Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 2608–2612, April 1982
Genetics

DNA sequence organization of IS*10*-right of Tn*10* and comparison with IS*10*-left

(transposon genetic organization/insertion sequence genetic organization/inverted repeat/regulation of transposition/transposon evolution)

S. M. HALLING*, R. W. SIMONS, J. C. WAY, R. B. WALSH†, AND N. KLECKNER

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138

Communicated by Mark Ptashne, December 22, 1981

ABSTRACT Tn*10* is 9,300 base pairs long and has inverted repeats of an insertion sequence (IS)-like sequence (IS*10*) at its ends. IS*10*-right provides all of the Tn*10*-encoded functions used for normal Tn*10* transposition. IS*10*-left can also provide these functions but at a much reduced level. We report here the complete nucleotide sequence of IS*10*-right and a partial sequence of IS*10*-left. From our analysis of this information, we draw the following conclusions. (*i*) IS*10*-right is 1,329 base pairs long. Like most IS elements, it has short (23-base pair) nearly perfect inverted repeats at its termini. We can divide these 23-base pair segments into at least two functionally distinct parts. IS*10*-right also shares with other elements the presence of a single long coding region that extends the entire length of the element. Genetic evidence suggests that this coding region specifies an essential IS*10* transposition function. A second, overlapping, coding region may or may not be important. (*ii*) The "outside" end of IS*10*-right contains three suggestively positioned internal symmetries. Two of these (A1 and A2) are nearly identical in sequence. Symmetry A1 overlaps the terminal inverted repeat; symmetry A2 overlaps the promoter shown elsewhere to be responsible for expression of IS*10* functions and lies very near a second characterized promoter that directs transcription outward across the end of IS*10*. Symmetries A1 and A2 may play a role in modulation of Tn*10* activity and are likely to function at least in part as protein recognition sites. We propose that the third symmetry (B) acts to prevent fortuitous expression of IS*10* functions from external promoters. The transcripts from such promoters can assume a stable secondary structure in which the AUG start codon of the long coding region is sequestered in a region of double-stranded mRNA formed by pairing between the two halves of symmetry B. (*iii*) IS*10*-left differs from IS*10*-right at many nucleotide positions in both the presumptive regulatory region and the long coding region. The available evidence suggests that Tn*10* may be older than other analyzed drug-resistance transposons and thus have had more time to accumulate mutational changes.

The tetracycline-resistance transposon Tn*10* is 9,300 base pairs (bp) long and has inverted repeats of ≈1400 bp at its ends (Fig. 1). Tn*10* is one of the original examples of a "composite" transposon in which two insertion sequence (IS)-like sequences cooperate to mediate transposition of intervening genetic material. Each of the inverted repeats of Tn*10* is a *structurally* intact IS-like element, which we have named IS*10*. All of the sites and functions required for Tn*10* transposition lie within these two IS*10* segments. However, genetic analysis has shown that the two IS*10* elements are not *functionally* identical. IS*10*-right (IS*10*-R) specifies all of the functions responsible for normal Tn*10* transposition: it can promote normal levels of transposition even when IS*10*-left (IS*10*-L) is specifically inactivated and provides the same level of complementing functions *in trans* to a mutant element as does wild-type Tn*10*. IS*10*-L, on the other hand, is functionally defective and can provide only 1–10% of the transposition activity of IS*10*-R. Genetic analysis has also established that all of the sites required for Tn*10* transposition lie within the outermost 70 bp at each end of the element (1–3).

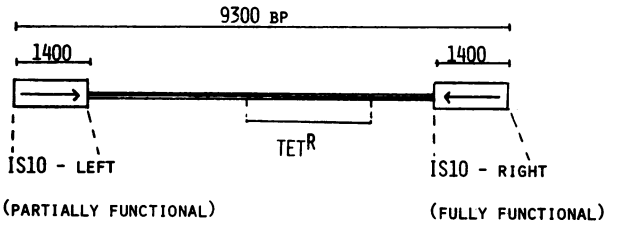

FIG. 1. Transposon Tn*10*.

We report here the entire nucleotide sequence of IS*10*-R and a partial sequence of IS*10*-L. This analysis provides information about the genetic organization of IS*10*-R and about the sequence divergence between IS*10*-R and IS*10*-L.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. Plasmids used as sources of IS*10*-R and IS*10*-L (pNK82, pNK83, pNK296, and pNK297) are described in Figs. 2–4 and Foster *et al.* (3). *Escherichia coli* strain MM294 (4) was used for preparation of plasmid DNA.

