5-29-2015

The Biosynthesis of Capuramycin-type Antibiotics: Identification of the A-102395 Biosynthetic Gene Cluster, Mechanism of Self-Resistance, and Formation of Uridine-5'-Carboxamide

Wenlong Cai  
*University of Kentucky*, wenlong.cai@uky.edu

Anwesha Goswami  
*University of Kentucky*, ago223@uky.edu

Zhaoyong Yang  
*Chinese Academy of Medical Sciences, China*

Xiaodong Liu  
*University of Kentucky*, xiaodong.liu@uky.edu

*See next page for additional authors*

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Follow this and additional works at: [https://uknowledge.uky.edu/ps_facpub](https://uknowledge.uky.edu/ps_facpub)

Part of the *Pharmacy and Pharmaceutical Sciences Commons*
The Biosynthesis of Capuramycin-type Antibiotics: Identification of the A-102395 Biosynthetic Gene Cluster, Mechanism of Self-Resistence, and Formation of Uridine-5'-Carboxamide

Digital Object Identifier (DOI)
https://doi.org/10.1074/jbc.M115.646414

Notes/Citation Information
Published in *The Journal of Biological Chemistry*, v. 290, no. 22, p. 13710-13724.

This research was originally published in *The Journal of Biological Chemistry*. Wenlong Cai, Anwesha Goswami, Zhaoyong Yang, Xiaodong Liu, Keith D. Green, Sandra Barnard-Britson, Satoshi Baba, Masanori Funabashi, Koichi Nonaka, Manjula Sunkara, Andrew J. Morris, Anatol P. Spork, Christian Ducho, Sylvie Garneau-Tsodikova, Jon S. Thorson, and Steven G. Van Lanen. The Biosynthesis of Capuramycin-type Antibiotics: Identification of the A-102395 Biosynthetic Gene Cluster, Mechanism of Self-Resistence, and Formation of Uridine-5'-Carboxamide. *The Journal of Biological Chemistry*. 2015; 290:13710-13724. © the American Society for Biochemistry and Molecular Biology.

The copyright holder has granted the permission for posting the article here.

Authors
Wenlong Cai, Anwesha Goswami, Zhaoyong Yang, Xiaodong Liu, Keith D. Green, Sandra Barnard-Britson, Satoshi Baba, Masanori Funabashi, Koichi Nonaka, Manjula Sunkara, Andrew J. Morris, Anatol P. Spork, Christian Ducho, Sylvie Garneau-Tsodikova, Jon S. Thorson, and Steven Van Lanen

This article is available at UKnowledge: https://uknowledge.uky.edu/ps_facpub/70
The Biosynthesis of Capuramycin-type Antibiotics

IDENTIFICATION OF THE A-102395 BIOSYNTHETIC GENE CLUSTER, MECHANISM OF SELF-RESISTANCE, AND FORMATION OF URIDINE-5’-CARBOXAMIDE

Received for publication, February 20, 2015, and in revised form, March 20, 2015. Published, JBC Papers in Press, April 8, 2015, DOI 10.1074/jbc.M115.646414

Wenlong Cai‡, Anwesha Goswami‡, Zhaoyong Yang‡, Xiaodong Liu‡, Keith D. Green‡, Sandra Barnard-Britson‡, Satoshi Baba‡, Masanori Funabashi‡, Koichi Nonaka**, Manjula Sunkara‡†, Andrew J. Morris‡†, Anatol P. Spork‡‡, Christian Ducho‡‡, Sylvie Garneau-Tsodikova‡, Jon S. Thorson‡†, and Steven G. Van Lanen‡†

From the ‡Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, Lexington, Kentucky 40506, the §Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 1000050, China, the ‡New Modality Research Laboratories, R&D Division, Daiichi Sankyo Co., Ltd., Tokyo 103-8426, Japan, the ‡Drug Discovery and Biomedical Technology Unit, Daiichi Sankyo RD Novare Co., Ltd., Tokyo, Japan, the **Biologics Technology Research Laboratories, R&D Division, Daiichi Sankyo Co., Ltd., Tokyo 103-8426, Japan, the §§Division of Cardiovascular Medicine and Gill Heart Institute, College of Medicine, University of Kentucky, Lexington, Kentucky 40506, and the ‡‡Department of Pharmacy, Pharmaceutical and Medicinal Chemistry, Saarland University, 66123 Saarbrücken, Germany

Background: Several nucleoside antibiotics contain a uridine-5’-carboxamide core of unclear origin.

Results: The A-102395 biosynthetic gene cluster was cloned, a genetic system was developed, and three enzymes were characterized in vivo and in vitro.

Conclusion: Uridine-5’-carboxamide originates from UMP and L-Thr by sequential reactions catalyzed by a dioxygenase and transaldolase.

Significance: The results provide the first opportunity to methodically interrogate the biosynthesis of these unusual antibiotics.

A-500359s, A-503083s, and A-102395 are capuramycin-type nucleoside antibiotics that were discovered using a screen to identify inhibitors of bacterial translocase I, an essential enzyme in peptidoglycan cell wall biosynthesis. Like the parent capuramycin, A-500359s and A-503083s consist of three structural components: a uridine-5’-carboxamide (CarU), a rare unsaturated hexuronic acid, and an aminocaprolactam, the last of which is substituted by an unusual arylamine-containing polyamide in A-102395. The biosynthetic gene clusters for A-500359s and A-503083s have been reported, and two genes encoding a putative non-heme Fe(II)-dependent α-ketoglutarate:UMP dioxygenase and an L-Thr:uridine-5’-aldehyde transaldolase were uncovered, suggesting that C–C bond formation during assembly of the high carbon (C6) sugar backbone of CarU proceeds from the precursors UMP and L-Thr to form 5’-C-glycuridine (C7) as a biosynthetic intermediate. Here, isotopic enrichment studies with the producer of A-503083s were used to indeed establish L-Thr as the direct source of the carboxamide of CarU. With this knowledge, the A-102395 gene cluster was subsequently cloned and characterized. A genetic system in the A-102395-producing strain was developed, permitting the inactivation of several genes, including those encoding the dioxygenase (cpr19) and transaldolase (cpr25), which abolished the production of A-102395, thus confirming their role in biosynthesis. Heterologous production of recombinant Cpr19 and CapK, the transaldolase homolog involved in A-503083 biosynthesis, confirmed their expected function. Finally, a phosphotransferase (Cpr17) conferring self-resistance was functionally characterized. The results provide the opportunity to use comparative genomics along with in vivo and in vitro approaches to probe the biosynthetic mechanism of these intriguing structures.

Capuramycin-type antibiotics include A-500359s from Streptomyces griseus SANK 60196 (1–4), A-503083s from Streptomyces sp. SANK 62799 (5), and A-102395 from Amycolatopsis sp. SANK 60206 (6). They were discovered using an activity-based screen to identify inhibitors of bacterial phospho-N-acetylmuramyl-pentapeptide translocase 1 (TL1; annotated as MraY), a potential, clinically unexploited antibiotic target that initiates the lipid cycle of peptidoglycan cell wall biosynthesis and is essential for bacterial survival (7, 8). Studies into the biological activity of A-500359 A, the major congener isolated from S. griseus SANK 60196, and several semisynthetic analogues have revealed a potential utility of these natural products as anti-tuberculosis antibiotics (9, 10). For example, SQ641 and SQ922, two leads under preclinical development by Sequella (Rockville, MD), have been shown to have several clinically desirable attributes, including in vitro activity against multiple-drug-resistant strains of Mycobacterium tuberculosis.

* This work was supported, in whole or in part, by National Institutes of Health, NIAID, Grant AI087849 (to S. V. L.). This work was also supported by National Center for Advancing Translational Sciences Grant UL1 TR000117, NSF CAREER award MCB-1149427 (to S. G.-T.), and the Kentucky Science and Education Foundation (to S. V. L.).

‡ This article contains supplemental Table 1.

† To whom correspondence should be addressed: Dept. of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, 789 S. Limestone, Lexington, KY 40506. Tel.: 859-323-6271; E-mail: svanlanen@uky.edu.

