Article

Revised structure of anthelvencin A and characterization of the anthelvencin biosynthetic gene cluster

Céline Aubry, Paolo Clerici, Claude Gerbaud, Laurent Micouin, Jean-Luc Pernodet, and Sylvie Lautru

ACS Chem. Biol., **Just Accepted Manuscript** • DOI: 10.1021/acschembio.9b00960 • Publication Date (Web): 04 Mar 2020

Revised structure of anthelvencin A and characterization of the anthelvencin biosynthetic gene cluster

Céline Aubry[a, b], Paolo Clerici[c], Claude Gerbaud[a], Laurent Micouin[c], Jean-Luc Pernodet[a], and Sylvie Lautru[a*]

[a] Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France

[b] present address: Unité Molécules de Communication et Adaptation des Microorganismes (MCAM, UMR 7245), Muséum National d'Histoire Naturelle (MNHN), Centre National de la Recherche Scientifique (CNRS), CP 54 57 rue Cuvier, 75005 Paris, France

[c] Université de Paris, Laboratoire de Chimie et Biochimie pharmacologiques et toxicologiques, LCBPT, UMR 8601 CNRS, F-75006 Paris, France

[*] Corresponding author: Sylvie LAUTRU, sylvie.lautru@i2bc.paris-saclay.fr

ABSTRACT

Anthelvencins A and B are pyrrolamide metabolites produced by *Streptomyces venezuelae* ATCC 14583 and 14585. Isolated in 1965, they were reported to exhibit anthelmintic and moderate antibacterial activities. In this study, we revise the structure of anthelvencin A and identify a third anthelvencin metabolite, bearing two N-methylated pyrrole groups, which we named anthelvencin C. We sequenced the genome of *S. venezuelae* ATCC 14583 and identified a gene cluster predicted to direct the biosynthesis of anthelvencins. Functional analysis of this gene cluster confirmed its involvement in anthelvencin biosynthesis and allowed us to propose a biosynthetic pathway for anthelvencins. In addition to a non-ribosomal peptide synthetase (NRPS), the assembly of anthelvencins involves an enzyme from the ATP-grasp ligase family, Ant23. We propose that Ant23 uses a PCP-loaded 4-aminopyrrole-2-carboxylate as substrate. As observed for the biosynthesis of the other pyrrolamides congocidine (produced by *Streptomyces ambofaciens* ATCC 25877) and distamycin (produced by *Streptomyces netropsis* DSM 40846), the NRPS assembling anthelvencins is composed of stand-alone domains only. Such NRPSs, sometimes called type II NRPSs, are less studied than the classical multimodular NRPSs. Yet, they constitute an interesting model to study protein-protein interactions in NRPSs and are good candidates for combinatorial biosynthesis approaches.

Anthelvencins A and B (Figure 1A) are specialized metabolites that were isolated in 1965 from cultures of *Streptomyces venezuelae* ATCC 14583 and 14585 and exhibit moderate antibacterial and anthelmintic activities[1]. They belong to the family of pyrrolamide metabolites, the best-characterized members of which are congocidine (also called netropsin, produced by *Streptomyces ambofaciens* ATCC 25877) and distamycin (produced by *Streptomyces netropsis* DSM 40846) (see Supplementary Figure S1). These metabolites are known to bind to the DNA minor groove with some sequence specificity for regions of four (or more) A or T bases[2]. During the last decade, the biosynthetic gene clusters of congocidine and distamycin have been identified and the biosynthesis of these metabolites has been elucidated[3–7]. One remarkable aspect of this biosynthesis is that it involves non-canonical non-ribosomal peptide synthetases (NRPSs), solely constituted of stand-alone modules or domains. Such NRPSs are sometimes called type II NRPSs. They participate in the biosynthesis of unusual precursors of non-ribosomal peptides (NRP)[8] and are involved in the assembly of some NRP such as phosphinothricin[9] and rhabdopeptide/xenortide peptides[10].

A structural analysis of anthelvencins shows that these metabolites most likely share two precursors with congocidine and distamycin: 4-acetamidopyrrole-2-carboxylate (**5**) and 3-aminopropionamidine (**6**). The remaining precursor is probably 5-amino-3,4-dihydro-2H-pyrrole-2-carboxylate (**4**), a precursor shared with other pyrrolamides such as kikumycins[11] or TAN 868A[12] (Figure S1). In fact, a limited number of precursors seems to be assembled in various combinations to give birth to the different members of the pyrrolamide family. Understanding how these precursors are assembled and combined may improve our comprehension of type II NRPS enzymatic mechanisms and help to design functional synthetic NRPSs using synthetic biology. For these reasons, we undertook to identify and characterize the anthelvencin biosynthetic gene cluster of *S. venezuelae* ATCC 14583. In this study, we revise the structure of anthelvencin A based on HR-MS[2] and NMR data. We show that *S. venezuelae* ATCC 14583 produces, in addition to the already known anthelvencins A and B, a third anthelvencin, methylated on the two pyrrole groups. We named this new anthelvencin

anthelvencin C. We also identify the gene cluster directing the biosynthesis of anthelvencins in *S. venezuelae* ATCC 14583 genome and we functionally characterize this gene cluster.

