JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

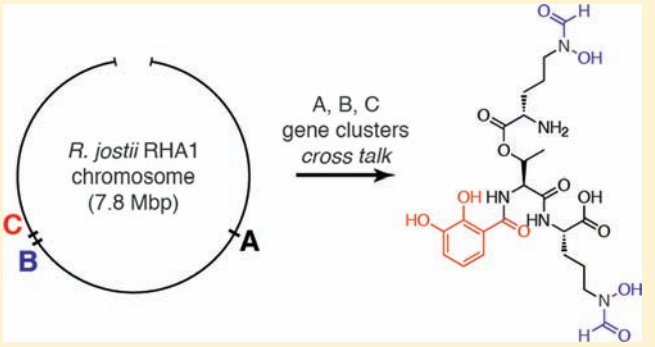
Biosynthesis of the Siderophore Rhodochelin Requires the Coordinated Expression of Three Independent Gene Clusters in *Rhodococcus jostii* RHA1

Mattia Bosello, Lars Robbel, Uwe Linne, Xiulan Xie, and Mohamed A. Marahiel*

Biochemistry, Department of Chemistry, Philipps-University Marburg, Hans-Meerwein-Strasse D-35043 Marburg, Germany

Supporting Information

ABSTRACT: In this work we report the isolation, structural characterization, and the genetic analysis of the biosynthetic origin of rhodochelin, a unique mixed-type catecholate-hydroxamate siderophore isolated from *Rhodococcus jostii* RHA1. Rhodochelin structural elucidation was accomplished via MS^n- and NMR-analysis and revealed the tetrapeptide to contain an unusual ester bond between an L-*δ*-*N*-formyl-*δ*-*N*-hydroxyornithine moiety and the side chain of a threonine residue. Gene deletions within three putative biosynthetic gene clusters abolish rhodochelin production, proving that the ORFs responsible for rhodochelin biosynthesis are located in different chromosomal loci. These results demonstrate the efficient cross-talk between distantly located secondary metabolite gene clusters and outline new insights into the comprehension of natural product biosynthesis.



INTRODUCTION

In most microbial habitats, soluble $Fe^{(II)}$ is spontaneously oxidized to $Fe^{(III)}$ which, in the presence of oxygen and water and at neutral pH, forms insoluble ferric oxide hydrate complexes, leading to a free $Fe^{(III)}$ concentration of only up to 10^{-18} M.[1] Because iron is an essential cofactor required in many enzymes of the primary and secondary metabolism of most living organisms, it clearly becomes a key limiting feature of microbial growth, in particular during the colonization and the development of virulence in animal and human hosts. In order to cope with iron-limiting conditions, microbes have developed mechanisms for highly selective metal uptake.[2] The secretion of low-molecular weight organic chelator compounds called siderophores is one of the main iron-mobilizing strategies used by environmental and pathogenic strains.[3]

Depending on the chemical nature of the moieties involved in the coordination of ferric iron, siderophores can be divided into three main classes: catecholates, hydroxamates, and (hydroxy)-carboxylates. However, continuous discovery of new siderophores led to a more complex classification, due to the presence of at least two different chemical features within one molecule, resulting in "mixed-type" siderophores.[1] In addition, regarding their different biosynthetic origins, siderophore biosynthesis can be classified as NRPS (nonribosomal peptide synthetase) dependent[4] and NRPS independent.[5] Since siderophores display the same biological function, the most interesting point about these molecules is their chemical diversity, which can thus reveal new mechanisms of their assembly or aid in the determination of the fate of iron after recognition of the iron−siderophore complex by specific importers and subsequent channeling of iron to intracellular targets.[6−8]

In recent years, increasing information derived by microbial genome sequencing projects and the application of bioinformatic tools for genome mining has enabled an improved rationale for natural product discovery. In this context, several siderophores have been described, starting from the in silico characterization of their respective biosynthetic pathways: coelichelin,[9] fuscachelin A,[10] and erythrochelin[11] **3** (Figure 1) are examples of the application of this strategy.[12]

Rhodococcus species are extensively studied as extraordinary biocatalysts for steroid production and fossil fuel biodesulfurization and as tools suitable for bioremediation purposes.[13,14] These widespread biotechnological and industrial interests derive from their diverse metabolic capability. Furthermore, genome sequencing information revealed *Rhodococcus* species to possess a vast genetic potential for secondary metabolite production.[15] However, only very few natural products have been isolated from this genus, among them two siderophores, heterobactin A **1** and rhodobactin **2** (Figure 1), isolated from *Rhodococcus erythropolis* IGTS8 and *Rhodococcus rhodochrous* OFS, respectively.[16,17] They belong to the hydroxamate−catecholate mixed type family, with the common presence of 2,3-dihydroxybenzoic acid (2,3-DHB) and differently modified ornithine residues within their structures. In both cases, no gene clusters responsible for siderophore

Received: December 6, 2010
Published: March 07, 2011

dx.doi.org/10.1021/ja1109453 | *J. Am. Chem. Soc.* 2011, 133, 4587–4595

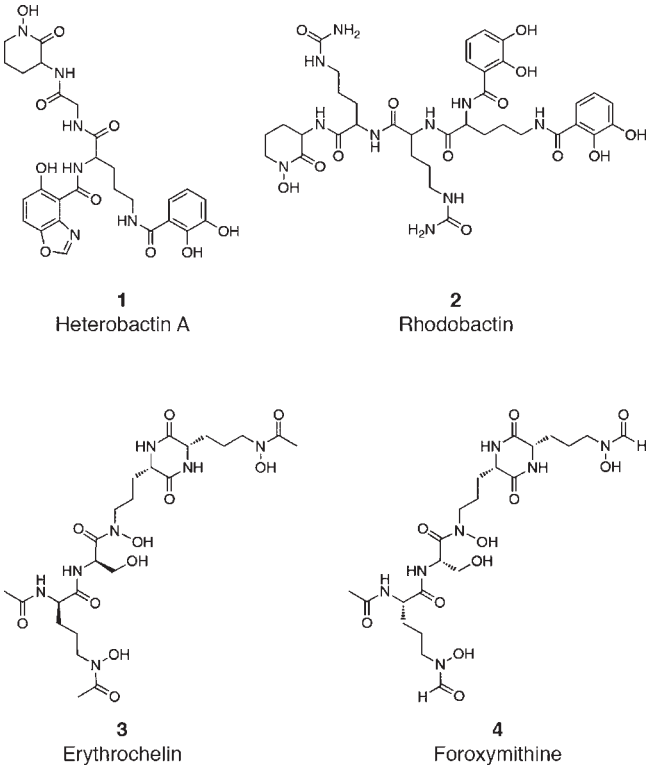


Figure 1. Chemical structures of representative siderophores isolated from *Rhodococcus* strains: heterobactin A **1** and rhodobactin **2** were isolated from *R. erythropolis* IGT8 and *R. rhodochrous* OFS, respectively. Comparison of erythrochelin **3** (*Saccharopolyspora erythraea* NRRL2338) and foroxymithine **4** (*Streptomyces nitrosporeus*) reveals a high degree of structural similarity.

biosynthesis have been identified. In addition, the presence of a coordinating aryl residue within the structure of the siderophore produced by the pathogenic *Rhodococcus equi* has been proposed, although the isolation and characterization of the molecule has not been completed.[18]

Rhodococcus jostii RHA1 was isolated from a lindane-contaminated soil and is known for its ability to transform polychlorinated biphenyls, and to utilize a wide range of aromatic compounds, carbohydrates, nitriles, and steroids as the sole energy source:[19] these features make it a species of significant industrial interest. In this context, its genome sequence has been published in 2006 and contains approximately 9.7 Mbps arranged in one linear replicon and three additional linear plasmids. *R. jostii* RHA1 also contains 24 NRPS genes, 6 of which exceed 25 kbp, and 7 PKS genes, providing evidence of an extensive and uncharacterized secondary metabolism.[15]

In this study, we report the isolation of rhodochelin, a new mixed-type hydroxamate−catecholate siderophore from *R. jostii* RHA1 and the elucidation of its chemical structure via NMR and MS^n studies. A genome mining and gene disruption approach allowed the identification of three distinct gene clusters shown to be essential for rhodochelin biosynthesis. These results provide the first example of functional cross-talk between three distantly located NRPS gene clusters.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Growth Conditions. *Rhodococcus jostii* RHA1 was maintained on LB agar slants medium at 30 °C. M9 minimal medium, supplemented with trace elements[20] (depleted of iron) and 4 g/L glucose, was used to induce rhodochelin production. *Escherichia coli* strains were grown in LB medium at 37 °C and 180 rpm. Strain TOP10 (Invitrogen) was used for cloning, strain S17-1 for conjugation purposes, and strain BL21(λDE3) (Novagen) for the production of recombinant proteins. Antibiotics were used where required, with the following concentrations: nalidixic acid 30 μg/mL, kanamycin 50 μg/mL for pET28a(+) (Novagen) and *R. jostii* transconjugants, and kanamycin 25 μg/mL for pK18mobsacB derivatives in *E. coli*.

