Identification and characterization of a type II peptidyl carrier protein from the bleomycin producer *Streptomyces verticillus* ATCC 15003

Liangcheng Du and Ben Shen

Background: Nonribosomal peptide synthetases (NRPSs) catalyze the assembly of a structurally diverse group of peptides by the multiple-carrier thiotemplate mechanism. All NRPSs known to date are exclusively type I modular enzymes that consist of domains, such as adenylation (A), peptidyl carrier protein (PCP) and condensation (C) domains, for individual enzyme activities. Although several A and PCP domains have been demonstrated to function independently, aminoacylation *in trans* has been successful only between PCPs and their cognate A domains.

Results: We have identified within the bleomycin-biosynthesis gene cluster from *Streptomyces verticillus* ATCC15003 the *blmI* gene that encodes a discrete PCP protein. We overexpressed the *blmI* gene in *Escherichia coli*, purified the BlmI protein, and demonstrated that apo-BlmI can be efficiently modified into holo-BlmI either *in vivo* or *in vitro* by PCP-specific 4′-phosphopantetheine transferases (PPTases). Unlike the PCP domains in type I NRPSs, BlmI lacks its cognate A domain and can be aminoacylated by Val-A, an A domain from a completely unrelated type I NRPS.

Conclusions: BlmI represents the first characterized type II PCP. The BlmI type II PCP, like the PCP domains of type I NRPSs, can be 4′-phosphopantetheinylated by PCP-specific PPTases but is biochemically distinct in that it can be aminoacylated by an A domain from a completely unrelated type I NRPS. Our results provide for the first time the genetic and biochemical evidence to support the existence of a type II NRPS, which might be useful in the combinatorial manipulation of NRPS proteins to generate novel peptides.

Address: Department of Chemistry, University of California, One Shields Avenue, Davis, CA 95616, USA.

Correspondence: Ben Shen
E-mail: shen@chem.ucdavis.edu

Key words: biosynthesis, bleomycin, peptidyl carrier protein, nonribosomal peptide synthetase, *Streptomyces verticillus*

Received: **18 March 1999**
Revisions requested: **14 April 1999**
Revisions received: **4 May 1999**
Accepted: **7 May 1999**

Published: **28 June 1999**

Chemistry & Biology August 1999, **6**:507–517
http://biomednet.com/elecref/1074552100600507

 ISSN 1074-5521

Introduction

Nonribosomal peptides form a large family of natural products that includes many clinically valuable drugs such as bleomycin (antitumor), cyclosporin (immunosuppressant) and vancomycin (antibacterial). In contrast to peptides of ribosomal origin, which are synthesized from gene-encoded peptide precursors and are restricted to the 21 proteinogenic amino acids (including selenocysteine), nonribosomal peptides are synthesized by large, multifunctional proteins called nonribosomal peptide synthetases (NRPSs) [1]. NRPSs catalyze the assembly of nonribosomal peptides from an exceedingly diverse group of precursors, which includes over 300 different residues known to date; the peptide products can be linear, cyclic or branched, or can be modified even further through acylation, glycosylation or heterocyclic ring formation, leading to remarkable structural and functional diversity [2–4].

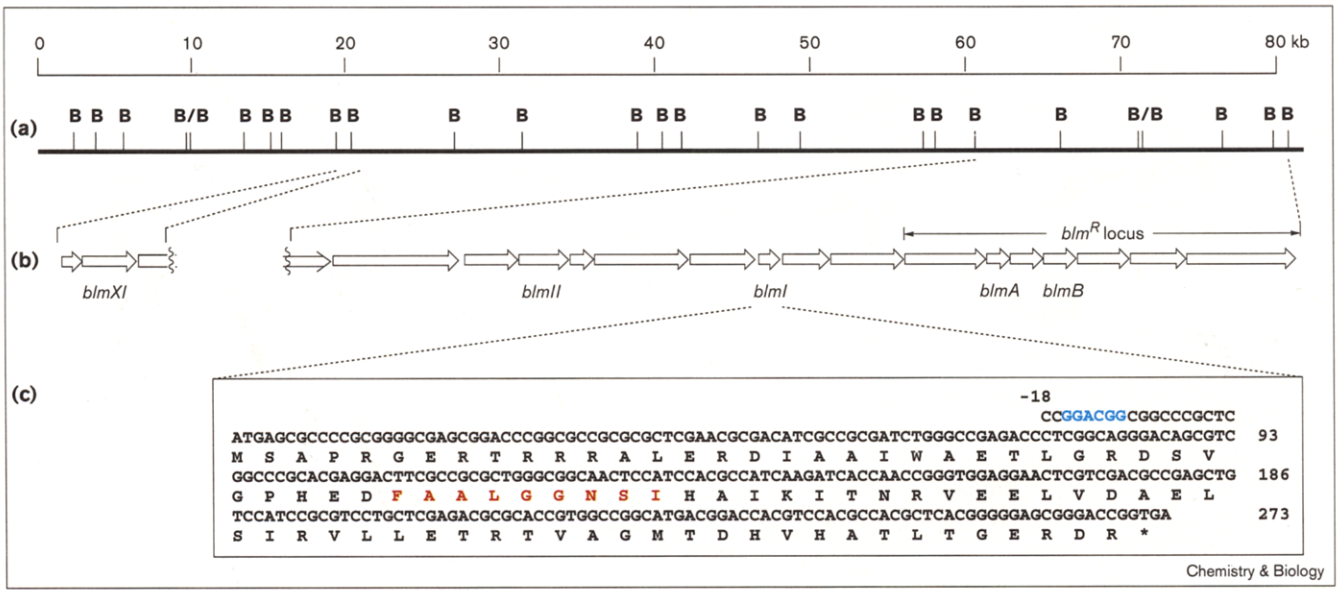
Following the classification of fatty acid synthases (FASs) [5] and polyketide synthases (PKSs) [6], NRPSs known to date are found exclusively to be the large multifunctional enzymes of the type I architecture with modular arrangement [2–4]. A module is defined as a set of distinctive domains that encodes all the enzyme activities necessary for one cycle of peptide chain elongation and the associated modifications. The peptidyl carrier protein (PCP) domain (also called the thiolation domain) is central to each module [7]. During the entire elongation process, the growing peptide intermediates remain tethered, in a thioester linkage, to the terminal -SH group of the 4′-phosphopantetheine prosthetic group of a PCP. The number and order of the modules and the type of domains within each module on the NRPS protein therefore determine the structural variations of the resulting peptide products by dictating the number, order and choice of the incorporated amino acids, and the modifications associated with a particular elongation cycle. This mode of peptide biosynthesis has been termed the multiple-carrier thiotemplate mechanism [8].

A typical NRPS module consists of an adenylation (A) domain [9], a PCP domain [7] and a condensation (C; also called elongation) domain [10], and each of these domains is characterized by one or a few highly conserved sequence

motifs (also called core sequences) [2–4,11]. The recognition of the motifs has greatly facilitated the cloning and identification of NRPS genes from genomic DNA using the polymerase chain reaction (PCR) with primers designed on the basis of the motifs [12] or from available genome sequences by directly searching the genebank using the motifs as signature sequences [13]. The A domain activates a specific amino acid as an aminoacyl adenylate at the expense of an ATP [9]. The aminoacyl adenylate is then transferred to its cognate PCP domain immediately downstream, forming an aminoacyl thioester with the terminal -SH of the 4′-phosphopantetheine prosthetic group [3]. Nucleophilic attack at the carbonyl group of the aminoacyl thioester of one module by the amino group of the aminoacyl thioester from the downstream neighboring module results in amide-bond formation, therefore effecting the amino to carboxyl terminus stepwise assembly of the peptide product [1,8]; this process is catalyzed by the C domain [10]. Additional domains have also been identified that are responsible for the modification of the peptide product, such as an epimerization domain for the conversion of L-amino acids to D-amino acids [3], an *N*-methyltransferase domain for the addition of a methyl group to the amide nitrogen [3] and a cyclization domain for the formation of heterocyclic rings [14]. A thioesterase domain is usually found at the distal carboxyl terminus of the NRPS, the function of which has been proposed to be the release of the full-length peptide product from the NRPS protein to terminate biosynthesis [3,15]. A distinct gene encoding a discrete thioesterase has also been found within most of the NRPS gene clusters. The latter thioesterase has been speculated to liberate mischarged NRPS that is blocked by an unspecific thioesterification at the PCP domain [3].