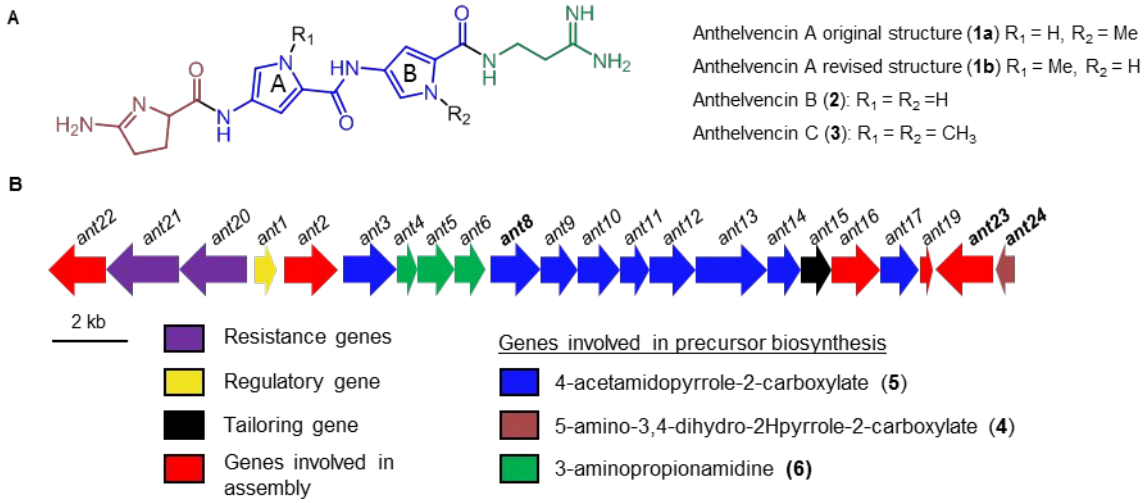


Figure 1. (A) Chemical structure of anthelvencins A, B and C; (B) Genetic organization of the anthelvencin biosynthetic gene cluster in *S. venezuelae* ATCC 14583. Genes in boldface are genes that were used as targets for genetic inactivation. The numbers in boldface (**1a**-**6**) are used to refer to the corresponding chemical structures.

RESULTS AND DISCUSSION

In silico identification of a gene cluster putatively involved in anthelvencin biosynthesis in* S. venezuelae *ATCC 14583. To identify the gene cluster directing anthelvencin biosynthesis, we sequenced the genome of the *S. venezuelae* ATCC 14583 strain by the Illumina technology, using a paired-end genomic library. The 5.45 million reads of 301 bps were assembled using Velvet v1.2.10, resulting in 63 contigs with a total length of 9.08 Mbps (180-fold coverage).

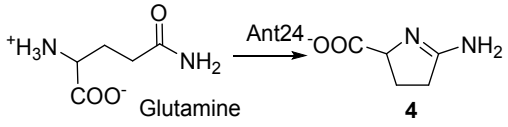
The gene cluster directing the biosynthesis of anthelvencins was identified by mining the genome of *S. venezuelae* ATCC 14583 for homologs of genes involved in the biosynthesis of congocidine[3]. We identified a gene cluster (*ant,* GenBank accession number MK483114) that spans 26 kb and contains 22 genes (Figure 1B). The deduced products of twenty of these genes exhibit a high amino acid sequence identity with Cgc proteins involved in congocidine biosynthesis (from 64 to 84 % sequence identity, Table 1) and they most likely have similar

Table 1. Sequence identities between Ant and Cgc proteins

Protein	Putative protein function	Cgc ortholog (Accession number)	% sequence identity
Ant1	Transcriptional regulator	Cgc1 (AKZ59702.1)	71
Ant2	NRPS, C domain	Cgc2 (AKZ59703.1)	66
Ant3	4-Acetamidopyrrole-2-carboxaldehyde dehydrogenase	Cgc3 (AKZ59704.1)	74
Ant4	Cytosine monophosphate hydrolase	Cgc4 (AKZ59705.1)	83
Ant5	Cytosine reductase	Cgc5 (AKZ59706.1)	77
Ant6	Dihydrocytosine hydrolase	Cgc6 (AKZ59707.1)	78
Ant8	Nucleotidyl N-acetylglucosamine dehydrogenase	Cgc8 (AKZ59709.1)	84
Ant9	Nucleotidyl-2-acetamido-2-deoxyglucopyranuronate decarboxylase	Cgc9 (AKZ59710.1)	84
Ant10	Glycosyltransferase-like enzyme	Cgc10 (AKZ59711.1)	81
Ant11	N-acetylglucosamine-1-phosphate nucleotidyltransferase	Cgc11 (AKZ59712.1)	76
Ant12	Nucleotidyl threo-2-acetamido-2-deoxy-pentopyran-4-ulose aminotransferase	Cgc12 (AKZ59713.1)	79
Ant13	Glycoside hydrolase	Cgc13 (AKZ59714.1)	78
Ant14	4-Acetamidopyrrole-2-carboxylate deacetylase	Cgc14 (AKZ59715.1)	80
Ant15	Methyltransferase	Cgc15 (AKZ59716.1)	84
Ant16	NRPS, C domain	Cgc16 (AKZ59717.1)	68
Ant17	4-Acetamidopyrrole-2-carboxaldehyde dehydrogenase	Cgc17 (AKZ59718.1)	83
Ant19	NRPS, PCP domain	Cgc19 (AKZ59720.1)	64
Ant20	ABC transporter	Cgc20 (AKZ59701.1)	81
Ant21	ABC transporter	Cgc21 (AKZ59700.1)	81
Ant22	Acyl co-A synthetase	Cgc22 (AKZ59699.1)	72
Ant23	ATP-grasp domain-containing protein	None	
Ant24	Ectoine synthase-like protein	None	

function to their Cgc homologs. Thus, the gene numbers attributed to the *ant* genes were chosen to follow the *cgc* nomenclature whenever possible. The genetic organization of the *ant*

