Heterologous expression in *Escherichia coli* of the first module of the nonribosomal peptide synthetase for chloroeremomycin, a vancomycin-type glycopeptide antibiotic

John W. Trauger and Christopher T. Walsh*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115

Contributed by Christopher T. Walsh, December 21, 1999

The gene cluster from *Amycolotopsis orientalis* responsible for biosynthesis of the vancomycin-type glycopeptide antibiotic chloroeremomycin was recently sequenced, indicating that this antibiotic derives from a seven-residue peptide synthesized by a three-subunit (CepA, CepB, and CepC) modular nonribosomal peptide synthetase. Expression of all or parts of the peptide synthetase in *Escherichia coli* would facilitate biochemical characterization of its substrate specificity, an important step toward the development of more potent glycopeptides by combinatorial biosynthesis. To determine whether CepA, a three-module 3,158-residue peptide synthetase expected to assemble the first three residues of the heptapeptide precursor, could be heterologously expressed in *E. coli* and converted to active, holo form by posttranslational priming with a phosphopantetheinyltransferase, we expressed two CepA fragments (CepA1-575 and CepA1-1596) as well as full-length CepA (CepA1-3158). All three constructs were expressed in soluble form. We find that the CepA1-575 fragment, containing adenylation and peptidyl carrier protein domains (A1-PCP1), specifically adenylates L-leucine and D-leucine in a 6:1 ratio, and it can be converted to holo form by the phosphopantetheinyltransferase Sfp; also, we find that holo-CepA1-575 can be covalently aminoacylated with L-leucine on the peptidyl carrier protein 1 domain. However, no amino acid-dependent adenylation or aminoacylation activity was detected for the larger CepA constructs with L-leucine or other expected amino acid substrates, suggesting severe folding problems in the multidomain proteins.

The vancomycin family of glycopeptide antibiotics are characterized by mono- or disaccharides attached to an aglycone core consisting of a heptapeptide in which aromatic amino acid side chains have been oxidatively cross-linked, forming several macrocycles that impart a rigid architecture (Fig. 1). The molecules function by binding to *N*-acyl-D-Ala-D-Ala termini in bacterial cell wall precursors. This interaction prevents cross-linking of peptidoglycan chains, which disrupts the cell wall and leads to bacterial cell death by osmotic lysis (1). Vancomycin and teicoplainin have been front line drugs against methicillin-resistant *Staphylococcus aureus* (2–4). The emergence of glycopeptide-resistant bacteria, first documented in 1988, has resulted in sometimes fatal infections that are untreatable by any current antibiotic regimen. The emergence of such resistant bacteria has generated great interest in the preparation of new semisynthetic glycopeptides (3, 4). However, the structural complexity of glycopeptides limits the range of possible modifications (4). The fact that vancomycin-type antibiotics are synthesized by a modular nonribosomal peptide synthase (NRPS) (5) suggests that engineered biosynthesis (6, 7) could complement chemical synthesis and provide a useful approach to new molecules of this class.

The genes for a number of NRPSs have been sequenced, revealing the presence of conserved modules (8, 9). The number and order of modules correspond directly to the amino acid sequence of the peptide product. Studies to date have revealed the basic sequence of events leading to peptide bond formation (8–10): (*i*) an adenylation domain (A) (≈550 aa) within each module adenylates a specific amino acid; (*ii*) the carboxyl-activated amino acid then forms a thioester with the thiol group of a phosphopantetheine cofactor located in an adjacent peptidyl carrier protein (PCP) domain (≈100 aa); and (*iii*) condensation (C) domains (≈430 aa) located between PCP domains catalyze peptide bond formation. Peptide synthesis occurs in the N to C direction. A thioesterase domain (≈250 aa) catalyzes cleavage of the completed peptide from the synthetase. Modules may also contain epimerization (for L- to D- conversions) and N-methylation modification domains.

The conversion of inactive, apo-PCP domains to the holo-, phosphopantetheine-containing form is catalyzed by phosphopantetheinyltransferases (PPTases) (11, 12). PPTases are generally specific for particular cognate carrier domains. An exception to this trend is Sfp, a PPTase involved in the biosynthesis of the lipoheptapeptide surfactin, which has been shown to transfer phosphopantetheine to a wide range of carrier proteins (13).

The biosynthetic gene cluster for the glycopeptide antibiotic chloroeremomycin from *Amycolotopsis orientalis* was recently sequenced (5). Chloroeremomycin and vancomycin share identical aglycone structures, and differ only in their glycosylation patterns (Fig. 1). The chloroeremomycin gene cluster includes three peptide synthetase proteins, CepA, CepB, and CepC, that encode seven amino acid-activating modules (Fig. 2*A*). Our first focus is on the initiating NRPS subunit CepA, which is expected to assemble the first three residues of the chloroeremomycin heptapeptide precursor. To determine whether fragments of CepA and full-length CepA could be expressed in *Escherichia coli* and converted to active, holo- form by using Sfp, we cloned and expressed three constructs: CepA1-575 (A1-PCP1), CepA1-1596 (A1-PCP1-C-A2-PCP2), and full-length CepA1-3158 (Fig. 2*B*). The CepA domains should permit determination of the timing of epimerization of Leu-1 and Tyr-2, the timing of N-methylation of D-Leu-1, and the timing of hydroxylation and chlorination of Tyr-2.