Materials. Restriction endonucleases and the Klenow fragment of *E. coli* DNA polymerase I were obtained from New England BioLabs. T4 polynucleotide kinase was purchased from Boehringer Mannheim. Radioactive nucleotides were purchased from New England Nuclear.

DNA Labeling and Sequence Determination. DNA fragments were labeled at their 5′ ends with polynucleotide kinase and at their 3′ ends with DNA polymerase I. Isolation, labeling, and sequence determination of DNA restriction fragments were as described by Maxam and Gilbert (5). Procedures for preparation, purification, and restriction endonuclease analysis of plasmid DNA are described by Foster *et al.* (3).

RESULTS

Our general approach to nucleotide sequence analysis of IS*10*-R and IS*10*-L is shown in Fig. 2. We have defined as our ref-

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: bp, base pair(s); IS, insertion sequence; IS*10*-L and IS*10*-R, IS*10*-left and -right, respectively.

* Present address: Molecular Genetics, Inc., 10320 Bren Road East, Minnetonka, MN 55343.
† Present address: Dept. of Microbiology and Molecular Genetics, Harvard Medical School, 25 Shattuck St., Boston, MA 02115.

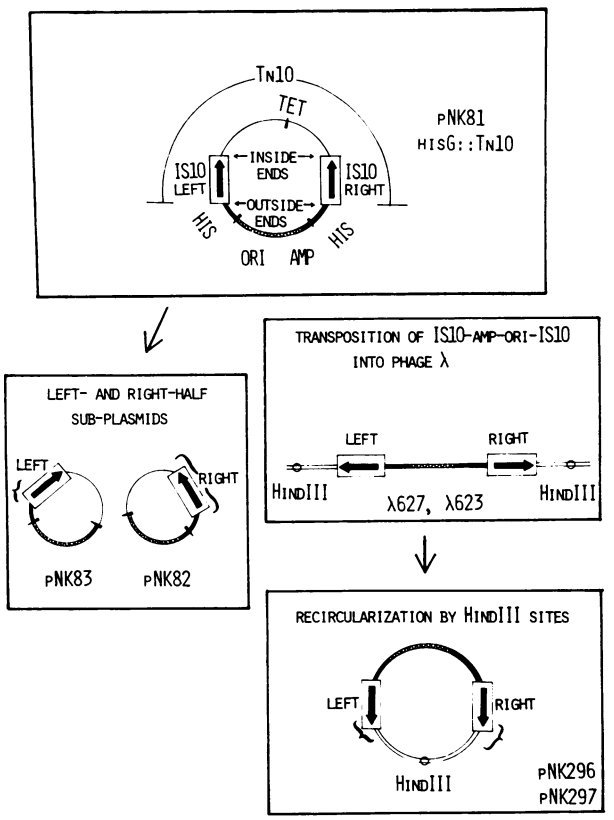

FIG. 2. General approach to sequence analysis of IS*10*-R and IS*10*-L. Tn*10* sequences are derived from a *Salmonella his*G9424::Tn*10* insertion. The sequences for all of IS*10*-R and the outer 515 bp of IS*10*-L were obtained by using a multicopy plasmid containing this insertion (pNK81) and *Eco*RI-generated right- and left-half subplasmids (pNK82 and pNK83). The inside ends of the IS*10* elements in Tn*10* were defined by using two independent transpositions of the IS*10*-*amp*^{R}-*ori*-IS*10* segment from pNK81 into the *nin* region of bacteriophage λ (λ623, λ627). Small plasmids containing only the *amp*^{R}, *ori*, and λ/IS*10* junction regions of these phages were constructed by cleavage and religation at *Hin*dIII sites that occur in λ on either side of the new inserts. The sequences across the λ/IS*10* junctions in pNK296 and pNK297 were compared with the sequence across the inside end of IS*10*-R in pNK81/82. Brackets indicate regions in pNK82, pNK83, pNK296, and pNK297 for which sequences were determined. pNK81, pNK82, pNK83, λ623, and λ627 are described in ref. 3.

erence wild-type element a particular Tn*10* insertion in the *his*G gene of *Salmonella* (3). The transposon in this insertion is separated by only two cycles of transposition from Tn*10* as originally isolated on plasmid R222. The entire sequence of IS*10*-R and the sequence of the outermost 515 bp of IS*10*-L were obtained by using plasmids containing all or part of this *his*G::Tn*10* insertion. The detailed strategy for sequence analysis of IS*10*-R is shown in Fig. 3. The outer portion of IS*10*-L was analyzed on both strands by a similar strategy (not shown).

