Chemistry & Biology
Article

Cell PRESS

An Iterative Nonribosomal Peptide Synthetase Assembles the Pyrrole-Amide Antibiotic Congocidine in *Streptomyces ambofaciens*

Maud Juguet,[1,5] Sylvie Lautru,[1,5] François-Xavier Francou,[1] Šárka Nezbedová,[1,2] Pierre Leblond,[3] Muriel Gondry,[4] and Jean-Luc Pernodet[1,*]
[1]Institut de Génétique et Microbiologie, University Paris Sud 11, CNRS, UMR 8621, 91405 Orsay, France
[2]Institute of Microbiology, Academy of Sciences of the Czech Republic, Vídeňská 1083, 142 20 Praha 4-Krč, Czech Republic
[3]Laboratoire de Génétique et Microbiologie, UMR, INRA 1128, IFR 110, Faculté des Sciences et Techniques, Université Henri Poincaré Nancy 1, BP239, 54506 Vandœuvre-lès-Nancy, France
[4]Commissariat à l'Energie Atomique, Institut de Biologie et de Technologies de Saclay, Service d'Ingénierie Moléculaire des Protéines (SIMOPRO), 91191 Gif-sur-Yvette, France
[5]These authors contributed equally to this work
*Correspondence: jean-luc.pernodet@igmors.u-psud.fr
DOI 10.1016/j.chembiol.2009.03.010

SUMMARY

Congocidine (netropsin) is a pyrrole-amide (oligopyrrole, oligopeptide) antibiotic produced by *Streptomyces ambofaciens*. We have identified, in the right terminal region of the *S. ambofaciens* chromosome, the gene cluster that directs congocidine biosynthesis. Heterologous expression of the cluster and in-frame deletions of 8 of the 22 genes confirm the involvement of this cluster in congocidine biosynthesis. Nine genes can be assigned specific functions in regulation, resistance, or congocidine assembly. In contrast, the biosynthetic origin of the precursors cannot be easily inferred from in silico analyses. Congocidine is assembled by a nonribosomal peptide synthetase (NRPS) constituted of a free-standing module and several single-domain proteins encoded by four genes. The iterative use of its unique adenylation domain, the utilization of guanidinoacetyl-CoA as a substrate by a condensation domain, and the control of 4-aminopyrrole-2-carboxylate polymerization constitute the most original features of this NRPS.

INTRODUCTION

Pyrrole-amides, sometimes also called oligopyrroles or oligopeptides, are a family of natural products containing one or several pyrrole-2-carboxamides. They are synthesized by *Streptomyces* and related actinobacteria and have a variety of biological activities including antiviral, antibiotic, antitumor, and antihelmintic activities. The two best-known members of the pyrrole-amide family are congocidine (also called netropsin) and distamycin (Figure 1). Isolated about 50 years ago, they have been extensively studied due to their ability to bind to the minor groove of the DNA double helix in a sequence-specific manner and have served as models for the study of DNA minor-groove binders. In particular, it has been shown that they bind DNA reversibly through hydrogen bonds, van der Waals contacts, and electrostatic interactions at sequences of four or more consecutive A-T pairs and strongly discriminate against G-C pairs (Neidle, 2001).

Molecules capable of binding specific DNA sequences, such as congocidine and distamycin, are attracting considerable interest because they could potentially be used to regulate the expression of particular genes involved in various diseases such as cancer. Consequently, many analogs and hybrids of congocidine and distamycin have been synthesized chemically in attempts to modify or increase the sequence specificity of the DNA binding, or to create molecules capable of binding to DNA covalently and possessing potential antitumor activities (Neidle, 2001). Congocidine and distamycin are not among the most efficient minor-groove DNA-binder anticancer agents, as they do not bind DNA covalently. However, it has been suggested that a recently discovered member of the pyrrole-amide family, pyrronamycin B, binds DNA in a covalent manner and exhibits antitumor activities (Asai et al., 2000). Moreover, congocidine and distamycin have been used as sequence-selective carriers of alkylating agents. Thus, for example, tallimustine, a benzoic acid mustard derivative of distamycin, has been in phase I and II clinical trials as an anticancer agent (Denny, 2001). Congocidine and distamycin analogs have also been coupled with known antitumor agents (or their active moiety) such as anthramycin (specific for GC-rich regions); the resulting hybrid molecules display antitumor activity (Baraldi et al., 2003).

Although the nature of the interactions of distamycin and congocidine with the minor groove of DNA has been extensively studied, little is known about the biosynthesis of these molecules: incorporation experiments with isotopically labeled precursors have shown that the guanidinium moiety of congocidine originates from arginine (Wildfeuer, 1964), and genes that confer resistance to congocidine (netropsin) have been characterized in *Streptomyces flavopersicus* NRRL 2820 (synonym: *Streptomyces netropsis*) (Stumpp et al., 2005). However, no gene cluster directing the biosynthesis of a pyrrole-amide has been isolated. *Streptomyces ambofaciens*, the industrial producer of spiramycin, has been known since 1952 to produce congocidine

 ©2009 Elsevier Ltd All rights reserved

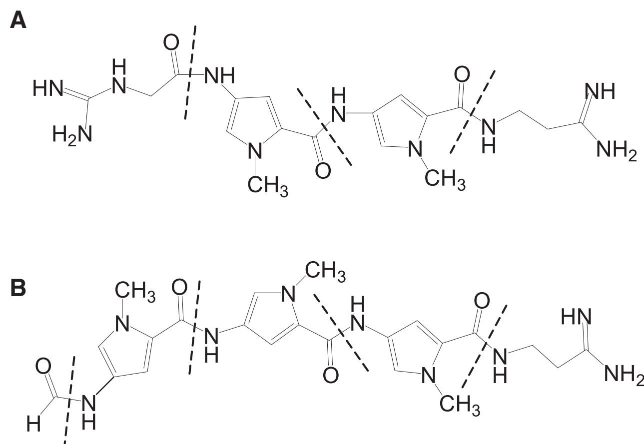

Figure 1. Chemical Structures of Congocidine and Distamycin
(A) Congocidine.
(B) Distamycin.
The dashed lines separate the putative precursors.

(Cosar et al., 1952). Recently, the extremities of the linear chromosome of this bacterium have been sequenced (Choulet et al., 2006). As gene clusters involved in secondary metabolites are mostly found in the arms of the *Streptomyces* chromosome, we wondered whether the gene cluster directing the biosynthesis of congocidine in *S. ambofaciens* was located in the sequenced extremities. We report here the identification, delimitation, and analysis of the congocidine biosynthetic gene cluster from *S. ambofaciens*. We show that congocidine is assembled by an atypical nonribosomal peptide synthetase (NRPS), constituted of a free-standing module and several single-domain proteins. This enzyme contains a unique adenylation domain that acts iteratively and a condensation domain that uses CoA-activated rather than peptidyl carrier protein (PCP)-phosphopantetheinyl (Ppant)-activated guanidinoacetate as a substrate. To our knowledge, this work constitutes the first characterization of a pyrrole-amide biosynthetic gene cluster and should facilitate the isolation and study of biosynthetic pathways of other known or new pyrrole-amides.

RESULTS AND DISCUSSION

In Silico Identification of a Candidate Congocidine Biosynthetic Gene Cluster in the *S. ambofaciens* Chromosome

About 1.3 Mb of each of the extremities of the *S. ambofaciens* ATCC23877 chromosome has been sequenced (Choulet et al., 2006). Sequence analysis of these regions reveals the presence of at least nine putative secondary metabolite gene clusters, including gene clusters directing the biosynthesis of alpomycin (Pang et al., 2004) and coelichelin (Barona-Gómez et al., 2006). One of these putative secondary metabolite gene clusters, located in the right arm of the chromosome, was identified by a clear break of synteny in a region otherwise syntenic with the chromosome of *Streptomyces coelicolor*, a species that does not synthesize congocidine. Between the regions of synteny, 24 genes (SAMR0898 to SAMR0921), encoding products with no or low similarity to *S. coelicolor* proteins, occupy a region of about 31 kb. Four of these genes code for nonribosomal peptide synthetases, enzymes typically involved in secondary metabolism; two other genes, SAMR0920 and SAMR0919, code for proteins homologous to NetP1 and NetP2 from *S. flavopersicus* (Stumpp et al., 2005). NetP1 and NetP2 confer resistance to netropsin (congocidine) on this bacterium. These various observations suggested that the gene cluster identified may direct the biosynthesis of congocidine.

