Heterologous Production of Epothilone C and D in Escherichia coli †

ARTICLE *in* BIOCHEMISTRY · FEBRUARY 2006

Impact Factor: 3.02 · DOI: 10.1021/bi052075r · Source: PubMed

Heterologous Production of Epothilone C and D in *Escherichia coli*[†]

Sarah C. Mutka, John R. Carney, Yaoquan Liu, and Jonathan Kennedy*

Kosan Biosciences, Inc., 3832 Bay Center Place, Hayward, California 94545

ABSTRACT: The epothilones are a family of polyketide natural products that show a high potential as anticancer drugs. They are synthesized by the action of a hybrid nonribosomal peptide synthetase/polyketide synthase in the myxobacterium *Sorangium cellulosum*. In this work, the genes encoding the entire cluster, *epoA*, *epoB*, *epoC*, *epoD*, *epoE*, and *epoF*, were redesigned and synthesized to allow for expression in *Escherichia coli*. The expression of the largest of the proteins, EpoD, also required the protein be separated into two polypeptides with compatible module linkers. Using a combination of lowered temperature, chaperone coexpression, and alternative promoters, we succeeded in producing a soluble protein from all genes in the epothilone cluster. The entire synthetic epothilone cluster was then expressed in a strain of *E. coli* modified to enable polyketide biosynthesis, resulting in the production of epothilones C and D. Furthermore, feeding a thioester of the normal substrate for EpoD to cells expressing the *epoD*, *epoE*, and *epoF* genes also led to the production of epothilones C and D. The design of the synthetic epothilone genes together with *E. coli* expression provides the ideal platform for both the biochemical investigation of the epothilone PKS and the generation of novel biosynthetic epothilone analogues.

The epothilones are potential anticancer agents currently undergoing evaluation in human clinical trials. These compounds are potent microtubule-stabilizing agents with a mechanism similar to paclitaxel and are effective against paclitaxel-resistant tumor cell lines (*1*, *2*). Epothilones A and B are naturally produced by the myxobacterium *Sorangium cellulosum* by the action of a hybrid nonribosomal peptide synthetase (NRPS)[1]/polyketide synthase (PKS) together with an epoxidase (*3*, *4*). Inactivation or deletion of the epoxidase gene results in the production of epothilones C and D as the direct products of the NRPS/PKS (Figure 1). The gene cluster encoding the NRPS/PKS enzyme complex that catalyzes epothilone C and D biosynthesis is comprised of six open-reading frames (ORFs), *epoA*, *epoB*, *epoC*, *epoD*, *epoE*, and *epoF*, spanning 55 kb and encoding a single NRPS module, eight PKS modules, and a C-terminal thioesterase (TE) (Figure 1) (*4*, *5*). We are interested in modifying the epothilone NRPS/PKS gene cluster to produce novel analogues, but *S. cellulosum* is slow-growing and not readily amenable to genetic manipulations. Heterologous expression of the entire epothilone gene cluster has been achieved in both *Streptomyces coelicolor* (*6*) and *Myxococcus xanthus* (*7*−*9*); however, the former provided a low yield of epothilones, and the latter is encumbered by its slow growth rate relative to *Escherichia coli*.

Recently, *E. coli* has been extensively engineered to serve as a host for polyketide production (*10*, *11*). The modified strain *E. coli* K207-3 is a derivative of BL21 (DE3) that has the propionate utilization operon (*prpRBCD*) and the methylmalonyl-CoA decarboxylase gene (*ygfG*) deleted and carries copies of *sfp* (phosphopantetheine transferase), *prpE* (propionyl-CoA ligase), and *accA1*/*pccB* (propionyl-CoA carboxylase) under the control of T7 promoters. Feeding propionate to this strain results in the accumulation of methylmalonyl-CoA, a polyketide synthase substrate not normally produced in *E. coli*. When optimized for the production of the erythromycin precursor, 6-deoxyerythronolide B, a remarkable yield of >1 g/L was achieved (*12*). Nevertheless, the use of *E. coli* as a generic host for polyketide production presents formidable challenges: (i) codons rarely used in *E. coli* are more common in PKS genes because of their higher G+C content, leading to difficulties in translation; and (ii) the ORFs of multimodular polyketide genes encode proteins far larger than those normally found in *E. coli*. Expression of active epothilone modules in *E. coli* has proven to be difficult, and previous attempts to express the *epoA* gene have been unsuccessful. To date, only *in vitro* reconstitution of the early steps of epothilone biosynthesis has been demonstrated (*13*, *14*). However, Boddy and co-workers have recently succeeded in expressing the last four modules of the epothilone gene cluster and producing epothilone C by exogenous feeding of a complex synthetic substrate comprising the *N*-acetyl-cysteamine thioester (SNAC) of the C9−C21 fragment of epothilone C [(*S*,4*Z*,8*E*)-7-hydroxy-8-methyl-9-(2-methylthiazol-4-yl)nona-4,8-dienoic acid] (*15*).

[†] This work was supported in part by the National Institute of Standards and Technology Advanced Technology Program Grant Award Number 70NANB2H3014.

* To whom correspondence should be addressed. Mailing address: Kosan Biosciences, Inc., 3832 Bay Center Place, Hayward, CA 94545. Telephone: (510) 732-5236. Fax: (510) 732-8401. E-mail: kennedy@kosan.com.

[1] Abbreviations: NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase; ORF, open-reading frame; SNAC, *N*-acetyl-cysteamine thioester; UDG−LIC, uracil DNA glycosylase−ligation independent cloning; LBS, ligation by selection; rbs, ribosome-binding site; TE, thioesterase; $_{SYN}$, synthetic version of gene; KS, ketosynthase; AT, acyltransferase; DH, dehydratase; KR, ketoreductase; ER, enoyl-reductase; ACP, acyl carrier protein.

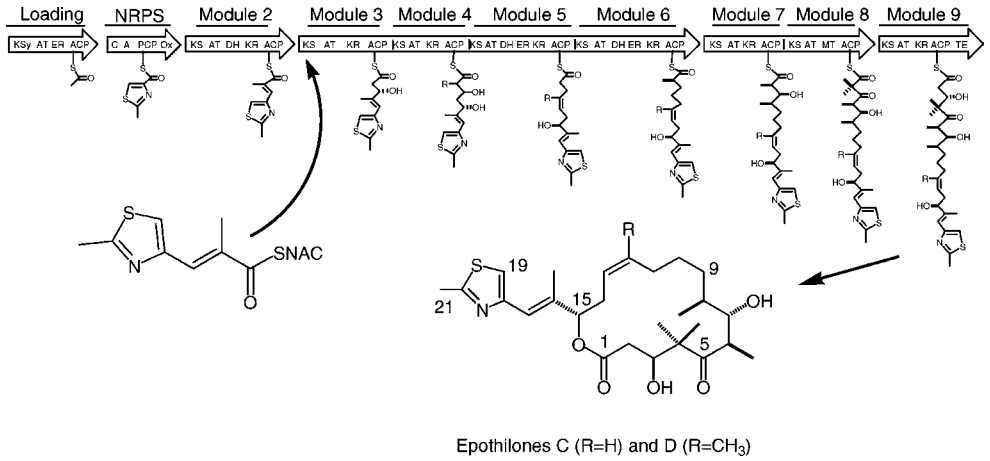
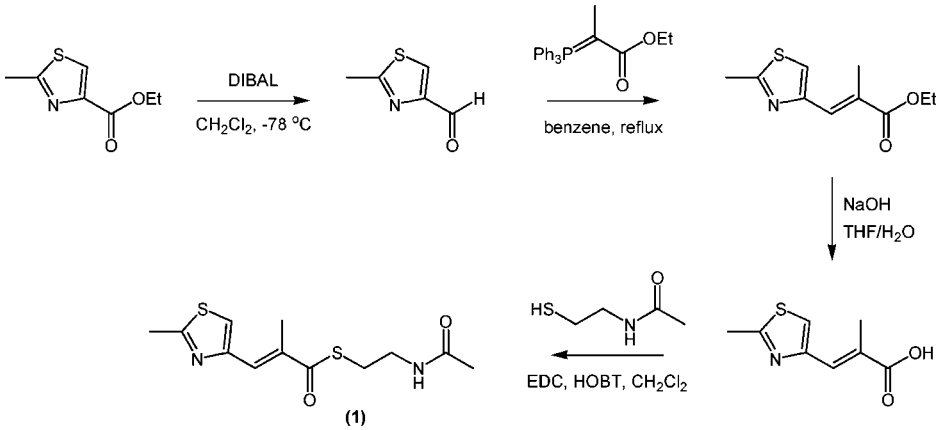


FIGURE 1: Epothilone biosynthesis. The epothilone NRPS/PKS consists of six separate polypeptides each containing the active sites as indicated, leading to the predicted biosynthetic pathway shown. Precursor-directed biosynthesis of epothilones KS3 with an appropriate acyl thioester is also shown. Enzymatic domains on the epothilone NRPS/PKS: KS, ketosynthase; AT, acyltransferase; DH, dehydratase; KR, ketoreductase; ER, enoylreductase; ACP, acyl carrier protein; MT, methyl transferase; TE, thioesterase; C, condensation; A, adenylation; PCP, peptidyl carrier protein; Ox, oxidase.