The exact nucleotide positions of the "inside" ends of IS*10*-R and IS*10*-L have been determined by using the strategy used to define the "outside" ends of Tn*10* itself. We have previously described an "inside-out" transposon derived from the reference *his*G::Tn*10* element, in which the plasmid ampicillin-resistance (*amp*^{R}) genes are flanked by IS*10*-R and IS*10*-L elements having relative orientations opposite to those in Tn*10*. The inside ends of the IS*10* sequences as they occur in Tn*10* now form the outside ends of the new transposon as shown in Fig. 2. We have isolated insertions of this inside-out transposon at two different positions in bacteriophage λ and analyzed the nucleotide sequences across their IS*10*/λ junctions. These results are shown in Fig. 4. On the IS*10* side of each junction is a stretch of sequence identical to that near the inside end of IS*10*-R of Tn*10*, which ends at a position corresponding to bp 1329 of the IS*10*-R sequence. We interpret this result to mean that the inside end of IS*10*-R, as it occurs in Tn*10*, is at bp 1329. Further support for this conclusion comes from the fact that the inserted IS*10* material in λ627, in which both junctions were analyzed, is flanked by a 9-bp direct repeat of non-IS*10* (i.e., λ) sequence. Insertions of Tn*10* are flanked by 9-bp direct repeats of target DNA (6). λ623 and λ627 also provided the DNA sequence for the inner 150 bp of IS*10*-L.

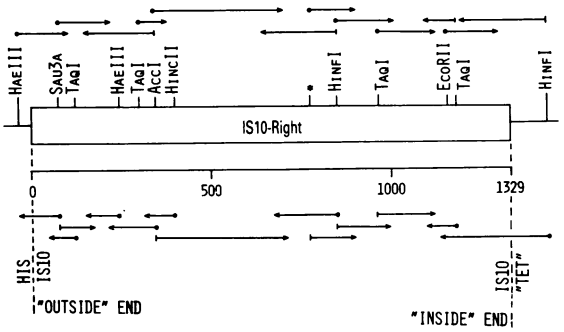

FIG. 3. Strategy for sequence analysis of IS*10*-R. A simplified restriction map of IS*10*-R and adjacent regions in pNK82 is shown in Fig. 2. Numbers are distances (bp) from the outside end of IS*10*-R. ●⟶ and ↦, Extents of sequences determined from fragments labeled at their 5′ and 3′ ends, respectively. Sequences on both strands across the *Hin*fI site at bp 850 were obtained by using a variant of pNK82 in which a *Bam*HI linker fragment had been introduced at the position indicated by * (unpublished).

The entire nucleotide sequence of IS*10*-R and the positions at which IS*10*-L differs from IS*10*-R in the regions analyzed are shown in Fig. 5. This figure also shows a number of sequence features discussed below: the 23-bp nearly perfect inverted repeats at the termini of IS*10*-R, the start and stop codons for a single long coding region, and the positions of three internal sequence symmetries.

Also shown in Fig. 5 are the positions of two promoters located near the outside end of IS*10*-R. The existence of these promoters and the definition of their exact transcription start points have been established *in vitro* by length analysis, 5′-γ-^{32}P labeling of transcripts, and abortive initiation/dinucleotide cycling analysis (unpublished results). Genetic analysis suggests that p-IN is the promoter responsible *in vivo* for expression of IS*10* transposition function(s) (unpublished results). The *in vivo* role of p-OUT is not yet known.

