JOURNAL OF BACTERIOLOGY, Feb. 1992, p. 1345–1351
0021-9193/92/041345-07$02.00/0

Vol. 174, No. 4

DNA Sequence of IS*91* and Identification of the Transposase Gene

M. VICTORIA MENDIOLA,* YOLANDA JUBETE, AND FERNANDO DE LA CRUZ

Departamento de Biología Molecular, Universidad de Cantabria, Cardenal Herrera Oria s/n, 39011 Santander, Spain

Received 10 September 1991/Accepted 10 December 1991

IS*91* is a 1,830-bp insertion sequence that inserts specifically at the sequence CAAG or GAAC of the target and does not duplicate any sequence upon insertion (23). By transposon mutagenesis, we have identified open reading frame 426 (ORF_{426}; bp 454 to 1731) as the putative ORF for the transposase. It displays a cysteine-rich, potential metal-binding domain in its N-terminal region. Adjacent to ORF_{426}, there is an ORF (ORF_{121}) which precedes and terminally overlaps ORF_{426} by one amino acid. Tn*1732* insertions in ORF_{121} do not affect the transposition frequency. IS*91* has sequence similarities to IS*801* from *Pseudomonas syringae*. Their putative transposases are 36% identical, including conservation of the cysteine-rich cluster. The information concerning IS*801* insertion specificity and target duplication has been reevaluated in the light of our results.

Prokaryotic transposable elements are discrete DNA segments that can insert with varying degrees of specificity in the bacterial chromosome or in a plasmid or virus genome. According to their structural and genetic organization, they have been assigned to three classes. Insertion sequences (ISs) correspond to the first class and are defined as transposable elements that do not carry accessory genes (15, 32).

A number of ISs have been described and analyzed in some detail. Galas and Chandler (15) list more than 50 different ISs for gram-negative bacteria. The complete nucleotide sequences of at least 30 ISs are available, and comparisons which suggest evolutionary relationships have been produced. Homologous ISs, detected mainly by conservation of specific motifs among their presumed transposases, have been grouped in families. The two most conspicuous families are the IS*3* family (13, 30) and the IS*4* family (32). The IS*3* family includes a number of ISs sharing extended homology, particularly in a 50-amino-acid segment of the C-terminal end of their transposases, the signature of which is DG2Y5 and 26 amino acids later N32E36K43 (13, 30). The best-known members of this family are IS*2*, IS*3*, IS*51*, IS*150*, IS*600*, IS*629*, IS*861*, IS*3411*, and IS*6110*. IS*21* is also related to this group since it possesses the N32E36K43 motif. They share not only the same signature but also a common genetic organization. All constituent members have two open reading frames (ORFs), with the start of the second one overlapping the terminus of the first in the −1 phase. It is thought that the active transposase is produced by slippage of the first ORF to the second during translation (25). The IS*4* family is also very widespread. Its diagnostic signature, YE7K14, is found in IS*4*, IS*10*, IS*50*, IS*186*, IS*231*, and IS*H1* and also in the more distantly related elements IS*26*, IS*52*, and IS*903* (32). In this case, there is a common genetic organization, since the members of the IS*4* family contain a long ORF covering almost all of the coding capacity of the element. However, there are still many well-known ISs, such as IS*1*, IS*5*, IS*30*, and IS*200*, etc., which do not display a significant relationship with either of these families or among themselves.

IS*91* was isolated from a hemolysin plasmid of *Escherichia coli* and found to occur in multiple copies in a number of Hly plasmids of different incompatibility groups (35). It shows the recombinational properties of an IS element (9) and is probably involved in the evolutionary spread of the *hly* genes (36). IS*91* is particularly interesting because it shows an absolute insertion specificity for the sequences GAAC and CAAG and does not duplicate the target DNA upon insertion (10, 23). We report the DNA sequence of IS*91* and identify the transposase gene by transposon insertion mutagenesis. We also show that IS*91* is similar to IS*801* and discuss the implications of this similarity for the previously proposed mechanism of IS*801* transposition.

* Corresponding author.

MATERIALS AND METHODS

Strains and plasmids. The strains used are listed in Table 1, and plasmids are listed in Table 2.

Standard genetic techniques. Transformations were carried out by the method of Chung and Miller (5). For mating experiments, donor and recipient strains were grown overnight in Luria-Bertani (LB) broth (20) with appropriate antibiotics for plasmid selection. Donors were then diluted 1:10 into fresh LB broth and grown for 2 h. Then 0.5 ml of recipient culture was mixed with 0.2 ml of donor culture, the mixture was centrifuged for 1 min, and the cells were placed onto a Millipore filter on a prewarmed AM3 plate for 1 h at 37°C. Bacteria were washed from the filters with LB broth, diluted, and plated on selective media. Transfer frequencies are expressed as the number of transconjugants per donor cell. Antibiotics were used at the following concentrations (milligrams per liter): chloramphenicol (CM, 25); kanamycin (KM, 25); streptomycin (SM, 300); tetracycline (TC, 10); and trimethoprim (TP, 25).

