JOURNAL OF BACTERIOLOGY, Dec. 1997, p. 7551–7558 Vol. 179, No. 23
0021-9193/97/$04.00+0

Terminal Inverted Repeats of Insertion Sequence IS*30* Serve as Targets for Transposition

FERENC OLASZ,[1,2]* TIBOR FARKAS,[2] JÁNOS KISS,[2] ACHILLE ARINI,[1] AND WERNER ARBER[1]

Department of Microbiology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland,[1] and Institute for Molecular Genetics, Agricultural Biotechnology Center, H-2101 Gödöllő, Hungary[2]

Received 12 May 1997/Accepted 22 September 1997

In the present study, we demonstrate that the terminal inverted repeats of the *Escherichia coli* insertion sequence IS*30* are functional target sites for the transposition of the $(IS30)_2$ dimer, which represents an intermediate structure in the transposition of IS*30*. Comparative analysis of various target regions revealed that the left and right ends differ in their "attractivity." In our experiments, the joined left and right ends, i.e., the $(IS30)_2$ intermediate structure, was found to be the most preferred target. It was also shown that flanking sequences can influence the target activity of the terminal repeats. The functional part of the target region was localized in the inverted repeats by means of mutational analysis, and it corresponds to the binding site of IS*30* transposase. Insertion of 1 bp into the right inverted repeat resulted in unusual target duplication accompanied by gene conversion. The choice of the terminal inverted repeats as targets in transposition leads to the reconstruction of the $(IS30)_2$ structure, which may induce a cascade of further rearrangements. Therefore, this process can play a role in the evolution of the genome.

The insertion sequence IS*30* (Fig. 1A) resides in the K-12 and C strains of *Escherichia coli*. The left and right terminal inverted repeats (IR) of the element are both 26 bp long, and they differ at three positions (2). IS*30* contains one large open reading frame (ORF-A) coding for the 44.3-kDa transposase protein (4). The 17-kDa N-terminal part of the transposase (ORF-A-N) was shown to specifically bind to the left and right IRs (LIR/RIR) of the element (31).

IS*30*, IS*21*, and IS*2* belongs to a group of insertion sequences (IS) in which the transposition of the element is known to occur via an $(IS)_2$ intermediate (3, 8, 22, 25–27, 32). The $(IS30)_2$ structure is formed by site-specific dimerization (Fig. 1B) and contains two elements that are joined to each other so that their LIR and RIR are separated by 2 bp or, occasionally, 1 or 3 bp (19). This intermediate is very active and plays a central role in the transpositional reactions characteristic of IS*30*: transpositional fusion (Fig. 1B), deletion, and inversion are all mediated by this structure (22).

Many insertion sequences (e.g., IS*4* [15], IS*5* [7], and IS*10* [9, 12]) have a pronounced target specificity; i.e., they frequently transpose into target regions that can be characterized by a consensus sequence (for a review, see reference 10). In the case of IS*30*, the palindromic 24-bp consensus sequence is unusually long and highly degenerate (see Fig. 6) (21). Although it was published previously that the inverted repeats of some elements show partial homology to their preferred target sequences (e.g., IS*1*/Tn*9* [11], IS*26*/IS*6* [17], IS*911* [24], Tn*5* [1, 14], and Tn*3* [33]) and thus might function as hot spots for transposition, this hypothesis has not yet been verified experimentally.

In this paper, we show that IS*30* transposes into sequences containing the left or right IR of the element. The analysis of various target sequences revealed that the 26-bp IRs themselves are functional targets and that their attractivity can be influenced by the bordering sequences. The core sequence, which is responsible for the target activity of the IRs, was determined by mutational analysis and corresponds to the binding site of IS*30* transposase.

* Corresponding author. Mailing address: Institute for Molecular Genetics, Agricultural Biotechnology Center, Szent-Györgyi A. u. 4, H-2101 Gödöllő, Hungary. Phone: (36) 28 430 600. Fax: (36) 28 430 416. E-mail: olasz@hubi.abc.hu.

MATERIALS AND METHODS

Bacterial strains, media, chemicals, and DNA techniques. *E. coli* JM109 (35) was used as the host in the transpositional fusion experiments. Luria-Bertani medium (16) was supplemented with antibiotics to final concentrations of 20 μg of kanamycin per ml and 150 μg of ampicillin per ml. Commonly used DNA techniques were performed by the method of Sambrook et al. (28). DNA sequences were determined by the dideoxynucleotide chain termination method (29) with modified T7 polymerase (Sequenase version 2.0; U.S. Biochemicals). Computer analysis was performed with the Genetics Computer Group software package (6). PCR (18) conditions consisted of 35 cycles of 94°C for 1 min, 45°C for 1 min, and 72°C for 1 min. The reaction was carried out with a Progene apparatus (Techne Ltd.). Chemicals were obtained from Promega, Madison, Wis.

Transpositional fusion. Transpositional fusion products were detected as described by Olasz et al. (22). In these experiments, the target plasmids and their rearranged derivatives were packed into infectious phage particles as single-stranded DNA. After infection of the host bacteria and appropriate selection (note that donor plasmids were Km^r and target plasmids were Ap^r), fusion products were detected as Km^r Ap^r transductants. The frequency of transpositional fusion was calculated as follows. First the ratio of Km^r Ap^r to Ap^r colonies was determined. Then Km^r Ap^r colonies were tested by restriction analysis, and the ratio of confirmed fusion products to total tested Km^r Ap^r colonies was used for correction of the frequency data.

Plasmids. In our experiments, pAW1039 (22) carrying a dimer $(IS30)_2$ structure is called the donor plasmid and the derivatives of pEMBL19 (5) that contain different target regions for IS*30* transposition are referred to as target plasmids (Fig. 2).

