Chemistry & Biology, Vol. 11, 1533–1542, November, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.chembiol.2004.08.017

Hybrid Nonribosomal Peptide-Polyketide Interfaces in Epothilone Biosynthesis: Minimal Requirements at N and C Termini of EpoB for Elongation

Fei Liu, Sylvie Garneau, and Christopher T. Walsh*
Department of Biological Chemistry
and Molecular Pharmacology
Harvard Medical School
240 Longwood Avenue
Boston, Massachusetts 02115

Summary

Epothilone (Epo) D, an antitumor agent currently in clinical trials, is a hybrid natural product produced by the combined action of nonribosomal peptide synthetases (NRPS) and polyketide synthases (PKS). In the epothilone biosynthetic pathway, EpoB, a 165 kDa NRPS is inserted into an otherwise entirely PKS assembly line, forming two hybrid NRPS-PKS interfaces. In light of the terminal linker effect previously identified in PKS, the N- and C-terminal sequences of EpoB were examined for their roles in propagating the incipient natural product. Eight amino acid residues at EpoB C terminus, in which six are positively charged, were found to be a key component of the C-terminal linker effect. A minimal sequence of 56 residues at EpoB N terminus was required for elongating the acetyl group from the acyl carrier protein (ACP) of EpoA to form methylthiazolyl-*S*-EpoB.

Introduction

Nonribosomal peptide synthethases (NRPS) and polyketide synthases (PKS) are multimodular enzymes that synthesize secondary metabolites from small monomeric building blocks in assembly-line format [1–4]. In both cases, each module, equipped with multiple catalytic domains, is responsible for incorporating one building block, receiving the growing acyl chain from an upstream module and elongating the incipient natural product for transfer to the next module down the assembly line. The fully elongated natural product after release from assembly lines can be further modified, decorated, or conformationally constrained by a variety of tailoring enzymes acting either in *trans* or in *cis*. The modular nature of these pathways has inspired the concept of combinatorial biosynthesis for generating "unnatural" natural products for therapeutic applications [5, 6]. Hybrid NRP/PK natural products, produced by pathways containing both types of enzymes, further expand the perspective of combinatorial biosynthesis. Representative examples of hybrid NRP-PK natural products with medicinal importance include rapamycin (immunosuppressant), bleomycin (anticancer), epothilone (microtubule binder), yersiniabactin (siderophore), and leinamycin (antibiotic) [7–14].

Questions concerning the biosynthesis of hybrid NRP-PK natural products are of current interest as their answers pertain to biosynthetic engineering efforts [15]. While both rely on thio-templated mechanism for monomer activation and acyl chain growth, these two types of pathways use different classes of building blocks, proteinogenic or modified amino acids for NRPS, and acyl-CoAs for PKS. The chemistry that drives iterative elongation is thus fundamentally different in these two systems. In an NRPS pathway, the elongation involves carbon-nitrogen bond formation between the carbonyl group of the upstream acyl donor and the amino group of the downstream amino-acyl acceptor, both attached to peptide carrier protein (PCP) domains. In a PKS pathway, the upstream and downstream acyl groups, attached to acyl carrier protein (ACP) domains, couple through Claisen-type carbon-carbon bond formation. In addition, these two systems have shown variant protein-protein interface characteristics. PKS proteins are dimeric [16, 17], in contrast to S proteins predominantly found as monomers [18–22].

Currently unknown are the details of molecular recognition events that allow hybrid NRP-PK pathways to functionally accommodate different assembly-line chemical logics with divergence in quaternary organization. Elucidation of molecular interactions that govern recognition at these interfaces therefore will be a prerequisite to sustainable pathway engineering strategies. Information gathered on NRP and PK systems so far provides invaluable leads in studying hybrid NRP-PK systems. One of the most significant discoveries is the shuffling of PKS subunits facilitated by N- and C-terminal linkers [23–27]. A similar linker hypothesis has been proposed for the hybrid NRP-PK systems based on primary sequence analysis [28]. The concept of portable docking units for proteins otherwise unrecognizable to each other is very attractive to engineering pathway reconstruction in hybrid -PK biosynthetic machineries as an approach particularly relevant to drug discovery [29–32].

The epothilone biosynthetic pathway, whose PKS/NRPS/PKS hybrid interfaces (EpoA/B/C subunits) have recently been reconstituted [33, 34], emerges as a good system for investigating the linker hypothesis (Figure 1). The two hybrid interfaces are formed by the initial three protein subunits of the pathway, EpoA, EpoB, and EpoC. EpoB, inserted between two PKS subunits, is the only NRPS protein in the otherwise entirely PKS pathway. EpoB is responsible for accepting the upstream acetyl group presented by EpoA-ACP and, as cysteinyl-*S*-EpoB, elongating it to the methylthiazolyl intermediate. This transformation involves multiple chemical steps controlled by the cyclization (Cy), adenylation (A), and oxidation (Ox) domains of EpoB. The A domain is responsible for activating L-cysteine (L-Cys) as an aminoacyl-AMP. The thiol group of the phosphopantetheinyl prosthetic arm of *holo*-PCP domain of EpoB reacts with Cys-AMP to form Cys-*S*-EpoB-PCP as the competent downstream amino-acyl group to couple with the incoming acetyl group. The resulting *N*-acetyl-Cys intermediate is cyclized and dehydrated to form the methylthiazolinyl ring by the action of the Cy domain before being oxidized

*Correspondence: christopher_walsh@hms.harvard.edu

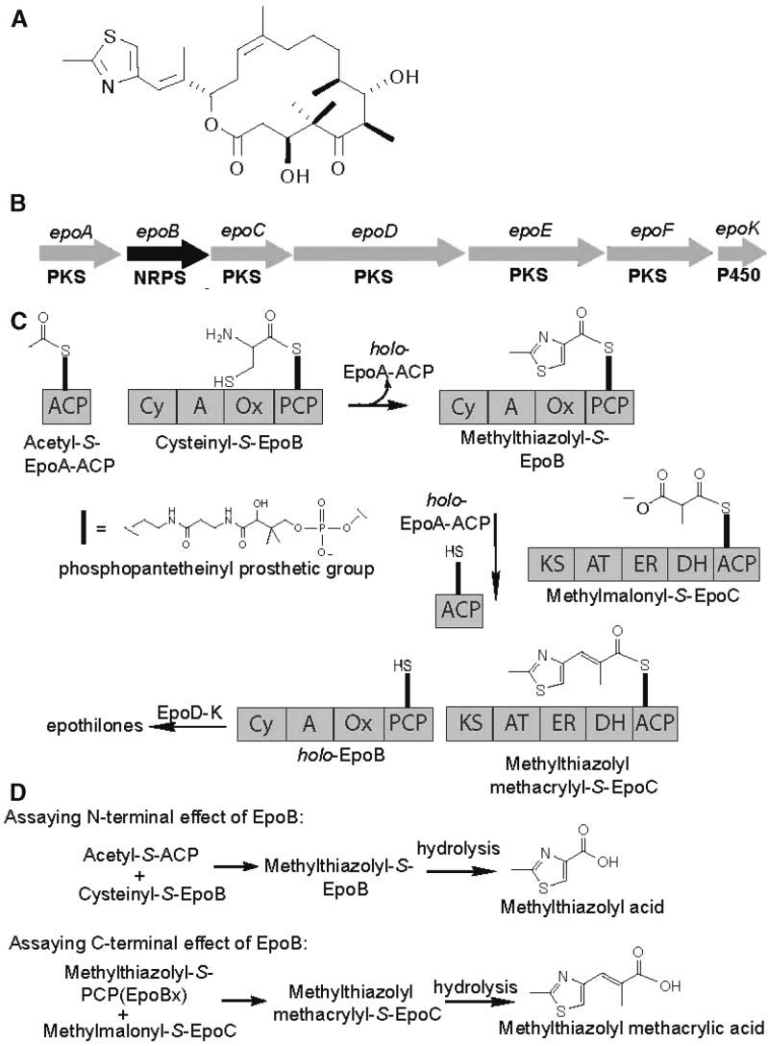


Figure 1. The Epothilone Biosynthesis

(A) Structure of epothilone D, a representative member of the epothilone family.

