# Enzymatic Assembly of Epothilones: The EpoC Subunit and Reconstitution of the EpoA-ACP/B/C Polyketide and Nonribosomal Peptide Interfaces†

Sarah E. O'Connor, Huawei Chen, and Christopher T. Walsh*

*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115*

*Received January 7, 2002; Revised Manuscript Received February 28, 2002*

ABSTRACT: The biosynthesis of epothilones, a family of hybrid polyketide (PK)/nonribosomal peptide (NRP) antitumor agents, provides an ideal system to study a hybrid PK/NRP natural product with significant biomedical value. Here the third enzyme involved in epothilone production, the five domain 195 kDa polyketide synthase (PKS) EpoC protein, has been expressed and purified from *Escherichia coli*. EpoC was combined with the first two enzymes of the epothilone biosynthesis pathway, the acyl carrier protein (ACP) domain of EpoA and EpoB, to reconstitute the early steps in epothilone biosynthesis. The acyltransferase (AT) domain of EpoC transfers the methylmalonyl moiety from methylmalonyl-CoA to the holo HS-acyl carrier protein (ACP) in an autoacylation reaction. The ketosynthase (KS) domain of EpoC decarboxylates the methylmalonyl-*S*-EpoC acyl enzyme to generate the carbon nucleophile that reacts with methylthiazolylcarboxyl-*S*-EpoB. The resulting condensation product can be reduced in the presence of NADPH by the ketoreductase (KR) domain of EpoC and then dehydrated by the dehydratase (DH) domain to produce the methylthiazolylmethylacrylyl-*S*-EpoC acyl enzyme intermediate that serves as the acyl donor for subsequent elongation of the epothilone chain. The acetyl-CoA donor can be replaced with propionyl-CoA, isobutyryl-CoA, and benzoyl-CoA and the acyl chains accepted by both EpoB and EpoC subunits to produce ethyl-, isopropyl-, and phenylthiazolylmethylacrylyl-*S*-EpoC acyl enzyme intermediates, suggesting that future combinatorial biosynthetic variations in epothilone assembly may be feasible. These results demonstrate in vitro reconstitution of both the PKS/NRPS interface (EpoA-ACP/B) and the NRPS/PKS interface (EpoB/C) in the assembly line for this antitumor natural product.

Both polyketide (PK)[1] and nonribosomal peptide (NRP) products are biosynthesized by parallel enzymatic assembly lines, a series of multimodular proteins that sequentially construct PK and NRP products from acyl-CoA or amino acid building blocks via a thiotemplated mechanism (*1*). The polyketide synthases (PKS) utilize acyl-CoA substrates, in which the acyltransferase (AT) domain autocatalytically transfers the acyl group to the phosphopantetheinyl group attached to the acyl carrier protein (ACP) domain (*2*). Analogously, in the nonribosomal peptide synthetases (NRPS), amino acids are activated as the AMP ester by the adenylation (A) domain and then transferred to the phosphopantetheinyl moiety of the peptidyl carrier protein (PCP) domain (*3*). A PKS also contains a ketosynthase (KS) domain that catalyzes both the decarboxylation of the acyl substrate and the subsequent condensation of the resulting nucleophile to the upstream polyketide chain. Analogously, the NRPS contain a condensation (C) domain that catalyzes amide bond formation between the downstream amino acid substrate and the upstream growing peptide chain. Additional domains that tailor the PK or NRP product may also be present; NRPS proteins, for example, may contain oxidase or epimerization domains, while PKS proteins may contain ketoreductase (KR), dehydratase (DH), or enoylreductase (ER) domains.

Several hybrid PK and NRP products have been observed in nature, indicating that in certain cases the NRPS and PKS assembly lines can intersect to produce PK/NRP hybrid products (*4*−*6*). Examples of hybrid products include yersiniabactin (*7*), myxothiazole (*8*), myxalamide (*9*), and, as described here, epothilone (*10*, *11*). Recent successes in PKS bioengineering to generate novel PK products (*12*, *13*) suggest that bioengineering efforts could also be extended toward the production of hybrid PK/NRP products.

We have recently focused on the biosynthesis of the epothilones, a family of hybrid PK/NRP products that are produced by the myxobacterium *Sorangium cellulosum* (*14*). Although many variants of the epothilones are produced in vivo, only epothilones A−D are produced in significant quantities (Figure 1A) (*15*). Like the well-known antitumor agent paclitaxel (Taxol), the epothilones have proven to be potent tubulin depolymerization inhibitors in vitro (*16*, *17*), and the anticancer efficacy of epothilone B and epothilone D is currently being assessed in clinical trials. The gene

† This work was supported by National Institutes of Health Grant GM20011. S.E.O. is an Irving S. Sigal Postdoctoral Fellow.

* To whom correspondence should be addressed: tel, (617) 432-1715; fax, (617) 432-0438; e-mail, christopher_walsh@hms.harvard.edu.

[1] Abbreviations: A, adenylation; ACP, acyl carrier protein; AT, acyl transferase; C, condensation; Cy, cyclization; DH, dehydratase; DME, ethylene glycol dimethyl ether; ER, enoyl reductase; HPLC, high-pressure liquid chromatography; KR, ketoreductase; KS, ketosynthase; MALDI, matrix-assisted laser desorption ionization; NADPH, nicotinamide adenine dinucleotide phosphate, reduced form; NMR, nuclear magnetic resonance; NRP, nonribosomal peptide; NRPS, nonribosomal peptide synthetase; Ox, oxidase; PCP, peptidyl carrier protein; PCR, polymerase chain reaction; PK, polyketide; PKS, polyketide synthase; *S*-NAC, *N*-acetylcysteamine; TCA, trichloroacetic acid; TCEP, tris(2-carboxyethyl)phosphine hydrochloride; TFA, trifluoroacetic acid.

10.1021/bi020006w CCC: $22.00 
Published on Web 03/29/2002

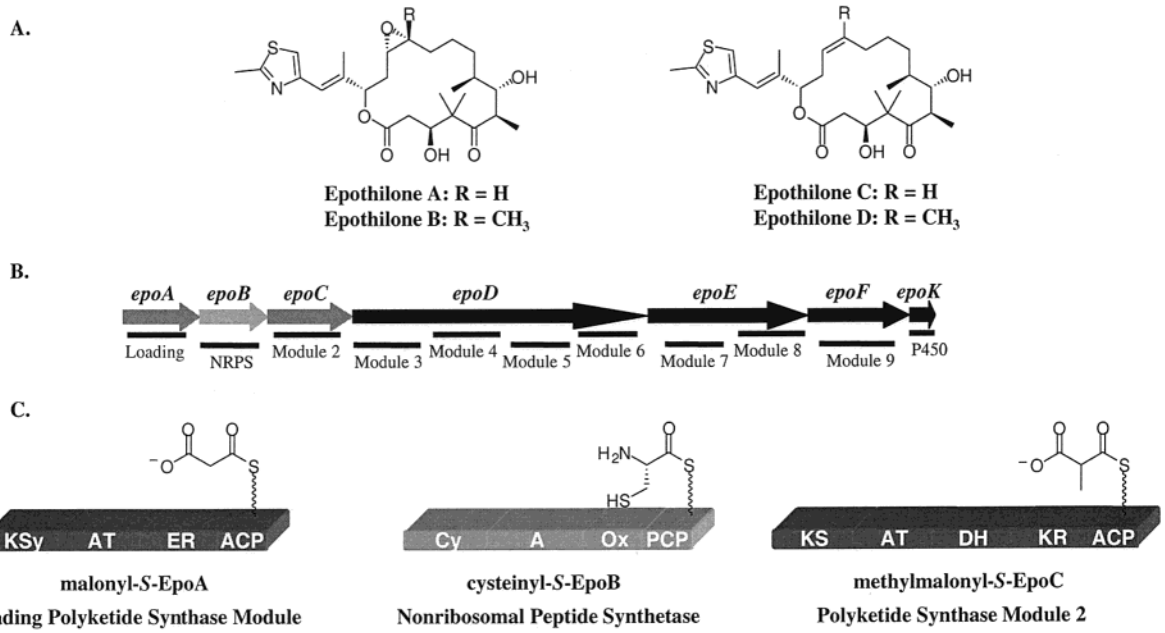



FIGURE 1: (A) Epothilones A−D. (B) Epothilone biosynthetic gene cluster. (C) EpoA, EpoB, and EpoC with thioester-linked substrates.

