Electronic Supplementary Information

Analysis of the cercosporin polyketide synthase CTB1 reveals a new fungal thioesterase function.

Adam G. Newman, Anna L. Vagstad, Katherine Belecki, Jonathan R. Scheerer and Craig A. Townsend*

Department of Chemistry, Johns Hopkins University, Baltimore, MD, 21218, USA.
E-mail: ctownsend@jhu.edu

Experimental Details

1. Cloning.

DNA manipulations were carried out in *E. coli* strains BL21 (DE3) and DH5α (EMD Biosciences, Inc., Madison, WI) by standard methods.[1] Protein expression constructs used in this study are presented in Table S1. Primers used for cloning are presented in Table S2. The SAT-KS-MAT tridomain construct (pECTB1-NKA6) was generated from three separate DNA fragments that were elongated together using overlap extension PCR.[2] The first fragment (containing the 3' end of exon 3) was amplified from genomic *C. nicotianae* DNA template with primers CTB1-ex3-5 and CTB1-ex3-3. The second fragment was an artificial synthon with optimized codons spanning exons 4 and 5 constructed through polymerase cycling assembly (PCA).[3] The primers and final sequence of this synthon are described in Table S3. The third fragment (containing the 5' end of exon 6) was amplified from genomic *C. nicotianae* DNA template with primers CTB1-ex6-5 and CTB1-MAT6-3. Overlap extension PCR of these three fragments with outside primers CTB1-ex3-5 and CTB1-MAT6-3 was used to generate the SAT-KS-MAT fragment. This product was digested with *HindIII* and *NotI* (all restriction enzymes were NEB, Ipswich, MA) and inserted into analogous sites in pECTB1-SAT[4] to generate pECTB1-NKA6.

The PT domain construct (pECTB1-PT) was made by overlap extension PCR of three DNA fragments. The first fragment was amplified from the genomic DNA template with primers CTB1-PT-5 and CTB1-ex6-3. The second fragment was amplified from the genomic DNA template with primers CTB1-ex7-5 and CTB1-ex7-3. The third fragment was amplified from *C. nicotianae* genomic DNA template with primers CTB1-ex8-5 and CTB1-PT-3. Overlap

extension PCR was carried out with these three fragments with primers CTB1-PT-5 and CTB1-PT-3 to generate the PT DNA fragment. The resulting fragment was digested with *NdeI* and *NotI* and inserted into corresponding restriction sites in pET24a (EMD Millipore) to make pECTB1-PT.

The ACP_2 domain construct (pECTB1-ACP) was made using *C. nicotianae* genomic DNA template with primers CTB1-ACP-5 and CTB1-ACP-3s. The amplified product was digested with *NdeI* and *NotI* and inserted into the respective restriction sites in pET28a (EMD Millipore) to generate pECTB1-ACP.

An ACP_2-TE didomain construct (pECTB1-TA) was made through overlap extension PCR of two DNA fragments. The first fragment was amplified from *C. nicotianae* genomic DNA with primers CTB1-ACP-5 and CTB1-ex8-3. The second fragment was amplified from *C. nicotianae* genomic DNA with primers CTB1-ex9-5 and CTB1-3. Overlap extension PCR of these two fragments with primers CTB1-ACP-5 and CTB1-3 generated the ACP_2-TE insert. This amplified product was digested with *NdeI* and *NotI* and inserted into corresponding restriction sites in pET24a to generate pECTB1-TA.

The TE domain construct (p28CTB1-TE) was created using pECTB1-TA template with primers CTB1-TE-5 and CTB1-3-Stop. The amplified product was digested with *NdeI* and *NotI* and inserted into pET28a to generate p28CTB1-TE. The TE-S2008A mutant (p28CTB1-TE-S2008A) was generated by site-directed mutagenesis with the p28CTB1-TE template using primers CTB1-S2008A-5 and CTB1-S2008A-3. The TE-H2171Q mutant (p28CTB1-TE-H2171Q) was similarly generated using the primers CTB1-H2171Q-5 and CTB1-H2171Q-3. All constructs were confirmed through DNA sequencing (Sequencing and Synthesis Facility, Johns Hopkins University, Baltimore, MD).