Involvement of the SAMR0898–09021 Gene Cluster in Congocidine Biosynthesis

To verify the involvement of the SAMR0898–0921 gene cluster in congocidine biosynthesis, we inactivated SAMR0901 (encoding an NRPS enzyme). SAMR0901 was disrupted by homologous recombination between the *S. ambofaciens* SPM110 chromosome and a suicide plasmid (pJMC101; see Table S1 available online) containing an internal fragment of SAMR0901. The antibacterial activity of the mutant strain, CGCA004, was then tested. Patches of *S. ambofaciens* SPM110 (in which production of spiramycin is abolished) and CGCA004 were grown on HT, a medium on which this species produces congocidine. After 4 days, plates were overlaid with soft nutrient agar containing the indicator microorganisms *Escherichia coli* or *Bacillus subtilis*. Growth inhibition was observed around SPM110, but not around CGCA004, suggesting that production of congocidine was abolished (data not shown). Both SPM110 and CGCA004 were also grown for 4 days in liquid MP5 medium, and culture supernatants were analyzed by HPLC. A peak around 14.9 min observed in the chromatogram of SPM110 culture supernatant was absent from the chromatogram of CGCA004 (Figures 2A and 2B). Comparison with authentic standard and LC-MS-MS analysis confirmed that this peak corresponds to congocidine (data not shown). Thus, SAMR0901 inactivation abolishes congocidine production and confirms the involvement of this gene in congocidine biosynthesis. As genes directing the biosynthesis of a given secondary metabolite are usually clustered, we hypothesized that the 24 nearby genes (SAMR0898–0921) constituted the congocidine biosynthetic gene cluster.

Heterologous Expression of the SAMR0898–0921 Gene Cluster

To determine whether the SAMR0898–0921 gene cluster is sufficient to direct congocidine biosynthesis in *Streptomyces*, we introduced the entire gene cluster into heterologous hosts.

For sequencing purposes, a library of *S. ambofaciens* ATCC23877 genomic DNA has been constructed in pBeloBAC11 (Choulet et al., 2006) and the ends of the inserts of about 5000 bacterial artificial chromosomes (BACs) from this library have been sequenced. Using these sequences, we identified one BAC, the insert of which contains all 24 genes of the putative cluster. The insert of this BAC, which we named pCGC001, includes the 43 kb region of the *S. ambofaciens* chromosome located between SAMR0888 and SAMR0926.

For conjugational transfer of pCGC001 and its integration into the *Streptomyces* chromosome, a DNA fragment containing the hygromycin resistance gene, the ΦC31 integrase gene, the corresponding *attP* sequence, and *oriT* (the transfer origin of the RK2 plasmid) was inserted into this BAC. The resulting BAC

 ©2009 Elsevier Ltd All rights reserved

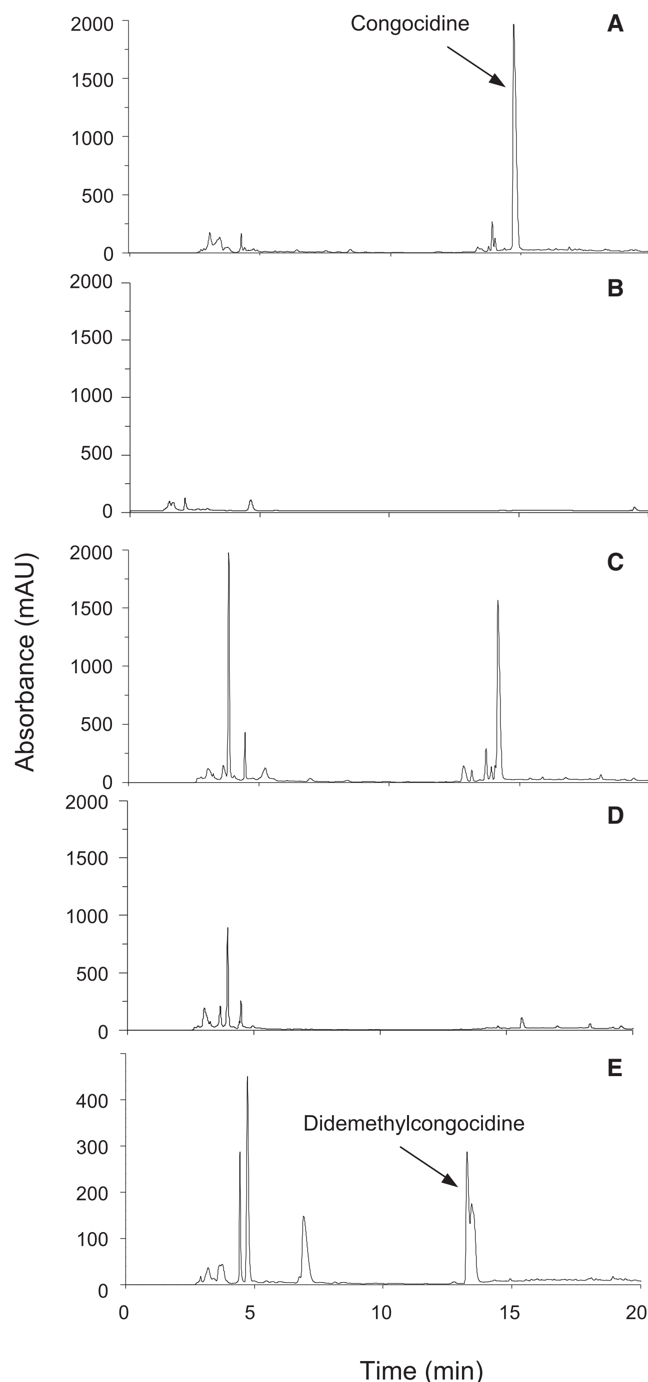

Figure 2. HPLC Analyses of Congocidine Production
(A) SPM110 (*S. ambofaciens* strain producing congocidine).
(B) CGCA004 (*S. ambofaciens* mutant strain with disruption of SAMR0901).
(C) CGCL006 (*S. lividans* TK23 harboring pCGC002 with the complete congocidine cluster).
(D) CGCL017 (*S. lividans* TK23 harboring pCGC213 [in-frame deletion of *cgc10* in the congocidine cluster]).
(E) CGCL031 (*S. lividans* TK23 harboring pCGC223 [in-frame deletion of *cgc15* in the congocidine cluster]).
Absorbance was monitored at 297 nm.

was named pCGC002 and introduced into the heterologous host *Streptomyces lividans* TK23, yielding the strain CGCL006. This strain shows antibacterial activity against *E. coli* and *B. subtilis* whereas the parental *S. lividans* TK23 does not, which suggests that it produces congocidine. This was confirmed by HPLC (Figure 2C) and LC-MS analyses of culture supernatants. Furthermore, CGCL006 produces a quantity of congocidine comparable to that produced by *S. ambofaciens* SPM110. Thus, it constitutes a convenient tool for studying the congocidine biosynthetic gene cluster.

Three other heterologous hosts (*S. coelicolor* M145, *Streptomyces fungicidicus* B-5477, and *Streptomyces albus*) were also tested. In all cases, integration of pCGC002 into their chromosome led to the production of congocidine, a molecule that they do not naturally produce (Figure S1). This suggests that the SAMR0888–0926 region of the *S. ambofaciens* chromosome might be sufficient for congocidine production in *Streptomyces*.

Identification of the Limits of the Congocidine Gene Cluster

To determine the ends of the congocidine biosynthetic gene cluster, genes upstream from SAMR0898 and downstream from SAMR0921 were deleted in two steps from the insert of pCGC002 (Figure S2). First, the DNA between SAMR0898 and pBeloBAC11 was replaced with the excisable apramycin cassette *att3aac* (similar to cassettes described by Raynal et al., 2006) by PCR targeting. The resulting BAC was named pCGC016. The *att3aac* cassette was then excised using the plasmid pOSInt3 to give a BAC named pCGC017 (Table S1). The same procedure was applied to delete the DNA between SAMR0921 and pBeloBAC11 from pCGC017, yielding pCGC018 (containing the apramycin cassette) and subsequently pCGC019 (cassette excised) (Table S1). Thus, pCGC019 contains only the 24 genes of the putative cluster. The BAC pCGC019 was introduced into *S. lividans* TK23 by conjugation. The resulting mutant strain was named CGCL010. HPLC analysis detected congocidine in CGCL010 culture supernatants, suggesting that the 24 genes SAMR0898–0921 are sufficient for congocidine production in *Streptomyces* (Figure S3A).

