Supporting Information

for

Synthetic Strategy of Nonreducing Iterative Polyketide Synthases and the Origin of the Classical "Starter Unit Effect"

Jason M. Crawford, Anna L. Vagstad, Karen P. Whitworth
Kenneth C. Ehrlich, and Craig A. Townsend*

Strains and Culture Conditions. *C. orbiculare* NRRL 28842 (*pks1*) was obtained from Dr. Kerry O'Donnell at the Northern Center for Agricultural Utilization Research (NCAUR), Peoria, IL. Cultures showing dark brown mycelia on potato dextrose (PD) agar plates were harvested in Triton X-100 (0.1%) and grown overnight with shaking at 30 °C in PD liquid medium. Mycelia were collected by filtration and ground in liquid nitrogen to a fine powder. DNA was prepared from an aliquot of the ground mycelia using the DNeasy method (Invitrogen). The gDNA PCR templates for *A. nidulans* RLMH67 (*wA*), *C. nicotianae* ATCC18366 (*ctb1*), and the imperfect form of *G. fujikuroi*, *Fusarium fujikuroi* M6884 (+) (*pks4*), were gifts from Dr. Nancy Keller, University of Wisconsin Madison, Dr. Margo Daub, North Carolina State University, and Dr. Robert Proctor, NCAUR, respectively.

Cloning and Mutagenesis. The predicted protein sequence data (NCBI) coded by exons 1 and 2 in *wA* (Q03149), *pks1* (BAA18956), *ctb1* (AAT69682), and *pks4* (CAB92399) were codon optimized for *E. coli*. The synthons (synthetic gene fragments) with a 5' NdeI site for cloning were assembled from standard 40-mer oligonucleotides using polymerase cycling assembly (PCA) followed by PCR amplification. The remaining portion of the SAT domains within exon 3 were amplified from the gDNA templates including an engineered 5' overlap to the codon optimized fragment and a 3' NotI site. Primer pairs include WA-ex3-5 / SAT-WA-3, *wA*; PKS1-ex3-5 / SAT-PKS1-3, *pks1*; CTB1-ex3-5 / SAT-CTB1-3, *ctb1*; and PKS4-ex3-5 / SAT-PKS4-3, *pks4*. The codon-optimized fragment was fused to the gDNA fragment by overlap extension PCR and amplified using the end primers for *wA*, *pks1*, and *ctb1*. The full-length PCR products were digested with NdeI and NotI and inserted into the corresponding sites in pET24a (Novagen) generating pEWA-SAT, pEPKS1-SAT, and pECTB1-SAT. The PKS4 overlapping region contained a unique EcoRI site for ligation of the two fragments. The exon 3 fragment was

inserted into the EcoRI/NotI sites in pET24a yielding pEPKS4-exon3. The synthon was subsequently digested with NdeI and EcoRI and inserted into the corresponding sites in pEPKS4-exon3 giving rise to the expression construct pEPKS4-SAT. All translated SAT domain sequences were identical to their corresponding protein accession numbers. As a control, the active site Cys119 codon within the PKS1 SAT domain was mutated to Ala using overlap extension PCR.

The expected GX**C**XG motif within PKS4 (CAB92399) was encoded within the formerly proposed intron, and expression of this SAT resulted in no soluble protein. Therefore, a gDNA fragment encompassing this region was amplified with oligonucleotides PKS4-ex1-5 and PKS4-ex2-3 for sequence verification. Products from two separate PCR reactions were cloned into pCR-ZeroBlunt (Invitrogen) and sequenced. A deletion mutation was discovered in the available sequence (CAB92399). Hidden Markov Model (HMM)-based gene structure prediction modeled on the similar fungus *Fusarium graminearum* yielded an alternative splicing pattern that translated through the GX**C**XG motif (FGENESH, www.softberry.com). The incorrect synthon sequence within the expression construct, pEPKS4-SAT, was repaired using both gene synthesis and overlap extension PCR to generate the revised expression construct, pEPKS4-SATn.

The SAT-KS-MAT tri-domain from PKS1 was constructed similarly to the SAT domain. The exon 3 fragment was amplified from the gDNA using PKS1-ex3-5 and NKA-PKS1-3 (NotI) and subsequently fused to the codon-optimized fragment. Insertion into pET24a generated pEPKS1-NKA. Like the SAT monodomain control, the active site Cys119 codon within the SAT domain was mutated to Ala.

