The clavulanic acid biosynthetic cluster of Streptomyces clavuligerus: genetic organization of the region upstream of the car gene

Encarnación Mellado,¹ Luis Miguel Lorenzana,² Marta Rodríguez-Sáiz,¹ Bruno Diez,¹ Paloma Liras² and José Luis Barredo¹

Author for correspondence: José L. Barredo. Tel: +34 987 895826. Fax: +34 987 895986. e-mail: jbarredo@antibioticos.it

The genetic organization of the region upstream of the car gene of the clavulanic acid biosynthetic gene cluster of Streptomyces clavuligerus has been determined. Sequence analysis of a 12.1 kb region revealed the presence of 10 ORFs whose putative functions, according to database searches, are discussed. Three co-transcriptional units are proposed: ORF10–11, ORF12–13 and ORF15–16–17–18. Potential transcriptional terminators were identified downstream of ORF11 (fd) and ORF15. Targeted disruption of ORF10 (cyp) gave rise to transformants unable to produce clavulanic acid, but with a considerably higher production of cephamycin C. Transformants inactivated at ORF14 had a remarkably lower production of clavulanic acid and similar production of cephamycin C. Significant improvements of clavulanic acid production, associated with a drop in cephamycin C biosynthesis, were obtained with transformants of S. clavuligerus harbouring multiple copies of plasmids carrying different constructions from the ORF10–14 region. This information can be used to guide strain improvement programs, blending random mutagenesis and molecular cloning, to optimize the yield of clavulanic acid.

Keywords: antibiotic biosynthesis, gene cluster, cytochrome P450, ferredoxin, PBP

INTRODUCTION

Streptomyces clavuligerus is used industrially for the production of clavulanic acid (CA), a β-lactam compound with β-lactamase-inhibitory properties, useful against β-lactamase-producing micro-organisms when combined with conventional β-lactam antibiotics (Reading & Cole, 1977). The biosynthetic pathway leading to cephamycin in S. clavuligerus has been well characterized and the corresponding genes have been identified (Paradkar et al., 1996a; Liras, 1999). However, the CA biosynthetic route remains only partially understood. A pathway comprising a putative sequence of biochemical reactions leading to CA and clavams has been proposed (Fig. 1) (Liras & Rodríguez-García, 2000), but additional work needs to be done in order to determine the complete biosynthetic pathway.

A condensation reaction of arginine with d-glycer-aldehyde 3-phosphate to form N2-(2-carboxyethyl)-arginine (CEA) has been recently elucidated as the initial step of the CA pathway (Khaleeli et al., 1999). This biosynthetic step is catalysed by a TPP-dependent carboxyethylarginine synthase encoded by the ceas (also named pyc) gene (Pérez-Redondo et al., 1999). The synthesis of the β-lactam ring from CEA is catalysed by β-lactam synthetase, a new type of ATP/Mg2+-dependent enzyme, encoded by the bls gene (Bachmann et al., 1998; Bachmann & Townsend, 2000). The crystal structure of this β-lactam synthetase has been recently reported, revealing evolutionary changes with reference to its homologous class B asparagine synthetases, which are essential for β-lactam biosynthesis (Miller et al., 2001). After β-lactam ring formation, only three enzymes have been characterized: clavaminate synthase (CAS),

Abbreviations: CA, clavulanic acid; PBP, penicillin-binding protein.

The GenBank accession number for the 12162 bp sequence reported in this paper is AY034175.
Fig. 1. Biosynthetic pathway leading to CA (Liras & Rodríguez-García, 2000). The genes and enzymic activities characterized are the following: ceaS (also named pyc) encoding \(N^2\)-carboxyethylarginine synthase (CEAS) (Khaleeli et al., 1999; Pérez-Redondo et al., 1999), bls encoding \(\beta\)-lactam synthetase (BLS) (Bachmann et al., 1998; Bachmann & Towsend, 2000), cas2 encoding clavaminic acid synthase (CAS) (Elson et al., 1987; Hodgson et al., 1995), pah encoding proclavaminic acid amidinohydrolase (PAH) (Wu et al., 1995) and car (also named cad) encoding clavulanate-9-aldehyde reductase (CAR) (Nicholson et al., 1994; Pérez-Redondo et al., 1998).