Insertional mutagenesis of IS*91* with Tn*1732*. Plasmid pSU234 DNA was sequentially treated with *Eco*RI, the Klenow fragment of DNA polymerase I to fill in the cohesive ends, and DNA ligase, and the resulting DNA was used to transform strain DH1 to a Tc^r phenotype. In this way, we obtained plasmid pSU2911, which had lost the *Eco*RI site and was Cm^s.

Strain RU4406 was transformed with pSU2911, streaked on LB agar plates, and grown at 30°C for 3 days. DNA was isolated from 80 single colonies and separately used to transform strain DH1 to a Km^r phenotype. The frequency of transposition (Km^r-to-Tc^r transformants) was roughly 10^{-4}. The sites of Tn*1732* insertion were mapped by restriction

TABLE 1. Strains

Strain	Relevant genotype	Reference
DH1	*supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	17
RU4406	MM294 (*thi endA hsdR*)::Tn*1732* (Km^r)	33
UB1637	F$^-$ *his lys trp rpsL recA56*	7

analysis and are correct to ±10 bp. In this way, we obtained 16 independent Tn*1732* insertions in IS*91*.

To delete Tn*1732*, each of the 16 pSU2911 (IS*91*::Tn*1732*) DNAs was separately digested with *Eco*RI, ligated, and used to transform DH1 to Tc^r. DNA was isolated from the Km^s transformants and analyzed with restriction endonucleases to verify that Tn*1732* had been deleted.

Transposition of IS*91*::Tn*1732*Δ*Eco*RI mutants. Derivatives of DH1 carrying R388 and each of the plasmids to be tested for transposition were mated with UB1637 as described above. Transconjugants were selected in LB agar plates containing either TP plus SM or TC plus SM. The ratio of Tc^r Sm^r to Tp^r Sm^r is the frequency of transposition.

DNA sequence of IS*91*. DNA was sequenced by the chain termination method of Sanger et al. (28) by using plasmid DNA as a substrate, as described by Korneluk et al. (18). The starting plasmid used was pSU240 (9), from which subclones covering the region to be sequenced were obtained by using pUR250 or pSU2719 as a vector (22). Commercial primers (M13 sequencing primer [17-mer] and M13 reverse sequencing primer [16-mer]) were used for sequencing both strands of inserts. All of the sequence shown in Fig. 1 was determined for both strands.

DNA sequence analysis. The GCGUW software (8) was used for DNA and protein sequence analyses. Data bank searches were carried out by using the FASTA and TFASTA programs (24); dot matrix analysis used the DOTPLOT program under conditions which, as a control, revealed the relationship among the ISs belonging to the IS*3* family (30); alignments used the GAP program; rho-independent transcriptional terminators were searched with the TERMINATOR program (4).

Nucleotide sequence accession number. The IS*91* sequence has been registered in the EMBL, GenBank, and DDBJ data bases under accession no. X17114.

RESULTS

Analysis of the DNA sequence. The complete DNA sequence of IS*91* is shown in Fig. 1. IS*91* is 1,830 bp long, with a G+C content of 52.9%. It contains four ORFs starting with ATG or GTG that would code for proteins longer than 100 amino acids; on the upper strand, ORF-I would encode a protein of 121 amino acids (bp 89 to 454) starting with ATG and containing a reasonable ribosome-binding site (RBS)

TABLE 2. Plasmids

Plasmid	Structure	Relevant phenotype	Reference
pSU234	pACYC184::IS*91*	Cm^r Tc^r Rep(p15A)	9
pSU240	pBR322(bla::IS*91*)	Tc^r Rep(pMB8)	9
pSU2911	pSU234Δ*Eco*RI	Tc^r Rep(p15A)	This work
pSU2719	Cloning vector	Cm^r LacZα Rep(p15A)	21
pUR250	Cloning vector	Ap^r LacZα Rep(pMB8)	27
R388	Natural plasmid	Tp^r Su^r Tra$^+$ Rep (W)	6

(AGGT) at bp 79 to 82. ORF-II has two possible translation start sites, GTG at bp 409, which will result in a protein of 441 amino acids (bp 409 to 1734), or ATG at bp 454, which will result in a protein of 426 amino acids (bp 454 to 1734). None of them contains a significant RBS. On the lower strand, ORF-III would encode a protein of 108 amino acids starting with GTG (bp 424 to 98) and ORF-IV would encode a protein of 106 amino acids starting with ATG (bp 1359 to 1039), neither being preceded by an identifiable RBS. There are also five ORFs that would encode proteins longer than 50 amino acids. There are reasonable matches to the consensus promoter sequence before both ORF_{121} and ORF_{426} (overlined in Fig. 1).