Genes located at each extremity of the putative cluster were deleted to map its ends more precisely. SAMR0921 was deleted in-frame from pCGC002 using the PCR targeting technology. The resulting BAC, pCGC221, was introduced into *S. lividans* TK23 to give CGCL030. HPLC analysis of CGCL030 culture supernatant showed that congocidine production is abolished by the deletion of SAMR0921 (Figure S3B). Therefore, this gene is essential for congocidine production and defines one of the extremities of the congocidine gene cluster.

SAMR0898 was similarly deleted in-frame from the other end of the cluster in pCGC019, yielding pCGC029. This BAC was introduced into *S. lividans* TK23 to give CGCL026. Congocidine was detected in CGCL026 culture supernatant (Figure S3C), indicating that this gene is not essential for congocidine production. SAMR0899 was then deleted in-frame from pCGC019, and the resulting BAC, pCGC030, was introduced into *S. lividans* TK23 to give CGCL027. CGCL027 culture supernatant contained congocidine, showing that SAMR0899 is not required for congocidine biosynthesis (Figure S3D). However, deletion of SAMR0900 using the same procedure yielded *S. lividans* TK23 strain CGCL028 that

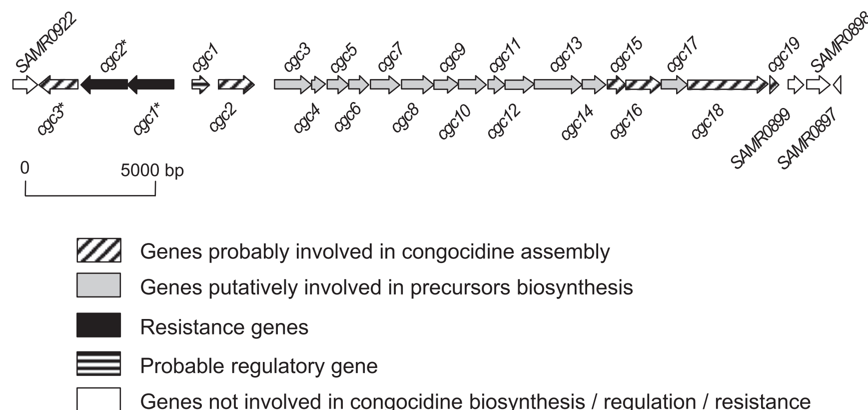

Figure 3. Genetic Organization of the Congocidine (*cgc*) Gene Cluster

does not produce any congocidine (Figure S3E). SAMR0900 is therefore essential for congocidine biosynthesis and delimits the other extremity of the congocidine gene cluster.

Overview of the Congocidine Biosynthetic Gene Cluster

We named the 22 genes constituting the congocidine biosynthetic gene cluster *cgc1** to *cgc3** (SAMR0919 to SAMR0921) and *cgc1* to *cgc19* (SAMR0918 to SAMR0900) (Figure 3). Table 1 presents an overview of the predicted function of each of the *cgc* gene products, and a possible role for these products in congocidine biosynthesis. Of these 22 *cgc* genes, two (*cgc1** and *cgc2**) are predicted to confer resistance to congocidine and one (*cgc1*) probably codes for a regulator of congocidine biosynthesis. The remaining 19 genes code for putative enzymes, with the possible exception of *cgc7*: the product of this gene does not share significant similarities with any known or hypothetical proteins present in databases. Eight of the 19 genes (*cgc3**, *cgc2*, *cgc7*, *cgc10*, *cgc15*, *cgc16*, *cgc18*, and *cgc19*) were each individually deleted in-frame in the BAC pCGC002 containing the complete cluster. None of the *S. lividans* TK23 strains harboring the deleted BACs produced congocidine (Figures 2D and 2E; Figures S3B and S3E; data not shown). In some cases (e.g., Figure S3E), new peaks were observed in the HPLC analyses of the culture supernatants of these strains. It is not known at the moment whether they correspond to biosynthetic intermediates or to shunt metabolites (except for the *cgc15* deletion mutant; see below). We propose that the 19 *cgc* genes not involved in congocidine resistance or regulation of biosynthesis direct either the biosynthesis of congocidine precursors (13 genes) or the assembly of the molecule (6 genes) (Figure 3).

Assembly of the Congocidine Molecule

A retrobiosynthetic analysis suggests that congocidine is assembled from three precursors: guanidinoacetate, 3-aminopropionamidine, and 4-aminopyrrole-2-carboxylate or 4-amino-1-methylpyrrole-2-carboxylate, depending on the timing of N-methylation of the pyrrole groups (Figure 1A). Moreover, amide bond formation between these putative precursors may be catalyzed by an NRPS, as previously suggested (Stumpp et al., 2005). NRPSs are large modular multienzymatic complexes responsible for the biosynthesis of most nonribosomal peptides. In canonical NRPSs, each module catalyzes the incorporation of one amino acid and consists of domains responsible for a particular stage of the reaction: adenylation (A) domains activate and transfer amino acids to the phosphopantetheinyl arm of peptidyl carrier protein (PCP) domains; condensation (C) domains catalyze the elongation of the peptidyl chain; and the hydrolysis of the chain from the enzyme is generally catalyzed by a thioesterase (Te) domain located at the end of the last module. Sequence analysis indicates the presence of four NRPS genes (*cgc2*, *cgc16*, *cgc18*, and *cgc19*) in the *cgc* cluster, and in-frame deletions of these genes confirmed their involvement in congocidine biosynthesis (Figure S3E; data not shown). The four NRPS genes, however, code for discrete modules or domains, not for large multimodular enzymes as usually observed. Thus, *cgc18* encodes a typical NRPS module containing an A domain, a PCP domain, and a C domain. In addition to *cgc18*, two genes—*cgc2* and *cgc16*—encode two stand-alone C domains and one gene—*cgc19*—a stand-alone PCP domain.

The assembly of congocidine from its putative precursors requires the formation of three amide bonds. This necessitates the activation of the carboxylate groups of two 4-aminopyrrole-2-carboxylate molecules and one guanidinoacetate molecule. However, other than the A domain of Cgc18 (Cgc18_A), only one enzyme encoded within the *cgc* gene cluster is likely to catalyze the activation of one of these groups. This enzyme, Cgc3*, belongs to the AMP-binding superfamily of proteins and resembles acyl-CoA synthetases. One of Cgc18_A and Cgc3* probably acts iteratively to activate the two pyrrole precursors, the other enzyme activating guanidinoacetate. The amino acids identified by established models (Challis et al., 2000; Stachelhaus et al., 1999) as likely to confer substrate specificity in Cgc18_A ($S_{235}V_{236}E_{239}Q_{278}V_{299}G_{301}E_{322}V_{330}S_{331}K_{517}$, GrsA-PheA numbering) do not allow the probable substrate for this domain to be predicted by comparison with databases (NRPSpredictor; Rausch et al., 2005). In particular, they differ from the amino acids ($D_{235}F_{236}H_{239}F_{278}I_{299}T_{301}H_{322}D_{330}A_{331}K_{517}$) conferring specificity on the CyrB_A domain, which has been proposed to activate guanidinoacetate in cylindrospermopsin biosynthesis (Mihali et al., 2008). No A domain is known to activate pyrrole substrates. Indeed, it has been suggested that activation of pyrrole-2-carboxylate may be hampered by the weak acidity of the carboxylate (Walsh et al., 2006). It is therefore possible that the molecule activated is not 4-aminopyrrole-2-carboxylate, but a precursor of this molecule (see discussion concerning pyrrole biosynthesis below). From this analysis, it appears to be difficult to predict the substrate of either Cgc3* or Cgc18_A. Nonetheless, if the two PCP domains are functional, they are likely to be loaded by the sole adenylation domain encoded in the *cgc* cluster, Cgc18_A. In this hypothesis, the same molecule is loaded onto the two PCP domains. As there are two pyrrole moieties but one guanidinoacetate moiety in congocidine, we propose that Cgc18_A activates two molecules of 4-aminopyrrole-2-carboxylate (or its precursor) and transfer the activated groups to the

 ©2009 Elsevier Ltd All rights reserved

Cell
PRESS

PCP domain of Cgc18 and to Cgc19 (Figure 4). Guanidinoacetate is therefore likely to be activated by Cgc3*, under the form of guanidinoacetyl-CoA.

