JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

The Amipurimycin and Miharamycin Biosynthetic Gene Clusters: Unraveling the Origins of 2-Aminopurinyl Peptidyl Nucleoside Antibiotics

Anthony J Romo, Taro Shiraishi, Hideo Ikeuchi, Geng-Min Lin, Yujie Geng, Yu-Hsuan Lee, Priscilla H. Liem, Tianlu Ma, Yasushi Ogasawara, Kazuo Shin-ya, Makoto Nishiyama, Tomohisa Kuzuyama, and Hung-wen Liu

J. Am. Chem. Soc., **Just Accepted Manuscript** • DOI: 10.1021/jacs.9b03021 • Publication Date (Web): 31 May 2019

The Amipurimycin and Miharamycin Biosynthetic Gene Clusters: Unraveling the Origins of 2-Aminopurinyl Peptidyl Nucleoside Antibiotics

Anthony J. Romo,[a,1,‖] Taro Shiraishi,[b,‖] Hideo Ikeuchi,[c] Geng-Min Lin,[d,2] Yujie Geng,[a] Yu-Hsuan Lee,[d] Priscilla H. Liem,[e] Tianlu Ma,[e,3] Yasushi Ogasawara,[a,4] Kazuo Shin-ya,[c,f] Makoto Nishiyama,[c,g] Tomohisa Kuzuyama,[b,g,*] and Hung-wen Liu,[a,d,e,*]

[a]Division of Chemical Biology and Medicinal Chemistry, College of Pharmacy, The University of Texas at Austin, Austin, TX 78712, USA; [b]Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan; [c]Biotechnology Research Center, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan; [d]Department of Chemistry; and [e]Department of Biochemistry, The University of Texas at Austin, Austin, TX 78712, USA; [f]National Institute of Advanced Industrial Science and Technology, 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan; [g]Collaborative Research Institute for Innovative Microbiology, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan.

KEYWORDS: Peptidyl nucleoside antibiotics, natural product biosynthesis, Streptomyces, nucleoside, ATP-grasp ligase

ABSTRACT: Peptidyl nucleoside antibiotics (PNAs) are a diverse class of natural products with promising biomedical activities. These compounds have tripartite structures composed of a core saccharide, a nucleobase, and one or more amino acids. In particular, amipurimycin and the miharamycins are novel 2-aminopurinyl PNAs with complex nine-carbon core saccharides and include the unusual amino acids (–)-cispentacin and N^5-hydroxyarginine, respectively. Despite their interesting structures and properties, these PNAs have heretofore eluded biochemical scrutiny. Herein is reported the discovery and initial characterization of the miharamycin gene cluster in *Streptomyces miharaensis* (*mhr*) and the amipurimycin gene cluster (*amc*) in *Streptomyces novoguineensis* and *Streptomyces* sp. SN-C1. The gene clusters were identified using a comparative genomics approach, and heterologous expression of the *amc* cluster as well as gene interruption experiments in the *mhr* cluster support their role in the biosynthesis of amipurimycin and the miharamycins, respectively. The *mhr* and *amc* biosynthetic gene clusters characterized encode enzymes typical of polyketide biosynthesis instead of enzymes commonly associated with PNA biosynthesis, which along with labeled precursor feeding studies, implies that the core saccharides found in the miharamycins and amipurimycin are partially assembled as polyketides rather than derived solely from carbohydrates. Furthermore, *in vitro* analysis of Mhr20 and Amc18 established their roles as ATP-grasp ligases involved in the attachment of the pendant amino acids found in these PNAs, and Mhr24 was found to be an unusual hydroxylase involved in the biosynthesis of N^5-hydroxyarginine. Finally, analysis of the *amc* cluster and feeding studies also led to the proposal of a biosynthetic pathway for (–)-cispentacin.

Introduction

Peptidyl nucleoside antibiotics (PNAs) are composed of a nucleobase, a core saccharide, and one or more appended amino acids. They comprise an underexploited, structurally heterogeneous family of natural products that has begun to garner more attention as a potential new source of antimicrobials.[1] However, the lack of structural conservation between PNAs and other more widely studied classes of natural products presents a formidable challenge to their biosynthetic study. Efforts to determine PNA biosynthetic pathways are further complicated by their use of extensively modified nucleobases, higher-carbon sugars, and non-proteinogenic amino acids in the assembly.[2,3] Several PNA biosynthetic gene clusters have been sequenced to date, and subsequent characterization has revealed that the core saccharides of these PNAs originate from only one of four different precursors: a nucleoside diphosphate (NDP)-activated deoxy sugar, uridine diphosphate (UDP)-activated glucuronic acid, an intact nucleoside from primary metabolism, or octosyl acid.[4,5] Despite the potential utility of such an inherent classification scheme, it is likely an oversimplification of the biosynthetic space occupied by this class of natural products given the paucity of sequenced PNA biosynthetic gene clusters compared to the number and diversity of known PNA structures.[2,3] Thus, investigation of how PNAs are constructed holds promise for both the development of useful therapeutics and the enrichment of our understanding of natural product biosynthesis.

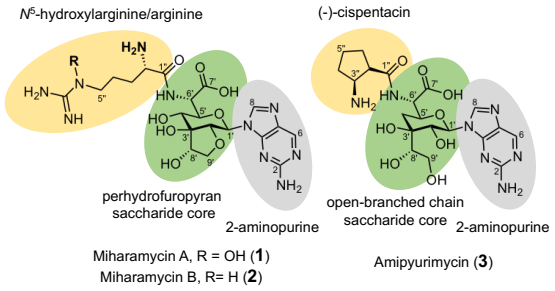


Figure 1. Comparison of the structural features of miharamycins (**1** and **2**) and amipurimycin (**3**).

The miharamycins (**1, 2**) and amipurimycin (**3**) are PNAs produced by the actinomycetes *Streptomyces miharaensis* and *Streptomyces novoguineensis*, respectively, exhibiting diverse antifungal, antibacterial, and antiviral activities, the mechanisms of which are currently undetermined.[6-9] The miharamycins and amipurimycin are each structurally composed of the rare nucleobase 2-aminopurine, a structurally complex core saccharide, and a single amino acid each.[10,11] Miharamycin A and amipurimycin are decorated by the non-proteinogenic amino acids N^5-hydroxyarginine and (1*R*,2*S*)-2-aminocyclopentane-1-carboxylic acid (also known as (–)-cispentacin), respectively, which themselves display mild antibiotic properties.[10-15] Besides their pendant amino acids, the miharamycins are distinguished from amipurimycin by their bicyclic perhydrofuropyran saccharide cores (Figure 1). The biosynthetic gene clusters and pathways for amipurimycin and the miharamycins have not been reported to date. Interestingly, while the total synthesis of amipurimycin was accomplished only recently,[16] some 40 years after the initial discovery of this PNA, a successful total synthesis of a miharamycin has yet to be reported.[17,18]

Herein, we identify and characterize the biosynthetic gene clusters of the miharamycins and amipurimycin from *S. miharaensis* and *S. novoguineensis*, respectively. The presence of genes encoding characteristic polyketide biosynthetic enzymes along with labeled precursor feeding studies imply that the core saccharides of these PNAs are likely partially polyketide-derived. This finding suggests a new paradigm for the biological production of unusual sugars, and allows us to propose a biosynthetic pathway for amipurimycin. Moreover, insight into the biosynthesis and ATP-grasp ligase-mediated attachment of the pendant amino acids decorating amipurimycin and the miharamycins is also provided.