Scheme 1: Synthesis of Thioester



We have undertaken a program directed toward developing *E. coli* as a production host for the epothilones. In the present work, we describe the total synthesis of the entire 54-kb epothilone PKS gene cluster to facilitate the expression of these large proteins in *E. coli*. When the entire NRPS/PKS synthetic cluster is expressed in an *E. coli* strain engineered for polyketide production, epothilone C and D were produced. Furthermore, we show that feeding the SNAC of a simple polyketide precursor comprising C15−C21 [(*E*)-2-methyl-3-(2-methylthiazol-4-yl)acrylic acid] (Scheme 1, **1**) to *E. coli* cells expressing the terminal seven epothilone modules encoded by *epoD*, *epoE*, and *epoF* results in PKS-catalyzed elongation of the chain and cyclization to form epothilone C and D.

EXPERIMENTAL PROCEDURES

Gene Design and Synthesis. Oligonucleotides for gene synthesis were purchased from Operon (Huntsville, AL). DNA manipulations were performed using standard methods (*16*). *E. coli* DH5α was used for routine plasmid DNA preparation. Restriction enzymes were purchased from New England Biolabs (Ipswich, MA), and other reagents were the highest quality commercially available. The epothilone PKS genes synthesized encoded for the same amino acid sequences reported for the natural proteins (*4*) with exceptions as noted in Table 1. Codon optimization, restriction-site placements, and oligonucleotide design were performed as described previously (*17*−*19*). Briefly, the sequences of each domain were separated into fragments of 129−781 bp in length, and overlapping 40-mer oligonucleotides that encompassed each of the fragments were prepared (*17*). The oligonucleotides were assembled and amplified by PCR using Expand High Fidelity (Roche, Indianapolis, IN) as described by others (*19*, *20*). The crude PCR fragments were cloned into uracil DNA glycosylase−ligation independent cloning/ligation by selection (UDG−LIC/LBS) vectors as described (*18*), and correct clones were identified by DNA sequencing. The fragments were assembled into individual modules by LBS (*18*), and the modules were assembled into the desired ORFs by cloning into conventional vectors. Each ORF contained a unique *Nde* I site at the start codon and an *Eco* RI site before the stop codon so that they could be readily mobilized to expression vectors, and the ORFs were moved as *Nde* I−*Eco* RI fragments into pET22b or pET26b vectors (Novagen, Madison, WI) and cloned into the various expression vectors with C-terminal His tags as described below.

The synthetic epoA gene ($epoA_{SYN}$) was cloned as a *Nde* I−*Eco* RI fragment into pET22b to give pKOS392-61-3 (P_{T7}, ColE1, Am^R). The $epoE_{SYN}$ gene was cloned as a *Nde* I−*Eco*

Table 1: Summary of Epothilone Gene Synthesis

ORF	GenBank accession number	component	size (bp)	amino acids encoded	5′ site	3′ site	designed substitutions	synthons: number	synthons: size range (bp)
epoA	DQ289491	load module	4155	1−1385	*Nde* I	*Xba* I	I889L; D890Q	10	349−628
		C-terminal linker	117	1386−1421 (+3)	*Spe*I	*Eco* RI	A1386S; T1387S	1	156
epoB	DQ289492	module 1, NRPS	4149	1−1383	*Nde* I	*Xba* I	none	9	283−568
		C-terminal linker	90	1384−1410 (+3)	*Spe* I	*Eco* RI	R1384S; R1385S	1	129
epoC	DQ289493	module 2	5379	1−1793	*Nde* I	*Xba* I	I884L; E885Q	11	367−648
		C-terminal linker	126	1794−1832 (+3)	*Spe* I	*Eco* RI	C1794S; E1795S	1	165
epoD	DQ289494	module 3	4518	1−1506	*Nde* I	*Xba* I	I909L; E910Q	9	349−781
		module 4	4497	1507−3005	*Spe* I	*Xba* I	L1507S; 1508S; P2402L; 2403Q	10	388−628
		module 5	6231	3006−5082	*Spe* I	*Xba* I	L3006S; K3007S	13	349−628
		module 6	6387	5083−7211	*Spe* I	*Xba* I	L5083S; 5084S; V5979L; E5980Q	13	388−648
		C-terminal linker	147	7212−7257 (+3)	*Spe* I	*Eco* RI	V7212S	1	147
epoE	DQ289495	module7	4509	1−1503	*Nde* I	*Xba* I	I906L; E907Q	9	388−778
		module 8	6726	1504−3745	*Spe* I	*Xba* I	L1504S; K1505S; I2398L; E2399Q	14	388−622
		C-terminal linker	168	3746−3798 (+3)	*Spe* I	*Eco* RI	F3746S; P3747S	1	168
epoF	DQ289496	module 9	6495	1−2165	*Nde* I	*Xba* I	I902L; E903Q	13	349−667
		TE	831	2166−2439 (+3)	*Spe* I	*Eco* RI	A2166S; V2167S	2	442−474
stiB	DQ289497	C-terminal linker	173	na	*Spe* I	*Eco* RI	na	1	173
stiC	DQ289498	N-terminal linker	109	na	*Nde* I	*Mfe* I	na	1	109

RI fragment into pET22b to give pKOS346-130 (P_{T7}, ColE1, Am^R). The *epoE*$_{SYN}$ gene was subsequently cloned as a *Nde* I−*Hin* DIII fragment from pKOS346-130 into pKOS392-97 (see below) to create the P_{BAD} expression vector pKOS346-159 (P_{BAD}, ColE1, Am^R). The native *epoD* module 3 was modified by PCR to contain a *Nde* I site at the start codon and a *Spe* I site at amino acids 1507−1508, analogous to the *Xba* I site present in the synthetic module 3 described below and cloned in frame with eryTE to create pKOS227-117A (J. Jez, unpublished plasmid). The *epoD* module 3 + eryTE *Nde* I−*Eco* RI fragment was then ligated with similarly digested pKOS392-97 to give pKOS486-59. pKOS392-97 is a derivative of P_{BAD}/Myc-HisA (Invitrogen, Carlsbad, CA), which was modified as follows: (i) the *Nde* I site at nucleotide 2846 was removed by cleavage, fill-in, and blunt-end ligation; and (ii) a *Nde* I site was engineered into the start codon, with concomitant destruction of the *Nco* I site.

The synthetic *epoD* module 3 + eryTE construct was made by combining a *Nde* I−*Xba* I *epoD*$_{SYN}$ module 3 fragment with a *Spe* I−*Eco* RI synthetic TE fragment (*19*, *20*) into a *Nde* I−*Eco* RI digested pKOS392-97 vector to create pKOS431-156-4.