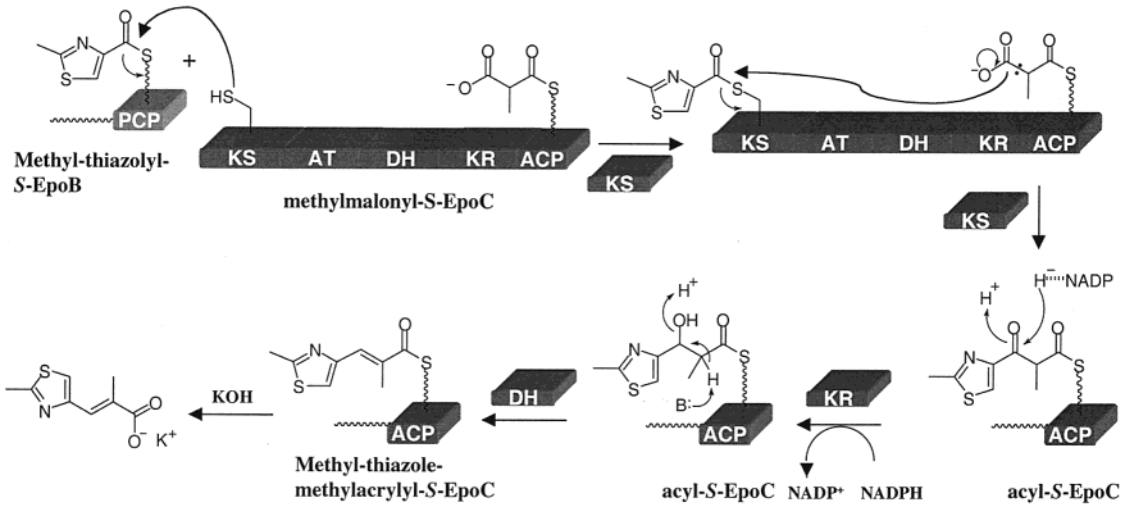



FIGURE 2: Proposed mechanism of methylthiazolylmethylacrylyl-*S*-EpoC formation from acetyl-*S*-EpoA-ACP, cysteinyl-*S*-EpoB, and methylmalonyl-*S*-EpoC.

cluster for the epothilones has been isolated and consists largely of PKS modules, with a single NRPS module, EpoB, located between the first (EpoA) and third (EpoC) PKS modules of the cluster (Figure 1B) (*10*, *11*). Epothilone biosynthesis provides an ideal opportunity to study a hybrid PKS/NRPS/PKS interface in a natural product with significant biomedical importance.

In earlier work the EpoA-ACP/EpoB PK-NRPS interface was characterized by monitoring the interaction of EpoB and the ACP domain of EpoA (*18*); in this work purification and characterization of EpoC, the third synthase of the pathway, are described. Furthermore, the reconstitution of the entire EpoA-ACP/EpoB/EpoC PKS/NRPS/PKS hybrid interface has been demonstrated. EpoC is a 195 kDa five domain PKS, containing ACP, AT, KS, KR, and DH domains (Figure 1C). The sequence of the AT domain is predicted to specifically autoacylate the ACP domain with methylmalonyl derived from methylmalonyl-CoA (*19*). The KS domain is proposed to then decarboxylate the methylmalonyl-*S*-EpoC acyl enzyme to generate the carbon nucleophile that reacts with the upstream donor. The resulting condensation product can be reduced in the presence of NADPH by the KR domain of EpoC and then dehydrated by the DH domain to produce the methylthiazolylmethylacrylyl-*S*-EpoC acyl enzyme (Figure 2). The experiments described here demonstrate in vitro reconstitution of both the PKS/NRPS interface (EpoA-ACP/B) and the NRPS/PKS interface (EpoB/C) in the assembly line for this antitumor natural product. Furthermore, EpoC has been shown to process several alkyl and aryl thiazoles presented by EpoB to the acrylate product, suggesting that biosynthesis of novel epothilone analogues with altered thiazole moieties via fermentation approaches may be feasible.

## EXPERIMENTAL PROCEDURES

*General.* Chemically competent cells of *Escherichia coli* strain Top10 and BL21(DE3) were purchased from Invitrogen. All restriction endonucleases, T4 DNA ligase, calf intestinal phosphatase, and T4 polynucleotide kinase were purchased from New England Biolabs. The DNA polymerase

Pfuturbo (Stratagene) was used for all PCR amplification. Oligonucleotides were ordered from Integrated DNA Technologies and used without further purification. The phosphopantetheinyl transferases Sfp and Svp were expressed and purified as previously described (*20*, *21*). [$^3$H]Acetyl-CoA, [$^{35}$S]cysteine, [$^{14}$C]methylmalonyl-CoA, and [$^{14}$C]-*N*-ethylmaleimide were purchased from New England Nuclear. All acyl-CoA derivatives were purchased from Sigma. DNA sequencing and MALDI mass spectrometry were performed at the Dana Farber Cancer Institute.

*Cloning of the EpoA-ACP Domain and the epoB and epoC Genes.* The EpoA-ACP domain and *epoB* were amplified and cloned as previously described (*18*). Briefly, residues 1286−1421 of *epoA* were amplified by PCR from a cosmid containing the *epoA/B/C* genes (provided by Kosan Biosciences) and inserted into pET28b vector using *Nde*I and *Eco*RI restriction sites to generate the *epoA-ACP* construct with an N-terminal 6× His tag. *epoB* was also amplified by PCR and cloned into pET28b using *Nde*I and *Not*I as restriction sites to generate the N-terminally His-tagged construct.

*epoC* was amplified in two parts from a cosmid containing the *epoA/B/C* genes. The N-terminal portion was amplified from the primers epoc1 (5′-CCCTTTCATATGGAAGAACAAGAGTCCTCCGC-3′) and epocmid1 (5′-CGCGGCTGGAAGTTTCGCCGCTGC-3′). The C-terminal portion was amplified from the primers epoc3 (5′-GGATACAAGCTTTCATGTAAGCGCCTTGAATT-3′) and epocmid2 (5′-GCGCCGGCGCGCTCTGCCGAGCTCTTC-3′). Restriction sites *Nde*I (epoc1) and *Hin*dIII (epoc3) are underlined. Prior to PCR amplification, the primers epocmid1 and epocmid2 were phosphorylated at the 5′ end with T4 polynucleotide kinase. PCR products were purified by agarose gel and were digested with *Nde*I and *Hin*dIII restriction enzymes. The restriction digest of the pET28b vector with *Nde*I and *Hin*dIII was followed by dephosphorylation with calf intestinal phosphatase. The two PCR fragments were simultaneously cloned into the digested, dephosphorylated pET28 by a blunt-end three-point ligation reaction. DNA sequencing confirmed that the appropriate insert had been ligated into the vector. The resulting plasmids were transformed into BL21(DE3) cells for protein overexpression.

*Overexpression and Purification of EpoA-ACP, EpoB, and EpoC.* EpoA-ACP and EpoB were expressed and purified as previously described (*17*). For overproduction of the EpoC protein, BL21(DE3) cells harboring the desired plasmid were grown in LB medium supplemented with 30 $\mu$g/mL kanamycin. Each liter of media was inoculated with 10 mL of overnight starter culture and cultured at 15 °C for 60 h. Protein expression was then induced by the addition of 100 $\mu$M isopropyl thiogalactoside, and cells were allowed to grow for an additional 10−12 h at 15 °C. Cells were harvested by centrifugation (10 min at 6000*g*), resuspended in lysis buffer (25 mM Tris, pH 8, 500 mM NaCl, 10% glycerol), and lysed by two passages through a French press at 10000 psi. Cell extracts were clarified by centrifugation (30 min at 15000*g*) and applied to nickel-NTA resin (1 mL of resin for 3 L of culture) (Qiagen). Cell lysate was allowed to bind in batch to the resin overnight at 4 °C and then decanted into a column. The resin was washed with 15 column volumes of lysis buffer and then eluted with a step gradient of lysis buffer containing increasing amounts of imidazole (5, 30, 60, 100, and 500 mM). EpoC eluted in the 100 and 500 mM imidazole fractions. Fractions containing the desired protein were dialyzed against 25 mM Tris, pH 7.5, 200 mM NaCl, and 10% glycerol and stored frozen at −80 °C.

*Characterization of EpoC. (A) Priming of EpoC with [$^3$H]-Phosphopantetheine.* To prime EpoC with the required phosphopantetheinyl moiety, EpoC (0.5 nmol) was incubated with [$^3$H]-HS-CoA (5 nmol, 74 Ci/mol) in 70 $\mu$L of buffer [50 mM Tris, pH 7, 5 mM $MgCl_2$, 5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 140 mM NaCl]. The priming reaction was initiated by addition of phosphopantetheinyl transferase Sfp or Svp (5 $\mu$M). After incubation for 1 h at 25 °C, the reaction was analyzed by SDS−PAGE (7%). The acrylamide gel was stained with Coomassie blue, destained, and soaked in Amplify (Amersham) for 30 min. The gel was dried and then exposed to film (Kodak Biomax) for 24 h at −78 °C.

*(B) Loading of EpoC with the [$^{14}$C]Methylmalonyl Moiety.* To prime any residual apo-EpoC, EpoC was primed as described above, using nonradiolabeled HS-CoA. Radiolabeled [$^{14}$C]methylmalonyl-CoA (1.7 nmol, 60 Ci/mol), [$^{14}$C]malonyl-CoA (1.7 nmol, 59 Ci/mol), and [$^3$H]acetyl-CoA (0.7 nmol, 2830 Ci/mol) were then added to three aliquots of holo-EpoC (0.5 nmol). After incubation at 25 °C for 15 min, the reactions were analyzed by SDS−PAGE (7%), stained, destained, soaked in Amplify, and dried. The gel was exposed to film for 24 h at −78 °C.

*(C) Incubation of EpoC with [$^{14}$C]-N-Ethylmaleimide.* Methylmalonyl-*S*-EpoC was prepared as described above, but nonradiolabeled methylmalonyl-CoA was used as a substrate for the EpoC AT domain. To the primed and loaded EpoC (1 nmol) was then added [$^{14}$C]-*N*-ethylmaleimide (0.5 nmol and 34.2 Ci/mol). The reaction was allowed to incubate at 25 °C for 15 min, at which point the reaction was subjected to analysis by SDS−PAGE. The gel was stained, amplified, and dried as described above and exposed to a phosphorimager (Fuji, TR2040) for 12 h.

