectrum (100 MHz, $CDCl_3$) of 5-(1-hydroxyhexylidene)-2,2-dimethyl-1,3-dioxane-4,6-dione.

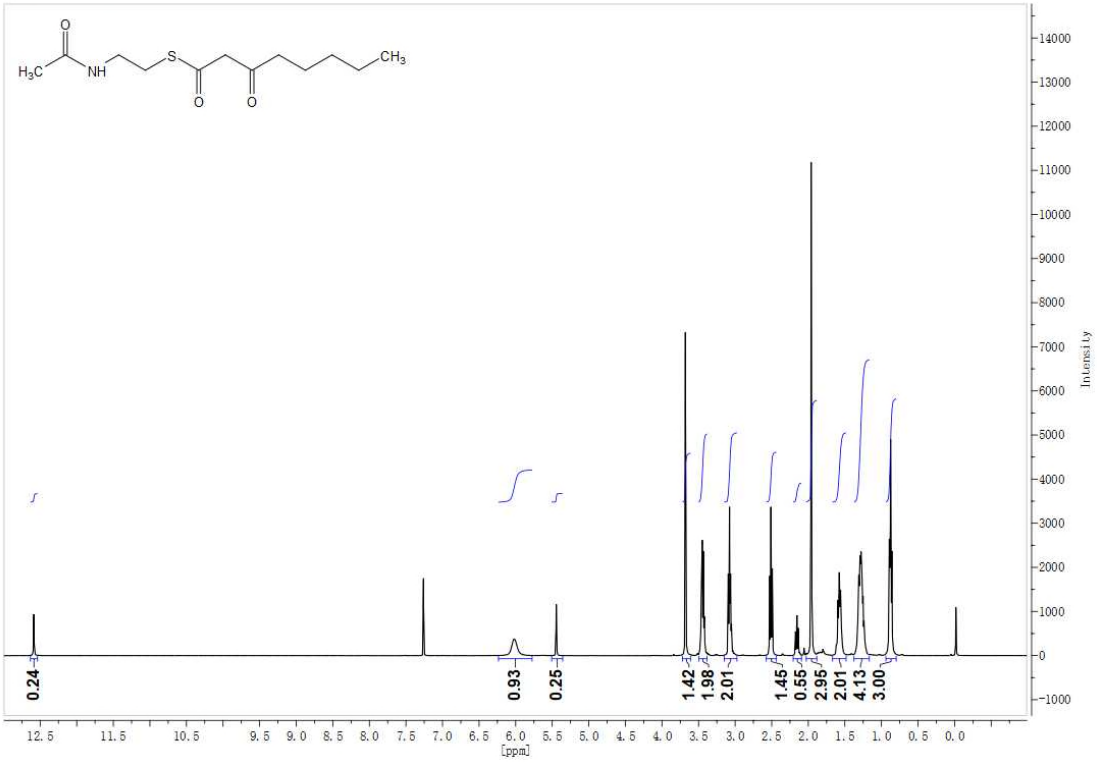

Fig. S7 ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **1**.

