Biosynthesis of elloramycin in *Streptomyces olivaceus* requires glycosylation by enzymes encoded outside the aglycon cluster

Angelina Ramos,1† Felipe Lombó,1 Alfredo F. Braña,1 Jürgen Rohr,2 Carmen Méndez1 and José A. Salas1

1Departamento de Biología Funcional e Instituto Universitario de Oncología del Principado de Asturias (IUOPA), Universidad de Oviedo, 33006 Oviedo, Spain

2Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, 725 Rose Street, Lexington, KY 40536-0082, USA

Elloramycin is an anthracycline-like antitumour drug produced by *Streptomyces olivaceus* Tü2353. Cosmid cos16F4 has been previously shown to direct the biosynthesis of the elloramycin aglycon 8-demethyltetracenomycin C (8-DMTC), but not elloramycin. Sequencing of the 24.2 kb insert in cos16F4 shows the presence of 17 genes involved in elloramycin biosynthesis (*elm* genes) together with another additional eight ORFs probably not involved in elloramycin biosynthesis. The 17 genes would code for the biosynthesis of the polyketide moiety, sugar transfer, methylation of the tetracyclic ring and the sugar moiety, and export. Four genes (*rhaA, rhaB, rhaC* and *rhaD*) encoding the enzymic activities required for the biosynthesis of the sugar L-rhamnose were also identified in the *S. olivaceus* chromosome. The involvement of this rhamnose gene cluster in elloramycin biosynthesis was demonstrated by insertional inactivation of the *rhaB* gene, generating a non-producer mutant that accumulates the 8-DMTC C aglycon. Coexpression of cos16F4 with pEM4RO (expressing the four rhamnose biosynthesis genes) in *Streptomyces lividans* led to the formation of elloramycin, demonstrating that both subclusters are required for elloramycin biosynthesis. These results demonstrate that, in contrast to most of the biosynthesis gene clusters from actinomycetes, genes involved in the biosynthesis of elloramycin are located in two chromosomal loci.

INTRODUCTION

A large variety of bioactive compounds has been isolated from different origins, including micro-organisms, plants and some animals. Structurally, these compounds belong to many different families and show diverse chemical structures. However, a large number of them share a common feature: they are glycosylated. Most of the sugars belong to the family of the 6-deoxysugars, of which more than 70 different variants have been reported in natural products from plants, fungi and bacteria (Weymouth-Wilson, 1997; Trefzer et al., 1999). These deoxysugars are synthesized from NDP-activated hexoses (mainly D-glucose) via 4-keto-6-deoxy intermediates (Liu & Thorson, 1994; Piepersberg, 1994). In recent years, an increasing number of gene clusters involved in 6-deoxysugar biosynthesis have been characterized from antibiotic-producing actinomycetes (Salas & Méndez, 2005, 2007). Usually, the complete set of genes for biosynthesis of deoxysugars is clustered together with other genes involved in the biosynthesis of the different compounds, including genes participating in the biosynthesis of the aglycon moiety, and regulatory, resistance and secretion genes.

Elloramycin is an anthracycline-like antitumour drug produced by *Streptomyces olivaceus* Tü2353 (Drautz et al., 1985). It belongs to the large and important family of the aromatic polyketides and its aglycon closely resembles tetracenomycin C, but has an additional C-12a-O-methyl group and, in contrast to tetracenomycin C, the C-8 hydroxyl group is not methylated but glycosylated with a permethylated L-rhamnose residue (Fig. 1A). From a cosmid library of the elloramycin producer *S. olivaceus* Tü2353, cosmid cos16F4 was isolated and expressed in *Streptomyces lividans*, resulting in the production of an elloramycin biosynthetic intermediate, 8-demethyltetracenomycin C (8-DMTC; Fig. 1B) (Decker et al., 1995). The lack of formation of elloramycin in these experiments suggested that cos16F4, although it should contain all genes...
necessary for the biosynthesis of the polyketide moiety of elloramycin, probably lacked some of the genes required for sugar biosynthesis and/or transfer.