Results and Discussion

Comparative genomics approach to gene cluster discovery. The *S. novoguineensis* and *S. miharaensis* genomes were sequenced and then annotated utilizing Rapid Annotation Using Subsystem Technology (RAST).[19-21] Scanning for shared gene clusters using antiSMASH[22] resulted in the identification of a single unique gene cluster shared between both strains at an approximately 30 kbp locus in each genome. The shared putative clusters contained 30 open reading frames (ORFs) in the 32.5 kbp conserved region found in *S. novoguineensis* (the *amc* cluster) and 27 ORFs in the 29.2 kbp conserved region found in *S. miharaenesis* (the *mhr* cluster). A duplicate, identical *amc* cluster was also found in *Streptomyces* sp. SN-C1 (see Supporting Information). The gene clusters are shown schematically in Figure S-3.3, and ORFs from both putative gene clusters are listed in Tables S-3.3a and S-3.3b along with their predicted products. Gene cluster boundaries were approximated as described in the Supporting Information (Section S-3.4).

Gene cluster validation. A bacterial artificial chromosome (BAC) containing the *amc* cluster was isolated (see Section S-4.1) in order to further characterize this gene cluster. The BAC clone, pKU503ampr1, was transformed into the non-producing host *Streptomyces albus* G153 via protoplast transformation, converting it to an amipurimycin-producing strain (see Figure 2A). While the same approach was not pursued with the *mhr* gene cluster, double crossover-mediated deletion of *mhr5*, *mhr7*, or *mhr17* in *S. miharaensis* completely ablated production of the miharamycins (see Figure 2B).

Additional evidence for the assignment of the *mhr* and *amc* clusters was obtained from an analysis of their putative regulatory elements. The ORFs *amc0* and *amc16* are annotated as encoding proteins that contain a SARP-responsive bacterial transcriptional activator domain, which indicated these proteins might increase expression of the *amc* biosynthetic gene cluster.[23] With this in mind, *amc0* and *amc16* were individually cloned and introduced into *S.* sp. SN-C1. Fermentation of the transformants showed significantly increased amipurimycin production, with up to 30 times the basal production level observed when *amc0* or *amc16* were individually overexpressed (Figures 2C and 2D). Altogether, the heterologous expression of the *amc* cluster and subsequent observation of the potent effect of the *amc0* and *amc16* gene products on amipurimycin overproduction, as well as the inactivating gene deletions performed in *S. miharaensis* strongly implicate a direct role of the *amc* and *mhr* clusters in the biosynthesis of amipurimycin and the miharamycins, respectively.

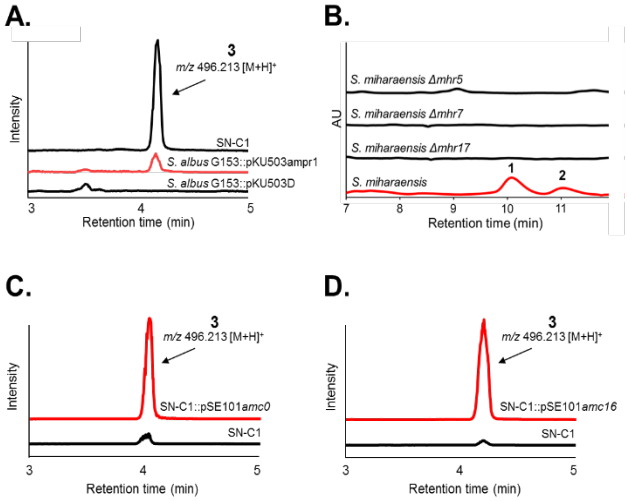


Figure 2. (A) Extracted ion current chromatogram comparing culture extracts from *S.* sp. SN-C1 (positive control) and *S. albus* G153 transformed with empty pKU503D vector (negative control) or pKU503ampr1, demonstrating heterologous expression of the *amc* cluster in *S. albus* G153; (B) HPLC chromatograms (λ= 305 nm) demonstrating loss of production of miharamycins in *S. miharaensis* as a result of selected gene deletions; (C) over-production of amipurimycin in *S.* sp. SN-C1 when *amc0* is supplemented; (D) over-production of amipurimycin in *S.* sp. SN-C1 when *amc16* is supplemented.

Proposed biosynthetic pathway based on feeding studies and gene cluster analyses. The biosyntheses of all of the PNAs discovered to date involve the incorporation of an intact sugar precursor molecule into the final structure, e.g., glucuronic acid for the cytosyl PNAs, frequently via the nucleoside diphosphate activated form of these sugars. Towards providing the first insight into the biosynthesis of amipurimycin and the miharamycins, we chose to focus on amipurimycin and administered uniformly ^{13}C-labeled glucose ([U-$^{13}C_6$]glucose, **4**) to the culture broth of *S.* sp.SN-C1/pSE101*amc0*. The labeled amipurimycin (**5**) was isolated and analyzed by ^{13}C NMR. The ^{13}C NMR spectrum revealed efficient ^{13}C incorporation of the core saccharide (Fig. S-5.1). Remarkably, carbon–carbon coupling analyses using 2D-INADEQUATE (incredible natural-abundance double-quantum transfer experiment) spectroscopy did not support incorporation of an intact sugar molecule (Fig. S-5.2–S-5.4). Instead, the observed coupling patterns

showed the core saccharide to be composed of three distinct fragments, C1'–C2', C3'–C4'–C5'–C6', and C8'–C9', with C7' labeled but uncoupled to any other carbon atom (see **5** in Fig. 3). Unlike the previously characterized PNA biosynthetic pathways, this unexpected pattern does not support the intermediacy of nucleoside diphosphate-activated sugars and suggests the biosyntheses of amipurimycin and the miharamycins deviate substantially from other PNA biosynthetic pathways.