Individual NRPS modules [15–19] and domains, such as A [9,14,19,20], PCP [7,19,21–26] and C domains [10], have also been heterologously expressed in *Escherichia coli* and biochemically characterized ***in vitro***. These studies not only proved directly the proposed functions for individual domains but also demonstrated that these domains can function independently and act ***in trans***, at least between the A domain and PCP. For example, the catalytic efficiency for aminoacylation ***in trans*** between the A and the PCP domains of SrfB1 (which activates valine in surfactin biosynthesis) was shown to be comparable to that for the undissected type I enzyme [7]. A similar result was also observed for aminoacylation ***in trans*** between the A domain of GrsA and the PCP domain of TycA — both the undissected GrsA and TycA enzymes activate phenylalanine in gramicidin S and tyrocidine biosynthesis, respectively [19]. Intriguingly, no ***in trans*** aminoacylation was detected between the A domain of SrfB1 and other PCP domains such as that of TycA, SrfB2 (which activates Asp in surfactin biosynthesis) or the aryl carrier protein (ArCP) domain of EntB (which activates salicylic acid in enterbactin biosynthesis). These results clearly suggested a specific protein–protein recognition between the A and PCP

Figure 1



domains; the A domain of a type I NRPS therefore appeared to be able to transfer the activated aminoacyl adenylate only to its cognate PCP domain [19]. Consequently, although the choice of an incorporating amino acid may be determined primarily by the A domain, the domain-specific aminoacylation between the A domain and its cognate PCP domains certainly could provide additional 'gating' against misincorporation of nonspecifically activated aminoacyl adenylate into the final peptide product.

The bleomycins (BLMs) are a family of glycopeptide-derived antitumor antibiotics, and the clinically used drug Blenoxane® contains BLM A2 and BLM B2 as the principal constituents, which have the same peptide/polyketide/peptide backbone and differ only at their carboxy-terminal amines. We have been studying the biosynthesis of BLM as a model system for a hybrid NRPS and PKS [27], and have cloned and sequenced a 80 kilobase (kb) gene cluster, including the two previously characterized *blmAB* resistance genes [28,29], for BLM biosynthesis from *Streptomyces verticillus* (*Sv*) ATCC15003 (L.D., C. Sanchez, M. Chen, D.J. Edwards, B.S., unpublished observations). Among the 36 open reading frames (orfs) identified within the *blm* gene cluster is a PKS gene flanked by multiple NRPS genes encoding a total of 12 putative NRPS modules, a genetic organization in agreement with the hybrid NRPS/PKS/NRPS hypothesis for BLM biosynthesis. Intriguingly, in addition to the type I NRPS and PKS genes, we also identified three orfs, *blmI*, *blmII* and *blmXI*, the deduced amino acid sequences of which are highly homologous to the PCP (BlmI) and C (BlmII and BlmXI) domains of type I NRPS, respectively (Figure 1b). As all NRPS domains known to date exist as an integrated part of a type I multifunctional NRPS module, the identification of these discrete PCP and C proteins suggests a novel type II organization of NRPS, reminiscent of type II FASs [5] and PKSs [6]. Here we report the heterologous expression of *blmI* in *E. coli*, *in vivo* 4′-phosphopantetheinylation of the BlmI protein by the Gsp phosphopantetheinyl transferase (PPTase) [26], and *in vitro* 4′-phosphopantetheinylation of the BlmI protein by the Sfp PPTase [18,23,24], validating BlmI as the first characterized type II PCP. Remarkably, in contrast to the type I PCP domains that accept only the aminoacyl adenylate from their cognate A domains [7,19], the BlmI type II PCP, which has no cognate A domain of its own, can be aminoacylated *in vitro* by an A domain from a completely unrelated type I NRPS module. The latter property of the BlmI protein implicates a distinct strategy for a type II NRPS to assemble peptides and can now be exploited directly in the combinatorial manipulation of NRPS proteins to generate novel peptides.

Results

Cloning and sequence analysis of the *blmI* gene

In our effort to clone the gene cluster responsible for BLM biosynthesis, we have determined an 80 kb DNA sequence from *Sv* ATCC15003 (Figure 1a) [27]. Among the 36 orfs identified within the *blm* gene cluster is a small orf of 273 base pairs (bp), *blmI*, which is located approximately 4 kb upstream of the previously characterized *blmAB* resistance locus [28,29] (Figure 1b). The *blmI* gene encodes a protein of 90 amino acids that has a molecular weight of 9957 and a pI of 6.52 (Figure 1c). Computer-assisted analysis [30] of the deduced amino acid sequence indicates that BlmI is very similar to various PCP domains of NRPSs (around 40% identity and 60% similarity, as shown in Figure 2). Like known PCP domains of NRPSs, BlmI has the highly conserved signature motif of LGGXS (using single-letter amino-acid code, where X is any amino acid). The serine residue in this motif is the site for 4′-phosphopantetheinylation [3,11], which is the post-translational modification essential for peptide biosynthesis because it converts the apo-PCP into the functional holo-PCP [3,31]. On the basis of sequence comparison, BlmI is most related to PCPs and not to other types of carrier proteins that also share the same LGGXS motif and undergo the same post-translational 4′-phosphopantetheinylation [31], such as the *E. coli* acyl carrier protein (ACP) [32], the ACP domain of type I PKS and the type II PKS ACP [33,34], the ArCP domain [17], and several nodulation-related ACP-like proteins [35,36].

Figure 2

Amino acid sequence comparison of BlmI with PCP domains of known type I NRPSs (Grs-2 [P14688], 36% identity, 58% similarity; Srfa-3 [Q08787], 40% identity, 64% similarity; Vir-S [CAA72310], 36% identity, 60% similarity; Saf-B [AAC44128], 40% identity, 54% similarity). Protein sequence accession numbers are in square parentheses. The blue-shaded letters indicate similar amino acids. Consensus residues are amino acids that are similar in more than three sequences. The signature motif for 4′-phosphopantetheinylation is underlined.

```
Grs-2      3045-ISIGTEYVAPRTMLEGKLEEIWKDVLGLQRVGIHDDFFTIGGHSL-3089
Srfa-3      960-DQLAEEWIGPRNEMEETIAQIWSEVLGRKQIGIHDDFFALGGHSL-1004
Vir-S       557-GRSVEGRGVPRTPQQEILASLFAEVLGLSKVGIHEDFFDLGGHSL- 601
Saf-B      1661-LDPGQDYLAPRNELEARIAAIWEGLLRRERVGVHDSFFDLGGNSL-1705
BlmI          1-MSAPRGERTRRRALERDIAAIWAETLGRDSVGPHEDFAALGGNSI-  45
Consensus     1-I  G EYVAPR  LE  IA IW EVLGR RVGIHDDFF LGGHSL-  45

Grs-2      3090-KAMAVISQVHKECQTEVPLRVLFETPTIQGLAKYIEETDTEQYMA-3134
Srfa-3     1005-KAMTAVPH.QQELGIDLPVKLLFEAPTIAGISAYLKNGGSDGLQD-1048
Vir-S       602-LATRLTSRIRTVLGAEIAVRDLFEAPTVEALAETLEEAREVRPAL- 646
Saf-B      1706-LATRLATRLAATLQVQAGVRTVFEHRTVAAQAAHFTQATKTHQAH-1750
BlmI         46-HAIKITNRVEELVDAELSIRVLLETRTVAGMTDHVHATLTGERDR-  90
Consensus    46-KAMRV SRV   L  EV VRVLFE PTVAGLA  I  G T     -  90
```