(B) Gene cluster of epothilone biosynthesis.

(C) The initial three enzyme subunits that form the PKS/NRPS/PKS hybrid interfaces of the epothilone biosynthetic pathway.

(D) Acyl transfer assays for examining the terminal effects of EpoB.

by the Ox domain with cofactor FMN to form methylthiazolyl-*S*-EpoB-PCP. EpoC, the downstream PKS partner, then further elongates the final EpoB intermediate to methylthiazolyl methacrylyl-*S*-EpoC (Figure 1C).

It has been proposed that the role of the terminal sequences of EpoB may very well resemble functionally the terminal linkers of PKS [28]. In this view, the pairwise interaction of the N- or C-terminal region of EpoB with the corresponding terminal region of its up- or downstream PKS partner would enable the productive transfer of the acyl intermediate down the assembly line. We have previously reported that a C-terminal linker on the PKS subunit EpoA in the form of the last 34 amino acid residues is portable, enabling noncognate ACP domains to interface with EpoB [35]. Analogously, the 27 C-terminal residues of EpoB were portable to enable EpoC recognition [35].

In this work we have examined the existence of an N-terminal linker in the NRPS subunit EpoB, analogous to N-terminal PKS linkers, via deletions of 30 or 56 of the N-terminal residues and assay of the consequences in acetyl transfer from the EpoA-ACP as donor. We have also examined the role of six basic residues at the C terminus of EpoB in constituting a functional EpoB/EpoC interface. To this end we used noncognate PCPs from the coumermycin or yersiniabactin biosynthetic clusters as the core domains on which the EpoB C-terminal residues were grafted. The assay was acyl transfer from the *holo* form of these constructs to methylmalonyl-*S*-EpoC (Figure 1D).

Results

Assay Design

To examine the EpoA/EpoB intersubunit communication as reflected by productive acyl transfer, terminally truncated forms of EpoA-ACP and EpoB were required. For EpoA-ACP, the C-terminal truncation needed to be suitable for the stability of the protein and its ability to be acylated as acetyl-*S*-ACP. The N-terminal deletion of EpoB, on the other hand, should not alter the loading of cysteine onto EpoB-PCP (Table 1). The acyl transfer reaction between a modified donor and acceptor pair was monitored by the formation of methylthiazolyl acid, from hydrolyzed methylthiazolyl-*S*-EpoB, under single-turnover conditions (Figure 1D).

Table 1. Terminal Sequences of the Mutant and Hybrid Proteins

EpoA-ACP (WT):
. . . *S*LMAVELRNRIEASLKLKLSTTFLSTSPNIALLTQNLLDA
LDALAT*ALSLERVAAENLRAGVQSDFVSSGADQDWEIIAL*

EpoB (WT):
TINQLLNELEHQGVKLAADGERLQIQAPKNALNPNLLARISEHKSTILTMLRQRLP
AESIVPAPAERHVPFPL
TDIQGS…(Cy-A-Ox-PCP)…
RRDSKDLEQRPNMQDRVEA *RRKGRRRS*

Constructs	Sequence Modification from Wild-Type
EpoB(N-30)	ALNPNLLARISEHKSTILTMLRQRLPAESIVPAPAERHVPFPLTDIQGS…
EpoB(N-56)	AESIVPAPAERHVPFPLTDIQGS…
EpoA-ACP(C-10)	…LDALATALSLERVAAENLRAGVQSDFVSSG
EpoA-ACP(C-20)	…LDALATALSLERVAAENLRA
EpoA-ACP(C-34)	…LDALAT
Ybt-PCP2(EpoB-8)	…RRDSKDLEQRPNMQDRVEA
CouN5(EpoB)	…RRDSKDLEQRPNMQDRVEARRKGRRRS
CouN5(EpoB-8)	…RRDSKDLEQRPNMQDRVEA
CouN5(EpoB6G)	…RRDSKDLEQRPNMQDRVEAGGGGGGGS

The phosphopantetheinylated serine residue is italicized. The proposed linker regions are bold. The regions found in this study to be essential for elongation are italicized. Residues underlined indicate potential helical regions as predicted by online PSIPRED prediction programs [46–47].

Comparably, to analyze recognition at the EpoB/EpoC interface by transfer of the acyl group on EpoB to methylmalonyl-*S*-EpoC, it was necessary to generate those requisite acyl-*S*-proteins and to generate C-terminally mutated EpoB variants (Table 1). For these EpoB/EpoC acyl transfer assays, the C terminus of EpoB-PCP was grafted onto noncognate PCP core domains to minimize recognition from elements that might be present at sites within EpoB other than at the C terminus. One was the PCP2 domain from the HMWP2 protein of yersiniabactin synthetase [37, 38], used by us previously [35] to study EpoB/EpoC communication, while the other was a free-standing PCP subunit in the coumermycin biosynthetic pathway, CouN5 [36]. CouN5 is stable and soluble when heterologously expressed in and purified from *E. coli*. It has only 12% sequence similarity to EpoB-PCP and lacks the stretch of basic residues seen at the C terminus of EpoB.

Construction of the Mutant and Hybrid Proteins

Possible boundary definitions for the N linker of EpoB were determined as guided by secondary structural predictions. Two predicted helical regions, one from the first residue to the 30th and the other from the 33rd to the 55th, were found. Break points, one after the 30th amino acid from the N terminus and the other the 56th, were therefore chosen to construct N-30 and N-56 mutant forms of EpoB. Both mutants were constructed using PCR techniques as N-terminally hexa-histidine (His_6)-tagged proteins in pET-28b vectors. The overexpression conditions of both mutants in BL21(DE3) *E. coli* cells followed that of the EpoB wild-type [33] and returned comparable yields at 2 mg/l for both.

To examine the effect of C-terminal deletions in EpoA-ACP, three C-terminally truncated mutants of EpoA-ACP were also prepared analogously. Two break points were chosen following the structural predictions that suggested two helices formed by the last 20 amino acids at the C terminus, which gave rise to EpoA-ACP(C-10) and EpoA-ACP(C-20). The third break point was chosen at the position after the 37th amino acid C-terminal to the conserved phosphopantetheinylation serine site, following a known boundary definition for a minimal carrier protein core to provide EpoA-ACP(C-34) [40]. All three mutants again overexpressed successfully in BL21(DE3) with yields in the range of 2–3 mg/l.

Wild-type CouN5 was subcloned from a cosmid by PCR techniques and overexpressed well in BL21(DE3) at 10 mg/l. To generate a CouN5-EpoB C-terminal hybrid, a prior boundary definition [35] of a portable C-linker region of EpoB, which includes the last 27 amino acids, was used to construct the CouN5(EpoB) hybrid. CouN5, with its C terminus shorter than that of EpoB by 18 residues, was treated as a naturally linkerless carrier protein. The EpoB C-linker was appended to the last amino acid by splice overhang extension [39] to produce the CouN5 (EpoB) hybrid. This protein was overexpressed in BL21 (DE3) at 10 mg/l. The truncation mutants removing the last eight residues, namely CouN5(EpoB-8) and YbtPC-P2(EpoB-8), were cloned using standard PCR techniques and overexpressed under similar conditions, and the yields were comparable at 10 mg/l. The length-invariant but charge-removed mutant CouN5(EpoB6G) was cloned in two sequential steps of mutating three charged residues at a time to glycines. The loss of charge again was not adverse to the overexpression with the yield being 12 mg/l (Figure 2).

All carrier proteins were constructed as N-terminally His_6-tagged proteins in pET-28 vectors and expressed in their *apo*-forms, since their heterologous overexpression in *E. coli* did not result in significant modification by endogenous phosphopantetheinyl transferases (PPTase) as determined by MALDI-TOF mass spectrometry (Table 2).