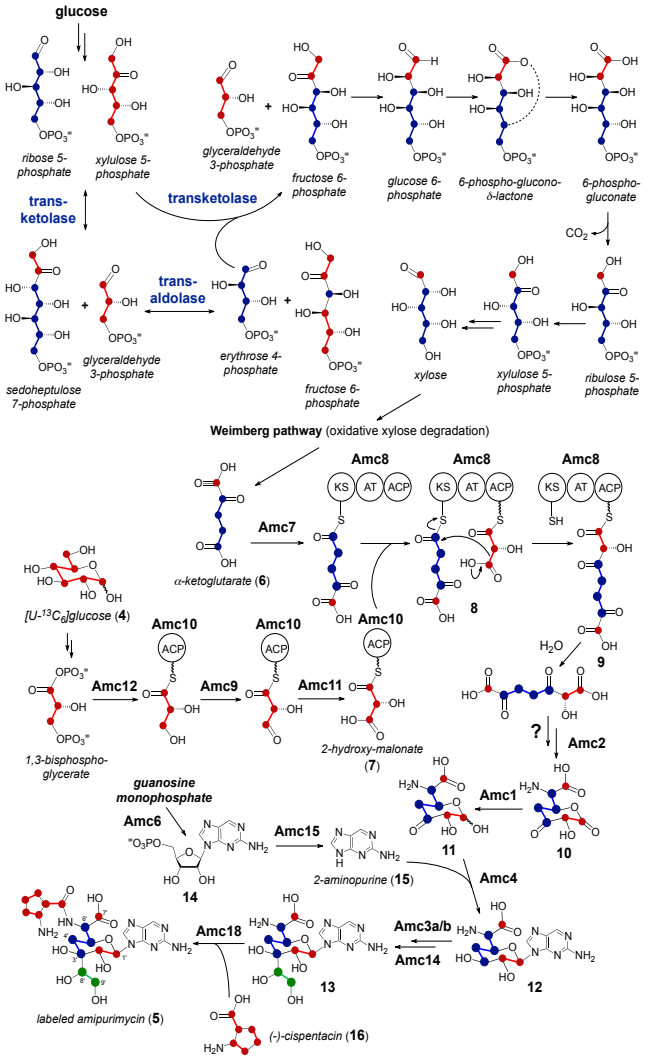


Figure 3. Proposed biosynthetic pathway for amipurimycin based on gene cluster analysis and results of feeding experiments. The labeled α-ketoglutarate is predicted to arise from [U-$^{13}C_6$]-labeled glucose via pentose phosphate interconversions and the Weimberg pathway, while the two-carbon fragment (C8'-C9', green) is predicted to originate from a ketose-5-phosphate, as described in the text.

Upon closer analysis of the contents of the *amc* gene cluster, we found it to encode a mono-modular minimal type I polyketide synthase (PKS) composed of one ketosynthase (KS), one acyltransferase (AT) and one thiolation (T) domain (Amc8) together with a separate ORF encoding a putative, unusual adenylation/thiolation-type loading module (Amc7). In addition, the *amc* cluster possesses a cassette (*amc9–12*) encoding four conserved enzymes (Amc9–12) implicated in the biosynthesis and ACP-thiolation of the atypical PKS extender unit hydroxymalonate (**7**) in polyketide biosynthetic pathways.[24a]

Given the simple composition of the Amc8 enzyme, this PKS enzyme would be expected to be capable of only one round of chain extension. In consideration of the hydroxymalonate cassette found in the *amc* cluster, as well as the predicted preference of the AT domain in Amc8 for hydroxymalonate via the ClusterCAD server,[24b] it can be reasoned that Amc8 uses a two-carbon fragment from ACP-linked hydroxymalonate (derived from **7**) in its extension reaction (**8** → **9**). We, thus, predict the hydroxymalonate-derived unit results in the labeled C1'–C2' fragment observed in the structure of amipurimycin. However, this only accounts for two of the nine carbons composing the core saccharide of amipurimycin. As the "acceptor" of the two-carbon extender unit (i.e., the PKS starter unit) is likely a linear five-carbon fragment (such as **8** shown in Figure 3), addition of two more carbons is needed to complete the assembly of the branched nine-carbon core. Such two-carbon extension could be a post-PKS event or may occur while the carbon chain is still attached to Amc8.

The C8'–C9' fragment which was labeled in one piece in the feeding experiments performed in *S.* sp. SN-C1/pSE101*amc0* fed with [U-$^{13}C_6$]glucose (**4**) constitutes the two-carbon branch in the structure of amipurimycin. Recent investigations have demonstrated PKSs can effect branching through the use of specialized PKS modules or a conserved set of "branching" enzymes.[25] However, Amc8 and the *amc* cluster lack these features. Instead, the *amc* cluster encodes a two-component transketolase, Amc3a/3b, which, by analogy to the biosynthesis of branched saccharides,[26] would be a good candidate to install the C8'–C9' two-carbon fragment (**12** → **13**), perhaps originating from ribulose- or xylulose-5-phosphate (see Figure 3).

In the course of analyzing the *amc* and *mhr* clusters, the discovery of several more homologous gene clusters allowed us to predict conserved characteristics of ostensibly 2-aminopurinyl (i.e., amipurimycinoid) PNAs (see section S-7). The newly identified amipurimycinoid clusters do not all contain a homolog of Amc3a/3b, but those that do possess a homologous transketolase also encode an oxidoreductase homologous to Amc14, which could serve to catalyze reduction of the C8' carbonyl of the Amc3a/3b-installed glycoaldehyde, thereby affording the C8' alcohol (as shown in **13**) found in the structure of amipurimycin.

If branching at C3' can be seen as a "tailoring" reaction, then the true Amc8 starter unit must be composed of the remaining five carbons, i.e., C3', C4', C5', C6' and C7'. As described above, C3' through C6' were coupled, while C7' was labeled but never coupled — a "4+1" labeling pattern. One compelling explanation for this distinctively labeled C_5 starter unit is that it originates from the transformation of glucose to ribulose to α-ketoglutarate via pentose phosphate interconversions and the Weimberg pathway (Figure 3).[27,28] Thus, while there are few reports exploring the linkage between primary (e.g., from the Weimberg pathway) and secondary metabolism in Streptomyces, it is plausible that the starter unit could be α-ketoglutarate (**6**, Figure 3). Overall, Amc8 is believed to catalyze the addition of two-carbon atoms derived from hydroxymalonate (**7**) to α-ketoglutarate to yield a seven-carbon chain (**8** → **9**), whose cyclization would generate the pyran portion of the amipurimycin core saccharide (**10**). Though its origin is not obvious, an alternative pathway starting from 2-keto-3-hydroxy-glutarate could also be possible (see Figure S-5.5). Then, Amc3a/3b catalyz-

es the addition of two more carbon atoms to complete the nine-carbon skeleton of amipurimycin (**12** → **13**).

