JOURNAL OF NATURAL PRODUCTS

Article

Cite This: J. Nat. Prod. XXXX, XXX, XXX−XXX

pubs.acs.org/jnp

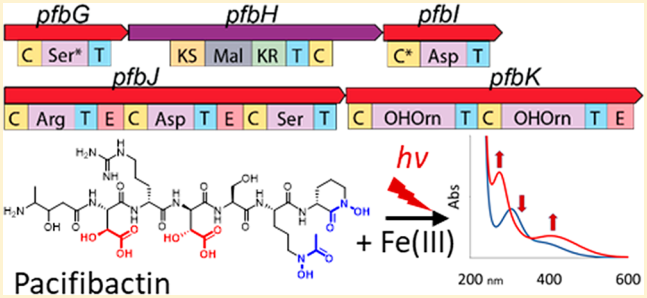
Ambiguity of NRPS Structure Predictions: Four Bidentate Chelating Groups in the Siderophore Pacifibactin

Clifford D. Hardy and Alison Butler*

Department of Chemistry & Biochemistry, University of California Santa Barbara, Santa Barbara, California 93106-9510, United States

Supporting Information

ABSTRACT: Identified through a bioinformatics approach, a nonribosomal peptide synthetase gene cluster in *Alcanivorax pacificus* encodes the biosynthesis of the new siderophore pacifibactin. The structure of pacifibactin differs markedly from the bioinformatic prediction and contains four bidentate metal chelation sites, atypical for siderophores. Genome mining and structural characterization of pacifibactin is reported herein, as well as characterization of pacifibactin variants accessible due to a lack of adenylation domain fidelity during biosynthesis. A spectrophotometric titration of pacifibactin with Fe(III) and ^{13}C NMR spectroscopy of the Ga(III)-pacifibactin complex establish 1:1 metal:pacifibactin coordination and reveal which of the bidentate binding groups are coordinated to the metal. The photoreaction of Fe(III)-pacifibactin, resulting from Fe(III) coordination of the *β*-hydroxyaspartic acid ligands, is reported.



Automated genome mining tools enable high-throughput scanning of bacterial genomes for gene clusters encoding biosynthetic machinery.[1,2] Natural products produced by nonribosomal peptide synthetases (NRPSs) are particularly amenable to discovery through bioinformatics approaches. NRPSs are multimodular proteins that work in an assembly line fashion to synthesize a peptidic natural product.[3−6] A typical NRPS consists of at least one adenylation domain, a thiolation domain, and a condensation domain. This triad of catalytic domains governs selection and activation of a substrate (generally an amino acid), anchors the activated substrate to the NRPS assembly, and incorporates the amino acid into the synthesized product through peptide bond formation. Domains responsible for tailoring reactions may also be present that carry out tailoring reactions such as epimerization and cyclization.[4] The organization of an NRPS into distinct domains with predictable functions and amino acid substrates is a key feature driving software such as antiSMASH and PRISM, which utilize sequence analysis to identify NRPS-encoding gene clusters and predict their functionality.[7,8]

Iron plays a central role in many enzymatic processes and protein functions across all organisms and, thus, is a required nutrient for virtually all life. The requirement for iron complicates bacterial growth, as readily bioavailable iron is limited by the low solubility of Fe(III) in aerobic environments. In response to low iron stress, many bacteria synthesize siderophores, small-molecule Fe(III) chelators produced by a bacterium to solubilize Fe(III) in the environment. Fe(III)-siderophore complexes are recognized by an outer membrane receptor protein and taken up by the cell. Many siderophores are synthesized by NRPSs, and genome mining has enabled the prediction and discovery of many new siderophore structures.[9−13] NRPS-directed siderophore biosynthesis often employs extensive tailoring of both the amino acid substrates and the assembled product to yield Fe(III)-chelating functional groups. These tailoring reactions may be carried out by standalone proteins not accounted for in the commonly utilized NRPS analysis tools.

The genomes of many species within the obligate hydrocarbon-degrading microbial genus *Alcanivorax* have been sequenced, yet siderophore biosynthesis has only been identified in *Alcanivorax borkumensis*.[14] This abundance of sequenced *Alcanivorax* genomes enables a genome mining approach for further characterization of siderophore production in the genus. To this end, fully sequenced genomes of *Alcanivorax* species not known to produce siderophores were screened for potential NRPS gene clusters encoding siderophore biosynthesis. Of the analyzed genomes, only *Alcanivorax pacificus* contains a candidate NRPS gene cluster for siderophore production.

We report herein that *A. pacificus* contains a biosynthetic gene cluster that encodes the synthesis of the previously unknown siderophore pacifibactin. Structural characterization of pacifibactin reveals limitations of current automated genome mining approaches, highlighting several tailoring steps as yet undetectable through existing software. Structural variants of pacifibactin resulting from substrate selection promiscuity by

Received: December 19, 2018

DOI: 10.1021/acs.jnatprod.8b01073
J. Nat. Prod. XXXX, XXX, XXX−XXX

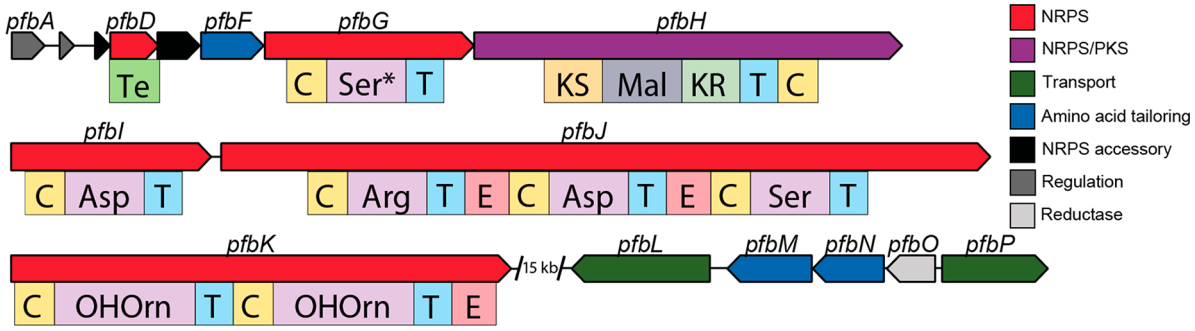


Figure 1. Pacifibactin biosynthetic gene cluster, annotated with NRPS and PKS domains as identified by antiSMASH and PRISM. Adenylation domains (lavender boxes) are labeled with their substrate specificity prediction. The adenylation domain of PfbG is predicted to incorporate Ser; however structural characterization of pacifibactin establishes the presence of Ala. Te, thioesterase domain; C, condensation domain; T, thiolation domain; KS, ketosynthase domain; Mal, malonic acid (substrate specificity); KR, ketoreductase domain; E, epimerization domain.

