Chemistry & Biology, Vol. 11, 927–938, July, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.chembiol.2004.04.012

# The Biosynthetic Gene Cluster for a Monocyclic β-Lactam Antibiotic, Nocardicin A

**Michele Gunsior,[1] Steven D. Breazeale,[2] Amanda J. Lind, Jacques Ravel,[3] James W. Janc,[4] and Craig A. Townsend***
Department of Chemistry
The Johns Hopkins University
3400 North Charles Street
Baltimore, Maryland 21218

## Summary

**The monocyclic β-lactam antibiotic nocardicin A is related structurally and biologically to the bicyclic β-lactams comprised of penicillins/cephalosporins, clavams, and carbapenems. Biosynthetic gene clusters are known for each of the latter, but not for monocyclic β-lactams. A previously cloned gene encoding an enzyme specific to the biosynthetic pathway was used to isolate the nocardicin A cluster from *Nocardia uniformis*. Sequence analysis revealed the presence of 14 open reading frames involved in antibiotic production, resistance, and export. Among these are a two-protein nonribosomal peptide synthetase system, *p*-hydroxyphenylglycine biosynthetic genes, an *S*-adenosylmethionine-dependent 3-amino-3-carboxypropyl transferase (Nat), and a cytochrome P450. Gene disruption mutants of Nat, as well as an activation domain of the NRPS system, led to loss of nocardicin A formation. Several enzymes involved in antibiotic biosynthesis were heterologously overproduced, and biochemical characterization confirmed their proposed activities.**

## Introduction

In the mid-1970s, researchers at Fujisawa Pharmaceuticals isolated a new antibiotic from the fermentation broth of an actinomycete [1], *Nocardia uniformis* subsp. *tsuyamanensis*, that showed moderate activity against a broad spectrum of Gram-negative bacteria. Structure determination of nocardicins A, **1**, and B, **2**, [2] revealed a monocyclic β-lactam moiety, as well as uncommon *p*-hydroxyphenylglycine (pHPG) and oxime units. Unlike the bicyclic β-lactams, nocardicin A exhibits some β-lactamase resistance as a result of its relatively more stable monocyclic β-lactam ring and low affinity by these enzymes [3]. Nocardicin A has also been isolated from the actinomycetes *Actinosynnema mirum* [4], *Nocardiopsis atra* [5], and *Microtetraspora caesia* [6]. The naturally occurring nocardicins A-G differ in amine oxidation state at C-2′ and the presence or absence of an ether-linked homoseryl terminus (Figure 1). The presence of the oxime, and its *syn* relation to the acylamino group, as well as the 3-amino-3-carboxypropyl side chain are critical for antimicrobial activity [7]. Interestingly, in contrast to other antibiotics such as carbenicillin, the activity of nocardicin A is higher in vivo than in vitro [8].

Precursor incorporation studies of **1** have shown its biosynthetic origin in L-serine and two units of the nonproteinogenic amino acid L-pHPG [9, 10], while L-methionine provides the homoseryl side chain [11, 12]. During fermentation L-pHPG is much more efficiently utilized by *N. uniformis* than the D-enantiomer [13]. In addition it was shown that the oxime is formed by amine oxidation [9]. Further biosynthetic studies demonstrated that the amino acid precursors of nocardicin A are assembled at the correct oxidation state [11], and the stereochemical course of β-lactam ring formation in vivo occurs with inversion of configuration at the seryl β-carbon [13, 14].

The most direct mechanistic rationale for the formation of the monocyclic β-lactam consistent with the above results is intramolecular nucleophilic displacement of a presumably activated seryl hydroxyl in an amide precursor. An attractive mechanism of serine activation, although not the only possibility, is in vivo phosphorylation. Phosphorylation can be mimicked in vitro by reaction of a protected serine-containing dipeptide under Mitsunobu conditions where the β-lactam product is formed rapidly at room temperature, thus supporting the proposed $S_Ni$ mechanism in a chemical model system [15].

The intermediacy of a tripeptide or a related derivative was suggested by the efficient and intact incorporation of nocardicin G (D,L,D) into nocardicin A [16]. However, additional feeding studies with a variety of activated and unactivated di- and tripeptides were unsuccessful [17]. This failure notwithstanding, it is hypothesized that in the initial step two units of L-pHPG and one unit of serine are assembled by a nonribosomal peptide synthetase to yield a D,L,D-tripeptide, or possibly nocardicin G, the simplest of the seven known nocardicins.

The in vitro characterization of two tailoring enzymes active later in the pathway provided details of nocardicin A biosynthesis. The protein Nat, responsible for addition of the homoseryl sidechain, catalyzes the biosynthetically rare transfer of a 3-amino-3-carboxypropyl group from *S*-adenosyl-L-methionine (AdoMet) to the substrates nocardicin E, F, and G, but significantly preferring the latter [18]. NocL (vide infra), a cytochrome P450, was demonstrated to carry out the formation of the oxime of nocardicin A in the presence of spinach ferrodoxin, spinach ferrodoxin-$NADP^+$ reductase, NADPH, and nocardicin C, **3**, [19]. While nocardicin C was a substrate for recombinant NocL, reaction of nocardicin G to nocardicin E or F could not be detected under the assay conditions used. The complementary outcomes of these two studies allow the preferred order of the biosynthetic steps to be assigned as shown in Figure 2

*Correspondence: ctownsend@jhu.edu
[1]Present address: National Cancer Institute, Building 37, 9000 Rockville Pike, Bethesda, Maryland 20892.
[2]Present address: Department of Biochemistry, Box 3711, Duke University Medical Center, Durham, North Carolina 17710.
[3]Present address: The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850.
[4]Present address: Celera Genomics, 180 Kimball Way, South San Francisco, California 94080.

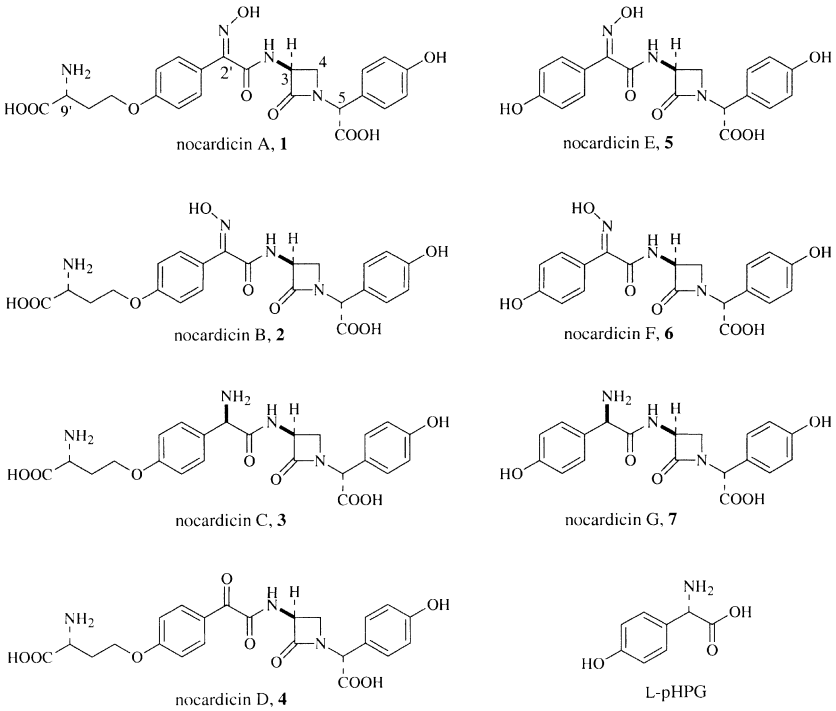



Figure 1. The Nocardicin Family of Antibiotics

[19]. The point at which C-9′ epimerization takes place is unknown.

We report in this paper the cloning and sequence analysis of the nocardicin A biosynthetic gene cluster from *N. uniformis* as well as the initial characterization of several enzymes involved in its biosynthesis. Sequence analysis revealed the presence of a pair of nonribosomal peptide synthetases and pHPG biosynthetic genes upstream of *nat*. Biochemical experiments supporting the proposed functions of these encoded biosynthetic enzymes are also described.

