Structure determination of a siderophore peucechelin from Streptomyces peucetius

Biometals, full paper

Title: Structure determination of a siderophore peucechelin from *Streptomyces peucetius*

Authors: Shinya Kodani[*,1], Hisayuki Komaki[2], Masahiro Suzuki[1], Fumiya Kobayakawa[1], Hikaru Hemmi[3]

Affiliations and addresses:

[1]Graduate School of Agriculture, Shizuoka University, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529, Japan

[2]Biological Resource Center, National Institute of Technology and Evaluation (NBRC), 2-5-8 Kazusakamatari, Kisarazu, Chiba 292–0818, Japan

[3]National Food Research Institute, National Agriculture and Food Research Organization (NARO), 2–1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan

***To whom correspondence should be addressed:** Shinya Kodani, Graduate School of Agriculture, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529, Japan, Tel/Fax; +81(54)238-5008, E-mail; kodani.shinya@shizuoka.ac.jp

Abstract

Previously, Park et al. isolated a new siderophore from *Streptomyces peucetius* ATCC 27952 based on information of the genome sequence. The structure of the siderophore was deduced based on MS/MS analysis. To clarify the complete structure of the siderophore, we cultured *S. peucetius* with iron deficient medium. Through several chromatographic procedures, the siderophore named peucechelin was isolated with the yield enough to perform NMR experiments. The planar structure of peucechelin was elucidated by the combination of ESI-MS experiment and NMR spectroscopic analyses of the gallium (III) complex. Unlike the previously deduced cyclic structure, the structure was determined to be a linear peptide which was similar to a known siderophore foroxymithine. The stereochemistries of amino acids constituting peucechelin were determined by applying modified Marfey method to the hydrolysate. Since the biosynthetic gene of peucechelin was formerly determined by Park et al., the similar genes were searched using genome data of other streptomycetes. As a result, the similar genes were found in the genome data of *S. venezuelae* and *S. purpureus*. Isolation of siderophore was performed from the iron deficient culture of *S. venezuelae*. The siderophore of *S. venezuelae* was identified to be known compound foroxymithine by analysis ESI-MS and NMR spectra in the similar manner with peucechelin. Production of foroxymithine was also observed in the iron deficient culture of *S. purpureus*. Based on the genome data, comparison of the biosynthetic genes of structurally related siderophores peucechelin and foroxymithine was accomplished in discussion.

Keywords: siderophore, *Streptomyces peucetius*, peptide, biosynthesis, foroxymithine

Introduction

Iron is essential for critical processes, such as respiration and DNA synthesis, in almost all life forms. Despite being one of the most abundant elements in the Earth's crust, the bioavailability of iron in many environments, such as the soil, is limited due to the very low aqueous solubility of the ferric ion. In iron deficient condition, some bacteria secrete siderophores which are defined as small molecular weight compounds with high-affinity of iron chelating (Ahmed and Holmstrom 2014). Previously, A new siderophore biosynthesized by non-ribosomal peptide synthetase (NRPS) was found and isolated from *Streptomyces peucetius* ATCC 27952 based on genome mining (Doroghazi et al. 2014; Park et al. 2013). The structure of the siderophore was deduced to be a cyclic peptide possessing three hydroxamate moieties by the analysis of MS/MS experiment. So far, many peptidic siderophore such as foroxymithine (Umezawa et al. 1985), coelichelin (Challis and Ravel 2000; Lautru et al. 2005), tsukubachelins A and B (Kodani et al. 2011; Kodani et al. 2013b), scabchelin (Kodani et al. 2013a), turgichelin, and albachelin (Kodani et al. 2015) have been isolated and structure-determined from streptomycetes. The common structural characteristic of these siderophores was a linear peptide structure which possessed three hydroxamate moieties. Considering similarity of the biosynthetic genes among the streptomycetes, there is the possibility that the siderophore produced by *S. peucetius* may also have the similar linear peptide structure. In this paper, isolation of siderophore named peucechelin was accomplished from iron deficient culture to clarify the complete structure of the siderophore of *S. peucetius*. As a result, enough amount of the siderophore was obtained for NMR and MS experiments to determine the chemical structure. Since the biosynthetic gene of peucechelin was formerly determined, the similar genes were searched using genome data of other streptomycetes. As a result, we found similar gene in the genome data of *S. venezuelae* (Pullan et al. 2011) and *S. purpureus*. Isolation and identification of siderophore foroxymithine

was performed from the iron deficeint culture of *S. venezuelae*. In addition, production of foroxymithine was also observed in the iron deficient culture of *S. purpureus.* Comparison of the biosynthetic genes of peucechelin and foroximithine was accomplished in discussion. Here we describe isolation of siderophore peucechelin from *S. peucetius* and discuss its biosynthetic genes.

Results

Following the method in previous paper (Kodani et al. 2015), cultivation of *S. peucetius* NBRC 14660 (=ATCC 27952) was performed using 2L of iron deficient media. To avoid the contamination of ferric ion, flasks and funnels made of polystyrene were used to culture and harvest. After cultivation for 7 days, the bacterial cells were removed from the culture media by filtration. 1M $FeCl_3$ solution (0.5 mL) was added to the spent culture media to generate the complex of siderophore with ferric ion. The culture media was concentrated to the aqueous solution up to 50 mL by rotary evaporator. The concentrated material was subjected to open column chromatography using hydrophobic resin CHP-20P with elution of 10% MeOH, 60% MeOH, and MeOH. The 60% MeOH fraction was subjected to HPLC purification to yield 7.2 mg of the ferri-peucechelin (Fe-peucechelin). The measurement of ESI-TOF mass spectrum of Fe-peucechelin gave an ion peak at *m/z* 698.3 (Fig. S2). Since the presence of ferric ion is not compatible with the NMR spectroscopy analysis, the conversion of the ferric siderophore into gallium ion complex via desferri-peucehcelin was performed following previous report (Kodani et al. 2015). As a result, desferri-peucechelin (**1** in Fig. 1) and gallium-complex of peucechelin (Ga-peucechelin) were obtained by HPLC purification. The presence of Ga-peucechelin was