Condensation with 2-aminopurine (**15**), transamination by Amc2, and attachment of (–)-cispentacin (**16**) are three further PKS tailoring steps (Fig. 3), but the sequence of these events is not apparent. It remains to be seen whether the same labeling patterns are observed in the miharamycins. Nevertheless, given the strong conservation of content between the *amc* and *mhr* clusters, the proposed biosynthetic pathway is expected to be operative in the assembly of the miharamycins as well. Further work to explore these biosynthetic pathways is in progress.

Following construction, off-loading, and cyclization of the nascent heptosyl core (**10**), it may then be converted to the nucleoside moiety by introduction of the 2-aminopurine nucleobase (**15**). The first step is likely reduction of the lactone (**10**) at the "anomeric" position catalyzed by the conserved *myo*-inositol 2-dehydrogenase/oxidoreductase Amc1. The *amc* cluster also contains a cistronic "glycosylation cassette" (e.g., *amc4–6*) that possess a gene annotated as encoding an *O*-kinase, *amc5*. The co-localization of these genes suggests that coupling of **11** with **15** may be initiated by C1 phosphorylation followed by substitution with 2-aminopurine. The latter reaction may be catalyzed by the product of *amc4*, a predicted cytosylglucuronic acid (CGA) synthase which mediates the condensation between UDP-glucuronic acid and cytosine[29] in the biosynthesis of UDP-glucuronic acid-derived PNAs, including blasticidin S, arginomycin, and gougerotin.[30]

The origin of the 2-aminopurine nucleobase itself (**15**) is unclear; however, it may involve the activities of the annotated flavoprotein Amc6 as well as Amc15, a homolog of the cytokinin maturating phosphoribohydrolase LONELY GUY-1 (LOG-1) from *Arabidopsis thaliana*.[31] Although mechanistically distinct from nucleoside 2′-deoxyribosyltransferases such as BlsM from the blasticidin S biosynthetic pathway,[32] the LOG-1 homolog Amc15 could conceivably accomplish the same end, namely, the release of a nucleobase from a donor nucleotide (**14** → **15**). Taken together, the coexistence of a phosphoribohydrolase and CGA synthase homolog in the *amc* cluster implies amipurimycin is derived from glycosylation of a nucleobase (e.g., **15**) with a sugar (e.g., **11**) and not via modification of a preformed nucleoside or nucleotide as is observed in the biosynthesis of some other PNAs such as puromycin.[33] Additional work is required to gain further insight into the features of the biosyntheses of amipurimycin and the miharamycins.

Characterization of Mhr20 and Amc18 as ATP-grasp ligases involved in the biosynthesis of the miharamycins and amipurimycin. By analogy to other known PNA biosynthetic pathways, the ATP-grasp ligases encoded by *amc18* and *mhr20* were identified as likely candidates for catalyzing amide bond formation in the biosynthesis of amipurimycin and the miharamycins, respectively. Therefore, their corresponding ORFs were cloned and overexpressed in *E. coli* as His_6-tagged constructs prior to purification by Ni(II) affinity chromatography. In order to prepare a substrate to assay the activity of Mhr20, commercially-available peptidases were screened for their ability to remove the pendant amino acids from the miharamycins. This led to the discovery of a protease from *Streptomyces griseus* that cleaves the L-arginine moiety of miharamycin B (**2**) to release the core nucleoside miharamycinine (**17**), a putative substrate for Mhr20. Likewise, single crossover mediated deletion of amc18 in *S.* sp. SN-C1 provided the mutant *S.* sp. SN-C1Δamc18 that produced amipurimycinine (**19**, i.e., the nucleoside scaffold of amipurimycin) as a putative substrate for Amc18. Structural and spectroscopic characterization of both miharamycinine and amipurimycinine is provided in the Supporting Information (Section S-6.2 and Figures S-6.2.3–S.6.2.8).

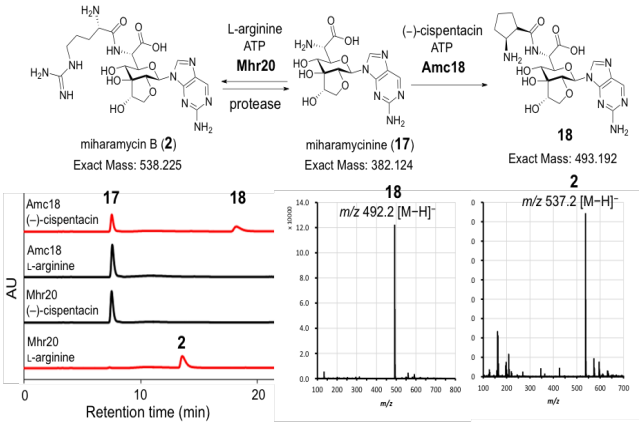


Figure 4. Overview of *in vitro* assays performed using miharamycinine (**17**) as the amino acid acceptor; HPLC chromatogram (λ= 305 nm) of assays, as well as mass spectra of the products of the assays are shown.

Miharamycinine (**17**) was consumed in the presence of Mhr20 only when co-incubated with both L-arginine and ATP (Figure 4). The identity of the product was confirmed to be miharamycin B (**2**) by co-elution with the standard compound and high-resolution mass spectrometry. Similarly, amipurimycin (**3**) was produced when Amc18 was incubated with (–)-cispentacin, amipurimycinine (**19**), and ATP (Figure 5). The results suggest that Mhr20 and Amc18 are indeed the ATP-dependent ligases responsible for the attachment of L-arginine and (–)-cispentacin to their respective core nucleosides and that amino acid ligation likely occurs late in the biosynthetic pathways of these PNAs. Moreover, Amc18 could catalyze the ligation of (–)-cispentacin to miharamycinine (Figure 4), and Mhr20 could catalyze ligation of L-arginine to amipurimycinine (Figure 5); however, no cross-reactivity was observed with respect to their amino acid substrates. These observations are consistent with previous reports that ATP-grasp ligases are highly specific with respect to their amino acid substrates while being more flexible regarding their acceptor substrates.[34,35]

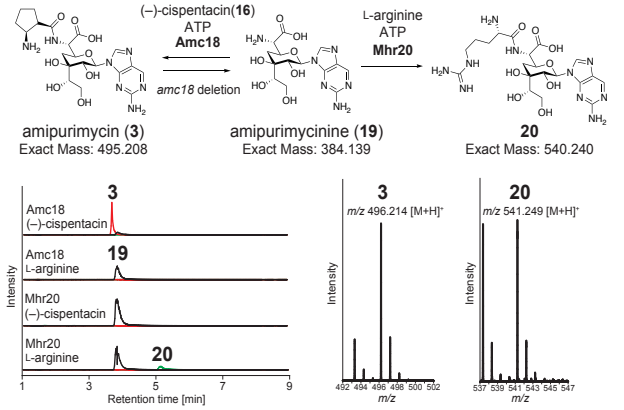


Figure 5. Overview of *in vitro* assays performed using amipurimycinine (**19**) as the amino acid acceptor; extracted ion current chroma-

tograms of *in vitro* assay substrate (black trace for **19**) and products (red and green traces for **3** and **20**, respectively) as well as mass spectra of the products of the assays are shown.