Materials and Methods

Materials and Recombinant DNA Techniques. Competent cells of *E. coli* strains DH5α and XL10-Gold were purchased from

Abbreviations: A, adenylation; C, condensation; E, epimerase; NRPS, nonribosomal peptide synthetase; PCP, peptidyl carrier protein; PPTase, phosphopantetheinyltransferase; TCA, trichloroacetic acid.

*To whom reprint requests should be addressed. E-mail: walsh@walsh.med.harvard.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.040560597. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.040560597

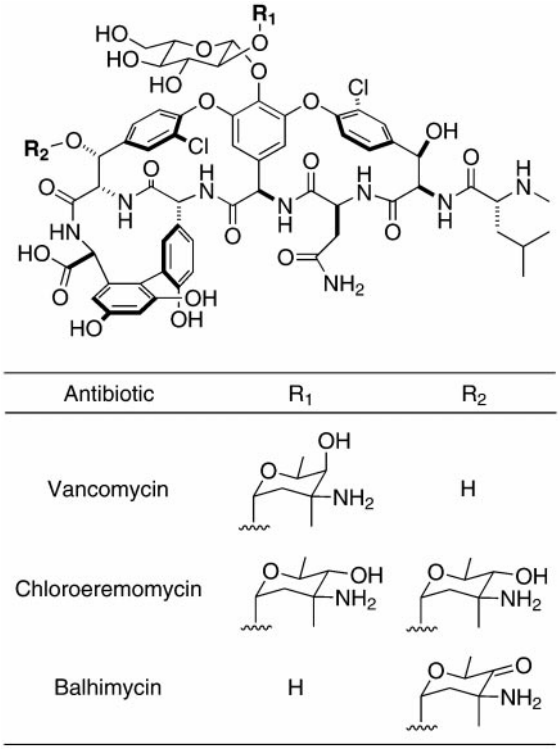

Fig. 1. Structures of the glycopeptide antibiotics vancomycin, chloroeremomycin, and balhimycin.

GIBCO/BRL and Stratagene, respectively. Competent cells of *E. coli* strains BL21(DE3) and BL21(DE3)pLysS were from Novagen. Sfp (13) and genomic DNA from *A. orientalis* strain NRRL 18098 (5) were prepared as described. CoA was purchased from Sigma. [^{3}H]CoA disulfide (≈70% label in the phosphopantetheinyl moiety) was prepared by DuPont/NEN. Restriction endonucleases and T4 DNA ligase were obtained from New England Biolabs. PCRs were carried out by using *Pfu* DNA polymerase (Stratagene) as described by the enzyme supplier, except for the addition of 10% DMSO. Plasmid DNA preparation, gel extraction of DNA fragments, and purification of DNA amplified by PCR were performed by using QIAprep and QIAquick kits from Qiagen (Chatsworth, CA). The sequences of PCR-amplified DNA fragments were determined by nucleotide sequencing after cloning into the expression vector pET22b (Novagen). DNA sequencing was performed by the Molecular Biology Core Facility of the Dana–Farber Cancer Institute (Boston). Oligonucleotide primers were obtained from Integrated DNA Technologies (Coralville, IA), and are listed in Table 1.

Cloning, Expression, and Purification of CepA Constructs. The CepA1-575 gene fragment was amplified from *A. orientalis* NRRL 18098 genomic DNA by using the primers AT1-F and AT1-R. The amplified fragment was digested with *Nde*I and *Xho*I, gel purified, and ligated to the large pET22b *Nde*I/*Xho*I fragment to generate the plasmid pCepA1-575-His$_6$. After isolation of the plasmid from DH5α cells, the plasmid was transformed into *E. coli* BL21(DE3)pLysS cells. For protein overproduction, the *E. coli* strain BL21(DE3)pLysS/pCepA1-575-His$_6$ was cultivated with shaking at 24°C in 1 liter of LB medium containing 100 μg/ml ampicillin and 34 μg/ml chloramphenicol and induced with 0.2 mM isopropyl β-D-thiogalactoside when the culture reached an OD_{600} of 0.5. After induction, incubation was continued for 5 h. The cells were then harvested by centrifugation (10 min at 2,000 $\times$ *g*) and resuspended in 20 ml of buffer A (20 mM Tris·HCl/0.3 M NaCl, pH 8). After resuspension, the cells were lysed by two passages through a French press (18,000 psi). Cellular debris was removed from the lysate by centrifugation (45 min at 50,000 $\times$ *g*). CepA1-575-His$_6$ was purified by affinity chromatography by using Ni-NTA Superflow resin (Qiagen). Fractions containing the desired protein were pooled and dialyzed against buffer B (50 mM Tris·HCl/100 mM NaCl/10 mM $MgCl_2$/1 mM DTT/1 mM EDTA/10% glycerol, pH 8). After dialysis, the protein was aliquoted, flash frozen in liquid nitrogen, and stored at −80°C. DNA corresponding to CepA1-1596 was amplified from *A. orientalis* genomic DNA by using the primers AT1-F and AT2-R. The fragment was digested with *Nde*I and *Hin*dIII, ligated to the