2. Protein expression and purification.

All protein expressions were carried out in *E. coli* strain BL21 (DE3) harboring the appropriate expression vector. Cultures were grown in LB media supplemented with 25 µg/mL kanamycin at 37 °C until 0.6-0.8 OD_{600}. Protein expression was induced with the addition of 1 mM IPTG and carried out at 19 °C for 16 h. Cells were harvested by centrifugation, flash frozen in liquid nitrogen, and stored at -80 °C until protein purification.

Induced cell pellets were thawed on ice and suspended in 10 mM imidazole lysis buffer (4 mL/g wet cells). Lysis was accomplished by sonication (Vibra-Cell Ultrasonic Processor, Sonics & Materials, Inc., Newtown, CT) using 10 pulses of 10 s each at 40% amplitude. Protein was purified from cleared lysates by Ni-NTA chromatography as follows: 1 mL/L culture Ni-NTA slurry (Gold Biotechnology, St. Louis, MO), 3 x 10 mL washes for each 1 mL of resin slurry used with 10 mM imidazole buffer, 1 × 10 mL wash for each 1 mL of resin slurry used with 20 mM imidazole buffer, and elution with 2 × 1 mL for each 1 mL of resin slurry used with 250 mM imidazole buffer; all buffers contained 50 mM potassium phosphate pH 8.0, 300 mM sodium chloride, 10% glycerol. Proteins were dialyzed against 100 mM potassium phosphate pH 7.0, 5% glycerol at 4 °C, and were analyzed by 12% SDS-PAGE with Coomassie stain visualization (Fig. S1). Protein concentrations were determined by Bradford assay using BSA as a standard.

3. [^{18}O]Acetyl-CoA synthesis and purification.

[^{18}O]Acetyl-CoA was prepared enzymatically as follows: 5 mM CoASH, 10 mM sodium acetate ($^{18}O_2$, 95%, Cambridge Isotope Labratories, Cambridge, MA), 10 mM ATP, 1 mM TCEP, and 0.4 unit/mL acetyl-CoA synthetase (Sigma) in 25 mM potassium phosphate pH 7.0 at 25 °C for 2 h. Protein was removed by filtration through an Amicon Ultra-15 Centrifugal Filter Unit (10 kDa molecular weight cut-off, Millipore, Billerica, MA). [^{18}O]acetyl-CoA was purified by HPLC on an Agilent 1200 (Santa Clara, CA) system fitted with a Prodigy 5u ODS-3 100Å (250 x 4.60 mm, Phenomenex, Torrance, CA) column using the following mobile phase gradient: 5% solvent A, 0-5 min; 5-35% solvent A over 5-25 min (where solvent A was acetonitrile + 0.1% TFA and solvent B was water + 0.1% TFA). Purified [^{18}O]acetyl-CoA was lyophilized until dry, dissolved in 100 mM potassium phosphate pH 7.0 to yield a 25 mM solution, and stored at -80 °C. High-resolution mass analysis conducted on a Shimadzu LCMS-IT-TOF (Columbia, MD) showed the [^{18}O]acetyl-CoA had an exact mass consistent with a single heavy atom incorporation (*m/z* 812.1328 [MH^+]). A mass indicative of no heavy atom incorporation was not detected.

4. ACP activation.

The ACP fragment needed to be activated with 4'-phosphopantetheine prior to reconstitution reactions. This was accomplished *in vitro* using Svp, a promiscuous phosphopantetheinyl transferase as previously described.[5] Briefly, the ACP activation reaction was as follows: 100 µM CTB1-ACP, 2 µM Svp, 400 µM CoASH, and 20 mM $MgCl_2$ in 100 mM potassium phosphate pH 7.0, 5% glycerol at 25 °C for 1 h.

5. Reconstitution reactions.

Proteins used in the reconstitution reactions were purified immediately before use. Reaction conditions were as follows: 10 µM each of the CTB1 protein fragments to be used, 0.5 mM acetyl-CoA, 2 mM malonyl-CoA, 1 mM TCEP in 100 mM potassium phosphate, 5% glycerol, pH 7.0. The CoA substrates were added simultaneously to initiate reactions. Reactions were conducted at room temperature for 4 h. To demonstrate the starter unit effect, acetyl-CoA was excluded. To test the mechanism of pyrone formation, [^{18}O]acetyl-CoA was used instead of the unlabeled acetyl-CoA.

Reactions were quenched with the addition of concentrated HCl, and the products extracted thrice with ethyl acetate. The organic extractions were combined, evaporated under vacuum, and dissolved in a 1:4 acetonitrile:water mixture to a final volume equivalent to the reaction volume. These extracts were filtered through a 0.2 um PTFE filter and analyzed by HPLC and LC-ESI-MS.

