Identification and Characterization of IS1381, a New Insertion Sequence in *Streptococcus pneumoniae*

ANA R. SÁNCHEZ-BEATO, ERNESTO GARCÍA, RUBENS LÓPEZ, AND JOSÉ L. GARCÍA*

Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, 28006 Madrid, Spain

Received 11 November 1996/Accepted 24 January 1997

A new insertion sequence (IS1381) was identified in the genome of *Streptococcus pneumoniae* R6 as an 846-bp segment containing 20-bp terminal inverted repeats and flanked by 7-bp direct repeats. The three sequenced copies of this element have two overlapping open reading frame (ORF) genes named *orfA* and *orfB*. However, significant variations between individual copies were found, suggesting that inactivating mutations have occurred in an original single ORF. Accordingly, the consensus IS1381 element derived from the comparison of the three available copies should contain a single ORF sufficient to encode a basic protein of 267 amino acids which exhibited high similarity to the putative transposases of ISL2 from *Lactobacillus helveticus* and of IS702 from the cyanobacterium *Calothrix* sp. strain PCC 7601. A minimum of five to seven copies were detected by hybridization experiments in the R6 genome. In remarkable contrast with the two previously reported pneumococcal insertion sequences, several copies of IS1381 have been detected in all of the clinical isolates tested so far. Interestingly, *Streptococcus oralis* NCTC 11427 (type strain), a close relative of pneumococcus, does not contain this element, but its occurrence in the type strain of *Streptococcus mitis* (NCTC 12261) suggests that this species has exchanged DNA with *S. pneumoniae* directly or through an intermediate species.

Streptococcus pneumoniae is an important human pathogen, causing invasive diseases such as pneumonia, meningitis, and bacteremia. Morbidity and mortality from pneumococcal infections remains high, even in regions where efficient antibacterial therapy is freely available. For most bacterial pathogens, virulence is a multifunctional process requiring two general classes of determinants, one represented by genes that participate in physiological processes necessary for survival which might be common to pathogenic and nonpathogenic organisms, and the other represented by specific virulence genes that are rarely detected in nonpathogenic organisms (7). Although the method used by bacteria to acquire the virulence genes is still an open question, it has been proposed that mobile DNA elements can contribute to the acquisition of such pathogenic phenotypes by promoting mutations and rearrangements within the chromosome as well as by favoring the horizontal exchange of virulence genes (7). In this sense, a small number of transposable elements have been characterized in *S. pneumoniae* so far: i.e., four conjugative transposons, Tn1545 (8), Tn3701 (4), and Tn5251 and Tn5252 that form the composite element Tn5253 (17). The role of these elements in the natural population dynamics of *S. pneumoniae* is entirely unknown, but it has been documented that they can contribute in many ways to its variability, particularly in *S. pneumoniae*, for which only a few cryptic plasmids have been reported (see reference 19 and references therein). Furthermore, a wide recognition of common insertion sequences (ISs) will be of great value in the design of new genetic tools in pneumococcus, since only two ISs, namely IS1202 (15) and IS1167 (23), and a 3'-terminal portion of a transposase-coding gene corresponding to a putative IS named IS-SP (16) have been described so far in this species.

In the course of a survey to identify genes encoding choline-binding domains (ChBDs) (6), we isolated a segment of pneu-

mococcal DNA that exhibited all of the hallmarks of a bacterial IS. One copy of this element was inserted upstream of the *pcpA* gene, which encodes a novel choline-binding protein. Here, we describe the sequence and characterization of IS1381, a new IS in *S. pneumoniae*.