Chemistry & Biology

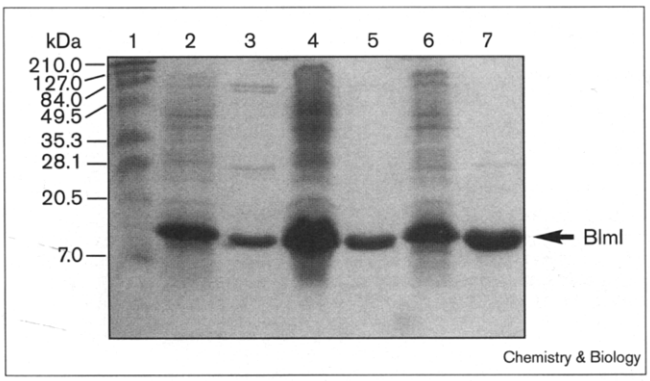
Overexpression of *blmI* in *E. coli*

To overexpress the *blmI* gene in *E. coli*, we directly amplified the *blmI* gene, using PCR, from the *Sv.* ATCC15003 genomic DNA and cloned it into the pQE-60 expression vector to give pBS1 so that BlmI could be produced as a protein with a native amino terminus and a His_6-tag at its carboxyl terminus. No production of the BlmI protein was detected, however, as judged using sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), upon introduction of pBS1 into *E. coli* M15(pREP4) under the standard overexpression conditions recommended by the manufacturer (Qiagen). We reasoned that the small BlmI protein with its native amino terminus may not be stable in the heterologous host, and therefore moved the *blmI* gene from pBS1 into pET-29a to yield the second overexpression construct of pBS2. In pBS2, BlmI should be produced as a fusion protein with 27 extra amino acid residues at its amino terminus, including an S-tag and the thrombin-cleaving site, in addition to the His_6-tag at its carboxyl terminus. Introduction of pBS2 into *E. coli* BL21(DE-3) under the standard overexpression conditions recommended by the manufacturer (Novagen) indeed resulted in overproduction of BlmI. In fact, the bulk of the soluble protein was the overproduced BlmI (Figure 3, lane 2), which was easily purified by affinity chromatography using Ni-NTA resin (Qiagen). Fusion of the additional 27 amino acids to the amino terminus of BlmI as in pBS2 and change of the expression system from *E. coli* M15(pREP4) (pBS1) to *E. coli* BL21(DE-3)(pBS2) therefore dramatically improved the expression level of *blmI*.

In vivo 4′-phosphopantetheinylation of the BlmI protein

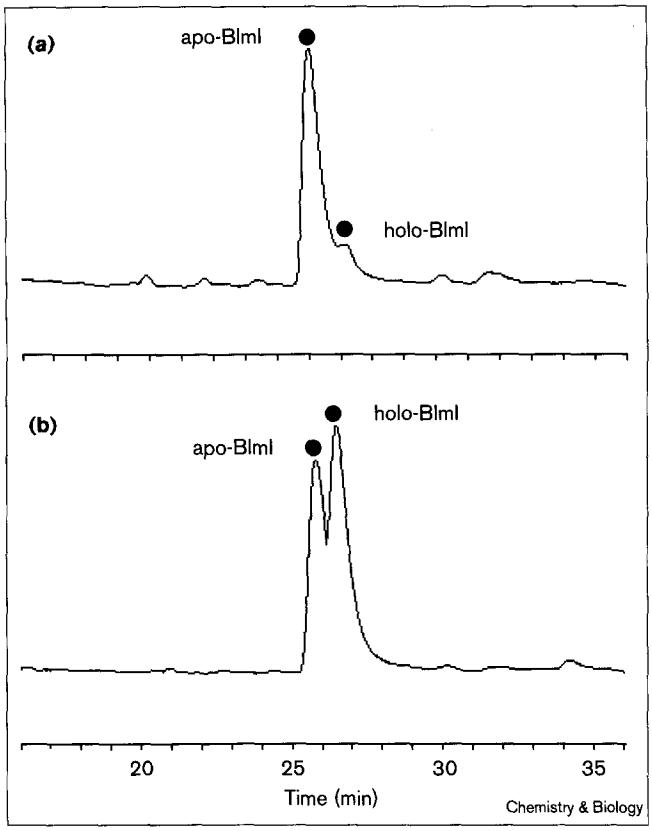
To establish BlmI as a type II PCP, we tested whether it could serve as a substrate for a PCP-specific 4′-PPTase. PPTases catalyze the post-translational modification of an apo-PCP into a holo-PCP by transferring the 4′-phosphopantetheine moiety from co-enzyme A (CoA) to the conserved serine residue of PCP, and this reaction has been developed recently into a general method to prepare various holo-PCP, holo-ACP or holo-ArCP from the corresponding apoproteins [7,17–19,23–25,31,35]. We therefore decided to investigate the 4′-phosphopantetheinylation of BlmI under both *in vivo* [26] and *in vitro* [18,23,24] conditions.

Figure 3



Overexpression of *blmI* in *E. coli* and purification of BlmI protein. Lane 1, molecular weight markers; lanes 2 and 3, total soluble proteins and purified BlmI from BL21(DE3)(pBS2); lanes 4 and 5, total soluble proteins and purified BlmI from OG7001(pBS2); lanes 6 and 7, total soluble proteins and purified BlmI from OG7001(pBS2/pDPT-Gsp).

Figure 4



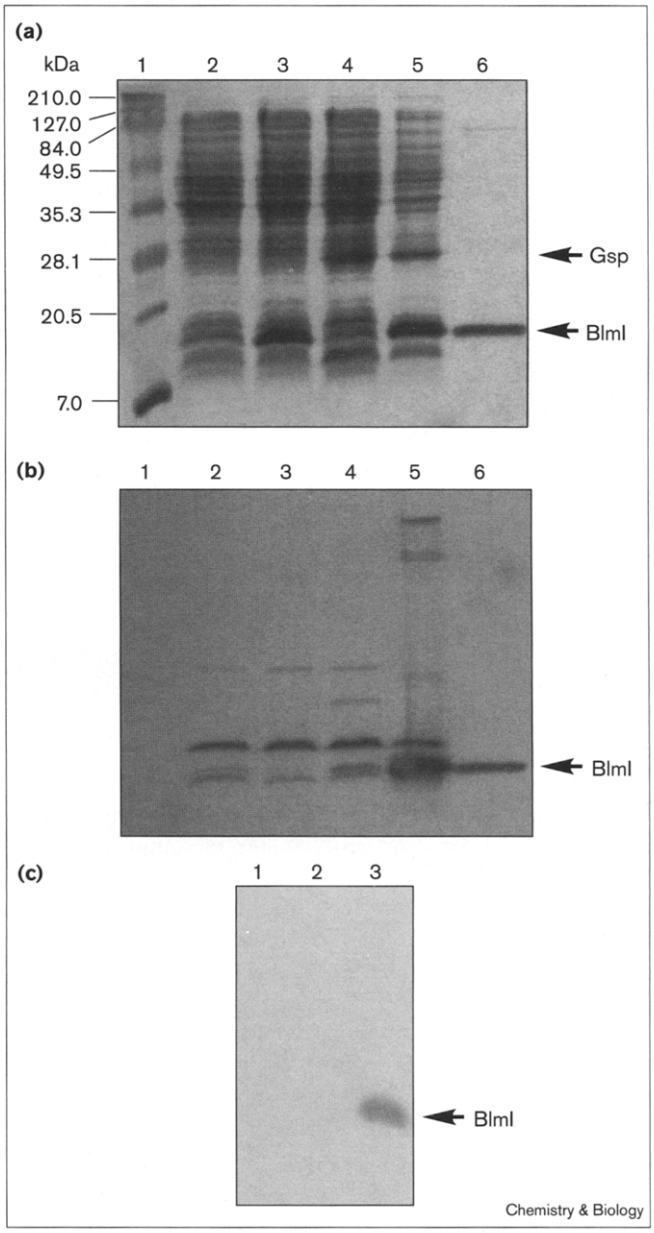
HPLC analysis of BlmI purified from **(a)** *E. coli* OG7001(pBS2) and **(b)** *E. coli* OG7001(pBS2/pDPT-Gsp).

To examine 4′-phosphopantetheinylation of BlmI *in vivo*, we chose *E. coli* OG7001 as the expression host, which is a β-alanine auxotroph derived from *E. coli* BL21(DE3) by P1 co-transduction of the *panD* mutation from *E. coli* SJ16 [35]. Upon introduction of pBS2 into *E. coli* OG7001, *blmI* was exceptionally well expressed (Figure 3, lane 4) and the overproduced BlmI protein was readily purified (Figure 3, lane 5). High-performance liquid chromatography (HPLC) analysis showed, however, that the purified BlmI was essentially in the apo-form (Figure 4a), indicating that apo-BlmI was a poor substrate for the *E. coli* endogenous PPTases, such as EntD and ACP synthase [23,31,32]. To circumvent the poor endogenous PPTase activity, we next co-expressed *blmI* with the *gsp* gene, which was isolated from the gramicidin S producer *Bacillus brevis*, and

encoded a PPTase that was known to 4′-phosphopantetheinylate heterologously produced PCPs in *E. coli* [23,26]. We co-transformed pDPT-Gsp, in which the expression of the *gsp* gene was under the control of the T5/Lac promoter [26], and pBS2 into *E. coli* OG7001. As shown in Figure 3 (lane 6), *blmI* was again very well expressed and the resulting BlmI protein was similarly purified (Figure 3, lane 7). HPLC analysis showed that at least 60% of overproduced BlmI was modified into the holo-BlmI protein (Figure 4b). (A PCP domain was similarly 4′-phosphopantetheinylated *in vivo* before by co-expressing *gsp* in *E. coli* using pDPT-Gsp, and approximately 80% of the PCP was produced in the holo-form [26].)