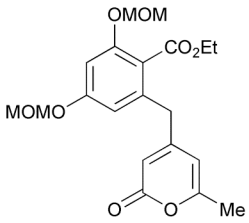
6. Analytical methods.

HPLC was carried out on an Agilent 1200 fitted with a Prodigy 5u ODS-3 100A (250 x 4.60 mm) column and UV-vis DAD detector. Reaction extracts were analyzed as 50 uL injections in a linear mobile phase gradient of 5-85% solvent A over 40 min at 1 mL/min (where solvent A was acetonitrile + 0.1% formic acid and solvent B was water + 0.1% formic acid). Polyketide products were detected by monitoring at 280 nm and full UV-vis spectra were collected for each product.

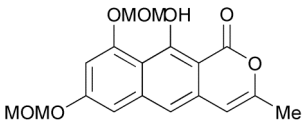
LCMS analyses were carried out on a Shimadzu LCMS-IT-TOF fitted with a Luna c18(2) 3µ (150 x 2.0 mm) column. Reaction extracts were analyzed by LC as described above, except with a flow rate of 0.2 mL/min. Mass detection was carried out in the positive ion mode.

We thank K. L. Fiedler and Prof. R. J. Cotter of the Middle Atlantic Mass Spectrometry Laboratory at the Johns Hopkins School of Medicine and the Mid-Atlantic Regional Office of Shimadzu Scientific Instruments, Inc. (Columbia, MD) for use of their Shimadzu LC-IT-TOF.

7. *Nor*-toralactone total synthesis.

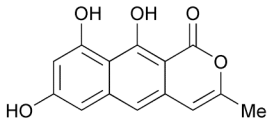


Ethyl 2,4-bis(methoxymethoxy)-6-((6-methyl-2-oxo-2H-pyran-4-yl)methyl)benzoate. A colorless solution of ethyl 2,4-bis(methoxymethoxy)-6-methylbenzoate[5] (100 mg, 0.35 mmol) in THF (0.7 mL) was added dropwise to a freshly prepared -78 °C solution of LDA in THF (0.5 M, 1.54 mL, 0.77 mmol). After allowing the resulting deep red solution to equilibrate at -78 °C for 20 min, a solution of 4-methoxy-6-methyl-2*H*-pyran-2-one[5] (50 mg, 0.35 mmol) in THF (0.7 mL) was added slowly. The resulting orange solution was allowed to react at -78 °C for 1 h, quenched (sat. NH_4Cl, 10 mL), diluted (Et_2O, 30 mL), washed (3 × sat. NH_4Cl), concentrated and purified by flash chromatography (4:1 hexane:EtOAc → 3:1 hexane:EtOAc); yielding a colorless oil (99 mg, 72%). ^{1}H NMR: (400 MHz, $CDCl_3$) δ 6.77 (d, J = 2.0 Hz, 1H), 6.48 (d, J = 2.0 Hz, 1H), 5.89 (s, 1H), 5.87 (s, 1H), 5.17 (s, 2H), 5.15 (s, 2H), 4.30 (q, J = 7.2 Hz, 2H), 3.66 (s, 2H), 3.47 (s, 3H), 3.44 (s, 3H), 2.17 (s, 3H), 1.29 (t, J = 7.2 Hz, 3H).



10-hydroxy-7,9-bis(methoxymethoxy)-3-methyl-1H-benzo[g]isochromen-1-one. To a 0 °C colorless solution of the ester above (117 mg, 0.30 mmol) in THF (5 mL), LiHMDS (1 M solution in hexane, 0.66 mL, 0.66 mmol) was added. The resulting dark brown solution was allowed to warm to room temperature over 3 h, quenched (sat. NH_4Cl), diluted (Et_2O), washed (3 × sat. NH_4Cl), concentrated and purified by flash chromatography on silica gel (4:1 hexane:EtOAc, R_f 0.4) to afford the tricyclic pyrone as a colorless solid (69 mg, 67 %). ^{1}H NMR:

(400 MHz, $CDCl_3$) δ 12.99 (s, 1H), 6.94 (s, 1H), 6.92 (d, J = 2.4 Hz, 1H), 6.77 (d, J = 2.4 Hz, 1H), 6.20 (s, 1H), 5.36 (s, 2H), 5.28 (s, 2H), 3.61 (s, 3H), 3.52 (s, 3H), 2.22 (s, 3H).