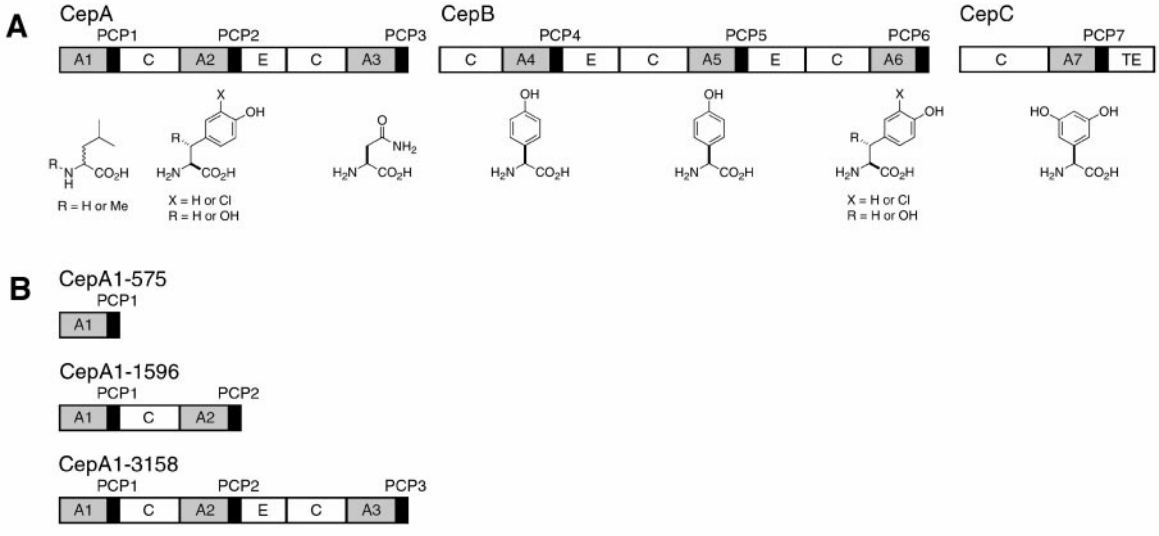

Fig. 2. (*A*) Subunit and domain arrangement of the chloroeremomycin nonribosomal peptide synthetase. The following domains are indicated: adenylation (A), peptidyl carrier protein (PCP), condensation (C), epimerization (E), and thioesterase (TE). The expected amino acid precursors for each A domain are indicated. The identities of the amino acids activated by the first, second, and sixth A domains (A1, A2, and A6) were not clear at the outset: L-Leu may be methylated and/or epimerized before adenylation, and L-Tyr may be hydroxylated and/or chlorinated before adenylation. (*B*) Proteins cloned and expressed in *E. coli* in this study.

Table 1. Oligonucleotide primers

Primer	Sequence
AT1-F	5′-GAATTCCATATGACGATGACTTCCGCA-3′
AT1-R	5′-TATTCTCGAGCCTGTCGGCCGGTACCGCA-3′
AT2-R	5′-CCCAAGCTTCTCCGACGGCGCCGATTCCA-3′
4300F-*Eco*RI	5′-GAGATCGAATTCGCCCTGTCCG-3′
4300R-*Eco*RI	5′-ACAGGGCGAATTCGATCTC-3′
7470F-*Nhe*I	5′-ACCGCCCAGGCTAGCGCCGTG-3′
7470R-*Nhe*I	5′-AGCACGGCGCTAGCCTGGGCGGT-3′
CepAR-Δ*Eco*RI	5′-CCCAAGCTTACGGTCGCTGTCGAACTCGGACTCGAACTCG-3′

large pET22b *Nde*I/*Hin*dIII fragment, and transformed into XL10-Gold cells affording the plasmid pCepA1-1596-His$_6$. The plasmids pHis$_6$-CepA1-1596-His$_6$, pMBP-CepA1-1596-His$_6$, and pCepA1-1596-CBD-His$_8$ were prepared by subcloning the pCepA1-1596-His$_6$ *Nde*I/*Hin*dIII insert into *Nde*I/*Hin*dIII-digested pET24b (Novagen), pIADL16 (14), and pET38b (Novagen), respectively. CepA1-3158 (full-length CepA) was amplified in three fragments, with a silent *Eco*RI site added at the first junction, a silent *Nhe*I site added at the second junction, and two *Eco*RI sites near the 3′ end of the gene removed by silent mutagenesis, by using the primers AT1-R and 4300R-*Eco*RI (for the first segment), 4300F-*Eco*RI and 7470R-*Nhe*I (for the second segment), and 7470F-*Nhe*I and CepAR-Δ*Eco*RI (for the third segment). The amplified fragments were digested with the appropriate enzymes and purified. The *Nde*I/*Eco*RI fragment (first segment) was ligated to the large pET22b *Nde*I/*Eco*RI fragment and transformed into XL10-Gold cells. The resulting plasmid was digested by using *Eco*RI and *Hin*dIII, ligated to the second and third segments in a three-part ligation reaction, and transformed into XL10-Gold cells affording the plasmid pCepA1-3158-His$_6$. For protein overproduction, the expression plasmids were transformed into *E. coli* strain BL21(DE3)pLysS and grown in 1 liter of LB medium at 18°C containing 100 μg/ml ampicillin (or 50 μg/ml kanamycin for pMBP-CepA1-1596-His$_6$) and 34 μg/ml chloramphenicol. The cells were induced with 0.2 mM isopropyl β-D-thiogalactoside at an OD_{600} of 0.5 and grown for 5 h at 18°C, and then harvested by centrifugation (10 min at 2,000 × *g*) and resuspended in 20 ml of buffer A. Cell lysis and protein purification were carried out as described above for CepA1-575 with the following exceptions: (*i*) for Ni^{2+} affinity purifications (all proteins except MBP-CepA1-1596), cell lysates were incubated with Ni-NTA resin for 12 h at 4°C and the purified protein concentrated by using a Centriprep 50 concentrator (Millipore), and (*ii*) MBP-CepA1-1596 was purified by using amylose resin (NEB, Beverly, MA).