Here we report the complete nucleotide sequence of the 24.2 kb insert in cos16F4, showing that the elloramycin cluster lacks all genes involved in the biosynthesis of the deoxyxugar L-rhamnose. Identification of L-rhamnose biosynthesis genes in this organism was also achieved; they are located outside the boundaries of the gene cluster. By heterologous expression of these genes we showed that they direct the biosynthesis of NDP-activated L-rhamnose. Involvement of these genes in elloramycin biosynthesis was proved by generating an elloramycin non-producing mutant through inactivation of the L-rhamnose biosynthesis gene rhaB.

METHODS

Micro-organisms, plasmids and culture conditions. Streptomyces olivaceus Tu2353 (Drautz et al., 1985) was used as donor of chromosomal DNA and Streptomyces lividans TK21 (Kieser et al., 2000) as host for gene expression. For sporulation they were grown on A medium and for antibiotic production on RSA medium (Fernández et al., 1998). Escherichia coli DH10B (Invitrogen) and E. coli ET12567(pUB307) (Flett et al., 1997) were used as hosts. Production, transformation, regeneration of protoplasts and conjugation experiments were performed following standard procedures (Kieser et al., 2000). Plasmids pOJ260 (Bierman et al., 1992) and pEM4 (Quirós et al., 1998) were used for gene inactivation and gene expression, respectively. pCR-Blunt (Invitrogen) was used for cloning of PCR products. pEFBA was used as donor of the apramycin-resistance gene aac(3) IV (Fernández Lozano et al., 2000). PHAM was used as a source of genes coding for enzymes involved in L-rhamnose biosynthesis (Rodriguez et al., 2000). pBluescript SK (Stratagene) and pUK21 (Vieira & Messing, 1991) were used for subcloning. When antibiotic selection of transformants was required, ampicillin (100 μg ml⁻¹), apramycin (25 μg ml⁻¹), thiostrepton (50 μg ml⁻¹) or kanamycin (50 μg ml⁻¹) were used.

DNA manipulation and sequencing. Total DNA isolation, plasmid DNA preparation, restriction endonuclease digestions, ligations and other DNA manipulations were performed according to standard procedures for E. coli (Sambrook & Russell, 2001) and Streptomyces (Kieser et al., 2000). DNA sequencing was performed on double-stranded DNA templates with the dideoxynucleotide chain-termination method (Sanger et al., 1977) and the Cy5 Autocycle Sequencing kit (GE Healthcare), using the Alph-Express automatic DNA sequencer (GE Healthcare). Computer-aided database searching and sequence analysis were carried out with the University of Wisconsin Genetics Computer Group software (Devereux et al., 1984) and the BLAST program (Altschul et al., 1990). In situ Southern hybridization was carried out according to standard procedures (Sambrook & Russell, 2001) and by using the DIG DNA Labelling and Detection kit (Roche). The 24.2 kb insert from cos16F4 and the 6.4 kb Splh–BamHI sequenced DNA fragment from the chromosome have been deposited in the EMBL database under accession numbers AM900040 and AM889123, respectively.

PCR amplification and insertional inactivation of rhaB. To locate DNA sequences coding for NDP-α-glucose-4,6-dehydratases, degenerate oligoprimers were used based on amino acid sequences conserved in these enzymes: dh-1 (5’-CSGGSGSGSGSGTTCAT-SGG-3’) and dh-2 (5’-GGGCGCTGGYRSGGCGATGG-3’) (Decker et al., 1996). PCR conditions used were 95 °C for 4 min; 30 cycles of 95 °C for 30 s, 65 °C for 30 s, and 68 °C for 1.30 min; and a final extension cycle at 68 °C for 5 min. The enzyme used for PCR amplification was Platinum Pfu (Invitrogen). The 546 bp amplified PCR product was cloned into pCR-Blunt, generating pCRDH.