There are two clusters of repeated sequences near the inverted repeats (IRs), particularly in IR_R (see Fig. 1). The termini of IS*91* form an 8-bp imperfect IR (Fig. 1, large open arrows), as reported previously (23). In the sequence proximal to IR_L, there are various possible combinations of direct repeats (DRs) and IRs between bp 8 and 27. A 6-6 IR sequence GCAGCC (bp 8 to 23) and a 6-6 DR sequence GCGGCT (bp 16 to 27) are shown in Fig. 1 (see Discussion). A 9-10 IR (bp 57 to 77) may also be significant, since it is positioned between the putative promoter and the RBS of ORF-I. This IR has the characteristics of a rho-independent transcriptional terminator. The sequence 3′ to ORF-II and proximal to IR_R displays three interesting features: (i) an imperfect IR (10-13) between bp 1758 and 1788 flanked by a 6-bp perfect DR of the sequence TGGTGC, (ii) a set of four runs of 4, 6, 4, and 3 Ts (bp 1784 to 1816) separated by 10 or 11 bp in a manner resembling bent DNA (16), and (iii) a 6-7 IR (bp 1810 to 1823).

Tn*1732* insertion mutagenesis of IS*91*: identification and sequence analysis of the *tnpA* gene. Transposon mutagenesis is an effective and widely used technique for the genetic analysis of bacterial traits. We used Tn*1732* as a mutagen since it contains *Eco*RI sites in its IRs so that the transposon can be easily deleted in vitro, leaving an insertion of 35 bp containing an *Eco*RI site. This *Eco*RI site allows easy mapping of the insertion site (33). We used strain RU4406(pSU2911) for mutagenesis. Of the 80 independent Tn*1732* insertions analyzed, 16 lay within IS*91*. DNAs from the 16 pSU2911 (IS*91*::Tn*1732*) insertions were digested with *Eco*RI, religated, and introduced by transformation into DH1 to delete the transposon. The resulting mutants were assayed for transposition; the results are shown in Fig. 2. All insertions in ORF_{426} were Tnp$^-$, and insertions outside remained Tnp$^+$. We conclude that ORF_{426} codes for the IS*91* transposase (the *tnpA* gene). Note that insertion no. 430 interrupts the longer ORF_{441} but is proficient in transposition. This strongly suggests that the downstream ATG at bp 454 is the translational start site of *tnpA*.

ORF_{426} is particularly rich in cysteine and tryptophan (4 and 3.5 times more abundant than in an average *E. coli* protein). Most of the cysteine residues are clustered in the N-terminal 100 amino acids of the protein. Eight are clustered in a region reminiscent of a Cx type zinc finger (C-C-X4-C-C-X5-C-X2-C-X4-C-X2-C), which might be implicated in metal binding; however, the structure in ORF_{426} does not correspond to any of the standard zinc finger motifs or other metal-binding domains (3, 12, 14).

Comparison of the IS*91* sequence with known ISs. The DNA sequence of IS*91* and the deduced amino acid sequences of its two putative gene products were compared with those of other IS elements. Only one IS appears to be similar to IS*91*, IS*801* of *Pseudomonas syringae* (26). Figure 3 shows a plot of the comparison of both DNA sequences.