cluster is remarkably similar to the one of the *cgc* cluster[3] from *S. ambofaciens* ATCC23877 (Figure S2). Two *cgc* genes (*cgc7* and *cgc18*) have no homologs in the *ant* gene cluster. These two genes are involved in the biosynthesis of the guanidinoacetate precursor (absent in anthelvencins) and its assembly. Based on the high level of sequence identity between Ant and Cgc proteins and on the previously published biosynthetic pathways for congocidine and distamycin[3,4,6,7], we propose the following roles for the Ant proteins. Ant8 to Ant13, Ant3 and Ant17 are involved in the biosynthesis of 4-acetamidopyrrole-2-carboxylate (**5**) and Ant14 is involved in its deacetylation. Ant4 to Ant6 are involved in the biosynthesis of 3-aminopropionamidine (**6**). Ant2, Ant16, Ant19, and Ant22 are involved in the assembly of anthelvencins. In addition to these, proteins are necessary for the biosynthesis and assembly of 5-amino-3,4-dihydro-2H-pyrrole-2-carboxylate (4). The *ant* cluster contains two genes, *ant24* and *ant23*, which have no homolog in the *cgc* cluster. We hypothesized that they might be involved in the biosynthesis of (4) and its assembly. Indeed, a protein blast[13] and a conserved domain searches[14] on the Ant24 sequence suggested that Ant24 belongs to the L-ectoine synthase (EC 4.2.1.108) family of enzymes. L-ectoine synthases catalyze the ring closure of Nγ-acetyl-L-2,4-diaminobutyric acid, yielding the osmolyte ectoine, a metabolite structurally related to (4). In 2011, Witt and collaborators reported that the ectoine synthase from *Halomonas elongata* can catalyze the intramolecular condensation of glutamine to form (**4**) as a side reaction[15]. This enzyme has 34% sequence identity and 51% sequence similarity with Ant24. Thus, it appears likely that Ant24 catalyzes the same reaction (Scheme 1).



Scheme 1: Proposed biosynthesis of 5-amino-3,4-dihydro-2H-pyrrole-2-carboxylate (**4**) by Ant24

Ant23 contains an ATP-grasp domain. ATP-grasp enzymes usually catalyze the ATP-dependent ligation of a carboxylate-containing molecule to an amino or thiol group-containing molecule[16]. Some of these ATP-grasp enzymes play a role in the biosynthesis of specialized metabolites[17]. They can function as an alternative to or in combination with non-ribosomal peptide synthetases (NRPS), to elongate a peptide chain[17,18]. Thus, it appears plausible that Ant23 catalyzes the amide bond formation between (**4**) and a PCP (Ant19)-bound 4-aminopyrrole-2-carboxylate.

Abolition of the production of four metabolites in a* S. venezuelae *ATCC 14583 mutant deleted for* ant8 *(coding for a putative nucleotidyl N-acetylglucosamine dehydrogenase). To verify that the *ant* gene cluster is involved in the biosynthesis of anthelvencins, we inactivated *ant8*. This gene is the ortholog of *cgc8,* the product of which is involved in the biosynthesis of the 4-acetamidopyrrole-2-carboxylate (**5**). The *ant8* gene was replaced by an *aac*(3)IV resistance cassette by homologous recombination using the pANT007 suicide plasmid, yielding the *S. venezuelae* ANT007 strain. This strain and the wild type *S. venezuelae* strain were cultivated for three days in MP5 liquid medium. The culture supernatants were then filtered and analyzed by HPLC. The chromatograms show that four metabolites present in the wild type strain supernatant (peaks I to IV) are absent in the supernatant of the ANT007 mutant strain (Figure 2). The first metabolite (peak I, retention time of 11.5 min) corresponds to 4-aminopyrrole-2-carboxylate (**5**), identified by its UV spectrum and by comparison with an authentic standard (Figure 2 and reference 7). The three peaks II (retention time of 13.3 min), III (retention time of 14.3 min) and IV (retention time of 15.5 min) have UV absorption spectra typical of pyrrolamides (Figure S3 and reference 4).

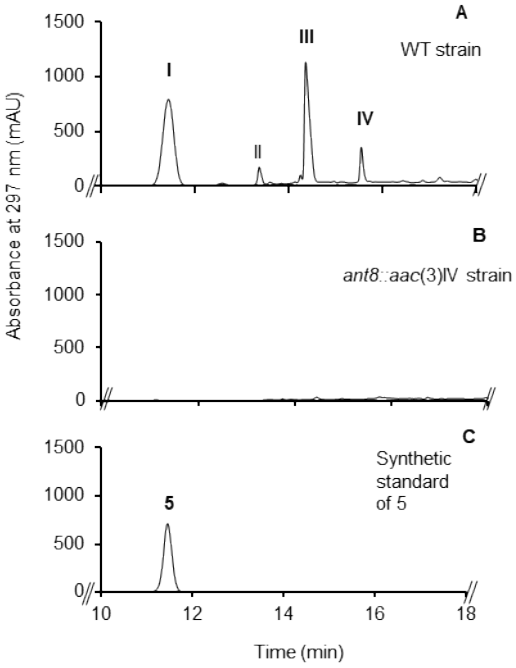


Figure 2: HPLC analysis of culture supernatants of A) *S. venezuelae* ATCC 14583 wild type and B) ANT007 (*S. venezuelae* ATCC 14583 *ant8::aac*(3)IV). C) Standard of 4-acetamidopyrrole-2-carboxylate (**5**). Peaks I to IV are compounds the biosynthesis of which is linked to the *ant* gene cluster. Peak I corresponds to (**5**).

Chemical nature of metabolites II, III and IV. To determine the chemical nature of the metabolites II, III and IV, we partially purified them. For that purpose, we used ANT012, a strain that expressed a second copy of the genes *ant23* and *ant24* under the control of the promoter *rpsL*(TP)[19], as this strain produces compounds III and IV in slightly higher titers (data not shown). The ANT012 culture supernatant was recovered after three days of culture in MP5 medium and the compounds of interest were partially purified on a XAD16 resin. The elution fraction was concentrated to dryness solution, resuspended in water and analyzed by LC-HR-MS².

