environmental microbiology

Environmental Microbiology (2012) doi:10.1111/1462-2920.12075

Identification and structural characterization of serobactins, a suite of lipopeptide siderophores produced by the grass endophyte *Herbaspirillum seropedicae*

Federico Rosconi,[1]* Danilo Davyt,[3] Verónica Martínez,[3] Marcela Martínez,[2] Juan Andrés Abin-Carriquiry,[2] Hannah Zane,[4] Alison Butler,[4] Emanuel M. de Souza[5] and Elena Fabiano[1]

[1]*Departamento de Bioquímica y Genómica Microbianas and* [2]*Plataforma HPLC, IIBCE, Montevideo, Uruguay.*
[3]*Cátedra de Química Farmacéutica, Facultad de Química, Montevideo, Uruguay.*
[4]*Department of Chemistry and Biochemistry, UCSB, Santa Barbara, USA.*
[5]*Departamento de Bioquimica e Biologia Molecular, UFPR, Curitiba, Brazil.*

Summary

***Herbaspirillum seropedicae* Z67 is a diazotrophic endophyte able to colonize the interior of many economically relevant crops such as rice, wheat, corn and sorghum. Structures of siderophores produced by bacterial endophytes have not yet been elucidated. The aim of this work was to identify and characterize the siderophores produced by this bacterium. In a screening for mutants unable to produce siderophores we found a mutant that had a transposon insertion in a non-ribosomal peptide synthase (NRPS) gene coding for a putative siderophore biosynthetic enzyme. The chemical structure of the siderophore was predicted using computational genomic tools. The predicted structure was confirmed by chemical analysis. We found that siderophores produced by *H. seropedicae* Z67 are a suite of amphiphilic lipopeptides, named serobactin A, B and C, which vary by the length of the fatty acid chain. We also demonstrated the biological activity of serobactins as nutritional iron sources for *H. seropedicae*. These are the first structurally described siderophores produced by endophytic bacteria.**

Received 2 May, 2012; revised 28 November, 2012; accepted 10 December, 2012. *For correspondence. E-mail federh@iibce.edu.uy; Tel. (+598) 24871616 ext. 146; Fax (+598) 24875461. The authors have declared that no competing interest exists.

Introduction

Endophytic bacteria colonize internal tissues of plants without causing disease to the host (Ryan *et al*., 2008; Reinhold-Hurek and Hurek, 2011). *Herbaspirillum seropedicae* Z67, a member of the *Betaproteobacteria* class, is a plant endophytic bacterium able to colonize and survive within tissues of many important agricultural crops such as rice, sorghum, wheat, corn and sugar cane (Baldani *et al*., 1986; James *et al*., 2002). This bacterium, which was isolated in Brazil (the name refers to the location of the EMBRAPA National Center for Agrobiology in Seropédica, Rio de Janeiro), is of biotechnological interest due to its nitrogen-fixing and plant growth-promoting capabilities (Gyaneshwar *et al*., 2002; Roncato-Maccari *et al*., 2003).

The nitrogen-fixation process is highly dependent on iron (Rees and Howard, 2000). We demonstrated that in *H. seropedicae* Z67 the activity and expression of nitrogenase respond to iron availability (Rosconi *et al*., 2006). Our results indicated that under iron sufficient conditions, nitrogenase activity and *nif* gene expression are optimal. Whereas when the iron level is reduced, *nifH* and *nifA* expression is depressed and under severe iron depletion the nitrogenase activity is completely lost. In addition to biological nitrogen fixation, numerous metabolic processes depend on iron including respiration, amino acid synthesis, and the Krebs cycle (Andrews *et al*., 2003). Despite its relevance, information about systems involved in iron homeostasis in *H. seropedicae* is scarce. Moreover, very little is known about the iron role in the interaction between plant and endophytes. With the exception of epichloënin, a ferrichrome-related siderophore produced by the fungal endophyte *Epichloë festucae* (Koulman *et al*., 2012), to our knowledge, no other siderophore produced by a plant endophyte has been structurally and functionally characterized.

Analysis of the published genome sequence of *H. seropedicae* SmR1 showed that at least 27 genes are

predicted to be involved in iron transport and metabolism. A putative gene (Hsero_2343) coding for a modular peptide synthase has been identified as a likely candidate in siderophore biosynthesis (Pedrosa *et al*., 2011).

Siderophores are low molecular weight metabolites with masses below 2000 Da that have a very high affinity for ferric ion (Schalk *et al*., 2011). Their main role is to provide the cell with nutritional iron. Siderophore production is widespread among bacteria and fungi and is found even in higher plants. The major Fe^{3+} ligands found in bacterial and fungal siderophores are catecholates, hydroxamic acids, and α-hydroxycarboxylic acids, although the overall structures differ widely among known siderophores (Miethke and Marahiel, 2007). Depending on the biosynthetic pathways, siderophores can be classified as Non-Ribosomal Peptide Synthetases (NRPS)-dependent or NRPS-independent (Challis, 2005).

NRPSs are the focus of intense investigation. These synthetases are modular enzymes containing consecutive arrays of tandem domains and are involved in the synthesis of peptide siderophores, antibiotics or surfactants (Strieker *et al*., 2010). Each module is responsible for the incorporation of a specific residue into the peptide backbone. Usually, modules are colinear with the amino acids they link according to the N-terminal-to-C-terminal colinearity rule (Grunewald and Marahiel, 2006). Each NRPS module contains at least three domains: an adenylation domain (A), a peptidyl-carrier protein (PCP) or T domain and a condensation domain (C). The A domain is responsible for the recognition of the amino acid and its activation as an acyl adenylate. The activated amino acid is then transferred to the thiol group of the 4′-phosphopantetheine cofactor present at the PCP domain where a thioester bond is produced. The C domain joins two substrates through a condensation reaction performing the peptide elongation. An additional domain, the thioesterase (TE) domain, located at the end of the module, catalyses the release of the peptide by either hydrolysis or macrocyclization (Strieker *et al*., 2010). This basic structure can be complemented by substrate-modifying domains, including substrate epimerization, α-hydroxylation, N methylation, and heterocyclic ring formation.

A large database of novel NRPS gene sequences is present in microbial genomes and metagenomes. *In silico* analyses of the amino acid sequences allow the substrate specificity of the A-domain to be predicted through particular sequence signatures or through machine learning methods (Jenke-Kodama and Dittmann, 2009; Rottig *et al*., 2011). Phylogeny and sequence examination also allow the prediction of functional subtypes of C domains ($^{L}C_{L}$, $^{D}C_{L}$, starter, heterocyclization, epimerization, and dual E/C domains), giving relevant information to aid structural determination, such as prediction of the amino acid isomeric form (D or L) in the nascent peptide, the presence of an β-hydroxy-carboxylic acid in the first amino acid, cyclization of cysteine, serine or threonine residues or amino acid epimerization, etc. (Rausch *et al*., 2007). Genome mining and bioinformatics approaches are promising strategies for the discovery and design of new peptides derived from NRPS (Winter *et al*., 2011).

In this work we identified an NRPS of the plant growth-promoting endophyte *H. seropedicae* Z67 responsible for synthesizing siderophores. Computational genomics of the NRPS sequence enabled prediction of the siderophore structure. The isolated siderophores, named the serobactins A–C, were structurally characterized, as reported herein. Furthermore, we showed that these siderophores can be used by *H. seropedicae* as an iron source. These results suggest that production of siderophores by *H. seropedicae* is an important trait in environments where iron is scarce.

Results

Identification of an NRPS gene implicated in siderophore biosynthesis in H. seropedicae *Z67*

Colonies of *H. seropedicae* Z67 produce a strong orange halo on chromo azurolsulfonate (CAS) plates indicative of siderophore production. We used this CAS assay in a genetic screen for mutants unable to produce siderophores. From random mutagenized derivatives of *H. seropedicae* Z67, a mutant named Hs 4.3 that did not produce a visible halo on CAS plates (Fig. 1A) was selected. Southern blot analysis confirmed that the Hs 4.3 mutant had a single transposon insertion. Supernatants from cultures of wild type and Hs 4.3 mutant were applied to Sep Pak cartridges and the presence of iron-binding compounds in the eluents was investigated by using the CAS assay. Only the supernatant of the wild-type strain grown in low iron media showed CAS activity (Fig. 1B). Thus, the interrupted locus was essential for siderophore production in the CAS medium, suggesting that *H. seropedicae* Z67 produces a unique siderophore or that different siderophores produced share a common biosynthetic gene.