Construction and screening of genomic libraries. ChBDs have been found in different pneumococcal proteins that bind to choline residues present in the teichoic acids, e.g., the major pneumococcal LytA autolytic amidase (13), the cell wall lytic enzymes of several pneumococcal bacteriophages (13), and the PspA protein (21). On the other hand, we had found a similar ChBD in the CspA protein of *Clostridium beijerinckii* NCIB 8052 (formerly *Clostridium acetobutylicum*), a strain that also contains choline in its cell wall (18). To investigate the origin of the CspA protein, we carried out several Southern blot analyses of pneumococcal DNA by using the ChBD-coding region of *cspA* of *C. beijerinckii* as a probe, and we found that at low stringency, this fragment of *cspA* hybridized with several pneumococcal DNA fragments different from those containing the *lytA* and *pspA* genes (data not shown). Hence, to isolate these fragments, a genomic library of *S. pneumoniae* R6 was constructed in *Escherichia coli* TG1 (Amersham) by digestion of the chromosomal DNA of R6 with *HincII* and ligation of the resulting fragments to *SmaI*-digested pUC18. Using a colony hybridization procedure and the *PvuII*-*SspI* fragment of pKDE543 which encodes the ChBD of CspA from *C. beijerinckii* (18) as the screening probe, we isolated the plasmid pEK1 (Fig. 1). Restriction and Southern blot analyses revealed that pEK1 contained an *EcoRI* fragment which was subcloned to generate pEK2, a plasmid that also reacted with the probe. In addition, we surprisingly observed that the 2-kb *HindIII* fragment subcloned to give plasmid pEK6 hybridized at high stringency (1) with multiple fragments of the pneumococcal genome, suggesting the presence of a highly repeated DNA sequence which was studied further.

The analysis of the nucleotide sequence of the 2-kb *HindIII* fragment showed the existence of two overlapping open reading frames (ORFs), ORFA and ORFB, flanked by two 20-bp imperfect inverted repeats (IRs [IRL and IRR]) and two short

* Corresponding author. Mailing address: Departamento de Microbiología Molecular, Centro de Investigaciones Biológicas (CSIC), Velázquez 144, 28006 Madrid, Spain. Phone: (341) 561-1800. Fax: (341) 562-7518. E-mail: cibj116@pinar1.csic.es.

FIG. 1. Restriction map of a region of the *S. pneumoniae* R6 chromosome containing IS1381. The positions of the pneumococcal DNA inserts on pEK1, pEK2, and pEK6 are shown. The location of IS1381 corresponds to that of *orf3* (*orfA* plus *orfB*). The gene *pcpA* encodes a choline-binding protein. The white bar between *orfA* and *orfB* represents the overlapping sequence. Restriction sites are indicated as follows: A, *Ava*I; E, *Eco*RI; Ev, *Eco*RV; H, *Hind*III; HII, *Hinc*II; N, *Nco*I; Ss, *Ssp*I; and X, *Xba*I.

(7-bp) imperfect direct repeats (DRs [DRL and DRR]) (Fig. 1 and 2). The pI values corresponding to the deduced amino acid sequences of *orfA* and *orfB* were 10.3 and 11.2, respectively, which are characteristic of IS transposases (5). Furthermore, using the GCG Program of the University of Wisconsin Molecular Biology Package (WIMP) to search for possible homologies in the GenBank database, we observed that the proteins encoded by *orfA* and *orfB* display significant similarities to the transposases of two ISs, i.e., ISL2 from *Lactobacillus helveticus* (23) and IS702 from the cyanobacterium *Calothrix* sp. strain PCC 7601 (14) (Fig. 3). All of these features strongly suggested that the pneumococcal DNA fragment under study encoded a new IS element, which was named IS1381 according to the nomenclature of transposable elements (2). Although *orfA* extends upstream of the IRL sequence, we favor the hypothesis that the putative transposase might start at the ATG codon located 81 bp downstream of IRL, and therefore, *orfA* would encode a protein with a size of 103 amino acids (Fig. 2). On the other hand, *orfB* would start at an ATG codon located 26 bp downstream of the TAA stop codon of *orfA* and would encode a protein with a size of 129 amino acids (Fig. 2). Nonconsensus ribosome-binding site (RBS) sequences were found at a reasonable distance from both translation initiation sites. The polypeptides encoded by *orfA* and *orfB* matched, respectively, the N- and C-terminal regions of the transposases from ISL2 or IS702 (Fig. 3), suggesting that a +1 frameshift could in principle lead to a fusion protein or that the sequence suffered some deletion and/or insertion mutations. In this sense, although it has been well established that several ISs contain transposases that are synthesized by a -1 frameshift, transposases translated by a +1 frameshift have not been reported. Moreover, the potential CUU-UGA frameshift site found in other +1 frameshifts was not observed in this case (20). Interestingly, a motif, AAAAAAG, followed by a hairpin loop containing the stop codon of *orfA* and the start codon of *orfB* was found (Fig. 2). Nevertheless, this arrangement does not exactly match with the "slippery sequence" A-AAA-AAG, which has been shown to be one of the most efficient heptanucleotide windows to authorize a -1 frameshifting (3). On the other hand, it has been shown that only one out of seven copies of the IS1221 from *Mycoplasma hyorhinis* contains a single, long ORF, whereas the others have different mutations rendering truncated transposases (22). Hence, the observation of overlapping frames does not necessarily indicate the existence of a frameshifting regulatory mechanism, since the shifting might have resulted because of nucleotide deletions or insertions rendering inactive transposases.

