

## Mixing and Matching Siderophore Clusters: Structure and Biosynthesis of Serratiochelins from *Serratia* sp. V4

Mohammad R. Seyedsayamdost,<sup>†,||</sup> Sara Cleto,<sup>‡,§,||</sup> Gavin Carr,<sup>†</sup> Hera Vlamakis,<sup>‡</sup> Maria João Vieira,<sup>§</sup> Roberto Kolter,<sup>\*,‡</sup> and Jon Clardy<sup>\*,†</sup>

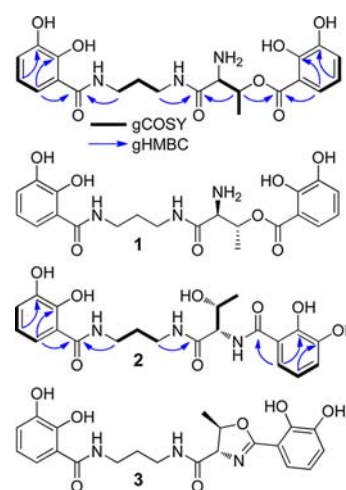
<sup>†</sup>Department of Biological Chemistry and Molecular Pharmacology, and <sup>‡</sup>Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts 02115, United States

<sup>§</sup>Institute for Biotechnology and Bioengineering, Center of Biological Engineering, University of Minho, 4710-057 Braga, Portugal

### S Supporting Information

**ABSTRACT:** Interrogation of the evolutionary history underlying the remarkable structures and biological activities of natural products has been complicated by not knowing the functions they have evolved to fulfill. Siderophores—soluble, low molecular weight compounds—have an easily understood and measured function: acquiring iron from the environment. Bacteria engage in a fierce competition to acquire iron, which rewards the production of siderophores that bind iron tightly and cannot be used or pirated by competitors. The structures and biosyntheses of “odd” siderophores can reveal the evolutionary strategy that led to their creation. We report a new *Serratia* strain that produces serratiochelin and an analog of serratiochelin. A genetic approach located the serratiochelin gene cluster, and targeted mutations in several genes implicated in serratiochelin biosynthesis were generated. Bioinformatic analyses and mutagenesis results demonstrate that genes from two well-known siderophore clusters, the *Escherichia coli* enterobactin cluster and the *Vibrio cholera* vibriobactin cluster, were shuffled to produce a new siderophore biosynthetic pathway. These results highlight how modular siderophore gene clusters can be mixed and matched during evolution to generate structural diversity in siderophores.

Iron, an essential element for almost all life forms, is abundant in the Earth's crust, but the limited solubility of ferric ions in most environments sets up a fierce competition for iron acquisition. Bacteria acquire iron using siderophores, small molecules that tightly bind ferric ions and transport them into cells through cognate receptors.<sup>1</sup> Most bacteria have receptors for siderophores that they themselves do not produce.<sup>2</sup> This allows for siderophore piracy, poaching siderophores from neighboring microbes. Some bacteria counter piracy by producing siderophores that have unusual structures and limited distribution.<sup>3</sup> The genes that encode for the production of such atypical siderophores likely represent recent evolutionary modifications of more widely distributed siderophore biosynthetic pathways.<sup>4</sup> Here, we present an example of a modified biosynthetic pathway in which the coordinated expression of two distinct gene clusters, related to enterobactin and vibriobactin biosynthetic genes, combine to produce an unusual siderophore.