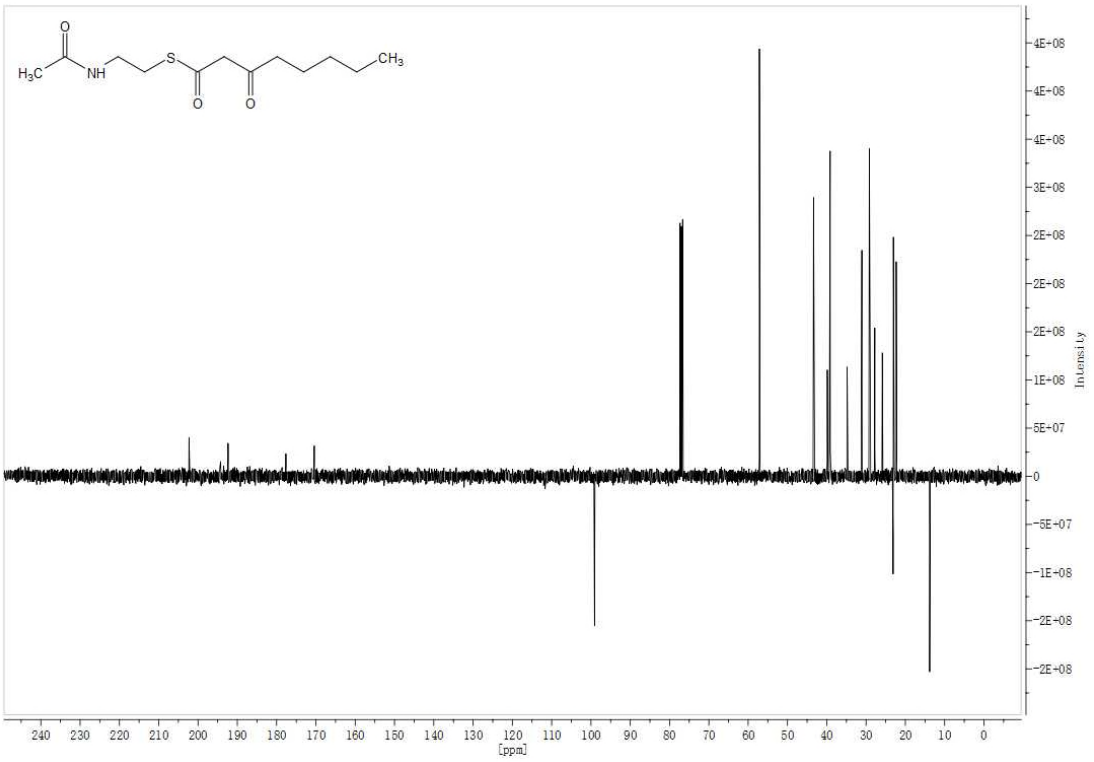

Fig. S8 ^{13}C NMR spectrum (100 MHz, $CDCl_3$) of compound **1**.

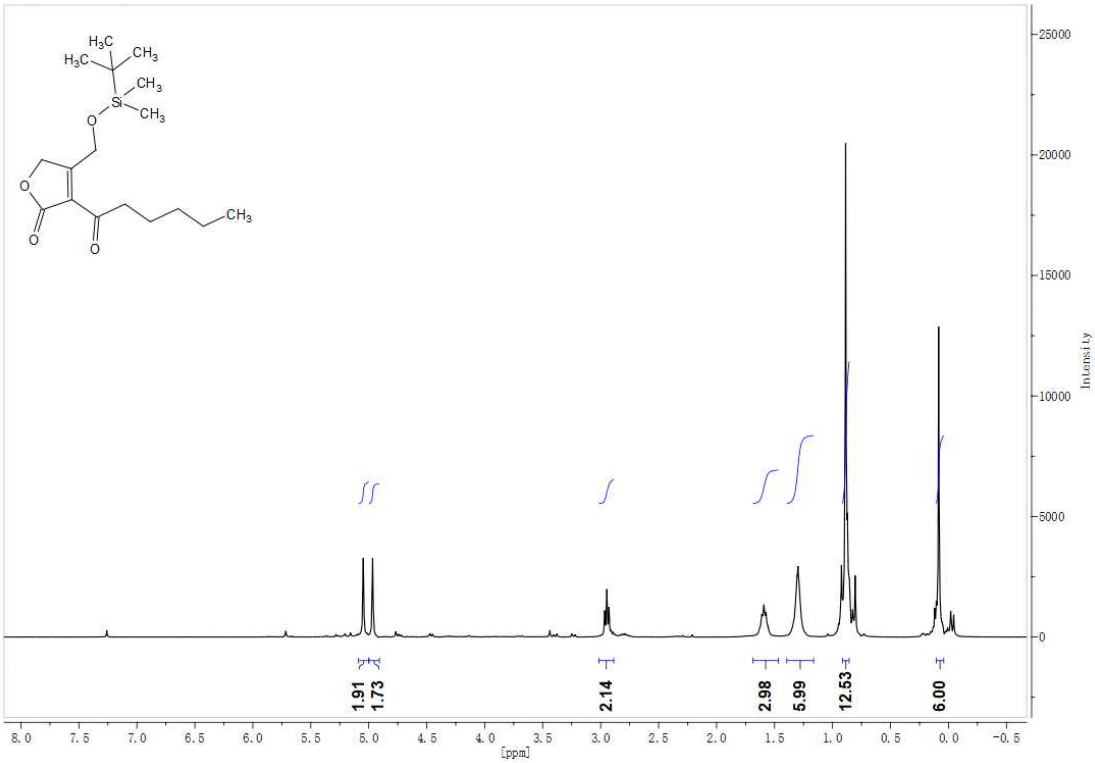

Fig. S9 ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **4**.