ATP-PP$_i$ Exchange Assays. ATP-PP$_i$ exchange was assayed as described with minor modifications (15). The assay mixture (100 μl) contained 2 μM CepA1-575-His$_6$, 2 μM CepA1-1596-His$_6$, or 0.2 μM CepA1-3158-His$_6$, and 2 mM amino acid/2 mM ATP/1 mM sodium [^{32}P]PP$_i$ (300,000 cpm)/75 mM Tris·HCl/100 mM NaCl/10 mM $MgCl_2$/5 mM DTT, pH 8. After incubation at 23, 30, or 37°C for 3 h, the reaction was quenched by the addition of 0.5 ml of a charcoal suspension [1.6% (wt/vol) activated charcoal/0.1 M sodium pyrophosphate/0.35 M perchloric acid]. The charcoal was pelleted by centrifugation, washed twice with 1 ml of 0.1 M sodium pyrophosphate/0.35 M perchloric acid, resuspended in 0.5 ml H_2O, and added to a scintillation vial containing 3.5 ml of scintillation fluid, and the bound radioactivity was determined by liquid scintillation counting. Optimal exchange activity for CepA1-575-His$_6$ was observed at 23°C, with two- and threefold lower activity at 30°C and 37°C, respectively.

CepA1-575-His$_6$ Phosphopantetheinylation. The radioassay for determination of PPTase activity has been described (11). The apoprotein (6 μM) and [^{3}H]CoA disulfide (50 μM; 74 μCi/μmol) were incubated for 90 min at 30°C with Sfp (65 nmol) in a 100-μl reaction mixture containing 75 mM Tris·HCl, 50 mM NaCl, 5 mM $MgCl_2$, 25 mM DTT, 5% glycerol at pH 8. Reactions were quenched by the addition of 0.8 ml of 10% trichloroacetic acid (TCA) and BSA (0.38 mg). Precipitated protein was pelleted by centrifugation, and the pellets were washed three times with 0.8 ml of 10% TCA, dissolved in 150 μl of formic acid, and mixed with 3.5 ml of liquid scintillation cocktail. Incorporated radioactivity was quantified by liquid scintillation counting.

CepA1-575-His$_6$ Aminoacylation. Purified apo-CepA1-575 was converted to the holo- form as described above, except that unlabeled CoA (0.1 mM) was used and DTT was excluded. ATP (to 10 mM) and L-[^{3}H]leucine (to 1.5 Ci/mmol, 0.2 mM) were then added, and the reaction was allowed to proceed at 23°C for 0.5–80 min. Reactions were quenched by the addition of 0.38 mg of BSA and 0.8 ml of 10% TCA. The amount of radioactivity incorporated was determined as described above.

Aminoacylation of CepA Proteins in Cell Lysates. Clarified cell lysates containing CepA1-575-His$_6$, CepA1-1596-His$_6$, or CepA1-3158-His$_6$ were prepared as described above, except that the cells were lysed into 50 mM sodium phosphate/0.3 M NaCl/20% glycerol, pH 7.5. For some experiments, BL21(DE3)/pGroES/EL cells were used for protein production. The cell lysate (50 μl), 65 nM Sfp, and 0.1 mM CoA were incubated at 30°C for 45 min in a 120-μl reaction containing 50 mM sodium phosphate, 125 mM NaCl, 5 mM $MgCl_2$, 10% glycerol at pH 7.5. ATP (to 10 mM) and L-[^{3}H]leucine (to 1.5 Ci/mmol; 0.2 mM) were added, and the reaction was incubated for 2 h at 23°C. The reaction was then quenched by the addition of 0.8 ml of 10% TCA, and the precipitated proteins were recovered by centrifugation. The cell pellets were washed three times with 0.8 ml of 10% TCA and then resuspended in 40 μl of SDS/PAGE loading buffer (without DTT) titrated to neutral pH by using 1 M Tris base. The mixture was run on a 10% Tris·glycine gel, and the gel was stained with Coomassie Blue, soaked in Amplify (Amersham International) for 30 min, and exposed to film for 5 days at −80°C.

Results and Discussion

CepA1-575 (A1-PCP1) L- and D-Leucine Activation. The DNA sequence encoding amino acid residues 1–575 of CepA was amplified from *A. orientalis* genomic DNA and cloned into the expression vector pET22b. The resulting construct encodes the first A and PCP domains (A1-PCP1) of CepA (Fig. 2) followed by a six-His purification tag. This construct overexpressed well to provide soluble CepA1-575-His$_6$ that was purified to >90% homogeneity by a single Ni-NTA purification step (Fig. 3*A*).