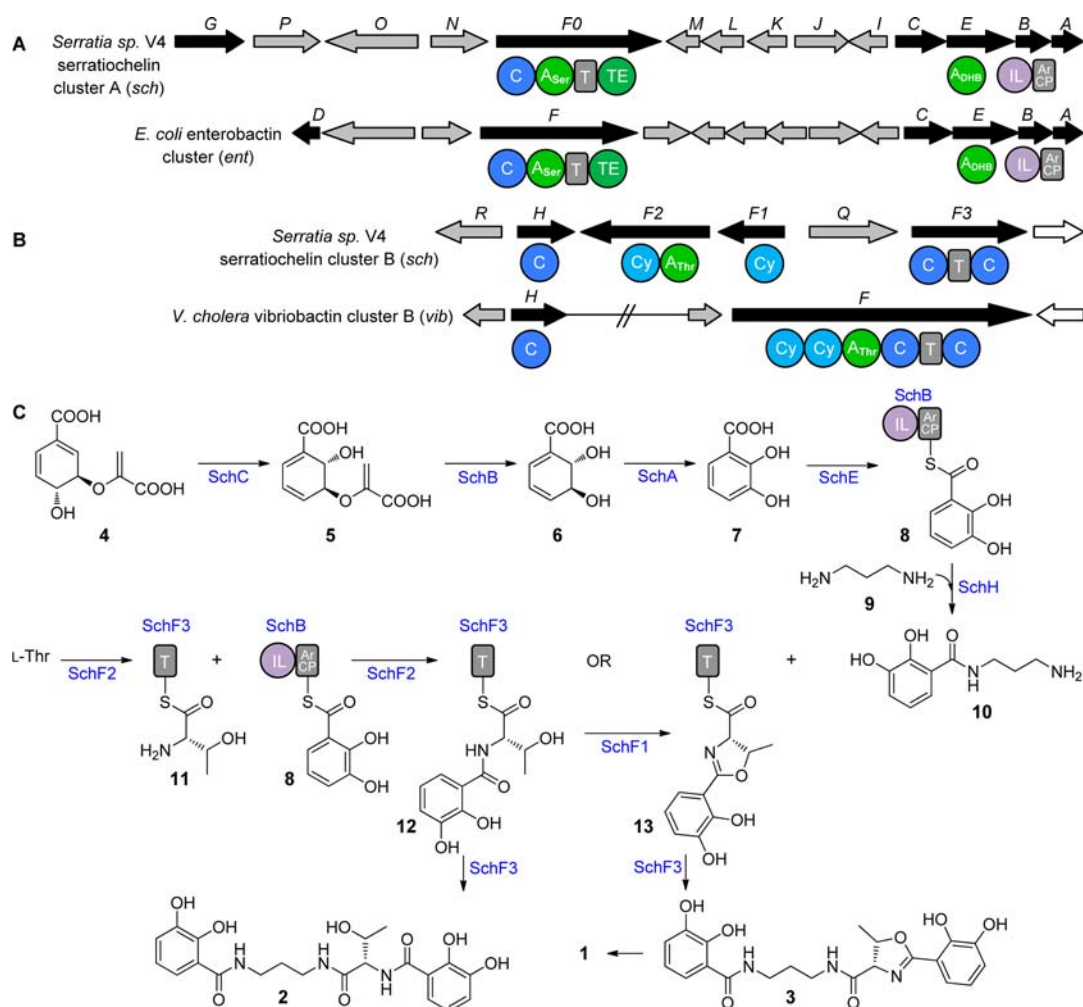
**Figure 1.** Serratiochelins produced by *Serratia* sp. V4: serratiochelins C (1), B (2), and A (3). Relevant spin systems and gHMBC correlations used to solve the structures of 1 and 2 are shown.

As part of a larger screen to identify atypical siderophores, we examined a panel of bacteria isolated from a milk processing plant. The chrome azurol S (CAS) solid medium assay<sup>5</sup> showed that one of these isolates produced large amounts of siderophore activity. The bacterial isolate belonged to the genus *Serratia*, as determined by 16S rRNA analysis. We designate this strain as *Serratia* sp. V4. Initial characterization suggested that the siderophore activity was due to a novel siderophore, which we called serratiochelin C (see below). Serratiochelin C was not detected when *Serratia* sp. V4 was cultured in the presence of iron (50  $\mu$ M, Figure S1), indicating its production is regulated by the concentration of iron in the medium.

