Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 1922–1926, March 1994
Genetics

Differential roles of the transposon termini in IS*91* transposition

(rolling-circle transposition/site-specific nicking)

M. VICTORIA MENDIOLA, IRANTZU BERNALES, AND FERNANDO DE LA CRUZ*

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

Communicated by Donald R. Helinski, October 18, 1993

ABSTRACT **Insertion sequence 91 (IS*91*) inserts specifically at GTTC or CTTG target sequences without duplication of the target. After insertion, the right inverted repeat (IR_R) lies adjacent to the 3′ end of the target sequences (or 5′ to the complementary sequence CAAG or GAAC). We have analyzed the effects of alteration of each terminus of IS*91* on transposition activity in *Escherichia coli*. IR_R is absolutely required for transposition. Deletion analysis indicates that a 14-bp segment is not sufficient, but an 81-bp sequence within the IR_R region is sufficient. Furthermore, the GTTC/CTTG target site is also required. The left inverted repeat (IR_L) of IS*91* is dispensable. Plasmid fusions originated by one-ended transposition of IS*91* derivatives lacking IR_L occur at about the same frequency as cointegrate formation observed for the wild-type element. In the one-ended-type fusions, the inserted fragment of donor DNA is flanked at one end (constant end) by IR_R and at the other end by a GTTC or CTTG sequence present in the donor (variable end) in a way that usually results in multiple tandem insertions of the donor plasmid in the target site. These results are easily accommodated by a rolling-circle replicative transposition mechanism. This model also draws support from the finding that the IS*91* transposase is related in sequence to the superfamily of rolling-circle replication proteins and the observation that IR_R shows some conservation in sequence and secondary structure with the origins of replication of some rolling-circle replication plasmids.**

Transposons are genetic elements able to move to many different genome loci in their hosts, therefore contributing significantly to genome diversity. They have been classified into subgroups based on their transposition mechanism or the sequence similarities between their transposition proteins (1). The mechanisms of the early steps in transposition of a few diverse well-studied transposons display a remarkable similarity in their basic three-step mechanism: nuclease cut, strand transfer, and replication or repair, depending on whether the transposition is replicative (Tn*3*, replicative Mu) or conservative (Tn*7*, Tn*10*, lysogenic Mu) (2).

Insertion sequence 91 (IS*91*) displays a number of characteristics that are unique among IS elements, suggesting an unusual transposition mechanism. First, it shows an absolute insertion specificity for the target sequence GAAC or CAAG; it inserts at the 5′ end of these sequences so that its right inverted repeat (IR_R) lies adjacent to the specific target site. Furthermore, the target DNA is not duplicated (3). Analysis of the IS*91* DNA sequence (4, 5) indicates that (*i*) its transposase is related to a family of replication proteins of plasmids that replicate by rolling-circle (RCR plasmids) (4), and (*ii*) the termini of IS*91* form a very short IR (7 bp with one mismatch). The latter finding suggests that a protein operating at the IRs will also recognize unique sequences at each end, raising the possibility that each terminus plays a different role in IS*91* transposition. This possibility is strengthened by the comparison between the termini of IS*91* and those of the related element IS*801* (6). The 23 terminal bp at IR_R is highly conserved between both elements, including a 9-bp internal palindromic sequence (IR_{CR}). Similarly, there is a 17-bp conserved sequence at the left termini of the elements (IR_L), also including a conserved 6-bp inverted repeat (IR_{CL}). However, IR_{CR} and IR_{CL} show no sequence homology.

Because of the unusual properties of the IS*91* element, its termini were manipulated and transposition of the wild-type and mutant ISs was examined. We provide evidence showing that each terminus of the element has a different role in transposition. The results of the study further suggest that IS*91* transposes by a RCR mechanism.

MATERIALS AND METHODS

Genetic Techniques. The *Escherichia coli* K-12 strains used were UB1637 (F⁻ *his lys trp rpsL recA56*) and UB5201 (F⁻ *pro met gyrA recA56*) (7). Plasmids used are listed in Table 1. Conditions for bacterial growth, plasmid selection, transformation, and conjugation were described (5). Antibiotics were used at the following concentrations: chloramphenicol (Cm), 25 mg/liter; kanamycin (Km), 50 mg/liter; streptomycin (Sm), 300 mg/liter; tetracycline (Tc), 10 mg/liter; and trimethoprim, 25 mg/liter.