In this model, Cgc18_A is proposed to act iteratively to load Cgc18_PCP in *cis* and Cgc19 in *trans*. Iterative use and *trans*-adenylation activity of A domains have already been reported (for a review, see Haynes and Challis, 2007). In yersiniabactin biosynthesis, for instance, a single adenylation domain loads three PCP domains (one in *cis*, two in *trans*) (Keating et al., 2000a).

Based on the above analysis and on gene deletion studies, we propose the following model for congocidine assembly (Figure 4): the activated pyrrole precursors are transferred to Cgc18_PCP and Cgc19, and condensation of these two molecules is then catalyzed by the condensation domain of Cgc18 to give a bipyrrole molecule covalently attached to Cgc19. The two free-standing condensation domains—Cgc2 and Cgc16—catalyze then the sequential addition of the activated guanidinoacetyl-CoA and 3-aminopropionamidine, yielding a demethylated form of congocidine. It is unclear which enzyme is responsible for the addition of which precursor. N-methylation of the pyrrole groups by the SAM-dependent methyltransferase Cgc15 is then the final step of congocidine biosynthesis.

Order of the Condensation and N-Methylation Reactions

The present analysis of the congocidine biosynthetic gene cluster does not allow us to unambiguously determine the order in which the different condensations and the methylation of the pyrrole groups occur. This order could be different from the one presented in Figure 4: the condensation reaction with guanidinoacetate may occur before (Figure 4) or after the condensation reaction with 3-aminopropionamidine; the methylation of the pyrrole groups is not necessarily the last step of the biosynthesis. HPLC analysis of CGCL031 (deletion mutant of *cgc15*) culture supernatants indicates that abolition of congocidine production is accompanied by the accumulation of a product also absorbing at 297 nm (Figure 2). LC-MS and MS-MS analyses demonstrate that this product is N,N-didemethylcongocidine (Figure S4). As LC-MS-MS analysis of the *S. ambofaciens* SPM110 also shows the presence of a small amount of N-demethylcongocidine in addition to congocidine (data not shown), these findings strongly suggest that N-methylation of the pyrrole groups is the last step in congocidine biosynthesis.

Specific Features of the NRPS Involved in Congocidine Biosynthesis

Several features of the NRPS involved in congocidine biosynthesis are worth considering. First, the aspartate residue (D_{235}, GrsA-PheA numbering) conserved in most adenylation domains and involved in binding the α-amino group of amino acids is replaced by a serine in Cgc18_A. This may indicate that the substrate of Cgc18_A has no α-amino group and is thus consistent with 4-aminopyrrole-2-carboxylate being the substrate. Second, guanidinoacetyl-CoA is the most likely substrate for one of the two condensation domains, Cgc2 and Cgc16 (see above). This activation of a condensation domain substrate under the form of a CoA thioester rather than a PCP-Ppant-thioester is unusual. Other examples occur in the biosynthesis of lipopeptides such as CDA (Hojati et al., 2002) or fengycin (Steller et al., 1999). In these biosynthetic systems, the lipidic chain is activated as an acyl-CoA thioester that serves as substrate for the first NRPS condensation domain. The use of CoA-activated substrates illustrates the flexibility of NRPS enzymes. It also opens possibilities for molecular engineering, as the use of CoA rather than PCP-Ppant eliminates some protein/protein interactions. Third, both Cgc2 and Cgc16 lack the glycine residue in the C3 HHxxxDG motif generally conserved in condensation domains (Cgc2 motif: **S**HLAAD**L**; Cgc16 motif: HHLVSD**A**) (Schwarzer et al., 2003). Cgc2 is also missing the first histidine of the motif. No particular role has been proposed for this histidine, but substitution of the glycine residue with a leucine in the EntF C domain (enterobactin biosynthesis) reduced activity by at least 1000-fold (Roche and Walsh, 2003). In VibH (vibriobactin biosynthesis), however, the same mutation had no effect on catalytic efficiency (Keating et al., 2002). The absence of congocidine production in *cgc2* and *cgc16* deletion mutants (CGCL035 and CGCL032, respectively) demonstrates that Cgc2 and Cgc16 are active despite not having a glycine residue in the C3 motif.

The organization of congocidine NRPS domains is unusual: half of these domains are free-standing domains (Cgc19, PCP domain; Cgc2 and Cgc16, C domains), and the organization of the only complete module, a free-standing module (Cgc18, A-PCP-C), differs from the canonical C-A-PCP organization. Free-standing A and PCP domains or A-PCP didomains are frequently found when an amino acid is modified prior to its incorporation into a natural product, such as in the biosyntheses of vancomycin, aminocoumarins, or prodigiosin (Fischbach and Walsh, 2006). Free-standing C domains, however, are less frequent: few have been reported, and they include those involved in the biosynthesis of vibriobactin (VibH; Keating et al., 2000b), viomycin (VioM; Thomas et al., 2003), or of the enediyne C-1027 (SgcC5; Liu et al., 2002). In these three biosyntheses, the condensation reactions catalyzed by the free-standing C domains involve a PCP-tethered biosynthetic intermediate and a nucleophilic substrate free in solution, as proposed in congocidine biosynthesis for the condensation of guanidinoacetate. Although unusual, the organization of congocidine NRPS is particularly suited to the synthesis of oligopyrrole molecules. Indeed, an iteration of the condensation reaction catalyzed by Cgc18_C between the bipyrrole attached to Cgc19 and a third pyrrole residue attached to Cgc18_PCP would yield a tripyrrole attached to Cgc19. This mechanism is in fact plausible for the biosynthesis of distamycin, a molecule closely related to congocidine (Figure 1B) that comprises three 4-aminopyrrole-2-carbonyl groups. This raises the issue of how pyrrole polymerization is controlled by the enzyme, as it has been shown that C domains exhibit substrate specificity mainly at the acceptor, or aminoacyl, site (i.e., pyrrole site), whereas more relaxed specificity is observed at the donor, or peptidyl, site (i.e., bipyrrole site) (Lautru and Challis, 2004). One possibility is that in the case of Cgc18_C, there is strict substrate specificity at the donor site.

No gene encoding the phosphopantetheinyl transferase required for the posttranslational modification of the NRPS is present in the congocidine biosynthetic gene cluster. A phosphopantetheinyl transferase is encoded by *alpN* in the *alp* cluster, located in the terminal inverted repeats of the *S. ambofaciens* chromosome (Pang et al., 2004), but it is not known whether this enzyme is the one involved in the congocidine NRPS modification.

 ©2009 Elsevier Ltd All rights reserved

Table 1. Components of the Congocidine Biosynthetic Gene Cluster and Their Proposed Function