Target plasmids pAW758 and pAW759 were derived from the transpositional fusion products of pAW1039 and pAW782, in which IS*30* had transposed into the target region LHS (lambda phage hot spot) (21, 22) similarly to the scheme in Fig. 1B. Likewise, pFOL701 and pFOL703 resulted from the fusion of pAW1039 and pAW1063, which contained the synthetic target region OHS (consensus oligonucleotide hot spot, 5′-CAAAAAATGCAGCTGCATTTTTTG-3′). In the constructs where the IS*30* ends are joined to LHS or OHS, only half of the original hot spot is present, since IS*30* is inserted into the center of the hot spot sequence. So as to render their comparison possible, these constructs always contained the same half of LHS or OHS.

Target plasmids pAW1005 and pAW1017 were cloned from pAW305 (31). pAW1059, which carries an intact IS*30* element, and pAW1105, which contains a truncated dimer structure, were cloned from pAW305 (31) and pAW136 (3), respectively.

pTFA303 and pTFA305 carried the 26-bp LIR and RIR as cloned PCR products, respectively. The LIR (oligonucleotide pair, 5′-GTTTTCACAGATC TGTTGCG-3′ and universal forward primer) was cloned as a *Bgl*II-*Eco*RI frag-

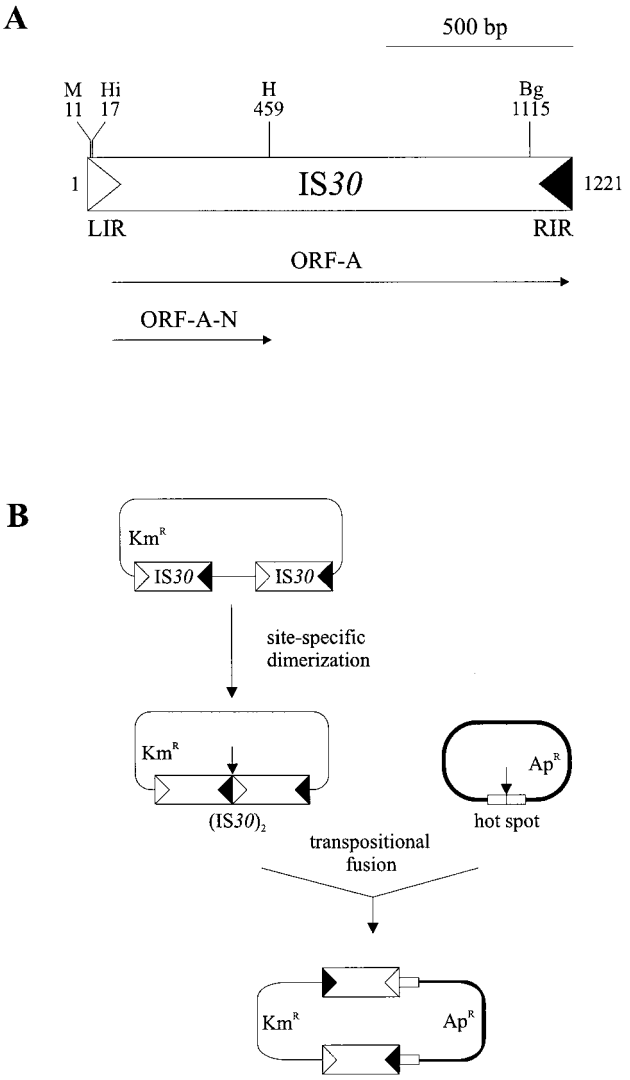

FIG. 1. (A) Structure of IS*30*. (B) Schematic representation of the site-specific dimerization and transpositional fusion mediated by the element. IS*30* is represented by a box containing the LIR (open triangle) and RIR (solid triangle). ORF-A and ORF-A-N are the ORFs of the intact IS*30* transposase and the truncated N-terminal part of the protein, respectively. The restriction cleavage sites are shown above the IS*30* sequence: *Bgl*II (Bg), *Hinc*II (Hi), *Hin*dIII (H), and *Mfe*I (M) (the numbers refer to the IS*30* positions) (4). Symbols in panel B: thin line, IS*30*-donor replicon; thick line, target replicon; small open box, hot spot for IS*30* insertion; $(IS30)_2$, intermediate structure (see the text); small arrow, the cleavage sites in transposition. Abbreviations: Ap^r, ampicillin resistance gene (β-lactamase); Km^r, kanamycin resistance gene (aminoglucoside-3′-phosphotransferase).

ment into the *Bam*HI-*Eco*RI site, and the RIR (oligonucleotide pair, 5′-TAAT TGAAAAGCTTGTTGCA-3′ and universal reverse primer) was cloned as a *Hin*dIII fragment into the *Hin*dIII site of pEMBL19. In both reactions, pAW1058, which differs from pAW1059 only by containing the intact IS*30* element in the opposite orientation, was used as the template. Plasmids with two target sites (Fig. 2) were constructed by multiple cloning steps from plasmids carrying one target region.

Origin of mutations. The insertion mutant mut300 was constructed in vitro by using the mutant oligonucleotide 5′-GACAGATTGAATTCTACAgcctgc-3′ (capital letters refer to the IS*30* sequence; the inserted 1 bp is underlined). Mutations mut2, mut4, mut73, and mut307 were deviant products from the PCR amplification of the ends. Deletion mutants were obtained from the digestion of the IRs and the subsequent cloning of the resulting two fragments into pEMBL19: d782 and d744 as *Eco*RI fragments derived from mutant mut300 into the *Eco*RI site; d643 and d645 as *Mfe*I fragments into the *Eco*RI site; d644 and d489 as *Hinc*II fragments into the *Hinc*II site; and d647 and d780 as *Hinc*II fragments into the *Sma*I site.

RESULTS

The terminal IRs of IS*30* serve as targets. In the following experiments, we took advantage of the fact that the $(IS30)_2$ intermediate structure undergoes transpositional fusions with appropriate target plasmids at relatively high frequencies. In our experimental setup, the plasmid containing the dimer structure (pAW1039) was used as the donor and the IRs were carried by target plasmids in various constructions. The products of transpositional fusion could be identified, after phage-mediated transduction, by their antibiotic resistance spectrum (donor, Km^r; target, Ap^r; fusion product, Km^r Ap^r [see Materials and Methods]). A typical reaction is shown in Fig. 3A and B.