Isolation and Purification of Rhodochelin from Culture Supernatant. *R. jostii* was grown for two days in LB medium. Cells were harvested, washed, and resuspended in an equal amount of M9 medium. A 1/100 aliquot was used to inoculate fresh minimal medium (in polycarbonate flasks), and cultures were grown for two days, until a CAS positive reaction[21] of the supernatant was observed. The culture supernatant was extracted with 5 g/L XAD16 resin (Sigma-Aldrich) for 2 h, and after a washing step, the adsorbed compounds were eluted with methanol and concentrated under vacuum to dryness. The eluate was resuspended in 2 mL of water and analyzed on a Nucleodur C_8(ec) column 125 × 2 mm (Macherey & Nagel) combined with an Agilent 1100 HPLC system, connected to an ESI-MS detector (Agilent 1100 MSD), utilizing the solvent gradient water + 0.05% formic acid (solvent A) and methanol + 0.04% formic acid (solvent B), with a linear gradient from 0% to 20% B within 40 min, followed by a linear increase to 95% B in 5 min, holding B for an additional 5 min. The flow rate was set to 0.3 mL/min and column temperature at 40 °C. The gradient was also used to analyze comparative extractions of *R. jostii* RHA1 mutants. Large-scale purification was carried out by scaling-up the described protocol for 5 L culture. The dried eluate was dissolved in 10 mL of water and separated on a preparative HPLC composed of a Nucleodur C_8(ec) 250 × 21 mm column combined with an Agilent 1100 HPLC system. Elution was performed with the same gradient previously described, using 215 and 280 nm as wavelengths for detection and a flow rate of 16 mL/min. Siderophore-containing fractions were analyzed via CAS assay and ESI-MS. Positive fractions were pooled according to their respective *m/z*, lyophilized, and subjected to further analysis.

MS^n and NMR Structure Elucidation. Rhodochelin MS-fragmentation experiments were carried on a LTQ-FT instrument (Thermo Fisher Scientific) by collision-induced dissociation (CID).

NMR-structure elucidation was carried out according to the following procedure. About 8 mg of rhodochelin was dissolved in 0.7 mL of H_2O/D_2O (9:1). Measurements were carried out on a Bruker AV600 spectrometer with an inverse broadband probe installed with z gradient. The one-dimensional spectra ^1H and ^{13}C, the homonuclear two-dimensional spectra DQF-COSY, TOCSY, NOESY, and ROESY, the $^1H-^{13}C$ HSQC and HMBC, and the $^1H-^{15}N$ HSQC spectra were recorded with standard pulse programs at 283 K. The TOCSY spectrum was recorded with a mixing time of 80 ms, whereas NOESY and ROESY spectra were taken at 150 and 300 ms mixing times. The 1D spectra were acquired with 65 536 data points, whereas 2D spectra were collected using 4096 points in the F_2 dimension and 512 increments in the F_1 dimension. For 2D spectra, 16−32 transients were used, while the ^{13}C spectrum was recorded with 12 K scans. The relaxation delay was 2.5 s. The ^1H chemical shifts were referenced to 4,4-dimethyl-4-silapentane sodium sulfonate (DSS) in H_2O/D_2O (9:1) at 283 K externally, whereas those of ^{13}C and ^{15}N were referenced with spectrometer default calibration. The spectra were processed with Bruker Topspin 2.1.

FT-IR and UV−vis Spectroscopic Analysis of Rhodochelin. Purified rhodochelin in a KBr disk was subjected to FT-IR spectroscopy on a Magna-IR 750 spectrometer (Nicolet). Main signals are the following: ν = 3367.2, 1749.5, 1660.5, 1586.1, 1534.1, 1448.0, 1381.8, 1204.0, 1137.4, 1066.9, 878.9, and 748.0 cm^{-1} (Supporting Information Figure 1).

UV−vis spectra were recorded on an Ultrospec 3000 (Pharmacia) spectrophotometer. Wavescan measurements were performed within a wavelength range of 200−800 nm and a scan rate of 750 nm/min.

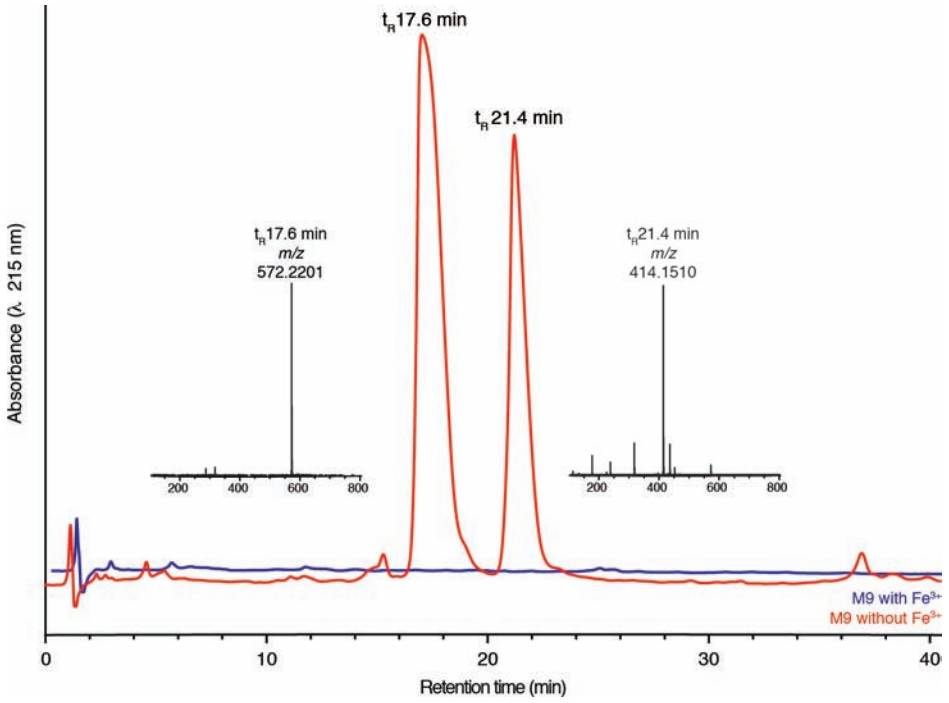


Figure 2. HPLC-MS profiles of extracted culture supernatants of *R. jostii* RHA1 grown in M9 minimal medium in the absence (red line) or presence (blue line) of Fe^{3+}. Detailed mass spectra corresponding to the UV signals are given in the insets.