primer	sequence 5′-3′
WA-ex3-5	GATAAATTTACCCGGCGTAACGTTATCCCTCCTTCGTCTCGGCC
SAT-WA-3	GTAAGCGGCCGC-TTCAGCCCTGCCTGTCTGGTTG
PKS1-ex3-5	GCGAACTTTCATGAATCGAATTACATCTCCGTCGCCAGCCAAG
SAT-PKS1-3	GTAAGCGGCCGC-GCCACTGGAACCGTGGTTTCC
CTB1-ex3-5	GAACAGGCGCTTGCATCCGTGTCCGTTCCAT
SAT-CTB1-3	GTAA-GCGGCCGCTGCCATCGGCTGCTGTTCC
PKS4-ex3-5	CTGAAAGAATTCTCAAAGGAAAAA AATCTTACATACAGTTCGCGTCCCTAC
SAT-PKS4-3	GTAAGCGGCCGC-ATCGCCAGGAGCAAGATTGGG
PKS4-ex1-5	CATATGGCGTCCTCAGCAGATGTGT
PKS4-ex2-3	CTGATGTAGGGACGCGAACTGTATGT
NKA-PKS1-3	GTAAGCGGCCGCTTTGGGTCTGGCAGGCTCGAT

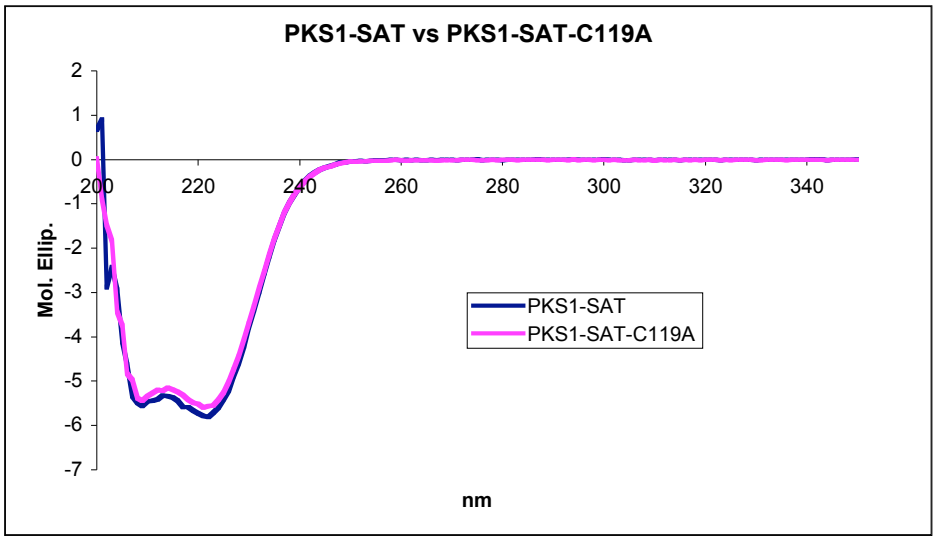
Expression and Purification of SAT monodomains. BL21(DE3) *E. coli* cells (Novagen) harboring pEWA-SAT, pEPKS1-SAT, pECTB1-SAT, or pEPKS4-SATn were grown to OD_{600} = 0.6 at 30 °C in 2×YT medium (200 mL) supplemented with kanamycin (25 μg/mL). The cultures were induced with IPTG (1 mM) and the temperature was lowered to 20 °C for a 16 hour expression. The cells were harvested at 4080 × *g* for 20 min and resuspended in resuspension buffer (5 mL, 50 mM potassium phosphate pH 7.6, 10% glycerol, 300 mM NaCl, 10 mM imidazole, and 1 mg/mL lysozyme). Cells were incubated on ice for 30 min, disrupted by sonication, and cleared by centrifugation for 30 min at 27000 × *g*. Protein was bound to nickel-nitrilotriacetic acid resin (Qiagen) and eluted with an imidazole step-gradient according to the manufacturer's instructions. The purified protein was dialyzed against 100 mM potassium phosphate pH 7.0 for assay, and enzyme concentrations were determined with the Bradford assay in triplicate using bovine albumin as a standard.

Expression and Purification of PKS1 SAT-KS-MAT tridomain. The SAT-KS-MAT tridomain was expressed in Rosetta2(DE3). The cells were grown to OD_{600} = 1.0 at 30 °C in 2×YT (1 L) supplemented with both kanamycin (25 μg/mL) and chloramphenicol (25 μg/mL).

The cells were induced with IPTG (1 mM) and grown for 16 additional hours at 20 °C. Harvested cells were resuspended in resuspension buffer (20 mL) including pepstatin A (1 mg/mL) and leupeptin (0.5 mg/mL) and the protein was purified similarly.

Kinetic Assay of SAT Transfer. Enzyme (0.5 μM) was reacted with acyl-CoAs (1 mM) and pantetheine (2.5 mM) at 28.0 °C in 100 mM potassium phosphate, pH 7.0, so that transfer could be followed chromatographically. Aliquots (100 μL) at 2,4,6, and 8 min were quenched in 8M urea (100 μL), frozen in liquid nitrogen, and stored at -80 °C. Quenched reactions (200 μL) were individually thawed, diluted with mobile phase (300 μL), and filtered. Part of the filtered sample (200 μL) was injected onto a Luna Phenyl-Hexyl 100-Å, 5-μm, 4.6- × 250-mm column (Phenomenex), and CoA and acyl-CoAs (258 nm absorbance) and pantetheine and acyl-pantetheines (232 nm) were eluted with an acetonitrile gradient.