Mhr24 is a hydroxylase involved in the biosynthesis of N^5-hydroxyarginine. While the hydroxylated guanidine of the N^5-hydroxyarginine found in the structure of miharamycin A (**1**) resembles a key intermediate in the reaction of nitric oxide synthase, no homolog of nitric oxide synthase could be identified in the *mhr* cluster. Recently, N^5-hydroxyarginine has also been found as a component of argolaphos A (**21**), a phosphonate natural product produced by *Streptomyces monomycini* NRRL B-24309,[36] and the *mhr* cluster contains Mhr24, whose product shows homology (24.7% similarity, 17.7% identity) to a predicted "YqcI/YcgG" protein (NCBI accession WP_078624154) encoded in the argolaphos biosynthetic gene cluster. Although N^5-hydroxyarginine and similarly hydroxylated guanidines have precedent in natural products,[14,37-39] only DcsA, a YqcI/YcgG homolog and putative N^{ω}-hydroxy-L-arginine hydroxylase from *Streptomyces lavendulae*, has been implicated in the oxidation of arginine in natural product biosynthesis (Figure 6A).[40] Thus, Mhr24 emerged as a reasonable candidate for the requisite *N*-hydroxylase activity in the biosynthesis of N^5-hydroxyarginine.

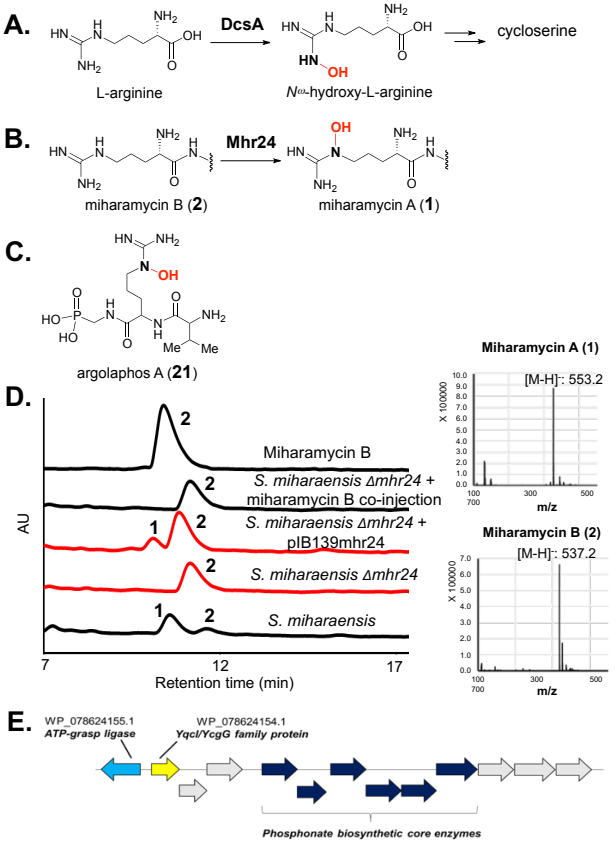


Figure 6. (A) DcsA is the proposed hydroxylase involved in arginine N^{ω}-oxidation in cycloserine biosynthesis; (B) proposed function of Mhr24 in the oxidation of arginine N^5 of miharamycin B (**2**) to yield miharamycin A (**1**); (C) argolaphos A (**21**) contains N^5-hydroxyarginine; (D) HPLC chromatograms (λ= 305 nm) of *in vivo* assays of Mhr24 demonstrate its role in the biosynthesis of miharamycin A with each peak verified by MS analysis; (E) the putative argolaphos gene cluster encodes a YqcI/YcgG protein.

To determine the function of Mhr24, its encoding gene was removed from the *S. miharaensis* genome through double-crossover mediated in-frame deletion. The resulting strain, *S. miharaensisΔmhr24*, was found to produce only miharamycin B (**2**). To ensure this observation was not due to polar effects resulting from the deletion, *mhr24* was cloned into vector pIB139[41] downstream from an *ermE* promoter sequence[42] to yield the complementation vector pIB139mhr24. When *S. miharaensisΔmhr24* was transformed with pIB139mhr24, albeit not to the extent of the wild-type strain, production of miharamycin A (**1**) was restored (Figure 6D). The relatively low sequence homology between Mhr24 and its homolog from the argolaphos gene cluster may reflect the structural dissimilarity between argolaphos (**21**) and miharamycin A (**1**). Considering, then, that hydroxylation of these compounds' respective pendant arginine moieties likely occurs late in their biosynthetic pathways, we predict miharamycin B (**2**) is generated first and is the substrate for Mhr24-mediated oxidation to yield miharamycin A (**1**).

Cispentacin biosynthesis mirrors biosynthesis of coronafacic acid. The introduction of (–)-cispentacin (**16**) to the saccharide core is a key tailoring step in amipurimycin biosynthesis. The construction of (–)-cispentacin is presently obscure despite its importance as a lead compound in the development of icofungipen, a potent antifungal compound that inhibits eukaryotic leucyl-tRNA synthetase.[43-45] While the cyclopentyl carbocycle of (–)-cispentacin is not well-represented in nature, it bears a striking similarity to 2-carboxy-2-cyclopentenone (**25**, CPC), an intermediate in the biosynthesis of coronafacic acid.[46]

In the *amc* cluster, flanking the core set of genes shared with the *mhr* cluster are *amcA–H*, whose products include the complete complement of enzymes comprising the conserved biosynthetic pathway of CPC (*cfa* cluster). Moreover, the *amcA–H* subcluster is augmented by the presence of an enoyl reductase (AmcA) and an acetylornithine-like aminotransferase (AmcC) that are not typically found in CPC gene clusters. It is therefore hypothesized that the biosynthesis of (–)-cispentacin (**16**) parallels the coronafacic acid biosynthetic pathway and commences with the formation of ACP-linked CPC (**25**) mediated by AmcH, B, F, G, and E, as previously reported for the analogous sequence catalyzed by Cfa5, 1, 3, 4, and 2,[46,47] respectively, and is concluded by the action of AmcA, AmcC, and the thioesterase AmcD (see Figure 7).

