Chemistry & Biology
Article

N-Carbamoylation of 2,4-Diaminobutyrate Reroutes the Outcome in Padanamide Biosynthesis

Yi-Ling Du,[1] Doralyn S. Dalisay,[1] Raymond J. Andersen,[1,2] and Katherine S. Ryan[1,*]
[1]Department of Chemistry
[2]Department of Earth, Ocean and Atmospheric Sciences
University of British Columbia, Vancouver, BC V6T 1Z1, Canada
*Correspondence: ksryan@chem.ubc.ca
http://dx.doi.org/10.1016/j.chembiol.2013.06.013

SUMMARY

Padanamides are linear tetrapeptides notable for the absence of proteinogenic amino acids in their structures. In particular, two unusual heterocycles, (*S*)-3-amino-2-oxopyrrolidine-1-carboxamide (*S*-Aopc) and (*S*)-3-aminopiperidine-2,6-dione (*S*-Apd), are found at the C-termini of padanamides A and B, respectively. Here we identify the padanamide biosynthetic gene cluster and carry out systematic gene inactivation studies. Our results show that padanamides are synthesized by highly dissociated hybrid nonribosomal peptide synthetase/polyketide synthase machinery. We further demonstrate that carbamoyltransferase gene *padQ* is critical to the formation of padanamide A but dispensable for biosynthesis of padanamide B. Biochemical investigations show that PadQ carbamoylates the rare biosynthetic precursor L-2,4-diaminobutyrate, generating L-2-amino-4-ureidobutyrate, the presumed precursor to the C-terminal residue of padanamide A. By contrast, the C-terminal residue of padanamide B may derive from glutamine. An unusual thioesterase-catalyzed cyclization is proposed to generate the *S*-Aopc/*S*-Apd heterocycles.

INTRODUCTION

Discovery of novel compounds with pharmaceutical potential from marine sources has received increased attention in recent years (Gerwick and Moore, 2012). The marine actinomycetes, in particular, have emerged as a promising resource, producing novel natural products with unusual chemical structures and various biological activities (Hughes and Fenical, 2010). Recently, *Streptomyces* sp. RJA2928 (Williams et al., 2011), a strain isolated from sediment in the Pacific Ocean, near Padana Nahua in Papua New Guinea, has emerged as an exciting source for new molecules. The strain produces a series of polyketides, including the nahuoic acids (Williams et al., 2013), along with padanamides A and B (Williams et al., 2011), which are highly modified linear tetrapeptides (Figure 1A). The padanamides lack any proteinogenic amino acids and possess substructures that have not been previously reported in the natural products literature. It is interesting that a compound identical to padanamide A, named actinoramide A (Nam et al., 2011), was independently reported from *Streptomyces* sp. CNQ-027. This actinomycete strain was isolated from sediment on the opposite side of the Pacific Ocean, near San Diego, CA, suggesting a potentially wide distribution of the padanamides/actinoramides. It is intriguing that, whereas padanamide A and actinoramide A are identical, the minor compounds (actinoramides B and C) co-isolated from *Streptomyces* sp. CNQ-027 are unique (Figure 1A). Total synthesis of padanamides A and B was recently reported (Long et al., 2013), confirming the previous structural elucidations.

The padanamides attracted our attention for their many unusual chemical features. In particular, the (*S*)-3-amino-2-oxopyrrolidine-1-carboxamide (*S*-Aopc) and (*S*)-3-aminopiperidine-2,6-dione (*S*-Apd) residues at the C termini of padanamides A and B, respectively, are unprecedented in naturally occurring peptides and lack a clear biosynthetic origin. Other unusual structural features for padanamides include an N-terminal methoxyacetyl group and a 4-amino-3-hydroxy-2-methyl-5-phenylpentanoic acid (Ahmpp) residue. 3-hydroxyleucine (Hleu) and piperazate (Piz) are also rare, nonproteinogenic peptidic building blocks. Padanamide B displays cytotoxicity against Jurkat T lymphocyte cells, with a half maximal inhibitory concentration value of 20 μg/ml, while padanamide A is roughly 3-fold less active in the same assay (Williams et al., 2011). A chemical genomics analysis using yeast deletion mutants suggests that padanamide A is involved in the inhibition of cysteine and methionine biosynthesis or that these amino acids are involved in the yeast's response to the peptide. By contrast, padanamide B does not exhibit the same activity at the concentration where padanamide A is active (Williams et al., 2011). As padanamides A and B differ only in their corresponding C-terminal residues, these biological data suggest that the unusual C-terminal residues of padanamides are quite important to their biological activities. Even more intriguing, the positioning of the *S*-Aopc and *S*-Apd residues at the C termini of the corresponding peptides suggests that elucidation of the responsible genes and enzymes might enable new strategies for combinatorial biosynthesis of linear peptides with new C-terminal modifications (Menzella and Reeves, 2007).

Here, we report the biosynthetic gene cluster of padanamides from *Streptomyces* sp. RJA2928. In our work, systematic gene inactivation mutants are used to delineate the padanamide gene cluster. Our results reveal a hybrid nonribosomal peptide synthetase/polyketide synthase (NRPS/PKS) system for assembly of the padanamide skeleton, along with a number of enzymes

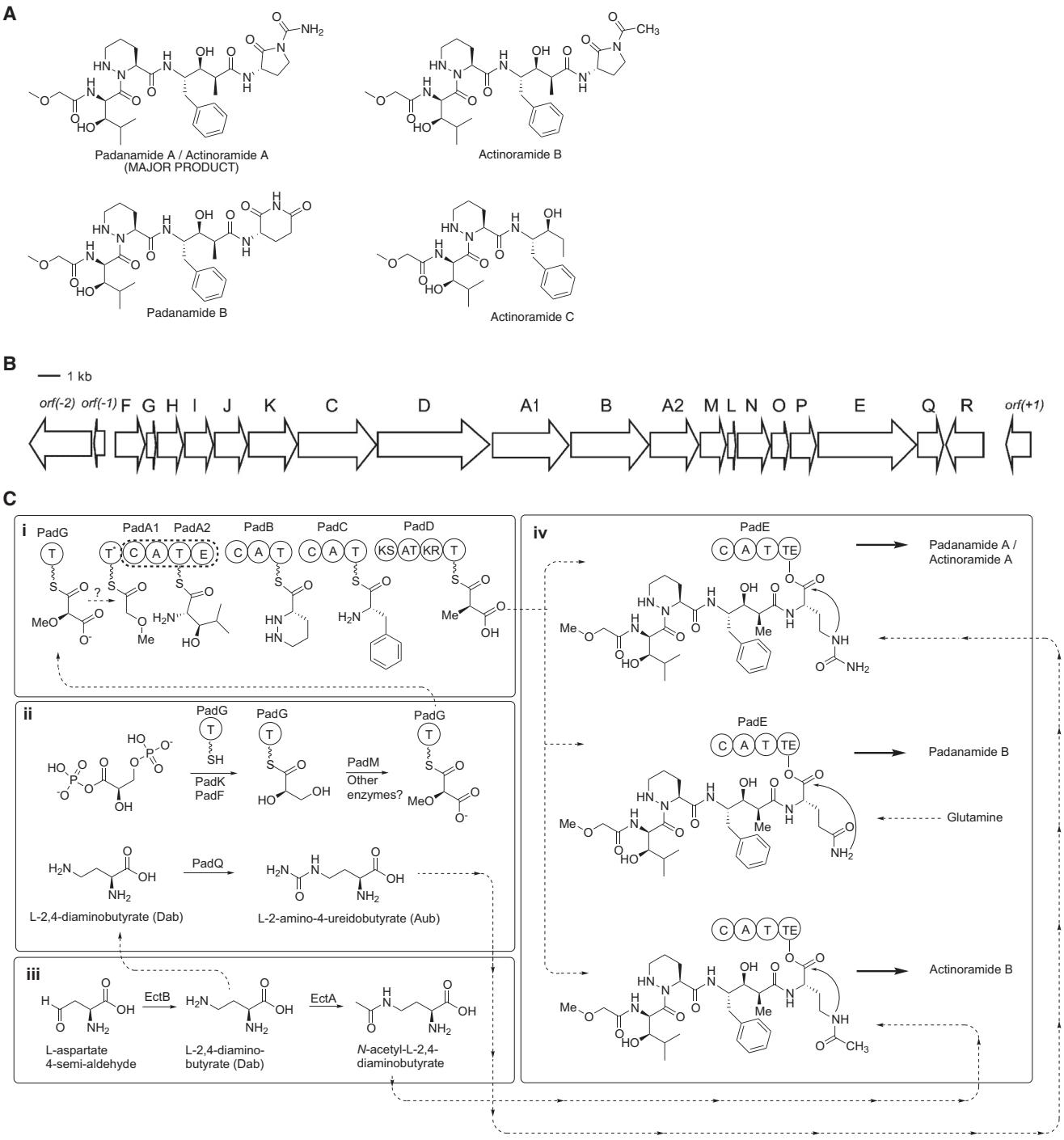

Figure 1. Padanamide Biosynthetic Pathway

(A) Structures of padanamides and actinoramides.