ORF			Best Blast Hit				
SAM Number	*cgc* Number	Product Size (aa)	GenBank Accession Number	Species	Identity/ similarity (%)	Putative Function	Proposed Role in Congocidine Biosynthesis
SAMR0922		321	CAB62672	*Streptomyces coelicolor* A3(2)	93/94	Pirin-like protein	Not involved
SAMR0921	*cgc3**	532	ABO65957	*Geobacillus thermodenitrificans* NG80-2	33/49	Acyl-CoA synthetase	Congocidine assembly
SAMR0920	*cgc2**	611	AAU04845	*Streptomyces netropsis*	75/84	ABC transporter, ATP-binding protein	Resistance
SAMR0919	*cgc1**	619	AAU04844	*Streptomyces netropsis*	74/83	ABC transporter, transmembrane protein	Resistance
SAMR0918	*cgc1*	223	AAU04843	*Streptomyces netropsis*	61/71	Two-component response regulator	Regulation
SAMR0917	*cgc2*	470	ABF87402	*Myxococcus xanthus* DK 1622	35/48	NRPS, C domain	Congocidine assembly
SAMR0916	*cgc3*	485	ABQ67968	*Sphingomonas wittichii* RW1	45/63	Aldehyde dehydrogenase	Precursor biosynthesis
SAMR0915	*cgc4*	181	CAE16290	*Photorhabdus luminescens subsp. laumondii* TTO1	54/71	Nucleoside 2-deoxyribosyltransferase	Precursor biosynthesis
SAMR0914	*cgc5*	297	CAE16289	*Photorhabdus luminescens subsp. laumondii* TTO1	50/67	Dihydroorotate dehydrogenase	Precursor biosynthesis
SAMR0913	*cgc6*	276	CAE16288	*Photorhabdus luminescens subsp. laumondii* TTO1	55/72	Creatininase	Precursor biosynthesis
SAMR0912	*cgc7*	377	No significant alignment			Hypothetical protein	Precursor biosynthesis
SAMR0911	*cgc8*	436	ABV96412	*Salinispora arenicola* CNS-205	52/65	Nucleotide sugar dehydrogenase	Precursor biosynthesis
SAMR0910	*cgc9*	327	ABC45959	*Salinibacter ruber* DSM 13855	44/58	Nucleoside diphosphate sugar epimerase/ dehydrogenase	Precursor biosynthesis
SAMR0909	*cgc10*	360	EDT60921	*Cyanothece sp.* PCC 8801	35/51	Glycosyltransferase	Precursor biosynthesis
SAMR0908	*cgc11*	254	ABE52488	*Methanococcoides burtonii* DSM 6242	43/65	Sugar nucleotidyltransferase	Precursor biosynthesis
SAMR0907	*cgc12*	381	CAN94946	*Sorangium cellulosum* “So ce 56”	47/61	Nucleotide-sugar aminotransferase	Precursor biosynthesis
SAMR0906	*cgc13*	641	AAM23368	*Thermoanaerobacter tengcongensis* MB4	25/44	Glycoside hydrolase	Precursor biosynthesis
SAMR0905	*cgc14*	302	EAQ78058	*Blastopirellula marina* DSM 3645	31/48	Amidohydrolase	Precursor biosynthesis
SAMR0904	*cgc15*	268	CAJ36215	Uncultured methanogenic archaeon RC-I	32/46	Methyltransferase	Methylation of pyrrole groups
SAMR0903	*cgc16*	445	ZP_00110264	*Nostoc punctiforme* PCC 73102	37/55	NRPS, C domain	Congocidine assembly
SAMR0902	*cgc17*	348	EAT61433	*Geobacter sp.* FRC-32	38/50	Alcohol dehydrogenase	Precursor biosynthesis
SAMR0901	*cgc18*	1067	CAO90637	*Microcystis aeruginosa* PCC 7806	36/53	NRPS, A-PCP-C domains	Congocidine assembly
SAMR0900	*cgc19*	121	ABX06605	*Herpetosiphon aurantiacus* ATCC 23779	37/50	NRPS, PCP domain	Congocidine assembly

 ©2009 Elsevier Ltd All rights reserved

Table 1. *Continued*

ORF			Best Blast Hit				
SAM Number	*cgc* Number	Product Size (aa)	GenBank Accession Number	Species	Identity/ similarity (%)	Putative Function	Proposed Role in Congocidine Biosynthesis
SAMR0899		221	BAG18463	*Streptomyces griseus subsp. griseus* NBRC 13350	72/80	RNA polymerase ECF subfamily σ factor	Not involved
SAMR0898		217	BAG18464	*Streptomyces griseus subsp. griseus* NBRC 13350	64/74	Hypothetical protein	Not involved
SAMR0897		100	CAB65634	*Streptomyces coelicolor* A3(2)	71/79	Membrane protein	Not involved

NRPS, nonribosomal peptide synthetase; C, condensation; A, adenylation; PCP, peptidyl carrier protein. Percentages of identity/similarity refer to deduced amino acid sequence comparisons.

Origin of Congocidine Precursors

As indicated above, 3-aminopropionamidine, guanidinoacetate, and 4-aminopyrrole-2-carboxylate are the putative precursors of congocidine. To our knowledge, there is no known biosynthetic pathway for 3-aminopropionamidine, and the origin of this molecule cannot easily be inferred from the analysis of the *cgc* gene cluster.

The second probable precursor of congocidine is guanidinoacetate. It is a known precursor in the biosynthesis of cylindrospermopsin (Burgoyne et al., 2000), in which it is biosynthesized via transamidination of glycine by an amidinotransferase, CyrA, with arginine being the most likely donor of the guanidino group (Mihali et al., 2008). No amidinotransferase is encoded within the *cgc* gene cluster, and there are no other obvious candidate genes for the biosynthesis of guanidinoacetate among the *cgc* genes. Nonetheless, incorporation experiments with isotopically labeled precursors have shown that the guanidinium moiety of congocidine originates from arginine (Wildfeuer, 1964). SAML1205, a protein predicted to belong to the amidinotransferase family (PF02274), is encoded elsewhere in the genome of *S. ambofaciens*. Although genes directing the biosynthesis of secondary metabolites are usually clustered in the chromosome, this protein may catalyze the synthesis of guanidinoacetate. Were this the case, however, all the heterologous hosts tested

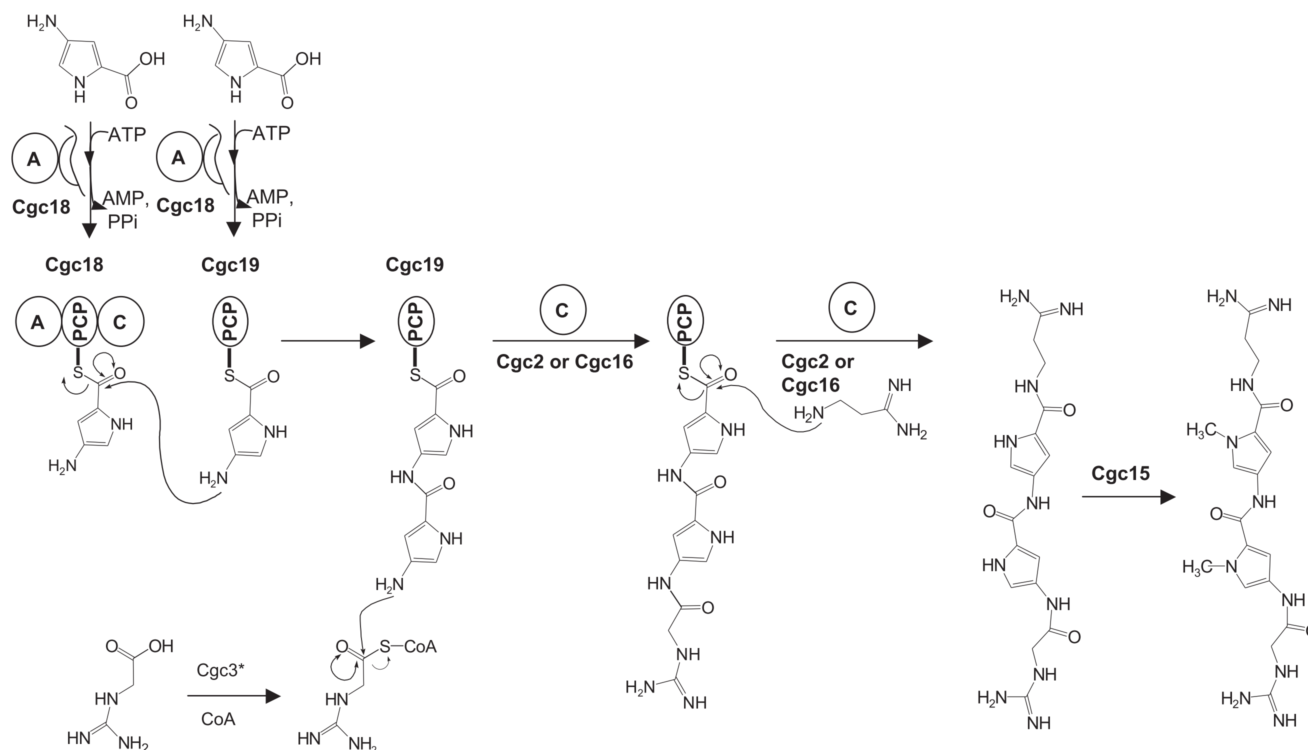

Figure 4. Proposed Pathway for Congocidine Assembly
A, adenylation domain; PCP, peptidyl carrier protein; C, condensation domain.

for producing congocidine (*S. lividans*, *S. coelicolor* M145, *S. fungicidicus* B-5477, and *S. albus*) would have to have a homolog of SAML1205 in their genomes. Such a homolog, SCO1222, is present in the *S. coelicolor* M145 genome and, as *S. coelicolor* and *S. lividans* are close relatives, it is likely that a homolog of SCO1022/SAML1205 is also present in the *S. lividans* genome.