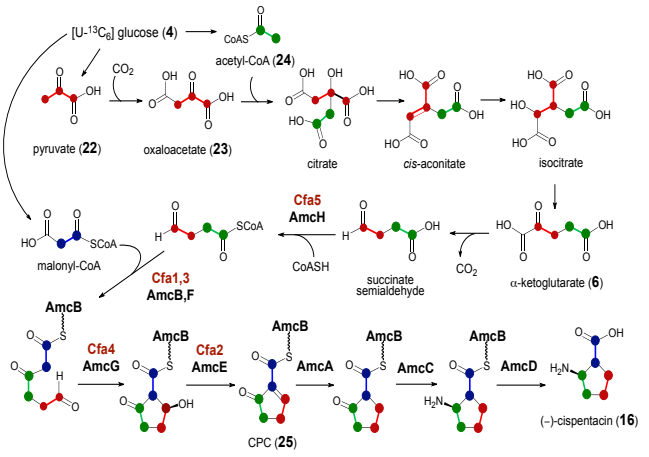


Figure 7. Proposed CPC-based (–)-cispentacin biosynthetic pathway. Colored carbon atoms reflect the observed ^{13}C-labeling

pattern. The corresponding enzymes from the coronafacic acid pathway (*cfa*) are shown in red.

In support of this pathway, ^{13}C NMR analysis of amipurimycin labeled with [U-$^{13}C_6$]glucose (**4**) showed efficient ^{13}C incorporation into the (–)-cispentacin moiety, with C1"–C2", C3"–C4" and C5"–C6" coupling observed. Interestingly, CPC biosynthesis begins with α-ketoglutarate (**6** in Figure 7), yet the labeling pattern seen in (–)-cispentacin does not match the portion of the amipurimycin core saccharide in which the α-ketoglutarate precursor (**6** in Figure 3) is predicted to yield a "4+1" labeling pattern (i.e., C3'-C4'-C5'-C6', and C7'). However, α-ketoglutarate can also be biosynthesized from anaplerotically generated oxaloacetate (**23**) using pyruvate (**22**) derived from [U-$^{13}C_6$]glucose. When this labeled oxaloacetate is combined with acetate (**24**) derived from the fed [U-$^{13}C_6$]glucose, the Krebs cycle could generate "2+2" labeled α-ketoglutarate (**6**), leading to the pattern observed in (–)-cispentacin (see Figure 7). Thus, it is plausible that the α-ketoglutarate used in the core saccharide and (–)-cispentacin are generated in different phases of growth of the producing bacteria. Therefore, the observed (–)-cispentacin labeling pattern in this work is consistent with previous feeding studies of the biosynthesis of coronafacic acid.[46,48]

Conclusion

Heterologous production of amipurimycin and gene-deletion studies support the assignment of the *mhr* and *amc* clusters as responsible for the biosynthesis of miharamycins and amipurimycin. These clusters encode enzymes for biosynthetic pathways that differ significantly from those of other, previously described PNAs and implicate a polyketide origin for the core saccharide. Indeed, we carried out labeled precursor feeding studies which implicate Amc8 in the assembly of the amipurimycin core saccharide. If correct, our proposed biosynthetic pathway would represent a new paradigm in carbohydrate biosynthesis. We also performed *in vitro* assays characterizing Mhr20 and Amc18 as functional ATP-grasp ligases with specificity for their predicted amino acid substrates. Finally, analysis of the flanking region of the *amc* cluster for amipurimycin has resulted in the identification of a set of genes likely to be responsible for (–)-cispentacin biosynthesis, which had previously remained elusive. These findings lay the groundwork for further analysis of each encoded enzyme necessary for definitive reconstruction of the biosynthetic pathways for these PNAs and possibly others,[49] and their eventual modification and repurposing for the development of useful antimicrobial agents.

ASSOCIATED CONTENT

Supporting Information. Genomic sequencing and analyses, isolation and characterization of compounds and intermediates, NMR methodologies and spectra, and enzyme cloning and purification. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*To whom correspondence may be addressed. E-mail: h.w.liu@mail.utexas.edu and utkuz@mail.ecc.u-tokyo.ac.jp

Present Addresses

[1]University of Texas M.D. Anderson Cancer Center, Investigational Pharmacy Services, Houston, TX 77030, USA.

[2]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[3]Department of Biochemistry and Molecular Biology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD 21205, USA

[4]Graduate School of Engineering, Hokkaido University, Sapporo, Hokkaido 060-8628, Japan

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

||These authors contributed equally.

Funding Sources

This work is supported in part by grants from the National Institutes of Health (GM035906 and GM040541) and the Welch Foundation (F-1511) to H.-w.L.; JSPS KAKENHI (16H06453) and the Japan Agency for Medical Research and Development (AMED) to T.K.

ACKNOWLEDGMENT

We thank Masato Tani and Makoto Ohyama of Meiji Seika Pharma Co., Ltd., for their kind gift of authentic samples of miharamycins and Dr. Mark Ruszczycky for his critical review of the manuscript.

REFERENCES

(1) Serpi, M.; Ferrari, V.; Pertusati, F. Nucleoside Derived Antibiotics to Fight Microbial Drug Resistance: New Utilities for an Established Class of Drugs? *J Med Chem.* **2016**, *59*, 10343.

(2) Isono, K. Nucleoside antibiotics: structure, biological activity, and biosynthesis. *J Antibiot (Tokyo).* **1988**, *41*, 1711.

(3) Isono, K. Current progress on nucleoside antibiotics. *Pharmacol Ther.* **1991**, *52*, 269.

(4) Rachakonda, S.; Cartee, L. Challenges in antimicrobial drug discovery and the potential of nucleoside antibiotics. *Curr Med Chem.* **2004**, *11*, 775.

(5) Niu, G.; Tan, H. Nucleoside antibiotics: biosynthesis, regulation, and biotechnology. *Trends Microbiol.* **2015**, *23*, 110.

(6) Iwasa, T.; Kishi, T.; Matsuura, K.; Wakae, O. Streptomyces novoguineensis sp. Nov., an amipurimycin producer, and antimicrobial activity of amipurimycin. *J Antibiot (Tokyo).* **1977**, *30*, 1.

(7) Noguchi, T.; Komoto, K.; Yasuda, Y.; Hashimoto, S.; Niida, T. New antibiotics, miharamycins A and B. III. Control activity of miharamycin against plant disease. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 11.