The N-terminal amino acid in chloroeremomycin (and vancomycin) is D-*N*-methylleucine. However, there are no internal epimerization or N-methylation domains located adjacent to the PCP1 domain of CepA, raising the question of when and how leucine epimerization and N-methylation occur during chloroeremomycin biosynthesis. The amino acid adenylation specificity of the A1 domain of CepA1-575 was studied by ATP-PP$_i$ exchange assays, revealing that A1 specifically activates both L-Leu and D-Leu, with a sixfold preference for L- versus D-Leu (Fig. 4*A*). No activation of L- or D-*N*-methylleucine is observed. We note that the observed exchange activity is very weak (1–5%) by catalytic efficiency estimate compared with many other NRPS A domains that have been biochemically characterized (16, 17). For example, in parallel reactions, the activity for CepA1-575-His$_6$ with its preferred substrate L-Leu was 2% that of the EntF protein (from enterobactin synthetase) with its preferred sub-

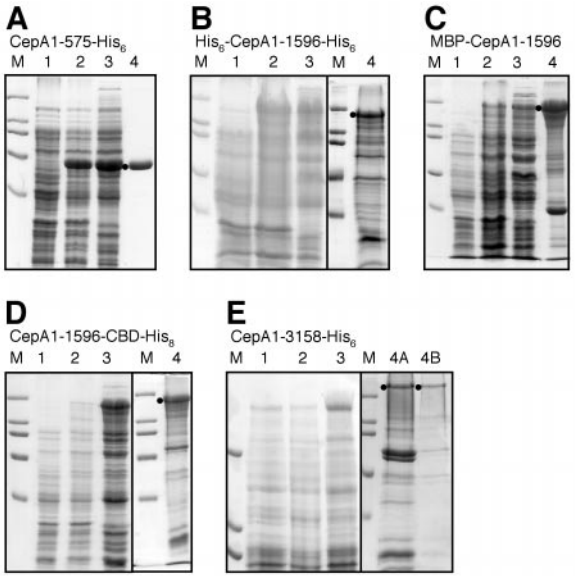

Fig. 3. Expression of CepA constructs in *E. coli*: (*A*) CepA1-575-His_6; (*B*) His6-CepA1-1596-His_6; (*C*) MBP-CepA1-575-His_6; (*D*) CepA1-1596-CBD-His_8; and (*E*) CepA1-3158-His_6. For all gels: lanes 1 and 2, whole-cell lysates from uninduced and isopropyl β-D-thiogalactoside-induced cultures, respectively, of *E. coli* strain BL21(DE3)pLysS; lane 3, soluble fraction from the induced cells; and lane 4, protein purifed by Ni^{2+}-affinity chromatography (or amylose affinity chromatography for MBP-CepA1-575-His_6). M, molecular mass markers (200, 116, 97, 66, and 45 kDa). Protein samples were resolved by 10% Tris·glycine SDS/PAGE except for *E Left*, for which a 5% gel was used.

strate L-Ser. This result suggests that most of the heterologously expressed CepA1-575 A1 domains are partially misfolded.

The observation that L- and D-*N*-methylleucine are not adenylated by the A1 domain indicates that Leu methylation occurs after adenylation. This result is consistent with recent studies by Jung *et al.* (18, 19) with a strain of the balhimycin (Fig. 1) producer *Amycolotopsis mediterranei* in which the oxidative cross-linking genes were deleted, resulting in the production of a linear heptapeptide precursor. Characterization of this heptapeptide showed that the Leu residue was unmethylated, indicating that Leu methylation occurs after heptapeptide synthesis and oxidative cyclization.

With respect to the timing of Leu L- to D- epimerization, Jung *et al.* (19) found that the balhimycin heptapeptide precursor contained D-Leu (19), indicating that Leu epimerization occurs before oxidative cyclization, either at the stage of free Leu, Leu-S-PCP, peptide-S-PCP, or free linear peptide. There are two precedents that may be relevant to the CepA A1 initiation module selection and utilization of D-Leu. One is the A domain of the first module of gramicidin S synthetase. It activates both D- and L-Phe and makes the D- and L-Phe-S-PCP domains (20); however, only L-Phe is available in the producing *Bacillus subtilis* cell, so the physiological flux is L-Phe to L-Phe-S-PCP, and then epimerization by an epimerase (E) domain (A-PCP-E module) to D-Phe-S-PCP before transfer by a D-selective C domain (21). Since there is no E domain in the first module of CepA, this L-Leu-S-PCP to D-Leu-S-PCP route seems unlikely although an E domain that acts in trans could be involved. Two candidates for such epimerization of L-Leu-S-PCP1 are ORFs 29 and 30 in the chloroeremomycin gene cluster, which have homology to carnitinoyl CoA epimerase (22). The other precedent is in cyclosporin A (CsA) synthetase where the first residue Ala-1 has the D configuration although there is no E domain in the first module of CsA synthetase (23). Instead, there is a separate alanine racemase that provides D-Ala to the A-1 domain (24).