The last identified precursor of congocidine is 4-aminopyrrole-2-carboxylate. Recently, Walsh and colleagues reviewed all known pathways for pyrrole biosynthesis, including pyrrole-2-carboxylate (Walsh et al., 2006). This molecule is a constituent of many secondary metabolites and is found in prodiginines, clorobiocin, and pyoluteorin, for example. In these metabolites, pyrrole-2-carboxylate is synthesized from proline: a proline-specific adenylation domain activates the amino acid and transfers it to a PCP domain; a flavin-dependent dehydrogenase then catalyzes the dehydrogenation of proline to yield pyrrolyl-2-carboxyl-S-PCP. A similar mechanism is plausible for congocidine biosynthesis, proline being activated by Cgc18_A and dehydrogenated as a prolyl thioester on Cgc18-PCP and Cgc19. However, no homolog of the flavin-dependent dehydrogenase is encoded within the *cgc* gene cluster. Moreover, amino acids in Cgc18_A predicted to confer substrate specificity differ from the amino acids known to confer proline specificity. Thus, it appears unlikely that 4-aminopyrrole-2-carboxylate is synthesized via this mechanism. No homolog of the enzymes required for other known biosyntheses of pyrrole molecules are encoded in the *cgc* gene cluster, so the biosynthetic pathway for 4-aminopyrrole-2-carboxylate in congocidine biosynthesis is probably novel.

Regulation and Resistance

Congocidine biosynthesis is probably controlled by *cgc1*, the only gene in the *cgc* cluster predicted to be involved in regulation. This gene codes for a putative transcriptional regulator. The predicted Cgc1 protein contains a C-terminal DNA-binding HTH domain and belongs to the two-component response regulator family. It is similar (61% identity, 71% similarity) to the product of the partial open reading frame (ORF1′) identified upstream from *netP1* in *S. flavopersicus* NRRL 2820 (Stumpp et al., 2005).

Two genes, *cgc1** and *cgc2**, code for transmembrane and ATP-binding proteins, respectively, constituting ABC transport systems. Cgc1* and Cgc2* are similar to NetP1 (74% identity, 83% similarity) and NetP2 (75% identity, 84% similarity), respectively, proteins that have been shown to confer resistance to netropsin (congocidine) in *S. flavopersicus* NRRL 2820 (Stumpp et al., 2005). It is therefore likely that Cgc1* and Cgc2* confer resistance to congocidine in *Streptomyces*. To test this, we inserted a 5.4 kb fragment containing *cgc1** and *cgc2** into pSET152, yielding pCGC502. This plasmid was introduced into *S. lividans* TK23 by protoplast transformation to give the strain CGCL001. CGCL001 grew on HT plates containing 20 μg/ml congocidine whereas the control did not (Figure S5), demonstrating that *cgc1** and *cgc2** are sufficient to confer resistance to congocidine.

Resistance mechanisms in antibiotic producer strains can involve several genes, some of which may not be clustered with the biosynthetic genes. To investigate whether genes located outside the congocidine cluster participate in congocidine resistance in *S. ambofaciens*, we used a strain (CGCA019) in which the chromosomal region encompassing *cgc3** to *cgc18* is deleted. The resistance of this mutant strain to congocidine was assessed as described above and compared to that of the parental SPM110 strain containing the complete *cgc* cluster (Figure S5). The CGCA019 strain is more sensitive than the SPM110 strain to congocidine and its sensitivity is similar to that of *S. lividans* TK23, suggesting that no other gene outside the cluster is involved in congocidine resistance.

Genes of Unknown Function in Congocidine Biosynthesis

Of the 22 genes constituting the congocidine biosynthetic gene cluster, only nine can be assigned a function, either in resistance to congocidine (*cgc1** and *cgc2**), biosynthesis regulation (*cgc1*), or congocidine assembly (*cgc3**, *cgc2*, *cgc15*, *cgc16*, *cgc18*, and *cgc19*). The role of the 13 other genes, most of which code for enzymes, in congocidine biosynthesis is unknown. It is likely that some of these genes are involved in the biosynthesis of the precursors of congocidine, as the origin of these precursors has not yet been identified. Nonetheless, the presence of six genes encoding enzymes similar to those involved in sugar biosynthesis and transfer (*cgc8* to *cgc13*) is puzzling, as congocidine does not contain any sugar moiety. One hypothesis is that a glycosylated form of congocidine is produced by *S. ambofaciens*, and that this putative glycosylated form has not yet been isolated. However, no such product has been detected in culture supernatants or in cell extracts of *S. ambofaciens* SPM110 (data not shown). The presence of a gene coding for a putative glycosyltransferase (Cgc10) and another coding for a putative glycoside hydrolase (Cgc13) might suggest the existence of a self-resistance mechanism involving inactivation of congocidine by intracellular glycosylation and extracellular reactivation by hydrolysis of the added sugar. This type of inactivation/reactivation mechanism by glycosylation/hydrolysis has been demonstrated in *Streptomyces antibioticus*, the producer of oleandomycin (Quirós et al., 1998). Nonetheless, the predicted intracellular localization of Cgc13 and the absence of congocidine production in CGCL017 (the *cgc10* deletion mutant; Figure 2D) do not support this model.

SIGNIFICANCE

We report here the isolation of the gene cluster directing the biosynthesis of the pyrrole-amide antibiotic congocidine from *S. ambofaciens*. A striking feature of this cluster is the absence of obvious candidate genes for the biosynthesis of the three putative congocidine precursors. This could have been anticipated for 3-aminopropionamidine, for which no biosynthetic pathway has yet been described. However, it is surprising that biosynthesis of the other two precursors, guanidinoacetate and 4-aminopyrrole-2-carboxylate, does not seem to involve any known pathway.

Congocidine is assembled by a nonribosomal peptide synthetase that displays some unusual features. Its single adenylation domain acts iteratively, and one of its condensation domains uses CoA- rather than PCP-Ppant-activated guanidinoacetate as a substrate. Moreover, it is constituted of a module and domains that are free standing. This organization and the proposed enzyme mechanism may also apply

 ©2009 Elsevier Ltd All rights reserved

to the synthesis of many oligo-pyrrole molecules, and especially of distamycin, which comprises three 4-aminopyrrole-2-carbonyl groups; however, it also raises the question of the control of pyrrole polymerization by the enzyme. These findings illustrate the versatility of the NRPS family of enzymes. Our work provides a basis for the elucidation of the molecular principles of the biosynthesis of antibiotics of the pyrrole-amide family and paves the way for the isolation and characterization of biosynthetic gene clusters of known or new pyrrole-amides. Once deciphered, the biosynthetic pathways of these molecules may well provide new tools for biological engineering and combinatorial biosynthesis.

EXPERIMENTAL PROCEDURES

Strains, Plasmids, and Culture Conditions

Strains and plasmids used are listed in Table S1. *E. coli* and *B. subtilis* strains were grown in LB or SOB media supplemented with 30 mM $MgSO_4$ and with appropriate antibiotics as necessary. HT or SFM media (Kieser et al., 2000) were used for genetic manipulations of *Streptomyces* strains and spore stock preparations. For congocidine production, *Streptomyces* strains were grown in MP5 medium (Pernodet et al., 1993).

Preparation and Manipulation of DNA

E. coli transformation and *E. coli*/*Streptomyces* conjugation were performed according to standard procedures (Sambrook and Russell, 2001; Kieser et al., 2000). The GC-rich PCR system (Roche) was used to amplify the *att3aac* and FRT*aac* cassettes used. Taq polymerase from QIAGEN was used for PCR for the verification of gene deletions in BACs or *Streptomyces* strains. DNA fragments were purified from agarose gels using the GFX-PCR and gel band purification kit from GE Healthcare. All oligonucleotides used in this study are listed in Table S2.

Construction of the *cgc18* Insertion Mutant

An internal fragment of *cgc18* was amplified by PCR from *S. ambofaciens* ATCC23877 genomic DNA using the primers Congo-D and Congo-R. The 815 bp product was purified and inserted into pGEM-T-Easy (Promega). The EcoRI/Klenow fragment from the resulting plasmid was ligated into EcoRV-digested pOSV202, yielding pJMC101. This plasmid was introduced into *S. ambofaciens* SPM110 by conjugation with *E. coli* S17.1. Transconjugants were selected on 50 μg/ml HT-apramycine, and integration of pJMC101 into the chromosome of the resulting strain, CGCA004, was verified by Southern blot.

