

Supporting Information

Methylation-dependent Acyl Transfer between Polyketide Synthase and Nonribosomal Peptide Synthetase Modules in Fungal Natural Product Biosynthesis

Yi Zou¹||, Wei Xu¹||, Yuta Tsunematsu², Mancheng Tang¹, Kenji Watanabe², Yi Tang^{1,3*}

¹Department of Chemical and Biomolecular Engineering, ³Department of Chemistry and Biochemistry, University of California, Los Angeles, ²Department of Pharmaceutical Sciences, University of Shizuoka, Shizuoka 422-8526.

Experimental Procedures

1. Strains and General DNA Manipulation Techniques

E. coli TOP10 (Invitrogen) and *E. coli* DH10B were used for cloning, following standard recombinant DNA techniques. DNA restriction enzymes were used as recommended by the manufacturer (New England Biolabs, NEB). PCR was performed using Phusion[®] High-Fidelity DNA Polymerase (NEB). PCR products were confirmed by DNA sequencing. *E. coli* BL21(DE3) (Novagen) was used for protein expression. *Saccharomyces cerevisiae* strain BJ5464-NpgA (*MATa ura3-52 his3-Δ200 leu2- Δ1 trp1 pep4::HIS3 prb1 Δ1.6R can1 GAL*) was used as the yeast expression host.^[1]

2. Plasmid construction

The *psoA*, *psoB*, *psoF*, *psoF-MT*, *psoF-FMO* and *psoG* genes were cloned from the genomic DNA or cDNA of *A. fumigatus* Af293 using the primers listed in Table S1. The 2μ-based yeast-*E. coli* shuttle plasmids with different auxotrophic markers (*URA3*, *TRP1* and *LEU2*) were used for construction of the yeast expression vectors.^[2]

Cloning of N-FLAG and C-His₆ PsoA *S. cerevisiae* expression vector

psoA was cloned from the genomic DNA or cDNA in 5 segments, using the primers listed in Table S1. All of them were cloned into pXW55 vector at *SpeI* and *PmlI* sites by yeast recombination to yield the plasmid pXW55-PsoA.

Cloning of N-His₆ PsoF *E. coli* expression vector

psoF was cloned using the primers for *E. coli* vector listed in Table S1. The PCR product was digested with *EcoRI* and *XhoI* and then inserted into pET-28a(+) to yield the N-His₆ PsoF *E. coli* expression plasmid pET28a-PsoF.

Cloning of PsoF *S. cerevisiae* expression vector

psoF was cloned using the primers for yeast vector listed in Table S1. The PCR product was ligated into pXW06 vector at *NdeI* and *PmeI* sites by yeast recombination to yield the plasmid pXW06-PsoF.

Cloning of PsoF-MT domain *S. cerevisiae* expression vector

psoF-MT domain was cloned using the primers for yeast vector listed in Table S1. The PCR product was ligated into pXW06 vector at *NdeI* and *PmeI* sites by yeast recombination to yield the plasmid pXW06-PsoF-MT.

Cloning of PsoF-FMO domain *S. cerevisiae* expression vector

PsoF-FMO domain was cloned using the primers for yeast vector listed in Table S1. The PCR product was ligated into pXW06 vector at *NdeI* and *PmeI* sites by yeast recombination to yield the plasmid pXW06-PsoF-FMO.

Cloning of N-His₆ PsoB *E. coli* expression vector

psoB was cloned using the primers for *E. coli* vector listed in Table S1. The PCR product was digested with *NdeI* and *NotI* and then inserted into pET-28a(+) to yield the N-His₆ PsoB *E. coli* expression plasmid pET28a-PsoB.

Cloning of PsoB *S. cerevisiae* expression vector

psoB was cloned using the primers for yeast vector listed in Table S1. The PCR product was digested with *NdeI* and *PmeI*, and then ligated into pXW02 vector at *NdeI* and *PmeI* sites to yield the plasmid pXW02-PsoB.

Cloning of PsoB and PsoG *S. cerevisiae* co-expression vector

psoG was cloned from the genomic DNA in 3 segments, using the primers listed in Table S1. All of them were ligated into pXW55 vector at *SpeI* and *PmlI* sites by yeast recombination to yield the plasmid pXW55-PsoG. pXW55-PsoG was digested with *SacI*, and then inserted into the same digested site of pXW02-PsoB to yield the PsoB and PsoG co-expression plasmid pXW02-PsoBG.

3. Protein expression and purification

Expression and purification of PsoF, PsoB and ApdA NRPS from *E. coli*

The PsoF and PsoB expression plasmids were transformed into *E. coli* BL21(DE3) strain for protein expression. The *E. coli* BL21(DE3) cell harboring the expression plasmid was cultured at 37°C and 220 rpm in LB medium supplemented with 50 µg mL⁻¹ kanamycin (final concentration) to OD₆₀₀ between 0.4 and 0.6. The cultures were then incubated on ice for 10 min before addition of 0.2 mM (final concentration) IPTG to induce protein expression. The cells were further cultured at 16°C for 18-20 hours, and then were harvested by centrifugation (3,500 rpm, 15 min, 4°C), re-suspended in 40 mL buffer A (50 mM Tris-HCl, pH 7.9, 0.5 M NaCl, and 10% glycerol) and lysed by sonication on ice for 40 min. Cellular debris was removed by centrifugation (14,000 rpm, 45 min, 4°C), and the supernatant was used to purify the protein by nickel-affinity chromatography using standard protocols. The protein was eluted with increasing gradient of buffer B (500 mM imidazole in buffer A). Purified proteins were concentrated and exchanged into Buffer C (50 mM Tris-HCl, pH 7.9, 50 mM NaCl, and

5% glycerol) with the centriprep filters (Amicon). The protein was stored in buffer C at -80°C . Protein concentration was determined by Bradford assay using bovine serum albumin as a standard. Expression and purification of ApdA NRPS from *E. coli* using protocols reported previously.^[2b]

Expression and purification of PsoA from *S. cerevisiae*

pXW55-PsoA was transformed into *S. cerevisiae* strain BJ5464-NpgA by using *S. c.* EasyComp™ Transformation Kit (Invitrogen). For 1L of yeast culture, the cells were grown at 28°C in YPD media for 72 hours. The cells were harvested by centrifugation (3,500 rpm, 20 minutes, 4°C), re-suspended in 20 mL lysis buffer (50 mM NaH_2PO_4 pH 8.0, 0.15 M NaCl, 10 mM imidazole) and lysed through sonication on ice. Cellular debris was removed by centrifugation (17,000 rpm, 60 min, 4°C). Ni-NTA agarose resin was added to the supernatant (2 mL per liter of culture) and the solution was stirred at 4°C at least 4 hours. Purify PsoA by nickel-affinity chromatography using standard protocols. The protein was eluted with increasing gradient of buffer B (500 mM imidazole in buffer A). Purified proteins were concentrated and exchanged into Buffer C (50 mM Tris-HCl, pH 7.9, 50 mM NaCl, 2 mM DTT and 5% glycerol) with the Centriprep filters (Amicon). The protein was stored in buffer C at -80°C . Protein concentration was determined by Bradford assay using bovine serum albumin as a standard.

4. In vitro characterization of PsoA, PsoF, PsoB and ApdA NRPS

1) Assays for the PsoA activity with various CoA substrates were performed at 100 μL scale, 25°C , 12 hrs. 12 μM PsoA, 1 mM malonyl-CoA, 4 mM NADPH with 0.5 mM acetyl-CoA, isobutyryl-CoA, hexanoyl-CoA, octanoyl-CoA, or propionyl-CoA and equal amounts of $\text{C}_2\text{-C}_8$ CoAs. Assays for ^{13}C labeling of the products, 25 mM MatB^[31] was used to replenish malonyl-CoA from 20 mM CoA and 100 mM [2- ^{13}C]-malonate.

