the FEBS Journal

Erythrochelin – a hydroxamate-type siderophore predicted from the genome of *Saccharopolyspora erythraea*

Lars Robbel, Thomas A. Knappe, Uwe Linne, Xiulan Xie and Mohamed A. Marahiel

Department of Chemistry, Philipps-University Marburg, Germany

Keywords
genome mining; nonribosomal peptide synthetase; radiolabeling; secondary metabolites; siderophore

Correspondence
M. A. Marahiel, Department of Chemistry, Philipps-University Marburg, D-35043 Marburg, Germany
Fax: +49 (0) 6421 282 2191
Tel: +49 (0) 6421 282 5722
E-mail: marahiel@staff.uni-marburg.de

(Received 4 October 2009, revised 10 November 2009, accepted 23 November 2009)

doi:10.1111/j.1742-4658.2009.07512.x

The class of nonribosomally assembled siderophores encompasses a multitude of structurally diverse natural products. The genome of the erythromycin-producing strain *Saccharopolyspora erythraea* contains 25 secondary metabolite gene clusters that are mostly considered to be orphan, including two that are responsible for siderophore assembly. In the present study, we report the isolation and structural elucidation of the hydroxamate-type tetrapeptide siderophore erythrochelin, the first nonribosomal peptide synthetase-derived natural product of *S. erythraea*. In an attempt to substitute the traditional activity assay-guided isolation of novel secondary metabolites, we have employed a dedicated radio-LC-MS methodology to identify nonribosomal peptides of cryptic gene clusters in the industrially relevant strain. This methodology was based on transcriptome data and adenylation domain specificity prediction and resulted in the detection of a radiolabeled ornithine-inheriting hydroxamate-type siderophore. The improvement of siderophore production enabled the elucidation of the overall structure via NMR and MS^n analysis and hydrolysate-derivatization for the determination of the amino acid configuration. The sequence of the tetrapeptide siderophore erythrochelin was determined to be D-α-*N*-acetyl-δ-*N*-acetyl-δ-*N*-hydroxyornithine-D-serine-cyclo(L-δ-*N*-hydroxyornithine-L-δ-*N*-acetyl-δ-*N*-hydroxyornithine). The results derived from the structural and functional characterization of erythrochelin enabled the proposal of a biosynthetic pathway. In this model, the tetrapeptide is assembled by the nonribosomal peptide synthetase EtcD, involving unusual initiation- and cyclorelease-mechanisms.

Introduction

Bacterial growth is strongly influenced by the availability of iron as an essential trace element employed as a cofactor [1]. The fact that the bioavailability of iron is challenging for most microorganisms because it is mostly found in the Fe(III) (ferric iron) redox state, forming insoluble $Fe(OH)_3$ complexes, has led to the evolutionary development of highly efficient iron uptake systems. In response to iron starvation, many microorganisms produce and secrete iron-scavenging compounds (generally < 1 kDa) termed siderophores, with a high affinity for ferric iron ($K_f = 10^{22}$ to 10^{49} M^{-1}) [2]. After the extracellular binding of iron, the siderophores are reimported into the cell after recognition by specific receptors and iron is released from the chelator complex and subsequently channelled to the intracellular targets [3–5]. Siderophores in general

Abbreviations
A, adenylation domain; ac-haOrn, α-*N*-acetly-δ-*N*-acetyl-δ-*N*-hydroxyornithine; C, condensation domain; CAS, chromazurol S; DKP, diketopiperazine; E, epimerization domain; FDAA, *N*-α-(2,4-dinitro-5-fluorophenyl)-L-alaninamide; haOrn, δ-*N*-acetyl-δ-*N*-hydroxyornithine; HMBC, heteronuclear multiple bond correlation; HSQC, heteronuclear single-quantum correlation; hOrn, δ-*N*-hydroxyornithine; NRP, nonribosomal peptide; NRPS, nonribosomal peptide synthetase; PCP, peptidyl carrier protein.

constitute a class of structurally diverse natural products that are classified into two main groups based on the mechanism of biosynthesis. Common structural features of siderophores are catecholate, hydroxamate or carboxylate functionalities conferring chelating properties for the octahedral coordination of ferric iron. Some siderophores are assembled via a template-directed manner by multimodular nonribosomal peptide synthetases (NRPSs). The class of nonribosomally assembled siderophores can be exemplified by enterobactin **1** (*Escherichia coli*), coelichelin **2** (*Streptomyces coelicolor*) and fuscachelin A **3** (*Thermobifida fusca* YX) (Fig. 1) [6–8]. The second class is known as NRPS-independent siderophores and involves a novel family of synthetases, represented by IucA and IucC, which are responsible for aerobactin (*E. coli* K-12) biosynthesis [9,10]. Siderophores of NRPS-independent origin encompass desferrioxamine E (*Streptomyces coelicolor* M145), putrebactin (*Shewanella putrefaciens*) and further compounds [11,12]. The biosynthetic genes of these secondary metabolites are usually clustered within one operon, showing coordinated transcriptional regulation [13].

Extensive bioinformatic analysis of these biosynthetic clusters allowed the prediction of the incorporated building blocks and the mechanism of iron coordination [14,15]. This genomics-based characterization of natural products has been successfully applied in the discovery of the siderophores coelichelin and fuscachelin A. Because siderophores often function as virulence factors in pathogens, the interest in the structural and functional characterization of these compounds is growing and may result in the synthesis of specific inhibitors based on the structure of the pathogen siderophore [16].

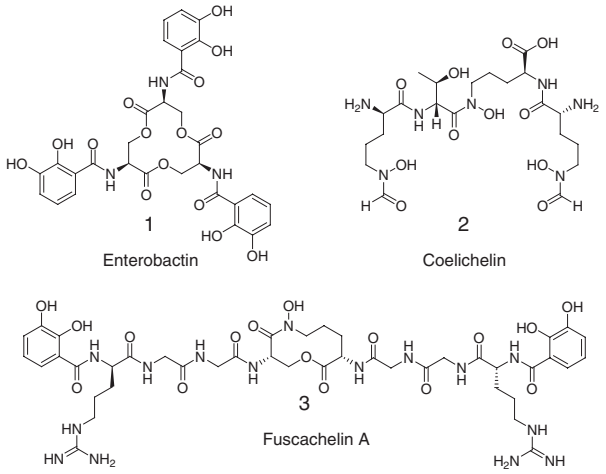

Fig. 1. Representatives of nonribosomally assembled oligopeptide siderophores: the catecholate siderophore enterobactin **1**, the hydroxamate siderophore coelichelin **2** and the decapeptide fuscachelin A **3**. The latter two siderophores were discovered via genome mining methodology.

A promising approach for the isolation of secondary metabolites, predicted from genome analysis, results from feeding experiments of a predicted precursor molecule in an isotopically labeled form to cultures of the target strains. Direct identification of the incorporated label either by NMR, if using ^{15}N-enriched precursors, or by radio-LC-MS, if employing ^{14}C-labeled building blocks, facilitates the identification of new natural products of the orphan pathway and has successfully been applied in the discovery of orfamide A [17]. The accurate prediction of adenylation domain specificity was found to be crucial for successful mining and structural prediction and is the basis of the methodology applied in the present study [7,8]. This approach was applied for the aerobic mesophilic Gram-positive filamentous actinomycete *Saccharopolyspora erythraea* NRRL 23338, the producer strain of the macrolide polyketide erythromycin. The recently sequenced and annotated genome comprises 8.2 mb and contains at least 25 biosynthetic operons for the production of known or predicted secondary metabolites, including two gene clusters for the biosynthesis of siderophores [18,19]. Transcriptome data for *S. erythraea* using GeneChip DNA microarrays, collected by Peano *et al.* [20], indicate an up-regulation of gene expression associated with siderophore assembly under specific conditions.

In the present study, we report the identification and isolation of erythrochelin, a hydroxamate-type siderophore produced by the industrially relevant strain *S. erythraea*, utilizing a novel radio-LC-MS-guided genome mining methodology. Structural and functional characterization was carried out relying on NMR and MS^n analysis and derivatization-based elucidation of the overall stereochemistry. Furthermore, the functional properties of erythrochelin acting as an iron-chelating compound were investigated. On the basis of the analysis of the *S. erythraea* genome, transcriptome and the structural characterization, an NRPS-dependent assembly of erythrochelin mediated by a tetramodular NRPS is proposed.