Absorption spectra of rhodochelin and *holo*-rhodochelin as well as the *apo*- and *holo*-rhodochelin tripeptide were recorded in water at a final concentration of 400 μM. *holo*-Complexes were obtained by incubating rhodochelin or the tripeptide (400 μM) with aqueous $FeCl_3$ (400 μM) for 10 min at RT prior to the scan. Extinction coefficients were derived from the UV−vis spectra. The following values were obtained. *holo*-Rhodochelin: λ_{max} 330 nm, ε = 3842.5 M^{-1} cm^{-1}, λ_{max} 395 nm, ε = 2205.0 M^{-1} cm^{-1}, λ_{max} 525 nm, ε = 1567.5 M^{-1} cm^{-1}. *holo*-Rhodochelin tripeptide: λ_{max} 327 nm, ε = 3372.5 M^{-1} cm^{-1}, λ_{max} 397 nm, ε = 1430.0 M^{-1} cm^{-1}, λ_{max} 580 nm, ε = 1567.5 M^{-1} cm^{-1} (Supporting Information Figures 2 and 3).

Amino Acid Stereochemistry Determination. Assignment of the amino acid stereochemistry was achieved by rhodochelin total acid hydrolysis followed by derivatization with Nα-(2,4-dinitro-5-fluorophenyl)-L-alaninamide (FDAA, Marfey's reagent, Sigma-Aldrich).[22] Purified rhodochelin (3.76 mg) was hydrolyzed in 400 μL of 6 M HCl at 99 °C, 1000 rpm for 24 h. The lyophilized hydrolysate was resuspended in 50 μL of 1 M $NaHCO_3$, and 10 μL of this solution was added to 170 μL of 1% FDAA solution in acetone. The derivatization reaction was carried out for 1 h at 37 °C and terminated by addition of 20 μL of 1 M HCl. FDAA standard derivatives of amino acids (L/D-Thr, L/D-Orn, L-hOrn) were prepared by incubation of 25 μL of 50 mM amino acid solution with 50 μL of 1% FDAA solution and 10 μL of 1 M $NaHCO_3$ for 1 h. After lyophilization, all the products were resuspended in 200 μL of 1:1 water:acetonitrile solution prior to injection (10 μL) into a HPLC-MS system equipped with a Synergi Fusion RP-80 250 × 2.0 mm column (Phenomenex) utilizing the following solvent gradient: 0−30 min, 0−30% buffer A (10 mM ammonium formate, 1% methanol, 5% acetonitrile, pH 5.2) into buffer B (10 mM ammonium formate, 1% methanol, 60% acetonitrile, pH 5.2), followed by a linear increase to 95% buffer B in 2 min and holding 95% buffer B for an additional 5 min. The flow rate was set at 0.3 mL/min and the column temperature at 25 °C. Elution was monitored in single ion mode. L-hOrn was synthesized according to a protocol described previously.[23]

Construction of *R. jostii* RHA1 Deletion Mutants. A markerless gene deletion approach was used in order to obtain mutants of *R. jostii* RHA1, maintaining the 5′ and the 3′ end, together with the ORF frame shift. Two ∼1500 bp PCR products flanking the desired gene were amplified from the genomic DNA using two pairs of primers P1, P2 and P3, P4, and Phusion High Fidelity DNA polymerase (Finnezymes). Because of P2 and P3 primers inserting overlapping flanks, the resulting in-frame deletion amplicon was obtained by joining the former PCR reactions by splicing overlap extension (SOE) PCR.[24] The amplified deletion fragment was subsequently cloned into pK18mobsacB plasmid[25] and used to transform the *E. coli* S17-1 strain[26] to give a system suitable for *Rhodococcus* conjugative transformation.

Rhodococcus mutants were generated according to the protocol described by Van der Geize et al.[27] Cells of *R. jostii* RHA1 were grown on LB plates supplemented with nalidixic acid for 5 days, harvested, and resuspended in 2 mL of fresh LB broth. The same procedure was repeated with the overnight plates of derivative mutagenic *E. coli* S17-1 strains, additionally grown at RT for further 24 h. A 750 μL amount of each cell suspension was mixed together, incubated briefly at RT, pelleted, and resuspended in 2 mL of LB broth. A 200 μL amount was spread on LB and incubated overnight at 30 °C. The following day, cells were harvested and resuspended in 2 mL of LB broth. Aliquots (150 μL each) were successively spread on LB plates supplemented with nalidixic acid and kanamycin and incubated at 30 °C for three days, until colonies appeared. Transconjugants were grown in liquid medium, and vector integration was checked by PCR of the kanamycin cassette and by replica plating on LB supplemented with kanamycin, or kanamycin and 10% sucrose (integrant strains acquired sucrose sensitivity). To force plasmid excision, in order to obtain the second homologous recombination event, single integrant colonies were inoculated in LB broth, plated on LB supplemented with 10% sucrose, and grown at 30 °C until new colonies appeared. To confirm correct plasmid excision, single clones were tested for kanamycin sensitivity by replica plating and by PCR using different primer pairs, for the kanamycin cassette, for a flanking region and for the deleted gene. The complete list of the primers and restriction sequences used in this study is available in Supporting Information Table 1.

Cloning, Expression, and Purification of DhbE. The *dhbE* gene was amplified from genomic DNA and cloned into the pET28a(+)

dx.doi.org/10.1021/ja1109453 |*J. Am. Chem. Soc.* 2011, 133, 4587–4595

vector as N-terminal His-tag fusion and used to transform BL21 cells. A 1/100 overnight inoculum was diluted in fresh LB medium and grown at 25 °C at 230 rpm until $OD_{600} \sim 0.5$ was reached. Protein expression was induced with IPTG (50 μM) for an additional 4 h before harvesting the cells. The resulting biomass was resuspended in HEPES A buffer (50 mM HEPES, 300 mM NaCl, pH 8) and disrupted via a French press. The cleared lysate was applied to a Ni-NTA column using an ÄktaPrime (Amersham Pharmacia Biotechnology) system and eluted using a linear gradient from 0% to 50% HEPES B buffer (50 mM HEPES, 250 mM imidazole, 300 mM NaCl, pH 8) over 30 min. Protein-containing fractions were analyzed by SDS-PAGE, pooled, dialyzed (25 mM HEPES, 100 mM NaCl, pH 7.5) and concentrated.

ATP/PP$_i$ Exchange Assay. A 100 μL reaction was composed of the following: 50 mM Tris HCl pH 7.5 buffer, 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 5 mM $Na_4P_2O_7$, 10 mM amino acid. Prior to initiation of the reaction with 2 μM recombinant DhbE, 20 μL of a $Na_4{}^{32}P_2O_7$ solution (Perkin-Elmer) (approx 100 000 counts) was added. The reaction was incubated at 25 °C for 30 min and subsequently quenched with 750 μL of charcoal suspension (100 mM $Na_4P_2O_7$, 600 mM $HClO_4$, 1.6% charcoal). After a washing step with water, the resuspended charcoal was transferred to 3 mL of scintillation fluid, prior to counting with a Packard Tri-carb 2100TR liquid scintillation analyzer. All reactions were performed in triplicate.

RESULTS

Isolation and Purification of Rhodochelin from *R. jostii* RHA1. In bacteria, the biosynthesis of siderophores is a tightly regulated event. In order to force the organism to produce these secondary metabolites, *R. jostii* was cultivated in M9 minimal medium under iron starvation conditions (see Figure 2). The production of an iron-scavenging compound was confirmed via chromeazurol S (CAS) liquid assay and could be observed two days after inoculum. The supernatant was subsequently extracted with XAD-16 resin. The application of iron-deficient conditions resulted in two major metabolites. The first one (retention time, t_R 17.6 min) showed a m/z of 572.2201 ($[M + H]^+$), whereas the second one (t_R 21.4 min) displayed a m/z of 414.1510 ($[M + H]^+$).

In order to obtain sufficient amounts of sample material for NMR structure elucidation studies, the bacterial culture was scaled-up to 5 L volume. Preparative HPLC showed a similar "two-peaks" profile, and single fractions were analyzed via CAS liquid assay and ESI-MS and pooled according to their m/z (Supporting Information Figure 4). The following studies were performed utilizing the compound with a m/z of 572.2, which was obtained with a yield up to 13 mg/L.