86 determined by ESI-TOF MS analysis, which gave ion peaks at *m/z* 711.3 and 713.3 with
87 intensity ratio of 6:4, which indicated gallium complex (Fig. S3).
88 Desferri-peucechelin (**1**) was isolated as a white powder after lyophilization, and the molecular
89 formula of **1** was determined to be $C_{25}H_{44}N_{10}O_{10}$ by HR ESI-TOF MS analysis (*m/z* calculated
90 for $C_{25}H_{45}N_{10}O_{10}^{+}$: 645.3320 found: 645.3322 for $[M+H]^{+}$). The analyses of NMR spectra
91 including ^{1}H, ^{13}C, DEPT-135, DQF-COSY, TOCSY, NOESY, HMBC, HSQC, and ^{1}H-^{15}N
92 HSQC were performed on Ga-peucechelin dissolved in 0.5 mL of DMSO-d_6. The ^{1}H NMR
93 spectrum gave the duplicate signals with intensity rate of 3:2, which indicated the possibility of
94 presence of isomer. For instance, amide protons gave duplicate signals (7.83/7.79 ppm for NH-
95 *N*-Ac hOrn1, 7.64/7.73 ppm for NH-Arg2, and 7.42/7.44 ppm for δ-NH-Arg2), as shown Fig. 2.
96 The related siderophore foroxymithine was recently isolated from *Streptomyces narbonensis* and
97 the NMR spectrum of Ga-foroxymithine was also reported to give duplicate signals (Ahmad
98 2011).
99 By the interpretation of DQF-COSY and TOCSY, proton spin system of each amino acid was
100 constructed as shown by bold line in Fig. 3a. The assignments of C-H spin system were
101 performed by the interpretation of HSQC data (Table 1), and revealed that peucechelin consisted
102 of 4 mole of amino acids including one mol each of Arg, *N*-α-acethyl-*N*-δ-hydroxy-*N*-δ-
103 formylornithine (*N*-Ac hfOrn), *N*-δ-hydroxyornithine (hOrn), and *N*-δ-hydroxy-*N*-δ-
104 formylornithine (hfOrn). The existence of an acetyl residue was confirmed by the HMBC
105 correlation from methyl protons (1.80 ppm) to carbonyl carbon (168.7 ppm). The HMBC
106 correlation from NH-*N*-Ac hfOrn1 to the same carbonyl carbon (168.7 ppm) indicated that acetyl
107 residue attached to amide proton of *N*-Ac hfOrn1. The presence of formyl residue was indicated
108 by HSQC correlation from singlet proton (8.06 ppm) to characteristic carbon chemical shift

109 value of 152.5 ppm. The formyl residue was indicated to be attached to δ position in *N*-Ac
110 hfOrn1, by the HMBC correlation from δ-protons of *N*-Ac hfOrn1 to formyl carbonyl carbon.
111 As shown in Fig. 3a by one end arrow, the HMBC correlations (α proton-*N*-Ac hfOrn1/CO-*N*-
112 Ac hfOrn1, NH-Arg2/CO-*N*-Ac hfOrn1) were used to establish the connections between *N*-Ac
113 hfOrn1/Arg2. The connection between Arg2 and hOrn3 was established HMBC correlations (α
114 proton-*N*-Ac hfOrn1/CO-*N*-Ac hfOrn1, δ proton-hOrn3/CO-*N*-Ac hfOrn1). Diketopiperazine
115 structure was confirmed by HMBC correlations (NH-hOrn3/CO-hfOrn4, α proton-hfOrn4/CO-
116 hfOrn4, NH-hfOrn4/CO-hOrn3, and α proton-hOrn3/CO-hOrn3). The hydroxyl residues in *N*-
117 Ac hfOrn1, hOrn3 and chOrn4 were not detected by ^{1}H NMR, however the fragmentation of
118 ESI-MS experiment supported the positions of hydroxyl residues as indicated in Fig. 3b. Above
119 all, the structure of peucechelin was determined as **1** in Fig. 1.
120 To elucidate the absolute stereochemistries of amino acids, the hydrolysate of ferri-peucechelin
121 was derivatized with *N*α-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide (L-FDLA), and the
122 derivative was subjected to HPLC analysis to compare with the standard amino acid derivatives
123 with L-FDLA or D-FDLA (Harada et al. 1996). To obtain Orn, hydrogen iodide (HI) was used
124 for hydrolysis with reduction (Stephan et al. 1993). Since only L-Orn was detected by HPLC
125 analysis, the streochemistryies of 3 mole of Orn was determined to be all L-form. The
126 stereochemistry of Arg was determined to be D-form in the same manner.
127 As the biosynthetic gene of peucechelin which possessed similar structure of foroxymithine
128 was identified, we searched for similar NRPS gene in the genome database. As a result, we
129 found two similar NRPS genes in the genome data of *Streptomyces venezuelae* ATCC 10712 and
130 *S. purpureus* KA281 as stated in Discussion. Isolation and identification of foroxymithine was
131 performed from iron deficient culture of *S. venezuelae*, in the same manner with peucechelin.

Briefly the molecular formula of foroxymithine derived from *S. venezuelae* was confirmed to be $C_{22}H_{37}N_7O_{11}$ by HR-ESI TOF-MS. In the same manner with peucechelin, Ga complex of foroxymithine was obtained by conversion from ferric foroxymithine. The ^1H NMR spectrum of Ga-foroxymithine was measured, and NMR chemical shift values of Gallium complex of foroxymithine from *S. venezuelae* were identical to the previous report (Figure S15). The assignment of chemical shifts was performed by analyzing 2D NMR spectra including DQF-COSY, NOESY, HSQC, and HMBC (Table S1). Production of foroxymithine in the iron deficient culture medium of *S. purpureus* was also confirmed by comparative HPLC (Figure S1) and ESI-analysis. Above all, we confirmed production of foroxymithine by *S. venezuelae* and *S. purpureus*.

Discussion

The biosynthetic gene cluster of peucechelin was formerly determined as shown in Fig. 4a (Park et al. 2013). A large NRPS encoded by sp970 comprised four modules composed of fourteen domains, each module containing a set of condensation (C), adenylation (A), and peptidyl carrier protein (PCP) domains essential for NRPS modules (Fig. 4b). We predicted amino-acid substrates recognized by A domains in the NRPS by comparing the substrate specificity-determining residues with those of NRPSs for amychelin, scabichelin, erythrochelin and coelichelin syntheses (Kodani et al. 2015). Consequently, A domains of modules 1 and 4 were predicted to recognize hfOrn based on the similarities (Table 2), that of module 3 was indicated to recognize hOrn as reported (Park et al. 2013), and that of module 2 is unpredictable because its substrate specificity-determining residues were not similar to the others. However,

154 since peucechelin is composed of hfOrn, Arg, hOrn and hfOrn molecules, the A domain of
155 module 2 must recognize Arg.

