Supporting Information: Stefanie B. Bumpus, Nathan A. Magarvey, Neil L. Kelleher, Christopher T. Walsh, and Christopher T. Calderone

Polyunsaturated Fatty Acid–Like *Trans*-Enoyl Reductases Utilized in Polyketide Biosynthesis

Stefanie B. Bumpus, Nathan A. Magarvey, Neil L. Kelleher, Christopher T. Walsh, and Christopher T. Calderone*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115 and Department of Chemistry, University of Illinois, 600 South Mathews Avenue, Urbana, Illinois 61801

Supporting Information

Contents

I. Supplementary Figures

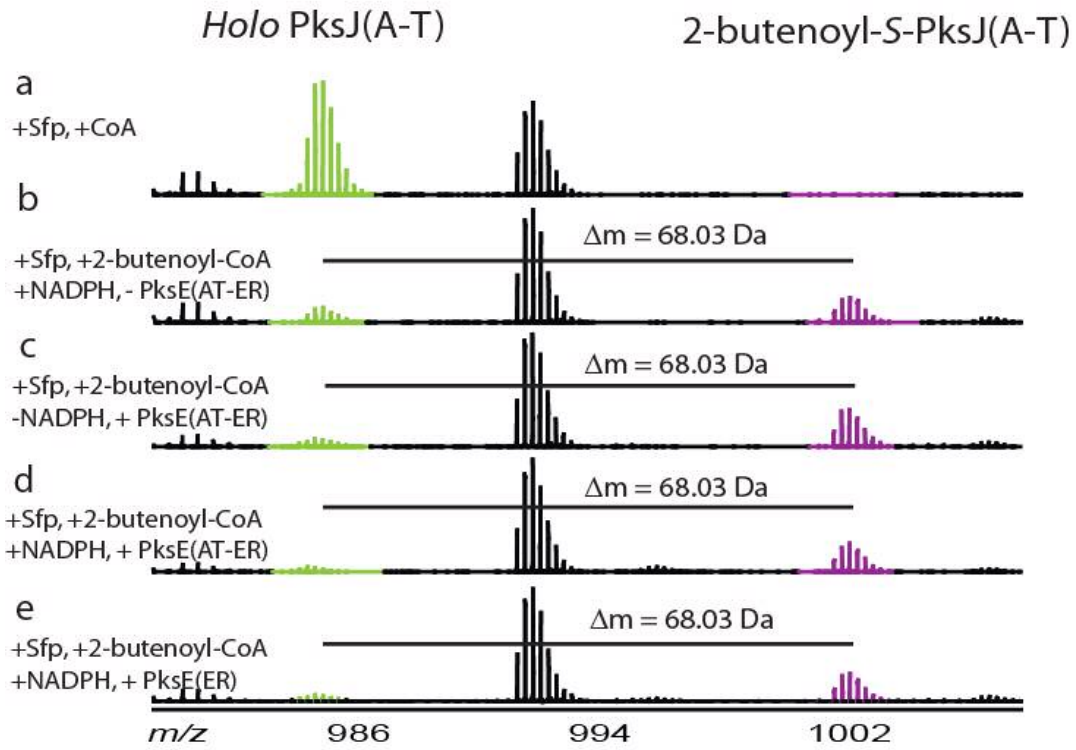
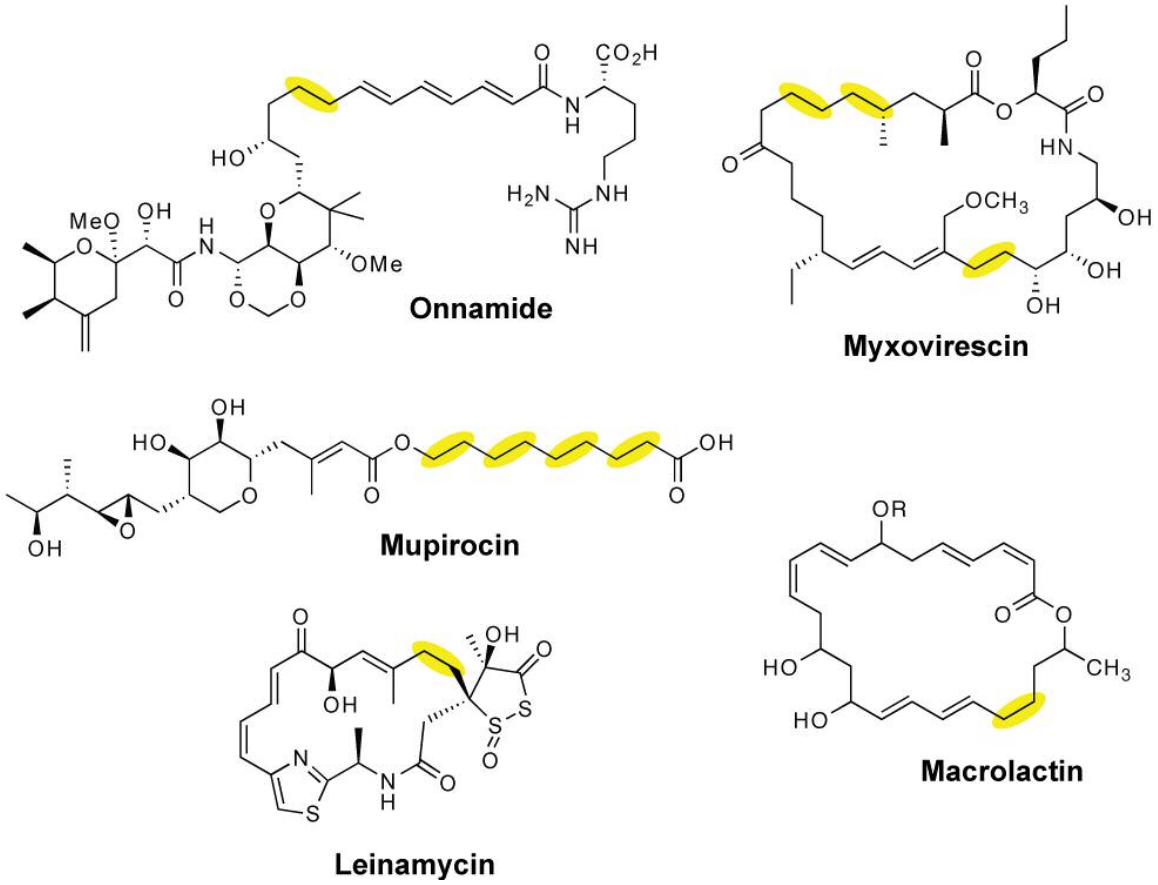