We next cultured *E. coli* OG7001(pBS2) and *E. coli* OG7001(pBS2/pDPT-Gsp) in the presence of [3-^{3}H]-β-alanine, a known biosynthetic precursor of 4′-phosphopantetheine [7,35]. Specific incorporation of [3-^{3}H]-β-alanine into the 4′-phosphopantetheine moiety of holo-BlmI was determined by autoradiographic analysis. Although fermentation of *E. coli* OG7001(pBS2) in the presence of [3-^{3}H]-β-alanine led to an IPTG-dependent overproduction of BlmI, little of the resulting BlmI protein was ^{3}H-labeled, indicating it was produced in the apo-form (Figure 5a,b; lane 2,3). In contrast, fermentation of *E. coli* OG7001(pBS2/pDPT-Gsp) in the presence of [3-^{3}H]-β-alanine resulted in a significant increase in IPTG-dependent incorporation of the ^{3}H-label into the overproduced BlmI protein (Figure 5a,b; lane 4,5), suggesting a specific incorporation of [3-^{3}H]-β-alanine into holo-BlmI, presumably in the 4′-phosphopanthetheine moiety. Several additional proteins were also weakly labeled by [3-^{3}H]-β-alanine. Both their expression and their incorporation of ^{3}H-label were independent from either IPTG induction or the presence of Gsp, however, hence these proteins were unrelated to BlmI. (Similar background labeling was reported before for *in vivo* 4′-phosphopantetheinylation of other PCPs [35].) We also purified the BlmI protein from *E. coli* OG7001(pBS2/pDPT-Gsp) and demonstrated that the holo-BlmI protein was specifically associated with the ^{3}H-activity (Figure 5a,b; lane 6). Finally, we confirmed the identity of holo-BlmI by subjecting the purified BlmI protein to matrix-assisted laser desorption ionization–time of flight (MALDI–TOF) mass spectral analysis [19]. BlmI produced in the absence of the Gsp PPTase yielded a single peak with a molecular weight of 13,952 Da, suggesting that the produced BlmI protein is in the apo-form (calc'd, 13,949). In contrast, BlmI produced in the presence of Gsp yielded two species with molecular weights of 13,969 and 14,303 Da, respectively. The species with the molecular weight of 13,969 Da represents apo-BlmI, whereas a molecular weight of 14,303 Da unambiguously confirmed the other protein as holo-BlmI (calc'd, 14,289). This result indicated that the purified BlmI consisted of both the apo- and holo-BlmI proteins, in agreement with the HPLC analysis results (Figure 4b).

Figure 5



In vivo incorporation of [3-^{3}H]-β-alanine and *in vitro* incorporation of [^{3}H-pantetheine]-CoA into holo-BlmI. **(a)** Coomassie-Blue-stained 15% SDS–PAGE gel and **(b)** its autoradiogram for *in vivo* labeling of BlmI by ^{3}H-β-alanine in *E. coli* OG7001. Lane 1, molecular weight markers; lanes 2 and 3, total proteins from OG7001(pBS2) without and with IPTG induction; lanes 4 and 5, total proteins from OG7001(pBS2/pDPT-Gsp) without and with IPTG induction; lane 6, BlmI purified from lane 5. **(c)** Autoradiogram for *in vitro* labeling of BlmI by [^{3}H-pantetheine]-CoA. Lane 1, apo-BlmI; lane 2, Sfp; lane 3, apo-BlmI with Sfp.

In vitro 4′-phosphopantetheinylation of the BlmI protein

To investigate 4′-phosphopantetheinylation of BlmI *in vitro*, we chose the Sfp protein as the preferred PPTase, which had been isolated before from the surfactin producer *Bacillus subtilis* [37]. (Overexpression of *gsp* in *E. coli*

using pDPT-Gsp resulted predominantly in an insoluble Gsp protein [26].) The Sfp PPTase was overproduced in *E. coli* MV1190(pUC8-Sfp) and purified to near homogeneity as described previously [24,37]. Upon incubation of the purified apo-BlmI with [^{3}H-pantetheine]-CoA in the presence of the Sfp PPTase, we examined the covalent incorporation of the [^{3}H-pantetheine]-4′-phosphopantetheine moiety from CoA into holo-BlmI using autoradiographic analysis. Indeed, the apo-BlmI was quantitatively labeled by [3H-pantetheine]-CoA (Figure 5c; lane 3), and no labeling was observed in the absence of either the apo-BlmI or the Sfp PPTase protein (Figure 5c; lane 1,2), demonstrating that the Sfp PPTase can recognize apo-BlmI as a substrate and specifically transfer the 4′-phosphopantetheine group from CoA into holo-BlmI.

In vitro aminoacylation of BlmI

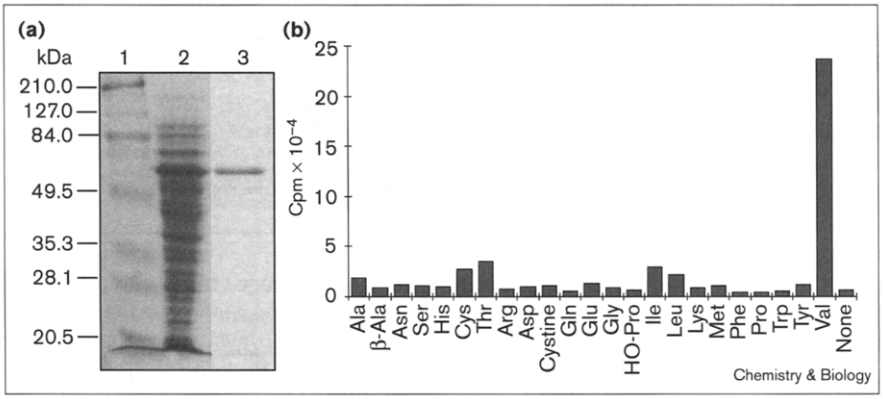
Once we established BlmI as a type II PCP that can be readily modified by PCP-specific PPTases into the holo-BlmI protein, we tested whether the holo-BlmI could be aminoacylated *in trans*, which requires an A domain. As BlmI has no cognate A domain of its own, we turned our attention to another putative biosynthesis gene cluster we cloned previously from *Sv* ATCC15003, which encodes at least three NRPS modules and one PKS module. We have established that this gene cluster is not clustered with the *blm* locus and is unrelated to BLM biosynthesis (L.D., C. Sanchez, M. Chen, S-M. Lee and B.S., unpublished observations). From this gene cluster, we amplified using PCR a 1579 bp fragment encoding an A domain, named Val-A, which we predicted to have a molecular weight of 56,581 Da and a pI of 7.39. We cloned *val*-A into pET-28a to yield pBS3, in which Val-A would be produced as a fusion protein with a His_6-tag at the amino terminus. Introduction of pBS3 into *E. coli* BL21(DE3) under the standard overexpression conditions recommended by the manufacturer (Novagen) resulted in good overproduction of Val-A, predominantly in soluble form, which was purified by affinity chromatography using Ni-NTA resin (Figure 6a). The purified Val-A protein was active as measured using the amino-acid-dependent ATP–PPi exchange assay [38]. Among the 23 amino acids tested, Val-A specifically activated valine, an amino acid that is not required for BLM biosynthesis (Figure 6b).

To carry out the aminoacylation *in trans*, we incubated the purified holo-BlmI and Val-A *in vitro* in the presence of L-[^{14}C(U)]valine and ATP [7,19]. The aminoacylated holo-BlmI-L-[^{14}C(U)]valine species was subjected to SDS–PAGE and specific attachment of L-[^{14}C(U)]valine to holo-BlmI was determined using autoradiographic analysis. Remarkably, the holo-BlmI was specifically labeled by L-[^{14}C(U)]valine in the presence of Val-A, indicative of the formation of the holo-BlmI-S-valine thioester (Figure 7; lane 5). The *in trans* aminoacylation between the holo-BlmI and Val-A proteins appeared to be very specific. Neither incubation of L-[^{14}C(U)]valine with Val-A (Figure 7; lane 1) , the apo-BlmI (Figure 7; lane 2) or the holo-BlmI protein (Figure 7; lane 3) alone, nor incubation of L-[^{14}C(U)]valine with the Val-A and apo-BlmI proteins (Figure 7; lane 4), resulted in the detection of ^{14}C-labeled BlmI protein.