To purify serratiochelin C for further characterization, *Serratia* sp. V4 was cultured in 1 L of an optimized liquid minimal medium. The siderophore was isolated by solid phase extraction and HPLC methods to yield 8 mg of the desferri-form. The structure of serratiochelin C was solved by <sup>1</sup>H and 2D (COSY, HSQC, HMBC) NMR and high-resolution (HR) ESI-MS analyses (1, Figures 1 and S2, Table S1). <sup>1</sup>H and COSY cross peaks revealed two aromatic spin systems, which based on coupling constants and HMBC spectra were assigned to two

Received: May 22, 2012

Published: July 25, 2012



**Figure 2.** Serratiochelin gene clusters and a biosynthetic model. (A) Comparison of serratiochelin cluster A with the *E. coli* enterobactin cluster. The sequences and synteny of these two clusters are highly homologous. (B) Comparison of serratiochelin cluster B with the *V. cholera* vibriobactin cluster. *VibH* and *vibF* are highly similar to *schH* and *schF1*, *F2* and *F3*, respectively. (C) Biosynthetic model for serratiochelin based on gene deletion data, bioinformatic analyses, and previous studies on vibriobactin biosynthesis.<sup>11,12</sup> A, adenylation; ArCP, aryl carrier protein; C, condensation; Cy, cyclization; IL, isochorismate lyase; T, thiolation.

distinct 2,3-dihydroxybenzoyl groups, consistent with the UV/vis spectrum of purified **1** (Figure S3). COSY cross peaks between an oxygenated methine-<sup>1</sup>H (5.65 ppm) and a methyl group (1.53 ppm) as well as an  $\alpha$ -<sup>1</sup>H (4.22 ppm) indicated the presence of a Thr moiety. An HMBC correlation from the oxygenated methine-<sup>1</sup>H to the carbonyl of one of the 2,3-dihydroxybenzoyl groups showed that these two fragments are linked through an ester bond. HMBC correlations also showed that the second aromatic system was connected to propane-1,3-diamine via an amide bond. The other amide group of propane-1,3-diamine showed an HMBC cross peak to the  $\beta$ -<sup>1</sup>H of Thr, linking the two main fragments and completing the assignment of **1** (Figure 1). The structure is consistent with HR-MS analysis (C<sub>21</sub>H<sub>25</sub>N<sub>3</sub>O<sub>8</sub>, [M+H]<sup>+</sup>: calcd 448.1720, expt 448.1720). The absolute configuration of L-Thr was determined by Marfey's analysis (Figure S4).

We also identified an analog of **1**, serratiochelin B (**2**), with an identical molecular formula but different migratory properties. Using NMR and HR-MS ([M+H]<sup>+</sup>: calcd 448.1720, expt 448.1721), we assigned this product to the amide form of **1** (**2**, Figure S5, Table S2). A third analog was the dehydrated form of **1** and **2**. Using NMR, HR-MS ([M+H]<sup>+</sup>: calcd 430.1614, expt

430.1610), and MS-MS (Figures S6, S7), we assigned it as the 2-methyloxazoline derivative (**3**). Compound **3** was discovered in *S. marcescens* in 1974 and named serratiochelin, but it managed to evade inclusion in almost all major reviews on siderophores since its original discovery.<sup>6</sup> It seemed likely that **1** and **2** were hydrolytic products of **3** that formed during the purification procedure, which employed formic acid. We examined untreated culture extracts and monitored the stability of **1**-**3** as a function of [formic acid] and incubation time at rt (Figure S8). **2** and **3**, but not **1**, were present in the untreated culture extracts (~1:6 ratio), and **3** converted to **1** upon incubation with formic acid, the rate of formation of **1** being commensurate with [formic acid]. We conclude that **2** and **3** are the intended products and that **3** hydrolyzes to give **1** during the purification.