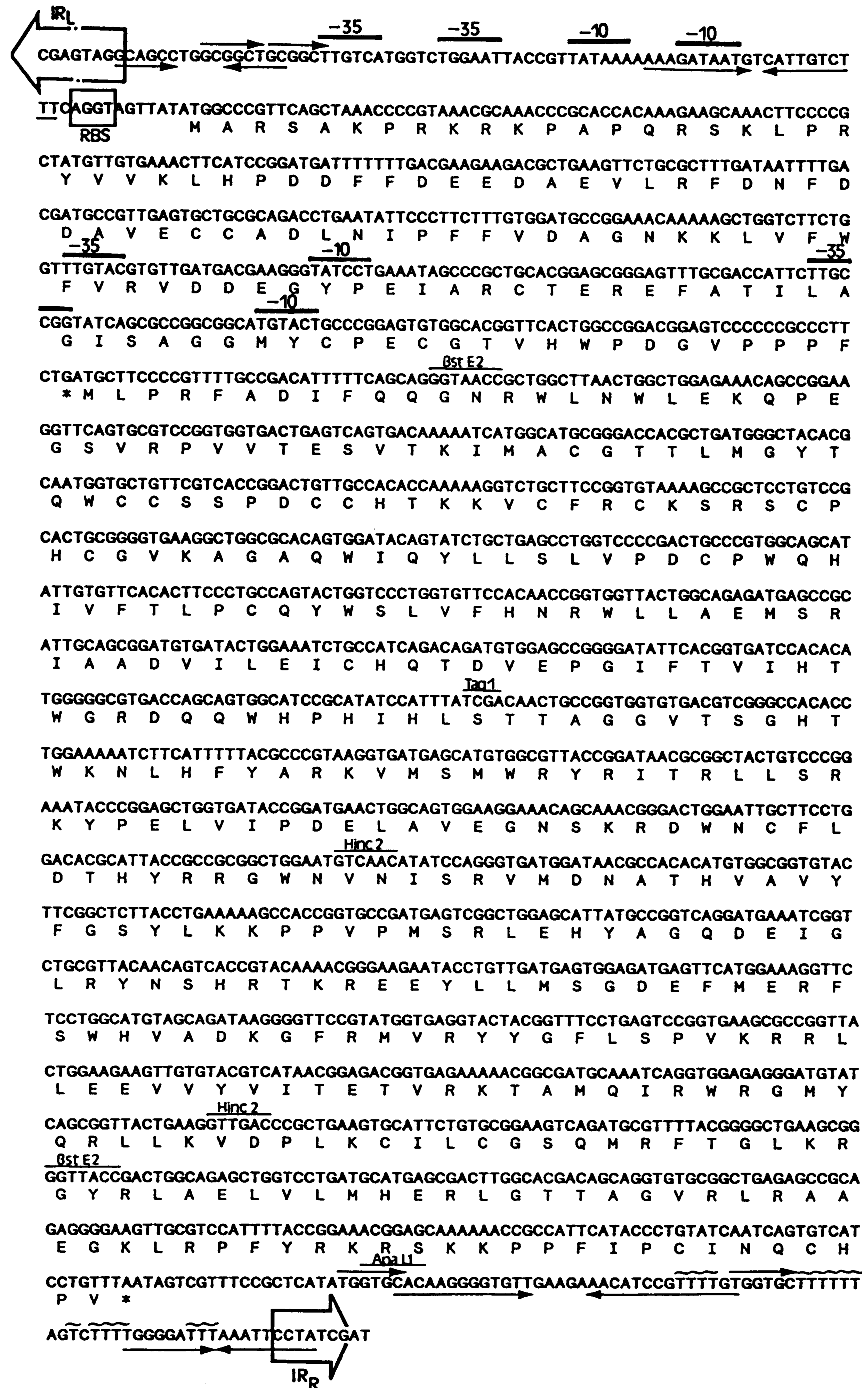

FIG. 1. DNA sequence of the upper strand of IS*91* and the amino acid sequences of the proposed ORF_{121} and ORF_{426} proteins (see text). The terminal IRs (IR_L and IR_R) are enclosed by arrows. The −35 and −10 regions of putative promoters are overlined, and the RBSs are boxed. DRs are shown by arrows above the nucleotide sequence, and IRs are shown by arrows below the sequence. The four runs of Ts suspected to provoke DNA bending are shown by wavy lines. Also shown are the endonuclease cleavage sites for *Apa*LI, *Bst*EII, *Hin*cII, and *Taq*I. The terminal ATCGAT sequence of IR_R is a *Cla*I and *Taq*I site (not shown). IS*91* does not contain sites for the common endonucleases *Eco*RI, *Hin*dIII, *Bam*HI, *Sal*I, and *Pst*I.

They can be aligned from bp 550 of IS*91* and bp 235 of IS*801* to the ends of their respective DNA sequences (data not shown). The DNA similarity in this region is about 45% and indicates similarity in the two transposases. It is obvious from this comparison that the sequence corresponding to ORF_{121} is missing in IS*801* (Fig. 3). ORF_{426} of IS*91* is 36% identical to the long ORF of IS*801*, and the similarities extend throughout the ORF (Fig. 4). It should be emphasized that both proteins show exact conservation of eight cysteines in the N-terminal cluster. If conservative amino acid