Discussion

Nonribosomal peptides and polyketides are two distinct classes of natural products yet they are assembled from amino acids and short carboxylic acids by NRPSs and PKSs, respectively, using strikingly similar strategies [1]. These fascinating multifunctional enzyme complexes have been classified into two types on the basis of their gene organization and enzyme architecture. Type I enzymes are multifunctional proteins that consist of domains for individual enzyme activities and type II enzymes are multienzyme complexes that consist of discrete proteins that are largely monofunctional. Although both type I and type II PKSs (Figure 8a,c) have been well characterized to account for the vast structural diversities found in polyketide biosynthesis [6], all NRPSs studied so far are exclusively

Figure 6

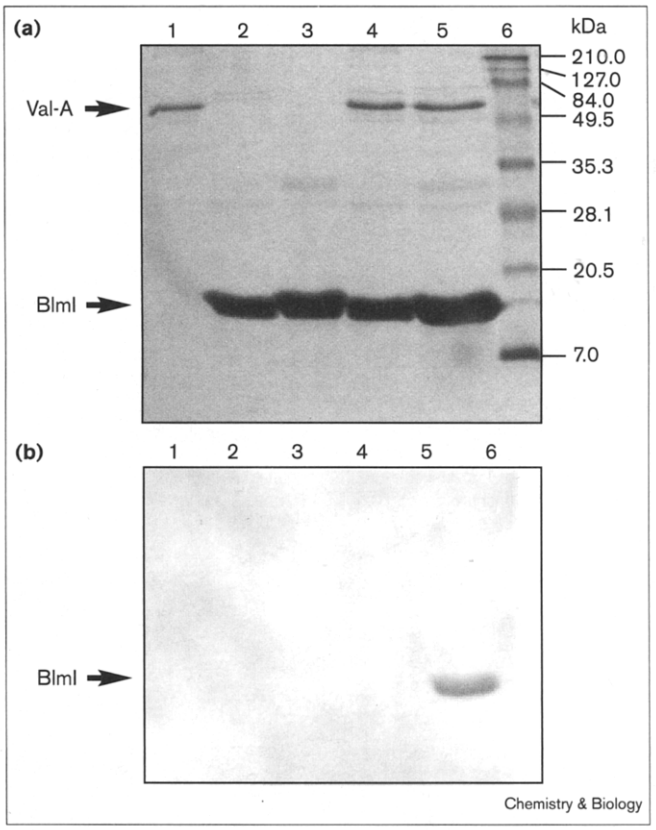


(a) Overexpression of *val*-A in *E. coli* and purification of Val-A. Lane 1, molecular markers; lanes 2 and 3, total soluble proteins and purified Val-1 from BL21(DE3)(pBS3), respectively. **(b)** Amino acid specificity of Val-A as determined using the amino-acid-dependent ATP–PPi assay (standard three-letter amino-acid designations are used and HO-Pro refers to 4-hydroxyproline).

the type I modular enzymes (Figure 8b) [2–4]. It is very tempting to speculate that a type II NRPS exists that, analogous to type II PKS [39–41], consists of discrete proteins possessing enzyme activities such as the A [9], the PCP [7] or the C [10] domains of type I NRPSs (Figure 8d). The fact that both the A [9,14,19,20] and the PCP [7,19,21–26] domains of type I NRPSs can act as independent enzymes supports the hypothesis of a type II NRPS.

We have now cloned and sequenced the *blmI* gene, overproduced and characterized the BlmI protein as a *bona fide* type II PCP, and demonstrated that holo-BlmI can be aminoacylated by a completely unrelated A domain, providing for the first time genetic and biochemical evidence for a type II NRPS enzyme. We concluded BlmI to be a type II PCP on the basis of the following criteria. First, the deduced amino acid sequence of the *blmI* gene is highly homologous to various PCP domains of known NRPSs, particularly at the signature motif of LGGXS within which the 4′-phosphopantetheine prosthetic group is covalently attached to the serine residue [3,11]. Although the current boundaries for a PCP domain in the literature were defined arbitrarily [7] and varied from one PCP to another [7,19,21–26], we can now re-define a PCP domain for the type I NRPS as a 90 amino acid peptide in which approximately 45 amino acids flank either side of the the essential serine residue in the LGGXS motif, in light of this discrete BlmI type II PCP (Figure 2). Second, the *blmI* gene has been successfully expressed in *E. coli*, and fusion of a short peptide to the amino terminus of BlmI dramatically improved its overproduction efficiency. Although we cannot exclude the effect of different systems on gene expression (i.e. *E. coli* M15(pREP4)(pBS1) compared with *E. coli* BL21(DE-3)(pBS2)) we attribute the increase in expression efficiency to the stability of BlmI as an amino-terminal fusion protein instead of the otherwise labile BlmI protein with its native amino terminus. As BlmI was produced predominantly in the apo-form in *E. coli*, apo-BlmI was apparently not a substrate for the endogenous PPTases, such as EntD or ACP synthase, excluding BlmI as an ArCP or ACP, respectively. EntD and ACP synthase are known to 4′-phosphopantetheinylate apo-ArCP and apo-ACP, respectively, to their holo-forms efficiently [23,31,32]. Third, the apo-BlmI protein serves as a substrate for PCP-specific PPTases that transfer the 4′-phosphopantetheine moiety from CoA to apo-BlmI to yield the holo-BlmI protein. We have demonstrated this post-translational modification for BlmI *in vivo* with the Gsp PPTase [26] and *in vitro* with the Sfp PPTase [18,23,24], both of which have been extensively used in preparing holo-PCPs. Fourth, the specific modification of apo-BlmI by 4′-phosphopantetheinylation has been monitored using HPLC analysis (Figure 4) [19] and by specific incorporation of [3-^{3}H]-β-alanine *in vivo* [7,26,35] and of [^{3}H-pantetheine]-CoA *in vitro* [18,23,24] into the

Figure 7

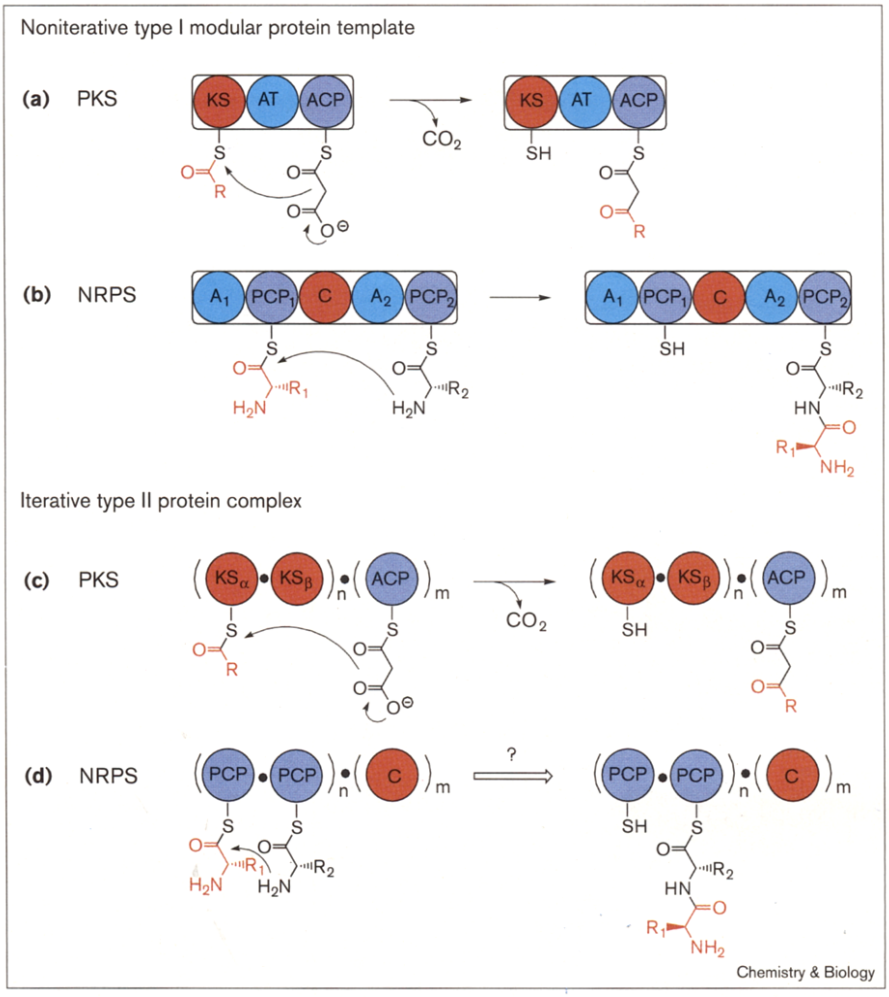


In vitro aminoacylation of holo-BlmI by Val-A. **(a)** Coomassie-Blue-stained 15% SDS–PAGE gel and **(b)** its autoradiogram. Lane 1, Val-A; lane 2, apo-BlmI; lane 3, holo-BlmI; lane 4, apo-BlmI with Val-A; lane 5, holo-BlmI with Val-A; lane 6, molecular weight markers.