‡‡ The abbreviations used are: TL1, translocase 1; CarU, uridine-5’-carboxamide; L-ACL, L-α-amino-ε-caprolactam; xKG, α-ketoglutarate; GlyU, 5’-glycuridine; UA, uridine-5’-aldehyde; PABA, para-aminobenzoic acid, ADC, 4-amino-4-deoxychorismate; ACP, acyl carrier protein.
(the primary causative agent of tuberculosis); high efficacy in a murine model of tuberculosis; rapid kill time in vitro and in vivo; and no toxicity to mice (11–17). Given the widespread documentation of extensively drug-resistant M. tuberculosis and the recent reality of totally drug-resistant M. tuberculosis (18), the development of drugs with novel targets, such as the capuramycin-type antibiotics, makes them attractive leads for tuberculosis chemotherapy (19, 20).

The namesake capuramycin, initially discovered in 1986 from an antibacterial screening program (21), consists of three structurally distinct, modular components: a uridine-5′-carboxamide (CarU), an unsaturated α-d-mannopyranuronate, and an L-α-amino-ε-caprolactam (L-ACL) (Fig. 1A) (22). Structural elucidation of the A-500359s revealed that A-500359 B 1a [IC50 = 18 nM against TL1] is identical to the original capuramycin, and A-500359 A 1b contains a 6′′-C-methylated L-ACL, a modification that has no effect on TL1 inhibition (IC50 = 17 nM) (1, 2). Two additional structural variants of note include the deaminocaprolactam (de-ACL) congeners that contain a free carboxylic acid (A-500359 F, 1c) or a methyl ester (A-500359 E, 1d), the former of which has a significantly decreased TL1 inhibitory activity (IC50 = 1.1 μM), whereas the latter has a slight decrease (IC50 = 27 nM) (3). The A-503083s 2a-2d have structures identical to those of A-500359s with one exception; in all cases, CarU is modified with a 2′′′-O-carbamoyl group that has minimal effect on TL1 inhibition and antibacterial activity (5). A-102395 3 significantly diverges from the other capuramycins in both structure, wherein the L-ACL is substituted with an unusual arylamine-containing polypeptide, and activity, with 3 being the most potent inhibitor of TL1 (IC50 = 11 nM) yet completely inactive against Mycobacteria sp. and other bacterial strains (6).

Shortly after the rediscovery of capuramycin, the biosynthesis of each component was interrogated using feeding experiments with the producing strain of 1 (4). As expected, high incorporation was observed for L-ACL using L-[1-13C]Lys (16-fold enrichment of C-1′′); α-d-mannopyranuronate using D-[1-13C]mannose (11-fold enrichment of C-1′′′); and CarU using D-[1-13C]ribose (17-fold enrichment of C-1′′′) (Fig. 1B). The metabolic origin of the 5′-carboxamide of CarU was less clear, although a modest 3-fold enrichment of C-6′′ using [3,13C]pyruvate led to a proposal that the biosynthesis involves an aldol reaction between phosphoenolpyruvate and uridine-5′-aldehyde (UA). This mechanism of ribose chain extension is analogous to that previously proposed for the biosynthesis of polyoxin and nikkomycin, two nucleoside antibiotics that also contain a high carbon sugar albeit with distinct chemical features (23). However, the specific enzymatic transformations required for the assembly of the high carbon sugar nucleoside of either the capuramycins or polyoxin and nikkomycin have not been biochemically established.

The biosynthetic gene cluster for 1 and 2 have been identified, and evidence for the involvement of 21 open reading frames (ORFs) was provided by gene expression profiling of the wild-type, 1-producing strain and several null mutants created by random chemical mutagenesis (24, 25). Additional evidence that the correct gene cluster was identified was provided by functional assignment of genes involved in 2 biosynthesis: one (capP) encoding an ATP-dependent phosphotransferase that modifies the 3′-OH of the unsaturated α-d-mannopyranuronate as a mechanism of self-resistance (26) and the other (capW) encoding an L-ACL:2d transacylase (25). Bioinformatic analysis of the gene clusters unexpectedly revealed two shared ORFs encoding putative proteins with sequence similarity to LipL: a non-heme Fe(II)-dependent α-ketoglutarate (αKG):uridine-5′-monophosphate (UMP) dioxygenase (27) and LipK, a pyridoxal-5′-phosphate-dependent L-Thr:UA transaldolase (28). These two enzymes sequentially convert UMP (5′S,6′S)-5′-C-glycyluridine (GlyU) during the biosynthesis of A-90289, a...
member of a distinct group of TL1 inhibitors that contain a GlyU component in the final product (Fig. 2) (29, 30). Additional members of the GlyU-containing nucleoside antibiotics with identified biosynthetic gene clusters include liposidomycins (31), caprazamycins (32), muraymycin (33), muraminomicin (34), and the recently discovered sphaerimicins (35). Thus, the uncovering of these two genes within the clusters for 1 and 2 suggested that the biosynthesis of CarU proceeds via condensation of L-Thr with UA to generate GlyU as a cryptic pathway intermediate. Herein we provide evidence through newly designed feeding experiments that this is indeed the case. The probability for the involvement of an L-Thr:UA transaldolase in CarU biosynthesis was subsequently exploited to clone and characterize the 3 gene cluster. The first genetic system for a producer of the capuramycin-type antibiotics was developed using the 3-producing strain, which, in addition to in vitro characterization of enzymes with sequence similarity to CapP (Cpr17), LipL (Cpr19), and LipK (CapH), has provided new insight into the resistance and biosynthetic strategies for the capuramycin-type antibiotics.

**Experimental Procedures**

*Standards—* UA and GlyU were prepared following previously described procedures (27, 28). 2a and 2c were isolated from *Streptomyces* sp. SANK 62799, and 3 was isolated from *Amycolatopsis* sp. SANK 60206 following the procedures described (5, 6). Mass, 1H, and 13C NMR spectroscopic analyses of 3 were identical with the prior report (Fig. 3).

*Isotopic Enrichment—* Fermentation media and growth conditions for *Streptomyces* sp. SANK 62799 were as described previously (5). A seed culture was incubated at 28 °C for 48 h, when 1.5 ml was used to inoculate fresh media (50 ml of liquid medium in a 250-ml flask). After fermentation for 70 h, 25 mg of filter-sterilized [1-13C]Gly, [2-13C]Gly, or L-[13C4,15N]Thr was
added to each flask. Fermentation was continued an additional 72 h. An equal volume of methanol was added directly to the culture and mixed vigorously prior to centrifugation to remove cell debris. The supernatant was lyophilized, and the dried powder was resuspended in water (100 mg/ml) for HPLC purification using a C-18 reverse phase semipreparative column. A series of linear gradients was developed from A (0.1% TFA, 2.5% acetonitrile) to B (0.1% TFA, 90% acetonitrile) in the following manner (beginning time and ending time with linear increase to percentage of B): 0–6 min, 0% B; 6–26 min, 100% B; 26–30 min, 100% B; 30–34 min, 0% B; and 34–35 min, 0% B. The flow rate was kept constant at 3.5 ml/min, and elution was monitored at 260 nm. Purified 2a and 2c were analyzed by NMR, and relative peak intensities were assigned based on C-2" and C-3 from the natural abundance 13C NMR spectra.

Cloning of the 3 Gene Cluster—Amycolatopsis sp. SANK 60206 genomic DNA was partially digested with Sau3AI to give ~40-kb DNA fragments that were dephosphorylated with bacterial alkaline phosphatase and ligated into BamHI-digested cosmid vector SuperCos1 (Stratagene, Cedar Creek, TX), which was dephosphorylated by bacterial alkaline phosphatase after XbaI digestion. The ligation products were packaged with Gigapack III Gold packaging extract as described by the manufacturer (Stratagene), and the resulting recombinant phage was used to transfect Escherichia coli XL-1 Blue MR. Approximately 8,000 colonies from the obtained genomic library were screened by colony hybridization using a digoxigenin-labeled lipK/capH homologous DNA fragment obtained by PCR using degenerate primers (supplemental Table 1) and genomic DNA from Amycolatopsis sp. SANK 60206. Hybridization was carried out using DIG Easy Hyb (Roche Applied Science) at 42 °C, and the resulting filter was washed under high stringency conditions (0.1× SSC including 0.1% SDS, 68 °C). Detection was performed using CDP-Star (Roche Applied Science) according to the manufacturer’s procedures.