CD Analysis of PKS1 SAT and PKS1 SAT C119A. The monodomains were purified as above and their CD spectra (JASCO J-710) were obtained as shown below.



Supporting Figure 1. CD analysis of PKS1 SAT domains.

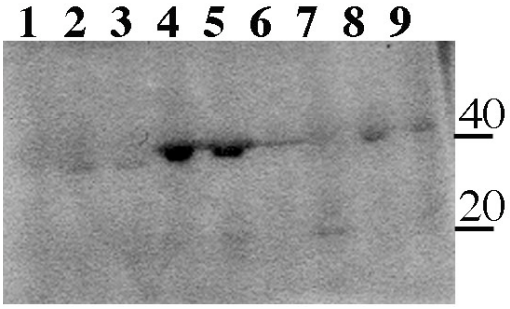
Method t(min)	A	B
acetyl/malonyl		
5	0	100
15	5	95
19	55	45
23	70	30
27		100
propionyl/ butanoyl		
5		100
7	19	81
9	19	81
14	29.5	70.5
15	35.2	64.8
20	75	25
25		100
hexanoyl		
5		100
7	23.8	76.2
9	23.8	76.2
15	40.8	59.2
17	76	24
22	65	35
28		100

HPLC methods. Samples saved for HPLC analysis were examined using a Luna Phenyl-Hexyl 100Å, 5µm, 4.6 × 250 mm column (Phenomenex, Torrance, CA) as follows. In all cases CoA was detected under initially isocratic conditions, (Solvent A: 5% acetonitrile and 25 mM KH_2PO_4, pH = 5.4 w/NaOH; 50 °C, 1.5 mL/min for 5 min) and then gradient elution with Solvent B (acetonitrile, 50 °C) was varied depending on the substrates and products to be quantified.

HPLC method	**Compound**	**Retention Time**
malonyl	pantetheine	14.5
	malonly-CoA	2.9 min
	malonyl-pant	5.7 min
acetyl	pantetheine	14.4 min
	acetyl-CoA	5.9 min
	acetyl-pant	19.1 min
propionyl	pantetheine	9.5 min
	propionyl-CoA	9.0 min
	propionyl-pant	11.7 min
butanoyl	pantetheine	9.5 min
	butanoyl-CoA	9.4 min
	butanoyl-pant	14 min
hexanoyl	pantetheine	9.3 min
	hex-CoA	9.8 min
	hex-pant	15.9 min
All methods	CoA	3.2 min

Radiochemical Assay of SAT Transfer

The SAT monodomains (10 μM) from WA, PKS1, CTB1, and PKS4 in 100 mM potassium phosphate buffer (pH 7.0) in the presence of 200 μM [1-^{14}C]acetyl-CoA (54 mCi/mmol, Sigma-Aldrich, St. Louis, MO), with or without PksA ACP (10 μM), were reacted at 25 °C for 5 min. The reactions were quenched with 5 × SDS gel loading buffer and separated over a 12% SDS-PAGE gel. The gels were dried between cellophane sheets and exposed to BioMax XAR film (Eastman, Kodak).



Supporting Figure 2. Radioactive transacylase experiments. The SAT and PksA *holo*-ACP combinations include: **1.** ACP component **2.** WA SAT **3.** WA SAT + ACP **4.** PKS1 SAT **5.** PKS1 SAT + ACP **6.** CTB1 SAT **7.** CTB1 SAT + ACP **8.** PKS4 SAT **9.** PKS4 SAT + ACP. All lanes including a SAT domain (~40 kDa) showed covalent loading of the acetyl group. Signals for the non-cognate ACP (~20 kDa) loading were faintly apparent, possibly due to hydrolysis during the analysis procedure, leading to the reduced intensity. Nevertheless, the PksA ACP was faintly labeled in all lanes including a SAT domain.

Synthon synthesis

PCA: Standard desalted oligonucleotides were dissolved as 200 μM stocks in EB buffer (Qiagen). An equal volume of each oligo was added together for a master mix solution. PCA was carried out in 50 μL reactions: oligos (2 μM each), cloned Pfu DNA polymerase buffer (Stratagene), dNTPs (0.2 mM each), and cloned Pfu DNA polymerase (0.5 μL, Stratagene). PCA cycle: denaturation, 94 °C (5 min), 1 cycle; extension, 94 °C (30 s), 55 °C (30 s), 72 °C (2 min), 25 cycles; clean up, 72 °C (10 min).