The target region was defined as a sequence where IS*30* inserted into the same position in independent experiments. Target regions consisted of two parts (Fig. 2), the IR itself and the flanking sequences: multicloning site (MCS), IS*30* hot spot (LHS or OHS [Fig. 2 and Materials and Methods]), or another IS*30* end, representing the truncated $(IS30)_2$ structure.

All constructs containing IRs served as targets for the transposition of $(IS30)_2$. The frequency of transpositional fusion between the donor and target replicons varied between 2×10^{-6} and 5×10^{-4} (Table 1). Restriction cleavage analysis of the rearranged products showed that the majority of the insertions were found in the target sequences, whereas in the remaining cases, multiple rearrangements (deletion and/or inversion in the fusion product) or illegitimate recombination events occurred. LHS and OHS themselves behaved as target regions in the transpositional fusion reactions (Table 1). In contrast, the multicloning site (MCS) of the vector pEMBL19 was not recognized as a target (negative control).

Both the left and right ends were functional targets for transpositional fusion (Table 1). In this reaction, a strong orientation effect was observed: the RIR always joined to the LIR and vice versa. A similar phenomenon was observed in earlier studies investigating site-specific dimerization (19, 22). The $(IS30)_2$ structure was reconstructed in the reaction due to the joining of the LIR and RIR (Fig. 3B). Due to the presence of an unstable $(IS30)_2$ intermediate, the fusion products took part in further reactions and segregated rapidly (the most common segregation product of pAW1072 is shown in Fig. 3C). This result is consistent with the previously observed high efficiency of site-specific deletion formation in analogous structures (22).

The fact that the targets composed of the left or right IS*30* end and a multicloning site as flanking sequence were also efficient in transpositional fusion (Table 1) indicates that specific flanking regions are not essential for target activity.

To investigate whether the IRs themselves are recognized as target sites, plasmids containing only the synthetic 26-bp long LIR or RIR (flanked by MCS sequence [Table 1]) were constructed. Both of them were functional targets in the transpositional fusion reactions, demonstrating that no other IS*30*-derived region is required in this process.

On the basis of these results, it was not surprising that the whole IS*30* element also served as a target for transpositional fusion (Table 1). In this reaction, insertions into the LIR were detected more often (17 of 24 insertions) than those into the RIR (7 of 24), suggesting that there is bias in the target activity of the two IRs. This finding was not in agreement with the data of Table 1, which might indicate that the experimental setup, where the target replicons contain one target region, is not

Plasmid	Schematic structure	Relevant properties	Reference of source
Donor plasmid			
pAW1039		Km^R pACYC177 derivative containing (IS*30*)$_2$ intermediate	(22)
Plasmids with one target site			
	Target region		
pEMBL19	MCS	Ap^R cloning vector	(5)
pAW782	LHS	Hot spot from lambda bacteriophage (33244-33498 bp) cloned into pEMBL19	(21, 22)
pAW1063	OHS	Synthetic oligonucleotide hot spot cloned into pEMBL19	(21)
pAW759	LHS-left end	Left part of LHS (33244-33290 bp) and the joining IS*30* left end (1-459 bp) cloned into pEMBL19	this work
pAW758	LHS-right end	Left part of LHS (33244-33290 bp) and the joining IS*30* right end (1221-1115 bp) cloned into pEMBL19	this work
pFOL703	OHS-left end	Half OHS (1-13 bp) and the joining IS*30* left end (1-459 bp) cloned into pEMBL19	this work
pFOL701	OHS-right end	Half OHS (1-13 bp) and the joining IS*30* right end (1221-1115 bp) cloned into pEMBL19	this work
pAW1017	MCS-left end	IS*30* left end (1-459 bp) cloned into pEMBL19	this work
pAW1005	MCS-right end	IS*30* right end (1221-1115 bp) cloned into pEMBL19	this work
pTFA303	MCS-LIR	IS*30* left inverted repeat (1-26 bp) cloned into pEMBL19	this work
pTFA305	MCS-RIR	IS*30* right inverted repeat (1221-1196 bp) cloned into pEMBL19	this work
pAW1105	(IS*30*)$_2$*	Joined right (1115-1221 bp) and left (1-459 bp) IS*30* ends cloned from pAW136 (3) into pEMBL19	this work
pAW1059	MCS-IS*30*-MCS	Intact IS*30* (1-1221 bp) cloned from pAW305 (31) into pEMBL19	this work

Plasmids with two target sites	1 2	Target region 1	Target region 2	
pAW1037		LHS	LHS-left end	this work
pFOL14		LHS	MCS-left end	this work
pFOL16		LHS	LHS-right end	this work
pAW494		OHS-left end	MCS-left end	this work
pAW434		right end-OHS	right end-MCS	this work
pAW428		left end-LHS	LHS-right end	this work
pAW433		left end-OHS	OHS-right end	this work
pAW429		left end-MCS	MCS-right end	this work
pAW431		(IS*30*)$_2$*	LHS-left end	this work

FIG. 2. Schematic structure of the donor and target plasmids. Symbols: ▽, pEMBL19 MCS flanking the IS*30* IRs; □, the LHS sequence (wild-type lambda phage genome positions bp 33244 to 33290); ■, OHS sequence (bp 1 to 24); ▷□, left IS*30* end as far as the *Hin*dIII site (bp 1 to 459); □◁, right IS*30* end from the *Bgl*II site (bp 1115 to 1221); ⊠, 26-bp LIR (bp 1 to 26); ◪, 26-bp RIR (bp 1196 to 1221). (IS*30*)$_2$* represents the truncated intermediate structure. Numbers in parentheses refer to IS*30* and hot spot positions. All target plasmids are pEMBL19 derivatives.

suitable for the exact comparison of target activities, since transpositional frequencies reflect only the major differences. However, it was obvious that the truncated (IS*30*)$_2$ intermediate structure was the most attractive target (Table 1), since it received 10 to 200 times more transpositional hits than did the constructs containing only a single IS*30* end (Table 1).