Figure S1. PksE does not show enoyl reductase activity at PksJ(A-T). PksJ(A-T) is not predicted to be a scaffold for enoyl reduction. (a) *Holo*-PksJ(A-T) (green) is observed when *apo*-PksJ(A-T) is incubated with Sfp and CoA. (b) 2-butenoyl-*S*-PksJ(A-T) (purple) is observed when *apo*-PksJ(A-T) is incubated with Sfp in the presence of 2-butenoyl-CoA. (c) No enoyl reduction is observed when 2-butenoyl-*S*-PksJ(A-T) is incubated with PksE in the absence of NADPH. (d) No enoyl reduction is observed when 2-butenoyl-*S*-PksJ(A-T) is incubated with PksE in the presence of NADPH. (e) No enoyl reduction is observed when 2-butenoyl-*S*-PksJ(A-T) is incubated with PksE(ER) in the presence of NADPH. The theoretical mass shift between *holo*-PksJ(A-T) and 2-butenoyl-*S*-PksJ(A-T) is 68.03 Da. The theoretical mass shift is 70.04 Da if enoyl reduction occurs.



Metabolite	PUFA-ER homologue	Homology to PfaD
Dihydrobacillaene	PksE (AT-ER)	47% identical/65% similar
Onnamide	PedB (ER)	53%/73%
Myxovirescin	TaN (ER)	52%/72%
Mupirocin	MmpIII (AT-ER)	50%/69%
Leinamycin	LnmG (AT-ER)	51%/71%
Macrolactin	MlnA (AT-ER)	47%/66%

Figure S2. Polyketide metabolites with PUFA-like enoyl reductases. The sites of predicted PUFA enoyl reduction are highlighted in yellow.

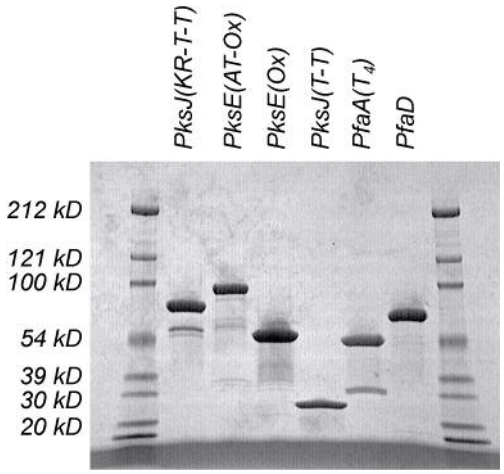


Figure S3. Gel electrophoresis of proteins utilized in this study.