2) Assay for the PsoA and PsoF activities was performed at 100 μL scale, 25°C , 12 hrs. 12 μM PsoA, 10 μM PsoF, 1 mM malonyl-CoA, 4 mM NADPH, 0.5 mM propionyl-CoA, and 2 mM SAM.

3) Assay for the PsoA, PsoF and PsoB activities was performed at 100 μ L scale, 25°C, 12 hrs. 12 μ M PsoA, 10 μ M PsoF, 10 μ M PsoB, 1 mM malonyl-CoA, 4 mM NADPH, 0.5 mM propionyl-CoA, 2 mM SAM, 1 mM L-Phe, 6 mM ATP, 0.5 mM MgCl₂ and 50 μ M FAD.

4) Assays for the ApdA_NRPS activity was performed at 100 μ L scale, 25°C, 12 hrs in a final 100 mM pH 7.5 sodium phosphate buffer. 10 μ M ApdA_NRPS, 1 mM Acyl-S-NAC, 1 mM L-Tyr, 10 mM ATP, 1 mM MgCl₂.

The reaction mixtures were quenched and extracted twice with equal volume of 99% ethyl acetate (EA)/1% acetic acid (AcOH). The resultant organic extracts were evaporated to dryness, re-dissolved in methanol, and then analyzed on LC-MS. LC-MS was conducted with a Shimadzu 2010 EV Liquid Chromatography Mass Spectrometer by using both positive and negative electrospray ionization, and a Phenomenex Luna 5 μ 2.0 x 100 mm C₁₈ reverse-phase column. Samples were separated on a linear gradient of 5 to 95% CH₃CN (v/v) in H₂O supplemented with 0.05% (v/v) formic acid at a flow rate of 0.1 mL/min.

5. Heterologous production of azaspirene and its intermediates in yeast

pXW55-PsoA, pXW06-PsoF or pXW06-PsoF-MT or pXW06-PsoF-FMO and pXW02-PsoB or pXW02-PsoBG plasmids were co-transformed into *S. cerevisiae* BJ5464-NpgA by using *S. c.* EasyComp™ Transformation Kit (Invitrogen). The cells were firstly grown in a minimal medium that lacks of uracil, tryptophan and leucine for 2 days. The seed culture (2.5 mL) was transferred into 25 mL YPD media, and then incubated at 28°C, 250 rpm for 3 days. The supernatant was extracted with equal volume of ethyl acetate twice. The organic extract was combined, concentrated and then dissolved in 100 μ L methanol for LC-MS analysis. LC-MS was conducted with a Shimadzu 2010 EV Liquid Chromatography Mass Spectrometer by using both positive and negative electrospray ionization, and a Phenomenex Luna 5 μ 2.0 x 100 mm C₁₈ reverse-phase column. Samples were separated on a linear gradient of 5 to 95% CH₃CN (v/v) in H₂O supplemented with 0.05% (v/v) formic acid at a flow rate of 0.1 mL/min.

6. Chemical synthesis of S-NACs

Preparations of the 3-oxo-acyl-S-NACs **14a-14c** follow similar procedures. The synthesis of **14a** is described here as an example. 2-methylheptanoyl chloride was synthesized from 2-methylheptanoic acid (Aldrich) using oxalyl chloride in methylene chloride. After removal

of solvent and excess oxalyl chloride in vacuo, the resulting 2-methylheptanoyl chloride was treated with Meldrum's acid according to a general procedures (Oikawa et al., 1978) to yield the acyl-Meldrum's acid. A solution of *N*-Acetylcysteamine (0.119 g, 1.0 mmol) in toluene (5 ml) was added to a solution of 2-methylheptanoic Meldrum's acid (0.270 g, 1.0 mmol) in toluene (5 ml) in a round-bottom flask equipped with a magnetic stir bar. The reaction mixture was then stirred at reflux condition for 4 hr. The solvent was removed in vacuo, and the crude product was purified by silica column chromatography (EtOAc). Further purification on HPLC using a C₁₈ reverse phase column affords the 4-methyl, 3-oxononanoyl-S-NAC as a waxy solid.

NMR data: **14a**: keto:enol=1.7:1, 1H NMR (500 MHz, CDCl₃) δ 12.61 (s, 0.6H, enol-OH), 6.03 (s, 1.5H, NH), 5.42 (s, 0.6H, enol-COH=CHCO), 3.71 (s, 2H, keto-COCH₂CO), 3.44 (m, 4.1H, NH-CH₂), 3.07 (m, 3.3H, S-CH₂), 2.58 (m, 1H, keto-CH(CH₃)CO), 2.15 (m, 0.6H, enol-CH(CH₃)COH), 1.97 (s, 4.9H, COCH₃), 1.2-1.4 (m, 11.8H, CH(CH₂)₃CH₃), 1.10 (d, 6.9Hz, 1.8H, enol-CH(CH₃)COH), 1.08 (d, 7.0 Hz, 3.4H, keto-CH(CH₃)CO), 0.86 (m, 5.9H, CH₃CH₂); ¹³C NMR (125 MHz, CDCl₃) δ 206.3, 194.6, 192.7, 189.7, 170.8, 170.6, 98.3, 55.5, 47.2, 40.1, 39.6, 39.3, 34.0, 32.4, 29.8, 29.6, 29.33, 29.3, 28.0, 23.34, 23.26, 18.1, 16.0, 14.1, 14.0. **14b**: keto:enol=1.7:1 1H NMR (500 MHz, CDCl₃) δ 12.62 (s, 0.6H, enol-OH), 5.91 (s, 1.6H, NH), 5.43 (s, 0.6H, enol-COH=CHCO), 3.72 (s, 2H, keto-COCH₂CO), 3.46 (m, 4.3H, NH-CH₂), 3.07 (m, 3.2H, S-CH₂), 2.61 (m, 1H, CH(CH₃)CO), 2.19 (m, 0.6H, CH(CH₃)COH), 1.97 (s, 4.8H, COCH₃), 1.2-1.4 (m, 5.3H, CH(CH₂)₂CH₃), 1.11 (d, 6.9Hz, 1.5H, enol-CH(CH₃)COH), 1.09 (d, 7.0 Hz, 3H, keto-CH(CH₃)CO), 0.89 (m, 5.0H, CH₃CH₂); ¹³C NMR (125 MHz, CDCl₃) δ 206.2, 194.6, 192.7, 181.7, 170.6, 170.4, 98.3, 55.5, 51.0, 47.0, 40.1, 39.41, 39.38, 36.5, 34.8, 29.4, 28.0, 23.4, 23.3, 20.6, 20.4, 18.1, 16.0, 14.2, 14.1. **14c**: keto:enol=1.8:1, 1H NMR (500 MHz, CDCl₃) δ 12.58 (s, 0.53H, enol-OH), 6.25 (s, 1.51H, NH), 5.39 (s, 0.56H, enol-COH=CHCO), 3.68 (s, 2H, keto-COCH₂CO), 3.44 (m, 3.34H, NH-CH₂), 3.03 (m, 3.19H, S-CH₂), 2.51 (m, 1H, keto-CH(CH₃)CO), 2.05 (m, 0.6H, enol-CH(CH₃)COH), 1.92 (s, 4.8H, COCH₃), 1.3-1.45 (m, 2H, CHCH₂CH₃), 1.07 (d, 6.9Hz, 1.5H, enol-CH(CH₃)COH), 1.05 (d, 6.9 Hz, 3H, keto-CH(CH₃)CO), 0.84 (m, 5H, CH₃CH₂); ¹³C NMR (125 MHz, CDCl₃) δ 206.1, 194.4, 192.5, 189.7, 181.2, 170.8, 170.6, 98.3, 55.5, 48.5, 41.1, 39.9, 39.2, 29.2, 27.9, 25.6, 23.23, 23.16, 17.6, 15.4, 11.8, 11.5.