Transposition Assays. Assays were carried out by a mating-out assay. Donor strains were derivatives of UB5201 (Sm sensitive, Sm^S) containing a conjugative plasmid [either R388 (trimethoprim resistant) or pSU4053 (ampicillin resistant)] plus a nonmobilizable plasmid that carried the transposon sequences. Thus, the antibiotic-resistance gene of the nonmobilizable plasmid [Tc resistant (Tc^r) in the case of pSU240 and pSU2602; Cm^r for the rest of them] can be transferred to the recipient strain only if there is physical linkage between both plasmids—that is, replicon fusion brought about by the transposon. Donor strains were mated with the recipient strain UB1637 (Sm^r) for 90 min at 37°C on the surface of L-broth agar plates, and suitable dilutions were plated onto selective medium. Sm was used to counterselect the donors. The transposition frequency is the number of transconjugants carrying the marker of the nonmobilizable plasmid divided by the number of transconjugants carrying the marker of the conjugative plasmid.

Plasmid Constructions. Molecular cloning techniques were carried out as described (12). pSU234 and pSU240 are spontaneous IS*91* insertions into pACYC184 and pBR322, respectively. pSU2600 is the result of the insertion of the 2228-bp *Xho* I/*Cla* I fragment from pSU240 that contains IS*91* in the multiple cloning site of the p15A replicon vector pSU20. pSU2601 was constructed by cloning the 1830-bp *Xho* I/*Cla* I fragment of pSU234 that contains IS*91* in the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: IS, insertion sequence; Cm, chloramphenicol; IR_R, right inverted repeat; IR_L, left inverted repeat; IR_{CR} and IR_{CL}, conserved IR_R and IR_L; Km, kanamycin; RCR, rolling-circle replication; Sm, streptomycin; Tc, tetracycline; Tnp^+/Tnp^-, transposition proficient/deficient.

*To whom reprint requests should be addressed.

Table 1. Plasmids

Plasmid	Structure	Relevant phenotype	Ref.
pSU20	Cloning vector	Cm^r LacZα Rep(p15A)	8
pSU40	Cloning vector	Km^r LacZα Rep(p15A)	8
pSU234	pACYC184::IS*91*	Cm^r Tc^r Rep(p15A)	9
pSU240	pBR322::IS*91*	Tc^r Rep(pMB8)	9
pSU256	Fig. 2	Cm^r Tc^r Rep(p15A)	*
pSU257	Fig. 2	Cm^r Tc^r Rep(p15A)	*
pSU258	Fig. 2	Cm^r Tc^r Rep(p15A)	*
pSU2572	Figs. 2 and 3	Cm^r Rep(p15A)	*
pSU2600	Fig. 1	Cm^r Rep(p15A)	*
pSU2601	Fig. 1	Cm^r Rep(p15A)	*
pSU2602	Fig. 1	Tc^r Rep(p15A)	*
pSU2606	pSU2600 (*Nde* I::*kan*)	Cm^r Km^r Rep(p15A)	*
pSU2607	Fig. 1	Cm^r Rep(p15A)	*
pSU2608	Fig. 1	Cm^r Rep(p15A)	*
pSU2615	pSU2600 (Δ*Xho* I/*Nde* I)	Cm^r Rep(p15A)	*
pSU2616	pSU2600 (Δ*Xho* I/*Bst*EII)	Cm^r Rep(p15A)	*
pSU4053	pHG327::TRA_W	Ap^r $Tra(W)^+$ Rep(pMB8)	10
R388	Wild type	Tp^r Su^r $Tra(W)^+$ IncW	11

Ap, ampicillin; Su, sulfadiazine.
*This work.

same sites of pSU20. pSU2602 contains an 18-bp oligonucleotide insertion at the single *Apa*LI site of IS*91* in pSU1911 (5). It was constructed in two steps: first, the Ω segment of pHP45Ω (13) containing a Sm^r gene was cut with *Eco*RI and its ends were made blunt with Klenow polymerase. This fragment was cloned at the single *Apa*LI site of pSU2911 (5), also made blunt with Klenow polymerase. The resulting plasmid (pSU2938) was deleted of the Sm^r segment by digestion with *Sma* I and self-ligation, resulting in the 18-bp insertion shown in Fig. 1. pSU2607 was constructed by digestion of pSU2600 with *Nde* I and *Swa* I, followed by filling in with Klenow polymerase and self-ligation. pSU2608 contains a 14-bp insertion at the *Nde* I site of IS*91* in pSU20. It was also constructed in two steps: first, the 1252-bp *Hin*cII fragment of plasmid pUC4K (Pharmacia) containing the Km^r gene from Tn*903* (14) was inserted at the *Nde* I site of pSU2600 previously made blunt with Klenow polymerase, producing the plasmid pSU2606. Second, the Km^r gene was deleted by *Pst* I digestion and self-ligation. Plasmids pSU2615 and pSU2616 are deletion derivatives of pSU2600. For construction of pSU2615, pSU2600 DNA was digested with *Xho* I and *Nde* I; the resulting ends were filled in with Klenow polymerase and self-ligated. pSU2616 was obtained in a similar way after *Xho* I plus *Bst*EII digestion. For construction of pSU256, pSU257, and pSU258, respectively, 10 μg of plasmid pSU234 DNA was linearized with *Xho* I and treated with 0.3 unit of exonuclease BAL-31 in 50 μl of BAL-31 buffer at 30°C. Samples taken at different times (1–5 min) were purified, filled in with Klenow polymerase, and self-annealed using T4 DNA ligase. Plasmid DNAs from Cm^r colonies obtained after transformation were characterized by restriction analysis and DNA sequencing. The three plasmids selected carry deletions of 14, 45, and 142 bp, respectively. pSU2572 was constructed as follows: pSU256 contains a new *Bgl* I site as a consequence of a deletion that originated it from pSU234. The *Bgl* I/*Hin*dIII fragment of pSU256 containing the Tc^r gene (2162 bp) was deleted after *Bgl* I and *Hin*dIII digestion, treatment with Klenow polymerase, and self-annealing, resulting in pSU2572.

