ELSEVIER

Chemistry & Biology 8 (2001) 899–912

Chemistry & Biology

www.elsevier.com/locate/chembiol

Research Paper

Epothilone biosynthesis: assembly of the methylthiazolylcarboxy starter unit on the EpoB subunit

Huawei Chen [a], Sarah O'Connor [a], David E. Cane [b], Christopher T. Walsh [a,*]

[a] *Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA*
[b] *Department of Chemistry, Box H, Brown University, Providence, RI 02912-9108 USA*

Received 11 April 2001; revisions requested 8 June 2001; revisions received 14 June 2001; accepted 19 June 2001
First published online 1 August 2001

Abstract

Background: Polyketides (PKs) and non-ribosomal peptides (NRPs) are therapeutically important natural products biosynthesized by multimodular protein assembly lines, termed the PK synthases (PKSs) and NRP synthetases (NRPSs), via a similar thiotemplate-mediated mechanism. The potential for productive interaction between these two parallel enzymatic systems has recently been demonstrated, with the discovery that PK/NRP hybrid natural products can be of great therapeutic importance. One newly discovered PK/NRP product, epothilone D from *Sorangium cellulosum*, has shown great potential as an anti-tumor agent.

Results: The chain-initiating methylthiazole ring of epothilone has been generated in vitro as an acyl-*S*-enzyme intermediate, using five domains from two modules of the polymodular epothilone synthetase. The acyl carrier protein (ACP) domain, excised from the EpoA gene, was expressed in *Escherichia coli*, purified as an apo protein, and then post-translationally primed with acetyl-CoA using the phosphopantetheinyl transferase enzyme Sfp. The four-domain 150-kDa EpoB subunit (cyclization–adenylation–oxidase–peptidyl carrier protein domains: Cy–A–Ox–PCP) was also expressed and purified in soluble form from *E. coli*. Post-translational modification with Sfp and CoASH introduced the HS-pantP prosthetic group to the apo-PCP, enabling subsequent loading with L-cysteine to generate the Cys-*S*-PCP acyl enzyme intermediate. When acetyl-*S*-ACP (EpoA) and cysteinyl-*S*-EpoB were mixed, the Cy domain of EpoB catalyzed acetyl transfer from EpoA to the amino group of the Cys-*S*-EpoB, generating a transient *N*-Ac-Cys-*S*-EpoB intermediate that is cyclized and dehydrated to the five-membered ring methylthiazolinyl-*S*-EpoB. Finally, the FMN-containing Ox domain of EpoB oxidized the dihydro heterocyclic thiazolinyl ring to the heteroaromatic oxidation state, the methylthiazolylcarboxy-*S*-EpoB. When other acyl-CoAs were substituted for acetyl-CoA in the Sfp-based priming of the apo-CP domain, additional alkylthiazolylcarboxy-*S*-EpoB acyl enzymes were produced.

Conclusions: These experiments establish chain transfer across a PKS and NRPS interface. Transfer of the acetyl group from the ACP domain of EpoA to EpoB reconstitutes the start of the epothilone synthetase assembly line, and installs and converts a cysteine group into a methyl-substituted heterocycle during this natural product chain growth.

Keywords: Epothilone biosynthesis; Polyketide synthase; Non-ribosomal peptide synthetase; polyketide synthase/non-ribosomal peptide synthetase hybrid

1. Introduction

Both polyketide (PK) and non-ribosomal peptide (NRP) natural products are constructed on thiotemplated multimodular protein assembly lines [1,2], termed the PK synthases (PKS) and NRP synthetases (NRPS), respectively. In these protein assembly lines the inventory and placement of catalytic and carrier domains in the modules dictate the selection and sequence of acyl and aminoacyl monomers incorporated into the growing natural product chains. Some of the most therapeutically interesting members of these natural product classes are PK/NRP hybrids,

Abbreviations: A domain, adenylation domain; ACP, acyl carrier protein; Cy domain, condensation/cyclization domain; Ox domain, oxidase domain; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; pantP, phosphopantetheine; PCP, peptidyl carrier protein; PKS, polyketide synthase; PPTase, phosphopantetheinyl transferase; NRPS, non-ribosomal peptide synthetase; rt, retention time; TCA, trichloroacetic acid; TCEP, tris(2-carboxyethyl)phosphine hydrochloride

* Corresponding author.
E-mail address: christopher_walsh@hms.harvard.edu (C.T. Walsh).

1074-5521/01/$ – see front matter
PII: S1074-5521(01)00064-3

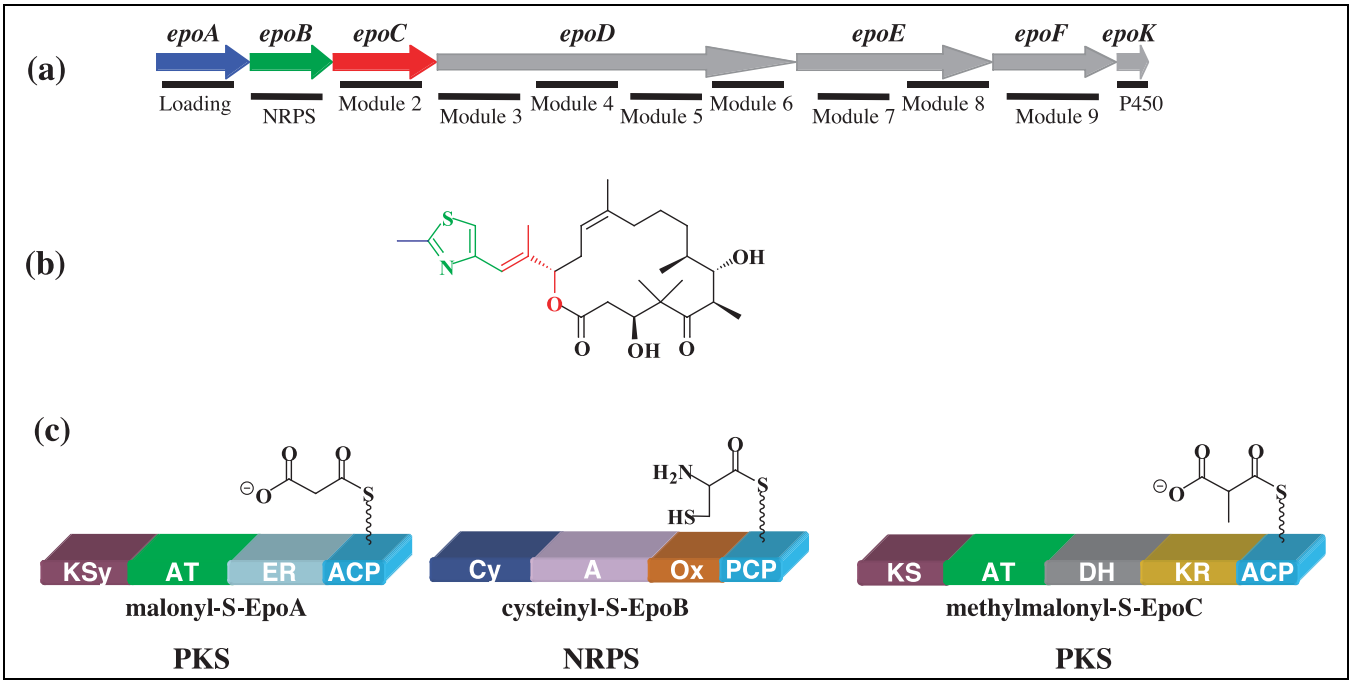


Fig. 1. (a) The epothilone biosynthetic gene cluster from *S. cellulosum*. Genes *epoA*, *epoB*, and *epoC* whose gene products represent the PKS/NRPS/PKS interfaces are highlighted in color. The colors of the genes are coded to follow the pieces they incorporate into epothilones. (b) Epothilone natural products include epothilone A, B, C, and D and among them epothilone D is the most promising anticancer drug candidate whose chemical structure is shown. (c) Domain organization of EpoA, EpoB, and EpoC with the putative substrates covalently docked onto the carrier protein moiety. KSy denotes a mutation of a critical active site residue cysteine to tyrosine in EpoA KS.

such as the immunosuppressant drugs FK506 and rapamycin [3,4], the widely used anticancer chemotherapeutic agent bleomycin [5] and, most recently discovered, epothilone D (Fig. 1b), now undergoing clinical trials for oncology. Sequence analysis of the epothilone biosynthetic gene cluster [6,7] from the producer organism *Sorangium cellulosum* reveals a hybrid assembly line (Fig. 1a), with the EpoA, C, D, E, and F subunits comprising a PKS array of a loading module (EpoA) followed by eight extender modules (EpoC–F). The second subunit EpoB has four domains reminiscent of an NRPS module, cyclization (Cy)–adenylation (A)–oxidase (Ox)-peptidyl carrier protein (PCP), and, interspersed between EpoA and EpoC PKS subunits, is appropriately placed to create the thiazole ring of epothilones [6,7]. It is anticipated that the interface between the EpoA and EpoB subunits is a switch point between PKS-type Claisen chemistry and NRPS-type peptide chemistry (Fig. 1c). The EpoB and EpoC interface correspondingly switches the assembly line logic back from NRPS to PKS for modules 2–9 for construction of the 16-membered macrolactone ring moiety (Fig. 1b).

In this work we have expressed and purified the acyl carrier protein (ACP) domain of the EpoA subunit and the full-length EpoB subunit in *Escherichia coli*, have enzymatically loaded the former with an acetyl group, the latter with a cysteinyl group, and detected chain transfer, condensation, cyclization/dehydration, and oxidation to build a methylthiazolylcarboxy-*S*-EpoB acyl enzyme intermediate reconstituting the PKS/NRPS switch point and the start of the epothilone synthetase assembly line. When various acyl-CoAs are used in place of acetyl-CoA, a series of alkylthiazoles are formed as covalent acyl-*S*-enzymes on the EpoB subunit.