PCR: The PCA reaction (4 μL) was directly used for PCR amplification in reactions (50 μL): End primers (0.4 μM each, +1 and -1 synthon primers), cloned Pfu DNA polymerase buffer, dNTPs (0.2 mM each), DMSO (5 %), and cloned Pfu DNA polymerase (0.5 μL). PCR cycle: denaturation, 98 °C (5 min), 1 cycle; extension 1, 98 °C (30 s), 45 °C (30 s), 72 °C (3 min), 5 cycles; extension 2, 98 °C (30 s), 55 °C (30 s), 72 °C (3 min), 40 cycles; clean up, 72 °C (10 min).

note: assembly and amplification of synthetic gene fragments could be carried out in halves for increased yield of PCR products.

WA SAT (Q03149) synthon construction

Amino acid sequence reverse translated into a codon optimized sequence for E. coli

gtaacatATGGAGGATCCGTACCGTGTTTACTTGTTCGGTGATCAAACTGGCGACTTTGA
AGTTGGCTTACGTCGTCTGCTTCAAGCTAAGAATCATAGCCTCTTGTCGTCCTTCCTG
CAGCGTTCCTACCACGCCGTACGCCAGGAAATTTCACACTTACCGCCAAGCGAACGT
TCTACGTTTCCACGCTTCACCTCGATCGGCGATTTACTGGCGCGCCATTGCGAGTCTC
CCGGTAACCCGGCGATTGAATCAGTCTTAACATGTATTTATCAACTGGGATGTTTTA
TCAATTATTATGGGGACCTCGGACACACCTTTCCTAGCCATAGCCAGAGCCAGCTCG
TAGGTCTGTGTACGGGTCTGCTTAGTTGCGCAGCCGTTTCTTGCGCTAGTAACATTGG
CGAATTGCTTAAGCCGGCGGTCGAAGTGGTGGTGGTGGCCCTGCGGCTGGGCCTGTG
CGTCTATCGCGTGCGCAAATTGTTCGGGCAGGACCAGGCAGCGCCTCTGAGCTGGTC
AGCACTGGTTTCTGGCCTGTCCGAGTCGGAAGGTACCAGTCTGATCGATAAATTTAC
CCGGCGTAAC

WA-SAT upper strand oligos #1-14

#	sequence
1	gtaacatATGGAGGATCCGTACCGTGTTTACTTGTTCGGTGATCAAA
2	CTGGCGACTTTGAAGTTGGCTTACGTCGTCTGCTTCAAGC
3	TAAGAATCATAGCCTCTTGTCGTCCTTCCTGCAGCGTTCC
4	TACCACGCCGTACGCCAGGAAATTTCACACTTACCGCCAA
5	GCGAACGTTCTACGTTTCCACGCTTCACCTCGATCGGCGA
6	TTTACTGGCGCGCCATTGCGAGTCTCCCGGTAACCCGGCG
7	ATTGAATCAGTCTTAACATGTATTTATCAACTGGGATGTT
8	TTATCAATTATTATGGGGACCTCGGACACACCTTTCCTAG
9	CCATAGCCAGAGCCAGCTCGTAGGTCTGTGTACGGGTCTG
10	CTTAGTTGCGCAGCCGTTTCTTGCGCTAGTAACATTGGCG
11	AATTGCTTAAGCCGGCGGTCGAAGTGGTGGTGGTGGCCCT
12	GCGGCTGGGCCTGTGCGTCTATCGCGTGCGCAAATTGTTC
13	GGGCAGGACCAGGCAGCGCCTCTGAGCTGGTCAGCACTGG

14	TTTCTGGCCTGTCCGAGTCGGAAGGTACCAGTCTGATCGA

WA-SAT lower strand oligos #1-14

#	sequence
-1	GTTACGCCGGGTAAATTTATCGATCAGACTGGTACCTTC
-2	CGACTCGGACAGGCCAGAAACCAGTGCTGACCAGCTCAGA
-3	GGCGCTGCCTGGTCCTGCCCGAACAATTTGCGCACGCGAT
-4	AGACGCACAGGCCCAGCCGCAGGGCCACCACCACCACTTC
-5	GACCGCCGGCTTAAGCAATTCGCCAATGTTACTAGCGCAA
-6	GAAACGGCTGCGCAACTAAGCAGACCCGTACACAGACCTA
-7	CGAGCTGGCTCTGGCTATGGCTAGGAAAGGTGTGTCCGAG
-8	GTCCCCATAATAATTGATAAAACATCCCAGTTGATAAATA
-9	CATGTTAAGACTGATTCAATCGCCGGGTTACCGGGAGACT
-10	CGCAATGGCGCGCCAGTAAATCGCCGATCGAGGTGAAGCG
-11	TGGAAACGTAGAACGTTCGCTTGGCGGTAAGTGTGAAATT
-12	TCCTGGCGTACGGCGTGGTAGGAACGCTGCAGGAAGGACG
-13	ACAAGAGGCTATGATTCTTAGCTTGAAGCAGACGACGTAA
-14	GCCAACTTCAAAGTCGCCAGTTTGATCACCGAACAAGTAA