DNA Sequencing. DNA was sequenced by the dideoxynucleotide chain-termination method using Sequenase II and plasmid DNA as a substrate (15). The primers used were commercial M13 and M13 reverse sequencing primers, the IS*91*-specific oligonucleotides IR_R and IR_L used previously (3), and two oligonucleotides synthesized specifically for this work: pAC 1 (corresponding to pACYC184 bp 1234–1218) and pAC 2 (pACYC184 bp 947–931) (16).

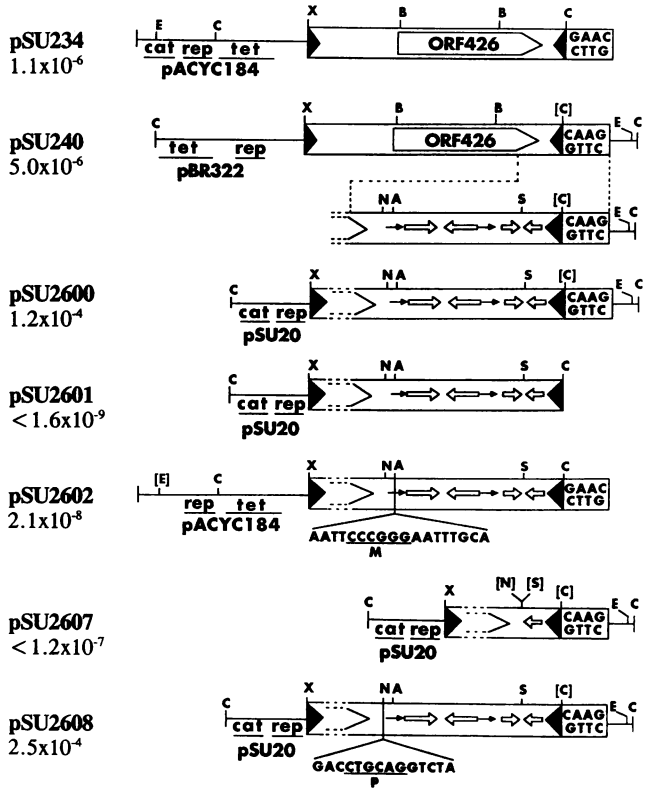

FIG. 1. Transposition of derivatives of IS*91* with alterations around IR_R. Derivatives of strain UB5201 containing R388 plus the plasmids shown were mated with UB1637. Numbers show the mean transposition frequency. Structures shown in the second column represent linear drawings of the plasmids broken at a *Cla* I site, except in the cases of pSU234 and pSU2602, in which the GAAC box has been situated to the right of the dividing *Cla* I site. For each plasmid, the thin horizontal line represents vector sequences and the large open box represents IS*91* sequences. The complete IS*91* is shown for pSU234 and pSU240. For the rest, only the region around IR_R is presented. Solid arrowheads represent the terminal inverted repeats of IS*91*. ORF426 is the coding region of the transposase (5). Small horizontal arrows adjacent to IR_R represent IR_{CR}, IR_T, and one direct repeat that can be relevant for IR_R function (5). The CAAG or GAAC boxes represent the target sequences that are essential for IR_R function. Plasmids pSU2602 and pSU2608 contain insertions of 18 and 14 bp, respectively, shown in capital letters. Genes: *cat*, chloramphenicol acetyltransferase; *tet*, Tc^r; *rep*, replication region of the plasmid. Restriction enzyme sites: A, *Apa*LI; B, *Bst*EII; C, *Cla* I; E, *Eco*RI; M, *Sma* I; N, *Nde* I; P, *Pst* I; S, *Swa* I; X, *Xho* I. A restriction site in brackets means that it was either lost in a fill-in Klenow reaction, or it is not cleaved because of adjacent Dam methylation (*Cla* I sites).