an adenylation domain are also described. Pacifibactin is unique among siderophores in that it contains two hydroxamic acid and two β-hydroxyaspartic acid functional groups. Four potential bidentate Fe(III) binding groups are rarely observed in siderophore structures and were not predicted from the genome mining analysis. The coordination chemistry of Fe(III)-pacifibactin is reported, as is the photoreactivity of the Fe(III)-siderophore complex due to coordination by β-hydroxyaspartic acid ligands.[15]

RESULTS AND DISCUSSION

Analysis of the *pfb* Gene Cluster for Siderophore Biosynthesis. The genomes of eight fully sequenced *Alcanivorax* species available through NCBI were analyzed using the bioinformatics software tools antiSMASH and PRISM,[7,8] which identify and annotate putative nonribosomal peptide synthetase genes. The genome of *A. pacificus* W11-5T, a species isolated from a bacterial consortium found within seafloor sediment in the Pacific Ocean,[16] contains a putative siderophore biosynthetic gene cluster centered around three NRPS-encoding genes and one polyketide synthase (PKS)-encoding gene, identified by both antiSMASH and PRISM (Figure 1). Putative genes involved in siderophore transport, Fe(III)-siderophore reduction, and amino acid tailoring are also present within the cluster (full annotation, Table S1). A siderophore structure comprising L-Ser, malonic acid (PKS), L-Asp, D-Arg, D-Asp, L-Ser, L-OH-Orn, and D-OH-Orn is predicted from the adenylation domain specificity and the location of epimerization domains within the NPRS/PKS assembly line (Figure 1). The presence of genes predicted to encode a TonB-dependent receptor protein and siderophore-iron reductase (Table S1), and the incorporation of aspartic acid and ornithine, which are commonly functionalized as Fe(III) binding groups, further supported siderophore production. Additionally, the genome of *A. pacificus* contains no other NRPS-encoding gene clusters, nor does it contain any biosynthetic gene clusters indicating currently known NRPS-independent siderophore biosynthesis.

Isolation and Structural Characterization of Pacifibactin. To induce siderophore production, *A. pacificus* was grown in an iron-deficient artificial seawater medium. *A. pacificus* grows readily under iron-starvation conditions, and aliquots of the culture tested positive in the liquid chrome azurol S (CAS) assay for strong Fe(III)-binding ligands.[17] UPLC-ESIMS analysis of the supernatant extract revealed a candidate compound with protonated molecule masses of m/z 923.41 $[M + H]^+$ and m/z 462.20 $[M + 2H]^{2+}$. The putative siderophore, named herein as pacifibactin, was purified by semipreparative RP-HPLC. HR-ESIMS of purified pacifibactin detects a protonated molecule of m/z 923.4081 (Figure S1). MS/MS peptide b/y fragmentation reveals the constituent amino acids of pacifibactin (Figure 2, Figure S2), with an

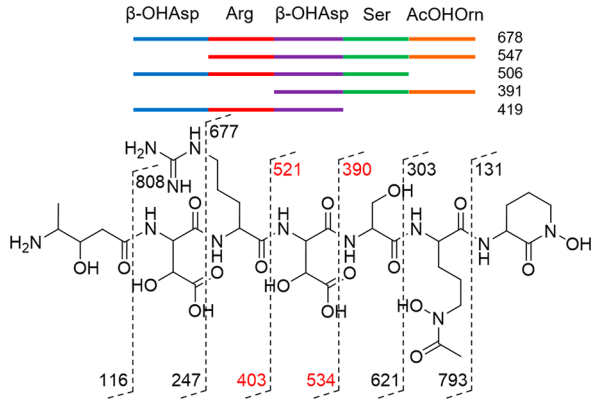


Figure 2. Structure of pacifibactin, with b/y peptide fragment masses (fragments with m/z in red were not observed). Observed internal fragments are shown above the structure.

amino acid sequence from the carboxylate terminus of cyclized N5-OH-Orn, N5-acetyl-N5-OH-Orn, Ser, β-OH-Asp, Arg, and β-OH-Asp. The two β-OH-Asp residues, N5-acetyl-N5-OH-Orn residue, and cyclized N5-OH-Orn residue yield in total four bidentate metal binding sites. The presence of four sites stands in contrast to almost all known siderophores, which generally contain no more than three bidentate metal binding sites.

MS/MS fragmentation suggested a mass for the N-terminal amino acid that did not match any amino acid previously found in peptidic siderophores. Analysis of the NRPS domains within the gene cluster suggested that Ser and malonic acid (presumably incorporated by the PKS enzyme PfbH) should be incorporated at the N-terminus of pacifibactin. Mixed NRPS-PKS siderophore biosyntheses involve a decarboxylative Claisen condensation between the carboxylate functional groups of an NRPS-recruited amino acid and a PKS-recruited substrate,[12,18] and indeed the mass of the N-terminal pacifibactin amino acid determined by MS/MS fractionation is consistent with a decarboxylative Claisen condensation of

DOI: 10.1021/acs.jnatprod.8b01073
J. Nat. Prod. XXXX, XXX, XXX−XXX

Ser and malonate followed by stepwise reduction of the β-keto group to an alkyl moiety.

Complete assignment of each ^{1}H and ^{13}C chemical shift of pacifibactin was accomplished through COSY, HSQC, and HMBC NMR techniques (Table 1, Figure 3). The NMR-

Table 1. NMR Spectroscopic Data (500 MHz, $(CD_3)_2SO$) for Pacifibactin (20 mg)

	δ_C, type	δ_H (J in Hz)	COSY	HMBC
1	164.5, C			
2	49.6, CH	4.30, m	3, N1	1, 3, 6
3	20.1, CH_2	1.85, m; 1.59, m	2, 4	2
4	27.5, CH_2	1.88, m	3, 5	2, 3, 5
5	51.2, CH_2	3.48, m	4	3, 4
N1		8.12, d (8.0)	2	2, 6
6	171.0, C			
7	52.9, CH	4.24, m	8, 9, N2	6, 8
8	29.1, CH_2	1.52, m	7, 9	7, 9, 10
9	23.0, CH_2	1.57, m	8, 10	8, 10
10	46.7, CH_2	3.46, m	9	8, 11
11	170.3, C			
12	20.4, CH_3	1.97, s		11
N2		8.22, d (7.5)	7	7, 13
13	170.0, C			14, 15
14	54.6, CH	4.38, m	15, N3	13, 15, 16
15	61.9, CH_2	3.48, m; 3.65, m	14	13, 14
N3		7.62, d (7.5)	14	14, 16
16	168.9, C			
17	55.5, CH	4.76, t (2.6)	18, N4	16, 18, 19, 20
18	70.3, CH	4.56, d (2.5)	17	16, 17, 19
19	173.0, C			17, 18
N4		8.18, d (8.8)	17	20
20	171.4, C			
21	51.9, CH	4.47, m	22, N6	20, 22
22	29.6, CH_2	1.74, m; 1.55, m	21	20, 21, 23
23	24.5, CH_2	1.44, m; 1.48, m	24	21, 22, 24
24	40.3, CH_2	3.07, m	23, N5	22, 23
N5		7.47, t (5.6)	24	24, 25
25	156.7, C			
N6		7.87, d (8.0)	21	21, 26
26	168.8, C			
27	55.5, CH	4.74, t (2.6)	28, N7	26, 28, 29, 30
28	70.1, CH	4.54, d (2.5)	27	26, 27, 29
29	172.8, C			
N7		8.03, d (9.0)	27	27, 30
30	170.2, C			
31	38.9, CH_2	2.32, dd (14.27, 5.49) 2.43, dd (14.33, 8.35)	32	30, 32, 33
32	67.6, CH	4.04, m	31, 33	31, 33, 34
33	50.0, CH	3.19, m	32, 34	34
34	11.9, CH_3	1.09, d	33	32, 33