156 We propose the biosynthetic pathway of peucechelin (Fig 4b), since it has not been proposed
157 (Park et al. 2013). For biosynthesis of L-hOrn and L-hfOrn, a lysine/ornithine *N*-
158 monooxygenases encoded by sp978 initially catalyzes *N*5-hydroxylation of L-Orn. The L-hOrn
159 molecules are not only directly incorporated into A domain of module 3 in the NRPS, but also
160 undergoes *N*5-formylation catalyzed by a formyl transferase encoded by sp980 yielding L-hfOrn.
161 Each amino-acid building block (L-hfOrn for modules 1 and 4, L-hOrn for module 3, Arg for
162 module 2) is converted to aminoacyl adenylate by each the A domain in the NRPS, and
163 transferred on to the adjacent PCP domain within each module to form the corresponding
164 aminoacyl thioesters. When L-hfOrn is loaded onto module 1, its C domain catalyzes acetylation
165 yielding *N*-Ac hfOrn by the mechanism as reported (Robbel et al. 2010). The L-Asp residue of
166 module 2 undergoes α-carbon epimerization catalyzed by the epimerization (E) domain within
167 module 2, and then C domains of modules from 2 to 4 catalyze three successive *N*-acylation
168 reactions to yield an L, D, L, L-tetrapeptidyl thioester attached to the PCP domain of module 4.
169 Interestingly, the NRPS lacks a C-terminal thioesterase (TE) domain but instead has a C domain
170 at the position as same as those of albacelin (Kodani et al. 2015) and erythrochelin (Lazos et al.
171 2010). In the case of biosynthesis of structurally related siderophore amychelin, a putative
172 standalone α,β-hydrolase (AmcB) was proposed to get involved in final release of peptide
173 (Seyedsayamdost et al. 2011). Since a possible hydrolase and a metal-dependent hydrolase are
174 encoded by sp994 and sp957-1, respectively, in the peucechelin biosynthetic gene cluster, these
175 may catalyze final peptide chain release.

Two NRPS genes similar to sp970 were also found in the genome sequences of *S. venezuelae* ATCC 10712 (SVEN_7059, accession number: CCA60345) and *S. purpureus* KA281 (STRPU_RS0128225, accession number: WP_019890436). Domain organizations of these two NRPSs were identical (C/A/PCP-C/A/PCP-C/A/PCP-C/A/PCP/C) each other, and were the same as that of sp970 (C/A/PCP-C/A/PCP/E-C/A/PCP-C/A/PCP/C) except for the lack of an E domain. Comparison among substrate specificity-determining residues suggested that these two NRPSs incorporate fhOrn, Ser, hOrn and fhOrn as the building blocks (Table 3). We had predicted these products are foroxymithines, which possess similar structure of peucechelin, because foroxymithine is a tetrapeptide siderophore including fhOrn, Ser, hOrn and fhOrn. As expected, we successfully demonstrated foroxymithine production by *S. venezuelae* NBRC 13096 (=ATCC 10712) and *S. purpureus* NBRC 13925 (=KA281) in the present study.

We next investigated the gene clusters encoding these foroxymithine NRPSs in *S. venezuelae* ATCC 10712 and *S. purpureus* KA281. Genes in the clusters were well conserved between the two strains as shown in Table S2 and interestingly orders of the encoded genes were the same as that of peucechelin-synthetic gene cluster (Fig 4a). Each foroxymithine gene cluster encoded a lysine/ornithine *N*-monooxygenase (SVEN_7057 or STRPU_RS0128215) and a formyltransferase (SVEN_7056 or STRPU_RS0128210), essential for hOrn and hfOrn syntheses, respectively, upstream of the NRPS gene, and encoded an MbtH (SVEN_7060 or STRPU_RS0128230) downstream of the NRPS gene. Proteins for ferric-siderophore export/uptake were also present near the NRPS. We proposed the biosynthetic pathway of foroxymithine as shown in Fig. 5b taking *S. venezuelae* pathway as an example, since foroxymithine-synthetic pathway has not been reported. L-hOrn, L-hfOrn and tetrapetide are synthesized in the same manner as those in peucechelin synthesis except module 2 of

199 peucechelin NRPS incorporates Arg whereas that of foroxymithine NRPS loads Ser. Above all,
200 we concluded that the NRPS gene clusters shown in Fig. 5 and Table S2 were responsible for the
201 synthesis of foroxymithine.

202 **Materials and Methods**

203 **Bacterial strain and culture condition**

204 *Streptomyces peucetius* NBRC 14660^T (=ATCC 27952, type strain), *Streptomyces venezuelae*
205 NBRC 13096^T (=ATCC 10712, type strain), and *S. purpureus* NBRC 13925^T (=KA281, type
206 strain) was obtained from the NBRC culture collection. The iron deficient medium was prepared
207 by adding 2g of K_2SO_4, 3g of K_2HPO_4, 1g of NaCl, 5g of NH_4Cl in 1 L of deionized water. To
208 remove ferric ions, the solution (1 L) was stirred with 50g of weakly acidic cation exchange resin
209 Chelex-100 sodium form (Bio-rad, CA, USA) for 2 h. The solution was filtrated with paper
210 filter (Whatman No.1, GE Healthcare Life Sciences, Buckinghamshire, England) and added with
211 80 mg of $MgSO_4$, followed by autoclaving. The separately sterilized solutions (10 mL each) of
212 $CaCl_2$•H_2O (10 mg/mL), glucose (250 mg/mL), and 0.5% yeast extract (Difco) were added to 1L
213 of sterile medium in clean bench. *S. peucetius* was cultured by total 2 L of iron-deficient media
214 with incubated rotary shaker (100 rpm, 30 °C) for 7 days.

215 **Isolation of peucechelin**

216 The culture medium of *S. peucetius* was harvested by filtering with paper filter to remove
217 bacterial cells (Whatman No.1). The medium was added with 0.5 mL of 1M $FeCl_3$ and
218 evaporated using a rotary evaporator to concentrate to 50 mL of the final volume. The
219 concentrated solution was subjected to open column chromatography with hydrophobic resin
220 CHP20P (Mitsubishi Chemical, Tokyo, Japan) eluted with 10% MeOH, 60% MeOH, and MeOH.
221 The 60% MeOH fraction was concentrated and repeatedly subjected to HPLC purification to

obtain 9.2 mg of ferri-peucechelin using C18 column (4.6× 250 mm, Wakopak Handy-ODS, WAKO, Osaka, Japan), eluted with 3% MeCN/97%water containing 0.05%TFA at a flow rate of 1 mL/min and monitored at UV-VIS absorbance 435 nm.