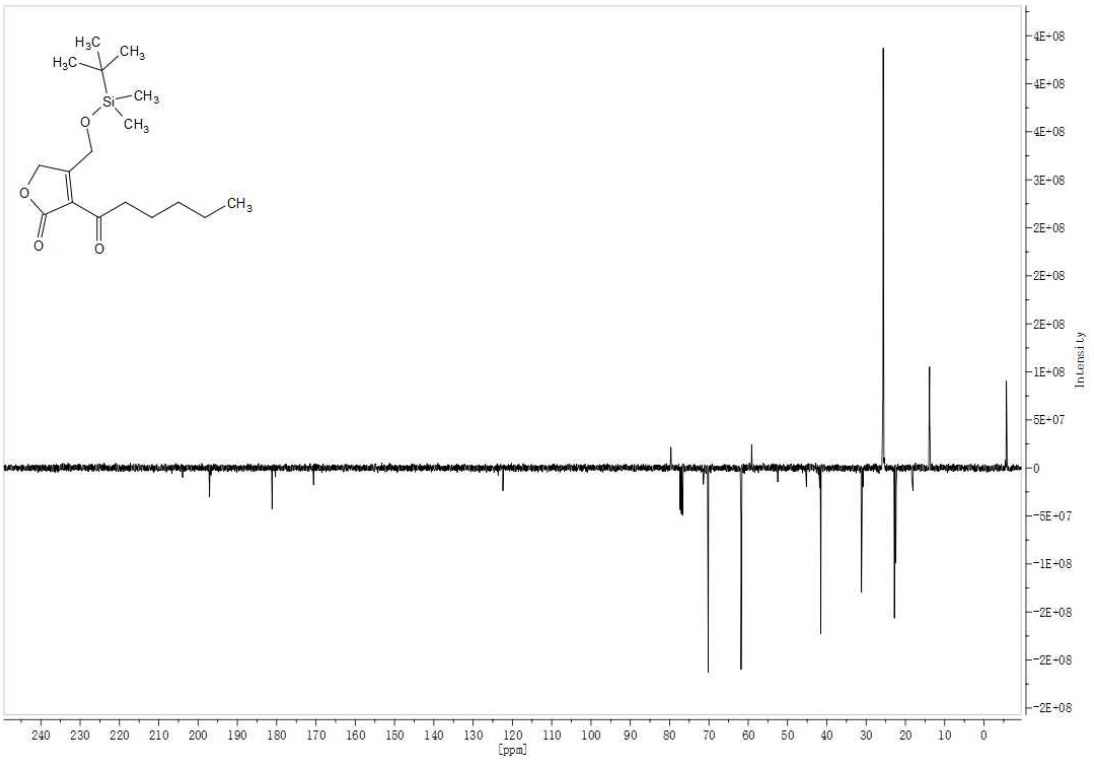

Fig. S10 ^{13}C NMR spectrum (100 MHz, $CDCl_3$) of compound **4**.