TARGET	9BP REPEAT	IS10	JUNCTION	INSERTION
...GACCACAAAACGCCA	CGCTCTGTA	CTGAGAGATCCCCTCATAAT...	λ/IS10-L	
...CTGGTGTTTTGCGGT	GCGAGACAT	GACTCTCTAGGGGAGTATTA...		λ627
TATACAGGCATTCGG	TACAGAGCG	CTGAGAGATCCCCTCATAAT...	λ/IS10-R	
...ATATGTCCGTAAGCC	ATGTCTCGC	GACTCTCTAGGGGAGTATTA...		
AAAGGA	GATTATGTA	CTGAGAGATCCCCTCATAAT...	λ/IS10-R	λ623
...TTTCCT	CTAATACAT	GACTCTCTAGGGGAGTATTA...		
ATTATGGGAATTGGA	GGCAATGCA	CTGAGAGATCCCCTCATAAT...	TET/IS10-R	pNK81
TAATACCCTTAACCT	CCGTTACGT	GACTCTCTAGGGGAGTATTA...		(HIS::TN10)
		1329* 1310*		

FIG. 4. DNA sequences of IS*10*-*amp*^{R}-*ori*-IS*10* insertion junctions that define the inside ends of the IS*10*s in Tn*10*. Sequences across both λ/IS*10* junctions of λ627 and the λ/IS*10*-R junction of λ623 (from pNK297 and pNK296; Fig. 2) are compared with the sequence across the inside end of IS*10*-R in Tn*10* (from pNK82; Fig. 2). Nucleotides of IS*10*-R are numbered as in Fig. 5.