2. PKS1 SAT (BAA18956) synthon construction

Amino acid sequence reverse translated into a codon optimized sequence for E. coli

gtaacatATGGCCGATACAATGAGCTATTTGCTGTTTGGAGACCAGTCGCTGGATACGCA
TGGTTTCTTAGCTGAATTTTGTCGTAACGGGAATCCGTCCATTTTGGCTAAGACTTTT
CTGGAGCAGGCCGGTCAGGCGCTGCGTGAAGAGATCGACGGTCTGGGTAAACTGGA
ACGCTCAAAACTGCCAACGTTCCAAACACTGCGGCAGCTCAACGAGCGCTATCACG
CACAGTCAATCAAACATCCGGGCATTGATAGTGCGCTGCTCTGCACCACTCAATTAG
CACACTACATTGATCGCACCGAAAAGGAGCCCCAGGACGCGTGTCTGCACGATCAT
ACGTTTTTTATGGGCCTGTGCACCGGATTGTTCGCCGCGGCGGCCATCGCAAGCACC
CCATCTGTGTCTACCTTAATTCCCCTCGCGGTGCAAGTAGTTCTTATGGCATTCCGCA
CTGGCTCCCACGTCGGCAGCTTAGCCGAACGTCTTAGCCCGCCGGTTGGCCAGAGCG
AACCTTGGACGCATATCCTTCCTGGGTTGAAAGAAAGTGATGCGAAGGAAGCTCTG
GCGAACTTTCATGAATCGAAT

PKS1-SAT upper strand oligos #1-15

#	sequence
1	gtaacatATGGCCGATACAATGAGCTATTT
2	GCTGTTTGGAGACCAGTCGCTGGATACGCATGGTTTCTTA
3	GCTGAATTTTGTCGTAACGGGAATCCGTCCATTTTGGCTA
4	AGACTTTTCTGGAGCAGGCCGGTCAGGCGCTGCGTGAAGA
5	GATCGACGGTCTGGGTAAACTGGAACGCTCAAAACTGCCA
6	ACGTTCCAAACACTGCGGCAGCTCAACGAGCGCTATCACG
7	CACAGTCAATCAAACATCCGGGCATTGATAGTGCGCTGCT
8	CTGCACCACTCAATTAGCACACTACATTGATCGCACCGAA

9	AAGGAGCCCCAGGACGCGTGTCTGCACGATCATACGTTTT
10	TTATGGGCCTGTGCACCGGATTGTTCGCCGCGGCGGCCAT
11	CGCAAGCACCCCATCTGTGTCTACCTTAATTCCCCTCGCG
12	GTGCAAGTAGTTCTTATGGCATTCCGCACTGGCTCCCACG
13	TCGGCAGCTTAGCCGAACGTCTTAGCCCGCCGGTTGGCCA
14	GAGCGAACCTTGGACGCATATCCTTCCTGGGTTGAAAGAA
15	AGTGATGCGAAGGAAGCTCTGGCGAACTTTCATGAATCGA

PKS1-SAT lower strand oligos #1-15

#	sequence
-1	ATTCGATTCATGAAAGTTCGCC
-2	AGAGCTTCCTTCGCATCACTTTCTTTCAACCCAGGAAGGA
-3	TATGCGTCCAAGGTTCGCTCTGGCCAACCGGCGGGCTAAG
-4	ACGTTCGGCTAAGCTGCCGACGTGGGAGCCAGTGCGGAAT
-5	GCCATAAGAACTACTTGCACCGCGAGGGGAATTAAGGTAG
-6	ACACAGATGGGGTGCTTGCGATGGCCGCCGCGGCGAACAA
-7	TCCGGTGCACAGGCCCATAAAAAACGTATGATCGTGCAGA
-8	CACGCGTCCTGGGGCTCCTTTTCGGTGCGATCAATGTAGT
-9	GTGCTAATTGAGTGGTGCAGAGCAGCGCACTATCAATGCC
-10	CGGATGTTTGATTGACTGTGCGTGATAGCGCTCGTTGAGC
-11	TGCCGCAGTGTTTGGAACGTTGGCAGTTTTGAGCGTTCCA
-12	GTTTACCCAGACCGTCGATCTCTTCACGCAGCGCCTGACC
-13	GGCCTGCTCCAGAAAAGTCTTAGCCAAAATGGACGGATTC
-14	CCGTTACGACAAAATTCAGCTAAGAAACCATGCGTATCCA
-15	GCGACTGGTCTCCAAACAGCAAATAGCTCATTGTATCGGC

PKS1 SAT-C119A mutagenesis

The active site Cys within the SAT (synthon region) was mutated (C119A) using overlap extension PCR with the mutant primers, PKS1-SAT-C119A-5 / PKS1-SAT-C119A-3, and the end primers, SAT-PKS1-3 / PKS1 upper strand oligo # 1. The PCR product was digested with NdeI and NotI and inserted into pET24a to generate the SAT monodomain mutant construct, pEPKS1-SAT-C119A. The PCR product was digested with NdeI and SacII and inserted into the corresponding sites in pEPKS1-NKA to create the SAT-KS-MAT tridomain mutant construct, pEPKS1-NKA-C119A.