(B) Padanamide biosynthetic gene cluster.

(C) Proposed biosynthetic route to the padanamides and actinoramides. (i) Hybrid NRPS/PKS assembly of the N-terminal aa of the padanamides. (ii) Biosynthesis of methoxymalonyl-PadG (top) and L-Aub (bottom). (iii) Biosynthesis of L-Dab and N-acetyl-L-2,4-diaminobutyrate from L-aspartate-4-semialdehyde. (iv) Formation of padanamide A, padanamide B, and actinoramide B using TE-catalyzed cyclization with amide nitrogens as nucleophiles.

See also Figure S1 and Table S2.

believed to be involved in the formation of unusual building blocks. Notably, further in vitro biochemical characterization reveals a carbamoyltransferase that catalyzes the transformation of L-2,4-diaminobutyrate (L-Dab) to L-2-amino-4-ureidobutyrate (L-Aub), which is a key step in the biosynthesis of the *S*-Aopc residue of padanamide A.

Table 1. Deduced Functions of ORFs of the Padanamide Biosynthetic Gene Cluster

Gene	Size (aa)	Homolog and Identity/ Similarity (%)[a]	Proposed Function
orf(−2)	3,363	SBI_09160 (YP_ 004967409), 78/82	lantibiotic biosynthesis protein LanM
orf(−1)	171	SBI_09159 (YP_ 004967408), 75/83	putative type 2 lantibiotic
padF	390	–	unknown
padG	104	Npun_F3159 (YP_ 001866560), 36/58	discrete ACP
padH	316	QncN (AFJ11257), 65/73	dehydrogenase complex, E1 component, alpha subunit
padI	348	QncL (AFJ11255), 55/68	dehydrogenase complex, E1 component, beta subunit
padJ	409	Tsn2 (ACR50769), 40/53	dehydrogenase complex, E2 component
padK	578	PAV_7c00910 (ZP_ 10866675), 29/44	FkbH-like protein
padC	1,101	MicvaDRAFT_0252 (ZP_08495279), 38/54	NRPS: C-A-T
padD	1,538	SRIM_23766 (ZP_ 20967056), 46/57	PKS: KS-AT-KR-T
padA1	1,116	–	NRPS: T-C-A
padB	1,136	AerG1 (ACM68690), 31/47	NRPS: C-A-T
padA2	605	–	T-E didomain protein
padM	311	TMO_a0589 (YP_ 006373576), 47/57	*O*-methyltransferase
padL	73	MbtH (YP_006267078), 78/83	MbtH-like protein
padN	448	KtzI (ABV56589), 48/61	ornithine *N*-hydroxylase
padO	233	Orf4 (ABC56600), 51/62	regulator
padP	281	Orf5 (AAO23330), 31/50	hydroxylase
padE	1,355	MicD (CAQ48261), 29/44	NRPS: C-A-T-TE
padQ	318	Zwa6 (ZP_03234165), 38/56	Asp/OTC
padR	597	SARP regulator (YP_ 004967407), 77/84	SARP family regulator
orf(+1)	142	Strvi_4981 (YP_ 004814808), 92/95	Asp aminotransferase

See also bioinformatic analysis of PadFKG and PadE-TE in Figure S3.
[a]Parenthetical codes are National Center for Biotechnology Information accession numbers.

RESULTS

Identification of the Padanamide Biosynthetic Gene Cluster

Retrobiosynthetic analysis of padanamide A (1) suggests that this tetrapeptide is assembled via a hybrid NRPS/PKS assembly line from six precursors: (1) 2-methoxyacetate, (2) 3-OH-D-leucine, (3) Piz, (4) phenylalanine, (5) methylmalonate, and (6) an unknown molecule that is ultimately converted to the *S*-Aopc residue. We thus hypothesized that the padanamide gene cluster would contain genes for NRPSs and PKSs, along with accessory genes. Accessory genes would be involved in the construction of the 2-methoxyacetyl group and might resemble the starter unit pathway for apoptolidins (Du et al., 2011). Additionally, an ornithine *N*-monooxygenase would likely generate N^5-OH-ornithine, the precursor to Piz, as observed in the biosynthesis of the kutznerides (Fujimori et al., 2007), sangliferhin A (Qu et al., 2011), himastatin (Ma et al., 2011), and other Piz-containing natural products.

To find the padanamide gene cluster, we initially explored isolation of adenylation (A) domain and *N*-oxygenase genes from RJA2928 genomic DNA using PCR with degenerate primers. However, we found that *Streptomyces* sp. RJA2928 genomic DNA contained many distinct A domain and *N*-oxygenase genes, making this approach inefficient. We thus turned to genome scanning, assembling ~11.3 megabase pairs of nonredundant sequences over 862 contigs with 236-fold read coverage. Bioinformatic analysis identified one putative biosynthetic gene cluster for padanamides biosynthesis (Figure 1B). This cluster spans 34.2 kb on a single contig and consists of 19 open reading frames (ORFs), whose deduced functions are consistent with padanamide biosynthetic pathway. Among these ORFs, there are four encoding NRPSs, one for a PKS, and a thiolation-epimerization (T-E) didomain protein, suggesting that an NRPS/PKS hybrid system is involved in padanamide biosynthesis. Several other genes potentially involved in regulation, starter unit construction, and building block formation are also observed (Table 1).

Determination of the Gene Cluster Boundaries

To test the necessity of various genes and provide insight into their roles in biosynthesis of padanamides, we conducted gene inactivation experiments by using the REDIRECT technique (Gust et al., 2003; Tables S1 and S2 available online). We first screened the cosmid genomic library of *Streptomyces* sp. RJA2928 and obtained four cosmids that cover part of the *pad* cluster. The coverage of each cosmid was identified by end sequencing or estimated by restriction enzyme digestion pattern analysis.

Initially, we generated an inactivation mutant in *padK* (Figure S1A), which is expected to encode an FkbH-like enzyme, potentially involved in starter unit construction. The *ΔpadK* mutant lost the ability to produce padanamides (Figure 2A). To verify that the metabolite profile change was solely due to the inactivation of this gene, we cloned the coding region of *padK* into the integrative vector pYLD20 under the control of the constitutive promoter *ermE*p* (Figure S2F). High-pressure liquid chromatography (HPLC) and liquid chromatography-mass spectrometry (LC-MS) analysis revealed that reintroduction of *padK* restored production of padanamides A and B (Figure S2A). These data confirm that *padK* and, by extension, the *pad* cluster we identified are responsible for padanamide biosynthesis.

The 3′ boundary of the *pad* cluster is determined by gene inactivation of *orf(+1)*, which is predicted to belong to the aspartate aminotransferase superfamily. This gene is located 642 base