Quantitative comparison of the target specificity of IRs. As indicated in the previous section, the attractivities of the left

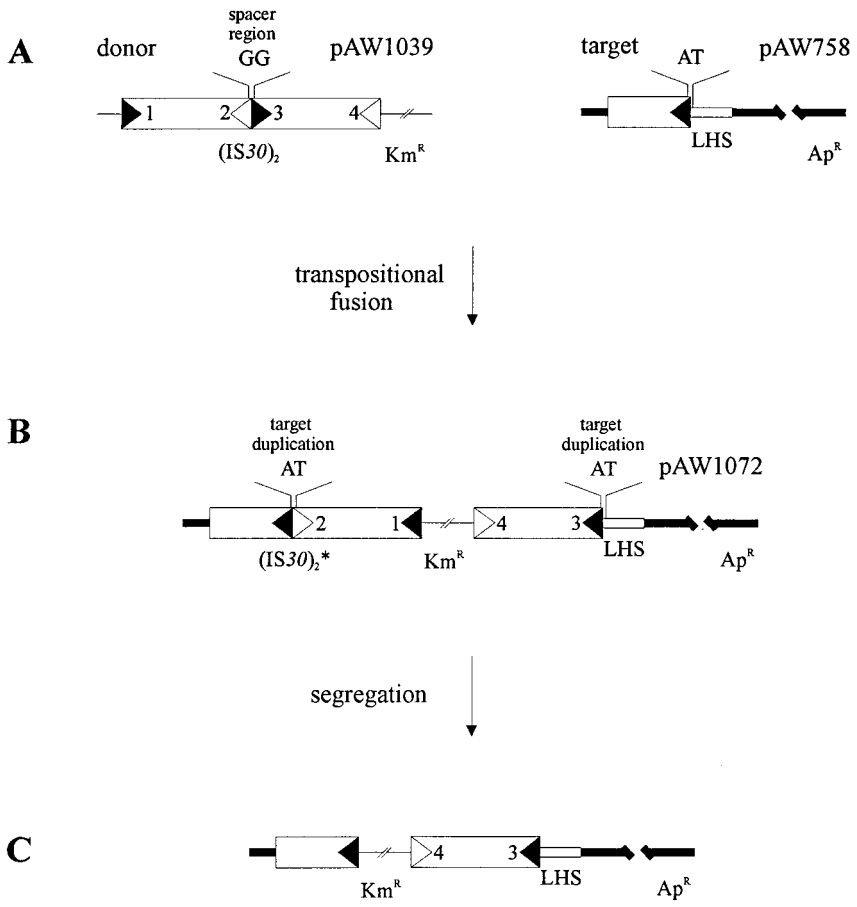

FIG. 3. Formation (A and B) and segregation (B and C) of a typical fusion product. (A) pAW1039 contains the $(IS30)_2$ intermediate structure, while pAW758 carries the right IS*30* end-LHS junction as a target region. (B) Fusion between the donor and target replicons leads to the formation of pAW1072. (C) Main segregation product of this unstable plasmid. The IS*30* ends of the donor are numbered. $(IS30)_2$* represents the newly formed, partially truncated intermediate structure. In pAW758, AT represents the bases flanking the IS*30* end that become duplicated upon transposition. Other symbols and abbreviations are the same as in Fig. 1.

and right ends of IS*30* were not equivalent when both were offered as target in the intact IS*30* element. This shows that differences in the activities of target regions become detectable when the regions are present on the same replicon.

To exploit the possibility of direct comparison, we modified our transpositional system so that the target plasmids contained two target regions. The distribution of insertions between the two sites reflects the difference in their attractivities, and the analysis of different target pairs allows ranking of the target regions.

The position and orientation effects were investigated by the following preliminary experiments. (i) The two target regions were replaced by each other. (ii) The orientation of one of the two target regions was changed. (iii) One of the two target regions was inserted into different restriction sites in the MCS. No significant alterations were observed in these experiments, indicating that the location and orientation of the target regions have no remarkable effect on their attractivity (data not shown).

In the first set of experiments, the attractivities of the left or right IS*30* ends (and their flanking regions) were compared to that of the LHS (Table 2). The distribution of transpositional events between the two target regions in pAW1037 and pFOL16, which were structurally identical except that they contained either the left (pAW1037) or the right (pFOL16) end of the element, revealed that the activity of target sequences decreased in the following order: LHS-left end > LHS > LHS-right end. The data supported our earlier conclusion that the left and right IS*30* ends are not equivalent as targets.

Likewise, pAW1037 and pFOL14 differed only in the flanking region of the left end (pAW1037, LHS-left end; pFOL14, MCS-left end). By comparing the target activities of the left end-LHS and left end-MCS sequences to that of LHS, the following order was established: left end-LHS > LHS > left end-MCS. This finding indicates that the non-IS*30*-derived bordering sequence can modify the target activity.

pAW494 and pAW434 were used to investigate the effect of flanking sequences on the attractivity of IRs (Table 2). The data showed that the OHS-left end target region was twice as active as the MCS-left end region. Similarly, OHS-right end was used approximately three times more frequently than right end-MCS. These results confirmed our previous observation that the bordering sequence has an influence on the target activity of the IS*30* ends (Table 2).