Serratiochelins are bis-catecholate siderophores and may be tetra- or hexadentate. In **3**, the ligands for octahedral coordination of Fe may be provided by the pair of catechols, affording four ligands, with the oxazoline N and the adjacent amide O of the L-Thr moiety supplying two additional ligands. The latter ligand pair is geometrically similar to  $\alpha$ -amino carboxylates,  $\alpha$ -hydroxy carboxylates, and hydroxamates, all well-known Fe-binding moieties in siderophores.<sup>1</sup> To begin to assess

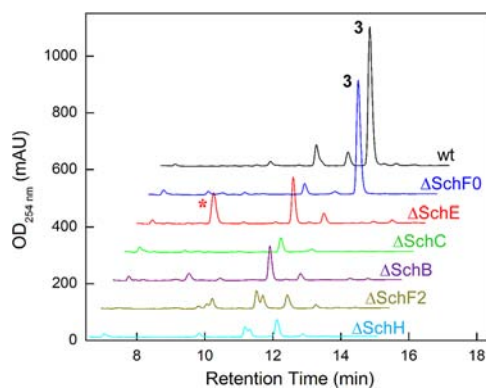
the mode of metal binding by **3**, we generated the corresponding Ga and Fe complexes (Figures S9, S10). Surprisingly, HPLC-MS analysis of these complexes was consistent with the loss of three protons in a 1:1 Ga or Fe siderophore complex, indicating tetradentate chelation by one of the catechols, as one pair of ligands, and by the oxazoline N and one hydroxyl group of the second catechol group, as a second pair of ligands. We also observed the hexacoordinate bis-aqua form of these Ga and Fe complexes. The preference for the oxazoline N over the second catechol hydroxyl as a ligand may be explained by the corresponding  $pK_a$ s of  $\sim 7$  and 11, respectively.<sup>4</sup>

A key and unusual feature of the serratiochelins is the propane-1,3-diamine linker. Incorporation of propanediamine into siderophores derived from a non-ribosomal peptide synthetase (NRPS) and the biosynthesis of serratiochelin have not been examined. To investigate these aspects further, we first predicted that the serratiochelin biosynthetic cluster would contain a 2,3-dihydroxybenzoate-AMP ligase (EntE), which is known to incorporate 2,3-dihydroxybenzoate into various siderophores.<sup>4b,7</sup> Using the *entE* sequences of three *Serratia* strains, we designed degenerate primers to find *entE* in *Serratia* sp. V4 and to locate the serratiochelin biosynthetic cluster. To obtain the full sequence of the cluster, the entire genome of *Serratia* sp. V4 was sequenced using 454 sequencing methods. The 21 kb serratiochelin cluster thus obtained, assigned *sch*, is shown in Figure 2A (Table S3). The *sch* cluster contains 14 open reading frames (ORFs); 8 (*schI*–*schP*) are involved in siderophore export or ferric-siderophore uptake and utilization. Of the remaining 6 ORFs, the first gene, *schG*, is homologous to acetolactate synthase, which catalyzes the first step in branched-chain amino acid biosynthesis. The other 5 genes share high sequence identity with the corresponding enterobactin biosynthetic genes and are assigned *schA*, *B*, *C*, *E*, *F0* accordingly.<sup>4b</sup> The sequences of the *sch* genes and the synteny of the *sch* cluster are highly homologous to the *E. coli* enterobactin cluster, suggesting an evolutionary relatedness (Figure 2A, Table S3). The putative functions of the *sch* genes, based on bioinformatic analyses, are shown in Figure 2C: SchC, 67% similar to EntC, is a putative isochorismate synthase that isomerizes chorismate (**4**) to isochorismate (**5**). SchB is 82% similar to EntB and likely a bifunctional isochorismate lyase-aryl carrier protein (ArCP), which generates **6** from isochorismate. The dehydrogenase SchA shares 74% similarity with EntA and would then convert this intermediate into 2,3-dihydroxybenzoate (**7**). SchE, 81% similar to EntE, likely adenylates **7** and transfers it to the ArCP domain of SchB (**8**). SchF0 is an NRPS with the domain architecture shown in Figure 2A<sup>8</sup> – it may be involved in incorporating L-Thr into serratiochelin (see below).