Results

The *etc* gene cluster in *S. erythraea*

Analysis of the sequenced and annotated genome of *S. erythraea* led to the discovery of two NRPS-gene clusters linked to siderophore biosynthesis and transport [18]. One of the two was predicted to encode for a mixed hydroxamate/catecholate-type siderophore

(Nrps3), whereas the second operon was envisaged to encode a tetramodular NRPS putatively capable of assembling a hydroxamate-type siderophore (Fig. 2). In this operon, 11 coding sequences are clustered in a region covering 28.8 kb, with an average GC content of 71.2%.

The NRP synthetase encoded by *etcD* (*sace_3035/nrps5*) comprises four modules, each containing the essential condensation (C), adenylation (A) and peptidyl carrier protein (PCP) domains. In addition, modules 1 and 2 contain an epimerization (E) domain each, which is responsible for stereoconversion of the accepted L-amino acids to D-isomers, indicating the presence of two D-configured residues in the assembled product. The N-terminal region of module 1 shares a high degree of homology to condensation domains, suggesting the function of an initiation module mediating the condensation of an external building block with the PCP-tethered substrate. Module 4 contains a C-terminal C-domain instead of a thioesterase domain commonly responsible for product release through hydrolytic cleavage or macrocyclization [21]. Upstream of *etcD*, a gene with high sequence homology to characterized L-ornithine hydroxylases (*etcB*) is located. On the basis of the proposed function of EtcB, the incorporation of δ-*N*-hydroxyornithine residues into the readily assembled oligopeptide was predicted [22]. Furthermore, genes present in the cluster encode for proteins traditionally associated with secondary metabolite biosynthesis and siderophore transport: a transcriptional regulator (*etcA*), MbtH-like protein (*etcE*) and proteins for siderophore export and uptake (*etcCFGK*). A bioinformatic overview of the encoded proteins and the corresponding functions is provided in Table S1.

The amino acid specificity of the synthetase was predicted by using a methodology comparing active-site residues of known NRPS adenylation domains with the adenylation domains found in EtcD (Table 1) [23–25]. The first adenylation domain (A_1) is predicted to activate L-arginine but reveals only 70% identity of the residues determining the specificity to MycC, suggesting the activation of a structurally analogous building block. MycC itself represents a NRPS-termination module involved in the assembly of microcystin by *Microcystis aeruginosa* PCC7806, predicted to activate L-arginine [26]. A_2 and A_3 are predicted to activate L-serine and L-δ-*N*-hydroxyornithine (L-hOrn), respectively, as found in the assembly of enterobactin and coelichelin [6,7]. The C-terminal adenylation domain A_4 again is predicted to activate L-arginine, displaying 60% identity to the characterized A-domain of MycC. Interestingly, A_1 and A_4 inherit a highly identical (90%) specificity-determining residue pattern, leading to the assumption that both activate the same substrate (Table S2A). On the basis of the bioinformatic analysis of the *etc* gene cluster, it was predicted that the assembled tetrapeptide consists of L-hOrn, L-Ser and two building blocks analogous to L-Arg.

Table 1. Comparison of active-site residues determining the adenylation domain specificity of EtcD with known adenylation domains. Variations in the residue pattern are highlighted in bold. EntF, enterobactin synthetase; CchH, coelichelin synthetase.

A-domain	Active site residues	Substrate	Product
A_1	D V W A L G A V N K		
MycC	D V W **T I** G A V **D** K	L-Arg	Microcystin
A_2	D V W H F S L V D K		
EntF	D V W H F S L V D K	L-Ser	Enterobactin
A_3	D M E N L G L I N K		
CchH-A_3	D M E N L G L I N K	L-hOrn	Coelichelin
A_4	D V F A L G A V N K		
MycC	D V **W T I** G A V **D** K	L-Arg	Microcystin

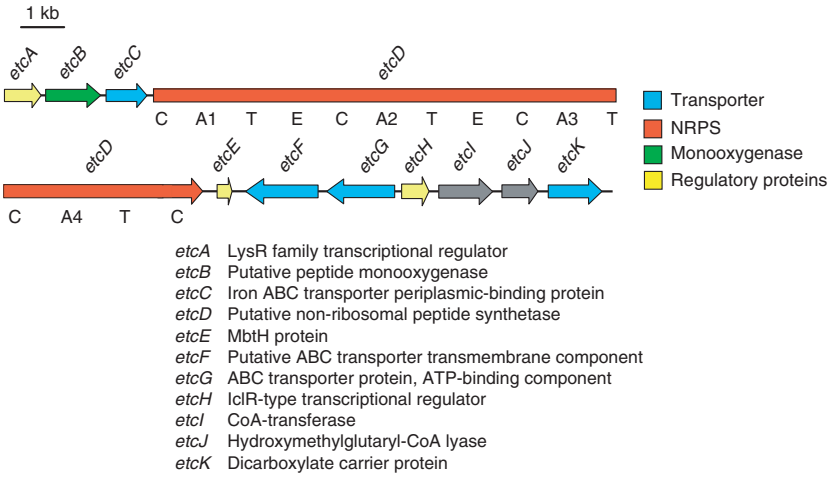

Fig. 2. Schematic overview of the *etc* gene cluster. Putative functions of the proteins encoded within the operon are shown based on BLAST analysis. Apart from the core components for siderophore biosynthesis, genes encoding for exporters and importers of the siderophore, as well as typical transcriptional regulators for secondary metabolism, are found, determining the boundaries of the cluster.

Identification and isolation of a hydroxamate-type siderophore via radio-LC-MS

On the basis of the transcriptome data for *S. erythraea* NRRL 23338 grown in SCM medium that clearly show an up-regulated gene expression of the NRPS encoding *etc* cluster, which is linked to siderophore biosynthesis, siderophore production was investigated throughout several growth phases [20]. Secondary metabolite identification and isolation is often challenging as a result of a high medium complexity or low amounts of the target compounds. To circumvent these challenges, a radio-LC-MS-guided genome mining approach was applied by feeding the nonproteinogenic amino acid ^{14}C-L-ornithine, as predicted to be incorporated into the tetrapeptide siderophore, to cultures of *S. erythraea*. These experiments were carried out in rich SCM medium, as previously employed in transcriptome analysis [20]. Extraction of the supernatant followed by radio-LC-MS analysis revealed the radiolabeling of a compound with a measured *m/z* of 604.27 $[M+H^+]$ (Fig. 3A). The incorporation of radiolabeled L-Orn was determined to be 2% of the total amount of radioactivity fed to the cultures employing the rich SCM medium. In addition, an extraction of the SCM medium supernatant after 4 days of growth, subsequent preparative HPLC fractionation and chromazurol S (CAS: an indicator of iron scavenging properties) liquid assay analysis of the fractions revealed a CAS-reactive compound (Fig. S1)