The exact mass and fragmentation pattern of compound II (Figure S4) are consistent with II being anthelvencin B (2) ($[M+H]^+$ m/z = 414.1998; calculated 414.1997). The exact mass

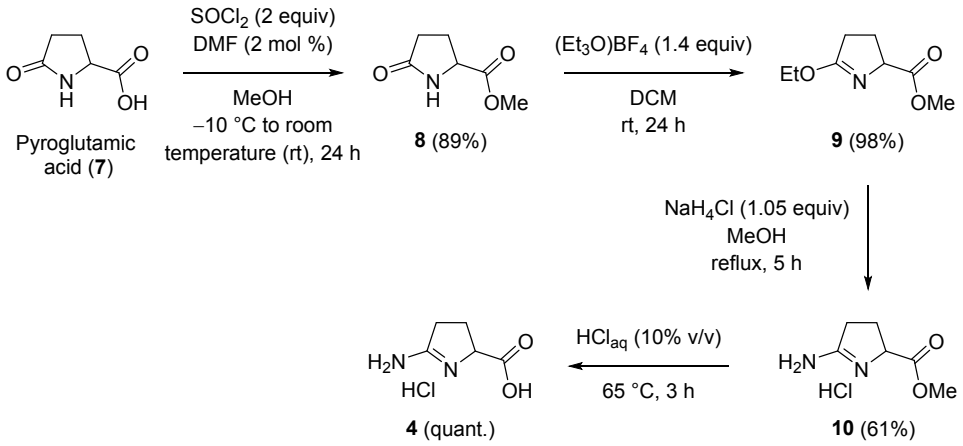
of compound III (Figure S5) is consistent with III being anthelvencin A ($[M+H]^+$ m/z = 428.2151; calculated 428.2153). The fragmentation pattern however (Figure S5, fragments 4 and 5), indicates that the position of the methyl group is not on the B pyrrole ring, as previously proposed (but never experimentally established[1]) but rather on the A pyrrole ring (Figure 1). To confirm the structure of anthelvencin A, we purified compound III and carried out NMR experiments. ^{1}H NMR of anthelvencin A hydrochloride proved to be very similar to the one reported for the authentic sample synthesized by Lee and coworkers[17] assuming the position of the methyl group on the B pyrrole ring. However, a combination of two-dimensional NMR experiments confirmed the position of this methyl on the A ring. We therefore propose a revised structure for anthelvencin A (Figure S6).

The exact mass and fragmentation pattern of compound IV (Figure S7) are consistent with IV being an anthelvencin metabolite methylated on both pyrrole groups ($[M+H]^+$ m/z = 442.2311; calculated 442.2310), a metabolite that we named anthelvencin C (Figure 1A, 3). We tried to purify anthelvencin C to confirm its chemical structure with NMR analyses but this metabolite turned out to be highly unstable, as already observed by Lee and coworkers[20].

Involvement of* ant24 *in 5-amino-3,4-dihydro-2H-pyrrole-2-carboxylate (4) biosynthesis. To verify that *ant24* is involved in the biosynthesis of anthelvencins, we replaced it by an *aac*(3)IV resistance cassette by homologous recombination, following the same procedure as described above. The culture supernatant of the resulting mutant strain, called ANT009, was analyzed by HPLC. No production of anthelvencins was observed, confirming that *ant24* is necessary for the production of these metabolites (Figure 3A). To ensure that the observed phenotype was due to the replacement of *ant24* by the *aac*(3)IV cassette, we genetically complemented the ANT009 strain using a plasmid expressing *ant23* and *ant24* under a constitutive promoter. The production of anthelvencins was restored in the complemented

strain, named ANT014 (Figure S8A), thus confirming that *ant24* is involved in anthelvencin biosynthesis.

To examine Ant24 putative function in the biosynthesis of (**4**), we undertook the chemical complementation of ANT009 by adding to the culture medium (**4**), the putative product of Ant24-catalyzed reaction. Thus, (**4**) was synthetized according to a previously described synthetic procedure[21] (Scheme 2).



Scheme 2: Synthesis of 5-amino-3,4-dihydro-2Hpyrrole-2-carboxylate (**4**)

We next fed the ANT009 strain with (**4**)**.** As shown in Figure 3B**,** this resulted in the restoration of the production of anthelvencins A and C, hence confirming the involvement of *ant24* in the biosynthesis of the anthelvencin precursor (**4**).

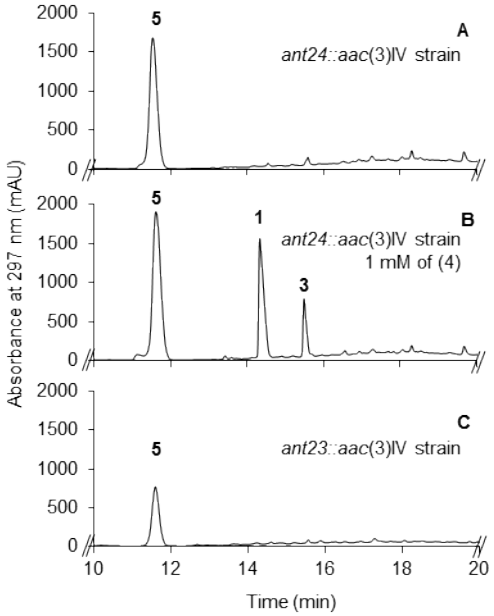


Figure 3: HPLC analysis of culture supernatants of (A) ANT009 (*S. venezuelae* ATCC 14583 *ant24*::*aac*(3)IV), (B) ANT009 (*S. venezuelae* ATCC 14583 *ant24*::*aac*(3)IV) cultivated in presence of 1mM of (**4**), and (C) ANT008 (*S. venezuelae* ATCC 14583 *ant23*::*aac*(3)IV). Numbers above peaks correspond to the metabolite numbers in the text.

Involvement of* ant23 *in the biosynthesis of anthelvencins. To confirm that Ant23 is involved in anthelvencin biosynthesis, we replaced *ant23* by the *aac*(3)IV resistance cassette following the previously described protocol. The resulting mutant strain was called ANT008. It was cultivated for three days in MP5 medium at 28°C and the culture supernatant was analyzed by HPLC. Figure 3C shows that no anthelvencin is produced by the ANT008 mutant. To ensure that the observed phenotype was due to the replacement of *ant23* by the *aac*(3)IV cassette, we genetically complemented the ANT008 strain using the plasmid expressing *ant23* and *ant24* under a constitutive promoter. The production of anthelvencins was restored in the complemented strain, named ANT013 (Figure S8B), thus confirming that *ant23* is involved in anthelvencin biosynthesis.