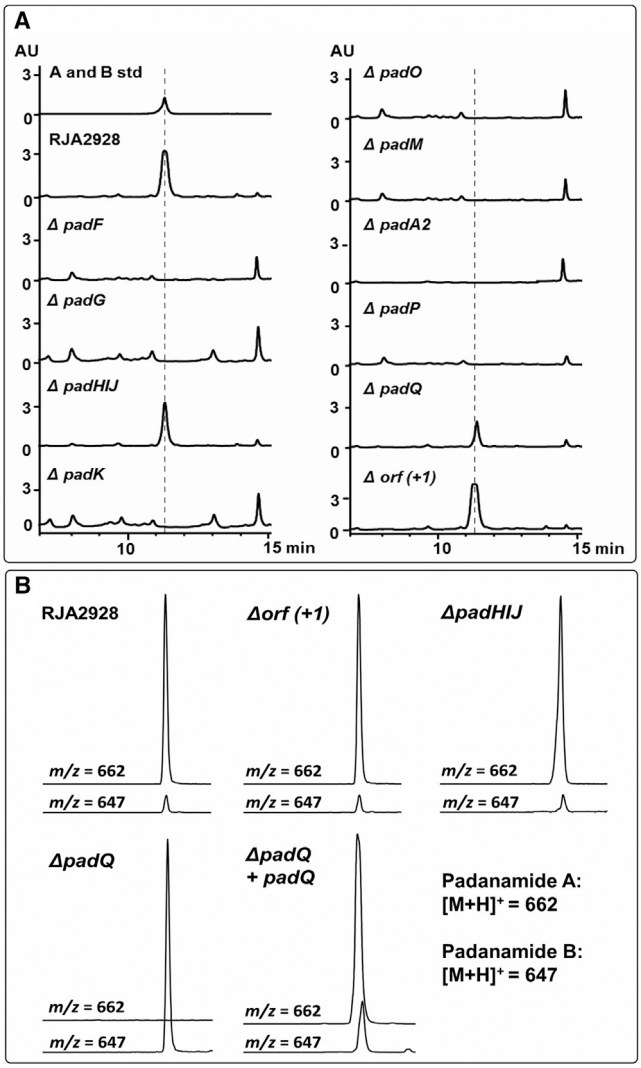

Figure 2. Metabolic Analysis of Engineered Strains

(A) Analysis of padanamide production of *Streptomyces* sp. RJA2928 and its mutants by HPLC (detection wavelength: 211 nm). The peak corresponding to the padanamides is marked with a dashed line (padanamides A and B are not well separated under the current HPLC method). Note that a peak at 14.6 min is present in all strains, including *Streptomyces* sp. RJA2928. Precursor metabolites may be funneled toward this competing biosynthetic pathway in the absence of a functional padanamide pathway; hence, the increase in intensity of this nonpadanamide peak when padanamide biosynthesis is eliminated. AU, absorbance units; std, standard.

(B) Padanamide production from *Streptomyces* sp. RJA2928 and *ΔpadHIJ*, *Δorf(+1)*, *ΔpadQ*, and *ΔpadQ* + *padQ* was further analyzed by LC-MS. Selective ion monitoring was used to detect the $[M+H]^+$ ion for padanamide A (m/z = 662) and padanamide B (m/z = 647). The results shows *ΔpadHIJ* and *Δorf(+1)* produce padanamides A and B normally, while *ΔpadQ* only produces padanamide B.

See also Figure S2 and Table S1.

pairs (bp) downstream of putative regulator gene *padR*, which encodes a putative SARP-family regulator, a type of regulator frequently associated with secondary metabolite gene clusters in *Streptomyces* (Liu et al., 2013). Our data show that the *Δorf(+1)* mutant strain produces padanamides at a titer comparable to that of the parental strain (Figure 2A), indicating that *orf(+1)* is outside of the *pad* cluster. Thus, we propose that *padR* is the 3′ boundary of the *pad* cluster. For the 5′ boundary of the *pad* cluster, we generated the *ΔpadF* mutant, which lost the ability to produce padanamides (Figure 2A). Three hundred bp upstream of *padF* are *orf(−1)* and *orf(−2)*, which show high sequence identity (75% and 78%, respectively) to putative type 2 lantibiotic precursor peptide (SBI_09159) and a lantibiotic biosynthesis protein LanM (SBI_09160) from *S. bingchenggensis* BCW-1. Given that the whole genome of RJA2928 exhibits a high degree of homology with *S. bingchenggensis* BCW-1, yet *S. bingchenggensis* BCW-1 lacks a *pad* gene cluster, we propose that *orf(−1)* and *orf(−2)* are involved in biosynthesis of a natural product distinct from padanamide. Thus, we propose that *padF* is the 5′ boundary of the *pad* cluster (Figure 1).

Assembly of the Padanamide Core

All of the four NRPS genes (*padA1*, *padB*, *padC*, and *padE*) found in the *pad* cluster were predicted to encode monomodular NRPSs. PadA1 contains three typical NRPS domains with the following arrangement: T domain, condensation (C) domain, and A domain. By contrast, most NRPS domains have the following characteristic ordering: C-A-T. Sequence analysis further reveals that the PadA1-T, which has a conserved serine residue, may be a truncated T domain. By contrast, the A domain of PadA1 appears typical, and the sequence of the A domain specificity "code" (DAWYIGAIVK) suggests that the PadA1 NRPS module will incorporate L-leucine. A T-E didomain protein PadA2 was also identified in the padanamide gene cluster, and the presence of an E domain in the cluster agrees with the occurrence of D-stereochemistry in the Hleu residue in the padanamides. Here, the T domain appears to be a typical E-domain-coupled T, with a well-conserved CoreT sequence, GGDSI (Linne et al., 2001). We propose that PadA1 and PadA2 collectively form the first NRPS module (C-A-T-E) through protein-protein interactions, where the initial PadA1-T domain may be inactive or may act to carry the putative methylacetyl starter unit (Figure 1C). Splitting of modules across two genes has been observed previously in zwittermicin A and himastatin biosynthesis (Kevany et al., 2009; Ma et al., 2011). The involvement of PadA2 in padanamide biosynthesis was confirmed by in vivo gene knockout and complementation experiments: padanamide biosynthesis is abolished in *ΔpadA2*, and padanamide production is restored by provision of *padA2* (Figure 2A; Figure S2E).

The A domain substrate signature of PadB (DVQYMAQVVK) is most similar to those for proline/pipecolic acid, which are structurally similar to Piz. We thus propose that PadB is responsible for Piz incorporation. The predicted substrate specificity (DAWTVAAVAK) of PadC is phenylalanine. PadD is a monomodular type I PKS containing four domains: ketosynthase (KS), acyltransferase (AT), ketoreductase (KR), and T domain. The predicted substrate for its AT domain is methylmalonyl-coenzyme A. The substrate selectivity and domain organization of *padC* and *padD* is consistent with the structure of Ahmpp residue present in padanamides. Finally, a thioesterase (TE) domain is identified at the C-terminal end of PadE, suggesting that PadE is the last module in the NRPS assembly line and is responsible for the peptidyl chain release. Sequence analysis of the A domain

of PadE reveals that its substrate signature (DVQDIGCVNK) is most similar to those of lysine/glutamine-activating A domains. The likely substrates of PadE are addressed later.

Additionally, *padL* encodes an MbtH-like protein, which has been found in many NRPS-encoding gene clusters and shown to be an integral component of NRPSs (Felnagle et al., 2010). Overall, this five-module NRPS/PKS assembly line supports the structure of padanamide core, but one module is split across two genes, and the arrangement of genes does not abide the co-linearity rule of classical NRPSs (Marahiel et al., 1997).

Starter Unit Biosynthesis

An N-terminal 2-methoxyacetyl group is present in padanamides. A number of methoxyacetyl-containing polyketides have been shown to use methoxymalonyl as a T-domain-protein-linked PKS extender unit to elongate the polyketide chain with a methoxylacetyl unit (Chan et al., 2006; Wu et al., 2000; Zhao et al., 2006), and recently, one report has described the use of methoxymalonyl as a starter unit that is ultimately converted to a methoxyacetyl group (Du et al., 2011). In each of the aforementioned pathways, five genes (*fkbGHIJK* homologs) are thought to be involved in the biosynthesis of the methoxymalonyl starter unit. At the first step, the enzyme FkbH dephosphorylates and covalently tethers the glycolytic substrate 1,3-bisphosphoglycerate to the T domain protein FkbJ, forming glyceryl-FkbJ (Chan et al., 2006). After the oxidation of glyceryl-FkbJ to hydroxymalonyl-FkbJ catalyzed by FkbK and FkbI, FkbG catalyzes the final *O*-methylation to generate methoxymalonyl-FkbJ (Chan et al., 2006). However, only a subset of these genes is present in the padanamide cluster. There is a free-standing T domain protein (PadG), an atypical FkbH-like protein (PadK) (Figure S3A), and an *O*-methyltransferase (PadM), which are functional equivalents to FkbJ, FkbH, and FkbG, respectively. Additionally, the padanamide pathway has an additional protein, PadF, whose N terminus has sequence similarity to N-terminal extensions of several enzymes that also contain FkbH domains, suggesting that PadF may play an accessory role to PadK (Figure S3B). However, FkbK and FkbI homologs are absent from the padanamide gene cluster, as is the KS^{Q} thought to be involved in conversion of the methoxymalonyl group to a methoxyacetyl group (Du et al., 2011).