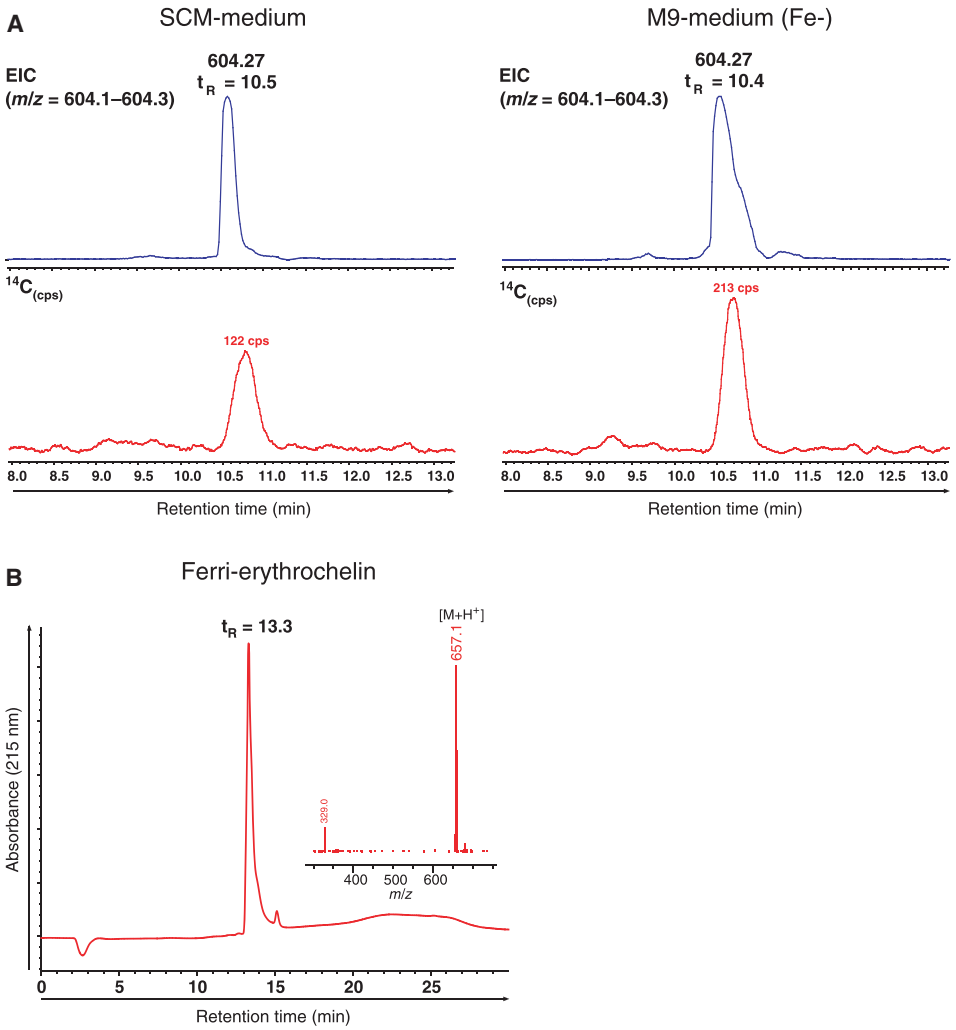

Fig. 3. (A) Radio-LC-MS profiles of radiolabeling experiments employing nonproteinogenic ^{14}C-L-Orn. In both cases, the incorporation of the radiolabel occurred (red trace), displaying a discrete *m/z* = 604.27 ($[M+H^+]$) in the extracted ion chromatogram (EIC). (B) ESI-MS analysis of ferri-erythrochelin; retention time = 13.2 min. Skimmer fragmentation was completely abolished when analyzing ferri-erythrochelin, which is indicative of a structurally rigid conformation induced by iron chelation.

[27]. The coelution of a multitude of compounds in the CAS assay positive fraction impeded the direct MS-based detection and isolation of the siderophore. To reduce media complexity and to facilitate the isolation procedure, a radiolabeling experiment was carried out in iron-deficient M9-minimal medium. The incorporation of the radiolabel increased from 2% to 4% (Fig. 3B), whereas coeluting compounds were reduced, as observed in the total ion chromatogram. To isolate the siderophore in sufficient amounts for NMR structure elucidation, a large-scale cultivation of *S. erythraea* in iron-deficient modified M9 medium was carried out, giving rise to siderophore production of 10.2 mg·L^{-1} culture (Fig. 4). The physiological function of the siderophore for iron uptake was confirmed by comparing supernatant extractions of *S. erythraea* cultures grown in the absence or presence of iron. The presence of iron in the medium completely supressed siderophore production (Fig. S2). UV/visible spectra of ferri-siderophore compared to the unloaded *apo*-form show the typical absorption spectrum for hydroxamate-type siderophores (λ_{max} = 440 nm), furthermore confirming the iron-chelating function of the product (Fig. S3). Additionally, the stochiometry of the Fe(III):siderophore-complex was determined to be 1 : 1 by UV/visible and MS analysis, indicating the presence of six Fe(III)-coordinating groups (Fig. 3C).

Structure elucidation by NMR

The amino acid sequence and the final structure of the siderophore were determined using NMR methodology (Fig. 5). The ^{1}H spectrum revealed the presence of four amide protons at 7.96, 7.74, 8.08 and 8.12 p.p.m. (Fig. S4). Four cross peaks were observed in the ^{1}H–^{15}N heteronuclear single-quantum correlation (HSQC) spectrum, which verified the presence of four amino acids in the sequence. TOCSY cross peaks confirmed the presence of three ornithines and one serine in the compound. Two strong singlets at 1.84 and 1.96 p.p.m. for three and six protons, respectively, revealed the presence of three acetyl groups, of which two are attached to very similar amino acids in the sequence. The observed long-range ^{1}H–^{13}C correlations showed the two acetyl groups to be connected to the δ-amino group of two δ-*N*-hydroxyornithines,

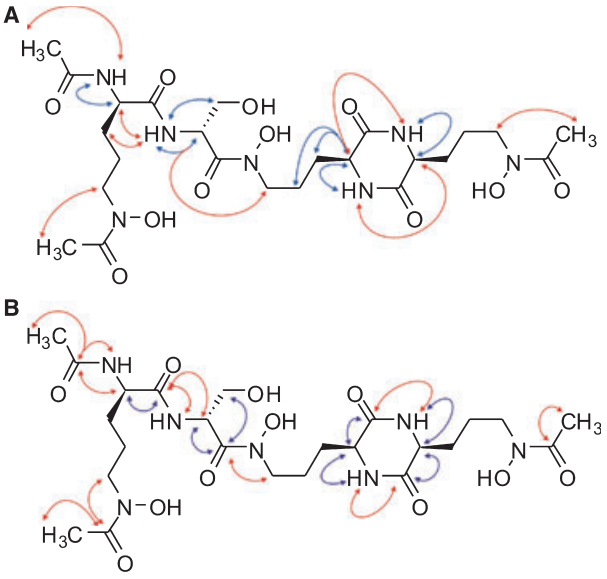

Fig. 5. The structure of erythrochelin as determined by NMR. NMR contacts are indicated by arrows. Blue arrows indicate intra-residue contacts; red arrows indicate long-range inter-residue contacts. (A) Long-range ^{1}H–^{13}C correlations observed in dimethylsulfoxide (300 K). (B) NOE contacts observed in dimethylsulfoxide (300 K). Sequential NOE contacts observed between $hOrn_3$ and $haOrn_4$ confirm the presence of a DKP moiety.

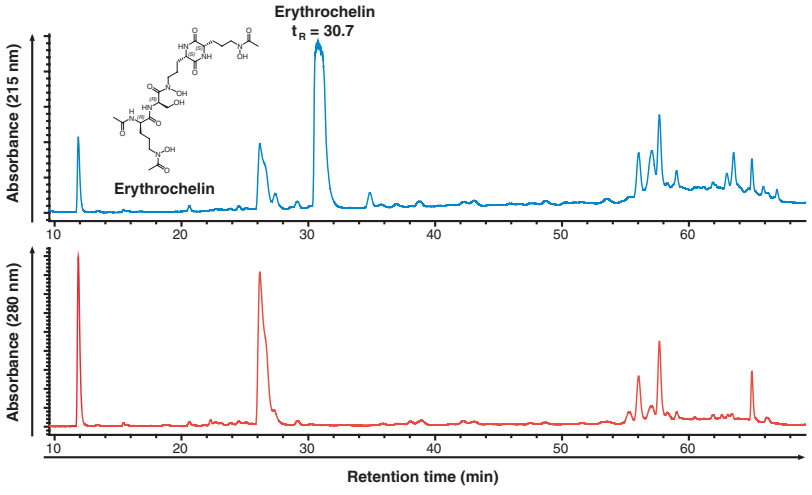

Fig. 4. Preparative HPLC profile of a XAD16 resin extraction of iron-depleted M9 minimal medium of *S. erythraea* cultures grown for 72 h. The absence of iron gives rise to an increased siderophore production of up to 10.2 mg·L^{-1} culture.

respectively, whereas the third one is attached to the α-amino group of one of the δ-*N*-acetyl-δ-*N*-hydroxyornithines (haOrn) resulting in α-*N*-acetly-δ-*N*-acetyl-δ-*N*-hydroxyornithine (ac-haOrn) (Fig. 5A). Three sequential NOE contacts were observed, one revealing a connection between the terminal ac-haOrn$_1$ and the Ser$_2$, whereas the other two were for a sequential connection between a δ-*N*-hydroxyornithine and a δ-*N*-acetyl-δ-*N*-hydroxyornithine and its reverse, respectively. Such double sequential connections can only be established through a diketopiperazine (DKP) unit, which is composed of a hOrn and a haOrn moiety. Furthermore, a long-range ^{1}H–^{13}C correlation was detected between the carbonyl carbon of the serine and the δ-CH_2 of the hOrn, which constitutes the DKP. Therefore, putting all these long-range connections together, we established a structure for the tetrapeptide siderophore, which is designated erythrochelin (Fig. 5). The assigned ^{1}H, ^{13}C and ^{15}N chemical shifts are listed in Tables S3–7. The observed NOE contacts and the long-range ^{1}H-^{13}C correlations verified the structure and are listed in listed in Tables S5 and S6. On the basis of the results obtained by NMR, the determined sequence for the peptide is ac-haOrn$_1$-Ser$_2$-cyclo (hOrn$_3$-haOrn$_4$). The corresponding DQF-COSY, ^{1}H–^{15}N HSQC, heteronuclear multiple bond correlation (HMBC) and ROESY spectra of erythrochelin are shown in Figures S5–S9.