*Transfer of [$^3$H]Acetyl from EpoA-ACP to EpoB and to EpoC: Analysis of Covalently Linked Products.* Apo-EpoA-ACP was primed with [$^3$H]acetyl-S-CoA. EpoA-ACP (0.45 nmol) in buffer (50 mM Tris, pH 7, 5 mM $MgCl_2$, 5 mM TCEP, 20 mM NaCl) was incubated with [$^3$H]acetyl-S-CoA (2 nmol, 28.3 Ci/mmol). Loading of the EpoA-ACP domain was initiated by the addition of Svp (5 $\mu$M) and allowed to incubate at ambient temperature for 30 min. EpoB (0.15 nmol) in buffer (50 mM Tris, pH 7, 5 mM $MgCl_2$, 5 mM TCEP, 20 mM NaCl, 5 mM ATP) was primed with HS-CoA (5 nmol) and then loaded with cysteine (5 nmol). The ACP domain of EpoC (0.5 nmol) was loaded with methylmalonyl as described above. The three proteins were then combined and allowed to incubate at room temperature for 60 min. Nonreducing protein sample buffer was then added to the solution, and the sample was loaded onto a 5% SDS−PAGE gel. The gel was stained with Coomasie, destained, soaked in Amplify for 30 min, and then dried. The gel was exposed to film or a phosphorimager for 48 h at −80 °C.

*Transfer of Acetyl from EpoA-ACP to EpoB and to EpoC: HPLC and Mass Spectrometric Analysis of Hydrolyzed Products. (A) Radio-HPLC Analysis of the EpoA-ACP/B/C Product.* Acetyl-*S*-EpoA-ACP, cysteinyl-*S*-EpoB, and methylmalonyl-*S*-EpoC were incubated together as described above. The radiolabel was incorporated either as the [$^3$H]-

acetyl-*S*-EpoA-ACP from [$^3$H]acetyl-CoA (28.3 Ci/mmol), as [$^{35}$S]cysteinyl-*S*-EpoB from [$^{35}$S]-L-cysteine (2.6 Ci/mmol), or as [$^{14}$C]methylmalonyl-*S*-EpoC from [$^{14}$C]methylmalonyl-CoA (60 Ci/mol). The reaction was quenched after 1 h by the addition of 10% trichloroacetic acid (TCA), and precipitated protein was pelleted by centrifugation and was then washed twice with 10% TCA. The protein pellet was then dissolved in KOH (0.1 M, 100 $\mu$L) and heated to 65 °C for 5 min. Trifluoroacetic acid (TFA) (50%, 5 $\mu$L) was then added, and the solution was centrifuged to remove precipitated proteins. A chemically synthesized standard was added to the radioactive enzymatic reaction and injected directly onto the HPLC (Vydac C18 reverse-phase column) (10−70% MeCN + 0.1% TFA in water + 0.1% TFA over 25 min). Dual on-line UV (254 nm) and radioisotope detectors (tuned for $^3$H, $^{14}$C, or $^{35}$S) were used to monitor the retention time of standards and radioactive enzymatic products, respectively.

*(B) Radio-HPLC Analysis of the NADPH Dependence of EpoA-ACP/B/C Product Formation.* To observe the nonreduced $\beta$-keto acid intermediate, reactions were performed in which NADPH was withheld from the reaction mixture. Acetyl-*S*-EpoA-ACP (1.4 nmol), [$^{35}$S]cysteinyl-*S*-EpoB (1 nmol), and methylmalonyl-*S*-EpoC (0.7 nmol) were prepared and combined together as described above. To one reaction mixture, NADPH (50 nmol) was added in excess, to a second reaction a substoichiometric amount of NADPH (0.5 nmol) was added, and to a third reaction NADPH was withheld. Reactions were incubated and hydrolyzed as described above and subjected to radio-HPLC analysis.

*(C) Mass Spectrometry of the EpoA-ACP/B/C Product.* To verify the molecular weight of the enzymatic product, acetyl-*S*-EpoA-ACP (1.4 nmol), cysteinyl-*S*-EpoB (1 nmol), and methylmalonyl-*S*-EpoC (1 nmol) were incubated together with NADPH (50 nmol) as described above. The proteins were precipitated and washed with 10% trichloroacetic acid (TCA), and the thioester-bound substrates and products were hydrolyzed with KOH as described above. The reaction was desalted (Millipore C18 zip-tip) and subjected to MALDI mass spectrometry. MALDI mass spectrometry analysis of an aliquot of the reaction mixture prior to base hydrolysis showed no product. The reaction mixture was also subjected to HPLC analysis with UV (254 nm) detection (Vydac C18 reverse-phase column) (0−100% MeCN + 0.1% TFA in water + 0.1% TFA over 25 min).

*Investigation of the Substrate Specificity of EpoC KS, KR, and DH Domains.* To investigate the substrate specificity of the EpoC KS, KR, and DH domains, various acyl groups were tested as substrates in the EpoA-ACP/B/C incubation. The EpoA-ACP domain was loaded with benzoyl-CoA, *n*-propionyl-CoA, and isobutyryl-CoA using the reaction conditions described above for loading with acetyl-CoA. The modified acyl-*S*-EpoA-ACP proteins were incubated with cysteinyl-*S*-EpoB and methylmalonyl-*S*-CoA using standard conditions. The radiolabel was positioned either on [$^{35}$S]-cysteinyl-*S*-EpoB or on [$^{14}$C]methylmalonyl-*S*-CoA. After being quenched with 10% TCA and hydrolyzed with KOH, the reaction products were analyzed by radio-HPLC.

*Chemical Synthesis of HPLC Standards.* All compounds and solvents were purchased from Aldrich and used without further purification, unless otherwise indicated. Anhydrous solvents were purchased from Aldrich and used as received.

*Methyl Thiazole Methyl Acrylic Acid. (A) Methyl Thiazole Aldehyde (**1**).* A solution of DIBAl-H in toluene (1.5 M, 1.6 mL, 2.46 mmol, 1.2 equiv) was added to a dry solution of methyl thiazole ethyl ester (Alfa Aesar) (305 mg, 2.05 mmol, 1 equiv) in toluene at −78 °C. After 1 h, excess aluminum hydride was quenched by the addition of methanol (1 mL) at −78 °C. After 5 min, the reaction mixture was allowed to warm to room temperature and partitioned between ether and an aqueous solution of Rochelle's salt (2 M, 8 mL), and the solution was stirred for 1.5 h to provide a clear bilayer. The aqueous layer was extracted once with ether (10 mL), and the combined organic layers were washed with brine, were dried with anhydrous sodium sulfate, and were concentrated. The resulting oil was found to be a mixture of product aldehyde and starting material ester (3:1) by $^1$H NMR analysis. The mixture was used without further purification. $^1$H NMR (for product **1**) (400 MHz, $CDCl_3$): 10.0 (s, 1H, C*H*O), 8.06 (s, 1H, SC(=C)*H*), 2.80 (s, 3H, C$H_3$CS).

*(B) Methyl Thiazole Methyl Acrylate Aldehyde (**2**).* To a solution of **1** (155 mg, 1.22 mmol) dissolved in anhydrous benzene (5 mL) was added $Ph_3P$=C(Me)CHO (464 mg, 1.46 mmol). The mixture was allowed to reflux at 80 °C under nitrogen for 1 h. The solution was then cooled to room temperature and diluted with ethyl acetate−hexanes (1:1, 20 mL). The resulting mixture was extracted with a saturated aqueous solution of sodium bicarbonate and brine, dried over anhydrous sodium sulfate, and concentrated. Purification of the residue by flash column chromatography on silica gel (eluant: ethyl acetate−hexanes, 1:4) provided aldehyde **2** as a solid (74%). $^1$H NMR (400 MHz, $CDCl_3$): 9.52 (s, 1H, CHC($CH_3$)C*H*O), 7.43 (s, 1H, SC(=C)*H*), 7.21 (s, 1H, C*H*C($CH_3$)CHO), 2.72 (d, 3H, CHC(C$H_3$)CHO), 2.16 (s, 3H, C$H_3$CS).

*(C) Methyl Thiazole Methyl Acrylate Carboxylic Acid (**3**).* Compound **2** (75 mg, 449 $\mu$mol) was dissolved in *t*-BuOH (10 mL) and 2-methyl-2-butene (2 M in THF, 2.5 mL, 4.98 mmol). $NaClO_2$ (536 mg, 4.76 mmol) and $NaH_2PO_4$ (483 mg, 3.50 mmol) in water (5 mL) were added. After 2 h at 25 °C, the yellow mixture was diluted with water (10 mL), a saturated solution of sodium chloride (10 mL), and ethyl acetate−hexanes (1:1, 20 mL). The organic layer was collected, and the aqueous layer was extracted with ethyl acetate−hexanes (1:1, 20 mL). The combined organic layers were dried over anhydrous sodium sulfate and were concentrated to provide the carboxcylic acid as a solid (99%). Analysis by $^1$H NMR revealed that the product was greater than 95% pure and was used without further purification. $^1$H NMR (400 MHz, $CDCl_3$): 7.32 (s, 1H, SC(=C)*H*), 7.74 (d, 1H, C*H*C($CH_3$)CHO), 2.75 (s, 3H, C$H_3$CS), 2.34 (d, 3H, CHC(C$H_3$)COOH). $^{13}$C NMR (200 MHz, $CDCl_3$): 174.39, 166.00, 151.96, 132.65, 128.17, 122.48, 19.56, 14.28. [M + H]$^+$: obsd 184.34, expected 184.23.