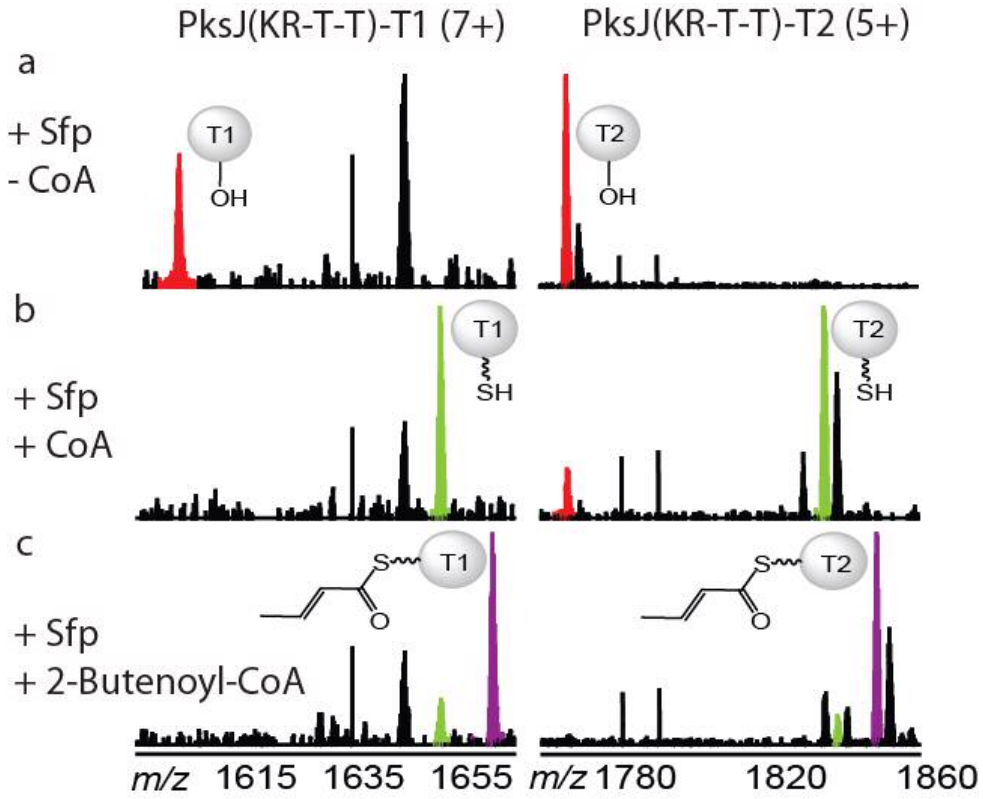


Figure S4. Identification of PksJ(KR-T-T) active-site peptides. Peptides were generated by digestion with the endoproteinase ArgC and analyzed by RPLC-FTMS. Left: PksJ-T1. Right: PksJ-T2. (a) *Apo* active-site peptides are observed in the presence of Sfp and absence of CoA (red). (b) *Holo* active-site peptides are generated in the presence of both Sfp and CoA (green). (c) 2-butenoyl-loaded active-site peptides observed in the presence of Sfp and 2-butenoyl-CoA (purple).

