CHEMBIOCHEM

Supporting Information

CHEMBIOCHEM

Supporting Information

for

Substrate Profile Analysis and ACP-Mediated Acyl Transfer in *Streptomyces coelicolor* Type III Polyketide Synthases

Sabine Grüschow, Tonia J. Buchholz, Wolfgang Seufert,
Jonathan S. Dordick, and David H. Sherman*

DNA manipulations: *S. coelicolor* genomic DNA was isolated using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI) according to the manufacturer's instructions. PCR was carried out on an iCycler Thermal Cycler (Bio-Rad, Hercules, CA) with *Pfu* Turbo DNA polymerase (Stratagene, La Jolla, CA). Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Beverly, CA) and used according to the manufacturer's recommendations. DNA sequencing was performed by the University of Michigan DNA Sequencing Core facility using an ABI Model 3730 sequencer.

Expression and purification of recombinant type III PKSs: Open reading frame (ORF) *sco*7221 was PCR-amplified from *S. coelicolor* genomic DNA using the primers 5'-aatcatatggccgcatacctgtg-3' and 5'aataagcttctacagccactccccttc-3' and ORF *sco*7671 was amplified using the primers 5'-aatcatatgaacgacatgcgtgtcc-3' and 5'-aataagctttcaccagcgcagcagga-3'. The PCR product was sub-cloned into pGem®-T Easy (Promega, Madison, WI) and the insert verified by sequencing. The *Nde*I/*Hind*III fragment containing *sco*7221 or *sco*7671 was transferred into linearized pET-28b(+) (Novagen, Madison, WI)

to give plasmids pDHS10019 and pDHS6596, respectively, which allow expression of SCO7221 and SCO7671 as N-terminal His_6-fusion proteins. BL21(DE3) harboring pDHS10019 was cultured in LB containing 50 μg/mL kanamycin until the OD_{600} reached 0.7. Protein expression was induced with 0.1 mM IPTG and culturing continued for 16-20 h at 18°C. Cells were harvested by centrifugation and resuspended in lysis buffer (250 mM NaCl, 50 mM sodium phosphate, 10 mM imidazole, pH 7.9). Cells were lysed by sonication and the cleared supernatant was loaded onto Ni-NTA resin. The resin was washed first with lysis buffer and then with lysis buffer containing 50 mM imidazole. SCO7221 was eluted with lysis buffer containing 0.2 M imidazole. Protein containing fractions were concentrated using Centriprep® YM-10 centrifugal filter units (Millipore, Billerica, MA) before SCO7221 was transferred into storage buffer (20% glycerol, 250 mM NaCl, 50 mM Na-phosphate, 1 mM TCEP, pH 7.4) using PD-10 columns (Amersham Biosciences). Aliquots of the enzymes were flash-frozen and stored at –80°C until needed. Approximately 50 mg SCO7221 could be purified from a 1 L culture. SCO7671 was expressed and purified in the same way except that the Ni-NTA resin was washed with lysis buffer containing 20 mM imidazole instead of 50 mM imidazole. Between 5-10 mg SCO7671 could be purified from a 1 L culture. Protein concentrations were determined using Bio-Rad Protein Assay with BSA as the standard. Proteins were estimated to be greater than 95% pure (SCO7221/Gcs) and greater than 90% pure (SCO7671) based on SDS-PAGE (Figure S1). No noticeable differences between batches of SCO7671 were noticed despite slight changes in purity level. Size exclusion chromatography of Gcs in 20 mM HEPES, pH 7.4, 150 mM Nacl, 10% glycerol was performed on an AKTA-explorer FPLC (GE Healthcare Life Sciences, Piscataway, NJ) outfitted with a HiLoad 16/60 S200 column. Calibration of the column was performed with molecular weight markers from Sigma-Aldrich (St. Louis, MO). Gcs eluted as a single peak at 74 mLs which is consistent with a dimeric complex in solution (Figure S3).