2. Results

2.1. Heterologous expression and purification of the ACP domain of the EpoA subunit and the full-length EpoB subunit

Initial efforts to express the 150-kDa four domain EpoA subunit of the *S. cellulosum* in *E. coli* produced inactive protein under various growth conditions. We therefore focused on the expression of the C-terminal ACP domain of EpoA (Fig. 1c). Sequence alignment with other ACP domains (EpoD, EpoE, MtaF, McyD, NysC) [6–10] using the BLAST algorithm led us to predict the domain boundary and express the ACP domain alone (residues 1286–1421) as a 136-residue fragment of 14.5 kDa. Included in the fragment was the proposed linker region connecting the ACP domain with the C-terminal portion by which the ACP domain is thought to interact with the N-terminal portion of EpoB protein [11]. Expressed with an N-terminal His6 tag, this protein was soluble and readily purified by nickel affinity chromatography as shown in Fig. 2A, with a yield of 5 mg/l. Relatively short induction times

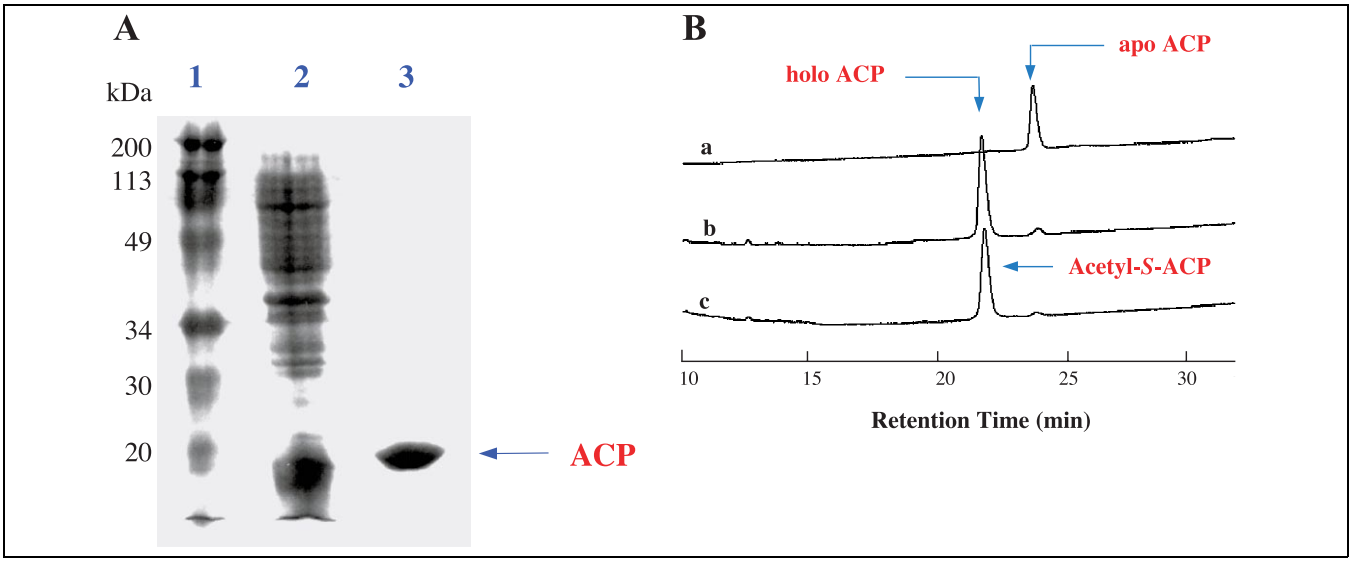


Fig. 2. (A) Analysis of the purification of EpoA-ACP by nickel affinity chromatography (15% SDS–PAGE). Lane 1, molecular weight markers; lane 2, crude cell extract; lane 3, purified ACP protein. (B) HPLC analysis of various EpoA-ACPs. (a) apo-ACP as purified from *E. coli*; (b) holo-ACP resulted from Sfp modification of the apo-ACP with free CoA. (c) Acetyl-*S*-ACP obtained by Sfp modification of the apo-ACP with acetyl-CoA.

(3 h at 24°C) minimized the degree of post-translational phosphopantetheine (pantP) modification of the ACP domain by the *E. coli* phosphopantetheinyl transferase (PPTase) ACPS [12,13] ensuring that the apo form of the ACP domain was isolated. HPLC analysis (Fig. 2B, trace a) and matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) (Table 1) confirmed pure apo-ACP protein ($[M+H]^+$ observed, 16 530; predicted, 16 527).

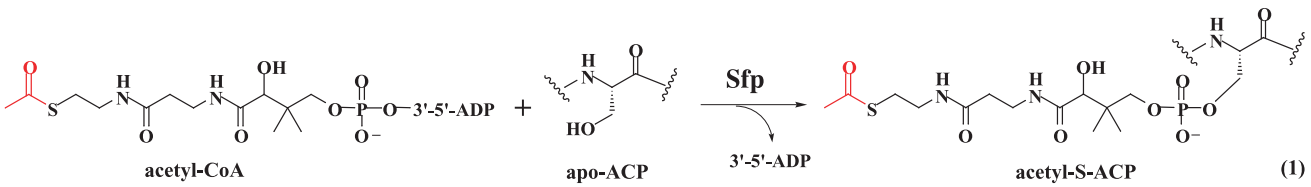
Purification of full-length 150-kDa four domain EpoB also involved an N-terminal His tag strategy. Expression conditions were optimized to give detectable soluble EpoB that could be purified on a Ni–NTA column, followed by MonoQ and size exclusion chromatographic steps (Fig. 3A) with a protein yield of 1.5 mg per liter of cell culture. The purified EpoB protein is noticeably yellow and shows characteristic absorbance of a flavin cofactor in its UV–vis spectrum (Fig. 3B). This was identified as FMN by coelution with authentic standard on HPLC analysis after heat denaturation of the protein. The $A_{279/441}$ ratio of 9.8/1 is consistent with essentially stoichiometric cofactor loading of the EpoB subunit. We could separately express just the 27.9-kDa Ox domain of EpoB and also purify it as an FMN-containing fragment (data not shown).

2.2. *Post-translational priming of the apo-ACP domain of EpoA by the PPTase Sfp with CoA or acetyl-CoA to yield HS-pantP-ACP or acetyl-S-pantP-ACP*

Ser1345 of the ACP domain is located within the consensus sequence GMDSLM for post-translational phosphopantetheinylation. In the absence of the *S. cellulosum* PPTase that would perform the normal post-translational priming reaction, we utilized the PPTase from the surfactin biosynthetic operon of *B. subtilis* as a versatile priming catalyst [14,15]. The HPLC trace b in Fig. 2B shows the quantitative conversion of the pure apo-ACP to the holo form and Table 1 validates the expected gain of 340 mass units of the pantP group. However, since we removed the ketosynthase (KS) and acyltransferase (AT) modules that normally derivatize the phosphopantetheinyl moiety with an acyl group, the HS-holo form of this ACP domain could not be utilized in the experiments described below. Instead, the acetyl-*S*-pantP-ACP derivative was generated using the strategy illustrated in Eq. 1:

Table 1
MALDI-TOF mass analysis of ACP loaded with different acyl groups

ACPs	Calculated $[M+H]^+$	Observed $[M+H]^+$
apo-ACP	16527	16530
holo-ACP	16867	16870
Acetyl-*S*-ACP	16909	16911
Propionyl-*S*-ACP	16925	16926
Butyryl-*S*-ACP	16939	16948
Isobutyryl-*S*-ACP	16939	16946
Isovaleryl-*S*-ACP	16953	16959
Hexanoyl-*S*-ACP	16967	16969
Benzoyl-*S*-ACP	16973	16977
Acetoacetyl-*S*-ACP	16953	16957



by taking advantage of our previous findings that Sfp will accept not only CoASH as substrate but also various acyl-CoAs [14,15].

In this case, the acetyl-*S*-pantP group is transferred onto the side chain of Ser1345 to generate the acyl thioester form of the holo-ACP. Thus, the autoradiography (Fig. 4A) shows that the apo-ACP can be covalently modified to the holo form by Sfp and [^{3}H]CoA (lane 3) or to the [^{3}H]acetyl-*S*-ACP (lane 5) by instead using [^{3}H]acetyl-CoA as priming substrate. Fig. 4B shows a time course for the covalent loading of [^{3}H]acetyl-*S*-pantP moiety onto the apo-ACP, with a calculated fractional stoichiometry of loading of 70%. Finally, when the acetyl-*S*-pantP priming was repeated with unlabeled acetyl-CoA, HPLC anal-

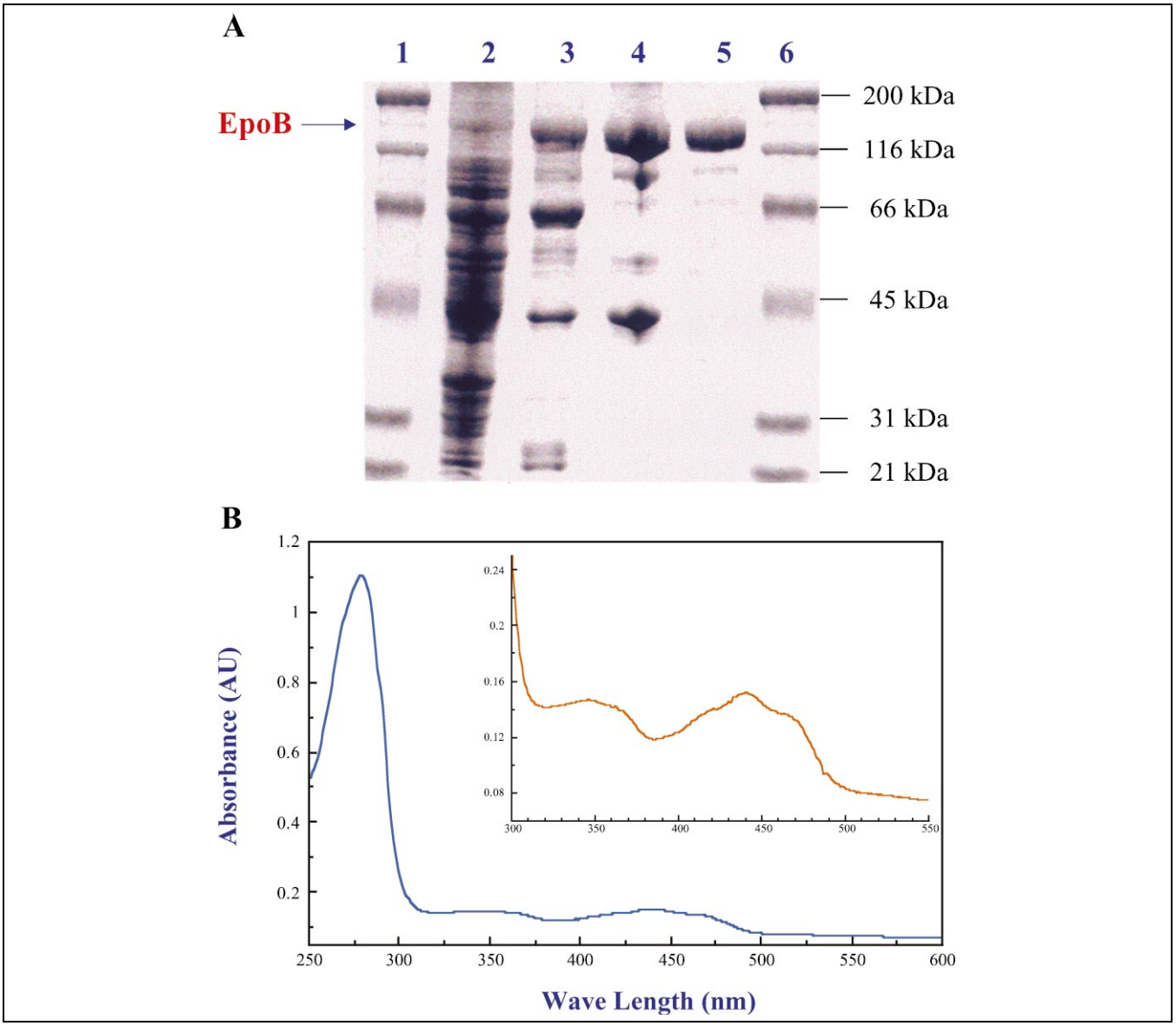


Fig. 3. (A) 10% SDS–PAGE analysis of the protein samples collected from different steps during EpoB purification. Lanes 1 and 6, molecular weight markers; lane 2, crude extract; lane 3, Ni–NTA affinity chromatography; lane 4, FPLC-MonoQ; lane 5, Size exclusion chromatography. (B) UV–vis spectrum of the purified EpoB protein. This spectrum displays three absorption maxima at 279 nm, 349 nm, and 441 nm, reminiscent of a flavin-containing protein. Inset shows the expanded 300–550 nm region to illustrate the fingerprint fine structure of the bound flavin cofactor.