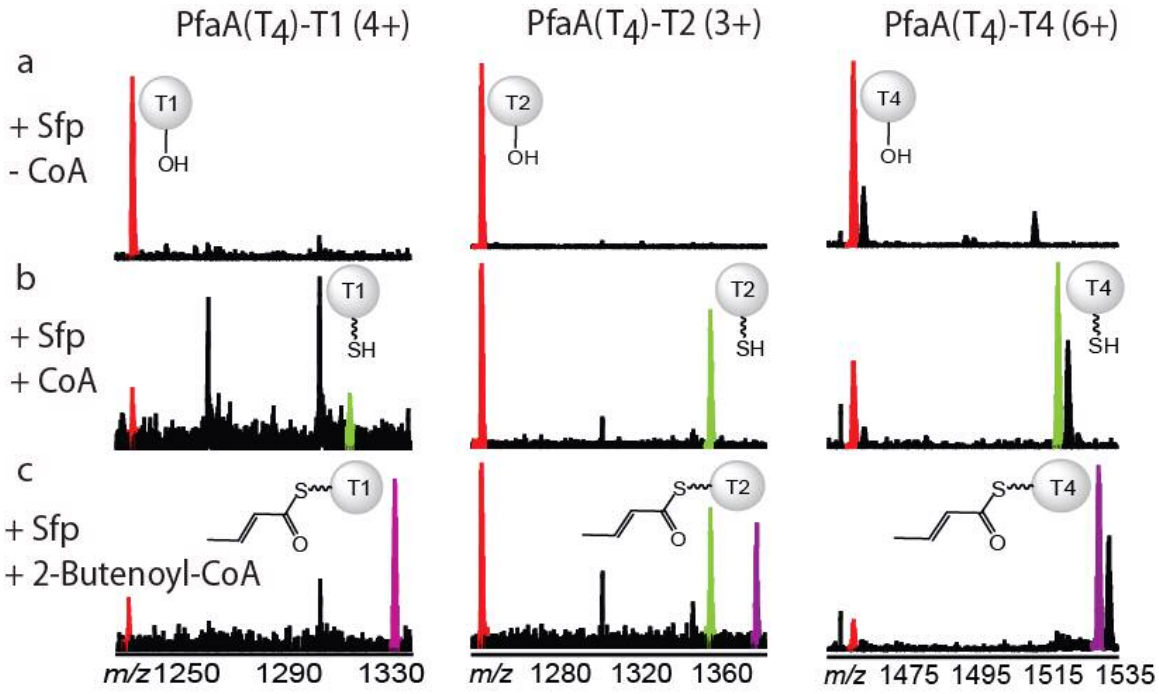


Figure S5. Identification of PfaA(T_4) active-site peptides. Peptides were generated by digestion with the endoproteinase ArgC and analyzed by RPLC-FTMS. Left: Active-site peptide from PfaA-T1. Middle: Active-site peptide from PfaA-T2. Right: Active-site peptide from PfaA-T4. (a) *Apo* active-site peptides observed in the presence of Sfp but absence of CoA (red). (b) *Holo* active-site peptides are observed when *apo*-PfaA(T_4) is incubated in the presence of Sfp and CoA (green). (c) 2-butenoyl-loaded active-site peptides are observed when *apo*-PfaA(T_4) is incubated with Sfp and 2-butenoyl-CoA (purple). Despite testing multiple digestion conditions, the active-site peptide from PfaA(T_4)-T3 was not observed.

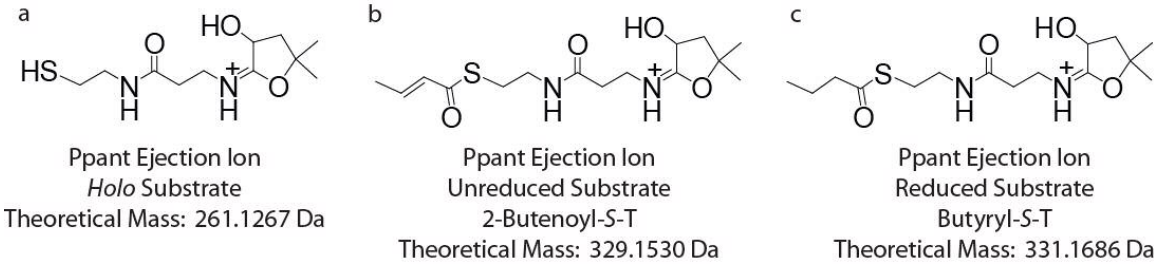


Figure S6. Theoretical phosphopantetheinyl (Ppant) ejection ions. (a) Ppant ejection ion from *holo* T domain. (b) Ppant ejection ion from T domain with unreduced 2-butenoyl substrate bound. (c) Ppant ejection ion from T domain with reduced butyryl substrate bound.

Table S1. Active-site peptides discussed in this study. All masses are reported in Da and as the monoisotopic mass of the peptide. All mass errors are reported in parts–per-million (ppm).