For inactivating the rhaB gene, pCRDH was digested with EcoRI, and the released fragment subcloned into pOJ260 to obtain pOJDH1. This construct was introduced into S. olivaceus by intergeneric conjugation from E. coli ET12567(pUB307) and transformants were selected with apramycin. Southern analysis of the mutant strain (S. olivaceus Δrhab) was performed using the 546 bp PCR product and the apramycin-resistance gene aac(3) IV as probes.

Construction of a plasmid directing the biosynthesis of L-rhamnose. Chromosomal DNA from S. olivaceus Δrhab was digested independently with BamHI and Splh. After religation and transformation of E. coli, plasmids pOJDHR1 and pOJDHR2 were obtained, containing the regions flanking the rhab gene. From pOJDHR1 a Pmbl–HindIII fragment was isolated and subcloned into pSL1180 digested with the same restriction enzymes, generating pSLRHAM. Then, a Pmbl–Splh fragment from pOJDHR2 was subcloned into the same sites of pSLRHAM, generating pSLRHAM, which contains six genes from the S. olivaceus chromosome.

In order to obtain a plasmid containing only the four L-rhamnose biosynthesis genes, an EcoRI–NotI DNA fragment from pOJDHR1 was subcloned into pUK21 using the same restriction sites. Then a Splh–EcoRI DNA fragment from pSLRHAM was subcloned into this pUK21 derivative, generating pUKR. This construct contains the four genes implicated in L-rhamnose biosynthesis controlled by its own divergent promoter: rhaA and rhaC on one side and rhab and rhaD on the other.

Fig. 1. Chemical structures of elloramycin (A) and 8-demethyltetracenomycin C (8-DMTC) (B).
Finally, the four rha genes were subcloned into pEM4. In order to do this, a SpeI DNA fragment obtained from pUKRO and containing the four rha genes was subcloned into the SpeI site of pBluescriptSK, and then released as a XbaI–HindIII fragment and subcloned into the same restriction sites of pEM4, generating pEM4RO. In this construct the four rha genes are expressed from their own bidirectional promoters.

**HPLC-MS analysis.** For detection of production of elloramycin or biosynthetic intermediates, *S. olivaceus* wild-type or mutants were grown on R5A medium. After extraction of the cultures with ethyl acetate, HPLC analysis was performed in a reversed-phase column (Symmetry C18, 4.6 × 250 mm, Waters) as previously described (Patallo et al., 2001). Detection and spectral characterization of peaks were done with a photodiode array detector and Millennium software (Waters), extracting two-dimensional chromatograms at 280 nm.

HPLC-MS analyses of the compounds were carried out by coupling the chromatographic equipment to a ZQ4000 mass spectrometer (Waters-Micromass), using electrospray ionization in the positive mode, with a capillary voltage of 3 kV and a cone voltage of 20 kV.

**RESULTS**

**Sequencing of the elloramycin gene cluster**

Previous studies in our laboratory identified a glycosyltransferase and three methyltransferase genes in cosmid cos16F4 (Blanco et al., 2001; Patallo et al., 2001). To achieve the complete characterization of the elloramycin gene cluster, we sequenced the entire insert in cos16F4. The analysis of the 24.2 kb DNA insertion revealed the presence of 23 complete ORFs and two incomplete ones (Fig. 2A). Comparison of the deduced products of the genes with proteins in databases allowed us to propose functions for the corresponding gene products (Table 1). Seventeen ORFs (covering a region over 17 kb) are transcribed in the same direction, and would probably code for enzyme activities involved in elloramycin biosynthesis. These genes are flanked on the upstream and downstream sides by three and five ORFs respectively, coding for enzymes potentially involved in primary metabolism to which no role in elloramycin biosynthesis could be assigned (Table 1). These eight ORFs showed, by BLAST analysis, significant similarities to related proteins in *S. coelicolor* A3(2) and *Streptomyces avermitilis* MA-4680, sharing the same genetic organization in these two streptomycetes as in the elloramycin producer. All these data strongly suggest that the limits of the elloramycin cluster reside in the elmD and elmMIII genes.