7,9,10-Trihydroxy-3-methyl-1*H*-benzo[*g*]isochromen-1-one. The di-MOM protected pyrone above (69 mg, 0.20 mmol) was dissolved in a solution of MeOH (10 mL) and conc. HCl (0.4 mL), and allowed to react at room temperature, under Ar for 4 h. The reaction mixture was diluted (EtOAc), washed (3 × sat. NH_4Cl), dried ($MgSO_4$) and concentrated to a light yellow solid (45 mg, 87%). ^{1}H NMR: (400 MHz, d_6-acetone) δ 7.07 (s, 1H), 6.73 (d, J = 2.2 Hz, 1H), 6.49 (d, J = 2.2 Hz, 1H), 2.26 (s, 3H); *m/z* 259.060 (MH^+).

8. References

(1) J. Sambrook, D.W. Russell, *Molecular Cloning: A laboratory Manual*; Cold Spring Harbor Laboratory Press, 2001.

(2) S.N., Ho, H.D. Hunt, R.M. Horton, J.K. Pullen, L.R. Pease, *Gene (Amst.),* 1989, **77,** 51-59.

(3) W.P.C. Stemmer, A. Crameri, K.D. Ha, T.M. Brennan, H.L. Heyneker, *Gene*, 1995, **164**, 49-53.

(4) J.M. Crawford, A.L. Vagstad, K.P. Whitworth, K.C. Ehrlich, C.A. Townsend, *ChemBioChem,*2008, **9**, 1019-1023.

(5) J.M. Crawford, P.M. Thomas, J.R. Scheerer, A.L. Vagstad, N.L. Kelleher, C.A. Townsend, *Science,* 2008, **320**, 243-246.

Table S1. Protein expression constructs used in this study.

Plasmid	Protein	A.A. Coverage	His_6-Tag
pECTB1-NKA6	SAT-KS-MAT	M1…S1293	C-terminal
pECTB1-PT	PT	S1293… I1654	C-terminal
pECTB1-ACP	ACP-ACP	S1637…K1909	N-terminal
p28CTB1-TE	TE	R1910…S2196	N-terminal
p28CTB1-TE-S2008A	TE-S2008A	R1910…S2196	N-terminal
p28CTB1-TE-H2171Q	TE-H2171Q	R1910…S2196	N-terminal

Table S2. Primers used in this study. Bold sequence indicates a restriction site used for cloning. Underlined sequence indicates a position of site-directed mutagenesis.

Primer	Sequence
CTB1-ex3-5	5'-GAACAGGCGCTTGCATCCGTGTCCGTTCCAT
CTB1-ex3-3	5'- CTTCACGCGGACTCATGTTGAAGTAAGC
CTB1-ex6-5	5'- TATCCCGGGTGGCAATCGCGCATTCATC
CTB1-MAT6-3	5'- TTAT**GCGGCCGC**TGACGCGGCAACAAC
CTB1-PT-5	5'- GGAATTC**CATATG**TCAAATTTGGAACTCTCTTCGTC
CTB1-ex6-3	5'- ATTAGCATGCTGCTCGACGATACTAGAGAACGTGCATGA
CTB1-ex7-5	5'- CACGTTCTCTAGTATCGTCGAGCAGCATGCTAAT
CTB1-ex7-3	5'- CTTTCTCAAGATCCACACCCTCGTTCGCGTTCAT
CTB1-ex8-5	5'- ATGAACGCGAACGAGGGTGTGGATCTTGAGAAAGAA
CTB1-PT-3	5'- GTAA**GCGGCCGC**AATCGAAGCATTCTTACGCTG
CTB1-ACP-5	5'- GTAA**CATATG**AGCGCGCCAGTCGCGCCCACG
CTB1-ACP3s	5'- GTAA**GCGGCCGC**CTACTTCTGCGCTGGGGGAGAATC
CTB1-ex8-3	5'- CTTGTGCGGTGTCTTCATGAAAGGCGAATTG
CTB1-ex9-5	5'- CCTTTCATGAAGACACCGCACAAGTTTGCTG
CTB1-3	5'- GTAA**GCGGCCGC**TGATGAAACAATCCCCAAACC
CTB1-TE-5	5'- GTAA**CATATG**AGGATACCATCCCCGCCCAAA
CTB1-3-Stop	5'- GTAA**GCGGCCGC**CTATGATGAAACAATCCCCAAACC
CTB1-S2008A-5	5'- GGCGGCTGGGCTGCTGGTGGT
CTB1-S2008A-3	5'- ACCACCAGCAGCCCAGCCGCC
CTB1-H2171Q-5	5'- AGCGAGCATCAGTTCAGCATG
CTB1-H2171Q-3	5'- CATGCTGAACTGATGCTCGCT

Table S3. Primers used for codon optimized synthon construction.