Conversion of ferri-peucechelin into Ga- peucechelin via desferri-peucechelin

Ferri- peucechelin (8.0 mg) was dissolved in 3 mL of water. The solution was mixed with 3 mL of 1M 8-quinolinol and stirred at room temperature for 30 min Total 2 times of the two layer partition were performed using 6 mL of CH_2Cl_2 each time to get rid of ferri-8 quinolinol. The water layer was immediately collected and lyophilized by freeze-dryer. After dissolving the dry material in 2 mL of water, HPLC purification was performed using C18 column (4.6 × 250 mm, Wakopak Handy-ODS), eluted with 4% MeCN/96%water containing 0.05%TFA at a flow rate of 1 mL/min, and monitored at UV-VIS absorbance 215 nm to yield 5.0 mg of desferri-peucechelin. Desferri-peucechelin (5.0 mg) was dissolved in 2 mL of distilled water, and 10 mg of Gallium chloride (Sigma Aldrich, MO, USA) was added to convert it to Ga-peucechelin. After HPLC purification in the same conditin as described above, 4.8 mg of Ga-peucechelin was obtained.

NMR experiments

A NMR sample was prepared by dissolving the purified peptide in 500 µl of DMSO-d_6 (Sigma Aldrich). 1D ^{1}H, ^{13}C, DEPT-135, and all 2D NMR spectra were obtained on Bruker Avance800 spectrometer with quadrature detection (Bruker BioSpin, MA, USA). The 1D ^{1}H, ^{13}C, DEPT-135 spectra were recorded at 25°C with 15 ppm for proton and 240 ppm for carbon. The following 2D ^{1}H-NMR spectra were recorded at 25°C with 8 ppm or 15 ppm spectral widths in *t1* and *t2* dimensions in the phase-sensitive mode by States-TPPI method: two-dimensional (2D) double quantum filtered correlated spectroscopy (DQF-COSY), recorded with 512 and 2048 complex

245 points in *t1* and *t2* dimensions; 2D homonuclear total correlated spectroscopy (TOCSY) with
246 MLEV-17 mixing sequence, recorded with mixing time of 80 ms, 256 and 1024 complex points
247 in *t1* and *t2* dimensions; 2D nuclear Overhauser effect spectroscopy (NOESY), recorded with
248 mixing time of 200 ms, 512 and 2048 complex points in *t1* and *t2* dimensions. Water suppression
249 was performed using presaturation method. 2D ^{1}H-^{13}C heteronuclear single quantum correlation
250 (HSQC) and heteronuclear multiple bond connectivity (HMBC) spectra were acquired at 25°C in
251 the echo-antiecho mode or in the absolute mode, respectively. The ^{1}H-^{13}C HSQC and HMBC
252 spectra were recorded with 1024 × 512 complex points for 12 ppm in the ^{1}H dimension and 170
253 ppm in the ^{13}C dimension or for 10 ppm in the ^{1}H dimension and 220 ppm in the ^{13}C dimension,
254 respectively, at a natural isotope abundance. 2D ^{1}H-^{15}N HSQC spectrum was also recorded at
255 25°C with 1024 × 64 complex points for 12 ppm in the ^{1}H dimension and 50 ppm in the ^{15}N
256 dimension in the phase-sensitive mode by States-TPPI method at natural isotope abundance. All
257 NMR spectra were processed using XWINNMR (Bruker). Before Fourier transformation, the
258 shifted sinebell window function was applied to *t1* and *t2* dimensions except for the HMBC
259 spectrum. All ^{1}H and ^{13}C dimensions were referenced to DMSO-d_6 at 25 °C.

260 **MS experiments**

261 All mass spectra were recorded in the positive-ion mode. ESI-TOF MS spectra of the peptides
262 were recorded using a JEOL JMS-T100LP mass spectrometer (JEOL Ltd., Tokyo, Japan). For
263 HR ESI-TOF MS alanysis, reserpine was used as internal standard.

264 **Modified Marfey method**

265 Ferri-peucechelin (1.0 mg) was subjected to acid hydrolysis at 105 °C for 16 h with
266 concentrated HI (0.5 mL, WAKO), and the hydrolysate was dried by a freeze-dryer and
267 resuspended in H_2O (200 μL). To the hydrolysate, 10 μL of a solution of *N*α-(5-fluoro-2,4-

dinitrophenyl)-L-leucinamide (L-FDLA, Sigma-Aldrich) or D-FDLA (Sigma-Aldrich) in acetone was added at the concentration of 10 mg/ml and 100 μL of 1 M $NaHCO_3$, after which the mixtures were heated to 80 °C for 3 min. The reaction mixtures were cooled, neutralized with 2 N HCl (50 μL), and diluted with MeCN (200 μL). About 20 μL of each solution of FDLA derivatives was subjected to HPLC analysis with C18 column (4.6 × 250 mm, Wakopak Handy-ODS, WAKO). The DAD detector (MD-2018, JASCO, Tokyo, Japan) was used for detection of the amino acid derivatives accumulating the data of the absorbance from 220nm to 420 nm. The HPLC analysis was performed at a flow rate of 1 mL/min using solvent A (distilled water containing 0.05% TFA) and solvent B (MeCN containing 0.05%TFA) with a linear gradient mode from 0 min to 50 min, increasing percentage of solvent B from 20% to 60% . The retention times (min) of L- or D-FDLA derivatized amino acids in this HPLC condition were following;

Acknowledgement

This study was supported by the Japan Society for the Promotion of Science by Grants-in-aids (grant number 25350964).

Conflict of interest

The authors had no conflict of interest in undertaking this project.