RESULTS

IR_R and the Flanking CAAG Sequence Are Absolutely Required for IS*91* Transposition. Fig. 1 shows the transposition frequencies of various derivatives of IS*91*, which have been modified in the IR_R region. The previously described plasmids pSU234 and pSU240 transpose at frequencies around 10^{-6}. pSU234 contains naturally occurring *Xho* I and *Cla* I sites that cut precisely at its ends (3). In pSU234, the target sequence adjacent to IR_R is 5′-GAAC [the terminal *Cla* I site is, therefore, not methylated (ATCGATGAAC) and, thus, can be cleaved by the enzyme], while in pSU240 the target sequence is 5′-CAAG [the terminal *Cla* I site is methylated by Dam methylase (ATCGATCAAG) and cannot

be cleaved]. The *Xho* I/*Cla* I fragments of pSU234 and pSU240 were inserted in the cloning vector pSU20, giving rise to pSU2600 and pSU2601, respectively. Therefore, pSU2600 retains the 5′-CAAG target sequence, while pSU2601 lost it. As shown in Fig. 1, this difference results in a dramatic difference in their transposition abilities: pSU2600 transposes at a frequency of 10^{-4}, while pSU2601 is markedly transposition deficient (Tnp$^-$) (frequency, $<10^{-9}$). We assume from this result that IS*91* transposition has an absolute requirement for the CAAG (or GAAC) target site.

The result of transposition is either the formation of plasmid cointegrates (donor and recipient plasmids are flanked by a direct repetition of IS*91*) or replicon fusions (see below). The relative proportions of fusions and cointegrates were analyzed for plasmids pSU234 and pSU2600. Usually cointegrates were more abundant, and fusions were obtained at a frequency of 1–10% with respect to cointegrates, depending on the plasmid. Direct transposition of IS*91* could be assayed by using plasmid pSU2606, a derivative of pSU2600 containing a Km^r gene inserted at the *Nde* I site. About 99% of the Km^r transposition products were direct transposition events, and only 1% were cointegrates or fusions, in agreement with previously published data (9).

We next analyzed how much of the IR_R terminus is required for transposition. pSU2607 (Fig. 1) has a 64-bp deletion near IR_R of IS*91* (from bp 1752 to 1816), which lies distal to the end of the transposase gene and leaves 14 bp of IR_R. This derivative of IS*91* is essentially Tnp$^-$. Upstream of IR_R, an 18-bp insertion at the single *Apa*LI site in IS*91* (between bp 1758 and 1759) results in a 50-fold decrease in the frequency of transposition (pSU2602; Fig. 1). On the other hand, pSU2608, which contains a 14-bp insertion at the *Nde* I site (between bp 1752 and 1753), shows no apparent alteration in its transposition ability. Therefore, it is likely that the IR_R sequence required for transposition is contained between the *Nde* I restriction site and the CAAG 5′-terminal adjacent sequence. This hypothesis was confirmed by construction and analysis of two new plasmids (pSU2615 and pSU2616), containing just 81 and 254 bp of the IR_R, respectively. When they were complemented in trans by a third plasmid providing the transposase, replicon fusions were obtained at a frequency of 4–8 × 10^{-7} in both cases. This result indicates that 81 bp of IR_R contains all the information required for one-ended transposition of IS*91* (see below) and, thus, for IR_R function.

Derivatives of IS*91* Lacking IR_L Are Capable of One-Ended Transposition. We also analyzed the effects of alterations in the sequence around the IR_L terminus on the frequency of IS*91* transposition. Surprisingly, derivatives of plasmid pSU234 that were deleted of IR_L were still Tnp proficient (Tnp$^+$) (Fig. 2). Plasmid pSU256, which is deleted for 11 bp, is still fully Tnp$^+$; pSU257, which is deleted for 20 bp, transposes at a frequency 10-fold lower; pSU258, which has a 106-bp deletion, is essentially Tnp$^-$. We assume that in pSU258 the transposase gene (*tnpA*) lost its promoter(s) and, thus, does not express the transposase. Other alternatives, such as a function of IR_{CL} or other potential structural elements near IR_L, were ruled out since, as indicated above, cloning of IR_R alone gives similar results. This one-ended transposition process was examined in more detail. We constructed pSU2572 (Fig. 2) and assayed one-ended transposition of IS*91*Δ16 (the element contained in pSU2572, which lacks 16 bp of IR_L) to pSU4053, a small Tra$^+$ plasmid of known DNA sequence. Eight representative independent insertions are shown in Fig. 3. As can be seen, the result of one-ended transposition is the formation of plasmid fusions by insertion of pSU2572 in pSU4053. In these fusion products, one junction is precisely the IR_R terminus of IS*91* (see below), and the other is variable so that the cointegrate contains from one to three complete copies of pSU2572 plus