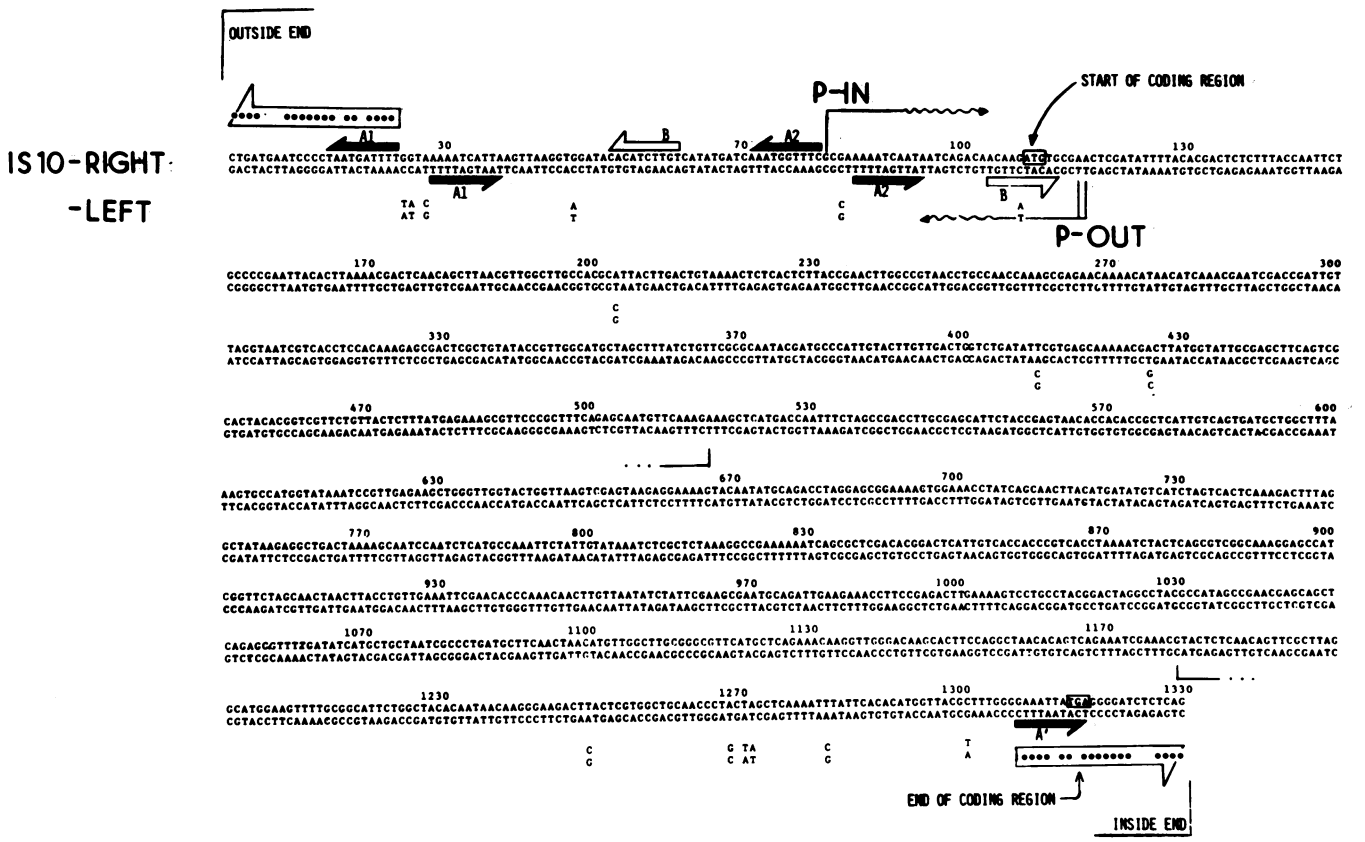

FIG. 5. DNA sequence of IS*10*-R and partial sequence of IS*10*-L. The sequence of IS*10*-R is numbered from 1 to 1,329 beginning with the *outside* end of the element. Thus, bp 1 is one of the termini of Tn*10* itself. The sequence of the top strand is written in the 5′ to 3′ direction from left to right. IS*10*-L differs from IS*10*-R at 15 positions in the regions analyzed (bp 1–515 and bp 1,179–1,329). Several features of IS*10*-R sequence are emphasized: (*i*) ⧋, the 23-bp nearly perfect homology between the inside and outside ends of IS*10*-R; ●, homologous base pairs; (*ii*) ◢, two related internal symmetries (A1 and A2) and a related 10-bp sequence at the inner terminus; (*iii*) ⧋, the two widely separated halves of symmetry B; and (*iv*) the start and stop codons of the long open coding region at bp 108 and bp 1,314, respectively. Also shown are the exact transcription start points for the promoters p-IN and p-OUT located near the outside end of IS*10*-R.

DISCUSSION

Coding Regions in IS*10*-R. IS*10*-R has a single long coding region that extends the entire length of the element from bp 108 to bp 1316 (Fig. 5). The polypeptide corresponding to this segment would be 402 amino acids long and have a M_r of 47,000. Several genetic observations suggest that this coding region specifies an essential IS*10* transposition function: (*i*) the start of the coding region is suitably positioned a few bp downstream from p-IN; (*ii*) insertion of a fragment bearing the *lac* UV5 promoter at bp 70, just upstream from the coding region, results in dramatic isopropyl-β-D-thiogalactoside-induced overproduction of *trans*-acting transposition functions (unpublished results); (*iii*) a mutation (*ins*1) created by "filling-in" the *Acc* I site at bp 340 results in a frame-shift mutation in the long coding region and produces both a recessive transposition defect and loss of the ability to provide complementing transposition functions *in trans* (ref. 3; unpublished results); and (*iv*) gene fusions between the coding region of β-galactosidase and (as yet uncharacterized) positions between bp 210 and bp 310 of IS*10*-R have been constructed, and some of these fusions exhibit β-galactosidase activity (unpublished results).

IS*10*-R contains four additional coding segments that correspond to polypeptides of >30 amino acids in length. One of these (bp 349 to bp 534; 62 amino acids) should be transcribed from p-IN; a genetic role for this segment has not been ruled out. The other three segments (bp 4 to bp 165, bp 62 to bp 208, bp 280 to bp 143) are less likely to specify relevant functions. They cannot be transcribed from p-IN or p-OUT, and extensive *in vitro* transcription and abortive initiation experiments have not detected other promoters that could transcribe these segments (unpublished results). There is no evidence bearing on the possibility that a transcript from p-OUT might be translated from one of the available AUG or GUG codons (bp 54, 56, or 64).

Short Terminal Inverted Repeats. The ends of IS*10*-R and IS*10*-L are short nearly perfect inverted repeats. The repeats are 23-bp long and have matches at 17 out of 23 positions (Fig. 5). The most important finding from inspection of these sequences is that the terminal inverted repeat sequence can be subdivided into at least two different parts.