MS Analysis of Rhodochelin Composition. In order to acquire additional information about the individual building blocks of rhodochelin and their connectivity, collision induced dissociation (CID) experiments were carried out. Rhodochelin MS^2 fragmentation revealed a major fragment ion with a m/z of 396.1402 ($[M + H]^+$, calculated 396.1401), due to the loss of a *δ-N*-formyl-*δ-N*-hydroxyornithine (fhOrn) moiety. MS^3 experiments of the aforementioned fragment ion led to the formation of two additional ions with a m/z of 177.0870 and m/z of 238.0711. The first one could easily be associated with a second fhOrn residue ($[M + H]^+$, calculated 177.0870), whereas the second one is indicative for a charged DHB-Thr dipeptide ($[M]^+$, calculated 238.0710) (Supporting Information Figure 5). In summary, the fragmentation studies led to the conclusion that rhodochelin is built up by two moieties of fhOrn, one moiety of 2,3-DHB, and one moiety of Thr (directly connected to the aryl residue). High resolution MS confirmed that rhodochelin exhibits an exact m/z of 572.2201 ($[M + H]^+$, calculated 572.2198), consistent with the molecular mass of a compound with a chemical formula of $C_{23}H_{33}N_5O_{12}$. Ferri-rhodochelin displayed a m/z of 625.1299 (calculated 625.1302).

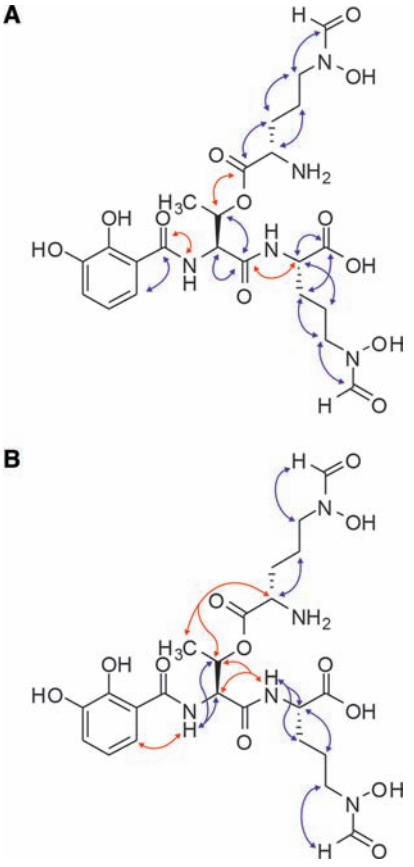


Figure 3. The structure of rhodochelin as determined by NMR. (A) Long-range $^1H-^{13}C$ correlations observed in H_2O/D_2O (9:1) at 283 K: blue arrows indicate intraresidue contacts, red arrows indicate long-range inter-residue contacts. (B) NOE contacts observed in H_2O/D_2O (9:1) at 283 K: blue arrows indicate intraresidue contacts; red arrows indicate long-range inter-residue contacts.

High resolution MS experiments of the second compound extracted from the culture supernatant revealed this compound to be a derivative of rhodochelin, lacking the second fhOrn moiety (m/z 414.1510 $[M + H]^+$, calculated 414.1507).

Structure Elucidation of Rhodochelin via NMR. On the basis of the knowledge that rhodochelin contains one 2,3-DHB, one threonine (Thr), and two *δ-N*-formyl-*δ-N*-hydroxyornithine (fhOrn) moieties, the final structure was determined by NMR spectroscopy (Supporting Information Figure 6). The assigned ^1H, ^{13}C, and ^{15}N chemical shifts are listed in Supporting Information Table 2. The ^1H spectrum showed two doublets at 9.169 and 8.505 ppm for the amide proton of Thr^1 and $fhOrn^2$, respectively. Two singlets at 7.893 and 7.826 ppm were observed for the formamide proton of $fhOrn^2$ and $fhOrn^3$, respectively. A second set of peaks was also observed in this region, which corresponds to a minor stereoisomer of the siderophore in water at 283 K. Only chemical shifts of the major stereoisomer are listed in Supporting Information Table 2. Two cross-peaks were observed in the $^1H-^{15}N$ HSQC spectrum, which verified the presence of two amide bonds in the sequence. TOCSY cross-peaks

dx.doi.org/10.1021/ja1109453 |*J. Am. Chem. Soc.* 2011, 133, 4587–4595

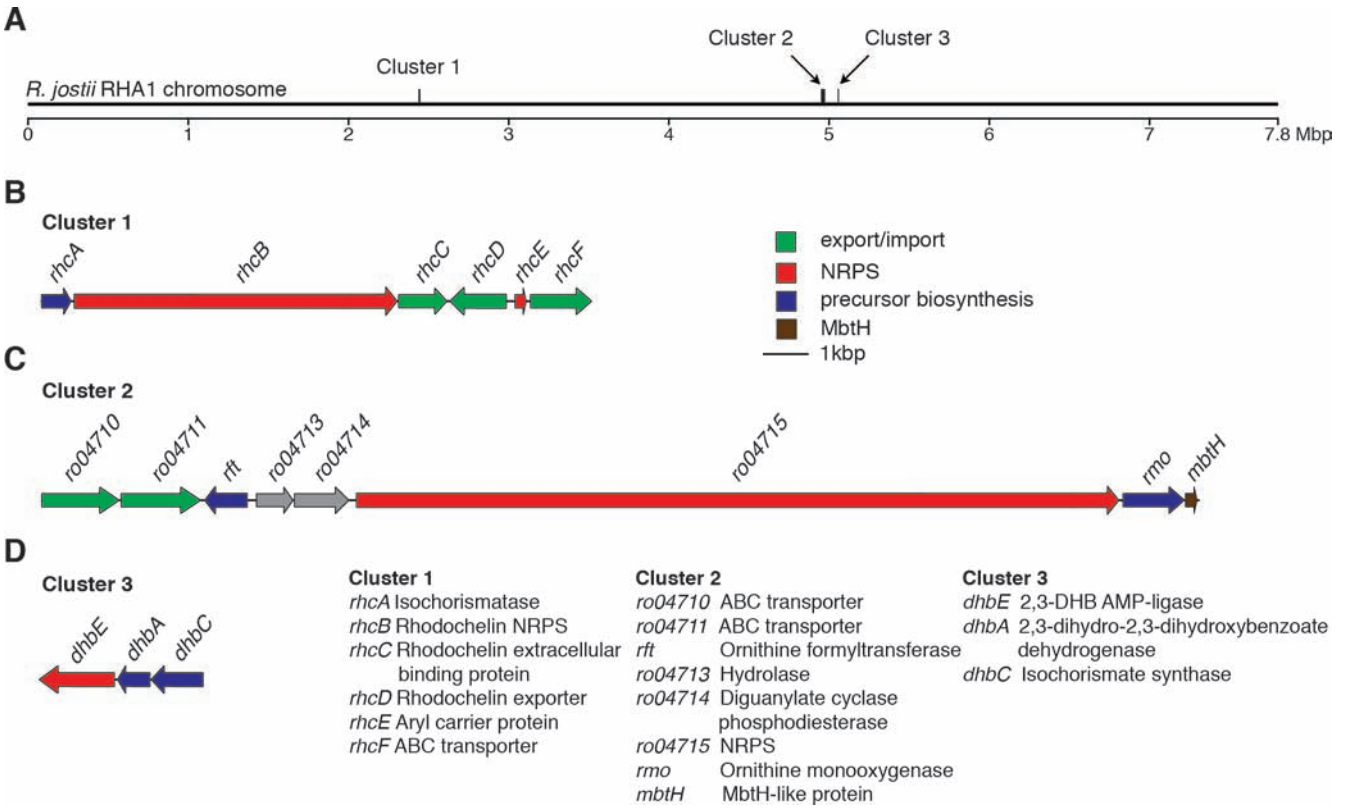


Figure 4. (A) Representation of the linear *R. jostii* RHA1 chromosome and localization of the three gene clusters involved in rhodochelin biosynthesis. (B) Schematic overview of the *rhc* gene cluster responsible of rhodochelin assembly. (C) Cluster 2 encodes for the tailoring enzymes required for the biosynthesis of L-fhOrn. (D) The *dhb* gene cluster responsible for the synthesis of 2,3-DHB. Genes are differently colored by proposed function.

confirmed the presence of one threonine and two ornithines in the compound. NOE contacts between the amide proton of Thr^1 (NH^1) and H^4 of 2,3-DHB and NH^2 and $H\alpha^1$ were observed, which revealed the partial structure 2,3-DHB-Thr^1-$fhOrn^2$. Furthermore, long-range NOE contacts of $H\beta^1$ and $H\gamma^1$ with $H\alpha^3$ were detected. A long-range $^1H-^{13}C$ correlation was observed between $H\beta^2$ and the carbonyl carbon of $fhOrn.^3$ Therefore, putting all these long-range connections together, we established an ester bond between the carbonyl of $fhOrn^3$ and the side chain hydroxyl of $Thr.^1$ The structure for our siderophore is thus presented in Figure 3 and Supporting Information Figure 6. The corresponding COSY, ROESY, and HMBC spectra are given as Supporting Information Figures 7−9.