Sfp-Catalyzed Acylation of the Hybrid Apo-Carrier Proteins with Acyl-CoAs

For studying either the EpoA/EpoB or EpoB/EpoC linker interface, the upstream acyl donor is consistently a car-

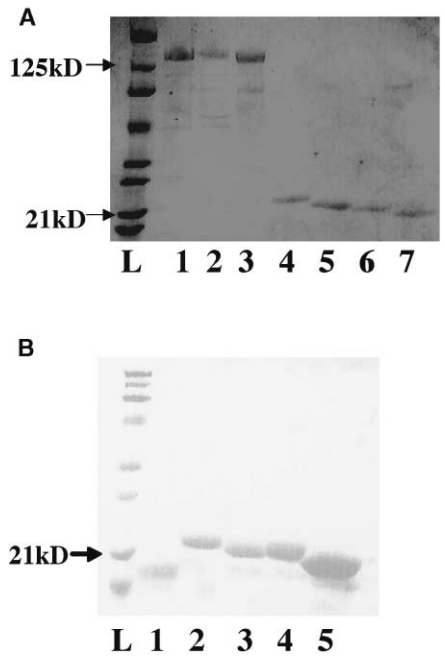


Figure 2. Purification of Carrier Proteins and EpoA and EpoB Mutants

(A) SDS-PAGE gel of EpoA and EpoB deletion mutants. L = ladder; 1 = EpoB WT; 2 = EpoB(N-30); 3 = EpoB(N-56); 4 = EpoA-ACP; 5 = EpoA-ACP(C-10); 6 = EpoA-ACP(C-20); 7 = EpoA-ACP(C-34). (B) SDS-PAGE gel of the carrier proteins with various C-terminal regions imported from EpoB. L = ladder; 1 = CouN5 WT; 2 = CouN5(EpoB); 3 = CouN5(EpoB-8); 4 = CouN5(EpoB6G); 5 = Ybt-PCP2(EpoB-8).

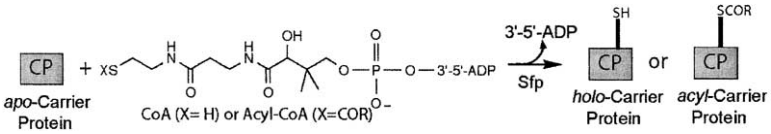
rier protein with an acyl group covalently attached through the phosphopantetheinyl prosthetic group. Normally, a carrier protein domain needs to be first primed into its *holo*-form by a PPTase that catalyzes covalent attachment of the prosthetic group through a phosphodiester linkage to the conserved serine site on the carrier protein. An additional catalytic domain then further converts the *holo*-form into the *acyl*-form. Sfp, isolated from *B. subtillis*, is a nonspecific PPTase found to have broad substrate specificity for both carrier proteins and acyl-CoAs and thus was used to directly modify the carrier proteins from the *apo*-form to the *acyl*-form [41]. The recognition of the carrier protein by Sfp has been shown to depend on the second helix of the four-helix bundle that typically defines the minimal core structure of a carrier protein [42]. Although it is not known structurally how the C terminus of a PCP from an NRPS would interact with the four-helix bundle core, recent characterizations of the C linker region of an ACP from the DEBS PKS pathway [43] and PCP-Sfp [44] recognition surface strongly suggest that the loading is unlikely to be influenced by the C-terminal linker. Indeed, Sfp successfully converted the *apo*-proteins to their *acyl*-forms, regardless of the C-terminal variations, as shown by MALDI-TOF mass spectrometry (Table 2). Specifically, the PCPs were loaded with either the methylthiazolyl-phosphopantetheinyl or picolinyl-phosphopantetheinyl group, resulting in an increase in mass by 464 or 470 units, respectively. The ACPs, after being loaded with the acetyl-phosphopantetheinyl group, provided an increase of 382 units in mass.

A-Domain Activity Assessment of the EpoB Mutants

The first domain with recognized function from the N terminus of EpoB is the cyclization domain, followed by the A domain that activates L-Cys to L-Cys-AMP to form cysteinyl-*S*-EpoB-PCP. This domain activity is typically assessed by the ATP-$^{32}PP_i$ exchange assay [45], which measures the formation of ^{32}P-ATP from aminoacyl-AMP with $^{32}PP_i$. It was necessary to examine if the removal of part of the N terminus of EpoB would change the activity of the A domain that is in proximity. A compromised A-domain activity would obscure the interpretations of the elongation efficiency at the EpoA/EpoB interface.

To compare the A-domain activity of EpoB-WT with the two N-terminal mutants, all three proteins were first modified with Sfp to convert from the *apo*-form to the *holo*-form using coenzyme A. The *holo*-EpoBs were then

Table 2. MALDI-TOF Mass Analysis of Carrier Proteins Loaded with Acyl Groups



Carrier Proteins (Acyl Group)	Calculated $[MH]^+$		Observed $[MH]^+$	
	Apo-	Acyl-	Apo-	Acyl-
EpoA-ACP(-10)(acetyl)	15,370	15,752	15,395	15,790
EpoA-ACP(-20) (acetyl)	14,406	14,788	14,434	14,822
EpoA-ACP(-34) (acetyl)	12,917	13,299	13,299	13,311
Ybt-PCP2(EpoB-8) (methylthiazolyl)	14,088	14,552	14,124	14,585
CouN5 (methylthiazolyl)	11,755	12,219	11,744	12,203
CouN5(EpoB) (methylthiazolyl)	15,576	16,040	15,556	16,034
CouN5(EpoB-8) (methylthiazolyl)	14,522	14,986	14,515	14,982
CouN5(EpoB6G) (methylthiazolyl)	15,007	15,471	15,031	15,504
CouN5 (picolinyl)	11,755	12,225	11,755	12,213
CouN5(EpoB) (picolinyl)	15,576	16,046	15,556	16,044

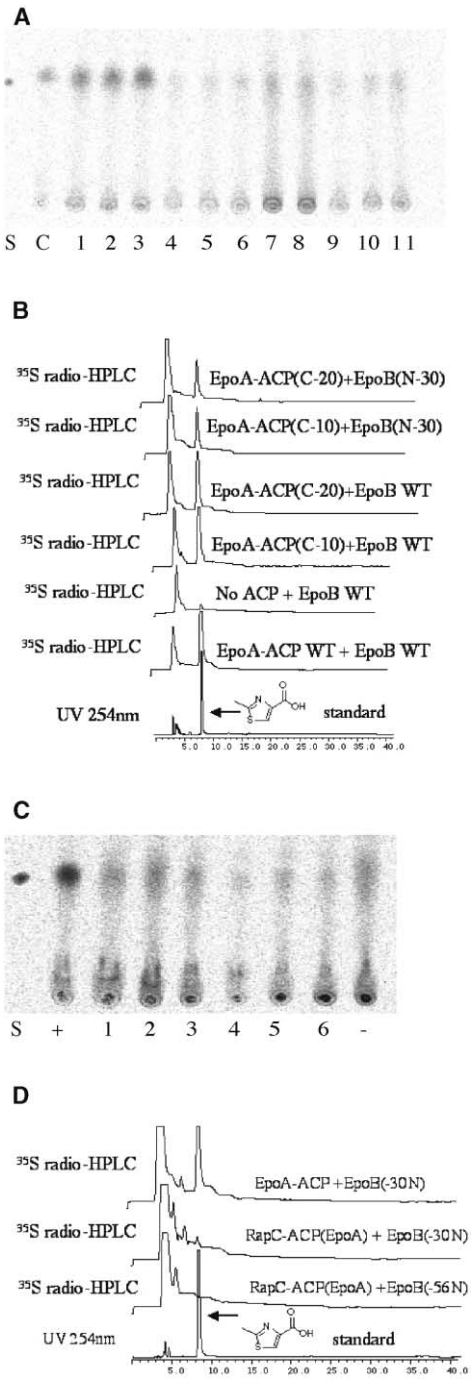


Figure 3. The N-Terminal Region of EpoB

(A) Time courses (three time points of 30, 90, and 250 time points) of acetyl transfer reactions between EpoA-ACP(C-34) and EpoB(N-56). S = methylthiazole acid standard; C = Cospotting of standard and EpoA-ACP WT and EpoB WT transfer reaction after 250 min. Lanes 1–3 are transfer reactions between EpoA-ACP WT and EpoB WT. Lanes 4–6 are transfer reactions between EpoA-ACP WT and EpoB(N-56). Lanes 9–11 are transfer reactions between EpoA-ACP (C-34) and EpoB(N-56). Lanes 7 and 8 are negative control reactions where the carrier protein was absent in the transfer reaction in two different acetyl-CoA concentrations of 50 and 200 μM, respectively.

