

## Supplementary Materials

# A Turnstile Mechanism for the Controlled Growth of Biosynthetic Intermediates on Assembly Line Polyketide Synthases

Brian Lowry<sup>‡a†</sup>, Xiuyuan Li<sup>‡a</sup>, Thomas Robbins<sup>‡a</sup>, David E. Cane<sup>b</sup>, Chaitan Khosla<sup>\*a</sup>

<sup>a</sup> Departments of Chemistry and Chemical Engineering, Stanford University, Stanford CA 94305, <sup>b</sup> Department of Chemistry, Brown University, Providence RI 02912-9108

### AUTHOR INFORMATION

#### Corresponding Author

\* To whom correspondence should be addressed. Email: [khosla@stanford.edu](mailto:khosla@stanford.edu). Telephone: (650) 723-6538.

#### Present Addresses

†AbbVie Inc., North Chicago, IL, 60064

#### Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally.

# Table of Contents

<b>Materials – Reagents and Chemicals</b>	<b>4</b>
<b>Molecular Biology and Plasmid Construction</b> .....	<b>4</b>
<b>Bacterial Cell Culture and Protein Purification</b> .....	<b>4</b>
<b>Enzymatic Assays</b> .....	<b>4</b>
UV Spectrophotometric Kinetic Assays of Bimodular PKSs .....	4
Preparation of Chemical Crosslinker (Crypto-CoA) for Protein Crosslinking .....	4
Preparation of (2 <i>S</i> ,3 <i>R</i> )-2-methyl-3-hydroxypentanoyl-Coenzyme A (NDK-CoA) .....	6
Propionyl-CoA Synthetase, PrpE, Activity Assays .....	7
[ <sup>14</sup> C]-SDS-PAGE Labeling Assays.....	7
Proteomic LC-MS Analysis of ACP-Derived Peptides .....	7
<b>Methods</b>	<b>7</b>
<b>Plasmid Construction</b> .....	<b>7</b>
<b>Bacterial Cell Culture and Protein Purification</b> .....	<b>9</b>
Proteins Used in Crosslinking Experiments.....	9
Proteins Used in Spectrophotometric Assays, [ <sup>14</sup> C]-SDS-PAGE and Proteomics Experiments .....	9
<b>Enzymatic Assays</b> .....	<b>10</b>
UV Spectrophotometric Kinetic Assays of Bimodular PKSs .....	10
Propionyl-CoA Synthetase, PrpE, Activity Assay .....	10
Protein Crosslinking.....	11
[ <sup>14</sup> C]-SDS-PAGE Labeling Assays.....	11
LC-MS Analysis of Proteolytically-Digested Acyl-ACP Peptides (Figures S7 and S9).....	13
<b>Supplementary Figures and Results</b>	<b>14</b>
<b>References</b>	<b>27</b>

## Supplementary Figures

Figure S1. <i>Alternative representations of the 6-Deoxyerythronolide B Synthase (DEBS)</i> .....	15
Figure S2. <i>Kinetic Analysis and Engineering of Bimodular PKSs</i> .....	16
Figure S3. <i><sup>14</sup>C-SDS-PAGE Protocol Summary</i> .....	17
Figure S4. <i>Standard [<sup>14</sup>C] Labeling Plots for Sfp-labeled DEBS Proteins</i> .....	18
Figure S5. <i>Use of E. coli Propionyl-CoA Synthetase, PrpE</i> .....	19
Figure S6. <i>[<sup>14</sup>C] Labeling of LDD(4) in the absence or presence of (5)M1(2)</i> .....	20
Figure S7. <i>Proteomic LC-MS Analysis of acyl and alkyl-ACP1 active site peptides</i> .....	21
Figure S8. <i>[<sup>14</sup>C] Labeling of LDD(4), (5)M1(2), and (3)M2+TE or (3)M2+TE<sup>0</sup></i> .....	22
Figure S9. <i>Proteomic LC-MS Analysis of the Triketide Acyl-ACP2 active site peptide</i> .....	23
Figure S10. <i>[<sup>14</sup>C] Labeling of LDD(4), (5)M1(2), (3)M2(2), (5)M5, and DEBS2</i> .....	24
Figure S11. <i>Energetic Coupling of the turnstile to polyketide chain elongation</i> .....	25
Table S1. <i>Plasmid Design and Primers</i> .....	26

# MATERIALS – REAGENTS AND CHEMICALS

## Molecular Biology and Plasmid Construction

All primers for cloning of gene constructs were synthesized by Elim Biopharmaceuticals. Phusion High Fidelity polymerase and Phusion Hot Start polymerase were from Thermo Scientific. T5 Exonuclease was from Epicentre. Taq DNA ligase was from New England Biolabs. For standard restriction-ligation procedures, restriction enzymes, NdeI and EcoRI, were from New England Biolabs and T4 DNA Ligase was from Invitrogen. For site-directed mutagenesis, the QuikChange II XL Site-Directed Mutagenesis Kit was from Agilent Technologies. DNA isolation and purification kits, GeneJET Plasmid Miniprep, PCR Purification, and Gel Extraction Kits, were from Thermo Scientific.

## Bacterial Cell Culture and Protein Purification

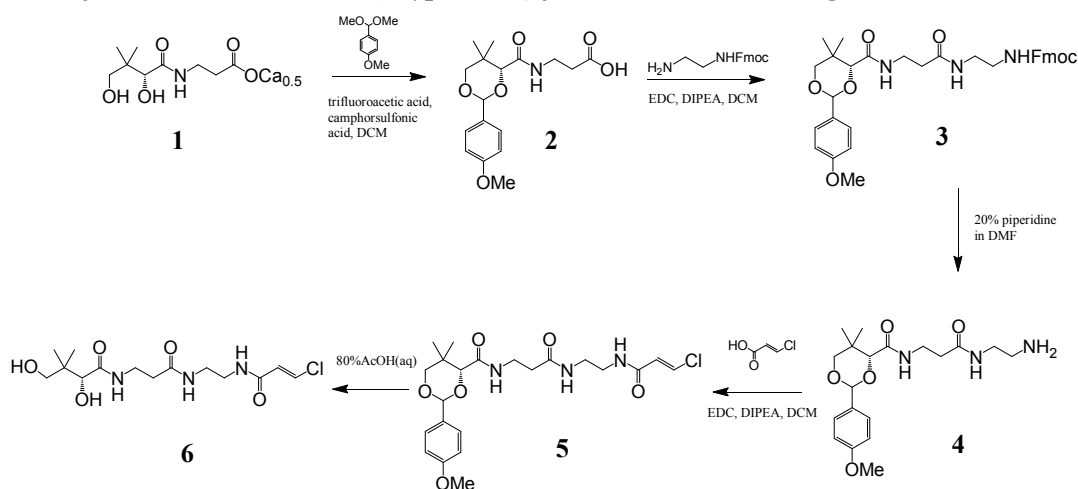
All chemicals for preparation of buffers were from Sigma-Aldrich. Luria-Bertani (LB) Miller Broth was from Fisher Scientific. Isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG), kanamycin sulfate, and carbenicillin were from Gold Biotechnology. Complete protease inhibitor cocktail tablets were from Roche. Ni-NTA affinity resin was from MC Lab. HiTrap Q anion exchange chromatography columns for protein purification were from GE Healthcare. SDS-PAGE Mini Protean TGX Precast Gels were from Bio-Rad Laboratories, and Amicon Ultra centrifugal filters were from Millipore. Mini-PROTEAN TGX precast polyacrylamide gels (4-20% polyacrylamide gradient) were from Bio-Rad Laboratories.

## Enzymatic Assays

### UV Spectrophotometric Kinetic Assays of Bimodular PKs

Coenzyme A (CoASH), reduced  $\beta$ -nicotinamide adenine dinucleotide 2'-phosphate (NADPH), sodium propionate, propionyl-CoA, methylmalonic acid, and magnesium chloride hexahydrate were purchased from Sigma Aldrich. Adenosine-5'-triphosphate (ATP) was purchased from Teknova. Bond-breaker tris-(2-carboxyethyl)-phosphine (TCEP) was purchased from Thermo Scientific. UVette cuvettes (2 mm x 10 mm path) were purchased from Eppendorf.

### Preparation of Chemical Crosslinker (Crypto-CoA) for Protein Crosslinking



**PMB-pantothenic acid (2)** - Pantothenate hemicalcium (**1**, 9.59 g, 40.2 mmol) and camphorsulfonic acid (937 mg, 4.0 mmol) were added to a 500 mL round-bottom flask equipped with a magnetic stir bar. With stirring on ice, trifluoroacetic acid (TFA) was slowly added to dissolve the solids. The residual TFA was then removed under vacuum. Dichloromethane (100 mL) was then added to dissolve the glassy/oily pantothenic acid. Then, p-anisaldehyde dimethyl acetal was added in three portions (in total 22.5 g, 21 mL, 123.5 mmol). The reaction mixture was stirred for 8 h at room temperature, and then slowly added to 100 mL 1 M Na<sub>2</sub>CO<sub>3</sub> (aq.) with vigorous stirring. The precipitate was removed by filtration, and the aqueous phase (pH 8-9) was washed three times with 100 mL dichloromethane, slowly acidified by addition of acetic acid to pH~4.5, and again extracted three times with 100 mL dichloromethane. The organic phase was dried on Na<sub>2</sub>SO<sub>4</sub> and filtered. Removal of solvent yielded **2** as a pale yellow solid (6.9g, 51%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.41 (d, 2H, *J*= 6.9 Hz), 7.03 (broad, 1H), 6.91 (d, 2H, *J*=6.6 Hz), 5.46 (s, 1H), 4.10 (s, 1H), 3.81 (s, 3 H), 3.68 (d, 2H, *J*= 7.2 Hz), 3.52 (m, 2H), 2.62 (t, 2H, *J*=6.3Hz), 1.10 (s, 3 H), 1.09 (s, 3 H).

**1,3-PMP-protected C11-Fmoc-amino-pantetheine analogue (3)** - PMB-pantothenic acid **2** (2.48 g, 7.35 mmol), mono-Fmoc-ethylenediamine hydrochloride (1.59 g, 4.99 mmol), 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) (1.17 g, 6.1 mmol) and 1-hydroxybenzotriazole (HOBT) (158 mg, 1.17 mmol) were added to a 100 mL round-bottom flask equipped with a magnetic stir bar. Dry dichloromethane (50 mL) was added and the flask was cooled on ice. N, N-Diisopropylethylamine (3.5 mL, 20.0 mmol) was added slowly, and the system was allowed to slowly warm up to room temperature and react for 13 h. The mixture was washed twice with 50 mL water and dilute acetic acid (aq, pH=6), and once with 50 mL saturated NaCl solution. The organic phase was dried on Na<sub>2</sub>SO<sub>4</sub>. Removal of solvent produced crude **3** as a white solid (1.86 g, 62%). <sup>1</sup>H NMR δ (300 MHz, CDCl<sub>3</sub>) δ 7.77 (d, 2H, *J*=7.2 Hz) 7.59 (d, 2H, *J*=7.8 Hz) 7.28-7.43 (m, 6H), 6.89 (d, 2H, *J*=9.0 Hz), 5.38 (s, 1H), 4.44 (d, 2H, *J*=7.8 Hz) 4.20 (t, 1H, *J*=6.6 Hz) 4.04 (s, 1H), 3.80 (s, 3H), 3.64 (d, 2H, *J*= 7.2 Hz), 3.55 (m, 4H) 2.41 (t, 2H, *J*=6.3 Hz), 1.09 (s, 3H), 1.07 (s, 3H).