Proclavaminic acid amidinohydrolase (PAH) and clavulanate-9-aldehyde reductase (CAR). *S. clavuligerus* has two CAS isoenzymes (CAS1 and CAS2), encoded by the genes cas1 and cas2, which are at least 20 kb apart in the chromosome of *S. clavuligerus* (Marsh et al., 1992). CAS2 catalyses the hydroxylation of deoxyguanidinoproclavaminic acid to guanidinoproclavaminic acid and the further oxidative cyclization of proclavaminic acid to clavaminic acid (Elson et al., 1987). PAH, which hydrolyses the guanidic residue of the guanidinoproclavaminic acid, is encoded by the pah gene (Wu et al., 1995). CAR is likely to be involved in the direct conversion of the highly reactive intermediate clavulanate-9-aldehyde into CA (Nicholson et al., 1994) and is encoded by the car (also named cad) gene (Pérez-Redondo et al., 1998).

A total of 11 genes involved in CA biosynthesis have been located in the genome of *S. clavuligerus* in a region immediately contiguous to the cephamycin biosynthetic gene cluster (Jensen et al., 2000; Li et al., 2000). In addition to the five above-mentioned genes (ceaS, bls, cas2, pah and car), ORF6, ORF7, claR, cyp, fd and ORF12 have been described (Fig. 2). ORF6 and ORF7 encode respectively a putative ornithine acyltransferase and a protein involved in peptide transport (Hodgson et al., 1995). The claR gene, which encodes a regulatory protein similar to transcriptional activators of the LysR family, is involved in the regulation of the late steps of the pathway (Pérez-Redondo et al., 1998; Paradkar et al., 1998). While the genes cyp (encoding a cytochrome P450) and fd (encoding a ferredoxin) may be involved in an oxidative reaction late in the pathway, the function of the product of ORF12 is still unknown (Li et al., 2000). The cyp gene was disrupted in *S. clavuligerus* by insertion of the tsr gene, leading to the complete loss of CA production (Li et al., 2000). The enzymic complex cytochrome P450/ferredoxin is a common component in antibiotic gene clusters of other actinomycetes (Fouces et al., 1999). Ferredoxins have been shown to supply electrons to multicomponent monooxygenases in Streptomyces spp. (O’Keefe et al., 1991). Amplification of the pyc or claR genes in *S. clavuligerus* results in an increase of CA production; however, amplification of the car gene has no significant effect on production (Pérez-Redondo et al., 1999).

Although preliminary studies in *S. clavuligerus* indicated the presence of the complete CA biosynthetic pathway in a discrete DNA region containing eight contiguous genes (Jensen et al., 1993), the existence of additional clustered genes required for CA production outside of this region is assumed (Li et al., 2000). Since the genes involved in antibiotic biosynthesis are clustered in most
actinomycetes, we worked on the isolation and characterizing of additional genes of the right region flanking the CA gene cluster. In this way, we have identified seven additional ORFs (from ORF13 to ORF19; Fig. 2). This information will contribute to clarifying the CA biosynthetic pathway and provides valuable tools for strain improvement of *S. clavuligerus*.

**METHODS**

**Micro-organisms, phages and plasmids.** The CA-producing strain *S. clavuligerus* ATCC 27064 (Reading & Cole, 1977) was used as source of DNA. *Escherichia coli* SURE (Stratagene) and *E. coli* LE392 (Sambrook et al., 1989) were the hosts for λGEM12 phage derivatives. *E. coli* DH5α (Sambrook et al., 1989) was the recipient for high-frequency plasmid transformation. Phage λGEM12 (Promega) was used for the construction of the genomic library of *S. clavuligerus*. pBlue-script I KS(+) and pBC KS(+) phagemids (Stratagene) were used for subcloning and sequencing. Single-stranded DNA for sequencing was produced with phagemids by standard techniques (Sambrook et al., 1989). pULVK99 (Chary et al., 1997) was used as *E. coli*-streptomycoses spp. shuttle vector and pHZ1331 (Pérez-Redondo et al., 1999) as targeted inactivation vector.