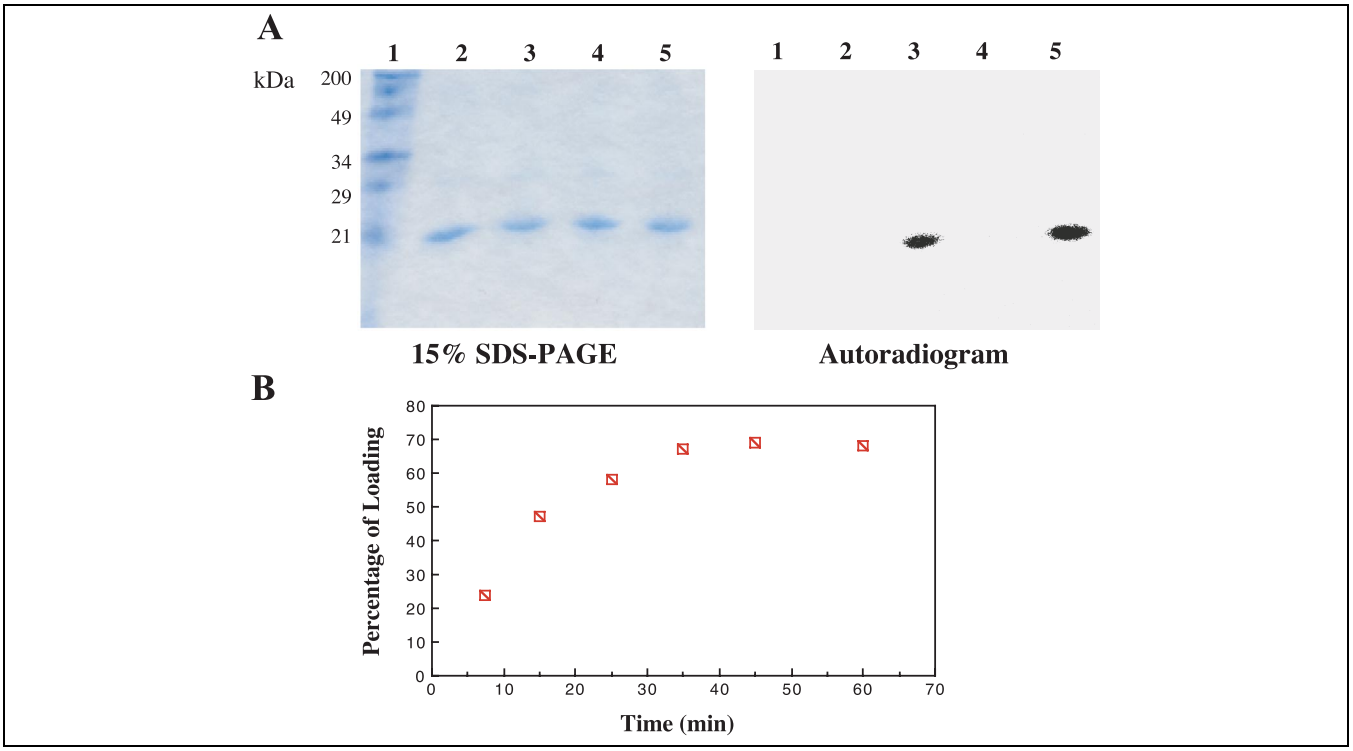


Fig. 4. (A) Coomassie brilliant blue-stained 15% SDS–PAGE gel analysis of Sfp-catalyzed post-translational modification and corresponding autoradiogram of EpoA-ACP samples. Lane 1, molecular weight markers; lane 2, ACP+[^{3}H]CoA−Sfp; lane 3, ACP+[^{3}H]CoA+Sfp; lane 4, ACP+[^{3}H]acetyl-CoA−Sfp, lane 5, ACP+[^{3}H]acetyl-CoA+Sfp. (B) Time course of the post-translational modification of ACP by PPTase Sfp with [^{3}H]acetyl-CoA.

ysis of the reaction mixture shows almost quantitative loading of acetyl group (Fig. 2B, trace c). MALDI-MS analysis of the reaction mixture showed an increase of 41 mass units (expected 42 for an acetyl group) compared to the HS-pantP-ACP, confirming the covalent loading of an acetyl group (Table 1).

2.3. *Characterization of the A and PCP domains of the EpoB subunit*

The NRPS-type EpoB subunit is predicted to have four domains, Cy–A–Ox–PCP, two of which, the A domain and the PCP domain, could be assayed for activity in the purified EpoB subunit alone. Given the presence of the thiazole in the epothilone natural product, the A domain is predicted to activate L-Cys to the L-Cys-AMP. This expectation is validated by the amino acid dependency and specificity of the ^{32}PPi–ATP exchange reaction, which is a classic measure of reversible formation of enzyme-bound aminoacyl-AMP [16]. The EpoB A domain was selective both for the chirality of α-carbon and properties of the side chain group. Results for the substrate-dependent PPi–ATP exchange reaction catalyzed by EpoB are shown in Fig. 5 and Table 2. Amino acids L-Ala, L-Asp, D-Cys, L-Cys, L-Leu, L-Lys, L-Phe, L-Ser, L-Thr, and (*S*)-2-aminobutyrate were examined. Several of these substrates failed to produce detectable signal. The relative selectivity of some representative substrates is shown in Fig. 5. Only L-Cys, L-Ser, and (*S*)-2-aminobutyrate were sufficiently active for the kinetic measurements and the results are shown in Table 2.

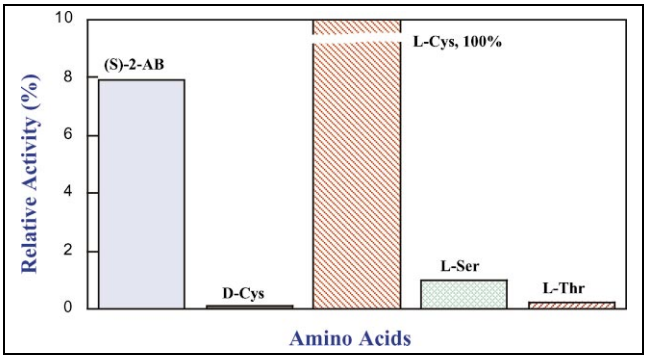


Fig. 5. Adenylation specificity of the EpoB A domain. The incorporation of [^{32}P]PPi into ATP was assessed by the ATP–PPi exchange assays. (*S*)-2-AB = (*S*)-2-aminobutyrate. L-Ala, L-Asp, L-Leu, and L-Lys, and L-Phe were also examined but all of them failed to show detectable activity (data not shown).

The activated Cys-AMP is expected to be captured intramolecularly by the HS-pantetheinyl arm of the holo-PCP domain of EpoB to covalently install the cysteinyl moiety. Conversion of apo to holo form of the PCP domain of EpoB was achieved by the same strategy of post-translational modification used on the apo-ACP domain of the EpoA subunit above with CoA and the Sfp PTTase (data not shown). Then, the covalent auto-aminoacylation

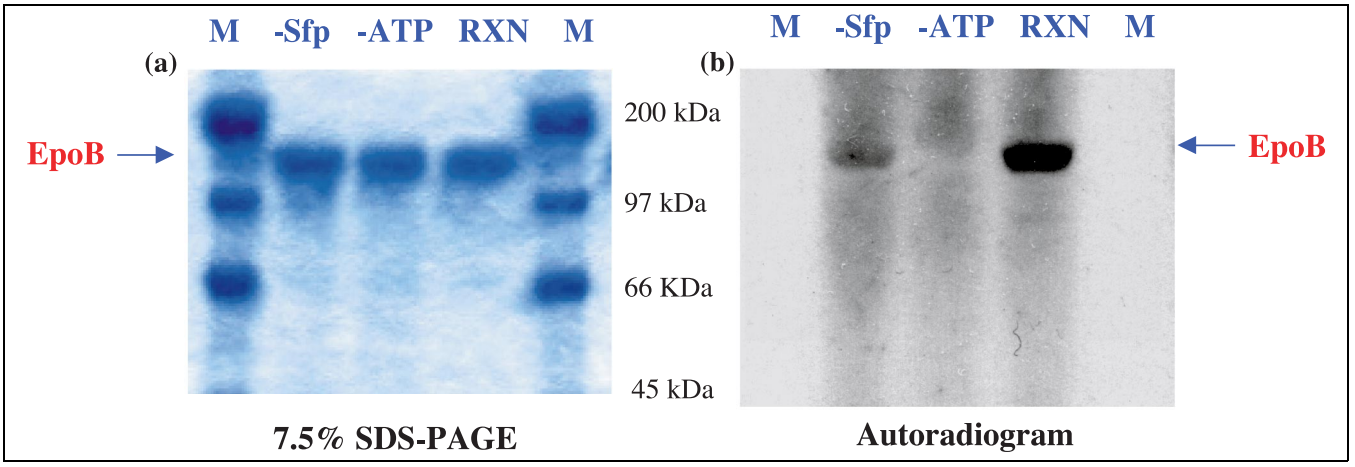


Fig. 6. Analysis of auto-aminoacylation of EpoB with [^{35}S]L-cysteine by SDS gel electrophoresis and autoradiography. (a) Coomassie brilliant blue-stained 10% SDS–PAGE gel. Lanes 1 and 5, molecular weight markers; lane 2, control reaction omitting Sfp; lane 3, control reaction omitting ATP; lane 4, full reaction. (b) Autoradiogram of this gel.

of holo EpoB subunit by [^{35}S]Cys was readily detected by autoradiography after SDS gel electrophoresis (Fig. 6). The Cys-*S*-enzyme formation is dependent on ATP (lanes 3 and 4 on the autoradiogram) and prior priming by Sfp (lane 2). A small amount of labeling in the absence of Sfp most likely reflects a small fraction of holo EpoB, presumably resulting from the action of an indigenous *E. coli* PPTase [13] under the particular overexpression conditions used here for EpoB. Trichloroacetic acid (TCA) precipitation assays were used to analyze the time course of auto cysteinylation, which was shown to be a rapid process and L-Cys reached its final stoichiometry of 82% within a minute (data not shown).