*Ethyl Thiazole Methyl Acrylic Acid. (A) Ethyl Thiazole Ethyl Ester (**4**).* To a mixture of thiopropionamide (TCI America) (1.0 g, 11.2 mmol) and $KHCO_3$ (8.97 g, 89.6 mmol) was added anhydrous ethylene glycol dimethyl ether (DME) (10 mL) and ethyl bromopyruvate (4.25 mL, 33.8 mmol). A white precipitate formed, and the reaction was allowed to stir under nitrogen at 25 °C for 2 h. The mixture was then cooled to 0 °C, and a solution of trifluoroacetic anhydride (6.43 mL, 45.5 mmol) and pyridine (7.75 mL, 95.8 mmol) in DME (3 mL) was added. The orange reaction was

then allowed to warm to ambient temperature and was stirred under nitrogen for an additional hour. The reaction was then concentrated, and the residue was dissolved in $CHCl_3$ (10 mL), extracted with brine, dried with anhydrous sodium sulfate, and concentrated. The residue was purified by flash column chromatography on silica gel (ethyl acetate−hexanes, 3:7); yield 79%. $^1$H NMR (200 MHz, $CDCl_3$): 8.01 (s, 1H, SC(=C)*H*), 4.42 (q, 2H, COC*H*$_2$CH$_3$), 3.10 (t, 3H, CH$_3$C*H*$_2$-CS), 1.42 (m, 3H, C*H*$_3$CH$_2$CS), 1.42 (m, 3H, COCH$_2$C*H*$_3$).

*(B) Ethyl Thiazole Aldehyde (**5**).* $^1$H NMR (200 MHz, $CDCl_3$): 9.73 (s, 1H, C*H*O), 7.93 (s, 1H, SC(=C)*H*), 2.85 (q, 2H, CH$_3$C*H*$_2$CS), 1.17 (t, 3H, C*H*$_3$CH$_2$CS).

*(C) Ethyl Thiazole Methyl Acrylate Aldehyde (**6**).* $^1$H NMR (200 MHz, $CDCl_3$): 9.46 (s, 1H, CHC(CH$_3$)C*H*O), 7.16 (d, 1H, C*H*C(CH$_3$)CHO), 7.40 (s, 1H, SC(=C)*H*), 2.97 (q, 2H, CH$_3$C*H*$_2$CS), 2.12 (d, 3H, CHC(C*H*$_3$)CHO), 1.33 (t, 3H, C*H*$_3$-CH$_2$CS).

*(D) Ethyl Thiazole Methyl Acrylate Carboxylic Acid (**7**).* $^1$H NMR (200 MHz, $CDCl_3$): 7.74 (d, 1H, C*H*C(CH$_3$)-COOH), 7.34 (s, 1H, SC(=C)*H*), 3.06 (q, 2H, CH$_3$C*H*$_2$CS), 2.36 (d, 3H, CHC(C*H*$_3$)CHO), 1.42 (t, 3H, C*H*$_3$CH$_2$CS). $[M + H]^+$: obsd 198.06, expected 198.05.

*Isopropyl Thiazole Methyl Acrylic Acid. (A) Isobutyryl Thioamide (**8**).* Lawesson's reagent (2.31 g, 5.72 mmol) was added to isobutyryl amide (1.00 g, 11.48 mmol) dissolved in DME (25 mL). After 6 h of stirring under nitrogen atmosphere at 25 °C, the reaction was concentrated, and the residue was dissolved in $CH_2Cl_2$ (30 mL) and washed with a 1% solution of sodium hydroxide. The organic layer was washed with brine, concentrated, and purified by flash chromatography on silica gel (ethyl acetate−hexanes, 2:8); yield 54%. $^1$H NMR (200 MHz, $CDCl_3$): 8.18, 7.36 (s, 1H, (CH$_3$)$_2$CHCSN*H*$_2$), 2.82 (septet, 1H, (CH$_3$)$_2$C*H*CSNH$_2$), 1.20 (s, 3H, (C*H*$_3$)$_2$CHCSNH$_2$), 1.16 (s, 3H, (C*H*$_3$)$_2$CHCSNH$_2$).

*(B) Isopropyl Thiazole Ethyl Ester (**9**).* $^1$H NMR (200 MHz, $CDCl_3$): 7.86 (s, 1H, SC(=C)*H*), 4.18 (q, 2H, COC*H*$_2$-CH$_3$), 3.21 (septet, 3H, (CH$_3$)$_2$C*H*CS), 1.22 (m, 3H, (C*H*$_3$)$_2$-CHCS), 1.19 (m, 3H, (C*H*$_3$)$_2$CHCS), 1.18 (m, 3H, COCH$_2$-C*H*$_3$).

*(C) Isopropyl Thiazole Aldehyde (**10**).* $^1$H NMR (200 MHz, $CDCl_3$): 9.86 (s, 1H, C*H*O), 7.99 (s, 1H, SC(=C)*H*), 3.25 (septet, 1H, (CH$_3$)$_2$C*H*CS), 1.33 (s, 3H, (C*H*$_3$)$_2$CHCS), 1.29 (s, 3H, (C*H*$_3$)$_2$CHCS).

*(D) Isopropyl Thiazole Methyl Acrylate Aldehyde (**11**).* $^1$H NMR (200 MHz, $CDCl_3$): 9.48 (s, 1H, CHC(CH$_3$)C*H*O), 7.18 (d, 1H, C*H*C(CH$_3$)CHO), 7.42 (s, 1H, SC(=C)*H*), 3.26 (septet, 3H, (CH$_3$)$_2$C*H*CS), 2.15 (d, 3H, CHC(C*H*$_3$)CHO), 1.38 (s, 3H, (C*H*$_3$)$_2$CHCS), 1.34 (s, 3H, (C*H*$_3$)$_2$CHCS).

*(E) Isopropyl Thiazole Methyl Acrylate Carboxylic Acid (**12**).* $^1$H NMR (200 MHz, $CDCl_3$): 7.74 (d, 1H, C*H*C(CH$_3$)-COOH), 7.35 (s, 1H, SC(=C)*H*), 3.34 (septet, 1H, (CH$_3$)$_2$-C*H*CS), 2.37 (d, 3H, CHC(C*H*$_3$)COOH), 1.44 (s, 3H, (C*H*$_3$)$_2$-CHCS), 1.42 (s, 3H, (C*H*$_3$)$_2$CHCS). $[M + H]^+$: obsd 2.12.04, expected 212.07.

*Phenyl Thiazole Methyl Acrylic Acid. (A) Phenyl Thiazole Ethyl Ester (**13**).* $^1$H NMR (200 MHz, $CDCl_3$): 8.14 (s, 1H, SC(=C)H), 8.00 (m, 3H, C$_6$H$_5$CS), 7.44 (m, 2H, C$_6$H$_5$CS), 4.43 (q, 2H, COCH$_2$CH$_3$), 1.42 (t, 3H, COCH$_2$CH$_3$).

*(B) Phenyl Thiazole Aldehyde (**14**).* $^1$H NMR (200 MHz, $CDCl_3$): 10.0 (s, 1H, CHO), 8.08 (s, 1H, SC(=C)H), 7.91 (m, 2H, C$_6$H$_5$CS), 7.37 (m, 3H, C$_6$H$_5$CS).

*(C) Phenyl Thiazole Methyl Acrylate Aldehyde (**15**).* $^1$H NMR (200 MHz, $CDCl_3$): 9.43 (s, 1H, CHC(CH$_3$)C*H*O), 8.00 (s, 1H, SC(=C)*H*), 7.81 (m, 2H, C$_6$*H*$_5$CS), 7.42 (d, 1H, C*H*C(CH$_3$)CHO), 7.30 (m, 3H, C$_6$*H*$_5$CS), 2.18 (d, 3H, CHC-(C*H*$_3$)CHO).

*(D) Phenyl Thiazole Methyl Acrylate Carboxylic Acid (**16**).* $^1$H NMR (200 MHz, $CDCl_3$): 7.78 (s, 1H, SC(=C)*H*), 8.00 (m, 2H, C$_6$*H*$_5$CS), 7.42 (d, 1H, C*H*C(CH$_3$)COOH), 7.45 (m, 3H, C$_6$*H*$_5$CS), 2.49 (d, 3H, CHC(C*H*$_3$)COOH). $[M + H]^+$: obsd 246.03, expected 246.05.