CTB1-S21	5'- TTCAACATGAGTCCGCGTGAAGCGCCGCAGGTTGATCCAG
CTB1-S22	5'- GCGTCTGGCCCTGCTGACCGCAACTGAAGCTCTGGAACAG
CTB1-S23	5'- GTGTTGTTCCGAACCGTACCTCTAGCACCCAGAAGAACCG
CTB1-S24	5'- GGCGTGTGGTACGGCGCTACTAGCAACGACTGGATGGAGA
CTB1-S25	5'- CTCTGCGCAGAACGTGGATACCTACTTTATCCCGGGTGGC
CTB1-S2-1	5'- TGCGCGATTGCCACCCGGGATAAAGTA
CTB1-S2-2	5'- TCCACGTTCTGCGCAGAGTTAGTCTCCATCCAGTCGTTGC
CTB1-S2-3	5'- AGCGCCGTACCACACGCCTACACGGTTCTTCTGGGTGCTA
CTB1-S2-4	5'- TACGGTTCGGAACAACACCTGCCTGTTCCAGAGCTTCAGT
CTB1-S2-5	5'- GTCAGCAGGGCCAGACGCTGTGCTGGATCAACCTGCGGCG

Synthon DNA Sequence:

CCGCGTGAAGCGCCGCAGGTTGATCCAGCACAGCGTCTGGCCCTGCTGACCGCAACTGAAGCTC
TGGAACAGGCAGGTGTTGTTCCGAACCGTACCTCTAGCACCCAGAAGAACCGTGTAGGCGTGTG
GTACGGCGCTACTAGCAACGACTGGATGGAGACTAACTCTGCGCAGAACGTGGATACCTACTTT
ATCCCGGGT

Synthon Protein Sequence:

PREAPQVDPAQRLALLTATEALEQAGVVPNRTSSTQKNRVGVWYGATSNDWMETNSAQNVDTYF
IPG

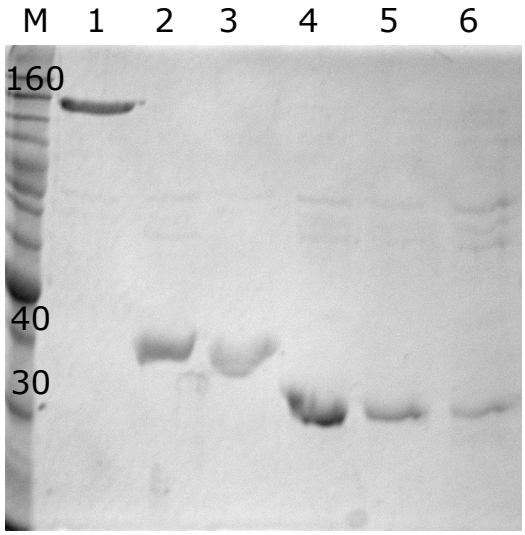


Fig. S1 12% SDS-PAGE for proteins used in this study. Lane M is Benchmark protein ladder. Lane 1 is CTB1-SAT-KS-MAT tridomain fragment. Lane 2 is CTB1-PT monodomain. Lane 3 is CTB1-ACP_2 monodomain. Lane 4 is CTB1-TE monodomain. Lane 5 is CTB1-TE-S2008A monodomain. Lane 6 is CTB1-TE-H2171Q monodomain.