4′-phosphopantetheine moiety of the holo-BlmI protein (Figure 5). The identity of BlmI was finally confirmed using MALDI–TOF mass spectral analysis [19] that determined the molecular weight for both the apo- and holo-BlmI proteins.

Although individual domains of type I NRPSs can function independently and several A [9,14,19,20] and PCP [7,19,21–26] domains have been overproduced, purified and biochemically characterized, aminoacylation *in trans* has been successful only between PCPs and their cognate A domains [7,19]. No aminoacylation between PCP and A domains from different NRPS modules has been observed. These results led to the conclusion that there is a specific protein–protein recognition between the A domain and its cognate PCP [19]. Such domain-specific aminoacylation, in fact, should be beneficial in maintaining the fidelity of a type I NRPS by providing additional 'gating' against misincorporation of nonspecifically activated aminoacyl adenylate into the final peptide product. As a type II PCP such as BlmI lacks its cognate A domain, we asked if BlmI could be aminoacylated by an unrelated

Figure 8



Enzyme architecture of **(a,b)** type I PKS and NRPS and **(c,d)** type II PKS and NRPS. A, adenylation domain; ACP, acyl carrier protein or ACP domain; AT, acyl transferase; C, condensation protein or C domain; KSα, β-ketoacyl synthase α subunit; KSβ, β-ketoacyl synthase β subunit; PCP, peptidyl carrier protein or PCP domain.

A domain of a type I NRPS. Although we have yet to determine the biochemical role of BlmI *in vivo*, the fact that the *blmI* gene is located in the middle of the *blm* gene cluster suggests that it may be involved in BLM biosynthesis. To avoid the ambiguity of selecting an A domain that may potentially interact with BlmI *in vivo*, we preferred not to choose any A domain from the *blm* gene cluster to test if it could aminoacylate BlmI *in trans*. We reasoned that an A domain that is unrelated to BlmI should come from a gene cluster independent from BLM biosynthesis and should activate an amino acid not required by BLM. We chose Val-A because it satisfied both requirements. Val-A is an A domain of a type I NRPS from a gene cluster we have cloned previously from *Sv* ATCC15003 that has proven to be unrelated to BLM biosynthesis, and it specifically activates valine among the 23 amino acids tested (Figure 6b). Remarkably, BlmI was efficiently aminoacylated by Val-A. The valine residue is specifically attached in a thioester linkage to the terminal -SH of the 4′-phosphopantetheine moiety of the holo-BlmI protein, as evidenced by the finding that the apo-BlmI was inactive under the identical conditions (Figure 7).

Aminoacylation of holo-BlmI by Val-A represents the first example in which an A domain aminoacylates a protein other than its cognate PCP domain. As it has been suggested that an A domain of a type I NRPS can transfer the activated aminoacyl adenylate only to its cognate PCP domain because of the specific protein–protein recognition between the two domains [19], the finding that BlmI is aminoacylated by Val-A revealed a distinct feature of a type II PCP. It is very tempting to speculate that type II PCPs such as BlmI may have broad intrinsic substrate specificity towards the aminoacyl adenylate, the A domain or both. In fact, the latter feature is reminiscent of the type II PKS ACPs, which have been shown to be interchangeable among different PKS complexes [39–41]. The biosynthesis of D-alanyl-lipoteichoic acid in *Bacillus subtilis* [42] and *Lactobacillus casei* [43] also involves a discrete ACP-like protein — the D-alanyl carrier protein — although it is clearly structurally and functionally different from PCPs.

We want to emphasize that the results we present here fall far short of formulating a model for a type II NRPS and that what we show in Figure 8d is, instead, a hypothesis modeled mainly on type II PKSs [39–41]. Nevertheless, our results strongly suggest the existence of a type II NRPS. In fact, we have already identified two additional genes within the *blm* gene cluster, *blmII* and *blmXI* (Figure 1b), that encode type II C proteins based on sequence analysis (L.D., C. Sanchez, M. Chen, DJ. Edwards, B.S., unpublished observations). We are currently investigating the interactions between these type II C enzymes and the BlmI type II PCP in an attempt to constitute a functional type II NRPS for peptide biosynthesis. These type II NRPS proteins and their distinct biochemical features should now be taken into consideration when combinatorial biosynthesis is used to make novel peptides.

Significance

Nonribosomal peptides are synthesized by nonribosomal peptide synthetases (NRPSs) from a large group of precursors to yield peptides of great structural and functional diversity. Many nonribosomal peptides have clinical applications as, for example, antibacterial, antitumor and immunosuppressive drugs. A greater understanding of peptide synthesis could lead to the manipulation of the process to create novel peptides. All NRPSs known to date are exclusively type I modular enzymes that are multifunctional proteins consisting of domains, such as adenylation (A) [9], peptidyl carrier protein (PCP) [7] and condensation (C) [10] domains, for individual enzyme activities [2–4]. These enzymes control the structural variations of the resulting peptide products by the multiple-carrier thiotemplate mechanism [1,8]. Although individual domains of type I NRPSs can function independently, aminoacylation *in trans* has been successful only between PCPs and their cognate A domains [7,19]. We have cloned and sequenced the *blmI* gene, overproduced and characterized the BlmI protein as a *bona fide* type II PCP, and demonstrated that the holo-BlmI can be aminoacylated by a completely unrelated A domain. Our results provide for the first time genetic and biochemical evidence to support the existence of a type II NRPS, setting the stage for formulating new research concepts to study peptide biosynthesis. Genetic manipulation of type I NRPS has already been successful in generating novel peptides [44]. An unprecedented type II NRPS should shed new light on engineering NRPS proteins, significantly increasing our ability to produce peptides with even greater structural diversities.

Materials and methods

General DNA manipulations

Plasmids preparation and DNA extraction were carried out using commercial kits (Qiagen, Santa Clarita, CA), and all other manipulations were carried out according to standard methods [45]. *E. coli* strain DH5α was used as the host for general DNA propagations.

Overexpression of BlmI *in* E. coli *and purification of the BlmI protein*

The *blmI* gene was amplified from *Sv* ATCC15003 by PCR using a forward primer of 5′-CCGC*CCATGG*GTCTCCGCGTGGCGAGC-GGACCCGGCGC-3′ (the *Nco*I site is in italics) and a reverse primer of 3′-CCT*AGATCT*CCGGTCCCGCTCCCCCGT-5′(the *Bgl*II site is in italics). In order to create the *Nco*I site, the original starting sequence of 'ATG AGC' has been changed to 'ATG GGT', which resulted in the change of the second amino acid from serine to glycine. The first five codons of *blmI* were also optimized for overexpression in *E. coli*. The PCR-amplified 0.3 kb *Nco*I-*Bgl*II fragment was cloned into the similar sites of pQE-60 (Qiagen) to form pBS1. Digestion of pBS1 with *Nco*I and *Hind*III and cloning the resulting 0.3 kb *Nco*I-*Hind*III fragment into the same sites of pET-29a (Novagen, Madison, WI) yielded pBS2.