Three genes, elmK, elmL and elmM, would code for the elloramycin polyketide synthase. The aromatase ElmNI and the cyclases ElmI (fourth ring) and ElmJ (third ring) are necessary for generation of early tetracyclic intermediates. Two genes coding for proteins catalysing oxygenation steps were also found: one corresponding to the monoxygenase ElmH and another one to the oxygenase ElmG. The sequence of the genes elmGHIJ has already been reported (Rafanan et al., 2001). The genes elmNII, elmP and elmD would code for O-methyltransferases. All these reactions end up with the production of the aglycon 8-DMTC, which is then glycosylated with an L-rhamnose by the action of the glycosyltransferase ElmGT (Blanco et al., 2001). Three O-methyltransferases coded by elmMI, elmMII and elmMIII act on this L-rhamnose once attached to the aglycon (Patallo et al., 2001). Finally, elmE would code for a membrane transporter for the secretion of elloramycin.

There is another gene in the cluster to which we cannot ascribe a function, elmF. Its deduced product shows similarity with several *N*-methyltransferases, particularly...
those involved in translation functions. Elloramic acid does not require such a type of methylation and therefore a role for ElmF cannot be proposed.

**Isolation of L-rhamnose biosynthesis genes**

For the biosynthesis of elloramycin, the existence of an L-rhamnose gene cluster in the *S. olivaceus* chromosome is required. The absence of L-rhamnose biosynthesis genes in the cos16F4 insert prompted us to search for these genes elsewhere in the chromosome of *S. olivaceus*. With this aim, we carried out a combined PCR-insertional inactivation strategy. Using oligoprimers based on consensus sequences from different NDP-D-glucose-4,6-dehydratases (Decker et al., 1996), we amplified a 550 bp DNA PCR fragment using *S. olivaceus* chromosomal DNA as template. This fragment was sequenced and confirmed to code for an amino acid sequence strongly resembling NDP-D-glucose dehydratases. By Southern analysis and chromosomal walking off both sides of the PCR fragment, we isolated and sequenced a larger region of 6468 bp. Six complete ORFs, designated *rhaD*, *rhaB*, *rhaA*, *rhaC*, *galE3* and *proB*, were identified (Fig. 2B). Comparison of the deduced gene products with proteins in databases showed that four of these ORFs would represent an L-rhamnose biosynthesis gene cluster (*rha* cluster), and for these enzymes a role in L-rhamnose biosynthesis could be proposed: RhaA as a D-glucose synthase, RhaB as a 4,6-dehydratase, RhaC as a 3,5-epimerase and RhaD as a 4-ketoreductase (Table 2). The four genes would be transcribed from two divergent promoters: *rhaA* and *rhaC* in one direction and *rhaB* and *rhad* in the other.

**Involvement of the rha genes in elloramycin biosynthesis**

The involvement of the *rha* gene cluster in elloramycin biosynthesis was demonstrated by insertional inactivation. The 550 bp PCR product internal to *rhaB* was subcloned into pOJ260. The resultant construct pOJDH1 was introduced into *S. olivaceus* by conjugation. Colonies resistant to apramycin (antibiotic marker in the vector) were isolated and tested for their ability to produce elloramycin by HPLC-MS analysis of the cultures. All apramycin-resistant colonies tested were found not to produce elloramycin but to accumulate 8-DMTC (Fig. 3B), showing the same retention time, absorption spectrum and mass spectrum as determined by comparison with a pure sample of this compound, used as standard for HPLC-MS and as determined by comparison with a pure sample of this compound, used as standard for HPLC-MS and showing the same retention time, absorption spectrum and m/z value. This demonstrated that the rhamnose cluster...
found was indeed involved in the biosynthesis of elloramycin.