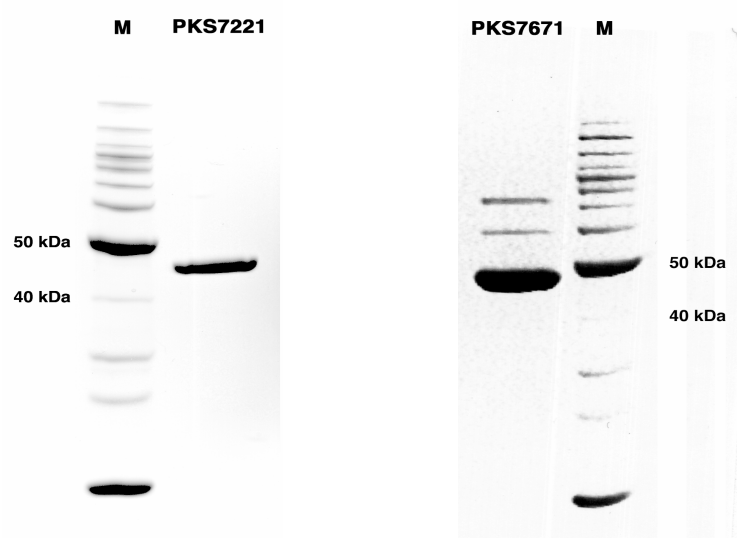
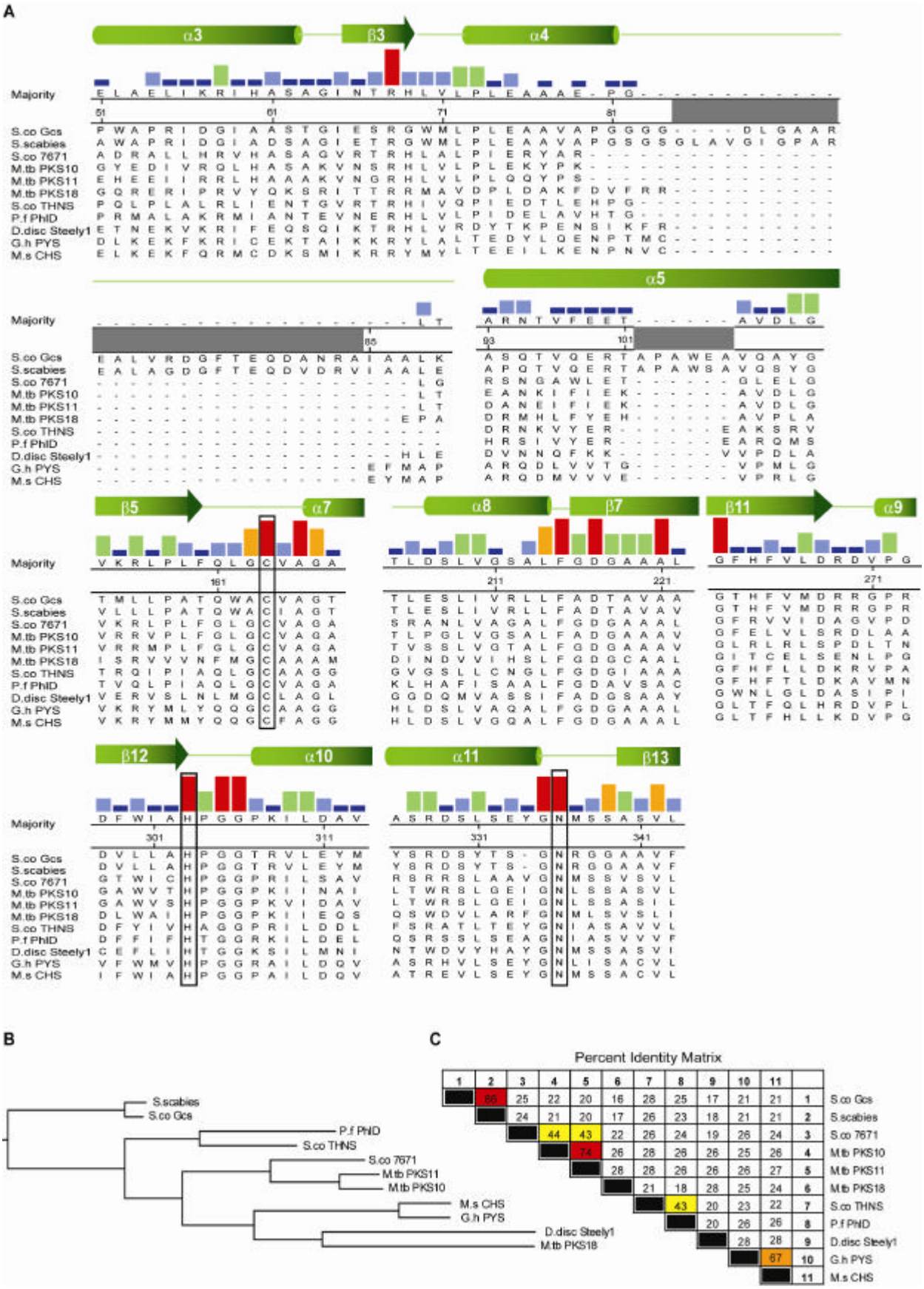