PKS1-SAT mutant oligos

primer	sequence
PKS1-SAT-C119A-5	ACGTTTTTTATGGGCCTG**GC**CACCGGATTGTTCGCC

PKS1-SAT-C119A-3	GGCGAACAATCCGGTGGCCAGGCCCATAAAAAACGT
SAT-PKS1-3	GTAAGCGGCCGC-GCCACTGGAACCGTGGTTTCC

3. CTB1 SAT (AAT69682) synthon construction

Amino acid sequence reverse translated into a codon optimized sequence for E. coli

gtaacatATGGAAGATGGCGCCCAAATGCGCGTGGTAGCCTTCGGGGACCAAACCTATG
ACTGCTCAGAAGCGGTGAGCCAGCTGTTGCGCGTACGTGACGATGCGATCGTGGTC
GATTTTTTAGAACGTGCACCGGCAGTTCTTAAAGCGGAACTGGCGCGCTTGTCCAGC
GAGCAGCAGGAAGAGACTCCGCGCTTTGCAACCCTGGCCGAGTTAGTCCCTCGTTAT
CGGGCCGGCACGCTGAATCCGGCCGTGAGTCAGGCTCTCACTTGTATTGCTCAACTG
GGTCTGTTCATCCGCCAGCATTCGAGCGGCCAAGAAGCGTATCCCACCGCCCACGAT
TCTTGCATTACCGGCGTTTGTACCGGCGCATTGACGGCGGTTGCGGTCGGCTCCGCA
AGCTCAGTCACAGCTCTGGTTCCATTAGCGCTGCATACGGTGGCAGTTGCGGTACGT
CTGGGAGCTCGTGCTTGGGAGATTGGTAGCTGCCTCGCTGATGCCCGTCGGGGTGCG
AACGGGCGTTACGCATCGTGGACAAGTGCCGTGGGTGGAATTTCTCCGCAGGATCTT
CAGGACCGCATCTCTGCGTACACCGCCGAACAGGCGcttgcatccgtgtcc

CTB1-SAT upper strand oligos #1-15

#	sequence
1	gtaacatATGGAAGATGGCGCCCAAATGCGCGTGGTAGCCTTCGG
2	GGACCAAACCTATGACTGCTCAGAAGCGGTGAGCCAGCTG
3	TTGCGCGTACGTGACGATGCGATCGTGGTCGATTTTTTAG
4	AACGTGCACCGGCAGTTCTTAAAGCGGAACTGGCGCGCTT
5	GTCCAGCGAGCAGCAGGAAGAGACTCCGCGCTTTGCAACC
6	CTGGCCGAGTTAGTCCCTCGTTATCGGGCCGGCACGCTGA
7	ATCCGGCCGTGAGTCAGGCTCTCACTTGTATTGCTCAACT
8	GGGTCTGTTCATCCGCCAGCATTCGAGCGGCCAAGAAGCG
9	TATCCCACCGCCCACGATTCTTGCATTACCGGCGTTTGTA
10	CCGGCGCATTGACGGCGGTTGCGGTCGGCTCCGCAAGCTC
11	AGTCACAGCTCTGGTTCCATTAGCGCTGCATACGGTGGCA
12	GTTGCGGTACGTCTGGGAGCTCGTGCTTGGGAGATTGGTA
13	GCTGCCTCGCTGATGCCCGTCGGGGTGCGAACGGGCGTTA
14	CGCATCGTGGACAAGTGCCGTGGGTGGAATTTCTCCGCAG
15	GATCTTCAGGACCGCATCTCTGCGTACACCGCCGAACAGG

CTB1-SAT lower strand oligos #1-15

#	sequence
1	ggacacggatgcaagCGCCTGTTCGGCGGTGTACGCA
2	GAGATGCGGTCCTGAAGATCCTGCGGAGAAATTCCACCCA
3	CGGCACTTGTCCACGATGCGTAACGCCCGTTCGCACCCCG
4	ACGGGCATCAGCGAGGCAGCTACCAATCTCCCAAGCACGA
5	GCTCCCAGACGTACCGCAACTGCCACCGTATGCAGCGCTA
6	ATGGAACCAGAGCTGTGACTGAGCTTGCGGAGCCGACCGC

