COMPARATIVE ANALYSIS OF THE REPLICATION REGIONS OF INCB, INCK, AND INCZ PLASMIDS

J. PRASZKIER, T. WEI,† K. SIEMERING, AND J. PITTARD*†

Department of Microbiology, University of Melbourne, Parkville, Victoria 3052, Australia

Received 12 September 1990/Accepted 31 January 1991

Minireplicons from the I-complex plasmids R387 (IncK) and pIE545 (IncZ) were constructed, and the nucleotide sequences of their replication regions were compared with that of the B plasmid, pMU720. The coding sequence of the putative replication protein, RepA, of each plasmid was located. RepA of K and B plasmids were homologous, whereas RepA of Z resembled RepA1 of FII plasmid. Sequences upstream of RepA were conserved in the three I-complex plasmids. Group B and Z plasmids were incompatible.

Plasmids belonging to incompatibility groups I1, I1, I1, B, K, and Z have been placed in the I complex on the basis of the morphological and serological similarities of their pili (2, 3). Subsequently, it was found that the incompatibility determinants of I1, I1, and B plasmids are small countertranscript RNA molecules (19). These RNAs (RNAI) are thought to regulate plasmid copy number by inhibiting the translation of the target RNA (RNAII) whose product, the RepA protein, is essential for replication (21). Thus, the mechanism of replication control of this group of plasmids is similar to that of IncFII plasmids (4, 16, 17, 23, 26, 29) and the RepFIC replicons from IncFII plasmids (24, 28). The replicons of the incompatibility group I1, I1, and B plasmids are closely related, with the sequences of their countertranscript RNAs showing >80% homology (19). Furthermore, this relationship appears to extend to group K plasmids, since DNA probes from the replication regions of group I1 and B plasmids cross-hybridize not only with each other but also with a group K plasmid (6). This raises the possibility that replicons of all the I-complex plasmids are closely related.

In this paper, we present sequence data which confirm the close relationship between group B and K plasmids and show that this relationship extends to a group Z plasmid. The replication region of the latter plasmid shares extensive homology with the replication regions of group B and K plasmids, but this homology does not extend into the coding region of the repA gene.

Construction of group K and Z minireplicons and the mapping of their inc loci. The IncK miniplasmid (pMU2209 [Fig. 1]) was constructed by ligating the products of a partial PstI digestion of R387 (12) DNA to a 5-kb PstI fragment encoding the enzymes for galactose fermentation (the Gal fragment [7]). This yielded a chloramphenicol-resistant replicon, which was made chloramphenicol sensitive by partial digestion with Sau3A and religation. The IncZ miniplasmid (pMU2200 [Fig. 1]) was made from pIE545 (27) by partial digestion of its DNA with PstI, followed by ligation to the Gal fragment.

Tests (7) of pBR322 derivatives carrying DNA fragments from the two minireplicons showed that the 1- and 1.5-kb PstI fragments of pMU2200 and pMU2209, respectively, expressed incompatibility against their parental plasmids. Deletion derivatives of these PstI fragments were generated with exonuclease III (13). Their analysis showed that the inc gene of pMU2200 was located within a 250-bp region and the inc gene of pMU2209 was located within a 300-bp region, immediately downstream of the BamHI site (Fig. 1).

Introduction of the pBR322 derivatives carrying these inc genes into strains carrying B, K, I1, or Z mimiplasmids showed that the Z plasmid was incompatible with B but compatible with K and I1 plasmids (Table 1). The K plasmid was compatible with the other 3 plasmids.

Tschäpe and Tietze (27), who first described the new incompatibility group which they named IncZ, also observed strong interaction between Z and B plasmids, but they believed this to be due to "dislodgement" rather than incompatibility. Our data suggest that at least on the basis of their incompatibility reactivities, IncB and IncZ plasmids belong in the same group.

Nucleotide sequence analysis. The nucleotide sequences of single stranded and double-stranded templates were deter-

* Corresponding author.
† Present address: Department of Enteric Infectious Disease Control, Anti-epidemic and Hygiene Center of Henan Province, 47 St Wei Wulu, Zhengzhou 45 0003, Henan, People’s Republic of China.

FIG. 1. Restriction maps of the minireplicons derived from pIE545 (pMU2200) and R387 (pMU2209). The 5-kb Gal fragment (7) is not shown. INC, Incompatibility region; B, BamHI; P, PstI; S, Sau3A. Arrows indicate the region whose sequence is presented in Fig. 2.
FIG. 2. Alignment, generated by using the CLUSTAL program (14, 15), of the nucleotide sequences of the replication regions of IncB (pMU720), IncK (pMU2209), and IncZ (pMU2200) plasmids. Nucleotides common in at least two of the sequences are capitalized. The nucleotide number is indicated to the right of each line. The $\sim$35 and $\sim$10 regions of the possible promoter sequences of RNAI (transcribed in the leftward direction) and RNAII (transcribed in the rightward direction) are marked above the line. In pMU720, RNAI initiates at nucleotide 420 or 423, whereas RNAII initiates at nucleotide 631 or 629 and terminates at nucleotide 561 or 559 (24). The start (GTG) and stop codons of the putative RepA proteins are indicated by 3 asterisks, and those of the putative RepB proteins are indicated by single asterisks. DnaA refers to the DnaA boxes. The alignment of the sequences of R100, pMU2200, and pMU720 in the region where the switches between pMU720-like and R100-like sequences occurs in pMU2200 is shown. The sequence of R100 is shown in small letters with nucleotides common to R100 and pMU2200 indicated by a line underneath the latter sequence. The sequence of R100 from nucleotide 1293 to nucleotide 2330, which is not shown, is highly homologous to the corresponding sequence of pMU2200.