## Results

### Identification of the Nocardicin A Cluster

A 6 kb PstI fragment of *N. uniformis* genomic DNA was obtained as described earlier through "reverse genetics" and determined to possess the entire *nat* gene [18]. Primer walking also revealed the presence of a partial gene showing unmistakable homology to nonribosomal peptide synthetases directly upstream of and overlapping the start codon of *nat*. An *N. uniformis* genomic DNA library was constructed in the *E. coli-Streptomyces* shuttle vector pOJ446 [20] and screened using *nat* and later *nocF*, which lies 17 kb upstream. Sixteen positive cosmids were identified from which a total of 54,100 sequenced bases were obtained. Included were approximately 33.3 kb encoding 14 genes predicted to be responsible for the assembly, self-resistance, and regulation of nocardicin A (Figure 3). No genes with apparent roles in the biosynthesis of nocardicin A were identified upstream of *nocN* (17 kb) or downstream of *nocE* (5 kb). A similar approach was used to isolate the nocardicin A biosynthetic cluster in *A. mirum*. The data indicated

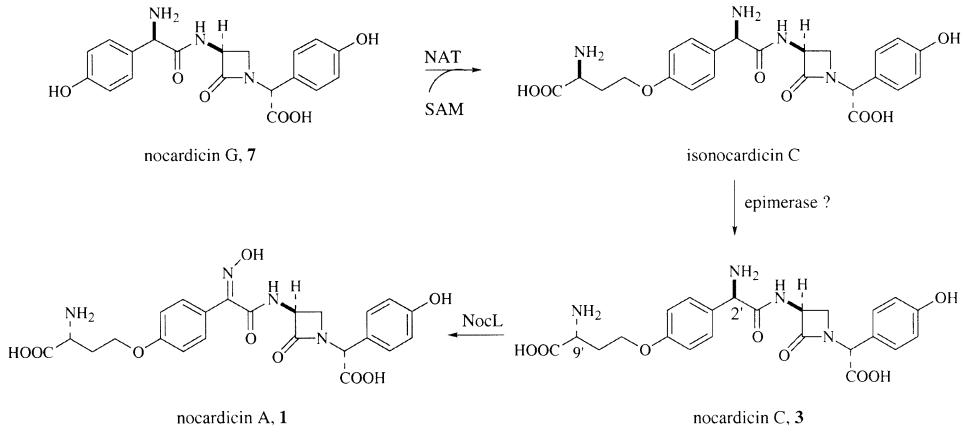



Figure 2. Preferred Biosynthetic Route to Nocardicin A

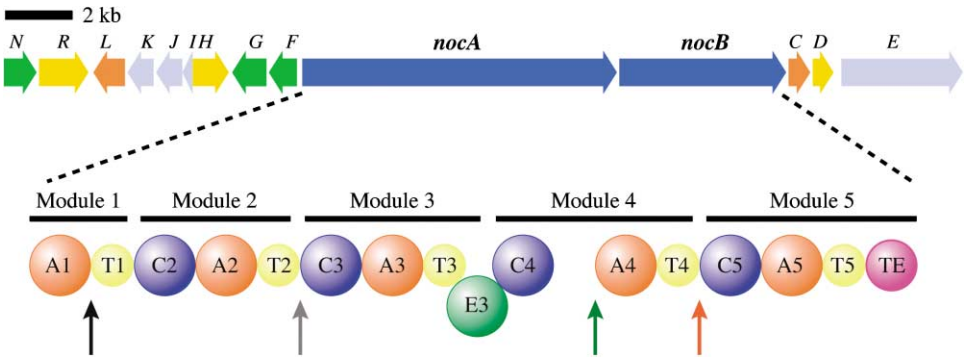



Figure 3. Schematic Diagram of the Nocardicin A Cluster

Genes colored in green denote pHPG biosynthesis; yellow, regulation and resistance; orange, tailoring enzymes; blue, nonribosomal peptide synthetases; and light blue, genes of unknown function. The domain architecture of NocA and NocB is shown enlarged beneath the cluster map. The black and gray arrows indicate the location of the repeat sequences in NocA. The green and red arrows indicate N termini of the 200 kDa and 150 kDa proteins isolated from wild-type *N. uniformis*, respectively.

that the organization of the nocardicin A biosynthetic cluster in *A. mirum* was identical to that of *N. uniformis*, and that sequence identity was nearly 100% (data not shown). Importantly, the sequence upstream of *nocN* and downstream of *nocE* was shown to be different in *A. mirum*, suggesting that both *nocN* and *nocE* mark the pathway boundaries. The genetic map of the nocardicin A biosynthetic cluster is shown in Figure 3.

The nocardicin A cluster sequence possessed an overall G+C content of 76.0%. In order to ascertain the putative function of each of the determined open reading frames (ORF), the deduced amino acid sequences were compared to those in the NCBI protein database using BLAST [21]. The proposed function for each ORF is compiled in Table 1.

## pHPG Biosynthetic Genes

The genes responsible for the catabolism of L-tyrosine to the nonproteinogenic amino acid *p*-hydroxyphenylglycine were found within the nocardicin A biosynthetic cluster. Genes *nocF*, *nocG*, and *nocN* encode a putative *p*-hydroxymandelate synthase, *p*-hydroxyphenylglycine transaminase, and *p*-hydroxymandelate oxidase, respectively. This set of three genes was homologous to that found in the chloroeremomycin biosynthetic pathway in *Amycolatopsis orientalis* [22], the calcium-dependent antibiotic (CDA) cluster in *Streptomyces coelicolor* A3(2) [23], and the complestatin biosynthetic cluster in *Streptomyces lavendulae* [24]. The nocardicin biosynthetic cluster does not possess a prephenate dehydrogenase gene proposed to be involved in pHPG biosynthesis, which is found in all other known clusters containing these genes.

The activities of NocF and NocN were characterized by expression in heterologous systems. Using chiral HPLC, NocF was determined to stereospecifically convert *p*-hydroxyphenylpyruvate to (*S*)-*p*-hydroxymandelate, and NocN was shown to oxidize (*S*)-pHMA to

Table 1. Homology and Deduced Function of ORFs in the Nocardicin A Biosynthetic Pathway

| Gene | Product Size (aa) | Protein Sequence Homolog[a] | Proposed Function |
|---|---|---|---|
| *nocE* | 1414 | no significant homology | unknown |
| *nocD* | 185 | probable acetyltransferase (AE001997) | resistance |
| *nat* | 301 | methyl transferase (AE005808) | 3-amino-3-carboxypropyl transferase [18] |
| *nocB* | 1925 | nonribosomal peptide synthetase (syringomycin, T14593) | module 4 C A T<br>module 5 C A T TE |
| *nocA* | 3692 | nonribosomal peptide synthetase (pyoverdin, F83345) | module 1 A T<br>module 2 C A T<br>module 3 C A T E |
| *nocF* | 345 | *p*-hydroxymandelate synthase, Hmas (CAA11761) | pHPG biosynthesis, *p*-hydroxymandelate synthase |
| *nocG* | 431 | *p*-hydroxyphenylglycine transaminase, HpgT (CAA11790) | pHPG biosynthesis, *p*-hydroxyphenylglycine transaminase |
| *nocH* | 408 | membrane transport protein (X86780) | transport |
| *nocI* | 74 | conserved hypothetical protein of unknown function (AE004668) | unknown |
| *nocJ* | 327 | putative 1-aminocyclopropane-1-carboxylate deaminase (P756316) | unknown |
| *nocK* | 344 | putative hydrolase/esterase (AL583919) | unknown |
| *nocL* | 398 | cytochrome P450 (S51594) | oxime-forming enzyme [19] |
| *nocR* | 582 | transcriptional activator protein (AF124138) | regulation |
| *nocN* | 376 | *p*-hydroxymandelate oxidase, Hmo (CAA11762) | pHPG biosynthesis, *p*-hydroxymandelate oxidase |

[a] GenBank protein accession numbers of selected homologs are given in parentheses.

*p*-hydroxybenzoylformate. We were unable to heterologously express soluble NocG, which prevented verification of this enzyme as a transaminase as demonstrated by Hubbard et al. [25] for the corresponding gene from *A. orientalis*.

### Regulatory/Resistance Genes

NocD shows approximately 30% identity and up to 60% similarity to the superfamily of acetyltransferases that includes *N*-, *O*-, and *S*-acetyltransferases of all types and function. Of acetyltransferases with known function, NocD shows apparent homology to *N*-acetylases, including those of ribosomal proteins from prokaryotic organisms such as *Thermoanaerobacter tengcongensis* and *Bacillus halodurans*. Acetyltransfer is the predominant mechanism of antibiotic resistance employed by pathogenic bacteria against aminoglycosides [26], and has also been implicated in bleomycin self-resistance [27].

The predicted product of *nocH* has 30% identity and 44% similarity to a membrane transport protein found in the rapamycin biosynthetic cluster from *Streptomyces hygroscopicus*. Other homologs include multidrug efflux permease and resistance proteins from a variety of organisms. It is likely that NocH functions by exporting the antibiotic outside the cell and might be a self-resistance mechanism.

The deduced product of *nocR* is a large protein of 582 amino acids with a predicted molecular weight of 61.8 kDa. A conserved domain (CD) search reveals a small region of fewer than 100 residues at the N terminus homologous to DNA binding domains of various regulators [28], in addition to a region corresponding to a bacterial transcriptional activator domain (BAD). NocR has homology to transcription regulatory proteins, most notably CdaR, a transcriptional activator found in the biosynthetic cluster of CDA from *S. coelicolor* (35% identity, 46% similarity), and AsfR, a global regulatory protein for secondary metabolite biosynthesis from the same organism [29]. Based on this analysis, NocR appears to be related to the growing family of pathway-specific activators known as SARPs (*Streptomyces* antibiotic regulatory proteins [30]).

### Genes of Unknown Function

Gene *nocI*, encodes a 73 amino acid protein that shows high homology to a *Pseudomonas aeruginosa* hypothetical protein (60% identity, 76% similarity). All of the sequences producing significant homology are small proteins ranging in size from 56–80 amino acids. The function of this small class of proteins is unknown, but they are typically present in NRPS and mixed NRPS/PKS systems.

The deduced amino acid sequence of NocJ has low homology (23% identity, 33% similarity) to 1-aminocyclopropane-1-carboxylate (ACC) deaminases, a PLP-dependent protein involved in the metabolism of ACC to $\alpha$-ketobutyrate and ammonia [31, 32]. An alignment of NocJ with putative and known ACC deaminases indicates that NocJ possesses a critical lysine that binds PLP within the conserved sequence VFGGCKTRALEF [33], which has been confirmed by experiment (W.L. Kelly and C.A.T., unpublished data).