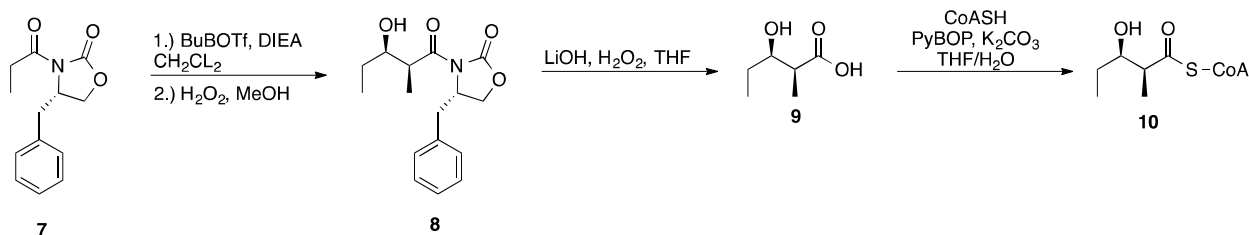
**1,3-PMP-protected amino-pantetheine analogue (4)** - 1,3-PMP-protected C11-Fmoc-amino-pantetheine **3** (423 mg, 703 μmol) was dissolved in 2mL 30% piperidine in DMF and stirred at room temperature for 3 h. The mixture was then concentrated under vacuum and recrystallized from chloroform and ethyl ether to yield **4** as a white solid (120 mg, 45%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.90 (broad, 1H), 7.42 (d, 2H, *J*=8.7 Hz), 7.13 (t, 1H, *J*=6 Hz), 6.89 (d, 2H, *J*=9.0 Hz), 5.42 (s, 1H), 4.04 (s, 1H), 3.78 (s, 3H), 3.62 (d, 2H, *J*=6.3 Hz), 3.47 (m, 4H), 3.13 (broad, 2H) 2.42 (t, 2H, *J*=6.0 Hz), 1.06 (s, 3H), 1.00 (s, 3H).

**1,3-PMP-(*R,E*)-N-(3-(2-(3-Chloroacrylamido) ethylamino)-3-oxopropyl)-2,4-dihydroxy-3,3-dimethylbutanamide (5)** - 1,3-PMP-protected amino-pantetheine **4** (87.6 mg, 231 μmol), E-3-chloroacrylic acid (87.6 mg, 822 μmol) and EDC (303.9 mg, 1.59 mmol) were dissolved in 2 mL ice-cold dichloromethane. N, N-Diisopropylethylamine (0.52 mL, 3 mmol) was added slowly and the system was allowed to slowly warm up to room temperature and react for 8 h. The crude was washed three times with water then once with brine. Then the crude was purified by flash chromatography (ethyl acetate to 10% methanol: ethyl acetate) to yield **5** (11.4 mg, 10%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.44 (dd, 2H, *J*=1.8 Hz and 6.6 Hz), 7.25 (d, *J*=13.2 Hz) 7.0 (t, 1H, *J*=6.0 Hz), 6.93 (dd, 2H, *J*=1.8 Hz and 6.9 Hz), 6.22 (d, 1H, *J*=13.2 Hz) 5.47 (s, 1H), 4.08 (s), 3.82 (s, 3H), 3.69 (d, 2H, *J*=7.8 Hz), 3.4-3.5 (m, 4H), 2.43 (t, 2H, *J*=6.0 Hz), 1.11 (s, 3H) 1.09 (s, 3H).

**(*R,E*)-N-(3-(2-(3-Chloroacrylamido) ethylamino)-3-oxopropyl)-2,4-dihydroxy-3,3-dimethylbutanamide (6)** - Methodology and spectroscopic data are consistent with a previously reported preparation<sup>1</sup>. 1,3-PMP-(*R,E*)-N-(3-(2-(3-Chloroacrylamido) ethylamino)-3-oxopropyl)-2,4-dihydroxy-3,3-dimethylbutanamide **5** (4.1 mg, 8 μmol) was dissolved in 2.5 mL 80% acetic acid in water and stirred for 1.5 h. The mixture was then dried in vacuum. The crude was washed three times with 1 mL chloroform to yield 2.6 mg of **6** (85%). <sup>1</sup>H NMR (300 MHz, CD<sub>3</sub>OD): δ 7.27 (d, 1H, *J*=13.2 Hz), 6.40 (d,

1H,  $J=13.2$  Hz), 3.89 (s, 1H), 3.45-3.50 (m), 2.41 (t, 2H,  $J=6.6$  Hz), 0.92 (s, 6H). LCMS (ESI) ( $m/z$ )  $[M+Na]^+$  calcd. for  $C_{14}H_{23}ClN_3O_5Na$ , 372.1 and 374.1, found 372.4 and 374.4.

**Preparation of (2S,3R)-2-Methyl-3-hydroxypentanoyl-Coenzyme A (NDK-CoA)**



**(S)-4-Benzyl-3-((2S,3R)-3-hydroxy-2-methylpentanoyl)oxazolidin-2-one (8)** – Methodology and spectroscopic data are consistent with a previously reported preparation.<sup>2</sup> (S)-4-Benzyl-3-propionyl-2-oxazolidinone (5.6 mmol, 1.1 eq) was dissolved in dry  $CH_2Cl_2$  (20 mL) under an argon atmosphere. Dibutylboron triflate (1M in  $CH_2Cl_2$ , 7.7 mmol, 1.5 eq) was added drop-wise at  $0^\circ C$  and the reaction left to stir for 10 minutes, at which point diisopropylethylamine (DIEA, 8.2 mmol, 1.6 eq.) was added. The reaction was left to stir for 45 min at which point the temperature was lowered to  $-78^\circ C$  followed by drop-wise addition of propanal (10.26 mmol, 2 eq). The reaction was stirred at  $-78^\circ C$  for 2 h, then brought to  $0^\circ C$  and allowed to stir for an additional hour, at which point it was quenched by the addition of phosphate buffer (1M, pH = 7.6, 20 mL). extracted with  $CH_2Cl_2$  (3x50 mL) and dried over  $Na_2SO_4$ . The solvent was removed *in vacuo* and the resulting orange oil was brought up in MeOH (3 mL).  $H_2O_2$  (30% w/v, 40 mmol, 8 equiv), was added dropwise at  $0^\circ C$  followed by 1 h of stirring. The reaction was extracted with  $CH_2Cl_2$  (3x50 mL), dried over  $Na_2SO_4$ , and solvent removed *in vacuo* to yield **1** as a white solid. Purification over silica gel (ethyl acetate in pentanes (30%)) yielded a white solid (41%,  $R_f=0.4$ , 30% ethyl acetate in pentanes).  $^1H$  NMR (300 MHz,  $CDCl_3$ )  $\delta$  7.65 (s, 1H),  $\delta$  7.37-7.19 (m, 5H),  $\delta$  4.72 (m, 1H),  $\delta$  4.23 (m, 2H),  $\delta$  3.86 (m, 1H),  $\delta$  3.78 (q, 1H),  $\delta$  3.27 (dd, 1H),  $\delta$  2.91 (s, 1H),  $\delta$  2.78 (dd, 1H),  $\delta$  1.50 (m, 2H),  $\delta$  1.25 (d, 3H),  $\delta$  0.98 (t, 3H).

**(2S,3R)-3-Hydroxy-2-methylpentanoic acid (9)** (8) Methodology and spectroscopic data are consistent with a previously reported preparation<sup>3</sup> (1 mmol, 1 eq, 300 mg) was dissolved in THF (10 mL) and brought to  $0^\circ C$ . LiOH (7.2 mmol, 7 eq, 172 mg) was added and the reaction was allowed to stir for 10 minutes, after which time  $H_2O_2$  (30% w/v, 15.5 mmol, 15 equiv), was added dropwise. The reaction was allowed to warm to room temperature over a period of 16 h, at which time 1M HCl (4 mL) was added. The reaction was extracted with  $CH_2Cl_2$  (3x50 mL) and dried over  $Na_2SO_4$  and solvent removed *in vacuo*. Purification over silica gel yielded a clear oil (25%,  $R_f=0.3$ , ethyl acetate in pentanes (70%)).  $^1H$  NMR (300 MHz,  $CDCl_3$ )  $\delta$  3.87 (m, 1H),  $\delta$  2.63 (m, 1H),  $\delta$  1.52 (m, 2H),  $\delta$  1.21 (d, 3H),  $\delta$  0.99 (t, 3H). LCMS (ESI) ( $m/z$ )  $[M+H]^+$  calcd. For  $C_6H_{13}O_3$ , 133.08, found, 133.15.

**(2S,3R)-2-Methyl-3-hydroxypentanoyl-Coenzyme A (NDK-CoA) (10)** Preparation of NDK-CoA is adapted from a previous study in this laboratory<sup>4</sup>. (9) (31  $\mu$ mol, 4.1 mg) was dissolved in 0.5 mL THF. Coenzyme-A (34.1  $\mu$ mol, 27.2 mg) and PyBOP (55.8  $\mu$ mol, 29 mg) were dissolved in 0.5 mL of 4% (w/v)  $K_2CO_3$  in  $H_2O$ . These were combined and stirred under argon at room temperature for 2 h, diluted with 3.5 mL of  $H_2O$ , and purified directly over a C8 column. Gradient (2-20% B over 35 min, 4 mL/min, A: 50

mM ammonium acetate, pH =4.2, B: Acetonitrile). Fractions were lyophilized to yield a white powder Yield: (5.1%). ESI-MS:  $[M+H]^+$  calcd. For  $C_{27}H_{47}N_7O_{18}P_3S$ , 882.19, found: 882.2. As proof of the identity of NDK-CoA, a phosphopantetheinylation reaction was catalyzed by the Sfp phosphopantetheine transferase under the following conditions: 50  $\mu$ M *apo*-DEBS ACP1, 20 equiv of either NDK-CoA or CoASH, 0.2 equiv Sfp in 25 mM  $NaH_2PO_4$  (pH=7.6), 10 mM  $MgCl_2$ , 5 mM TCEP, 5% glycerol at 37 °C for 30 min. Protein masses were confirmed by ESI-MS (*holo*-ACP1 expected, 13830.09, observed mass: 13831.0, *holo*-ACP1-NDK expected, 13944.17, observed mass: 13944.0).

### ***Propionyl-CoA Synthetase, PrpE, Activity Assays***

5,5'-Dithio-*bis*-(2-nitrobenzoic acid) (Ellman's Reagent) and ethylenediaminetetraacetic acid (EDTA) were from Sigma-Aldrich. Costar flat bottom 96-well microtiter UV plates were from Corning.

### ***[<sup>14</sup>C]-SDS-PAGE Labeling Assays***

[1-<sup>14</sup>C]-Propionic acid, sodium salt, was from Moravек Biochemicals. Methylmalonic acid, sodium propionate,  $\beta$ -nicotinamide adenine dinucleotide 2'-phosphate tetrasodium salt (NADPH), and Coenzyme A trilithium salt were from Sigma-Aldrich. Adenosine-5'-triphosphate (ATP) was from Teknova. Bond-breaker tris-(2-carboxyethyl)-phosphine (TCEP) was from Thermo Scientific. Mini-PROTEAN TGX precast polyacrylamide gels (7.5% polyacrylamide) were from Bio-Rad.

### ***Proteomic LC-MS Analysis of ACP-Derived Peptides***

Sodium propionate, methylmalonic acid, sodium propionate,  $\beta$ -nicotinamide adenine dinucleotide 2'-phosphate tetrasodium salt (NADPH), Coenzyme A trilithium salt, and N-ethylmaleimide were from Sigma-Aldrich. Adenosine-5'-triphosphate (ATP) was from Teknova. Bond-breaker tris-(2-carboxyethyl)-phosphine (TCEP) was from Thermo Scientific. Urea was from J.T. Baker Chemicals. Mass spectrometry grade Trypsin/Lys-C protease mix was from Promega. HPLC grade acetonitrile was from Fisher Scientific, and formic acid was from Fluka.

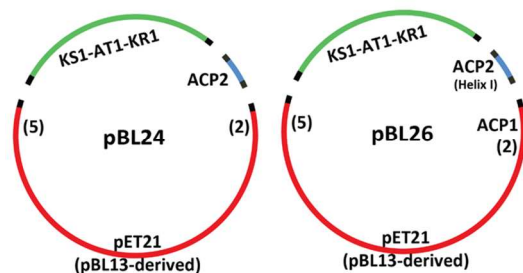
## **METHODS**

### **Plasmid Construction**

DNA constructs used to express the proteins of interest were prepared using either a standard restriction digestion – ligation protocol, site-directed mutagenesis, or the Gibson Assembly protocol<sup>5</sup>. In this work, multiple plasmids were used to express various DEBS proteins or auxiliary enzymes (see

Table S1 for a complete list). Plasmids constructed specifically for this study are pBL48 (expressing *E. coli* PrpE), pBL10 (expressing module 5 with its own N-terminal docking peptide and lacking a C-terminal docking peptide; M5), pBL53 (expressing the loading didomain harboring the C-terminal docking peptide from module 4 as well as an inactivated AT domain; LDD/AT<sub>L</sub><sup>0</sup>), pBL60 (expressing module 2 fused to the N-terminal docking peptide from module 3 and a C-terminal TE domain harboring an active site S→A mutation; M2+TE<sup>0</sup>), pBL24 (expressing module 1, where ACP1 has been exchanged for ACP2, with the N-terminal docking domain from module 5 and the C-terminal docking domain from module 2), and pBL26 (expressing module 1, where helix I of ACP1 has been exchanged for its counterpart from ACP2, with the N-terminal docking domain from module 5 and the C-terminal docking domain from module 2).