**Transformation and fermentation.** Protoplast transformation of *Streptomyces lividans* and *S. clavuligerus* was as previously described (Hopwood et al., 1985; García-Domínguez et al., 1987). Chloramphenicol (30 µg ml⁻¹) or ampicillin (100 µg ml⁻¹) were used for selection of transformants in *E. coli*. Transformants of *S. clavuligerus* were selected with thioestrepton (30 µg ml⁻¹). Insertional inactivation was done as described by Pérez-Redondo et al. (1999). CA production was tested by flask fermentation in TSB (tryptic soy broth 3%; Difco) every 12 h (Romero et al., 1984) and further quantified by derivatization with imidazole (Bird et al., 1982) and HPLC analysis (Mosher et al., 1999).

**Construction and screening of a genomic library of *S. clavuligerus*.** The DNA of *S. clavuligerus* was purified as described by Hopwood et al. (1985). Fragments of 17–22 kb were purified from Sau3AI partially digested DNA and ligated to λGEM12 according to standard procedures (Sambrook et al., 1989). Ligation products were packaged in vitro with the Gigapack II Gold kit (Stratagene), obtaining around 2 × 10⁸ p.f.u. For the screening, the library was amplified in *E. coli* LE392, plated to obtain about 3 × 10⁴ p.f.u., and hybridized with digoxigenin-labelled probes according to standard methods (Sambrook et al., 1989). Recombinant phages were amplified in liquid medium to purify their DNA (Sambrook et al., 1989).

**Nucleic acid hybridization and sequencing.** Southern hybridization was carried out according to standard procedures (Sambrook et al., 1989) using digoxigenin-labelled probes. Sequencing clones were constructed with the Erase a base kit (Promega) and sequenced by the dideoxynucleotide method using Sequenase 2.0 (Amersham). The DNA sequence was analysed using the Dnastar and Winstar packages. Comparisons of the deduced polypeptides against the databases...
were accomplished using the BLASTP 2.2.1 program (Altschul et al., 1997). Protein alignments were done with the CLUSTAL V algorithm (Winstar).

RESULTS AND DISCUSSION
Cloning and sequence determination of the region upstream of the car gene

A region of DNA adjacent to the cephamycin gene cluster including a total of 11 genes (from ceaS to ORF12; Fig. 2) has been demonstrated to be involved in CA biosynthesis in S. clavuligerus (Jensen et al., 2000; Li et al., 2000). In order to analyse the genomic region upstream of the car gene, a λGEM12 library of S. clavuligerus was screened using as a probe a 600 bp EcoRI fragment containing the 5′ flanking region of the car gene. Three recombinant phages (fALCL1, fALCL2 and fALCL3) were isolated. Since these phages covered only around 9 kb upstream of the car gene (Fig. 2), a further screening was performed with an 800 bp NotI probe located at the right edge of fALCL3. Three new recombinant phages (fALCL6, fALCL7 and fALCL11) were purified covering a region of around 25 kb upstream of the car gene (Fig. 2).

To construct ordered sets of deletion subclones, the NotI fragments of the recombinant phages previously purified were subcloned in pBluescript KS(+) or pBC KS(+), generating the following plasmids: pALCL23 (insert of 12.5 kb), pALCL25 (insert of 12 kb) and pALCL26 (insert of 4.5 kb included in pALCL23) (Fig. 2). A nucleotide sequence of 12162 bp was obtained using pALCL23 and pALCL25 as template. Computer analysis with the Geneplot program, using the codon preference algorithm, revealed the presence of 10 ORFs designated from ORF10 to ORF19 (Fig. 3; Table 1). All of them showed the typical biased codon usage of Streptomyces and a mean G+C content of 72.5 mol%. Whereas from ORF14 to ORF18 were transcribed in the same direction, from ORF10 to ORF13 and ORF19 were oppositely oriented. The arrangement of the genes within the cluster suggests the possibility of three co-transcriptional units (Fig. 3): ORF10-ORF11 (spaced by 5 bp), ORF12-ORF13 (spaced by 0 bp), ORF15-ORF16-ORF17-ORF18 (spaced by 20 bp between ORF16-