MS analysis of erythrochelin and determination of overall stereochemistry

On the basis of the observed NMR spectra, the presence and connectivity of δ-*N*-acetyl-δ-*N*-hydroxyornithine, δ-*N*-hydroxyornithine and serine in the sequence was determined. Erythrochelin itself shows an exact *m/z* of 604.2938 ([M + H$^+$]; calculated 604.2937) and a molecular formula of $C_{24}H_{41}N_7O_{11}$ and a *m/z* of 657.2056 ([M + H$^+$]; calculated 657.2051) as ferri-erythrochelin. To confirm the structural assignment obtained by NMR, MS^3 fragmentation studies were conducted (Fig. 6). An intense fragment with an *m/z* of 390.1979 ([M + H$^+$]; calculated 390.1983) corresponded to the C-terminal tripeptide comprised of serine and the DKP moiety built up by hOrn and haOrn residues (Fig. 6A). The loss of the N-terminal serine residue gave rise to a dipeptidyl DKP fragment with a *m/z* of 303.1662 ([M + H$^+$]; calculated 303.1663). This fragment was furthermore subjected to MS^3 fragmentation (Fig. 6B). The resulting fragments revealed the presence of hydroxylated and acetylated ornithine residues. In addition, an intense fragment with an *m/z* of 145.0869 ([M + H$^+$]; calculated 145.0971) was observed. This result provided strong evidence for the presence of the DKP moiety because such fragmentation behaviour is characteristic for DKP-containing compounds and has been detected during fragmentation of an albonoursin intermediate (Fig. S10) [28].

Determination of overall stereochemistry of erythrochelin was carried out utilizing Marfey's reagent [29]. Prior to the *N*-α-(2,4-dinitro-5-fluorophenyl)-L-alaninamide (FDAA) derivatization of the amino acids resulting from total hydrolysis of erythrochelin, the hydrolysate was analyzed via LC-MS to determine hydrolysate composition, revealing solely the presence of Ser- and hOrn-residues (Fig. S11). LC-MS analysis of the derivatized hydrolysate compared to synthetic standards indicated the presence of D-Ser, L-hOrn and D-hOrn in a 1 : 2 : 1 ratio (Figs S12 and S13), as expected from bioinformatic analysis of EtcD. To determine the connectivity of the amino acids, as well as their stereoconfiguration, a partial hydrolysis-derivatization approach was carried out. The C-terminal hOrn-hOrn-dipeptide was isolated, hydrolytically cleaved and derivatized (Fig. S14). Solely the presence of L-hOrn residues was observed, confirming the stereochemistry to be in full agreement with the proposed biosynthetic model (Fig. S15).

Discussion

The advance in sequencing technologies, ranging from whole genome shotgun sequencing to high-throughput pyrosequencing, has proliferated over 500 sequenced and annotated microbial genomes, revealing a multitude of gene clusters related to natural product biosynthesis [30,31]. The isolation of the corresponding products of these cryptic clusters is often challenging as a result of either a low rate of production or unknown conditions for secondary metabolite biosynthesis. In addition, bioactivity-guided natural product isolation is often impeded by unpredictable biological activities of the target compounds and a lack of appropriate screening methods. To circumvent the problem of a low rate of biosynthesis and unknown biological activity, we describe a genome mining approach relying on bioinformatic genome analysis and transcriptome data combined with radiolabeled precursor feeding studies for NRPS-derived natural products. In this methodology, transcriptome analysis provides information on the growth conditions leading to gene cluster expression, whereas A-domain specificity prediction defines the radiolabeled precursor.

Initial detection of erythrochelin was performed by cultivation of *S. erythraea* in a complex SCM medium utilizing a radio-LC-MS methodology, and confirmed

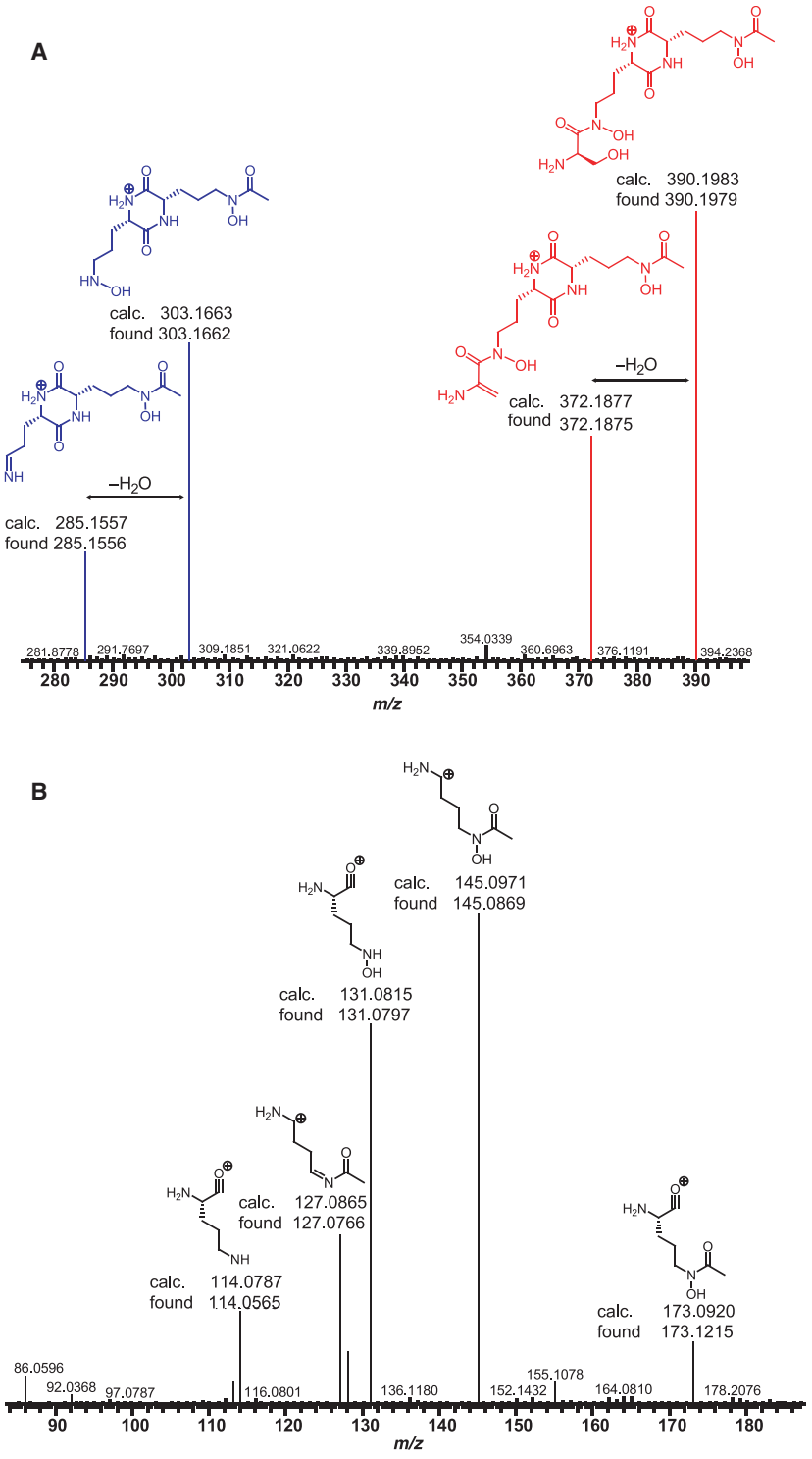

Fig. 6. MS/MS fragmentation studies of erythrochelin. (A) MS^2 fragmentation of the title compound. (B) MS^3 fragmentation pattern of the C-terminal DKP moiety m/z = 303.1662 ([M+H$^+$]). Calculated and observed m/z values for the fragments are given.

the DNA microarray gene expression profiles obtained for *S. erythraea* [20]. Feeding of the nonproteinogenic amino acid ^{14}C-L-Orn prior to expression of the *etc* gene cluster gave rise to radiolabeled erythrochelin, which could be clearly identified on an analytical scale. The sensitivity of radioactivity detection and sophisticated analytical separation proved to be advantageous in this approach. The iron-chelating properties of the

radiolabeled compound were confirmed by CAS assay-guided fractionation of medium-scale fermentation extractions. A comparison of the masses found in the CAS-reactive fraction and the *m/z* of the labeled product revealed erythrochelin to be an ornithine inheriting siderophore. Due to media complexity and coeluting impurities, which prevented rapid MS-based single compound identification, this radio-LC-MS methodology was utilized to identify a minimal medium enabling erythrochelin production. Cultivation of *S. erythraea* under iron-depleted conditions induced the production of erythrochelin compared to iron-rich media cultivations. Interestingly, the amount of ^{14}C-L-Orn incorporation was increased from 2% to 4% (based on the total amount of radioactivity fed) when switching to minimal media. It is likely that the decelerated growth in iron-depleted minimal media combined with an increase in siderophore production leads to the increased incorporation of ^{14}C-L-Orn into the main secondary metabolite erythrochelin. In conclusion, the described approach, solely based on A-domain specificity prediction and the available transcriptome data, can be applied for the initial detection and isolation of NRPs [20]. Furthermore, this approach substitutes the CAS assay-guided fractionation and enabled the scale-down of NRP discovery from a preparative to analytical scale. In addition, this approach can be utilized to substitute the detection and isolation of NRPs based on their biological activity, which is often challenging to predict. The utilization of radiolabeled proteinogenic amino acids, which can be channelled to ribosomal synthesis of peptides, remains to be elucidated.