The DNA sequence of the 5′ transposon flanking region in the Hs 4.3 mutant was determined. Results obtained by BLASTx searches in the *H. seropedicae* SmR1 genome (Pedrosa *et al*., 2011) indicated that transposon insertion was in a locus identical (100% identity) to Hsero_2343. The genome sequence of *H. seropedicae* Z67 has been sequenced by the Parana State Genome Programme (Genopar Consortium) using the next generation platform ABI-Solid, and only single nucleotide polymorphisms were found in comparison with the genome of *H. seropedicae* SmR1 strain (E. M. de Souza, results unpubl.). Therefore we retrieved the entire sequence of Hsero_2343 for further analyses. The deduced product

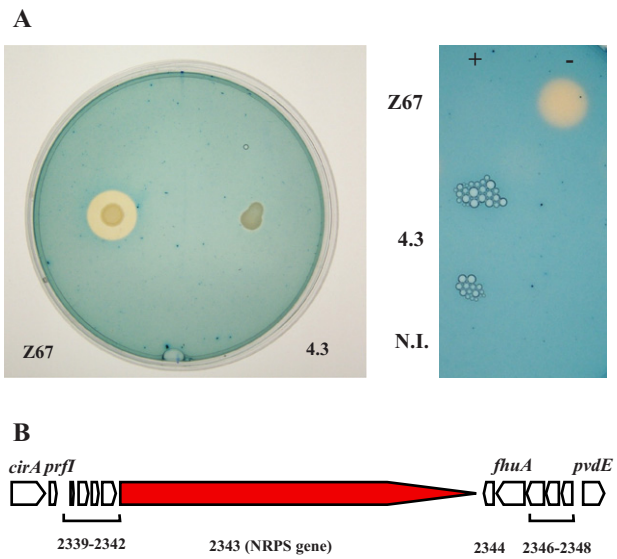

Fig. 1. Siderophore production of *H. seropedicae* Z67 wild-type strain and Hs 4.3 mutant strain grown on solid or in liquid medium.
A. Ten microlitres of a late exponential phase cultures containing 1×10^6 cfu were spotted on the CAS solid medium and incubated 48 h at 30°C.
B. Supernatants obtained from 50 ml of early stationary phase cultures (containing about 1×10^8 cfu ml^{-1}) of wild-type *H. seropedicae* Z67 and Hs 4.3 mutant strains grown in NfbHP-GG media with (+) or without (–) 37 μM $FeCl_3$ were purified with Sep Pak cartridges as described in *Experimental procedures*. Five microlitres of each purified supernatants were spotted on CAS solid medium. N.I. refers to non-inoculated media.
C. Genomic context of the NRPS interrupted gene (red) in the Hs 4.3 mutant. Positions of predicted ORFs are indicated by arrows and correspond to the following Locus tags: Hsero_2337 (*cirA* gene), Hsero_2338 (*prfI* gene), Hsero_2339 to Hsero_2343 (NRPS gene), Hsero_2344, Hsero_2345 (*fhuA* gene), Hsero_2346 to Hsero_2348 and Hsero_2349 (*pvdE* gene) from left to right respectively. Information of each ORF is summarized in Table 1.

of Hsero_2343 is a NRPS of 9160 amino acids (993 755 Da). The genetic context of Hsero_2343 in *H. seropedicae* SmR1 is depicted in Fig. 1C and information of each locus summarized in Table 1. This context includes genes encoding proteins with homology to NRPS tailoring enzymes, those involved in siderophore biosynthesis and processing, regulators of siderophore production and siderophore transporters (Agnoli *et al.*, 2006; Cornelis, 2010).

Prediction of the siderophore structure

To obtain insight into the chemical structure of the *H. seropedicae* Z67 siderophore, we applied a computational analysis of the NRPS coded by Hsero_4323. Knowledge of the probable structure is a useful tool for the design and analysis of purification and chemical characterization approaches. The domain architecture of the enzyme encoded by the Hsero_2343 locus and the specificity of A domains were predicted using publically available software (Ansari *et al.*, 2004; Bachmann and Ravel, 2009; Finn *et al.*, 2010; Markowitz *et al.*, 2010; Rottig *et al.*, 2011) and the predicted architecture is shown in Fig. 2. The predicted biosynthetic cluster comprises six classical NRPS modules (C-A-T), and a terminal TE. The first domain of the predicted protein presented homology with AMP-dependent synthetase/ligase (Pfam 00501, $e = 2.1^{-91}$), acyl-CoA synthetase related enzymes (PTHR 22754:SF4 $e = 7.4^{-136}$) and to members of the fatty acyl-AMP ligase family (FAAL, cd05931, $e = 0$). Taking into account the 30 and 31% identity with the first domain of NRPSs involved in iturin and mycosubtilin lipopeptide biosynthesis respectively (Hansen *et al.*, 2007; Chooi and Tang, 2010) and the 52% identity with the first domain of PvdL responsible for pyoverdin acylation (Drake and Gulick, 2011), we hypothesized that the first residue in the siderophore was a fatty acid. *In silico* analysis of the adenylation domains indicated that the subsequent five residues were aspartic acid, serine, aspartic acid, threonine and serine, in that order. The last residue could not be clearly predicted, as different software analysis predicted different amino acids. Probable residues were glutamine, lysine or N^5-hydroxyornithine according to the procedure described by Ansari and colleagues (2004) and

Table 1. Genomic context of *H. seropedicae* NRPS gene.

Locus tag	ORF predicted function	Gene annotation	GenBank Accession
Hsero_2337	TonB-dependent siderophore receptor protein	*cirA*	YP_003775744
Hsero_2338	Ferric regulator sigma-24 subunit	*pfrI*	YP_003775745
Hsero_2339	MbtH-like protein		YP_003775746
Hsero_2340	thioesterase involved in non-ribosomal peptide biosynthesis protein		YP_003775747
Hsero_2341	4′-phosphopantetheinyl transferase (EC:2.7.8.-)		YP_003775748
Hsero_2342	clavaminic acid synthetase like protein		YP_003775749
Hsero_2343	non-ribosomal peptide synthetase		YP_003775750
Hsero_2344	esterase/lipase		YP_003775751
Hsero_2345	TonB-dependent receptor protein	*fhuA*	YP_003775752
Hsero_2346	L-ornithine N^5-monooxygenase (EC:1.13.12.-)		YP_003775753
Hsero_2347	acetyltransferase-like protein		YP_003775754
Hsero_2348	Iron reductase		YP_003775755
Hsero_2349	siderophore (cyclic peptide) ABC transporter protein	*pvdE*	YP_003775756

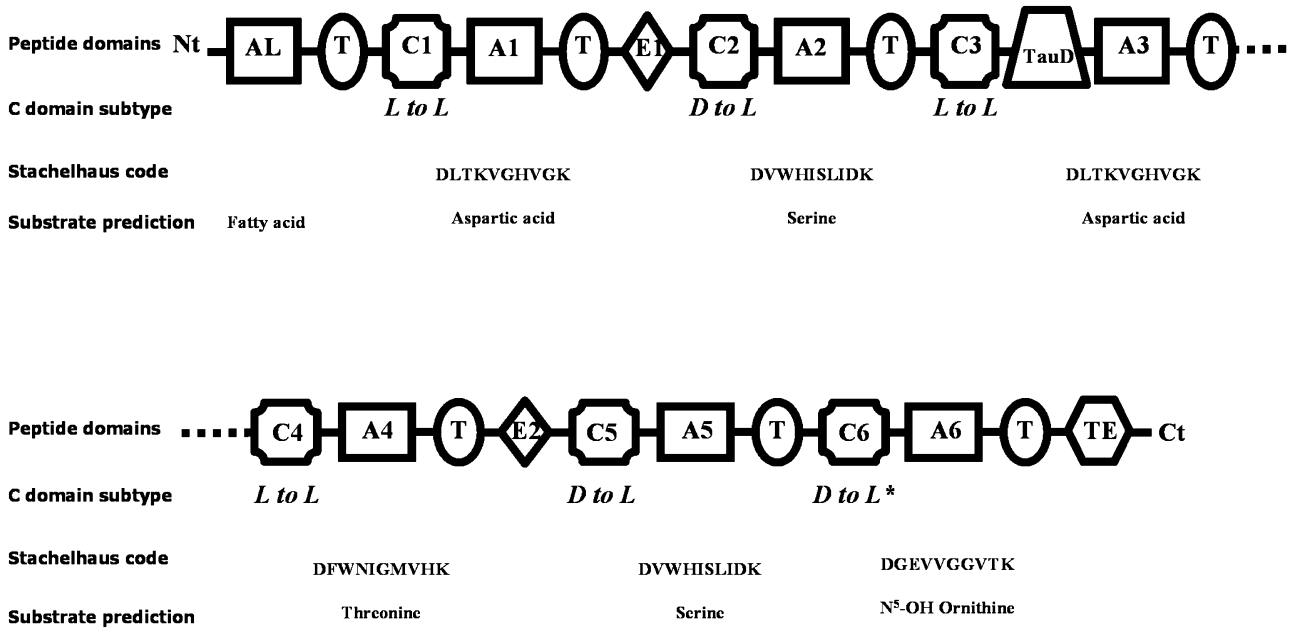

Fig. 2. Domain architecture of the NRPS protein codified by the Hsero_2343 locus. Below each A domain (A1 to A6) is shown the Stachelhaus '8 amino acid code' (delimited by the conserved aspartate and lysine residues) and the predicted substrate. Below each C domain (C1 to C6) is shown the predicted subtype according to the phylogenetic and functional classification described by Rausch and colleagues (2007). *Although the sixth C domain clustered within the $^{D}C_{L}$ branch, biochemical characterization showed that it corresponds to a $^{L}C_{L}$ subtype. Domains symbols: AL, acyl ligase; A, adenylation; T, thiolation; C, condensation; E, epimerization; TauD, pfam02668.