Based on restriction digest analysis, three positive cosmids, pNCap01, pNCap02, and pNCap03, were isolated and sequenced using a Roche Applied Science GS FLX system (Operon Biotechnologies). The terminal region of pNCap03 was used as a probe to identify a 10-kb DNA fragment following BamHI digestion of genomic DNA, which was cloned into pUC19 to yield pUC/B10k. This insert was subsequently used as a probe to rescreen the library to identify a fourth cosmid, pNCap04, which was also sequenced to complete the genetic locus. Potential open reading frames were defined using Frameplot version 4.0, and database comparison for sequence homology was performed with BLAST search tools using the National Center for Biotechnology Information (Bethesda, MD).

Gene Inactivation and Production Analysis—Genes were inactivated using REDIRECT technology (36). In short, pNCap03 was introduced into E. coli BW25141/pKD78, and the linear PCR fragment obtained with template pJJ773 was introduced by electroporation. Following confirmation of the genotype using PCR, the modified cosmids were introduced into Amycolatopsis sp. SANK 60206 by conjugation using E. coli ET12567 (pUZ8002). Apramycin resistance was used to select for a double-crossover event, the genotype of which was subsequently confirmed by both PCR and Southern blot analysis.

For comparative analysis of 3 production in the wild-type and mutant strains, 2 liters of culture broth was centrifuged at 2,600 × g for 20 min to remove mycelia. The pH of the supernatant was adjusted to 3.0 with concentrated HCl followed by the addition of an equal volume of methanol. After vigorous mixing for 1 h, the insoluble material was removed by centrifugation (6,000 × g for 40 min), and the solvent from the recovered supernatant was removed by rotary evaporation and lyophilization. Dried bacterial extracts (250–400 mg) were dissolved in 1.6 ml of 1:1 water/methanol, and the pH was adjusted to 3–4 with concentrated HCl if necessary prior to the addition of 6.4 ml of cold acetonitrile. The resulting mixture was vortexed for 5 min, followed by sonication for 1 min, and the insoluble material was removed by centrifugation (2,600 × g for 10 min). The supernatant was recovered and dried under a stream of nitrogen. The dried sample was resuspended in 50 μl of methanol for analysis by LC-MS-MS.

LC-MS-MS analysis of 3 was carried out using a Shimadzu UFLC equipped with an Apollo C18 column (250 × 4.6 mm, 5 μm) coupled to an AB Scieix 4000-Qtrap hybrid linear trap triple quadrupole mass spectrometer in multiple-reaction-monitoring mode. The mobile phase consisted of 10 mM ammonium formate with 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). A series of linear gradients was developed from A to B in the following manner (beginning time and ending time with linear increase to percentage of B): 0 min, 2% B; 0–8 min, 2% B; 8–27 min, 90% B; 27–30 min, 90% B; and 30–32 min, 2% B. The flow rate was kept constant at 0.5 ml/min with a column temperature of 30 °C, and elution was monitored at 260 nm. The mass spectrometer was operated in the negative electrospray ionization mode with optimal ion source settings determined with 3 standard with a declustering potential of −135 V, entrance potential of −10 V, collision energy of −60 V, collision cell exit potential of −7 V, curtain gas of 20 p.s.i., ion spray voltage of −4,200 V, ion source gas 1/gas 2 of 40 p.s.i., and temperature of 550 °C. Multiple-reaction-monitoring transitions monitored were as follows: 763.316/189.9 and 763.316/120. Negative electrospray ionization mode was used to analyze samples because it was more sensitive compared with positive electrospray ionization mode for 3.

Reverse Transcriptase-PCR Analysis—All procedures were performed according to the provided protocol. A TRizol Max bacterial RNA isolation kit (Ambion) was used to extract total RNA from wild-type and mutant strains. The SuperScript III first-strand synthesis system for RT-PCR (Invitrogen) was used to synthesize the total cDNA from total RNA. Takara Taq was used to amplify target genes using total cDNA as template.

Characterization of Recombinant Enzymes—Procedures for cloning of genes, heterologous production and purification of proteins, and enzyme assays used standard protocols (26–28).

The primers used for cloning were as follows: cpr17, 5'-GGT-ATTGAGGGTGCTGCATGACCGAGACGAGATCGAC-3' (forward)/5'-AGAGAGAGTTAGGGCCCATATCCGGTGGAATCGGGCCAG-3' (reverse); cpr19, 5'-GGTATTGGAGGGTGCTGCATGCAGCAGCTGCAAGCCG-3' (forward)/5'-AGAGAGAGTTAGGGCCCATATCCGGTGGAATCGGGCCAG-3' (reverse); capH, 5'-GGTATTGGAGGGTGCTGCATGCAGCAGCTGCAAGCCG-3' (forward)/5'-
Results

GlyU Is a Biosynthetic Intermediate for CarU—An ORF encoding a protein with high amino acid sequence similarity to the transaldolase LipK is located within the two gene clusters that have been identified for the CarU-containing nucleoside antibiotics: 1 (ORF14, 55% identity/81% similarity) and 2 (CapH, 47% identity/79% similarity) (35). Similarly, an ORF encoding a protein with moderate sequence similarity to the dioxygenase LipL is found within the 1 and 2 gene clusters (Table 1). Both LipK and LipL were previously biochemically characterized to establish that C-6/\(^{13}C\)H\(_{11032}\) and C-7/\(^{13}C\)H\(_{11032}\) of GlyU originates from C-2 and C-1, respectively, of L-Thr through a transaldol-like reaction using UA as the aldol acceptor (Fig. 2) (27, 28). LipK, the catalyst for the transaldol-like reaction, was completely inactive when L-Thr was replaced with Gly or L-Ser despite having the closest sequence homology with proteins annotated as serine hydroxymethyltransferases that utilize tetrahydrofolic acid to interconvert Gly and L-Ser (28). With this information in hand, the metabolic origin of the 5/\(^{13}C\)H\(_{11032}\)-carboxamide of CarU was re-investigated by undertaking feeding experiments with isotopically enriched Gly and L-Thr using the 2-producing strain.

To initially probe the metabolic origin of the 5’-carboxamide, either [1-\(^{13}C\)]Gly or [2-\(^{13}C\)]Gly was added to the culture both of the 2-producing strain. Under the growth conditions employed, 2a was the major congener produced and, as predicted from the bioinformatics analysis, no significant enrichment at C-6’ was observed based on comparative analysis of the \(^{13}C\) NMR spectra (data not shown). However, a 6% enrichment of the 3’-O-methyl group was observed using [2-\(^{13}C\)]Gly, which is similar to incorporation results that have been reported for other actinomycete-derived O-methylated metabolites (37). In contrast to the Gly enrichment experiments, the de-ACL derivative 2c was the major congener produced when fed with L-[1\(^{13}C\)\(_4\), 1\(^{15}N\)]Thr. Comparative analysis of the \(^{13}C\) NMR spectra of purified 2c revealed a 15% enrichment at C-6’ (Fig. 4), which is substantially greater than that previously reported with pyruvate (4). This high level of enrichment is comparable with the levels observed for enrichment of L-ACL with L-Lys, which has been supported by biochemical studies (25). Thus, the data are consistent with direct incorporation of L-Thr via a transaldol-like mechanism. Additionally, a \(^{13}C\)-\(^{15}N\) heteronuclear spin coupling at C-6’ of the \(^{13}C\) NMR spectrum (J\(_{CN}\) = 18 Hz) was detected that, along with analysis of the LC-MS spectrum (data not shown), is consistent with concomitant \(^{15}N\) incorporation from L-Thr during the biosynthesis of CarU.

Transaldolase as a Genetic Fingerprint for Identifying the 3 Gene Cluster—After providing evidence that CarU biosynthesis proceeds with direct incorporation of L-Thr, the probable involvement of an L-Thr:UA transaldolase was exploited as a genetic fingerprint to identify the biosynthetic gene cluster for 3. We have previously designed degenerate primers based on

\[\text{AGAGGAGAGTAGAGGAGAGACGAACTCGACGAAATAGAAGGGAG} \]

\[\text{3'} \text{ (reverse). DNA templates for PCR}\]

cloning were pNCap02 for cpr17 and cpr19 and cosmid N-4 for capH (25). The capH gene was subcloned into pUWL20pw and expressed as described (25).