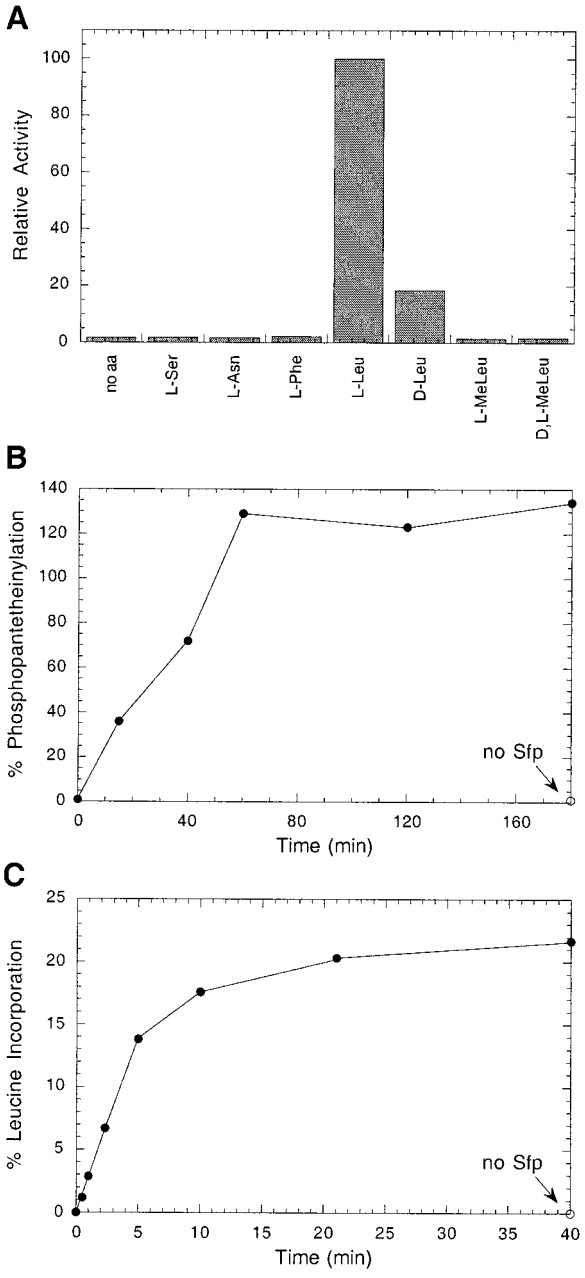

Fig. 4. Characterization of CepA1-575-His_6. (*A*) Amino acid activation specificity determined by ATP-PP_i exchange assay. Reactions contained 2 μM CepA1-575-His_6, 2 mM amino acid, 75 mM Tris·HCl, 2 mM ATP, 1 mM sodium [^{32}P]pyrophosphate (300,000 cpm), 100 mM NaCl, 10 mM $MgCl_2$, 5 mM DTT at pH 8 (2 h at 23°C). (*B*) Covalent incorporation of [^{3}H]phosphopantetheine catalyzed by the PPTase Sfp. The reactions contained 6 μM CepA1-575-His_6, 65 nM Sfp, 50 μM [^{3}H]CoA disulfide (75 mCi/mmol), 40 mM Tris·HCl, 50 mM NaCl, 5 mM $MgCl_2$, 25 mM DTT, 5% glycerol at pH 8. (*C*) Covalent incorporation of L-[^{3}H]Leu into holo-PCP1. Reactions contained final concentrations of 6 μM CepA1-575-His_6, 65 nM Sfp, 0.1 mM CoA, 0.2 mM L-[^{3}H]Leu (1.5 Ci/mmol), 10 mM ATP, 60 mM Tris·HCl, 50 mM NaCl, 10 mM $MgCl_2$, 5% glycerol at pH 8.

This route would seem to leave CepA with only one-sixth of its A1 domains loaded with the correct Leu-AMP isomer, even when free D-Leu was made available, so the first route seems more likely but needs to be established.

CepA1-575 Covalent Acyl-S-PCP Formation. NRPS proteins heterologously expressed in *E. coli* have been inefficient at covalent

aminoacyl thioester formation on PCP domains because of the lack of efficient phosphopantetheinylation by endogenous *E. coli* PPTases. The identification of the PPTase family by Lambalot *et al.* (11), coupled with the finding that the PPTase Sfp from surfactin synthetase has broad substrate specificity, allows production of holo-PCP domains after expression in *E. coli*. Treatment of CepA1-575 with CoA and Sfp resulted in efficient phosphopantetheine incorporation (Fig. 4*B*). Upon addition of L-Leu and ATP to holo-CepA1-575, covalent attachment of L-Leu to the protein is observed (Fig. 4*C*). A stoichiometry of 20% acylation is observed. No L-[^{3}H]leucine incorporation was detected in the absence of Sfp, indicating that the PCP domain is not phosphopantetheinylated during expression in *E. coli*. Radioactive D-Leu is not available to test the prediction that it should also be transferred from the A1 domain, but it did inhibit covalent loading of labeled L-Leu (data not shown).

Expression of CepA1-1596 and CepA1-3158. By either mechanism for L-Leu- versus D-Leu-S-PCP1 formation, the prediction is that the first C domain of CepA (located between PCP1 and A2) should be D-selective both for the donor D-Leu (on PCP1) and the acceptor D-Tyr (on PCP2). To test this prediction, DNA constructs encoding the CepA1-1596 fragment and full-length CepA were amplified by PCR from *A. orientalis* genomic DNA and cloned into pET22b, which appends a C-terminal six-His tag. Both constructs were overexpressed in *E. coli*, resulting in the production of soluble proteins with the expected molecular weights. CepA1-3158-His$_6$ was partially purified by Ni-NTA affinity chromatography (Fig. 3*E*). In the case of CepA1-1596-His$_6$, we were unable to purify the protein by Ni-NTA chromatography, despite the use of a batch method in which the protein was incubated with Ni-NTA resin overnight at 4°C, presumably because of inaccessibility of the six-His tag. To facilitate purification of this fragment, constructs encoding CepA1-1596 with the following purification tags were prepared: (*i*) an N-terminal six-His tag; (*ii*) an N-terminal maltose-binding protein tag; and (*iii*) a C-terminal chitin-binding domain-His$_8$ tag. Overexpression of all three constructs followed by Ni-NTA purification, or amylose column purification for the maltose-binding protein-tagged protein, provided partially purified CepA1-1596 proteins (Fig. 3 *B–D*).