Rottig and colleagues (2011) or by Bachmann and Ravel (2009). Analysis of the genomic context in *H. seropedicae* SmR1 indicates that Hsero_2346 and Hsero_2347 loci encode proteins with homology to L-lysine/ornithine-N-monooxygenase and acetyl transferase-like proteins respectively (Table 1). It has been reported that the concerted action of both enzymes is responsible for the insertion of an hydroxamic acid group into other siderophores (Challis, 2005). Thus we propose that the last residue is converted to hydroxamic acid. The presence of a TauD domain in the NRPS (Fig. 2) and the proximity to a putative TauD homologue (encoded by Hsero_2342) (Table 1) suggested hydroxylation of the aspartic acid residues (Singh *et al.*, 2008). The presence and location of two epimerase domains (Fig. 2) suggested that the first aspartic acid (β-hydroxyaspartic acid) and threonine (fifth residue) have a D configuration. Moreover, the second C-domain (presumably responsible for the condensation of the first β-hydroxyaspartic acid with serine), and the fifth C-domain (probably involved in the condensation of threonine and serine) were phylogenetically related to $^{D}C_{L}$-domain subtypes (Fig. S1) (Rausch *et al.*, 2007) involved in the condensation of D-amino acid in the nascent peptide to the L-amino acid of the next NRPS module. The sixth C-domain (predicted to be responsible for the condensation of serine with the last residue) also clustered within the $^{D}C_{L}$ branch, although no associated epimerase domain was observed (Figs 2 and S1).

Purification and structural characterization of the siderophores produced by H. seropedicae

In order to proceed with the chemical characterization and purification of the siderophore, supernatants of *H. seropedicae* Z67 cells grown in NfbHP-GG media were collected and subjected to purification by analytical and preparative HPLC (high-performance liquid chromatography) (see elution profiles in Fig. S2). Three metabolites with CAS activities were isolated corresponding to retention times of 16.5 min, 19.9 min and 23.7 min. These compounds were named serobactin A, serobactin B and serobactin C respectively. The amounts recovered of each compound were less than one milligram. The MALDI-TOF mass spectrum of the major component, serobactin B, exhibited a peak of m/z 872.37 Da (Fig. S3A) which was determined to be the $[M + Na]^{+}$ ion by MALDI-TOF MS/MS. Masses of 'y' fragments (Roepstorff and Fohlman, 1984) obtained by MALDI-TOF tandem MS/MS indicated that the second to the sixth residues were β-hydroxyaspartic acid, serine, β-hydroxyaspartic acid, threonine and serine (Fig. S3B). This peptide sequence is in agreement with the serobactin-predicted structure obtained by computational genomics (Fig. 3A). The mass obtained for the seventh residue (153.07 Da) did not correspond to N^{5}-OH-ornithine, but rather to a N^{5}-OH-cycloornithine associated with a Na^{+} ion. This result supports one of the *in silico*

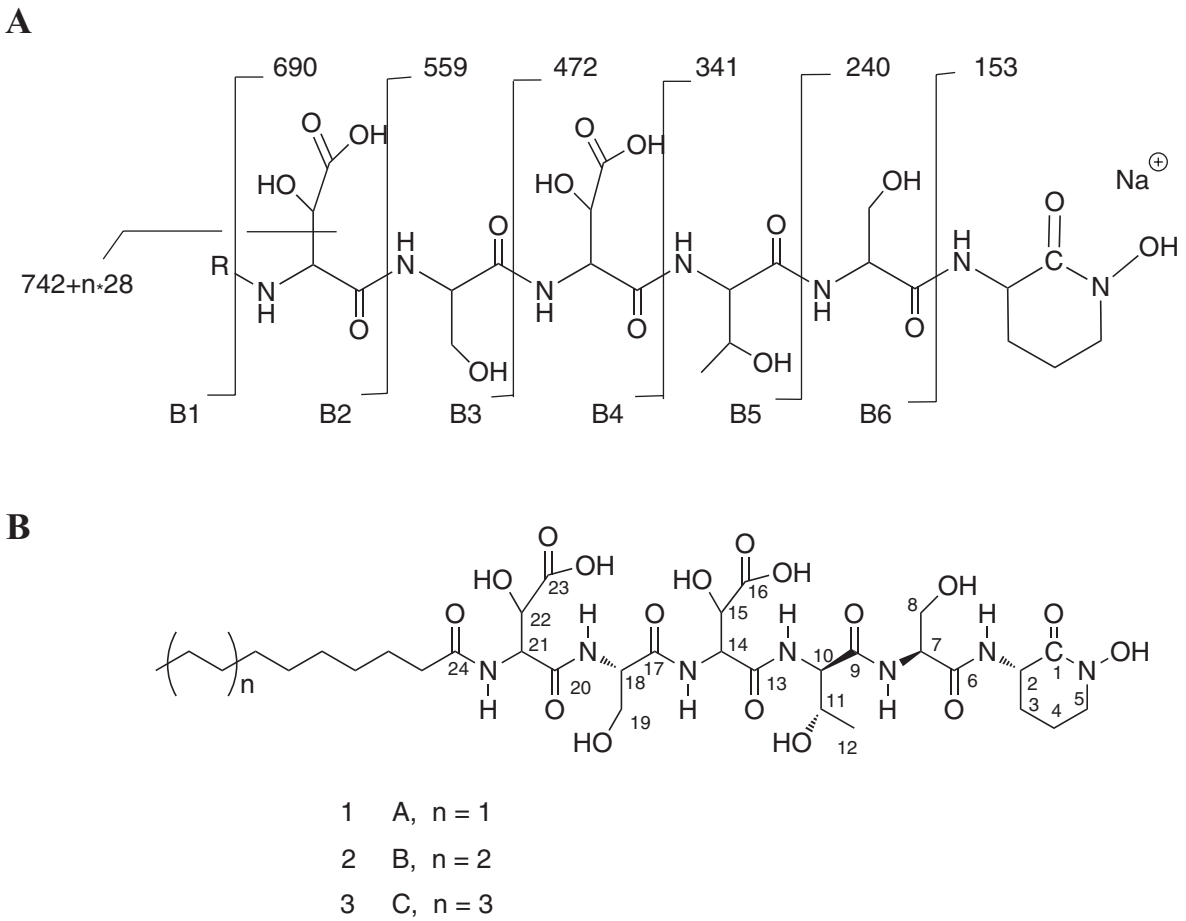

Fig. 3. Chemical structures of serobactins A, B and C.
A. Structure of fragments obtained by MALDI-TOF MS/MS. Vertical lines show the masses $[M + Na + H]^+$ of the 'y' fragments (Roepstorff and Fohlman, 1984) (see Fig. S3B). These fragments are common for serobactin A, B and C. R represents the fatty acid tail appendages of serobactins A, B and C.
B. Overall structure of serobactins A, B and C.

predictions, suggesting that the protein encoded by Hsero_2346 (a probable L-ornithine-N^5-monooxygenase) could be involved in the hydroxylation of the N^5 of ornithine, followed by cyclization by the last TE domain. As a result of the action of these two enzymes, a hydroxamic acid group is generated. The cyclic nature of the last residue was later confirmed by analysing the methylated derivatives of serobactin B by MALDI-TOF mass spectrometry, which showed an increase of only 28 Da (m/z 900.42 Da) indicating dimethylation of serobactin B (Fig. S4A), as opposed to 42 Da if a terminal carboxylate had been present. Moreover, results obtained by MALDI-TOF tandem MS/MS confirmed that methylation occurs at the β-hydroxyaspartic residues (Fig. S4B). Based on the data obtained by MS and MS/MS (not shown), the structures of serobactin A and serobactin C are homologous of serobactin B, with the decrease of 28 Da or increase of 28 Da, respectively, at the N-terminus of the peptide (Fig. 3B).