The propionyl-CoA synthetase, PrpE, from *E. coli* was amplified by PCR from the genomic DNA of *E. coli* BL21(DE3) cells with introduction of flanking 5'-*Nde*I and 3'-*Eco*RI sites. This PCR product and a corresponding pET28b expression vector were digested with *Nde*I and *Eco*RI at 37°C for 2 h. Following purification (GeneJET Gel Extraction Kit), the *Nde*I-*Eco*RI fragments were ligated at room temperature for 1 h using T4 DNA ligase to afford the expression construct pBL48. DEBS Module 5 possessing its natural N-terminal docking domain, M5, was amplified by PCR from pFW100<sup>6</sup>. This PCR product and a corresponding pET21b expression vector were digested with *Nde*I and *Eco*RI at 37°C for 2 h. Following purification (GeneJET Gel Extraction Kit), the *Nde*I-*Eco*RI fragments were ligated at room temperature for one hour using T4 DNA ligase to afford the expression construct pBL10. Inactivation of the AT domain within LDD was accomplished using site-directed mutagenesis. A S196A mutation was introduced into the plasmid, pBL12 – LDD (construction described previously<sup>6</sup>), using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent). Inactivation of the TE domain within M2TE was also accomplished using site-directed mutagenesis. A S1597A mutation was introduced into the plasmid, pB16 – M2+TE (construction described previously<sup>6</sup>), using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent). The resulting mutations afforded LDD/AT<sub>L</sub><sup>0</sup>, pBL53, and M2+TE<sup>0</sup>, pBL60.



Using the Gibson Assembly protocol, plasmid pBL24, encoding the gene for the recombinant M1/ACP2 protein (module 1, where ACP1 has been exchanged for ACP2, with the N-terminal docking domain from module 5 and the C-terminal docking domain from module 2), was assembled from three fragments: the module 5 N-terminal docking domain and the module 2 C-terminal docking domain with a pET21 vector backbone (red), module 1 (green), and ACP2 (blue). The pET21-containing vector fragment was amplified from pBL13<sup>6</sup>, module 1 was amplified from pBL13, and ACP2 was amplified from pNW7<sup>7</sup>. Plasmid pBL26, encoding the gene for the recombinant M1/ACP2-Helix I protein (module 1, where ACP1's helix I has been exchanged for ACP2's helix I, with the N-terminal docking domain from module 5 and the C-terminal docking domain from module 2), was assembled from three fragments in a similar manner. Fragments were added to the assembly mixture to synthesize the final circular plasmid.



## Bacterial Cell Culture and Protein Purification

### *Proteins Used in Crosslinking Experiments*

Apo-ACPs, KS3AT3, DEBS Module 3 and DEBS Module 3+TE were each expressed as previously described<sup>8,9</sup>. Plasmids encoding ACP2(2), ACP3 and KS3AT3 were introduced into *E. coli* BL21(DE3) by electroporation, whereas plasmids expressing Module 3 and Module 3+TE were introduced into *E. coli* BAP1 cells to ensure phosphopantetheinylation of their ACP domains<sup>10</sup>. Overnight seed cultures (5 mL) were used to inoculate a 1-L shake flask culture containing the appropriate antibiotic. Cells were grown to an approximate OD<sub>600</sub> of 0.6 at 37 °C and then induced with 250 µL of 1 M IPTG (for modules) or 1 mL of 1 M IPTG (for ACPs) and incubated at 18°C for 12-15 h. Cell pellets were harvested by centrifugation at 4420 g and lysed by sonication in lysis buffer (50 mM sodium phosphate, 10 mM imidazole, 450 mM NaCl, 20% glycerol, pH = 7.6). The lysate was clarified by centrifugation at 25,000 g, filtered through 0.2 µm filter, and then injected into HisTrap column (GE Life Sciences, Buffer A - 150 mM phosphate, 40 mM NaCl 10 mM imidazole, pH = 7.6, 10% glycerol and Buffer B - 150 mM phosphate, 40 mM NaCl, 500 mM imidazole, pH = 7.6, 10% glycerol). For ACPs, eluted fractions were further purified by hydrophobic interaction chromatography on a HiTrap Phenyl Sepharose column (Buffer A: 50 mM phosphate, 1M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH = 7.6, 10% glycerol and Buffer B: 50 mM phosphate, pH = 7.6, 10% glycerol) on an ÄKTA Pure FPLC system (GE Life Sciences). For KSAT3, Module 3, and Module 3+TE, Ni-NTA eluates were further purified using anion exchange chromatography (HiTrap Q HP 5Ml column, Buffer A: 50 mM phosphate, pH = 7.6, 10% glycerol and Buffer B: 50 mM phosphate, 500 mM NaCl, pH = 7.6, 10% glycerol). FPLC fractions were analyzed by SDS-PAGE, and fractions containing the desired protein were pooled and concentrated using an Amicon Ultra filter. Protein concentrations were measured using the Pierce™ BCA Protein Assay Kit (Thermo Scientific). Protein samples were aliquoted and stored at -80 °C.

### *Proteins Used in Spectrophotometric Assays, [<sup>14</sup>C]-SDS-PAGE and Proteomics Experiments*

Expression plasmids for LDD/AT<sub>L</sub><sup>0</sup>, M1, M2+TE<sup>0</sup>, and DEBS2 were introduced into *E. coli* BL21 (DE3) cells to produce the corresponding *apo*- proteins lacking a phosphopantetheine post-translational modification. The expression plasmids for LDD, M1, M2+TE<sup>0</sup>, M2+TE, M2, DEBS2, M5, BL24 (see Plasmid Construction for details), BL26 (see Plasmid Construction for details), and M3+TE were introduced into *E. coli* BAP1 cells to allow phosphopantetheinyl modification of ACP domains<sup>10</sup>. 5 mL overnight seed cultures were used to inoculate a 1 L shake flask culture containing the appropriate antibiotic. Cells were grown to an approximate O.D. of 0.6 and then induced with 250 µL of 1M IPTG. After 12-15 h, cells were harvested by centrifugation at 4420 g and lysed by sonication in lysis buffer (50 mM sodium phosphate, 10 mM imidazole, 450 mM NaCl, 20% glycerol, pH = 7.6). The lysate was clarified by centrifugation at 25,000 g, and the supernatant was added to Ni-NTA agarose resin (2 mL of resin per L of culture) and allowed to incubate at 4 °C for 1 h. The incubating resin was applied to a Kimble-Kontes Flex column and washed with 20 column volumes of lysis buffer, 10 column volumes of wash buffer (50 mM phosphate, 25 mM imidazole, 300 mM NaCl, 10% glycerol, pH = 7.6), and eluted with 8 column volumes of elution buffer (75 mM phosphate, 500 mM imidazole, 20 mM NaCl, 10% glycerol, pH = 7.6). The eluent was further purified using anion exchange chromatography (HiTrap Q column, Buffer A: 50 mM phosphate, pH = 7.6, 10% glycerol and Buffer B: 50 mM phosphate, 500 mM NaCl, pH = 7.6, 10% glycerol) on an ÄKTA Pure FPLC system (GE Life Sciences). FPLC fractions were

analyzed by SDS-PAGE, and fractions containing the desired protein were pooled and concentrated using an Amicon Ultra filter. Protein concentrations were measured using the Pierce™ BCA Protein Assay Kit (Thermo Scientific). Protein samples were aliquoted and stored at -80°C. The purification and isolation of the following DEBS-derived proteins used in this study have been described previously: LDD, M1, M2, M2+TE, and DEBS2 <sup>6</sup>.

The *E. coli* propionyl-CoA synthetase, PrpE, was also purified using the above protocol but using the following buffer system – lysis buffer: 100 mM HEPES, 500 mM NaCl, 20% glycerol, pH = 7.5, wash buffer: 100 mM HEPES, 300 mM NaCl, 15 mM imidazole 20% glycerol, pH = 7.5, elution buffer: 100 mM HEPES, 300 mM NaCl, 300 mM imidazole 20% glycerol, pH = 7.5, Buffer A: 50 mM HEPES, 10% glycerol, pH = 7.5, Buffer B: 50mM HEPES, 500 mM NaCl, 10% glycerol, pH = 7.5. After the protein was concentrated by Amicon filtration, 5 mM DTT was added to the final volume prior to freezing.

## Enzymatic Assays

### *UV Spectrophotometric Kinetic Assays of Bimodular PKSs*

Assays to measure the initial rates of polyketide synthesis were performed on a 70- $\mu$ L scale in 400 mM sodium phosphate (pH = 7.2) containing 5 mM TCEP, 10 mM MgCl<sub>2</sub>, 500  $\mu$ M Coenzyme A, and 4 mM ATP. Enzymes MatB (2  $\mu$ M) and methylmalonyl-CoA epimerase (4  $\mu$ M) were added to convert methylmalonic acid into racemic methylmalonyl-CoA <sup>11</sup>. PrpE (1  $\mu$ M) was added to convert propionate into propionyl-CoA (see below). The concentrations of these enzymes were selected to ensure that acyl-CoA supply was not rate limiting, as described above. The DEBS proteins (LDD, M1, M2+TE or M3+TE as shown in Figure S2B; LDD, M3+TE, and M1 or BL24 or BL26 as shown in Figure S2D) were then added to a final concentration of 4  $\mu$ M each. Reactions were initiated upon simultaneous addition of propionate (1 mM), methylmalonic acid (1 mM), and NADPH (750  $\mu$ M), and the rate was monitored over 30 min in an Eppendorf brand UVette cuvette at 340-nm (depletion of NADPH) using a Lambda-25 UV-Vis Spectrophotometer (Perkin-Elmer).

### *Propionyl-CoA Synthetase, PrpE, Activity Assay*

The enzymatic activity of the propionyl-CoA synthetase, PrpE, was evaluated using the colorimetric DTNB assay to monitor CoASH depletion. Initial rate measurements were performed on a 120  $\mu$ L scale in 400 mM sodium phosphate (pH = 7.2), 10 mM MgCl<sub>2</sub>, 1 mM CoA, and 4 mM ATP. PrpE (0.5  $\mu$ M) was added to each reaction, and reactions were initiated upon addition of sodium propionate (titration from 50-1600  $\mu$ M). Samples of 25  $\mu$ L were taken at regular intervals over a 20 minute range and quenched with 250  $\mu$ L of reaction buffer containing 5  $\mu$ L of Ellman's reagent solution, prepared as a 4 mg/mL solution of solid Ellman's reagent dissolved in reaction buffer (0.1 M sodium phosphate, pH = 8.0, containing 1 mM EDTA). The amount of free sulfhydryl at each time point was measured spectrophotometrically at 412 nm in a flat bottom Costar 96-well microtiter UV plate using a Synergy H4 Hybrid Multi-Mode Microplate Reader (BioTek), and converted to a concentration of CoASH using a CoASH standard. Data demonstrating CoASH depletion by PrpE at different concentrations of propionate are shown in Figure S5A. The initial rates obtained from each curve were used to estimate the  $k_{cat}$  and  $K_M$  of propionyl-CoA synthesis (fitting performed using MATLAB).