supported structure of pacifibactin predicts an exact mass of m/z 923.4065 $[C_{34}H_{59}N_{12}O_{18}]^+$, which is within 2 ppm of the measured mass of m/z 923.4081 (Figure S1). NMR analysis unambiguously assigns the N-terminal amino acid as a γ-amino acid methylated at the γ-carbon and hydroxylated at the β-carbon. This unusual amino acid is consistent with a Claisen condensation between Ala (presumably incorporated by PfbG) and malonate (presumably incorporated by PfbH) followed by reduction of the β-keto group to a hydroxy group, instead of the predicted condensation of Ser and malonate. Analysis of the PKS-encoding gene identifies a ketoreductase domain that could carry out this reduction to a hydroxyl functionality (Figure 1).

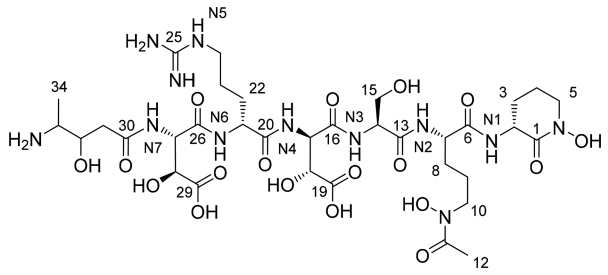


Figure 3. Structure of pacifibactin.

Comparison of the structure of pacifibactin to the bioinformatic prediction reveals several surprises. The adenylation domain of PfbG is predicted by both antiSMASH and PRISM to incorporate Ser, yet Ala appears to be incorporated into pacifibactin at the N-terminus instead. The ensemble algorithm SANDPUMA, which applies several different adenylation domain specificity predictors and has outperformed any individual method in accuracy,[19] also predicts Ser incorporation by PfbG. The incorporation of Ala highlights the need for continuous refinement of adenylation domain predictor tools as more experimental data are generated.

Perhaps even more intriguing in the characterized structure is the presence of four bidentate binding groups in pacifibactin, as siderophores generally have three bidentate binding groups to satisfy hexadentate coordination to Fe(III). The bioinformatic analysis predicts incorporation of two Asp and two OH-Orn by the NRPS assembly line; however neither Asp nor OH-Orn acts as a bidentate metal chelator without further tailoring. While these residues were expected to provide the metal chelation sites typical of siderophores, unmodified Asp and OH-Orn residues have been identified in other siderophore structures;[20−23] thus not all four residues were expected to be modified as bidentate metal chelators. An expanded manual analysis of the *pfb* gene cluster helps explain this unexpected result. The amino acid sequence of PfbN exhibits homology to N-acetyl transferases implicated in the acetylation of Lys and Orn in other siderophores.[24,25] N5-acetyl-N5-OH-Orn is known to be synthesized prior to adenylation;[25] thus further development of predictor tools is needed to distinguish this substrate from OH-Orn. The mechanism of N5-OH-Orn cyclization is as yet unknown; however in consideration of the mechanism of NRPS peptide bond formation cyclization must occur after incorporation of OH-Orn, making prediction through bioinformatics tools difficult.

Both Asp residues in pacifibactin are hydroxylated at the β-carbon to form β-OH-Asp. Accordingly, PfbF exhibits homology to the Fe(II)/α-ketoglutarate-dependent Asp β-hydroxylase SyrP of *Pseudomonas syringae*.[26] Additionally, a domain exhibiting homology with SyrP is found in the PKS protein PfbH. Notably, neither antiSMASH nor PRISM picked up this domain as an Asp β-hydroxylase. Putative Asp β-hydroxylating domains have been observed in other siderophore gene clusters.[9,12,23,27] In the biosynthetic gene clusters of serobactin and cupriachelin, both Asp β-hydroxylating NRPS domains and discrete Asp β-hydroxylating enzymes have

DOI: 10.1021/acs.jnatprod.8b01073
J. Nat. Prod. XXXX, XXX, XXX−XXX

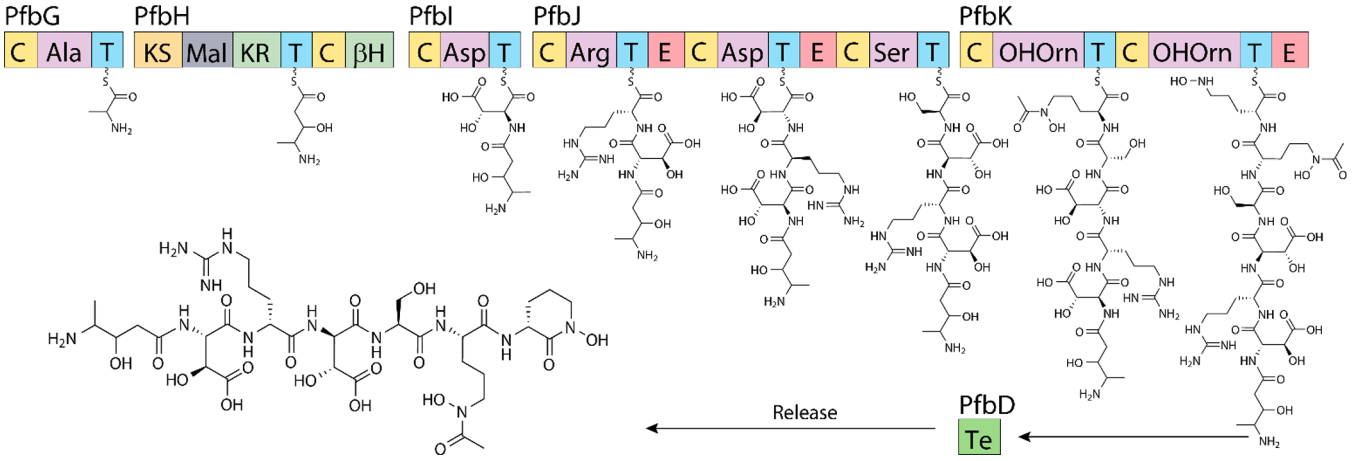


Figure 4. Proposed biosynthesis of pacifibactin, based on bioinformatics analysis of the putative biosynthetic gene cluster.