The attractivities of the ends were compared in an experiment where each target plasmid carried the left and right end flanked by the same sequence: LHS, OHS, or MCS (pAW428, pAW433, and pAW429, respectively [Table 2]). In each case, the left IS*30* end was significantly more active than the right end, and the flanking sequences increased the difference between the attractivities of the left and right IS*30* ends in the

TABLE 1. Analysis of transcriptional fusions between the donor and target plasmids containing IRs in the target region

Target region[a]	Frequency of transpositional fusion (10^{-6})[b]	No. of transpositional events in target region	No. of total tested Ap^r Km^r transductants[c]
LHS-left end	18	11	12
LHS-right end	26	12	12
OHS-left end	68	25	30
OHS-right end	65	27	30
MCS-left end	2	18	22
MCS-right end	3	9	12
MCS-LIR	2	19	30
MCS-RIR	7	17	30
$(IS30)_2$*	534	48	50
MCS-IS*30*-MCS	75	24	30
MCS	<0.01	0	28
LHS	9	76	94
OHS	2	101	134

[a] For the relevant structures and abbreviations, see Fig. 2.
[b] Calculation of these values is described in Materials and Methods.
[c] Each tested Ap^r Km^r transductant colony resulted from independent experiments.

following order: MCS < OHS < LHS. It should be mentioned that the same order was determined among these sequences when their own target activities were investigated (21).

The $(IS30)_2$ structure proved to be a very effective target (Table 1). When the joined ends were compared to the left end-LHS, the most preferred target region in the experiment presented in Table 2, $(IS30)_2$ was found to be 22 times more active (Table 2). The result indicated that the $(IS30)_2$ intermediate structure is the most attractive target sequence investigated so far.

Fidelity of target duplication. In earlier experiments, two characteristics of IS*30* transposition were found: the 2-bp duplication of the target sequence and the loss of the spacer region that separates the two elements in the intermediate structure (2, 22). When IS*30* ends were offered as target, the same phenomena were observed. In the target plasmid, pAW758, AT bases flanked the IS*30* end, while in the donor plasmid, pAW1039, a GG spacer region (Fig. 3A) separated the two IS*30* copies. Upon transpositional fusion, the AT bases became duplicated and were found in both junctions of the product (pAW1072) in all five cases sequenced (Fig. 3B). Meanwhile, the GG bases of the spacer region were lost, proving that the fusion of the donor and target plasmids was a transpositional reaction and that IS*30* ends are indeed transpositional targets.

The following experimental setup was used to study the fidelity of target duplication at the IRs (Fig. 4A): the donor plasmid was the same as that used above (pAW1039), while the target plasmid pFOL703 contained a left end-OHS region with a *Pvu*II recognition site in the junction. The transpositional fusion of the two plasmids creates *Pvu*II sites at both junctions, provided that the two bases flanking the IS*30* end are precisely duplicated. Therefore, the number of *Pvu*II sites reflects the fidelity of target duplication: if there is any alteration in its length, sequence, or site, no *Pvu*II recognition site is formed in the new junctions. In 32 of 33 independent fusion products, the target duplication was precise; just one product contained a single *Pvu*II site. In this exceptional case, the bordering sequence was GG instead of GC (Fig. 4A). This analysis revealed the high fidelity of target duplication mediated by the transposition of the element.

Mutational analysis of the IS*30* ends. Mutational analysis of the IS*30* ends was performed to determine the functional part of the IS*30* IRs responsible for the target activity. Different types of mutations were investigated: deletions (d782, d643, d647, d489, d780, d644, d645, d744, mut4, and mut307), point mutations (mut73 and mut2) and an insertion (mut300) (Fig. 5A).

Deletion mutants affected in the IRs between positions 6 and 26 lost their target function. Mutation d744, in which the first 5 bases of the IR were removed (Fig. 5A), retained its activity, indicating that positions 6 to 26 themselves are able to determine the target function of the IS*30* IR. These observations were confirmed by the analysis of mutations affecting only 1 base (mut73, mut2, mut307, and mut300) or 3 bases (mut4). Alterations in the first 5 bases (mut73, mut2, and mut300) had no detectable influence on the attractivity, whereas mutations affecting the sequence between positions 12 and 22 (mut307 and mut4) inactivated the IR as a target. On the basis of these data, the borders of the functional core region of IS*30* IRs were located between positions 6 and 10 and positions 18 and 26.

The mutant mut300 contained a 1-bp insertion after position 5 of the IR and was an active target in the transpositional reactions. Since the length of the IR also changed in this case, it was interesting to study the fidelity of the target duplication mediated by mut300. When both junctions of the rearranged product were sequenced, the following dramatic alterations were observed (Fig. 4B): (i) in one junction, the length of the target duplication was reduced to 1 bp (three cases) or 0 bp (two cases), whereas (ii) in the other junction, the length of the target duplication was 2 bp, similar to the wild type. Moreover, the 1-bp insertion of mut300, originally present in the RIR of the target replicon, was found in both junctions of the fusion product. This phenomenon can be interpreted as gene conversion.

DISCUSSION

In this study, we verified that the IRs of IS*30* serve as targets for the transposition of the element. We demonstrated that the

TABLE 2. Number of transpositional events in the two target regions[a]

Target region 1		Target region 2		Plasmid
Sequence	No. of events	Sequence	No. of events	
LHS	0	LHS-left end	48	pAW1037
LHS	37	MCS-left end	2	pFOL14
LHS	38	LHS-right end	12	pFOL16
OHS-left end	22	MCS-left end	10	pAW494
OHS-right end	20	MCS-right end	7	pAW434
Left end-LHS	47	Right end-LHS	0	pAW428
Left end-OHS	25	Right end-OHS	2	pAW433
Left end-MCS	37	Right end-MCS	6	pAW429
$(IS30)_2$	43	LHS-left end	2	pAW431

[a] Target plasmids contained two sites for transpositional fusion (target regions 1 and 2). Each tested transpositional event resulted from independent experiments. In each event, only one of the target regions was involved in the fusion reaction.