To ascertain whether the existence of two overlapping frames can be ascribed to a putative +1 frameshifting regulatory mechanism or to an inactivating mutation, we cloned and sequenced two additional copies of IS1381 (Fig. 2). The copy contained in plasmid pIS2 was isolated from a *Hind*III pneu-

mococcal genomic library of strain R6 constructed in pUC18 and screened with the 0.5-kb *Eco*RV-*Ava*I fragment of plasmid pEK6 as a probe (Fig. 1). We could determine the precise extent of the IR sequences which provided the information needed to synthesize the primer OIS1 (5'-CGGAATTCAAT AGCTTTCTGCGAAAC-3') corresponding to a consensus IR sequence. This primer was used to amplify by PCR all of the copies of IS1381 contained in the genome of R6. The resulting fragments (about 900 bp) containing the IS1381 sequences were purified with GeneClean (Bio 101, Inc., La Jolla, Calif.), digested with *Eco*RI, and ligated to *Eco*RI-digested pUC18. Several recombinant plasmids were isolated, and the copy contained in one of them, named pIS3, was sequenced (Fig. 2).

The IS1381 copies of plasmids pIS2 and pIS3 exhibited significant nucleotide differences between them and with respect to the first isolated copy of pEK6 and also showed truncated reading frames, although at different positions. Comparison of the three sequences revealed that the IS1381 copy from plasmid pEK6 had suffered a 22-bp deletion just downstream of IRL. This sequence extends the N-terminal amino acid sequence of the polypeptide encoded by *orfA* that perfectly matches the N-terminal sequences of the ISL2 and IS720 transposases (Fig. 2 and 3).

It is worth noting that the high similarity (59% identity) found between the transposases of the IS1381 and ISL2 elements is also found at the nucleotide level (64% identity). Both elements have essentially the same nucleotide length and IR sequences. Interestingly, the IR sequences of IS1381 are 4 bp (CAAT) longer than the corresponding sequence of ISL2, suggesting that the IRs of IS1381 have apparently embodied the proposed target sequence (AAT) of ISL2. However, one of the IS1381 copies contains the AAT sequence in both ends of the IRs in the same position found in ISL2, suggesting that the IRs and the target sequences could be identical in both ISs. Nevertheless, it must be pointed out that the analysis of the DRs and IRs does not always allow determination of the specific target site, since it has been demonstrated that the sequence that is duplicated can vary between two insertion events for a given IS (5).