To assess which gene(s) are responsible for formation of the starter unit, we constructed in-frame deletion mutants of *padF*, *padG*, *padM*, and *padK*. No padanamides were observed from HPLC analysis of the culture extracts from *ΔpadF*, *ΔpadG*, *ΔpadK,* and *ΔpadM* mutant strains (Figure 2A), demonstrating that each of the encoded enzymes are necessary for padanamide construction. The aforementioned bioinformatic analysis, as well as in vivo gene inactivation results, suggests that the incorporation of methoxyacetyl unit may share some of the steps for biosynthesis of methoxymalonyl unit previously described for PKS systems. In particular, glyceryl-PadG formation by PadK and PadF could be the initial step of padanamide biosynthesis, and PadM could catalyze the *O*-methylation step. However, our current data give no clue about how the intermediate step of conversion of glyceryl-ACP to hydroxymalonyl-PadG is carried out. It is possible that genes outside the *pad* cluster participate in this process, as two additional sets of *fkbGHIJK* homologs are identified in the genome. However, in this case, the *padK* mutant itself may not be complemented, as observed, as PadK is an "atypical" FkbH homolog (Figure S3A), which differs from the classic FkbH family members with an additional N-terminal extension (~260 amino acids [aa]). In the tetronomyin pathway, Sun et al. showed that PadK's homolog Tmn16 needs its N-terminal domain for functional integrity and the Tmn16-Tmn7a (cognate T domain protein) interaction is specific (Sun et al., 2008). PadK may thus require a specific cognate T domain protein (PadG) and accessory protein (PadF). Furthermore, the timing of methylation by PadM is unknown. Hence, the *padG*, *padF*, and *padM* mutants may not be complemented by exogenous *fkbGHIJK* genes. Finally, conversion of the final methoxymalonyl-PadG to methoxyacetyl-PadG is unknown, as no genes could be linked to this step. Another possible scenario to account for precursor construction is that a previously undescribed pathway gives rise to the methoxyacetyl precursor. All our efforts to obtain soluble PadK were unsuccessful, and details of the initial step remain to be established.

Embedded between *padG* (encoding a T domain protein) and *padK* (encoding an FkbH-like enzyme) are *padH*, *padI*, and *padJ* (*padH-I-J*). PadH, PadI, and PadJ are not related to Fkb enzymes: PadH and PadI are instead highly related to recently characterized QncN and QncL (Table 1), respectively, from the quinocarcin biosynthetic pathway (Peng et al., 2012). QncN and QncL have been shown to catalyze the transfer of a C2 unit from ketose to a cognate T domain protein QncM to form glycolicacyl-ACP as an extender unit for an NRPS. The transcription of *padFGHIJKCD* appears to be coupled due to overlapping stop and start codons. It is unexpected to observe genes that are potentially involved in the biosynthesis of the N-terminal methoxylacetyl group from two different pathways for incorporating two-carbon units. We constructed a deletion mutant of the entire *padH-I-J*, and we observed that production of padanamides in *ΔpadH-I-J* mutants remained nearly unaffected (Figure 2), demonstrating that *padH-I-J* are not involved in padanamide production.

Additional Precursor Biosynthesis and Regulation

The gene product of *padP* (281 aa) belongs to the 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily (Pfam ID 13640) and shows 27% sequence identity with LdoA, recently characterized as an Fe(II)/α-ketoglutarate-dependent L-leucine 5-hydroxylase (Hibi et al., 2013). Sequence alignment reveals that both LdoA and PadP contain a conserved histidine $(His)_2$-aspartate (Asp) facial triad (H167, D169, and H251 for PadP) involved in Fe(II) binding as well as the basic residue (R262 for PadP) that binds to α-ketoglutarate. PadP is likely to catalyze a related reaction, i.e., 3-hydroxylation of L-leucine. We constructed *ΔpadP* and confirmed by HPLC analysis of the crude extracts that padanamide production was eliminated (Figure 2A), definitively linking PadP to the biosynthesis of padanamides. Furthermore, selective ion monitoring by MS failed to detect the target ion for dehydroxypadanamide A or B for the *ΔpadP* mutant, indicating that the β-hydroxylation on leucine of padanamide is likely to occur on the NRPS assembly line rather than after release of the peptide chain from the NRPS.

The gene product of *padN* is homologous to various lysine/ornithine monooxygenases, and its homologs have been found in the biosynthetic gene clusters of kutznerides, himastatin, and

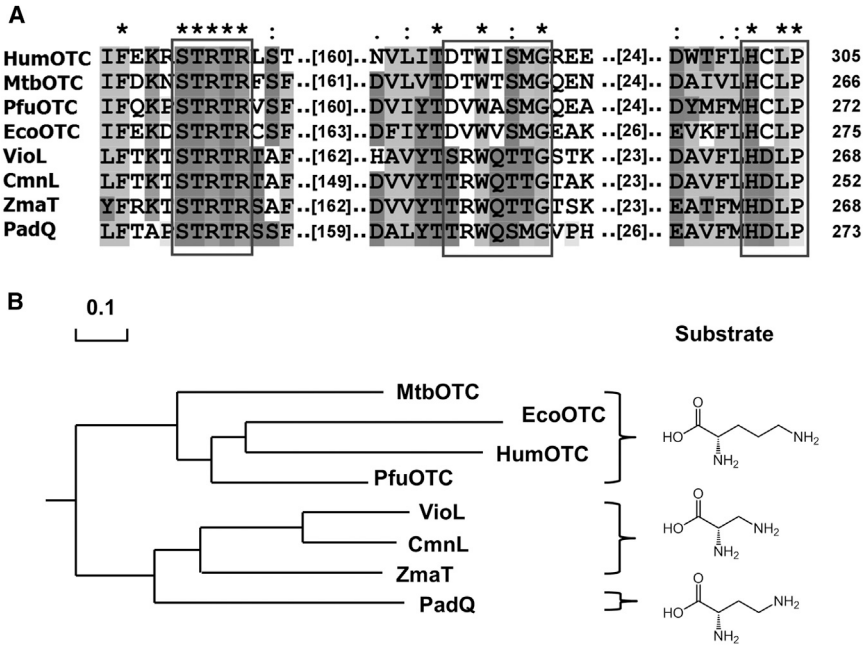

Figure 3. Bioinformatic Analysis of PadQ

(A) Sequence alignment of PadQ with ornithine and L-Dap carbamoyltransferases. HumOTC, human OTC (AAI07154); MtbOTC, *Mycobacterium tuberculosis* OTC (UniProtKB/Swiss-Prot: P0A5M8.1); PfuOTC, *Pyrococcus furiosus* OTC (X99225); and EcoOTC, *Escherichia coli* OTC (YP_003937659). Boxes indicate the conserved carbamoyl phosphate binding motif STRTR and the ornithine-binding motifs DXXXSMG and HCLP for OTC. VioL (AAP92502), CmnL (ABR67755), and ZmaT (ACM79819) are the putative L-Dap carbamoyltransferases from viomycin, capreomycin, and zwittermicin A biosynthetic pathways, respectively.

PfuOTC, *Pyrococcus furiosus* OTC (X99225).

(B) Phylogenetic analysis of PadQ with its homologs. Their substrates are indicated at the right side. See also Figure S5.

sanglifehrin, all of which contain Piz (Fujimori et al., 2007; Ma et al., 2011; Qu et al., 2011). KtzI, the homolog from the kutzneride gene cluster, has been recently defined as an ornithine *N*-hydroxylase, which catalyzes the conversion of ornithine to N^5-OH-ornithine (Neumann et al., 2012). Furthermore, N^5-OH-ornithine was identified as an early intermediate in Piz biosynthesis by isotopic-labeling experiments (Neumann et al., 2012). These data suggest that PadN could be involved in the production of the Piz residue of padanamides in a similar manner.

The gene *padR* encodes a putative SARP family regulator, and SARP family regulators are frequently found to be associated with a secondary metabolite gene clusters in *Streptomyces* (Liu et al., 2013). Additionally, there appears to be a second regulator in the padanamide pathway, encoded by *padO*. PadO contains an FMN binding domain (Pfam ID 04229), showing homology to PaiB, which is a transcriptional regulator negatively controlling sporulation in *Bacillus subtilis* (Honjo et al., 1990). Although the detailed mechanism for this family of regulators is still unknown, recent structural analysis of a homolog from *Geobacillus stearothermophilus* revealed a likely C-terminal helix-turn-helix DNA binding domain, suggesting that these regulators may act as DNA-binding transcriptional regulators (Filippova et al., 2011). We constructed *ΔpadO* and its complemented mutant (Figure 2A; Figure S2C), and the HPLC results showed that the *ΔpadO* mutant abolished padanamide production, indicating that it probably acts as a pathway activator. PadO also has sequence identity of 51% to Orf4 of the kutzneride gene cluster (Fujimori et al., 2007) and 45% to SfaC of the sanglifehrin gene cluster (Qu et al., 2011). Both Orf4 and SfaC are annotated as putative regulators; however, no genetic evidence is available. Our genetic results indicate that PadO and its homologs may represent a novel family of transcriptional regulators that control the biosynthesis of some secondary metabolites.