### Table 1

| Gene Product | LipL | Cpr15 | LpmM | Mur16 | SphE | Mra24 | ORF7 | CapA | Cpr19 |
|--------------|------|-------|------|-------|------|-------|------|------|-------|
| GlyU         | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |
| Cpr15        | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |
| LpmM         | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |
| Mur16        | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |
| SphE         | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |
| Mra24        | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |
| ORF7         | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |
| CapA         | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |
| Cpr19        | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |

\(^{a}\) LipL is encoded within the biosynthetic gene cluster for A-90289, Cpr15 within the cluster for caprazamycin, LpmM within the cluster for liposidomycin, Mur16 within the cluster for muramycin, SphE within the cluster for sphaerimicin, Mra24 within the cluster for muraminomicin, ORF7 within the cluster for A-500359, CapA within the cluster for A-503083, and Cpr19 with the cluster for A-102935.

\(^{b}\) Mra24 appears truncated and inactive, missing 57 amino acids at the C terminus when aligned with LipL.

\(^{c}\) GlyU, 5’-C-glycyluridine-containing nucleoside antibiotics.

\(^{d}\) CarU, uridine-5’-carboxamide-containing nucleoside antibiotics.
distinctive sequence blocks found within the functionally assigned L-Thr:UA transaldolases (35), and these primer sets were utilized to amplify DNA fragments of the expected size from the genomic DNA of the producing strain of 3 (Fig. 5). The pattern and size of the PCR product were similar to those previously reported using template DNA from the producing strains of 1 and 2, and, as expected, no PCR product was obtained using genomic DNA isolated from Streptomyces lividans TK21, whose whole genome has been sequenced and does not harbor the homologous gene (Fig. 5).

The PCR products amplified using the genomic DNA from the producing strain of 3 were subsequently utilized as probes to identify three overlapping cosmids (pNCap01–03) encompassing ~57-kb DNA, which were sequenced and analyzed to reveal 49 complete ORFs (Fig. 6A). One of the ORFs (cpr25) encoded the expected L-Thr:UA transaldolase with moderate (43–50%) to high (70–79%) amino acid sequence identity to transaldolases encoded in gene clusters for GlyU-containing and CarU-containing nucleoside antibiotics, respectively. Due to the novel chemical structure of the polyamide component of 3 and hence uncertainty regarding its mechanism of biosynthesis and incorporation, the cluster was extended by first cloning a BamHI-digested, 10-kb DNA fragment that overlapped with pNCap03. The cloned DNA fragment was subsequently used as a probe to identify a fourth cosmid, pNCap04. Following sequencing and bioinformatic analysis, a total of ~85 kb of contiguous DNA was identified, including 69 complete ORFs and an additional ORF (cpr51) in pNCap04 that has a homolog within the 1 and 2 biosynthetic gene cluster, thus bringing the total shared ORFs for CarU-containing nucleoside antibiotics to 16 (Fig. 6, A and B). The closest homologs and putative functions of the ORFs are listed in supplemental Table 1; ORFs predicted to be involved in the biosynthesis of 3 were annotated as cpr17–cpr57, for capuramycin-related nucleoside.

**Biosynthesis of Capuramycin-type Antibiotics**

The genetic architecture of the clusters for 1 and 2 is virtually identical, excluding a single ORF encoding a truncated 2’-O-carbamoyltransferase within the 1 gene cluster that leads to the sole structural variation between these two groups of capuramycin-type antibiotics. Conversely, the 16 conserved ORFs within the 3 gene cluster have a genetic organization different from that found in 1 and 2 (Fig. 6B). Nonetheless, individual comparison of the gene products revealed a relatively high amino acid sequence identity, which includes the aforementioned transaldolase (Cpr 25; 79% identity with CapH) and non-heme Fe(II)-dependent aKg:UMP dioxygenase (Cpr19; 76% identity with CapA). Additional shared gene products of note include a phosphotransferase Cpr17, probably involved in self-resistance (53% identity with CapP), a carboxymethyltransferase (Cpr27; 57% identity with CapS), and a putative amide bond-forming N-transacylase (Cpr51; 65% identity with CapW). CapS and CapW have been shown to function sequentially to activate the carboxylic acid of 2c in the form of the methyl ester 2d, followed by transacylation to incorporate the l-ACL to generate 2a (Fig. 6C). The uncovering of genes for Cpr27 and Cpr51 suggests that the unusual arylamine-containing polyamide of 3 is incorporated using a comparable mechanism (i.e. carboxylic acid 1c → methyl ester 1d → amide 1a) (Fig. 6C), which was unexpected given the unique chemical nature of the acyl acceptors.

**Phosphorylation by Cpr17 as a Mechanism of Self-resistance**—Heterologous expression of cpr17 in Streptomyces albus yielded a recombinant strain that was, in contrast to the strain harboring the empty vector, resistant to 1a at 500 μg/ml (Fig. 7A). The cpr17 gene was subsequently cloned and expressed in E. coli to yield soluble protein for activity assessment. CapP, which was previously shown to confer resistance by catalyzing regiospecific phosphorylation at the 3’-OH of 2a and 2d (26), was used as a control. Activity of Cpr17 was initially assessed by HPLC using 2a, 2c, and 2d as potential substrates, and new peaks appeared with 2a and 2d. The retention time of the new peak was identical when Cpr17 was substituted with CapP (Fig. 7, B and C), consistent with the formation of 3’-phospho-2a and -2d, respectively. CapP utilized 2c with a very low turnover of ≤0.2 min⁻¹ (26), and accordingly no appreciable activity was detected using Cpr17 or CapP with 2c under the reaction conditions employed here (data not shown). The activity of Cpr17 and CapP was tested with 3, and HPLC analysis revealed a new peak in both cases with an (M – H)⁻ ion at m/z = 843.0 (Fig. 7, D–F), consistent with the molecular formula for 3’-phospho-3 (C₃₁H₄₇N₆O₂₀P; expected (M – H)⁻ ion at m/z = 843.2). Although the limited availability of 3 precluded structural elucidation and product bioactivity analysis at the current time, the results suggest that Cpr17, like CapP, functions as capuramycin-type antibiotic phosphotransferase involved in self-resistance.

After identifying the function of Cpr17, single-substrate kinetic experiments were performed using 2a as a substrate. Using nearly saturating ATP and variable 2a, standard Michaelis-Menten kinetics were observed, yielding a \( K_m = 146 \pm 13 \) μM and \( k_{cat} = 5.8 \pm 0.2 \) min⁻¹ (Fig. 6G), the former constant similar to and the latter 5-fold lower than that reported for CapP (Table 2). Single-substrate kinetic experiments with sat-
urating 2a and variable ATP yielded a $K_m = 1.4 \pm 0.1$ mM and $k_{cat} = 10.1 \pm 0.4$ min$^{-1}$ (Fig. 7G). The high observed $K_m$ for ATP prompted us to test GTP as an alternative phosphate donor, which served as an excellent substrate for Cpr17, yielding a $K_m = 9.2 \pm 1.7$ mM and $k_{cat} = 3.3 \pm 0.1$ min$^{-1}$ (Fig. 7G), a 50-fold improvement in catalytic efficiency, suggesting that GTP is the in vivo substrate for Cpr17. A similar nucleotide preference has been reported recently for certain aminoglycoside phosphotransferases (38–40).