FIG. 2. Transposition of derivatives of IS*91* with deletions around IR_L. Transposition experiments were carried out, and the frequencies were calculated as described in the legend of Fig. 1. Results are the mean of six experiments. Plasmids pSU256, pSU257, and pSU258 are deletion derivatives of pSU234 (Fig. 1). DNA sequences of the plasmids around the deletion endpoints are shown. Vector sequences are in lowercase letters, and IS*91* sequences are in capital letters. Coordinates for the vector pACYC184 (numbers after dots) are from Rose (16). Coordinates for IS*91* start at +1, which marks the beginning of IR_L (shaded) in pSU234. In the other plasmids, the + sign shows the beginning of the IS*91* sequences. Arrows below the pSU234 sequence show the position of IR_{CL}, thought to be involved in IR_L function. P(−35) shows the position of the −35 sequence of the most plausible promoter for the transposase gene. Relevant restriction enzyme sites are shown below the sequences.

an additional segment of variable length. The sequences at the junctions between donor and recipient DNAs were determined for these recombinants and are shown in Fig. 4 (the DNA sequences represented show the complementary strand of IS*91*, so that IR_R is located to the left; in this orientation, wild-type IS*91* inserts 3′ of either CTTG or GTTC). The following conclusions can be drawn from the resulting sequence analysis: (*i*) One junction (constant end) is always formed by recombination of IR_R with the target sequence CTTG or GTTC. As is the case with wild-type transposition, both sequences are used with roughly equal probability (3). One target site (bp 831) was selected twice, and in a third

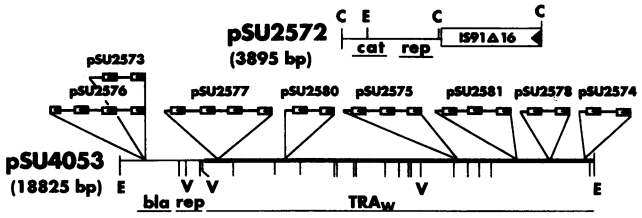

FIG. 3. Plasmid recombinants obtained by one-ended transposition of IS*91*Δ16 to pSU4053. Strain UB5201 (pSU4053 + pSU2572) was mated with UB1637, and recombinant plasmids were selected on Cm and Sm plates. (*Upper*) Physical map of pSU2572 (IS*91*Δ16 is the 16-bp deletion derivative of IS*91* present in pSU2572; see Fig. 2). (*Lower*) Map of eight different recombinants that were analyzed in detail. Each recombinant is represented by an insertion fragment in a given site of pSU4053; within the insert, open boxes represent IS*91*Δ16 sequences and lines represent vector sequences. Structures of the recombinants were determined as follows: pSU4053 contains 18 *Bgl* I sites, while pSU2572 contains none. The recombinants always contained one missing *Bgl* I site of pSU4053, which was replaced by a larger fragment. *Bgl* I digestion profiles gave us a first approximate size of the inserted DNA. A second digestion (*Eco*RI and *Pvu* I) produced 0, 1, or 2 copies of a 3.9-kb fragment that represents a complete sequence of pSU2572. Assuming that one junction site is always IR_R (as shown in Fig. 4), the copy number of this segment equals the number of times that the donor sequence is repeated in the recombinants. Genes: *bla*, β-lactamase; *cat*, chloramphenicol acetyltransferase; *rep*, replication region of the plasmid; TRA_W, conjugal transfer region of the IncW plasmid R388. Restriction enzyme sites: C, *Cla* I; E, *Eco*RI; V, *Pvu* I; vertical lines without a symbol, *Bgl* I.

instance (bp 16243) the target site also starts with ATC—that is, the same sequence as IR_R. It is possible, therefore, that target site selection is influenced by similarity to IR_R. (*ii*) The other junction (variable end) always occurs 3′ to a CTTG/GTTC sequence in the donor. In pSU2572, there are only 9 GTTC sequences and 10 CTTG sequences in the right orientation. Therefore, the divergence seen in the six recombinants analyzed suggests that the variable end is probably selected with no more specificity than the recipient target site. It is perhaps significant that the GTTC sequence adjacent to IR_R was selected only once (pSU2573) and that an alternative site with obviously less similarity to IR_R (bp 772) was selected twice. Due to the scarcity of potential variable ends, this double occurrence probably represents just a coincidence. (*iii*) No duplications or deletions are observed in the target DNA.