By using a preparative purification protocol (Fig. S2B), 10 mg of serobactin A (retention time of 25.5 min) and 16 mg of serobactin B plus serobactin C (which eluted together at 29.0 min) were obtained, along with a minor peak with a retention time of 27.5 min. This minor peak had a mass of 866.37 Da, and the MALDI-TOF MS/MS analysis indicated that it had the same peptide core of serobactins A, B and C with a different fatty acid residue. Unfortunately, this minor compound could not be further characterized. Serobactin B and serobactin C were successfully separated by an alternative preparative purification protocol performed under isocratic conditions (not shown).

The final structure of serobactin B was elucidated by NMR (nuclear-magnetic resonance). The ^1H and ^{13}C chemical shift assignments are summarized in Fig. 4, Table 2 and Fig. S5. The results confirmed the amino acid sequence indicated by mass spectrometry. Analyses of NMR spectra, as well as the absence of free amines

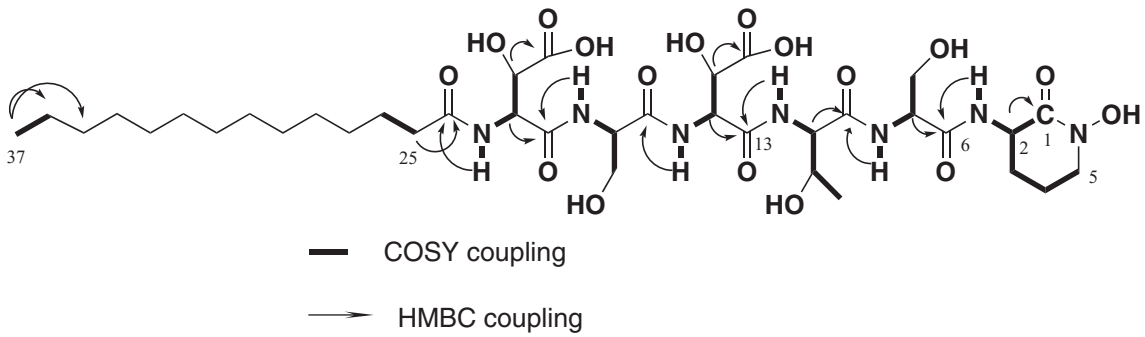

Fig. 4. Selected correlations in 2D NMR spectra of serobactin B.

determined by EttanCaf (not shown), support results obtained by MALDI-TOF analysis, demonstrating that the C-terminal residue is a cyclic N^5-OH-ornithine.

The ^{1}H-NMR spectrum of serobactin B displayed signals distinctive of a lipid chain (16 protons at 1.24 ppm coupled to a triplet signal at 0.86 ppm and to methylenes at 1.47 and 2.15 ppm). The ^{1}H-NMR spectrum presented six signals characteristic of amides, while the 2D spectra (TOCSY and HMBC) indicated that one resonance corresponded to an amide coupled with a fatty acid chain (Fig. 4).

Finally, to confirm that the serobactins are lipopeptides distinguished by different fatty acids, the serobactins A, B and C were partially hydrolysed and their methyl esters were analysed by GC-MS to identify the fatty acids. Decanoic, lauric and myristic acids were found in serobactins A, B and C respectively. The GC-MS spectrum of serobactin B is shown in Fig. S6. In conclusion, *H. seropedicae* Z67 produces a new suite of amphiphilic siderophores with a common peptide core substituted by fatty acids of different lengths (Fig. 3B).

Amino acid chirality analysis

As stated above, the domain architecture of the NRPS responsible for the biosynthesis of serobactin, predicted that the first β-hydroxyaspartic acid and the threonine have D configurations. In order to confirm this prediction, the amino acid analysis was carried out using Marfey's reagent, as previously carried out for characterization of other siderophores (Martinez *et al*., 2000; Bhushan and Bruckner, 2004; Homann *et al*., 2009; Vraspir *et al*., 2011; Kreutzer *et al*., 2012). The amino acid content of serobactin was revealed through both reductive hydrolysis with 2M HI and non-reductive hydrolysis with 6M HCl (Fig. S7A and B, Table S1). Stereochemical assignments of peaks were made by co-injections with Marfey's derivatized standards (data not shown). Results obtained demonstrate that the molecule contains D-threonine, both DL-*threo*-β-hydroxyaspartic acids, L-serine and

Table 2. ^{1}H (400 MHz) and ^{13}C NMR (100 MHz) data for serobactin B (12 mg) in d_6-DMSO.

	δ_H (*J* in Hz)	δ_C (*J* in Hz)	HMBC
	Cyclic, *N*-OH Ornithine		
1	–	170.5	1
2	4.31, m	49.8	
3	1.67,m; 1.92,bs	27.6	
4	1.92, bs	20.7	
5	3.46, m	51.2	
N2	8.06, d (8.6)		6
	Serine		
6	–	169.9	
7	4.31, m	55.2	
8	3.60, bs	62.0	
N3	7.83, d (7.9)		9
	Threonine		
9	–	170.3	
10	4.21, m	58.8	9, 11
11	3.80, m	68.9	
12	1.04, d (6.0)	20.5	10, 11
N4	7.73, d (7.9)		13
	β-OH Aspartic Acid		
13	–	169.2	
14	4.70, d (8.6)	55.6	15
15	4.51, d (8.6)	70.0	16
16	–	173.5	
N5	8.06, d (8.6)		17
	Serine		
17	–	170.1	
18	4.39, m	54.6	
19	3.42, m; 3.66, m	61.8	
N6	7.67, d (7.2)		20
	β-OH AsparticAcid		
20	–	169.9	
21	4.70, d (8.6)	56.1	22, 23
22	4.51, d (8.6)	70.4	23
23	–	173.5	
N7	7.87, d (8.6)		24
	FattyAcid Tail		
24	–	173.1	
25	2.15, t (6.6)	35.7	24, 26
26	1.47, bs	25.7	
27	1.15–1.35 bs	29.0	
28	1.15–1.35 bs	29.2	
29	1.15–1.35 bs	29.3	
30	1.15–1.35 bs	29.5	
31	1.15–1.35 bs	29.4	
32	1.15–1.35 bs	29.3	
33	1.15–1.35 bs	31.7	
34	1.15–1.35 bs	22.5	
35	0.86, t (6.4)	14.4	35, 36

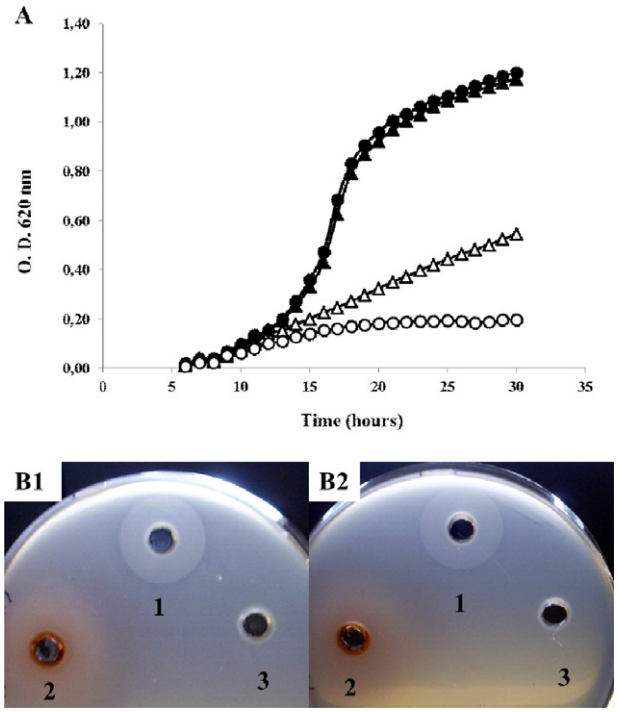

Fig. 5. Growth assays.
A. Growth of *H. seropedicae* Z67 wild-type strain (triangles) and of Hs 4.3 mutant strain (circles) in NfbHP-GG medium containing 37 μM FeCl3 (solid symbols), and in NfbHP-GG medium plus 25 μM 2,2′-dipyridyl (open symbols).
B. Bioassays of *H. seropedicae* Z67 wild-type strain (B1) and of Hs 4.3 mutant strain (B2). Bacteria were grown in TY solid medium plus 500 μM EDDHA. Twenty microlitres of a 300 nM Fe^{3+}-serobactin (well 1), 37 mM $FeCl_3$ (well 2) and purified supernatant of Hs 4.3 mutant cells (well 3) were added to a 5 mm hole in the agar.