All of the observations given above suggest that the three IS1381 copies sequenced here contain different mutations rendering inactive transposases. However, a comparison with the transposases of ISL2 and IS720 allows us to propose a consensus sequence for the IS1381 transposase, which should be essentially composed of the sequences *orfA* and *orfB* and which should be equivalent to the product predicted for a +1 translational frameshift at the last codon of *orfA* (Fig. 2). The consensus transposase sequence starts at an ATG codon located 23 bp downstream of IRL and stops just in the first nucleotide of IRR, as reported for the transposase of ISL2. The consensus IS1381 transposase encodes a protein of 267 residues with a high proportion of basic amino acids (19.4%

```

DRL
----->
c1 CTGGGAAACCTACTCAGCATTTCGAAATACAAATGGCGGAAAAAATTCGCGGATAGATGAAAAAAGATTGACACCTACCAAAGAGT
c2 TGAATCTT*G*GA*ATTTA**TT*CC*GA**AATA*CACC*CGTTCA*AATTC*TCTAAT*GCCTCAGC*TCA*TA*T*TT*CTTTC

IRL
----->
c1 CAATAGACTCCCTCGAAACAAAA-----TTATGAAGCAACTAAACAACCAACTGATGCACGATTAAAGCGTC
c2 *****T*****TATGGTATAGTAGTTCTATGAA*G*****T*****T*****
c3 <--- primer --->***TATGGTATAGTAGTTCTATGAA*G*****G*****T*****
m1 TT*****A**C*A**T*****T*****T*G*****TT*****A*****C*
cs M N D E A T K Q L T D A R F K R

orfa ---> M L A V L K T A Y Q L K H A K G G R K
c1 TTGTGGTGTTCACGCGCAGCACTTTTGAAGAGATGTTAGCTGTATTAAAAACAGCTTATCAACTTAAACACGCAAAAGGTGGACGAAAAAC
c2 *****G*****G*****
c3 *****A*****A*****
m2 *****C*****T*****A*A**T*****T*****A**TC*
cs L V G V Q R T T F E E M L A V L K T A Y Q L K H A K G G R K

P K L S L E A F L W P L F N M C E N I E L M N K L R L I F G
c1 CTAAATTAAGCCTAGAAG-CCTTCTTATGCGCACTCTTCAATATGTCGAGAATATCGAACTTATGAACAAATGCGGGCTGATTTTGGT
c2 *****A*****T*****G*****
c3 *****A**T*****G*****
m1 *****A*****A**AA**C**C*****T*****C*****CT**A**C**C*****T**A*****G**-**C
cs P K L S L E D L L M A T L Q Y V R E Y R T Y E E I A A D F G

I H E S N L I R R S Q W V E V T L V Q S G V T I S R T P L S
c1 ATTCACGAAAGCAACTTAATCCGTCGAGCGCAATGGGTGAAGTAACCTCTGTTCAAAGTGGTGTACGATTCAAGAACTCCTCTCAGT
c2 *****C*****G*****
c3 **C*****G*****
m1 *****A*****T*****CC*****A*****CT**A*****A*****T**T**
cs I H E S N L I R R S Q W V E V T L V Q S G V T I S R T P L S

orfb ---> k s i a l k k e l a n n s q k k k
S E D T V M I I D A T E V K I N R P K K R I S E
c1 TCTGAGGACACGGTAATGATTATTGATGCGACGGAAGTAAAAATCAATCGCCCTAAAAAAGAAATTAGCGAATAATTCGTGTAAGAAAGAA
c2 *****G*****
c3 *****TC*****
m1 G*****T*****G**TG*G---*****A**A**G*****TC**C**C**T*****T*****
cs S E D T V M I I D A T E V K I N R P K K E L A N D S G K K K

f h a M K A Q A I V T S Q G R I V S L D I A V N Y C H D M K
c1 ATTTACCGCTATGAAGGCTCAAGCGATTGTCAAGTCAAGGAGAAATTGTTCTTGGATATCGCTGTGAACCTATAGTCATGATATGAA
c2 *****A*****
c3 **GC*****A*****T*****
m1 **G**T*****G*****C*****T**A*****T*****
cs F H A M K A Q A I V T S Q G R I V S L D I T V N Y C H D M K

L F K M S R R N I E Q A G K I L A D S G Y Q G L M K I Y P Q
c1 GTTGTCAAAATGAGTCGTAGAAATATCGAACAGCTGGTAAAACTTGGCTGACAGTGGTTATCAAGGCTCATGAAGATATATCCTCA
c2 *****C*****
c3 *****C*****G*****
m2 *****C*****G*****A*****A*****G**T**A**
cs L F K M S R R N I G Q A G K I L A D S G Y Q G L M K I Y P Q

V Q T P R K S S K L K P L T A E N K A Y N H A L S K E R S K
c1 AGTACAACTCCAGTAAATCCAGCAAACTCAAGCCGCTAACAGCTGAAAAATAAGCCTATAACCTGCGCTATCTAAGGAAAGAAAGCAA
c2 **C*****A*****G*****G*****G*****
c3 **C*****A*****T**G*****G*****G*****
m1 **CG***T**GA*G*****A*****T**T**A*****TCT**G*****T**A**T**T*****A**G**T**
cs A Q T P R K S S K L K P L T V E D K A C N H A L S K E R S K

V E N I F A K V K T F K I F S T T Y R N H R K R F G L R M N
c1 GGTGAGAACATCTTTGCCAAAGTAAAAACGTTAAAAATATTTCAACAACCTATCGAAATCATCGTAAACGCTTCGGAATACGAATGAA
c2 *****G*****A*****
c3 *****G*****
m2 *****T**T**G**T*****G*****GA*****A**T*****
cs V E N I F A K V K T F K M F S T T Y Q N H R K R F G L R M N

L S A G I I N H E L G F <----->
c1 TTTGAGTGCTGGTATTATCAATCATGAAGTCTAGTTTTCGAGGAAGTCTATTGAAATACTAAGCCTACACAGACTGATTTCAA
c2 *****T*****A*****A**T**T**C*AG**GTGAGTT*ATTA*GATT
c3 *****<--- primer --->
m1 *****T*****A*****CTG*****A*****
cs L S A G I I N H E L G F