**Development of a Genetic System and Biochemical Interrogation of CarU Biosynthesis**—Unlike the producing strains for 1 and 2, which have not been amenable to genetic manipulation, gene inactivation was possible with the 3-producing strain, thus enabling the first in vivo analysis of the gene clusters for the capuramycin-type antibiotics. Both $cpr19$ and $cpr25$ were individually targeted for inactivation by double crossover homologous recombination, and successful inactivation of each was confirmed by PCR and Southern blot analysis (Fig. 8, A–C). The resulting $\Delta cpr19$ and $\Delta cpr25$ mutant strains were unable to produce 3 (Fig. 8D), consistent with an essential role in the biosynthesis of 3. Gene complementation has yet to be successful in this strain, and thus to exclude possible polar effects, the expression levels of the upstream and downstream genes were analyzed by RT-PCR. Using mRNA extracted from the $\Delta cpr25$ mutant strain at the onset of 3 production, expression of $cpr25$ was not detected, whereas genes flanking $cpr25$ were clearly expressed (Fig. 8E). In contrast, all three genes were detected within the wild-type strain, suggesting that there is no polar
Characterization of the phosphotransferase Cpr17. A, resistance to 1a conferred to S. albus upon heterologous expression of cpr17 using variable amounts of 1a (0 µg/ml (column 1), 100 µg/ml (column 2), and 500 µg/ml (column 3)). The top two rows (pWHM3 (row 1) and pWHM3-ermEp-cpr17 (row 2)) consist of mycelia diluted 1:10 following homogenization of a liquid culture of S. albus, whereas the bottom two rows (pWHM3 (row 3) and pWHM3-ermEp-cpr17 (row 4)) consist of mycelia diluted 1:100 prior to spotting on ISP2 agar plates. B–D, HPLC traces with the indicated substrate incubated without enzyme (i), with Cpr17 (ii), and with CapP (iii). E, mass spectrum of the peak at retention time 7.7 min (D, trace i) corresponding to authentic 3. F, representative mass spectrum of the peak at retention time 7.6 min (D, trace ii or iii) corresponding to 3’-phospho-3. All mass data are reported in negative ion mode. G, plots for single-substrate kinetic analysis of Cpr17.

TABLE 2
Michaelis-Menten kinetic parameters

| Enzyme | Substrate | $K_M$ (µM) | $k_{cat}$ (min$^{-1}$) | $k_{cat}/K_M$ | Relative efficiency |
|--------|-----------|------------|----------------------|--------------|---------------------|
| CapP   | 2a        | 170 ± 40   | 34 ± 4               | 2.0 x 10$^1$ |                     |
|        | 2a        | 146 ± 13   | 5.8 ± 0.2            | 4.0 x 10$^2$ |                     |
|        | ATP       | (1.4 ± 0.1) x 10$^2$ | 10.1 ± 0.4          | 7.2 x 10$^5$ | 1                   |
|        | GTP       | 9.2 ± 1.7  | 3.3 ± 0.1            | 3.6 x 10$^1$ | 50                  |
| LipL$^b$ | UMP      | 14 ± 3     | 76 ± 8               |              |                     |
|        | αKG       | 7.5 ± 2.4  | 92 ± 9               |              |                     |
| Cpr19  | UMP       | 25 ± 4     | 95 ± 10              |              |                     |
|        | αKG       | 6.2 ± 1.0  | 78 ± 7               |              | 1                   |
|        | Adipic Acid | (1.1 ± 0.2) x 10$^2$ | 17 ± 6             | 0.012        |                     |
|        | Pyruvic Acid | (8.2 ± 1.0) x 10$^2$ | 9.0 ± 2.0          | 0.0009       |                     |
| LipK$^c$ | UA       | (29 ± 10) x 10$^3$ | 40 ± 4              |              |                     |
|        | l-Thr     | 25 ± 8     | 32 ± 2               |              |                     |
| CapH   | UA        | (19 ± 6) x 10$^3$ | 44 ± 5              |              |                     |
|        | l-Thr     | 31 ± 6     | 45 ± 3               |              |                     |

$^a$ Data taken from Ref. 26.
$^b$ Data taken from Ref. 27.
$^c$ Data taken from Ref. 28.
effect due to \( cpr25 \) inactivation. Ambiguous results were obtained with expression analysis within the \( H9004 \) mutant strain, and hence we turned our attention to \textit{in vitro} characterization for direct evidence of the involvement of \( cpr19 \) in the biosynthesis of 3.

The gene encoding \( cpr19 \) was heterologously expressed in \textit{E. coli} to obtain pure, recombinant protein (Fig. 9A). Based on prior characterization of LipL, Cpr19 was expected to convert UMP to UA in a reaction that is dependent upon Fe(II), O$_2$, and \( KG \). HPLC analysis of reactions with all of these components revealed a new peak that co-eluted with authentic UA, and as expected, the formation of the product was absolutely dependent on the inclusion of O$_2$ and \( KG \) (Fig. 9B). Inductively coupled plasma MS revealed a fraction (4%) of Cpr19 co-purified with Fe(II), thus explaining the residual activity in the absence of additional Fe(II). Optimal activity for Cpr19 was observed between 100 and 1000 \( \mu \)M FeCl$_2$ (Fig. 9C). Similar to other enzymes of this dioxygenase superfamily, including \textit{E. coli} taurine dioxygenase (41), ascorbic acid enhanced the activity, reaching an optimum between 1 and 5 mM (Fig. 9D). No tested divalent or transition metals were able to substitute for Fe(II), and similarly to LipL, Zn$^{2+}$ inhibited the reaction (Fig. 9E).

Finally, single-substrate kinetic analysis yielded a \( K_m \) of 25 ± 4 \( \mu \)M and \( k_{cat} \) of 95 ± 10 min$^{-1}$ with respect to UMP and \( KG \) of 6.2 ± 1.0 \( \mu \)M and \( k_{cat} \) of 78 ± 7 min$^{-1}$ with respect to \( KG \) (Fig. 9F) (Table 2). The overall biochemical and kinetic properties of Cpr19 were comparable with those previously reported for LipL excluding two of note (27). First, although LipL is able to catalyze uncoupled oxidative decarboxylation of \( KG \), the rate of
formation of succinate in the absence of UMP was moderately greater for Cpr19 (k\text{cat} = 1.6 \text{ min}^{-1} \text{ compared with } k\text{cat} = 0.4 \text{ min}^{-1} ), corresponding to a relative turnover of 2.5% to the UMP-containing reaction (Fig. 9G). Second, in contrast to the high specificity of LipL, Cpr19 catalyzed UA formation with alternative α-keto acids pyruvate and α-ketoadipate. Single substrate kinetic analysis yielded constants of Km = 8.2 \times 10^{-2} \pm 90 \text{ mM} and k\text{cat} = 9.0 \pm 2.0 \text{ min}^{-1} with respect to pyruvate and K\text{m} = 1.1 \times 10^{-2} \pm 16 \text{ mM} and k\text{cat} = 17 \pm 6 \text{ min}^{-1} with respect to α-ketoadipate (Table 2), a relative efficiency of 8.7 \times 10^{-4} and 1.2 \times 10^{-2}, respectively, compared to the preferred substrate αKG. Similarly to LipL, however, Cpr19 was inactive with α-ketobutyrate, α-ketovalerate, and oxaloacetate.

After demonstrating that UA is a biosynthetic intermediate of CarU-containing nucleosides, the probable L-Thr:UA transaldolase, Cpr25, was targeted for characterization; however, soluble protein could not be obtained using multiple hosts and a variety of expression conditions. Instead, Cpr25 was substituted with CapH that is encoded within the 2 biosynthetic gene cluster (5). Similarly to LipK (28), soluble CapH was only produced in S. lividans TK64 (Fig. 10A), and UV-visible spectroscopic analysis revealed trace amounts of PLP co-purified with CapH (Fig. 10B). Enzyme assays monitored by HPLC revealed identical traces to LipK (Fig. 10C), and the product peak was characterized by spectroscopic analysis to demonstrate that CapH is an L-Thr:UA transaldolase. The biochemical properties of CapH were similar to those of LipK, and kinetic analysis following acetaldehyde production (28) yielded kinetic constants of Km = 19 \pm 6 \text{ mM} and k\text{cat} = 44 \pm 5 \text{ min}^{-1} with respect to UA and Km = 25 \pm 8 \mu \text{M} and k\text{cat} = 32 \pm 2 \text{ min}^{-1} with respect to L-Thr (Fig. 10D), values comparable with those for LipK (Table 2).

CprS1 Is Essential for 3 Biosynthesis—To determine whether cprS1, encoding a putative N-transacylase with a proposed functional role comparable with that of CapW, was involved in the biosynthesis of 3, the gene was inactivated by double-cross-over homologous recombination, which was confirmed by PCR and Southern blot analysis (data not shown). Utilizing mRNA
extracted from the Δcpr51 mutant strain at the onset of 3 production, expression of cpr51 was not detected, whereas genes flanking cpr51 were clearly expressed (Fig. 8F). As expected, the resulting Δcpr51 mutant strain was unable to produce 3 (Fig. 8D), consistent with an essential role in 3 biosynthesis.