Expressions of *blmI* in *E. coli* M15 (pREP4)(pBS1) and in *E. coli* BL-21(DE-3)(pBS2) and purification of the resulting BlmI protein by affinity chromatography on Ni-NTA resin were carried out under the standard conditions recommended by Qiagen and Novagen, respectively. The incubation temperature was lowered to 30°C to improve the solubility. The purification of BlmI was monitored by SDS–PAGE on 15% gel. The final pure BlmI protein was desalted on PD-10 column (Sephadex G-25, Pharmacia Biotech, Piscataway, NJ) into 50 mM sodium phosphate buffer, pH 7.8, containing 200 mM NaCl, 10 mM $MgCl_2$, 2 mM dithiothreitol (DTT), 1 mM EDTA, 10% glycerol, and stored at −80 °C for *in vitro* assays.

HPLC analysis and MALDI–TOF mass spectral determination

Samples of BlmI (30–70 μg) purified from *E. coli* OG7001(pBS2) or *E. coli* OG7001(pBS2/pDPT-Gsp) were analyzed on a Nova-Pak C18 column (5 mm × 10, Waters, Milford, MA) using a Rainin DMAX HPLC unit. The column was developed by a linear gradient of 0–50% acetonitrile in 0.1% trifluoroacetic acid in 25 min, followed by additional 5 min at 50% acetonitrile, with a flow rate of 0.6 ml/min and detection at 280 nm. MALDI-TOF mass spectral determination was performed on a Bruker Biflex IIII spectrometer at the Facility for Advanced Instrumentation of University of California, Davis.

In vivo *labeling of BlmI with [3-3H]-β-alanine*

The β-alanine auxotroph *E. coli* strain OG7001 [35] was transformed with pBS2 and cultured under the same conditions as for *E. coli* BL21(DE3) (Novagen). For co-expression of *blmI* with *gsp*, pDPT-Gsp [26] was similarly transformed into *E. coli* OG7001(pBS2) and the transformants were cultured in 2xYT [43] in the presence of kanamycin (25 μg/ml) and chloramphenicol (50 μg/ml). For *in vivo* labeling experiment, cells from 2 ml overnight culture of either *E. coli* OG7001(pBS2) or *E. coli* OG7001(pBS2/pDPT-Gsp) were harvested, washed with M9 minimal medium [43], and re-suspended in 2 ml of M9 minimal medium. The latter were used as seed cultures (20 μl) to inoculate 1 ml M9 medium with kanamycin (25 μg/ml) or kanamycin (25 μg/ml) and chloramphenicol (50 μg/ml) for *E. coli* OG7001(pBS2) or *E. coli* OG7001(pBS2/pDPT-Gsp), respectively. The resulting culture was incubated at 30°C, 250 rpm to OD_{600nm} 0.6 and to this was added 10 μCi of [3-^{3}H]-β-alanine (50 Ci/mmol, American Radiolabeled Chemicals Inc., St. Louis, MO) with or without IPTG (1 mM). Total proteins were resolved by SDS–PAGE on 15% gels that were Coomassie blue-stained. To determine ^{3}H-labeling of the overproduced holo-BlmI protein, gels were soaked in Amplifier (Amersham, Arlington Heights, Il) for 20 min, dried between two sheets of cellulose membrane (KOH Development Inc., Ann Arbor, MI), and visualized by autoradiography on X-ray films (Fuji Medical Systems, Stamford, CT).

In vitro *labeling of BlmI with [3H-pantetheine]-CoA*

Expression of *sfp* in *E. coli* MV1190(pUC8-Sfp), purification of the Sfp PPTase to homogeneity, and 4′-phosphopantetheinylation of apo-BlmI by Sfp *in vitro* were carried out essentially according to literature procedures [24,37]. A typical 100 μl assay solution contained 26 μM apo-BlmI, 2.9 μM Sfp, 25 μM [^{3}H-pantetheine]-CoA (0.9 μCi, 40 Ci/mM), 10 mM $MgCl_2$, and 5 mM DTT, in 75 mM MES/NaOAc buffer, pH 6.0.

After 30 min incubation at 37°C, the assays were stopped by addition of 5 µl of bovine serum albumin (0.2 mg/ml) and 0.9 ml of cold 10% (v/v) trichloroacetic acid (TCA). The precipitated proteins were collected by centrifugation at 14,000 rpm, 20 min, 4°C (Eppendorf 5415C centrifuge), washed with 10% TCA three times, and resolved using SDS–PAGE on 15% gel. The ^{3}H-activity incorporated into holo-BlmI was similarly determined by autoradiography as described for *in vivo* labeling of holo-Blm with [3-^{3}H]-β-alanine.

Overexpression of val-A *in* E. coli *and purification and assay of the Val-A protein*

The *val-A* fragment was amplified from *Sv* ATCC15003 by PCR using a forward primer of 5′-GGAATTC*CATATG*GGCACCACCGTCGC-CGCG-3′ (the *Nde*I site is in italics), and a reverse primer of 3′-GGC*A-AGCTT*GGGACCGGGCGTGGAGCGC (the *Hind*III site is in italics). The PCR-amplified 1.6 kb *Nde*I-*Hind*III fragment was cloned in the similar sites of pET-28a (Qiagen) to yield pBS3. Expression of *val-A* in *E. coli* BL-21(DE-3)(pBS3) and purification of the resulting Val-A protein using affinity chromatography on Ni-NTA resin were carried out under the standard conditions recommended by Novagen.

Amino acid-dependent ATP–PPi assays were performed essentially according to the literature procedures [26,38]. A typical 100 µl assay solution contained 180 nM Val-A, 1 mM ATP, 0.1 mM PPi with 0.2 µCi of ^{32}P-PPi (11.75 Ci/mmol, NEN Life Science Products, Inc., Boston, MA), 1 mM $MgCl_2$, 0.1 mM EDTA, and 1 mM L-amino acid in 50 mM sodium phosphate buffer, pH 7.8. After 30 min incubation at 30 °C, the assays were stopped by addition of 0.9 ml of cold 1% (w/v) activated charcoal in 3% (v/v) perchloric acid. The precipitates were collected on glass fiber filters (2.4 cm, G-4, Fisher, Pittsburgh, PA), washed successively with 10 ml of 0.2 M sodium phosphate buffer, pH 8.0, 4 ml water, and 1 ml of ethanol, and dried in air. The filters were mixed with 7 ml of scintillation fluid (ScintiSafe Gel, Fisher) and counted on a Beckman LS-6800 scintillation counter to determine the radioactivity.

In vitro *aminoacylation of holo-BlmI by Val-A*

The aminoacylation of holo-BlmI was carried out essentially according to literature methods [7,19]. A typical 100 µl assay solution contained 180 nM Val-A, 1.5–2.8 µM apo- or holo-BlmI, 35 µM L-[^{14}C(U)]-valine (283 mCi/mmol, NEN Life Science Products, Inc., Boston, MA), 5 mM ATP, 10 mM $MgCl_2$, and 5 mM DTT in 75 mM Tris–HCl buffer, pH 8.0. The reactions were started by the addition of ATP and, after incubation at 37°C for 30 min, were stopped by addition of 0.9 ml of cold 7% (v/v) TCA. The precipitated proteins were collected by centrifugation at 14,000 rpm, 20 min, 4°C (Eppendorf 5415C centrifuge) and resolved by SDS–PAGE on a 15% gel. The radioactivity incorporated into the holo-BlmI-L-[^{14}C(U)]valine species was similarly determined using autoradiography as described for *in vivo* labeling of holo-BlmI with [3-^{3}H]-β-alanine.

Accession numbers

The *blmI* gene sequence has been deposited in GenBank under the accession number AF149091.

Acknowledgements

We thank Daniel V Santi at University of California, San Francisco for the pDPT-Gsp construct; Otto Geiger at Technical University Berlin, Germany, for the *E. coli* OG7001 strain; Peter Zuber at Oregon Graduate Institute for the pUC8-Sfp construct; Frank Schmidt at University of Missouri, Columbia for clones of the *blmAB* locus; and Cesar Sanchez for critical reading of the manuscript. Mass spectral analysis was performed at Facility for Advanced Instrumentation of University of California, Davis. This work was supported in part by an Institutional Research Grant from the American Cancer Society and the School of Medicine, University of California, Davis, the National Institutes of Health grant AI40475, and the Searle Scholars Program/The Chicago Community Trust.