Biosynthesis of *S*-Aopc and *S*-Apd

1 is the major padanamide (~92%) relative to **2** (~8%) in wild-type *Streptomyces* sp. RJA2928. The difference between the molecules is in the C-terminal residue, where the *S*-Aopc residue in **1** differs from the *S*-Apd in **2** via both an altered cyclization and the formal addition of an "NH." Both substructures are unique in the natural products literature. From a biosynthetic point of view, two pathways could be envisioned to account for these two molecules. A first possibility is that padanamide B could represent a biosynthetic intermediate that undergoes further conversions to give padanamide A. The second possibility is that alternate aa are loaded and cyclized by a nonribosomal synthetase module, with preference given to the precursor that leads to *S*-Aopc.

The gene *padQ*, putatively encoding a carbamoyltransferase, attracted our attention. In the *ΔpadQ* mutant, padanamide B production increased more than 5-fold, with a complete loss in padanamide A production. Production of padanamide A can be restored by complementation in trans (Figures 2A and 2B). This result unambiguously connects *padQ* to the formation of the unique *S*-Aopc residue in padanamide A.

PadQ is predicted to be a putative ornithine carbamoyltransferase (OTC), containing two putative domains: the N-terminal carbamoyl phosphate-binding domain and the C-terminal ornithine-binding domain. OTCs catalyze the carbamoylation of L-ornithine to form L-citrulline. There is a carbamoyl group in *S*-Aopc of padanamide A; however, L-citrulline itself cannot be converted to *S*-Aopc, as it has an additional methylene group relative to *S*-Aopc. Instead, carbamoylation of L-Dab, with one less methylene than L-ornithine, could sensibly give this structure. We speculated that the putative carbamoylated product of L-Dab, i.e., L-Aub, could be activated and cyclized to give rise to *S*-Aopc. Bioinformatic analysis of PadQ shows that PadQ has the characteristic STRTR sequence for carbamoyl phosphate binding seen in OTCs (Sankaranarayanan et al., 2008). However, unlike OTCs that have ornithine-binding motifs DXXXSMG and HCLP, PadQ shows deviations in this strictly conserved region: DXXXSMG is replaced by TRWQSMG and HCLP is replaced by HDLP (Figure 3A), suggesting that the

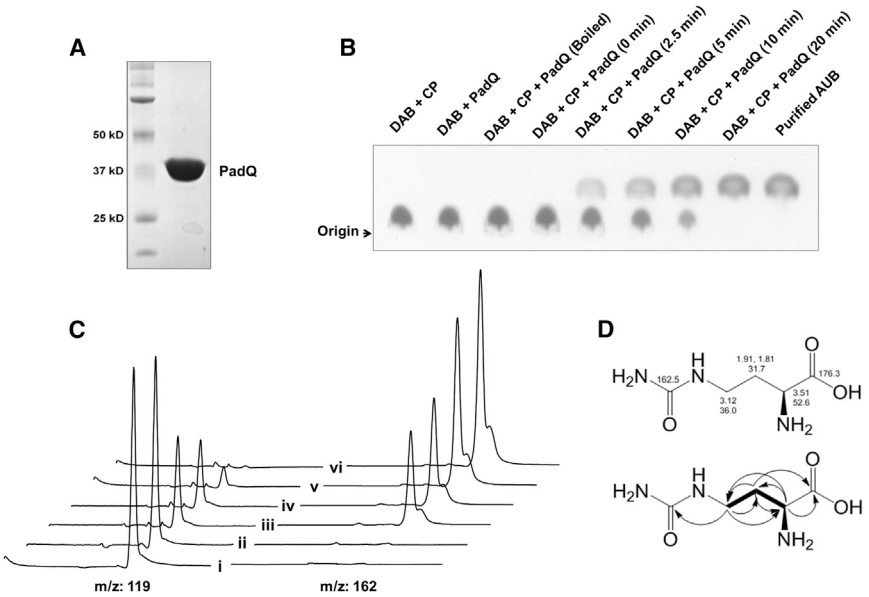

Figure 4. In Vitro Analysis of Carbamoyltransferase PadQ

(A) Purification of PadQ from *Rhodococcus* sp. RHA1.

(B) Monitoring of PadQ-catalyzed carbamoylation of L-Dab by TLC.

(C) Analysis of PadQ-catalyzed carbamoylation of L-Dab (indicated by DAB) by LC-MS analysis monitoring the $[M+H]^+$ ion for L-Dab (*m/z* = 119) and L-Aub (*m/z* = 162). (i) L-Dab + CP + Boiled PadQ (20 min), (ii) L-Dab + CP + PadQ (0 min), (iii) L-Dab + CP + PadQ (2.5 min), (iv) L-Dab + CP + PadQ (5 min), (v) L-Dab + CP + PadQ (10 min), and (vi) L-Dab + CP + PadQ (20 min). AUB, L-Aub.

(D) NMR analysis of purified L-Aub.

See also Figure S4.

substrate for PadQ may not be ornithine. Furthermore, we also highlight related changes in these two sites for putative L-2,3-diaminopropionate (L-Dap) carbamoyltransferases (Barkei et al., 2009). In these enzymes, DXXXSMG is replaced by (S/T) RWQTTG and HCLP is replaced again by HDLP. L-Dap is two methylene groups shorter than L-ornithine (Figure 3B).

Initially, we attempted to purify PadQ from recombinant expression of *padQ* in *E. coli*, but we were unable to obtain soluble protein. Instead, we overexpressed PadQ as a His-tagged fusion protein in *Rhodococcus* sp. RHA1 and purified it to homogeneity (Figure 4A). We first tested L-Dab as a substrate for PadQ. We incubated L-Dab with carbamoyl phosphate (CP) and PadQ, and the reaction mixture was analyzed by thin-layer chromatography (TLC) and visualized with ninhydrin. As shown in Figure 4B, a new alpha amino acid appears that depends on CP and active PadQ. Full conversion of L-Dab to the new product is observed after 20 min of incubation. LC-MS analysis was then used to monitor the $[M+H]^+$ ion for L-Dab (mass-to-charge ratio [*m/z*] = 119) and L-Aub (*m/z* = 162), which is the expected product of PadQ-catalyzed carbamoylation of L-Dab (Figure 4C). The result shows that the ion signal for protonated L-Dab declines over incubation time, while a new, PadQ-dependent ion signal (*m/z* 162) increases, in a manner highly consistent with the TLC data. To further confirm that the new compound observed on TLC plate is L-Aub, we purified this enzymatic product to homogeneity by repeated preparative TLC on silica gel and subjected to HRESIMS and tandem MS (MS/MS) analysis. The purified compound gave an $[M+H]^+$ ion in the HRESIMS at *m/z* 162.0879, consistent with the molecular formula of $C_5H_{11}N_3O_3$ for L-Aub (calculated for $C_5H_{12}N_3O_3$, 162.0879) (Figure S4A). MS/MS analysis produced an ion at *m/z* 145, which could be due to the loss of ammonia from the $[M+H]^+$ ion, and higher fragmentation energy caused full conversion of *m/z* 162 to *m/z* 145. Further fragmentation of the ion at *m/z* 145 gave both an ion at *m/z* 99 and 101, likely corresponding to the loss of, respectively, formic acid and CO_2 from the $[M+H\text{-}NH_3]^+$ ion (Figure S4A). The structure of L-Aub is consistent with all the ion species arising from the fragmentation. Using nuclear magnetic resonance (NMR) experiments (Figures S4B–S4E), the structure of this compound was unambiguously elucidated to be L-Aub (Figure 4D). Overall, these data support the conclusion that PadQ catalyzes the transfer of the carbamoyl group of CP to L-Dab, resulting in L-Aub. To further test the substrate selectivity for PadQ, we instead tested L-ornithine, L-lysine, L-asparagine, and L-glutamine in the reaction mixture; however, no carbamoyltransferase activity was noted for any of these substrates in the reaction conditions we used (Figure S5).