Supplementary Tables and Figures

1. Supplementary Tables

Table S1: Primers used in this study

| Primers code | Sequences of primers (5'-3') |
|------------------------|---|
| <i>PsoA seg1f</i> | CTCGACGTACCCCAGTACTGTTGGCGCCGGAAGCTGGATATTCCGGGGATCCGTCGACC |
| <i>PsoA seg1r</i> | CTGTGTCATCTCTTTTGGCTTTCC |
| <i>PsoA seg2f</i> | CGTTGTGCCTCCTGACACACCGCTAC |
| <i>PsoA seg2r</i> | GAGGAGACTCTCGCCAATAGTGCGCAC |
| <i>PsoA seg3f</i> | GACTCCTCGACCAGTGGGCTGACAAG |
| <i>PsoA seg3r</i> | CTCGGGCTCAGGGTAATGCTCTGAG |
| <i>PsoA seg4f</i> | GCTTGC GGAGGCTTTGTCCATTATTC |
| <i>PsoA seg4r</i> | CTCGACATCAGAGGAGGCACTGACAC |
| <i>PsoA seg5f</i> | CACTCATT CAGGTGGTG CAGCGTCAC |
| <i>PsoA seg5r</i> | AAATTTGTCATTTAAATTAGTGATGGTGATGGTGATGCACCTACTCGACATGAGGACCTC |
| <i>PsoF 28af</i> | CCG <u>GAATTC</u> ATGACCAGAATACGTCCAGACTATG (<i>EcoRI</i> site underlined) |
| <i>PsoF 28ar</i> | CCG <u>CTCGAG</u> TCAATGCCATTAGTAGTAACGGAG (<i>XhoI</i> site underlined) |
| <i>PsoF pXW06f</i> | ATCAACTATCAACTATTA ACTATATCGTAATACCATCATATGACCAGAATACGTCCAGAC |
| <i>PsoF pXW06r</i> | TGATAATGGAAACTATAAATCGTGAAGGCATGTTTAAACTTAGCAGCCGGATCTCAGTGG |
| <i>PsoF-MT pXW06f</i> | ATCAACTATCAACTATTA ACTATATCGTAATACCATCATATGCTGGAGTCCGTGCAAATA |
| <i>Pso-MT pXW06r</i> | TTGATAATGGAAACTATAAATCGTGAAGGCATGTTTAAACTCAATGCCATTAGTAGTAA |
| <i>PsoF-FMO pXW06f</i> | ATCAACTATCAACTATTA ACTATATCGTAATACCATCATATGACCAGAATACGTCCAGAC |
| <i>PsoF-FMO pXW06r</i> | TGATAATGGAAACTATAAATCGTGAAGGCATGTTTAAACTCACTCGGGCGTGACATGGAC |
| <i>PsoB 28af</i> | GGTACTC <u>CATATG</u> GCAACGATACACAAGCTGTTC (<i>NdeI</i> site underlined) |
| <i>PsoB 28ar</i> | AGAAT <u>GCGGCCGC</u> CTACAAACGGCTCACCAGATCTTC (<i>NotI</i> site underlined) |
| <i>PsoB pXW02f</i> | GGTACTC <u>CATATG</u> GCAACGATACACAAGCTGTTC (<i>NdeI</i> site underlined) |
| <i>PsoB pXW02r</i> | AGCTTT <u>GTTTAAAC</u> CTACAAACGGCTCACCAGATC (<i>PmeI</i> site underlined) |
| <i>PsoG seg1f</i> | CATATGGCTAGCGATTATAAGGATGATGATGATAAGACTAGTATGATGGCCTCGTTATTC |
| <i>PsoG seg1r</i> | GAGATTTACGAAAGCCCCTCCCGGCAGTCCTCACATCGCCAATAC |
| <i>PsoG seg2f</i> | GTATTGGCGATGTGAGGACTGCCGGGAGGGCTTTTCGTGAAATCTCGG |
| <i>PsoG seg2r</i> | CGCGTCGCATCCTGGCCTTGATGTATCACAATCATTTC AAGAATC |
| <i>PsoG seg3f</i> | GATTCTTGAAATGATTGTGATACATCAAGGCCAGGATGCGACGCGTCC |
| <i>PsoG seg3r</i> | TTTGTCATTTAAATTAGTGATGGTGATGGTGATGCACGTGCTAACACTTCCAGAGTGTTA |

2. Supplementary Figures

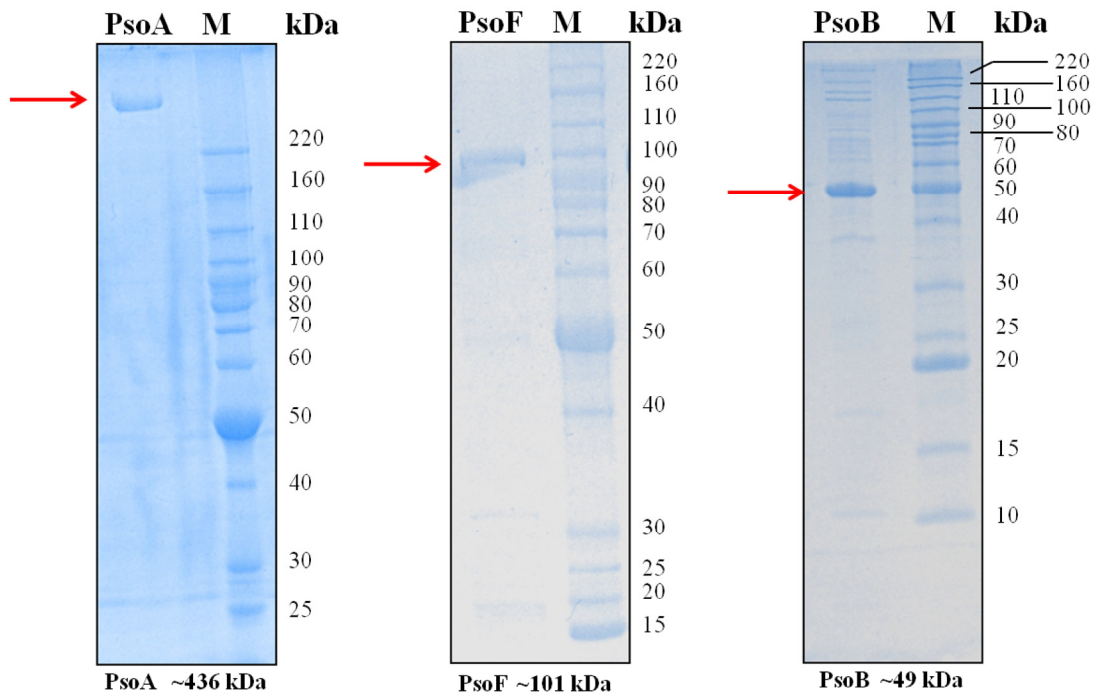


Figure S1. SDS PAGE of the heterogeneously expressed proteins in this study. PsoA is with C-terminal His₆-Tag, PsoF and PsoB are with N-terminal His₆-Tag. All these enzymes were purified by Ni-NTA agarose affinity resin. PsoA (~436 kDa) was expressed from *S. cerevisiae* BJ5464-NpgA; PsoF (~101 kDa) was expressed from *E. coli* BL21; PsoB (~49 kDa) was expressed from *E. coli* BL21.