*Methyl Thiazole β-Keto Ester (17).* Methyl thiazole ethyl ester (800 mg, 4.67 mmol) was suspended in wet methanol (1 mL). Sodium methoxide (252 mg, 4.67 mmol) and methyl propionate (540 μL, 5.60 mmol) were added. The reaction was allowed to reflux under nitrogen for 3 h. Three additional aliquots of methyl propionate (540 μL, 5.60 mmol) were added at 1 h intervals. The reaction was allowed to cool to ambient temperature and was then cooled to 0 °C. The reaction was diluted with ether (4 mL) and added all at once to an ice-cold mixture of acetic acid−water (1:1, 4 mL). The organic layer was separated, and the aqueous layer was extracted with ether. The combined organic layers were washed with water, neutralized by washing with saturated sodium bicarbonate, dried over anhydrous sodium sulfate, and concentrated. The reaction mixture was purified over silica gel (ethyl acetate−hexanes, 2:8); yield 25%. $^1$H NMR (200 MHz, $D_2O$): 8.08 (s, 1H, SC(=C)*H*), 3.75 (s, 3H, C(O)-C*H*$_3$COOH), 2.54 (s, 3H, C*H*$_3$CS). $[M + H]^+$: obsd 237.48, expected 237.99.

*Methyl Thiazole-S-NAC.* To a solution of methyl thiazole carboxylic acid (10 mg, 70 μmol) in dimethylformamide (0.5 mL) was added benzotriazol-1-yloxytris(pyrrolidino)phosphonium hexafluorophosphate (110 mg, 210 μmol, 3 equiv) and *N*-acetylcysteamine (25 mg, 210 μmol, 3 equiv). After being shaken at ambient temperature for 1.5 h, the mixture was purified by preparative HPLC (0−100% MeCN over 25 min in water with 0.1% TFA). $[M + H]^+$: obsd 244.98, expected 245.03.

*Phenyl Thiazole-S-NAC.* $[M + H]^+$: obsd 307.70, expected 307.41.

## RESULTS

*Heterologous Expression of EpoA-ACP/B/C in E. coli.* A strategy has been developed to overcome the difficulties associated with expression of the starter module, EpoA. Since efforts to produce EpoA in active form have not yet succeeded, the ACP domain of EpoA has been expressed in apo form and then reacted with acetyl-CoA and Sfp to generate the acetyl-*S*-EpoA-ACP protein (*18*). This truncated version of EpoA serves as a functional substitute for EpoA. Cloning, expression, and purification of the EpoA-ACP domain and EpoB were performed as previously described (*18*). Expression of the 195 kDa (1833 amino acid) EpoC protein in *E. coli* under a variety of conditions revealed that low-temperature expression was essential for production of soluble, active protein. The cell culture at 15 °C for 60 h, followed by 12 h induction with IPTG, produced approximately 2.5 mg of EpoC/L of cell culture. Purification was facilitated by the incorporation of a histidine affinity tag at the N-terminus of the construct, enabling nickel resin affinity chromatography. The purity of the protein is illustrated in Figure 3A.

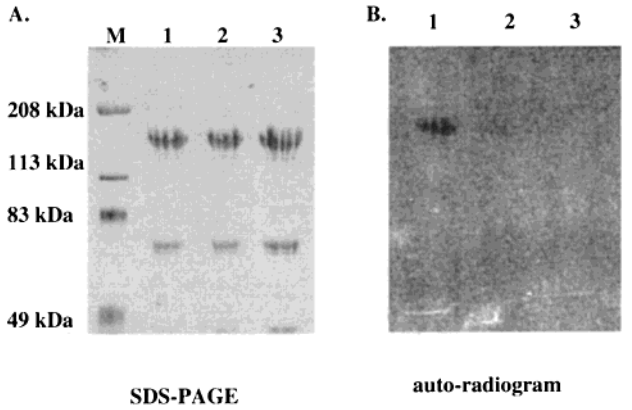



FIGURE 3: (A) SDS−PAGE and (B) autoradiogram of purified EpoC loaded with (1) [$^{14}$C]methylmalonyl-CoA, (2) [$^{14}$C]malonyl-CoA, and (3) [$^{3}$H]acetyl-CoA.

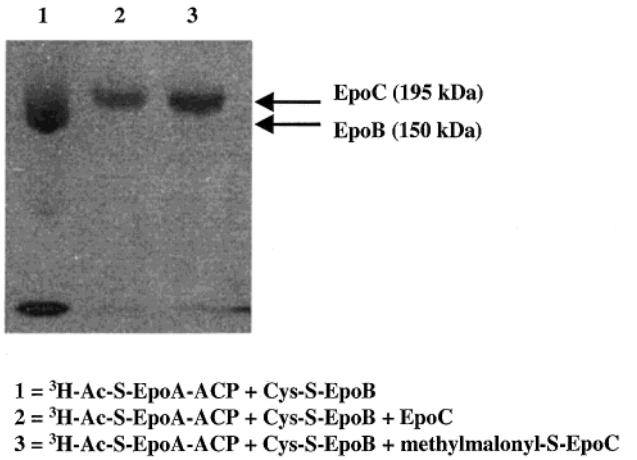



FIGURE 4: Autoradiogram of the reaction of acetyl-*S*-EpoA-ACP, cysteinyl-*S*-EpoB, and methylmalonyl-*S*-EpoC.

*Characterization of the AT and ACP Domains of EpoC.* The ACP domain of EpoC was primed with the phosphopantetheinyl transferase Sfp (*20*) or Svp (*21*) and HS-CoA to generate the holo, phosphopantetheinylated form of the protein. Analysis of EpoC primed with [$^{3}$H]-HS-CoA by SDS−PAGE and autoradiography, however, revealed no modification of EpoC with the CoA-derived pantetheine moiety, suggesting that EpoC is fully primed in vivo (data not shown). Due to the lengthy expression conditions, EpoC was most likely primed in vivo by endogenous *E. coli* phosphopantetheinyl transferases.

Holo-EpoC was incubated with radiolabeled acetyl-CoA, malonyl-CoA, and methylmalonyl-CoA to assay the activity and substrate specificity of the AT domain (Figure 3B). As expected, autoradiography analysis of the reaction reveals that the AT domain preferentially loads methylmalonyl onto the ACP domain of EpoC.

*KS Domain of EpoC: Transfer of [$^{3}$H]Methylthiazole from EpoB to EpoC.* In a preliminary analysis of KS domain activity, methylmalonyl-*S*-EpoC was labeled with [$^{14}$C]-*N*-ethylmaleimide as evidenced by analysis by SDS−PAGE and autoradiography (data not shown), a result consistent with the presence of a functional active site cysteine in the KS domain (*22*). To more specifically assess the activity of the KS domain, EpoC was assayed in the presence of the upstream partner proteins in the epothilone assembly line, EpoA-ACP and EpoB. A mixture of [$^{3}$H]acetyl-*S*-EpoA-ACP, cysteinyl-*S*-EpoB, and methylmalonyl-*S*-EpoC was allowed to incubate and then subjected to SDS−PAGE. Analysis of the SDS−PAGE by autoradiography revealed that the radioactive band shifted to a slightly higher molecular weight, suggesting that the tritiated acetyl group was transferred from the 150 kDa [$^{3}$H]methylthiazolylcarboxyl-*S*-EpoB to the 195 kDa EpoC protein (Figure 4). The radiolabeled thiazole produced by EpoA-ACP/B is presumably transferred to the active site cysteine of the EpoC KS domain. After KS-mediated decarboxylation of methylmalonyl-*S*-EpoC, condensation of the resulting propionate nucleophile with the methylthiazolylcarboxyl-*S*-EpoB generates methylthiazole-$\beta$-keto-$\alpha$-acyl-*S*-EpoC (Figure 2). Transfer of the tritium label from [$^{3}$H]acetyl-*S*-EpoA-ACP to EpoC was also observed in the absence of the methylmalonyl substrate (Figure 4, lane 2). The EpoA-ACP/B methyl thiazolyl carboxylic acid product may be transferred to the active site cysteine of the KS domain, accumulating as an acyl enzyme intermediate in the absence of the downstream acceptor substrate, methylmalonyl-CoA. However, the timing of decarboxylation of the acyl substrate relative to loading of the upstream acyl chain onto the KS domain active site cysteine has not yet been established (*23*). Alternatively, the methyl thiazole carboxylic acid may be transferred to the holo-ACP domain in a nonspecific thiol exchange reaction. Since EpoC is purified in largely holo form, it is difficult to distinguish between these two possibilities using this assay.

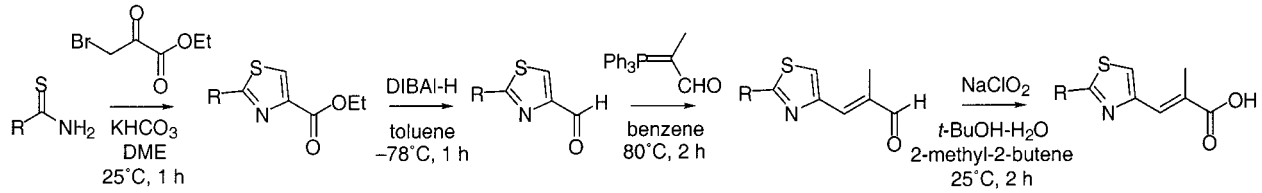
No transfer of radioactivity from [$^{3}$H]acetyl-*S*-EpoA-ACP to methylmalonyl-*S*-EpoC was observed in the absence of cysteinyl-*S*-EpoB; furthermore, no product formation was observed from cysteinyl-*S*-EpoB and methylmalonyl-*S*-EpoC in the absence of an acyl-*S*-EpoA-ACP (data not shown). These observations suggest a level of specificity and directionality present in the assembly line.