After having identified the CAS-reactive and ^{14}C-L-Orn incorporating erythyrochelin, a large-scale isolation was conducted affording 10 mg·L^{-1} erythrochelin. The overall structure of erythrochelin was determined by NMR and MS analysis as well as hydrolysate derivatization for determination of amino acid configuration. The peptide sequence is composed of D-ac-haOrn$_1$-D-Ser$_2$-cyclo(L-hOrn$_3$-L-haOrn$_4$). Erythrochelin represents a hydroxamate-type tetrapeptide siderophore containing three ornithine residues, of which two are δ-*N* acetylated and δ-*N* hydroxylated. In addition, the N-terminal α-amino group of haOrn$_1$ is acetylated. A local symmetry in erythrochelin is attained by a DKP structure consisting of two cyclodimerized L-Orn residues. The mode of Fe(III) chelation by erythrochelin remains to be elucidated, although we postulate an iron-binding mode analogous to gallium-binding by coelichelin (Fig. S16). MS analysis of ferri-erythrochelin reveals an abolished skimmer fragmentation compared to erythrochelin, being indicative of an induced rigidification of the siderophore upon iron binding. Erythrochelin shows an absorption spectrum typical of ferri-hydroxamate siderophores with λ_{max} = 440 nm.

Erythrochelin shares a high degree of structural similarity to the angiotensin-converting enzyme inhibitor and siderophore foroxymithine isolated from cultures of *Streptomyces nitrosporeus* (Fig. S17) [32–34]. In contrast to erythrochelin, the δ-amino groups of ac-hOrn$_1$ and hOrn$_4$ are formylated, suggesting that a formyltransferase is involved in biosynthesis, analagous to coelichelin assembly [7]. In an attempt to chemically obtain foroxymithine, a total synthesis was established by Dolence and Miller [35] that resulted in a compound exhibiting the same NMR spectroscopic properties as the isolated natural product. All residues within the peptide chain showed an L-configuration. This stereochemistry differs from erythrochelin, in which two residues show a D-configured stereocenter, thus suggesting a similar NRPS-based assembly of foroxymithine by a synthetase lacking all E-domains. The lack of sequence information for the *S. nitrosporeus* genome impeded the identification of a biosynthetic machinery governing foroxymithine assembly. Future work will focus on the investigation of erythrochelin-mediated angiotensin-converting enzyme inhibition, aiming to assign a bioactivity going beyond iron chelation.

On the basis of the results obtained in the present study, a model for erythrochelin biosynthesis by the tetramodular NRPS EtcD in combination with EtcB and an acetyltransferase was established (Fig. 7). In contrast to the second NRPS gene cluster associated with siderophore production (*nrps*3), which putatively encodes for a catecholate-type compound, the *etc* gene cluster is congruent with the structure of erythrochelin (Fig. S18). The domain organization and the predicted substrate specificities of the A-domains do not reflect in the structure of erythrochelin and exclude its biosynthesis by Nrps3. The extraction of culture supernatants of *S. erythraea*, cell pellets and lysed cells with a variety of organic solvents did not lead to the identification of the second siderophore (data not shown). We therefore assume that either the extraction conditions were inadequate for the isolation of the natural product, or that the gene cluster is silent under the conditions employed. The irrevocable evidence for EtcD-mediated erythrochelin assembly would result from targeted gene deletion of *etcD* followed by LC-MS analysis of culture supernatants. Erythrochelin biosynthesis by EtcD follows a linear enzymatic logic, in which the number of A-domains located within the template directly correlates with the number of amino acids found in the

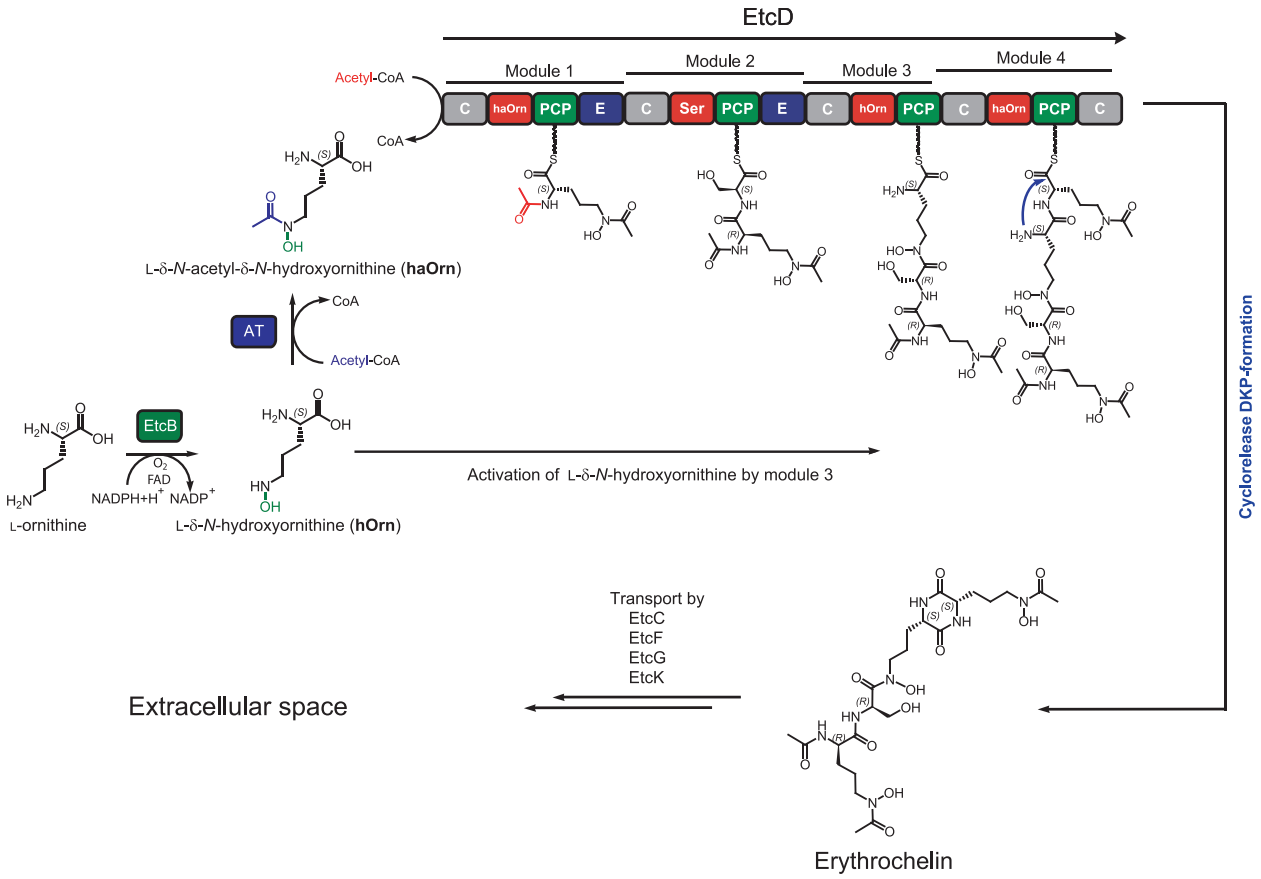

Fig. 7. Proposed biosynthesis of erythrochelin by the tetramodular nonribosomal peptide synthetase EtcD. δ-*N*-hydroxylation of L-ornithine is putatively mediated by the peptide monooxygenase EtcB. δ-*N*-acetylation of L-hydroxyornithine is putatively carried out by an external *N*-acetyltransferase not encoded in the *etc* gene cluster. The N-terminal C-domain of the NRPS catalyzes the α-*N*-acetylation of haOrn$_1$ *in cis*. Cyclorelease of the assembled tetrapeptide mediated by the C-terminal C-domain of EtcD results in the formation of a DKP moiety.