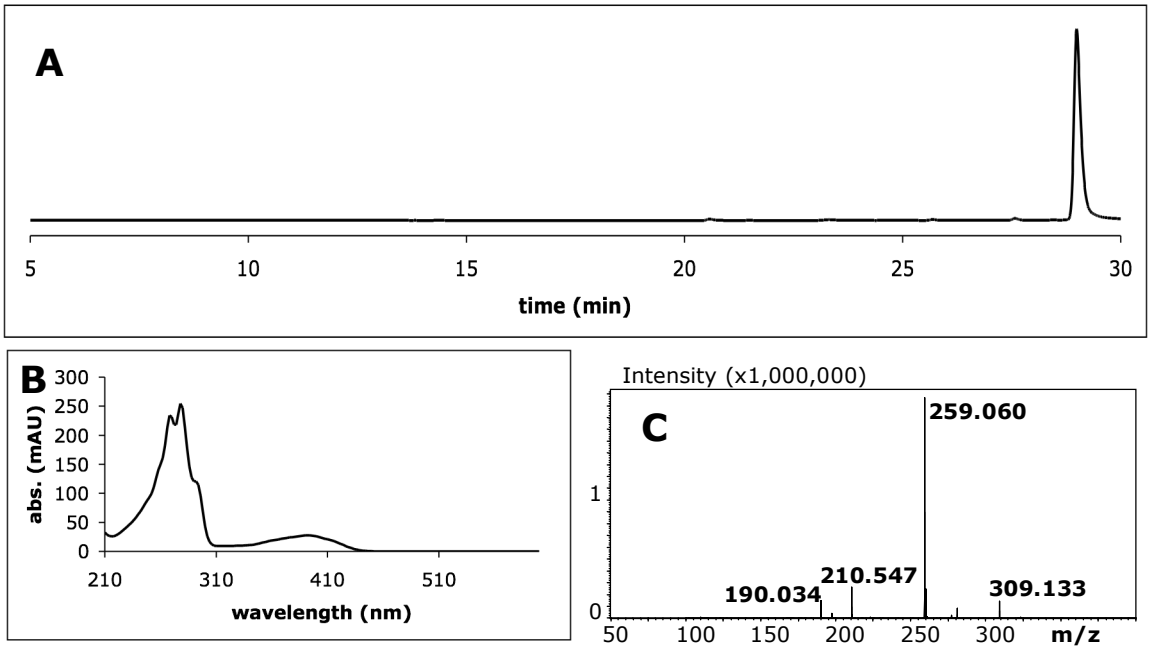


Fig. S2 LCMS data for *nor*-toralactone standard. (A) LC trace for *nor*-toralactone standard under identical experimental condition. (B) UV-vis spectrum for *nor*-toralactone. (C) MS for *nor*-toralactone.