NocK has high homology to a putative secreted protein from *S. coelicolor* A3(2), and LpqC (34% and 29% identity respectively), an *M. tuberculosis* hypothetical protein thought to act as an esterase or a poly(3-hydroxyalkanoate) depolymerase [34]. NocK also shows lower homology (22% identity) to poly(3-hydroxybutyrate) depolymerases, a type of serine esterase that hydrolyzes polyhydroxyalkanoates to water soluble products [35].

The translated product of *nocE* is a large protein of 1414 residues. No homology or conserved domains were discovered when analyzed by BLAST or any of several motif search engines.

### Tailoring Enzymes

Studies in this laboratory have previously shown that Nat (NocC) transfers the 3-amino-3-carboxypropyl moeity of AdoMet to nocardicin G, forming isonocardicin C (the C-9′ epimer of **3**) in an unusual use of the cofactor [18]. Nat possesses weak homology to bacterial AdoMet-utilizing enzymes.

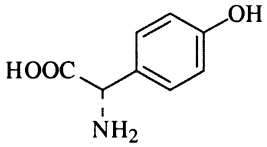
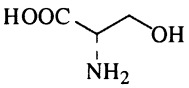
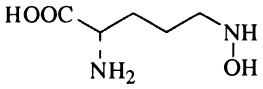
*nocL* encodes a cytochrome P450 that possesses both the conserved heme binding region FGHGxHxCLG [19, 36] centered around the invariant cysteine residue, and a region homologous to an oxygen binding motif LLxAGHET [37, 38]. The nocardicin A cluster does not possess an NAD(P)H-dependent reductase or ferrodoxin gene, two additional enzymes involved in electron transport and required in bacterial P450 catalysis. It has been generally observed that all three genes are rarely found in biosynthetic clusters, and it is assumed that electron transport proteins are recruited from elsewhere in cellular metabolism [39]. NocL was demonstrated to carry out formation of the oxime moiety of nocardicin A in the presence of spinach ferrodoxin, spinach ferrodoxin-NADP$^+$ reductase, NADPH, and nocardicin C, **3**, [19].

Another enzyme not found in the nocardicin A biosynthetic cluster is the phosphopantetheinyl transferase (PPTase), required in the conversion of apo-NocA/B to holo-NocA/B. Holo-NRPS enzymes possesses 4′-phosphopantetheine arms, covalently bound to a conserved serine residue in the thiolation domain responsible for shuttling enzyme bound intermediates between modules. Genes encoding PPTase enzymes are highly homologous and often found adjacent to biosynthetic clusters containing peptide synthetases. However, when an organism contains two or more NRPS systems, one PPTase seems sufficient to modify all systems [40]. This seems to be the case for nocardicin A. Accordingly, nearly identical genes encoding Sfp-subfamily PPTases were found elsewhere in the genomes of *N. uniformis*, and *A. mirum* (J.R. and C.A.T., unpublished data). This family of enzymes has the ability to modify both type I and type II acyl carrier proteins and peptidyl carrier proteins.

### NRPS Biosynthetic Genes

Sequence analysis showed that NocA and NocB belong to the growing family of nonribosomal peptide synthetases. The coding sequence of the peptide synthetases spans 17 kb corresponding to 3692 and 1925 amino acids, respectively. The protein sequence was exam-

Table 2. Prediction of Amino Acid Substrate for NocA and NocB Adenylation Domains

| A Domain | Homologous Sequence | Substrate Prediction |
|---|---|---|
| CepB-A1 | D-I-F-H-L-G-L-L |  L-PHPG |
| Cda1-A6 | D-V-Y-H-L-G-L-L | |
| BpsB-A4 | D-I-F-H-L-G-L-L | |
| **NocA-A1** | **D-I-W-H-V-G-A-I** | |
| **NocA-A3** | **D-I-W-H-V-G-T-I** | |
| **NocB-A5** | **D-I-W-H-L-G-A-I** | |
| SyrE-A3 | D-V-W-H-L-S-L-I |  L-Ser |
| NosA-A2 | D-V-W-H-I-S-L-I | |
| PvdD-A1 | D-V-W-H-I-S-L-I | |
| **NocB-A4** | **D-V-W-H-L-S-L-I** | |
| ApdB-A2 | D-V-E-N-A-G-V-V |  L-5NhOrn |
| FxbC-A1 | D-M-E-N-L-G-L-I | |
| **NocA-A2** | **D-V-E-N-V-G-A-I** | |

Abbreviations are as follows: CepB, chloroeremomycin synthetase; Cda1, calcium-dependent antibiotic peptide synthetase 1; BspB, balhimycin synthetase; SyrE, syringomycin synthetase; NosA, nostopeptilide synthetase, PvdD, pyoverdine synthetase; ApdB: anabaenopeptilide synthetase B; FxbC: exochelin synthetase; L-5NhOrn: L-$N^5$-hydroxyornithine.

ined for the presence of standard NRPS motifs [41]. The result of this analysis indicated a departure from the canonical domain organization, in which the enzyme(s) responsible for producing a D,L,D-tripeptide would be expected to be composed of three modules, the first and third of which would also contain epimerization domains. In contrast, NocA and NocB together possess five modules and only one epimerization domain, as shown in Figure 3. Furthermore, NocB begins with an adenylation domain rather than a conventionally observed condensation domain, which is located at the C terminus of NocA. Thus, the two enzymes involved in nocardicin A biosynthesis can each be classified as a Type C NRPS [42].

It is now possible to predict A-domain substrate specificity by comparison of eight critical residues lining the amino acid binding pocket to other known A-domains [43, 44]. Table 2 shows the eight residue motifs for each of the NocA and NocB A-domains with their predicted substrate.

Alignments of the substrate recognition residues of putative serine-activating A-domains show a universally conserved H278/S301 dyad believed to play a functional role in hydrogen bonding of the substrate hydroxyl [43]. NocB-A4 shows high homology to domains activating serine and possesses both H278/S301. pHPG-activating domains are remarkably similar to serine domains, but contain a S301→G301 substitution [43]. The specificity-conferring residues of NocA-A1, NocA-A3, and NocB-A5 are similar to other pHPG domains, and all three possess a glycine residue at position 301. Recently an A-domain from the balhimycin antibiotic biosynthetic cluster (BpsB-A4) in *Amycolatopsis mediterranei* [45] was overexpressed and shown to activate L-pHPG. According to our analysis, the BpsB-A4 substrate specificity motif closely matches those of all three predicted nocardicin A pHPG A-domains, and other pHPG A-domains (Table 2).

The specificity-conferring residues of NocA-A2 show highest homology to the substrate binding pocket of anabaenopeptilide synthetase B-A2 (ApdB-A2) [46], L-$N^5$-hydroxyornithine (5NhOrn) activating FxbC-A1 and A3 of exochelin biosynthesis [47], and CchH-A3 of coelichelin synthetase [48]. Known ornithine-activating domains have a universally conserved E278/S322 dyad believed to form hydrogen bonds with the terminal amine/ammonium. In 5NhOrn A-domain substrate binding pockets, E278 is replaced by asparagine, and S322 by leucine or glycine. NocA-A2 and ApdB-A2 do not possess the E278/S322 dyad conserved in the binding pocket of ornithine-activating A domains. The differences between NocA2-A2 and ApdB-A2 are conservative changes, and given the close homology to 5NhOrn-activating domains, it is possible that both activate this substrate or possibly ornithine.

Both *nocB*-A4 and *nat* were disrupted with a thiostrepton gene cassette resulting in the complete loss of nocardicin A production. Disappointingly, no detectable accumulation of nocardicin E, **5**, or nocardicin G, **7**, was observed in the *nat* disruption mutants when grown under production conditions (Figure 4). Complementation of the disrupted *nat* gene, however, with a full-length transcript expressed under the ermE* constitutive promoter on shuttle vector pULVK2A [49] restored production of nocardicin A to at least a low level, confirming the critical role of Nat in its biosynthesis. The HPLC peak arising from complementation was shown to be nocardicin A by ESI-MS and MS/MS comparisons to authentic material. We can conclude, therefore, that the gene cluster isolated and characterized is, indeed, responsible for nocardicin A biosynthesis.