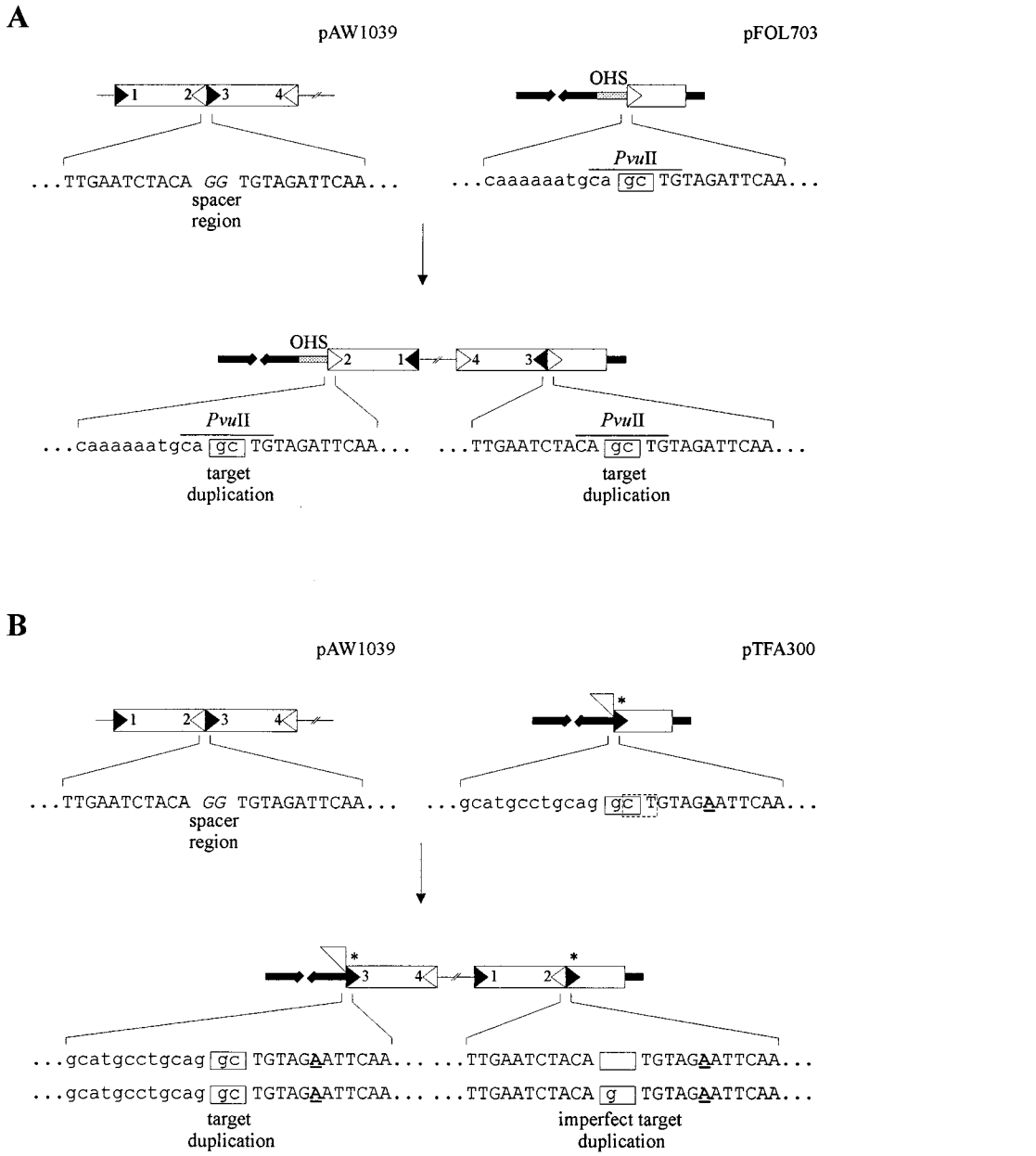

FIG. 4. Target duplication in the transpositional fusion products of the donor plasmid pAW1039 and the target plasmids pFOL703 (A) and pTFA300 (B). Capital letters indicate IS*30* sequences, lowercase letters indicate the sequence of the flanking region OHS (A) or MCS (standing triangle) (B), and italic letters indicate the spacer region between the IS*30* elements of the dimer structure. The *Pvu*II recognition sites are shown above the sequence. The 2-bp target duplications (small box) are separated by gaps. (B) The inserted A nucleotide is in bold and underlined and mutant IRs are marked by asterisks. The standing triangle represents the pEMBL19 MCS flanking the IS*30* IR. The cT sequence of the hypothetical target duplication is shown in a dotted box. Other symbols and abbreviations are the same as in Fig. 1 and 3.

26-bp IRs alone are functional targets but their attractivities can be influenced by the bordering sequences. In our direct comparative system, the left and right ends differed in their target activities: the left end was more frequently used than the right end. This finding can be explained by the three mismatches between the LIR and RIR and/or by the differences in the flanking IS*30*-derived sequences.

Experimental evidence showed that non-IS-derived flanking regions modify the attractivity of the offered IS*30* end. It was also demonstrated, however, that the presence of a bordering hot spot was not required, since it could be replaced by a random DNA sequence represented by a multicloning site lacking any target activity (Table 1).

Observations suggest that IS*30* has two different kinds of target specificity: it can transpose to its own IRs and into "natural" hot spot regions. This hypothesis is based on the fact that in our experiments the intermediate structure was the most preferred target sequence, although no significant similarities can be observed between the $(IS30)_2$ structure and the consensus target sequence. Note that the LHS was less efficient, despite its strong homology to the consensus sequence (Fig. 6). Furthermore, half of the $(IS30)_2$ junction structure,

A

Mutation	Sequence of mutant inverted repeats (positions 10, 20)	Activity as target
wild-type LIR	TGTAGATTCAATTGGTCAACGCAACA	+
wild-type RIR	TGTAGATTCAATGTGTCAATGCAACA	+
d782	TGTAGA....................	–
d643	TGTAGATTCAATT.............	–
d647	TGTAGATTCAATTGGTC.........	–
d489	TGTAGATTCAATTGGTC**G***AC......	–
d780	..................ACGCAACA	–
d644	..............GTC**G***ACGCAACA	–
d645	.........AATTGGTCAACGCAACA	–
d744	.....ATTCAATGTGTCAATGCAACA	+
mut4	TGTAGATTCAA.TGGTCAAC..AACA	–
mut307	TGTAGATTCAATTGGT.AACGCAACA	–
mut73	**C***GTAGATTCAATTGGTCAACGCAACA	+
mut2	TGT**C***GATTCAATTGGTCAACGCAACA	+
mut300	TGTAG**<u>A</u>***ATTCAATGTGTCAATGCAACA	+