Figure S1: 4-20% Tris Glycine SDS-PAGE gels. Gel 1: molecular weight marker and PKS7221/ Gcs; Gel 2: PKS7671/SCO7671 and molecular weight marker.

Figure S2 (over): Sequence comparisons among selected prokaryotic and eukaryotic type III polyketide synthases. A) Sequence alignments are shown for selected residues. Numbering and secondary structure assignment is based on Chalcone Synthase (CHS) from *Medicago sativa* (alfalfa). Catalytic triad residues are boxed. Similarity bars above the majority sequence are colored from red to blue for each position; red represents 100% identity. B) Cladogram of selected type III PKSs. C) Percent identities above 70% are highlighted in red, those between 50-69% in orange, and identities between 30-49% are colored yellow. Alignments, cladogram and percent identities were calculated using the ClustalW method using Lasergene MegAlign software from DNASTAR. Sequences: S.co Gcs, S.co 7671, S.co THNS = *Streptomyces coelicolor* genes *sco7221, sco7671, sco1206*, www.sanger.ac.uk; S.scabies = *Streptomyces scabies*, www.sanger.ac.uk; M.tb PKS10, M.tb PKS11, M.tb PKS18 = *Mycobacterium tuberculosis* CHS-like proteins, accessions CAB06631, CAB09101, A70958; P.f PhlD = *Pseudomonas fluorescens* PhlD, accession AAB48106; D.disc Steely1 = *Dictyostelium discoideum* C-terminal Type III PKS domain of Steely1, pdb 2H84; G.h PYS = *Gerbera hybrida* (daisy) 2-pyrone synthase, accession CAA86219; M.s CHS = *Medicago sativa* (alfalfa) chalcone synthase, accession P30074.


A
α3
β3
α4
α5
β5
α7
α8
β7
β11
α9
β12
α10
α11
β13
Majority
S.co Gcs
S.scabies
S.co 7671
M.tb PKS10
M.tb PKS11
M.tb PKS18
S.co THNS
P.f PhlD
D.disc Steely1
G.h PYS
M.s CHS
B
C
Percent Identity Matrix