Discussion

We have performed isotope enrichment experiments, identified the 3 gene cluster, characterized a phosphotransferase involved in resistance, and characterized two biosynthetic enzymes to support a shared biosynthetic paradigm for the CarU-containing nucleoside antibiotics 1-3. Interestingly, the biosynthesis of CarU occurs through the morphing of the primary metabolites UMP and L-Thr and not uridine and phosphoenolpyruvate as previously suggested. The pathway to CarU proceeds through the unusual -hydroxy-L-amino acid GlyU as an intermediate following sequential catalysis by a non-heme Fe(II), KG:UMP dioxygenase, and an L-Thr:UA transaldolase. The function of the dioxygenase Cpr19/LipL, the generation of UA from the mononucleotide, is novel for enzymes of the non-heme Fe(II), KG-dependent dioxygenases and thus mechanistically intriguing (42–45). Based on the mechanism for taurine dioxygenase, which catalyzes the transformation of taurine to aminoacetaldehyde and sulfite and is perhaps the best characterized enzyme of the superfamily (46–48), Cpr19/LipL probably catalyze insertion of one atom of O2 into KG and the other at C-5 of UMP, leading to a geminal hydroxyl-phosphoester intermediate. This labile intermediate is proposed to undergo phosphate elimination to yield UA.

In addition to being mechanistically interesting, the identification of the new activity of Cpr19/LipL has biosynthetic implications for multiple chemical entities. Not only is this dioxygenase activity required to initiate the biosynthesis of the high carbon sugar nucleosides CarU and GlyU, but it is also required for biosynthesis of the aminoribosyl moiety found in the GlyU-containing nucleoside antibiotics (Fig. 2) (49). During the assembly of this unusual sugar, an L-Met:UA aminotransferase uses the Cpr19/LipL product to form 5-amino-5-deoxyuridine. The discovery of the dioxygenase and aminotransferase activities involved in aminoribose biosynthesis has subsequently led to the uncovering of the biosynthetic gene cluster for jawsamycin, whose structure contains a 5-amino-5-deoxyuridine (50). Finally, a survey of the whole genomes of diverse actinomycetes reveals a high prevalence of genes encoding proteins of the dioxygenase superfamily within other putative natural product biosynthetic operons, suggesting the potential for discovery of comparable dioxygenase activities.

Following the reaction by Cpr19, the transaldolase Cpr25 generates a new β-hydroxy-L-amino acid that is ultimately converted to the carboxamide functionality of CarU. Carboxamides in primary metabolites are made by one of four enzymatic processes: an ATP-dependent enzymatic process from a carboxylic acid using an NH3/Gln as the amine source, a hydrolytic process, lyase chemistry, or (in the unusual case of L-Arg catabolism in certain microorganisms) an oxidative decarboxylation catalyzed by a flavin-dependent L-Arg 2-monooxygenase (51). Carboxamides are found in some secondary metabolites, such as phenazines (52), oxytetracycline (53), AB-400 (54, 55), and sarubicin A (56), and genetic or biochemical evidence suggests that they are produced from the carboxylic acid by an ATP-dependent process. A gene encoding a protein with similarity to those catalyzing ATP-dependent carboxamide formation or any alternative mechanism is not apparent within the 1–3 gene
clusters. The revelation of a $^{13}$C-$^{15}$N heteronuclear spin coupling in the $^{13}$C NMR spectrum upon feeding labeled L-Thr suggests that the C–N bond remains intact during CarU biosynthesis (57), providing additional support for a different, potentially novel mechanism of carboxamide formation.

Because the conversion of GlyU to CarU is likely to involve decarboxylation, hydroxylation, and/or oxidation, it is possible that the PLP-dependent enzyme Cpr23, a second non-heme Fe(II)-dependent dioxygenase Cpr18, and/or the putative oxidoreductase Cpr20 are involved.

The formation of the unsaturated α-L-mannopyranuronate that is attached to the CarU core is not obvious from bioinformatics analysis, excluding the attachment of the sugar by the glycosyltransferase, Cpr24. GDP-α-L-mannuronic acid, expected to be an intermediate based on prior isotopic enrichment studies, is a building block of alginate, a major anionic polysaccharide found in some strains of bacteria (58). During alginate biosynthesis, GDP-α-L-mannuronic acid is formed from GDP-α-L-mannose by a GDP-mannose 6-dehydrogenase, yet the two remaining shared proteins encoded in the 1–3 gene clusters with similarity to sugar biosynthetic enzymes are Cpr21, a putative NDP-hexose 2,3-dehydratase, and Cpr22, a putative NDP-α-L-xylulose 5-epimerase/dehydratase. Although Cpr21 and Cpr22 are very different in sequence and domain architecture from the carbohydrate-containing component of the polyamide is predicted to be assembled from chorismic acid and malonyl-CoA by minimally eight ORFs, cpr12 and cpr32–cpr38 (Fig. 11). The pathway initially proceeds by the conversion of chorismic acid to p-aminobenzoic acid, a transformation that occurs during the de novo synthesis of essential folates in plants, fungi, and bacteria and requires three enzyme activities: an L-Gln amidotransferase, a 4-amino-4-deoxychorismate (ADC) synthase, and ADC lyase (59). In E. coli, these three activities, which have been biochemically confirmed, are found on proteins encoded by three distinct genes (pabA, pabB, and pabC, respectively) located at remote locations within the chromosome. A survey of 100 Streptomycetes genomes revealed that pabB and pabC are located in a conserved, probable three-gene operon (the third gene encodes a protein with sequence similarity to acetyltransferases of the GNAT family), and pabA is located elsewhere within the chromosome. Twenty of these 100 genomes contain at least one additional gene encoding a PabB but, contrastingly, as part of a bidomain protein containing an N-terminal PabA domain. Although not yet biochemically confirmed, a protein with identical bidomain architecture has been uncovered for the biosynthesis of amicetin (60), pristinamycin (61), aureothin (62), candicidin/FR-008 (63), chloramphenicol (64–66), the recently identified albidicidin (67), and now 3. Thus, cpr38 encoding this bidomain protein is proposed to catalyze a two-step reaction involving amidohydrolsy of L-Gln with ammonia channeled and incorporated into chorismic acid to generate ADC. Cpr12, encoded by a gene that was initially believed to be outside the gene cluster, has similarity to proteins annotated as ADC lyase or aminotransferase class IV, of which ADC lyase belongs (59), suggesting that this protein catalyzes the elimination of pyruvate to form PABA.

The remaining steps require C–C bond formation between the carboxylic acid of PABA and a C-2 extender unit. This is putatively initiated by Cpr37-catalyzed activation of PABA as the acyl-adenylate with loading to the free-standing carrier protein Cpr36 to form the thioester-linked PABA. Homologous genes are found within the actinomycin gene cluster (68), wherein biosynthesis is proposed to be initiated by activation.
and loading of a 4-methyl-3-hydroxyanthranilic acid starter unit to a free-standing carrier protein. Adjacent to cpr36 are two genes encoding proteins with similarity to a ketosynthase and chain length factor (Cpr34 and Cpr35, respectively) that work in concert to catalyze decarboxylative condensation between a malonyl-S-acetyl carrier protein (ACP) and recipient thioester during aromatic polyketide biosynthesis catalyzed by type II polyketide synthases (69). In several instances, type II polyketide synthase systems recruit fatty acid biosynthetic machinery during polyketide assembly, and it is probably the same scenario here, wherein an ACP and malonyl-CoA:ACP transacetylase from fatty acid biosynthesis generate the co-substrate for decarboxylative condensation by the Cpr34/35 heterodimer. A comparable mechanism of condensation has previously been proposed during hygromycin B biosynthesis upon identification of its gene cluster (70).