Proposed biosynthetic pathway for anthelvencin biosynthesis. Based on the results presented above and on previous characterizations of pyrrolamide biosyntheses[3,4,6], we

proposed that anthelvencins are assembled from 3-amidinopropionamidine, 4-aminopyrrole-2-carboxylate and 5-amino-3,4-dihydro-2H-pyrrole-2-carboxylate following the biosynthetic pathway presented in Figure 4. As already observed for the biosynthesis of other pyrrolamides (congocidine, distamycin), the non-ribosomal peptide synthetase involved in anthelvencins is constituted solely of stand-alone domains (C domains Ant2 and Ant16 and PCP domain Ant19). No adenylation domain is involved in the activation of the carboxylate groups of the precursors. Instead, activation of the carboxylate group of the pyrrole precursor (**5**) and the covalent attachment of the activated precursor to the PCP domain Ant19 is catalyzed by Ant22, which belongs to the family of acyl-CoA synthetases. The formation of the first amide bond between (**4**) and Ant19-bound (**5**) is likely catalyzed by Ant23, an enzyme from the ATP-grasp ligase family, which form acylphosphate intermediates. Two stand-alone condensation domains, Ant16 and Ant2, catalyze the formation of the other amide bonds, adding respectively a second pyrrole precursor and the 3-aminopropionamidine to Ant19-bound intermediates. Thus, during anthelvencin assembly, the stand-alone PCP domain, Ant19, interacts with two different families of enzymes (condensation domain and ATP-grasp ligase).

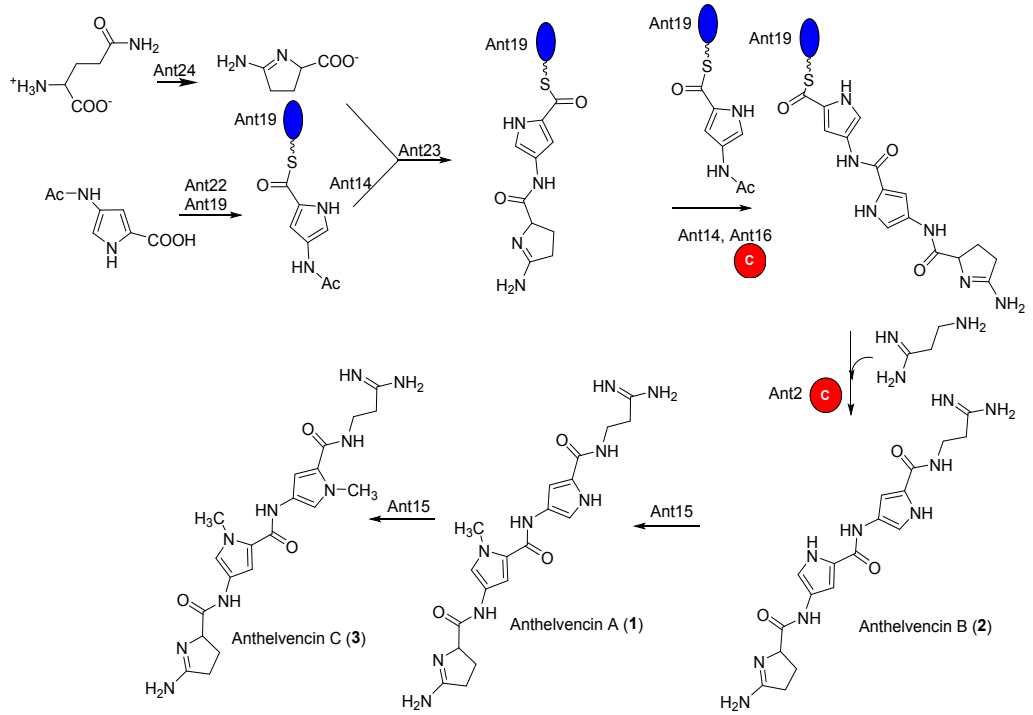


Figure 4: Proposed biosynthetic pathway for anthelvencins A, B and C

The blue ellipses represent PCP domains, and the red circles represent C domains

In conclusion, we have identified and characterized the gene cluster directing the biosynthesis of anthelvencins in *Streptomyces venezuelae* ATCC 14583. We showed that this cluster directs the biosynthesis of two known metabolites, anthelvencin A, for which we propose a revised structure, anthelvencin B, and new anthelvencin that we named anthelvencin C. As already observed for the assembly of other pyrrolamides (congocidine, distamycin), the assembly of anthelvencins involves a non-canonical NRPS constituted of stand-alone domains only. Such structural organization, rarely encountered in NRPS, appears to be characteristic of pyrrolamide NRPSs. Characterizing biosynthetic pathways involving type II NRPSs is of interest for the production of novel specialized metabolites by synthetic biology as the architecture of these NRPSs is particularly suited for combinatorial biosynthesis approaches.

MATERIALS AND METHODS

Bacterial strains, plasmids and growth conditions. Strains and plasmids used in this study are listed in Table S1 and S2. *Escherichia coli* strains were grown at 37 °C in LB or SOB complemented with $MgSO_4$ (20 mM final), supplemented with appropriate antibiotics as needed[22]. The Soya Flour Mannitol (SFM) medium[23] was used for genetic manipulations of *Streptomyces* strains and spore stocks preparations at 28°C. *Streptomyces* strains were grown for three days at 28°C in MP5[24] for anthelvencins (1-3) production.

DNA preparation and manipulations. All oligonucleotides used in this study were purchased from Eurofins and are listed in Table S3. The High fidelity DNA polymerase Phusion (Thermo Fisher Scientific) was used to amplify the DNA fragments for the construction of the suicide plasmids. DreamTaq polymerase (Thermo Fisher Scientific) was used for PCR verification of plasmids and of the replacement of the targeted genes by the resistance cassette *aac*(3)IV.

DNA fragments were purified from agarose gels using the Nucleospin Gel and PCR clean-up kit from Macherey-Nagel. *E. coli* transformations and *E. coli*/*Streptomyces* conjugations were performed according to standard procedures[25,23].