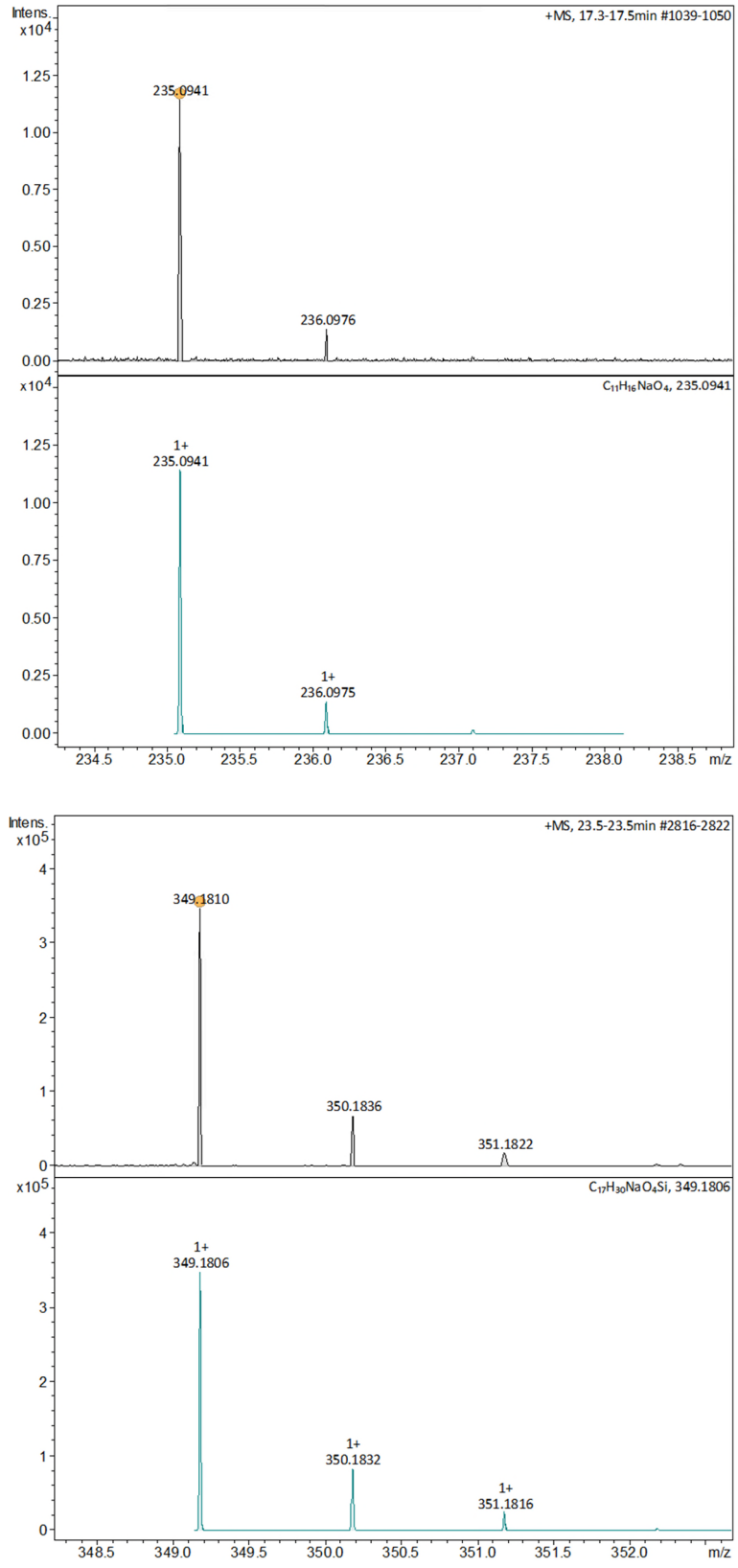

Fig. S11 Comparison of measured (top panel) and simulated (bottom panel) mass spectra for compounds **3** (top) and **4** (bottom).

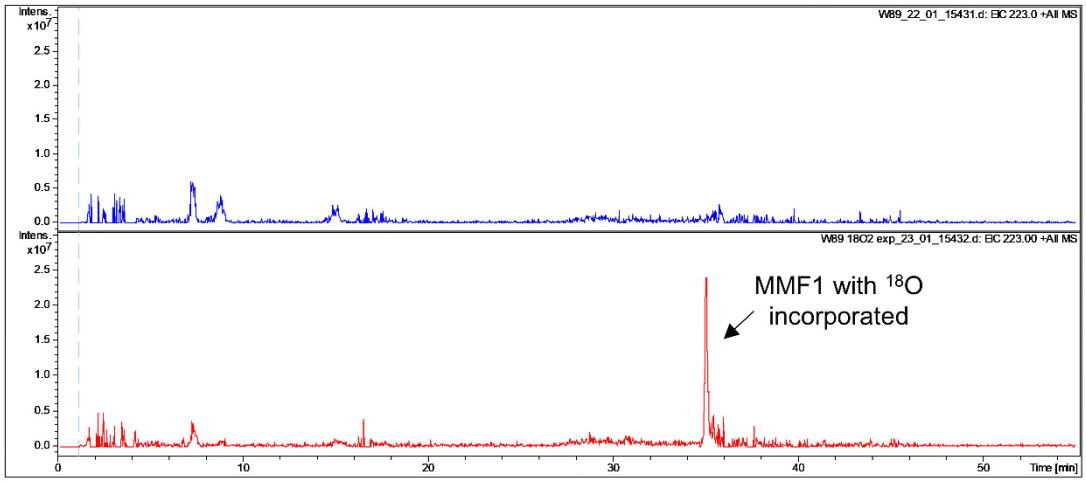

Fig. S12 Extracted ion chromatograms (EICs) for *m/z* = 223, corresponding to $[M+Na]^+$ for singly ^{18}O-labelled MMF1, from LC-MS analyses of culture supernatants of *S. coelicolor* W89 grown in air (top panel) and under an $^{18}O_2$ atmosphere (bottom panel).