7	AACCGCCGTCAATGCGCCGGTACAAACGCCGGTAATGCAA
8	GAATCGTGGGCGGTGGGATACGCTTCTTGGCCGCTCGAAT
9	GCTGGCGGATGAACAGACCCAGTTGAGCAATACAAGTGAG
10	AGCCTGACTCACGGCCGGATTCAGCGTGCCGGCCCGATAA
11	CGAGGGACTAACTCGGCCAGGGTTGCAAAGCGCGGAGTCT
12	CTTCCTGCTGCTCGCTGGACAAGCGCGCCAGTTCCGCTTT
13	AAGAACTGCCGGTGCACGTTCTAAAAAATCGACCACGATC
14	GCATCGTCACGTACGCGCAACAGCTGGCTCACCGCTTCTG
15	AGCAGTCATAGGTTTGGTCCCCGAAGGCTACCACGCGCAT

4. PKS4 SAT (CAB92399) synthon construction

Amino acid sequence reverse translated into a codon optimized sequence for E. coli

gtaacatATGGCGAGCAGCGCCGACGTGTATGTGTTCGGCGATCAGAGCACCCCAGTAC
TGGACAAATTACAGGCATTGGTTCGCGTGAAGGATAACGCTTTACTGACATCGTTTT
TGGGCGAAGCCTTTTTGGCCGTTCGCCGTGAAATCGTGTCCCTCTCGTCTCTGGAGC
GTAAGAGTATTCCGGAAGCGGAGTCTTTAAGCCTCCTGCTTGAAGGCGTACGGCGCT
CCGAGCCCCACGCAGCTCTGGATAGCGCGTTTGTCTGTATTTACGAAATTGGTTATT
ACATCGACCGTTGTTGCCATTGCGCAAAAGATGTTTTTGAAATTAGTCGCCTGGGAG
TTGAGGCTGCCACGGTCGCGTTCCGTCTGGGTATGCATGTCCGTCGCCGCGCAGAAA
ATCTTGGGTATTCGACTCCTTCATCCTGGAGTATGATCCTGTCATCTAACCAAGAGG
AACTGGTGAGCGAAGCGCTGAAAGAATTCTCAAAGGAAAAA

PKS4-SAT upper strand oligos #1-12

#	sequence
1	gtaacatATGGCGAGCAGCGCCGACGTGTATGTGTTCGGCGAT
2	CAGAGCACCCCAGTACTGGACAAATTACAGGCATTGGTTC
3	GCGTGAAGGATAACGCTTTACTGACATCGTTTTTGGGCGA
4	AGCCTTTTTGGCCGTTCGCCGTGAAATCGTGTCCCTCTCG
5	TCTCTGGAGCGTAAGAGTATTCCGGAAGCGGAGTCTTTAA
6	GCCTCCTGCTTGAAGGCGTACGGCGCTCCGAGCCCCACGC
7	AGCTCTGGATAGCGCGTTTGTCTGTATTTACGAAATTGGT
8	TATTACATCGACCGTTGTTGCCATTGCGCAAAAGATGTTT
9	TTGAAATTAGTCGCCTGGGAGTTGAGGCTGCCACGGTCGC
10	GTTCCGTCTGGGTATGCATGTCCGTCGCCGCGCAGAAAAT
11	CTTGGGTATTCGACTCCTTCATCCTGGAGTATGATCCTGT
12	CATCTAACCAAGAGGAACTGGTGAGCGAAGCGCTGAAAGA

PKS4-SAT lower strand oligos #1-12

#	sequence
1	TTTTTCCTTTGAGAATTCTTTCAGCGCTTCGCTCAC
2	CAGTTCCTCTTGGTTAGATGACAGGATCATACTCCAGGAT
3	GAAGGAGTCGAATACCCAAGATTTTCTGCGCGGCGACGGA
4	CATGCATACCCAGACGGAACGCGACCGTGGCAGCCTCAAC

5	TCCCAGGCGACTAATTTCAAAAACATCTTTTGCGCAATGG
6	CAACAACGGTCGATGTAATAACCAATTTCGTAAATACAGA
7	CAAACGCGCTATCCAGAGCTGCGTGGGGCTCGGAGCGCCG
8	TACGCCTTCAAGCAGGAGGCTTAAAGACTCCGCTTCCGGA
9	ATACTCTTACGCTCCAGAGACGAGAGGGACACGATTTCAC
10	GGCGAACGGCCAAAAAGGCTTCGCCCAAAAACGATGTCAG
11	TAAAGCGTTATCCTTCACGCGAACCAATGCCTGTAATTTG
12	TCCAGTACTGGGGTGCTCTGATCGCCGAACACATACACGT