(8) Shomura, T.; Hamamoto, K.; Ohashi, T.; Amano, S.; Yoshida, J.; Moriyama, C.; Niida, T. New antibiotics, miharamycins A and B. II. Some biological characteristics of miharamycin. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 5.

(9) Tsuruoka, T.; Yumoto, H.; Ezaki, N.; Niida, T. New antibiotics, miharamycins A and B. I. Isolation and characterization of miharamycins A and B. *Meiji Seika Kenkyu Nenpo.* **1967**, *9*, 1.

(10) Seto, H.; Koyama, M.; Ogino, H.; Tsuruoka, T.; Inouye, S.; Otake, N. The structures of novel nucleoside antibiotics, miharamycin A and miharamycin B. *Tetrahedron.* **1983**, *24*, 1805.

(11) Goto, T.; Toya, Y.; Kondo, T. Structure of amipurimycin, a new nucleoside antibiotic produced by Streptomyces novoguineensis. *Nucleic Acids Symp Ser.* **1980**, s73.

(12) Konishi, M.; Nishio, M.; Saitoh, K.; Miyaki, T.; Oki, T.; Kawaguchi, H. Cispentacin, a new antifungal antibiotic. I. Production, isolation, physico-chemical properties and structure. *J Antibiot (Tokyo).* **1989**, *42*, 1749.

(13) Oki, T.; Hirano, M.; Tomatsu, K.; Numata, K.; Kamei, H. Cispentacin, a new antifungal antibiotic. II. In vitro and in vivo antifungal activities. *J Antibiot (Tokyo)*. **1989**, *42*, 1756.

(14) Fischer, B.; Keller-Schierlein, W.; Kneifel, H.; Konig, W. A.; Loeffler, W.; Muller, A.; Muntwyler, R.; Zahner, H. Metabolic products of microorganisms. 118. Delta-N-hydroxy-L-arginine, an amino acid antagonist from Nannizzia gypsea. *Archiv fur Mikrobiologie.* **1973**, *93*, 203.

(15) Perlman, D.; Vlietinck, A. J.; Matthews, H. W.; Lo, F. F. Microbial production of vitamin B12 antimetabolites. I. N5-hydroxy-L-arginine from Bacillus cereus 439. *J. Antibiot.* **1974**, *27*, 826.

(16) Wang, S.; Sun, J.; Zhang, Q.; Cao, X.; Zhao, Y.; Tang, G.; Yu, B. Amipurimycin: total synthesis of the proposed structures and diastereoisomers. *Angew Chem Int Ed Engl.* **2018**, *57*, 2884.

(17) Markad, P. R.; Kumbhar, N.; Dhavale, D. D. Synthesis of the C8'-epimeric thymine pyranosyl amino acid core of amipurimycin. *Beilstein J Org Chem.* **2016**, *12*, 1765.

(18) Cachatra, V.; Almeida, A.; Sardinha, J.; Lucas, S. D.; Gomes, A.; Vaz, P. D.; Florencio, M. H.; Nunes, R.; Vila-Vicosa, D.; Calhorda, M. J.; Rauter, A. P. Wittig reaction: domino olefination and stereoselectivity DFT Study. Synthesis of the miharamycins' bicyclic sugar moiety. *Org Lett.* **2015**, *17*, 5622.

(19) Aziz, R. K.; Bartels, D.; Best, A. A.; DeJongh, M.; Disz, T.; Edwards, R. A.; Formsma, K.; Gerdes, S.; Glass, E. M.; Kubal, M.; Meyer, F.; Olsen, G. J.; Olson, R.; Osterman, A. L.; Overbeek, R. A.; McNeil, L. K.; Paarmann, D.; Paczian, T.; Parrello, B.; Pusch, G. D.; Reich, C.; Stevens, R.; Vassieva, O.; Vonstein, V.; Wilke, A.; Zagnitko, O. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics.* **2008**, *9*, 75.

(20) Overbeek, R.; Olson, R.; Pusch, G. D.; Olsen, G. J.; Davis, J. J.; Disz, T.; Edwards, R. A.; Gerdes, S.; Parrello, B.; Shukla, M.; Vonstein, V.; Wattam, A. R.; Xia, F.; Stevens, R. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **2014**, *42*, D206.

(21) Brettin, T.; Davis, J. J.; Disz, T.; Edwards, R. A.; Gerdes, S.; Olsen, G. J.; Olson, R.; Overbeek, R.; Parrello, B.; Pusch, G. D.; Shukla, M.; Thomason, J. A., 3rd; Stevens, R.; Vonstein, V.; Wattam, A. R.; Xia, F. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep.* **2015**, *5*, 8365.

(22) Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de Los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **2017**, *45*, W36.

(23) Cuthbertson, L.; Nodwell, J. R. The TetR family of regulators. *Microbiol Mol Biol Rev.* **2013**, *77*, 440.

(24) (a) Ray, L.; Moore, B. S. Recent advances in the biosynthesis of unusual polyketide synthase substrates. *Nat Prod Rep.* **2016**, *33*, 150. (b) Eng, C. H.; Backman, T. W. H.; Bailey, C. B.; Magnan, C.; Martin, H. G.; Katz, L.; Baldi, P.; Keasling, J. D. ClusterCAD: a computational platform for type I modular polyketide synthase design. *Nucleic Acids Res.* **2018**, *46*, D509.

(25) Hertweck, C. Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem Sci.* **2015**, *40*, 189.

(26) Lin, C. I.; McCarty, R. M.; Liu, H. W. The biosynthesis of nitrogen-, sulfur-, and high-carbon chain-containing sugars. *Chem Soc Rev.* **2013**, *42*, 4377.

(27) Kim, D.; Woo, H. M. Deciphering bacterial xylose metabolism and metabolic engineering of industrial microorganisms for use as efficient microbial cell factories. *Appl Microbiol Biotechnol.* **2018**, *102*, 9471.

(28) Valdehuesa, K. N. G.; Ramos, K. R. M.; Nisola, G. M.; Banares, A. B.; Cabulong, R. B.; Lee, W. K.; Liu, H.; Chung, W. J. Everyone loves an underdog: metabolic engineering of the xylose oxidative pathway in recombinant microorganisms. *Appl Microbiol Biotechnol.* **2018**, *102*, 7703.

(29) Gould, S. J.; Guo, J. Cytosylglucuronic acid synthase (cytosine: UDP-glucuronosyltransferase) from Streptomyces griseochromogenes, the first prokaryotic UDP-glucuronosyltransferase. *J Bacteriol.* **1994**, *176*, 1282.