Operons were constructed using a synthetic linker. The linker was created by hybridizing two complementary oligos: 431-151-4 (5′-AATTGTAGCTCCGTCGACAAGCT-TGCGGCCGCACTCGAA**CACCACCACCACCACCAC**-TGAACTAGTTGC) and 431-151-5 (5′-TCGAGCAACTA-GTTCAGTGGTGGTGGTGGTGGTGTTCGAGTGCGGC-CGCAAGCTTGTCGACGGAGCTAC). This linker was designed to encode a polyhistidine (bold) tag followed by a stop codon, include a *Spe* I site (underlined), and have sticky ends compatible with *Eco* RI/*Mfe* I and *Xho* I. The genes were combined sequentially into the appropriate P_{BAD}-driven vectors by digesting the first vector with *Eco* RI and *Xho* I and ligating the hybridized linker as an *Mfe* I−*Xho* I fragment. The resulting intermediate vectors were digested with *Spe* I and *Xho* I, and the second gene was ligated from a pET-derived intermediate as an *Xba* I−*Xho* I fragment including the ribosome-binding site (rbs). The process was then repeated to add a third gene to the operon if needed. The *epoB*$_{SYN}$-rbs-*epoA*$_{SYN}$-rbs-*epoC*$_{SYN}$ operon was constructed with the synthetic linker described above. Briefly, the *epoB*$_{SYN}$ gene was cloned into a derivative of P_{BAD}/Myc-HisA (Invitrogen) into which a *Nde* I site was engineered at the start codon, with concomitant destruction of the *Nco* I site. Then, the *Nde* I/*Hin* DIII fragment was replaced by a *Nde* I−(*Eco* RI−*Xho* I)−*Hin* DIII fragment from a pET22b *epoB*$_{SYN}$ plasmid. The resulting vector, pKOS431-155C, was digested with *Eco* RI and *Xho* I, and the operon linker was added to create pKOS431-172-2. The *epoA*$_{SYN}$ gene, a second operon linker, and the *epoC*$_{SYN}$ gene were sequentially added to pKOS431-172-2 to create pKOS486-2-1 (P_{BAD}-*epoB*$_{SYN}$-rbs-*epoA*$_{SYN}$-rbs-*epoC*$_{SYN}$, ColE1, ampicillin^r). The *epoF*$_{SYN}$-rbs-*epoE*$_{SYN}$ operon was created by first cloning the *Nde* I/*Eco* RI *epoF*$_{SYN}$ gene into pKOS392-97 described above. The *araC*-P_{BAD}-*epoF*$_{SYN}$ expression cassette was then transferred as a *Bst* Z17I−*Eco* RI fragment into a modified pCDF-1b (Novagen) expression vector, pKOS431-39-2. The pCDF-1b modifications included (i) the *Xba* I site at nucleotide 2206 that was eliminated by cleavage, fill-in, and blunt-end ligation and (ii) the *Mlu* I/*Blp* I fragment of the polylinker that was replaced with that from pET26b. The *epoE*$_{SYN}$ gene was then added as an operon using the linker system described above to create pKOS431-188 (P_{BAD}-*epoF*$_{SYN}$-rbs-*epoE*$_{SYN}$, CloDF13, streptomycin^r).

For construction of the *epoD*$_{SYN}$ modules 3,4-rbs-*epoD*$_{SYN}$ modules 5,6 operon, incorporating synthetic linker regions from the stigmatellin *stiB* and *stiC* genes (*21*), the *stiB* C-terminal linker region was first added to a *epoD*$_{SYN}$ modules 3,4 fragment as a *Spe* I/*Eco* RI fragment. The *epoD*$_{SYN}$ modules 3,4 fragment was then cloned as a *Nde* I/*Eco* RI fragment into pKOS431-39-2 described above to give pKOS455-113A. A version of the *epoD*$_{SYN}$ modules 5,6 gene was constructed in which the N-terminal linker region was replaced with a synthetic N-terminal linker region from the *stiC* gene as a *Nde* I/*Mfe* I fragment. The two synthetic *epoD* gene fragments, incorporating the stigmatellin PKS linker fragments, were then combined as an operon using the linker system described above to create pKOS455-

152 (P_{BAD}-*epoD*$_{SYN}$ modules 3,4-rbs-*epoD*$_{SYN}$ modules 5,6, CloDF13, streptomycin^r). pRSF-1b (Novagen) was modified as follows to give pKOS431-39-1: (i) the *Xba* I site at nucleotide 2254 was eliminated by cleavage, fill-in, and blunt-end ligation; and (ii) the *Mlu* I/*Blp* I fragment of the polylinker was replaced with that from pET26b. The *Eco* RV−*Xho* I fragment from pKOS455-152, containing the operon, was then transferred to pKOS431-39-1 to give pKOS455-166 (P_{BAD}-*epoD*$_{SYN}$ modules 3,4-rbs-*epoD*$_{SYN}$ modules 5,6, RSF1030, kanamycin^r). Chaperone coexpression plasmids were purchased as a set (Takara Mirus Bio, Madison, WI) consisting of the chloramphenicol^r p15A plasmids pG-KJE8 (P_{BAD} *dnaK-dnaJ-grpE,* Pzt-1$_P$ *groES-groEL),* pGro7 (P_{BAD} *groES-groEL),* pKJE7 (P_{BAD} *dnaK-dnaJ-grpE),* pG-Tf2 (Pzt-1$_P$ *groES-groEL-tig*), and pTf16 (P_{BAD} *tig*) (*22*, *23*) and were used as recommended.

Strain Construction. The *E. coli* polyketide production strain, K207-3, has been described (*11*). K431-37-2A is a modified version of K207-3 in which arabinose utilization genes were eliminated by P1 transduction from strains RMK12 (*24*) and WM2923 (a generous gift from Professor J. Cronan) to create strain K332-176 [K207-3 Δ*araD139* (*araBAD-leu*); Δ*araFGH2908*]. The arabinose transporter gene, *araE*, was replaced in two steps by pKO3-mediated (*25*) insertion of a constitutive, mutant *lacY* gene [*lacYA177C* (*26*)] (reported to efficiently transport arabinose) at the *araE* locus, causing a total deletion of the *araE* coding region in *E. coli* strain, ET505 (*E. coli* Genetic Stock Center at Yale University). The disrupted *araE* locus was then transferred to K332-176 by P1 transduction to create strain K431-37-2A [K207-3 Δ*araD139* (*araBAD-leu*); Δ*araFGH2908*; *lysA::Tn10*, *araE::lacYA177C*]. The inability to utilize arabinose coupled with constitutive expression of an efficient arabinose transport system made this strain ideal for consistent arabinose-inducible gene expression.

Protein Expression. Culture tubes containing 2xYT (1.6% tryptone, 1.0% yeast extract, and 0.5% NaCl) and appropriate antibiotics were inoculated with an overnight culture (1:50 dilution) and grown at 30 °C to mid-log phase (0.4−0.6 OD_{600}/mL). Gene expression was induced with 0.5 mM IPTG, 2 mg/mL L-arabinose, and/or 20 ng/mL tetracycline, as appropriate, and cultures were grown overnight at 15−22 °C. Samples (3 OD_{600} units) of each culture were resuspended in 1 mL 20 mM Tris and 150 mM NaCl at pH 7.5, containing 1 tablet of Complete EDTA-free protease inhibitor cocktail (Roche) per 50 mL, lysed by sonication, and centrifuged at 14000*g* for 3 min. Pellets were resuspended in 0.5 mL of lysis buffer, and soluble and insoluble fractions equivalent to 0.03 OD_{600} cell suspension were analyzed on NuPAGE Novex 3−8% Tris-Acetate gels (Invitrogen). Gels were stained using Simply Blue SafeStain (Invitrogen) and analyzed using a UVP Bioimaging System with Labworks 4.0 software (UVP, Inc., Upland, CA).

Assay for EpoD-mod3-TE Activity. E. coli strains K486-63-1 [K431-37-2A + pG-TF2 + pKOS486-59] and K486-63-2 [K431-37-2A + pG-TF2 + pKOS431-156-4] were subcultured at a 1:50 dilution from an overnight culture into LB medium plus appropriate antibiotics and grown at 37 °C. After 1.75 h, 20 ng/mL tetracycline was added to induce chaperone expression and growth was continued at 37 °C to 0.4 OD_{600}/mL. The cultures were then induced and supplemented to 5 mM propionate, 50 mM succinate, 50 mM glutamate, 0.2% arabinose, 0.5 mM IPTG, and 0.05% thiazole diketide-SNAC (**1**). Cultures were grown at 22 °C. After 1 day, aliquots were removed for protein expression analysis as described above. After 6 days, supernatants were clarified by centrifugation and analyzed by LC/MS/MS as described below.