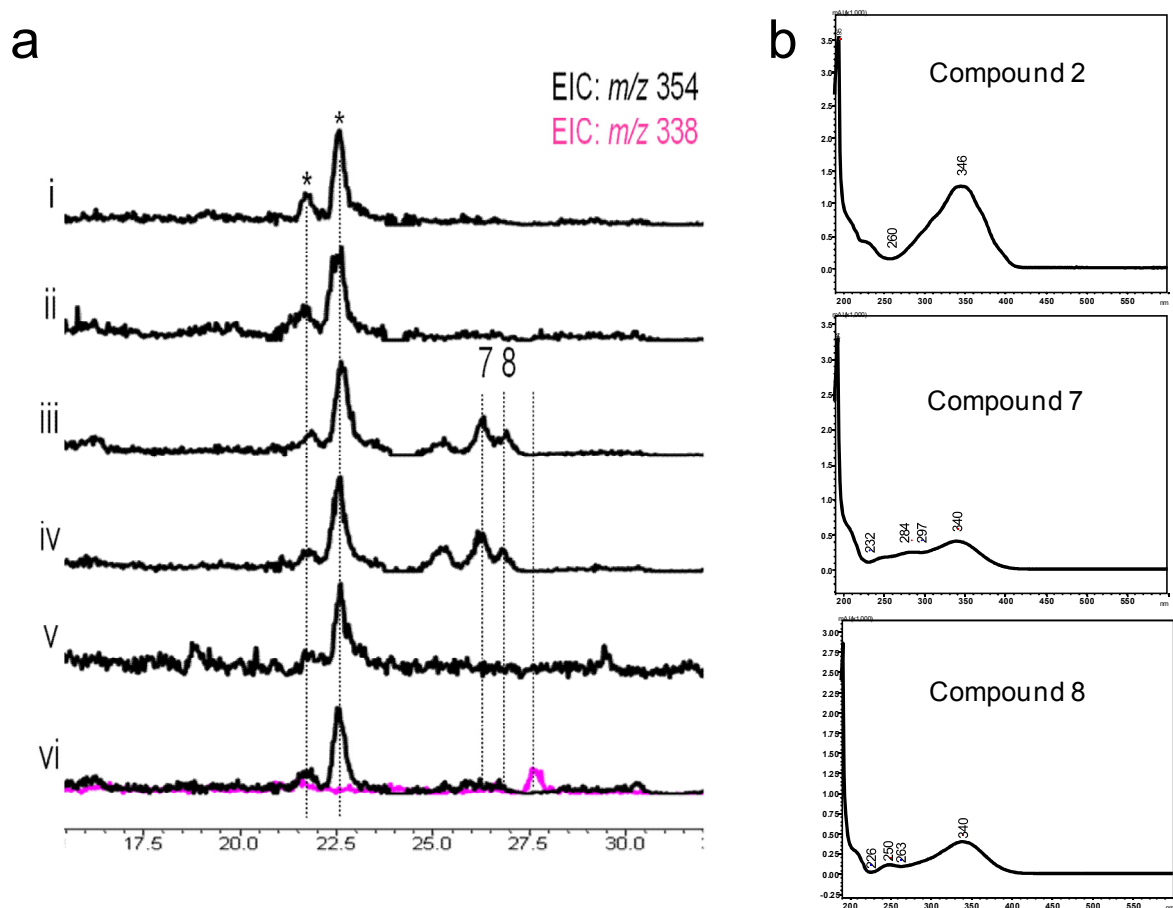


Figure S2. a. ESI-MS analyses of heterogeneous production of azaspirene pathway intermediates in *S.c* BJ5464-NpgA. i, *S.c* BJ5464-NpgA control; ii, *S.c* BJ5464-NpgA transformed with PsoA; iii, *S.c* BJ5464-NpgA transformed with PsoA and PsoF; iv, *S.c* BJ5464-NpgA transformed with PsoA, PosF and PsoB; v, *S.c* BJ5464-NpgA transformed with PsoA, PsoF-FMO and PsoB. vi, *S.c* BJ5464-NpgA transformed with PsoA, PsoF-MT and PsoB. * indicates that compounds with m/z 354 $[M+H]^+$ are produced by *S.c* BJ5464-NpgA itself. Compound with m/z 338 $[M+H]^+$ in (vi) is proposed to be the biosynthetic intermediate of azaspirene (**2**), which is shown in **Figure 1** of main text.

b. UV spectra of azaspirene (**2**), compound **7** and **8**.

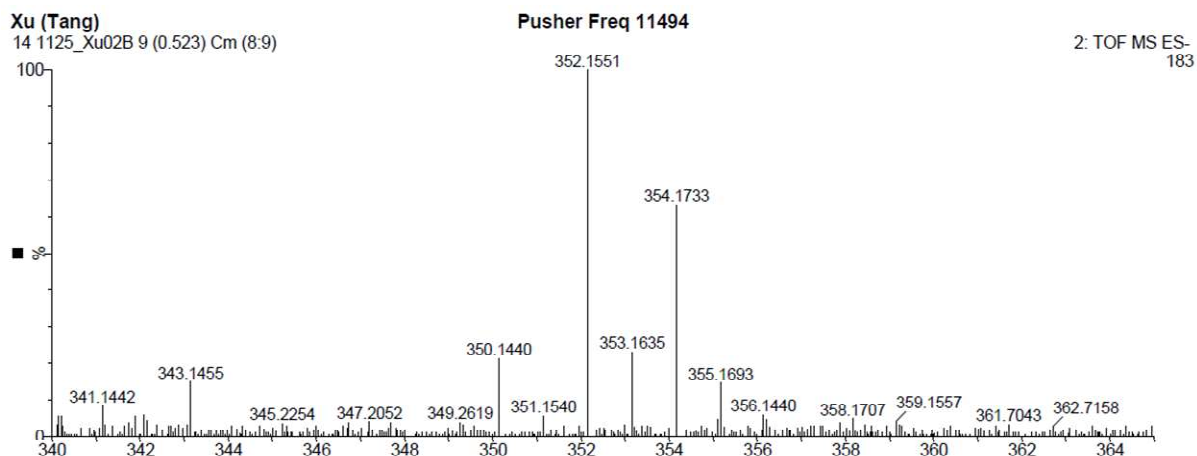
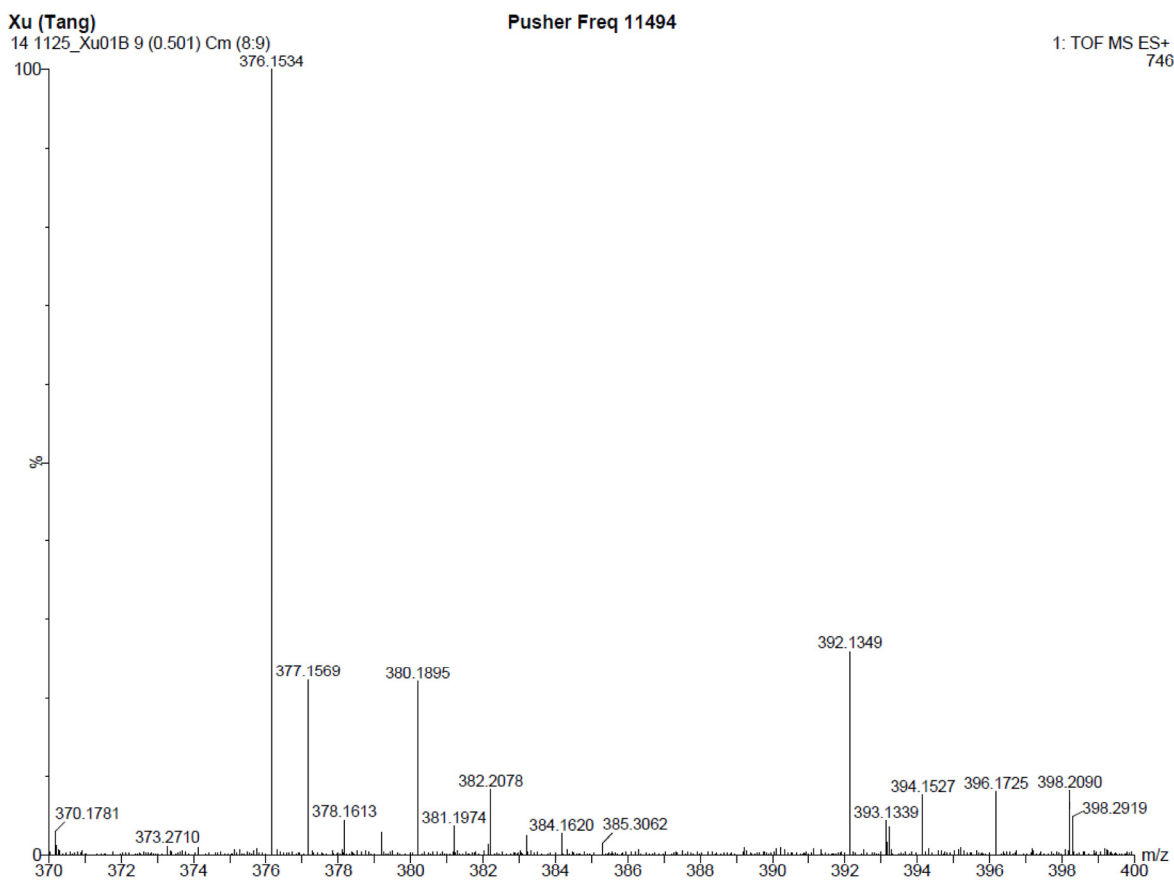