product. Initiation of erythrochelin assembly requires δ-*N*-hydroxylation of L-Orn by the flavin-dependent monooxygenase EtcB, analogous to the CchB-catalyzed oxygenation of L-Orn during coelichelin biosynthesis [22]. L-hOrn itself represents a branching point in erythrochelin synthesis. This building block is either directly recognized by A_3 or further modified by means of δ-*N*-acetylation. In this model, acetyltransferase-catalyzed acetylation of L-hOrn gives rise to L-haOrn, which is recognized by A_1 and A_4, and is activated and covalently tethered to the 4′-Ppant cofactors of the corresponding PCPs as aminoacyl thioester. We propose that acetyltransferases of the IucB- or VbsA-type, as involved in ornithine acetylation in aerobactin and vicibactin biosynthesis, are associated with L-haOrn synthesis [10,36]. These results are consistent with the bioinformatic analysis of EtcD adenylation domain specificity, resulting in the less accurate prediction of L-Arg as substrate for both A_1 and A_4. Differences in the specificity-determining residue pattern are likely to be the result of minimal structural differences between L-Arg and L-haOrn (Fig. S1B). When comparing the active site residues of A_1 and A_4, a high degree of identity (90%) is found, indicating L-haOrn as the common substrate. This model would exclude the online δ-*N*-hydroxylation and δ-*N*-acetylation of the NRPS-bound substrates as seen in the hydroxylation of PCP-bound Glu in kutzneride biosynthesis [37]. Prior to incorporation of haOrn$_1$ into the growing peptide chain, the α-*N*-acetylation is likely to be carried out by the C_1-domain located at the N-terminus of EtcD, recognizing acetyl-CoA as the substrate. A similar mechanism was shown to be adopted in the initiation reaction during surfactin biosynthesis, with β-hydroxymyristoyl-CoA being the substrate for NRPS-catalyzed acyl transfer [38]. Epimerization of the α-stereocenters of L-ac-haOrn$_1$ and L-Ser is

mediated by the E-domains located in modules 1 and 2, being in full agreement with the experimental determination of overall stereochemistry. The C-domain catalyzed condensation of the four unique building blocks follows a linear NRPS assembly line logic. In the first step, the C_2 domain catalyzes the nucleophilic attack of the Ser_1 α-amino group onto the PCP_1-bound ac-$haOrn_1$ resulting in a PCP_2-bound dipeptide. C_3-catalyzed isopeptide bond formation between the δ-amino group of L-$hOrn_3$ and the PCP_2-bound D-ac-$haOrn_1$-D-Ser_2 dipeptide results in the translocation of the tripeptide to PCP_3. A nucleophilic attack of the L-$haOrn_4$ α-amino group onto the PCP_3-bound tripeptide thioester functionality results in the fully assembled tetrapeptide consisting of D-ac-$haOrn_1$-D-Ser_2-L-$hOrn_3$-L-$haOrn_4$. The release of the assembled NRP is generally mediated by C-terminal thioesterase or reductase domains located in the termination module of the NRPS assembly line [21,39]. In contrast, we propose that the cyclorelease of erythrochelin through DKP formation is carried out by the C-terminal C_5-domain, catalyzing the intramolecular nucleophilic attack of the L-$hOrn_3$ α-amino group onto L-$haOrn_4$. Taking into account that the synthetases involved in the biosynthesis of the DKP-inheriting toxins thaxtomin and fumitremorgin also contain a C-terminal condensation domain, this C-domain catalyzed cyclorelease appears to be feasible [40,41]. *Apo*-erythrochelin is then exported into the extracellular space to scavenge iron. The import of ferri-erythrochelin is likely to be mediated by the FeuA homolog EtcC, which is responsible for periplasmic binding [4]. In combination with EtcF, the ABC-transporter transmembrane component and EtcG, the corresponding ATP-binding component, ferri-erythrochelin, is actively reimported into the cell [42].

Materials and methods

Strains and general methods

S. erythraea NRRL 23338 was obtained from the ARS (Agricultural Research Service, Peoria, IL, USA) Culture Collection. Chemicals were obtained from commercial sources and were used without further purification, unless noted otherwise.

Radio-LC-MS-guided genome mining

Radiolabeling studies were performed by cultivating *S. erythraea* in 100 mL of SCM medium (10 $g \cdot L^{-1}$ soluble starch, 20 $g \cdot L^{-1}$ soytone, 10.5 $g \cdot L^{-1}$ Mops, 1.5 $g \cdot L^{-1}$ yeast extract, 0.1 $g \cdot L^{-1}$ $CaCl_2$) or iron-deficient M9 medium (2 $g \cdot L^{-1}$ glucose, 6.78 $g \cdot L^{-1}$ Na_2HPO_4, 3 $g \cdot L^{-1}$ KH_2PO_4, 0.5 $g \cdot L^{-1}$ NaCl, 1.2 $g \cdot L^{-1}$ NH_4Cl, 120 $mg \cdot L^{-1}$ $MgSO_4$, 14.7 $g \cdot L^{-1}$ $CaCl_2$, 0.1 $g \cdot L^{-1}$ glycerol, 50 $\mu g \cdot L^{-1}$ biotin, 200 $\mu g \cdot L^{-1}$ thiamin). After 48 h of growth, 5 μCi of L-ornithine (Hartmann Analytic, Braunschweig, Germany) was added. The supernatants were extracted with XAD16 resin after an additional 2 days of growth. The dried eluate was dissolved in 10% methanol and analyzed on a Nucleodur C_{18}(ec) column 125 × 2 mm (Macherey & Nagel, Düren, Germany) combined with an Agilent 1100 HPLC system (Agilent, Waldbronn, Germany), connected to a FlowStar LB513 radioactivity flow-through detector (Berthold, Bad Wildbad, Germany) equipped with a YG-40-U5M solid microbore cell and a QStar Pulsar i (Applied Biosystems, Foster City, CA, USA), utilizing the solvent gradient: water/0.05% formic acid (solvent A) and methanol/0.05% formic acid (solvent B) at a flow rate of 0.3 $mL \cdot min^{-1}$: linear increase from 0% B to 50% within 20 min followed by a linear increase to 95% B in 5 min, holding B for an additional 5 min. This gradient was also used to analyze comparative extractions of *S. erythraea* cultures and erythrochelin and ferri-erythrochelin.

Isolation of erythrochelin from SCM medium

S. erythraea NRRL 23338, maintained on SCM agar slants, was used to inoculate 30 mL of SCM liquid culture. The cells were grown for 4 days at 30 °C and 250 r.p.m. and subsequently used to inoculate 1 L of SCM medium. The cells were grown for 5 days at 30 °C. The production phase of the strain was monitored via LC-MS and the CAS assay [27]. The culture supernatant was extracted with XAD16 resin (4.0 $g \cdot L^{-1}$). The resin was collected by filtration, washed twice with water and the absorbed compounds were eluted with methanol. The eluate was evaporated to dryness, dissolved in 10% acetonitrile and applied onto a RP-HPLC preparative Nucleodur C_{18}(ec) 250 × 21 mm column combined with an Agilent 1100 HPLC system. Elution was performed by application of the solvent gradient of water/0.05% formic acid (solvent A) and methanol/0.05% formic acid (solvent B) at a flow rate of 16 $mL \cdot min^{-1}$: linear increase from 0% B to 50% within 50 min followed by a linear increase to 95% B in 5 min, holding B for an additional 5 min. The wavelengths chosen for detection were 215 and 280 nm, respectively. Siderophore containing fractions were confirmed by using the CAS liquid assay and subjected to LC-MS analysis.

Large-scale purification of erythrochelin from M9 medium

S. erythraea, maintained on SCM agar slants, was used to inoculate 30 mL of SCM liquid culture. The cells were

grown for 4 days at 30 °C and 250 r.p.m. The cells were exchanged from the SCM medium into the iron-deficient M9 medium by repeated centrifugation and resuspension of the cells in the target medium. Subsequently, the cells were used to inoculate 5 L of iron-deficient M9 medium in PC-flasks until D_{600} of 0.01 was reached. After 4 days of cultivation, the cells were harvested by centrifugation at 6084 ***g*** and 4 °C for 30 min. The supernatant was separated from the cell pellet and incubated with XAD16 resin (4.0 $g \cdot L^{-1}$). The resin was collected by filtration, washed twice with water and the absorbed compounds were eluted with methanol. The eluate was evaporated to dryness, dissolved in 10% acetonitrile and applied onto a RP-HPLC preparative Nucleodur C_{18}(ec) 250 × 21 mm column combined with an Agilent 1100 HPLC system. Elution was performed by application of the solvent gradient of water/0.05% formic acid (solvent A) and methanol/0.05% formic acid (solvent B) at a flow rate of 16 $mL \cdot min^{-1}$: linear increase from 0% B to 50% within 50 min followed by a linear increase to 95% B in 5 min, holding B for an additional 5 min. The wavelengths chosen for detection were 215 and 280 nm, respectively. Siderophore containing fractions were confirmed by using the CAS assay. Positive fractions were lyophilized and subjected to further analysis. The retention time of erythrochelin was 30.7 min.