PrpE was used to generate propionyl-CoA *in situ* during the DEBS activity assays. Using reaction conditions identical to those previously reported for measuring the activity of the reconstituted bimodular DEBS<sup>6</sup>, propionyl-CoA was replaced with propionate and PrpE was titrated into the reaction from 0 to 1  $\mu\text{M}$ . As before, the rate was monitored in a UVette cuvette at 340-nm (depletion of NADPH) using a Lambda-25 UV-Vis Spectrophotometer (Perkin-Elmer). Initial rates of polyketide biosynthesis for different concentrations of PrpE and that for a propionyl-CoA control are presented in Figure S5B. All subsequent assays reported in this work have incorporated this strategy into the experimental design for *in situ* generation of propionyl-CoA.

### **Protein Crosslinking**

The *crypto*-ACPs were generated by a procedure adapted from a previously reported one-pot chemo-enzymatic reaction using a mixture of CoASH biosynthetic enzymes (CoaA, CoaD, CoaE) and the Sfp phosphopantetheinyl transferase (PPTase)<sup>12</sup>. A representative reaction mix contained: 75 mM sodium phosphate (pH 7.2), 12.5 mM ATP, 20 mM  $\text{MgCl}_2$ , 2  $\mu\text{M}$  CoaA, 2  $\mu\text{M}$  CoaD, 4  $\mu\text{M}$  CoaE, 15  $\mu\text{M}$  Sfp (native), 400  $\mu\text{M}$  apo-ACP, and 1.6 mM crosslinker (4 equivalents). After addition of all components, the reaction mixtures were incubated at 37°C for 6 h. Formation of *crypto*-ACP was confirmed by observation of a 410 Da mass increase compared with the starting *apo*-ACP on Waters 2795 HPLC system with dual wavelength UV detector and ZQ single quadrupole MS with electrospray ionization source.

All crosslinking reactions were performed at room temperature with 250  $\mu\text{M}$  *crypto*-ACP, 50  $\mu\text{M}$  KS3AT3 or DEBS module 3 (M3) or DEBS module 3+TE (M3TE), 50 mM  $\text{MgCl}_2$  in 100 mM sodium phosphate (pH 7.2). Time points were taken by mixing sample containing 5  $\mu\text{g}$  of KS3AT3 or M3 or M3TE with 5%  $\beta$ -mercaptoethanol in SDS loading dye. Samples were then separated on a 4-20% SDS-PAGE gel at 200 V for 30 min, and stained with SimplyBlue SafeStain (Invitrogen).

Acylation of the active site thiol of KS3AT3, M3 or M3TE was achieved by incubating 250  $\mu\text{M}$  protein with 5 mM compound 1 (Figure 2) in 100 mM phosphate, pH 7.2, for 30 min. The mixture was buffer exchanged to 400 mM sodium phosphate, pH 7.2 by using a Zeba<sup>TM</sup> Spin Desalting Column (MWCO 7k), immediately prior to crosslinking. To simulate the post-elongation state of DEBS module 3, the acylated protein after desalting, was incubated with 1 mM methylmalonyl-CoA for 1 min before subjecting it to crosslinking with the upstream ACP.

### **[<sup>14</sup>C]-SDS-PAGE Labeling Assays**

**General Procedure for [<sup>14</sup>C] Labeling of the DEBS Modules:** Assays to measure the total occupancy of DEBS modules were performed using a radio-SDS-PAGE protocol. The general procedure (depicted in Figure S3) for occupancy analysis was as follows: 1) perform an *in vitro* time course assay of reconstituted derivatives of DEBS; 2) subject samples drawn at individual time points to SDS-PAGE; 3) determine the amount of radioactivity within each SDS-PAGE protein band using phosphorimaging; and 4) convert radioactive counts into protein concentrations using a <sup>14</sup>C-labeling standard curve, prepared as detailed below. While this general procedure has been reported previously<sup>13</sup>, the current procedure contains revisions to better suit the analysis of larger proteins. Assays were performed on a 45  $\mu\text{L}$  scale containing 400 mM sodium phosphate (pH = 7.2), 5 mM TCEP, 10 mM  $\text{MgCl}_2$ , 2.5 mM CoASH, 2 mM ATP, 1 mM methylmalonate, and 1 mM [<sup>14</sup>C]-propionate. Enzymes MatB (2  $\mu\text{M}$ ) and methylmalonyl-CoA epimerase (4  $\mu\text{M}$ ) were added to convert methylmalonate into racemic methylmalonyl-CoA,<sup>11</sup> and PrpE (1  $\mu\text{M}$ ) was added to convert propionate into propionyl-CoA. The concentration of both these

enzymes was selected to ensure that supply of acyl-CoA was not rate limiting. This mixture was pre-incubated at room temperature for 30 min to ensure full conversion to the acyl-CoA substrates. Reactions were initiated upon addition of the previously mentioned components directly to the DEBS proteins (protein concentrations were 2  $\mu$ M for reactions involving LDD and M1 and a concentration of 1  $\mu$ M for proteins in all other reactions) and were quenched at designated time-points by mixing 10  $\mu$ L of the reaction mixture with 10  $\mu$ L of Laemmli buffer lacking a reducing agent. The quenched samples were applied to a Mini-PROTEAN TGX precast 7.5% polyacrylamide gel (Bio-Rad), which was developed at 200 V for 40 min in an ice bath. The gel was washed with distilled water for 10 min, stained with SimplyBlue SafeStain (Invitrogen) for 10 min, and washed again with distilled water for 10 min. The washed gel was mounted onto filter paper and dried *in vacuo* using a Bio-Rad 543 Gel Dryer at for 2 h. The final gel was imaged lane by lane for quantification of  $^{14}$ C in a Rita Star TLC Analyzer (Raytest).

**[ $^{14}$ C]-Labeled Protein Standard Curves (Table 1 and Figures S6, S8, and S10):** Radioactive counts bound to a given protein were converted into a concentration of this  $^{14}$ C-labeled protein using a standard curve (counts versus concentration). Known amounts (in  $\mu$ M) of  $^{14}$ C-labeled DEBS proteins were generated using a well-established ACP loading procedure featuring the Sfp phosphopantetheinyl transferase<sup>4,14</sup>. By supplying  $^{14}$ C-propionyl-CoA to Sfp, a single equivalent of  $^{14}$ C can be loaded onto the ACP domain of an *apo*-DEBS protein. The Sfp assays were performed on a 10  $\mu$ L scale containing 400 mM sodium phosphate (pH = 7.2), 5 mM TCEP, 10 mM MgCl<sub>2</sub>, 500  $\mu$ M CoASH, and 4 mM ATP, and 1 mM [1- $^{14}$ C]-propionate (specific activity of 57 mCi / mmol). PrpE (1  $\mu$ M) was added to convert propionate into propionyl-CoA. The concentration of CoASH was selected to be limiting (500  $\mu$ M) to assure that all CoASH was converted to [1- $^{14}$ C]-propionyl-CoA. This mixture was pre-incubated at room temperature for 30 min to ensure full conversion to the acyl-CoA substrates. Sfp labeling reactions were initiated upon addition of the mixture components directly to the *apo*-DEBS modules of interest (over a range of 0.2, 1, 3, and 6  $\mu$ M) and allowed to incubate at room temperature for 12 h to assure full pantetheinylation. Reactions were quenched by mixing 10  $\mu$ L of the reaction mixture with 10  $\mu$ L of Laemmli buffer lacking a reducing agent. The quenched samples were applied to the wells of a Mini-PROTEAN TGX precast 7.5% polyacrylamide gel (Bio-Rad) which was developed at 200 V for 40 min. The gel was washed with distilled water for 10 min, stained with SimplyBlue SafeStain (Invitrogen) for 10 min, and washed again with distilled water for 10 min. The washed gel was mounted onto filter paper and dried *in vacuo* using a Bio-Rad 543 Gel Dryer at for 2 h. The final gel was imaged lane by lane for  $^{14}$ C quantification in a Rita Star TLC Analyzer (Raytest). This procedure was applied to 4 *apo*-modules of interest: *apo*-LDD/AT<sub>L</sub><sup>0</sup>, *apo*-M1, *apo*-M2+TE<sup>0</sup>, and *apo*-DEBS2. Imager counts were plotted against the  $^{14}$ C-labeled protein concentration in each case, and the resulting data was fitted using a linear regression in MATLAB. Standard curves and their conversion functions are presented in Figure S4. Converted labeling data (for 1-, 2-, 4-, and 5-module DEBS variants) is presented in Figure S6, Figure S8, and Figure S10. Steady state labeling data is tabulates in Table 1.

**Radiolabeling assay to demonstrate that polyketide chain elongation is required to restrict premature access of the KS by the upstream ACP (Figure 4):**

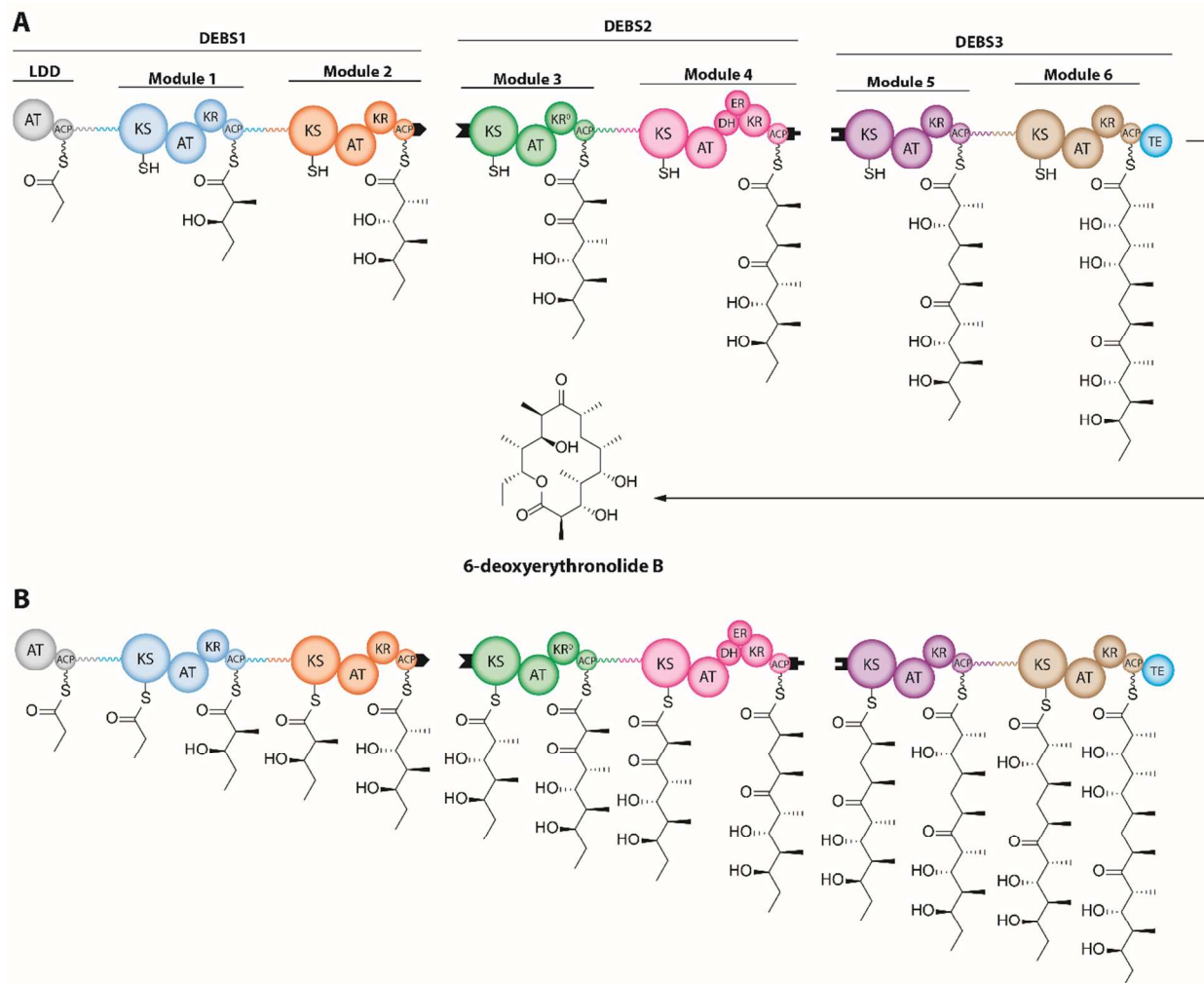
The Coenzyme A thioester analog of diketide **1** was used to post-translationally modify either *apo*-module 1 (Figure 4, Panel B) or *apo*-module 2+TE<sup>0</sup> (Figure S11, Panel B) using the Sfp phosphopantetheinyl transferase under the following conditions: either 4  $\mu$ M loading di-domain and 4  $\mu$ M *apo*-module 1 (Panel A), or 2  $\mu$ M loading di-domain, 2  $\mu$ M *holo*-module 1, and 2  $\mu$ M *apo*-module 2+TE<sup>0</sup> was incubated with 0.5 mM CoA thioester, 1  $\mu$ M Sfp in 400 mM sodium phosphate (pH 7.2), 10 mM MgCl<sub>2</sub>, and 5 mM TCEP. The Sfp reaction was allowed to proceed at 37 °C for 30 min. Meanwhile