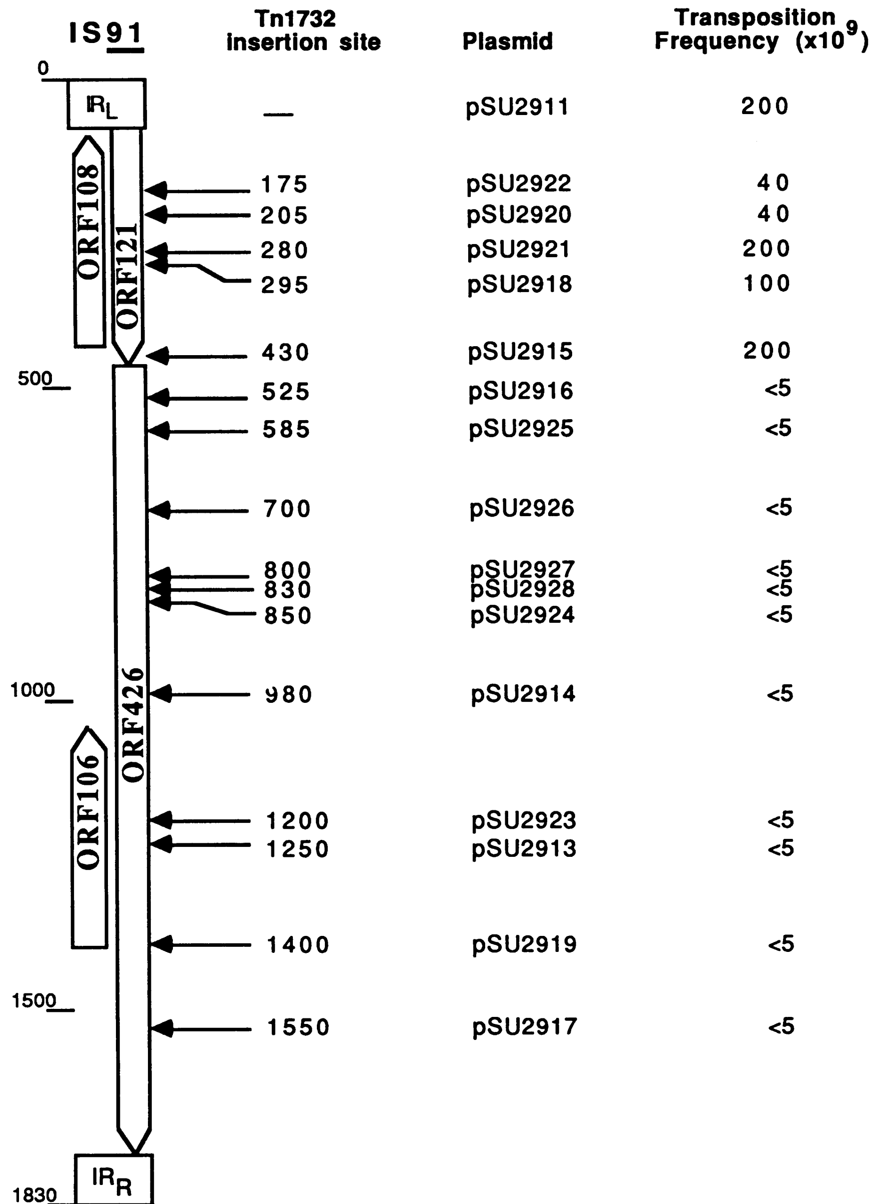

FIG. 2. Transposition frequencies of IS*91*::Tn*1732*Δ*Eco*RI mutants. At left is a map of IS*91* with the localization of the Tn*1732*Δ*Eco*RI insertions, the names of the donor plasmids containing the insertions, and their frequency of transposition (the number of Cm^r Sm^r transconjugants divided by the number of Tp^r Sm^r transconjugants). The results shown are the averages of at least four different experiments.

changes are accepted, the similarity of the transposases is 54%.

We have analyzed ORF_{426} for a number of signatures diagnostic of different families of proteins known to interact with DNA. It does not contain the helix-turn-helix domain common to many bacterial DNA-binding proteins (11), the leucine zipper motif (19), the nucleotide-binding site motif GX4GK(S/T) (29, 34), or the IS*3* or IS*4* family signatures. It is also unrelated to the integrase and resolvase families of site-specific recombinases (1, 31).

DISCUSSION

Genetic organization of IS*91*. IS*91* is an unusual IS element. It transposes with high specificity to the target sequence GAAC or CAAG without duplicating any base pairs at the target (10, 23). Surprisingly, it can promote replicon fusion with plasmids containing only IR_R in a manner similar to the one-ended transposition described previously for class II transposons (2, 23a). To provide us with the necessary background information for a more-detailed analysis of the mechanism of IS*91* transposition, we decided to determine the sequence of the entire element.

Superficial analysis of the coding capacity of IS*91* suggests a gene organization analogous to the IS*3* family of ISs, that is, one whereby two ORFs translated from the same DNA strand and overlapping by a single nucleotide cover the coding capacity of the element (25). The results presented in this work strongly suggest that ORF_{426} codes for the transposase of IS*91*. This was indicated by the abolition of IS*91*