Heterologous Expression of the *cgc* Gene Cluster in *S. lividans* TK23

The BAC pCGC001 was engineered to allow conjugational transfer from *E. coli* and integration into the *Streptomyces* chromosome. First, the *aac*(3)IV-oriT, BamHI/Klenow/XbaI fragment from pPAOI6 was replaced with the *hyg*-oriT, HindIII/Klenow/XbaI fragment from pOSV010 (Karray, 2005) to give pOSV201. Then, the 5.3 kb PvuII fragment containing Ω*hyg*, *oriT*, the ΦC31 integrase gene, and the *attP* site was purified and introduced into HpaI-digested pCGC001, yielding pCGC002. This BAC was introduced into *S. lividans* TK23 by conjugation with *E. coli* S17.1. Transconjugants were selected on 50 μg/ml HT-hygromycin, and the presence of Ω*hyg* was verified by PCR using the primers OMEGA A1 and OMEGA HYG2. The *S. lividans* TK23 strain containing pCGC002 integrated into its chromosome was named CGCL006.

Construction of Strains CGCL009, CGCL010, CGCL026, CGCL027, CGCL028, CGCL029, and CGCL035

PCR targeting (Yu et al., 2000) was used to delete the sequences between SAMR898 and pBeloBAC11, between *cgc3** and pBeloBAC11, as well as the SAMR898 and SAMR899 genes, generating the strains CGCL009, CGCL010, CGCL026, and CGCL027, respectively.

In pCGC002, the region between SAMR0898 and pBeloBAC11 was replaced with the *att3aac* (Karray, 2005) cassette amplified with the primers Seq_amontF and Seq_amontR. The resulting BAC, pCGC016, was introduced into *E. coli* DH5α containing the plasmid pOSInt3 (Raynal et al., 1998) to excise the *att3aac* cassette, yielding the BAC pCGC017. This BAC was verified by BamHI digestion and introduced into *S. lividans* TK23 by conjugation with *E. coli* S17.1 (pCGC017). Mutants were selected on 50 μg/ml hygromycin and verified by PCR using the primers OMEGA A1 and OMEGA HYG2. The resulting deletion strain was named CGCL009.

This protocol was repeated to delete the region between SAMR0921 and pBeloBAC11 from pCGC017, yielding successively pCGC018 (containing the *att3aac* cassette) and pCGC019 (cassette excised). The *S. lividans* TK23 strain containing pCGC018 obtained after conjugation was named CGCL010.

SAMR0898 and SAMR0899 were deleted from pCGC019 and *cgc2*, *cgc19*, and *cgc18* from pCGC002 by application of the same procedure. Replacements of the genes by the *att2aac* cassette (for *cgc2*) and by the *att3aac* cassette (for *cgc19* and *cgc18*) yielded the BACs pCGC025, pCGC026, pCGC049, pCGC035, and pCGC039, and excisions of the cassettes resulted in the BACs pCGC029, pCGC030, pCGC057, pCGC038, and pCGC040, respectively. Replacements and excisions were verified by PCR. The BACs were integrated into the *S. lividans* chromosome as previously described and the constructs verified by PCR. The resulting strains were named CGCL026, CGCL027, CGCL035, CGCL028, and CGCL029.

Construction of Strains CGCL016, CGCL017, CGCL030, CGCL031, and CGCL032

PCR-based REDIRECT technology was used to construct the *cgc7*, *cgc10*, *cgc3**, *cgc15*, and *cgc16* deletion mutants (Gust et al., 2004). In pCGC002, *cgc7*, *cgc10*, *cgc3**, *cgc15*, and *cgc16* were replaced, respectively, with the FRT*aac* cassette from pW60, yielding the BACs pCGC208, pCGC212, pCGC220, pCGC222, and pCGC224. The BACs pCGC208, pCGC212, pCGC220, pCGC222, and pCGC224 were introduced into *E. coli* BT340 to excise the disruption cassettes, resulting in pCGC209, pCGC213, pCGC221, pCGC223, and pCGC225, respectively. Gene deletions were confirmed by PCR. The mutagenized BACs pCGC209, pCGC213, pCGC221, pCGC223, and pCGC225 were introduced into *S. lividans* TK23 by RP4-based conjugation from *E. coli* ET12567 containing pUZ8002 and selected for hygromycin. The presence of the BACs in *S. lividans* TK23 was verified by PCR using the same oligonucleotides as used to confirm gene deletions. The resulting *cgc7*, *cgc10*, *cgc3**, *cgc15*, and *cgc16* deletion mutant strains were named CGCL016, CGCL017, CGCL030, CGCL031, and CGCL032.

Construction of Strain CGCL001

The BAC pCGC001 was digested with TfiI. A 5.4 kb fragment containing *cgc1**, *cgc2**, the 5′ extremity of *cgc3**, and the intergenic region between *cgc1** and *cgc1*, likely to contain the promoter region of *cgc1**, was blunt ended by Klenow enzyme and ligated into EcoRV-digested pSET152. The resulting plasmid, pCGC502, was introduced into *S. lividans* TK23 by protoplast transformation. Transformants were selected on 50 μg/ml apramycin. Resistance to congocidine of the resulting strain, CGCL001, of *S. lividans* TK23, and of *S. lividans* TK23 carrying the empty vector pSET152 was determined on HT plates containing 20 μg/ml congocidine after 4 days of incubation at 30°C.

Construction of CGCA019

Twenty-one *cgc* genes, from *cgc3** to *cgc18*, were replaced with a hygromycin-resistance cassette in the strain SPM110, already mutated in the spiramycin biosynthetic cluster. To delete the *cgc* genes, two fragments, a 1.2 kb fragment internal to *cgc3** and a 1.5 kb fragment overlapping *cgc18* and *cgc19*, were generated by PCR using the primers CG3+F, CG3+R, CG18_SNF, and CG18_SNR. An insert consisting of the *att3*Ω*hyg* cassette flanked by these two fragments was inserted into the suicide vector pOJ260 (Bierman et al., 1992) to give pCGC516. This plasmid was introduced into *S. ambofaciens* SPM110 by conjugation, and hygromycin-resistant apramycin-sensitive clones were verified by PCR for the replacement of the *cgc3**–*cgc18* region by the hygromycin-resistance cassette.

Congocidine Bioassay

The congocidine bioassay was carried out on HT plates. A patch of about 1 cm^2 of the *Streptomyces* strain to be tested was grown at the center of the plate for 5 days at 30°C. A 3 ml top layer of soft nutrient agar (Difco) containing the indicator microorganisms *E. coli* or *B. subtilis* was then added onto the plates. Growth inhibition zones were observed after overnight incubation at 37°C.

 ©2009 Elsevier Ltd All rights reserved

LC and LC-MS-MS Analyses

After 4 days of culture at 30°C, supernatants were filtered through ultrafree-MC (0.1 μm; Millipore) and analyzed on an Atlantis dC18 column (250 mm × 4.6 mm, 5 μm, column temperature 30°C) using an Agilent 1200 HPLC instrument equipped with a quaternary pump. Samples were eluted with isocratic 0.1% HCOOH in H_2O (solvent A):0.1% HCOOH in CH_3CN (solvent B) (95:5) at 1 ml/min for 7 min followed by a gradient to 40:60 A:B over 23 min. Congocidine was detected by monitoring absorbance at 297 nm. A Bruker Daltonik Esquire HCT ion trap mass spectrometer equipped with an orthogonal atmospheric pressure interface-electrospray ionization (AP-ESI) source was used for LC-MS analyses. Elution from the Atlantis dC18 was split into two flows: 1/10 was directed to an ESI-mass spectrometer for MS and MS-MS measurements, and the remaining 9/10 to a diode array UV detector. In the ESI process, nitrogen served as the drying and nebulizing gas, and helium gas was introduced into the ion trap both for efficient trapping and cooling of the ions generated by the ESI and for fragmentation processes. Ionization was carried out in positive/negative mode with a nebulizing gas set at 241 kPa, a drying gas set at 8 μl/min, and a drying temperature set at 340°C for optimal spray and desolvation. Ionization and mass analysis conditions (capillary high voltage, skimmer and capillary exit voltages, and ion transfer parameters) were tuned for optimal detection of compounds in the range of masses between 100 and 600 m/z. For structural characterization by mass fragmentation, an isolation width of 1 mass unit was used for isolating the parent ion. A fragmentation energy ramp was used for automatically varying the fragmentation amplitude to optimize the MS-MS fragmentation process.

SUPPLEMENTAL DATA

Supplemental data include five figures and two tables and can be found with this article online at http://www.cell.com/chemistry-biology/supplemental/S1074-5521(09)00113-6.