In order to demonstrate that the four genes rhaA, rhaB, rhaC and rhaD were able to direct the biosynthesis of L-rhamnose, they were expressed in S. lividans. In control experiments where pEM4 (control vector) and cos16F4 were coexpressed, no elloramycin was detected but rather 8-DMTC (Fig. 3C), indicating that the S. lividans host was not providing L-rhamnose from its own biochemical machinery. To detect the formation L-rhamnose mediated by pEM4RO, we coexpressed this plasmid in S. lividans simultaneously with cos16F4 (for aglycon biosynthesis). HPLC-MS analysis of cultures of the resultant recombinant strain showed the presence of elloramycin (Fig. 3D), with the same HPLC-MS retention time and m/z value as from a pure sample. Interestingly, coexpression of pEM4RO and cos16F4 originated, in addition to elloramycin, in a variety of elloramycin biosynthetic intermediates or shunt products with different degrees of sugar methylation as detected by HPLC-MS (Fig. 3D). It is worth mentioning that these compounds were either not detected or detected in very small amounts in cultures of the elloramycin producer S. olivaceus (Fig. 3A), indicating that biosynthesis of elloramycin in this organism is an efficiently regulated process. Apparently, expression in another host such as S. lividans causes an unbalancing of the process, producing this range of incompletely biosynthesized compounds.

### DISCUSSION

On the basis of the sequence information included in this paper and some previous reports (Shen & Hutchinson, 1994; Blanco et al., 2001; Patallo et al., 2001), we propose a biosynthetic pathway for elloramycin (Figs 4 and 5), which would involve the participation of the two independently chromosomally located gene clusters. The polyketide synthase proteins (coded by elmK, elmL and elmM) are highly similar to the tetracenomycin C ketoacyl synthases α and β and acyl carrier protein, respectively (Bibb et al., 1989) and would generate the initial decaketide. The aromatase ElmNI would be in charge of the first two ring cyclizations. In contrast with its homologue in the

### Table 2. Deduced functions of genes involved in L-rhamnose biosynthesis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Position</th>
<th>aa</th>
<th>Deduced function</th>
<th>Closest match*</th>
</tr>
</thead>
<tbody>
<tr>
<td>rhaD</td>
<td>429–1311c</td>
<td>294</td>
<td>4-Ketoreductase</td>
<td>StrL, S. griseus subsp. griseus (60/68)</td>
</tr>
<tr>
<td>rhaB</td>
<td>1333–2296c</td>
<td>321</td>
<td>4,6-Dehydratase</td>
<td>NovT, S. caeruleus (68/81)</td>
</tr>
<tr>
<td>rhaA</td>
<td>2399–3281</td>
<td>294</td>
<td>NDP-D-glucose synthase</td>
<td>NovV, S. caeruleus (69/83)</td>
</tr>
<tr>
<td>rhaC</td>
<td>3298–3904</td>
<td>202</td>
<td>3,5-Epimerase</td>
<td>NovW, S. caeruleus (63/75)</td>
</tr>
<tr>
<td>galE3</td>
<td>3997–4987</td>
<td>330</td>
<td>UDP-D-Glucose 4-epimerase</td>
<td>BAC72798, S. avermitilis (69/78)</td>
</tr>
<tr>
<td>proB</td>
<td>5042–6185</td>
<td>381</td>
<td>γ-Glutamyl kinase</td>
<td>Q82C83, S. avermitilis (87/95)</td>
</tr>
</tbody>
</table>

*Values in parentheses indicate percentage identity/similarity.
tetracenomycin C pathway TcmN, in which the C-terminus also includes an O-methyltransferase domain (Summers et al., 1992), ElmNI appears to be a monofunctional protein. A frameshift mutation could be responsible for the existence of two contiguous and separate genes in the elloramycin cluster, elmNI (aromatase) and elmNII (O-methyltransferase; see below). The action of the cyclases ElmI and ElmJ and the monooxygenase ElmH would generate the tetracyclic biosynthetic intermediate tetracenomycin D3. The next two steps in elloramycin biosynthesis would be catalysed by the