```

FIG. 2. Nucleotide sequence of individual copies of IS1381. The three isolated pneumococcal copies of IS1381, c1, c2, and c3, correspond to those of the plasmids pEK5, pIS2, and pIS3, respectively, whereas m1 corresponds to the sequence of the copy isolated by PCR from *S. mitis*. Nucleotides identical to the c1 copy are represented by asterisks, whereas gaps are indicated by dashes. The amino acid translation of *orfA* and *orfB* corresponding to the c1 copy is represented by the single-letter code above the c1 nucleotide sequence, whereas the amino acid translation of the putative consensus sequence (cs) is represented below the m1 sequence. The amino acid sequence corresponding to the putative overlapping region between *orfA* and *orfB* is underlined and shown by lowercase letters. The heptanucleotide AAAAAAG and a contiguous putative hairpin loop contained in this overlapping region are doubly underlined. The left and right terminal IRs (IRL and IRR) and the DRs (DRL and DRR) are indicated by arrows. The putative -35 and -10 promoter regions located upstream of the IS sequence are also doubly underlined.

```

ISL2   MKYETAKNLNNTFRKRLIGVAKPVFDEMVKVLAKEYQVKHARGGRKPKLA 50
       | . . . . . | . . . . . | . . . . . | . . . . . |
IS1381 MNDEATKQLTDARFKRLVGVQRTTFEEMLAVLKTAYQLKHAKGGRKPKLS 50
       | . . . . . | . . . . . | . . . . . | . . . . . |
IS702  MKIEKAKQLTARKFKRMSGVSRQTFNYMVDVVK-ADEKKKKKFGRRPKLY 49
       | . . . . . | . . . . . | . . . . . | . . . . . |

ISL2   IEDLLLATLQYLKEYRTYEQIAADYGVDHNSLIRSHWAEETLVKH---G 97
       : : : : : | : : : : : | : : : : : | : : : : : |
IS1381 LEDLLMATLQYVREYRTYEETAADFGIHESNLIRSQWVEVTLVQS---G 97
       : : : : : | : : : : : | : : : : : | : : : : : |
IS702  IEDQVLMVIQYVREYRTYHIGLDWGLSESAVCRVYKIKENILISSRKFS 99
       : : : : : | : : : : : | : : : : : | : : : : : |

ISL2   FNIGKQEI---KPDDVVLIDATEVKIQRPKKDKQLIIPARKSSSTVLKAQA 144
       . : : . : . : : . : : | : : : : : | : : : . : :
IS1381 VTISRTPL---SSEDVTMIDATEVKINRPKKELAN-DSGKKKFHAMKAQA 143
       : . : . | : : : : : | : : : : : | : : : : : |
IS702  LFGKKELLKMPSEQENLVVMDVTESPIERPCKKSQKFFSGKAGEHTLTKQL 149
       : . : . | : : : : : | : : : : : | : : : : : |

ISL2   ITDIT-GRIIHLSDCQAYRDMRLRLRESRSLHRSGLILADSGYQGLDKI 193
       | . . . | : : : | . . | : : : : : | : : : . : :
IS1381 IVTSQ-GRIVSLDITVNYCHDMKLFKMSRRNIGQAGKILADSGYQGLMKI 192
       : : . : : | : : . : : | : : : : : | : : : : : |
IS702  VIHQKTSQIICLGHGKRIHDFRLFKTSGVKFSELLKVIADKGYQGITKI 199
       : : : : : | : : : : : | : : : : : | : : : : : |

ISL2   YFQAQTPVKSSKKKPLTQDDRELNLISSIRIKVEHVFGVKYKAYKIFSTT 243
       | : : . | : : : | : : . | : : . | : : : : : | : : :
IS1381 YPQAQTPRKSSKLLKPLTVEDKACNHALSKERSKVENIFAKVKTFFKMFSTT 242
       . . : : | : : | . : . | : : : : : | : : : : : |
IS702  HKLSETPIKKPKGKLAKKEQKEYNRELNRRLRVVEHVNRRLKIFNLSNQ 249
       : : : : | : : : | : : : : : | : : : : : |

ISL2   YRNHRRFNLRLMNLICGIIQELAI 268
       | : : : | : : : | : : : : : |
IS1381 YQNHRRFGLRLMNLISAGIINHELGF 267
       | : : : | : : : | : : : : : |
IS702  YRNHRRFGLRLMNLISAGIINYLALKA 276
       | : : : | : : : | : : : : : |