*Characterization of EpoC: Methyl Thiazole Methyl Acrylic Acid Formation.* The activity of the EpoC domains can be further assessed by analyzing the chemical structure of the product that is enzymatically generated from EpoA-ACP/B/C. Although both substrates and products are covalently attached to PKS and NRPS proteins through a thioester bond, the thioester linkage can be easily hydrolyzed by base and the free products subjected to standard methods of analysis such as HPLC and mass spectrometry.

To characterize the EpoA-ACP/B/C reconstituted product, acetyl-*S*-EpoA-ACP, cysteinyl-*S*-EpoB, methylmalonyl-*S*-EpoC, and NADPH were incubated together. A variety of EpoA-ACP/B/C ratios were used, with each reaction giving similar results. The enzyme-bound products were hydrolyzed from the protein with aqueous potassium hydroxide and subjected to radio-HPLC analysis. Individual incubations of EpoA-ACP/B/C with the radiolabel positioned at [$^{3}$H]acetyl-*S*-EpoA-ACP, [$^{35}$S]cysteinyl-*S*-EpoB, and [$^{14}$C]methylmalonyl-*S*-EpoC were performed. In each of the three cases, the enzymatically generated product coeluted with a chemically synthesized (Scheme 1) authentic standard (Figure 5) (*24*, *25*). In the absence of methylmalonyl-CoA, only methyl thiazole carboxylic acid formation was observed (data not shown).

To further characterize the enzymatic product, the reaction was performed using nonradiolabeled substrates and was analyzed by MALDI mass spectrometry after hydrolysis of the products with potassium hydroxide. The expected mass

Scheme 1 [a]




[a] R = methyl, ethyl, isopropyl, or phenyl.

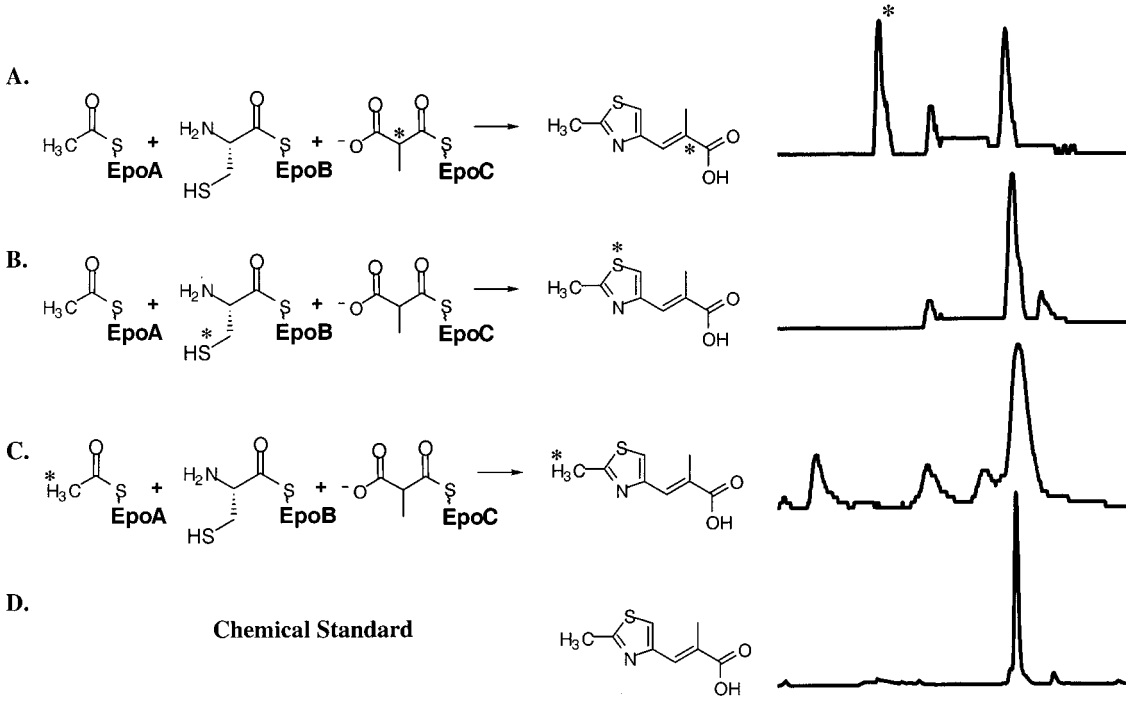



FIGURE 5: Radio-HPLC analysis of the EpoA/B/C enzyme product using radiolabel in (A) [$^{14}$C]methylmalonyl-*S*-EpoC (the * peak is unreacted methylmalonate), (B) [$^{35}$S]cysteinyl-*S*-EpoB, and (C) [$^{3}$H]acetyl-*S*-EpoA-ACP. (D) UV trace (254 nm) of the chemically synthesized standard.

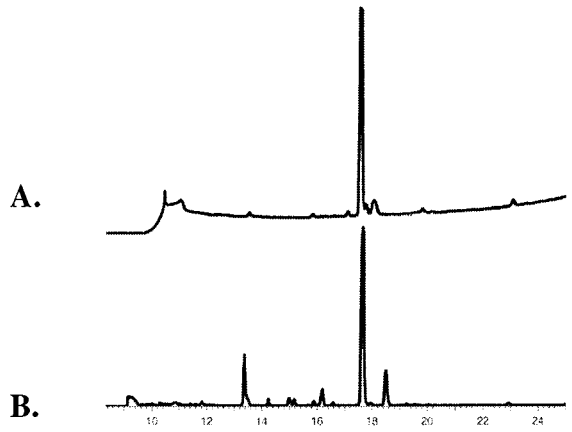
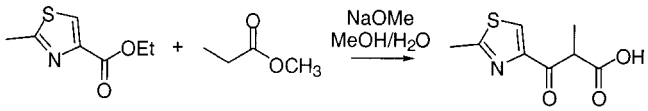



FIGURE 6: UV HPLC (254 nm) of the (A) methyl thiazole methylacrylic acid standard and (B) enzymatic product formation from acetyl-*S*-EpoA-ACP, cysteinyl-*S*-EpoB, and methylmalonyl-*S*-EpoC.

of the potassium salt of the methyl thiazole methylacrylic acid was clearly observed (predicted [M$^-$H] 221.33, observed 221.46). HPLC analysis of this reaction mixture (UV detection at 254 nm) revealed relatively clean conversion to the desired product (Figure 6).

*KR and DH Domains of EpoC: Formation of Methylthiazolyl-β-keto-α-methyl Acid.* After formation of the methylthiazolyl-β-keto-α-methyl-acyl-*S*-EpoC intermediate, the KR domain is expected to reduce this intermediate using NADPH, followed by dehydration to the methyl thiazole methylacrylyl-*S*-EpoC by the DH domain of EpoC (Figure 2). Removal of NADPH from the reaction mixture and analysis of the resulting base-hydrolyzed product by radio-HPLC demonstrated the formation of a new reaction product that comigrated with the chemically synthesized (Scheme 2) methyl thiazole methyl β-keto acid authentic standard (Figure 7) (*26*). Addition of a substoichiometric amount of NADPH resulted in a mixture of β-keto acid and dehydrated product. Although the reduced hydroxyl intermediate has not been explicitly observed, it is predicted that the DH domain dehydrates the β-OH-acyl-*S*-EpoC intermediate to the methyl thiazole methylacrylyl-*S*-EpoC product (Figure 2).

Scheme 2




*Substrate Specificity of EpoC: Acceptance of Various Alkyl/Aryl Thiazoles.* Previous work has demonstrated that the cyclization (Cy) domain and oxidase (Ox) domains of cysteine-*S*-EpoB recognize and process a variety of acyl-*S*-EpoA-ACP derivatives to the corresponding alkyl or aryl thiazole carboxylic acids (*18*). To characterize the specificity

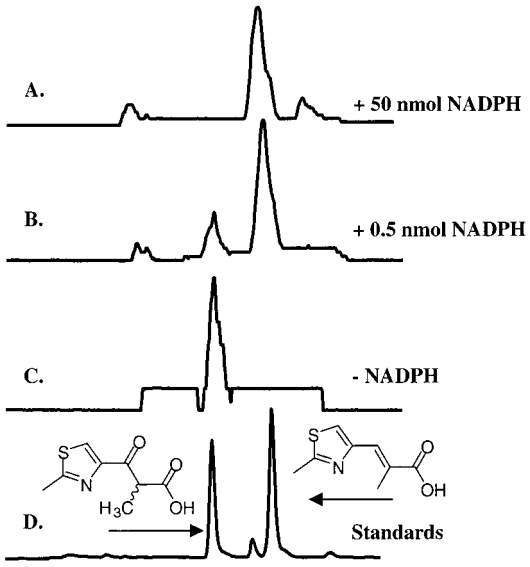



FIGURE 7: Accumulation of [$^{35}$S]methyl thiazole $\beta$-keto-$\alpha$-methyl acyl product in the absence of NADPH. (A) Radio-HPLC trace of the enzymatic reaction with 50 nmol of NADPH. (B) Reaction with 0.5 nmol of NADPH. (C) Reaction with no NADPH. (D) UV HPLC trace of chemically synthesized standards.

of the EpoC KS domain and KR and DH tailoring domains, several substituted thiazoles were enzymatically synthesized using various aryl- and alkyl-*S*-EpoA-ACP proteins and cysteine-*S*-EpoB. By reaction of EpoA-ACP with Svp and *n*-propionyl-CoA, isobutyryl-CoA, or benzoyl-CoA, ethyl-*S*-EpoA-ACP, isopropyl-*S*-EpoA-ACP, or phenyl-*S*-EpoA-ACP was generated. These modified acyl-*S*-ACP proteins were incubated with [$^{35}$S]cysteinyl-*S*-EpoB and methylmalonyl-*S*-EpoC. The enzyme-bound products were base hydrolyzed and then subjected to radio-HPLC analysis. Co-injection with chemically synthesized authentic standards (Scheme 1) (*24*, *25*, *27*) revealed that formation of the alkyl/aryl thiazole methyl acrylic acid product was observed in all three cases (Figure 8). *Reaction of [$^{14}$C]Methylmalonyl-S-EpoC with Methylthiazolyl-S-NAC.* PKS proteins can often utilize soluble acyl-*S*-*N*-acetylcysteamine (*S*-NAC) substrates in addition to substrates presented on a carrier protein via a phosphopantetheinyl linkage (*28*−*30*). [$^{14}$C]Methylmalonyl-*S*-EpoC was incubated with methylthiazolyl-*S*-NAC, and the resulting product was hydrolyzed by base, analyzed by radio-HPLC, and found to coelute with an authentic standard of the expected product. The reaction was also performed using nonradiolabeled methylmalonyl substrate, and subsequent analysis by MALDI mass spectrometry revealed that the expected molecular weight of the product was observed.