The inner 10-bp of each repeat can be defined as one subsection. As shown in Fig. 5, this particular 10-bp sequence occurs five times near the ends of IS*10*-R. Four copies of the sequence occur in two symmetrical inverted repeats near the outside end of the element (symmetries A1 and A2) with one of these repeats (A1) containing the innermost 10 bp of the terminal inverted repeat. A fifth copy of the sequence occurs by itself in the inside terminus of the element. Because the sequence occurs in a single isolated copy at the inside terminus, without opportunity for secondary structure interactions with nearby copies, we favor the possibility that this sequence functions within the terminal inverted repeats as a protein recognition site. An alternative would be that it participates in direct pairing between the termini of the element during transposition. There is no genetic evidence as to the mechanistic or reg-

ulatory roles of this subsection of the terminal inverted repeat.

The remaining outer 13 bp of the two terminal inverted repeats constitute a second subsection. The 13-bp segments at the inside and outside ends share two continuous stretches of homology (6 bp and 4 bp) separated by a 3-bp gap. Direct evidence for the role of the 13-bp segment in Tn*10* transposition is provided by the finding that two mutations that confer a *cis*-dominant transposition-negative phenotype on Tn*10* have altered nucleotides in the 6-bp homology region (unpublished results). The 4-bp homology, at the very ends of the element, is identical in sequence to the very ends of the IS*50* elements of Tn5. The 13-bp subsection may itself be further divisible.

Role of Symmetries Near the Outside Ends of IS*10*-R. The outside end of IS*10*-R contains three suggestively located internal symmetries. Symmetries A1 and A2 contain the same (or closely related) 10-bp half-sites; the two half-sites in each symmetry are separated by 4 bp. Symmetry A1 is centered about bp 25/26; its outer half overlaps the terminal inverted repeat. Symmetry A2 is centered about bp 82/83; it is located directly on top of the transcription start point for the promoter p-IN at bp 81. Symmetry B involves a second, unrelated, 10-bp sequence. Two copies of this sequence form the widely separated halves of an inverted repeat whose center of symmetry is bp 82. The two halves of this inverted repeat flank symmetry A2 and the inner half contains the AUG start codon for the long open coding region.

We suggest a particular role for symmetry B. Because IS elements insert more-or-less randomly into DNA, they may sometimes be subject to fortuitous activation by strong external promoters. We suggest that symmetry B, in combination with symmetry A2, acts to protect IS*10*-R from such fortuitous activation. A transcript extending into IS*10*-R from an outside promoter can, by using symmetries A2 and B, fold into the stable secondary structure shown in Fig. 6. Since symmetry B includes the AUG start codon of the long coding region, such folding would sequester the start codon of the long coding region, render it inaccessible to ribosomes, and thus block gene expression at the translational level. (Symmetry A1 would also be present in such a transcript but it is irrelevant to the interaction of symmetries A2 and B). By contrast, when transcription is initiated from the normal p-IN promoter, there will be no opportunity for sequestering of the start codon in mRNA secondary structure since the resulting transcript will contain only half of symmetries A2 and B. Evidence that mRNA secondary structure suppresses translation by sequestering initiation codons has been presented for RNA phages MS2 and QB and in several other cases (8–10). IS elements may be one case in which evolution of such a mechanism has been particularly favored, since each transposition of an IS sequence presents the element with an entirely new chromosomal environment. Analogous protection mechanisms involving sequestering of AUG or Shine–Delgarno sequences have also been suggested for IS4 and IS5 (11, 12).

We suspect that both the inner half of symmetry A1 and symmetry A2 play a role in regulating or modulating IS*10* transposition. The position of symmetry A2, directly over p-IN and near p-OUT and the long coding regions, is suggestive of such a role, and an essential mechanistic role for this sequence has been directly excluded by the finding that deletion mutants lacking symmetry A2 are fully transposition proficient when complementing functions are provided *in trans* (3). Genetic experiments have also established that the two ends of IS*10*-R are functionally symmetrical. Although it has not been directly shown, this symmetry makes it likely that both inside and outside ends contain all of the sites essential for transposition (3). If so, the fact that symmetry A1 occurs only at the outside end