Assignment of Rhodochelin Stereochemistry. The assignment of amino acid stereochemistry was carried out by derivatization of a rhodochelin acid hydrolysate with FDAA. The derivatized mixture was then subjected to HPLC-MS and the comparison with synthetic amino acid standards revealed the sole presence of L-Thr and L-hOrn as siderophore constituents (Supporting Information Figure 10).

Genome Mining Approach To Identify Genes Involved in Rhodochelin Biosynthesis. In order to associate a gene cluster to rhodochelin biosynthesis, a genome-mining analysis of *R. jostii* RHA1 identified three gene clusters putatively responsible for siderophore assembly, export, and uptake (Figure 4).

The first cluster is composed of six genes and is located in a region approximately covering 12 kbp between ORFs *RHA1_ro02318* and *RHA1_ro02323*: these genes have been renamed from *rhcA* to *rhcF*. *rhcB* is the largest one (7.1 kbp) and encodes for a typical modular NRPS, composed of two complete modules and a terminal thioesterase (TE) domain. Substrate-specificity prediction for the adenylation (A) domains[28] (Table 1) proposed a preference for L-Thr for the first module, and L-Asp for the second one.

Table 1. Substrate Specificity Prediction for the Adenylation Domains of the NRPS RhcB

A domain	active site residues	substrate	product
RhcB-A_1	DFWNVGMVHK		
CDA PSI-A_2	DFWNVGMVHK	L-Thr	CDA
RhcB-A_2	DLWGMGAVNK		
CDA PSI-A_4	DLTKIGAVNK	L-Asp	CDA

The protein sequences of *rhcA* and *rhcE* showed sequence homology to the single domains of DhbB found in the bacillibactin (*Bacillus subtilis*) gene cluster. RhcA has been annotated as an isochorismatase (ICL), resembling the N-terminal domain of DhbB, whereas RhcE is homologous to the C-terminal aryl carrier protein (ArCP) domain.[29] The remaining three ORFs located in the cluster are involved in siderophore export (*rhcD*) and uptake: *rhcC* has been proposed being an extracellular Fe^{3+}/hydroxamate binding protein, whereas *rhcF* encodes for an ABC-2 type transporter.

The absence of genes encoding for tailoring enzymes involved in the biosynthesis of the fhOrn moiety led to the identification of a second gene cluster responsible for siderophore biosynthesis located in another genomic region. It is composed of eight genes, spanning a DNA region of approximately 25 kbp, located between ORFs *RHA1_ro04710* and *RHA1_ro04717*. The largest

dx.doi.org/10.1021/ja1109453 |*J. Am. Chem. Soc.* 2011, 133, 4587–4595

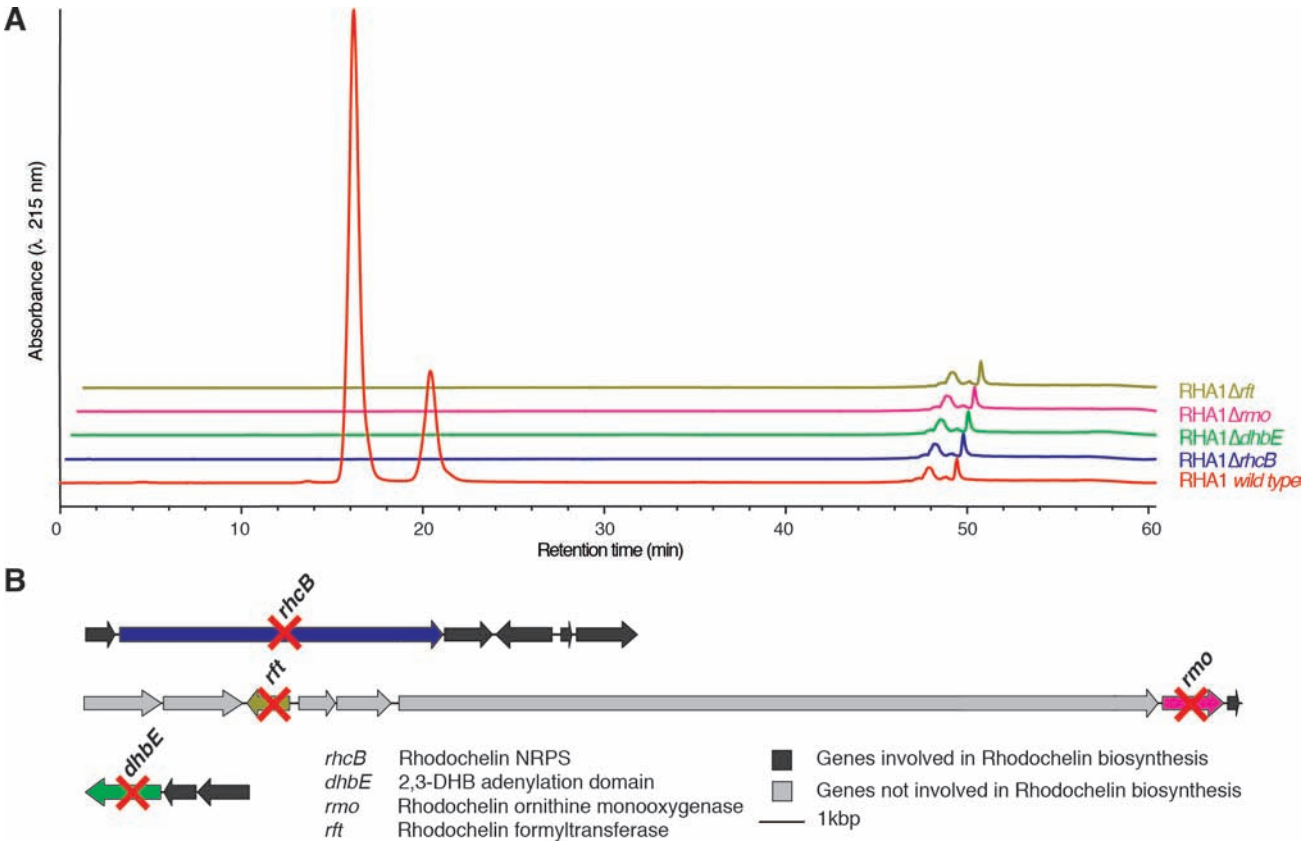


Figure 5. (A) Comparative HPLC-MS profiles obtained for *R. jostii* RHA1 wild type strain and nonproducing rhodochelin mutant strains. (B) Schematic overview of the genes involved in the siderophore biosynthesis that have been deleted.