While the *sch* cluster provides candidate genes for generation of intermediate **8** (Figure 2C), it does not reveal any genes that would incorporate propane-1,3-diamine. As mentioned above, insertion of this group has not been studied, but the pathway for incorporation of norspermidine [ $H_2N-(CH_2)_3-NH-(CH_2)_3-NH_2$ ] into vibriobactin in *Vibrio cholera* has been established. In this case, Keating et al. demonstrated that an unusual free-standing amide synthase, VibH, condenses norspermidine with the ArCP-bound 2,3-dihydroxybenzoyl thioester.<sup>9</sup> To examine if a VibH homologue was involved in serratiochelin biosynthesis, we searched the genome of *Serratia* sp. V4 using the VibH sequence as a query and obtained a single hit. This gene, assigned *schH*, has 45% sequence similarity to VibH and is embedded in a cluster of 7 genes (Figure 2B, Table S4).<sup>10</sup> SchQ and SchR are siderophore uptake genes, homologous to those in *V. cholera*

(60% and 56% similar to *viuA* and *sapA*, respectively). SchF1 is a putative free-standing NRPS cyclization (Cy) domain and bears 60% sequence similarity to the N-terminal Cy-domain of VibF, which is essential in vibriobactin biosynthesis. SchF2 is putatively a Cy-A didomain and is also highly similar (64%) to the corresponding domains in VibF, with the A-domain coding for Thr (Table S4). SchF3 shares 56% sequence similarity with the C-terminal portion of VibF and contains a C-T-C domain arrangement (Figure 2B). Remarkably, SchF1, F2, and F3 recapitulate the domain architecture of VibF entirely (Figure 2B). The homology of the *sch* genes in this second cluster suggests an evolutionary link to the vibriobactin biosynthetic cluster. Together the bioinformatic analyses imply that modification and combination of enterobactin and vibriobactin clusters may have led to evolution of a new siderophore.

To examine this model and the proposed roles of some of the enzymes in serratiochelin biosynthesis, a series of targeted gene interruption studies were carried out (Tables S5–S7). SchB, C, E, and F0 in *sch* cluster A, and SchF2 and H in cluster B, were genetically interrupted by Campbell insertions; the resulting mutants were examined for serratiochelin production (Figure 3, Tables S6, S7). Interruption of *schB* and *schC* completely abolished serratiochelin synthesis, consistent with the assigned roles for these genes. Interruption of *schE* not only impaired serratiochelin production but also led to accumulation of **7**, as identified by HPLC-MS analysis ( $[M-H]^-$ : calcd 153.0, expt 153.0), entirely consistent with the assigned role for SchE. Surprisingly, *schF0* was not required for siderophore production. SchF0 is an EntF homologue; in enterobactin biosynthesis, EntF condenses a T-domain-bound Ser with an ArCP-tethered **7**.<sup>4b</sup> In this case, the *schF0* mutant produced wt levels of **3** (Figure 3). A



**Figure 3.** HPLC-MS analysis of wt *Serratia* sp. V4 and *sch* mutants. Traces are offset in both axes for clarity. Peaks corresponding to **3** in the wt and  $\Delta$ SchF0 traces are marked. \*, 2,3-dihydroxybenzoic acid (**7**).

mutant lacking SchF2, which also contains an A-domain (Thr-specific),<sup>8</sup> was deficient in serratiochelin synthesis. Further, interruption of the *schH* gene also abolished serratiochelin production. Thus, gene interruption studies establish that products from two distinct clusters, one homologous to an *E. coli* enterobactin cluster and another to the *V. cholera* vibriobactin cluster, are required for biosynthesis of **2** and **3**.