![Geneplot analysis revealing the ORFs in the 12162 bp sequenced region. The putative transcriptional units are indicated by arrows. Co-transcription of ORFs 10–11, ORFs 12–13 and ORFs 15–16–17–18 is proposed. The numbering above the ORFs indicates spacing in bp between each pair.](image-url)
Table 1. Putative products and deduced functions of the ORFs

<table>
<thead>
<tr>
<th>ORF</th>
<th>Size (bp)</th>
<th>No. of amino acids</th>
<th>Mol. mass (Da)</th>
<th>Proteins with highest similarity</th>
<th>Accession no.</th>
<th>Similarity index (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF12</td>
<td>1398</td>
<td>466</td>
<td>50607</td>
<td>ORF12 of S. clavuligerus (Li et al., 2000)</td>
<td>AAF85940</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hypothetical protein of M. tuberculosis</td>
<td>AAK48037</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LPQF of M. tuberculosis</td>
<td>F70552</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Secreted protein of M. leprae</td>
<td>CAC30878</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>β-Lactamase of D. radiodurans</td>
<td>B75577</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>β-Lactamase of P. stuartii</td>
<td>AAC05975</td>
<td>26</td>
</tr>
<tr>
<td>ORF13</td>
<td>1020</td>
<td>340</td>
<td>35707</td>
<td>Export pump for cysteine metabolites of E. coli</td>
<td>BAB35563</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Integral membrane protein of S. coelicolor</td>
<td>CAB66295</td>
<td>36</td>
</tr>
<tr>
<td>ORF14</td>
<td>1017</td>
<td>339</td>
<td>36292</td>
<td>Acetyltransferase for tabtoxin resistance of P. syringae</td>
<td>P16966</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hypothetical protein of D. radiodurans</td>
<td>F75594</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hypothetical protein of S. coelicolor</td>
<td>T35933</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phosphinothricin acetyltransferase of D. radiodurans</td>
<td>B75444</td>
<td>21</td>
</tr>
<tr>
<td>ORF15</td>
<td>1686</td>
<td>562</td>
<td>61987</td>
<td>Peptide transport protein of S. clavuligerus (Jensen et al., 2000)</td>
<td>AAF86623</td>
<td>47</td>
</tr>
<tr>
<td>ORF16</td>
<td>1179</td>
<td>393</td>
<td>42628</td>
<td>Hypothetical protein of S. coelicolor</td>
<td>T34757</td>
<td>35</td>
</tr>
<tr>
<td>ORF17</td>
<td>1587</td>
<td>529</td>
<td>57127</td>
<td>Carboxylase of Rhizobium sp.</td>
<td>P55641</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Carboxylase of S. coelicolor</td>
<td>CAB92560</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Carboxylase of M. loti</td>
<td>NP_106830</td>
<td>27</td>
</tr>
<tr>
<td>ORF18</td>
<td>1482</td>
<td>494</td>
<td>53011</td>
<td>PBP type A of S. coelicolor</td>
<td>CAB32353</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PBP type A of S. griseus</td>
<td>AAF61246</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PBP of S. coelicolor</td>
<td>T36716</td>
<td>46</td>
</tr>
<tr>
<td>ORF19</td>
<td>&gt;270</td>
<td>&gt;90</td>
<td>&gt;10117</td>
<td>PBP type 2 of S. coelicolor</td>
<td>CA75389</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PBP of S. coelicolor</td>
<td>CA41600</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PBP type 2 of S. clavuligerus</td>
<td>CAA04971</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PBP type 2 of P. multocida</td>
<td>AAK04008</td>
<td>34</td>
</tr>
</tbody>
</table>

* Estimated with the program MegAlign (Winstar) by the Lipman-Pearson protein alignment method.