2.4. *Transfer of the [^{3}H]acetyl group from [^{3}H]acetyl-S-ACP to cysteinyl-S-EpoB: [^{3}H]methylthiazolylcarboxy-S-enzyme formation*

With [^{3}H]acetyl-*S*-ACP and cysteinyl-*S*-EpoB acyl enzyme forms in hand, the intersubunit transfer of the acetyl group, across the PKS/NRPS interface, could be examined by monitoring transfer of the radioactive acetyl group from the EpoA-ACP to the EpoB subunit. As shown in Fig. 7A, an autoradiogram clearly indicates [^{3}H]acetyl-containing EpoB after incubation with [^{3}H]acetyl-ACP of EpoA. This transfer is dependent on ATP (lane 3) and on cysteine (lane 4) and is specific for the acetyl-*S*-ACP as donor, since acetyl-CoA will not substitute up to concentration of 5 mM (data not shown). Although this substrate transfer experiment clearly establishes the condensation function of the Cy domain, the cyclization/dehydration function of the Cy domain of EpoB remained unclear, requiring a separate assay to establish whether the acylated form of EpoB was the acyclic *N*-acetyl-Cys-*S*-EpoB, or the heterocyclized methylthiazolinyl-*S*-EpoB. Also, the function of a putative Ox domain in EpoB would be confirmed if the final product was the two-electron oxidized heterocyclic methylthiazolylcarboxy-*S*-EpoB. The reactions were repeated with [^{35}S]cysteine as the labeled EpoB substrate, the acyl thioester linkage(s) in the acyl-EpoB enzyme were hydrolyzed by 0.1 M KOH treatment, and the released small molecule was subjected to HPLC analysis and the results are shown in Fig. 7B. The [^{35}S]Cys peak (retention time (rt) = 4.5 min) was almost undetectable, and instead a single radioactive peak (rt = 14.3 min) was observed. When coinjected with the chemically prepared standards, this ^{35}S-associated product peak comigrated with 2-methylthiazolyl-4-carboxylic acid. Neither the acyclic *N*-acetyl-Cys nor the methylthiazolinyl intermediate was detected, indicating that cyclization/dehydration activity of the Cy domain of EpoB as well as the FMN-dependent Ox domain acting as a thiazoline oxidase are fully functioning. The identity of the enzymatic product was further confirmed by mass analysis, MALDI-time-of-flight (-TOF) of the HPLC purified material from the EpoB incubation to be methylthiazolyl carboxylate ([M+H]$^+$, $C_5H_5NO_2S$, calculated, 144.01; observed, 143.98).

Table 2
Kinetic parameters of the ATP-[^{32}]PPi exchange assays of the EpoB A domain

Substrate	k_{cat} (min^{-1})	K_m (mM)	k_{cat}/K_m ($mM^{-1}min^{-1}$)
L-Cysteine	54.9 ± 1.3	0.46 ± 0.04	119.3
S-2-Aminobutyrate	29.4 ± 0.4	8.0 ± 0.3	3.7
L-Serine	11.0 ± 0.2	45.1 ± 1.2	0.24

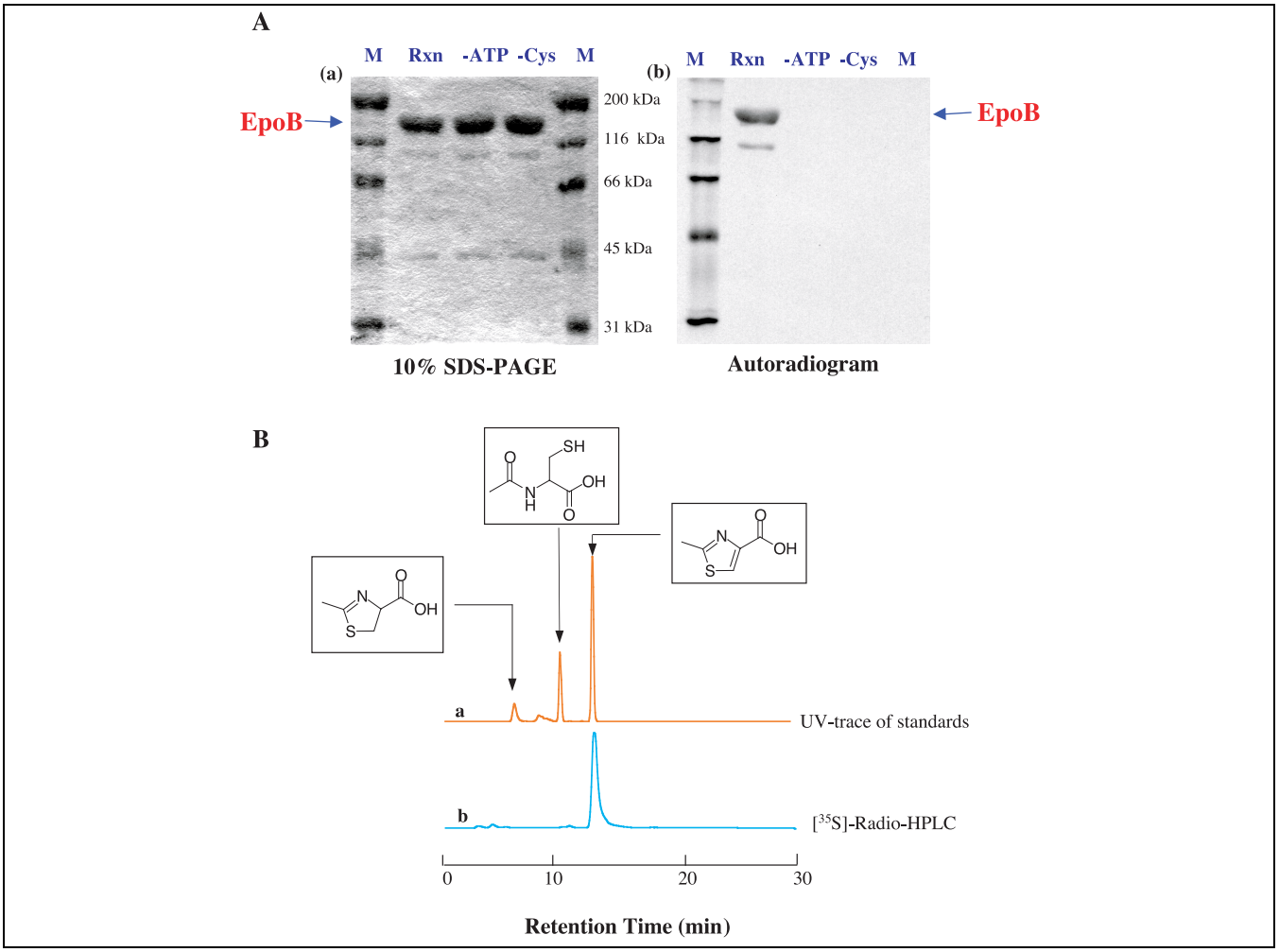


Fig. 7. (A) Analysis of the [^{3}H]acetyl group transfer from [^{3}H]acetyl-*S*-ACP to EpoB by SDS gel electrophoresis and autoradiography. (a) 10% SDS–PAGE of EpoB samples from various transfer experiments. Lanes 1 and 5, molecular weight markers; lane 2, full reaction containing [^{3}H]acetyl-*S*-ACP, ATP, L-Cys, holo EpoB; lane 3, control sample omitting ATP; lane 4, control sample omitting L-Cys. (b) Corresponding autoradiogram. (B) HPLC analysis and identification of acetyl-*S*-ACP and [^{35}S]Cys-*S*-EpoB incubation product released from the PCP domain of EpoB by KOH treatment. (a) HPLC trace of the authentic standards added to radioactive reaction mixture as cold carriers which were monitored by an UV detector at 220 nm. (b) Radio-HPLC trace of the same sample recorded simultaneously by an on-line radioactivity detector tuned for ^{35}S.

2.5. *Recognition of different acyl-S-ACPs by the Cy domain of EpoB: Aryl- and alkylthiazolylcarboxy-S-EpoB acyl enzyme formation*

Given the successful use of the Sfp PPTase to load the acetyl-*S*-pantetheinyl moiety onto the apoACP domain of EpoA, we then tested several additional acyl-CoA derivatives to produce the corresponding acyl-*S*-ACP species as potential acyl donors in the condensation/heterocyclization/oxidation reactions carried out by the EpoB subunit. As detailed in Table 1, the broad specificity of Sfp allowed installation of propionyl, butyryl, and hexanoyl straight chain acyl-*S*-pantP moieties as well as the branched chain isobutyryl and isovaleryl thioesters. The β-keto-acetoacetyl-*S*-ACP and the aryl benzoyl-*S*-ACP were likewise converted to a single species as demonstrated by HPLC analysis (data not shown) and by MALDI-MS analysis (Table 1). These acyl-*S*-ACP derivatives were then used singly as donors to [^{35}S]cysteinyl-*S*-EpoB. The small molecule products were cleaved from the acyl-*S*-EpoB and subjected to HPLC analysis. Figure 8A shows the HPLC traces for seven such reactions, where in every case a new radioactive peak containing [^{35}S]cysteine-derived label was detected. To validate that the alkyl or aryl aromatic heterocycles were generated, six standards of 2-substituted thiazolyl-4-carboxylic acids have been synthesized according to the reaction scheme outlined in Fig. 8B. HPLC coinjection of these chemically prepared standards with the radioactive products released from EpoB confirmed their identities and some selected chromatograms are shown in Fig. 8C. Clearly the Cy and the Ox domains of the EpoB subunit will tolerate wide variety in the acyl donor presented on the upstream partner ACP, allowing a set of substituted thiazolyl caps to be made in the epothilone assembly line.