polyketide chain elongation substrates were generated in a separate reaction under the following conditions: 1  $\mu$ M PrpE, 1  $\mu$ M MatB, and 4  $\mu$ M SCME were allowed to incubate in 400 mM sodium phosphate (pH 7.2), 10 mM  $MgCl_2$ , and 5 mM TCEP, 1 mM NADPH, 3 mM ATP, 2 mM CoASH, 1 mM [ $^{14}C$ ]-labeled propionic acid, and 1 mM methylmalonate. The substrate generation reaction was allowed to proceed for 30 min at room temperature. The reaction was initiated upon combining equal volumes of the reaction containing DEBS proteins with that containing elongation substrates and were quenched at designated time-points by mixing 10  $\mu$ L of the reaction mixture with 10  $\mu$ L of Laemmli buffer lacking a reducing agent. The quenched samples were applied to a Mini-PROTEAN TGX precast 7.5% polyacrylamide gel (Bio-Rad), which was developed at 200 V for 45 min in an ice bath. The gel was washed with distilled water for 10 min, stained with SimplyBlue SafeStain (Invitrogen) for 10 min, and washed again with distilled water for 10 min. The washed gel was mounted onto filter paper and dried *in vacuo* using a Bio-Rad 543 Gel Dryer at for 2 h. The final gel was imaged lane by lane for quantification of  $^{14}C$  in a Rita Star TLC Analyzer (Raytest). For Panel A of Figure 4 and Panel A of Figure S11, data points were taken from Figure S6-D and Figure S8-B, respectively.

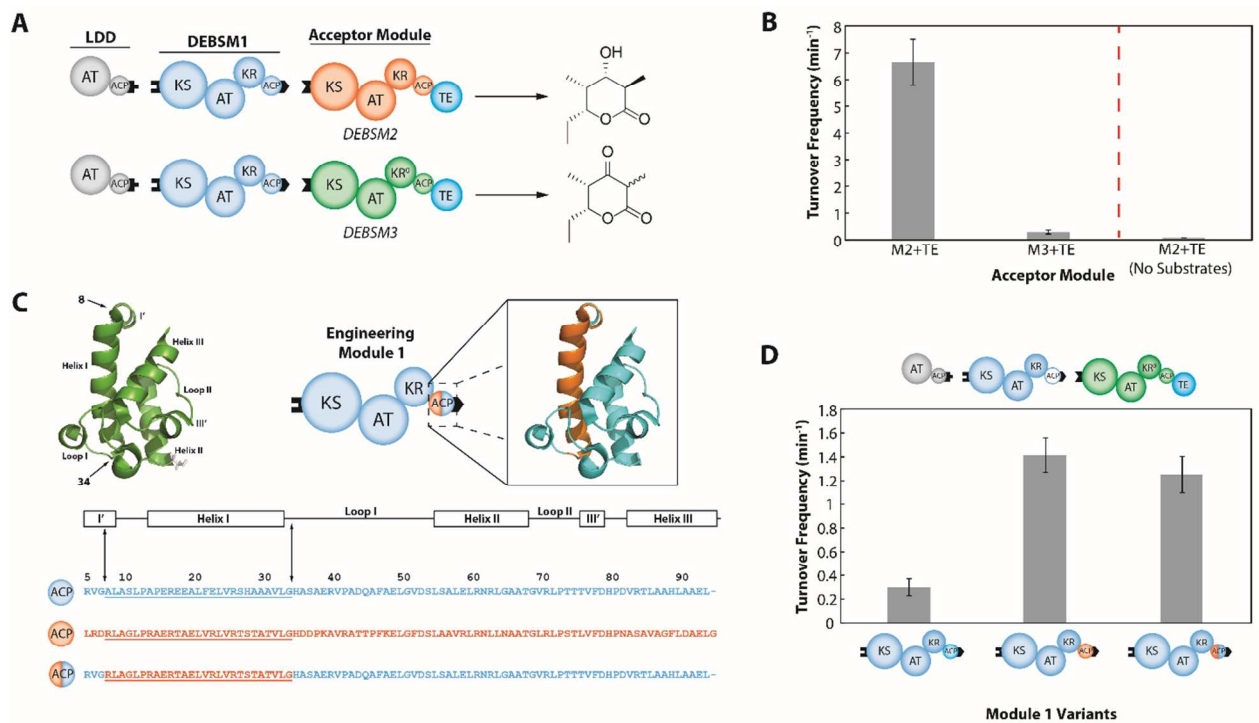
### ***LC-MS Analysis of Proteolytically-Digested Acyl-ACP Peptides (Figures S7 and S9)***

Proteomic analysis of DEBS reactions were performed on a 10  $\mu$ L scale (approximately 10  $\mu$ g of total protein) containing 400 mM sodium phosphate (pH = 7.2), 5 mM TCEP, 10 mM  $MgCl_2$ , 2.5 mM CoASH, 2 mM ATP, 1 mM methylmalonate, and 1 mM [ $^{14}C$ ]-propionate. Enzymes MatB (2  $\mu$ M) and methylmalonyl-CoA epimerase (4  $\mu$ M) were added to convert methylmalonate into racemic methylmalonyl-CoA<sup>11</sup>, and PrpE (1  $\mu$ M) was added to convert propionate into propionyl-CoA. The concentration of both these enzymes was selected to ensure that acyl-CoA supply was not rate limiting. This mixture was pre-incubated at room temperature for 30 min to ensure full conversion to the acyl-CoA substrates. Reactions were initiated upon addition of the previously mentioned components directly to the DEBS proteins (2  $\mu$ M) and were quenched after 15 min upon addition of 10  $\mu$ L of 8 M urea in 50 mM sodium phosphate, pH = 7.2. The denatured proteins were alkylated upon addition of N-ethylmaleimide (15 mM final concentration) in the dark for 30 min. The alkylated proteins were buffer exchanged into 50 mM sodium phosphate by using a Zeba<sup>TM</sup> Spin Desalting Column (eluted using 100  $\mu$ L of sodium phosphate). Trypsin/Lys-C (1 mg/mL stock) was added to the eluent in a protein: protease ratio of 5:1, and the digestion was incubated at 37  $^{\circ}C$  for 30 min. The digests were quenched upon addition of 5  $\mu$ L of formic acid. Peptides were purified using the well-established STAGE-tip procedure<sup>15</sup> and eluted into 80% (v/v) acetonitrile, 5% (w/v) formic acid in water. The eluent was lyophilized and reconstituted into 5% (w/v) formic acid in water to give an approximate peptide concentration of 500 fmol/ $\mu$ L. A 2- $\mu$ L portion of each sample was loaded onto a HPLC (Eksigent<sup>®</sup> ekspert<sup>TM</sup> nanoLC 425) coupled to an Orbitrap Elite Hybrid Ion Mass Spectrometer (Thermo Scientific). Peptides were separated using an 18 cm C18 HPLC column and eluted with a linear acetonitrile gradient of 2-98% over 40 min. MS-MS fragmentation was performed using collision induced dissociation (CID). Peptide mass location/identification was performed using algorithms within the GFY online proteomics analysis software (Harvard Medical School) and subsequently confirmed using the Thermo XCalibur Qual-Browser software (Thermo Fisher Scientific). Plots generated using the Thermo XCalibur Qual-Browser software are reported in **Error! Reference source not found.**Figure S7 and Figure S9.

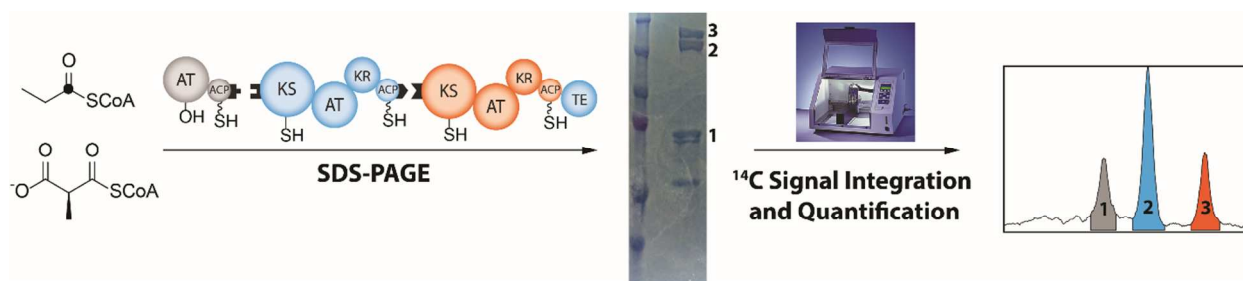
## SUPPLEMENTARY FIGURES AND RESULTS



**Figure S1.** *Alternative representations of the 6-Deoxyerythronolide B Synthase (DEBS).* DEBS1, DEBS2, and DEBS3 are the three constituent polypeptides of DEBS. Each protein harbors two sets of homodimeric modules containing the appropriate combination of catalytic domains that serve as catalysts of polyketide chain elongation and modification. Modules are shown in distinct colors; black tabs indicate peptide sequences enabling recognition between successive modules on distinct polypeptides. Chain initiation is catalyzed by the loading didomain (LDD), while chain release is catalyzed by the thioesterase (TE). Each monomeric subunit of the homodimeric module has two distinct sulfhydryl sites for covalent attachment of biosynthetic intermediates – one corresponding to the core active site cysteine of the ketosynthase (KS) domain and the other provided by the terminal thiol of the pantetheinyl side chain of the acyl carrier protein (ACP) domain. **(A)** shows the assembly line according to the standard convention of tethering the final product of each module to its ACP, whereas **(B)** shows the same assembly line but with the maximum theoretical occupancy under steady state turnover conditions. AT = acyl transferase; KR = ketoreductase; KR<sup>0</sup> = inactive ketoreductase; DH = dehydratase; ER = enoyl reductase.