MS analysis

The MS characterization of erythrochelin was performed with an LTQ-FT instrument (Thermo Fisher Scientific, Langenselbold, Germany) connected to a microbore Agilent 1100 HPLC system. *Apo*- and *holo*-erythrochelin were analyzed on a Nucleodur C_{18}(ec) 125 × 2 mm column utilizing the solvent gradient: 0–30 min, 0–100% acetonitrile into water, both supplemented with 0.1% trifluoroacetic acid. The column temperature was 45 °C and the flow rate was 0.3 $mL \cdot min^{-1}$. Collision induced dissociation fragmentation studies within the linear ion trap were carried out using online LC-MS.

NMR structure elucidation

Approximately 16 mg of the title compound was dissolved in 0.7 mL of dimethylsulfoxide-d_6. Measurements were carried out on a AV600 (Bruker, Madison, WI, USA) spectrometer with an inverse broadband probe installed with *z*-gradient. The 1D spectra ^{1}H and ^{13}C; the homonuclear 2D spectra DQF-COSY, TOCSY, NOESY and ROESY; the ^{1}H–^{3}C HSQC and HMBC; and the ^{1}H–^{15}N HSQC spectra were recorded at room temperature using standard PULSE software [43]. The phase-sensitive HMBC spectrum focused on the carbonyl region with high resolution in the ^{13}C dimension was recorded by using PULSE software with a semi-selective ^{13}C pulse built into an HMBC experiment with sensitivity enhancement [44,45]. The TOCSY spectrum was recorded with mixing time of 200 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. The relaxation delay was 2.5 s. Chemical shifts of ^{1}H and ^{13}C were referenced to the solvent signals, whereas that of ^{15}N was referenced to the urea signal, externally. The spectra were processed using TOPSPIN, version 2.1 (Bruker).

Amino acid analysis by FDAA (Marfey's reagent) derivatization

Five hundred micrograms of erythrochelin were completely hydrolyzed by the addition of 400 μL of 6 M HCl and incubation at 110 °C for 24 h. The solution was lyophilized and the remaining residue dissolved in 10 μL of 1 M $NaHCO_3$. One hundred and seventy microliters of 1% FDAA (Sigma-Aldrich, Munich, Germany) in acetone were added and the solution was heated at 37 °C for 1 h. The derivatization reaction was terminated by the addition of 20 μL of 1 M HCl. After lyophilization, the derivatized amino acids were resolubilized by the addition of 1 : 1 water : acetonitrile solution and 0.1% trifluoroacetic acid to obtain a final volume of 400 μL. Products of derivatization were analyzed by RP-LC-MS on a Synergi Fusion-RP 80 250 × 2.0 mm column (Phenomenex, Aschaffenburg, Germany) utilizing the 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 wavelength chosen for detection was 340 nm and the flow rate was 0.3 $mL \cdot min^{-1}$. [29]. Ten microliters of sample was added to 90 μL of water prior to the injection of 10 μL.

To determine the stereochemistry of the present amino acids, amino acid standards (D/L-Ser and L-hOrn) were prepared to compare retention times and MS spectra, as well as to perform coelution experiments. The FDAA-derivatized amino acids were synthesized by incubation of 25 μL of 50 mM amino acid in water, 50 μL of 1% FDAA in acetone and 10 μL of 1 M $NaHCO_3$ at 37 °C for 1 h. The solution was lyophilized, and the dried products resolubilized in 1 : 1 water : acetonitrile solution and 0.1% trifluoroacetic acid to obtain 200 μL. L-hOrn was synthesized chemically according to an established protocol [46]. Coelution experiments were conducted by mixing 10 μL of derivatized erythrochelin hydrolysate with 1 μL of derivatized D-Ser amino acid standard and 3 μL of derivatized L-hOrn standard. RP-LC-MS analysis was performed as described above.

Determination of amino acid connectivity via partial hydrolysis of erythrochelin

Three milligrams of erythrochelin were partially hydrolyzed in 200 μL of 6 M HCl at 110 °C for 20 min. The resulting solution was lyophilized and resolubilized in 1 : 1 water : acetonitrile solution and 0.1% trifluoroacetic acid to a final volume of 200 μL and analyzed via an LTQ-FT instrument to a microbore Agilent 1100 HPLC system. Products were analyzed on a Nucleodur C_{18}(ec) 125×2 mm column, utilizing the solvent gradient: 0–30 min, 0–100% acetonitrile into water, both supplemented with 0.1% trifluoroacetic acid followed by a linear increase to 95% acetonitrile in 5 min and holding 95% acetonitrile for an additional 5 min. The column temperature was 45 °C and the flow rate was 0.3 mL·min^{-1}. Collision induced dissociation fragmentation studies within the linear ion trap were carried out using online LC-MS. The target fragment was isolated from the mixture with an Agilent 1100 HPLC system connected to an AnalytFC fraction collector (Agilent) on a Hypercarb 100×2.1 mm column (Thermo, Waltham, MA, USA) utilizing the solvent gradient: 0–30 min, 0–70% acetonitrile into water, both supplemented with 20 mM nonafluoro-1-pentanoic acid, each followed by a linear increase to 95% acetonitrile in 2 min and holding 95% acetonitrile for an additional 5 min. The wavelength chosen for detection was 215 nm, with a column temperature of 20 °C and a flow rate of 0.3 mL·min^{-1}. Product containing fractions were identified by RP-LC-MS. Positive fractions were lyophilized, hydrolyzed and derivatized with FDAA as described above. Analysis of the derivatized amino acids was performed by RP-LC-MS.

Acknowledgements

We would like to thank Antje Schäfer and Anke Botthof for their excellent support during this project. We gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft (M.A.M.).

References

1 Miethke M & Marahiel MA (2007) Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev* **71**, 413–451.
2 Winkelmann G & Drechsel H (1997) *Microbial Siderophores*. Wiley-VCH, Weinheim.
3 Miethke M, Klotz O, Linne U, May JJ, Beckering CL & Marahiel MA (2006) Ferri-bacillibactin uptake and hydrolysis in *Bacillus subtilis*. *Mol Microbiol* **61**, 1413–1427.
4 Peuckert F, Miethke M, Albrecht AG, Essen LO & Marahiel MA (2009) Structural basis and stereochemistry of triscatecholate siderophore binding by FeuA. *Angew Chem Int Ed Engl* **48**, 7924–7927.
5 Andrews SC, Robinson AK & Rodriguez-Quinones F (2003) Bacterial iron homeostasis. *FEMS Microbiol Rev* **27**, 215–237.
6 Gehring AM, Bradley KA & Walsh CT (1997) Enterobactin biosynthesis in *Escherichia coli*: isochorismate lyase (EntB) is a bifunctional enzyme that is phosphopantetheinylated by EntD and then acylated by EntE using ATP and 2,3-dihydroxybenzoate. *Biochemistry* **36**, 8495–8503.
7 Lautru S, Deeth RJ, Bailey LM & Challis GL (2005) Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nat Chem Biol* **1**, 265–269.
8 Dimise EJ, Widboom PF & Bruner SD (2008) Structure elucidation and biosynthesis of fuscachelins, peptide siderophores from the moderate thermophile *Thermobifida fusca*. *Proc Natl Acad Sci USA* **105**, 15311–15316.
9 Kadi N, Oves-Costales D, Barona-Gomez F & Challis GL (2007) A new family of ATP-dependent oligomerization-macrocyclization biocatalysts. *Nat Chem Biol* **3**, 652–656.
10 de Lorenzo V, Bindereif A, Paw BH & Neilands JB (1986) Aerobactin biosynthesis and transport genes of plasmid ColV-K30 in *Escherichia coli* K-12. *J Bacteriol* **165**, 570–578.
11 Barona-Gomez F, Wong U, Giannakopulos AE, Derrick PJ & Challis GL (2004) Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces coelicolor* M145. *J Am Chem Soc* **126**, 16282–16283.
12 Ledyard KM & Butler A. (1997) Structure of putrebactin, a new dihydroamate siderophore produced by *Shewanella putrefaciens*. *J Am Chem Soc* **2**, 93–97.
13 Lautru S, Oves-Costales D, Pernodet JL & Challis GL (2007) MbtH-like protein-mediated cross-talk between non-ribosomal peptide antibiotic and siderophore biosynthetic pathways in *Streptomyces coelicolor* M145. *Microbiology* **153**, 1405–1412.
14 Challis GL (2008) Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology* **154**, 1555–1569.
15 Zerikly M & Challis GL (2009) Strategies for the discovery of new natural products by genome mining. *Chembiochem* **10**, 625–633.
16 Schaible UE & Kaufmann SH (2004) Iron and microbial infection. *Nat Rev Microbiol* **2**, 946–953.
17 Gross H, Stockwell VO, Henkels MD, Nowak-Thompson B, Loper JE & Gerwick WH (2007) The genomisotopic approach: a systematic method to isolate products of orphan biosynthetic gene clusters. *Chem Biol* **14**, 53–63.
18 Oliynyk M, Samborskyy M, Lester JB, Mironenko T, Scott N, Dickens S, Haydock SF & Leadlay PF (2007) Complete genome sequence of the erythromycin-