Figure S3. HR-MS analyses of compound **7** and **8**, compound **7** (obsd m/z 376.1534, calcd m/z 376.1525 for $C_{21}H_{23}NO_4Na$, 2.4 ppm), compound **8** (obsd m/z 352.1551, calcd m/z 352.1549 for $C_{21}H_{22}NO_4^-$, 0.6 ppm). The data were collected on the Waters LCT Premier with ACQUITY UPLC at the UCLA Molecular Instrumentation Center.

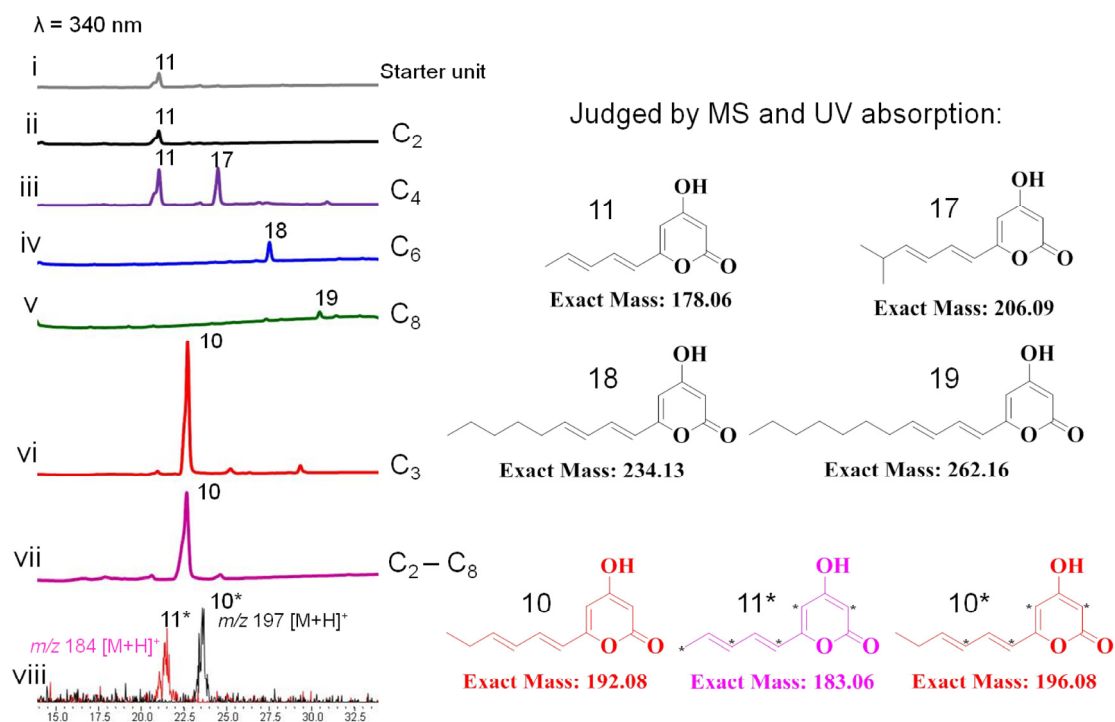
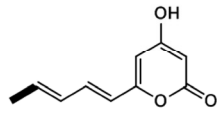
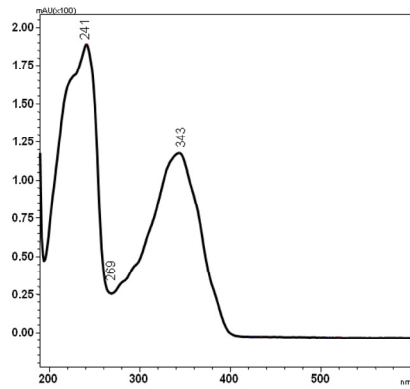
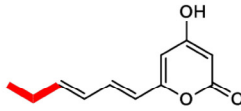
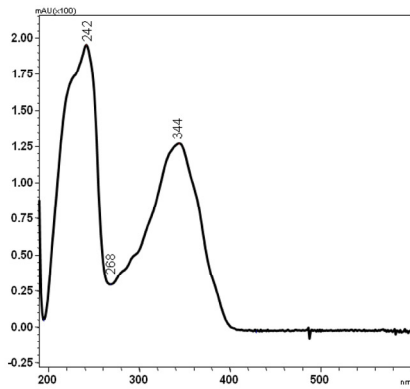
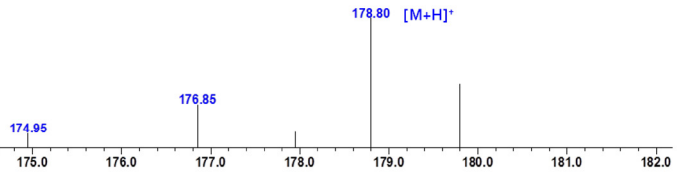


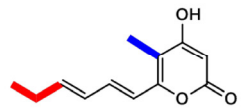
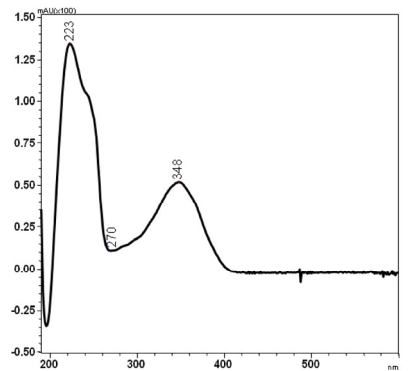
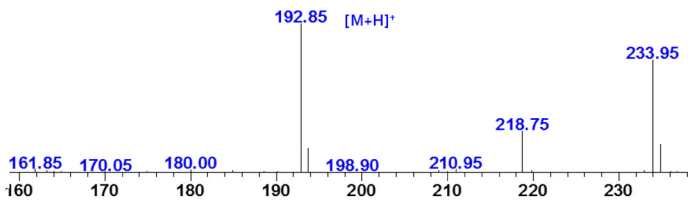
Figure S4. LC-MS analyses of PsoA HR-PKS module activity with different starter units. i, 1 mM malonyl-CoA, 4 mM NADPH, 12 μM PsoA; ii, (i)+0.5 mM acetyl-CoA; iii, (i)+0.5 mM isobutyryl-CoA; iv, (i)+0.5 mM hextanoyl-CoA; v, (i)+0.5 mM octanoyl-CoA; vi, (i)+0.5 mM propionyl-CoA and vii, (i)+equal amounts (0.5 mM) of C₂-C₈ CoAs; viii, 4 mM NADPH, 12 μM PsoA, 0.5 mM propionyl-CoA, 100 mM [2-¹³C] malonate, 25 μM MatB, 20 mM CoA, and 25 mM ATP. * in compound **10*** and **11*** indicates the incorporation of [2-¹³C] malonate.