	Apo	*Apo*	*Apo*	*Holo*	*Holo*	*Holo*
Active-site Peptide	Theor	Exp	Error	Theor	Exp	Error
PksJ(KR-T-T)-T1[#]	11240.57	11241.51	83.63	11580.67	11581.71	89.80
PksJ(KR-T-T)-T2	8850.35	8850.35	0.00	9190.45	9190.45	0.00
PfaA(T_4)-T1	4958.40	4958.40	0.00	5298.49	5298.49	0.00
PfaA(T_4)-T2[$]	3758.80	3758.79	-2.66	4098.89	4098.88	-2.44
PfaA(T_4)-T3	8503.24	N.D.	-	8843.33	N.D.	-
PfaA(T_4)-T4[%]	8790.33	8790.32	-1.14	9130.42	9130.40	-2.19

	2-butenoyl-*S*-T	2-butenoyl-*S*-T	2-butenoyl-*S*-T
Active-site Peptide	Theor	Exp	Error
PksJ(KR-T-T)-T1[#]	11648.70	11649.68	84.31
PksJ(KR-T-T)-T2	9258.48	9258.45	-3.24
PfaA(T_4)-T1	5366.52	5366.50	-3.73
PfaA(T_4)-T2[$]	4166.91	4166.91	0.00
PfaA(T_4)-T3	8911.35	N.D.	-
PfaA(T_4)-T4[%]	9198.45	9198.43	-2.17

[#]Active-site peptide with one N-terminal missed cleavage.
[$]Active-site peptide could correspond to PfaA(T_4)-T4 as well, but is identified at PfaA(T_4)-T2 due to observation of authentic PfaA(T_4)-T4 peptide and observation in a truncated form of PfaA (data not shown).
[%]Active-site peptide with one N-terminal missed cleavage and one C-terminal missed cleavage.
N.D. – peptide not detected; Theor – theoretical monoisotopic mass; Exp – experimental monoisotopic mass.

II. Methods and materials

a. Gene cloning

Genes were amplified from genomic DNA from *B. subtilis* or *S. oneidensis* MR-1 (ATCC 700550) with Phusion High-Fidelity DNA polymerase (Finnzymes) using the manufacturer's protocol and the primers shown in Table S2. Reaction mixtures (50 µL) were subjected to the following conditions and amplicons were purified using Illustra GFX DNA purification kits (GE Healthcare):

1x	2 min	98˚C
30x	10 s	98˚C
	30 s	60˚C
	1 min	72˚C
1x	10 min	72˚C

Amplicons were digested with the appropriate restriction endonucleases (New England Biolabs), gel-purified, and ligated (T4 DNA ligase, New England Biolabs) into similarly digested and purified pET-24b [*pfaD*, *pksE(AT-ER)*, and *pksE(ER)*] or pET-28a [*pfaA(T$_4$)*, *pksJ(KR-T-T)*, and *pksJ(T-T)*] at 16˚C. 5 µL of the ligation mixtures were used to transform chemically competent TOP10 cells (Invitrogen). Identities of individual clones were verified by sequencing (Dana-Farber Cancer Institute Molecular Biology Core Facility). Sequence-confirmed plasmids were used to transform BL21*(DE3) chemically competent cells (Invitrogen) for heterologous overexpression.

Table S2. Primers used in this study. Engineered restriction sites are underlined; engineered stop codons are shown in bold.

Amplicon		Primer
pfaA(T$_4$)	*Nde*I	GGAATTCCATATGGCCTTGCAGCAAATCGAGCACGCTATGC
	*Xho*I	GGTTAACCTCGAG**TTA**ACCAGCATAGAGGGCGACAATTTCACC
pfaD	*Nde*I	GGTTAACCATATGACGAATACCACACTCGATAATAACGC
	*Xho*I	GGTTAACCTCGAGGCAGCGTTGCAGAGGTTTCCAACGTTG
pksE	*Nhe*I	GATCGCTAGCATTACATATGTCTTTCCAGGGCAAG
	*Xho*I	GATCCTCGAGAACTTTTATTTTTATGTTGTCTGTTTCATG
pksE(ER)	*Nhe*I	GTACGCTAGCGGGATTACGGCAGAATCTTTAGGAAGTG
	*Xho*I	GATCCTCGAGAACTTTTATTTTTATGTTGTCTGTTTCATG
pksJ(KR-T-T)	*Nde*I	GGAATTCCATATGTCAACACGGCACGAACGCTTAATGC
	*Xho*I	GATCCTCGAG**TCA**GGTCAAATACTCAGCCAGGCTGTGG
pksJ(T-T)	*Nde*I	GGAATTCCATATGCGGAACGATCAGAGAAAGGCGGATC
	*Xho*I	GATCCTCGAG**TCA**GGTCAAATACTCAGCCAGGCTGTGG

b. Protein expression

Cultures for overexpression were grown in LB media containing 50 µg/mL kanamycin overnight at 37˚C. Cultures were diluted 100x, incubated at 37˚C 2 h and cooled to 15˚C. Incubation continued until $OD_{600} \approx 0.5$, and protein expression was induced by addition

of 400 μM isopropyl-β-D-thiogalactopyranoside (IPTG) and incubation was continued overnight.