been identified.[9,27] Each of these siderophores, like pacifibactin, contain two *β*-OH-Asp residues. *β*-Hydroxylation of Asp occurs after tethering to the assembly line; thus adenylation domain analysis alone is insufficient to predict incorporation of *β*-OH-Asp residues in an NRPS product. However, development of future bioinformatics tools that pair adenylation domain analysis with identification of these putative *β*-hydroxylases could successfully predict incorporation of *β*-OH-Asp.

The *pfb* gene cluster contains three epimerization domains within the NRPS assembly line (Figure 4). Given the placement of the epimerization domains, incorporation of (starting from the N-terminus) L-Ala, L-*β*-OH-Asp, D-Arg, D-*β*-OH-Asp, L-Ser, L-N5-acetyl-N5-OH-Orn, and D-N5-OH-Orn is expected. Hydrolysis of pacifibactin with HCl or HI (for reductive hydrolysis to yield unfunctionalized ornithine) and derivatization of the resultant hydrolysates with Marfey's reagent (1-fluoro-2,4-dinitrophenyl-5-L-alanine amide, FDAA) identifies L-*threo*-*β*-OH-Asp, D-Arg, D-*threo*-*β*-OH-Asp, L-Ser, L-Orn, and D-Orn in the hydrolysate through co-injections with the corresponding derivatized amino acid standards (Figures S9, S10). The identified amino acids are consistent with the bioinformatic prediction. Co-injections with D,L-*threo*-*β*-OH-Asp confirmed the presence of the *threo* diastereomers in pacifibactin. Moreover, co-injections with a D,L-*erythro*-*β*-OH-Asp establish that pacifibactin does not incorporate the *erythro* diastereomers (Figure S9). The incorporation of L-Ala was not confirmed, as the PKS-governed reaction forms a carbon−carbon bond between Ala and malonate that is not broken under the hydrolytic conditions.

Variations in Biosynthetic Substrate Incorporation. The unexpected incorporation of Ala into pacifibactin at the N-terminal position, despite a consensus bioinformatic prediction of Ser incorporation, calls into question the substrate specificity of the corresponding adenylation domain. To probe whether L-Ser could be incorporated into pacifibactin in place of L-Ala, *A. pacificus* was cultured in media supplemented with L-Ser. ESIMS analysis of the resultant culture extracts identified a coeluting compound of *m*/*z* 470.2 (*z* = 2), 16 amu higher than the mass of pacifibactin. This mass increase matches the expected mass increase of Ser incorporation over Ala incorporation (Figure S11). The fragmentation pattern observed by ESIMS/MS confirms that this mass corresponds to a variant of pacifibactin, with the mass increase localized to the N-terminal amino acid (Figure 5, Figure S12).

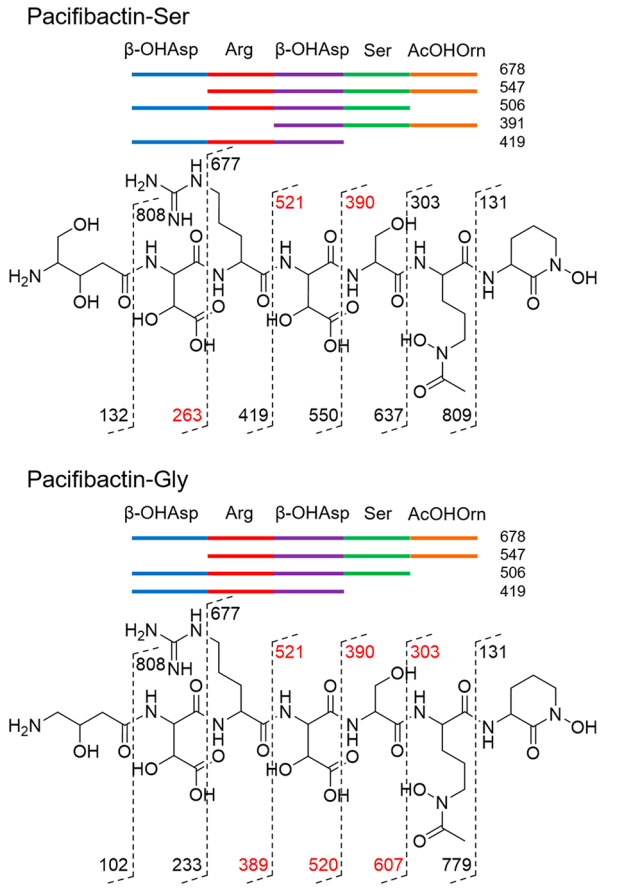


Figure 5. MS/MS b/y fragmentation patterns for pacifibactin Ser (470.2 *m*/*z* $[M + 2H]^{2+}$) and Gly (455.2 *m*/*z* $[M + 2H]^{2+}$) variants (fragments with *m*/*z* in red were not observed). Observed internal fragments are shown above the structure.

To further test substrate variation at the pacifibactin N-terminus, *A. pacificus* was cultured with supplements of Gly, L-Thr, and L-Val. While no incorporation of L-Thr or L-Val was observed, *A. pacificus* produced a compound of *m*/*z* 455.2 (*z* = 2) when supplemented with Gly, 14 amu lower than the mass of pacifibactin (Figure S11). The difference is again localized

DOI: 10.1021/acs.jnatprod.8b01073
J. Nat. Prod. XXXX, XXX, XXX−XXX

to the N-terminal amino acid, based on MS/MS fragmentation (Figure 5, Figure S13). The interchange between incorporation of L-Ala, L-Ser, and Gly in pacifibactin biosynthesis is reminiscent of the moanachelin siderophores produced by *Vibrio* sp. NT1, which incorporate either Ala or Gly as the third amino acid residue.[28] The N-terminus of pacifibactin is presumably biosynthesized by PfbG based on bioinformatics analysis; thus further investigation into structural differences between the adenylation domain of PfbG and known Ser-incorporating domains could deepen the understanding of adenylation domain selectivity.