B

```
                10        20        30
           .    .    .    .    .    .    .
IS30 left IR  TGTAGATTCAATTGGTCAACGCAACAgttatgtgaaaa
              ACATCTAAGTTAACCAGTTGCGTTGTcaatacactttt
```

FIG. 5. Localization of the functional part of IS*30* IRs responsible for target activity (A) and its correlation with the result of the transposase-binding assay (B). (A) Data of mutational analysis. In the right-hand column, the plus sign refers to the functional targets and the minus sign means that no transpositional fusion was detected in at least 20 independent experiments. Numbers above the sequences are LIR positions; boldface letters with asterisks are point mutations, and dots are deletions. The inserted A nucleotide in mut300 is underlined and marked by an asterisk. (B) Sequence of the left IS*30* end that was protected by the N-terminal part of IS*30* transposase in DNase I footprint experiments (31). Capital and lowercase letters indicate the LIR and other IS*30* sequences, respectively. The protected DNA sequence is marked with lines.

```
(IS30)2                     TTGAATCTACA nn TGTAGATTCAA
                            .  ::    :       :   ::  .
Consensus target sequence   YAAAAAWnGCn RY nGCnWTTTTTR
                            .:::: . ::  ..  :  . ::: .
LHS                         TAAAAGTGGCG AT AGTGAATTTAG
```

FIG. 6. Comparison of the (IS*30*)$_2$ structure and LHS to the consensus target sequence of IS*30*. Symbols: n, any base; R, A or G; Y, C or T; W, A or T. The consensus sequence for insertion of IS*30* is based on the alignment of more than 10 different natural target sites (21).

i.e., the LIR or RIR, is still a functional target, whereas this is not true for the natural half-target sequences (data not shown).

The fact that the (IS*30*)$_2$ structure was rebuilt in the transpositional reactions of the ends (Fig. 3B) may have an evolutionary significance at the genomic level. The newly formed intermediate may choose from two types of potential targets. (i) The reaction with an IS*30* end leads again to the formation of a very active (IS*30*)$_2$ structure that results in subsequent transpositional processes and therefore in various types of rearrangements (22). Since the (IS*30*)$_2$ structure is formed again and again in each case when the target is an IR, the choice of the IRs as targets may induce a cascade of mutations. The IS*30*-mediated burst of transpositional rearrangements in old cultures that was observed by Naas et al. (20) can represent such a cascade. (ii) The insertion into a natural hot spot yields far less active structures, with a lower probability of further rearrangements, since the active intermediate structure disappears. In this case, the final stable state could be a rearranged genome without any (IS*30*)$_2$ structure.

The cascade of mutations mediated by IS elements may increase the variability of the bacterial population harboring them. When the inverted repeats serve as targets for the transposition of the intermediate, further rearrangements and mutations are generated, and it may provide the possibility for the bacterial population to adapt to the new circumstances.

The mutant analysis of the IRs resulted in the following findings. First, the core target region of the IR corresponds to the sequence which was protected by the N-terminal part of IS*30* transposase in a DNase I footprint experiment (Fig. 5B) (31). This observation suggests that the DNA-binding domain of IS*30* transposase may play a role in the recognition of the ends as targets. Second, mutations or even removal of the first 5 bp of the IR end did not affect its target activity (Fig. 5A). This could suggest that this segment has only a spacer-like function, i.e., determines the cleavage site with its length. However, this limited role appears to be unlikely, since the comparison of various IS*30* isolates and its relatives shows a pronounced conservation in these positions (data not shown). The second argument against the simple spacer-like function was provided by the analysis of the mutation mut300, which was active as a target but whose fidelity in the target duplication altered notably (Fig. 4B and 5A). If the spacer hypothesis were true, the 1-bp insertion at position 6 should cause an altered target duplication: a CT sequence should flank the IS on both sides (dotted box in Fig. 4B). However, in the transpositional fusion product, one junction contains the original GC target sequence and in the other junction only 1 bp (G [three cases]) or 0 bp (two cases) was found instead of the wild-type 2 bp. These facts are inconsistent with the "simple-spacer" hypothesis: not only was no CT duplication detected, but also the two junctions were not uniform and differences were observed even within the same junction of the five samples. These data were in contrast to the high fidelity of target duplication observed for the wild-type IRs. This phenomenon and the gene conversion detected in the fusion product mediated by the insertional mutant mut300 are expected to open new insights into the transpositional mechanism of IS*30* transposition.

According to the dimer IS model (22) the (IS*30*)$_2$ structure and the circular form of the element, which has been recently isolated and analyzed (13), are generated during site-specific dimerization by the joining of the LIR and RIR of two or one element, respectively. The circularization of insertion sequences IS*1* (34), IS*3* (30), and IS*911* (23) and the dimerization of IS*21* (25) and IS*2* (32) raise the possibility that the IRs are also utilized as transpositional targets in these cases, since they all result in analogous structures (i.e., the joined IR ends). However, there are experimental results contradicting the above hypothesis; namely, in our experiments, the length of the spacer in the newly formed (IS*30*)$_2$-like structure was consistently 2 bp, while during site-specific dimerization it varied between 1 and 3 bp (19). Our experimental setup was not really suitable for the study of this question because the IRs involved in transpositional fusion were in *trans* while upon site-specific dimerization or circularization the reacting ends are located in *cis*; furthermore, instead of the reaction of one or two single elements, in our case it was the (IS*30*)$_2$ structure that inter-

acted with an IR. Further experiments are needed to confirm this hypothesis.