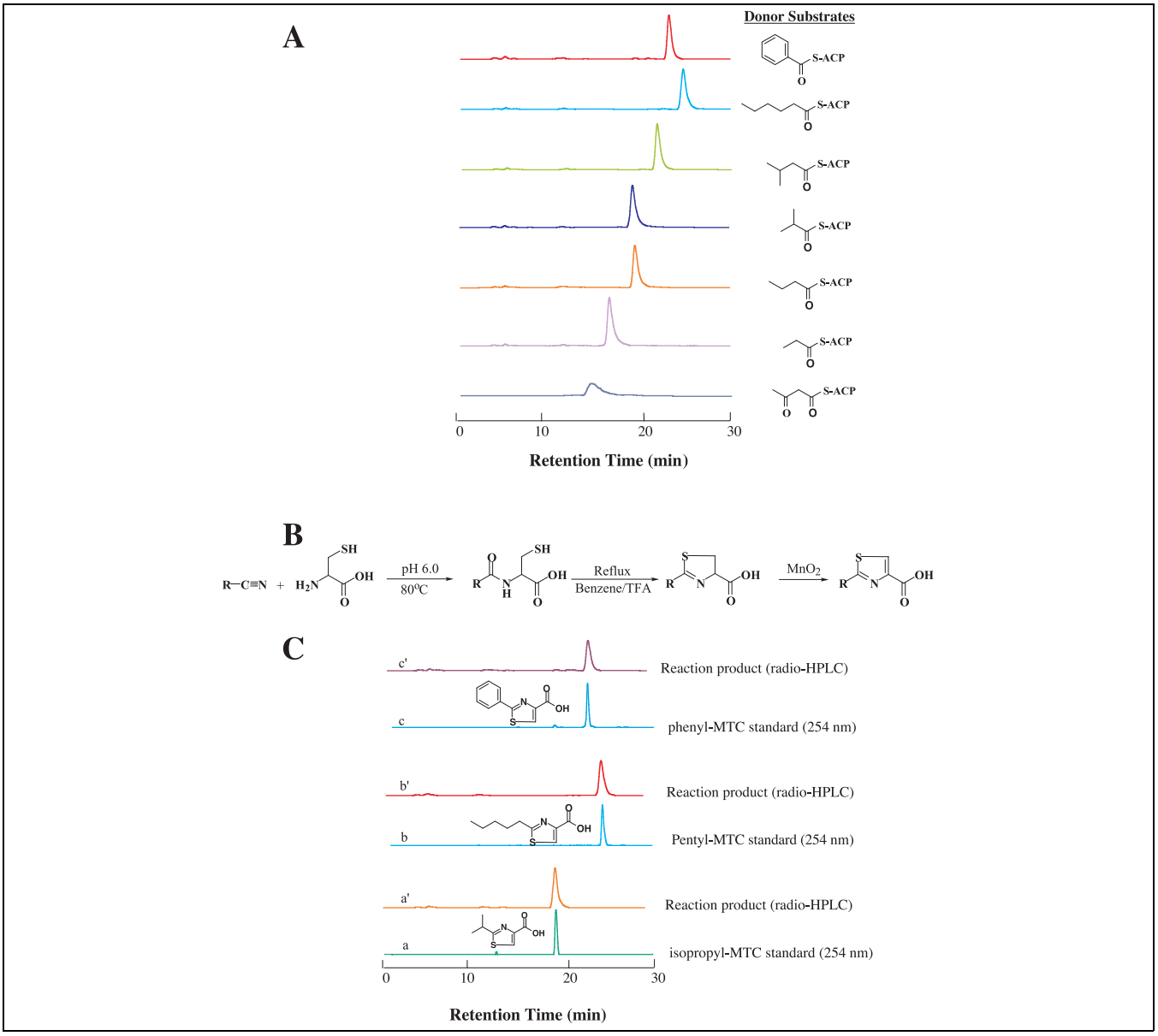


Fig. 8. (A) HPLC analysis of the acyl transfer reactions using various acyl groups loaded onto EpoA-ACP to serve as donor substrates for [^{35}S]Cys-*S*-EpoB. The reactions were quenched with 10% TCA and the pelleted proteins were treated with 0.1 M KOH to hydrolyze the thioester linkage to release the small molecule products. The supernatant was then injected into HPLC to resolve radioactive species. The [^{35}S]Cys peak (rt = 4.5 min) was almost non-detectable and a single new [^{35}S]Cys derived peak was observed for all the acyl donors tested. (B) General chemical synthetic scheme to obtain acyl-Cys, heterocyclized thiazolinyl carboxylic acids, and the two electron oxidized thiazolyl carboxylic acids. (C) HPLC coinjections of some selected synthetic standards with radioactive enzymatic reaction product. Dual on-line UV (at 254 nm) and radioactive detectors (tuned for ^{35}S) were used to monitor the standard and radioactive product simultaneously. UV-HPLC traces a, b, and c show the retention property of three selected standards 2-isopropyl-, 2-pentyl-, and 2-phenylthiazolyl-4-carboxylic acids, respectively, while radio-HPLC traces a′, b′, and c′ show the corresponding enzymatic reaction products which are coeluted with the standards. The rest enzymatic products were also confirmed to be the substituted thiazoles by same method (data not shown) except for the incubation of acetoacetyl-*S*-EpoA and EpoB whose standard has not been synthesized.

3. Discussion

The epothilone family of natural products has attracted interest for its potent anti-tumor effects [17,18], acting analogously to taxol in blocking mitotic spindle progression by stabilization of microtubules [19,20]. The maintenance of anti-tumor activity in multi-drug resistant and taxol-resistant cell lines has highlighted interest in clinical evaluation of epothilones and focused attention on the epothilone biosynthetic pathway. The epothilone family is a hybrid of PK and NRP building blocks, with one NRPS-type subunit, EpoB, embedded in a five-subunit, nine-module EpoA, C, D, E, F assembly line that creates the 16-membered macrolactone PK. It seemed likely that EpoB would be the key subunit in creation of the thiazole moiety of the epothilones, given the existence of compa-

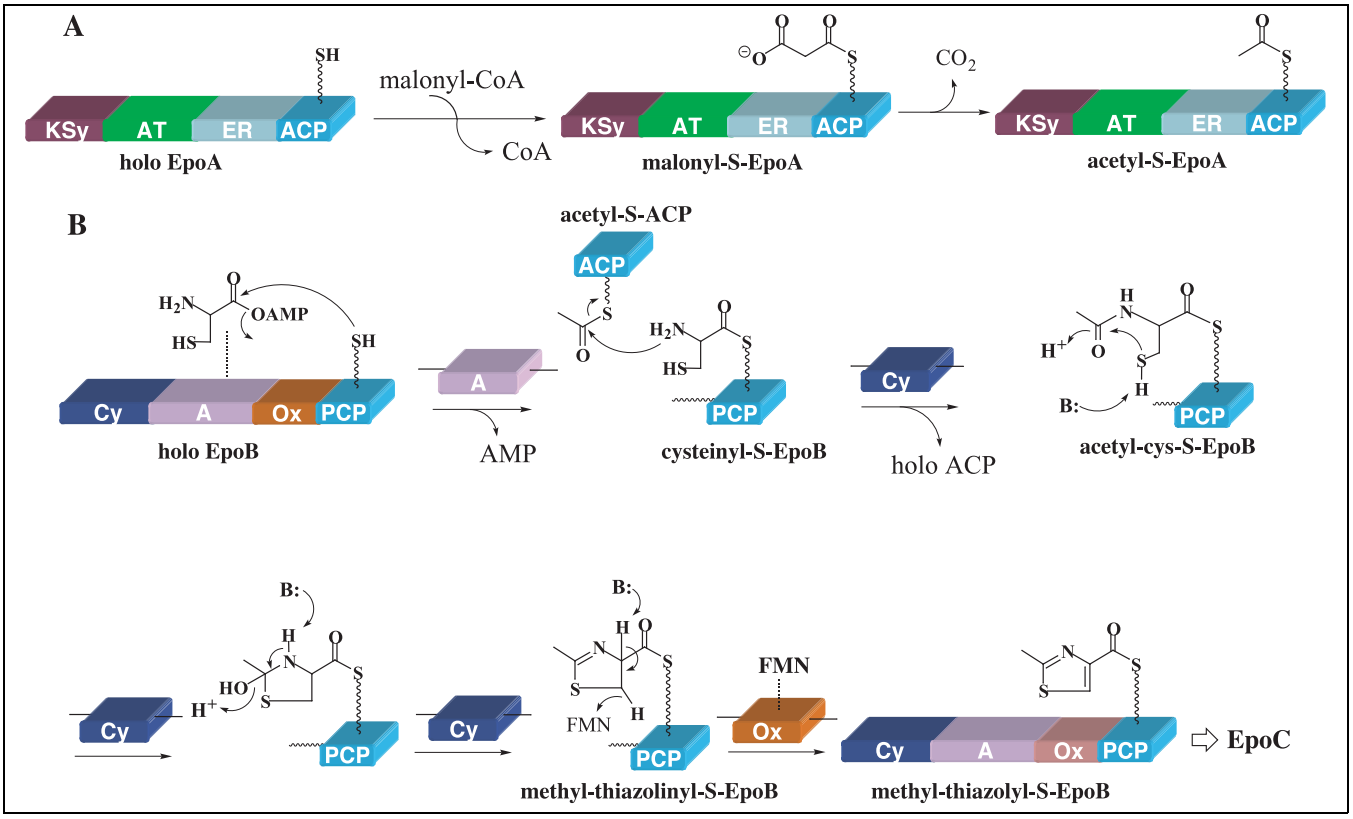


Fig. 9. (A) The EpoA is a PKS loading module of the epothilone biosynthetic machinery to provide the acetyl starter unit. The AT domain is proposed to specifically load a malonyl group on the C-terminal ACP domain, and the KSy domain in which the conserved active site cysteine is mutated to tyrosine would promote the decarboxylation reaction to give acetyl-*S*-EpoA. The ER domain is most likely to be non-functional. (B) Detailed reaction mechanism of these transformations effected by the NRPS module EpoB. The Cy domain is bifunctional catalyzing the acetyl group transfer across the first PKS/NRPS interface for C–N bond formation and the subsequent cyclization/dehydration to create the thiazolinyl thioester intermediate. The FMN-containing Ox domain acts as an acyl 2,3-desaturase to give the heteroaromatic moiety which is ready to be passed on to the following PKS module EpoC.

rable Cy and A domains in synthetase modules for thiazolines and thiazoles formation in other NRPS biosynthetic clusters (pyochelin [21]) and NRP/PKS hybrids (yersiniabactin [22], bleomycin [5]). Given its placement in the assembly line EpoB should accept an acyl donor from the EpoA subunit and in turn donate the methylthiazole moiety to the downstream PKS module in EpoC (Fig. 9B).

In this work we have focused on the EpoA and EpoB subunits and the acyl chain transfer between the subunits that constitutes the switch from PKS to NRPS assembly line logic. The four domain EpoA subunit has attributes of a prototypic initiation or loading module in a PKS assembly line, KS–AT–enoyl reductase (ER)–ACP domains, but two of the four, KS and ER, have mutations in consensus sequences likely to cause deactivation. In particular, the active site cysteine of the KS domain is missing, typical for a chain-initiating KS that is defective in condensation function but retains its ability to decarboxylate the malonyl-*S*-ACP [23]. The AT domain is predicted to specify malonyl CoA for loading onto the holo HS-pant-ACP domain [6,7]. The net result of the three functional domains should be to generate an acetyl-*S*-ACP on the C-terminal domain of EpoA as donor to the downstream EpoB module (Fig. 9A).

Our efforts to generate the *Sorangium* full-length EpoA in soluble, active form by heterologous expression in *E. coli* have not yet met with success, so we elected to bypass the KS and AT domain requirements and isolate the ACP domain of EpoA. A 14.5-kDa ACP fragment expresses well in *E. coli* as a soluble protein. To enact the bypass strategy and create the kinetically labile, thermodynamically activated acetyl-*S*-ACP fragment, two issues needed to be solved. The first issue was to produce the apo form of ACP rather than the phosphopantetheinylated holo form and the second was to concurrently install the pantP prosthetic arm and the acetyl moiety efficiently and quantitatively. In our prior experience in expression of heterologous ACP and PCP domains in *E. coli* [13,15,24], we have repeatedly noted that the *E. coli* PPTases are slow to post-translationally prime the foreign ACP and PCP proteins, such that relatively short induction times yield largely apo form, as was the case with the ACP domain excised from the EpoA subunit. The requirement for the apo rather than the HS-pantP-ACP holo form was at the core of the strategy for the second step:

to introduce the acyl-*S*-pantP moiety intact in a concerted enzymatic transfer step. This was possible because our earlier characterization of the *Bacillus subtilis* Sfp PPTase [14,15] indicated that this enzyme would utilize a variety of acyl-*S*-CoA derivatives and transfer the acyl-*S*-pantP moiety intact to the carrier protein. Subsequent X-ray analysis of Sfp [8] confirms that the cysteamine end of CoASH pointing out into solution when bound by the PPTase, suggesting that the acyl group is not involved in Sfp recognition. Indeed, the EpoA apo-ACP domain could be quantitatively primed not only with acetyl-*S*-pantP but with a variety of other acyl groups from acyl-*S*-CoA substrates.