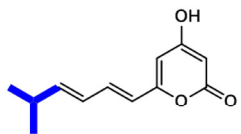
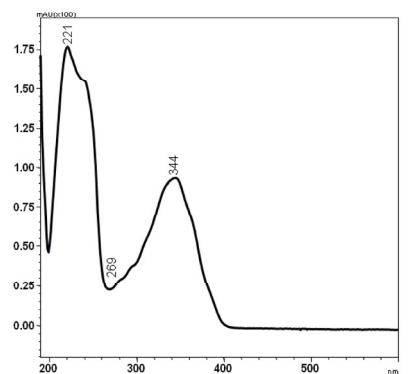
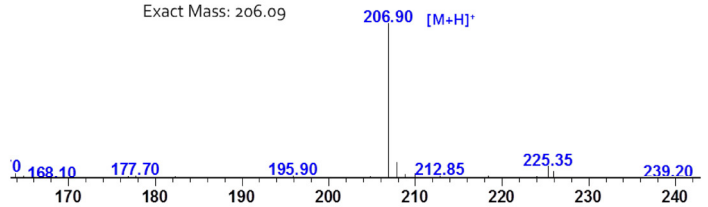
Chemical Formula: $C_{10}H_{10}O_3$
Exact Mass: 178.06



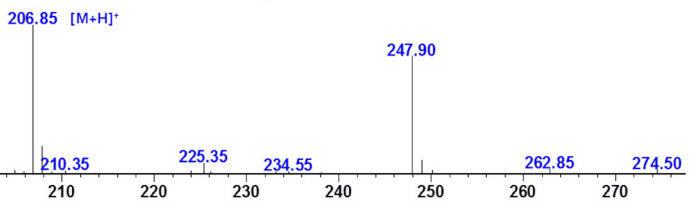
Chemical Formula: $C_{11}H_{12}O_3$
Exact Mass: 192.08



Chemical Formula: $C_{12}H_{14}O_3$
Exact Mass: 206.09



Chemical Formula: $C_{12}H_{14}O_3$
Exact Mass: 206.09



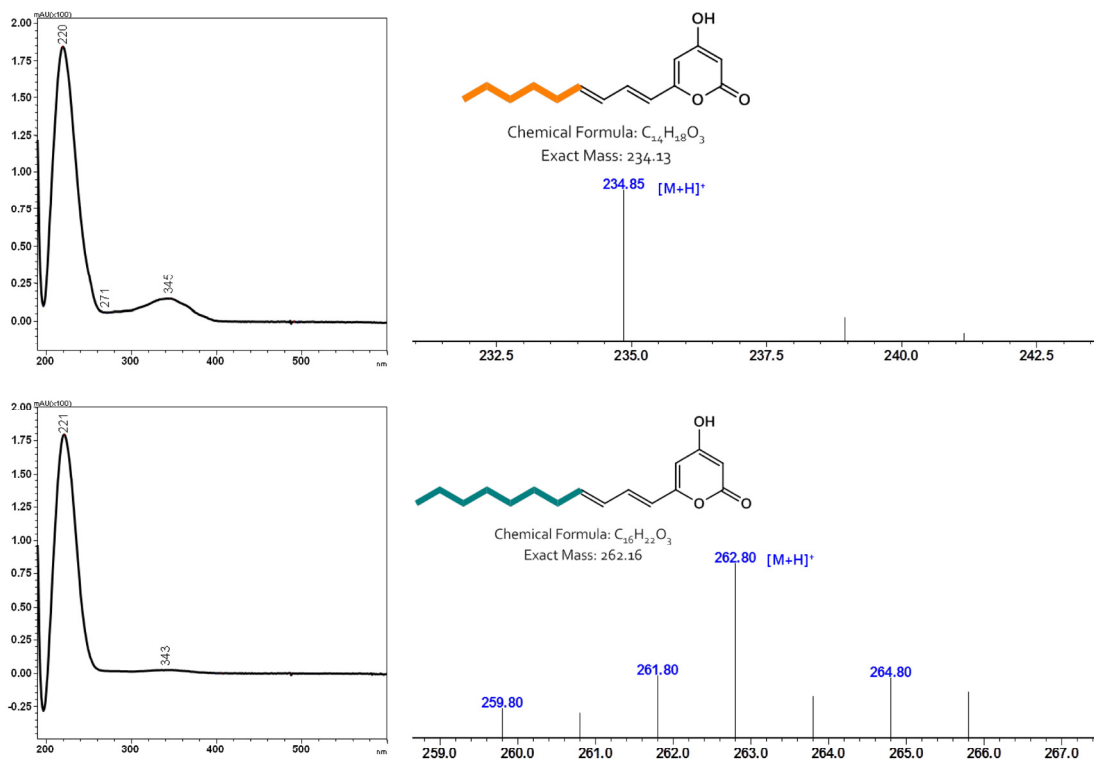


Figure S5. LC-MS analyses of the reaction products of PsoA with various CoAs.

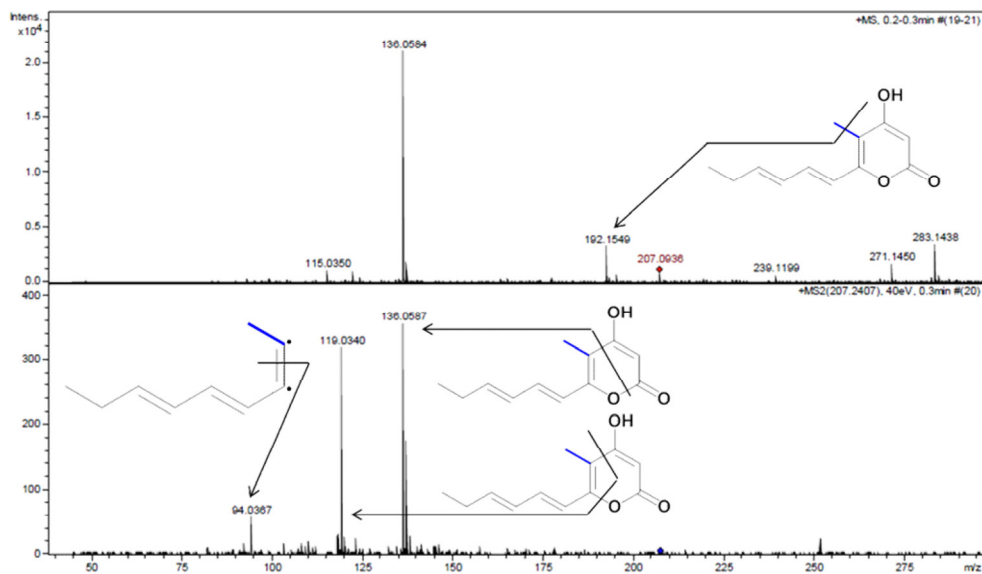


Figure S6. MS-MS analysis of 4-methyl pyrone.