L-ornithine. The absence of a D-serine showed that the sixth condensation domain in *H. seropedicae* NRPS corresponds to a $^{L}C_{L}$ subtype.

The iron complex of serobactin

Iron(III) complexation by serobactins is expected to occur through coordination of both β-hydroxyaspartic acid groups and the hydroxamic acid group present in the cyclic N^5-hydroxyornithine, in an hexadentate manner. Thus, a 1:1 stoichiometry is expected for the Fe^{3+}-serobactins complexes. Additionally, the presence of β-hydroxyaspartic acid suggested that the Fe^{3+}-serobactin complex would be photoreactive, as has been described for other siderophores containing α-hydroxycarboxilic acids (Barbeau *et al.*, 2001; Vraspir and Butler, 2009; Butler and Theisen, 2010; Kreutzer *et al.*, 2012). To test these hypotheses, equimolar quantities of serobactin A and $FeCl_3$ were mixed and the product was then purified by HPLC and analysed by ESI-MS (Fig. S8A and B). The peaks m/2z (436.12 Da) and m/z (873.28 Da) correspond to the Fe^{3+}-serobactin A complex confirming a 1:1 stoichiometry. The UV-Vis spectrum (Fig. S8C) shows two absorption bands at *c.* 300 nm and 440 nm, consistent with coordination by both β-hydroxyaspartic acid groups and the hydroxamic acid group to Fe^{3+} (Barbeau *et al.*, 2001; Homann *et al.*, 2009). After 180 min of UV radiation, the charge-transfer band from hydroxamic acid-Fe^{3+} (440 nm) remained while the charge-transfer band from β-hydroxyaspartatic acids-Fe^{3+} (300 nm) was lost indicating that the Fe^{3+}-serobactin complexes are photoreactive.

Biological activity of serobactins

Growth of the wild type and the Hs 4.3 mutant strains were evaluated in liquid NfbHP-GG medium supplemented with 37 μM $FeCl_3$ or with 25 μM 2,2′-dipyridyl. Both the mutant and the wild type grew at slower rate in the iron depleted medium, but the former was more sensitive to the presence of the chelator (Fig. 5A). In order to assess the use of serobactins as iron nutritional source, we performed bioassays as previously described (Amarelle *et al.*, 2010). As shown in Fig. 5B1 and B2, ferric serobactins were able to restore the growth of the wild type and the Hs 4.3 mutant strains in the presence of the iron chelator. Moreover, the Sep Pak-extracted supernatant of the Hs 4.3 mutant could not restore the growth. Both, growth and bioassay experiments, demonstrated the biological role of serobactins as nutritional iron sources.

Discussion

In this work we structurally characterized the serobactins, a new suite of amphiphilic siderophores produced by the grass endophyte *H. seropedicae* Z67. In addition we demonstrate that serobactins are the only siderophores produced by *H. seropedicae* Z67 under the conditions assayed. This conclusion is based on the fact that the Hs 4.3 mutant strain did not produce a visible halo on CAS plates and no detectable fractions with siderophore activity were detected in Hs 4.3 culture supernatants. Serobactins share two structural features common to many siderophores produced by aquatic bacteria discovered to date: (i) production of families of amphiphilic compounds, consisting of a central Fe^{3+}-binding headgroup that is appended by one of a series of fatty acids and; (ii) the presence of an α-hydroxycarboxylic acid moiety (β-hydroxyaspartic acid or citric acid), which is photoreactive when coordinated to Fe^{3+} (Sandy and Butler, 2009). Siderophores produced by soil bacteria with these characteristics include corrugatin and ornicorrugatin from *Pseudomonas* spp. (Matthijs *et al.*, 2008), ornibactins from *Burkholderia* species (Stephan *et al.*, 1993; Agnoli *et al.*, 2006), and rhizobactin 1021, from *Sinorhizobium*

meliloti 1021 (Smith *et al*., 1985; Nogales *et al*., 2010). Corrugatin and ornibactins have shorter fatty acid tails (no more than 8 carbons in length) and, in the case of ornibactins, they are attached to the N^5 of the first ornithine residue while in corrugatin they are attached to the N^2 amine group. In turn, rhizobactin 1021 is a non-peptide NRPS-independent siderophore. The recently described siderophore cupriachelin produced by *Cupriavidus necator* H16, a bacterium isolated from sludge in a creek, is structurally very similar to serobactins (Kreutzer *et al*., 2012). Moreover, the genomic context of the cupriachelin biosynthetic genes is conserved in comparison with the *H. seropedicae* SmR1 genome. The two β-hydroxyaspartic acid and the cyclic N^5 OH-ornithine residues are responsible for iron coordination. These residues are also present in other siderophores such as the pyoverdines and loihichelins (Visca *et al*., 2007; Homann *et al*., 2009).

The biological advantages conferred by amphiphilicity and photoreactivity in marine bacteria may be related to diffusion gradients and iron recycling in the water column reached by sunlight (Sandy and Butler, 2009; Vraspir and Butler, 2009; Butler and Theisen, 2010). Amphiphilicity of rhizobactin 1021 produced by *S. meliloti* 1021, as well as other non-siderophore lipopeptides produced by *Pseudomonas* species have been demonstrated to be involved in swarming motility (Berti *et al*., 2007; Nogales *et al*., 2010). Natural niches of *H. seropedicae* include rhizosphere and internal tissues of tropical grasses, which are protected from sunlight. However, *H. seropedicae* was found in high numbers in the xylem vessels of leaves of rice, sorghum, maize, and wheat plants (James *et al*., 1997; Gyaneshwar *et al*., 2002; Roncato-Maccari *et al*., 2003) and even on the surface of rice leaves (Gyaneshwar *et al*., 2002), which is exposed to sunlight, suggesting that serobactin photoreactivity may important for survival of the bacterial in these niches. Also, the affinity of α-hydroxycarboxylates siderophores for ferric ion under low pHs makes them efficient under these conditions (Miethke and Marahiel, 2007).

During the last decade, in depth studies on NRPSs were undertaken by many groups (Baltz *et al*., 2005; Challis, 2008; Nikolouli and Mossialos, 2012), due mainly to the interest of discovery, combinatorial synthesis and exploitation of novel and already described non-ribosomal peptides for diverse environmental and pharmaceutical applications. The increase in whole-genome and metagenome sequencing projects, and bioinformatics advances allow the prediction with accuracy of the number and type of building blocks that are incorporated into a final NRPS product, as well as the prediction of modifications from tailoring domains (Winter *et al*., 2011). This work is another example of the successful application of genomics, bioinformatics and biochemical approaches in NRPS studies. Structural characterization of NRPS products also gives feedback on *in silico* techniques, in addition to increasing biochemical characterized databases on which prediction approaches are based. As an example, in this work we found that the sixth NRPS condensation domain clustered phylogenetical close to the $^{D}C_{L}$ subtype, but it acts as an $^{L}C_{L}$, possibly due to the absence of an associated epimerase domain (Rausch *et al*., 2007).

The relevance of iron uptake systems mediated by siderophores is clearly established for animal pathogens (Miethke and Marahiel, 2007). In phytopathogens the importance is less clear, depending on the host–pathogen system studied (Pandey and Sonti, 2010; Jones and Wildermuth, 2011), while the importance of siderophores in the interaction of bacterial endophytes with its hosts, is largely unknown. The complete genomes of 12 endophytic bacteria and a rice endophyte metagenome show numerous copies of genes coding putative TonB-dependent receptors and iron storage proteins, suggesting that iron acquisition systems have a relevant role in part or all of the life cycle of endophytes (Reinhold-Hurek and Hurek, 2011; Sessitsch *et al*., 2012).

In this work, we demonstrated the biological importance of serobactins in iron acquisition by *H. seropedicae* Z67, suggesting that the production of the siderophores by this bacterium is an important trait in environments where iron is scarce. Additionally the high number of putative TonB dependent receptors genes present in the *H. seropedicae* genome may suggest that siderophore 'piracy' (Traxler *et al*., 2012) or utilization of other iron nutritional sources, such as hemin or citrate, are also possible strategies used by this bacterium to circumvent iron deficiency.

Experimental procedures

Bacteria and media

Herbaspirillum seropedicae Z67 (ATCC 35892) is a nalidixic acid resistant, wild-type strain. Siderophore production of *H. seropedicae* Z67 can be clearly evaluated as the formation of an orange halo on a blue background using the CAS agar plate method (Schwyn and Neilands, 1987). Strain Hs 4.3 is a derivative mutant of *H. seropedicae* Z67. This mutant was obtained by random mutagenesis performed with the mini-transposon mTn5gusA-O-pgfp as previously described (Rosconi *et al*., 2006). The Hs 4.3 mutant was selected by screening transconjugants for the absence of an orange halo on CAS plates. Southern analysis of Pvu I-digested genomic DNA using biotinylated mTn5gusA-O-pgfp as a probe was performed in order to confirm a single insertion of the mini-transposon in the genome of the Hs 4.3 mutant.