Coordination Chemistry of Pacifibactin. The structure of pacifibactin contains four bidentate metal binding sites, an unusual feature for siderophores only observed among some desferrioxamines and malleobactin D.[20,29] The presence of four bidentate ligands in the structure of pacifibactin calls into question which groups coordinate Fe(III). Titrating a buffered solution of apo pacifibactin with Fe(III) under neutral pH conditions indicates that despite the extra binding group, pacifibactin coordinates Fe(III) in a 1:1 ratio, as observed for hexadentate siderophores (Figure 6). The UV−visible absorption spectrum of Fe(III)-pacifibactin features a peak at 305 nm characteristic of Fe(III)-α-hydroxycarboxylate ligand-to-metal charge transfer (LMCT) and a broad shoulder at 400 nm indicative of Fe(III)-hydroxamate LMCT (Figure 6). The spectrum resembles that of Fe(III)-serobactin, a siderophore with two β-OH-Asp residues and one cyclized N5-OH-Orn residue for Fe(III) coordination.[9] This resemblance suggests that pacifibactin coordinates one Fe(III) through both β-OH-Asp residues and either the cyclized N5-OH-Orn or the N5-acetyl-N5-OH-Orn residue.

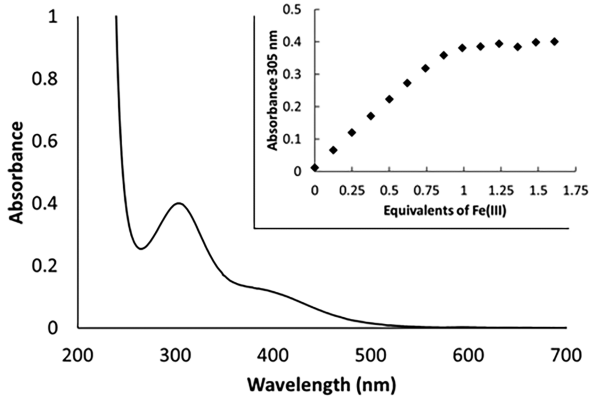


Figure 6. UV−visible absorption spectrum of 0.1 mM Fe(III)-pacifibactin in 100 mM MOPS buffer pH 7.1. Inset: Spectrophotometric titration of apo pacifibactin with Fe(III). Apo pacifibactin (0.1 mM, in 100 mM MOPS pH 7.1) was titrated with Fe(III) stock (2.14 mM in 40 mM HNO_3). A break point in absorbance is observed at 0.9 equiv of Fe(III), suggesting a 1:1 coordination mode.

To identify the binding groups involved in Fe(III) coordination in pacifibactin, the Ga(III)-pacifibactin complex was prepared as an NMR-compatible mimic to Fe(III)-pacifibactin. After Ga(III) complexation, ^{13}C chemical shifts in both β-OH-Asp residues and the N5-acetyl-N5-OH-Orn residue showed significant changes relative to apo pacifibactin, while no significant differences were observed in the ^{13}C chemical shifts of the cyclized N5-hydroxyornithine residue (Table 2). The observed chemical shifts indicate that Ga(III) likely coordinates to pacifibactin through both β-OH-Asp residues and N5-acetyl-N5-OH-Orn, but not the cyclized N5-OH-Orn residue (Figure 7).

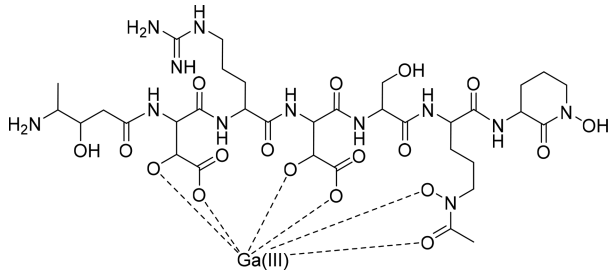


Figure 7. Proposed coordination mode of Ga(III)-pacifibactin. Ga(III) is bound by both β-OH-Asp residues and the N5-acetyl-N5-OH-Orn residue, while the cyclized N5-OH-Orn residue remains an open coordination site.

Fe(III) complexes of α-hydroxycarboxylate siderophores, including β-OH-Asp siderophores, are photoreactive.[12,15,27,30−34] To probe photoreactivity, Fe(III)-pacifibactin was photolyzed with a 450 W mercury-arc lamp and monitored by UV−visible spectrophotometry. Through 8 h of continuous photolysis, clear shifts in the UV−visible spectrum are observable (Figure 8). Upon photolysis, the Fe(III)-hydroxamate absorbance band around 400 nm nearly doubles in intensity and red shifts slightly, while two near-isosbestic points are present, suggesting photolysis of Fe(III)-pacifibactin initially yields one Fe(III)-coordinating photoproduct. Continued photolysis after 8 h leads to a loss of isosbestic points and eventual elimination of the Fe(III)-α-hydroxycarboxylate charge transfer band around 300 nm (Figure S21). After photolysis, UPLC-ESIMS/MS analysis of the reaction mixture identifies a photoproduct resulting from photooxidative cleavage of the peptide backbone at the C-16 to C-19 β-

Table 2. ^{13}C Chemical Shifts of Ga(III)-Pacifibactin (in D_2O) Compared to Apo Pacifibactin[a]

		CO (ppm)	Cα (ppm)	Cβ (ppm)	Cγ (ppm)	Cδ (ppm)	C acetyl (ppm)
CyOHOrn	Apo	166.41	50.32	26.56	19.93	51.57	
	Ga(III)	166.18	49.97	26.83	19.83	51.65	
AcOHOrn	Apo	173.49	53.62	27.91	22.36	**47.20**	**173.84**
	Ga(III)	173.78	54.28	26.11	22.99	**49.87**	**164.61**
βOHAsp	Apo	**170.37**	**55.76**	**70.27**	**174.21**		
(C#16−19)	Ga(III)	**175.94**	**61.23**	**73.18**	**182.65**		
βOHAsp	Apo	**171.56**	**55.66**	**70.07**	**174.09**		
(C#26−29)	Ga(III)	**173.16**	**59.09**	**74.10**	**180.05**		

[a]Bolded resonances indicate a significant change in the ^{13}C chemical shift after Ga(III) coordination.

DOI: 10.1021/acs.jnatprod.8b01073
J. Nat. Prod. XXXX, XXX, XXX−XXX

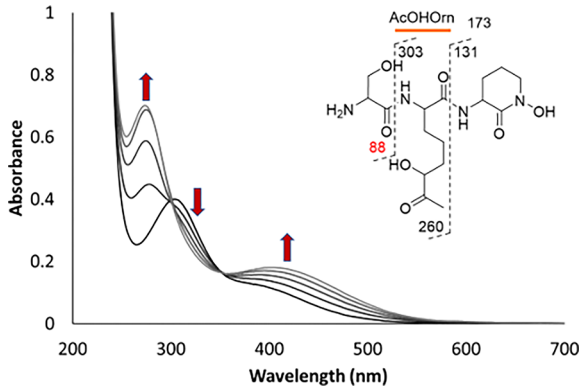


Figure 8. UV−visible absorbance spectra of 0.1 mM Fe(III)-pacifibactin in 100 mM MOPS buffer pH 7.1 subjected to 8 h of continuous photolysis with a 450 W UV mercury-arc lamp, collected from 220 to 700 nm. Scans are taken at 2 h time points, and lighter gray represents increased time. Arrows indicate increases and decreases in absorbance over time. Inset: Proposed structure of photoproduct detected by UPLC-ESIMS/MS (390.2 m/z, z = 1), with MS/MS b/y fragmentation pattern. The observed internal fragment is shown by the orange bar above the structure.