### Overexpression of NocA/B Adenylation Domains

Many attempts were made to overexpress the adenylation domains of NocA/B in *E. coli* in an effort to biochemically validate the substrate predictions and correlate to earlier whole-cell incorporation experiments. A-domains of NocA/B A1-A5 were cloned into a variety of expression vector systems. Despite repeated efforts varying the expression system and host, the affinity tags, the size of the constructs, the temperature for expression, and the addition of the *E. coli* chaperones GroES/EL,[50] no soluble, active proteins were obtained. Insolubility and low expression of adenylation domains is not limited

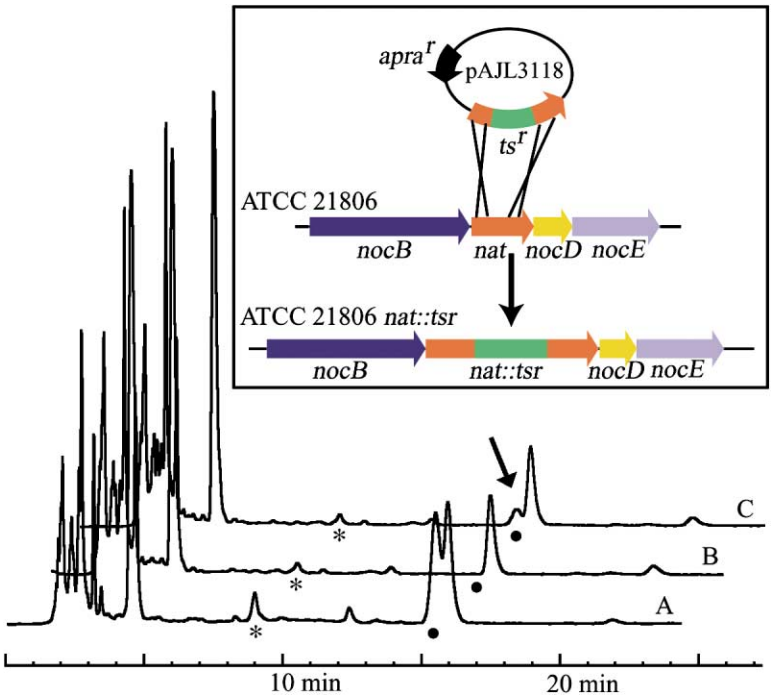



Figure 4. Disruption and Complementation of the *nat* Gene and HPLC Analysis

Chromatograms of supernatants from (A) wild-type *N. uniformis*, (B) disruption mutant *nat::tsr*, and (C) complementation mutant pAJL3248 *nat::tsr* are shown. Nocardicin G (*) and nocardicin A (•) peaks were confirmed by coinjection of authentic material. Inset: Disruption of *nat* in the plasmid pAJL3118 with the thiostrepton resistance cassette.

to NocA/B; these problems have been encountered in NRPSs involved in enterobactin, bleomycin, and mycobactin biosynthesis, among others [51, 52].

### NRPS Characterization

Two proteins, 200 kDa and 150 kDa in lesser, variable amounts as estimated by SDS-PAGE, were purified to near homogeneity from *N. uniformis* based on their ability to catalyze ATP/PPi exchange in the presence of L-pHPG. PPi exchange in the presence of L-pHPG was 30 times higher than for D-pHPG, and ten times higher than L-phenylalanine indicating that a Phe-tRNA synthetase had not been isolated. Exchange in the presence of L-serine was less efficient, but nonetheless was reproducibly 3-4 times higher than that observed for the negative control. Importantly, the 150 kDa enzyme failed to give exchange detectable above background when assayed with L-serine. Failure to stimulate exchange was also observed with L-tyrosine, L-tryptophan, L- and D-phosphoserine in both the 200 and 150 kDa enzymes.

The 200 kDa enzyme was inactivated in the presence of sulfhydryl modifying reagents such as *p*-chloromercuribenzoate, *N*-ethylmaleimide, or iodoacetamide, suggesting that a reactive sulfhydryl group is central to the chemistry performed by this enzyme. Additionally, the presence of 4′-phosphopantetheine in this enzyme was shown by release of D-pantothenate through the action of alkaline phosphatase and biochemical assay [53].

Further experiments assayed the ability of the 200 kDa protein to promote peptide formation and epimerization. Incubation with radiolabeled substrates L-[2-$^{3}$H,1-$^{14}$C]-pHPG and D-[2-$^{3}$H,1-$^{14}$C]-pHPG, was monitored by HPLC and scintillation counting. The absence of a decrease in ($^{3}$H/$^{14}$C) ratio in each of the substrates indicated that the protein did not catalyze in vitro epimerization. In addition, neither the (D,L,D)-tripeptide, nor nocardicin G was detected when the 200 kDa protein was incubated with the required precursors, in keeping with the colinearity rule.

The N-terminal sequences of the 200 kDa and 150 kDa proteins were obtained following transfer to a PVDF membrane. Sequencing of the former yielded approximately 20 N-terminal amino acids that did not show homology to any known nonribosomal peptide synthetase in the database, but corresponded exactly to the first 20 residues of NocB. The N-terminal sequence of the 150 kDa protein yielded 11 amino acids that corresponded to the beginning of module 5 in NocB, as indicated by the red arrow in Figure 3.

### Discussion

It was postulated from precursor incorporation experiments and the presence of D-pHPG residues that nocardicin A would be biosynthesized by way of NRPS enzymes. Sequencing of the nocardicin A biosynthetic gene cluster confirmed this hypothesis and also shed light on other aspects of its formation, including some of the unusual structural features of the molecule such as the oxime, the homoseryl side chain, and the two pHPG units. Disruptions of nocB-A4 and *nat* resulted in complete loss of nocardicin A production indicating that the isolated gene cluster was responsible for formation of the monocyclic β-lactam. Nocardicin A production was restored by complementation with a full-length copy of the late-stage biosynthetic gene *nat*. Comparison of the virtually identical clusters found in *N. uniformis* and *A. mirum* suggested that *nocN* and *nocE* delineate the pathway boundaries.

The genes involved in pHPG biosynthesis were found clustered in the nocardicin A biosynthetic pathway, similar to other systems [22, 23], with the exception of pre-

phenate dehydrogenase (PDH), an enzyme normally involved in aromatic amino acid biosynthesis. PDH, which catalyzes the formation of *p*-hydroxyphenylpyruvate (pHPP) from prephenate and found in other pHPG pathways, is absent from the nocardicin A cluster, suggesting that pHPP is obtained from primary metabolism and not a separate, dedicated enzyme.

Two large enzymes of 200 kDa and 150 kDa with the ability to catalyze L-pHPG-dependent ATP/PPi exchange were purified from *N. uniformis*. Experiments with the larger of the two confirmed the presence of a reactive sulfhydryl group in the enzyme, as well as independent demonstration of a phosphopantetheine prosthetic group. N-terminal sequence of the 200 kDa protein mapped to the first 20 amino acids of NocB (green arrow, Figure 3). The observed molecular weight corresponded well with the calculated mass of 206 kDa for NocB. NocB possesses only two modules, and, therefore, cannot catalyze tripeptide synthesis alone.

N-terminal sequence data of the 150 kDa protein showed identity to a region of NocB four amino acids upstream of the condensation domain motif C1 of module 5 (shown by the red arrow in Figure 3), indicating that the smaller enzyme is a proteolytic fragment of NocB of 140.5 kDa. The site of proteolysis accounts for experimentally determined substrate specificity of module 5 and allows the function of module 4 to be deduced. The 150 kDa enzyme catalyzed PPi exchange with L-pHPG, but not the D-isomer or several other amino acids including L-serine. These experiments prove that module 5 contains a pHPG A-domain, and, therefore, module 4 likely activates L-serine in keeping with predictions [43]. Attempts to determine the substrate specificity of the other adenylation domains of NocA through expression of autonomous enzymes were unsuccessful. However, the binding pockets of three A-domains in NocA and NocB proposed to activate pHPG are highly homologous; therefore, it is likely that NocA-A1 and NocA-A3, like NocB-A5, recognize and adenylate pHPG.

The mechanism by which the core peptide structure is formed remains unclear. NocA and NocB do not follow the NRPS colinearity paradigm, and the presence of five modules instead of the three expected for tripeptide formation complicates the interpretation. Furthermore, the earliest known intermediate nocardicin G contains two D-pHPG units, while NocA and NocB contain only one epimerization domain. In the absence of biochemical data, the formation of a product of the correct stereochemistry can only be speculated.

Based on the configurations of the amino acids comprising nocardicin G (D,L,D), and experiments that determined the stereochemistry of the substrate amino acids to be the L-enantiomers, a peptide synthetase system comprising three modules, the first and third of which contain epimerization domains, was anticipated. Sequence analysis indicates modules 3, 4, and 5 might be responsible for formation of a tripeptide intermediate as these adenylation domains activate pHPG, serine, and pHPG, respectively. Module 3 (M3) has an epimerization domain (Figure 3) while M5 does not. Several possibilities could account for the absence of an M5 epimerization domain: (1) The position is epimerized following completion of the peptide product by a separate enzyme. Such a peptide isomerase has been purified from funnel web spider venom and shown to interconvert the configuration of a single serine residue in a 48-amino-acid peptide [54]. (2) An epimerase similar to that found in the biosynthesis of cyclosporin A [55] inverts the stereochemistry of L-pHPG prior to activation. This possibility is unlikely as biochemical experiments with NocB indicate that in the presence of L-pHPG ATP/$PP_i$ exchange occurs 30 times more efficiently than with D-pHPG, and is also inconsistent with our experiments using [2-$^{3}$H,1-$^{14}$C]-pHPG. (3) The C-5 position is benzylic and known to be base-labile [56]. The pKa of this hydrogen is similar to that of an amide and epimerization at this center could occur in tandem with deprotonation of the adjacent amide and $\beta$-lactamization.