Production of Epothilone in E. coli. For precursor directed biosynthesis of epothilone in *E. coli*, strain K486-138 [K431-37-2A + pGTf-KJE + pKOS486-131 + pKOS431-188] was subcultured at a 1:50 dilution from an overnight culture into 2xYT medium plus appropriate antibiotics and grown at 30 °C to 0.4 OD_{600}/mL. The cultures were then induced and supplemented with 5 mM propionate (required for methylmalonyl-CoA production via the propionyl-CoA carboxylase route), 0.2% arabinose, 0.5 mM IPTG, 50 μg/mL chlortetracycline, and 0.05% thiazole diketide-SNAC (**1**). Cultures were grown at 15 °C for 6 days and extracted with an equal volume of ethyl acetate. The extracts were dried in vacuo, dissolved in methanol, and analyzed by LC/MS/MS as described below. Authentic epothilone C and D were isolated from *M. xanthus* as described (*27*) for use as standards.

For *de novo* synthesis of epothilone in *E. coli*, strain K486-62-1 #3 [K431-37-2A + pGTf-KJE + pKOS486-2-1 + pKOS455-166 #52 + pKOS431-188] was subcultured at a 1:50 dilution from an overnight culture into 2xYT medium plus appropriate antibiotics and grown at 30 °C to 0.4 OD_{600}/mL. The cultures were then induced and supplemented with 5 mM propionate, 0.2% arabinose, 0.5 mM IPTG, and 50 μg/mL chlortetracycline. Cultures were grown with XAD at 15 °C for 6 days, and the resin was extracted twice with methanol. The extracts were dried in vacuo, dissolved in methanol, and analyzed by LC/MS/MS as described below.

Analysis of the Production of Epothilones C and D and Intermediates. Samples (10 μL) were analyzed by LC/MS/MS on a system comprised of an HTC PAL autosampler, Agilent 1100 HPLC pump and an Applied Biosystems API-3000 triple quadrupole mass spectrometer equipped with a Turboion spray source operated in positive-ion mode. For detection of epothilones C and D, the analyte was injected onto a Zorbax Eclipse XDB-C18 column (2.1 × 150 mm, 3.5 μm, Agilent) and eluted with a linear gradient of 35% acetonitrile (0.1% formic acid) to acetonitrile (0.1% formic acid) at 0.25 mL/min over 12 min. The eluate was subjected to multiple reaction monitoring of *m*/*z* 478 → 290 for epothilone C and 492 → 304 for epothilone D.

To analyze samples for (*E*)-3-hydroxy-4-methyl-5-(2-methylthiazol-4-yl)pent-4-enoic acid (see Figure 4, **2**) chromatography was performed on a Polarity dC18 column (2.1 × 100 mm, 3 μm, Waters), using a linear gradient of 20% acetonitrile (0.1% formic acid) to acetonitrile (0.1% formic acid) at 0.25 mL/min over 6 min. For detection, two parent−daughter pairs, *m*/*z* 228 → 210 and 228 → 164, were monitored.

*Synthesis of Thioester **1** (Scheme 1).* To a cold solution (−78 °C) of ethyl 2-methylthiazole-4-carboxylate (2.0 g, 11.7 mmol) in CH_2Cl_2 (40 mL) was added slowly diisobutylaluminum hydride (DIBAL-H, 1 M solution in hexane, 20 mL, 20 mmol). The solution was stirred at −78 °C for 2 h before a small amount of methanol was added to quench the reaction. The mixture was warmed to room temperature, and saturated aqueous NH_4Cl was added and extracted 3

times with EtOAc. The combined organic layers were dried over $MgSO_4$ and concentrated to give 1.64 g of crude aldehyde.

The crude aldehyde was mixed with (carbethoxyethylidene) triphenylphosphorane (4.0 g, 1.2 equiv) in benzene (20 mL) and refluxed for 3 h. The solvent was removed under reduced pressure, and the residue was filtered through a silica gel disk, eluting with 3:1 hexane/EtOAc. The resulting material was further purified with silica gel chromatography (10−20% acetone in hexane) to afford 1.6 g (65% yield for two steps) of product. ^{1}H NMR ($CDCl_3$) δ: 7.51 (s, 1H), 7.17 (s, 1H), 4.14 (q, 2H, J = 7.1), 2.61 (s, 3H), 2.22 (s, 3H), 1.22 (t, 3H, J = 7.1). ^{13}C NMR ($CDCl_3$) δ: 168.4, 165.0, 151.7, 130.2, 128.5, 121.0, 60.5, 19.0, 14.0 (2C).

A solution of the unsaturated ester (1.6 g, 2.8 mmol) in 1 M NaOH solution (9 mL, 9.0 mmol) and THF (20 mL) was stirred at room temperature for 2 h. The majority of THF was removed under reduced pressure, and the concentrated mixture was stirred at room temperature overnight. The solution was acidified to pH 1 with 1 N HCl, while being cooled in an ice bath, extracted with CH_2Cl_2 (4×), and dried over $MgSO_4$. The product (1.34 g) was obtained after the removal of the solvent. ^{1}H NMR ($CDCl_3$) δ: 7.74 (s, 1H), 7.33 (s, 1H), 2.76 (s, 3H), 2.35 (s, 3H). ^{13}C NMR ($CDCl_3$) δ: 174.1, 165.7, 151.8, 132.4, 127.9, 122.3, 19.3, 14.0.

To the acid in CH_2Cl_2 (50 mL) were added solutions of 1-[3-(dimethylamino)-propyl]-3-ethylcarbodiimide hydrochloride (EDC, 1.97 g, 1.4 equiv) and 4-(dimethylamino)-pyridine (70 mg) in CH_2Cl_2 (25 mL), and *N*-acetylcysteamine (1.31 g, 1.5 equiv) and 1-hydroxybenzotriazole (HOBT, 1.39 g, 1.4 equiv) in CH_2Cl_2 (25 mL) simultaneously via two syringes. After the mixture was stirred at room temperature overnight, it was diluted with CH_2Cl_2, washed with AcOH/NaOAc buffer (2×) at pH 4, saturated $NaHCO_3$ (2×), and dilute $CuSO_4$ (2×), and dried over $MgSO_4$. The crude product (2.0 g) was purified with silica gel chromatography (10−30% acetone in hexane) to provide 1.68 g (78% for two steps) of thioester **1**. ^{1}H NMR ($CDCl_3$) δ: 7.52 (s, 1H), 7.31 (s, 1H), 6.27 (br s, 1H), 3.44 (q, 2H, J = 6.2), 3.09 (t, 2H, J = 6.4), 2.70 (s, 3H), 2.34 (s, 3H), 1.94 (s, 3H). ^{13}C NMR ($CDCl_3$) δ: 194.4, 170.3, 165.6, 151.3, 135.7, 129.4, 122.5, 39.5, 28.7, 23.1, 19.3, 14.3.

RESULTS

Synthesis and Expression of Epothilone Genes. The DNA sequences of the ORFs corresponding to *epoA*$_{SYN}$ (loading module; 4272 bp), *epoB*$_{SYN}$ (module 1; 4239 bp), *epoC*$_{SYN}$ (module 2; 5505 bp), *epoD*$_{SYN}$ (modules 3, 4, 5, and 6; 21 780 bp), *epoE*$_{SYN}$ (modules 7 and 8; 11 403 bp), and *epoF*$_{SYN}$ (module 9, TE; 7326 bp) were redesigned to contain preferred *E. coli* codons and restriction sites convenient for cloning (see Table 1). For DNA synthesis, PCR assembly/amplification (*17*) was used to assemble 40-mer overlapping oligonucleotides into 283−731-bp fragments encompassing these ORFs. Fragments with confirmed sequences were then ligated together to prepare the individual modules and the TE. The component modules and the TE were then combined to form the *epoA*$_{SYN}$, *epoB*$_{SYN}$, epoC$_{SYN}$, *epoD*$_{SYN}$, *epoE*$_{SYN}$, and *epoF*$_{SYN}$ ORFs. The epothilone PKS genes synthesized encoded for the same amino acid sequences reported for the natural proteins (*4*), with exceptions made for cloning convenience as noted in Table 1.