ORF17 and 0 bp between the rest). Moreover, two sequences recognized as potential transcriptional terminators with free energies of $-75.6 \text{kcal} (-316.3 \text{kJ}) \text{mol}^{-1}$ and $-58.6 \text{kcal} (-245.2 \text{kJ}) \text{mol}^{-1}$ were identified, with the aid of the mfold program (Mathews et al., 1999), downstream of ORF11 (fd) and ORF15, respectively (Fig. 4). The position of these terminators agrees to the above-proposed organization of the ORFs.

Deduced products of the identified ORFs

The proteins deduced from the previously identified ORFs were used for searching the databases to obtain information about their putative functions. The results are summarized in Table 1.

ORF10 and ORF11. These ORFs correspond to the previously described cyp and fd genes encoding a cytochrome P450 and a ferredoxin, respectively (Jensen et al., 2000; Li et al., 2000). Although the role of this cytochrome/ferredoxin enzymic complex in the biosynthesis of CA is not clear, it could potentially be involved in some of the oxidation steps. Li et al. (2000) suggested a role for this complex in the oxidative reaction between clavaminic acid and clavaldehyde. However, other oxidative steps in the CA pathway, i.e. the transformation of deoxyguanidine proclavaminic acid to guanidine proclavaminic acid, the conversion of proclavaminic acid to dihydroclavaminic acid, and the conversion of dihydroclavaminic acid to clavaminic acid, have been described to be catalysed by the versatile enzyme CAS (Jensen & Paradkar, 1999). Further experiments need to be done to demonstrate the exact role of this oxidative complex in the CA biosynthetic pathway.

ORF12. This encodes a protein of 466 amino acids, with...
a molecular mass of 50607. A protein of 430 residues and 47081 Da was proposed by Li et al. (2000). Using the translation start codon suggested by Geneplot analysis according to the codon bias, ORF12 starts at a GTG triplet located 108 nucleotides upstream of the start point proposed by Li et al. (2000). The difference between the two sequences consists in the lack of a cytosine (CGACGACTCCCATG) 7 bp upstream of the translation start codon in the sequence under accession number AF200819. This means 36 additional residues at the N-terminal end of the protein. The predicted protein showed similarity to two uncharacterized proteins of Mycobacterium tuberculosis (LPQF protein and a hypothetical protein), to a secreted protein of Mycobacterium leprae and to β-lactamases from Deinococcus radiodurans, Providencia stuartii, Streptomyces cacaoi, Streptococcus pyogenes and Pseudomonas aeruginosa. Although the protein includes a highly conserved SDN motif (residues 242–244), which plays a crucial role in the catalytic activity of class A β-lactamases, other residues critical to the catalytic activity of β-lactamases (STFK, EPELN and KGT) are absent.

**ORF13.** The deduced polypeptide presented significant similarity to an export pump for cysteine and other metabolites of the cysteine pathway such as N-acetyl-Ser and O-acetyl-Ser from E. coli and to an integral membrane protein from Streptomyces coelicolor. ORF13 could be involved in the transport of metabolites of the CA pathway, e.g. the N-acyl derivatives of clavaminic acid excreted into the medium (Elson et al., 1988). Whether the possible co-transcription of ORF12-ORF13 has any biological meaning is unclear.

**ORF14.** The deduced protein showed similarity to the following proteins: an acetyltransferase from Pseudomonas syringae responsible for tabtoxin resistance, a hypothetical protein from Deinococcus radiodurans, a hypothetical protein from Streptomyces coelicolor, and a phosphinothricin acetyltransferase from D. radio-
*Xylella fastidiosa* the hypothetical protein SF1447 of rapamycin, from aminoglycoside antibiotics (Zhu et al., 1999) are conferred by acetyltransferases. Since (i) apart from CA, *S. clavuligerus* produces other antibiotics, e.g. cephemycin C (Nagarajan et al., 1971), derivatives of olivamic acid (Brown et al., 1979), holomycin and an antibiotic related to tunicamycin (Kening & Reading, 1979), and (ii) the clustering of biosynthetic and resistance genes for the same antibiotic has been reported in several microorganisms (Fouces et al., 1999), ORF14 could be involved in antibiotic resistance. Whether the product of ORF13, similar to an export pump for N-acetyl-l-serine and O-acetyl-l-serine, could be involved in the transport of the acetylated metabolites generated by the acetyltransferase encoded by ORF14 will be investigated.