The key catalytic subunit in fashioning the heteroaromatic methyl thiazole ring system in the epothilones is the four domain EpoB. Two of the four domains, A and PCP, are conventional NRPS modules. The remaining two domains, Cy and Ox domains, are not commonly found in the NRPS assembly line. The enzymatic function of the Ox domain has never been validated in any NRPS system before, although a similar Ox domain from the bleomycin cluster has been expressed and purified as an FMN-containing protein [25]. Our ability to overexpress and isolate from *E. coli* the soluble, full-length 150-kDa EpoB subunit, loaded with FMN in the Ox domain, enabled validation of all four domain functions. The A domain, as expected [6,7], is specific for L-cysteine, making the Cys-AMP and then tethering it in a thioester linkage to the adjacent holo-PCP domain, to yield the Cys-*S*-EpoB acyl enzyme that is the acceptor for acetyl-*S*-ACP. The holo form of the PCP domain of the EpoB protein was generated after purification by using the Sfp for a second time, but in this case with its natural cosubstrate CoASH, to yield the phosphopantetheinylated PCP domain.

The functions of the Cy and Ox domains of EpoB could then be assessed in the acetyl transfer/condensation assays. Covalent transfer of the [^{3}H]acetyl label from [^{3}H]acetyl-*S*-ACP to Cys-*S*-EpoB (but not to holo EpoB, see Fig. 7A) validated that the Cy domain was functional, at least in amide bond-forming condensation mode. We have elsewhere established that such Cy domains in pyochelin synthetase [21] and yersiniabactin synthetase [26–28] are also cyclodehydration catalysts, producing thiazoline rings from acyl-Cys-*S*-enzyme intermediates. Given the evidence that acetyl-*S*-ACP and [^{35}S]Cys-*S*-EpoB incubation accumulate the methylthiazolylcarboxy-*S*-EpoB (Fig. 7B), then this Cy domain is also a heterocyclization catalyst on the presumed initial *N*-Ac-Cys-*S*-EpoB condensation product (Fig. 9B). The remaining domain of EpoB, the C-terminal oxidase with its FMN cofactor should act as an olefin-forming desaturase, akin to flavoprotein acyl-CoA desaturases or dihydroorotate dehydrogenases [9,29]. In this case such desaturase action will convert the methylthiazolinylcarboxy-*S*-EpoB to the heteroaromatic thiazolylcarboxy-*S*-EpoB, completing the morphing transformation of the initial Cys-*S*-EpoB acyl enzyme (Fig. 9B).

This condensation, cyclization, dehydration, oxidation chemical sequence on the acyl-*S*-EpoB intermediates is stoichiometric in the absence of the downstream assembly line components, e.g. EpoC, so structural analysis required release of the acyl thioesters by alkaline hydrolysis to cleave the small molecules from the 150-kDa EpoB protein. HPLC analysis of the radioactive product and separate mass spectroscopic evaluation established that only 2-methylthiazolyl-4-carboxylate and no 2-methylthiazolinyl-4-carboxylate or *N*-acetyl-cysteinyl intermediate was detected (Fig. 7B). Separate controls established that authentic 2-methylthiazolinyl-4-carboxylate was largely intact under these work up conditions but did open partially to *N*-Ac-Cys. Since no *N*-Ac-Cys was detected in release from the EpoB incubations, we assume there was no methylthiazolinylcarboxy-*S*-enzyme species. These results confirm that the Ox domain is functional, and suggests that cognate Ox domains in the myxothiazol and bleomycin synthetase clusters [5,10] will have comparable conjugated olefinic thioester-forming activities.

This work establishes that reconstitution of the PKS/NRPS switch point in the epothilone assembly line has been achieved in the acetyl-*S*-ACP/cysteinyl-*S*-EpoB incubations and sets the stage for subsequent kinetic and mechanistic analysis of the catalytic parameters for this PKS/NRPS interface to make PK/NRP hybrid molecules. The kinetic analysis of the ACP-EpoB incubation was not being performed due to the rapid nature of acetyl group transfer (data not shown), indicating that future kinetic study will require chemical rapid-quench techniques. One initial constraint is that the acetyl donor moiety must be presented on the holo-ACP platform since acetyl-SNAC (a very poor substrate, data not shown) and acetyl-CoA did not serve as competent substrates, supporting a subunit/subunit interface recognition for catalysis of chain transfer. Subsequent work will evaluate whether heterologous ACP domains can serve as the acyl donor platform which will have implications for domain swapping in combinatorial reconstruction of the epothilone assembly line with different acyl starter units. Initial results indicate that EpoB will not recognize an ACP domain excised from the rapamycin gene cluster [30], suggesting that a level of specificity exist in the protein-protein interaction across the PKS/NRPS interface. Our strategy of bypassing the constraints of the KS and AT domains of the EpoA subunit by priming the purified apo-ACP domain with Sfp and different acyl-*S*-pantP moieties gives an initial evaluation of the permissivity of the EpoB Cy domain for acyl group variation in condensation and heterocyclization and of the Ox domain for alkylthiazolinyl-*S*-enzyme desaturation tolerance. It has been reported that the phenylthiazolyl variant of epothilone, made by total synthesis retains 25% of its cell killing activity [31], so combinatorial biosynthesis to produce altered heterocyclic caps to the PK moiety of epothilones may be a viable approach for

generating production quantities of novel epothilone hybrids.

4. Significance

In this report we have reconstituted the initial steps for the biogenesis of epothilone D, a promising anticancer drug with a similar mode of action as Taxol but possesses some superior pharmacological properties. We have investigated the acyl group transfer across the first PKS/NRPS interface and subsequent reactions effected by the NRPS module EpoB leading to the methylthiazole moiety formation. The putative oxidase was validated for the first time to be a FMN-dependent acyl 2,3-desaturase to convert thiazoline to thiazole, a function relevant to the biosyntheses of other thiazole-containing natural products such as bleomycin. Our findings that the Cy and the Ox domains in EpoB have relaxed substrate specificity toward upstream acyl donors would promote future work to create epothilone analogs with engineered loading modules through genetic manipulation.

5. Materials and methods

5.1. General

Competent cells of *E. coli* strain top10 and BL21(DE3) were purchased from Invitrogen. All restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs. Oligonucleotides were ordered from IDT and used without further purification. The PPTase Sfp was expressed and purified as previously described [14]. [^{3}H]acetyl-CoA and [^{35}S]L-cysteine was purchased from New England Nuclear. All acyl-CoA derivatives were purchased from Sigma. DNA sequencing and MALDI-MS was performed at the Dana Farber Cancer Institute.

5.2. Cloning of ACP domain of epoA (residues 1286–1421) and full-length epoB gene

The ACP domain of *epoA* and full-length *epoB* genes were amplified by PCR from a plasmid containing the EpoA,B,C genes (provided by Kosan Biosciences) [32] with the following primer pairs (acpNterm*Nde*I, 5′-G GAA TTC **CAT ATG** GCA GGC GCC GGC CCG TCC ACC-3′, and acpCterm*Eco*RI, 5′-TTT **GAA TTC** TCA TAG GGC AAT GAT TTC CCA-3′; epoBNterm*Nde*I, 5′-G GAA TTC **CAT ATG** ACC ATC AAC CAG CTT CTG AA-3′, and epoBCterm*Not*I, 5′-G GGA ATT A**GC GGC CGC** TTA GCT ACG TCT CCT GCC-3′). Restriction sites introduced into the N- and C-terminal portions of the genes are highlighted in the primer sequence in bold. Amplified products were purified by agarose gel, digested with the appropriate restriction enzymes, and ligated into pET28b to provide the N-terminal His6-tagged construct. DNA sequencing confirmed the correct insertion of each gene into the plasmid, and the resulting plasmids were transformed separately into BL21(DE3) cells for protein overexpression.

5.3. Overexpression and purification of EpoA-ACP protein

For overproduction of ACP protein, cells harboring the desired plasmid were grown in LB medium supplemented with 30 μg/ml of kanamycin. Cultures (3×1 l) were grown to an OD_{600} of 0.55 at 37°C, cooled to 24°C, and protein expression induced by the addition of 100 μM IPTG. Cultures were grown for an additional 3 h. The harvested cells were then resuspended in 30 ml of lysis buffer (25 mM Tris, pH 8.0, 200 mM NaCl, 10% (v/v) glycerol) and lysed by French Press. The lysate was clarified by centrifugation (30 min at 35 000×*g*) and promptly applied to 2 ml of Ni–NTA resin (Qiagen). After binding in batch for 2 h at 4°C, the resin was decanted into a column, washed with 15 column volumes of lysis buffer, and then eluted with a step gradient of 30, 60, 100, 250 and 500 mM imidazole. After analysis by SDS–PAGE, fractions containing protein were pooled, dialyzed against 25 mM Tris, pH 8.0, 200 mM NaCl, 10% (v/v) glycerol, and stored at −80°C. ACP protein concentration was assessed using the Bradford assay (Bio-Rad).

5.4. Overexpression and purification of EpoB protein

To overexpress EpoB protein, BL21(DE3) strains carrying the desired plasmid were cultivated (6×1 l) in 2×YT broth containing 50 μg/ml of kanamycin at 15°C without adding IPTG to induce protein production. After 72 h of growth, cells were harvested by centrifugation (10 min at 6000×*g*), resuspended in 100 ml of lysis buffer (25 mM Tris, 200 mM NaCl, 2.5 mM imidazole, pH 8.0), disrupted by two passages of French Press. Cell lysate was clarified by centrifugation (30 min at 35 000×*g*). The EpoB protein was first extracted from the crude cell lysate using 25 ml Ni–NTA resin in batch binding fashion at 4°C for 4 h. After packing the Ni–NTA resin slurry into a column, the column was washed with 10 bed volumes of wash buffer (25 mM Tris, 200 mM NaCl, 10 mM imidazole, pH 8.0) and EpoB protein was eluted with 200 mM imidazole. The second chromatographic step was FPLC-MonoQ (Pharmacia) with a linear elution profile of 50–500 mM NaCl salt gradient over 270 ml. The flow rate was set at 3 ml/min and fractions of 3 ml were collected. Fractions containing EpoB, as judged by SDS–PAGE, were pooled, concentrated, and subjected to a size exclusion chromatographic step with a HiLoad 16/60 Superdex 200 prep grade column (Pharmacia). The buffer was 25 mM Tris, 100 mM NaCl, pH 7.5 and the flow rate was 0.8 ml/min. Fractions (0.8 ml in size) containing EpoB were pooled, dialyzed against 20 mM Tris, pH 7.5, 50 mM NaCl, 1 mM DTT, and 10% glycerol, concentrated, and stored at −80°C until use.