In contrast to our results with CepA1-575, no ATP-PP$_i$ exchange activity significantly above background was detected with partially purified CepA1-1596 proteins and CepA1-3158 with any of the expected amino acid substrates that were tested (L-Leu, L-Tyr, L-3-chlorotyrosine, L-β-hydroxytyrosine, and L-Asn). This result is consistent with the unusually low activity observed for CepA1-575, although it is not clear why activity was detectable for the smaller construct but not for the larger constructs. In the case of CepA1-3158-His$_6$, the protein concentration used in the ATP-PP$_i$ exchange assay was 10-fold lower (0.2 μM) than that used for CepA1-575-His$_6$, which may account for our inability to detect activity for this protein. However, further concentration of the partially purified protein would result in high background exchange activity because of contaminating proteins. The lack of detectable activity may result from misfolding, partial unfolding, or by formation of soluble aggregates during protein expression or purification. Similar results (soluble, yet inactive NRPS proteins) were recently observed in studies of mycobactin biosynthesis (L. E. N. Quadri and C.T.W., unpublished results). Until the problems are solved, we cannot ascertain the D, L-stereoselectivity of the C domains in CepA for aminoacyl and peptidyl donors and acceptors.

One possibility for the lack of activity observed with larger CepA proteins is loss of activity caused by lability during the purification procedure. In addition, the presence of an N-terminal affinity tag may interfere with the folding or activity of the protein. To assay CepA proteins without N-terminal tags immediately after cell lysis, we attempted to detect L-Leu aminoacylation of CepA1-575-His$_6$, CepA1-1596-His$_6$, and CepA1-3158-His$_6$ in cell lysates. The CepA proteins were produced in *E. coli* and lysed into glycerol phosphate buffer (50 mM sodium phosphate/0.3 M NaCl/20% glycerol, pH 7.5). The clarified lysate was treated first with Sfp and CoA, and then with L-[^{3}H]Leu and ATP. The proteins were then precipitated by using TCA and separated by SDS/PAGE. Whereas incorporation of L-Leu into CepA1-575-His$_6$ was clearly detected, no L-Leu incorporation above background into CepA1-1596-His$_6$ or CepA1-3158-His$_6$ was detected (Fig. 5). Coexpression of the *E. coli* chaperone GroES/EL has been shown in one case to facilitate the expression of large multidomain NRPS proteins (25). To assess whether GroES/EL coexpression would affect the activity of CepA proteins, the cell lysate-labeling experiment was repeated after expression of the constructs in BL21(DE3)/pGroES/EL (26). As in the initial experiment, L-Leu incorporation was detected for CepA1-575-His$_6$ but not for CepA1-1596-His$_6$ or CepA1-3158-His$_6$ (data not shown).

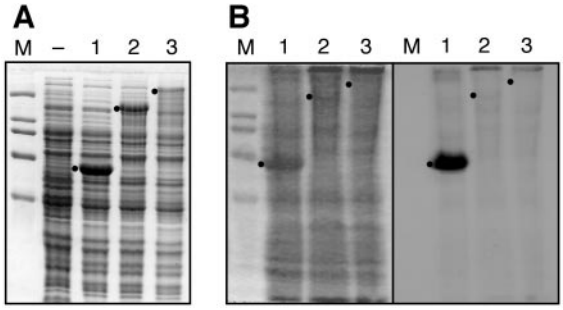

Fig. 5. Covalent incorporation of L-[^{3}H]Leu into CepA proteins present in *E. coli* cell lysates. (*A*) Coomassie-stained gel showing expression of CepA1-575-His$_6$ (lane 1); CepA1-1596-His$_6$ (lane 2); and CepA1-3158-His$_6$ (lane 3). (*B Left*) Coomassie-stained gel of cell lysates after treatment with Sfp, CoA, $MgCl_2$, ATP, and L-[^{3}H]Leu; and (*Right*) an autoradiogram of the same gel. L-Leu incorporation is detected for CepA1-575-His$_6$-containing lysate (lane 1) but not for the larger proteins CepA1-1596-His$_6$ (lane 2) and full-length CepA1-3158-His$_6$ (lane 3). The expected or observed positions of the three proteins on the gels are indicated. Reactions contained (in a total volume of 120 μl) 50 μl lysate, 65 nM Sfp, 0.1 mM CoA, 10 mM ATP, 0.2 mM L-[^{3}H]Leu (1.5 Ci/mmol), 50 mM sodium phosphate, 125 mM NaCl, 5 mM $MgCl_2$, 10% glycerol at pH 7.5.