Furthermore, a phenylthiazolyl-*S*-NAC derivative was prepared and incubated with [$^{14}$C]methylmalonyl-*S*-EpoC. Analysis by radio-HPLC revealed that the resulting product coeluted with an authentic sample of the phenyl thiazole methyl acrylic acid product, indicating that alternate substrates do not necessarily need to be presented by a carrier protein to be processed by EpoC.

## DISCUSSION

The four major epothilone metabolites, epothilones A−D (Figure 1A), are produced via a cascade of elongating acyl *S*-enzyme intermediates on the proteins EpoA−F, which

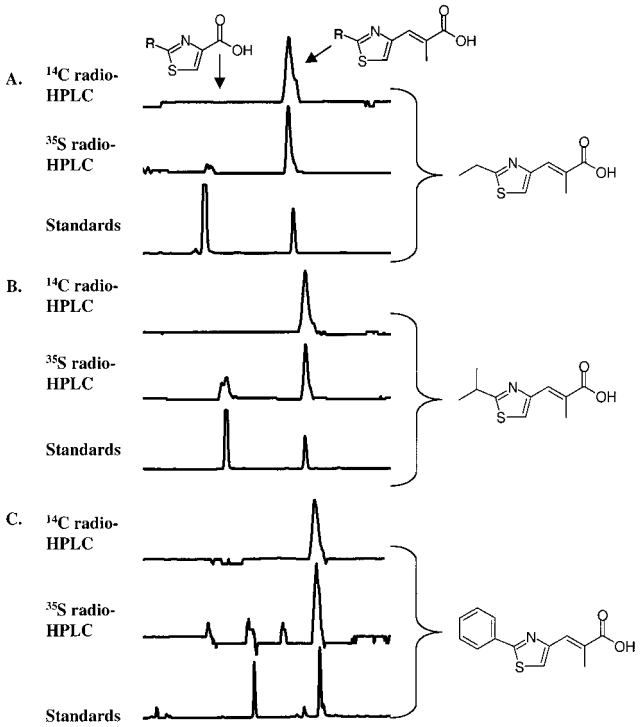



FIGURE 8: Use of various acyl-*S*-EpoA derivatives to obtain novel alkyl thiazole methyl acrylic acid products. Product formation is observed by radio-HPLC analysis using a radiolabel in both [$^{35}$S]-cysteine and [$^{14}$C]methylmalonyl. Enzymatic products are shown to coelute with chemically synthesized standards. (A) Radio-HPLC traces and UV HPLC traces of ethyl thiazole methyl acrylic acid products and standard. (B) Radio-HPLC traces and UV HPLC traces of isopropyl thiazole methyl acrylic acid products and standard. (C) Radio-HPLC traces and UV HPLC traces of phenyl thiazole methyl acrylic acid products and standard.

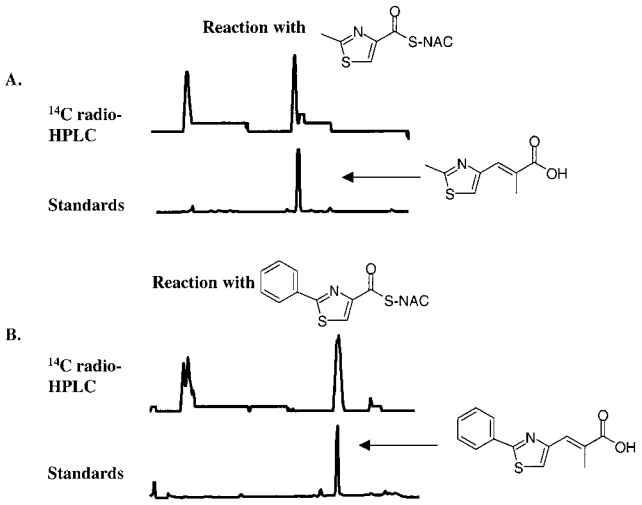



FIGURE 9: Use of (A) methylthiazolyl-*S*-NAC to obtain methyl thiazole methyl acrylic acid and (B) phenylthiazolyl-*S*-NAC to obtain the phenyl thiazole methyl acrylic acid product. Product formation is observed by radio-HPLC analysis using a radiolabel in [$^{14}$C]methylmalonyl. Enzymatic products are shown to coelute with chemically synthesized standards.

constitute a mixed PKS/NRPS assembly line. There are 10 modules distributed over the six proteins, with each module containing one carrier protein domain; nine ACP domains are found in the PKS modules and one PCP is found in the single NRPS module (Figure 1B). Each ACP/PCP domain is posttranslationally primed at a conserved serine residue

with a phosphopantetheine arm to introduce the terminal SH group that serves as the tether for the elongating acyl chain. The EpoD, EpoE, and EpoF subunits provide a conventional set of seven PKS modules responsible for construction of the 16-membered ring macrolactone scaffold, but the first three subunits, EpoA, EpoB, and EpoC, constitute a PKS/NRPS/PKS set that represents a switch from PKS to NRPS chain elongation logic and back. The net chemical outcome is the introduction of a cysteinyl residue between a starter acetyl moiety and a downstream propionyl group, in which the cysteine becomes cyclized and embedded within the acyl product as a heterocyclic thiazole ring.

In prior work we have reconstituted the first interface (PKS to NRPS) between the ACP domain of EpoA and the Cy domain of EpoB (Figure 1C) (*18*). In this study we reconstitute the second interface (NRPS to PKS) between the PCP domain of EpoB and the KS domain of the EpoC proteins. Successful reconstitution of this interface has depended on heterologous expression of soluble EpoC, a five domain 195 kDa PKS module, in *E. coli* and its subsequent posttranslation priming on the apo ACP domain by a phosphopantetheinyl transferase. After purification, we explicitly assayed three of the four remaining catalytic domains of EpoC. The activity of the AT domain was validated by observing transfer of [$^{14}$C]methylmalonyl to the HS-phosphopantetheinyl-ACP domain *in cis* from methylmalonyl-CoA. The AT domain of EpoC was observed to be specific for methylmalonyl-CoA, in agreement with both amino acid sequence predictions of the AT specificity (*19*) and the structure of the epothilones. The KS domain activity was assessed by monitoring acyl transfer of the methylthiazolyl moiety from EpoB to EpoC. The activity of the KR domain was assessed by observing the NADPH-dependent reduction of the $\beta$-ketoacyl-*S*-EpoC intermediate. Although the DH domain of EpoC could not be directly assayed in the reaction format described here, the dehydratase activity was indirectly confirmed by isolation of the base-hydrolyzed methyl thiazole methyl acrylic acid product. In the presence of NADPH, the KR-mediated reduction of the $\beta$-keto to the $\beta$-OH-acyl *S*-enzyme sets up the intermediate for the DH domain to generate the $\alpha,\beta$-enoyl-*S*-EpoC (Figure 2), which then serves as the substrate for the next subunit, EpoD, in the epothilone assembly line.

The EpoC subunit, the third module in the assembly line, offers several challenges for enzymatic assay and characterization that apply to all elongation domains in PKS and NRPS assembly lines. One point is that no catalytic turnover is observed in the absence of the seven downstream modules EpoD−F; only stoichiometric formation of the covalent acyl *S*-enzyme intermediates is observed. In other studies of PKS elongation domains, e.g., in the DEBS assembly line, constructs appending the chain-terminating thioesterase (TE) domain to upstream modules have created modules where the acyl *S*-enzymes are discharged hydrolytically to generate a multiple turnover system (*31, 32*). Although appending a TE domain to EpoC may be a viable approach for future studies, here we have used sufficient quantities of purified EpoC protein to detect stoichiometric acyl *S*-enzymes. After base-catalyzed hydrolysis of the acyl thioester linkages, the free acids were released for characterization by HPLC, mass spectrometry, and comparison to authentic standards. A second challenge in this system is to provide the authentic upstream acyl donor for the EpoC catalytic domains, methylthiazolylcarboxyl-*S*-EpoB. In other systems, acyl-*S*-NACs have been used as soluble surrogates of the donor acyl *S*-enzymes (*28*−*30*). Here we have generated the authentic acyl-*S*-EpoB donors, as reported in our prior efforts on EpoA-ACP/B interface reconstitution, although we have also shown that EpoC accepts both the methyl- and phenylthiazolyl-*S*-NAC derivatives.