ACKNOWLEDGMENTS

The authors would like to thank Robert Thaï for his help with the mass spectrometer. This work was supported in part by the European Union through the Integrated Project ActinoGEN (CT-2004-0005224), by grants from Région Ile-de-France (to M.J.), and by a fellowship from the French government (to S.N.).

Received: December 8, 2008
Revised: February 16, 2009
Accepted: March 6, 2009
Published: April 23, 2009

REFERENCES

Asai, A., Sakai, Y., Ogawa, H., Yamashita, Y., Kakita, S., Ochiai, K., Ashizawa, T., Mihara, A., Mizukami, T., and Nakano, H. (2000). Pyrronamycin A and B, novel antitumor antibiotics containing pyrrole-amide repeating units, produced by *Streptomyces* sp. J. Antibiot. (Tokyo) *53*, 66–69.

Baraldi, P.G., Tabrizi, M.A., Preti, D., Fruttarolo, F., Avitabile, B., Bovero, A., Pavani, G., del Carmen Nunez Carretero, M., and Romagnoli, R. (2003). DNA minor-groove binders. Design, synthesis, and biological evaluation of ligands structurally related to CC-1065, distamycin and anthramycin. Pure Appl. Chem. *75*, 187–194.

Barona-Gómez, F., Lautru, S., Francou, F.-X., Leblond, P., Pernodet, J.-L., and Challis, G.L. (2006). Multiple biosynthetic and uptake systems mediate siderophore-dependent iron acquisition in *Streptomyces coelicolor* A3(2) and *Streptomyces ambofaciens* ATCC 23877. Microbiology *152*, 3355–3366.

Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N., and Schoner, B.E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene *116*, 43–49.

Burgoyne, D.L., Hemscheidt, T.K., Moore, R.E., and Runnegar, M.T.C. (2000). Biosynthesis of cylindrospermopsin. J. Org. Chem. *65*, 152–156.

Challis, G.L., Ravel, J., and Townsend, C.A. (2000). Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem. Biol. *7*, 211–224.

Choulet, F., Aigle, B., Gallois, A., Mangenot, S., Gerbaud, C., Truong, C., Francou, F.-X., Fourrier, C., Guérineau, M., Decaris, B., et al. (2006). Evolution of the terminal regions of the *Streptomyces* linear chromosome. Mol. Biol. Evol. *23*, 2361–2369.

Cosar, C., Ninet, L., Pinnert-Sindico, S., and Preud'homme, J. (1952). Activité trypanocide d'un antibiotique produit par un *Streptomyces*. C. R. Hebd. Seances Acad. Sci. *234*, 1498–1499.

Denny, W.A. (2001). DNA minor groove alkylating agents. Curr. Med. Chem. *8*, 533–544.

Fischbach, M.A., and Walsh, C.T. (2006). Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery and mechanisms. Chem. Rev. *106*, 3468–3496.

Gust, B., Chandra, G., Jakimowicz, D., Yuqing, T., Bruton, C.J., and Chater, K.F. (2004). λRed-mediated genetic manipulation of antibiotic-producing *Streptomyces*. Adv. Appl. Microbiol. *54*, 107–128.

Haynes, S.W., and Challis, G.L. (2007). Non-linear enzymatic logic in natural product modular mega-synthases and -synthetases. Curr. Opin. Drug Discov. Devel. *10*, 203–218.

Hojati, Z., Milne, C., Harvey, B., Gordon, L., Borg, M., Flett, F., Wilkinson, B., Sidebottom, P.J., Rudd, B.A.M., Hayes, M.A., et al. (2002). Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from *Streptomyces coelicolor*. Chem. Biol. *9*, 1175–1187.

Karray, F. (2005). Etude de la biosynthèse de l'antibiotique spiramycine par *Streptomyces ambofaciens*. PhD thesis, Université Paris Sud 11, Orsay, France.

Keating, T.A., Suo, Z., Ehmann, D.E., and Walsh, C.T. (2000a). Selectivity of the yersiniabactin synthetase adenylation domain in the two-step process of amino acid activation and transfer to a holo-carrier protein domain. Biochemistry *39*, 2297–2306.

Keating, T.A., Marshall, C.G., and Walsh, C.T. (2000b). Vibriobactin biosynthesis in *Vibrio cholerae*: VibH is an amide synthase homologous to nonribosomal peptide synthetase condensation domains. Biochemistry *39*, 15513–15521.

Keating, T.A., Marshall, C.G., Walsh, C.T., and Keating, A.E. (2002). The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. Nat. Struct. Biol. *9*, 522–526.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, D.A. (2000). Practical *Streptomyces* Genetics (Norwich, UK: The John Innes Foundation).

Lautru, S., and Challis, G.L. (2004). Substrate recognition by nonribosomal peptide synthetase multi-enzymes. Microbiology *150*, 1629–1636.

Liu, W., Christenson, S.D., Standage, S., and Shen, B. (2002). Biosynthesis of the enedyine antitumor antibiotic C-1027. Science *297*, 1170–1173.

Mihali, T.K., Kellmann, R., Muenchhoff, J., Barrow, K.D., and Neilan, B.A. (2008). Characterization of the gene cluster responsible for cylindrospermopsin biosynthesis. Appl. Environ. Microbiol. *74*, 716–722.

Neidle, S. (2001). DNA minor-groove recognition by small molecules. Nat. Prod. Rep. *18*, 291–309.

Pang, X., Aigle, B., Girardet, J.-M., Mangenot, S., Pernodet, J.-L., Decaris, B., and Leblond, P. (2004). Functional angucycline-like antibiotic gene cluster in the terminal inverted repeats of the *Streptomyces ambofaciens* linear chromosome. Antimicrob. Agents Chemother. *48*, 575–588.

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.

Quirós, L.M., Aguirrezabalaga, I., Olano, C., Méndez, C., and Salas, J.A. (1998). Two glycosyltransferases and a glycosidase are involved in oleandomycin modification during its biosynthesis by *Streptomyces antibioticus*. Mol. Microbiol. *28*, 1177–1185.

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 (TSVM). Nucleic Acids Res. *33*, 5799–5808.

Raynal, A., Tuphile, K., Gerbaud, C., Luther, T., Guérineau, M., and Pernodet, J.-L. (1998). Structure of the chromosomal insertion site for pSAM2: functional analysis in *Escherichia coli*. Mol. Microbiol. *28*, 333–342.

 ©2009 Elsevier Ltd All rights reserved

Raynal, A., Karray, F., Tuphile, K., Darbon-Rongère, E., and Pernodet, J.-L. (2006). Excisable cassettes: new tools for functional analysis of *Streptomyces* genomes. Appl. Environ. Microbiol. *72*, 4839–4844.

Roche, E.D., and Walsh, C.T. (2003). Dissection of the EntF condensation domain boundary and active site residues in nonribosomal peptide synthesis. Biochemistry *42*, 1334–1344.

Sambrook, J., and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual, Third Edition (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Schwarzer, D., Finking, F., and Marahiel, M.A. (2003). Nonribosomal peptides: from genes to products. Nat. Prod. Rep. *20*, 275–287.

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.

Steller, S., Vollenbroich, D., Leenders, F., Stein, T., Conrad, B., Hofemeister, J., Jacques, P., Thonart, P., and Vater, J. (1999). Structural and functional organization of the fengycin synthetase multienzyme system from *Bacillus subtilis* b213 and A1/3. Chem. Biol. *6*, 31–41.

Stumpp, T., Himbert, S., and Altenbuchner, J. (2005). Cloning of the netropsin resistance genes from *Streptomyces flavopersicus* NRRL 2820. J. Basic Microbiol. *45*, 355–362.

Thomas, M.G., Chan, Y.A., and Ozanick, S.G. (2003). Deciphering tuberactinomycin biosynthesis: isolation, sequencing and annotation of the viomycin biosynthetic gene cluster. Antimicrob. Agents Chemother. *47*, 2823–2830.

Walsh, C.T., Garneau-Tsodikova, S., and Howard-Jones, A.R. (2006). Biological formation of pyrroles: nature's logic and enzymatic machinery. Nat. Prod. Rep. *23*, 517–531.

Wildfeuer, M.E. (1964). The biosynthesis of netropsin. Thesis, University of Delaware, Newark, Delaware.

Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., and Court, D.L. (2000). An efficient recombination system for chromosome engineering in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *97*, 5978–5983.