(B) One time point assessment (120 min) of acetyl transfer reaction using EpoA-ACP(C-10), EpoA-ACP(C-20), and EpoB (N-30). The standard is methylthiazole acid.

(C) Time courses (3 time points of 30, 90, and 250 time points) of acetyl transfer reactions between EpoA-ACP(C-34) and EpoB, compared to that between RapC-ACP and EpoB (N-56). S = methylthiazole acid standard. The positive control is the acetyl transfer reaction between EpoA-ACP WT and EpoB WT after 250 min. Lanes 1–3 are transfer reactions between EpoA-ACP(C-34) and EpoB-WT. Lanes 4–6 are acetyl transfer reactions between RapC-ACP WT and EpoB(N-56). Lane 7 is the negative control reaction where the carrier protein was absent in the transfer reaction with 50 μM of acetyl-CoA.

(D) One time point assessment (120 min) of acetyl transfer reaction using RapC-ACP(EpoA), EpoB(N-30) and EpoB(N-56). The standard is methylthiazole acid.

incubated with L-Cys in the presence of ATP and $^{32}PP_i$ for five minutes. The ^{32}P-ATP formed was isolated and subjected to radioactivity counting. The measured k_{cat} for the A domain activity of wild-type EpoB was 20/min. The two mutants, one at 157 kDa (N-30) and the other at 154 kDa (N-56), both displayed robust A-domain activity comparable to that of EpoB wild-type with k_{cat} values of 33/min and 45/min, respectively.

The First 56 Amino Acids at the N Terminus of EpoB

To test if there is a sequence requirement at the N terminus of EpoB for the functional constitution of the EpoA/EpoB interface, the elongation efficiencies of EpoB (N-56) were examined in the presence of either EpoA-ACP WT or EpoA-ACP(C-34). It has previously been shown that the last thirty-eight amino acid residues of EpoA-ACP is a portable recognition sequence that enables productive elongation between a noncognate ACP and EpoB to form methylthiazolyl-*S*-EpoB [35]. Elongation efficiencies of three interfaces, acetyl-*S*-EpoA-ACP WT/cysteinyl-*S*-EpoB WT, acetyl-*S*-EpoA-ACP WT/cysteinyl-*S*-EpoB(N-56), and acetyl-*S*-EpoA-ACP(C-34)/cysteinyl-*S*-EpoB(N-56), were assessed by measuring the formation of ^{35}S-labeled methylthiazole carboxylic acid hydrolyzed from [^{35}S]-methylthiazolyl-*S*-EpoB after three time points of 30, 90, and 250 min. Reactions were also repeated at two concentrations of acetyl-CoA of 50 μM and 200 μM in the absence of the carrier protein as the background for carrier protein independent product formation. The wild-type interface provided the anticipated elongation product over the time course (Figure 3A). Truncating the first fifty-six amino acids of EpoB resulted in only background product formation. The same effect was confirmed by the result from the interface formed by acetyl-*S*-EpoA-ACP(C-34) and cysteinyl-*S*-EpoB(N-56).

Partial terminal truncation mutants of the EpoA/EpoB interface, EpoA-ACP(C-10) and EpoA-ACP(C-20), and EpoB(N-30), were also constructed. The elongation assessment was repeated under identical conditions for these mutants for 120 min. In the case with EpoB WT, deleting the first 10 or 20 amino acids from the C terminus of EpoA-ACP provided nearly identical transfer efficiencies to that with wild-type EpoA-ACP. Removing the N-terminal 30 amino acids of EpoB, on the other hand, resulted in substantial loss of elongation but not yet reduced to the background level (Figure 3B).

To confirm that the N terminus EpoB is not the only

factor involved in productive elongation at the EpoA/EpoB interface, a C-linker removed truncation of EpoA-ACP, EpoA-ACP(C-34), was examined in the presence of the EpoB wild-type. The elongation efficiency of this interface was reduced to the background level again. This is consistent with the previous report showing that a noncognate interface formed between RapC-ACP, an ACP from the rapamycin pathway, with EpoB requires the presence of the EpoA-ACP C-linker [35]. To further examine if the N terminus of EpoB is merely a steric-gating element, acetyl-*S*-RapC-ACP was incubated with cysteinyl-*S*-EpoB(N-56) to test if the lack of the N terminus would permit an otherwise rejected upstream donor to regain recognition. This noncognate interface again did not provide protein-dependent methylthiazole carboxylic acid formed as indicated by ^{35}S radio-HPLC (Figure 3C).

The studies with N-truncated EpoB were conducted with another form of ACP acyl donor: a hybrid between the RapC ACP domain and the last 38 residues of EpoA [35]. In this case, RapC can only interface with EpoB when it has the EpoA C-terminal linker attached. Incubations were repeated with this potential acetyl donor to Cys-*S*-EpoB. Cysteinyl-*S*-EpoB(N-30), while able to productively interface with acetyl-*S*-EpoA-ACP, showed severe reduction in elongation efficiency when reacted with acetyl-*S*-RapC-ACP(EpoA). In the case of Cysteinyl-*S*-EpoB(N-56), elongation was entirely lost with acetyl-*S*-RapC-ACP(EpoA) (Figure 3D), indicating the elongation effect by importing the C-linker of EpoA requires the presence of the N terminus of EpoB.

The Basic Charge Cluster at the C Terminus of EpoB

To test if the basic residues at the C terminus of EpoB are sufficient and necessary to interface with EpoB's downstream acceptor EpoC, CouN5, a noncognate carrier protein, was chosen to minimize contributions from non-linker regions. CouN5 was loaded with the methylthiazolyl group in the presence of Sfp and methylthiazolyl-CoA [35], yielding methylthiazolyl-*S*-CouN5 as the upstream acyl donors for EpoC. The formation of the elongated product methylthiazolyl methacrylyl-*S*-EpoC, as detected by the formation of ^{14}C-labeled methylthiazolyl methyl acrylic acid after hydrolyzing the small molecule off EpoC, was monitored for several hours. Product formation was not observed over the time course, indicating that CouN5 by itself is not capable of interacting with EpoC productively.

The twenty-seven amino acid EpoB C-linker was then appended to the C terminus of CouN5, providing a noncognate hybrid carrier protein CouN5(EpoB). This hybrid protein was again successfully loaded with the methylthiazolyl group and combined with methylmalonyl-*S*-EpoC. The formation of methylthiazolyl methacrylyl-*S*-EpoC was observed shortly after initiation of the transfer reaction. To show that the acyl group being transferred does not influence the linker-dependent elongation, another heterocyclic acyl group was examined as well. The C-terminally modified picolinyl-*S*-CouN5 acyl donors were prepared by using Sfp and picolinyl-CoA and subjected to identical transfer reactions. In agreement with transferring the methylthiazolyl group, the formation of picolinyl methacrylyl-*S*-EpoC required the C-terminal linker of EpoB (Figure 4A).