Analysis by HPLC reveals that formation of the methyl thiazole methyl acrylic acid is observed, as validated by comigration with an authentic standard and by observation of the expected mass by MALDI mass spectrometry. In the absence of NADPH, product analysis indicates that the KR domain fails to function, effectively stopping the reaction at the $\beta$-keto acid intermediate (Figure 2). In the absence of cysteinyl-*S*-EpoB, no radioactivity was transferred from [$^{3}$H]-*S*-EpoA-ACP to methylmalonyl-*S*-EpoC, suggesting that these two proteins cannot productively interact without the presence of EpoB. In the absence of acetyl-*S*-EpoA-ACP, no radioactive product was formed from [$^{35}$S]cysteinyl-*S*-EpoB and methylmalonyl-*S*-EpoC, suggesting that the cysteinyl-*S*-EpoB substrate is not recognized by EpoC.

We have detected methylmalonyl-*S*-EpoC from AT domain action, the methylthiazolyl-$\beta$-keto-$\alpha$-methyl-acyl-*S*-EpoC from KS domain C−C bond formation, and, after addition of NADPH, the methylthiazolyl-$\alpha$-methyl-$\alpha,\beta$-enoyl-*S*-EpoC from KS/KR/DH tandem action. The acyl-*S*-EpoC enzyme forms are stable enough to withstand analysis by SDS−PAGE and subsequent detection by autoradiography (Figure 4). Furthermore, radioactive labels from all three substrates, the [$^{3}$H]acetyl group from the EpoA-ACP, the [$^{35}$S]cysteinyl group from EpoB, and the [$^{14}$C]methylmalonyl from EpoC autoacylation, are each incorporated into the methyl thiazole methyl acrylic acid released from EpoC by base hydrolysis (Figure 5). A preliminary kinetic analysis indicated that product formation in the epothilone system is too fast to measure by conventional assay and will require rapid quench approaches to deconvolute the rates of the single turnover acylations in the EpoA-ACP, EpoB, and EpoC proteins. However, the results presented here validate reconstruction of the EpoB/EpoC interface. Coupled to our prior efforts of reconstitution of the A/B interface, both switch points, PKS to NRPS and NRPS to PKS, are operational with the purified domains in vitro and can be utilized to evaluate specificity and mechanism of the first 13 domains of the epothilone synthetase assembly line.

In particular, studies that probed the EpoA-ACP/EpoB interface interrogated the specificity of the Cy domain of EpoB to accept various acyl donors tethered in a thioester linkage on the EpoA-ACP domain and uncovered the tolerance for acyl substituents in the amide bond-forming step and the heterocyclization and oxidation step. Analogously, the EpoB/C interface reconstitution allows evaluation of the KS domain of the EpoC subunit as a C−C bond-forming, Claisen condensation catalyst. In the initial studies here we note that alternate acyl groups that are accepted by the Cy domain of EpoB can also be tolerated by the KS domain of EpoC to yield variant alkyl and aryl substituents in the thiazole moiety of the acrylate chains lodged on EpoC. This result suggests that although the AT domain of EpoC is specific for methylmalonyl-CoA, a variety of substrates can be recognized and processed by the KS, KR, and DH

domains. Although these experiments establish that acrylate product formation is observed in all cases, further studies are required to quantitatively assess the efficiency of these alternate substrates and also to monitor partitioning among intermediates. These results suggest a strategy for further evaluation of the tolerance of EpoA, EpoB, and EpoC to substrates other than acetyl, cysteinyl, and methylmalonyl groups, which may require mutagenesis to alter monomer selection specificity.

## CONCLUSION

Epothilone, a hybrid PK/NRP natural product, is proving to be a promising anticancer agent with potentially superior pharmacological properties to Taxol in multi-drug-resistant cell lines (*16*). The studies of the early steps of epothilone biosynthesis represent an opportunity to investigate the construction of a hybrid PK/NRP product. The EpoA/B/C system, the first three enzymes involved in epothilone production, constitutes a switch from polyketide to peptide biosynthesis and then back to polyketide biosynthesis chemistry. Furthermore, we have established that the Cy and Ox domains of EpoB and the KS, KR, and DH domains of EpoC can process a variety of acyl groups presented by EpoA-ACP, and thus the possibility of fermentation of epothilone analogues with altered thiazole moieties appears to be a promising one.

## ACKNOWLEDGMENT

We thank Kosan Biosciences for providing the *epoA*, *epoB*, and *epoC* genes. We also thank Prof. Ben Shen for providing the *svp* construct and Dr. Steve Bruner for preparing Svp.

## REFERENCES

1. Keating, T. A., and Walsh, C. T. (1999) *Curr. Opin. Chem. Biol. 3*, 598−606.
2. Khosla, C. (1997) *Chem. Rev. 97*, 2577−2590.
3. Mootz, H. D., and Marahiel, M. A. (1997) *Curr. Opin. Chem. Biol. 1*, 543−551.
4. Cane, D. E., and Walsh, C. T. (1999) *Chem. Biol. 6*, R319−R325.
5. Du, L., and Shen, B. (2001) *Curr. Opin. Drug Discuss. Dev. 4*, 215−228.
6. Du, L., Sanchez, C., and Shen, B. (2001) *Metab. Eng. 3*, 78−95.
7. Pelludat, C., Rakin, A., Jacobi, C. A., Schubert, S., and Heesemann, J. (1998) *J. Bacteriol. 180*, 538−546.
8. Silakowski, B., Schairer, H. U., Ehret, H., Kunze, B., Weinig, S., Nordsiek, G., Brandt, P., Blocker, H., Hofle, G., Beyer, S., and Muller, R. (1999) *J. Biol. Chem. 274*, 37391−37399.
9. Silakowski, B., Nordsiek, G., Kunze, B., Blocker, H., and Muller, R. (2001) *Chem. Biol. 8*, 59−69.
10. Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C., and Julien, B. S. (2000) *Science 287*, 640−642.
11. Molnar, I., Schupp, T., Ono, M., Zirkle, R. E., 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) *Chem. Biol. 7*, 97−109.
12. Staunton, J. (1998) *Curr. Opin. Chem. Biol. 2*, 339−345.
13. McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M., Betlach, M., and Ashley, G. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 1846−1851.
14. Nicolaou, K. C., Roschangar, F., and Vourloumis, D. (1998) *Angew. Chem., Int. Ed. Engl. 37*, 2014−2045.
15. Hardt, I. H. (2001) *J. Nat. Prod. 64*, 847−856.
16. Bollag, D. M., McQueney, P. A., Zhu, J., Hensens, O., Koupal, L., Liesch, J., Goetz, M., Lazarides, E., and Woods, C. M. (1995) *Can. Res. 55*, 2325−2333.
17. Muhlradt, P. F., and Sasse, F. C. R. (1997) *Can. Res. 57*, 3344−3346.
18. Chen, H., O'Connor, S., Cane, D. E., and Walsh, C. T. (2001) *Chem. Biol. 8*, 899−912.
19. Haydock, S. F., Aparicio, J. F., Molnar, I., Schwecke, T., Khaw, L. E., Konig, A., Marsden, A. F. A., Galloway, I. S., Staunton, J., and Leadlay, P. F. (1995) *FEBS Lett. 374*, 246−248.
20. Quadri, L. E. N., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) *Biochemistry 37*, 1585−1595.
21. Sanchez, C., Du, L., Edwards, D. J., Toney, M. D., and Shen, B. (2001) *Chem. Biol. 8*, 725−738.
22. Riordan, J. F., and Vallee, B. L. (1967) *Methods Enzymol. 11*, 541−548.
23. Jez, J. M., Ferrer, J.-L., Bowman, M. E., Dixon, R. A., and Noel, J. P. (2000) *Biochemistry 39*, 890−902.
24. Yang, Z., He, Y., Vourloumis, D., Vallberg, H., and Nicolaou, K. C. (1997) *Angew. Chem., Int. Ed. Engl. 36*, 166−168.
25. Taylor, R. E., and Haley, J. D. (1997) *Tetrahedron Lett. 38*, 2061−2064.
26. Royals, E. E. (1948) *J. Am. Chem. Soc. 70*, 489−491.
27. Bredenkamp, M. W., Holzapfel, C. W., and van Zyl, W. J. (1990) *Synth. Commun. 20*, 2235−2249.
28. Staunton, J., and Sutkowski, A. C. (1991) *J. Chem. Soc., Chem. Commun.*, 1110−1112.
29. Li, Z., Martin, F. M., and Vederas, J. C. (1992) *J. Am. Chem. Soc. 114*, 1531−1533.
30. Cane, D. E., Tan, W., and Ott, W. R. (1993) *J. Am. Chem. Soc. 115*, 527−535.
31. Brown, M. J. B., Cortes, J., Cutter, A. L., Leadlay, P. F., and Staunton, J. (1995) *J. Chem. Soc., Chem. Commun.*, 1517−1518.
32. Pieper, R., Ebert-Khosla, S., Cane, D., and Khosla, C. (1996) *Biochemistry 35*, 2054−2060.

BI020006W