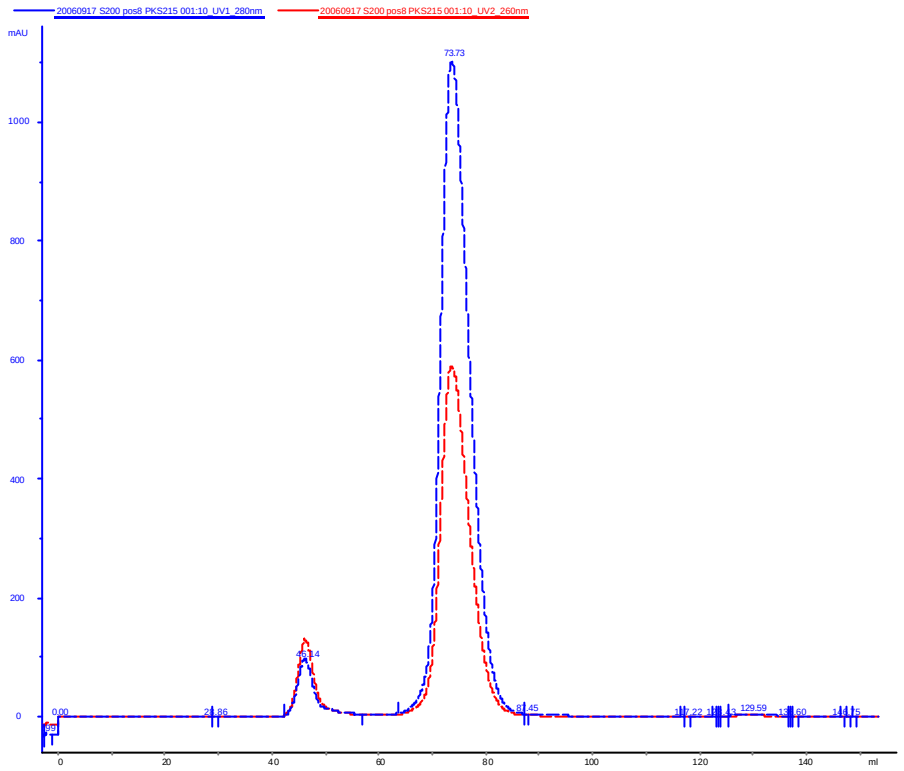


Figure S3: Size Exclusion Chromatography of Gcs in 20 mM HEPES, pH 7.4, 150 mM NaCl, 10% glycerol. The peak around 45 mLs represents material eluting near the void volume. The major protein peak eluted at 74 mLs. To calculate the apparent molecular weight, this elution volume was fit to the equation, mLs = -12.447(LnMW) + 129.52 (MW in kDaltons). An apparent molecular weight of 90,000 Daltons is consistent with a dimer for the His-tagged protein whose calculated molecular weight is 44 000 Daltons.

Expression and purification of recombinant ACPs: Expression plasmids pET-TcmM and pH8-Otc were obtained from Michael Burkart (University of California, San Diego). Expression plasmids for MmcB and scoACP were obtained through PCR amplification of the gene from genomic DNA, followed by enzymatic restriction and ligation into pET-28b(+). The following primers were used for PCR amplification (restriction enzyme recognition sites are shown in italics): gtatttctg*ccatgg*agaccctgacgac and aatat*ctcgag*ttcggcgggtgcggtgacgccgttc for amplification of *mmcB*, gattagc*catatg*ccgtccaccgccgacgaacgaca and tacc*aagctt*cacgcggtgccgccggcgg for amplification of *sco*0549. MmcB and TcmM were expressed with C-terminal His_6-tag, Otc with N-terminal His_8-tag, and scoACP with N-terminal His_6-tag. Active-site mutants of the ACPs (ACP S→A) were obtained through

primer-directed mutagenesis according to the Stratagene QuikChange protocol. The following primers were used (changed bases are shown in bold, amino acid numbering refers to wild-type sequence): gctgggcctggac**g**cgatcaacaccatc and gatggtgttgatcg**c**gtccaggcccagc for the MmcB S43A expression plasmid, ctcggctacgac**g**cgctggcgctgc and gcacgccagcg**c**gtcgtagccgag for the Otc S41A expression plasmid, cctcggctacgac**g**ccatcgccctgc and gcagggcgatgg**c**gtcgtagccgagg for the TcmM S41A expression plasmid, cctgggcgtcgac**g**cgctggcccggct and agccgggccagcg**c**gtcgacgcccagg for the scoACP S44A expression plasmid. BL21 (DE3) was used as the expression host except for TcmM where BL21 (DE3) CodonPlus RP cells were used. Culturing, induction of protein expression and purification was carried out essentially as described for the type III PKSs.

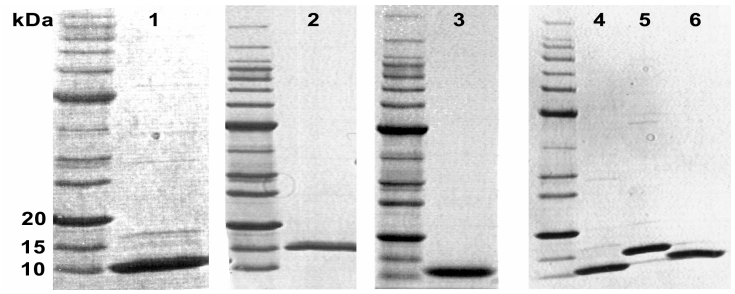


Figure S4: 4-20% Tris Glycine SDS-PAGE gels. Lane 1: *apo*-MmcB, lane 2: *apo*-Otc, lane 3: *apo*-TcmM, lane 4: MmcB S43A, lane 5: Otc S41A, lane 6: TcmM S41A.