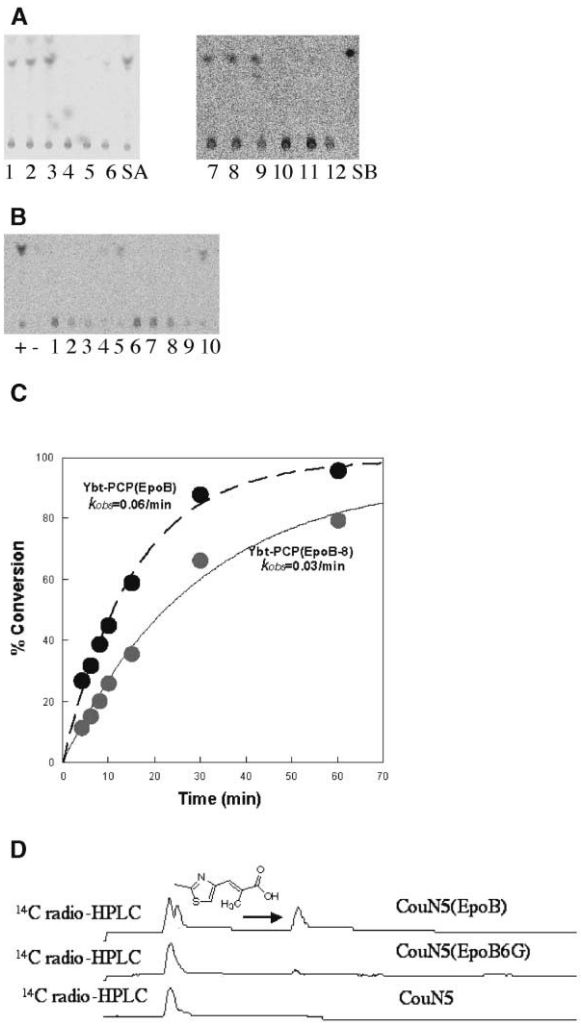


Figure 4. The Charge Cluster of the C-Terminal Linker of EpoB

(A) Time courses of CouN5 and CouN5(EpoB) transferring methylthiazolyl and picolinyl acyl groups to EpoC over three time points (20, 60, and 180 min). Lanes 1–3, methylthiazolyl-*S*-CouN5(EpoB) and methylmalonyl-*S*-EpoC; lanes 4–6, methylthiazolyl-*S*-CouN5 and methylmalonyl-*S*-EpoC; lanes 7–9, picolinyl-*S*-CouN5(EpoB) and methylmalonyl-*S*-EpoC; lanes 10–12, picoliny-*S*-CouN5 and methylmalonyl-*S*-EpoC. SA = methylthiazolyl methyl acrylic acid standard; SB = picolinyl methyl acrylic acid standard.

(B) Time courses of CouN5, CouN5(EpoB), and CouN5(EpoB-8) transferring methylthiazolyl group to EpoC (five time points of 30, 90, 150, 300, and 480 min). The positive control was the transfer reaction between CouN5(EpoB) and EpoC for 480 min. The negative control was the transfer reaction in the absence of the carrier protein. Lanes 1–5: methylthiazolyl-*S*-CouN5(EpoB-8) and methylmalonyl-*S*-EpoC; lanes 6–10: methylthiazolyl-*S*-CouN5 and methylmalonyl-*S*-EpoC.

(C) The charge-cluster truncated mutant, Ybt-PCP2(EpoB-8), of previously reported Ybt-PCP2(EpoB) hybrid carrier protein reduced the pseudo-first order rate of product formation by 2-fold. Carrier proteins were at 3 μM and EpoC 15 μM.

(D) Comparison of elongation efficiencies after 60 min of three constructs: CouN5(EpoB), CouN5(EpoB6G), and CouN5.

Truncation mutagenesis was applied to CouN5(EpoB) to give CouN5(EpoB-8), removing the last eight amino acids. The new construct was loaded with methylthiazolyl group and subjected to the transfer reaction to EpoC. In the absence of the charge cluster in those last eight residues, the methylthiazolyl methacrylyl-*S*-EpoC formation was abolished (Figure 4B). To confirm the effect of the charge truncation without the possibility of false negative outcome, Ybt-PCP2(EpoB-8), which is a charge-truncation mutant using the Ybt-PCP2 core domain, was also examined. It has been shown that Ybt-PCP2(NL), a C-linker truncated version of the native PCP2, can still interface productively with EpoC for elongation, although at half of the efficiency of PCP2(EpoB), a construct of PCP2 that contains the EpoB C-linker [35]. Ybt-PCP2(EpoB-8) exhibited a pseudo-first order rate of 0.03/min in forming methylthiazolyl methacrylyl-*S*-EpoC, which is half of that observed for PCP2(EpoB) at 0.06/min, indicating that the effect of charge truncation on elongation is equivalent to that of linker removal. Although only 20% of all loaded methylmalonyl-EpoC was converted to methylthiazolyl methacrylyl-*S*-EpoC due to competing hydrolysis of methylmalonyl-EpoC during the acyl transfer reaction, the captured methylmalonyl-EpoC was nearly all converted to the stable methylthiazolyl methacrylyl-*S*-EpoC product (Figure 4C).

To ensure that the effect of charges is not due to the shortening of the linker length, a length-invariant construct CouN5(EpoB6G), in which the six basic residues were uniformly mutated to glycine, was subjected to the acyl transfer and elongation reaction to EpoC. The removal of the charges rendered CouN5(EpoB6G) a poor acyl donor to EpoC (Figure 4D).

Discussion

The findings in this report establish the crucial role and minimal sequence definition of the terminal regions of EpoB at the PKS/NRPS/PKS hybrid interfaces formed by EpoA, EpoB, and EpoC, the initial three subunits of the epothilone pathway. As an initial investigation on the intersubunit linkers of an NRPS module in a hybrid environment, it begins to address the important question of how protein-protein interactions must occur to ensure elongation as the assembly line switches first from a PKS logic to an NRPS one, and then back again to the PKS logic. PKS subunits are dimeric while NRPS ones are generally monomeric, a discrepancy that makes hybrid NRPS and PKS interfaces more difficult to decipher. The previously defined linker effect of purely PKS systems [24] served as an initial guide to the existence of intersubunit linkers in hybrid NRPS subunits. The experimental evidence presented here for the crucial roles of the N-terminal 56 amino acids and C-terminal 8 amino acids of EpoB in enabling a functional PKS/NRPS/PKS hybrid line corroborates this idea at the primary sequence level.

For the EpoB/EpoC (S/PKS) interface, we have established previously that the C-terminal 27 amino acids can serve as a portable C-linker [35]. The generality of this effect was supported in this report by the example of CouN5, a PCP that required this C-linker of EpoB to elongate either methylthiazolyl or picolinyl acyl groups to EpoC. This noncognate hybrid interface also allowed for identification of the basic residues clustered in the last 8 amino acids as a key factor for enabling the noncognate elongation. Truncation or removal of the charge residues negates the positive effect of the imported C-linker. Since an excised EpoB-PCP itself could not be accessed as a stable domain, Ybt-PCP2(EpoB) has been used and reported as a homologous surrogate. The C-linker of EpoB, while sufficient in enabling noncognate hybrid interfaces for PCPs nonhomologous to EpoB-PCP, is not the sole factor in constituting the native EpoB/EpoC interface, as the linkerless Ybt-PCP2(NL) could still enable acyl transfer and chain elongation to EpoC [35], albeit less efficiently. However, the findings from the Ybt-PCP2(EpoB-8) indicated that the enhancing effect of the linker requires the charges, confirming again the observation made from the case of CouN5. This charge cluster-controlled linker effect is different from that reported for the PKS linkers where interactions of both hydrophilic and hydrophobic nature are spread through extended coiled-coiled surfaces formed by the linkers [27]. Potential C-linkers at other S/PKS interfaces, as seen in the bleomycin and yersiniabactin pathways, also contain basic charge clusters [28]. The generality of this type of charge-cluster effect, if established, will help formulating future linker engineering or evolution strategies for hybrid NRPS/PKS systems.

