An Evolutionary Relationship between the ColE5-099 and the ColE9-J Plasmids Revealed by Nucleotide Sequencing

By MARK D. CURTIS, RICHARD JAMES* AND ALAN CODDINGTON
Molecular Biology Sector, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

(Received 22 May 1989; revised 27 June 1989; accepted 6 July 1989)

The nucleotide sequence of a 1124 bp fragment of the ColE5-099 plasmid which encodes colicin E5 immunity, a lys gene involved in colicin release from the host cell, and the 3' end of the colicin E5 structural gene has been determined. Open reading frames corresponding to the three genes have been located by analogy with similar sequences from other E colicin plasmids. The location of these open reading frames corresponds with the position of the genes as determined by subcloning and transposon mutagenesis. The amino acid sequence of the carboxy-terminal 107 amino acid residues of the colicin E5 gene shows no homology with any other E colicin, suggesting a different mode of action in killing sensitive cells. A comparison of the nucleotide sequence of this region of the ColE5-099 plasmid with that of the equivalent region of the ColE9-J plasmid suggests a close evolutionary relationship between these two plasmids.

INTRODUCTION

Colicins are plasmid-encoded protein antibiotics (bacteriocins) produced by strains of Escherichia coli. They are classified into groups on the basis of the cell surface receptor to which they bind, e.g. the E group colicins all bind to the product of the chromosomal \textit{btuB} gene of \textit{E. coli} (Di Masi et al., 1973). Each E colicin plasmid codes for the production of a specific immunity protein, which protects colicin-producing cells against colicin of the same immunity type. The E group colicins have been subdivided into colicins E1 to E9 on the basis of immunity tests (Watson \textit{et al.}, 1981; Cooper & James, 1984). The catalytic activity of several of the E colicins (which is located in the carboxy-terminal trypsin fragment) has been identified, and falls into one of at least three classes; e.g. colicin E1 is a membrane depolarizing agent, colicin E2 is a DNAase, whilst colicin E3 is an RNAase (Konisky, 1982). \textit{E. coli} K12 strains carrying an E colicin plasmid are sensitive to mitomycin C (MC) compared to plasmid-free isogenic strains, due to the 'SOS' induction of a plasmid-encoded gene (\textit{lys}) which is involved in colicin release from the producing cell, and which causes partial cell lysis upon strong induction.

The discovery of two new members of the E group colicins, E8 and E9 (Cooper & James, 1984), was particularly fortuitous as the use of \textit{E. coli} strains carrying these plasmids in immunity tests led to the novel finding that some E colicin plasmids code for two different immunity genes. The ColE3-CA38 plasmid was demonstrated to encode both E3 immunity, as expected, and E8 immunity (Chak & James, 1984), whilst the ColE9-J plasmid was demonstrated to encode both E9 immunity and E5 immunity (Chak & James, 1986). The nucleotide sequence of a 1500 bp fragment of the ColE9-J plasmid, including both the E colicin immunity genes, was determined previously (James \textit{et al.}, 1987). The colicin E9 \textit{imm} gene is 69% homologous to the colicin E2 \textit{imm} gene and to the colicin E8 \textit{imm} gene, at both the nucleotide and amino acid levels, but is only 37% homologous with the colicin E3 \textit{imm} gene at the nucleotide level and 7% at the amino acid level. The complete nucleotide sequence of the colicin E9 structural gene has recently been

Abbreviation: MC, mitomycin C.
sequence (Fig. 2). Inspection of the sequence data obtained with pKC19 and pMC33 reveals a single open reading frame which terminates 3' to the EcoRI site at base 518. This is consistent with the location of the colicin E5 structural gene inferred from the subcloning experiments (Fig. 1) and is the only significant open reading frame in this region of the ColE5-099 plasmid.