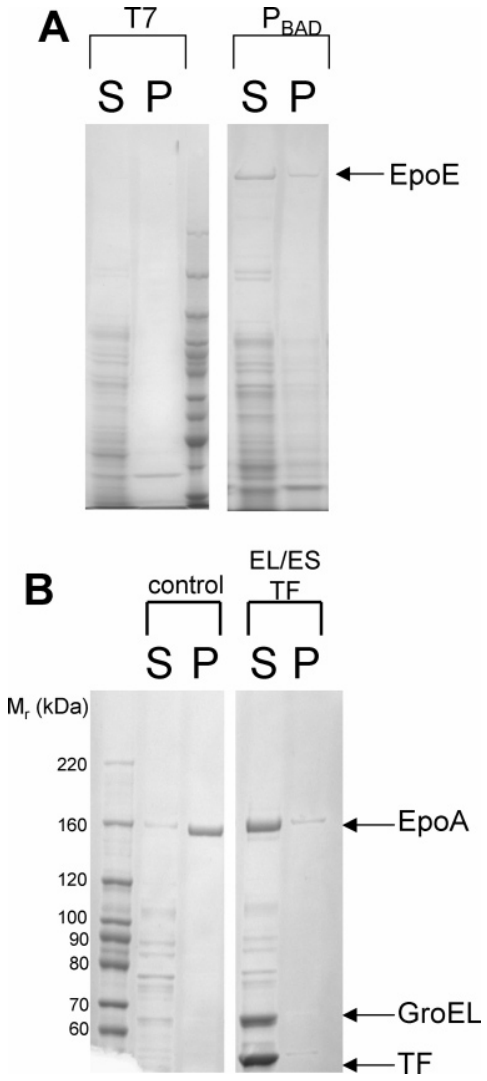


FIGURE 2: Improvements in solubility of heterologously expressed epothilone proteins. (A) EpoE was expressed from the T7 promoter (pKOS346-130) or the P_{BAD} promoter (pKOS346-159). (B) EpoA was expressed in K207-3 from the T7 promoter (pKOS392-61-3) at 15 °C in the presence (EL/ES/TF) or absence (control) of groEL/groES/trigger factor coexpression. S, soluble fraction; P, pellet (insoluble fraction).

Initial attempts to express active epothilone PKS genes in *E. coli* were hampered by low expression levels and limited solubility. We used a combination of several methods to overcome these problems. Low-temperature expression has previously been observed to enhance *E. coli* soluble expression of large PKS proteins in a number of laboratories (*10*, *14*, *28*). This approach also improved the solubility of the epothilone PKS proteins, as judged by SDS−PAGE of soluble and insoluble fractions (data not shown). Furthermore, we found switching from a T7 promoter system to the arabinose-induced P_{BAD} promoter in a strain of *E. coli* unable to utilize arabinose not only improved protein solubility but also could improve total expression levels. As shown in Figure 2A, when *epoE*$_{SYN}$ was expressed from the T7 promoter, no protein was detectable. However, expression from the P_{BAD} promoter produced significant levels of protein, most of which was soluble. Finally, it has been previously demonstrated that chaperone coexpression can improve the solubility of heterologously expressed proteins

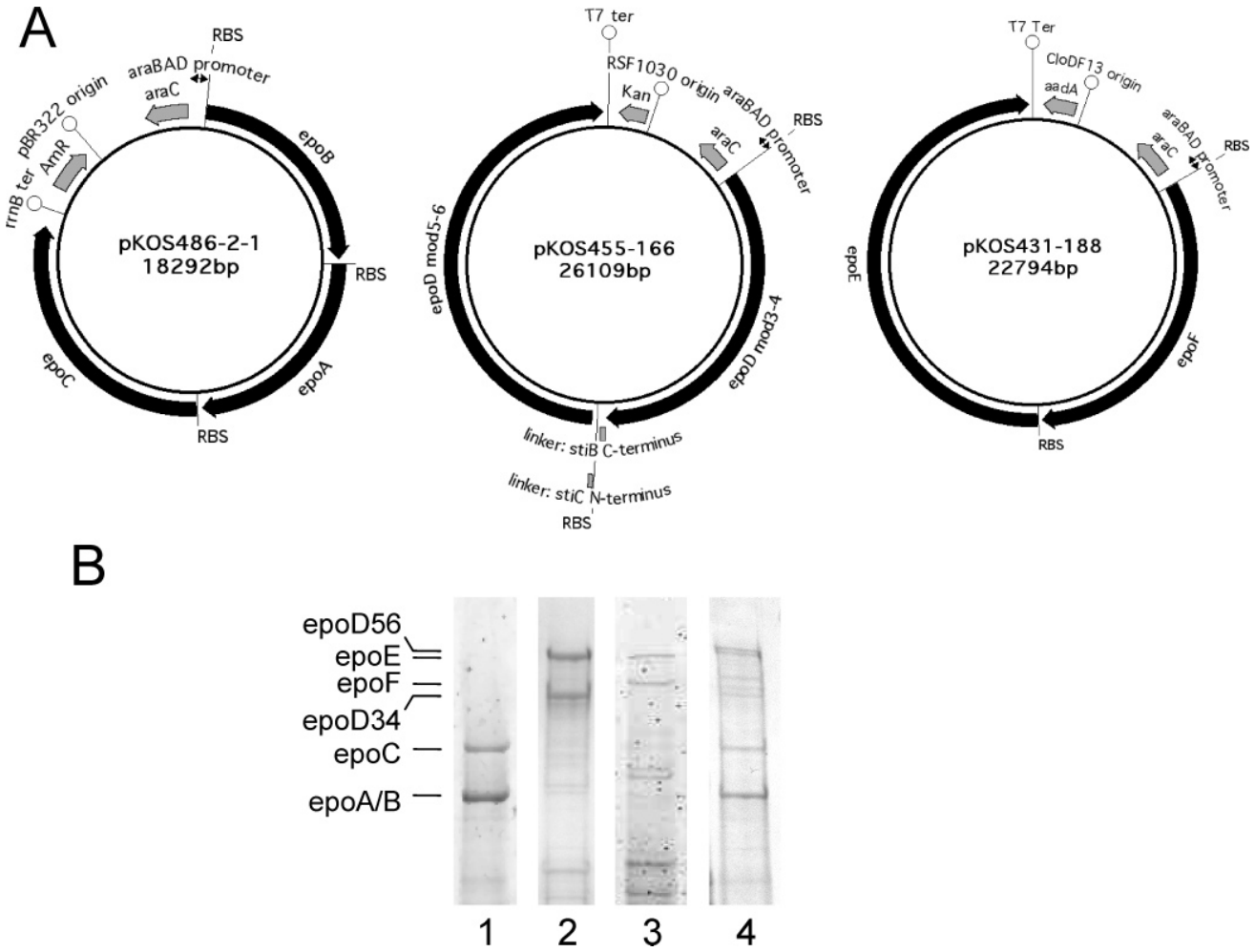


FIGURE 3: SDS−PAGE analysis of heterologously expressed epothilone NRPS/PKS. (A) Three plasmids used for the expression of the synthetic epothilone NRPS/PKS in *E. coli*. The plasmids encode synthetic operons with pKOS486-2-1 encoding *epo*B, *epo*A, and *epo*C, pKOS455-166 encoding *epo*D(mod3−4) and *epo*D(mod5−6), and pKOS431-188 encoding *epo*F and *epo*E. (B) SDS−PAGE analysis of the expression of the epothilone NRPS/PKS from these plasmids in *E. coli*. The gel is stained with Coomassie. Lane 1, extract from an *E. coli* strain containing plasmid pKOS486-2-1; lane 2, extract from an *E. coli* strain containing plasmid pKOS455-166; lane 3, extract from an *E. coli* strain containing plasmid pKOS431-188; and lane 4, extract from an *E. coli* strain bearing all three plasmids, expressing the entire epothilone NRPS/PKS.

in *E. coli* (*22*, *23*), and we found this technique necessary for the soluble expression of epothilone PKS proteins. Prior attempts, by other laboratories, to express EpoA in *E. coli* were unsuccessful with only the ACP domain expressing as a soluble protein (*13*, *14*). Here, we show that coexpression of the *E. coli* chaperone proteins GroES/EL and trigger factor resulted in a large increase in the level of soluble EpoA protein (Figure 2B). For other epothilone biosynthesis enzymes, we found that expression of DnaKJ and GrpE also aided in soluble protein expression (data not shown). The use of chaperone proteins and alternative promoter systems to aid in the heterologous expression of NRPS/PKS has not previously been reported.