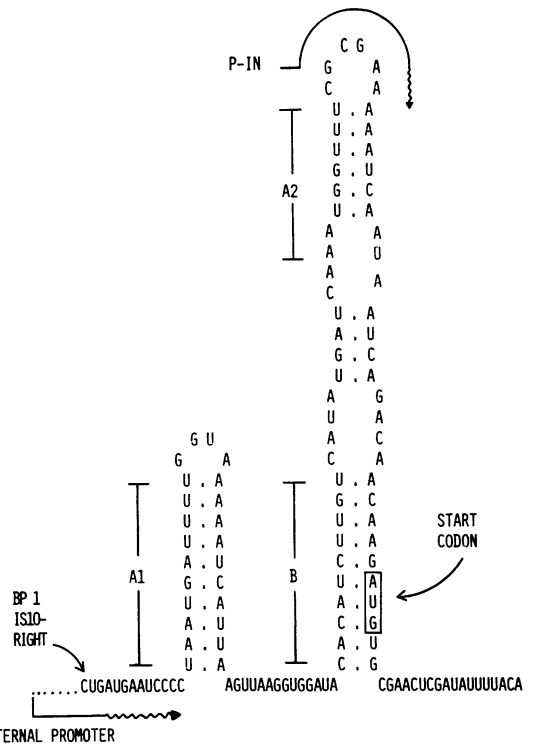

FIG. 6. Mechanism for protection of IS*10*-R from fortuitous activation by external promoters. Transcripts extending from an external promoter into and across the outside end of IS*10*-R can assume a stable secondary structure. The stability of the large hairpin involving symmetries A2 and B is −17.2 kcal (1 cal = 4.18 J) as calculated by the rules of Tinoco *et al.* (7). Symmetry B is directly responsible for sequestering the start codon, but pairing at symmetry A2 also stabilizes pairing at symmetry B: the stability of symmetry B in the absence of any contribution from symmetry A2 would be −9.7 kcal. Pairing at symmetry A1 (stability, −7.6 kcal) is irrelevant to the formation of the large hairpin.

of IS*10* would argue against an essential mechanistic role for the inner half of symmetry A1 as well. If the outer 10 bp half of symmetry A1 is in fact a protein recognition site as suggested above, then the same protein(s) might act at the inner half of that symmetry or of symmetry A2 (or both) as well. Internal symmetries have been identified near one or both ends of IS*4* and IS5; their importance is not yet known (11, 12).

Differences Between IS*10*-R and IS*10*-L. Six hundred and sixty-five nucleotides of IS*10*-L sequence have been determined, 515 at the outside end of the element and 150 at the inside end. IS*10*-L differs from IS*10*-R at 15 out of these 665 positions (≈2.5%). As shown in Fig. 5, six of these changes occur upstream of the ATG start codon for the long coding region. Of these, four occur in sequences that lie between the symmetrical halves of symmetries A1 and A2, one occurs between symmetries A1 and B, and one occurs in the last base upstream from the ATG start codon. Nine differences between IS*10*-L and IS*10*-R have been identified within the long coding region. Notably, only three of these differences change amino acids in presumptive protein. These three (bp 1249, leucine to serine; bp 1271, alanine to threonine; and bp 1300, alanine to valine) all lie in the carboxy-terminal 22 amino acids of the polypeptide; the entire amino-terminal portion of the protein, 135 amino acids analyzed thus far, is identical in IS*10*-R and IS*10*-L. There are no differences between the two IS*10* elements in the short terminal inverted repeat sequences. We do not yet know which of the observed sequence differences contributes to the functional differences between IS*10*-R and IS*10*-L.

It seems most likely that IS*10*-L and IS*10*-R were originally identical in sequence and have since diverged from one another. Simple mechanisms by which a composite transposon containing two IS elements can arise by replicative transposition from a single IS sequence have been suggested (13). There is no need to postulate formation of Tn*10* by independent insertion of two related but distinct sequences to positions flanking the tetracycline-resistance determinant. IS*10*-L mediates transposition at such a low frequency that it is unlikely to have transposed independently in its present form. IS*10*-R, on the other hand, mediates transposition at approximately the same frequency as elements such as IS*1* and IS*2*, which have been observed as independent entities in the chromosomes of *E. coli* and other bacteria. Thus, we favor the idea that IS*10*-R is very close to the original IS*10* element that gave rise to Tn*10* and that IS*10*-L has degenerated to its present nonfunctional state.