L-Dab is a rare biosynthetic precursor; however, it is generated in halophilic bacteria as an intermediate on the route to ectoine, which serves as compatible solute and is used to counteract high osmolarity (Louis and Galinski, 1997). The first conversion in this three-enzyme pathway is the conversion of L-aspartate-4-semialdehyde to L-Dab, catalyzed by EctB (Louis and Galinski, 1997). A number of nonribosomal peptides have been found to use L-Dab as a direct building block for NRPSs, such as cupriachelin (Kreutzer et al., 2012), pyoverdines (Mossialos et al., 2002), syringopeptin (Scholz-Schroeder et al., 2003), and polymyxin (Choi et al., 2009). In polymyxin production, a single *ectB* gene is found in the polymyxin producer, outside of the polymyxin gene cluster, and heterologous production of polymyxin requires provision of either *ectB* or L-Dab itself. We observed that padanamide producer *Streptomyces* sp. RJA2928 has a typical *ectABC* gene cluster, as well as an additional orphan *ectB* gene outside the *pad* cluster. One or both of these EctBs is likely to provide endogenous L-Dab to the padanamide pathway (Figure 1C), where it is converted to L-Aub by PadQ (Figure 1C).

Once generated, L-Aub can be loaded on to the T domain of PadE, the terminal NRPS in the padanamide pathway. An intramolecular, TE-catalyzed cyclization, using the secondary amide nitrogen of L-Aub as a nucleophile, could release the nascent padanamide A, forming the unusual *S*-Aopc residue. By contrast, generation of padanamide B could instead arise from a similar cyclization of glutamine. Here, intramolecular nucleophilic attack of the N^{ε} of glutamine on the carbonyl of a tethered linear intermediate could release the nascent padanamide B, forming *S*-Apd (Figure 1C). The presence of different C-terminal

residues in padanamides A and B suggests a relaxed amino acid substrate specificity of the A domain of PadE.

DISCUSSION

In this work, we isolate the padanamide gene cluster. Using comprehensive gene inactivation experiments, we were able to delineate the boundaries of the padanamide gene cluster and propose a biosynthetic pathway for padanamides A and B.

Our work primarily addresses the origin of the unusual amino acid *S*-Aopc in padanamide A. Targeted inactivation of *padQ* abolishes production of padanamide A and promotes the production of padanamide B. These results unequivocally link *padQ* to padanamide A production and demonstrate that it is dispensable to production of padanamide B. Initial bioinformatic annotation suggested that *padQ* encodes an OTC. Nonetheless, we observe that conserved residues related to specificity for ornithine were altered in PadQ, suggesting that a unique substrate may be taken up by PadQ. We further pursued biochemical analysis of this enzyme, revealing that the substrate accepted by PadQ is L-Dab, which is carbamoylated by PadQ to give L-Aub. We propose that this unusual amino acid could then be activated as a thioester intermediate on the corresponding NRPS PadE and undergo an unusual intramolecular cyclization, with an amide nitrogen on the loaded amino acid acting as a nucleophile. Finally, we address the origin of the PadQ substrate L-Dab: the genome of *Streptomyces* sp. RJA2928 has two *ectB* homologs. Encoded enzymes putatively convert L-aspartate-4-semialdehyde to L-Dab, suggesting a reasonable route for provision of L-Dab to the padanamide pathway.

Prior to our study, two pathways might have been envisioned to account for production for the *S*-Aopc in padanamide A and the *S*-Apd in padanamide B. The first possible pathway is one where padanamide B is a metabolic intermediate, which is off-loaded from the NRPS and undergoes further enzyme-catalyzed conversions to generate padanamide A. In this case, unusual enzymatic conversions would be required. The second possible route to the padanamides is the loading of alternate aa to give padanamides A and B on a promiscuous NRPS. This scenario could reasonably give rise to padanamide B by loading and cyclization of glutamine to give *S*-Apd, an idea that is partially supported by the predicted A domain substrate specificity of PadE. However, the problem with this scenario was that there is no common amino acid likely to give padanamide A. Our work has demonstrated that the second scenario, i.e., a promiscuous NRPS, is used in padanamide biosynthesis and that the C-terminal residue of padanamide A derives from L-Aub. Based on our results, we support the hypothesis that PadE is a promiscuous enzyme, which has a preference for L-Aub but will also accept glutamine. Rather than being a metabolic intermediate or shunt product, padanamide B is then an alternate product of the same biosynthetic pathway.

Padanamides were reported almost simultaneously with the actinoramides (Nam et al., 2011), which were isolated from *Streptomyces* sp. CNQ-027, another strain derived from sediment of the Pacific Ocean. Actinoramide A, the major actinoramide, is identical to padanamide A. By contrast, the minor actinoramides are distinct: actinoramide B has another unique C-terminal residue, (*S*)-1-acetyl-3-aminopyrrolidin-2-one, and actinoramide C is a truncated structure compared to all other structures. Actinoramide C presumably derives from an unusual, premature off-loading of an actinoramide intermediate from the PKS related to PadD in the actinoramide pathway. By contrast, actinoramide B may derive from loading and cyclization of another amino acid. The (*S*)-1-acetyl-3-aminopyrrolidin-2-one residue in actinoramide B is the acetylated derivative of *S*-Aopc in padanamide A, and like the *S*-Aopc in padanamide A, this residue may derive from an ectoine biosynthetic intermediate. In the *ectA-B-C*-encoded pathway for ectoine, L-Dab is further processed to *N*-acetyl-L-2,4-diaminobutyrate by EctA (Figure 1C) (Louis and Galinski, 1997). Loading and cyclization of *N*-acetyl-L-2,4-diaminobutyrate by a PadE-homolog could generate the (*S*)-1-acetyl-3-aminopyrrolidin-2-one residue in actinoramide B in a similar fashion to that proposed for *S*-Aopc in padanamide A (Figure 1C). Thus, actinoramide B represents a possible third final product of the padanamide biosynthetic machinery. However, the actinoramide gene cluster in *Streptomyces* sp. CNQ-027 has not yet been identified. Whether this gene cluster is indeed highly related, and whether there is an exogenous ectoine pathway in *Streptomyces* sp. CNQ-027, remains to be established.

Our biosynthetic proposal suggests that padanamide A, padanamide B, and actinoramide B derive from an unusual, TE-catalyzed cyclization. Typically, TE-catalyzed chain release mechanisms include hydrolysis (intermolecular release) and macrocyclization (intramolecular release) (Du and Lou, 2010). For hydrolysis, water acts as the nucleophile to attack the carbonyl carbon of the acyl or peptidyl-*O*-TE and release a linear product. For macrocyclization, a hydroxyl or an amino group of the acyl or peptidyl chain acts as an internal nucleophile to attack the carbonyl carbon of the acyl or peptidyl-*O*-TE and a cyclic product is released. Typically, the nucleophile is in an amino acid earlier in the peptide chain (Du and Lou, 2010), but recently, a number of siderophores have been isolated that have either a C-terminal cyclic N^{δ}-OH-ornithine (Kodani et al., 2013; Kreutzer and Nett, 2012; Mazzei et al., 2012; Seyedsayamdost et al., 2011) or a C-terminal cyclic N^{ε}-OH-lysine (Hoshino et al., 2011). In these cases, the hydroxylamine of the C-terminal N^{δ}-OH-ornithine or N^{ε}-OH-lysine is thought to be used as a nucleophile to generate small, *N*-heterocycles and release linear peptides. The padanamides represent a further extension of this modification. Again, small heterocycles are generated through the use of a nucleophile from within the last appended amino acid. In padanamide biosynthesis, however, we speculate that an amide nucleophile—rather than an amine nucleophile—is used. Amide-based nucleophiles have been previously proposed in the biosynthesis of spirolactams in sanglifehrin A (Qu et al., 2011) but have not been proposed for release of linear peptides from NPRS machinery. Nonetheless, we favor this unusual use of an amide nucleophile, as it nicely accounts for provision of padanamide A, padanamide B, and actinoramide B. However, the details of the proposed TE-catalyzed cyclization remain to be investigated: neither full-length PadE nor truncated PadE-T-TE could be purified from heterologous expression systems (i.e., *E. coli*, *Rhodococcus*), including coexpression with PadL, the MbtH-like protein. Additionally, soluble PadE-A could not be obtained from either *E. coli* or *Rhodococcus*. Thus, both the complete inventory of amino acid preferences of PadE and the

biochemical details of the proposed TE-catalyzed cyclization remain to be established. Bioinformatic analysis of the PadE-TE does not reveal any obvious differences from other TE domains, except its relatively large size (Figures S3C and S3D).