S. venezuelae *ATCC 14583 sequencing and assembly.* Total DNA of *S. venezuelae* ATCC 14583 was extracted following the salting out procedure[23] except that the mycelium was ground in liquid nitrogen. A paired-end library of the whole genome was constructed and sequenced at the high throughput sequencing core facility of I2BC with a MiSeq M01342 instrument (Illumina), generating 5.45 million 301 bp reads that were assembled using Velvet v1.2.10. The GenBank accession number of the anthelvencin gene cluster is MK483114.

Construction of the replacement mutants. The suicide plasmid pANT007 was constructed to replace the *ant8* gene by an *aac*(3)IV resistance cassette in *S. venezuelae*. This vector was constructed by assembling the three following inserts in the backbone pOSV400 (bearing a hygromycin resistance gene)[26]: a 1.8 kb fragment upstream of *ant*8, the resistance cassette *aac*(3)IV and a 2.0 kb DNA fragment downstream of *ant8*. The 1.8 kb and 2.0 kb DNA fragments from *S. venezuelae* ATCC 14583 were amplified by PCR with the primers CEA001/CEA002 and CEA003/CEA004 respectively. The PCR products were purified and ligated into pCR® Blunt, yielding pANT001 (containing the upstream DNA fragment) and pANT002 (containing the downstream DNA fragment). Both plasmids were verified by sequencing (Genewiz). The *aac*(3)IV resistance cassette was obtained by digestion of pW60[27] by *Hin*dIII. The 1.8 kb *Hin*dIII-*Xho*I fragment from pANT001, the 1.0 kb *Hin*dIII *aac*(3)IV fragment, and the 2.0 kb *Hin*dIII-*Spe*I fragment from pANT002 were then ligated into the *Xho*I-*Spe*I-digested pOSV400, yielding pANT007. The pANT007 plasmid was verified by restriction digestion and was introduced into *S. venezuelae* ATCC 14583 by intergeneric conjugation from the *E. coli* ET12567/pUZ8002/pANT007 strain. Double-recombinant mutants were selected on SFM plates with 50 µg/mL apramycin and screened for hygromycin sensitivity. DNA from the resulting strain, ANT007, was isolated and analyzed by PCR using the primers

A5, A6, and CEA013-CEA016 (See Figure S8). The same protocol was used for the construction of the ANT008 mutant strain (replacement of *ant*23), using the plasmid pANT008, and the ANT009 mutant strain (replacement of *ant*24), using the plasmid pANT009 (see Tables S2 and S3 for plasmid names and for primer sequences).

***Construction of the ANT012 strain overexpressing* Ant23 *and* Ant24.** The DNA region containing *ant23-ant24* was amplified by PCR from *S. venezuelae* ATCC 14583 genomic DNA using the primers CEA034/CEA035. The PCR product was purified and cloned into pCR® Blunt, yielding pANT011. The sequence of the insert was confirmed by sequencing. The 2.0 kb *Nhe*I/*Afl*II fragment from pANT011 was ligated into the *Spe*I/*Afl*II-digested pCEA005[28]. The obtained plasmid, containing *ant23* and *ant24* under the rpsL(TP) promoter, was named pANT012 and confirmed by restriction digestion. This plasmid was introduced into *S. venezuelae* ATCC 14583 by intergeneric conjugation. The correct integration of pANT012 was verified on the extracted DNA by PCR using the primers CEA_vec_seq14 and CEA_vec_seq15 and the strain was named ANT012.

Genetic complementation of ANT008. As the ANT008 strain bears the *aac*(3)IV resistance marker, the previously constructed pANT012 plasmid, bearing an *aac*(3)IV resistance gene cassette, could not be used for the genetic complementation of the strain. Thus, the 2.4 kb *Nsi*I/*Afl*II DNA fragment of pANT012 containing *ant23* and *ant24* under the control of the *rpsL*(TP) promoter was ligated into the *Nsi*I/*Afl*II-digested pOSV806[28]. The resulting plasmid was named pANT013 and was introduced into ANT008 by intergeneric conjugation. The strain obtained was named ANT013.

Chemical synthesis of 5-amino-3,4-dihydro-2H-pyrrole-2-carboxylate (4). Compound (4) was prepared according to a previously described synthetic procedure[29] (Scheme 2). Commercially available DL-pyroglutamic acid (7) was converted into the corresponding methyl ester (8) by treatment with thionyl chloride (2 equiv.), and N,N-dimethylformamide (DMF) (2

mol %) in methanol. Derivative (8) was then submitted to a reaction with triethyloxonium tetrafluoroborate (Meerwein's salt, 1.4 equiv.) in dichloromethane (DCM) to form carboximidate (9) in quantitative yield. This compound subsequently reacted with ammonium chloride (1.05 equiv.) in refluxing methanol to provide product (10) in 61% yield. Hydrolysis of the ester moiety of compound (10) finally afforded the desired acid (4) in a quantitative yield. A detailed synthesis protocol is available in the Supporting Information.

Chemical complementation of ANT009. *S. venezuelae* ANT009 strain was cultivated in 50 mL of MP5. After 24 h, the cultures were separated in two 25 mL cultures, and 1 mM of (4) (final concentration) was added to one of the cultures. After a total of 72 h of cultivation, culture supernatants were analyzed by HPLC as described below.

HPLC analysis of culture supernatants. *S. venezuelae* ATCC 14583 and its derivatives were cultivated in MP5 medium for three days at 28°C. The supernatants were filtered using Mini-UniPrep syringeless filter devices (0.2 µm, Whatman). The samples were analyzed on an Atlantis C_{18} T3 column (250 mm x 4.6 mm, 5 µm, column temperature 28°C) using an Agilent 1200 HPLC instrument with a quaternary pump. Samples were eluted in isocratic conditions with 0.1% v/v HCOOH in H_20 (solvent A)/ 0.1% v/v HCOOH in CH_3CN (solvent B) (95:5) at 1 $ml.min^{-1}$ for 7 min, followed by a gradient to 40:60 A/B over 23 min. Anthelvencins were detected by monitoring absorbance at 297 nm.