gene of the cluster is the NRPS encoding *RHA1_ro04715* (16.7 kbp). It is composed of four modules (each of the first two encompass an epimerization domain) and lacks the terminal TE domain, which is substituted by a condensation domain. This large NRPS gene shows high sequence homology (51% identity, 64% similarity) to EtcD, the NRPS responsible for erythrochelin biosynthesis, the siderophore **3** isolated from *Saccharopolyspora erythraea* (Figure 1).[11] The modular organization of the NRPS exactly resembles EtcD, and, in addition, the comparison of the adenylation domain substrate specificity prediction (Supporting Information Table 3), together with the presence of two tailoring enzymes annotated as a formyltransferase (ORF *RHA1_ro04712*, renamed *rft*) and an ornithine monooxygenase (ORF *RHA1_ro04716*, renamed *rmo*), respectively, led to the hypothesis that this cluster could be responsible for the synthesis of a foroxymithine (**4**) derivative (Figure 1).[30] In addition, the 3′ end of *rmo* is four nucleotides overlapping with the 5′ end of the following ORF (*RHA1_ro04717*). Due to the high sequence identity of the latter ORF with MbtH-like family proteins, it was renamed *mbtH*.[31] Interestingly, this gene is present as a single copy in the *R. jostii* genome. ORFs *RHA1_ro04710* and *RHA1_ro04711* are predicted to encode ABC-transporters, while *RHA1_ro04713* and *RHA1_ro04714* are not to be directly involved in rhodochelin biosynthesis and/or transport, as they share homology with hydrolases and diguanilate cyclase phosphodiesterase. Taken together these results support the hypothesis that while the *rhc* cluster is responsible for the synthesis of the catecholate−hydroxamate siderophore rhodochelin and its export and uptake, the second one hosts tailoring enzymes required for the synthesis of the L-fhOrn building block. In addition, cluster 2 contains the MbtH-like protein that was shown to be an essential cofactor for the amino acid activation in some other systems, especially under severe growth conditions, such as iron starvation.[32]

It is important to note that the incorporation of a 2,3-DHB building block into the NRPS assembly line of bacillibactin requires its activation as adenylate:[33] this reaction is carried out by the aryl acid-activating domain DhbE.[34] In addition, previous studies reported that the presence of 2,3-DHB moieties within catecholate siderophores requires additional genes necessary for its biosynthesis from the chorismate precursor.[4] In this study, no genes encoding enzymes involved in the chorismate pathway have been annotated in the surroundings of both gene clusters. A BLAST search using DhbE as a query led to the identification of a homologue (*RHA1_ro04793*) in another locus of the *R. jostii* RHA1 genome. This gene has been renamed *dhbE*. In addition, two more genes located upstream from the adenylation domain, previously annotated as an isochorismate synthase (*dhbC*) and a 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (*dhbA*), have been found, identifying all the enzymes necessary for the biosynthesis of the 2,3-DHB moiety from its chorismate precursor.[35]

A complete bioinformatic overview of the gene clusters involved in rhodochelin biosynthesis is presented in Supporting Information Table 4.

Gene Deletion Studies in *R. jostii* RHA1. In order to verify the hypothesis of distantly located genes involved in the biosynthesis of the same natural product, gene deletion studies in *R.*

dx.doi.org/10.1021/ja1109453 |*J. Am. Chem. Soc.* 2011, 133, 4587–4595

jostii RHA1 were performed. *rhcB*, *dhbE*, *rmo*, and *rft* were singularly deleted from the wild type strain, generating four new strains: RHA1Δ*rhcB*, RHA1Δ*dhbE*, RHA1Δ*rmo*, and RHA1Δ*rft*, respectively (Supporting Information Figure 11). The use of a markerless in-frame gene deletion approach has the advantage to avoid any polar effects on downstream genes.[36] The new strains grew in minimal medium, similar to the wild type. Culture supernatants were analyzed for CAS activity, but no strain was capable of producing an iron-chelating compound (Supporting Information Figure 12). HPLC-MS traces of the supernatants were lacking the "two-peak" profile, typical for the rhodochelin production (Figure 5).

RHA1Δ*rmo* and RHA1Δ*rft* extracts were furthermore analyzed via extracted ion chromatograms (EIC) for masses corresponding to rhodochelin derivatives lacking the *N*-hydroxy or *N*-formyl groups. MS analysis of these supernatant extractions confirmed an abolished production of rhodochelin. In addition, no foroxymithine derivative has been detected. These results clearly demonstrate that these four genes are directly associated with rhodochelin biosynthesis, even if they are located in three different regions within the bacterial chromosome.

DhbE ATP/PP$_i$ Exchange. In order to verify the adenylating activity and the substrate specificity of DhbE, the corresponding gene was cloned as an N-terminal His-tag fusion in pET28a(+), heterologously produced in *E. coli* and purified via affinity chromatography. Several amino acids and two aryl acids were used to analyze the substrate specificity of the enzyme. In ATP/PP$_i$ exchange assay, the recombinant DhbE showed a distinct preference for 2,3-DHB, validating the bioinformatic prediction (Supporting Information Figure 13).

■ DISCUSSION

In the recent years, increasing amount of sequenced microbial genomes has revealed the presence of an impressive number of secondary metabolite gene clusters, most of them considered "orphan" with respect to their natural product.[12] Furthermore, since the discovery of the possibility of single strains to produce many natural products, the concept "One Strain - MAny Compounds" (OSMAC) has been introduced; thus, the interest in known species to uncover new secondary metabolites is increasing.[37] Despite the use of genome mining and its successful application, the identification of new natural products still remains challenging. In fact, without any experimental proof, it is difficult to define if an orphan cluster is silent, because it is not functional, the metabolite cannot be detected due to analytical detection limits, or the laboratory cultivation methods are inappropriate for metabolite production.[38]

The continuous interest in the *Rhodococcus* genus as a bioremediation and bioconversion tool has shifted the focus to regard these species as natural product producing strains.[14] As streptomycetes, rhodococci belong to the actinomycetal order and thus are inclined to a putative extensive secondary metabolism. Genome mining methods aim at the discovery of secondary metabolites;[39] therefore, the complete genome sequence of the industrially relevant *Rhodococcus jostii* RHA1 provides an excellent opportunity for new natural product discovery.[15]

In this work, we report the isolation of rhodochelin, a new siderophore from *R. jostii* RHA1, also known to be the first secondary metabolite isolated from this strain. NMR and MS^n studies disclosed the branched tetrapeptidic structure, which is composed of a linear assembly of 2,3-DHB, L-Thr, and L-fhOrn. The fourth building block (an additional L-fhOrn moiety) is attached to the main tripeptide scaffold through an unusual and characteristic ester bond with the side chain hydroxyl group of L-Thr. In addition, structure comparison of rhodochelin with heterobactin A and rhodobactin highlights the fact that these two molecules are composed of 2,3-DHB and modified ornithine residues, suggesting the presence of these two building blocks as a common motif present among the rhodoccocal siderophores.[16,17]

The complete set of gene clusters for the biosynthesis of rhodochelin has been identified by a genome mining approach. The *rhc* cluster (cluster 1) contains the complete bimodular NRPS synthetase *rhcB*, the genes involved in rhodochelin export and import (*rhcC*, *rhcD*, and *rhcF*) and two additional ORFs (*rhcA* and *rhcE*), homologues to the two distinct domains of DhbB, ICL, and ArCP, respectively.[40] Interestingly, the first three genes of this cluster (*rhcA*, *rhcB*, and *rhcC*) have previously been annotated to belong to an orphan siderophore gene cluster.[39]

The presence of 2,3-DHB within the rhodochelin structure led to the investigation of the RHA1 genome to identify all the genes involved in the biosynthesis of the aryl moiety.[35] One of the enzymes, an isochorismatase, has already been found to be encoded in the *rhc* cluster (*rhcA*). Since the biosynthesis of the aryl-capped siderophore bacillibactin requires the activation of 2,3-DHB by the stand-alone A-domain DhbE prior to the NRPS-catalyzed assembly,[34,41,42] a gene homologous to *dhbE* (cluster 3) was identified in a different genomic region, along with two other genes involved in 2,3-DHB biosynthesis, namely *dhbC* and *dhbA*. These three genes are arranged in an operon-like way and, together with *rhcA*, cover the entire 2,3-DHB pathway, from the chorismate precursor through its activation as adenylate.