Phylogenetic analyses of several genes in the *sch* cluster are consistent with a model in which two gene clusters combine to yield a new siderophore. These analyses show that *schF1* and *schF2* fall into the same clade as the corresponding *vib* genes in many *Vibrio* species (Figure S11). *SchF3* and *schH* are also phylogenetically related to their *Vibrio* counterparts, while *schB*



and *schE* show close similarity to the corresponding *E. coli* genes. Further, insertional mutations in some of the genes responsible for serratiochelin biosynthesis result in a CAS-negative phenotype, indicating that serratiochelins are the dominant siderophores produced by *Serratia sp.* V4 and that the *sch* cluster does not produce enterobactin or vibriobactin.

Using results from gene interruption studies and bioinformatic analyses, we propose a biosynthetic model for serratiochelin A and B (Figure 2C). In this model, coordinated expression and catalysis by enzymes within two clusters lead to serratiochelin production.<sup>4b,11</sup> In analogy with enterobactin biosynthesis, SchA, B, C, and E generate an ArCP-bound thioester (**8**), as described above. In enterobactin biosynthesis, EntF would carry out the next biosynthetic step. At this point, serratiochelin biosynthesis appears to diverge from the enterobactin pathway. We propose that SchH, the free-standing amide synthase from a separate gene cluster, catalyzes the next step by forming an amide bond between **8** and a nontethered propane-1,3-diamine (**9**) to generate **10**, in contrast to conventional NRPS assembly line logic that involves condensation of two tethered substrates.<sup>9,10b</sup> We further propose that the A-domain in SchF2 activates L-Thr and installs it onto the T-domain of SchF3 (**11**). The Cy-domains in SchF1 and F2 then likely condense the tethered L-Thr and 2,3-dihydroxybenzoyl groups to give **12** and also cyclize the Thr side chain into a methyloxazoline to give **13**.<sup>12a</sup> In analogy with *in vitro* studies in the vibriobactin system, we propose that one of the C-domains in SchF3 (likely the C-terminal C-domain) can use both **12** and **13** as tethered substrates for amide bond formation, with the free amino group of soluble substrate **10** completing the production of **2** and **3**.<sup>12</sup> The N-terminal C-domain may be catalytically inactive and required for proper orientation of the Cy-domains with the C-T-C domains, as previously demonstrated for VibF.<sup>13</sup> Further, in vibriobactin biosynthesis, it has been shown that the second Cy-domain of VibF (corresponding to *schF2*) is responsible for amide bond formation and the N-terminal Cy-domain (corresponding to *schF1*) catalyzes methyloxazoline formation.<sup>12a</sup> Separation of the two Cy-domains in the serratiochelin gene cluster generates siderophore diversity by providing both the open and cyclized forms of the Thr side chain, resulting in the synthesis of **2** and **3**.

In summary, we have shown that serratiochelin is a tetradentate siderophore produced from two distinct gene clusters and proposed a model for its biosynthesis. The structures of **2** and **3** are unusual in that they contain a propane-1,3-diamine backbone. Our results provide a model for how this moiety is incorporated into serratiochelin: mutant analyses confirmed that five biosynthetic enzymes are provided by a cluster that bears high similarity, both in sequence and in synteny, to the *E. coli* enterobactin gene cluster. Four additional biosynthetic genes, which are homologous to those found in vibriobactin clusters in *V. cholerae*, are provided by a separate gene cluster. Thus, elements from two disparate clusters are mixed and matched to yield a new biosynthetic pathway. Combination of two genetic loci to generate a novel siderophore highlights the ability of siderophore biosynthetic clusters to evolve.<sup>4</sup> Not only is the modular nature of siderophore production well-suited for evolution of new structures, but the fierce competition for iron exerts a greater selective pressure for idiosyncratic siderophores that cannot be pirated by competing organisms. Elucidating the evolution of siderophore biosynthetic pathways has important implications for pathogenic bacteria, such as *Pseudomonas aeruginosa*, septicemic *E. coli*, and other human pathogens.<sup>14</sup>

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

Experimental procedures, spectra, and annotation of the *sch* cluster. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

rkolter@hms.harvard.edu; jon\_clardy@hms.harvard.edu

### Author Contributions

||M.R.S. and S.C. contributed equally.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (Grants GM82137 to R.K., and AI057159 and GM086258 to J.C.). M.R.S. acknowledges support from the NIH Pathway to Independence Award (Grant 1K99 GM098299-01). S.C. and M.J.V. acknowledge support from the Portuguese Foundation for Science and Technology (PhD Grant SFRH/BD/38298/2007 to S.C.; Project PTDC/EBB-EBI/104263/2008 to M.J.V.).