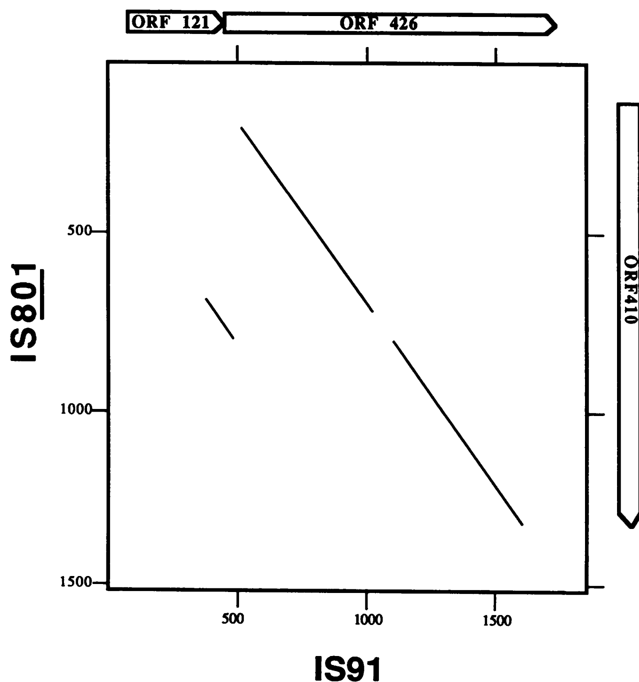

FIG. 3. Comparison of the DNA sequences of IS*91* and IS*801*. The DOTPLOT program of GCG was used, and the parameters were set to a length of 100 bp and mismatch of 57 bp. The positions of the relevant ORFs (see text) are shown for reference.

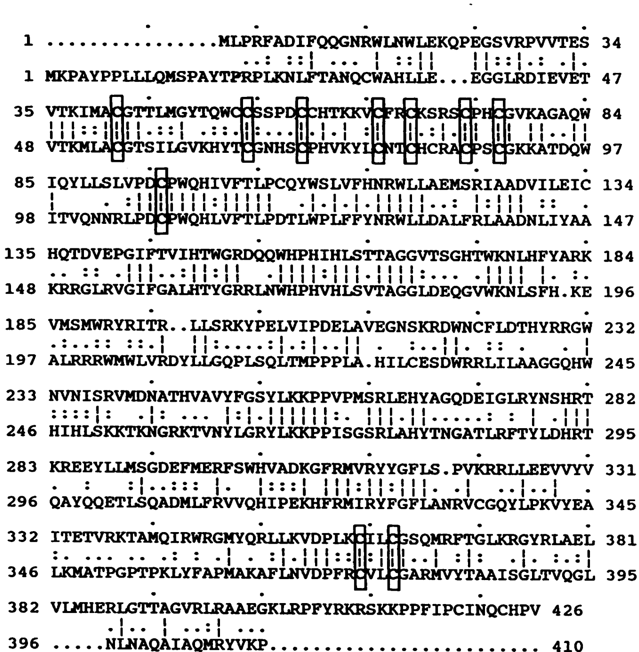

FIG. 4. Comparison of the amino acid sequences of ORF_{426} of IS*91* and ORF_{410} of IS*801*. The GAP program was used. Parameters: gap weight, 3.0 bp; length weight, 0.1 bp. Conserved cysteines are boxed.

transposition by Tn*1732* insertion mutations in ORF_{426} (Fig. 2).

Comparison of IS*91* and IS*801* protein-coding capacities. The insertion mutagenesis experiments summarized in Fig. 2 show that ORF_{426} is the transposase of IS*91*. This protein is homologous to ORF_{410} from IS*801*. These proteins are 36% identical, with the homology extending over their entire sequences (Fig. 4). When analyzing ORF_{410}, Romantschuk et al. (26) suggested Val-398 as the most plausible translational start site, on the basis of a better RBS. However, the region between Met-101 and Val-398 is homologous to the N terminus of ORF_{426} of IS*91*, including the highly conserved cysteine-rich domain. On the basis of this homology, we think that Met-101 is the most likely start for the IS*801* transposase. Perhaps significantly, there are good RBSs in front of both transposases (GAGG for IS*801* and GGAG for IS*91*), but they are too far (16 bp) from the proposed translational starts.

Interestingly, IS*801* does not have any sizeable ORF preceding ORF_{410}. This fact further suggests that ORF_{121} of IS*91* (or ORF_{108}, whichever is translated, if any) is probably irrelevant for transposition. The small differences in the transposition frequencies of the insertions within ORF_{121} (Fig. 2) are most likely due to statistical variation and are not considered significant.

In IS*801*, Romantschuk et al. (26) refer to a second ORF (ORF_{199}) which overlaps ORF_{410} in the complementary DNA strand. ORF-IV of IS*91* is 26% identical to the N-terminal 111 amino acids of ORF_{199} (data not shown). With the data reported here, we cannot comment on the potential function of these proteins. However, it may be significant that many IS elements contain surprisingly long ORFs overlapping the transposase in the complementary strand (15). They could have a regulatory function in expression of the transposase or in transposition itself.

The termini of IS*91* and IS*801*: specificity of insertion and target duplication. Figure 5 shows the alignment we propose for the termini of IS*91* and IS*801*. Romantschuk et al. (26) suggested that the 5-bp sequence TGAAC (underlined in Fig. 5) is duplicated upon insertion of IS*801*. We previously showed that IS*91* inserts specifically in GAAC (or CAAG) without target DNA duplication (10, 23). By analogy to IS*91*, we believe that the results of Romantschuk et al. (26) are more-adequately explained by assuming that GAAC is an integral part of IS*801* IR_L and that IS*801* inserts specifically in the sequence GAAC of the recipient (recipient sequences are shown in brackets in Fig. 5). In favor of this hypothesis is the fact that, in one of the three insertions sequenced, the target sequence is TGGAC. This sequence does not appear duplicated upon insertion, but (not surprisingly according to our interpretation) TGAAC appears at the other junction. We suggest that a more-thorough analysis of the IS*801* specificity of insertion will probably demonstrate that IS*801* inserts specifically in the target sequence GAAC (or GGAC) and does not produce any target duplication.