For the EpoA/EpoB (PKS/NRPS) interface, the N-terminal 56 amino acids were found to be required for elongating the natural acetyl group from EpoA-ACP to EpoB. This part of the N terminus seems modular, since the removal of this region did not alter the A-domain activity as judged by the ATP-PP_i exchange assay. Given a previously reported portable C-linker of EpoA-ACP containing 34 amino acids, it is now evident that, at the EpoA/EpoB interface, the intersubunit terminal regions show comparability in length and modularity to the linkers from the PKS systems. Truncations less complete, as shown by EpoA-ACP(C-10), EpoA-ACP(C-20), and EpoB(N-30), could measurably reduce the elongation efficiency but not prevent it entirely. A complete C-linker truncation on EpoA-ACP duplicated the detrimental effect on elongation seen for EpoB(N-56), suggesting the equally important role of the C terminus of EpoA.

The definition at the primary sequence level for the functional role of the N terminus of EpoB for elongation sets the stage for investigating the secondary, tertiary, and quaternary characteristics of the EpoA/EpoB terminal interface. The question remains if EpoB needs to be a dimer or monomer, under the constraint of differing quaternary models previously described for the PKS and NRPS proteins [16–22]. Similar investigation is needed for the EpoB/EpoC interface, although the scenario is different from that of EpoA/EpoB and other typical PKS interfaces, in that the EpoB C-linker effect appears to be concentrated on one basic charge cluster. The strategy of covalently joining linker pairs for direct structural characterization as reported for the DEBS PKS system [43] will likely be applicable for EpoB with the availability of some information at the quaternary level to help design the construct. As the first NRPS from a hybrid

NRPS-PKS pathway to show defined modular terminal regions essential for elongation, EpoB serves as a prototype to test if these characteristics may be extendable to other proteins from hybrid S-PKS systems. From the combinatorial biosynthesis point of view, modification or optimization in these crucial regions will be a key variable leading to additional noncognate hybrid pathways for producing epothilone variants.

Significance

NRPS and PKS biosynthetic pathways have been found to converge and form mixed assembly lines to produce hybrid natural products comprised of NRP and PK features. Many of these hybrid natural products, for example the epothilones, are medicinally important. Although the pure NRP or PK biosynthetic pathways have received much attention, the hybrid/PK assembly lines have just begun to emerge as complex systems to be examined. Given the divergent quaternary characteristics of NRPS and PKS, it is essential to understand how the hybrid systems operate at the molecular level, a prerequisite to optimizing future protein engineering or evolution effort for combinatorially synthesizing "unnatural" variants of natural products. This study defines minimal terminal regions of EpoB required for constituting the functional hybrid PKS/NRPS/PKS interface formed by EpoA, EpoB, and EpoC, the initial three protein subunits of the epothilone biosynthetic pathway. Although modular like the PKS linkers, these regions were found to have different themes for recognition, in that the C-terminal linker effect of EpoB concentrates on one basic charge cluster while the elongation effect of the N terminus of EpoB is spread over 56 amino acids.

Experimental Procedures

Cloning of the Carrier Proteins and EpoB Mutants

The pET-28 vectors (Novagen) were used for subcloning all proteins. Restriction enzymes and T4 DNA ligase were purchased (New England Biolabs). A vector containing Ybt-PCP2(EpoB) [35] was used for as a template for subcloning Ybt-PCP2(EpoB-8). PCR amplified genes were digested with NdeI and NotI restriction enzymes and ligated into pET-28b. Clones were characterized by DNA sequencing (Dana Farber Cancer Institute) and confirmed. For Ybt-PCP2(EpoB-8), the primers were 5′-GAA TTC **CAT ATG** GCA GAG CGC TCC CCG CGC GTA TGC-3′ (forward) and 5′-G GGA ATT A**GC GGC CGC** TTA AGC CTC CAC TCG GTC-3′ (reverse).

CouN5 was subcloned using NdeI and EcoRI with primers 5′-GTC ATA **CAT ATG** GCA ACC CAG ACC CGA GAA GAC-3′ and 5′-CAG TGC **GAA TTC** TTA CTT CCG CGA CAG GCT CTC ATG C-3′. CouN5(EpoB) was cloned using splicing by overlap extension. The two templates used for the first round of PCR were plasmids containing CouN5 and EpoB. For CouN5, the primers were 5′-GTC ATA **CAT ATG** GCA ACC CAG ACC CGA GAA GAC-3′ (N) and 5′-CAA ACC GGA CGC CAG CTT CCG CGA CAG GCT-3′(N′). For EpoB, the primers were 5′-AGC CTG TCG CGG AAG CTG GCG TCC GGT TTG-3′ (C) and 5′-CAG TGC **GAA TTC** TTA GCT ACG TCT CCT GCC-3′ (C′). For the second round of PCR, the primers were N and C′ with the two PCR products from the first round of PCR combined as the template. CouN5(EpoB-8) was subcloned analogously as described for Ybt-PCP2(EpoB-8), using primers N and 5′-CAG TGC **GAA TTC** TTA AGC CTC CAC TCG GTC CTG-3′. CouN5(EpoB6G) was constructed by two sequential PCR steps, using the same forward primer N and two reverse primers in tandem: first 5′-G GGA ATT A**GC GGC CGC** TTA GCC GCC GCC AGC CTC CAC TCG GTC-3′ and then 5′-GGA ATT A**GC GGC CGC** TTA GCT GCC GCC GCC GCC GCC GCC AGC-3′. The insert generated from the second round PCR was digested and ligated into a pET-28a vector.

All EpoA-ACP mutants were cloned in pET-28b using NdeI and BamHI restriction sites, sharing the same forward primer 5′-G GAA TTC **CAT ATG** GCA GGC GCC GGC CCG TCC ACC. The reverse primers for EpoA-ACP(C-10), (C-20), and (C-24) were 3′-CGC **GGA TCC** TTA GCC CGA TGA GAC GAA GTC GCT, 3′-CGC **GGA TCC** TTA TGC CCG TAG GTT CTC CGC CGC, and 3′-CGC **GGA TCC** TTA TGT GGC GAG AGC ATC CAA CAG, respectively.

A full-length EpoB gene in pET-28b was used as the template for cloning the EpoB mutants and amplified using primers containing NdeI and AvrII restriction sites. Both EpoB mutants shared the same reverse primer 3′-C AGG CTG CCT AGG TCA ACG TTA ATG AGA TCG. The forward primers were 5′-G GAA TTC **CAT ATG** GCC CTG AAC CCG AAC CTG CTC for EpoB(N-30) and 5′-G GAA TTC **CAT ATG** GCA GAG TCC ATC GTG CCC GCC for EpoB(N-56). The PCR products were then digested and ligated with the full-length EpoB vector double-digested using NdeI and AvrII restriction enzymes.

Overexpression and Purification

BL21(DE3) cells (Invitrogen) were used in all overexpressions. The procedure for the carrier proteins is in general as follows: cells were grown in LB medium in the presence of 50 μg/ml of kanamycin at 37°C to an OD_{600} of 0.5–0.6, cooled to 22°C, induced with 1 mM IPTG for 3.5–12 hr, and centrifuged. The harvested cells were resuspended in a lysis buffer (20 mM Tris [pH = 7.6], 500 mM NaCl, 5% glycerol, and 5 mM imidazole) and lysed by passing twice through French Press. The lysates were centrifuged at 4°C (30 min at 35,000 × g), and to the supernatant was added immediately 1 ml of Ni-NTA resin (Qiagen). Binding was allowed for 4 hr at 4°C, after which the resin was washed with 20 ml of the lysis buffer and eluted with increasing concentrations of imidazole (Sigma, 20 mM, 30 mM, 100 mM, and 500 mM) in the lysis buffer. The proteins were dialyzed (3 × 1 liter of 20 mM Tris [pH = 7.6], and 5% glycerol) and stored at −80°C. The protein concentrations were determined using the Bradford assay (Bio-rad). MALDI-TOF mass spectroscopy (Dana Faber Cancer Institute) was performed for all carrier proteins, which were found to be in the demethioninated *apo*-form.