A possibility to account for tripeptide formation in this five module NRPS system is that M1 and M2 are inactive. Related examples of inactive domains have been documented in the condensation domain of VibF in vibriobactin biosynthesis, [57] and the adenylation domain of module NRPS-0 in bleomycin formation [52, 58]. Primary sequence alignments of NocA indicate no anomalies in the core catalytic domains but does reveal two regions of additional sequence containing short amino acid repeats that are not present in other NRPSs. The first, approximately 40 residues in length, is located just before the thiolation domain of M1 (shown by the black arrow in Figure 3) and contains a disproportionate number of glycine and proline residues. The second region is larger, spanning 90 residues following the thiolation domain of M2 (gray arrow, Figure 3), and also contains a high number of glycine, alanine, valine, and proline residues. At the DNA level, the third position GC bias in this region drops to 60% from nearly 100%. The role of these repeats is unknown, but it might affect the tertiary structure of these modules, hence inactivating them. If this were the case, a dipeptidyl product could be formed, or the substrate amino acids of both modules could be activated individually, but spatial deformities would prevent condensations from taking place by interfering with transfer(s) to M3. The proposed substrates of the remaining modules 3 through 5, when linked by amide bonds, lead to the core nocardicin tripeptide, pHPG-Ser-pHPG. The role of these short repeats probably differs from that of the unique and quite faithfully duplicated hexapeptide motifs at the C terminus of BlmIV. Recent studies of bleomycin biosynthesis have unequivocally demonstrated the essential function of these C-terminal repeats exert in protein-protein interactions with BlmIII to enable bithiazole formation to occur [58].

In a second possible mechanism for the formation of the core tripeptide, if all five modules were active, the resulting pentapeptide could be processed down to a tripeptide. Many bacterial systems possess proteolytic enzymes whose substrates are small peptides composed of naturally occurring L-amino acids, while peptides containing D-epimers are more stable [59]. The first two modules likely incorporate L-residues, and consequently the first two residues of the pentapeptide could be enzymatically cleaved to give a tripeptide, which is then resistant to further cleavage by virtue of a D-pHPG configuration at the N terminus.

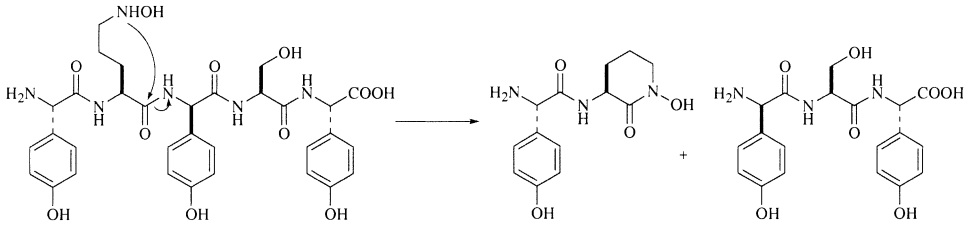


Figure 5. Proposed Formation of Hydroxamate and Tripeptide

Finally, it is predicted that M2 activates and incorporates $N^5$-hydroxyornithine or ornithine. A 6-membered ring might be formed in a favored 6-*exo* trig cyclization [60] through intramolecular attack of the $N^5$-hydroxyornithine or ornithine side chain, on the amide carbonyl, resulting in a hydroxamic acid/δ-lactam and a tripeptide (Figure 5). A similar mechanism has been proposed in the biosynthesis of coelichelin [48] to effect cleavage of the putative tripeptide from the thiolation domain of CchH. Such a self-editing event could occur in nocardicin G biosynthesis while covalently attached to NocA so that only D-pHPG, the substrate of M3, is translocated and elongated; or intrapeptide cleavage might occur after the pentapeptide is free in solution. If the two repeat sequences do affect tertiary structure and, hence, the native processivity, the likelihood of self-cleavage while linked to NocA is much greater. A postulated mechanism for hydroxamate formation is shown in Figure 5. Early release of a growing peptide in NRPS synthesis has been shown to occur; incorporation of ornithine in an engineered surfactin synthetase I NRPS altered the conformation of the surfactin product and resulted in premature cyclization, although the products obtained in this experiment were not those expected based on mass spectrometric characterization. The lower m/z ions observed suggested more extensive modifications had taken place, but the exact structures are not known [61].

The mechanism and the enzyme(s) responsible for the formation of the monocyclic β-lactam ring are still unresolved in nocardicin A biosynthesis. Enzymes catalyzing the formation of other β-lactam structures have been found and characterized. Isopenicillin N synthase (IPNS) catalyzes β-lactam closure in a tripeptide by an oxidative path [62]. β-Lactam synthetase (β-LS) [63] and carbapenam synthetase (CPS) [64], active in the biosynthesis of clavams and carbapenams/ems, respectively, require ATP to close a β-amino acid precursor. Ring closure in nocardicin A is thought to involve a mechanism distinct from either IPNS, or β-LS/CPS. Incorporation and stereochemical studies suggest a simple $S_Ni$ mechanism, supported by biomimetic model reactions [15], in which the amide nitrogen of a hypothetical peptide precursor would displace a presumably activated seryl hydroxyl. Ring closure to the β-lactam could be catalyzed by the NRPS itself, or by a separate set of proteins to activate the serine β-hydroxy and cyclize to the 4-membered ring.

## Significance

**The β-lactam antibiotics and β-lactamase inhibitors remain of central importance to the battle against infectious diseases. Rapid advances are being made currently to both elucidate and engineer natural product biosynthetic pathways at the genetic level for the production of new antibiotics. In this paper, we have described the isolation and initial characterization of the nocardicin A biosynthetic gene cluster from *N. uniformis* and *A. mirum*, the first example of a monocyclic member of the β-lactam family. This result marks the completion of at least one representative gene cluster sequence from each of the four known subclasses of these antibiotics. The *noc* cluster contains 14 genes for formation, resistance, and export of nocardicin A. Among these are the genes necessary for synthesis of the nonproteinogenic amino acid pHPG, a P450 involved in oxime formation, and an AdoMet-dependent transferase that attaches an aminocarboxypropyl group to the phenolic oxygen of pHPG. As is frequently observed in natural product biosynthetic pathways, some of the encoded biosynthetic proteins possess primary sequence similarity to known enzymes, but have evolved to perform new roles in the creation of natural products. There are also gene products whose function is wholly unknown. A pair of NRPS enzymes thought to be responsible for peptide formation does not follow the colinearity paradigm and may catalyze unknown transformations, such as epimerization in the absence of a specific domain or possibly β-lactam formation. Formation of the 4-membered ring likely involves an $S_Ni$ mechanism in a serine-containing peptide that is distinct from the oxidative formation of penicillin and cephalosporin, or the ATP-driven β-lactam closure in clavam and carbapenem biosynthesis. The bioengineering of combinatorial NRPS modules has exciting long-range prospects in the production of new antibiotics, especially if modules with novel activities can be exploited.**

### Experimental Procedures

#### Bacteria and Plasmids

Wild-type *N. uniformis* subspecies *tsuyamanesis* (ATCC 21806) and *A. mirum* (ATCC 29888) were purchased from the American Type Culture Collection (Manassas, VA). *Escherichia coli* DH5α was used for routine cloning and plasmid DNA preparation. *E. coli* BL21 (DE3) expression cells and pET28b+ vector were purchased from Novagen (Madison, WI). *E. coli* strain TB1 and pMAL-c2 expression vector were obtained from New England BioLabs (Beverly, MA). $Ni^{2+}$-NTA resin and the pQE60 vector were obtained from Qiagen (Valencia, CA). Cosmid libraries were amplified in *E. coli* XL1-Blue MRF′ (Stratagene, La Jolla, CA). Cosmid pOJ446 was previously described [20], and phagemid pBSII SK– (Stratagene), was purchased from the manufacturer. TOPO-TA cloning vector was obtained from Invitrogen (Carlsbad, CA). Nocardicin A was a gift from Fujisawa Pharmaceuticals Co., Ltd. Osaka, Japan.

### General Procedures

The QIAprep-Spin plasmid Kit (Qiagen) was used to prepare plasmid DNA. Restriction and modifying enzymes were purchased from New England Biolabs and used according to the manufacturer's recommendations. Platinum HiFi Taq (Invitrogen) was used for all PCR amplifications. All molecular biology techniques were performed according to standard methods [65]. Radioactivity was measured using a Beckman LS 5801 scintillation counter. HPLC analyses were performed on a Perkin-Elmer Series 410 LC Pump equipped with a Perkin-Elmer 235C diode array detector and an Agilent Technologies series 1100 fitted with a diode array detector.

### DNA Sequencing and Analysis

DNA sequencing was performed at the Biosynthesis & Sequencing Facility, Johns Hopkins Medical School, Baltimore, MD. Sequence data were edited, assembled, and analyzed with the Sequencher Software (Gene Codes Co., Ann Arbor, MI) and GCG software package (Version 9, Genetics Computer Group, Madison, WI). Start codons were assigned by choosing the ATG or GTG furthest upstream in the ORFs, in conjunction with a high G+C bias in the third position of the codons and probable ribosomal binding sites. Database searches were performed online using BLAST at the National Center for Biotechnology Information [21].

### Growth of *N. uniformis* and Isolation of Genomic DNA

Fermentation and DNA isolation were carried out as previously described [18].

### Preparation of Cosmid Libraries

*N. uniformis* and *A. mirum* genomic DNA (gDNA) were partially digested with *Sau*3AI to give fragments of 30–40 kb in size. These fragments were dephosphorylated with shrimp alkaline phosphatase (SAP) and ligated into the HpaI- and BamHI-digested cosmid vector pOJ446 [20]. The ligation products were packaged with Gigapack III Gold Packaging Extract (Stratagene) as described by the manufacturer, and the resulting phage was used to infect *E. coli* XL1-Blue MRF′ cells. The *N. uniformis* library, consisting of 2500 apramycin resistant colonies, was screened by hybridization with the 3-amino-3-carboxytransferase (*nat*) gene, [18], and later with *nocF*. Six positive colonies were identified, and the presence of *nat* in each was confirmed by restriction mapping, sequencing, and hybridization. Screening with the *nocF* primer resulted in identification of 10 additional cosmid clones.