Analysis of ACP acylation: Hexanoyl-ACPs were analyzed by electrospray mass spectrometry using a ThermoFinnigan LTQ linear ion trap instrument (capillary temperature 250 °C, capillary voltage 32V, tube lens 95 V). Mass spectra were deconvoluted using ProMass™ for Xcalibur (Novatia, Princeton).

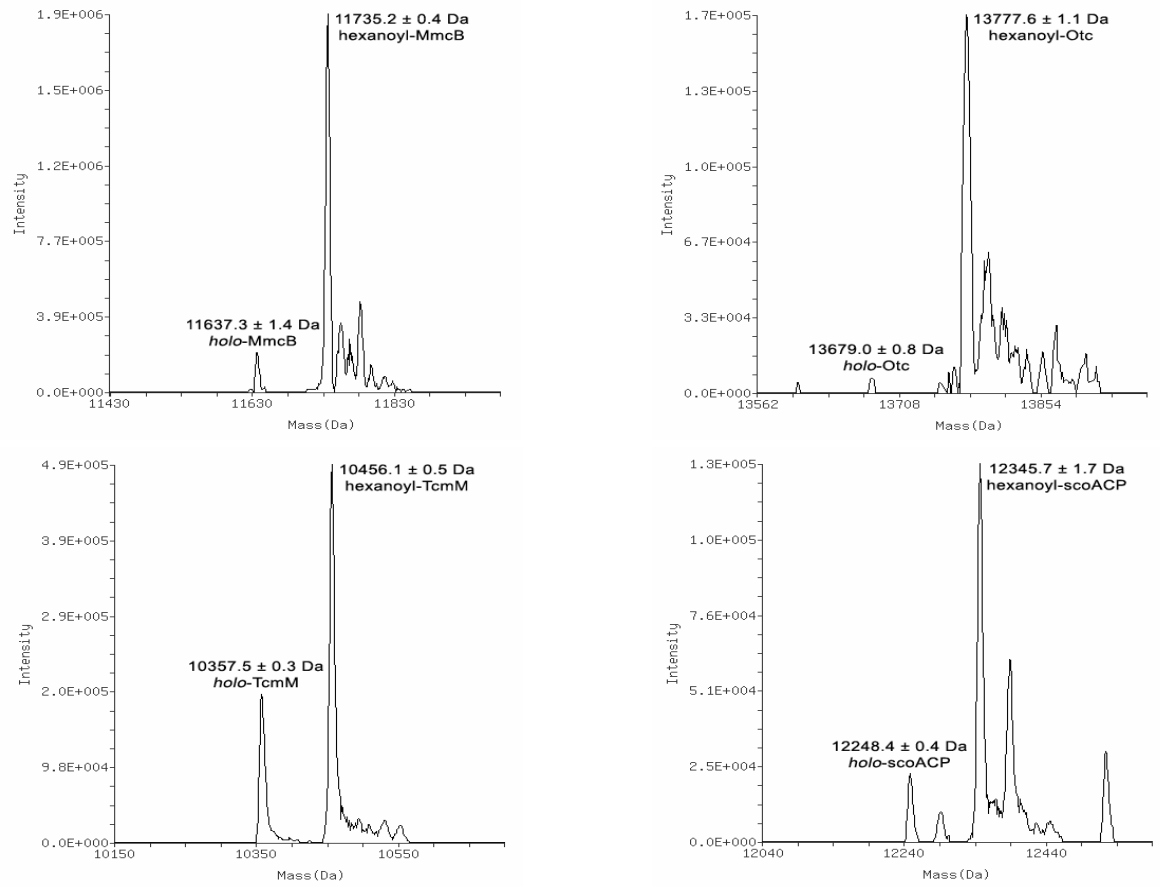


Figure S5: Mass spectrometry of *in vitro* acylated wild-type ACPs. Calculated molecular weights: 11637.1 Da (*holo*-MmcB), 11735.2 Da (hexanoyl-MmcB), 13679.2 Da (*holo*-Otc), 13777.6 Da (hexanoyl-Otc), 10357.5 Da (*holo*-TcmM), 10455.6 Da (hexanoyl-TcmM), 12248.4 Da (*holo*-scoACP), 12346.5 Da (hexanoyl-scoACP); *apo*-ACPs were not detected. Percentage of hexanoyl-ACP: 91% (MmcB), 96% (Otc), 72% (TcmM), 85% (scoACP).