Combining the three tactics of lowered temperature, chaperone coexpression, and use of the P_{BAD} promoter produced soluble protein from all genes with the exception of the 21.8-kb *epoD*$_{SYN}$. This synthetic gene, encoding a protein greater than 760 kDa in size containing modules 3−6, could not be expressed in *E. coli*. Truncated versions of *epoD*$_{SYN}$, consisting of modules 3, 4, and 5 or modules 4, 5, and 6, however, could be expressed (data not shown), indicating that there were no intrinsic obstacles to expression of this protein in *E. coli* and that the expression problems were likely to be due to the size of the protein. After an approach used for the production of an ansamycin polyketide precursor (*28*), the *epoD*$_{SYN}$ gene was expressed as two smaller polypeptides, each consisting of two modules, *epoD*$_{SYN}$ modules 3,4 and *epoD*$_{SYN}$ modules 5,6 (lane 2 in Figure 3). The C- and N-terminal regions of modular polyketide synthases have been shown to be involved in providing the protein−protein interactions necessary for the efficient transfer of enzyme-bound intermediates from one PKS module to another. To facilitate interactions between epoD modules 4 and 5, now expressed on separate polypeptides, compatible linker pairs from related polyketide synthases were added to the C terminus of *epoD*$_{SYN}$ modules 3,4 and to the N terminus of *epoD*$_{SYN}$ modules 5,6. Several linker pairs were evaluated, and preliminary testing showed that the linker pair from *stiB* and *stiC* in the stigmatellin PKS (*21*) allowed functional interaction between epoD module 4 and epoD module 5 while maintaining soluble expression (data not shown). For soluble expression of the entire synthetic gene cluster, the P_{BAD} promoter, reduced

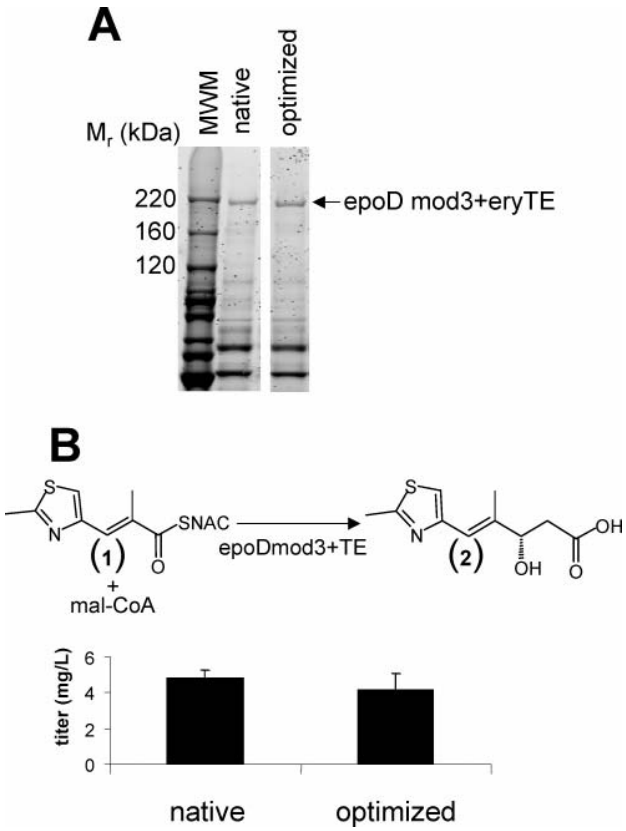


FIGURE 4: Activities of EpoD module 3 proteins from native and codon-optimized genes are equivalent. (A) Expression levels of soluble EpoD mod3 protein from *E. coli* strains with native and codon-optimized *epo*D. The conversion of thioester **1** to the product **2** from these strains was determined by LC/MS with titers shown in B ($n = 4$).

growth temperature, and chaperone coexpression were used (Figure 3).

Activity of Proteins from Synthetic versus Native Genes. To determine if there were deleterious effects caused by amino acid substitutions in the synthetic genes, we compared the *in vivo* activity of EpoD module 3 translated either from the native gene sequence or the codon-optimized gene. The single modules were fused to the TE from 6-deoxyerythronolide B synthase (DEBS) and expressed from the P_{BAD} promoter. DEBS TE has been frequently used to terminate polyketide synthesis and has been shown to be reasonably nonspecific for the polyketide chain that it hydrolyzes (*29*). As shown in Figure 4, levels of soluble EpoD module 3 protein and product formation are similar for the native and optimized genes. This demonstrates that the codon optimization and limited amino acid substitutions introduced in the synthetic gene do not result in a detectable loss of activity for this module. While this does not prove that all of the synthetic epothilone genes are similarly active, it does demonstrate that the amino acid substitutions (I909L and E910Q) at the AT-KR interface are well-tolerated. Similar substitutions immediately upstream of the KR are present in seven of the other synthetic epothilone modules. While other amino acid substitutions occur in the synthetic epothilone modules, these are limited to the intermodular regions. Work by Menzella et al. (*29*) has demonstrated that analogous changes in these regions are less likely to interfere with enzymatic activity.

Production of Epothilones C and D in E. coli. To provide a catalytically active holo-PKS, each of the subunits must undergo appropriate folding and form a protein complex of approximately 2 MDa (monomeric mass). To assess functionality of the epothilone NRPS/PKS complex, we first tested the activity of a truncated version using precursor-directed biosynthesis (*30*). Here, a SNAC of a polyketide resembling the structure of the normal substrate of a module is fed to cells harboring the PKS. In a catalytically active system, a thioester exchange occurs between the SNAC and the active-site Cys of the KS domain in the initial module and chain elongation proceeds through the remaining modules (*30*) (Figure 1).

We synthesized the thioester (**1**) of the polyketide chain that is normally formed at EpoD−module 2 and transferred to module 3 for elongation. When the precursor SNAC was fed to cells expressing the EpoD−EpoE−EpoF complex, LC/MS/MS analysis showed the presence of a compound with a retention time and mass spectrum identical to authentic epothilone C (Figure 5).

Through synthesis of the codon-optimized version of each epothilone ORF, we were able to circumvent the difficulties inherent in expressing genes with high G+C content in *E. coli.* Soluble expression of the entire epothilone NRPS/PKS gene cluster in *E. coli* was achieved by further modification to expression plasmids, host strains, and growth conditions (Figure 3). When the entire synthetic epothilone gene cluster was expressed, LC/MS/MS analysis again showed the presence of compounds with retention times identical to authentic epothilones C and D (Figure 6). Thus, we conclude with certainty that we can express the active epothilone NRPS/PKS complex in *E. coli* and produce authentic epothilones C and D either *de novo* or via precursor-directed biosynthesis.

DISCUSSION

Using the previously reported gene synthesis and design technology (*18*, *29*), we have redesigned and synthesized the six ORFs encoding the entire epothilone NRPS/PKS gene cluster. Expression of these large proteins in a soluble form required lowering the incubation temperature, coexpression of molecular chaperones, and splitting the largest protein into two polypeptides. Chaperone coexpression had not previously been reported to aid the folding of these large PKS proteins, and this may be of general utility for heterologous expression of PKSs. Furthermore, chaperone systems from native polyketide-producing hosts may be better adapted for folding of large, multimodular PKS proteins, and coexpression of the cognate chaperones in *E. coli* could result in more efficient folding. Expression of the tetramodular EpoD protein (765 kDa) proved to be inefficient in *E. coli.* A protein of this size may be above the upper limit of the capabilities of the *E. coli* translational apparatus; the largest known natural polypeptide in *E. coli* is the large subunit of glutamate synthase (166 kDa) (*32*), although ORFs encoding larger polypeptides have also been found (e.g., *yeeJ* encodes a hypothetical 248 kDa polypeptide) (*33*). To overcome the EpoD production problem, the synthetic *epoD* gene was split into two bimodular components, and these were expressed together with appropriate linker regions from the related stigmatellin PKS. This strategy resulted in reconstitution of