287 **References**

288 Ahmad M (2011) Investigations into the isolation, structure elucidation and biosynthesis of
289 bioactive natural products. Doctoral dissertation, University of Warwick

290 Ahmed E, Holmstrom SJ (2014) Siderophores in environmental research: roles and applications.
291 Microb Biotechnol 7:196-208 doi:10.1111/1751-7915.12117

292 Challis GL, Ravel J (2000) Coelichelin, a new peptide siderophore encoded by the *Streptomyces*
293 *coelicolor* genome: structure prediction from the sequence of its non-ribosomal peptide
294 synthetase. FEMS Microbiol Lett 187:111-114

295 Doroghazi JR et al. (2014) A roadmap for natural product discovery based on large-scale
296 genomics and metabolomics. Nat Chem Biol 10:963-968 doi:10.1038/nchembio.1659

297 Harada KI, Fujii K, Hayashi K, Suzuki M, Ikai Y, Oka H (1996) Application of D,L-FDLA
298 derivatization to determination of absolute configuration of constituent amino acids in
299 peptide by advanced Marfey's method. Tetrahedron Lett 37:3001-3004

300 Kodani S et al. (2013a) Structure and biosynthesis of scabichelin, a novel tris-hydroxamate
301 siderophore produced by the plant pathogen *Streptomyces scabies* 87.22. Org Biomol
302 Chem 11:4686-4694 doi:10.1039/c3ob40536b

303 Kodani S, Kayameya MO, Yoshida M, Ochi K (2011) A new siderophore isolated from
304 *Streptomyces* sp. TM-34 with potent inhibitory activity against angiotensin-converting
305 enzyme. Eur J Org Chem 17:3191-3196

306 Kodani S, Kobayakawa F, Hidaki M (2013b) Isolation and structure determination of new
307 siderophore tsukubachelin B from *Streptomyces* sp. TM-74. Nat Prod Res 27:775-781
308 doi:10.1080/14786419.2012.698412

309 Kodani S, Komaki H, Suzuki M, Hemmi H, Ohnishi-Kameyama M (2015) Isolation and
310 structure determination of new siderophore albachelin from *Amycolatopsis alba*.
311 Biometals 28:381-389 doi:10.1007/s10534-015-9842-z
312 Lautru S, Deeth RJ, Bailey LM, Challis GL (2005) Discovery of a new peptide natural product
313 by *Streptomyces coelicolor* genome mining. Nat Chem Biol 1:265-269
314 doi:10.1038/nchembio731
315 Lazos O, Tosin M, Slusarczyk AL, Boakes S, Cortes J, Sidebottom PJ, Leadlay PF (2010)
316 Biosynthesis of the putative siderophore erythrochelin requires unprecedented crosstalk
317 between separate nonribosomal peptide gene clusters. Chem Biol 17:160-173
318 doi:10.1016/j.chembiol.2010.01.011
319 Park HM et al. (2013) Genome-based cryptic gene discovery and functional identification of
320 NRPS siderophore peptide in *Streptomyces peucetius*. Appl Microbiol Biotechnol
321 97:1213-1222 doi:10.1007/s00253-012-4268-9
322 Pullan ST, Chandra G, Bibb MJ, Merrick M (2011) Genome-wide analysis of the role of GlnR in
323 Streptomyces venezuelae provides new insights into global nitrogen regulation in
324 actinomycetes. BMC genomics 12:175 doi:10.1186/1471-2164-12-175
325 Robbel L, Knappe TA, Linne U, Xie X, Marahiel MA (2010) Erythrochelin--a hydroxamate-type
326 siderophore predicted from the genome of *Saccharopolyspora erythraea*. FEBS J
327 277:663-676 doi:10.1111/j.1742-4658.2009.07512.x
328 Seyedsayamdost MR, Traxler MF, Zheng SL, Kolter R, Clardy J (2011) Structure and
329 biosynthesis of amychelin, an unusual mixed-ligand siderophore from *Amycolatopsis* sp.
330 AA4. J Am Chem Soc 133:11434-11437 doi:10.1021/ja203577e

Stephan H, Freund S, Meyer I, Winkelmann G, Jung C (1993) Structure elucidation of the gallium-ornibactin complex by 2D-NMR spectroscopy. Lieb Anna Chem 43-48

Umezawa H et al. (1985) Foroxymithine, a new inhibitor of angiotensin-converting enzyme, produced by actinomycetes. J Antibiot (Tokyo) 38:1813-1815

Figure legends

Fig. 1 Chemical structure of desferri-peucechelin (**1**) and desferri-foroxymithine (**2**)

Fig. 2 Enlarged ^{1}H-NMR spectrum of Gallium-complex of peucechelin over 7.40-7.85 ppm

Fig. 3 a) Key TOCSY, NOESY, and HMBC correlations of **1**, b) ESI-MS/MS analyses of **1**

Fig. 4 Putative peucechelin biosynthetic gene cluster and proposed pathway for peucechelin synthesis. (a) Organization of the peucechelin biosynthesis gene cluster in *S. peucetius* ATCC 27952 (Park et al. 2013). (b) Proposed role of enzymes encoded by sp970 to sp980 in the biosynthesis of peucechelin. Capital letters A, C, E represent adenylation, condensation, and epimerization domains, respectively. Black filled circle represents a peptidyl carrier protein (PCP) domain.

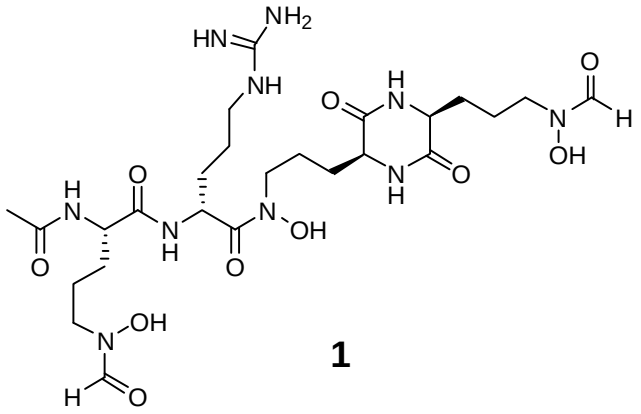
Fig. 5 Putative foroxymithine biosynthetic gene cluster and proposed pathway for foroxymithine synthesis. (a) Organization of the foroxymithine biosynthesis gene cluster in *S. venezuelae* ATCC 10712. The proposed functions of the protein encoded by the gene cluster are summarized in Table S2. (b) Proposed role of enzymes encoded by STRPU_RS0128210 to STRPU_RS0128225 in the biosynthesis of foroxymithine. Capital letters A and C represent adenylation and condensation domains, respectively. Black filled circle represents a peptidyl carrier protein (PCP) domain.