For the EpoB mutants, growth was maintained at 15°C for 72 hr without induction, and the purification procedure was followed as described previously [33]. Anion-exchange (MonoQ, Pharmacia) was used to further purify the 157 kDa and 154 kDa mutants after the Ni-column. The eluent system was 25 mM Tris buffer at pH 8.0 (A) in 50 mM NaCl and 25 mM Tris buffer at pH 8.0 in 500 mM NaCl (B). The gradient was from 0%–100% B in 90 min at 4 ml/min flow rate. The fractions containing pure proteins were pooled, concentrated, dialyzed against 20mM Tris (pH 7.5), 50 mM NaCl, 1 mM TCEP, and 10% glycerol, and stored at −80°C. Protein concentrations were determined using the Bradford assay.

Enzymic Synthesis of Acyl-*S*-Carrier Proteins

A typical reaction (20 μl) contained 10 μM PCP or ACP, 100 μM acyl-CoA, 50 mM Tris (pH 7), 5 mM $MgCl_2$, and 0.5 μM Sfp. After an hour at 37°C, the proteins were purified using ZipTip (C4, Millipore). MALDI-TOF mass spectrometry was performed using α-sinapinic acid (10 mg/ml in 60% acetonitrile/water) as the matrix.

ATP-PP_i Exchange Essay for A Domain Activity of EpoB

As described before [33, 45], reactions (100 μl) were initiated by adding the EpoB protein in a final solution containing 100 nM EpoB protein, 75 mM Tris (pH 7.5), 10 mM $MgCl_2$, 5 mM TCEP, 3 mM ATP, 1.5 mM L-cysteine, and 1 mM [^{32}P]pyrophosphate with specific activity at 302 μCi/mmol made from a mixture of pyrophosphate and [^{32}P]pyrophosphate from Perkin Elmer (55 Ci/mmol). After five minutes at room temperature, reactions were quenched with 500 μl of charcoal suspensions (1.6% [w/v] activated charcoal, 4.5% [w/v] tetrasodium pyrophosphate, and 3.5% perchloric acid in water). Centrifugation was applied to provide the charcoal pellet. The pellet was then washed with the wash solution (4.5% [w/v] tetrasodium pyrophosphate and 3.5% perchloric acid in water) twice, resuspended in water and counted in a scintillation counter (Beckman LS6500.)

Single-Turnover Elongation Reactions from Carrier Proteins to EpoC

In a typical reaction, the carrier protein was first subjected to the methylthiazolyl or picolinyl loading experiment described above [35]. EpoC and NADPH (Sigma) were then added, and the transfer reaction was initiated by adding [^{14}C]-methylmalonyl-CoA (50 mCi/mmol, Perkin Elmer). The final transfer reaction contained 3–5 μM loaded PCP, 5–15 μM EpoC, 1.2 mM NADPH, 10–30 μM [^{14}C]-methylmalonyl-CoA, 100 μM methylthiazolyl-CoA, 25 mM Tris (pH 7), 2 mM $MgCl_2$, and 0.3 μM Sfp. For each time point an aliquot of 100 μl was quenched by 10% TCA (2 $\times$ 500 μl) precipitation. The precipitated protein pellet was treated with 100 μl of 0.1 N KOH at 65°C for 10 min and acidified with 10 μl 6 N HCl. For reverse-phase (C_{18} Vydac Column) radio-HPLC analysis, the acidified solution was spiked with nonradioactive standard synthetically prepared [35] and analyzed by dual detection with a UV detector and an online ^{14}C radioisotope detector. The gradient was 0%–70% acetonitrile in water with 0.1% trifluoroacetic acid over 25 min. For radio-TLC analysis, the acidified solution was first dried in vacuo and then resuspended in 10 μl of 0.1 N aqueous HCL. The resuspension was then spotted onto a normal phase silica-gel coated plate (Sigma) and developed with an eluting system composed of chloroform: acetic acid: ethanol (15:1:1) for reactions transferring the methylthiazolyl group or ethyl acetate: ethanol (10:1) for transferring the picolinyl group. The developed plate was wrapped and exposed to a BAS-IIIS image plate (Fuji) for 12 hr and read by a Bio-Imaging Analyzer BAS1000 (Fuji). The image was analyzed densitometrically using Image Gauge 3.0.

Single-Turnover Elongation Reactions from Carrier Proteins to EpoB

In a typical reaction, the carrier protein was loaded with the acetyl group using Sfp and acetyl-CoA in similar conditions as described in the previous section. EpoB was first primed at room temperature for one hour with Sfp and CoA in a solution containing 50 mM Tris (pH 7), 3 mM ATP, 300 μM CoA, 5 mM $MgCl_2$, 5 mM TCEP, and 0.5 μM Sfp. The loading of EpoB was then initiated by adding a mixture of L-cysteine (Sigma) and radioactive L-cysteine (1000 Ci/mmol, Perkin Elmer) with the final concentration of L-cysteine at 100 μM (specific activity at 250 mCi/mmol). The loading reaction was allowed to proceed at room temperature for half an hour and then combined with the solution containing the acylated carrier protein to give the final transfer reaction. The final transfer reaction typically contained 3 μM ACP, 1 μM EpoB, 50 μM acetyl-CoA, 150 μM CoA, 25 mM Tris (pH 7), 1.5 mM ATP, 5 mM $MgCl_2$, 5 mM TCEP, 50 μM L-cysteine, and 0.5 μM Sfp. For each time point an aliquot of 200 μl was quenched by 10% TCA (2 $\times$ 500 μL) precipitation. The precipitated protein pellet was treated with 100 μl of 0.1 N KOH at 65°C for 10 min and acidified with 10 μl 6 N HCL. The radio-HPLC analysis was performed using methyl thiazolyl acid [33] as the synthetically prepared standard. A ^{35}S online radioisotope detector was used to monitor radio-labeled product formation. A similar radio-TLC analysis protocol as described in the previous section was followed except that the eluting system composed of butanol: acetic acid: water (14:3:3).

Acknowledgments

Funding for this work was provided by the NIH (GM21643, C.T.W.). F.L. is supported by a National Institutes of Health postdoctoral fellowship (GM66456). Gratitude is extended to Professor Heide for providing a cosmid containing CouN5.

Received: June 18, 2004
Revised: August 14, 2004
Accepted: August 23, 2004
Published: November 29, 2004