Translation of the nucleotide sequence of the 3' end of the colicin E5 gene reveals a region of 68 amino acid residues which is almost completely homologous to the equivalent region of all E. coli colicins which have been sequenced (Fig. 3). The published sequences of colicins E2, E3 and E9 in fact reveal almost complete amino acid sequence homology from the amino-terminus of these three proteins to the 3' end of the 68 amino acid sequence shown in Fig. 3. Therefore, it was not considered worthwhile to obtain the nucleotide sequence of the rest of the colicin E5 structural gene. This homology presumably reflects the suggested role of this region of the protein in binding to the BtuB receptor protein on the surface of susceptible E. coli cells. The remaining 107 amino acid residues up to the carboxy-terminus of the colicin E5 gene show no
**Sequence of the pColE5 col, imm and lys genes**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>20</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>60</td>
</tr>
</tbody>
</table>

**Fig. 3.** Amino acid sequences at the carboxy-terminus of E. coli colicin genes. The amino acid sequence of five different E. coli colicins is indicated, using single-letter code. Padding spaces are included to optimize the alignment. The region of almost complete amino acid sequence homology is boxed. The sequences of E2 and E3 are taken from the EMBL computer database, E8 is from Uchimara & Lau (1987), whilst that of E9 is from Eaton & James (1989).

**Fig. 4.** Possible evolutionary origin of the ColE9-J plasmid. It is suggested that the upper DNA fragment (of unknown origin) carrying a truncated E9 lys gene, the E9 imm gene and at least 73 amino acid residues at the carboxy-terminus of the colicin E9 gene has been inserted into the ColE9-099 plasmid, at a site some 48 amino acid residues from the carboxy-terminus of the colicin E5 structural gene, to give an in-frame fusion which allows production of colicin E9.

significant homology to any other E. coli colicin which has been sequenced. From physiological studies, which indicate a rapid inhibition of protein synthesis in colicin E5 treated *E. coli* cells (Mock & Pugsley, 1982), it has been suggested that colicin E5 is an RNAase type E. coli colicin (Toba et al., 1988). From the sequencing data obtained in this study, it can be concluded that there is no homology between the carboxy-terminus of colicin E5 and colicin E3, or the closely related bacteriocin cloacin DF13, the only other RNAase-type colicins which have been identified. A computer homology search of both the NBRF and Swisprot protein sequence databases with this 107 amino-acid sequence of colicin E5 has been carried out and no significant homology to any protein has been found. The evolutionary origin and the precise mechanism of killing by colicin E5 has thus to be established. It is interesting that in the other E. coli colicin plasmids in which two immunity genes are found, ColE3-CA38 and ColE6-CT14, the immunities are also against a DNAase and an RNAase type E. coli colicin (Uchimara & Lau, 1987; Toba et al., 1988).

**Evolutionary relationship**

The nucleotide sequence of the colicin E5 immunity gene from the ColE5-099 plasmid is virtually identical to that of the same gene from the ColE9-J plasmid (Fig. 2; James et al., 1987). The intergenic sequence between the 3' end of this immunity gene and the 5' end of the lys gene,
together with the nucleotide sequence of the lys genes from these two plasmids is also virtually identical. An evolutionary relationship between the ColE5-099 and the ColE9-J plasmid is suggested from a comparison of the nucleotide sequences. The nucleotide sequence which encodes the carboxy-terminal 48 amino acid residues of colicin E5 is also found in the ColE9-J plasmid in an identical location, 5' to the E5 immunity gene. It appears likely that some genetic exchange has occurred in which a segment of DNA which encodes at least the last 73 amino acid residues of the colicin E9 structural gene, the E9 imm gene and the E9 lys gene, has been transferred into the colicin E5 structural gene at a site 48 amino acid residues from the carboxy-terminus (Fig. 4). This event was presumably accompanied by an in-frame fusion of the carboxy-terminus of the colicin E9 structural gene and a disruption of the original ColE9 lys gene, resulting in a truncated gene (James et al., 1987). Whether the presence of two immunity genes on a single plasmid offers some advantage to the host cell, or whether such plasmids are simply intermediates in the evolution of novel E colicins, is an intriguing question. Further sequence data with colicin E4 and E7 may shed further light on the evolution of this family of plasmids.

We are grateful to Jill Debbage for excellent technical assistance. This work was supported by an SERC project grant. M.D.C. is the recipient of an SERC studentship.

REFERENCES