![Diagram of stem-and-loop structures](image)

Energy = -23.7 kcal
Energy = -30.4 kcal
Energy = -26.4 kcal
Energy = -25.7 kcal

FIG. 3. Sequences of the stem-and-loop structures predicted for RNAI molecules of 1-complex plasmids by using the computer program of Zuker and Stiegler (30) with the energy files based on the thermodynamic parameters of Freier et al. (8) (1 kcal = 4,184 J).
predicted to have an 8-base loop in which 7 bases are the same as those in the loops of B and Z countertranscript RNAs. The stems of the four RNA molecules show differences both in length and sequence. The difference between the loop of K RNAI and the loops of the other three plasmids is consistent with the incompatibility data (Table 1) since it is the loop which is thought to be the site of the interaction between RNAI and its target. Compatibility between plasmids such as Z, B, and I1 where the loops of RNAI are identical has been noted previously, underscoring the complexity of the incompatibility interactions (19, 28).

There is only one long open reading frame (ORF1) in each sequence and no other long open reading frame in the entire plasmid replicon (data not shown). This, together with the finding that in pMU720 the translation of ORF1 is negatively regulated, in trans, by RNAI (21) makes it a likely candidate for an essential replication protein, RepA. Although in pMU720, ORF1 has two potential GTG start codons, at positions 677 and 722 (21), the first one of these does not appear to be active in vivo, because fusion of the first 718 bp of pMU720 DNA with lac so that ORF1 is in phase with the lacZ gene but the second GTG codon is absent (pMU673) resulted in no detectable β-galactosidase activity (data not shown). ORF1 of pMU2200 is probably also encoding a RepA protein, since the DNA sequence of this plasmid is identical to that of pMU720 between bases 624 and 924, i.e., a region starting 56 bp upstream of the first GTG of ORF1 and ending 199 bp downstream of its second GTG. The GTG start codon of ORF1 of pMU2200 corresponds to the second GTG of ORF1 of pMU720, and the 20 bases immediately preceding this codon are identical in all three plasmids. Within these 20 bases lies a sequence, TAAGCGA, which may act as a ribosome binding site since it has six bases complementary to the 3' end of the 16S ribosomal RNA. Preceding and overlapping the start of ORF1 in each plasmid is an open reading frame of 29 codons. This sequence is completely conserved in B and K as well as in I1 (11, 19) and I2 (19) plasmids, but there are five differences at the amino acid level in the Z plasmid. Gene fusion studies indicate that in pMU720 the translation of this small open reading frame (repB) is a prerequisite for the translation of ORF1 (data not shown). Such translational coupling has been also shown in the I1 plasmid ColII-P9 (10).

The putative RepA proteins of pMU720 and pMU2209 (Fig. 4) have calculated molecular weights of 40,421 and 40,266, respectively. They are highly homologous not only with each other (sharing 325 of their 343 amino acids) but also with the RepZ protein of the I1 plasmid ColII-P9 (11). By contrast, the putative RepA protein of the group Z plasmid, pMU2200, has very little homology to the corresponding protein of the other I-complex plasmids, showing instead close sequence similarity to the RepA protein of the FII plasmid R100 (22), with which it shares 250 out of its 290 amino acids (Fig. 4).

The switch from pMU720-like in R100-like sequence in pMU2200 occurs at the beginning of ORF1 and is preceded by a region showing only a low level of homology with each of these plasmids (Fig. 2). The sequence upstream of this region shows high homology with pMU720 and low homology with R100. The switch back to pMU720-like sequence occurs in a region where the three plasmids share a 32-bp stretch of high-level homology (Fig. 2). By contrast, the coding sequences of the Rep proteins of the RepFIC and FII replicons, which are dissimilar to each other, are flanked by long regions of high homology (24).

<table>
<thead>
<tr>
<th>TABLE 1. Incompatibility tests between multicopy derivatives encoding inclI, inclB, inclK, and inclZ genes and the corresponding Gal minireplicons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incoming plasmid (pBR322 derivative)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>pMU1532 (I)</td>
</tr>
<tr>
<td>pMU749 (B)</td>
</tr>
<tr>
<td>pMU2210 (K)</td>
</tr>
<tr>
<td>pMU2202 (Z)</td>
</tr>
</tbody>
</table>

* 1, 100% loss of resident plasmid; WI, at least 50% of colonies showing loss or partial loss (sectoring) of resident plasmid; C, 100% coexistence of plasmids.

FIG. 4. Comparison of the putative RepA proteins of pMU720, pMU2209, and pMU2200 and the RepA1 protein of R100 (22). The predicted amino acid sequences of the four Rep proteins were aligned by using the CLUSTAL program (14, 15). Residues common to pMU720 and pMU2209 on the one hand, and to pMU2200 and R100 on the other hand, are indicated by a line between the appropriate pairs of sequences. Residues common to all four sequences are marked with an asterisk, and conservative substitutions are indicated by a dot.
other mechanism may be required to explain the origin of the Z sequence.

The GenBank accession numbers of the nucleotide sequences of pMU720, pMU2209, and pMU2200 are M38521, M38522, and M38523, respectively.

We thank W. Pemberton, S. Potter, and D. L. Presidente for technical assistance.

This project was supported by a grant from the National Health and Medical Research Council.

REFERENCES