```

FIG. 3. Alignment of transposases of IS1381, ISL2, and IS702. Alignment was performed with the BESTFIT program. Dashes indicate gaps introduced to maximize similarity. Identical residues are indicated by vertical lines. Conserved substitutions are denoted by dots or colons.

Lys+Arg+His, and only 10.1% Glu+Asp) and a pI of 10.7, which is consistent with the idea that it corresponds to a DNA-binding protein. A search for promoter sequences revealed the presence of a putative -35 box (TaGACT) within the IRL sequence and three putative -10 boxes (TATggT, TggtAT, and TATAgT) separated by 15, 17, and 20 bp, respectively. A putative RBS sequence (AtAGtAGtT) that imperfectly matches the 16S rRNA pneumococcal sequence (3'-UUUCC UCCA-5') was found 2 bp upstream of the ATG codon overlapping the most consensus -10 sequence. However, the absence of an RBS does not preclude translation, since it has been shown that the mRNA encoding the pneumococcal DNA polymerase I begins only two nucleotides before the first codon and lacks a typical RBS (12). Interestingly, the transposases of ISL2 also have no obvious RBS (24). Nevertheless, we cannot rule out the possibility that the active copy of IS1381 might have an additional insertion in this region containing more efficient transcription and/or translation signals.

The G+C content of the three copies of IS1381 is 38%, which is characteristic of pneumococcal genes. This finding suggests that this element originated in the pneumococcus or that its acquisition was not a recent event. Alternatively, taking into account the capacity of *S. pneumoniae* to undergo natural transformation, we cannot exclude that IS1381 may be recently acquired from bacteria showing similar G+C contents, as is the case of *L. helveticus* and its ISL2 element (38.5% G+C content) (24). On the other hand, IS1381 appears to be unrelated to any previously known pneumococcal ISs.

Distribution of IS1381 among pneumococcal isolates and related species. Pulsed-field gel electrophoresis (1) and Southern blot hybridization were used to investigate the copy number and distribution of IS1381 among different pneumococcal isolates belonging to different serotypes (Fig. 4). In spite of the noticeable polymorphism of the different pneumococcal iso-

FIG. 4. IS1381 distribution among strains of *S. pneumoniae* belonging to different serotypes. The Southern blot was performed with *Sma*I-digested genomic DNA from the indicated independent strains separated by pulsed-field gel electrophoresis. The hybridization was carried out at 65°C, and the *EcoRV-Ava*I fragment of plasmid pEK6 labeled with [α -³²P]dCTP was used as a probe. Strains D39 and 406 belong to serotypes 2 and 3, respectively. R6 and M24 are unencapsulated pneumococcal mutants (1).

lates, a common feature of this species (11), the minimum apparent copy numbers of IS1381 appear to range from 5 to 7.