LC-HR-MS-MS analyses. The resuspended elution fraction obtained above was analyzed by LC-HR-MS^2. The analysis was performed using a Dionex Ultimate 3000 HPLC system coupled with a Maxis II™ QTOF mass spectrometer (Bruker, MA, USA) fitted with an electrospray ionization (ESI) source.

Chromatographic analysis was performed using a C_{18} AcclaimTM RSLC PolarAdvantage II (2.1 x 100 mm, 2.2 µm pore size) column (Thermo Scientific, MA, USA).

Column temperature was set at 40 °C and 2 µL of each sample was injected via an autosampler cooled to 4 °C. A flow rate of 0.3 mL/min was used and the eluent was introduced directly into the MS for ion detection. Elution was conducted with a mobile phase consisting of 0.1% v/v HCOOH in H_20 (solvent A) and 0.1% v/v HCOOH in CH_3CN (solvent B) following the gradient elution profile: 0 min, 5% solvent B; 2 min, 5% solvent B; 9 min, 50% solvent B; 15 min 90% solvent B; 17 min 90% solvent B; 19 min 5% solvent B; 21 min 5% solvent B. In the first half minute of each run, a sodium formate solution was injected directly as an internal reference for calibration. The acquisition parameters of the ESI source were set up as follows: electrospray voltage for the ESI source: 3500V, nebulising gas (N2) pressure: 35 psi, drying gas (N_2) flow: 8 L/min, and drying temperature: 200°C. Mass spectra were recorded over the m/z range 100-1300 at a frequency of 2 Hz, in positive ion mode. For MS/MS analysis, the cycle time was of 3 sec. Mass spectra were recorded over the m/z range 100-1300 at a frequency of 2 Hz, in positive ion mode. Selected parent ion at m/z 442.23 was fragmented at a fixed collision energy value of 40 eV and an isolation window of 0.5 amu.

Accession codes. The GenBank accession number of the anthelvencin gene cluster is MK483114. Its MiBig[30] accession number is BGC0002042.

Supporting Information Available: The protocol for the synthesis of 5-amino-3,4-dihydro-2H-pyrrole-2-carboxylate (4), the strains, plasmids and oligonucleotides used in this study and seven supplementary figures are available in the Supporting Information. This material is available free of charge via the internet at http://pubs.acs.org

Acknowledgements

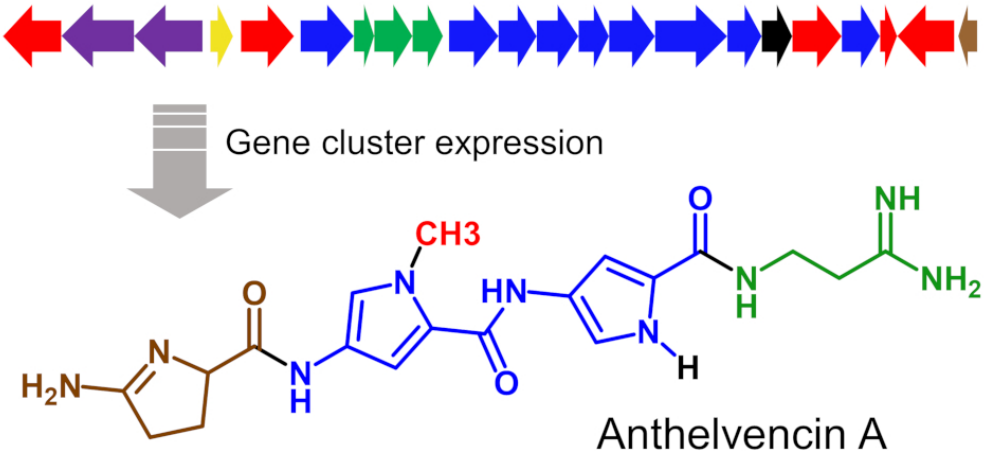
We acknowledge the High-throughput sequencing facility of I2BC for its sequencing and bioinformatics expertise (Centre de Recherche de Gif – http://www.i2bc.paris-saclay.fr/). We thank Z. Hong, Y. Li, S. Prado and D. Touboul for their help with the LC-HRMS[2] analysis.

The research received funding from ANR-14-CE16-0003-01. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