**ORF15.** The ATG start codon overlaps the last codon of ORF16. The deduced polypeptide displayed a high similarity index to the product of the ORF7 also located in the CA biosynthetic cluster of *S. clavuligerus*. ORF7 encodes a putative peptide transport protein (Hodgson et al., 1995; Jensen et al., 2000). Moreover, it showed similarity to the oligopeptide-binding lipoprotein encoded by the SC2A11.11 gene of *Streptomyces coelicolor*. The polypeptides encoded by ORF15 and ORF7 share a conserved domain characteristic of bacterial extracellular solute-binding proteins (family 5): ^85LEKGSSEDGRVWYRLREGLY*106 (ORF7) and ^46GEVGEGEGEVEGRTWYRLREGLY*109 (ORF15). The biological significance of the presence of two peptide transport proteins in the biosynthetic cluster is intriguing. The existence of two sets of paralogous genes, encoding a functionally equivalent protein, involved in the early steps of CA biosynthesis has been described (Jensen et al., 2000). While paralogous genes, expressed only in soy-based culture medium, are assumed for ccaS, bls, pab, cas2 and ORF6, individual disruption of ORF7, claR, car and cyp resulted in transformants defective in CA production, even in soy medium. However, all of them were able to produce other clavams (Jensen et al., 2000). The existence of ORF15 opens the question whether it is really a paralogous gene of ORF7 or whether it could be involved in clavam biosynthesis. The lack of CA production by the transformants inactivated in ORF7 described by Jensen et al. (2000) suggests that the second hypothesis may be more reasonable. Nevertheless, gene replacement of ORF15 resulted in loss of CA production, indicating that the protein encoded belongs to the CA biosynthetic pathway (L. M. Lorenzana & P. Liras, unpublished results).

**ORF16.** The product of ORF16 showed slight similarity to the hypothetical protein D, involved in the biosynthesis of rapamycin, from *Streptomyces bygrosoporus* and to the hypothetical protein SF1447 of *Xylella fastidiosa.* The polypeptide encoded by ORF16 includes the conserved hexapeptide ^275LPRTGE*290, which has been proposed to be responsible for a post-translational modification necessary for the proper anchoring of the proteins to the cell wall.

**ORF17.** The deduced polypeptide presented sequence similarity to the Y4RH protein (similar to biotin carboxylases) from *Rhizobium* sp., to a carboxylase from *S. coelicolor* and to a carboxylase and an arginino-succinate lyase from *Mesorhizobium loti*. It includes the carbamoyl phosphate synthase subdomain signature 2 ^234IEANPRF*331}. It is also noteworthy that ORF15, ORF16, ORF17 and ORF18 seem to be co-transcribed.

**Two PBP-encoding genes are present at the right end of the sequenced region**

**ORF18.** This ORF encodes a product with significant similarity to type A penicillin-binding proteins (PBP-A) from different bacteria, including *Streptomyces coelicolor, S. griseus* and *Mycobacterium tuberculosis*. A transpeptidase domain (Pfam00905) and an ATP/GTP binding site motif A (^220GVTVGGKT*245) are conserved in these proteins. The protein includes the conserved motifs of class A β-lactamases 167^KTG169, 220^STKF225, 245^STN256 (instead of SDN) and 428^KTC438, but the consensus EPELN is absent. The name *pbpA* is proposed for ORF18. Two other genes (*pbpR* and *pbp74*) encoding PBPs have been described in the cephamycin cluster of *S. clavuligerus* (Paradkar et al., 1996b; Pérez-Llarena et al., 1998). The similarity index of the protein encoded by *pbpA* to the products of *pcbR* (PBP type B involved in β-lactam resistance) and *pbp74* (high-molecular-mass PBP) was 22% and 35%, respectively, but lower in any case than the above-mentioned similarity to the PBPs from *S. coelicolor* and *S. griseus*.