5.5. Modification of EpoA-ACP with Acyl-CoA

Reaction conditions were identical for each of the acyl-CoA derivatives. In a representative example, EpoA-ACP (0.75 nmol) was incubated with acetyl-CoA (3–5 fold molar excess) in 50 μl buffer (50 mM Tris, pH 7.0, 5 mM $MgCl_2$, 5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 40 mM NaCl). The reaction was initiated by the addition of Sfp (70 nM) and allowed to incubate at ambient temperature (24°C). Reaction progresses were monitored either by TCA precipitation and liquid scintillation counting assays for the radiolabeled acyl-CoAs or by HPLC. In the case of radiolabeled [^{3}H]acetyl-CoA (28.3 Ci/mmol), the reaction was quenched by the addition of 10% TCA. BSA (100

μg) was added as a carrier protein, and the precipitated proteins were centrifuged, and washed twice with 10% TCA. The protein pellet was resuspended in 100 μl of formic acid and the amount of radiolabel incorporation into the ACP was quantified by liquid scintillation counting. Alternatively, the reaction was analyzed by autoradiography to assure the specific labeling of the ACP protein. A fraction of the reaction mixture (10 μl) was used for electrophoresis on a 15% SDS–PAGE. A control lacking Sfp was also carried out. The gel was stained with Coomassie blue solution, destained, and soaked in Amplify (Amersham) for 15 min. The dried gel was exposed to film for 14 h at −80°C. Reactions performed with unlabeled acetyl-CoA were analyzed by injecting a small amount of sample into HPLC with a Vydac C18 reverse phase column (linear gradient from 40% to 60% MeCN in water containing 0.1% TFA over 30 min).

5.6. Characterization of the A and PCP domains of EpoB to generate a L-cysteinyl-S-EpoB intermediate

5.6.1. ATP–PPi exchange assay for A domain substrate specificity of EpoB

Reactions (100 μl) containing 75 mM Tris (pH 7.5), 10 mM $MgCl_2$, 5 mM TCEP, 3 mM ATP, 1.5 mM of amino acid substrate, 200 nM EpoB, and 1 mM [^{32}P]pyrophosphate (3.6 Ci/mol) were carried out at 24°C. Following the initiation by the addition of EpoB, reactions were allowed to proceed for 7 min, and then quenched with charcoal suspensions (500 μl of 1.6% (w/v) activated charcoal, 4.5% (w/v) tetrasodium pyrophosphate, and 3.5% perchloric acid in water). The charcoal was pelleted by centrifugation, washed twice with quenching buffer without added charcoal, resuspended in 0.8 ml of water, and subjected to liquid scintillation counting. The amount of bound radioactivity was converted to reaction rate using the specific radioactivity of the [^{32}P]pyrophosphate and the length of reaction time. To determine kinetic parameters of L-Cys, reaction were carried out at 24°C in total volume of 100 μl that contained 75 mM Tris–HCl (pH 7.5), 10 mM $MgCl_2$, 5 mM TCEP, 3 mM ATP, 50 nM of EpoB, 1 mM [^{32}P] pyrophosphate, and various amount of L-Cys for 5 min. For the incubations of (*S*)-2-aminobutyrate, 0.5 μM EpoB was used and the reactions were allowed to continue for 10 min at 24°C, while 0.5 μM EpoB was used for 10 min incubations for L-Thr.

5.6.2. Assays for auto-aminoacylation of [^{35}S]L-Cys to the PCP domain of EpoB

EpoB was first primed with Sfp to install the pantP group on the PCP domain. The specific loading of [^{35}S]L-Cys to the PCP domain of EpoB was first investigated by autoradiography. Incubation mixture (50 μl) included 75 mM Tris (pH 7.5), 10 mM $MgCl_2$, 0.25 mM [^{35}S]Cys (10 Ci/mmol), 5 mM TCEP, 3 mM ATP, and 3 μM of EpoB. The reaction was allowed to proceed for 30 min at 24°C and 15 μl was used for electrophoresis on 10% SDS–PAGE. Control reactions were carried out in parallel in which either the ATP was omitted or EpoB was used without Sfp treatment. For visualization, the gel was stained with Coomassie blue solution, destained, and soaked in Amplify for 15 min. The dried gel was exposed to film for 15 h before developing. The time course of the auto-aminoacylation was then examined using TCA precipitation assay. Reactions of 100 μl in volume contained 75 mM Tris (pH 7.5), 10 mM $MgCl_2$, 5 mM TCEP, 0.3 mM CoA, 0.2 μM Sfp, 3 μM EpoB, 0.25 mM [^{35}S]L-Cys (90 Ci/mol), and 3 mM ATP, and were incubated for 30 min at 24°C to allow the phosphopantetheinylation of the PCP domain prior to initiation by the addition of ATP. After addition of ACP, aliquots of 10 μl were withdrawn at various time points, quenched with 0.5 ml of 10% TCA, and the precipitated protein was pelleted, washed, and counted for radioactivity. Percent aminoacylation was calculated from the specific activity of the [^{35}S]L-Cys and the EpoB concentration.

5.7. Characterization of the Cy and Ox domains of EpoB

5.7.1. Analysis of the acetyl group transfer from EpoA-ACP to EpoB by autoradiography

EpoA-ACP (0.45 nmol) was primed with [^{3}H]acetyl-CoA (28.3 Ci/mmol) as described above. EpoB (0.15 nmol) was primed with free CoA and then loaded with L-Cys as described. The two reaction solutions were then combined and allowed to incubate at 24°C for 60 min. Control reactions in which either ATP or L-Cys was omitted were also performed. Non-reducing protein sample buffer was then added to the solution and a fraction of the sample was loaded and resolved on a 10% SDS–PAGE gel. The gel was stained with Coomassie, destained, soaked in Amplify for 15 min and then dried. The gel was exposed to film for 48 h at −80°C.

5.7.2. Identification EpoA-ACP and EpoB reaction product by HPLC and mass spectrometric analysis

Acetyl-*S*-ACP and [^{35}S]Cys-*S*-EpoB (2.6 Ci/mmol) were incubated together as described above. The reaction was quenched by the addition of 10% TCA, and precipitated protein was pelleted by centrifugation and was then washed twice with 10% TCA. The protein pellet was then dissolved in 50 μl KOH (0.1 M) and heated to 65°C for 5 min. Trifluoroacetic acid (50%, 10 μl) was then added, the solution centrifuged to remove precipitated proteins, and the supernatant injected directly onto a HPLC equipped with a Vydac C18 reverse phase column using a linear gradient of 0–75% acetonitrile in water (containing 0.1% TFA) over 30 min. For coinjection experiments, non-radiolabeled standard compounds were mixed with the reaction eluate and then subjected to HPLC analysis. Dual on-line UV (at 220 nm) and radioactivity detectors (tuned for ^{35}S) were used to monitor the retentions of the standards and the radioactive enzymatic product, respectively.

A larger scale incubation (2 ml) contained 75 mM Tris (pH 7.5), 10 mM $MgCl_2$, 3 mM ATP, 5 mM TCEP, 0.5 mM L-Cys, 5 μM holo EpoB, 5 μM acetyl-*S*-ACP. The incubation proceeded for 30 min at 24°C for aminoacylation of EpoB with L-Cys prior to the addition of acetyl-*S*-ACP to initiate acetyl group transfer. The work up procedure noted above for radioactive incubations was followed. A peak having the same rt as the authentic methylthiazolyl carboxylate was collected from HPLC, lyophilized, and subjected to MALDI-TOF-MS analysis.

5.7.3. Investigation of the substrate specificity of Cy and Ox domains in EpoB

To investigate the substrate specificity of the Cy and Ox domains in EpoB, various acyl groups were loaded onto EpoA-ACP and tested as potential substrates for acyl transfer reaction across the PKS/NRPS interface and subsequent cyclization/dehydration and oxidation reactions. Some selected acyl-*S*-ACPs, acetoacetyl, propionyl, butyryl, isobutyryl, isovaleryl, hexanoyl, and benzoyl

were each mixed with [^{35}S]Cys-EpoB under similar conditions described above for the acetyl derivative. After 60 min incubation at 24°C, reactions were quenched, worked up, and analyzed by radio-HPLC as described above.

5.7.4. Chemical syntheses of HPLC standards

Since EpoB reaction is a single turnover event without the presence of the downstream modules, standards were prepared to elucidate the structure of the enzymatic products. There are three possible outcomes for every single acyl-EpoA-ACP substrate, i.e. the open form *N*-acyl-cysteine, the cyclized alkylthiazoline carboxylate, and the oxidized alkylthiazole carboxylate, depending on which functions of EpoB are operating. A general synthetic route was envisioned to produce all three structures along the way. A reported procedure to make 2-(2′-hydroxyphenyl)-2-thiazolines from 2-cyanophenol and L-cysteine [33] was modified to make *N*-acyl-cysteines, which can be readily converted to cyclized thiazoline, and further to the oxidized 2-alkylthiazolyl-4-carboxylic acids [34]. A representative case is described here for synthesizing *N*-acetyl-cysteine, 2-methylthiazolinyl-4-carboxylate, and 2-methylthiazolyl-4-carboxylate. L-cysteine hydrochloride (6.5 g, 41.2 mmol) was dissolved in 250 ml of degassed solution of MeOH: 0.1 M phosphate buffer. After the pH of the solution was adjusted to ~6, 20 ml of acetonitrile was added and the resulting solution was refluxed for 24 h under N_2 atmosphere. MeOH was removed under reduced pressure, the resulting solution was cooled in an ice-bath and the pH adjusted to ~2.5 with concentrated phosphoric acid. Then the product was extracted with CH_2Cl_2, dried ($MgSO_4$), and concentrated in vacuo to afford *N*-acetyl-cysteine as a white solid (65% yield, 4.4 g). ^{1}H NMR (DMSO, 200 MHz) δ 1.85 (3H, s), 2.38 (1H, t, J = 8.4 Hz, -SH), 2.63–2.86 (2H, m), 4.36 (1H, ddd, J = 7.8, 5.2, 4.8 Hz), 8.14 (1H, J = 7.8 Hz, N-H). MALDI-TOF calculated for $C_5H_9NO_3S$ (M+H$^+$) 164.03; found 164.04.