(30) Feng, J.; Wu, J.; Gao, J.; Xia, Z.; Deng, Z.; He, X. Biosynthesis of the beta-methylarginine residue of peptidyl nucleoside arginomycin in Streptomyces arginensis NRRL 15941. *Appl Environ Microbiol.* **2014**, *80*, 5021.

(31) Mortier, V.; Wasson, A.; Jaworek, P.; De Keyser, A.; Decroos, M.; Holsters, M.; Tarkowski, P.; Mathesius, U.; Goormachtig, S. Role of LONELY GUY genes in indeterminate nodulation on Medicago truncatula. *New Phytol.* **2014**, *202*, 582.

(32) Cone, M. C.; Yin, X.; Grochowski, L. L.; Parker, M. R.; Zabriskie, T. M. The blasticidin S biosynthesis gene cluster from Streptomyces griseochromogenes: sequence analysis, organization, and initial characterization. *Chembiochem.* **2003**, *4*, 821.

(33) Saugar, I.; Sanz, E.; Rubio, M. A.; Espinosa, J. C.; Jimenez, A. Identification of a set of genes involved in the biosynthesis of the aminonucleoside moiety of antibiotic A201A from Streptomyces capreolus. *Eur J Biochem.* **2002**, *269*, 5527.

(34) Kino, K.; Nakazawa, Y.; Yagasaki, M. Dipeptide synthesis by L-amino acid ligase from Ralstonia solanacearum. *Biochem Biophys Res Commun.* **2008**, *371*, 536.

(35) Suzuki, M.; Takahashi, Y.; Noguchi, A.; Arai, T.; Yagasaki, M.; Kino, K.; Saito, J. The structure of L-amino-acid ligase from Bacillus licheniformis. *Acta Crystallogr D Biol Crystallogr.* **2012**, *68*, 1535.

(36) Ju, K. S.; Gao, J.; Doroghazi, J. R.; Wang, K. K.; Thibodeaux, C. J.; Li, S.; Metzger, E.; Fudala, J.; Su, J.; Zhang, J. K.; Lee, J.; Cioni, J. P.; Evans, B. S.; Hirota, R.; Labeda, D. P.; van der Donk, W. A.; Metcalf, W. W. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. *Proc Natl Acad Sci U S A.* **2015**, *112*, 12175.

(37) Dobashi, K.; Naganawa, H.; Takahashi, Y.; Takita, T.; Takeuchi, T. Novel antifungal antibiotics octacosamicins A and B. II. The structure elucidation using various NMR spectroscopic methods. *J Antibiot (Tokyo).* **1988**, *41*, 1533.

(38) Makarieva, T. N.; Ogurtsova, E. K.; Denisenko, V. A.; Dmitrenok, P. S.; Tabakmakher, K. M.; Guzii, A. G.; Pislyagin, E. A.; Es'kov, A. A.; Kozhemyako, V. B.; Aminin, D. L.; Wang, Y. M.; Stonik, V. A. Urupocidin A: a new, inducing iNOS expression bicyclic guanidine alkaloid from the marine sponge Monanchora pulchra. *Org Lett.* **2014**, *16*, 4292.

(39) Nemoto, A.; Hoshino, Y.; Yazawa, K.; Ando, A.; Mikami, Y.; Komaki, H.; Tanaka, Y.; Grafe, U. Asterobactin, a new siderophore group antibiotic from Nocardia asteroides. *J. Antibiot.* **2002**, *55*, 593.

(40) Kumagai, T.; Takagi, K.; Koyama, Y.; Matoba, Y.; Oda, K.; Noda, M.; Sugiyama, M. Heme protein and hydroxyarginase necessary for biosynthesis of D-cycloserine. *Antimicrob Agents Chemother.* **2012**, *56*, 3682.

(41) Wilkinson, C. J.; Hughes-Thomas, Z. A.; Martin, C. J.; Bohm, I.; Mironenko, T.; Deacon, M.; Wheatcroft, M.; Wirtz, G.; Staunton, J.; Leadlay, P. F. Increasing the efficiency of heterologous promoters in actinomycetes. *J Mol Microbiol Biotechnol.* **2002**, *4*, 417.

(42) Bibb, M. J.; Janssen, G. R.; Ward, J. M. Cloning and analysis of the promoter region of the erythromycin resistance gene (ermE) of Streptomyces erythraeus. *Gene.* **1985**, *38*, 215.

(43) Hasenoehrl, A.; Galic, T.; Ergovic, G.; Marsic, N.; Skerlev, M.; Mittendorf, J.; Geschke, U.; Schmidt, A.; Schoenfeld, W. In vitro activity and in vivo efficacy of icofungipen (PLD-118), a novel oral antifungal agent, against the pathogenic yeast Candida albicans. *Antimicrob Agents Chemother.* **2006**, *50*, 3011.

(44) Kiss, L.; Mandity, I. M.; Fulop, F. Highly functionalized cyclic beta-amino acid moieties as promising scaffolds in peptide research and drug design. *Amino Acids.* **2017**, *49*, 1441.

(45) Yeates, C. Icofungipen (PLIVA). *Curr Opin Investig Drugs.* **2005**, *6*, 838.

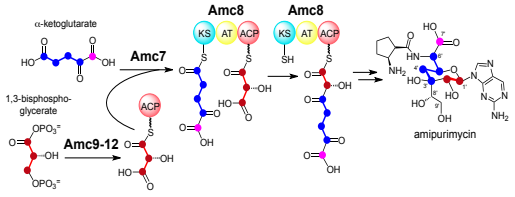
(46) Rangaswamy, V.; Jiralerspong, S.; Parry, R.; Bender, C. L. Biosynthesis of the Pseudomonas polyketide coronafacic acid requires monofunctional and multifunctional polyketide synthase proteins. *Proc Natl Acad Sci U S A.* **1998**, *95*, 15469.

(47) Strieter, E. R.; Koglin, A.; Aron, Z. D.; Walsh, C. T. Cascade reactions during coronafacic acid biosynthesis: elongation, cyclization, and functionalization during Cfa7-catalyzed condensation. *J Am Chem Soc.* **2009**, *131*, 2113.

(48) Rangaswamy, V.; Mitchell, R.; Ullrich, M.; Bender, C. Analysis of genes involved in biosynthesis of coronafacic acid, the polyketide component of the phytotoxin coronatine. *J Bacteriol.* **1998**, *180*, 3330.

(49) Lin, C.-I.; McCarty, R. M.; Liu, H.-w. "The Biosynthesis of Nitrogen-, Sulfur-, and High-carbon Chain-Containing Sugars." *Chem. Soc. Rev.* **2013**, *42*, 4377.

Table Of Contents artwork image:



Insert Table of Contents artwork here