Strains were grown at 30°C in the TY media (tryptone 5 g l^{-1}, yeast extract 3 g l^{-1}, $CaCl_2$ 0.1 g l^{-1}) (Beringer, 1974) or in the NfbHP-GG medium for siderophore production. The NfbHP-GG media is a modification of the NfbHP-malate media (Klassen *et al*., 1997), with the addition of 11 mM glucose, as a carbon source instead of malate, 3 mM

glutamate, 20 mM NH_4Cl and 0.4 mM biotin. Different iron availability conditions were obtained with the addition of 37 μM $FeCl_3$, the metal-chelators 2,2′-dipyridyl (DP) or ethylenediamine-di-*o*-hydroxyphenylacetic acid (EDDHA) as indicated for each assay. Bioassay experiments were carried out as previously described (Amarelle *et al*., 2010).

For growth assays, *H. seropedicae* Z67 and Hs 4.3 mutant strain were grown in Cell Star®, Greiner Bio One 96-well plates, and incubated at 30°C in a Varioskan Flash® (Thermo). Growth was monitored by absorbance at 620 nm. Different media were inoculated with cells in late exponential phase previously grown in NfbHP-GG media containing 37 μM $FeCl_3$ and diluted to an initial cell density of 1×10^5 colony-forming units (cfu) ml^{-1}.

Bioinformatic analyses of NRPS sequences

Substrate specificity of A-domains was determined using the '8 amino acid code' method (Stachelhaus *et al*., 1999) and transductive support vector machines (TSVMs) (Rausch *et al*., 2005). Software was obtained from three different web servers (Ansari *et al*., 2004; Bachmann and Ravel, 2009; Rottig *et al*., 2011). The amino acid sequence of a NRPS (YP_003775750.1) present in the *H. seropedicae* SmR1 genome (Pedrosa *et al*., 2011) was used as a query for these bioinformatic analysis.

Phylogenetic analysis of NRPS condensation and epimerase domains was performed as described (Rausch *et al*., 2007). Briefly, a subsets of at least ten different domains from each subtype (L to L, D to L, Dual, Starter, Epimerase, Heterocyclization) together with the predicted condensation and epimerase domains obtained from the bioinformatic analysis of the YP_003775750.1 amino acid sequence, were used to build a phylogenetic tree using MEGA, Neighbour-joining method (NJ), Jones-Taylor-Thornton (JTT) amino acid substitution matrix and gamma distributed rate variation with four categories. The support values were based on 100-fold bootstrapping.

Siderophore purification

For small-scale purification procedures, *H. seropedicae* Z67 strain was grown at 30°C for 24 h in 50 ml NfbHP-GG medium without added iron. Cells were removed by centrifugation at 10 000 *g* for 10 min and the supernatant was filtered through a 0.45 μm pore filter. The pH of the clear supernatant was adjusted to 3 with 10% TFA and loaded onto a C18 Sep Pak cartridge (Waters) previously equilibrated in 0.01% TFA. The cartridge was washed first with 10 ml of 0.01% TFA and then with 10 ml of a mixture of methanol and 0.01% TFA (20:80). Samples were eluted with 5 ml of methanol : 0.01% TFA (80:20), the eluted fractions were concentrated by vacuum and suspended in 1/10 initial volume of a mixture of methanol : DMSO : H_2O (60:25:15). Siderophores were purified by HPLC after loading 200 μl of the concentrated sample (diluted in an equal volume of 0.1% formic acid) in a C18 Phenomenex Luna analytical HPLC column using 10% of solvent B (0.01% formic acid in acetonitrile) and 90% solvent A (0.01% formic acid in water) for 40 min and then 90% solvent B and 10% solvent A for 7 min. The elution was monitored at 210 nm and siderophore activity was evaluated by the production of orange halos of 10 μl spots on CAS plates.

For large-scale purification of *H. seropedicae* Z67 siderophore, the wild-type strain was grown in 1 l NfbHP-GG medium at 30°C for 24 h. The supernatant was recovered, filtered and adjusted to pH 3 as described before. The total sample (1 l) was loaded in a 50 ml medium pressure column of C18 reverse phase. The loaded column was washed with 0.01% TFA and methanol : 0.01% TFA (20:80) and the elution was achieved with methanol : 0.01% TFA (80:20). Fractions with siderophore activity were pooled, concentrated by rotary vacuum evaporation and lyophilized. Samples were solubilized in methanol and further purified on a C4 preparative HPLC column. Fractions showing siderophore activity were concentrated using a rotary vacuum evaporator and then lyophilized. Lyophilized samples were suspended in 3 ml of 25% DMSO, loaded on a Sep Pak cartridge and eluted as described above. Finally, samples were concentrated using a rotary vacuum evaporator, lyophilized and stored at 4°C.

For bioassays, supernatants obtained from 50 ml of cultures (1×10^8 cfu ml^{-1}) in stationary phase of *H. seropedicae* Z67 and Hs 4.3 mutant cells grown in NfbHP-GG media without $FeCl_3$ added were purified by Sep Pak cartridges as described above. Samples were incubated with 300 nM $FeCl_3$ and diluted half in Tris 10 mM pH 7.5. Free iron was removed by incubation with Chelex 100® (Bio-Rad), according to manufacturer's instructions.

Mass spectrometry

The masses and fragmentation patterns of the HPLC-purified compounds were analysed in a 4800 MALDI-TOF/TOF mass spectrometer (AbiSciex) and in a Micromass Q-TOF2 (Waters Corp.). For MALDI-TOF/TOF analysis, aliquots of the samples were previously mixed with the same volume of a matrix-saturated solution of alpha-cyano-4-hydroxycinnamic acid (CHA) in 50% acetonitride–0.1% trifluoroacetic acid. The mixture was deposited directly onto the MALDI target and allowed to crystallize at room temperature for at least 1 min before mass analysis. All spectra were acquired by operating the mass spectrometer in the reflector mode with an accelerating voltage of 20 kV, and 250 laser shots per spectrum. External calibration was done by using a peptide standard mixture spotted immediately next to the sample. The purified Fe^{3+}-serobactin A complex was directly injected in a ESI-MS Bruker Micromass QTOF2 (Waters Corp) using N_2 as the collision gas and analysed in the negative mode.

To obtain methyl ester derivatives, 1 μl of air-dried samples were incubated for 30 min with 50 μl of a solution of methanol and acetyl chloride (5:1) at room temperature in the dark. Samples were then concentrated to dryness using a vacuum centrifuge, washed twice with methanol and solubilized in 0.5 μl of 0.1% TFA. Derivatives were analysed by MALDI-TOF.

Detection of free amines was carried out by using the Ettan CAF MALDI Sequencing Kit (Amersham) according to the manufacturer's instruction.

For the identification of fatty acids present in serobactin variants, samples were incubated for 3 h with methanolic HCl at 110°C. Fatty acids were extracted with n-hexane, dried

over $MgSO_4$, and analysed using a Varian 4000 GC-MS equipped with a DB-225-MS column (J&W) (Sassaki *et al*., 2008).

Nuclear-magnetic resonance spectroscopy

Approximately 12 mg of serobactin B was dissolved in 0.6 ml of dimethylsulfoxide-d_6 (d_6-DMSO, 99.9% Cambridge Isotopes). The 1D ^1H and ^{13}C spectra; the homonuclear 2D COSY and TOCSY spectra, and the ^1H–^{13}C *g*HSQC and *g*HMBC spectra were recorded on a Bruker Avance 400 instrument at room temperature using standard pulse software. The TOCSY spectrum was recorded by using a mixing time of 60 ms. The 1D spectra were acquired with 65 536 data points, whereas 2D spectra were collected using 2048 points in the F2 dimension and 256 increments in the F1 dimension. Chemical shifts of ^1H and ^{13}C were referenced to the solvent signals. The spectra were processed using topspin, version 2.0.

Amino acid analysis

Configuration of the amino acids present in serobactins was determined using Marfey's reagent (Bhushan and Bruckner, 2004). Purified serobactin A was hydrolysed with 2M HI or 6M HCl at 100°C for 24 h. Samples were then derivatized with Marfey's reagent (1-fluoro-2,4-dinitrophenyl-5-L-alanine amide; 1% w/v in acetone) as described (Martinez *et al*., 2000; Bhushan and Bruckner, 2004; Homann *et al*., 2009; Vraspir *et al*., 2011; Kreutzer *et al*., 2012) and loaded onto a C4 reverse phase HPLC analytical column. Derivatized serobactins were eluted with a linear gradient of 10% solvent B (acetonitrile) and 90% solvent A (50 mM triethylamine phosphate, pH 3.0) to 40% solvent B and 60% solvent A, for 45 min. Elution was monitored at 340 nm.