Following formation of the β-ketothioester, a putative 3-oxoacyl-ACP reductase, Cpr33, catalyzes reduction to the β-hydroxythioester intermediate, and the luciferase-like monooxygenase, Cpr32, catalyzes α-hydroxylation to give 3-(4-aminoaryl)-2,3-dihydroxypropanoic acid. These last two enzymatic steps to generate the vicinal diol of 3-(4-aminoaryl)-2,3-dihydroxypropanoic acid are of particular interest because the stereochemistry in 3 has not been established. Sequence analysis of Cpr33 suggests the closest similarity to 3-oxoacyl-ACP reductase (FabG) from bacterial type II fatty acid synthases, which catalyze hydride addition to the si face to generate the R configuration (71). It is thus reasonable to assume the same relative orientation of acyl-ACP binding and reduction, in this case hydride addition to the re face to generate the S configuration (the PABA starter unit changes priorities of the carbonyl carbon; Fig. 11). Stereoselective reduction of other members of the short chain dehydrogenase family, to which 3-oxoacyl-ACP reductases belong, has been the subject of several studies, culminating in sequence signatures to predict the stereochemical outcome, which is consistent with that proposed for Cpr33 (72, 73). Introduction of the C-2 hydroxyl is probably catalyzed by Cpr32 that has sequence similarity to proteins annotated as FMN-dependent monooxygenases of the bacterial luciferase family, which includes a variety of FMN- or coenzyme F420-dependent oxidoreductases that act on structurally diverse substrates. It is not possible to predict the stereochemistry of α-hydroxylation based solely on this information, but assuming that catalysis does not involve epimerization at C-3, C-2-hydroxylation leads to a change in substituent priority at C-3 so that the ultimate product is proposed to have the 3R configuration.

The remaining steps in polyamide biosynthesis involve formation of two amidic bonds, and as expected, there are putative proteins encoded within ORFs cpr47–cpr57 that are associated with amide bond-forming events, including an adenylating-like domain protein with predicted specificity to a hydrophilic amino acid (Cpr54), two carrier proteins (Cpr48 and Cpr55), a condensation domain protein (Cpr47), and three transglutaminase-like proteins (Cpr49, Cpr50, and Cpr57). Based on the functional assignment of an unrelated transglutaminase-like protein as a condensation catalyst that uses two substrate thioesters during andimide biosynthesis (74), the putative transglutaminases, all of which have similarity to uncharacterized MitI in the mitomycin C gene cluster (75), potentially catalyze amide bond formation. It is likely that this component is derived from a Gly and l-Asp (or l-Asn) unit, although a detailed proposal for polyamide biosynthesis at this stage would be highly speculative without isotopic enrichment or genetic evidence. Finally, components of Cpr41–Cpr46 have similarity to a pyruvate dehydrogenase complex or acetolactate synthase, and their role in 3 biosynthesis, if any, is unclear, but possibly they are utilized for pyruvate and by-product recycling that is generated by Cpr12.

The proposed final step in 3 biosynthesis is the coupling of the arylamine-containing polyamide to the capuramycin core via amide bond formation. Of the aforementioned natural products containing PABA, only the aromatic amine of PABA from amecitin is modified to an amide as also observed in 3. In the case of amecitin, gene inactivation of avIN encoding a putative malonyl-CoA:ACP transacylase revealed it as the catalyst for amide bond formation (60). Contrastingly, the formation of amide bond for 2a is orchestrated by an unconventional mechanism featuring sequential catalysis by CapS and CapW (25). Despite the stark structural contrast of the amine acceptor (aliphatic amine of aminocaprolactam versus the arylamine of the polyamide), genes encoding putative CapS and CapW homologues (cpr27 and cpr51, respectively) were identified in the 3 gene cluster, and gene inactivation demonstrated that cpr51 is essential for 3 biosynthesis. Although no pathway intermediates could be identified, Cpr51 is proposed to catalyze the coupling reaction, and future studies will be aimed at defining the function of Cpr51 and its role and catalytic properties relative to CapW.

In summary, we have established that UMP and, unexpectedly, 1-Thr serve as the metabolic precursors of the CarU nucleoside found in the capuramycin-type antibiotics. Using this information, the biosynthetic gene cluster for 3 was identified to enable a thorough, comparative genomic approach to propose a biosynthetic pathway for the capuramycin-type antibiotics. A mechanism of self-resistance to 3 was also defined, and the phosphotransferase Cpr17 conferring this resistance was shown to prefer GTP as the phosphate donor. Finally, the first genetic system for the capuramycin-type antibiotic producers was developed, a critical advancement for interrogating the biosynthesis of the unusual chemical components of this family of antibiotics.

References
1. Muramatsu, Y., Muramatsu, A., Ohnuki, T., Ishii, M. M., Kizuka, M., Enokita, R., Tsutsu, S., Arai, M., Ogawa, Y., Suzuki, T., Takatsu, T., and Inukai, M. (2003) Studies on novel bacterial translocase I inhibitors, A-500359s: I. Taxonomy, fermentation, isolation, physico-chemical properties and structure elucidation of A-500359 A, C, D, and G. J. Antibiot. 56, 243–252.
2. Muramatsu, Y., Ishii, M. M., and Inukai, M. (2003) Studies on novel bacterial translocase I inhibitors, A-500359s: II. Biological activities of A-500359 A, C, D and G. J. Antibiot. 56, 253–258.
3. Muramatsu, Y., Miyakoshi, S., Ogawa, Y., Ohnuki, T., Ishii, M. M., Arai, M., Takatsu, T., and Inukai, M. (2003) Studies on novel bacterial translocase I inhibitors, A-500359s: III. Deaminocaprolactam derivatives of capuramycin: A-500359 E, F, H, M-1 and M-2. J. Antibiot. 56, 259–267.
4. Ohnuki, T., Muramatsu, Y., Miyakoshi, S., Takatsu, T., and Inukai, M. (2003) Studies on novel bacterial translocase I inhibitors, A-500359s: IV.
Biosynthesis of Capuramycin-type Antibiotics

1. Shah, N. S., Wright, A., Bai, G.-H., Barrera, L., Boulahbal, F., Martín, M., Reddy, V. M., Einck, L., Andries, K., and Nacy, C. A. (2010) Anti-tuberculosis drug therapy in mice of different inbred strains. *Infect. Genet. Evol.* 54, 357–364.

2. Reddy, V. M., Bogatcheva, E., Einck, L., and Nacy, C. A. (2011) Nanoemulsion formulation enhances intracellular activity of capuramycin analogs. *Antimicrob. Agents Chemother.* 55, 3138–3139.

3. Reddy, V. M., Bogatcheva, E., Einck, L., and Nacy, C. A. (2011) Synthesis and antimycobacterial activity of capuramycin analogs. Part 2: Acylated derivatives of capuramycin-related compounds. *Bioorg. Med. Chem. Lett.* 13, 2823–2826.

4. Reddy, V. M., Einck, L., and Nacy, C. A. (2008) In vitro antimycobacterial activities of capuramycin analogues. *Antimicrob. Agents Chemother.* 52, 719–721.

5. Nikonenko, B. V., Reddy, V. M., Protopopova, M., Bogatcheva, E., Einck, L., and Nacy, C. A. (2009) Activity of SQ641, a capuramycin analog, in a murine model of tuberculosis. *Antimicrob. Agents Chemother.* 53, 3138–3139.

6. Reddy, V. M., Bogatcheva, E., Einck, L., and Nacy, C. A. (2011) Nanoemulsion formulation enhances intracellular activity of capuramycin analogues against *Mycobacterium tuberculosis*. *Drug Delivery Lett.* 10. 2174/2210304X11101020150.

7. Reddy, V. M., Dubuisson, T., Einck, L., Wallis, R. S., Jakubiec, W., Ladukto, L., Campbell, S., and Nacy, C. A. (2012) SQ109 and PNU-100480 interact to kill *Mycobacterium tuberculosis* in vitro. *J. Antimicrob. Chemother.* 67, 1163–1166.

8. Bogatcheva, E., Hanrahan, C., Nikonenko, B., de los Santos, G., Reddy, V., Chen, P., Barbosa, F., Einck, L., Nacy, C., and Protopopova, M. (2011) Identification of SQ609 as a lead compound from a library of dipiperidines. *Bioorg. Med. Chem. Lett.* 21, 5535–5537.

9. Nikonenko, B. V., Einck, L., and Nacy, C. A. (2010) Anti-tuberculosis drug therapy in mice of different inbred strains. *Infect. Genet. Evol.* 10, 1151–1154.