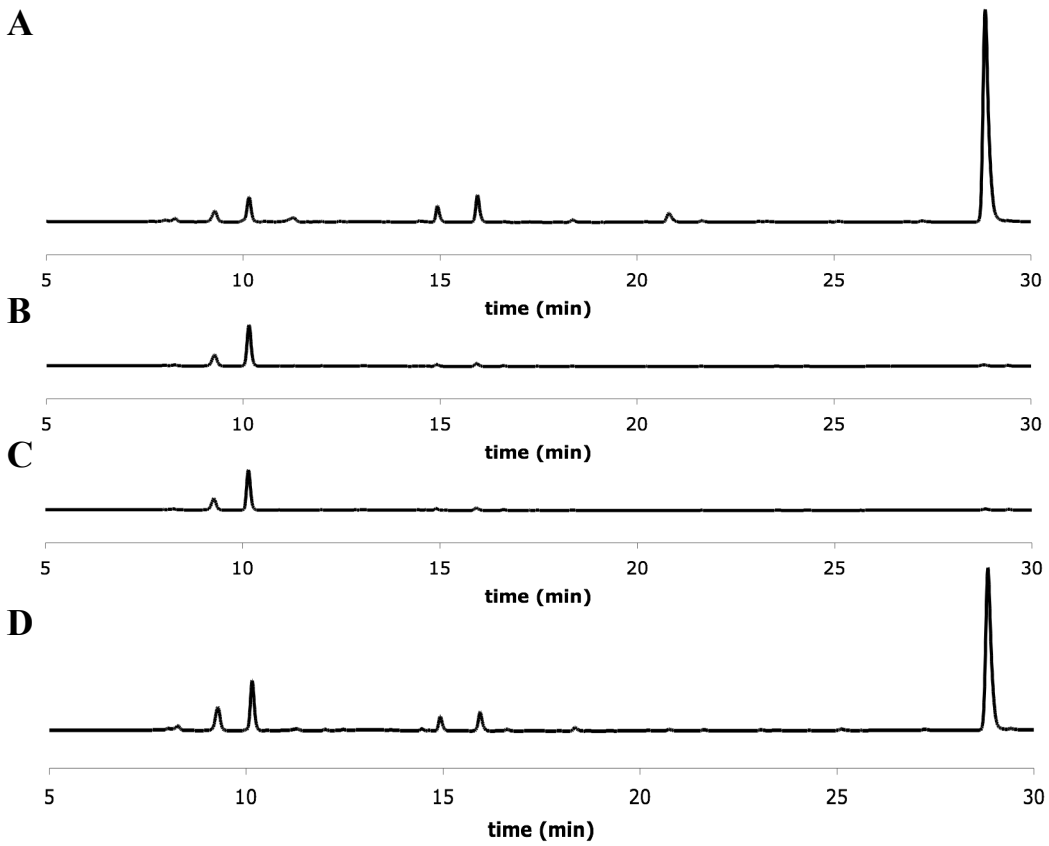


Fig. S3 Product profiles for reactions with mutant CTB1-TE. (A) Fully reconstituted CTB1 (SAT-KS-MAT + PT + ACP_2 + TE) produces *nor*-toralactone as the major product. (B) Reactions in which the CTB1 TE is not included (SAT-KS-MAT + PT + ACP_2) show reduced activity and nearly no production of *nor*-toralactone. (C) Fully reconstituted CTB1 with the TE-S2008A mutant showed no production of *nor*-toralactone. (D) Fully reconstituted CTB1 with the TE-H2171Q mutant showed reduced but significant production of *nor*-toralactone.

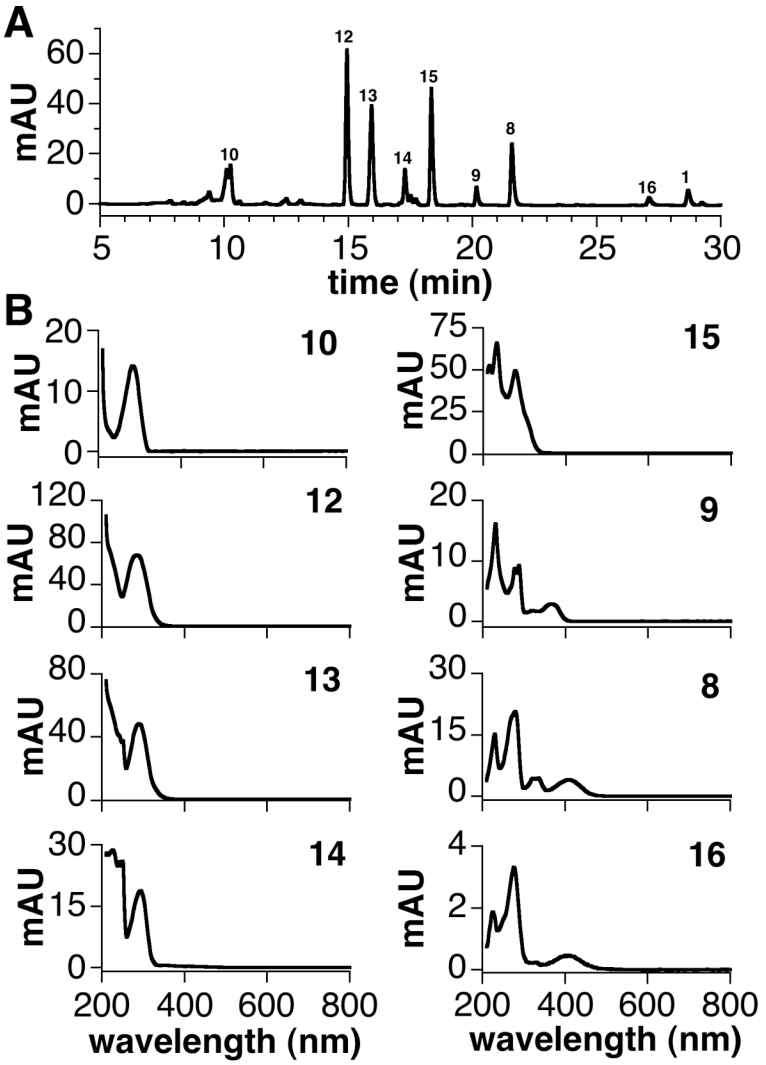


Fig. S4 UV-vis for derailment products of CTB1. (A) HPLC chromatogram at 280 nm for a minimally reconstituted CTB1. (B) UV-vis spectra for derailment products. Products **10**, **9** and **8** were identified as triacetic acid lactone, pannorin and YWA1, respectively.