producing bacterium *Saccharopolyspora erythraea* NRRL23338. *Nat Biotechnol* **25**, 447–453.
19 Oliveira PH, Batagov A, Ward J, Baganz F & Krabben P (2006) Identification of erythrobactin, a hydroxamate-type siderophore produced by *Saccharopolyspora erythraea*. *Lett Appl Microbiol* **42**, 375–380.
20 Peano C, Bicciato S, Corti G, Ferrari F, Rizzi E, Bonnal RJ, Bordoni R, Albertini A, Bernardi LR, Donadio S *et al*. (2007) Complete gene expression profiling of *Saccharopolyspora erythraea* using GeneChip DNA microarrays. *Microb Cell Fact* **6**, 37.
21 Kohli RM & Walsh CT (2003) Enzymology of acyl chain macrocyclization in natural product biosynthesis. *Chem Commun (Camb)* **3**, 297–307.
22 Pohlmann V & Marahiel MA (2008) Delta-amino group hydroxylation of L-ornithine during coelichelin biosynthesis. *Org Biomol Chem* **6**, 1843–1848.
23 Stachelhaus T, Mootz HD & Marahiel MA (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol* **6**, 493–505.
24 Challis GL, Ravel J & Townsend CA (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem Biol* **7**, 211–224.
25 Rausch C, Weber T, Kohlbacher O, Wohlleben W & Huson DH (2005) Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res* **33**, 5799–5808.
26 Tillett D, Dittmann E, Erhard M, von Dohren H, Borner T & Neilan BA (2000) Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system. *Chem Biol* **7**, 753–764.
27 Schwyn B & Neilands JB (1987) Universal chemical assay for the detection and determination of siderophores. *Anal Biochem* **160**, 47–56.
28 Gondry M, Sauguet L, Belin P, Thai R, Amouroux R, Tellier C, Tuphile K, Jacquet M, Braud S, Courcon M *et al*. (2009) Cyclodipeptide synthases are a family of tRNA-dependent peptide bond-forming enzymes. *Nat Chem Biol* **5**, 414–420.
29 Bhushan R & Bruckner H (2004) Marfey's reagent for chiral amino acid analysis: a review. *Amino Acids* **27**, 231–247.
30 Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z *et al*. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–380.
31 Donadio S, Monciardini P & Sosio M (2007) Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics. *Nat Prod Rep* **24**, 1073–1109.
32 Umezawa H, Aoyagi T, Ogawa K, Obata T, Iinuma H, Naganawa H, Hamada M & Takeuchi T (1985) Foroxymithine, a new inhibitor of angiotensin-converting enzyme, produced by actinomycetes. *J Antibiot (Tokyo)* **38**, 1813–1815.
33 Aoyagi T, Wada T, Iinuma H, Ogawa K, Kojima F, Nagai M, Kuroda H, Obayashi A & Umezawa H (1985) Influence of angiotensin-converting enzyme inhibitor, foroxymithine, on dynamic equilibrium around the renin-angiotensin system *in vivo*. *J Appl Biochem* **7**, 388–395.
34 Imoto M, Umezawa K, Komuro K, Sawa T, Takeuchi T & Umezawa H (1987) Antitumor activity of erbstatin, a tyrosine protein kinase inhibitor. *Jpn J Cancer Res* **78**, 329–332.
35 Dolence EK & Miller MJ (1991) Synthesis of foroxymithine, a microbial fermentation product and angiotensin 1 converting enzyme inhibitor. *J Org Chem* **56**, 492–499.
36 Heemstra JR, Walsh CT & Sattely ES (2009) Enzymatic tailoring of ornithine in the biosynthesis of the rhizobium cyclic trihydroxamate siderophore vicibactin. *J Am Chem Soc* **131**, 15317–15329.
37 Strieker M, Nolan EM, Walsh CT & Marahiel MA (2009) Stereospecific synthesis of threo- and erythro-beta-hydroxyglutamic acid during kutzneride biosynthesis. *J Am Chem Soc* **131**, 13523–13530.
38 Steller S, Sokoll A, Wilde C, Bernhard F, Franke P & Vater J (2004) Initiation of surfactin biosynthesis and the role of the SrfD-thioesterase protein. *Biochemistry* **43**, 11331–11343.
39 Kopp F, Mahlert C, Grunewald J & Marahiel MA (2006) Peptide macrocyclization: the reductase of the nostocyclopeptide synthetase triggers the self-assembly of a macrocyclic imine. *J Am Chem Soc* **128**, 16478–16479.
40 Healy FG, Wach M, Krasnoff SB, Gibson DM & Loria R (2000) The txtAB genes of the plant pathogen *Streptomyces acidiscabies* encode a peptide synthetase required for phytotoxin thaxtomin A production and pathogenicity. *Mol Microbiol* **38**, 794–804.
41 Maiya S, Grundmann A, Li SM & Turner G (2006) The fumitremorgin gene cluster of *Aspergillus fumigatus*: identification of a gene encoding brevianamide F synthetase. *Chembiochem* **7**, 1062–1069.
42 Hollenstein K, Dawson RJ & Locher KP (2007) Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol* **17**, 412–418.
43 Berger S & Braun S (2004) *200 and More NMR Experiments*. A Practical Course. Wiley-VCH, Weinheim.
44 Kessler H, Schmieder P, Köck M & Kurz M (1990) Improved resolution in proton-detected heteronuclear long-range correlation. *J Magn Reson* **88**, 615–618.

45 Cicero DO, Barbato G & Bazzo R (2001) Sensitivity enhancement of a two-dimensional experiment for the measurement of heteronuclear long-range coupling constants, by a new scheme of coherence selection by gradients. *J Magn Reson* **148**, 209–213.
46 Lin Y & Miller MJ (1999) Practical synthesis of hydroxamate-derived siderophore components by an indirect oxidation method and syntheses of a DIG-siderophore conjugate and a biotin-siderophore conjugate. *J Org Chem* **64**, 7451–7458.

Supporting information

The following supplementary material is available:
Fig. S1. HPLC-profile of SCM medium extraction.
Fig. S2. LC-MS traces of comparative extractions.
Fig. S3. UV/visible absorption spectra of erythrochelin.
Fig. S4. 1D ^{1}H-NMR spectrum of erythrochelin.
Fig. S5. DQF-COSY spectrum of erythrochelin.
Fig. S6. 1H-15N HSQC spectrum of erythrochelin.
Fig. S7. HMBC spectrum of erythrochelin; amide protons.
Fig. S8. HMBC spectrum of erythrochelin; side chain protons.
Fig. S9. ROESY spectrum of erythrochelin.
Fig. S10. Fragmentation pattern of C-terminal fragment.
Fig. S11. LC-MS analysis of erythrochelin hydrolysate.
Fig. S12. LC-MS trace of FDAA-derivatized standards.
Fig. S13. LC-MS trace of FDAA-derivatized hydrolysate.
Fig. S14. HRMS analysis of C-terminal dipeptidyl-fragment.
Fig. S15. LC-MS trace of FDAA-derivatized C-terminal fragment.
Fig. S16. Proposed Fe(III)-binding modes.
Fig. S17. Structural comparison of erythrochelin and foroxymithine.
Fig. S18. Schematic overview of Nrps3.
Table S1. Bioinformatic overview of *etc* gene cluster.
Table S2. (A) Comparison of A_1 and A_4. (B) Structures of L-Arg and L-haOrn.
Table S3. ^{1}H chemical shifts.
Table S4. ^{13}C chemical shifts.
Table S5. ^{15}N chemical shifts.
Table S6. Observed NOE contacts.
Table S7. Long-range ^{1}H-^{13}C correlations.

This supplementary material can be found in the online version of this article.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.