357 Figure 1

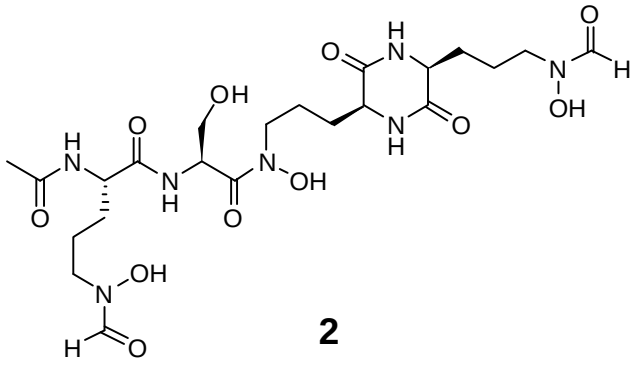
358

359

360



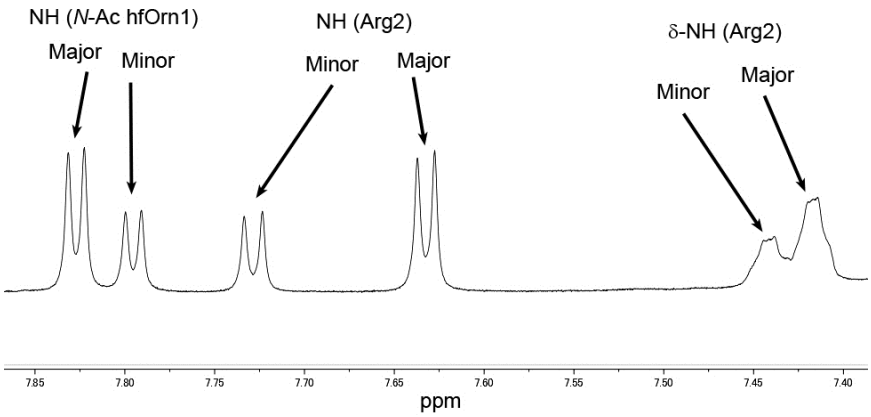
1



2

361 Figure 2

362



363 Figure 3

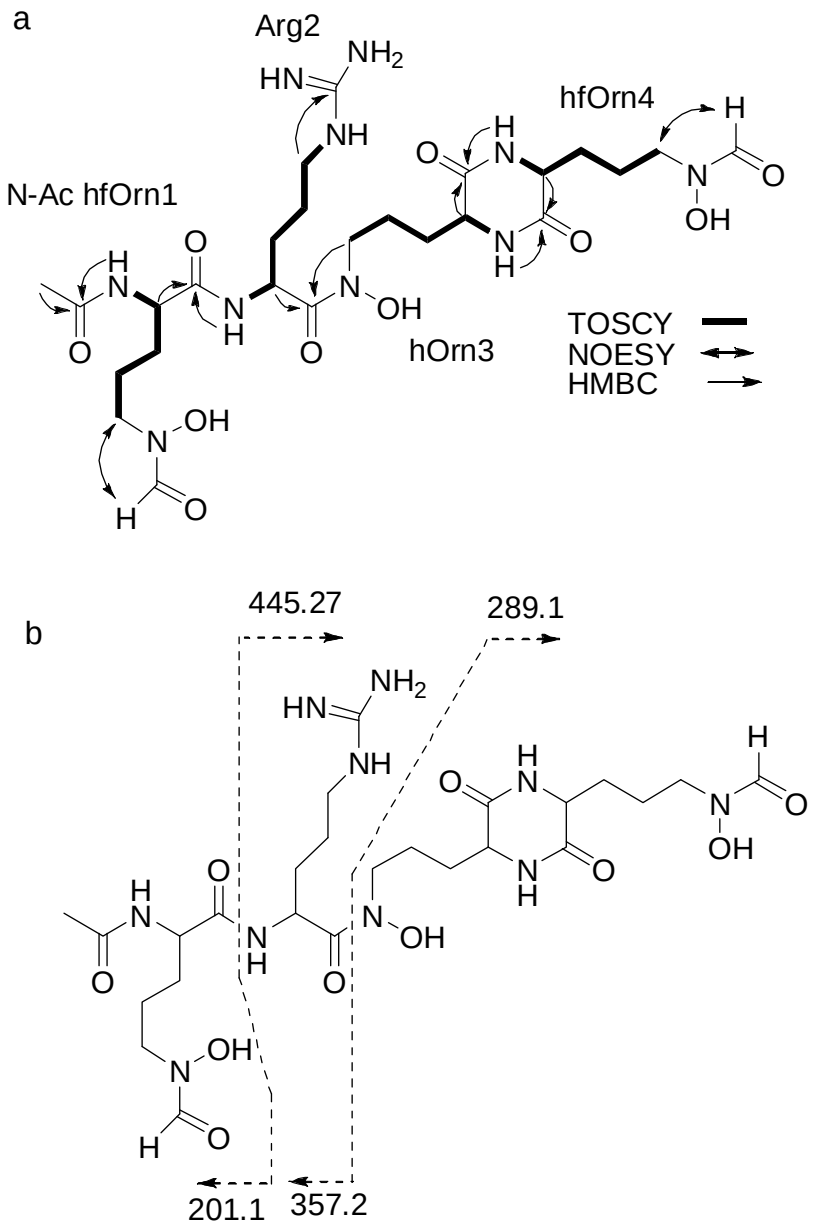
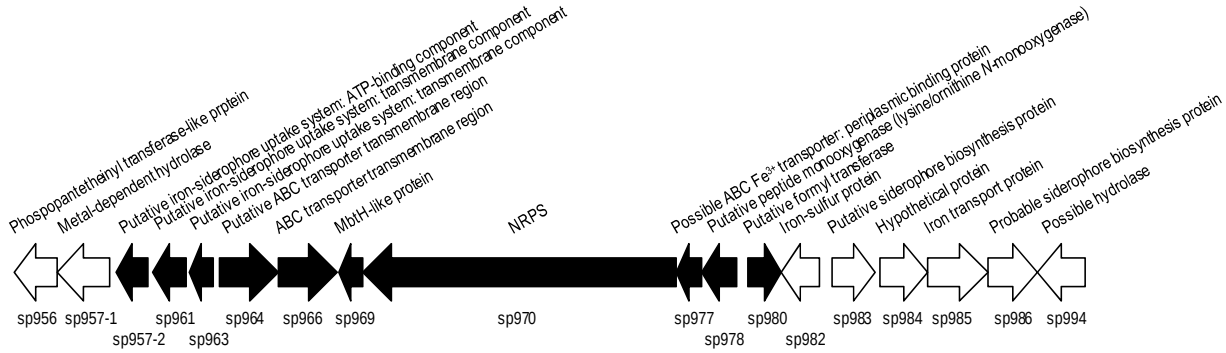
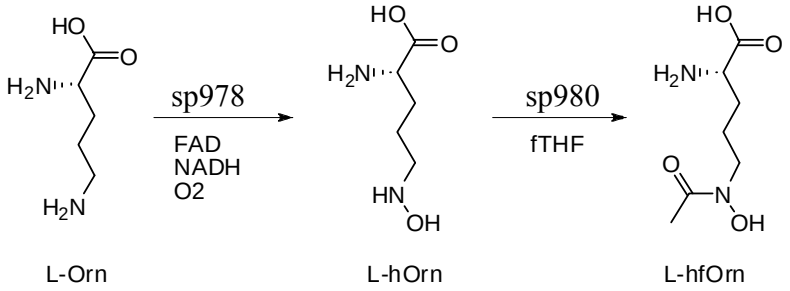