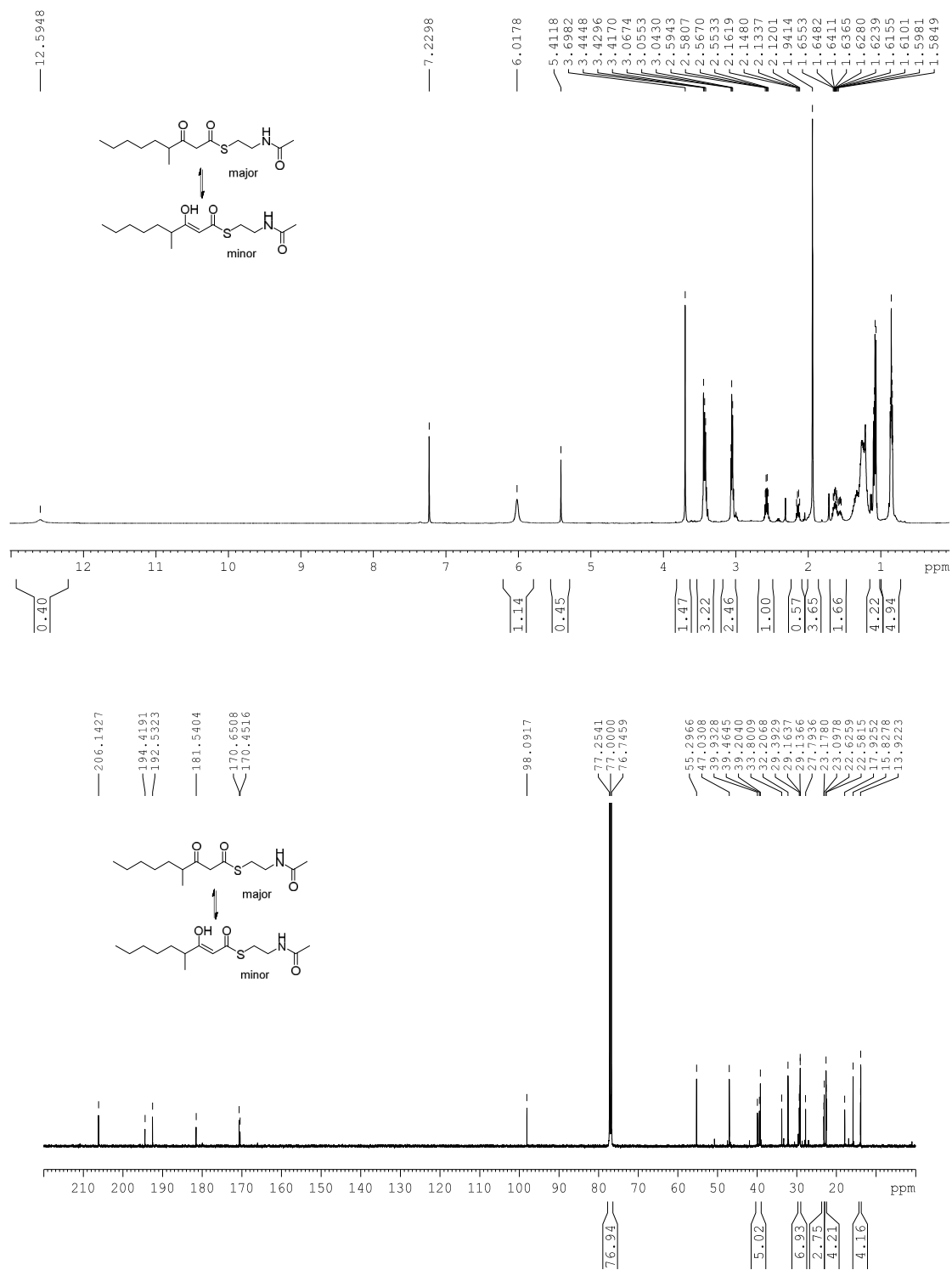


Figure S7. ¹H (500 MHz) and ¹³C (125 MHz) NMR spectra of **14a** (CDCl₃).

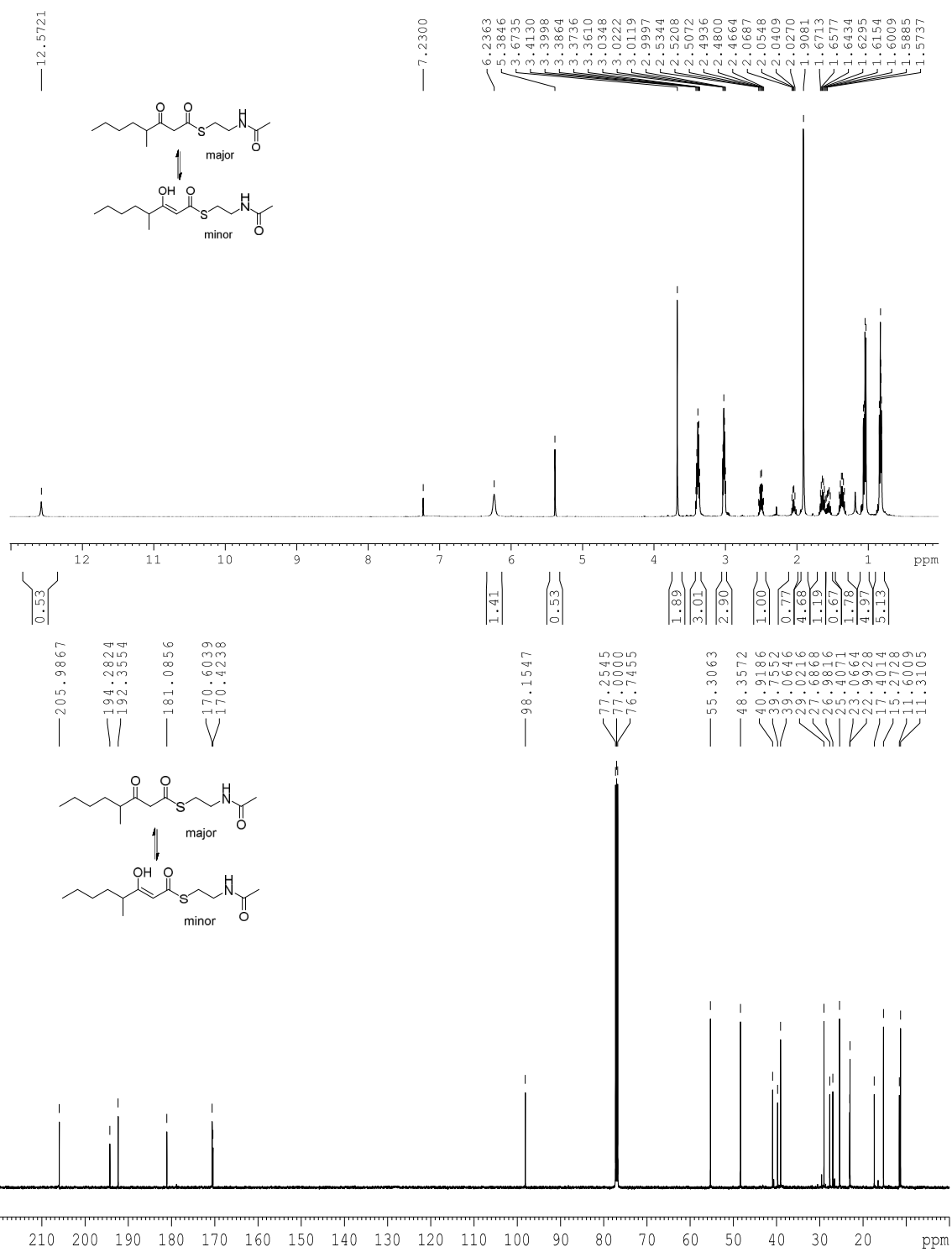


Figure S8. ¹H (500 MHz) and ¹³C (125 MHz) NMR spectra of **14b** (CDCl₃).