Table S4. Mass spectral data for derailment products of CTB1.

Peak	*m/z*	Predicted Formula
10	127.0395 (MH^+, 100%)	$C_6H_6O_3$, triacetic acid lactone
12	277.0703 (MH^+, 100%), 299.0515 (MNa^+, 40.99)	$C_{14}H_{12}O_6$
13	277.0705 (MH^+, 100%)	$C_{14}H_{12}O_6$
14	233.0802 (MH^+, 100%)	$C_{13}H_{12}O_4$
15	277.0703 (MH^+, 100%), 299.0528 (MNa^+, 12.45)	$C_{14}H_{12}O_6$
9	259.0597 (MH^+, 100%)	$C_{14}H_{10}O_5$, pannorin
8	277.0690 (MH^+, 100%)	$C_{14}H_{12}O_6$, YWA1
16	Not detected	

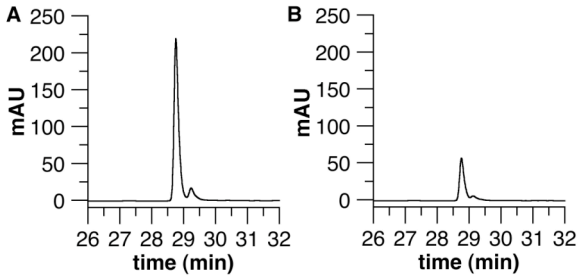


Fig. S5 "Starter-unit effect" in CTB1. (A) Fully reconstituted CTB1 with acetyl-CoA and malonyl-CoA. (B) Fully reconstituted CTB1 with only malonyl-CoA shows reduced *nor*-toralactone production.

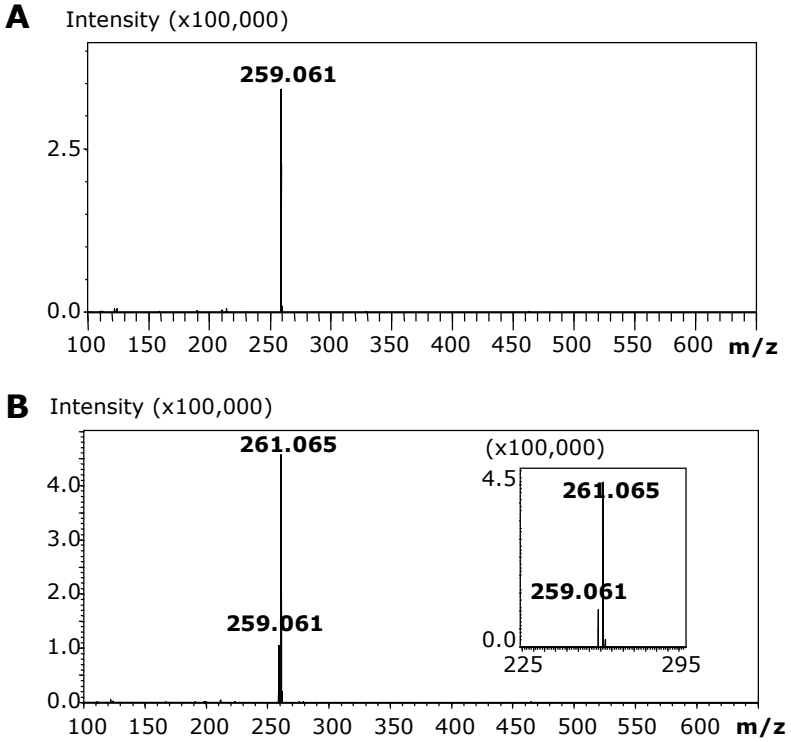


Fig. S6 A mass shift is observed for *nor*-toralactone when [^{18}O]acetyl-CoA is used as a substrate in a fully reconstituted CTB1 reaction. (A) The observed mass for *nor*-toralactone produced by a fully reconstituted CTB1 (SAT-KS-MAT + PT + ACP_2 + TE) with unlabeled acetyl-CoA as determined by LC-ESI-MS is indicative of no heavy atom incorporation. (B) The observed mass for the *nor*-toralactone produced by a fully reconstituted CTB1 (SAT-KS-MAT + PT + ACP_2 + TE) with [^{18}O]acetyl-CoA as determined by LCMS is indicative of heavy atom incorporation. Unlabeled *nor*-toralactone makes up a minor portion of the total amount of product.