It is important to note that an additional putative siderophore gene cluster has been identified in the *Rhodococcus* genome (cluster 2). Because of the overall homology of the tetramodular NRPS RHA1_ro04715 with EtcD[11] (supported by the comparison of the adenylation substrate specificity prediction, and the contextual presence of two tailoring enzymes Rmo and Rft), it is suggested that this cluster could be responsible for the synthesis of a foroxymithine derivative, which is structurally related to erythrochelin.[11,30] Bioinformatic analysis of the NRPS RHA1_ro04715 revealed that the first adenylation domain harbors the same specificity-conferring code found in the first A-domain of CchH, involved in coelichelin biosynthesis.[9] On the basis of these results, it is therefore suggested that RHA1_ro04715 also activates L-fhOrn as a cognate substrate, which is subsequently incorporated into the putative natural product.

Accordingly to the results of the genome mining, the existence of three distantly located gene clusters involved in rhodochelin assembly was proposed. This hypothesis was verified employing a gene disruption approach, showing that single key gene deletions in different clusters (*rhcB*, *dhbE*, *rmo*, and *rft* were targeted) were sufficient to inhibit siderophore biosynthesis. Until now, only one example of similar cross-talk between different NRPS gene clusters has been reported for a single natural product. It was shown that two gene clusters govern the assembly of the siderophore erythrochelin.[11,43] The *erc*/*etc* cluster encodes the enzymes responsible for precursor biosynthesis and siderophore assembly and export, whereas the *δ*-*N*-L-acetyltransferase *mcd* is located in the nonfunctional *nrps1* cluster.[44] The present results clearly demonstrate that all the genes required for the biosynthesis of rhodochelin building blocks are dispersed within three gene clusters. Interestingly, to

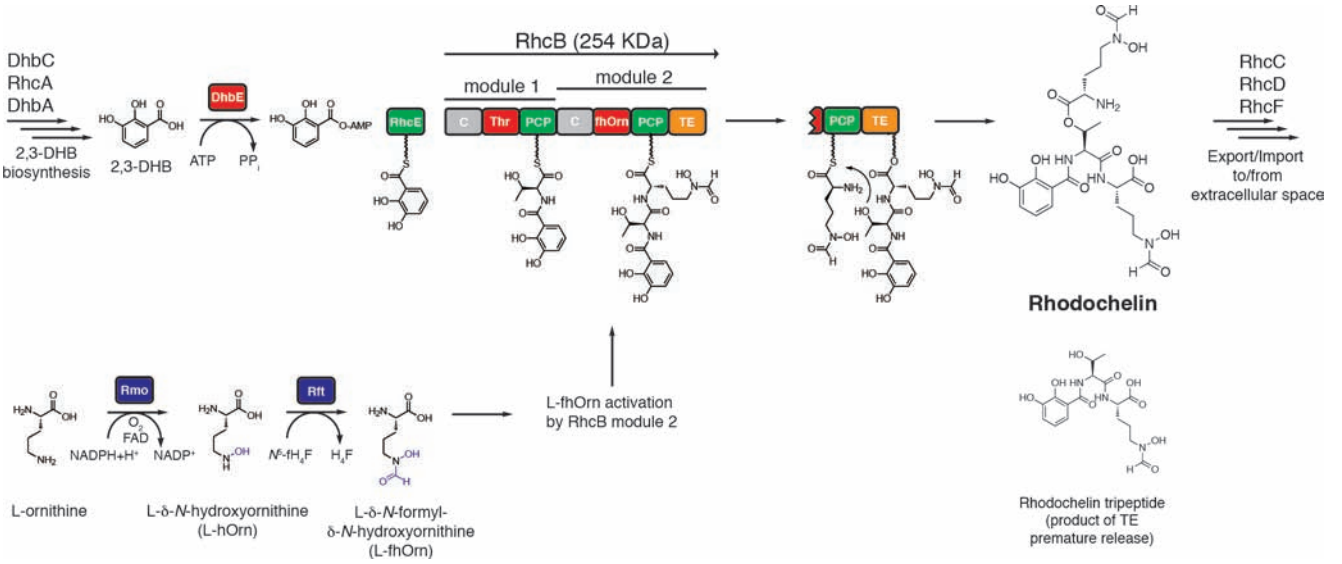


Figure 6. Proposed biosynthetic pathway for rhodochelin assembly. The two nonproteinogenic building blocks, 2,3-DHB and L-fhOrn are synthesized by the corresponding pathways and channeled to the synthetase RhcB. A detail of the ester bond formation between L-fhOrn and the side chain of L-Thr, catalyzed by the iterative TE domain is presented.

our knowledge, this is also the first example of a catecholate siderophore where the genes involved in the biosynthetic pathway of the aryl moiety are not clustered together within the same genetic locus.[4] In addition, the fact that no foroxymithine derivative was detected under the current growth conditions (even as a rescue strategy in the rhodochelin defective strains) does not exclude the NRPS RHA1_ro04715 to be functional. It is possible that *R. jostii* RHA1 could have evolved two different siderophores as a response to different environmental iron-depletion conditions, with the alternative gene cluster down-regulated under the conditions applied.

Genome comparison of *R. jostii* RHA1 with other sequenced *Rhodococcus* strains revealed that only *R. opacus* B4 shares an identical gene arrangement within the three dispersed gene clusters required for rhodochelin biosynthesis, implying an identical cross-talk mechanism for the biosynthesis of the siderophore. Interestingly, the additional sequenced strains analyzed (*R. equi* 103S and *R. erythropolis* PR4) share the contextual presence of cluster 2 and cluster 3, despite minor genetic rearrangements (gene integrations and deletions) (Supporting Information Table 5).[45] Besides, the RhcB NRPS synthetase and its *R. opacus* B4 homologue revealed 100% identity for the A-domain specificity-conferring codes, suggesting that the latter also mediates rhodochelin assembly (Supporting Information Table 6). In addition, similar results were obtained for the silent synthetase RHA1_ro04715 and its homologues in all the three strains analyzed (Supporting Information Table 7).

The results obtained in this study allow the postulation of a model for rhodochelin assembly (Figure 6). Rhodochelin assembly is initiated by DhbE, which activates 2,3-DHB that is subsequently transferred to its cognate stand-alone aryl carrier protein RhcE. RhcB assembles the tripeptide DHB-Thr-fhOrn following the classical NRPS enzymatic linear logic, similar to the fashion in which the tripeptide DHB-Gly-Thr of bacillibactin is formed.[46] From the RhcB-PCP_2 the newly assembled tripeptide is transferred to the conserved catalytic Ser of the TE domain. Subsequently, a second L-fhOrn is proposed to be activated by the respective domain (following module skipping mechanism, similar to the coelichelin assembly) and tethered to RhcB-PCP_2.[9] With the tripeptide and the monomer lying in these adjacent positions, the nucleophilic attack from the hydroxyl group of the L-Thr side chain on the L-fhOrn thioester could occur. This mechanism is consistent with an iterative TE domain following a "forward" mechanism like the proposed lactonization mechanism of DHB-Gly during enterobactin biosynthesis.[47] Therefore, the branched tetrapeptide, still attached to the TE, is hydrolytically released and exported to the extracellular space to carry out its biological function. The tripeptide found in the culture supernatants also displays iron-scavenging properties (Supporting Information Figures 2 and 3). With regards to the postulated "forward" mechanism, it is suggested that the tripeptide is the result of premature release from the TE domain during assembly. On the other hand, it cannot be excluded that it is the result of a spontaneous hydrolysis of the labile ester bond in the culture medium.