OH-Asp residue (Figure 8), consistent with photoproducts of other β-OH-Asp siderophores.[12,15,27] Further investigations are focused on elucidating the mechanism of pacifibactin photooxidation, including identification of intermediates in the reaction.

CONCLUSIONS

Characterization of pacifibactin revealed a highly unusual siderophore with marked structural differences from the predictions inferred through initial bioinformatic analysis. Two β-OH-Asp functional groups combined with two Orn-derived hydroxamate functional groups yield four bidentate metal binding sites, a very uncommon feature in siderophores. UV−visible spectrophotometry of Fe(III)-pacifibactin and ^{13}C NMR spectroscopy of Ga(III)-pacifibactin are consistent with metal coordination by both β-OH-Asp residues and the N5-acetyl-N5-OH-Orn residue, thus leaving a free cyclized N5-OH-Orn residue. Owing to the coordination of Fe(III) by β-OH-Asp, UV irradiation of Fe(III)-pacifibactin triggers a photooxidative breakdown of the ligand, as evinced through shifts in UV−visible absorbance.

The structure of pacifibactin highlights several limitations to existing NRPS cluster analysis programs. While programs such as antiSMASH identify and annotate NRPS domains within a gene cluster, both noncanonical NRPS domains and tailoring enzymes acting externally from the NRPS proteins elude identification. Initial bioinformatic analysis predicted incorporation of two Asp residues into pacifibactin; however a more exhaustive manual analysis of the gene cluster identified an NRPS domain and a standalone enzyme that are each likely responsible for aspartic acid β-hydroxylation. Similarly, antiSMASH predicted incorporation of two OH-Orn residues, while manual analysis identified an N-acetyl transferase likely responsible for N5-acetylation of N5-OH-Orn. This extended manual analysis of the gene cluster allowed a refinement of the initial structure prediction of the corresponding natural product. Structural characterization of pacifibactin then confirmed the accuracy of the refinements, as pacifibactin indeed incorporates two β-OH-Asp residues and one N5-acetyl-N5-OH-Orn. The characterization of pacifibactin makes clear that automated tools such as antiSMASH remain limited in their ability to generate accurate structure predictions of NRPS-synthesized siderophores, yet also details a path to improving these predictions. While antiSMASH by design takes a conservative approach to the chemistry prediction of NRPS and PKS products,[7] we propose that the identification of tailoring enzymes and domains associated with siderophore production could be incorporated into automated genome mining tools. The accurate prediction of OH-Orn tailoring enzymes and Asp β-hydroxylases within NRPS gene clusters could not only improve the accuracy of structure predictions but also aid in distinguishing NRPS clusters as siderophore producers.

EXPERIMENTAL SECTION

General Experimental Procedures. Optical rotation was measured on a Rudolph Autopol III polarimeter with a 50 mm microcell (1.2 mL). UV−visible absorbance was measured on an Agilent Cary 300 UV Vis spectrophotometer using 3 mL quartz cuvettes. NMR spectroscopy was carried out on 500 MHz (^1H, ^{13}C) and 600 MHz (COSY, HSQC, HMBC) Varian Unity Inova spectrometers. Chemical shifts were referenced through residual solvent peaks [^1H (DMSO-d_6) 2.50 ppm, ^{13}C (DMSO-d_6) 39.51 ppm] or an external reference for samples dissolved in D_2O [^1H, ^{13}C (TMS) 0.0 ppm]. Mass spectrometry analysis was carried out on a Waters Xevo G2-XS QTof with positive mode electrospray ionization coupled to an ACQUITY UPLC H-Class system with a Waters BEH C18 column. Culture extracts were analyzed with a linear gradient of 0% to 30% CH_3CN (0.1% formic acid) in ddH_2O (0.1% formic acid) over 10 min. For MS/MS analysis, a collision energy profile of 20, 25, and 30 kEV was employed.

Genome Mining and Gene Cluster Annotation. The genome of *Alcanivorax pacificus* W11-5[35] was accessed through NCBI and analyzed with the NRPS cluster-predicting software PRISM and antiSMASH.[7,8] Genes within the pacifibactin cluster and their corresponding amino acid sequences were analyzed using BLAST and the PFAM database to predict function of proteins encoded by the cluster.

Bacterial Growth and Siderophore Isolation. *Alcanivorax pacificus* W11-5^T, obtained from Dr. Zongze Shao (Marine Culture Collection of China, Third Institute of State Oceanic Administration, P. R. China), was cultured on Difco 2216 Marine medium agar plates amended with sodium pyruvate. Single colonies were inoculated in a liquid artificial seawater medium ASW+Py (10 g CAS amino acids L^{-1}, 1 g NH_4Cl L^{-1}, 1 g glycerol phosphate L^{-1}, 12.35 g $MgSO_4$ L^{-1}, 1.45 g $CaCl_2$ L^{-1}, 16.55 g NaCl L^{-1}, 0.75 g KCl L^{-1}, 5 g sodium pyruvate L^{-1} in ddH_2O, amended with 10 mL of 1.0 M HEPES L^{-1}, 2 mL of 1.0 M $NaHCO_3$ L^{-1}, and 6 mL of glycerol L^{-1}) or in a liquid single carbon source medium (24.6 g NaCl L^{-1}, 0.67 g KCl L^{-1}, 1.36 g $CaCl_2$ L^{-1}, 6.29 g $MgSO_4$ L^{-1}, 4.66 g $MgCl_2$ L^{-1}, 0.18 g $NaHCO_3$ L^{-1}, 10.0 g sodium pyruvate L^{-1}, 2.0 g NH_4Cl L^{-1}, 0.2 g Na_2HPO_4 L^{-1}, 10 mM of Ser/Gly) for amino acid amendment, with microbial growth monitored by OD_{600}. Cultures of 2 L and 500 mL (for amino acid amendment) volumes were grown. Cultures were harvested in the late log phase of growth by centrifugation (SLA-3000 rotor, ThermoScientific) at 6000 rpm for 30 min at 4 °C. Culture supernatants were decanted and shaken with 100 g of XAD-2 polystyrene resin for 3 h at 4 °C to adsorb organics. The resin was filtered from the supernatant, washed with 250 mL of 90/10% ddH_2O/MeOH, and eluted with 300 mL of 10/90% ddH_2O/MeOH. The eluent was concentrated under vacuum to 40 mL and stored at 4 °C for analysis. Eluent was further purified by semipreparative HPLC on a YMC 20 × 250 mm C18-AQ column, with a linear gradient of 10% MeOH in ddH_2O (+0.1% trifluoroacetic acid) to 30% MeOH in ddH_2O (+0.1% trifluoroacetic acid) over 40 min, yielding pure product (30.02 mg from 2 L of culture).