**ORF19.** The truncated sequence of ORF19 starts at an ATG codon at position 11893 and encodes a protein which presented similarity to type 2 PBPs (PBP-2) from *Streptomyces coelicolor, S. clavuligerus*, *Pasteurella multocida*, *Haemophilus influenzae* and *Rickettsia prowazeki*. The name *pbp2* is proposed for ORF18. The similarity index of the product encoded by *pbp2* to the PBPs encoded by *pcbR* and *pbp74* was 31% and 35%, respectively, being poorer than the above-stated similarity to the PBPs from *S. coelicolor*. The biological meaning of the presence of four PBP-encoding genes (*pbpR, pbp74, pbpA* and *pbp2*) in the contiguous biosynthetic clusters of CA and cephamycin could be related either to β-lactam resistance (*pcbR*) or to cell wall synthesis. Remarkably, *pcbR* is located at one end (left in Fig. 2) of the CA cluster and *pbpA-pbp2* are at the opposite end.

**Targeted disruption of ORF10 and ORF14**

To elucidate whether ORF10 and ORF14 were involved in CA biosynthesis, the plasmids pHZORF10: *aph* (Fig. 3a) and pHZORF14: *aph* (Fig. 3b), were constructed and transformed into *S. clavuligerus*. Both plasmids...
Fig. 5. Restriction maps of the plasmids pHZF10:aph (a) and pHZF14:aph (b) used for inactivation of ORF10 and ORF14, respectively. (c-e) Genetic organization of the CA cluster in (c) the wild-type strain *S. clavuligerus* ATCC 27064, (d) *S. clavuligerus*/pHZF10:aph and (e) *S. clavuligerus*/pHZF14:aph. (f) Southern hybridization analysis of *S. clavuligerus*/pHZF10:aph (NcoI, lane 1, and NruI, lane 2) and the parental untransformed strain (NcoI, lane 3, and NruI, lane 4) using as a probe the apha gene. (g) Southern hybridization analysis of the parental untransformed strain (BstXI, lane 1, and ApaLI, lane 3) and *S. clavuligerus*/pHZF14:aph (BstXI, lane 2, and ApaLI, lane 4) using the apha gene as a probe. Expected patterns were obtained, confirming targeted disruption.

To verify the targeted inactivation, chromosomal DNA of both disrupted transformants was purified, digested with appropriate restriction enzymes and separately hybridized to apha-, ORF10- and ORF14-specific probes. Both disrupted transformants showed hybridizing bands of the expected size with the apha probe: (i) *S. clavuligerus* ORF10:aph gave two bands of 1–6 kb and 1–9 kb when it was digested with NcoI and a single band of
Clavulanic acid gene cluster from *S. clavuligerus*

5.0 kb with *NruI* (Fig. 5d, f), and (ii) *S. clavuligerus*
ORF14: *aph* showed single bands of 3.5 kb with *BstXI* and of 2.5 kb with *ApaLI* (Fig. 5e, g). The wild-type strain did not show any hybridizing signal (Fig. 5f, g). In the wild-type genome, the probe corresponding to ORF10 (6.7 kb *SphI*-NcoI) showed two hybridizing bands of 5.0 kb and 2.0 kb when it was digested with *NruI* and a single band of 3.4 kb with *NcoI*. Additionally, the ORF14-*aph* probe (1.0 kb *ApaLI*-BstEII) gave a band of 2.6 kb with *BstXI* and of 1.6 kb with *ApaLI* in the wild-type strain (not shown). These results confirmed the expected targeted integration patterns.

Once the targeted gene inactivation had been verified, the parental strain and the disrupted transformants were fermented in TSB medium. The ORF10-disrupted transformant failed to produce CA, whereas the inactivation of ORF14 generated a transformant with a remarkably lower production (33 % of the parental strain; Fig. 6a). Although cephamycin C production was not affected by ORF14 gene disruption (data not shown), it was notably increased in the transformant inactivated in ORF10 (Fig. 6b). In this case, a part of the CA precursors could have been channeled toward cephamycin C. Insertional inactivation of ORF10 (*cyp*) and ORF12 was previously described (Li *et al.*, 2000), leading to the complete loss of CA production and demonstrating the involvement of these genes in CA biosynthesis.