References

1. Cane, D.E., Walsh, C.T. & Khosla, C. (1998). Harnessing the biosynthetic code: combinations, permutations, and mutations. *Science* **282**, 63-68.
2. Kleinkauf, H. & von Döhren, H. (1996). A nonribosomal system of peptide biosynthesis. *Eur. J. Biochem.* **236**, 335-351.
3. Marahiel, M.A., Stachelhaus, T. & Mootz, H.D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* **97**, 2651-2673.
4. von Döhren, H., Keller, U., Vater, J. & Zocher, R. (1997). Multifunctional peptide synthetases. *Chem. Rev.* **97**, 2675-2705.
5. Wakil, S.J. (1989). Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry* **28**, 4523-4530.
6. Hopwood, D.A. (1997). Genetic contributions to understanding polyketide synthases. *Chem Rev.* **97**, 2465-2497.
7. Stachelhaus, T., Hüser, A. & Marahiel, M.A. (1996). Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chem. Biol.* **3**, 913-921.
8. Stein, T., *et al.*, & Morris, H.R. (1996). The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates. *J. Biol. Chem.* **271**, 15428-15435.
9. Stachlhaus, T., & Marahiel, M.A. (1995). Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *J. Biol. Chem.* **270**, 6163-6169.
10. Stachlhaus, T., Mootz, H.D., Bergendahl, V., & Marahiel, M.A. (1998). Peptide bond formation in nonribosomal peptide biosynthesis, catalytic role of the condensation domain. *J. Biol. Chem.* **273**, 22773-22781.
11. Stachelhaus, T., & Marahiel, M.A. (1995). Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. *FEMS Microbiol. Lett.* **125**, 3-14.
12. Turgay, K. & Marahiel, M.A. (1994). A general approach for identifying and cloning peptide synthetase genes. *Pept. Res.* **7**, 238-241.
13. Quadri, L.E.N., Sello, J., Keating, T.A., Paul, H.W. & Walsh, C.T. (1998). Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chem. Biol.* **5**, 631-645.
14. Konz, D., Klens, A., Schorgendorfer, K. & Marahiel, M.A. (1997) The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. *Chem. Biol.* **4**, 927-937.
15. de Ferra, F., Rodriguez, F., Tortora, O., Tosi, C. & Grandi, G. (1997). Engineering of peptide synthetases. *J. Biol. Chem.* **272**, 25304-25309.
16. Elsner, A., *et al.*, & Bernhard, F. (1997). Substrate specificity of hybrid modules from peptide synthetases. *J. Biol. Chem.* **272**, 4814-4819.
17. Gehring, A.M., Mori, I. & Walsh, C.T. (1998). Reconstitution and characterization of the *Escherichia coli* enterobactin synthetase from EntB, EntE, and EntF. *Biochemistry* **37**, 2648-2659.
18. Gehring, A.M., Mori, I., Perry, R.D. & Walsh, C.T. (1998). The nonribosomal peptide synthetase HMWP2 forms a thiozoline ring during biogensis of yersiniabactin, an iron-chelating virulence factor of *Yersinia pestis*. *Biochemistry* **37**, 11637-11650.
19. Weinreb, P.H., Quadri, L.E.N., Walsh, C.T. & Zuber, P. (1998). Stoichiometry and specificity of *in vitro* phosphopantetheinylation and aminoacylation of the valine-activating module of surfactin synthetase. *Biochemistry* **37**, 1575-1584.
20. Mootz, H.D. & Marahiel, M.A. (1997). The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* **179**, 6843-6850.
21. Pfeifer, E., Pavela-Francis, M., von Döhren, H. & Kleinkauf, H. (1995). Characterization of tyrocidine synthetases 1 (TY1): requirement of posttranslational modification for peptide biosynthesis. *Biochemistry* **34**, 7450-7459.
22. Haese, A., Pieper, R., von Ostrowski, T. & Zocher, R. (1994). Bacterial expression of catalytically active fragments of the multifunctional enzyme enniatin synthetase. *J. Mol. Biol.* **243**, 116-122.
23. Lambalot, R.H., *et al.*, & Walsh, C.T. (1996). A new enzyme superfamily – the phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923-936.
24. Quadri, L.E.N., *et al.*, & Walsh, C.T. (1998). Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **37**, 1585-1595.
25. Gehring, A.M., Lambalot, R.H., Vogel, K.W., Drueckhammer, D.G. & Walsh, C.T. (1996). 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**, 17-24.
26. Ku, J., Mirmira, R.G., Liu, L., & Santi, D.V. (1997). Expression of a functional non-ribosomal peptide synthetase module in *Escherichia coli* by coexpression with a phosphopantetheinyl transferase. *Chem. Biol.* **4**, 203-207.

27. Shen, B., Du, L., Sanchez, C., Chen, M. & Edwards, D.J. (1999). Bleomycin biosynthesis in *Streptomyces verticillus* ATCC15003: a model of hybrid peptide and polyketide biosynthesis. *Bioorg. Chem.* **27**, 155-171.
28. Sugiyama, M., *et al.*, & Davies, J. (1994). Characterization by molecular cloning of two genes from *Streptomyces verticillus* encoding resistance to bleomycin. *Gene* **151**, 11-16.
29. Calcutt, M.J. & Schmidt, F.J. (1994). Gene organization in the bleomycin-resistance region of the producer organism *Streptomyces verticillus*. *Gene* **151**, 17-21.
30. Altschul, S.F., *et al.*, & Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402.
31. Walsh, C.T., Gehring, A.M., Weinreb, P.H., Quadri, L.E.N. & Flugel, R.S. (1997). Post-translational modification of polyketide and nonribosomal peptide synthases. *Curr. Opin. Chem. Biol.* **1**, 309-315.
32. Lambalot, R.H. & Walsh, C.T. (1995). Cloning, overproduction, and characterization of the *Escherichia coli* holo-acyl carrier protein synthase. *J. Biol. Chem.* **270**, 24658-24661.
33. Cox, R.J., *et al.*, & Simpson, T.J. (1997). Post-translational modification of heterologously expressed *Streptomyces* type II polyketide synthase acyl carrier proteins. *FEBS Lett.* **405**, 267-272.
34. Carreras, C.W., Gehring, A.M., Walsh, C.T., & Khosla, C. (1997). Utilization of enzymatically phosphopantetheinylated acyl carrier proteins and acetyl-acyl carrier proteins by the actinorhodin polyketide synthase. *Biochemistry* **36**, 11757-11761.
35. Epple, G., van der Drift, K.M.G.M., Thomas-Oates, J.E. & Geiger, O. (1998). Characterization of a novel acyl carrier protein, RkpF, encoded by an operon involved in capsular polysaccharide biosynthesis in *Sinorhizobium meliloti*. *J. Bacteriol.* **180**, 4950-4954.
36. Spaink, H.P., *et al.*, & Lugtenberg, B.J.J. (1991). A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of rhizobium. *Nature* **354**, 125-130.
37. Nakano, M.M., Corbell, N., Besson, J. & Zuber, P. (1992). Isolation and characterization of *sfp*: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*. *Mol. Gen. Genet.* **232**, 313-321.
38. Lee, S. & Lipmann, F. (1970). Tyrocidine synthetase system. *Methods Enzymol.* **43**, 585-602.
39. Shen, B. & Hutchinson, C.R. (1993). Enzymatic synthesis of a bacterial polyketide from acetyl and malonyl coenzyme A. *Science* **262**, 1535-1540.
40. Bao, W., Wendt-Pienkowski, E. & Hutchinson, C.R. (1998). Reconstitution of the iterative type II polyketide synthase for tetracenomycin F2 biosynthesis. *Biochemistry* **37**, 8132-8138.
41. Carreras, C.W. & Khosla, C. (1998). Purification and *in vitro* reconstitution of the essential protein components of an aromatic polyketide synthase. *Biochemistry* **37**, 2084-2088.
42. Perego, M., Glaser, P., Minutello, A., Strauch, M.A., Leopold, K. & Fischer, W. (1995). Incorporation of *D*-alanine into lipoteichoic acid and wall teichoic acid in *Bacillus subtilis*. *J. Biol. Chem.* **270**, 15598-15606.
43. Debabov, D.V., Heaton, M.P., Zhang, Q., Stewart, K.D., Lambalot, R.H., & Neuhaus, F.C. (1996). The *D*-alanyl carrier protein in *Lactobacillus casei*: cloning, sequencing, and expression of *dltC*. *J. Bacteriol.* **178**, 3869-3876.
44. Stachlhaus, T., Schneider, A., & Marahiel, M.A. (1995). Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* **269**, 69-72.
45. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual* (2nd ed). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

Because *Chemistry & Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cmb – for further information, see the explanation on the contents pages.