References

(1) Probst, G. W., Hoehn, M. M., and Woods, B. L. (1965) Anthelvencins, new antibiotics with anthelmintic properties. *Antimicrob. Agents Chemother. 5*, 789–795.
(2) Neidle, S. (2001) DNA minor-groove recognition by small molecules. *Nat. Prod. Rep. 18*, 291–309.
(3) Juguet, M., Lautru, S., Francou, F.-X., Nezbedová, S., Leblond, P., Gondry, M., and Pernodet, J.-L. (2009) An iterative nonribosomal peptide synthetase assembles the pyrrole-amide antibiotic congocidine in *Streptomyces ambofaciens*. *Chem. Biol. 16*, 421–431.
(4) Vingadassalon, A., Lorieux, F., Juguet, M., Le Goff, G., Gerbaud, C., Pernodet, J.-L., and Lautru, S. (2015) Natural combinatorial biosynthesis involving two clusters for the synthesis of three pyrrolamides in *Streptomyces netropsis*. *ACS Chem. Biol. 10*, 601–610.
(5) Hao, C., Huang, S., Deng, Z., Zhao, C., and Yu, Y. (2014) Mining of the pyrrolamide antibiotics analogs in *Streptomyces netropsis* reveals the amidohydrolase-dependent "iterative strategy" underlying the pyrrole polymerization. *PloS One 9*, e99077.
(6) Al-Mestarihi, A. H., Garzan, A., Kim, J. M., and Garneau-Tsodikova, S. (2015) Enzymatic Evidence for a revised congocidine biosynthetic pathway. *Chembiochem Eur. J. Chem. Biol. 16*, 1307–1313.
(7) Lautru, S., Song, L., Demange, L., Lombès, T., Galons, H., Challis, G. L., and Pernodet, J.-L. (2012) A sweet origin for the key congocidine precursor 4-acetamidopyrrole-2-carboxylate. *Angew. Chem. Int. Ed Engl. 51*, 7454–7458.
(8) Jaremko, M. J., Davis, T. D., Corpuz, J. C., and Burkart, M. D. (2020) Type II non-ribosomal peptide synthetase proteins: structure, mechanism, and protein-protein interactions. *Nat. Prod. Rep.* https://doi.org/10.1039/C9NP00047J
(9) Schinko, E., Schad, K., Eys, S., Keller, U., and Wohlleben, W. (2009) Phosphinothricin-tripeptide biosynthesis: an original version of bacterial secondary metabolism? *Phytochemistry 70*, 1787–1800.
(10) Cai, X., Nowak, S., Wesche, F., Bischoff, I., Kaiser, M., Fürst, R., and Bode, H. B. (2017) Entomopathogenic bacteria use multiple mechanisms for bioactive peptide library design. *Nat. Chem. 9*, 379–386.
(11) Takaishi, T., Sugawara, Y., and Suzuki, M. (1972) Structure of Kikumycin A and B. *Tetrahedron Lett. 13*, 1873–1876.
(12) Takizawa, M., Tsubotani, S., Tanida, S., Harada, S., and Hasegawa, T. (1987) A new pyrrole-amidine antibiotic TAN-868 A. *J. Antibiot. (Tokyo) 40*, 1220–1230.
(13) Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol. 215*, 403–410.
(14) Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., Geer, L. Y., and Bryant, S. H. (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res. 45*, D200–D203.
(15) Witt, E. M. H. J., Davies, N. W., and Galinski, E. A. (2011) Unexpected property of ectoine synthase and its application for synthesis of the engineered compatible solute ADPC. *Appl. Microbiol. Biotechnol. 91*, 113–122.

(16) Galperin, M. Y., and Koonin, E. V. (1997) A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. *Protein Sci. Publ. Protein Soc. 6*, 2639–2643.
(17) Goswami, A., and Van Lanen, S. G. (2015) Enzymatic strategies and biocatalysts for amide bond formation: tricks of the trade outside of the ribosome. *Mol. Biosyst. 11*, 338–353.
(18) Hollenhorst, M. A., Clardy, J., and Walsh, C. T. (2009) The ATP-dependent amide ligases DdaG and DdaF assemble the fumaramoyl-dipeptide scaffold of the dapdiamide antibiotics. *Biochemistry 48*, 10467–10472.
(19) Shao, Z., Rao, G., Li, C., Abil, Z., Luo, Y., and Zhao, H. (2013) Refactoring the Silent Spectinabilin Gene Cluster Using a Plug-and-Play Scaffold. *ACS Synth. Biol. 2*, 662–669.
(20) Lee, M., Coulter, D. M., and Lown, J. W. (1988) Total synthesis and absolute configuration of the antibiotic oligopeptide (4S)-(+)-anthelvencin A and its 4R-(-) enantiomer. *J. Org. Chem. 53*, 1855–1859.
(21) Lee, M., and Lown, J. W. (1987) Synthesis of (4S)- and (4R)-methyl 2-amino-1-pyrroline-5-carboxylate and their application to the preparation of (4S)-(+)- and (4R)-(-)-dihydrokikumycin B. *J. Org. Chem. 52*, 5717–5721.
(22) Sambrook, J., and Russell, D. (2000) Molecular Cloning: A Laboratory Manual 3rd Revised edition. Cold Spring Harbor Laboratory Press,U.S.
(23) Kieser, T., Bibb, M., Buttner, M., and Hopwood, D. A. (2000) Practical *Streptomyces* Genetics, John Innes Foundation, Norwich NR47UH, UK.
(24) Pernodet, J. L., Alegre, M. T., Blondelet-Rouault, M. H., and Guérineau, M. (1993) Resistance to spiramycin in *Streptomyces ambofaciens*, the producer organism, involves at least two different mechanisms. *J. Gen. Microbiol. 139*, 1003–1011.
(25) Sambrook, J., and Russell, D. W. (2001) Molecular cloning: A Laboratory Manual, Third edition. CSHL Press, Cold Spring Harbor, NY.
(26) Boubakri, H., Seghezzi, N., Duchateau, M., Gominet, M., Kofroňová, O., Benada, O., Mazodier, P., and Pernodet, J.-L. (2015) The Absence of pupylation (prokaryotic ubiquitin-like protein modification) affects morphological and physiological differentiation in *Streptomyces coelicolor*. *J. Bacteriol. 197*, 3388–3399.
(27) Corre, C., Song, L., O'Rourke, S., Chater, K. F., and Challis, G. L. (2008) 2-Alkyl-4-hydroxymethylfuran-3-carboxylic acids, antibiotic production inducers discovered by *Streptomyces coelicolor* genome mining. *Proc. Natl. Acad. Sci. U. S. A. 105*, 17510–17515.
(28) Aubry, C., Pernodet, J.-L., and Lautru, S. (2019) Modular and Integrative Vectors for Synthetic Biology Applications in *Streptomyces* spp. *Appl. Environ. Microbiol. 85*.
(29) Lee, M., and Lown, J. W. (1987) Synthesis of (4S)- and (4R)-methyl 2-amino-1-pyrroline-5-carboxylate and their application to the preparation of (4S)-(+)- and (4R)-(-)-dihydrokikumycin B. *J. Org. Chem. 52*, 5717–5721.
(30) Kautsar, S. A., Blin, K., Shaw, S., Navarro-Muñoz, J. C., Terlouw, B. R., van der Hooft, J. J. J., van Santen, J. A., Tracanna, V., Suarez Duran, H. G., Pascal Andreu, V., Selem-Mojica, N., Alanjary, M., Robinson, S. L., Lund, G., Epstein, S. C., Sisto, A. C., Charkoudian, L. K., Collemare, J., Linington, R. G., Weber, T., and Medema, M. H. (2020) MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res. 48*, D454–D458.



80x39mm (300 x 300 DPI)