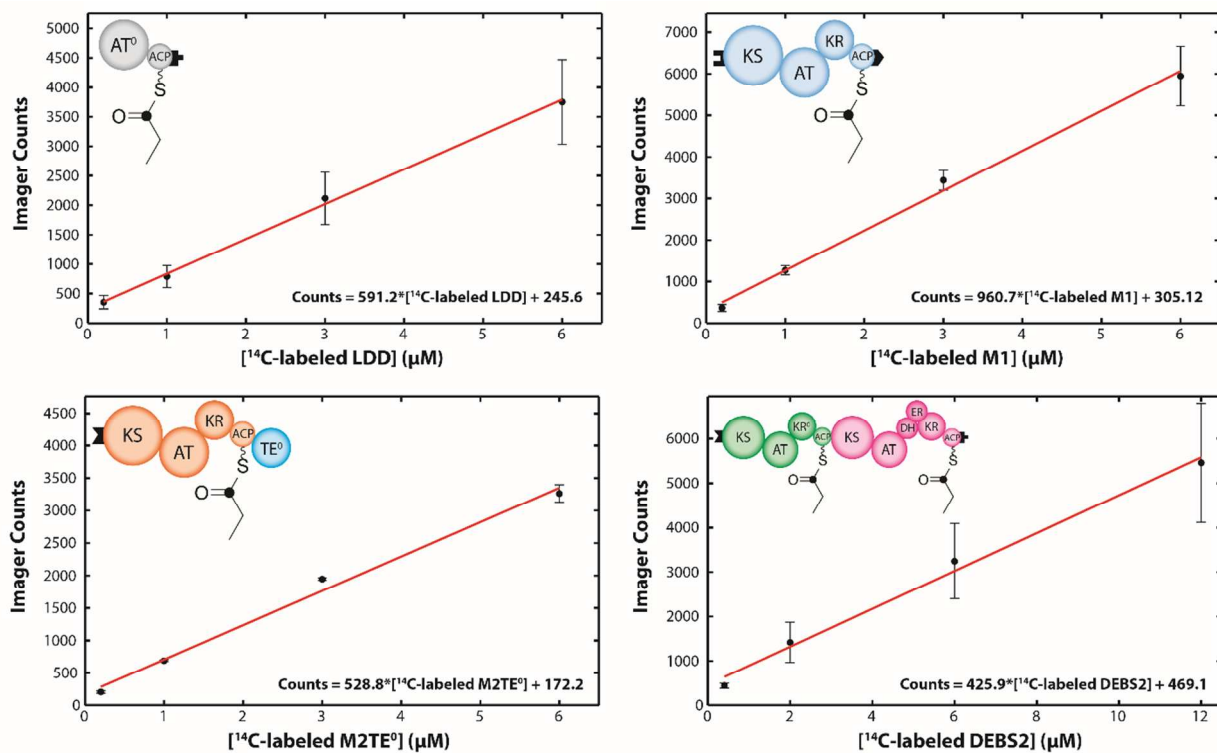


**Figure S2. Kinetic Analysis and Engineering of Bimodular PKSs.** (A) Two bimodular PKSs harboring the loading didomain (LDD), module 1, and either module 2+TE (orange) or module 3+TE (green) of DEBS are shown with their corresponding triketide lactone products. (B) The turnover frequency of each bimodular system was calculated using a spectrophotometric assay. A control reaction lacking CoA-linked substrates is shown on the right as a reference. (C) An engineering strategy for improvement of bimodular turnover. (Top Left) The NMR structure of DEBS ACP2 is shown with structural annotations. Fusion site residue numbers for the chain translocation epitope<sup>16</sup> are identified as black arrows. (Bottom) A linear map of the ACP secondary structure is shown along with multiple sequence alignment of DEBS ACP1 and ACP2. Numbering corresponds to that of the ACP2 NMR structure, PDBID: 2JU1<sup>17</sup>. The chain translocation epitope residues are underlined. The sequence of the engineered ACP is also indicated. (Top Right) The engineered DEBS module 1, in which the helix I residues 8-34 of ACP1 have been exchanged for the corresponding residues in ACP2, is shown as schematic and colored according to the alignments. (D) The rate of turnover was calculated for each LDD-module 1-module 3+TE system using a spectrophotometric assay: (left bar) module 1, (middle bar) module 1 variant where the entire ACP1 domain has been exchanged for ACP2, and (right bar) module 1 variant where a portion of the ACP1 residues have been exchanged for those from ACP2, as described in (C).

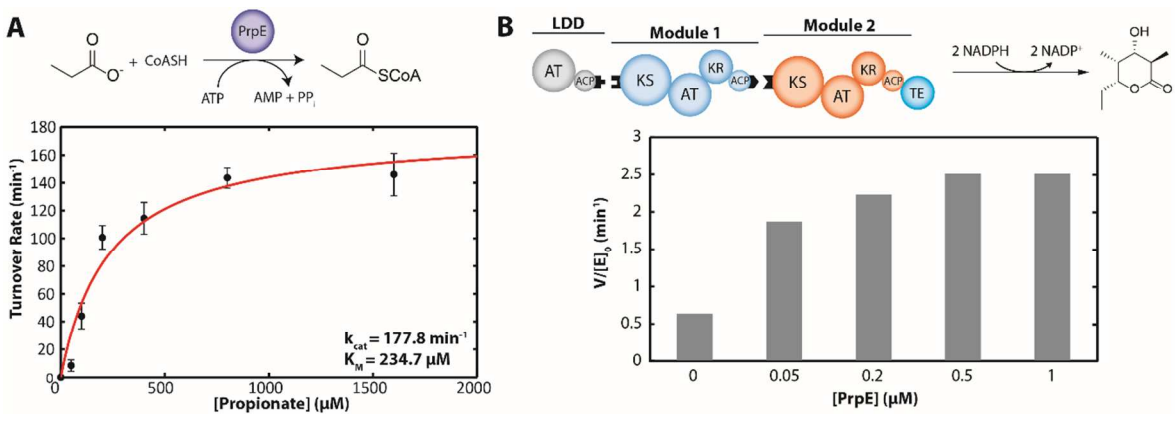


**Figure S3.**  $^{14}\text{C}$ -SDS-PAGE Protocol Summary. Radiolabeling of the DEBS proteins was performed using  $[1\text{-}^{14}\text{C}]$ -propionyl-CoA (derived in situ from propionate) as the sole radiolabel source. Proteins were incubated with their small molecule substrates,  $^{14}\text{C}$ -propionyl-CoA, methylmalonyl-CoA, and NADPH, and subjected to SDS-PAGE. SDS-PAGE gels were imaged for  $^{14}\text{C}$  counts. Each gel band signal was integrated and converted to concentration of labeled protein, as described below.

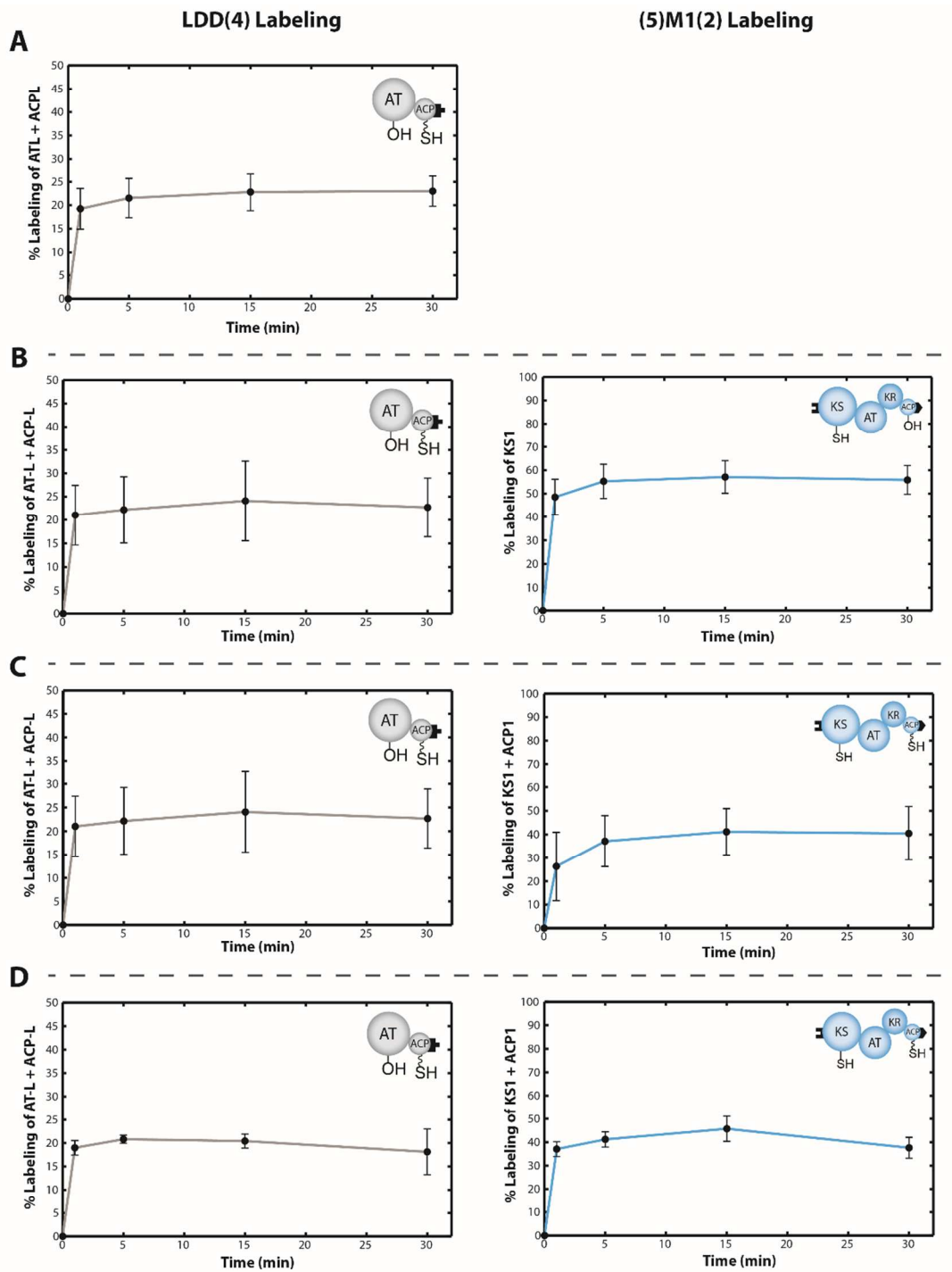




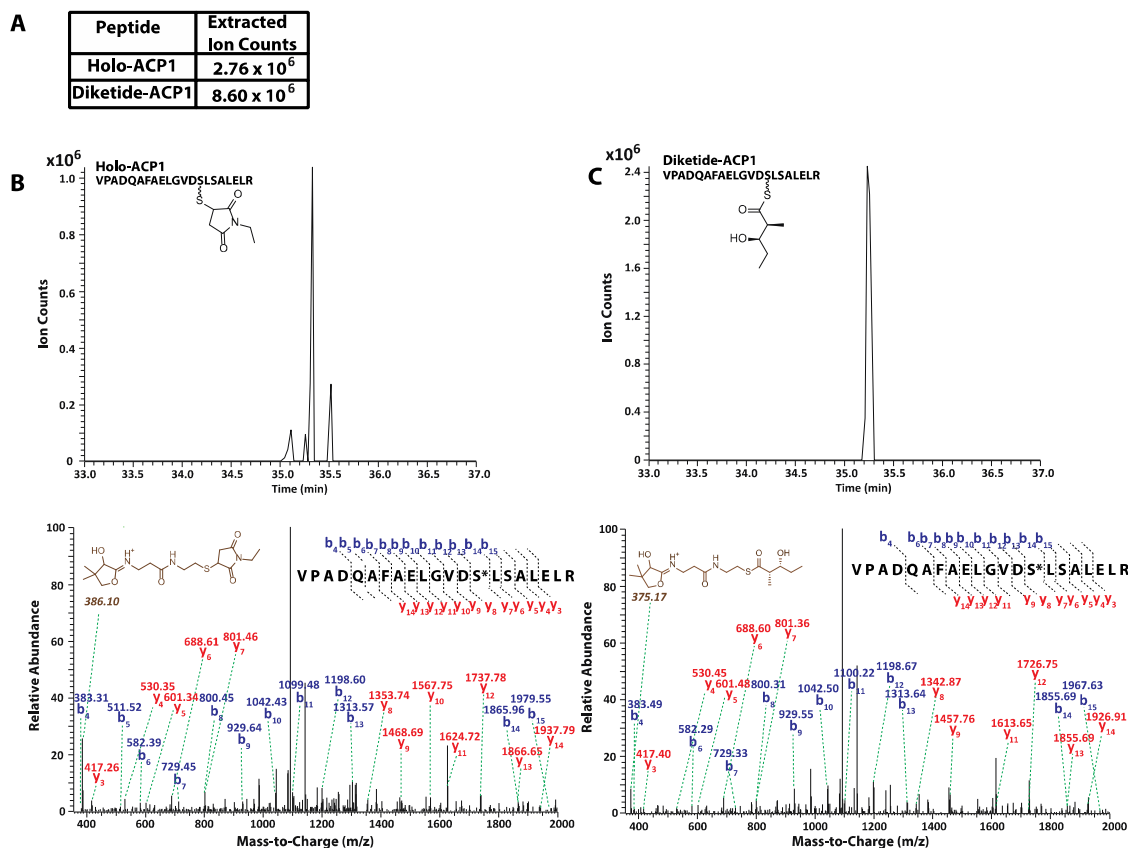
**Figure S4.** Standard [ $^{14}\text{C}$ ] Labeling Plots for *Sfp*-labeled DEBS Proteins. Each graph plots Imager Counts vs. the concentration of  $^{14}\text{C}$ -labeled protein. Note that for DEBS2, there are 2  $\mu\text{M}$  of labeling sites (ACP3 and ACP4) for every 1  $\mu\text{M}$  of DEBS2.