### *nat* Disruption Construct

*nat* was PCR amplified from plasmid pSDB05-152 [18] using the primers NatF–5′ GGAATTCCATATGACCGCGCTGTCACGCGTG–3′ and NatR–5′ AACAACGCTCTTCCGCACCCCTTGCGCAGCACGA TGC–3′ where underlined bases represent NdeI or SapI restriction sites. A thiostrepton resistance cassette was inserted into the cloned *nat* using a unique NcoI site. The resulting plasmid, pJRI-44, was digested with BamHI and HindIII and ligated into the corresponding sites of pULVK2A to create pAJL3118. Disruption constructs were confirmed by restriction digests and sequencing, and transformed into JM110.

### A4 Disruption Construct

The A4 adenylation domain was PCR amplified from *N. uniformis* gDNA using the primers BA4F, 5′–GAAGAACCATGGCCTTTGCAGT GCGCGACCTG–3′ and BA4R 5′–GAAGAACTCGAGGGTCTCGGGC GCGGCGG–3′, in which the underlined bases indicate NcoI and XhoI restriction sites, respectively, and the bolded base indicates a G to A substitution for cloning purposes. The purified amplicon was ligated into corresponding sites in pBSKII–. The thiostrepton resistance cassette was inserted at the unique BclI site of the cloned fragment. The disruption construct was then digested with ApaI and ligated into the pULVK2A vector to produce pAJL3126. Disruption constructs were confirmed by restriction digests and sequencing, and transformed into JM110 cells.

### *nat* Complementation Construct

The following primers were used to PCR amplify *nat* from *N. uniformis* gDNA, NatFXba, 5′–TATATCTAGAGTGACCGCGCTGTCAC GCGTG–3′ and NatRSph, 5′–TATAGCATGCTCACCCCTTGCGCA GCAC–3′ in which the underlined bases indicate XbaI and SphI restriction sites, respectively. The 905 bp amplicon was ligated into pIJ4070 downstream of the *ermE** promoter and the resulting *nat* cassette was excised with BglII and inserted into pULVK2A to produce pAJL3248, which was transformed into JM110 cells.

### Transformation of *N. uniformis* by Electroporation

A 10 ml culture of *N. uniformis* in Tryptic Soy Broth (TSB) was incubated overnight at 28°C with vigorous shaking. The overnight culture (1 mL) was used to inoculate 100 ml of #148G medium [66] supplemented with 0.85% glycine and 1.5 mg/L isonicotinic hydrazide, and grown for 24 hr at 28°C. The mycelium was harvested by centrifugation (16,000 × g, 15 min), washed with 10% sucrose and 15% glycerol, and finally resuspended in 600 μl of 30% PEG, 10% glycerol solution. The cells were stored in aliquots at –80°C until use. In each transformation, 2 μl of plasmid DNA was added to 20 μl mycelium. The mixture was transformed by electroporation on a Gibco BRL Cell-Porator (280V, 4Ω, fast charge rate), and allowed to recover in R2L liquid medium at 28°C with shaking (300 rpm) for 3 hr and plated as described [67].

### Disruption Procedure

Antibiotic resistant colonies of *N. uniformis* were grown in 100 ml of S27M medium for 72 hr. An aliquot (1 mL) of the culture was used to inoculate fresh S27M medium (100 mL). The procedure was repeated three times and 1 ml was plated on S27M plates containing thiostrepton. Colonies were picked and plated on S27M apramycin or thiostrepton plates. Colonies exhibiting *apra*$^s$ and *thio*$^r$ were verified by PCR amplification with primers designed outside the disrupted gene of interest. The band, which contained the gene and the thiostrepton resistance cassette, was cloned into pCR4 TOPO (Invitrogen) vector and sequenced.

### HPLC Analysis of Disruption Mutants

Starter cultures (10 mL) of disruption mutants were grown in TSB medium with antibiotic overnight at 28°C and 300 rpm, and 1 ml of this preculture was used to inoculate 50 ml of F2 medium [7]. After growing 3–5 days, the culture supernatants were analyzed by HPLC using a Phenomenex Luna 5μ C18 column (250 × 4.60 mm; Torrance, CA) under isocratic conditions (10% acetonitrile/90% water /0.01% TFA, 1 mL/min).

### Purification of NocB and 150 kDa Protein

*N. uniformis* cells (60 g) were suspended in lysis buffer (150 mM Tris [pH 7.5], 150 mM KCl, 10 mM DTT, 10 mM $MgCl_2$,1 mM EDTA, 1 mM benzamidine, 1 μM leupeptin, 0.1 μM pepstatin A, 5 mg/L soybean trypsin inhibitor, and 45% glycerol). Cells were ruptured at 4°C by sonication (Bransonic, model W-225R) at 50% duty, power level 7, for 5 min and centrifuged (14,600 × g, 20 min, 4°C). Nucleic acids were precipitated with streptomycin sulfate and the protein pellet obtained between 20% and 60% saturation with $(NH_4)_2SO_4$ was dissolved in the minimum amount of buffer A (100 mM Tris [pH 7.5], 5 mM DTT, 2 mM $MgCl_2$, 0.1 mM EDTA, and 15% glycerol) and dialyzed overnight against 2 liters of the same. The dialyzed extract was diluted with buffer A and applied to a Q-Sepharose column, which was then washed with buffer A and subjected to a linear gradient of 0–500 mM KCl. The fractions containing NocB (as determined by ATP/PPi exchange activity for L-pHPG) were pooled and concentrated in an Amicom ultrafiltration cell equipped with a PM30 membrane. The concentrated Q-Sepharose pool was applied to a Sephacryl S300-HR column and eluted with buffer A. Again, the fractions catalyzing ATP/PPi exchange in the presence of L-pHPG were pooled and applied to a Pharmacia MONO Q HR 5/5 FPLC column as a final purification step to separate NocB and the 150 kDa enzyme. The mobile phase consisted of 30 mM Tris (pH 7.4), 1 mM DTT, and 10% glycerol. Elution occurred under a shallow gradient of 300 mM to 475 mM KCl at 0.5 mL/min.

### ATP/PPi Exchange Assay

The reaction mixture (100 μL) consisted of 72.5 mM HEPES (pH 7.5), 7.3% glycerol, 14.5 mM $MgCl_2$, 7.3 mM ATP, 2.9 mM DTT, 0.29 mM EDTA, 1.5 mM substrate, and 20 μCi [$^{32}$P]-PPi. The assay was

initiated with 25 μl enzyme sample and incubated 30 min at 37°C. The reaction was quenched by the addition of 400 μl 0.5 M $HClO_4$ and 400 μl 100 mM unlabeled PPi. A 4% w/v suspension of activated charcoal in $dH_2O$ was added to the sample followed by vortexing and centrifugation. The supernatant was removed and the charcoal pellet washed three times with $dH_2O$, and transferred to a scintillation vial for analysis. Later analyses of the expressed A-domains used a simplified procedure [68, 69].

### N-Terminal Sequence Determination of NocB and 150 kDa Protein

Purified NocB and 150 kDa protein were analyzed on a 6.5% SDS-PAGE gel, transblotted onto a polyvinylidene fluoride membrane according to a standard protocol [65]. The bands corresponding to NocB and the 150 kDa protein were excised for N-terminal sequencing.

### Cloning L-pHPG Biosynthetic Genes

The coding regions for NocF, NocN, and NocG were PCR amplified from *N. uniformis* gDNA with primers *nocF*-NcoI (5′-TATACCATGGCGGCACAGGCAGGC–3′) and *nocF*-Hind (5′-TTATAAGCTTGCGCTCGGTCCGGTGGC-3′) for NocF, *nocN*-EcoRI (5′-AGAATTCGGCGTCCGCAACTCCGCAG-3′) and *nocN*-CtermHind (5′-AATAAGCTTTCAGGGCGCACCTCGCC-3′) for NocN, *nocG*-EcoRI (5′-AGAATTCGACCCGCTGCGCGCCGGTG-3′) and *nocG*-CtermHind (5′-AATAAGCTTTCAGGGCGCACCTCGCC-3′) for NocG. The introduced restriction sites NcoI, HindIII, and EcoRI are underlined. The purified amplicons were digested and ligated independently into the appropriate restriction sites on the plasmids pET29b for *nocF*, and pMAL-c4 for *nocG* and *nocN* and confirmed by sequencing. pET29-nocF possessed in-frame N-terminal S-tag and C-terminal His-6 tag while NocG and NocN were obtained appended to an N-terminal maltose binding protein (MBP). pET29-nocF was transformed into *E. coli* strain BL21 (DE3), while pMAL constructs were transformed by electroporation into *E. coli* TB1 cells. pMAL-nocG was overexpressed at 37, 28, and 15°C, but was not soluble under these expression conditions. Attempts to resolubilize NocG from inclusion bodies were performed using the Novagen Protein Refolding Kit (Madison, WI).