Results of Southern blot analysis of the species most closely related to pneumococci, the type strains *Streptococcus mitis* NCTC 12261 and *Streptococcus oralis* NCTC 11427, revealed that the latter does not contain IS1381, whereas the former contained one copy of this mobile element (Fig. 5). The copy present in *S. mitis* was isolated by PCR with the OIS1 primer corresponding to the consensus IR sequence. The sequence of this IS revealed that it contains a truncated transposase similar to the pneumococcal copies (Fig. 2). It has been shown that *S. mitis*, *S. oralis*, and *S. pneumoniae* exhibit more than 99% 16S rRNA sequence homology, although the DNA-DNA similarity

FIG. 5. Presence of IS1381 in *S. oralis* NCTC 12261 and *S. mitis* NCTC 11427. The Southern blot was performed with *Hinc*II (lanes 1 to 4)- or *Hind*III (lanes 5 to 8)-digested genomic DNA from *S. pneumoniae* R6 (lanes 1 and 5), *S. mitis* NCTC 12261 (type strain [lanes 2 and 6]), *S. pneumoniae* S19 (lanes 3 and 7), and *S. oralis* NCTC 11427 (type strain [lanes 4 and 8]). The hybridization was carried out at 65°C, and the *EcoRV-Ava*I fragment of plasmid pEK6 labeled with [α -³²P]dCTP was used as a probe. The migration of size markers is indicated on the left.

values for their total chromosome DNAs were lower than 60% (10). The existence of a single and apparently inactive copy of IS1381 in *S. mitis* strongly suggests that *S. mitis* and *S. pneumoniae* might have exchanged DNA. Exchange of genetic determinants (i.e., penicillin resistance genes) between *S. pneumoniae* and *S. mitis* has been well documented (9). Nevertheless, we cannot rule out the possibility that such DNA exchange could be produced through an intermediate species. Moreover, although the *S. mitis* and *S. oralis* strains used in these experiments are the representative type strains of these species, some differences could be expected from other strains of the same species.

Finally, it is interesting that whereas the other two pneumococcal insertion sequences, IS1202 (15) and IS1167 (23), have not been found in many of the serotypes tested (i.e., serotypes 2, 3, 14, and 23), IS1381 appears to be present in all of the serotypes assayed so far. Thus, IS1381 might represent a new widespread and highly polymorphic strain marker and, consequently, might be a good candidate for epidemiological investigations.

Nucleotide sequence accession number. The GenBank accession numbers for the IS1381 copies shown in Fig. 2 are Z77725 (c1 [pEK5]), Z77726 (c2 [pIS2]), Z77727 (c3 [pIS3]), and Z82003 (m1).

We thank P. García for critical reading of the manuscript. The artwork of A. Hurtado and the technical assistance of E. Cano and M. Carrasco are gratefully acknowledged.

This work was supported by grant PB93-115-CO2-01 from the Programa Sectorial de Promoción General del Conocimiento. A. R. Sánchez-Beato was supported by a fellowship from the Comunidad Autónoma de Madrid.