DISCUSSION

Transposons studied in detail move by one of two proposed mechanisms (2): phage Mu and class II transposons move by a replicative concerted mechanism; Tn*7* and IS*10* move by a cut and paste conservative mechanism. We would like to propose that IS*91* is the first representative of a third strategy: a replicative, sequential mechanism that can be called

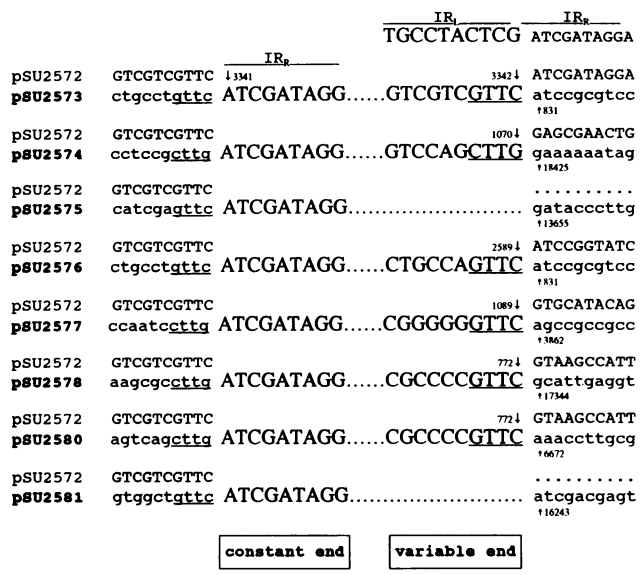

FIG. 4. DNA sequences at junctions between pSU2572 and pSU4053 in the plasmid recombinants shown in Fig. 3. The following strategy was used for sequencing. Plasmid recombinants were initially digested with *Cla* I. Those recombinants in which the *Cla* I site at the constant end junction was methylated (pSU2574, pSU2577, pSU2578, and pSU2580) were self-annealed and the constant end junction was sequenced by using an IR_R primer. The remainder were cloned in pSU40 and sequenced using the M13 primers. In this way, we obtained the sequence of the IR_R junction and the target site. From the known sequence of the target, a suitable restriction fragment could be chosen for subcloning and sequencing IR_L in the six cases shown. Capital letters represent pSU2572 sequences; lowercase letters represent pSU4053 sequences. For each recombinant, the DNA sequences in pSU2572 immediately flanking the insertion sites (and absent in the recombinant) are shown above for comparison. pSU2572 and pSU4053 sequences are numbered starting at the single *Eco*RI site in each plasmid (see Fig. 3). The DNA sequences corresponding to the complementary strand of IS*91* are represented, so that the junction starting at IR_R (constant end) lies on the left. The constant end is always adjacent to a GTTC or CTTG of the recipient DNA sequence (underlined; note that these are the sequences complementary to CAAG or GAAC). Variable end finishes with GTTC or CTTG. Sequences of IR_R and IR_L are shown above these sequences to allow inspection of possible similarities.

RC transposition. Three lines of evidence suggest it: (*i*) differential roles of IR_L and IR_R in IS*91* transposition, (*ii*) multiple tandem insertions of plasmids containing derivatives of IS*91* with deletions of IR_L, and (*iii*) sequence similarities between IS*91* and related transposases and RCR proteins, as well as between the IR_Rs and RCR plasmid origins of replication (see ref. 17 for references on RCR plasmids).

The work reported here shows that IR_R plus the flanking GTTC/CTTG sequence is absolutely required for transposition. A deletion that leaves 14 bp of IR_R is completely Tnp^-. All or most of an 80-bp sequence of the IR_R are required for full activity. On the other hand, IR_L is dispensable. When it is missing, one-ended transposition occurs at roughly the same frequency as cointegrate formation by the wild-type element. The products of one-ended transposition are plasmid fusions that start precisely at IR_R but apparently fail to terminate correctly, resulting in multiple tandem insertions of the donor plasmid. These structures were never seen with other transposons. The sequences at the "termination sites" are simple GTTC or CTTG sequences—that is, they look like recipient target sites or relaxed initiation sites. Interestingly, very similar products are observed in a system of forced deletion formation in the case of the RCR plasmid pC194 (18). Deletions shared both a common endpoint (the nick site in pC194 *ori*) and a variable endpoint whose DNA sequence was less homologous to the nick site.

It is of interest to compare RC transposition of IS*91* with one-ended transposition mediated by Mu and class II transposons. One-ended transposition has not been reported for IS elements or for class I transposons (1), but it has been described for Mu and class II transposons (19, 20). In these cases, the frequency of one-ended transposition is low, and the products fail to terminate at a fixed site. Furthermore, attempts were made at constructing variants of the IRs of Tn*21* (a class II transposon that shows a relatively high frequency of one-ended transposition) that could function in transposition termination but were unable to initiate (i.e., Tn*21* variants that give normal transposition when associated with a wild-type IR but do not produce one-ended transposition). These attempts were unsuccessful; all variant IRs tested were similarly active in both assays, which suggested that both IRs are involved in the rate-limiting step in normal transposition (21). In the case of IS*91*, the crucial difference is that IR_L is incapable of one-ended transposition but it is a very good transposition "terminator." This is likely due to the basic difference between a concerted replicative transposition mechanism and a truly sequential mechanism of replicative transposition.