Comparison of the left termini of IS*91* and IS*801* (IS*91L* and IS*801L* in Fig. 5) shows almost perfect conservation of a 16-bp sequence (bp 7 to 23) including a 6-bp IR (which will be called IR_{CL}). Since the transposases of both elements are homologous, they presumably recognize similar DNA sequences, so IR_{CL} may be important for transposase recognition. The DNA sequences of IS*91* and IS*801* beyond this point are not significantly homologous up to bp 550 in IS*91* (bp 235 in IS*801*), as shown in Fig. 3.

Comparison of the right termini (IS*91R* and IS*801R*) also shows almost perfect sequence conservation in the terminal 27 bp. Again in this case, the conserved region includes a

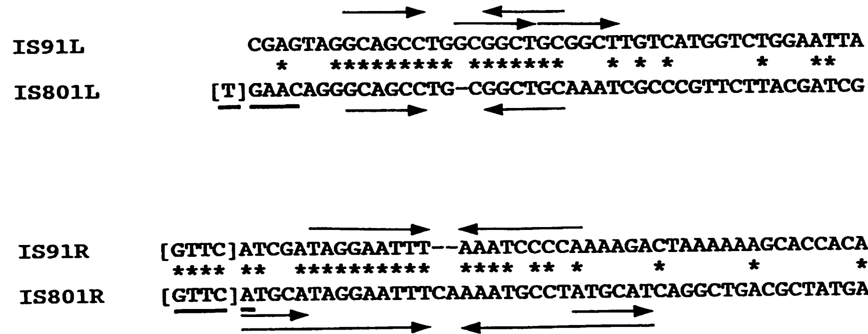

FIG. 5. Comparison of the DNA sequences at the termini of IS*91* and IS*801* (L and R refer to the left and right termini as shown in Fig. 1 for IS*91* and as described previously by Romanschuk et al. [24] for IS*801*). Identical nucleotides between each pair of sequences are shown by asterisks. Arrows show the IRs and DRs beside each sequence. The underlined TGAAC (and its complement GTTCA) sequences in IS*801* refer to the 5 bp presumably duplicated at IS*801* insertion sites (26). In our alternative explanation (discussed in the text), only those base pairs in brackets come from the target DNA.

conspicuous IR (which will be called IR_{CR}). Surprisingly, IR_{CR} and IR_{CL} are not similar to one another. If we assume that both termini of an IS element should be specifically recognized by the transposase, IR_R and IR_L of IS*91* and IS*801* must be recognized by different domains of their respective transposases. Alternatively, IR_{CR} and IR_{CL} could correspond to binding sites for different host factors. These possibilities are attractive for further work.

ACKNOWLEDGMENTS

We are thankful to Julian Davies from the Institut Pasteur (Paris) for critical reading and correction of the manuscript, to Luis Blanco from the C.B.M. (Madrid) for introducing us to the GCGUW package, to Juan C. Zabala for discussion and sharing of techniques, and to Matilde Cabezas for technical assistance.

Y.J. was supported by a P.F.P.I. from the Spanish Ministry of Education. This work was supported by a grant (BIO 661/90) from the C.I.C.Y.T. (Spain).