References

1. Schwarzer, D., Finking, R., and Marahiel, M.A. (2003). Nonribosomal peptides: from genes to products. Nat. Prod. Rep. *20*, 275–287.
2. Staunton, J., and Weissman, K.J. (2001). Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. *18*, 380–416.
3. Cane, D.E., and Walsh, C.T. (1999). The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. Chem. Biol. *6*, R319–R325.
4. Khosla, C. (1997). Harnessing the biosynthetic potential of modular polyketide synthases. Chem. Rev. *97*, 2577–2590.
5. Tsoi, C.J., and Khosla, C. (1995). Combinatorial biosynthesis of "unnatural" natural products: the polyketide example. Chem. Biol. *2*, 355–362.
6. Cane, D.E., Walsh, C.T., and Khosla, C. (1998). Harnessing the biosynthetic code: combinations, permutations, and mutations. Science *282*, 63–68.
7. Konig, A., Schwecke, T., Molnar, I., Bohm, G.A., Lowden, P.A., Staunton, J., and Leadlay, P.F. (1997). The pipecolate-incorporating enzyme for the biosynthesis of the immunosuppressant rapamycin–nucleotide sequence analysis, disruption and heterologous expression of rapP from Streptomyces hygroscopicus. Eur. J. Biochem. *247*, 526–534.
8. Du, L., Sánchez, C., Chen, M., Edwards, D.J., and Shen, B. (2000). 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*, 623–642.
9. Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C., and Julien, B. (2000). Cloning and heterologous expression of the epothilone gene cluster. Science *287*, 640–642.
10. Molnar, I., Schupp, T., Ono, M., Zirkle, R., Milnamow, M., Nowak-Thompson, B., Engel, N., Toupet, C., Stratmann, A., Cyr, D.D., et al. (2000). The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90. Chem. Biol. *7*, 97–109.
11. Bollag, D.M., McQueney, P.A., Zhu, J., Hensens, O., Koupal, L., Liesch, J., Goetz, M., Lazarides, E., and Woods, C.M. (1995). Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action. Cancer Res. *55*, 2325–2333.
12. Nicolaou, K.C., Roschangar, F., and Vourloumis, D. (1998). Chemical biology of epothilones. Angew. Chem. Int. Ed. Engl. *37*, 2014–2045.
13. Pelludat, C., Rakin, A., Jacobi, C., Schubert, S., and Heesemann, J. (1998). The yersiniabactin biosynthetic gene cluster of *Yersinia enterocolitica*: organization and siderophore-dependent regulation. J. Bacteriol. *180*, 538–546.
14. Tang, G.L., Cheng, Y.Q., and Shen, B. (2004). Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase. Chem. Biol. *11*, 33–45.
15. Du, L., and Shen, B. (2001). Biosynthesis of hybrid peptide-polyketide natural products. Curr. Opin. Drug. Discov. *4*, 215–228.
16. Kao, C.M., Pieper, R., Cane, D.E., and Khosla, C. (1996). Evidence for two catalytically independent clusters of active sites in a functional modular polyketide synthase. Biochemistry *35*, 12363–12368.
17. Staunton, J., Caffrey, P., Aparicio, J.F., Roberts, G.A., Bethell, S.S., and Leadlay, P.F. (1996). Evidence for a double-helical structure for modular polyketide synthases. Nat. Struct. Biol. *3*, 188–192.
18. Sieber, S.A., Linne, U., Hillson, N.J., Roche, E., Walsh, C.T., and Marahiel, M.A. (2002). Evidence for a monomeric structure of nonribosomal peptide synthetases. Chem. Biol. *9*, 997–1008.
19. Schwecke, T., Aharonowitz, Y., Palissa, H., von Dohren, H., Kleinkauf, H., and van Liempt, H. (1992). Enzymatic characterisation of the multifunctional enzyme delta-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase from Streptomyces clavuligerus. Eur. J. Biochem. *205*, 687–694.
20. Konig, A., Schwecke, T., Molnar, I., Bohm, G.A., Lowden, P.A., Staunton, J., and Leadlay, P.F. (1997). The pipecolate-incorporating enzyme for the biosynthesis of the immunosuppressant rapamycin: nucleotide sequence analysis, disruption and heterologous expression of rapP from Streptomyces hygroscopicus. Eur. J. Biochem. *247*, 526–534.
21. Nielsen, J.B., Hsu, M.J., Byrne, K.M., and Kaplan, L. (1991). Biosynthesis of the immunosuppressant immunomycin: the en-

zymology of pipecholate incorporation. Biochemistry *30*, 5789–5796.
22. Hillson, N.J., and Walsh, C.T. (2003). Dimeric structure of the six-domain VibF subunit of vibriobactin synthetase: mutant domain activity regain and ultracentrifugation studies. Biochemistry *42*, 766–775.
23. Gokhale, R.S., Tsuji, S.Y., Cane, D.E., and Khosla, C. (1999). Dissecting and exploiting intermodular communication in polyketide synthases. Science *16*, 482–485.
24. Gokhale, R.S., and Khosla, C. (2000). Role of linkers in communication between protein modules. Curr. Opin. Chem. Biol. *4*, 22–27.
25. Tsuji, S.Y., Cane, D.E., and Khosla, C. (2001). Selective protein-protein interactions direct channeling of intermediates between polyketide synthase modules. Biochemistry *40*, 2326–2331.
26. Wu, N., Cane, D.E., and Khosla, C. (2002). Quantitative analysis of the relative contributions of donor acyl carrier proteins, acceptor ketosynthases, and linker regions to intermodular transfer of intermediates in hybrid polyketide synthases. Biochemistry *41*, 5056–5066.
27. Kumar, P., Li, Q., Cane, D.E., and Khosla, C. (2003). Intermodular communication in modular polyketide synthases: structural and mutational analysis of linker mediated protein-protein recognition. J. Am. Chem. Soc. *125*, 4097–4102.
28. Du, L., Sanchez, C., and Shen, B. (2001). Hybrid peptide-polyketide natural products: biosynthesis and prospects toward engineering novel molecules. Metab. Eng. *3*, 78–95.
29. Walsh, C.T. (2004). Combinatorial biosynthesis of antibiotics: challenges and opportunities. Chembiochem *3*, 125–134.
30. Khosla, C., and Keasling, J.D. (2003). Metabolic engineering for drug discovery and development. Nat. Rev. Drug Discov. *2*, 1019–1025.
31. McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M., and Ashley, G. (1999). Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. Proc. Natl. Acad. Sci. USA *96*, 1846–1851.
32. Mootz, H.D., and Marahiel, M.A. (1999). Design and application of multimodular peptide synthetases. Curr. Opin. Biotechnol. *10*, 341–348.
33. Chen, H., O'Connor, S.E., Cane, D.E., and Walsh, C.T. (2001). Epothilone biosynthesis: assembly of the methylthiazolylcarboxy starter unit on the EpoB subunit. Chem. Biol. *8*, 899–912.
34. O'Connor, S.E., Chen, H., and Walsh, C.T. (2002). Enzymatic assembly of epothilones: the EpoC subunit and reconstitution of the EpoA-ACP/B/C polyketide and nonribosomal peptide interfaces. Biochemistry *41*, 5685–5694.
35. O'Connor, S.E., Walsh, C.T., and Liu, F. (2003). Biosynthesis of epothilone intermediates with alternate starter unit: engineering polyketide-nonribosomal interfaces. Angew. Chem. Int. Ed. Engl. *42*, 3917–3921.
36. Wang, Z.X., Li, S.M., and Heide, L. (2000). Identification of the coumermycin A(1) biosynthetic gene cluster of Streptomyces rishiriensis DSM 40489. Antimicrob. Agents Chemother. *9*, 3040–3048.
37. Miller, D.A., Luo, L., Hillson, N., Keating, T.A., and Walsh, C.T. (2002). Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*. Chem. Biol. *9*, 333–344.
38. Keating, T.A., Miller, D.A., and Walsh, C.T. (2000). Expression, purification, and characterization of HMWP2, a 229 kDa, six domain protein subunit of Yersiniabactin synthetase. Biochemistry *39*, 4729–4739.
39. Horton, R.M., Cai, Z.L., Ho, S.N., and Pease, L.R. (1990). Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. Biotechniques *8*, 528–535.
40. Weber, T., Baumgartner, R., Renner, C., Marahiel, M.A., and Holak, T.A. (2000). Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases. Struct. Fold. Des. *8*, 407–418.
41. Quadri, L.E., Weinreb, P.H., Lei, M., Nakano, M.M., Zuber, P., and Walsh, C.T. (1998). Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. Biochemistry *37*, 1585–1595.
42. Finking, R., Solsbacher, J., Konz, D., Schobert, M., Schafer, A., Jahn, D., and Marahiel, M.A. (2002). Characterization of a new type of phosphopantetheinyl transferase for fatty acid and siderophore synthesis in Pseudomonas aeruginosa. J. Biol. Chem. *277*, 50293–50302.
43. Broadhurst, R.W., Nietlispach, D., Wheatcroft, M.P., Leadlay, P.F., and Weissman, K.J. (2003). The structure of docking domains in modular polyketide synthases. Chem. Biol. *10*, 723–731.
44. Mofid, M.R., Finking, R., Essen, L.O., and Marahiel, M.A. (2004). Structure-based mutational analysis of the 4′-phosphopantetheinyl transferases Sfp from Bacillus subtilis: carrier protein recognition and reaction mechanism. Biochemistry *43*, 4128–4136.
45. Rusnak, F., Sakaitani, M., Drueckhammer, D., Reichert, J., and Walsh, C.T. (1991). Biosynthesis of the Escherichia coli siderophore enterobactin: sequence of the entF gene, expression and purification of EntF, and analysis of covalent phosphopantetheine. Biochemistry *30*, 2916–2927.
46. McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server. Bioinformatics *16*, 404–405.
47. Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. *292*, 195–202.