REFERENCES

- Arrecubieta, C., R. López, and E. García. 1994. Molecular characterization of *cap3A*, a gene from the operon required for the synthesis of the capsule of *Streptococcus pneumoniae* type 3: sequencing of mutations responsible for the unencapsulated phenotype and localization of the capsular cluster on the pneumococcal chromosome. *J. Bacteriol.* **176**:6375–6383.
- Campbell, A., D. E. Berg, D. Botstein, R. P. Novick, P. Starlinger, and W. Szybalski. 1979. Nomenclature of transposable elements in prokaryotes. *Gene* **5**:197–206.
- Chandler, M., and O. Fayet. 1993. Translational frameshifting in the control of transposition in bacteria. *Mol. Microbiol.* **7**:497–503.
- David, F., G. De Céspedes, F. Delbos, and T. Horaud. 1993. Diversity of chromosomal genetic elements and a gene identification in antibiotic-resistant strains of *Streptococcus pneumoniae* and *Streptococcus bovis*. *Plasmid* **29**:147–153.
- Galas, D. J., and M. Chandler. 1989. Bacterial insertion sequences, p. 109–162. In D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
- García, P., J. L. García, E. García, J. M. Sánchez-Puelles, and R. López. 1990. Modular organization of the lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. *Gene* **86**:81–88.
- Groisman, E. A., and H. Ochman. 1996. Pathogenicity islands: bacterial evolution in quantum leaps. *Cell* **87**:791–794.
- Guild, W., S. Hazum, and M. D. Smith. 1981. Chromosomal location of conjugative R determinants in strain BM4200 of *Streptococcus pneumoniae*, p. 610. In S. B. Levy, R. C. Clowes, and E. L. Koenig (ed.), *Molecular biology, pathogenicity, and ecology of bacterial plasmids*. Plenum, New York, N.Y.
- Hakenbeck, R. 1995. Target-mediated resistance to β -lactam antibiotics. *Biochem. Pharmacol.* **50**:1121–1127.
- Kawamura, Y., X.-G. Hou, F. Sultana, H. Miura, and T. Ezaki. 1995. Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus*. *Int. J. Syst. Bacteriol.* **45**:406–408.
- Lefevre, J.-C., G. Faucon, A. M. Sicard, and A. M. Gasc. 1993. DNA fingerprinting of *Streptococcus pneumoniae* strains by pulsed-field gel electrophoresis. *J. Clin. Microbiol.* **31**:2724–2728.
- López, P., S. Martínez, A. Díaz, M. Espinosa, and S. Lacks. 1989. Characterization of the *polA* gene of *Streptococcus pneumoniae* and comparison of the DNA polymerase I it encodes to homologous enzymes from *Escherichia coli* and phage T7. *J. Biol. Chem.* **264**:4255–4263.
- López, R., J. L. García, E. García, C. Ronda, and P. García. 1992. Structural analysis and biological significance of the cell wall lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. *FEMS Microbiol. Lett.* **100**:439–448.
- Mazel, D., C. Bernard, R. Schwarz, A. Castets, J. Hourmard, and N. Tandeau de Marsac. 1991. Characterization of two insertion sequences IS701 and IS702, from the cyanobacterium *Calothrix* species PCC 7601. *Mol. Microbiol.* **5**:2165–2170.
- Morona, J. K., A. Guidolin, R. Morona, D. Hansman, and J. C. Paton. 1994. Isolation, characterization, and nucleotide sequence of IS1202, an insertion sequence of *Streptococcus pneumoniae*. *J. Bacteriol.* **176**:4437–4443.
- Pan, X. S., and L. M. Fisher. 1996. Cloning and characterization of the *parC* and *parE* genes of *Streptococcus pneumoniae* encoding DNA topoisomerase IV: role in fluoroquinolone resistance. *J. Bacteriol.* **178**:4060–4069.
- Povvedi, R., R. Manganelli, and G. Pozzi. 1996. Characterization of conjugative transposon Tn5251 of *Streptococcus pneumoniae*. *FEMS Microbiol. Lett.* **135**:231–236.
- Sánchez-Beato, A. R., C. Ronda, and J. L. García. 1995. Tracking the evolution of the bacterial choline-binding domain: molecular characterization of the *Clostridium acetobutylicum* NCIB 8052 *cspA* gene. *J. Bacteriol.* **177**:1098–1103.
- Sibold, C., Z. Markiewicz, C. Latorre, and R. Hakenbeck. 1991. Novel plasmids in clinical strains of *Streptococcus pneumoniae*. *FEMS Microbiol. Lett.* **77**:91–96.
- Weiss, R. B. 1991. Ribosomal frameshifting, jumping and readthrough. *Curr. Opin. Cell Biol.* **3**:1051–1055.
- Yother, J., and J. M. White. 1994. Novel surface attachment mechanism of the *Streptococcus pneumoniae* protein PspA. *J. Bacteriol.* **176**:2976–2985.
- Zheng, J., and M. A. McIntosh. 1995. Characterization of IS1221 from *Mycoplasma hyorhinis*: expression of its putative transposase in *Escherichia coli* incorporates a ribosomal frameshift mechanism. *Mol. Microbiol.* **16**:669–685.
- Zhou, L., F. M. Hui, and D. Morrison. 1995. Characterization of IS1167, a new insertion sequence in *Streptococcus pneumoniae*. *Plasmid* **33**:127–138.
- Zwahlen, M. C., and B. Mollet. 1994. ISL2, a new mobile genetic element in *Lactococcus helveticus*. *Mol. Gen. Genet.* **245**:334–338.