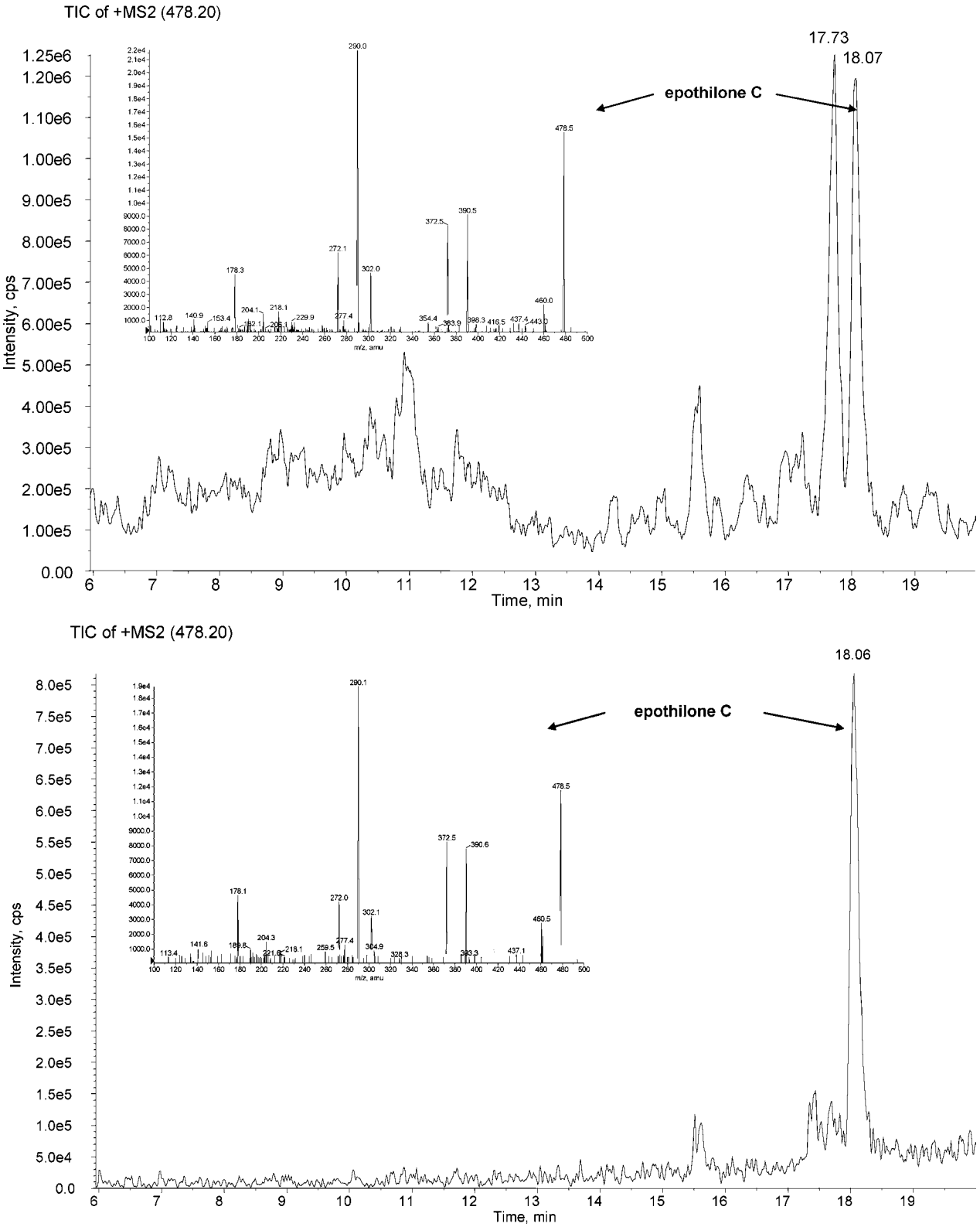


FIGURE 5: Biosynthesis of epothilone C in *E. coli* by thioester feeding. LC/MS/MS analysis of *E. coli* extracts. (Top chromatogram) TIC of product−ion of *m*/*z* 478 with an inset of the mass spectrum of the peak at 18.07 min of a sample from thioester feeding. The peak at 17.73 min is an unknown not related to epothilones. Estimated titers of epothilone C are 10 μg/L. (Bottom chromatogram) TIC of product−ion of *m*/*z* 478 with an inset of the mass spectrum of the peak at 18.06 min of a standard of epothilone C.

the EpoD activity. Previously, attempts to express epothilone biosynthesis genes in *E. coli* gave mixed results. While EpoB, EpoC, EpoE, and EpoF expression has been achieved with native genes, attempts to produce full-length EpoA and EpoD failed (*13*−*15*). In this work, the use of synthetic genes, with codon usage optimized for *E. coli*, has allowed the simultaneous expression of the full epothilone NRPS/PKS in *E. coli*, thus validating the synthetic approach.

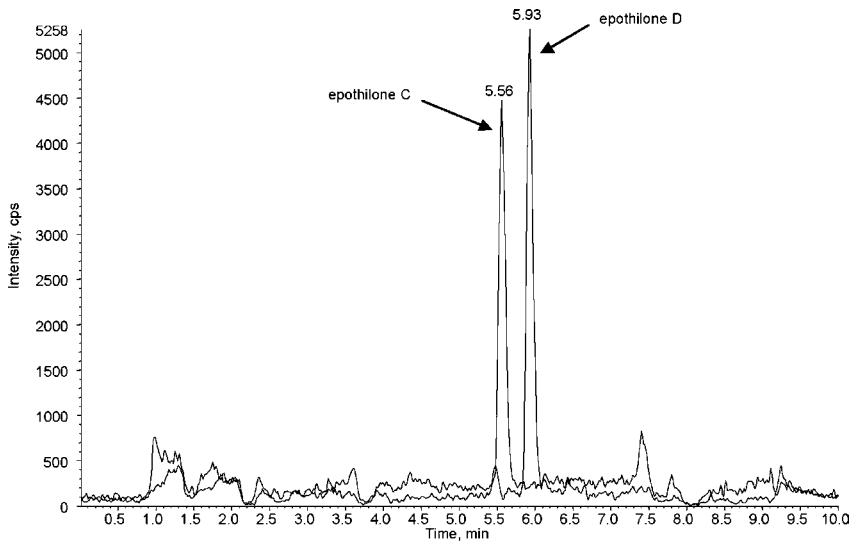


FIGURE 6: Complete biosynthesis of epothilones C and D in *E. coli*. LC/MS/MS analysis of *E. coli* extracts. Samples were subjected to multiple reaction monitoring of *m*/*z* 478 → 290 for epothilone C and 492 → 304 for epothilone D. Titers of epothilone C and D from this fermentation are estimated at <1 μg/L.

Expression of the entire ~2 MDa epothilone NRPS/PKS complex in *E. coli* resulted in the formation of authentic epothilones C and D. This demonstrates that the PKS is active and able to catalyze the many independent enzymatic steps necessary for the synthesis of epothilones. Furthermore, as demonstrated by the remarkable improvement in the production of 6dEB in *E. coli*, from 5 mg/L to >1 g/L (*11*, *12*), application of metabolic engineering and process science technologies should lead to a much improved production system for the epothilones in *E. coli*.

Precursor-directed biosynthesis is another attractive method for the production of epothilone analogues. Work by Boddy et al. (*15*) demonstrated the feasibility of feeding a complex aminoacyl-SNAC to a strain of *E. coli* that was engineered to express the final four modules of the epothilone PKS for the production of epothilone C. Here, when all of the modules of EpoD are expressed together with EpoE and EpoF in *E. coli*, we were able to demonstrate precursor-directed biosynthesis of epothilones C and D with a greatly simplified precursor. Feeding simple analogues of the precursor used in this study will allow the production of epothilone analogues inaccessible by biosynthetic engineering alone. The *in vitro* work of Walsh and co-workers using the initial modules of the epothilone NRPS/PKS, EpoB and EpoC, suggests that the epothilone NRPS/PKS has a great deal of flexibility toward different substrates (*14*, *34*). Therefore, precursor-directed biosynthesis in *E. coli* using modified polyketide thioesters will allow us to exploit the system for production of novel epothilone analogues.