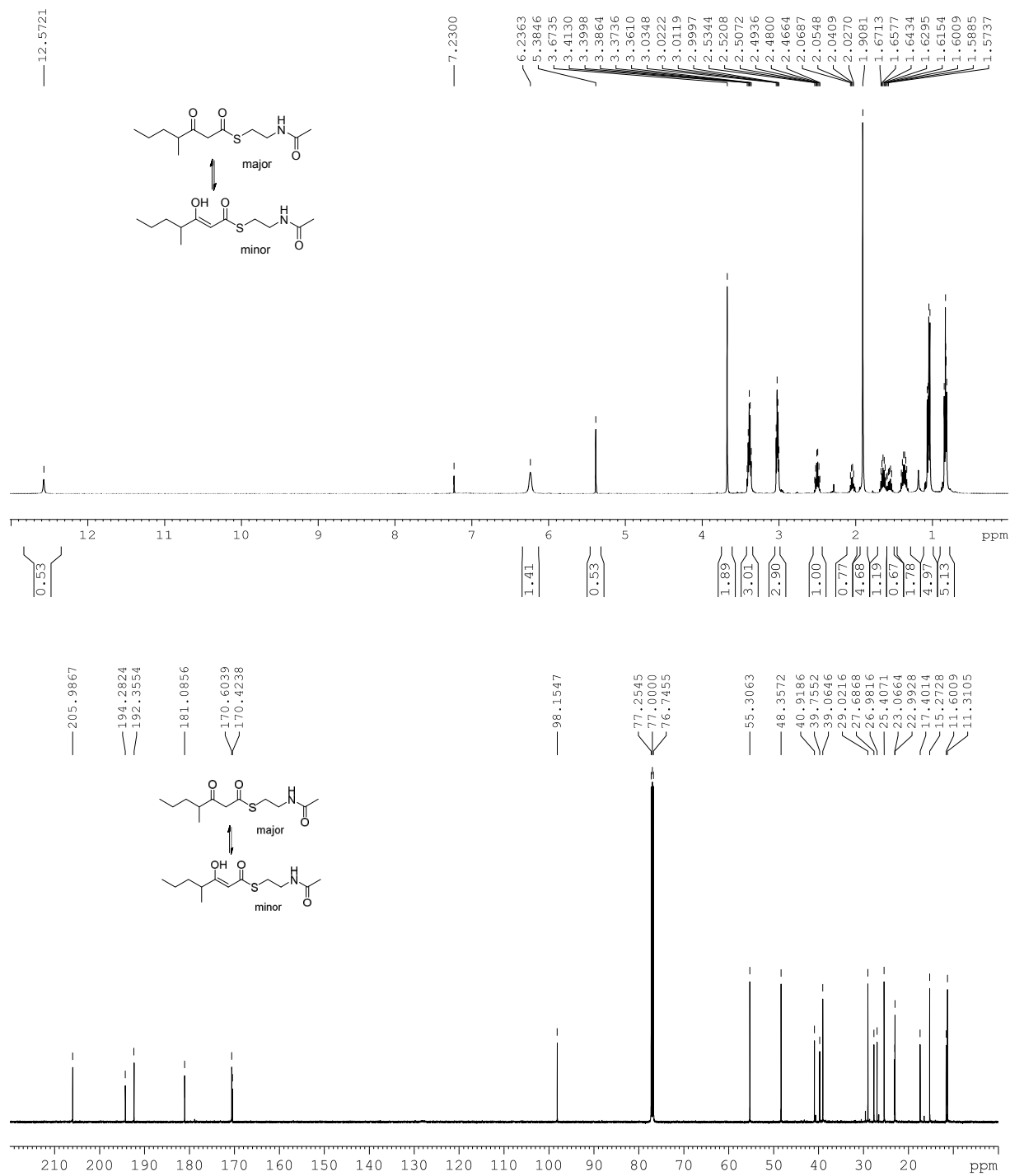


Figure S9. ¹H (500 MHz) and ¹³C (125 MHz) NMR spectra of **14c** (CDCl₃).

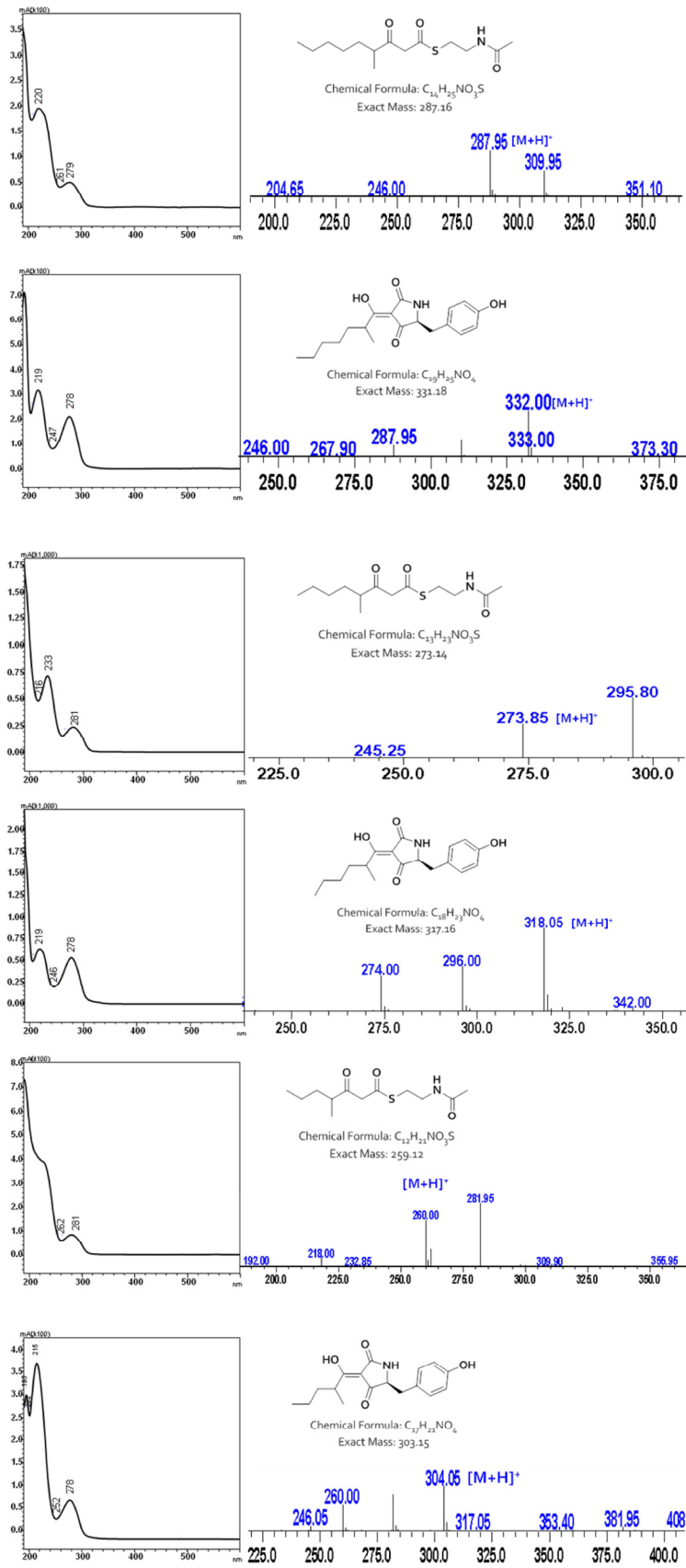


Figure S10. LC-MS analyses of 14a-c and 15a-c.

References

- [1] a)K. K. Lee, N. A. Da Silva, J. T. Kealey, *Anal Biochem* **2009**, *394*, 75; b)S. M. Ma, J. W. Li, J. W. Choi, H. Zhou, K. K. Lee, V. A. Moorthie, X. Xie, J. T. Kealey, N. A. Da Silva, J. C. Vederas, Y. Tang, *Science* **2009**, *326*, 589.
- [2] a)S. C. Mutka, S. M. Bondi, J. R. Carney, N. A. Da Silva, J. T. Kealey, *FEMS Yeast Res* **2006**, *6*, 40; b)W. Xu, X. Cai, M. E. Jung, Y. Tang, *J Am Chem Soc* **2010**, *132*, 13604.
- [3] J. H. An, Y. S. Kim, *Eur J Biochem* **1998**, *257*, 395.