ACKNOWLEDGMENTS

We are grateful to Ágota Bakos-Nagy, Ilona Keresztúri-Könczöl, and Rosmarie Hiestand-Nauer for their highly skilled technical assistance. We thank Géza Dallmann, Ferenc Müller, and László Orbán for reading the manuscript.

This work was supported by grants T019365, T6054, F016426, and F017090 of the Hungarian Research Foundation (OTKA).

REFERENCES

1. **Bossi, L., and M. S. Ciampi.** 1981. DNA sequences at the sites of three insertions of the transposable element Tn*5* in the histidine operon of *Salmonella*. Mol. Gen. Genet. **183:**406–408.
2. **Caspers, P., B. Dalrymple, S. Iida, and W. Arber.** 1984. IS*30* a new insertion sequence of *Escherichia coli* K-12. Mol. Gen. Genet. **196:**68–73.
3. **Dalrymple, B.** 1987. Novel rearrangements of IS*30* carrying plasmids leading to the reactivation of gene expression. Mol. Gen. Genet. **207:**413–420.
4. **Dalrymple, B., P. Caspers, and W. Arber.** 1984. Nucleotide sequence of the procaryotic mobile genetic element IS*30*. EMBO J. **3:**2145–2149.
5. **Dente, L., G. Cesareni, and R. Cortese.** 1983. pEMBL: a new family of single stranded plasmids. Nucleic Acids Res. **11:**1645–1655.
6. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
7. **Engler, J. A., and M. P. van Bree.** 1981. The nucleotide sequence and protein coding capability of the transposable element IS*5*. Gene **14:**155–163.
8. **Farkas, T., J. Kiss, and F. Olasz.** 1996. The construction and characterization of an effective transpositional system based on IS*30*. FEBS Lett. **390:** 53–58.
9. **Foster, T. J.** 1977. Insertion of the tetracycline resistance translocation unit Tn*10* in the *lac* operon of *Escherichia coli* K-12. Mol. Gen. Genet. **154:**305–309.
10. **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.
11. **Galas, D. J., M. P. Calos, and J. H. Miller.** 1980. Sequence analysis of Tn*9* insertions in the *lacZ* gene. J. Mol. Biol. **144:**19–41.
12. **Halling, S. M., and N. Kleckner.** 1982. A symmetrical six-base-pair target site sequence determines Tn*10* insertion specificity. Cell **28:**155–163.
13. **Kiss, J., and F. Olasz.** Unpublished data.
14. **Lupski, J. R., P. Gershon, L. S. Ozaki, and G. N. Godson.** 1984. Specificity of Tn*5* insertions into a 36-bp DNA sequence repeated in tandem seven times. Gene **30:**99–106.
15. **Mayaux, J. F., M. Springer, M. Graffe, M. Fromant, and G. Fayat.** 1984. IS*4* transposition in the attenuator region of the *Escherichia coli phe*S, T operon. Gene **30:**137–146.
16. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Mollet, B., S. Iida, and W. Arber.** 1985. Gene organization and target specificity of the prokaryotic mobile genetic element IS*26*. Mol. Gen. Genet. **201:**198–203.
18. **Mullis, K. B., and F. A. Faloona.** 1987. Specific synthesis of DNA *in vitro* via a polymerase chain reaction. Genetics **120:**621–623.
19. **Naas, T., A. Arini, and W. Arber.** Unpublished data.
20. **Naas, T., M. Blot, W. Fitch, and W. Arber.** 1994. Insertion sequence-related genetic variation in resting *Escherichia coli* K-12. Genetics **136:**721–730.
21. **Olasz, F., J. Kiss, P. König, Z. Búzás, R. Stalder, and W. Arber.** Unpublished data.
22. **Olasz, F., R. Stalder, and W. Arber.** 1993. Formation of the tandem repeat $(IS30)_2$ structure and its role in IS*30* mediated transpositional DNA rearrangements. Mol. Gen. Genet. **239:**177–187.
23. **Polard, P., M. F. Prere, O. Fayet, and M. Chandler.** 1992. Transposase-induced excision and circularization of the bacterial insertion sequence IS911. EMBO J. **11:**5079–5090.
24. **Prere, 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.
25. **Reimmann, C., and D. Haas.** 1987. Mode of replicon fusion mediated by the duplicated insertion sequence IS*21* in *Escherichia coli*. Genetics **115:**619–625.
26. **Reimmann, C., and D. Haas.** 1990. The *ist*A gene of insertion sequence IS*21* is essential for cleavage at the inner 3′ ends of tandemly repeated IS*21* elements. EMBO J. **9:**4056–4063.
27. **Reimmann, C., R. Moore, S. Little, A. Savioz, N. S. Willets, and D. Haas.** 1989. Genetic structure, function and regulation of the transposable element IS*21*. Mol. Gen. Genet. **215:**416–424.
28. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
30. **Sekine, Y., N. Eisaki, and E. Ohtsubo.** 1994. Translational control in production of transposase and in transposition of insertion sequence IS*3*. J. Mol. Biol. **235:**1406–1420.
31. **Stalder, R., P. Caspers, F. Olasz, and W. Arber.** 1990. The N-terminal domain of the IS*30* transposase is able to interact specifically with the terminal inverted repeats of the element. J. Biol. Chem. **265:**3757–3762.
32. **Szeverényi, I., T. Bodoky, and F. Olasz.** 1996. Isolation, characterization and transposition of an $(IS2)_2$ intermediate. Mol. Gen. Genet. **251:**281–289.
33. **Tu, C. P. D., and S. N. Cohen.** 1980. Translocation specificity of the Tn*3* element: characterization of sites of multiple insertions. Cell **19:**151–160.
34. **Turlan, C., and M. Chandler.** 1995. IS*1*-mediated intramolecular rearrangements: formation of excised transposon circles and replicative deletions. EMBO J. **14:**5410–5421.
35. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.