### Overexpression and Purification of NocF

NocF was overproduced and purified according to the pET expression protocol (Novagen). Cell free extract was incubated on ice for 1 hr with Qiagen $Ni^{2+}$-NTA resin, and purified according to manufacturer's recommendations. The homogenous protein was desalted on a Bio-RAD Econo-Pac 10DG desalting column equilibrated with Assay Buffer (50 mM Tris [pH 8], 10% glycerol).

### Overexpression and Purification of NocN

NocN was overproduced and purified according to the pMAL expression protocol (New England Biolabs). The desired protein was obtained in ≥ 80% purity by elution with buffer (20 mM Tris-HCl [pH 7.4], 200 mM NaCl, and 1 mM EDTA) containing 10 mM maltose. Fractions containing NocN-pMAL were pooled, concentrated, and applied to HQH anion exchange column on a BioCAD 700E Perfusion Chromatography Workstation (PerSeptive Biosystems, Framingham, MA). A salt gradient of 0-1 M NaCl (50 mM Tris [pH 8], 10% glycerol) was used to elute the enzyme. Homogenous, active enzyme was obtained following application to PE20 hydrophobic interaction column and elution with a gradient of 1.8–0 M $(NH_4)_2SO_4$.

### Assay of NocF Activity

The reaction mixture (500 μL) contained enzyme (100 μL), 50 mM Tris (pH 8), 0.25 mM ferrous ion, 0.5 mM ascorbate (as ferrous ascorbate), 0.1 mg/mL bovine liver catalase, and 1 mM substrate, *p*-hydroxyphenylpyruvate. The reaction was started by addition of substrate and allowed to proceed for 1 hr at room temperature. The reaction was stopped by addition of acetonitrile (15% final concentration) and the solution was centrifuged 20 min at 16,000 × g. Products were analyzed by HPLC on a Phenomenex Luna C18 (5μ) analytical column (250 × 4.6 mm). The solvents used were $ddH_20$, 0.01% TFA (A) and acetonitrile, 0.01% TFA (B). 100 μl in each reaction was analyzed under the following gradient conditions: 5 min 100% A/0% B, 20 min linear gradient from 100% A/0% B to 80% A/20% B, 5 min 80% A/20% B, 10 min 100% A/0% B. The elution of components was monitored at 270 nm.

### Assay of nocF Product Stereochemistry

A 10-times scale NocF activity assay was performed and the components were separated by HPLC using a Phenomenex Luna C18 (5μ) semipreparative column (250 × 10 mm) using the gradient described above. The peak corresponding to pHMA was collected for each injection, lyophilized, and resuspended in 200 μl $ddH_2O$. The compound was analyzed on an Astec Chirobiotic T (250 × 4.6 mm) chiral column (Whippany, NJ) using an isocratic gradient of 20% methanol, 1% triethylammonium acetate (pH 4.5).

### Assay of NocN Activity

The assay mixture (500 μL) contained 50 mM Tris buffer (pH 8) and 2 mM (*R,S*) *p*-hydroxymandelic acid, or *S*-pHMA alone. The reaction was initiated with the addition of partially purified NocN at a final concentration of 0.1 mg/mL (50 μL). Formation of *p*-hydroxybenzylformate (pHBF) was monitored spectrophotometrically at 332 nm. Alternately, the production of pHBF was monitored by HPLC under the conditions used above for NocF activity.

### Acknowledgments

We gratefully acknowledge financial support of this work by the National Institute of Health (AI014937 and GM020757 to J.R.). We thank the Biosynthesis & Sequencing Facility (The Johns Hopkins University School of Medicine, Baltimore, MD) for their tireless efforts, Prof. Juan F. Martin (University of Leon, Leon, Spain) for generously providing the pULVK2 shuttle vector, and Dr. Mervin J. Bibb (John Innes Center, Norwich, UK) for the gift of pIJ4070.

Received: February 9, 2004
Revised: March 29, 2004
Accepted: April 21, 2004
Published: July 23, 2004

### References

1. Aoki, H., Sakai, H., Kohsaka, M., Konomi, T., and Hosoda, J. (1976). Nocardicin A, a new monocyclic β-lactam antibiotic. I. Discovery, isolation and characterization. J. Antibiot. (Tokyo) *29*, 492–500.
2. Hashimoto, M., Komori, T.A., and Kamiya, T. (1976). Letter: Nocardicin A and B, novel monocyclic β-lactam antibiotics from a *Nocardia* species. J. Am. Chem. Soc. *98*, 3023–3025.
3. Kojo, H., Mine, Y., Nishida, M., Goto, S., and Kuwahara, S. (1988). Nature of monocyclic β-lactam antibiotic nocardicin A to β-lactamases. Microbiol. Immunol. *32*, 119–130.
4. Watanabe, K., Okuda, T., Yokose, K., Furumai, T., and Maruyama, H.B. (1983). *Actinosynnema mirum*, a new producer of nocardicin antibiotics. J. Antibiot. (Tokyo) *36*, 321–324.
5. Celmer, W.D., Huang, L.H., Jefferson, M.T., Maeda, H., Inoue, K., and Shibakawa, R. (1980). Process for producing nocardicin A. U.S. patent 4,212,944.
6. Tomita, K., Tsukiura, H., and Kawaguchi, H. (1982). Fermentation process for producing nocardicins A and B. U.S. patent 4,320,199.
7. Hosoda, J., Konomi, T., Tani, N., Aoki, H., and Imanaka, H. (1977). Biosynthesis of nocardicin A by *Nocardia uniformis* subsp. *tsuyamanensis*. 2. Isolation of new nocardicins from *Nocardia uniformis* subsp. *tsuyamanensis*. Agric. Biol. Chem. *41*, 2013–2020.
8. Kamiya, T. (1977). Studies on the new monocyclic β-lactam antibiotics, nocardicins. In Proceedings of International Symposium of the Chemical Society, J. Elks, ed. (London), pp. 281–294.
9. Townsend, C.A., and Salituro, G.M. (1984). Fate of [N-15]-(*p*-hydroxyphenyl)glycine in nocardicin A biosynthesis. J. Chem. Soc. Chem. Commun. 1631–1632.
10. Hosoda, J., Tani, N., Konomi, T., Ohsawa, S., Aoki, H., and Imanaka, H. (1977). Biosynthesis of nocardicin A by *Nocardia uniformis* subsp. *tsuyamanensis*. 1. Incorporation of C-14-