The sequences of IS*10*-R and IS*10*-L differ to a much greater extent than do those of the flanking IS elements of two other composite transposons. The IS*903* elements of the kanamycin-resistance transposon Tn*903* are exactly identical in sequence, and the IS*50* elements of the kanamycin-resistance transposon Tn5 differ at a single nucleotide position (14–16). The greater divergence between the IS elements of Tn*10* may mean that Tn*10* is older than Tn5 and Tn*903* and has had more time to accumulate changes. Tetracycline has been heavily used to treat both human and animal populations since before 1950. Resistance to tetracycline was observed in clinical isolates as early as 1952, and Tn*10*-like elements account for an important fraction of tetracycline-resistance determinants isolated from nature. In contrast, kanamycin was first identified as an antibiotic in 1956, and kanamycin resistance was not observed until 1964 and is still less common in clinical isolates than tetracycline resistance (ref. 17; and references therein).

IS*10*-L and IS*10*-R can have diverged from some common progenitor only in ways that have not interfered with the ability of IS*10*-R to promote Tn*10* transposition. Neither element can have suffered a lesion in an essential site; mutations that inactivate IS*10*-R would have been excluded, as would mutations in IS*10*-L that result in production of dominant deleterious functions. Such constraints may account for the fact that only three of nine coding region differences result in altered amino acids, for sequence conservation in the terminal inverted repeats, and for the nature of changes upstream from the coding region. The exact nature of these constraints remains to be established.

Compactness of IS*10* Genetic Organization. Important genetic determinants are compactly organized into the 1329 bp of IS*10*-R. A single long coding region extends almost the entire length of the element, beginning at bp 108 and ending within the terminal inverted repeat at the inside end. This compact structure is typical of IS-like elements. With the notable exception of IS*1*, all IS elements contain a long coding region extending the length of the element and ending within or very near the distal terminal inverted repeat. In some cases, a second overlapping coding region may also be important (for review, see ref. 18). In the case of IS*10*, we know that the 150-bp region at the outer end of the element contains three suggestively located inverted repeats, two promoters, the start of the long coding region, and all of the sites essential for transposition. There would appear to be almost no nucleotide sequences in IS*10*-R that are not part of some important genetic determinant; in several cases, two or more important determinants overlap one another.

We thank Richard Tizard for assistance in DNA sequence analysis techniques and Anna Ferri and Laura Cerf for help in preparation of the manuscript. This work was supported by grants to N.K. from the National Science Foundation, the National Institutes of Health, and the Camille and Henri Dreyfus Foundation. R.W.S. was supported by a post-doctoral grant from the Damon Runyon–Walter Winchell Cancer Fund.

1. Kleckner, N., Chan, R. K., Tue, B. K. & Botstein, D. (1975) *J. Mol. Biol.* **97**, 561–575.
2. Ross, D. G., Swan, J. & Kleckner, N. (1979) *Cell* **16**, 721–732.
3. Foster, T. J., Davis, M. A., Takeshita, K., Roberts, D. E. & Kleckner, N. (1981) *Cell* **23**, 201–213.
4. Backman, K., Ptashne, M. & Gilbert, W. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 4174–4178.
5. Maxam, A. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
6. Kleckner, N. (1979) *Cell* **16**, 711–720.
7. Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246**, 40–41.
8. Lodish, H. F. (1970) *J. Mol. Biol.* **50**, 689–702.
9. Steege, D. A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4163–4167.
10. Iserentant, D. & Fiers, W. (1980) *Gene* **8**, 1–12.
11. Klaer, R., Kuhn, S., Tillman, E., Fritz, H. J. & Starlinger, P. (1981) *Mol. Gen. Genet.* **181**, 169–175.
12. Hobom, G., Kroger, M., Kac, B. & Lusky, M. (1981) in *Structural and DNA–Protein Interactions of Replication Origin*, ed. Ray, D. S. (Academic, New York), pp. 209–222.
13. Shapiro, J. A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1933–1937.
14. Oka, A., Sugisaki, H. & Takanami, M. (1980) *J. Mol. Biol.* **147**, 217–226.
15. Auserswald, E. A. & Schaller, H. (1981) *Cold Spring Harbor Symp. Quant. Biol.* **45**, 107–114.
16. Rothstein, S. J. & Reznikoff, W. S. (1981) *Cell* **23**, 191–200.
17. Falkow, S. (1975) *Infectious Multiple Drug Resistance* (Arrowsmith, Bristol, England).
18. Kleckner, N. (1981) *Annu. Rev. Genet.* **15**, 341–404.