Implications for Expression of Glycopeptide NRPS Genes in *E. coli*. Our results with the CepA1-575 fragment demonstrate that, in principle, the *A. orientalis* NRPS Cep proteins can be expressed in *E. coli* and PCP domains converted to active, holo- form by using the PPTase Sfp. However, the relatively low activity observed with CepA1-575 and the lack of detectable activity for CepA1-1596 and CepA1-3158 (even for the A1-PCP1 pair) suggests that the chloroeremomycin NRPS proteins, despite efficient expression and solubility, are incorrectly folded or are aggregated after expression in *E. coli*. Expression of the chloroeremomycin NRPS genes will require either improved procedures for protein expression and purification in *E. coli* or the use of an alternate heterologous host such as *Streptomyces*. Alternatively, it may be that NRPS proteins from another vancomycin-type glycopeptide producer such as the balhimycin producer *A. mediterranei* may be more amenable to expression and purification in *E. coli*.

We gratefully acknowledge Luis Quadri for providing Sfp and Il-Seon Park for providing *A. orientalis* genomic DNA. This work was supported by grants to C.T.W. from the National Institutes of Health (GM-49338) and Schering-Plough. J.W.T. was supported by a National Institutes of Health Postdoctoral Fellowship.

1. Barna, J. C. J. & Williams, D. H. (1984) *Annu. Rev. Biochem.* **38,** 339–357.
2. Walsh, C. T., Fisher, S. L., Park, I. S., Prahalad, M. & Wu, Z. (1996) *Chem. Biol.* **3,** 21–28.
3. Malabarba, A., Nicas, T. I. & Ciabatti, R. (1997) *Eur. J. Med. Chem.* **32,** 459–478.
4. Nicas, T. I., Mullen, D. L., Flokowitsch, J. E., Preston, D. A., Snyder, N. J., Zweifel, M. J., Wilkie, S. C., Rodriguez, M. J., Thompson, R. C. & Cooper, R. D. (1996) *Antimicrob. Agents Chemother.* **40,** 2194–2199.
5. Wageningen, A. M. A., Kirkpatrick, P. N., Williams, D. H., Harris, B. R., Kershaw, J. K., Lennard, N. J., Jones, M., Jones, S. J. M. & Solenberg, P. J. (1998) *Chem. Biol.* **5,** 155–162.
6. Stachelhaus, T., Schneider, A. & Marahiel, M. A. (1995) *Science* **269,** 69–72.
7. Khosla, C. (1997) *Chem. Rev. (Washington, D.C.)* **97,** 2577–2590.
8. Marahiel, M. A., Stachelhaus, T. & Mootz, H. D. (1997) *Chem. Rev. (Washington, D.C.)* **97,** 2651–2673.
9. von Doehren, H., Keller, U., Vater, J. & Zocher, R. (1997) *Chem. Rev. (Washington, D.C.)* **97,** 2675–2705.
10. Stein, T., Vater, J., Kruft, V., Otto, A., Wittmann-Liebold, B., Franke, P., Panico, M., McDowell, R. & Morris, H. R. (1996) *J. Biol. Chem.* **271,** 15428–15435.
11. Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., Lacelle, M., Marahiel, M. A., Reid, R., Khosla, C. & Walsh, C. T. (1996) *Chem. Biol.* **3,** 923–936.
12. Walsh, C. T., Gehring, A. M., Weinreb, P. H., Quadri, L. E. N., Flugel, R. S. & Walsh, C. T. (1998) *Curr. Opin. Chem. Biol.* **1,** 309–315.
13. Quadri, L. E. N., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P. & Walsh, C. T. (1998) *Biochemistry* **37,** 1585–1595.
14. McCafferty, D. G., Lessard, I. A. D. & Walsh, C. T. (1997) *Biochemistry* **36,** 10498–10505.
15. Rusnak, F., Sakaitani, M., Druekhammer, D., Reichert, J. & Walsh, C. T. (1991) *Biochemistry* **30,** 2916–2927.
16. Mootz, H. D & Marahiel, M. A. (1997) *J. Bacteriol.* **179,** 6843–6850.
17. Reichert, J., Sakaitani, M. & Walsh, C. T. (1992) *Protein Sci.* **1,** 549–556.
18. Pelzer, S., Suessmuth, R., Heckmann, D., Recktenwald, J., Huber, P., Jung, G. & Wohlleben, W. (1999) *Antimicrob. Agents Chemother.* **43,** 1565–1573.
19. Suessmuth, R., Pelzer, S., Nicholson, G., Walk, T., Wohlleben, W. & Jung, G. (1999) *Angew. Chem. Int. Ed. Engl.* **38,** 1976–1979.
20. Kraetzschmar, J., Krause, M. & Marahiel, M. A. (1989) *J. Bacteriol.* **171,** 5422–5429.
21. Gevers, W., Kleinkauf, H. & Lipmann, F. (1969) *Proc. Natl. Acad. Sci. USA* **63,** 1335–1342.
22. Eichler, K., Bourgis, F., Buchet, A., Kleber, H. P. & Mandrand-Berthelot, M. A. (1994) *Mol. Microbiol.* **13,** 775–786.
23. Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. & Katz, L. (1991) *Science* **252,** 675–679.
24. Hoffman, K., Schneider-Scherzer, E., Kleinkauf, H. & Zocher, R. (1993) *J. Biol. Chem.* **269,** 12710–12714.
25. Symmank, H., Saenger, W. & Bernhard, F. (1999) *J. Biol. Chem.* **274,** 21581–21588.
26. Cole, P. A. (1996) *Curr. Biol.* **4,** 239–242.