Our work has investigated the unusual biosynthetic pathway to the highly modified linear peptides padanamides. We isolated the padanamide gene cluster, established the cluster boundaries, and carried out systematic gene inactivation experiments. Our work pinpointed *padQ*, encoding a carbamoyltransferase, which we biochemically characterized and demonstrated is critical to the outcome of the padanamide biosynthetic pathway. Furthermore, most intriguing is the enzyme PadE. PadE apparently catalyzes an intramolecular cyclization on multiple substrates using amide-based nucleophiles within the last appended amino acid, generating small, 5- or 6-membered ring heterocycles and releasing linear peptides. Although PadE itself could not yet be purified, PadE homologs mined from genomic data may represent exciting starting points for further biochemical investigation of this potentially unusual TE-catalyzed enzymology. Furthermore, *padE* or its engineered derivatives may find application in the combinatorial generation of new linear peptides with C-terminal heterocyclic residues.

SIGNIFICANCE

Secondary metabolites are the source for some of the most unusual structures in the chemistry literature. One biological strategy to generate novel molecules is to bring together enzymatic machinery from disparate sources. Here, we identify the biosynthetic machinery responsible for production of the padanamides. The padanamides possess C-terminal heterocycles that are unprecedented in the literature. Through our systematic gene inactivation, metabolic profiling, and biochemical experiments, we identify a three-step strategy to generate the unusual C-terminal heterocycle found in padanamide A. This strategy involves, first, interception of L-Dab, an intermediate from the ectoine pathway; second, *N*-carbamoylation of L-Dab; and third, putatively, a unique NRPS-catalyzed loading and cyclization mechanism. The interception of L-Dab by an *N*-carbamoyltransferase has not been previously observed, and the resulting product L-Aub has not been previously encountered in a natural product. Small changes in the amino acid sequence of the *N*-carbamoyltransferase PadQ, compared to more common OTCs, have apparently enabled this enzyme to react with L-Dab rather than ornithine. Yet, the presence or absence of this enzyme reroutes the outcome of the entire padanamide biosynthetic pathway. Our work provides molecular insights into the biosynthesis of two unprecedented heterocycles, and new genes identified from this pathway will enrich the toolbox for combinatorial biosynthesis.

EXPERIMENTAL PROCEDURES

See Supplemental Experimental Procedures for bacterial strains, plasmids, and culture conditions; genome scanning and in silico analysis; construction and screening of a *Streptomyces* sp. RJA2928 genomic cosmid library; generation of mutant strains of *Streptomyces* sp. RJA2928; cloning, expression, and purification of recombinant PadQ; and structure elucidation of enzymatically produced L-Aub.

Analysis of Padanamides Production in *Streptomyces* sp. RJA2928

Streptomyces sp. RJA2928 was cultured on 350 ml of solid MM1 agar (Williams et al., 2011). After 10 days of growth at 30°C, the agar was cut into small pieces, combined, and extracted with ethyl acetate. The extract was dried under vacuum and dissolved in 2 ml of acetonitrile. HPLC analysis was carried out with an Agilent 1260 HPLC apparatus with a diode array UV detector, fitted with an Agilent Zorbax 300SB-C18 column (5 μm, 4.6 mm × 250 mm). Elution was performed at 1 ml/min with a mobile-phase mixture consisting of a linear gradient of water and acetonitrile (85:15 to 60:40, 0 to 15 min; 60:100, 15 to 18 min; 0:100, 18 to 20 min), both of which contain 0.05% (v/v) trifluoroacetic acid. Padanamides A and B standard compounds were used as a control. LC-MS was performed under the same conditions on an Agilent 6120 Quadruple LC/MS system operated in both positive and negative ion electrospray modes.

In Vitro Enzyme Assay of PadQ

In a typical reaction, 1 μM of recombinant PadQ was reacted with 5 mM carbamoyl phosphate and 2 mM L-Dab at 30°C in 50 mM Tris-HCl buffer (pH 8.3) in a total volume of 1 ml. The reactions were initiated by the addition of carbamoyl phosphate, and quenched at different times by the addition of 1 ml cold methanol. As a control, PadQ that had been boiled for 10 min was used under identical reaction condition. After being passed through a 0.22-μm filter, reaction mixtures were spotted on TLC plates (Silica gel, Sigma-Aldrich) and developed with butanol/acetic acid/water (6:2:1). Visualization of the spots on TLC plate was performed by spraying with ninhydrin reagent (0.5% (w/v) ninhydrin and 1% (v/v) acetic acid in butanol). For LC-MS, 10 μl of each reaction solution was loaded on Agilent 6120 Quadruple LC/MS system (column: Econosil C8, 4.6 mm × 250 mm, 5 μm) with the following gradient: solvent A: water (0.05% (v/v) trifluoroacetic acid), solvent B: acetonitrile (0.05% (v/v) trifluoroacetic acid), flow rate: 0.6 ml/min, gradient: 100:0, 0 to 5 min; 90:10, 5 to 10 min, 0:100, 10 to 12 min; 100:0, 12 to 15 min. Protonated L-Dab has an *m/z* of 119, and protonated L-Aub has an *m/z* of 162.

ACCESSION NUMBERS

The sequence of the padanamide gene cluster is accessible at GenBank under accession number KC915040. Additionally, sequences of the ectoine gene cluster (accession number KF177908), stand-alone *ectB* gene (accession number KF177909), and two *fkbGHIJK* clusters (accession numbers KF177906 and KF177907) in *Streptomyces* sp. RJA2928 are accessible at GenBank.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.chembiol.2013.06.013.

ACKNOWLEDGMENTS

We are indebted to David Williams for padanamide chemical standards and helpful discussions. We thank Zhan Zhou for the phylogenetic analysis. We are grateful to Tomohiro Tamura for providing pTip-QC1. We thank the Natural Sciences and Engineering Council of Canada, the Canada Foundation for Innovation, and the British Columbia Knowledge Development Fund for funding.

Received: April 30, 2013
Revised: June 5, 2013
Accepted: June 25, 2013
Published: August 1, 2013

REFERENCES

Barkei, J.J., Kevany, B.M., Felnagle, E.A., and Thomas, M.G. (2009). Investigations into viomycin biosynthesis by using heterologous production in Streptomyces lividans. ChemBioChem *10*, 366–376.

Chan, Y.A., Boyne, M.T., 2nd, Podevels, A.M., Klimowicz, A.K., Handelsman, J., Kelleher, N.L., and Thomas, M.G. (2006). Hydroxymalonyl-acyl carrier

protein (ACP) and aminomalonyl-ACP are two additional type I polyketide synthase extender units. Proc. Natl. Acad. Sci. USA *103*, 14349–14354.

Choi, S.K., Park, S.Y., Kim, R., Kim, S.B., Lee, C.H., Kim, J.F., and Park, S.H. (2009). Identification of a polymyxin synthetase gene cluster of Paenibacillus polymyxa and heterologous expression of the gene in Bacillus subtilis. J. Bacteriol. *191*, 3350–3358.

Du, L., and Lou, L. (2010). PKS and NRPS release mechanisms. Nat. Prod. Rep. *27*, 255–278.

Du, Y., Derewacz, D.K., Deguire, S.M., Teske, J., Ravel, J., Sulikowski, G.A., and Bachmann, B.O. (2011). Biosynthesis of the Apoptolidins in Nocardiopsis sp. FU 40. Tetrahedron *67*, 6568–6575.

Felnagle, E.A., Barkei, J.J., Park, H., Podevels, A.M., McMahon, M.D., Drott, D.W., and Thomas, M.G. (2010). MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases. Biochemistry *49*, 8815–8817.

Filippova, E.V., Brunzelle, J.S., Cuff, M.E., Li, H., Joachimiak, A., and Anderson, W.F. (2011). Crystal structure of the novel PaiB transcriptional regulator from Geobacillus stearothermophilus. Proteins *79*, 2578–2582.

Fujimori, D.G., Hrvatin, S., Neumann, C.S., Strieker, M., Marahiel, M.A., and Walsh, C.T. (2007). Cloning and characterization of the biosynthetic gene cluster for kutznerides. Proc. Natl. Acad. Sci. USA *104*, 16498–16503.

Gerwick, W.H., and Moore, B.S. (2012). Lessons from the past and charting the future of marine natural products drug discovery and chemical biology. Chem. Biol. *19*, 85–98.