Previous work engineering the epothilone NRPS/PKS gene cluster has led to unanticipated results when certain modifications were introduced into modules 4 and 5 of the PKS (*9*, *31*). The design of the synthetic epothilone genes together with *E. coli* expression provides the ideal platform for both the biochemical investigation of the epothilone PKS and the generation of biosynthetic epothilone analogues. As demonstrated by Menzella et al. (*29*), the generic design of the synthetic PKS genes allows rapid interchange of modules and domains, an advance that will facilitate the generation of epothilone analogues in *E. coli*. When the power of gene synthesis is combined with precursor-directed biosynthesis, complex analogues previously obtainable only by total chemical synthesis could be produced by direct fermentation in *E. coli*.

ACKNOWLEDGMENT

The authors thank Ralph Reid for designing the synthetic genes, Sarah Kodumal for technical assistance with gene synthesis and cloning, Jinghua Tan and Zhiyuan Ma for DNA-sequencing work, and Drs. Leonard Katz, Bryan Julien, James Kealey, and Gary Ashley for critical review of the manuscript.

REFERENCES

1. Bollag, D. M., McQueney, P. A., Zhu, J., Hensens, O., Koupal, L., Liesch, J., Goetz, M., Lazarides, E., and Woods, C. M. (1995) Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action, *Cancer Res. 55*, 2325−2333.
2. Kowalski, R. J., Giannakakou, P., and Hamel, E. (1997) Activities of the microtubule-stabilizing agents epothilones A and B with purified tubulin and in cells resistant to paclitaxel (Taxol), *J. Biol. Chem. 272*, 2534−2541.
3. Gerth, K., Bedorf, N., Höfle, G., Irschik, H., and Reichenbach, H. (1996) Epothilones A and B: Antifungal and cytotoxic compounds from *Sorangium cellulosum* (Myxobacteria). Production, physico-chemical, and biological properties, *J. Antibiot. 49*, 560−563.
4. Julien, B., Shah, S., Ziermann, R., Goldman, R., Katz, L., and Khosla, C. (2000) Isolation and characterization of the epothilone biosynthetic gene cluster from *Sorangium cellulosum*, *Gene 249*, 153−160.
5. Molnar, I., Schupp, T., Ono, M., Zirkle, R., Milnamow, M., Nowak-Thompson, B., Engel, N., Toupet, C., Stratmann, A., Cyr, D. D., Gorlach, J., Mayo, J. M., Hu, A., Goff, S., Schmid, J., and Ligon, J. M. (2000) The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90, *Chem. Biol. 7*, 97−109.
6. Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C., and Julien, B. (2000) Cloning and heterologous expression of the epothilone gene cluster, *Science 287*, 640−642.
7. Julien, B., and Shah, S. (2002) Heterologous expression of epothilone biosynthetic genes in *Myxococcus xanthus*, *Antimicrob. Agents Chemother. 46*, 2772−2778.
8. Starks, C. M., Zhou, Y., Liu, F., and Licari, P. J. (2003) Isolation and characterization of new epothilone analogues from recombinant *Myxococcus xanthus* fermentations, *J. Nat. Prod. 66*, 1313−1317.

9. Tang, L., Chung, L., Carney, J. R., Starks, C. M., Licari, P., and Katz, L. (2005) Generation of new epothilones by genetic engineering of a polyketide synthase in *Myxococcus xanthus*, *J. Antibiot. 58*, 178−184.
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. Murli, S., Kennedy, J., Dayem, L. C., Carney, J. R., and Kealey, J. T. (2003) Metabolic engineering of *Escherichia coli* for improved 6-deoxyerythronolide B production, *J. Ind. Microbiol. Biotechnol. 30*, 500−509.
12. Lau, J., Tran, C., Licari, P., and Galazzo, J. (2004) Development of a high cell-density fed-batch bioprocess for the heterologous production of 6-deoxyerythronolide B in *Escherichia coli*, *J. Biotechnol. 110*, 95−103.
13. Chen, H., O'Connor, S., Cane, D. E., and Walsh, C. T. (2001) Epothilone biosynthesis: Assembly of the methylthiazolylcarboxy starter unit on the EpoB subunit, *Chem. Biol. 8*, 899−912.
14. O'Connor, S. E., Chen, H., and Walsh, C. T. (2002) Enzymatic assembly of epothilones: The EpoC subunit and reconstitution of the EpoA-ACP/B/C polyketide and nonribosomal peptide interfaces, *Biochemistry 41*, 5685−5694.
15. Boddy, C. N., Hotta, K., Tse, M. L., Watts, R. E., and Khosla, C. (2004) Precursor-directed biosynthesis of epothilone in *Escherichia coli*, *J. Am. Chem. Soc. 126*, 7436−7437.
16. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
17. Hoover, D. M., and Lubkowski, J. (2002) DNAWorks: An automated method for designing oligonucleotides for PCR-based gene synthesis, *Nucleic Acids Res. 30*, e43.
18. Kodumal, S. J., and Santi, D. V. (2004) DNA ligation by selection, *Biotechniques 37*, 34−40.
19. Kodumal, S. J., Patel, K. G., Reid, R., Menzella, H. G., Welch, M., and Santi, D. V. (2004) Total synthesis of long DNA sequences: Synthesis of a contiguous 32-kb polyketide synthase gene cluster, *Proc. Natl. Acad. Sci. U.S.A. 101*, 15573−15578.
20. Stemmer, W. P., Crameri, A., Ha, K. D., Brennan, T. M., and Heyneker, H. L. (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides, *Gene 164*, 49−53.
21. Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blocker, H., Hofle, G., and Muller, R. (2002) The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase, *J. Biol. Chem. 277*, 13082−13090.
22. Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., and Yura, T. (1998) Chaperone coexpression plasmids: Differential and synergistic roles of DnaK−DnaJ−GrpE and GroEL−GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*, *Appl. Environ. Microbiol. 64*, 1694−1699.
23. Nishihara, K., Kanemori, M., Yanagi, H., and Yura, T. (2000) Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*, *Appl. Environ. Microbiol. 66*, 884−889.
24. Morgan-Kiss, R. M., Wadler, C., and Cronan, J. E., Jr. (2002) Long-term and homogeneous regulation of the *Escherichia coli* araBAD promoter by use of a lactose transporter of relaxed specificity, *Proc. Natl. Acad Sci. U.S.A. 99*, 7373−7377.
25. Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P., and Kushner, S. R. (1989) New method for generating deletions and gene replacements in *Escherichia coli*, *J. Bacteriol. 171*, 4617−4622.
26. Goswitz, V. C., and Brooker, R. J. (1993) Isolation of lactose permease mutants which recognize arabinose, *Membr. Biochem. 10*, 61−70.
27. Lau, J., Frykman, S., Regentin, R., Ou, S., Tsuruta, H., and Licari, P. (2002) Optimizing the heterologous production of epothilone D in *Myxococcus xanthus*, *Biotechnol. Bioeng. 78*, 280−288.
28. Watanabe, K., Rude, M. A., Walsh, C. T., and Khosla, C. (2003) Engineered biosynthesis of an ansamycin polyketide precursor in *Escherichia coli*, *Proc. Natl. Acad. Sci. U.S.A. 100*, 9774−9778.
29. Menzella, H. G., Reid, R., Carney, J. R., Chandran, S. S., Reisinger, S. J., Patel, K. G., Hopwood, D. A., and Santi, D. V. (2005) Combinatorial polyketide biosynthesis by *de novo* design and rearrangement of modular polyketide synthase genes, *Nat. Biotechnol.*
30. Jacobsen, J. R., Hutchinson, C. R., Cane, D. E., and Khosla, C. (1997) Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase, *Science 277*, 367−369.
31. Tang, L., Ward, S., Chung, L., Carney, J. R., Li, Y., Reid, R., and Katz, L. (2004) Elucidating the mechanism of cis double bond formation in epothilone biosynthesis, *J. Am. Chem. Soc. 126*, 46−47.
32. Oliver, G., Gosset, G., Sanchez-Pescador, R., Lozoya, E., Ku, L. M., Flores, N., Becerril, B., Valle, F., and Bolivar, F. (1987) Determination of the nucleotide sequence for the glutamate synthase structural genes of *Escherichia coli* K-12, *Gene 60*, 1−11.
33. Blattner, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12, *Science 277*, 1453−1474.
34. Schneider, T. L., Walsh, C. T., and O'Connor, S. E. (2002) Utilization of alternate substrates by the first three modules of the epothilone synthetase assembly line, *J. Am. Chem. Soc. 124*, 11272−11273.

BI052075R