To a solution of dry benzene (20 ml) and TFA (1 ml) was added *N*-acetyl-cysteine (1 g) and the resulting mixture was heated to reflux under N_2 atmosphere overnight. The solvent was then removed by rotaevaporator, the residue was taken up in 200 ml CH_2Cl_2, the organic phase washed with 100 mM phosphate buffer (pH 2.5), dried ($MgSO_4$), and concentrated in vacuo to afford 2-methylthiazolinyl-4-carboxylic acid in quantitative yield. MALDI-TOF calculated for $C_5H_7NO_2S$ (M+H$^+$) 146.01; found 146.04. Oxidation was achieved by adding five equivalents of MnO_2 to the methylene chloride solution of the thiazoline compound and the reaction was allowed to proceed at 24°C with vigorous magnetic stirring. After 48 h, the mixture was filtered through a pad of Celite and concentrated in vacuo to give the oxidized product. When the purity was not satisfactory, the final product was further purified by HPLC with a prep C18 column using a linear gradient of acetonitrile in water (contained 0.1% TFA) from 0 to 70% over 30 min. ^{1}H NMR ($CDCl_3$, 200 MHz) δ 2.79 (3H, s), 8.15 (1H, s), 10.91 (1H, s, –COOH), MALDI-TOF calculated for $C_5H_5NO_2S$ (M+H$^+$) 144.01; found 144.01. NMR and mass data for 2-phenylthiazolyl-4-carboxylic acid, ^{1}H NMR ($CDCl_3$, 200 MHz) δ 7.51 (3H, m), 7.98 (2H, m), 8.26 (1H, s), MALDI-TOF $C_{10}H_7NO_2S$ (M+H$^+$) calculated 206.03; found 206.08. MALDI-TOF mass data of other thiazolyl carboxylic acids are listed below: 2-pentylthiazolyl-4-carboxylic acid, $C_9H_{13}NO_2S$ (M+H$^+$) calculated 200.07; found 199.99. 2-Isobutylthiazolyl-4-carboxylic acid, $C_8H_{11}NO_2S$ (M+H$^+$) calculated 186.06; found 186.96. 2-Isopropylthiazolyl-4-carboxylic acid, $C_7H_9NO_2S$ (M+H$^+$) calculated 172.04; found 186.01. 2-Propylthiazolyl-4-carboxylic acid, $C_7H_9NO_2S$ (M+H$^+$) calculated 172.04; found 186.00. 2-Ethytlthiazolyl-4-carboxylic acid, $C_6H_7NO_2S$ (M+H$^+$) calculated 158.03; found 157.97.

Acknowledgements

We thank KOSAN Biosciences for clones containing *epoA* and *EpoB* genes. We also thank Mr. Toby Zaleski of the Department Chemistry, Brown University for helpful discussions. S.O.E. acknowledges the financial support from the Irving S. Sigal Postdoctoral Fellowship. This work was supported by NIH grants GM20011 (for CTW) and GM22172 (for DEC).

References

[1] D.E. Cane, C.T. Walsh, C. Khosla, Harnessing the biosynthetic code: combinations, permutations, and mutations, Science 282 (1998) 63–68.

[2] D. Cane, Editor of a special issue of Chemical Review, polyketide and nonribosomal peptide synthetase, Chem. Rev. 97 (1997) 2463–2699.

[3] G. Weber, K. Schörgendorfer, E. Schneider-Scherzer, E. Leitner, The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame, Curr. Genet. 26 (1994) 120–125.

[4] J. Staunton, B. Wilkinson, Biosynthesis of erythromycin and rapamycin, Chem. Rev. 97 (1997) 2611–2629.

[5] L. Du, C. Sánchez, M. Chen, D.J. Edwards, B. Shen, The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase, Chem. Biol. 7 (2000) 623–642.

[6] L. Tang, S. Shah, L. Chung, J. Carney, L. Katz, C. Khosla, B. Julien, Cloning and heterologous expression of the epothilone gene cluster, Science 287 (2000) 640–642.

[7] I. Molnár, T. Schupp, M. Ono, R.E. Zirkle, M. Milnamow, B. Nowak-Thompson, N. Engel, C. Toupet, A. Stratmann, D.D. Cyr, J. Gorlach, J.M. Mayo, A. Hu, S. Goff, J. Schmid, J.M. Ligon, The biosynthetic cluster for the microtubule-stabilizing agents epothilones A and B from *Soragium cellulosum* So ce90, Chem. Biol. 7 (2000) 97–109.

[8] K. Reuter, M.R. Mofid, M.A. Marahiel, R. Ficner, Crystal structure of the surfactin synthetase-activating enzyme Sfp: a prototype of the 4′-phosphopantetheinyl transferase superfamily, EMBO J. 18 (1999) 6823–6831.

[9] P. Rowland, F.S. Nielsen, K.F. Jensen, S. Larsen, The crystal structure of the flavin containing enzyme dihydroorotate dehydrogenase A from *Lactococcus lactis*, Structure 5 (1997) 239–252.

[10] B. Silakowski, H.U. Schairer, H. Ehret, B. Kunze, S. Weinig, G. Nordsiek, P. Brandt, H. Blocker, G. Hofle, S. Beyer, R. Muller, New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1, J. Biol. Chem. 274 (1999) 37391–37399.

[11] R.S. Gokhale, S.Y. Tsuji, D.E. Cane, C. Khosla, Dissecting and exploiting intermodular communication in polyketide synthases, Science 284 (1999) 482–485.

[12] R.H. Lambalot, A.M. Gehring, R.S. Flugel, P. Zuber, M. LaCelle, M.A. Marahiel, R. Reid, C. Khosla, C.T. Walsh, A new enzyme

superfamily – the phosphopantetheinyl transferases, Chem. Biol. 3 (1996) 923–936.

[13] A. Gehring, R. Lambalot, K. Vogel, D. Drueckhammer, C.T. Walsh, Ability of *Streptomyces* spp acyl carrier proteins and Coenzyme A analogs to serve as substrates in vitro for *E. coli* holo-ACP synthase, Chem. Biol. 4 (1997) 17–24.

[14] L.E.N. Quadri, P.H. Weinreb, M. Lei, M.M. Nakano, P. Zuber, C.T. Walsh, Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases, Biochemistry 37 (1998) 1585–1595.

[15] P.J. Belshaw, C.T. Walsh, T. Stachelhaus, Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis, Science 284 (1999) 486–489.

[16] F. Rusnak, M. Sakaitani, D. Drueckhammer, J. Reichert, C.T. Walsh, Biosynthesis of the *Escherichia coli* siderophore enterobactin: sequence of the entF gene, expression and purification of EntF, and analysis of covalent phosphopantetheine, Biochemistry 30 (1991) 2916–2927.

[17] K. Gerth, N. Bedorf, G. Hofle, H. Irschik, H. Reichenbach, Epothilones A and B: antifungal and cytotoxic compounds from *Sorangium cellulosum* (Myxobacteria). Production, physico-chemical and biological properties, J. Antibiot. 49 (1996) 560–563.

[18] D. Bollag, P. McQueney, J. Zhu, O. Hensens, L. Koupal, J. Liesch, M. Goetz, E. Lazarides, C. Woods, Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action, Cancer Res. 55 (1995) 2325–2333.

[19] R.J. Kowalski, P. Giannakakou, E. Hamel, Activities of the microtubule-stabilizing agents epothilones A and B with purified tubulin and in cells resistant to paclitaxel (Taxol(R)), J. Biol. Chem. 272 (1997) 2534–2541.

[20] D.-S. Su, A. Ballog, D. Meng, P. Bertinato, S. Danishevsky, Y.H. Zheng, T.C. Chou, L. He, S. Horowitz, Structure–activity relationships of the epothilones and the first in vivo comparison with paclitaxel, Angew. Chem. Int. Ed. Engl. 36 (1997) 2093–2096.

[21] L.E.N. Quadri, T.A. Keating, H.M. Patel, C.T. Walsh, Assembly of the *Pseudomonas aeruginosa* nonribosomal peptide siderophore pyochelin: In vitro reconstitution of aryl-4,2-bis-thiazoline synthetase activity from PchD, PchE, and PchF, Biochemistry 38 (1999) 14941–14954.

[22] A.M. Gehring, E. DeMoll, J.D. Fetherston, I. Mori, G.F. Mayhew, F.R. Blattner, C.T. Walsh, R.D. Perry, Iron acquisition in plague: modular logic in enzymatic biogenesis of yersiniabactin by *Yersinia pestis*, Chem. Biol. 5 (1998) 573–586.

[23] C. Bisang, P.F. Long, J. Cortes, J. Westcott, J. Crosby, A.L. Matharu, R.J. Cox, T.J. Simpson, J. Staunton, P.F. Leadlay, A chain initiation factor common to both modular and aromatic polyketide synthases, Nature 401 (1999) 502–505.

[24] Z. Suo, C.C. Tseng, C.T. Walsh, Purification, priming, and catalytic acylation of carrier protein domains in the polyketide synthase and nonribosomal peptidyl synthetase modules of the HMWP1 subunit of yersiniabactin synthetase, Proc. Natl. Acad. Sci. USA 98 (2001) 99–104.

[25] L. Du, M. Chen, C. Sanchez, B. Shen, An oxidation domain in the BlmIII non-ribosomal peptide synthetase probably catalyzing thiazole formation in the biosynthesis of the anti-tumor drug bleomycin in *Streptomyces verticillus* ATCC15003, FEMS Microbiol. Lett. 189 (2000) 171–175.

[26] A.M. Gehring, I. Mori, R.D. Perry, C.T. Walsh, The nonribosomal peptide synthetase HMWP2 forms a thiazoline ring during biogenesis of yersiniabactin, an iron-chelating virulence factor of *Yersinia pestis*, Biochemistry 37 (1998) 11637–11650.

[27] T.A. Keating, D.A. Miller, C.T. Walsh, Expression, purification, and characterization of HMWP2, a 229 kDa, six domain protein subunit of yersiniabactin synthetase, Biochemistry 39 (2000) 4729–4739.

[28] Z. Suo, C.T. Walsh, D.A. Miller, Tandem heterocyclization activity of the multidomainal 230 kDa HMWP2 Subunit of *Yersinia pestis* yersiniabactin synthetase: interaction of the 1–1382 and 1383–2035 fragments, Biochemistry 38 (1999) 14023–14035.

[29] T. Osumi, T. Hashimoto, Acyl-CoA oxidase of rat liver: a new enzyme for fatty acid oxidation, Biochem. Biophys. Res. Commun. 83 (1978) 479–485.

[30] T. Schwecke, J.F. Aparicio, I. Molnar, A. Koenig, L.E. Khaw, S.F. Haydock, M. Oliynyk, P. Caffrey, J. Cortes, J.B. Lester, G.A. Boehm, J. Staunton, P.F. Leadlay, The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin, Proc. Natl. Acad. Sci. USA 92 (1995) 7839–7843.

[31] K.C. Nicolaou, F. Roschangar, D. Vourloumis, Chemical biology of epothilones, Angew. Chem. Int. Ed. 37 (1998) 2014–2045.

[32] B. Julien, S. Shah, R. Ziermann, R. Goldman, L. Katz, C. Khosla, Isolation and characterization of the epothilone biosynthetic gene cluster from *Sorangium cellulosum*, Gene 249 (2000) 153–160.

[33] H.R. Haveyda, V. Karunaratne, C. Orvig, Design and synthesis of multidentate 2-(2′-hydroxyphenyl)-2-thiazolines for biomedical application, Tetrahedron 48 (1992) 5219–5226.

[34] J.A. Sowinski, P. Toogood, Synthesis of an enantiomerically pure serine-derived thiazole, J. Org. Chem. 61 (1996) 7671–7676.