On the basis of these observations we propose that IR_R is an origin of RC transposition. In this context, there are commonalities in overall structure between RC transposon origins and origins of replication of RCR plasmids. First, the nicking site has the conserved sequence GTTC/AT (in the transposon IRs it can be also CTTG/AT). Second, adjacent to the 3′ end of the nicking site, there is a conserved inverted repeat. The replication termination signal of pC194 lies 5′ to the nicking site (22). The RC transposons should, however, differ in this respect in that IR_L is likely to contain the transposition termination signals. The functional similarity between the nicking/replication origins is strengthened by the previously reported sequence similarities between the transposases of IS*91*, IS*801*, and the replication proteins of the pC194/pUB110 family of RCR plasmids (4).

Our observations can be integrated in a model for the mechanism of IS*91* transposition (Fig. 5). We propose that IS*91* transposition shares a number of steps with the synthesis of single-stranded DNA from the replicative form of phage øX174, and that IS*91* transposase is a functional analog of the nicking protein gpA of the phage (see ref. 23 for a review on the biochemical activities of gpA). Therefore, the trans-

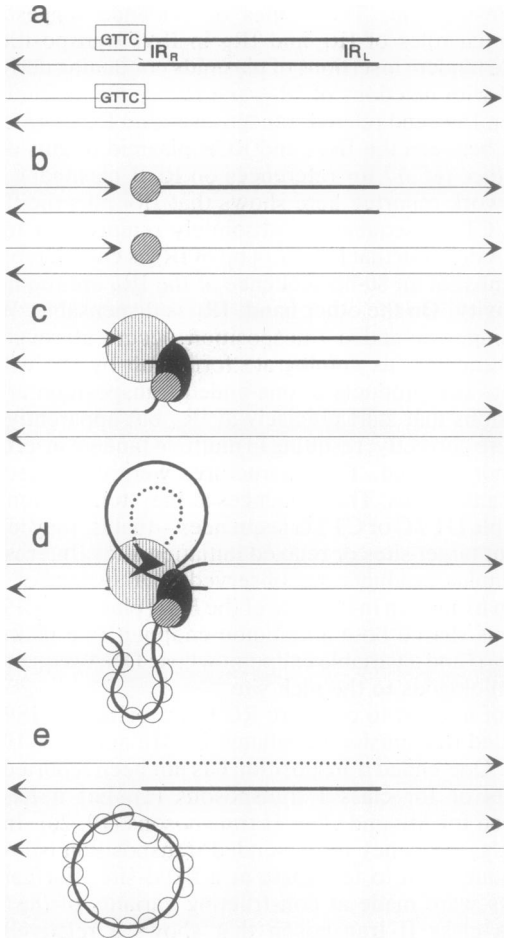

FIG. 5. A model for IS*91* RC transposition. For each step of the reaction, both strands of the donor (top lines) and recipient (bottom lines) DNAs are shown as linear DNAs, with arrowheads pointing to the 3′ ends of the molecules. Thick continuous lines represent preexisting IS*91* sequences, dotted lines represent newly synthesized IS*91* sequences, and thin lines represent nontransposon sequences. Two transposase molecules (small hatched circles) bind, one to the IR_R of IS*91* and the other to the target site (step a). Both cut a single strand at the 3′ site of their respective GTTC (or CTTG) sequences and presumably bind covalently to the resulting 5′ ends (step b). The transposase molecule bound to IR_R catalyzes the first strand transfer event, which binds the 5′ end of the nicked IR_R to the 3′ end of the nicked target DNA. The transposase molecule bound to the target DNA complexes to the Rep helicase (ellipse), which starts unwinding the 5′ → 3′ strand of the donor IS*91* sequences (step c). The displaced strand is coated by the *E. coli* single-stranded DNA-binding protein (small open circles). DNA unwinding allows the loading of the DNA polymerase III holoenzyme (large shaded circle), which synthesizes the leading strand of IS*91* by RCR (step d). Upon arrival to IR_L, the transposase verifies the second strand transfer event by nicking the 3′ end of IR_L and linking it to the 5′ end of the nicked target site. Overall result of the transposition reaction is the transfer of one single preexisting strand of IS*91* to the target site of the recipient DNA (step e). This looped DNA can be passively replicated by the bacterial replication machinery.