Cells were harvested by centrifugation and resuspended in 15 mL of lysis buffer (20 mM Tris, 500 mM NaCl, 10% glycerol, pH 8.0) per liter of culture, with added protease inhibitors (Compleat, Roche). Cells were lysed by passage through a French press and the lysate was clarified by ultracentrifugation. The supernatant was added to 50% Ni-NTA agarose slurry (Qiagen) pre-washed with lysis buffer (1 mL slurry per liter culture) and batch bound 90 min at 4˚C. The resin was subjected to step-gradient elution: 2x 7.5 mL lysis buffer + 5 mM imidazole followed by 2.5 mL volumes of lysis buffer + 25 mM, 50 mM, 75 mM, 100 mM, 125 mM, 150 mM, and 200 mM imidazole. Protein-containing fractions were identified by gel electrophoresis and combined.

PfaD-containing fractions were dialyzed against 2 L 20 mM Tris, 50 mM NaCl, 10% glycerol, pH 8.0, for 4 h at 4˚C. The protein solution was transferred to 2 L of fresh dialysis buffer, followed by overnight dialysis at 4˚C. All other proteins were subjected to gel filtration on a Superdex S200 column (Akta, GE Life Sciences), eluting with 20 mM Tris, 50 mM NaCl, 10% glycerol, pH 8.0. Where necessary, buffer-exchanged proteins were concentrated using an Amicon Ultrafree concentrator (Millipore). Protein concentrations were determined using the Bradford assay and bovine serum albumin as a standard.

Samples of PfaD and PksE were heated to 95˚C for 5 min, followed by centrifugation. HPLC analysis of the supernatant (Vydac 250 x 4.6 mm small-pore C18 column) using a mobile phase of 10-40% acetonitrile in 0.1% aqueous trifluoroacetic acid over 15 minutes identified the flavin cofactor in both cases as flavin mononucleotide (FMN).

c. Determination of malonyl transferase activity

Holo-PksJ(T-T), AcpK, and PksL(T-T) were generated by incubating the *apo* proteins (55 μM) with 3.5 μM Sfp and 1 mM CoA in 50 mM HEPES, 10 mM $MgCl_2$, 1 mM TCEP, pH 7.5, for 1 h at room temperature. The *holo* proteins (17.5 μM final concentration) were then combined in the same buffer (18 μL) including 100 μM 2-^{14}C-malonyl-CoA (52 Ci/mol, GE Life Sciences) in the presence or absence of 2 μM PksC or PksE and incubated 5 min before subjecting to SDS-PAGE. The gel was visualized by Coomassie staining, dried, and subjected to autoradiography.

d. Active site mapping of PksJ(KR-T-T)

For enoyl reduction assays on PksJ, the protein included a ketoreductase (KR) domain. Though the KR played no functional role in the assay, its inclusion led to improved protein expression. *Holo*-PksJ(KR-T-T) was generated by incubation of *apo*-PksJ(KR-T-T) (35.6 μM) with 12.5 μM Sfp and 0.45 mM CoA in 50 mM HEPES, 10 mM $MgCl_2$, pH 7.8, in a reaction volume of 30 μL for 2 h at 30°C. *Apo*-PksJ(KR-T-T) (in a control reaction omitting CoA) and *holo*-PksJ(KR-T-T) were digested by addition of endoproteinase ArgC (Worthington Biochemical Company) in a ratio of 1:10 w:w ArgC:total protein and incubated at 30°C for 20 min. Digestion reactions were quenched by addition of an equal reaction volume of 10% formic acid and frozen at -80°C. The

same digestion and quenching protocol was performed after all reactions described below unless noted. Mass spectra of active-site peptides are shown in Figure S4 and masses of detected peptides are provided in Table S1.