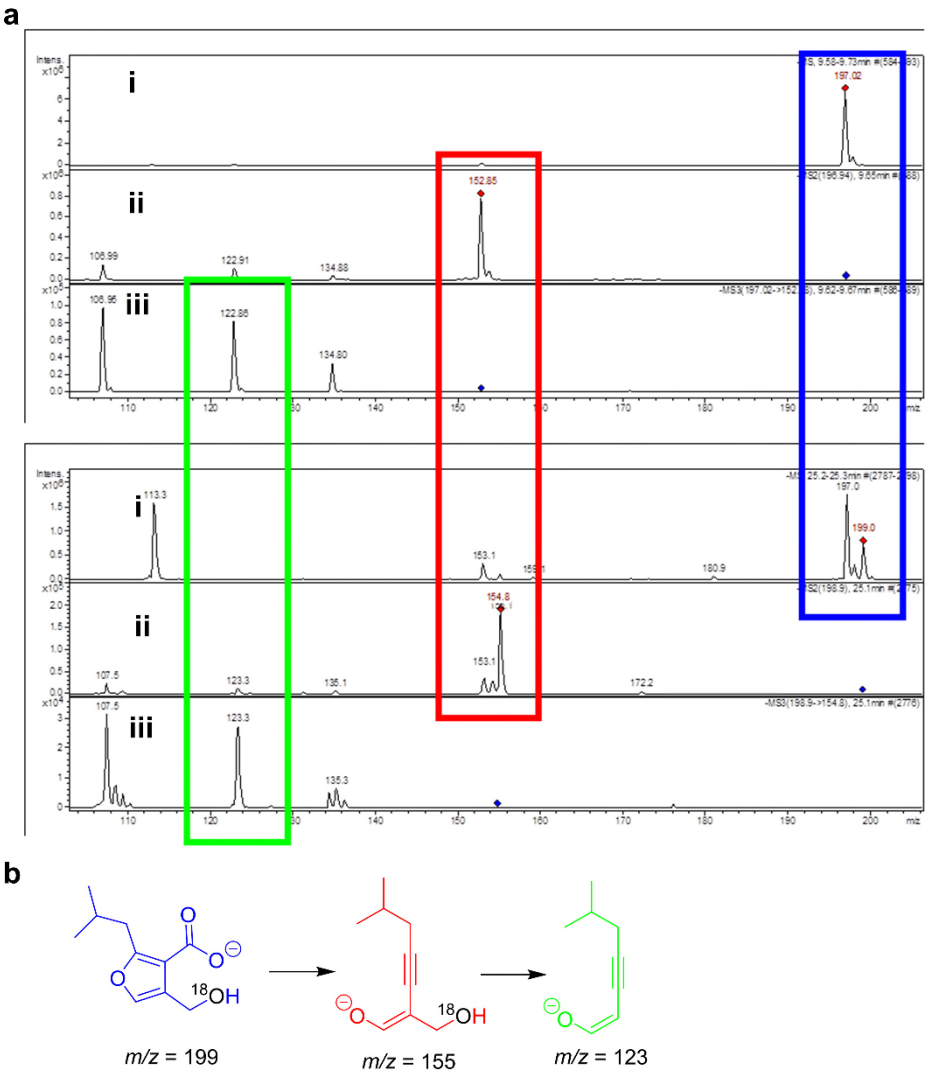

Fig. S13 (a) Comparison of fragment ions observed for unlabelled MMF1 (top panel) and ^{18}O-labelled MMF1 (bottom panel). **i**) Parent ion; **ii**) Daughter ion resulting from fragmentation of the parent ion; **iii**) Second daughter ion from resulting fragmentation of the first daughter ion. **(b)** Proposed structures of ions observed.[7]

References:

1. C. Corre, L. Song, S. O'Rourke, K. F. Chater and G. L. Challis, *Proc. Natl. Acad. Sci. U. S. A.,* 2008, **105**, 17510.
2. J. Schröer and P. Welzel, *Tetrahedron,* 1994, **50**, 6839.
3. C. Arthur, R. J. Cox, J. Crosby, M. M. Rahman, T. J. Simpson, F. Soulas, R. Spogli, A. E. Szafranska, J. Westcott and C. J. Winfield, *ChemBioChem,* 2002, **3**, 253.
4. B. Aigle and C. Corre, *Methods Enzymol.* 2012, **517**, 343.
5. E. Takano, R. Chakraburtty, T. Nihira, Y. Yamada and M.J. Bibb, *Mol. Microbiol.* 2001, **41**, 1015.
6. W. Wang, J. Zhang, X. Liu, D. Li, Y. Li, Y. Tian and H. Tan, *J. Biol. Chem.,* 2018, **293**, 20029.
7. C. Corre, S. W. Haynes, N. Malet, L. Song and G. L. Challis, *Chem. Commun.* 2010, **46**, 4079.