Analysis of reaction products: 5-10 µM type III PKS, 0.2 mM acyl-CoA, 0.5 mM malonyl-CoA in 0.1 M HEPES (pH 7.4) were incubated at 30 °C for 2-5 h. The reaction was stopped with concentrated HCl and extracted with two volumes of ethyl acetate. The products were purified by reversed phase HPLC (Alltech Econosil C18, 10 µm, 10 x 250 mm, flow rate: 3 ml/min, solvent A: 0.1% formic acid or trifluoroacetic acid in ddH_2O, solvent B: 0.1% formic acid or trifluoroacetic acid in acetonitrile) using either of the following gradients: 1) 5% B for 2 min, 5-70% B in 20 min, 70% B for 5 min, 70-5% B in 2 min, 5% B for 10 min; 2) 5% B for 2 min, 40-100% B in 30 min, 100-40% B in 2 min,

40% B for 10 min; 3) 5% B for 2 min, 5-80% B in 20 min, 80% B for 5 min, 80-5% B in 2 min, 5% B for 10 min. Elution of products was monitored using a photodiode array detector (Beckman Coulter, System Gold® 168 Detector, Fullerton, CA). Product containing fractions were concentrated *in vacuo* and the water was removed by freeze-drying. NMR spectra were recorded on a Bruker DPX300 or DRX500 using [D_6]DMSO as solvent. GC-MS analysis was performed by James Windak (Department of Chemistry, University of Michigan) on a Finnigan Trace GC 2000 with a DB-1 capillary column installed in a split injector (split ratio 1:100). Electrospray ionization mass spectrometry (ESI-MS) was performed on a ThermoFinnigan LCQ by direct injection or on a ThermoFinnigan LTQ instrument for LC-MS using a Waters XBridge™ (C18 3.5 µm, 2.1 x 150 mm) column and the following gradient (solvent A: 0.1% formic acid in water, solvent B: 0.1% formic acid in acetonitrile): 5% B for 1.2 min, 5-10% B in 6 min, 10-70% B in 12 min (gradient 4).

Analytical data for polyketide products:

4-Hydroxy-6-methyl-2-pyrone (**2a**)[1]: ^{1}H NMR (500 MHz, d_6-DMSO): δ = 1.93 (s, 3H), 4.23 (s, 1H), 5.28 (s, 1H), 10.56 (s, 1H). MS (GC-MS, EI, +ve): *m/z* = 126.0 ($M^{\bullet+}$, 30), 111.0 (14), 98.0 (79), 85.0 (27), 69.0 (79), 55.2 (19), 43.3 (100); calculated for $C_6H_6O_3$ 126.0. HPLC (gradient 1): t_R = 13.0 min. UV-vis: λ_{max} = 286 nm.

4-Hydroxy-6-ethyl-2-pyrone (**2c**)[1]: MS (ESI, +ve): *m/z* = 141.2 ($M+H^+$); MS/MS: *m/z* = 141.1 (100), 113.0 (15), 98.8 (8); calculated for $C_7H_8O_3$ 140.1. HPLC (Gradient 1): t_R = 15.7 min. UV-vis: λ_{max} = 286 nm.

4-Hydroxy-6-propyl-2-pyrone (**2d**)[1]: ^{1}H NMR (500 MHz, d_6-DMSO): δ = 0.90 (t, *J* = 7.3, 3H), 1.52 (dt, *J* = 7.3, 7.5, 2H), 2.18 (t, *J* = 7.5, 2H), 4.23 (s, 1H), 5.26 (s, 1H). MS (GC-MS, EI, +ve): *m/z* = 154.1 ($M^{\bullet+}$,39), 139.1 (1), 126.1 (48), 111.0 (96), 97.1 (21), 84.1 (39), 69.0 (100), 55.0 (19), 43.0 (29); calculated for $C_8H_{10}O_3$ 154.1. HPLC (gradient 1): t_R = 17.8 min. UV-vis: λ_{max} = 286 nm.