DOI: 10.1021/acs.jnatprod.8b01073
J. Nat. Prod. XXXX, XXX, XXX−XXX

Pacifibactin: white solid; $[\alpha]^{18}_{D}$ −63 (*c* 0.100, MeOH); ^{1}H and ^{13}C NMR data, Table 1; HRESIMS *m/z* 923.4081 $[M + H]^+$ (calcd for $C_{34}H_{59}N_{12}O_{18}$, 923.4065).

Amino Acid Analysis of Pacifibactin with Marfey's Reagent. Purified apo pacifibactin (2 mg) was dissolved in 6 M HCl, sealed in an ampule under argon, and heated at 80 °C for 8 h to hydrolyze the siderophore. The hydrolysis mixture was evaporated to dryness to remove HCl and redissolved in ddH_2O. After two additional cycles of evaporation and dissolution in ddH_2O, the hydrolysis mixture was derivatized with 1-fluoro-2−4-dinitrophenyl-5-L-alanine amide (Marfey's reagent) using standard procedures.[36] The hydrolysis procedure was also performed as described using 45% HI in place of 6 M HCl to reduce any N5-acetyl-N5-hydroxyornithine and N5-hydroxyornithine to ornithine to aid in analysis. Derivatized hydrolysis products were separated by HPLC on a YMC 4.6 × 250 mm C18-AQ column with a gradient from 15% CH_3CN in ddH_2O (0.05% trifluoroacetic acid) to 50% CH_3CN in ddH_2O (0.05% trifluoroacetic acid) over 60 min. Derivatized hydrolysis products were co-injected with standards of Marfey's derivatized amino acids to determine the constituent amino acids of pacifibactin: D,L-*threo*-β-OH-Asp (Sigma-Aldrich), L-Ser (Alfa-Aesar), D-Arg (Alfa-Aesar), D-Orn (Sigma-Aldrich), L-Orn (Sigma-Aldrich). D,L-*erythro*-β-OH-Asp was synthesized through treatment of 2,3-*trans*-expoxysuccinic acid (50 mg) with 375 μL of concentrated aqueous NH_4OH (28%).[33] The reaction was sealed in a glass ampule and heated for 20 h at 50 °C. The crude mixture was dried, then dissolved in 1.5 mL of ddH_2O. The product was then derivatized with Marfey's reagent, and the formation of derivatized D,L-*erythro*-β-OH-Asp as the dominant product was confirmed by UPLC-ESIMS, noting the mass of the derivatized amino acid and the difference in retention time in comparison to the D,L-*threo*-β-OH-Asp-derivatized standard.

Fe(III) Titration of Pacifibactin. A 2.12 mM stock solution of Fe(III) was prepared by diluting a 1 mg/mL $Fe(NO_3)_3$ atomic absorption standard solution with ddH_2O and standardized spectrophotometrically with 1,10-phenanthroline using established procedures.[37] A stock solution of apo pacifibactin was prepared by dissolving freeze-dried siderophore in ddH_2O. To standardize the pacifibactin stock solution, a 400 μL aliquot of apo pacifibactin stock solution was lyophilized then dissolved with 2.77 mg of dried maleic acid in 700 μL of 99% $(CD_3)_2SO$. ^{1}H NMR peak integrations of pacifibactin and the maleic acid internal standard were then taken, establishing a stock concentration of 3.5 mM. A solution of apo pacifibactin (2 mL, 0.1 mM in ddH_2O buffered with 100 mM MOPS pH 7.1) was prepared in a 3 mL quartz cuvette and titrated with the standardized Fe(III) stock solution (2.14 mM in 40 mM HNO_3). After each aliquot of Fe(III) was added, the solution was allowed to equilibrate for 24 h, a period determined by monitoring changes in the UV−vis absorption spectrum of the solution after Fe(III) addition. After equilibration, the UV−vis absorbance spectrum of the solution was measured.

Preparation of Ga(III)-Pacifibactin. A stock solution of Ga(III) was prepared by dissolving a gallium metal ingot in boiling 30% HNO_3 for 24 h. The resultant $Ga(NO_3)_3$ stock solution was standardized by colorimetric titration with EDTA (pyrocatechol violet indicator).[38] Ga(III)-bound pacifibactin was prepared by adding Ga(III) stock solution to a solution of apo pacifibactin at a 4:3 molar ratio in ddH_2O (pH adjusted to 7.0 with NaOH addition) and equilibrating for 48 h. The solution was then purified by solid-phase extraction using a Waters C18 SepPak to remove any excess Ga(III) and salts.

Photolysis of Fe(III)-Pacifibactin. Fe(III)-bound pacifibactin was prepared by adding Fe(III) stock solution to a solution of apo pacifibactin at a 1:1 ratio for a final concentration of 0.1 mM Fe(III)-pacifibactin in 100 mM MOPS pH 7.1 buffer (2 mL) in a 3 mL quartz cuvette. The solution was equilibrated for 24 h before photolysis. Solutions were photolyzed using a 450 W mercury arc lamp (Ace Glass, 40−48% output in UV) as the light source, with cuvettes placed on a stand approximately 15 cm away from the lamp. UV−vis absorbance spectra of the solutions were measured at timed intervals. After photolysis, the reaction mixture was analyzed by UPLC-ESIMS/MS to detect and characterize any photoproducts.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jnatprod.8b01073.

Gene cluster annotation, ESIMS/ESIMS/MS spectra, NMR spectra, UV−visible spectra from titration and photolysis, HPLC chromatograms from amino acid analysis (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: butler@chem.ucsb.edu.

ORCID

Alison Butler: 0000-0002-3525-7864

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful for support from NSF CHE-1710761. C.D.H. thanks the UCSB Mellichamp Academic Initiative in Sustainability for a fellowship in support of his research. This work made use of the MRL Shared Experimental Facilities supported by the MRSEC Program of the NSF, DMR 1121053. We thank R. Behrens (UPLC-MS) and H. Zhou (NMR) for technical help.