Identification of the D- and L-amino acids was accomplished by co-injection of amino acid standards that were derivatized with Marfey's reagent (Sigma-Aldrich).

Photorreactivity of the ferric-serobactin A complex

Purified serobactin A and $FeCl_3$ were mixed in equimolar quantities (2.4 mM, 3 ml total volume), the pH was adjusted to 9 and complex formation was allowed to proceed overnight at 4°C. The complex was purified by HPLC using a C4 reverse phase column under the aforementioned conditions. Fractions containing the Fe^{3+}-serobactin A complex were collected, concentrated via rotary vacuum evaporation, lyophilized and resuspended in 3 ml of double distilled H_2O. The UV-Vis spectrum (300–550 nm) was recorded before and after 3 h of continuous irradiation to UV light with a 200 W mercury arc lamp.

Acknowledgements

This work was partially supported by ANII and PEDECIBA, Uruguay and PROSUL-CNP-q, Brasil, and NSF CHE-1059067 (Alison Butler) grant.

We thank Rosario Durán (UByPA, Institute Pasteur Montevideo, Uruguay) and Luciano Huergo (Departamento de Bioquimica e Biologia Molecular, UFPR, Brazil) for MS-MS analyses, Arquimedes Paixão de Santana Filho (Laboratório de Química de Carboidratos, UFPR, Brazil) and Cecilia Callejas (Unidad de Microbiología Molecular, IIBCE, Uruguay) for GC-MS analyses. We also thank M. G. Yates for critical reading of the manuscript.

References

Agnoli, K., Lowe, C.A., Farmer, K.L., Husnain, S.I., and Thomas, M.S. (2006) The ornibactin biosynthesis and transport genes of *Burkholderia cenocepacia* are regulated by an extracytoplasmic function sigma factor which is a part of the Fur regulon. *J Bacteriol* **188:** 3631–3644.

Amarelle, V., Koziol, U., Rosconi, F., Noya, F., O'Brian, M.R., and Fabiano, E. (2010) A new small regulatory protein, HmuP, modulates haemin acquisition in *Sinorhizobium meliloti*. *Microbiology* **156:** 1873–1882.

Andrews, S.C., Robinson, A.K., and Rodriguez-Quinones, F. (2003) Bacterial iron homeostasis. *FEMS Microbiol Rev* **27:** 215–237.

Ansari, M.Z., Yadav, G., Gokhale, R.S., and Mohanty, D. (2004) NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. *Nucleic Acids Res* **32:** W405–W413.

Bachmann, B.O., and Ravel, J. (2009) Chapter 8. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. *Methods Enzymol* **458:** 181–217.

Baldani, J., Baldani, V.L.D., Seldin, L., and Döbereiner, J. (1986) Characterization of *Herbaspirillum seropedicae* gen. Nov., a root-asociated nitrogen-fixing bacterium. *Int J Syst Bacteriol* **36:** 86–93.

Baltz, R.H., Miao, V., and Wrigley, S.K. (2005) Natural products to drugs: daptomycin and related lipopeptide antibiotics. *Nat Prod Rep* **22:** 717–741.

Barbeau, K., Rue, E.L., Bruland, K.W., and Butler, A. (2001) Photochemical cycling of iron in the surface ocean mediated by microbial iron(III)-binding ligands. *Nature* **413:** 409–413.

Beringer, J.E. (1974) R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* **84:** 188–198.

Berti, A.D., Greve, N.J., Christensen, Q.H., and Thomas, M.G. (2007) Identification of a biosynthetic gene cluster and the six associated lipopeptides involved in swarming motility of *Pseudomonas syringae* pv. tomato DC3000. *J Bacteriol* **189:** 6312–6323.

Bhushan, R., and Bruckner, H. (2004) Marfey's reagent for chiral amino acid analysis: a review. *Amino Acids* **27:** 231–247.

Butler, A., and Theisen, R.M. (2010) Iron(III)-siderophore coordination chemistry: reactivity of marine siderophores. *Coord Chem Rev* **254:** 288–296.

Challis, G.L. (2005) A widely distributed bacterial pathway for siderophore biosynthesis independent of nonribosomal peptide synthetases. *Chembiochem* **6:** 601–611.

Challis, G.L. (2008) Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology* **154:** 1555–1569.

Chooi, Y.H., and Tang, Y. (2010) Adding the lipo to lipopeptides: do more with less. *Chem Biol* **17:** 791–793.

Cornelis, P. (2010) Iron uptake and metabolism in pseudomonads. *Appl Microbiol Biotechnol* **86:** 1637–1645.

Drake, E.J., and Gulick, A.M. (2011) Structural characterization and high-throughput screening of inhibitors of PvdQ, an NTN hydrolase involved in pyoverdine synthesis. *ACS Chem Biol* **6:** 1277–1286.

Finn, R.D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., *et al*. (2010) The Pfam protein families database. *Nucleic Acids Res* **38:** D211–D222.

Grunewald, J., and Marahiel, M.A. (2006) Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides. *Microbiol Mol Biol Rev* **70:** 121–146.

Gyaneshwar, P., James, E.K., Reddy, P.M., and Ladha, J.K. (2002) *Herbaspirillum* colonization increases growth and nitrogen accumulation in aluminium-tolerant rice varieties. *New Phytol* **154:** 131–145.

Hansen, D.B., Bumpus, S.B., Aron, Z.D., Kelleher, N.L., and Walsh, C.T. (2007) The loading module of mycosubtilin: an adenylation domain with fatty acid selectivity. *J Am Chem Soc* **129:** 6366–6367.

Homann, V.V., Sandy, M., Tincu, J.A., Templeton, A.S., Tebo, B.M., and Butler, A. (2009) Loihichelins A-F, a suite of amphiphilic siderophores produced by the marine bacterium *Halomonas* LOB-5. *J Nat Prod* **72:** 884–888.

James, E.K., Olivares, F.L., Baldani, J.I., and Döbereiner, J. (1997) *Herbaspirillum*, an endophytic diazotroph colonizing vascular tissue *Sorghum bicolor* L. Moench. *J Exp Bot* **48:** 785–798.

James, E.K., Gyaneshwar, P., Mathan, N., Barraquio, W.L., Reddy, P.M., Iannetta, P.P., *et al*. (2002) Infection and colonization of rice seedlings by the plant growth-promoting bacterium *Herbaspirillum seropedicae* Z67. *Mol Plant Microbe Interact* **15:** 894–906.

Jenke-Kodama, H., and Dittmann, E. (2009) Bioinformatic perspectives on NRPS/PKS megasynthases: advances and challenges. *Nat Prod Rep* **26:** 874–883.

Jones, A.M., and Wildermuth, M.C. (2011) The phytopathogen *Pseudomonas syringae* pv. tomato DC3000 has three high-affinity iron-scavenging systems functional under iron limitation conditions but dispensable for pathogenesis. *J Bacteriol* **193:** 2767–2775.

Klassen, G., Pedrosa, F.O., Souza, E.M., Funayama, S., and Rigo, L.U. (1997) Effect of nitrogen compounds on nitrogenase activity in *Herbaspirillum seropedicae* SMR1. *Can J Microbiol* **43:** 887–891.

Koulman, A., Lee, T.V., Fraser, K., Johnson, L., Arcus, V., Lott, J.S., *et al*. (2012) Identification of extracellular siderophores and a related peptide from the endophytic fungus *Epichloe festucae* in culture and endophyte-infected *Lolium perenne*. *Phytochemistry* **75:** 128–139.

Kreutzer, M.F., Kage, H., and Nett, M. (2012) Structure and biosynthetic assembly of cupriachelin, a photoreactive siderophore from the bioplastic producer *Cupriavidus necator* H16. *J Am Chem Soc* **134:** 5415–5422.

Markowitz, V.M., Chen, I.M., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., *et al*. (2010) The integrated microbial genomes system: an expanding comparative analysis resource. *Nucleic Acids Res* **38:** D382–D390.

Martinez, J.S., Zhang, G.P., Holt, P.D., Jung, H.T., Carrano, C.J., Haygood, M.G., and Butler, A. (2000) Self-assembling amphiphilic siderophores from marine bacteria. *Science* **287:** 1245–1247.