posase will have nicking and closing activities and will direct DNA strand displacement upon binding to the *E. coli* Rep helicase. Our model can be compared with the replicative sequential transposition model of Galas and Chandler (24). The basic difference is that in our model the transposon target DNA is nicked in only one strand, which leads to a different geometry for the second strand transfer event. One simple possibility is that a dimer of transposase forms the initial complex that brings together the donor and recipient DNAs. Two nicks are produced, and the first strand transfer event occurs early in the reaction. The transposase molecule bound to the 5′ end of the nick of the recipient DNA directs strand displacement as well as the second strand transfer event, which involves the same donor and recipient DNA strands as the first one. One-ended transposition occurs by the same series of steps except that in step d the displaced DNA strand is nicked at the 3′ end of an arbitrary GTTC (or CTTG) sequence, and this 3′ end is ligated to the 5′ target end bound to the transposase.

In summary, site-specific nicking proteins involved in initiation of RCR appear to be very plastic enzymes that have been used in nature in a variety of DNA replication processes (25): vegetative replication of RC plasmids from Gram-positive and Gram-negative bacteria and of single-stranded DNA phages in many different bacteria, eukaryotic viruses infecting cells of different kingdoms, conjugal transfer and mobilization of many bacterial plasmids, and, as our study suggests, IS*91* transposition. Replication, conjugation, and transposition are therefore processes that can be mechanistically compared. Further study of the structure–function relationships of the nickase proteins responsible for these specificities is a very exciting field to be explored.

We thank Matilde Cabezas for excellent technical assistance. This work was supported by a grant (BI093-0371) from the Comision Interministerial de Ciencia y Tecnologia (Spain).

1. Berg, D. E. & Howe, M. (1989) *Mobile DNA* (Am. Soc. Microbiol., Washington).
2. Mizuuchi, K. (1992) *Annu. Rev. Biochem.* **61,** 1011–1051.
3. Mendiola, M. V. & de la Cruz, F. (1989) *Mol. Microbiol.* **3,** 979–984.
4. Mendiola, M. V. & de la Cruz, F. (1992) *Nucleic Acids Res.* **20,** 3521 (abstr.).
5. Mendiola, M. V., Jubete, Y. & de la Cruz, F. (1992) *J. Bacteriol.* **174,** 1345–1351.
6. Romantschuk, M., Richter, G. Y., Mukhopadhyay, P. & Mills, D. (1991) *Mol. Microbiol.* **5,** 617–622.
7. de la Cruz, F. & Grinsted, J. (1982) *J. Bacteriol.* **151,** 222–228.
8. Bartolomé, B., Jubete, Y., Martínez, E. & de la Cruz, F. (1991) *Gene* **102,** 75–78.
9. Díaz-Aroca, E., de la Cruz, F., Zabala, J. C. & Ortiz, J. M. (1984) *Mol. Gen. Genet.* **193,** 493–499.
10. Bolland, S., Llosa, M., Avila, P. & de la Cruz, F. (1990) *J. Bacteriol.* **172,** 5795–5802.
11. Datta, N. & Hedges, R. W. (1972) *J. Gen. Microbiol.* **72,** 349–355.
12. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
13. Prentki, P. & Krisch, H. M. (1984) *Gene* **29,** 303–313.
14. Oka, A., Sugisaki, H. & Takanami, M. (1981) *J. Mol. Biol.* **147,** 217–226.
15. Korneluk, R. G., Quan, F. & Gravel, R. A. (1985) *Gene* **40,** 317–323.
16. Rose, R. E. (1988) *Nucleic Acids Res.* **16,** 355 (abstr.).
17. del Solar, G., Moscoso, M. & Espinosa, M. (1993) *Mol. Microbiol.* **8,** 789–796.
18. Ballester, S., Lopez, P., Espinosa, M., Alonso, J. C. & Lacks, S. A. (1989) *J. Bacteriol.* **171,** 2271–2277.
19. Groenen, M. A. M., Kokke, M. & van de Putte, P. (1986) *EMBO J.* **5,** 3687–3690.
20. Motsch, S., Schmitt, R., Avila, P., de la Cruz, F., Ward, E. & Grinsted, J. (1985) *Nucleic Acids Res.* **13,** 3335–3342.
21. Martín, C., Grinsted, J. & de la Cruz, F. (1989) *J. Bacteriol.* **171,** 3996–4001.
22. Gros, M. F., te Riele, H. & Ehrlich, S. D. (1987) *EMBO J.* **6,** 3863–3869.
23. Kornberg, A. & Baker, T. (1992) *DNA Replication* (Freeman, New York), 2nd Ed.
24. Galas, D. J. & Chandler, M. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 4858–4862.
25. Ilyina, T. V. & Koonin, E. V. (1992) *Nucleic Acids Res.* **13,** 3279–3285.