PKS4-SAT revised gDNA sequence

ATGGCGTCCTCAGCAGATGTGTATGTCTTTGGCGATCAGAGCACACCCGTTCTTGAT
AAATTGCAAGCCCTTGTGCGCGTGAAAGATAATGCGCTTTTGACTTC**T**TTTCTTGGC
GAGGCGTTTTTGGCTGTGCGAAGAGAGATTGTTTCGCTCTCGTCTTTAGAGAGAAAA
TCGATTCCCGAGGCTGAGAGCTTGAGTTTGCTTTTGGAGGGCGTTCGGAGGAGTGAG
CCTCATGCGGCGCTGGATAGTGCTTTTGTGTGCATTTATGAGATTGGATACTACATTG
A*gtaagtctgatggaaacgagcaagttct***A***ggtttgctgatgagattcaag*CTATTTGGCTCGTTCGGATAAGCA
ACACCCGCCGGCAGCTCCATCTTTGCTACTCGGTATCTGCACTGGTAGCATTGCAGC
CGCTGCTGTCA**G**CTGCGCGAAGGATGTTTTCGAGATCTCGAGACTGGGCGTTGAAG
CTGCGACAGTTGCCTTCCGACTTGGAATGCACGTTCGCAGAAGAGCCGAGAACCTTG
GGTACTCGACACCTTCGAGCTGGTCCATGATTCTGTCTTCGAATCAGGAGGAGCTTG
TTTCAGAAGCATTAAAAGAATTCTCCAAGGAAAAA*gtaagttcgagagctttggtattcccaatttgagta*
*atgctcattcttgtttcag*AATCTTACATACAGTTCGCGTCCCTACATCAGCGCCACTGGTCCTGG
TTTCACAACAATCAGCGGTCCCCCATCTATCCTCGAGTCAGTGAAGTCTTGCGATAC
TTTCTCTGGAAAGAGGTTATATCCAGCTCCCATCTATGGTCCCTACCATAACTCTTCA
TCATACTCCGAATCCAGCCTGGAACATGGCCTTGCATCAATCCTTGAAGACGTTGGG
TTCCTAGAGAACGAGATGCTCATTCCCATCATCTCCTGCGCCTCCGGATCTCGCTTA
GACCAGTTGTCTTTCGGGAACCTCCTCAAGAACGTTTTGAGCAGCGCTCTGTCTCAG
CAAATTCGTATGGATCTCGTCACCGACGCTCTAGTCGAAACCGTTTCTGGCACTGAA
GCTACCTTGATACCGGTCAATGCCCAAACAACTGTTTGCAGTCTCGCTGATTGGTTG
GCTAAGAGAGGAGCTACAACTCGCATTGGGCCAACCTTGGAGAGCCTGACCAAAGA
CCGAGCAGAACCCAATCTTGCTCCTGGCGAT . . .

PKS4 revised gDNA sequence. Capitol letters represent the revised exon sequence as predicted by FGENESH (www.softberry.com). Bold letters represent differences found in exons 1 and 2 compared to the CAB92399 sequence, which led to this alternative splicing pattern: first mutation C→**T**, silent Ser codon; C→**A**, noncoding mutation; and **G** additional base pair found in the revised sequence.

PKS4-SAT revised amino acid sequence

CIYEIGYYID-"YLARSDKQHPPAAPSLLL(**GI**C**G**)SIAAAAVS"-CAKDVFEISR

Amino acid sequence reverse translated into a codon optimized sequence for E. coli

"TACCTGGCTCGTTCCGACAAACAGCATCCACCGGCAGCGCCTTCATTGCTGTTAGG
TATTTGTACTGGCTCGATCGCCGCCGCGGCAGTCAGC"

PKS4-SAT revised upper and lower strand oligos #1-2

#	sequence
1	GAAATTGGTTATTACATCGAC-TACCTGGCTCGTTCCGACAAACAGCATCCACCGGCAG
2	CGCCTTCATTGCTGTTAGGTATTTGTACTGGCTCGATCGC
-1	ATTTCAAAAACATCTTTTGCGCA- GCTGACTGCCGCGGCGGCGATCGAGCCAGTACAAAT
-2	ACCTAACAGCAATGAAGGCGCTGCCGGTGGATGCTGTTTG

The revised PKS4-SAT sequence was constructed using pEPKS4-SAT as the template in a 3-part overlap extension PCR. The 5'-part was amplified using the first forward synthon upper strand oligo #1 and PKS4-3-n. The central portion was constructed from the revised synthetic fragment. The 3'-part was amplified using PKS4-5-n and SAT-PKS4-3 primers. The three fragments were fused and amplified using the end primers prior to insertion into pET24a to generate the revised construct, pEPKS4-SATn.

PKS4-SAT revised oligos for overlap extension PCR

primer	sequence
PKS4-3-n	GTCGATGTAATAACCAATTTCG
PKS4-5-n	TGCGCAAAAGATGTTTTT