e. Determination of PksE enoyl reductase activity

2-butenoyl-*S*-PksJ(KR-T-T) was generated by incubation of *apo*-PksJ(KR-T-T) (35.6 μM) with 12.5 μM Sfp and 0.45 mM 2-butenoyl-CoA in 50 mM HEPES, 10 mM $MgCl_2$, pH 7.8, in a reaction volume of 30 μL for 1 h at 30°C. In a volume of 40 μl, PksE (1:5 mol:mol PksE(AT-ER or ER):PksJ(KR-T-T); PksJ(KR-T-T) final concentration of 15 μM) was added with 4 mM NADPH or NADH, 10 μM FMN and 50 mM HEPES, pH 7.8, (final reaction volume of 70 μl) and incubated for 1 h at 30°C.

f. Determination of PksE activity at PksJ(A-T)

2-butenoyl-*S*-PksJ(A-T) was generated in a 50 μl reaction by incubation of *apo*-PksJ(A-T) (66 μM) with 12.5 μM Sfp and 0.50 mM 2-butenoyl-CoA in 50 mM HEPES, 10 mM $MgCl_2$, pH 7.8, for 1 h at 30°C. In a volume of 50 μl, PksE (1:5 mol:mol PksE(AT-ER):PksJ(A-T); PksJ(A-T) final concentration of 33 μM) was added with 4 mM NADPH, 10 μM FMN and 50 mM HEPES, pH 7.8, and incubated for 1 h at 30°C. 30 μg of sequencing grade trypsin (Promega) was resuspended in 30 μl of trypsin resuspension buffer and incubated at 30°C for 15 min. 100 μl of 0.1 M NH_4HCO_3, pH 7.8, and the activated trypsin were added to the reaction and incubated at 30°C for 20 min. The digestion was quenched by addition of an equal reaction volume of 10% formic acid and frozen at -80°C.

g. Determination of *S. oneidensis* MR-1 PfaD enoyl reductase ability

For active site mapping experiments, *holo*-PfaA(T_4) was generated by incubating 14 μM *apo*-PfaA(T_4) with 5 μM Sfp and 0.35 mM CoA in a 80 μl reaction containing 50 mM HEPES and 10 mM $MgCl_2$, pH 7.8, for 2 h at 30°C. 2-butenoyl-*S*-PfaA(T_4) was generated by incubating 28 μM *apo*-PfaA(T_4) with 10 μM Sfp and 0.70 mM CoA in a 50 μl reaction containing 50 mM HEPES and 10 mM $MgCl_2$, pH 7.8, for 1 hr at 30°C. PfaD was added (1:5 mol:mol PfaD:PfaA(T_4); PfaA(T_4) final concentration of 14 μM) with 4 mM NADPH, 10 μM FMN and 50 mM HEPES, pH 7.8, in a volume of 30 μl and incubated for 1 h at 30°C. Mass spectra of active-site peptides are shown in Figure S5 and masses of detected peptides are provided in Table S1.

h. Reverse phase liquid chromatography (RPLC)-Fourier Transform Mass Spectrometry (FTMS) analysis

All RPLC analyses were conducted using a Jupiter C4 column (Phenomenex, 1 mm x 150 mm, 5 μ d_p, 300 Å). All FTMS analyses were conducted using a LTQ-FT hybrid ion trap-FTMS (ThermoFisher Scientific) operating at 12 Tesla. All digested samples were injected directly onto the RPLC column and eluted into the mass spectrometer using a linear gradient starting at 95% buffer A/5% buffer B and ramping to 5% buffer A/95%

buffer B over 1 h, where buffer A was water + 0.1% formic acid and buffer B was acetonitrile + 0.1% formic acid. All FTMS analysis methods included a full scan (detect *m/z* 400-2000), source induced dissociation (SID) for the phosphopantetheinyl ejection assay[1] (SID = 75V, detect *m/z* 250-600), and data dependent tandem MS (MS/MS) using collision induced dissociation (CID) with FT detection on the top three peptide species in each full scan. MS/MS data were analyzed using ProSight PC (ThermoFisher Scientific). All other data were analyzed using the Qualbrowser software suite provided with LTQ-FT instrumentation (ThermoFisher Scientific). Theoretical and experimental masses of all peptides discussed are provided in Table S1. Theoretical structures and masses of Ppant ejection ions are shown in Figure S6.

1. Dorrestein, P. C.; Bumpus, S. B.; Calderone, C. T.; Garneau-Tsodikova, S.; Aron, Z. D.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Kelleher, N. L., Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry. *Biochemistry* **2006,** 45, (42), 12756-66.