## ■ REFERENCES

- (1) (a) Hider, R. C.; Kong, X. *Nat. Prod. Rep.* **2010**, *27*, 637. (b) Miethke, M.; Marahiel, M. A. *Microbiol. Mol. Biol. Rev.* **2007**, *711*, 413. (c) Sandy, M.; Butler, A. *Chem. Rev.* **2009**, *109*, 4580.
- (2) (a) Pool, K.; McKay, G. A. *Front. Biosci.* **2003**, *8*, d661. (b) Jurkevitch, E.; Hadar, Y.; Chen, Y. *Appl. Environ. Microbiol.* **1992**, *58*, 119.
- (3) (a) Seyedsayamdost, M. R.; Traxler, M. F.; Zheng, S.-L.; Kolter, R.; Clardy, J. *J. Am. Chem. Soc.* **2011**, *133*, 11434. (b) Challis, G. L.; Hopwood, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14555.
- (4) (a) Fischbach, M. A.; Walsh, C. T.; Clardy, J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4601. (b) Crosa, J. H.; Walsh, C. T. *Microbiol. Mol. Biol. Rev.* **2002**, *66*, 223.
- (5) Schwyn, B.; Neilands, J. B. *Anal. Biochem.* **1987**, *160*, 47.
- (6) Ehlert, G.; Taraz, K.; Budzikiewicz, H. *Z. Naturforsch.* **1974**, *49c*, 681.
- (7) (a) Luke, R. K.; Gibson, F. *J. Bacteriol.* **1971**, *107*, 557. (b) Gehring, A. M.; Mori, I.; Walsh, C. T. *Biochemistry* **1998**, *37*, 2648.
- (8) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. *Chem. Biol.* **1999**, *6*, 493.
- (9) Keating, T. A.; Marshall, C. G.; Walsh, C. T. *Biochemistry* **2000**, *39*, 15513.
- (10) (a) Wyckoff, E. E.; Stoebner, J. A.; Reed, K. E.; Payne, S. M. *J. Bacteriol.* **1997**, *179*, 7055. (b) Keating, T. A.; Marshall, C. G.; Walsh, C. T. *Biochemistry* **2000**, *39*, 15522.
- (11) (a) Lazos, O.; Tosin, M.; Slusarczyk, A. L.; Boakes, S.; Cortes, J.; Sidebottom, P. J.; Leadlay, P. F. *Chem. Biol.* **2010**, *17*, 160. (b) Bosello, M.; Robbel, L.; Linne, U.; Xie, X.; Marahiel, M. A. *J. Am. Chem. Soc.* **2011**, *133*, 4587.
- (12) (a) Marshall, C. G.; Hillson, N. J.; Walsh, C. T. *Biochemistry* **2002**, *41*, 244. (b) Marshall, C. G.; Burkart, M. D.; Keating, T. A.; Walsh, C. T. *Biochemistry* **2001**, *40*, 10655.
- (13) (a) Hillson, N. J.; Walsh, C. T. *Biochemistry* **2003**, *42*, 766. (b) Hillson, N. J.; Balibar, C. J.; Walsh, C. T. *Biochemistry* **2004**, *43*, 11344.
- (14) (a) Lamont, I. L.; Beare, P. A.; Ochsner, U.; Vasil, A. I.; Vasil, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7072. (b) Caza, M.; Lepine, F.; Dozois, C. M. *Mol. Microbiol.* **2011**, *80*, 266. (c) Ratledge, C.; Dover, L. *G. Annu. Rev. Microbiol.* **2000**, *54*, 881.