Taking into account that ORF12 and ORF13 could share the same transcript, it is likely that ORF13 belongs to CA cluster. Additionally, the reduction (but not loss) of CA production obtained by ORF14 disruption indicates its relationship to the biosynthetic pathway. In this case, the possibility to produce CA could be explained by the presence of a paralogous gene of ORF14 expressed in soy-based medium (TSB), as previously described for other genes of the CA cluster (Jensen *et al.*, 2000). Moreover, both (i) the characterization of ORF15 as similar to ORF7, and (ii) the co-transcription of ORF15-ORF16-ORF17-ORF18, point to their likely association with the CA cluster. Assuming this hypothesis, the CA gene cluster could be located from the *ceaS* gene up to, at least, ORF18.

**Overexpression of ORFs in *S. clavuligerus***

In order to determine whether any of the previously described ORFs has a positive effect on CA production, different plasmid constructions were introduced by transformation into *S. clavuligerus* ATCC 27064, testing CA and cephamycin C production in the recombinant strains.

The DNA fragments isolated from this region were cloned in pULVK99 (7.8 kb), generating the following plasmids (Fig. 2): pULVK-6-cyp (1.9 kb *EcoRI*-NcoI insert including ORF10-ORF11), pULVK0R12 (1.8 kb *SphI*-NruI insert including ORF12), pALCL66 (3.4 kb *BglII*-SalI insert including ORF12-ORF13), pULVK-ORF14 (1.7 kb *NruI*-BstEII insert including ORF14), pALCL36 (6.6 kb *EcoRI*-NotI insert including ORF10-ORF11-ORF12-ORF13-ORF14).

The effect of these genes on CA and cephamycin C production was tested by fermentation of a selected transformant of each plasmid (pULVK99, pULVK-cyp, pULVKORF12, pALCL66, pULVKORF14 and pALCL36) in triplicate flasks of TSB medium. In order to avoid interferences originated by the different rate of growth of the transformants, the production is shown as µg antibiotic per mg DNA (Fig. 7). Those transformants harbouring multiple copies of ORF10–11, ORF12 and ORF14 showed the higher improvements of CA production. At 34 h of fermentation, where the production level was maximum, increases of higher than 100 % were obtained. Curiously, only a slight production increase was obtained with the transformant

**Fig. 6.** CA (a) and cephamycin C (b) production by the parental untransformed strain *S. clavuligerus* ATCC 27064 (●) and transformants disrupted in ORF10 (▲) or ORF14 (▼). Inactivation of ORF10 or ORF14 gave rise to the lack of or a notable drop in CA production, respectively. Cephamycin C production was notably increased in the transformant disrupted in ORF10, but was not affected in the other transformant (this last result not shown). Fermentations were carried out in triplicate flasks. Vertical bars indicate standard deviations.
corresponding to pALCL36 (carrying from ORF10 to ORF14) (Fig. 7, top row). These results point to the involvement of the analysed genes in CA biosynthesis.

Cephamycin C production was higher in the transformant harbouring pULVK99 (without CA genes) (Fig. 7, bottom row), indicating that those transformants containing additional CA genes could have displaced their secondary metabolism toward CA production.

The complete characterization of the CA biosynthetic gene cluster will provide a very useful tool for the improvement of CA production. The development of some antibiotic-producing strains by increasing the copy number of the biosynthetic genes has been reported (Diez et al., 1997). This suggests that transforming CA-producing micro-organisms with these biosynthetic genes would improve CA productivity.

ACKNOWLEDGEMENTS

The authors thank M. Sandoval, J. A. González, P. Merino, M. Mediavilla and M. T. García for their excellent technical assistance. L. M. Lorenzana was supported by a fellowship of the University of León. This work was partially supported by grant 1FD97-1419-CO2-02.

REFERENCES


Pérez-Redondo, R., Rodríguez-García, A., Martín, J. F. & Liras, P. (1999). Deletion of the *pyc* gene blocks clavulanic acid bio-


Received 14 August 2001; revised 12 November 2001; accepted 7 January 2002.