4-Hydroxy-6-pentyl-2-pyrone (**2e**)[1,2]: ^{1}H NMR (500 MHz, d_6-DMSO): δ = 0.88 (t, *J* = 6.8, 3H), 1.25-1.32 (m, 4H), 1.51 (quint, *J* = 7.3, 2H), 2.24 (t, *J* = 7.4, 2H), 4.40 (s, 1H), 5.39 (s, 1H), 11.67 (s, 1H). MS (GC-MS, EI, +ve): *m/z* = 182.1 ($M^{\bullet+}$,17), 139.0 (11), 126.0 (85), 111.0 (86), 98.0 (72), 84.0 (77), 69.0 (100), 55.2 (41), 43.4 (35); calculated for

$C_{10}H_{14}O_3$ 182.1. MS (ESI, +ve): *m/z* = 183.1 ($M+H^+$); MS/MS: *m/z* = 183 (92), 165 (48), 155 (52), 126 (100), 99 (70). HPLC (gradient 1): t_R = 22.1 min. UV-vis: λ_{max} = 286 nm.

4-Hydroxy-6-heptyl-2-pyrone (**2f**)[1,2]: MS (ESI, +ve): *m/z* = 211.3 ($M+H^+$); MS/MS: *m/z* = 211.2 (100), 183.2 (24), 127.0 (8); calculated for $C_{12}H_{18}O_3$ 210.1. HPLC (Gradient 2): t_R = 22.0 min. UV-vis: λ_{max} = 286 nm.

4-Hydroxy-6-nonyl-2-pyrone (**2g**)[2]: MS (ESI, +ve): *m/z* = 239.4 ($M+H^+$); MS/MS: *m/z* = 239.2 (100), 211.1 (30), 197.2 (24), 183.2 (18), 127.0 (8); calculated for $C_{14}H_{22}O_3$ 238.2. HPLC (Gradient 2): t_R = 25.8 min. UV-vis: λ_{max} = 286 nm.

4-Hydroxy-6-phenyl-2-pyrone (**2i**)[3]: ^{1}H NMR (500 MHz, d_6-DMSO): δ = 5.42 (s, 1H), 6.79 (s, 1H), 7.54 (os, 3H), 7.86-7.88 (m, 2H). MS (ESI, +ve): *m/z* = 189.1 ($M+H^+$); MS/MS: *m/z* = 189.3 (5), 171.1 (40), 161.3 (90), 133.3 (30), 111.1 (25), 105.2 (100); calculated for $C_{11}H_8O_3$ 188.1. HPLC (gradient 3): t_R = 18.6 min. UV-vis: λ_{max} = 320 nm.

4-Hydroxy-6-*iso*-propyl-2-pyrone (**2j**)[3]: ^{1}H NMR (500 MHz, d_6-DMSO): δ = 1.08 (d, *J* = 6.8, 6H), 2.56 (os, 1H), 4.20 (s, 1H), 5.23 (s, 1H). MS (GC-MS, EI, +ve): *m/z* = 154.1 ($M^{\bullet+}$,17), 126.1 (8), 111.1 (100), 69.0 (95), 55.0 (8), 43.1 (23.1); calculated for $C_8H_{10}O_3$ 154.1. HPLC (gradient 1): t_R = 17.5 min. UV-vis: λ_{max} = 286 nm.

4-Hydroxy-6-(2'-hydroxypropyl)-2-pyrone (**2m**): LC-MS (gradient 4): *m/z* = 171.1 ($M+H^+$); MS/MS: *m/z* = 111.2 (10), 127.1 (30), 153.1 (100); calculated for $C_8H_{10}O_4$ 170.1; t_R = 12.4 min, λ_{max} = 285 nm.

4-Hydroxy-6-(2'-hydroxy-1'-methylpropyl)-2-pyrone (**2p**): LC-MS (gradient 4): *m/z* = 198.1 ($M+H^+$); MS^2: *m/z* = 181.2; MS^3: *m/z* = 111.2 (100), 139.2 (40); t_R = 14.2 min, λ_{max} = 285 nm.

References for known compounds

[1] I. Abe, T. Watanabe, H. Noguchi, *Phyochemistry* **2004**, *65*, 2447-2453.

[2] J. C. Jeong, A. Srinivasan, S. Grüschow, H. Bach, D. H. Sherman, J. S. Dordick, *J. Am. Chem. Soc.* **2005**, *127*, 64-65.

[3] A. R. Katritzky, Z. Wang, M. Wang, C. D. Hall, K. Suzuki, *J. Org. Chem.* **2005**, *70*, 4854.