■ REFERENCES

(1) Jensen, P. R. *Trends Microbiol.* **2016**, *24* (12), 968−977.
(2) Ziemert, N.; Alanjary, M.; Weber, T. *Nat. Prod. Rep.* **2016**, *33* (8), 988−1005.
(3) Koglin, A.; Walsh, C. T. *Nat. Prod. Rep.* **2009**, *26* (8), 987−1000.
(4) Sattely, E. S.; Fischbach, M. A.; Walsh, C. T. *Nat. Prod. Rep.* **2008**, *25* (4), 757−793.
(5) Crosa, J. H.; Walsh, C. T. *Microbiol. Mol. Biol. Rev.* **2002**, *66* (2), 223−249.
(6) Samel, S. A.; Marahiel, M. A.; Essen, L.-O. *Mol. BioSyst.* **2008**, *4* (5), 387−393.
(7) 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. *Nucleic Acids Res.* **2017**, *45* (W1), W36−W41.
(8) Skinnider, M. A.; Dejong, C. A.; Rees, P. N.; Johnston, C. W.; Li, H.; Webster, A. L. H.; Wyatt, M. A.; Magarvey, N. A. *Nucleic Acids Res.* **2015**, *43* (20), 9645−9662.
(9) Rosconi, F.; Davyt, D.; Martínez, V.; Martínez, M.; Abin-Carriquiry, J. A.; Zane, H.; Butler, A.; de Souza, E. M.; Fabiano, E. *Environ. Microbiol.* **2013**, *15* (3), 916−927.
(10) Bosello, M.; Zeyadi, M.; Kraas, F. I.; Linne, U.; Xie, X.; Marahiel, M. A. *J. Nat. Prod.* **2013**, *76* (12), 2282−2290.
(11) Zane, H. K.; Naka, H.; Rosconi, F.; Sandy, M.; Haygood, M. G.; Butler, A. *J. Am. Chem. Soc.* **2014**, *136* (15), 5615−5618.
(12) Kurth, C.; Schieferdecker, S.; Athanasopoulou, K.; Seccareccia, I.; Nett, M. *J. Nat. Prod.* **2016**, *79* (4), 865−872.
(13) Hermenau, R.; Ishida, K.; Gama, S.; Hoffmann, B.; Pfeifer-Leeg, M.; Plass, W.; Mohr, J. F.; Wichard, T.; Saluz, H.-P.; Hertweck, C. *Nat. Chem. Biol.* **2018**, *14* (9), 841−843.
(14) Kem, M. P.; Zane, H. K.; Springer, S. D.; Gauglitz, J. M.; Butler, A. *Metallomics* **2014**, *6* (6), 1150−1155.
(15) Barbeau, K.; Rue, E. L.; Bruland, K. W.; Butler, A. *Nature* **2001**, *413*, 409−413.

DOI: 10.1021/acs.jnatprod.8b01073
J. Nat. Prod. XXXX, XXX, XXX−XXX

(16) Lai, Q.; Wang, L.; Liu, Y.; Fu, Y.; Zhong, H.; Wang, B.; Chen, L.; Wang, J.; Sun, F.; Shao, Z. *Int. J. Syst. Evol. Microbiol.* **2011**, *61* (6), 1370−1374.
(17) Schwyn, B.; Neilands, J. B. *Anal. Biochem.* **1987**, *160* (1), 47−56.
(18) Quadri, L. E. N.; Sello, J.; Keating, T. A.; Weinreb, P. H.; Walsh, C. T. *Chem. Biol.* **1998**, *5* (11), 631−645.
(19) Chevrette, M. G.; Aicheler, F.; Kohlbacher, O.; Currie, C. R.; Medema, M. H. *Bioinformatics* **2017**, *33* (20), 3202−3210.
(20) Franke, J.; Ishida, K.; Hertweck, C. *Chem. - Eur. J.* **2015**, *21* (22), 8010−8014.
(21) Franke, J.; Ishida, K.; Ishida-Ito, M.; Hertweck, C. *Angew. Chem., Int. Ed.* **2013**, *52* (32), 8271−8275.
(22) Johnston, C. W.; Skinnider, M. A.; Wyatt, M. A.; Li, X.; Ranieri, M. R. M.; Yang, L.; Zechel, D. L.; Ma, B.; Magarvey, N. A. *Nat. Commun.* **2015**, *6*, 8421.
(23) Kreutzer, M. F.; Nett, M. *Org. Biomol. Chem.* **2012**, *10* (47), 9338−9343.
(24) Heydel, P.; Plattner, H.; Diekmann, H. *FEMS Microbiol. Lett.* **1987**, *40* (2−3), 305−309.
(25) Robbel, L.; Helmetag, V.; Knappe, T. A.; Marahiel, M. A. *Biochemistry* **2011**, *50* (27), 6073−6080.
(26) Singh, G. M.; Fortin, P. D.; Koglin, A.; Walsh, C. T. *Biochemistry* **2008**, *47* (43), 11310−11320.
(27) Kreutzer, M. F.; Kage, H.; Nett, M. *J. Am. Chem. Soc.* **2012**, *134* (11), 5415−5422.
(28) Gauglitz, J. M.; Butler, A. *JBIC, J. Biol. Inorg. Chem.* **2013**, *18* (5), 489−497.
(29) Feistner, G. J.; Stahl, D. C.; Gabrik, A. H. *Org. Mass Spectrom.* **1993**, *28* (3), 163−175.
(30) Küpper, F. C.; Carrano, C. J.; Kuhn, J.-U.; Butler, A. *Inorg. Chem.* **2006**, *45* (15), 6028−6033.
(31) Amin, S. A.; Green, D. H.; Küpper, F. C.; Carrano, C. J. *Inorg. Chem.* **2009**, *48* (23), 11451−11458.
(32) Butler, A.; Theisen, R. M. *Coord. Chem. Rev.* **2010**, *254* (3), 288−296.
(33) Robertson, A. W.; McCarville, N. G.; MacIntyre, L. W.; Correa, H.; Haltli, B.; Marchbank, D. H.; Kerr, R. G. *J. Nat. Prod.* **2018**, *41*, 858.
(34) Hardy, C. D.; Butler, A. *JBIC, J. Biol. Inorg. Chem.* **2018**, *23*, 957.
(35) Lai, Q.; Shao, Z. *J. Bacteriol.* **2012**, *194* (24), 6936.
(36) Marfey, P. *Carlsberg Res. Commun.* **1984**, *49* (6), 591.
(37) Fortune, W. B.; Mellon, M. G. *Ind. Eng. Chem., Anal. Ed.* **1938**, *10*, 60−64.
(38) Loomis, L. D.; Raymond, K. N. *J. Coord. Chem.* **1991**, *23* (1−4), 361−387.

DOI: 10.1021/acs.jnatprod.8b01073
J. Nat. Prod. XXXX, XXX, XXX−XXX