Gust, B., Challis, G.L., Fowler, K., Kieser, T., and Chater, K.F. (2003). PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc. Natl. Acad. Sci. USA *100*, 1541–1546.

Hibi, M., Kawashima, T., Sokolov, P.M., Smirnov, S.V., Kodera, T., Sugiyama, M., Shimizu, S., Yokozeki, K., and Ogawa, J. (2013). L-leucine 5-hydroxylase of Nostoc punctiforme is a novel type of Fe(II)/α-ketoglutarate-dependent dioxygenase that is useful as a biocatalyst. Appl. Microbiol. Biotechnol. *97*, 2467–2472.

Honjo, M., Nakayama, A., Fukazawa, K., Kawamura, K., Ando, K., Hori, M., and Furutani, Y. (1990). A novel Bacillus subtilis gene involved in negative control of sporulation and degradative-enzyme production. J. Bacteriol. *172*, 1783–1790.

Hoshino, Y., Chiba, K., Ishino, K., Fukai, T., Igarashi, Y., Yazawa, K., Mikami, Y., and Ishikawa, J. (2011). Identification of nocobactin NA biosynthetic gene clusters in Nocardia farcinica. J. Bacteriol. *193*, 441–448.

Hughes, C.C., and Fenical, W. (2010). Antibacterials from the sea. Chemistry *16*, 12512–12525.

Kevany, B.M., Rasko, D.A., and Thomas, M.G. (2009). Characterization of the complete zwittermicin A biosynthesis gene cluster from Bacillus cereus. Appl. Environ. Microbiol. *75*, 1144–1155.

Kodani, S., Kobayakawa, F., and Hidaki, M. (2013). Isolation and structure determination of new siderophore tsukubachelin B from Streptomyces sp. TM-74. Nat. Prod. Res. *27,*, 775–781. Published online June 19, 2012. http://dx.doi.org/10.1080/14786419.14782012.14698412.

Kreutzer, M.F., and Nett, M. (2012). Genomics-driven discovery of taiwachelin, a lipopeptide siderophore from Cupriavidus taiwanensis. Org. Biomol. Chem. *10*, 9338–9343.

Kreutzer, M.F., Kage, H., and Nett, M. (2012). Structure and biosynthetic assembly of cupriachelin, a photoreactive siderophore from the bioplastic producer Cupriavidus necator H16. J. Am. Chem. Soc. *134*, 5415–5422.

Linne, U., Doekel, S., and Marahiel, M.A. (2001). Portability of epimerization domain and role of peptidyl carrier protein on epimerization activity in nonribosomal peptide synthetases. Biochemistry *40*, 15824–15834.

Liu, G., Chater, K.F., Chandra, G., Niu, G., and Tan, H. (2013). Molecular regulation of antibiotic biosynthesis in streptomyces. Microbiol. Mol. Biol. Rev. *77*, 112–143.

Long, B., Tang, S., Chen, L., Qu, S., Chen, B., Liu, J., Maguire, A.R., Wang, Z., Liu, Y., Zhang, H., et al. (2013). Total synthesis of padanamides A and B. Chem. Commun. (Camb.) *49*, 2977–2979.

Louis, P., and Galinski, E.A. (1997). Characterization of genes for the biosynthesis of the compatible solute ectoine from Marinococcus halophilus and osmoregulated expression in Escherichia coli. Microbiology *143*, 1141–1149.

Ma, J., Wang, Z., Huang, H., Luo, M., Zuo, D., Wang, B., Sun, A., Cheng, Y.Q., Zhang, C., and Ju, J. (2011). Biosynthesis of himastatin: assembly line and characterization of three cytochrome P450 enzymes involved in the post-tailoring oxidative steps. Angew. Chem. Int. Ed. Engl. *50*, 7797–7802.

Marahiel, M.A., Stachelhaus, T., and Mootz, H.D. (1997). Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. Chem. Rev. *97*, 2651–2674.

Mazzei, E., Iorio, M., Maffioli, S.I., Sosio, M., and Donadio, S. (2012). Characterization of madurastatin C1, a novel siderophore from Actinomadura sp. J. Antibiot. *65*, 267–269.

Menzella, H.G., and Reeves, C.D. (2007). Combinatorial biosynthesis for drug development. Curr. Opin. Microbiol. *10*, 238–245.

Mossialos, D., Ochsner, U., Baysse, C., Chablain, P., Pirnay, J.P., Koedam, N., Budzikiewicz, H., Fernández, D.U., Schäfer, M., Ravel, J., and Cornelis, P. (2002). Identification of new, conserved, non-ribosomal peptide synthetases from fluorescent pseudomonads involved in the biosynthesis of the siderophore pyoverdine. Mol. Microbiol. *45*, 1673–1685.

Nam, S.J., Kauffman, C.A., Jensen, P.R., and Fenical, W. (2011). Isolation and Characterization of Actinoramides A-C, Highly Modified Peptides from a Marine Streptomyces sp. Tetrahedron *67*, 6707–6712.

Neumann, C.S., Jiang, W., Heemstra, J.R., Jr., Gontang, E.A., Kolter, R., and Walsh, C.T. (2012). Biosynthesis of piperazic acid via N5-hydroxy-ornithine in Kutzneria spp. 744. ChemBioChem *13*, 972–976.

Peng, C., Pu, J.Y., Song, L.Q., Jian, X.H., Tang, M.C., and Tang, G.L. (2012). Hijacking a hydroxyethyl unit from a central metabolic ketose into a nonribosomal peptide assembly line. Proc. Natl. Acad. Sci. USA *109*, 8540–8545.

Qu, X., Jiang, N., Xu, F., Shao, L., Tang, G., Wilkinson, B., and Liu, W. (2011). Cloning, sequencing and characterization of the biosynthetic gene cluster of sanglifehrin A, a potent cyclophilin inhibitor. Mol. Biosyst. *7*, 852–861.

Sankaranarayanan, R., Cherney, M.M., Cherney, L.T., Garen, C.R., Moradian, F., and James, M.N. (2008). The crystal structures of ornithine carbamoyltransferase from Mycobacterium tuberculosis and its ternary complex with carbamoyl phosphate and L-norvaline reveal the enzyme's catalytic mechanism. J. Mol. Biol. *375*, 1052–1063.

Scholz-Schroeder, B.K., Soule, J.D., and Gross, D.C. (2003). The sypA, sypS, and sypC synthetase genes encode twenty-two modules involved in the nonribosomal peptide synthesis of syringopeptin by Pseudomonas syringae pv. syringae B301D. Mol. Plant Microbe Interact. *16*, 271–280.

Seyedsayamdost, M.R., Traxler, M.F., Zheng, S.L., Kolter, R., and Clardy, J. (2011). Structure and biosynthesis of amychelin, an unusual mixed-ligand siderophore from Amycolatopsis sp. AA4. J. Am. Chem. Soc. *133*, 11434–11437.

Sun, Y., Hong, H., Gillies, F., Spencer, J.B., and Leadlay, P.F. (2008). Glyceryl-S-acyl carrier protein as an intermediate in the biosynthesis of tetronate antibiotics. ChemBioChem *9*, 150–156.

Williams, D.E., Dalisay, D.S., Patrick, B.O., Matainaho, T., Andrusiak, K., Deshpande, R., Myers, C.L., Piotrowski, J.S., Boone, C., Yoshida, M., and Andersen, R.J. (2011). Padanamides A and B, highly modified linear tetrapeptides produced in culture by a Streptomyces sp. isolated from a marine sediment. Org. Lett. *13*, 3936–3939.

Williams, D.E., Dalisay, D.S., Li, F., Amphlett, J., Maneerat, W., Chavez, M.A., Wang, Y.A., Matainaho, T., Yu, W., Brown, P.J., et al. (2013). Nahuoic acid A produced by a Streptomyces sp. isolated from a marine sediment is a selective SAM-competitive inhibitor of the histone methyltransferase SETD8. Org. Lett. *15*, 414–417.

Wu, K., Chung, L., Revill, W.P., Katz, L., and Reeves, C.D. (2000). The FK520 gene cluster of Streptomyces hygroscopicus var. ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. Gene *251*, 81–90.

Zhao, C., Ju, J., Christenson, S.D., Smith, W.C., Song, D., Zhou, X., Shen, B., and Deng, Z. (2006). Utilization of the methoxymalonyl-acyl carrier protein biosynthesis locus for cloning the oxazolomycin biosynthetic gene cluster from Streptomyces albus JA3453. J. Bacteriol. *188*, 4142–4147.