a
Arg2
hfOrn4
N-Ac hfOrn1
hOrn3
TOSCY
NOESY
HMBC
b
445.27
289.1
201.1
357.2

Figure 4

a


Phospopantetheinyl transferase-like prptein
Metal-dependent hydrolase
Putative iron-siderophore uptake system: ATP-binding component
Putative iron-siderophore uptake system: transmembrane component
Putative iron-siderophore uptake system: transmembrane component
Putative ABC transporter transmembrane region
ABC transporter transmembrane region
MbtH-like protein
NRPS
Possible ABC Fe3+ transporter: periplasmic binding protein
Putative peptide monooxygenase (lysine/ornithine N-monooxygenase)
Putative formyl transferase
Iron-sulfur protein
Putative siderophore biosynthesis protein
Hypothetical protein
Iron transport protein
Probable siderophore biosynthesis protein
Possible hydrolase
sp956
sp957-1
sp957-2
sp961
sp963
sp964
sp966
sp969
sp970
sp977
sp978
sp980
sp982
sp983
sp984
sp985
sp986
sp994

b


sp978
FAD
NADH
O2
sp980
fTHF
L-Orn
L-hOrn
L-hfOrn

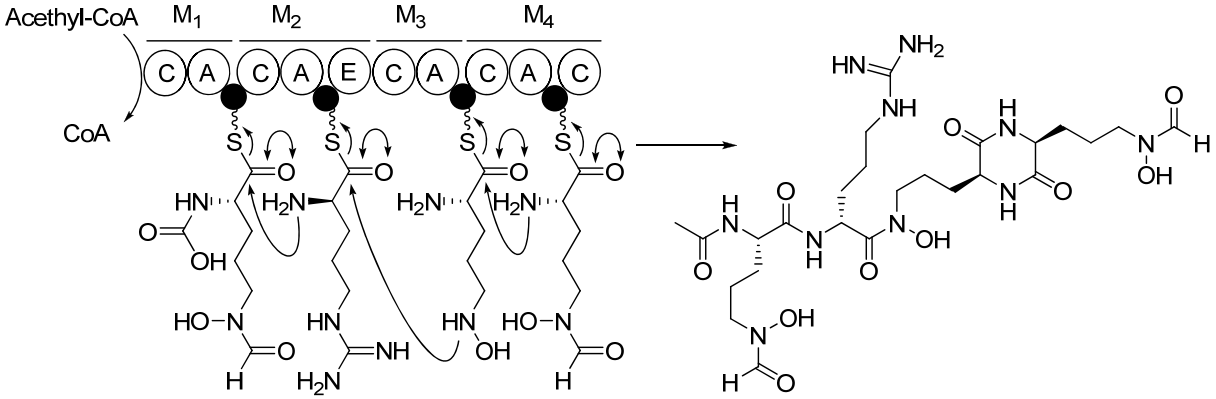
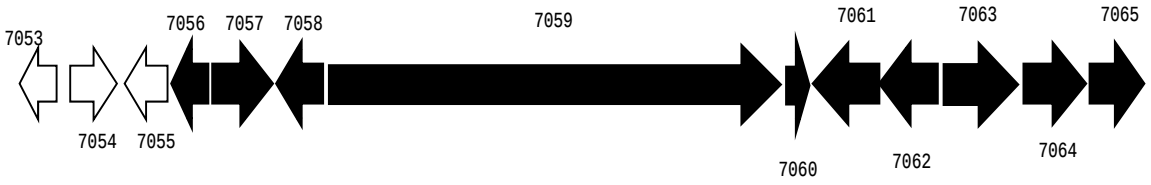
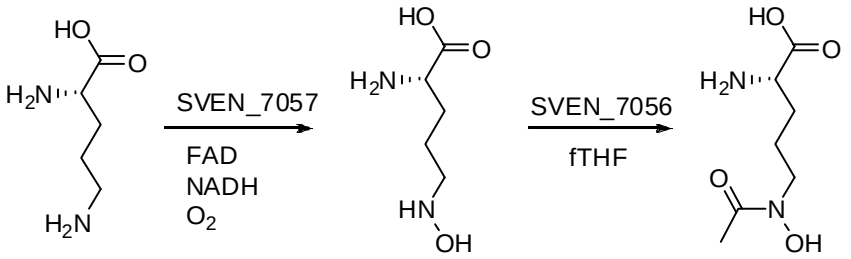

Acethyl-CoA
M1
M2
M3
M4
C A C A E C A C A C
CoA

Figure 5

a.