---

**Fig. 4.** Proposed biosynthetic steps for elloramycin biosynthesis.

---

**Fig. 5.** Proposed biosynthetic steps for L-rhamnose biosynthesis.
O-methyltransferase ElmNII, which would generate a methoxy group at C-3, and by the O-methyltransferase ElmP, acting on the carboxyl group attached at C-9. These two enzymes are highly similar to TcmN (C terminus only) and to TcmP (Decker et al., 1993) from the tetracenomycin C pathway, respectively. Finally, the oxygenase ElmG would introduce the hydroxyl groups at C-4, C-4a and C-12a, generating 8-DMTC, which is the last non-glycosylated intermediate during elloramycin biosynthesis. The glycosyltransferase ElmGT would now transfer an L-rhamnose moiety at the C-8 position of 8-DMTC (Blanco et al., 2001). As shown in this paper, the sequenced region does not contain genes encoding enzymic functions required for the biosynthesis of L-rhamnose. This glycosylated intermediate would be further permethylated by the action of the three O-methyltransferases ElmMI, ElmMII and ElmMIII (Patallo et al., 2001). The last biosynthetic step would be the generation of a methoxy group at C-12a. Most probably this gene function is encoded by elmD. ElmD contains the same O-methyltransferase domain as ElmP, and it shows high similarity with several antibiotic biosynthesis methyltransferases.

In the last decade a good number of antibiotic biosynthesis gene clusters have been characterized from actinomycetes. These clusters have been sequenced and functions have been assigned to the different genes (and their products) mainly on the basis of database comparisons, gene inactivation and gene expression. A common feature to most of these gene clusters is that usually all the genes involved in the biosynthesis of the compound are linked together in the chromosome of the producer strain. The absence of some genes in the cluster is the exception rather than the rule. However, it seems that antibiotics containing L-rhamnose in their structures constitute such an exception. The gene clusters for spinosad (Waldron et al., 2001), steffimycin (Gullón et al., 2006), aranciamycin (Luzhetskyy et al., 2007) and elloramycin (this paper) do not contain any L-rhamnose biosynthesis genes, and so probably they use L-rhamnose present in the cytoplasm for other cellular functions or structures. In two cases, spinosin (Madduri et al., 2001) and elloramycin (this paper), a rhamnose biosynthesis cluster has been identified. In the case of Saccharopolyspora spinosa (spinosin producer), it has been demonstrated that this rhamnosyl biosynthesis gene cluster is required both for the biosynthesis of spinosin and for the correct formation of cell wall (Madduri et al., 2001). In the case of Streptomyces olivaceus, disruption of the rha genes apparently only affects the biosynthesis of elloramycin, with no morphological changes in the microorganism. Using primary metabolism genes for secondary biosynthetic purposes may represent an energetic advantage, but on the other hand this requires some kind of common regulation involving such distinct cellular processes.

To get further insight into the possible presence of L-rhamnose biosynthesis genes in other actinomycetes, we carried out a search using the deduced products of the four S. olivaceus rha cluster genes against five actinomycete genome sequences available (S. coelicolor, S. avermitilis, Saccharopolyspora erythraea, Salinospora arenicola and Salinospora tropica). This search showed that only Sac. erythraea and Sal. arenicola seem to possess homologues to the four genes while the others appear to contain only an incomplete rhamnose cluster: S. coelicolor contains only a rhaB-like gene (SCO0749); S. avermitilis and Sal. tropica contain rhaB- (SAV946 and Strop_2222) and rhaC-like genes (SAV949 and Strop_2217); S. avermitilis also contains a rhaA-like gene (SAV947). rhaD-like genes were not found in these three organisms. These genes, when present, are grouped in a small cluster as usually occurs for deoxxygenated biosynthesis genes (Méndez & Salas, 2001). The absence of an L-rhamnose cluster in S. coelicolor explains why upon introduction of cos16F4 in this organism formation of elloramycin is not detected (data not shown). We can conclude that a rhamnose gene cluster seems not to be present in all actinomycetes and it is not essential for normal growth and development of these organisms.

ACKNOWLEDGEMENTS

This research was supported by the Spanish Ministry of Education and Science (BFU2006-00404) and Red Temática de Investigación Cooperativa de Centros de Cáncer (Ministry of Health, Spain; ISCIII-RETIC RD06/0020). We thank Obra Social Cajastur for financial support to F.L.

REFERENCES


Edited by: C. W. Chen