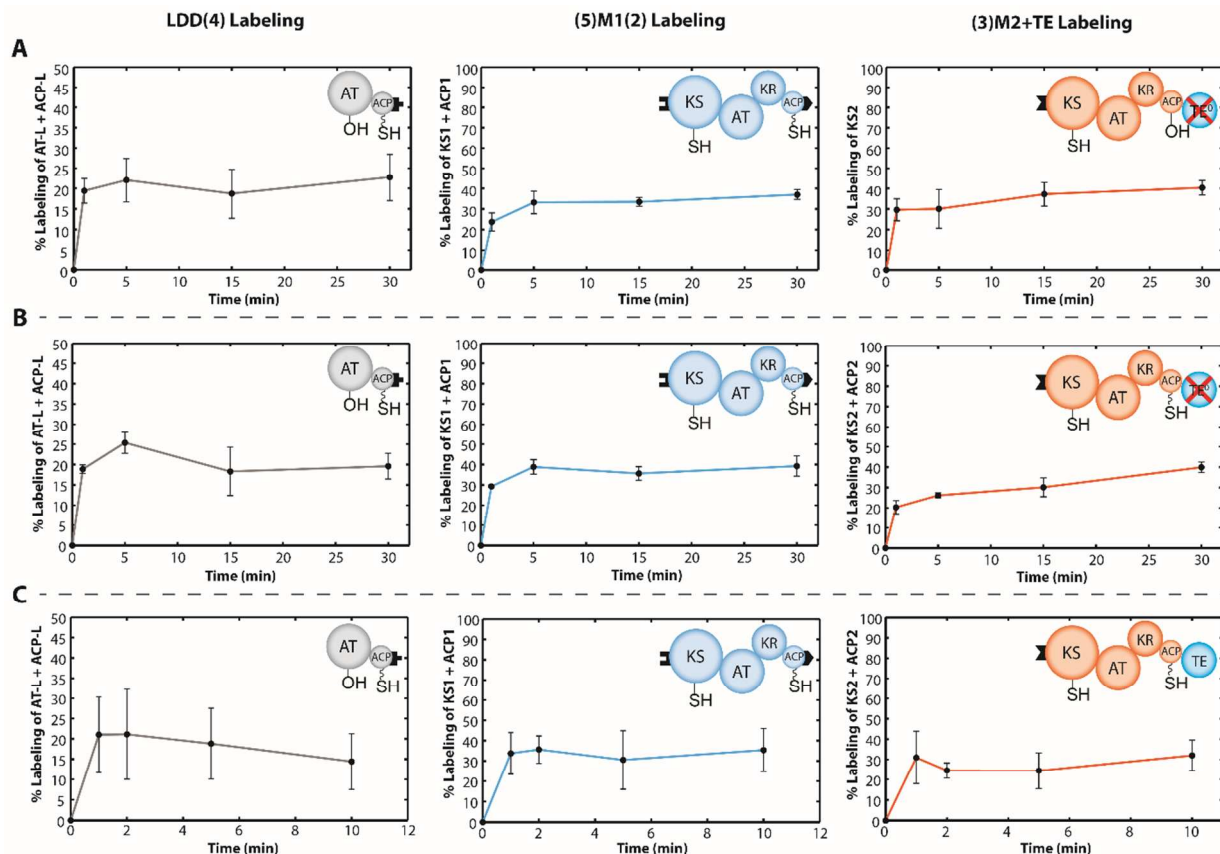
**Figure S5.** Use of the *E. coli* Propionyl-CoA Synthetase, PrpE. **(A)** PrpE turnover was measured by monitoring the depletion of CoASH for different propionate concentrations. The resulting data was fitted to a hyperbolic equation to determine  $k_{cat}$  and  $K_M$ . **(B)** Turnover of bimodular DEBS was monitored with PrpE used to generate propionyl-CoA *in situ*. PrpE does not limit DEBS turnover at or above 500 nM.



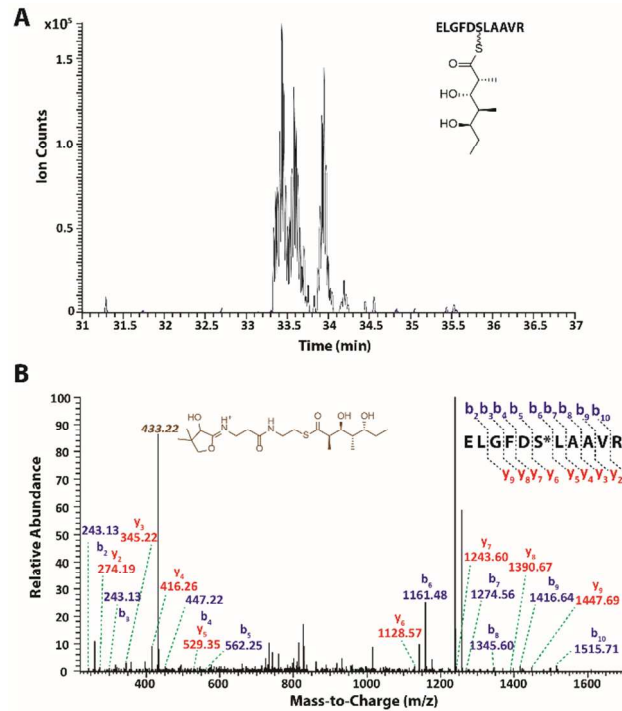
**Figure S6.**  $[^{14}\text{C}]$  Labeling of LDD(4) in the absence or presence of (5)M1(2). 2  $\mu\text{M}$  of each DEBS protein was added to the assays, and the percentage occupancy reported on the y-axis is based on a theoretical maximum labeling. **(A)** LDD(4) and  $[^{14}\text{C}]$ -propionyl-CoA. **(B)** LDD(4), apo-(5)M1(2), and  $[^{14}\text{C}]$ -propionyl-CoA. **(C)** LDD(4), holo-(5)M1(2), and  $[^{14}\text{C}]$ -propionyl-CoA. **(D)** LDD(4), holo-(5)M1(2), NADPH,  $[^{14}\text{C}]$ -propionyl-CoA, and methylmalonyl-CoA.



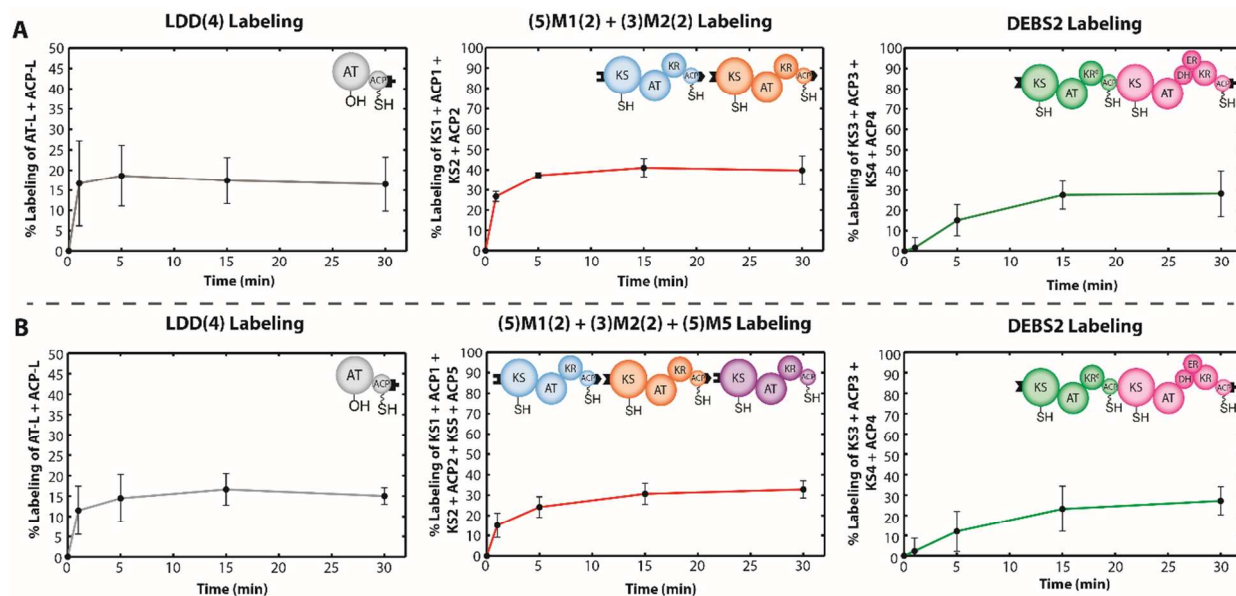
**Figure S7: Proteomic LC-MS Analysis of acyl and alkyl-ACP1 active site peptides.** Proteomic LC-MS analysis of LDD(4), holo-(5)M1(2), NADPH, methylmalonyl-CoA, and propionyl-CoA (generated as in Figure S6D) (A) Extracted ion counts for the alkylated holo-ACP1 peptide (sequence: VPADQFAELGVDS\*LSALELR; theoretical  $m/z = 1333.6484$ ,  $z = +2$ , where \* denotes the alkylation of the phosphopantetheine arm by N-ethylmaleimide) and diketide-ACP1 peptide (sequence: VPADQFAELGVDS\*LSALELR; theoretical  $m/z = 1328.1571$ ,  $z = +2$ , where \* denotes the acylation of the phosphopantetheine arm by module 1's natural diketide product). (B) Extracted ion chromatogram for alkylated holo-ACP1 peptide (top) and MS-MS fragmentation by CID. B and Y ions are indicated in blue and red, respectively. A putative phosphopantetheine ejection fragment<sup>18</sup> corresponding to phosphopantetheine alkylated with N-ethylmaleimide is identified in brown (theoretical  $m/z = 386.17$ ; observed  $m/z = 386.10$ ). (C) Extracted ion chromatogram for acylated diketide-ACP1 peptide (top) and MS-MS fragmentation by CID. B and Y ions are indicated in blue and red, respectively. A putative phosphopantetheine ejection fragment corresponding to phosphopantetheine alkylated with N-ethylmaleimide is identified in brown (theoretical  $m/z = 375.20$ ; observed  $m/z = 375.17$ ). The acyl-ACP1 species identified in (C) could not be found in experiments involving LDD(4), holo-(5)M1(2), and propionyl-CoA but lacking NADPH and methylmalonyl-CoA.



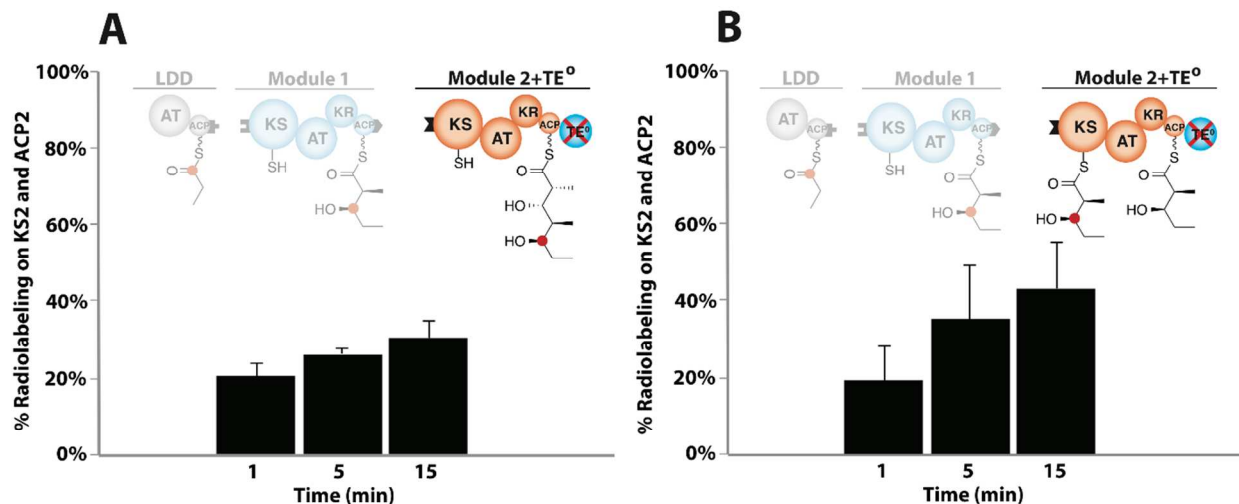
**Figure S8.**  $[^{14}\text{C}]$  Labeling of LDD(4), (5)M1(2), and (3)M2+TE or (3)M2+TE<sup>0</sup>. 1  $\mu\text{M}$  of each DEBS protein was added to the assays, and the percent occupancy reported on the y-axis is based on a theoretical maximum labeling. **(A)** LDD(4), (5)M1(2), *apo*-(3)M2+TE<sup>0</sup>, NADPH,  $[^{14}\text{C}]$ -propionyl-CoA, and methylmalonyl-CoA. **(B)** LDD(4), (5)M1(2), *holo*-(3)M2+TE<sup>0</sup>, NADPH,  $[^{14}\text{C}]$ -propionyl-CoA, and methylmalonyl-CoA. **(C)** LDD(4), (5)M1(2), *holo*-(3)M2+TE, NADPH,  $[^{14}\text{C}]$ -propionyl-CoA, and methylmalonyl-CoA.