amino acids into nocardicin A by growing cells. Agric. Biol. Chem. *41*, 2007–2012.
11. Townsend, C.A., and Brown, A.M. (1981). Biosynthetic studies of nocardicin A. J. Am. Chem. Soc. *103*, 2873–2874.
12. Townsend, C.A., and Brown, A.M. (1983). Nocardicin A: biosynthetic experiments with amino acid precursors. J. Am. Chem. Soc. *105*, 913–918.
13. Townsend, C.A., Brown, A.M., and Nguyen, L.T. (1983). Nocardicin A: stereochemical and biomimetic studies of monocyclic β-lactam formation. J. Am. Chem. Soc. *105*, 919–927.
14. Townsend, C.A., and Brown, A.M. (1982). Nocardicin A biosynthesis - stereochemical course of monocyclic β-lactam formation. J. Am. Chem. Soc. *104*, 1748–1750.
15. Townsend, C.A., and Nguyen, L.T. (1981). Asymmetric, biogenetically modeled synthesis of (-)-3-aminonocardicinic acid. J. Am. Chem. Soc. *103*, 4582–4583.
16. Townsend, C.A., and Wilson, B.A. (1988). The role of nocardicin G in nocardicin A biosynthesis. J. Am. Chem. Soc. *110*, 3320–3321.
17. Wilson, B.A. (1989). Studies on the biosynthesis of nocardicin A. PhD thesis, The Johns Hopkins University, Baltimore, Maryland.
18. Reeve, A.M., Breazeale, S.D., and Townsend, C.A. (1998). Purification, characterization, and cloning of an *S*-adenosylmethionine- dependent 3-amino-3-carboxypropyltransferase in nocardicin biosynthesis. J. Biol. Chem. *273*, 30695–30703.
19. Kelly, W.L., and Townsend, C.A. (2002). Role of the cytochrome P450 NocL in nocardicin A biosynthesis. J. Am. Chem. Soc. *124*, 8186–8187.
20. Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N., and Schoner, B.E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene *116*, 43–49.
21. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.
22. van Wageningen, A.M., Kirkpatrick, P.N., Williams, D.H., Harris, B.R., Kershaw, J.K., Lennard, N.J., Jones, M., Jones, S.J., and Solenberg, P.J. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. Chem. Biol. *5*, 155–162.
23. Chong, P.P., Podmore, S.M., Kieser, H.M., Redenbach, M., Turgay, K., Marahiel, M., Hopwood, D.A., and Smith, C.P. (1998). Physical identification of a chromosomal locus encoding biosynthetic genes for the lipopeptide calcium-dependent antibiotic (CDA) of *Streptomyces coelicolor* A3(2). Microbiol. *144*, 193–199.
24. Chiu, H.T., Hubbard, B.K., Shah, A.N., Eide, J., Fredenburg, R.A., Walsh, C.T., and Khosla, C. (2001). Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. Proc. Natl. Acad. Sci. USA *98*, 8548–8553.
25. Hubbard, B.K., Thomas, M.G., and Walsh, C.T. (2000). Biosynthesis of L-p-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics. Chem. Biol. *7*, 931–942.
26. Shaw, K.J., Rather, P.N., Hare, R.S., and Miller, G.H. (1993). Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes. Microbiol. Rev. *57*, 138–163.
27. Sugiyama, M., Kumagai, T., Shionoya, M., Kimura, E., and Davies, J.E. (1994). Inactivation of bleomycin by an N-acetyltransferase in the bleomycin-producing strain *Streptomyces verticillus*. FEMS Microbiol. Lett. *121*, 81–85.
28. Martinez-Hackert, E., and Stock, A.M. (1997). The DNA-binding domain of OmpR: crystal structures of a winged helix transcription factor. Structure *5*, 109–124.
29. Horinouchi, S., Kito, M., Nishiyama, M., Furuya, K., Hong, S.K., Miyake, K., and Beppu, T. (1990). Primary structure of AfsR, a global regulatory protein for secondary metabolite formation in *Streptomyces coelicolor* A3(2). Gene *95*, 49–56.
30. Wietzorrek, A., and Bibb, M. (1997). A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. Mol. Microbiol. *25*, 1181–1184.
31. Penrose, D.M., and Glick, B.R. (1997). Enzymes that regulate ethylene levels–1-aminocyclopropane-1-carboxylic acid (ACC) deaminase, ACC synthase and ACC oxidase. Indian J. Exp. Biol. *35*, 1–17.
32. Minami, R., Uchiyama, K., Murakami, T., Kawai, J., Mikami, K., Yamada, T., Yokoi, D., Ito, H., Matsui, H., and Honma, M. (1998). Properties, sequence, and synthesis in *Escherichia coli* of 1-aminocyclopropane-1-carboxylate deaminase from *Hansenula saturnus*. J. Biochem. (Tokyo) *123*, 1112–1118.
33. Yao, M., Ose, T., Sugimoto, H., Horiuchi, A., Nakagawa, A., Wakatsuki, S., Yokoi, D., Murakami, T., Honma, M., and Tanaka, I. (2000). Crystal structure of 1-aminocyclopropane-1-carboxylate deaminase from *Hansenula saturnus*. J. Biol. Chem. *275*, 34557–34565.
34. Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., 3rd, et al. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature *393*, 537–544.
35. Jendrossek, D., Frisse, A., Behrends, A., Andermann, M., Kratzin, H.D., Stanislawski, T., and Schlegel, H.G. (1995). Biochemical and molecular characterization of the *Pseudomonas lemoignei* polyhydroxyalkanoate depolymerase system. J. Bacteriol. *177*, 596–607.
36. Wright, R.L., Harris, K., Solow, B., White, R.H., and Kennelly, P.J. (1996). Cloning of a potential cytochrome P450 from the archaeon *Sulfolobus solfataricus*. FEBS Lett. *384*, 235–239.
37. Molnar, I., Aparicio, J.F., Haydock, S.F., Khaw, L.E., Schwecke, T., Konig, A., Staunton, J., and Leadlay, P.F. (1996). Organisation of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of genes flanking the polyketide synthase. Gene *169*, 1–7.
38. Poulos, T.L., Finzel, B.C., and Howard, A.J. (1987). High-resolution crystal structure of cytochrome P450cam. J. Mol. Biol. *195*, 687–700.
39. O'Keefe, D.P., and Harder, P.A. (1991). Occurrence and biological function of cytochrome P450 monooxygenases in the actinomycetes. Mol. Microbiol. *5*, 2099–2105.
40. Doekel, S., and Marahiel, M.A. (2001). Biosynthesis of natural products on modular peptide synthetases. Metab. Eng. *3*, 64–77.
41. Marahiel, M.A., Stachelhaus, T., and Mootz, H.D. (1997). Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. Chem. Rev. *97*, 2651–2674.
42. Mootz, H.D., Schwarzer, D., and Marahiel, M.A. (2002). Ways of assembling complex natural products on modular nonribosomal peptide synthetases. Chembiochem *3*, 491–504.
43. Challis, G.L., Ravel, J., and Townsend, C.A. (2000). Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem. Biol. *7*, 211–224.
44. Stachelhaus, T., Mootz, H.D., and Marahiel, M.A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem. Biol. *6*, 493–505.
45. Recktenwald, J., Shawky, R., Puk, O., Pfennig, F., Keller, U., Wohlleben, W., and Pelzer, S. (2002). Nonribosomal biosynthesis of vancomycin-type antibiotics: a heptapeptide backbone and eight peptide synthetase modules. Microbiol. *148*, 1105–1118.
46. Rouhiainen, L., Paulin, L., Suomalainen, S., Hyytiainen, H., Buikema, W., Haselkorn, R., and Sivonen, K. (2000). Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in *Anabaena* strain 90. Mol. Microbiol. *37*, 156–167.
47. Yu, S., Fiss, E., and Jacobs, W.R., Jr. (1998). Analysis of the exochelin locus in *Mycobacterium smegmatis*: biosynthesis genes have homology with genes of the peptide synthetase family. J. Bacteriol. *180*, 4676–4685.
48. Challis, G.L., and Ravel, J. (2000). Coelichelin, a new peptide siderophore encoded by the *Streptomyces coelicolor* genome: structure prediction from the sequence of its non- ribosomal peptide synthetase. FEMS Microbiol. Lett. *187*, 111–114.
49. Kumar, V., de la Fuente, J.L., Leitao, A.L., Liras, P., and Martin, J.F. (1996). Effect of amplification or targeted disruption of the beta-lactamase gene of *Nocardia lactamdurans* on cephamycin biosynthesis. Appl. Microbiol. Biotechnol. *45*, 621–628.

50. Cole, P.A. (1996). Chaperone-assisted protein expression. Structure *4*, 239–242.
51. Quadri, L.E., Sello, J., Keating, T.A., Weinreb, P.H., and 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.
52. Du, L., Sanchez, 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.
53. Gilhuus-Moe, C.C., Kristensen, T., Bredesen, J.E., Zimmer, T., and Laland, S.G. (1970). The presence and possible role of phosphopantothenic acid in gramicidin S synthetase. FEBS Lett. *7*, 287–290.
54. Shikata, Y., Watanabe, T., Teramoto, T., Inoue, A., Kawakami, Y., Nishizawa, Y., Katayama, K., and Kuwada, M. (1995). Isolation and characterization of a peptide isomerase from funnel web spider venom. J. Biol. Chem. *270*, 16719–16723.
55. Hoffmann, K., Schneider-Scherzer, E., Kleinkauf, H., and Zocher, R. (1994). Purification and characterization of eucaryotic alanine racemase acting as key enzyme in cyclosporin biosynthesis. J. Biol. Chem. *269*, 12710–12714.
56. Salituro, G.M., and Townsend, C.A. (1990). Total syntheses of (-)-nocardicins a-G—a biogenetic approach. J. Am. Chem. Soc. *112*, 760–770.
57. Marshall, C.G., Hillson, N.J., and Walsh, C.T. (2002). Catalytic mapping of the vibriobactin biosynthetic enzyme VibF. Biochemistry *41*, 244–250.
58. Du, L., Chen, M., Zhang, Y., and Shen, B. (2003). BlmIII and BlmIV nonribosomal peptide synthetase-catalyzed biosynthesis of the bleomycin bithiazole moiety involving both in cis and in trans aminoacylation. Biochemistry *42*, 9731–9740.
59. Hamamoto, K., Kida, Y., Zhang, Y., Shimizu, T., and Kuwano, K. (2002). Antimicrobial activity and stability to proteolysis of small linear cationic peptides with D-amino acid substitutions. Microbiol. Immunol. *46*, 741–749.
60. Baldwin, J.E. (1976). Rules for ring-closure. J. Chem. Soc. Chem. Comm. 734–736.
61. Schneider, A., Stachelhaus, T., and Marahiel, M.A. (1998). Targeted alteration of the substrate specificity of peptide synthetases by rational module swapping. Mol. Gen. Genet. *257*, 308–318.
62. Roach, P.L., Clifton, I.J., Hensgens, C.M., Shibata, N., Schofield, C.J., Hajdu, J., and Baldwin, J.E. (1997). Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. Nature *387*, 827–830.
63. Bachmann, B.O., Li, R., and Townsend, C.A. (1998). β-Lactam synthetase: a new biosynthetic enzyme. Proc. Natl. Acad. Sci. USA *95*, 9082–9086.
64. Gerratana, B., Stapon, A., and Townsend, C.A. (2003). Inhibition and alternate substrate studies on the mechanism of carbapenam synthetase from *Erwinia carotovora*. Biochemistry *42*, 7836–7847.
65. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1995). Current Protocols in Molecular Biology, Third Edition (Boston: John Wiley and Sons).
66. Schupp, T., and Divers, M. (1986). Protoplast preparation and regeneration in Nocardia mediterranei. FEMS Microbiol. Lett. *36*, 159–162.
67. Madon, J., and Hutter, R. (1991). Transformation system for *Amycolatopsis (Nocardia) mediterranei*: direct transformation of mycelium with plasmid DNA. J. Bacteriol. *173*, 6325–6331.
68. Du, L., and Shen, B. (1999). Identification and characterization of a type II peptidyl carrier protein from the bleomycin producer *Streptomyces verticillus* ATCC 15003. Chem. Biol. *6*, 507–517.
69. Ku, J., Mirmira, R.G., Liu, L., and 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.

**Accession Numbers**

Accession numbers will be released on January 1, 2005. Please contact the corresponding author for more information.