Matthijs, S., Budzikiewicz, H., Schafer, M., Wathelet, B., and Cornelis, P. (2008) Ornicorrugatin, a new siderophore from *Pseudomonas fluorescens* AF76. *Z Naturforsch C* **63:** 8–12.

Miethke, M., and Marahiel, M.A. (2007) Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev* **71:** 413–451.

Nikolouli, K., and Mossialos, D. (2012) Bioactive compounds synthesized by non-ribosomal peptide synthetases and type-I polyketide synthases discovered through genome-mining and metagenomics. *Biotechnol Lett* **34:** 1393–1403.

Nogales, J., Dominguez-Ferreras, A., Amaya-Gomez, C.V., van Dillewijn, P., Cuellar, V., Sanjuan, J., *et al*. (2010) Transcriptome profiling of a *Sinorhizobium meliloti fadD* mutant reveals the role of rhizobactin 1021 biosynthesis and regulation genes in the control of swarming. *BMC Genomics* **11:** 157.

Pandey, A., and Sonti, R.V. (2010) Role of the FeoB protein and siderophore in promoting virulence of *Xanthomonas oryzae* pv. *oryzae* on rice. *J Bacteriol* **192:** 3187–3203.

Pedrosa, F.O., Monteiro, R.A., Wassem, R., Cruz, L.M., Ayub, R.A., Colauto, N.B., *et al*. (2011) Genome of *Herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grasses. *PLoS Genet* **7:** e1002064.

Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W., and Huson, D.H. (2005) Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res* **33:** 5799–5808.

Rausch, C., Hoof, I., Weber, T., Wohlleben, W., and Huson, D.H. (2007) Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evol Biol* **7:** 78.

Rees, D.C., and Howard, J.B. (2000) Nitrogenase: standing at the crossroads. *Curr Opin Chem Biol* **4:** 559–566.

Reinhold-Hurek, B., and Hurek, T. (2011) Living inside plants: bacterial endophytes. *Curr Opin Plant Biol* **14:** 435–443.

Roepstorff, P., and Fohlman, J. (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom* **11:** 601.

Roncato-Maccari, L.D., Ramos, H.J., Pedrosa, F.O., Alquini, Y., Chubatsu, L.S., Yates, M.G., *et al*. (2003) Endophytic *Herbaspirillum seropedicae* expresses *nif* genes in gramineous plants. *FEMS Microbiol Ecol* **45:** 39–47.

Rosconi, F., Souza, E.M., Pedrosa, F.O., Platero, R.A., Gonzalez, C., Gonzalez, M., *et al*. (2006) Iron depletion affects nitrogenase activity and expression of *nifH* and *nifA* genes in *Herbaspirillum seropedicae*. *FEMS Microbiol Lett* **258:** 214–219.

Rottig, M., Medema, M.H., Blin, K., Weber, T., Rausch, C., and Kohlbacher, O. (2011) NRPSpredictor2 – a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res* **39:** W362–W367.

Ryan, R.P., Germaine, K., Franks, A., Ryan, D.J., and Dowling, D.N. (2008) Bacterial endophytes: recent developments and applications. *FEMS Microbiol Lett* **278:** 1–9.

Sandy, M., and Butler, A. (2009) Microbial iron acquisition: marine and terrestrial siderophores. *Chem Rev* **109:** 4580–4595.

Sassaki, G.L., Souza, L.M., Serrato, R.V., Cipriani, T.R., Gorin, P.A., and Iacomini, M. (2008) Application of acetate derivatives for gas chromatography-mass spectrometry: novel approaches on carbohydrates, lipids and amino acids analysis. *J Chromatogr A* **1208:** 215–222.

Schalk, I.J., Hannauer, M., and Braud, A. (2011) New roles for bacterial siderophores in metal transport and tolerance. *Environ Microbiol* **13:** 2844–2854.

Schwyn, B., and Neilands, J.B. (1987) Universal chemical assay for the detection and determination of siderophores. *Anal Biochem* **160:** 47–56.

Sessitsch, A., Hardoim, P., Doring, J., Weilharter, A., Krause, A., Woyke, T., *et al*. (2012) Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *Mol Plant Microbe Interact* **25:** 28–36.

Singh, G.M., Fortin, P.D., Koglin, A., and Walsh, C.T. (2008) beta-Hydroxylation of the aspartyl residue in the phytotoxin syringomycin E: characterization of two candidate hydroxylases AspH and SyrP in *Pseudomonas syringae*. *Biochemistry* **47:** 11310–11320.

Smith, M.J., Shoolery, J.N., Schwyn, B., Holden, I., and Neilands, J.B. (1985) Rhizobactin, a structurally novel siderophore from *Rhizobium meliloti*. *J Am Chem Soc* **107:** 1739–1743.

Stachelhaus, T., Mootz, H.D., and Marahiel, M.A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol* **6:** 493–505.

Stephan, H., Freund, S., Beck, W., Jung, G., Meyer, J.M., and Winkelmann, G. (1993) Ornibactins – a new family of siderophores from *Pseudomonas*. *Biometals* **6:** 93–100.

Strieker, M., Tanovic, A., and Marahiel, M.A. (2010) Nonribosomal peptide synthetases: structures and dynamics. *Curr Opin Struct Biol* **20:** 234–240.

Traxler, M.F., Seyedsayamdost, M.R., Clardy, J., and Kolter, R. (2012) Interspecies modulation of bacterial development through iron competition and siderophore piracy. *Mol Microbiol* **86:** 628–644.

Visca, P., Imperi, F., and Lamont, I.L. (2007) Pyoverdine siderophores: from biogenesis to biosignificance. *Trends Microbiol* **15:** 22–30.

Vraspir, J.M., and Butler, A. (2009) Chemistry of marine ligands and siderophores. *Ann Rev Mar Sci* **1:** 43–63.

Vraspir, J.M., Holt, P.D., and Butler, A. (2011) Identification of new members within suites of amphiphilic marine siderophores. *Biometals* **24:** 85–92.

Winter, J.M., Behnken, S., and Hertweck, C. (2011) Genomics-inspired discovery of natural products. *Curr Opin Chem Biol* **15:** 22–31.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1. Identification of D and L configurations. Reductive and non-reductive hydrolysis refers to HPLC chromatograms of Fig. S7A and B respectively.

Fig. S1. Phylogenetic tree of different subtypes of condensation domains of *H. seropedicae* NRPS (Hsero_2343), including the condensation (C_DOMAIN_1 to 6, Nt to Ct, red boxes) and epimerase (E_DOMAIN_1 and 2, blue boxes). Nomenclature: C Domain Subtype_NCBI accession number.

Fig. S2. HPLC chromatograms obtained from analytical (A) or preparative (B) purifications of *H. seropedicae* Z67 siderophores.

A. HPLC from supernatant of a bacterial culture grown in 50 ml of a low-content iron media.

B. HPLC from supernatant of a bacterial culture grown in 1 l of low-content iron media.

Fig. S3. A. MALDI-TOF analyses of the CAS positive eluent with a retention time of 19.9 min obtained by analytical HPLC (Fig. S2A).

B. MALDI-TOF MS/MS anslysis of the 872 m/z peak. The vertical lines through the structure show the mass to charge $[M + Na + H]^+$ for the 'y' fragments (Roepstorff and Fohlman, 1984).

Fig. S4. Methanol:acetyl chloride derivatization of the 872.32 Da metabolite. Mass (A) and tandem mass spectra (B) of the methyl ester derivative of the 872.32 Da metabolite.

Fig. S5. A. ^{1}H-NMR of 12 mg of serobactin B in 600 µl d_6-DMSO.

B. ^{13}C-NMR of 12 mg of serobactin B in 600 µl d_6-DMSO.

C. HSQC of 12 mg of serobactin B in 600 µl d_6-DMSO.

Fig. S6. Identification of the first residue of serobactin B as dodecanoic acid.

A. GC chromatogram of derivatized fatty acids present in a hydrolysed serobactin B sample.

B. MS-MS spectra of the peak of retention time 5 min 18 s, marked with a red circle in the chromatogram.

Fig. S7. Identification of D and L configurations in the residues of Serobactin.

A. HPLC chromatogram of serobactin A hydrolysed with 2M HI and derivatized with Marfey's reagent.

B. Similar to A, but the hydrolysis was performed with 6M HCl. Identification of peaks is shown in Table S1.

Fig. S8. Photoreactivity of the Fe^{3+}-serobactin complex.

A. HPLC chromatogram of Fe^{3+}-serobactin A complex.

B. ESI-MS in negative mode of the purified Fe^{3+}-serobactin A complex.

C. UV-Vis spectra of Fe^{3+}-serobactin A before (filled line) and after 3 h of continuous irradiation with a 200 W mercury arc lamp (dot line).