**Figure S9. Proteomic LC-MS Analysis of the Triketide Acyl-ACP2 Active Peptide.** Extracted ion chromatograms for the triketide acyl-ACP2 (sequence: ELGFDS\*LA AVR; theoretical  $m/z = 845.4155$ ,  $z = +2$  where \* denotes the acylation of the phosphopantetheine arm with triketide) for (A) LDD(4), *holo*-(5)M1(2), *holo*-(3)M2+TE<sup>0</sup>, NADPH, methylmalonyl-CoA, and propionyl-CoA (as in Figure S8B). The acyl-ACP2 peptide elutes between 33.3-34.1 min. (B) MS-MS fragmentation by CID of the acyl-ACP2 peptide located in (A). B and Y ions are indicated in blue and red, respectively. A putative phosphopantetheine ejection fragment<sup>18</sup> corresponding to phosphopantetheine acylated with triketide is identified in brown (theoretical  $m/z = 433.22$ ; observed  $m/z = 433.24$ ).



**Figure S10.**  $[^{14}\text{C}]$  Labeling of LDD(4), (5)M1(2), (3)M2(2), (5)M5, and DEBS2. 1  $\mu\text{M}$  of each DEBS protein was added to the assays, and the percent occupancy reported on the y-axis is based on a theoretical maximum labeling. **(A)** LDD(4), (5)M1(2), apo-(3)M2+TE<sup>0</sup>, NADPH,  $[^{14}\text{C}]$ -propionyl-CoA, and methylmalonyl-CoA. **(B)** LDD(4), (5)M1(2), holo-(3)M2+TE<sup>0</sup>, NADPH,  $[^{14}\text{C}]$ -propionyl-CoA, and methylmalonyl-CoA. **(C)** LDD(4), (5)M1(2), holo-(3)M2+TE, NADPH,  $[^{14}\text{C}]$ -propionyl-CoA, and methylmalonyl-CoA.



**Figure S11. Energetic coupling of the turnstile to polyketide chain elongation.** In Panel A, the ACP-bound diketide product of DEBS module 2+TE<sup>0</sup> was generated by mixing *holo*-module 2+TE<sup>0</sup> with DEBS loading didomain (LDD), DEBS module 1, 1-<sup>14</sup>C-propionyl-CoA, methylmalonyl-CoA, and NADPH. Consistent with data shown in Table 1B and Figure S8, radiolabel accumulated rapidly on module 2+TE<sup>0</sup> and attained steady state corresponding to ~30% occupancy of the ACP+KS sites. In contrast, in Panel B the ACP domain of *apo*-module 2+TE<sup>0</sup> was directly loaded with the formal diketide product of the chain elongation reaction of module 1. This was accomplished using the Coenzyme A thioester analog of unlabeled diketide **1** in the presence of Sfp phosphopantetheinyl transferase. When the resulting module was mixed with LDD, module 1, 1-<sup>14</sup>C-propionyl-CoA, methylmalonyl-CoA, and NADPH, it was rapidly radiolabeled to a comparable steady state occupancy level, indicative of efficient translocation of <sup>14</sup>C-propionyl units onto the KS domain of module 2+TE<sup>0</sup>. In both Panels the occupancy values are reported as the mean ± SD (n = 3) of the ACP+KS occupancy for DEBS module 2+TE<sup>0</sup>.



**Table S1. Plasmid Design and Primers**

Plasmid Encoded Protein	Cloning Method	PCR Fragment	Primer Name	Primer Sequence	PCR Template
pBL48 <i>E. coli</i> Propionyl-CoA Synthetase (PrpE)	Restriction Digestion / Ligation	PrpE Insert	pBL48_Nterm	AAAAAACATATGATGTCTTTT AGCGAATTTATCAGCG	<i>E. coli</i> BL21 (DE3) Genomic DNA
			pBL48_Cterm	AAAAAAGAATTCCTACTCTTC CATCGCCTGGC	
pBL10 M5	Restriction Digestion / Ligation	M5 Insert	pBL10_Nterm	AAAAAACATATGAGCGGTGA CAACGGCATG	pFW100 <sup>6</sup>
			pBL10_Cterm	AAAAAAGAATTCGCGACGAG CCGCTCCAGGTAG	
pBL60 M2+TE <sup>0</sup>	Site-Directed Mutagenesis		pBL60_QC_fwd	GTGGCCGGTCACGCGGCGGG GGCACTG	pBL16 <sup>6</sup>
			pBL60_QC_rev	CAGTGCCCCCGCCGCGTGACC GGCCAC	
pBL53 LDD/ AT-L <sup>0</sup>	Site-Directed Mutagenesis		pBL53_QC_fwd	GTGGTGGGCCACGCGATCGG CGAGCTG	pBL12 <sup>6</sup>
			pBL53_QC_rev	CAGCTCGCCGATCGCGTGGCC CACCAC	
pBL24 M1/ACP2	3-fragment Gibson Assembly	1. M1	pBL24_M1_Nterm	GATATACATATGAGCGGTGAC AACGGCATGAC	pBL13 <sup>6</sup>
			pBL24_M1_Cterm	CGCCAGCCGGTCGCGCAG CGGTTGCGCCCGCCTGCG	
		2. ACP2	pBL24_ACP2_Nterm	CGCAGGCGGGCGCCGAACCG CTGCGGACCGGCTGGCG	pNW7 <sup>4</sup>
			pBL24_ACP2_Cterm	CTCCCCCGGACCTCGGT GCCGAGCTCGGCGTCGAG	
		3. M2 DD and pET21 Vector	pBL24_vector_Nterm	CTCGACGCCGAGCTCGGC ACCGAGTCCGGGGGGAG	pBL13 <sup>6</sup>
			pBL24_vector_Cterm	GTCATGCCGTTGTCACCGCTC ATATGTATATC	
pBL26 M1/	3-fragment Gibson	1. M1	pBL24_M1_Nterm	GGAATTGTGAGCGGATAACA ATTCCCC	pBL13 <sup>6</sup>

ACP2-Helix I	Assembly		pBL26_M1_Cterm	GGCAGACCCGCCAGCCG GCCACCCGCGGTTTCGGC	
		2. ACP2-HI	pBL26_ACP2HI_Nterm	GCCGAACCGCGGGTGGGC CGGCTGGCGGGTCTGCC	pNW7 <sup>4</sup>
			pBL26_ACP2HI_Cterm	GGCCGAGGCGTGGCC CAGCACGGTCGCGGTG	
		3. M2 DD and pET21 Vector	pBL26_vector_Nterm	CACCGCGACCGTGCTG GGCCACGCCCTCGGCC	pBL13 <sup>6</sup>
			pBL24_vector_Cterm	GGGGAATTGTTATCCGCTCAC AATTC	

## REFERENCES

- (1) Meier, J. L., and Burkart, M. D. (2009) Chapter 9. Synthetic probes for polyketide and nonribosomal peptide biosynthetic enzymes. *Methods Enzymol.* 458, 219–254.
- (2) Cane, D. E., Lambalot, R. H., Prabhakaran, P. C., and Ott, W. R. (1993) Macrolide biosynthesis. 7. Incorporation of polyketide chain elongation intermediates into methymycin. *J. Am. Chem. Soc.* 115, 522–526.
- (3) Harris, R. C., Cutter, A. L., Weissman, K. J., Hanefeld, U., Timoney, M. C., and Staunton, J. (1998) Enantiospecific Synthesis of Analogues of the Diketide Intermediate of the Erythromycin Polyketide Synthase (PKS). *J. Chem. Res.* 6, 283.
- (4) Wu, N., Tsuji, S. Y., Cane, D. E., and Khosla, C. (2001) Assessing the balance between protein-protein interactions and enzyme-substrate interactions in the channeling of intermediates between polyketide synthase modules. *J. Am. Chem. Soc.* 123, 6465–6474.
- (5) Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345.
- (6) Lowry, B., Robbins, T., Weng, C.-H., O'Brien, R. V, Cane, D. E., and Khosla, C. (2013) In vitro reconstitution and analysis of the 6-deoxyerythronolide B synthase. *J. Am. Chem. Soc.* 135, 16809–16812.
- (7) Wu, N., Cane, D. E., and Khosla, C. (2002) Quantitative analysis of the relative contributions of donor acyl carrier proteins, acceptor ketosynthases, and linker regions to intermodular transfer of intermediates in hybrid polyketide synthases. *Biochemistry* 41, 5056–5066.
- (8) Chen, A. Y., Schnarr, N. A., Kim, C.-Y., Cane, D. E., and Khosla, C. (2006) Extender unit and acyl carrier protein specificity of ketosynthase domains of the 6-deoxyerythronolide B synthase. *J. Am. Chem. Soc.* 128, 3067–3074.
- (9) Edwards, A. L., Matsui, T., Weiss, T. M., and Khosla, C. (2014) Architectures of whole-module and bimodular proteins from the 6-deoxyerythronolide B synthase. *J. Mol. Biol.* 426, 2229–2245.
- (10) Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., and Khosla, C. (2001) Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* 291, 1790–1792.
- (11) Hughes, A. J., and Keatinge-Clay, A. (2011) Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem. Biol.* 18, 165–176.
- (12) Worthington, A. S., Rivera, H., Torpey, J. W., Alexander, M. D., and Burkart, M. D. (2006) Mechanism-based protein cross-linking probes to investigate carrier protein-mediated biosynthesis. *ACS Chem. Biol.* 1, 687–691.
- (13) Wong, F. T., Chen, A. Y., Cane, D. E., and Khosla, C. (2010) Protein-protein recognition between acyltransferases and acyl carrier proteins in multimodular polyketide synthases. *Biochemistry* 49, 95–102.

- (14) Charkoudian, L. K., Liu, C. W., Capone, S., Kapur, S., Cane, D. E., Togni, A., Seebach, D., and Khosla, C. (2011) Probing the interactions of an acyl carrier protein domain from the 6-deoxyerythronolide B synthase. *Protein Sci.* 20, 1244–1255.
- (15) Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption / Ionization , Nanoelectrospray , and LC / MS Sample Pretreatment in Proteomics. *Anal. Chem.* 75, 663–670.
- (16) Kapur, S., Chen, A., Cane, D., and Khosla, C. (2010) Molecular recognition between ketosynthase and acyl carrier protein domains of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U. S. A.* 107, 22066–22071.
- (17) Alekseyev, V. Y., Liu, C. W., Cane, D. E., Puglisi, J. D., and Khosla, C. (2007) Solution structure and proposed domain domain recognition interface of an acyl carrier protein domain from a modular polyketide synthase. *Protein Sci.* 16, 2093–2107.
- (18) Dorrestein, P. C., Bumpus, S. B., Calderone, C. T., Garneau-Tsodikova, S., Aron, Z. D., Straight, P. D., Kolter, R., Walsh, C. T., and Kelleher, N. L. (2006) Facile detection of acyl and peptidyl intermediates on thioester carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry. *Biochemistry* 45, 12756–12766.