■ ASSOCIATED CONTENT

Supporting Information. FT-IR spectrum, UV−vis spectra, and HPLC-MS profiles of rhodochelin and rhodochelin tripeptide, rhodochelin purification, MS-fragmentation studies, 2D NMR spectra, HPLC-MS of rhodochelin stereochemistry assignment, PCR and CAS assay comparison of the gene deletion strains, SDS-PAGE of purified recombinant DhbE and ATP/PP_i exchange assay, list of oligonucleotide primers used, NMR chemical shifts tables, adenylation substrate specificity prediction of the NRPS RHA1_ro04715, bioinformatic table of genes annotation, genome comparison of rhodococci, adenylation substrate specificity prediction for RhcB and RHA1_ro04715 homologues, complete refs 15 and 45. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
marahiel@staff.uni-marburg.de

■ ACKNOWLEDGMENT

The authors thank Prof. Lindsay D. Eltis (University of British Columbia, Vancouver, Canada) for providing *R. jostii* RHA1, and Dr. Robert van der Geize (University of Groningen, The Netherlands) for the strain S17-1 and the plasmid pK18mobsacB. We gratefully acknowledge Deutsche Forschungsgemeinschaft and the International Max-Planck Research School (MPI for Terrestrial Microbiology, Marburg) for financial support.

■ REFERENCES

(1) Miethke, M.; Marahiel, M. A. *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 413.
(2) Wandersman, C.; Delepelaire, P. *Annu. Rev. Microbiol.* **2004**, *58*, 611.
(3) Schaible, U. E.; Kaufmann, S. H. *Nat. Rev. Microbiol.* **2004**, *2*, 946.
(4) Crosa, J. H.; Walsh, C. T. *Microbiol. Mol. Biol. Rev.* **2002**, *66*, 223.
(5) Challis, G. L. *Chembiochem* **2005**, *6*, 601.
(6) Miethke, M.; Klotz, O.; Linne, U.; May, J. J.; Beckering, C. L.; Marahiel, M. A. *Mol. Microbiol.* **2006**, *61*, 1413.
(7) Peuckert, F.; Miethke, M.; Albrecht, A. G.; Essen, L.-O.; Marahiel, M. A. *Angew. Chem., Int. Ed.* **2009**, *48*, 7924.
(8) Andrews, S. C.; Robinson, A. K.; Rodríguez-Quiñones, F. *FEMS Microbiol. Rev.* **2003**, *27*, 215.
(9) Lautru, S.; Deeth, R. J.; Bailey, L. M.; Challis, G. L. *Nat. Chem. Biol.* **2005**, *1*, 265.
(10) Dimise, E. J.; Widboom, P. F.; Bruner, S. D. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 15311.
(11) Robbel, L.; Knappe, T. A.; Linne, U.; Xie, X.; Marahiel, M. A. *FEBS J.* **2010**, *277*, 663.
(12) Zerikly, M.; Challis, G. L. *Chembiochem* **2009**, *10*, 625.
(13) Banerjee, A.; Sharma, R.; Banerjee, U. C. *Appl. Microbiol. Biotechnol.* **2002**, *60*, 33.
(14) van der Geize, R.; Dijkhuizen, L. *Curr. Opin. Microbiol.* **2004**, *7*, 255.
(15) McLeod, M. P.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 15582.
(16) Carrano, C. J.; Jordan, M.; Drechsel, H.; Schmid, D. G.; Winkelmann, G. *Biometals* **2001**, *14*, 119.
(17) Dhungana, S.; Michalczyk, R.; Boukhalfa, H.; Lack, J. G.; Koppisch, A. T.; Fairlee, J. M.; Johnson, M. T.; Ruggiero, C. E.; John, S. G.; Cox, M. M.; Browder, C. C.; Forsythe, J. H.; Vanderberg, L. A.; Neu, M. P.; Hersman, L. E. *Biometals* **2007**, *20*, 853.
(18) Miranda-CasoLuengo, R.; Prescott, J. F.; Vázquez-Boland, J. A.; Meijer, W. G. *J. Bacteriol.* **2008**, *190*, 1631.
(19) Seto, M.; Kimbara, K.; Shimura, M.; Hatta, T.; Fukuda, M.; Yano, K. *Appl. Environ. Microbiol.* **1995**, *61*, 3353.
(20) Bauchop, T.; Elsden, S. R. *J. Gen. Microbiol.* **1960**, *23*, 457.
(21) Schwyn, B.; Neilands, J. B. *Anal. Biochem.* **1987**, *160*, 47.
(22) Bhushan, R.; Bruckner, H. *Amino Acids* **2004**, *27*, 231.
(23) Lin, Y.; Miller, M. *J. Org. Chem.* 1999, *64*, 7451–7458.
(24) Horton, R. M.; Cai, Z. L.; Ho, S. N.; Pease, L. R. *BioTechniques* **1990**, *8*, 528.
(25) Schäfer, A.; Tauch, A.; Jäger, W.; Kalinowski, J.; Thierbach, G.; Pühler, A. *Gene* **1994**, *145*, 69.
(26) Simon, R.; Priefer, U.; Pühler, A. *Nat. Biotechnol.* 1983, *1*, 784–791.
(27) van der Geize, R.; Hessels, G. I.; van Gerwen, R.; van der Meijden, P.; Dijkhuizen, L. *FEMS Microbiol. Rev.* **2001**, *205*, 197.
(28) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. *Chem. Biol.* **1999**, *6*, 493.
(29) Gehring, A. M.; Mori, I.; Walsh, C. T. *Biochemistry* **1998**, *37*, 2648.
(30) Umezawa, H.; Aoyagi, T.; Ogawa, K.; Obata, T.; Iinuma, H.; Naganawa, H.; Hamada, M.; Takeuchi, T. *J. Antibiot.* **1985**, *38*, 1813.
(31) Quadri, L. E.; Sello, J.; Keating, T. A.; Weinreb, P. H.; Walsh, C. T. *Chem. Biol.* **1998**, *5*, 631.
(32) Felnagle, E. A.; Barkei, J. J.; Park, H.; Podevels, A. M.; McMahon, M. D.; Drott, D. W.; Thomas, M. G. *Biochemistry* **2010**, *49*, 8815.
(33) Miethke, M.; Bisseret, P.; Beckering, C. L.; Vignard, D.; Eustache, J.; Marahiel, M. A. *FEBS J.* **2006**, *273*, 409.
(34) May, J. J.; Kessler, N.; Marahiel, M. A.; Stubbs, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12120.
(35) Rowland, B. M.; Grossman, T. H.; Osburne, M. S.; Taber, H. W. *Gene* **1996**, *178*, 119.
(36) Raynal, A.; Karray, F.; Tuphile, K.; Darbon-Rongere, E.; Pernodet, J. L. *Appl. Environ. Microbiol.* **2006**, *72*, 4839.
(37) Bode, H. B.; Bethe, B.; Hofs, R.; Zeeck, A. *Chembiochem* **2002**, *3*, 619.
(38) Challis, G. L. *Microbiology (Reading, U.K.)* **2008**, *154*, 1555.
(39) Nett, M.; Ikeda, H.; Moore, B. S. *Nat. Prod. Rep.* **2009**, *26*, 1362.
(40) Gehring, A. M.; Bradley, K. A.; Walsh, C. T. *Biochemistry* **1997**, *36*, 8495.
(41) Finking, R.; Marahiel, M. A. *Annu. Rev. Microbiol.* **2004**, *58*, 453.
(42) Marahiel, M. A. *J. Pept. Sci.* **2009**, *15*, 799.
(43) Lazos, O.; Tosin, M.; Slusarczyk, A. L.; Boakes, S.; Cortés, J.; Sidebottom, P. J.; Leadlay, P. F. *Chem. Biol.* **2010**, *17*, 160.
(44) Oliynyk, M.; Samborskyy, M.; Lester, J. B.; Mironenko, T.; Scott, N.; Dickens, S.; Haydock, S. F.; Leadlay, P. F. *Nat. Biotechnol.* **2007**, *25*, 447.
(45) Letek, M. et al. *PLoS Genet. [Online]* **2010**, *6*. http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1001145, accessed on January 26, 2011.
(46) May, J. J.; Wendrich, T. M.; Marahiel, M. A. *J. Biol. Chem.* **2001**, *276*, 7209.
(47) Shaw-Reid, C. A.; Kelleher, N. L.; Losey, H. C.; Gehring, A. M.; Berg, C.; Walsh, C. T. *Chem. Biol.* **1999**, *6*, 385.