7053
7054
7055
7056
7057
7058
7059
7060
7061
7062
7063
7064
7065

b


SVEN_7057
FAD
NADH
O2
SVEN_7056
fTHF
L-Orn
L-hOrn
L-hfOrn

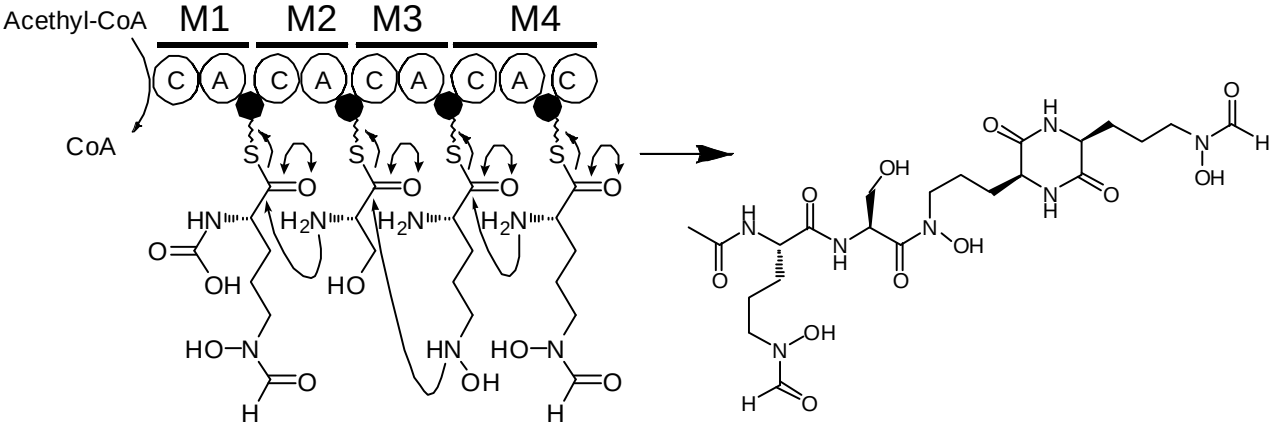

Acethyl-CoA
M1
M2
M3
M4
CoA

Table 1. NMR chemical shift values of Ga-peucechelin in DMSO-d_6

Major component				Minor component			
Residue	Position	δH (*J* inHz)	δC	Residue	Position	δH (*J* inHz)	δC
N-Ac hfOrn	*N*-Ac-CO		168.7	*N*-Ac hfOrn	*N*-Ac-CO		168.7
	N-Ac-Me	1.80 (s)	22.6		*N*-Ac-Me	1.80 (s)	22.6
	CO		172.7		CO		172.7
	NH	7.83 (d, 7.0)			NH	7.79 (d, 7.0)	
	α	3.74 (m)	54.6		α	3.74 (m)	54.6
	β	1.25 (m)	28.7		β	1.25 (m)	28.6
		1.46 (m)				1.46 (m)	
	γ	1.35(m)	20.2		γ	1.35(m)	20.3
		1.84 (m)				1.82 (m)	
	δ	3.46 (m)	50.6		δ	3.37 (m)	50.6
	formyl	8.06 (s)	152.5			3.47 (m)	
Arg	CO		162.0		formyl	8.09 (s)	152.8
	NH	7.64 (d, 7.7)		Arg	CO		162.0
	α	4.45 (m)	47.2		NH	7.73 (d, 7.7)	
	β	1.54 (m)	27.7		α	4.52 (m)	46.3
		1.69 (m)			β	1.54 (m)	27.7
	γ	1.32 (m)	24.7			1.62 (m)	
		1.58 (m)			γ	1.32 (m)	24.8
	δ	3.10 (m)	40.6			1.44 (m)	
	δ–NH	7.42(br)			δ	3.06 (m)	40.3
	C=N		156.5		δ–NH	7.44(br)	
hOrn	CO		167.7		C=N		156.5
	NH	8.24 (d, 3.3)		hOrn	CO		168.0
	α	3.67(m)	54.7		NH	8.09 (d, 3.3)	
	β	1.58 (m)	32.3		α	3.76(m)	54.9
		1.66 (m)			β	1.84 (m)	30.7
	γ	1.53 (m)	24.1			1.94 (m)	
		1.69 (m)			γ	1.71 (m)	22.9
	δ	3.46 (m)	50.5		δ	3.59 (m)	50.0
hfOrn	CO		166.9			3.99 (m)	
	NH	8.38 (d, 3.6)		hfOrn	CO		166.9
	α	3.53 (m)	54.5		NH	8.37 (d,3.6)	
	β	1.74 (m)	34.1		α	3.61 (m)	55.1
		1.99 (m)			β	1.67 (m)	32.2
	γ	1.54 (m)	26.2			1.77 (m)	
		1.67 (m)			γ	1.53 (m)	24.1
	δ	3.34 (m)	50.5			1.69 (m)	
		4.35 (m)			δ	3.64 (m)	50.6
	formyl	8.18(s)	152.7		formyl	8.09(s)	152.8

Table 2. A-domain specificity-determining residues and amino-acid substrates

A-domain	Residue (according to Grs numbering)								Substrate
	235	236	239	278	299	301	322	330	
sp970-M1	D	V	W	I	L	V	A	V	fhOrn
SVEN 7059-M1	D	V	W	I	L	V	A	V	fhOrn
STRPU_RS0128225-M1	D	V	W	I	L	V	A	V	fhOrn
AMYAL_RS0130210-M1	D	V	W	I	L	V	A	V	fhOrn
SCAB_85471-M1	D	V	W	I	L	G	A	T	fhOrn
sp970-M4	D	V	F	I	V	V	A	V	fhOrn
SVEN 7059-M4	D	V	F	I	V	V	A	V	fhOrn
STRPU_RS0128225-M4	D	V	F	V	V	V	A	V	fhOrn
sp970-M2	D	P	L	D	F	G	V	I	Arg
SVEN 7059-M2	D	V	W	H	F	S	L	V	Ser
STRPU_RS0128225-M2	D	V	W	H	F	S	L	V	Ser
AMYAL_RS0130210-M2	D	V	W	H	F	S	L	V	Ser
SCAB_85471-M2	D	V	W	H	L	S	L	V	Ser
ErcD-M2	D	V	W	H	F	S	L	V	Ser
CchH-M2	D	F	W	N	I	G	M	V	Ser
sp970-M3	D	M	W	N	L	G	L	I	hOrn
SVEN 7059-M3	D	M	W	N	L	G	L	I	hOrn
STRPU_RS0128225-M3	D	M	W	N	V	L	L	I	hOrn
AMYAL_RS0130210-M3	D	M	W	N	L	G	L	I	hOrn
SCAB_85471-M3	D	M	W	N	L	G	L	I	hOrn
ErcD-M3	D	M	W	N	L	G	L	I	hOrn
CchH-M3	D	M	W	N	L	G	L	I	hOrn

AMYAL_RS0130210, SCAB_85471, ErcD and CchH are NRPSs for amychelin, scabichelin, erythrochelin and coelichelin, respectively, sp970 is that for peucechelin, and SVEN 7059 and STRPU_RS0128225 are those for foroxymithines. M1, module 1; M2, module 2; M3, module 3; M4, module 4.