10. Reddy, V. M., Einck, L., Andries, K., and Nacy, C. A. (2010) In vitro interactions between new antituberculosis drug candidates SQ109 and TMCC207. *Antimicrob. Agents Chemother.* 54, 2840–2846.

11. Shah, N. S., Wright, A., Bai, G.-H., Barrera, L., Boualahal, F., Martin-Casabona, N., Drobniewski, F., Gilpin, C., Havelková, M., Lepe, R., Lumb, R., Metchock, B., Portaels, F., Rodrigues, M. F., Rüsch-Gerdes, S., Van Deun, A., Vincent, V., Laserson, K., Wells, C., and Cegelski, J. P. (2007) Worldwide emergence of extensively drug-resistant tuberculosis. *Emerg. Infect. Dis.* 13, 380–387.

12. Wong, E. B., Cohen, K. A., and Bishai, W. R. (2013) Rising to the challenge: new therapies for tuberculosis. *Trends Microbiol.* 21, 149–501.

13. Jackson, M., McNeil, M. R., and Brennan, P. J. (2013) Progress in targeting cell envelope biogenesis in *Mycobacterium tuberculosis*. *Future Microbiol.* 8, 855–875.

14. Yamaguchi, H., Sato, S., Yoshida, S., Takada, K., Itoh, M., Seto, H., and Otake, N. (1986) Capuramycin, a new nucleoside antibiotic: taxonomy, fermentation, isolation and characterization. *J. Antibiot.* 39, 1047–1053.

15. Seto, H., Otake, N., Sato, S., Yamaguchi, H., Takada, K., Itoh, M., Lu, H. S. M., and Clardy, J. (1988) The structure of a new nucleoside antibiotic, capuramycin. *Tetrahedron Lett.* 10.1016/S0040-4020(88)80555-4.

16. Reddy, V. M., Dubuisson, T., Einck, L., Wallis, R. S., Jakubiec, W., Ladukto, L., Campbell, S., and Nacy, C. A. (2012) SQ109 and PNU-100480 interact to kill *Mycobacterium tuberculosis* in vitro. *J. Antimicrob. Chemother.* 67, 1163–1166.

17. Reddy, V. M., Bogatcheva, E., Einck, L., and Nacy, C. A. (2011) Nanoemulsion formulation enhances intracellular activity of capuramycin analogues against *Mycobacterium tuberculosis*. *Drug Delivery Lett.* 10. 2174/2210304X11101020150.

18. Reddy, V. M., Dubuisson, T., Einck, L., Wallis, R. S., Jakubiec, W., Ladukto, L., Campbell, S., and Nacy, C. A. (2012) SQ109 and PNU-100480 interact to kill *Mycobacterium tuberculosis* in vitro. *J. Antimicrob. Chemother.* 67, 1163–1166.

19. Bogatcheva, E., Hanrahan, C., Nikonenko, B., de los Santos, G., Reddy, V., Chen, P., Barbosa, F., Einck, L., Nacy, C., and Protopopova, M. (2011) Identification of SQ609 as a lead compound from a library of dipiperidines. *Bioorg. Med. Chem. Lett.* 21, 5535–5537.

20. Nikonenko, B. V., Einck, L., and Nacy, C. A. (2010) Anti-tuberculosis drug therapy in mice of different inbred strains. *Infect. Genet. Evol.* 10, 1151–1154.
Biosynthesis of Capuramycin-type Antibiotics

40. Toth, M., Chow, J. W., Mobashery, S., and Vakulenko, S. B. (2009) Source of phosphate in the enzymic reaction as a point of distinction among aminoglycoside 2'-phosphotransferases. J. Biol. Chem. 284, 6690–6696

41. Eichhorn, E., Van Der Ploeg, I. R., Kertesz, M. A., and Leisinger, T. (1997) Characterization of α-ketoglutarate-dependent taurine dioxygenase from Escherichia coli. J. Biol. Chem. 272, 23031–23036

42. Bollinger, J. M., Jr., and Krebs, C. (2006) Stalling intermediates in oxygen activation by iron enzymes: motivation and method. J. Inorg. Biochem. 100, 586 – 605

43. Straganz, G. D., and Nidetzky, B. (2006) Variations of the 2-His-1-carboxylate theme in mononuclear non-heme FeII oxygenases. ChemBioChem 7, 1536 – 1548

44. Mantri, M., Zhang, Z., McDonough, M. A., and Schofield, C. J. (2012) Autocatalysed oxidative modifications to 2-oxoglutarate dependent oxygenases. FEBS J. 279, 1563–1575

45. Hausinger, R. P. (2004) Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes. Crit. Rev. Biochem. Mol. Biol. 39, 21 – 68

46. Price, J. C., Barr, E. W., Tirupati, B., Bollinger, J. M., Jr., and Krebs, C. (2006) Stalking intermediates in oxygen activation by iron enzymes: motivation and method. J. Inorg. Biochem. 100, 586 – 605

47. Toomey, R. E., and Wakil, S. J. (1966) Studies on the mechanism of fatty acid biosynthesis: 4. Details of the tryptophan metabolism. J. Biol. Chem. 231, 1809 – 1819

48. Ye, Q. Z., Liu, J., and Walsh, C. T. (1990) p-Aminobenzoate synthesis in Escherichia coli: purification and characterization of PabB as aminoadoxorhodamine synthase and enzyme X as aminoadoxorhodamine lyase. Proc. Natl. Acad. Sci. U.S.A. 87, 9391 – 9395

49. Zhang, G., Zhang, H., Li, S., Xiao, J., Zhang, G., Zhu, Y., Niu, S., Ju, I., and Zhang, C. (2012) Characterization of the amicetin biosynthesis gene cluster from Streptomyces vinaceae NRRL B3107. J. Antibiot. 64, 191–202

50. He, J., and Hertweck, C. (2003) Iteration as programmed event during polyketide assembly: molecular analysis of the aureothin biosynthesis gene cluster. Chem. Biol. 10, 1225–1232

51. Grzyska, P. K., Ryle, M. J., Monterosso, G. R., Liu, J., Ballou, D. P., and Hausinger, R. P. (2005) Kinetic dissection of the catalytic mechanism of taurine-α-ketoglutarate dioxygenase (Taud) from Escherichia coli. Biochemistry 44, 8139 – 8147

52. Yanai, K., Sumida, N., Okakura, K., Moriya, T., Watanabe, M., and Marakami, T. (2004) Para-position derivatives of fungal anthelmintic cyclodepsipeptides engineered with Streptomyces venezuelae antibiotic biosynthetic genes. Nat. Biotechnol. 22, 848 – 855

53. Cocianchich, S., Pesic, A., Petras, D., Uhlmann, S., Kretz, J., Schubert, V., Vieweg, L., Duplan, S., Marguetrazet, M., Noël, J., Pieretti, I., Hügell, M., Kemper, S., Mainz, A., Rott, P., Royer, M., and Süssmuth, R. D. (2015) The gyrase inhibitor alibixin consists of α-p-aminobenzoic acids and cya-noalanine. Nat. Chem. Biol. 11, 195–197

54. Keller, U., Lang, M., Crnovcic, I., Pfennig, F., and Schauwecker, F. (2010) The actinomycin biosynthetic gene cluster of Streptomyces chrysomallus: a genetic hall of mirrors for synthesis of a molecule with mirror symmetry. J. Bacteriol. 192, 2583–2595

55. Toomey, R. E., and Wakil, S. J. (1966) Studies on the mechanism of fatty acid biosynthesis: XV. Preparation and general properties of β-ketoacyl acyl carrier protein reductase from Escherichia coli. Biochim. Biophys. Acta 116, 189 – 197

56. Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D. V., Hutchinson, C. R., and McDaniel, R. (2003) A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. Biochemistry 42, 72–79

57. Fortin, P. D., Walsh, C. T., and Magarvey, N. A. (2007) A transglutaminase homologue as a condensation catalyst in antibiotic assembly lines. Nature 448, 654 – 657

58. Mao, Y., Varoglu, M., and Sherman, D. H. (1999) Molecular characterization and analysis of the biosynthetic gene cluster for the antitumor antibiotic mitomycin C from Streptomyces lavendulae NRRL B3156. Chem. Biol. 6, 251–263