REFERENCES

1. **Argos, P., A. Landy, K. Abremski, J. B. Egan, L. E. Haggard, R. H. Hoess, M. L. Kahn, B. Kalionis, S. V. L. Narayana, L. S. Pierson, N. Sternberg, and J. M. Leong.** 1986. The integrase family of site-specific recombinases. EMBO J. **5:**433–440.
2. **Avila, P., F. de la Cruz, E. Ward, and J. Grinsted.** 1984. Plasmid containing one inverted repeat of Tn*21* can fuse with other plasmids in the presence of Tn*21* transposase. Mol. Gen. Genet. **195:**288–293.
3. **Berg, J. M.** 1990. Zinc fingers and other metal-binding domains. J. Biol. Chem. **265:**6513–6516.
4. **Brendel, V., and E. N. Trifonov.** 1984. A computer algorithm for testing potential prokaryotic terminators. Nucleic Acids Res. **12:**4411–4427.
5. **Chung, C. T., and R. H. Miller.** 1988. A rapid and convenient method for the preparation and storage of competent bacterial cells. Nucleic Acids Res. **16:**3580.
6. **Datta, N., and R. W. Hedges.** 1972. Trimethoprim resistance conferred by W plasmid in *Enterobacteriaceae*. J. Gen. Microbiol. **72:**349–355.
7. **de la Cruz, F., and J. Grinsted.** 1982. Genetic and molecular characterization of Tn*21*, a multiple resistance transposon from R100.1. J. Bacteriol. **151:**222–228.
8. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
9. **Díaz-Aroca, E., F. de la Cruz, J. C. Zabala, and J. M. Ortiz.** 1984. Characterization of the new insertion sequence IS*91* from an alpha-hemolysin plasmid of *Escherichia coli*. Mol. Gen. Genet. **193:**493–499.
10. **Díaz-Aroca, E., M. V. Mendiola, J. C. Zabala, and F. de la Cruz.** 1987. Transposition of IS*91* does not generate a target duplication. J. Bacteriol. **169:**442–443.
11. **Dodd, I. B., and J. B. Egan.** 1990. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. Nucleic Acids Res. **18:**5019–5026.
12. **Evans, R. M., and S. M. Hollenberg.** 1988. Zinc-fingers: guilt by association. Cell **52:**1–3.
13. **Fayet, O., P. Ramond, P. Polard, M. F. Prère, and M. Chandler.** 1990. Functional similarities between retroviruses and the IS*3* family of bacterial insertion sequences. Mol. Microbiol. **4:**1771–1777.
14. **Freemont, P. S., I. M. Hanson, and J. Trowsdale.** 1991. A novel cysteine-rich sequence motif. Cell **64:**483–484.
15. **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.
16. **Hagerman, P. J.** 1990. Sequence-directed curvature of DNA. Annu. Rev. Biochem. **59:**755–781.
17. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
18. **Korneluk, R. G., F. Quan, and R. A. Gravel.** 1985. Rapid and reliable dideoxy sequencing of double-stranded DNA. Gene **40:**317–323.
19. **Landschulz, W. H., P. F. Johnson, and S. L. McNight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240:**1759–1764.
20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Martinez, E., B. Bartolomé, and F. de la Cruz.** 1988. pACYC184-derived cloning vectors containing the multiple cloning site and *lac*Zalpha reporter gene of pUC8/9 and pUC18/19 plasmids. Gene **68:**159–162.
22. **Mendiola, M. V.** 1989. Ph.D. thesis. Universidad Complutense de Madrid, Spain.
23. **Mendiola, M. V., and F. de la Cruz.** 1989. Specificity of insertion of IS*91*, an insertion sequence present in alpha-hemolysin plasmids of *Escherichia coli*. Mol. Microbiol. **3:**979–984.
23a. **Mendiola, M. V., and F. de la Cruz.** Unpublished report.
24. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
25. **Prère, M. F., M. Chandler, and O. Fayet.** 1990. Transposition in *Shigella dysenteriae*: isolation and analysis of IS*911*, a new member of the IS*3* group of insertion sequences. J. Bacteriol. **172:**4090–4099.
26. **Romantschuk, M., G. Y. Richter, P. Mukhopadhyay, and D.**

Mills. 1991. IS*801*, an insertion sequence element isolated from *Pseudomonas syringae* pathovar *phaseolicola*. Mol. Microbiol. **5:**617–622.

27. **Ruther, U.** 1982. pUR250 allows rapid chemical sequencing of both DNA strands of its inserts. Nucleic Acids Res. **10:**5765–5772.
28. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
29. **Saraste, M., P. R. Sibbald, and A. Wittinghofer.** 1990. The P-loop—a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. **15:**430–434.
30. **Schwartz, E., M. Kroeger, and B. Rak.** 1988. Distribution nucleotide sequence and philogenetic relationship of a new *E. coli* insertion element. Nucleic Acids Res. **16:**6789–6802.
31. **Sherratt, D. J.** 1989. Tn*3* and related transposable elements: site-specific recombination and transposition, p. 163–184. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
32. **Syvanen, M.** 1988. Bacterial insertion sequences, p. 331–356. *In* R. Kecherlapati and G. R. Smith (ed.), Genetic recombination. American Society for Microbiology, Washington, D.C.
33. **Ubben, D., and R. Schmitt.** 1986. Tn1721 derivatives for transposon mutagenesis, restriction mapping and nucleotide sequence analysis. Gene **41:**145–152.
34. **Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay.** 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1:**945–951.
35. **Zabala, J. C., F. de la Cruz, and J. M. Ortiz.** 1982. Several copies of the same insertion sequence are present in alpha-hemolytic plasmids belonging to four different incompatibility groups. J. Bacteriol. **151:**472–476.
36. **Zabala, J. C., J. M. García-Lobo, E. Díaz-Aroca, F. de la Cruz, and J. M. Ortiz.** 1984. *Escherichia coli* alpha-hemolysin synthesis and export genes are flanked by a direct repetition of IS*91*-like elements. Mol. Gen. Genet. **197:**90–97.