The EMBO Journal vol.8 no.7 pp.2101 - 2109, 1989

Mutational analysis of IS10's outside end*

O.Huisman[1,2], P.R.Errada[1], L.Signon[2] and N.Kleckner[1]

[1]Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, USA, and [2]Département de Biotechnologie, Institut Pasteur, 28 rue de Docteur Roux, 75724 Paris Cedex 15, France

Communicated by A.Toussaint

*This paper is dedicated to the memory of Olivier Huisman, who passed away unexpectedly in early 1988.

We present the genetic analysis of a large number of mutations in the outside end of insertion sequence IS10. (i) The terminal inverted repeat sequence is probably the primary site of transposase binding. Mutations in this region fall into phenotypic classes which correspond to their map locations, suggesting that this region may consist of several distinct functional segments. Similarities between the organization of IS10's inverted repeat and those of other transposable elements are discussed. (ii) Base pairs 23–42 include a consensus binding sequence for one of the IS10 transposition host factors, IHF. The phenotypes of mutations in this region suggest that IHF is the major host factor for outside-end transposition activity *in vivo* and that base pairs throughout this region are important for the IHF interaction. (iii) Mutations in bp 43–61 do not affect outside-end transposition activity but do affect, in expected ways, previously identified determinants involved in expression and regulation of transposase. (iv) Some mutations in bp 23–42 also affect transposase expression; the possibility that IHF negatively regulates transcription initiation is discussed.

Key words: IHF/IS10/mutational analysis/transposase binding

Introduction

Insertion sequence IS10 occurs as an inverted repeat at the ends of transposon Tn*10* (Kleckner, 1988). The two termini of IS10 are named 'outside' and 'inside' with respect to their positions in Tn*10*. The outside end is structurally complex (Figure 1). Base pairs 1–42 of IS10 are required for full outside-end activity in the transposition reaction (Way and Kleckner, 1984; Morisato and Kleckner, 1987; see below). The terminal 22 bp occur as a nearly perfect inverted repeat at both ends of IS10 (Halling *et al.*, 1982). This region is presumed to comprise the primary site of interaction with IS10 transposase protein, because the inside-end inverted repeat sequence alone is sufficient for nearly full inside-end tranposition activity (D.Ahmann, D.Morisato, L.Signon, P.R.Errada and N.Kleckner, unpublished).

Transposition *in vitro* by a substrate containing either two IS10 outside ends (O × O, in inverted orientation, analogous to transposon Tn*10*) or one outside and one inside end (O × I, inverted, analogous to IS10) requires both transposase protein and an accessory host protein, which can be either IHF or HU (Morisato and Kleckner, 1987). DNase I protection analysis reveals a classical IHF footprint (Craig and Nash, 1984; Gamas *et al.*, 1987) at the outside end of IS10: an extensive region from bp 7 to 50 is protected, and positions near the ends of this segment exhibit enhanced cleavage (Figure 1; D.Morisato and N.Kleckner, unpublished). An IHF consensus sequence is present at bp 30–42; this is presumably the site of the strongest sequence-specific contacts between IHF and DNA. HU binding is generally non-specific (Berthold and Geider, 1976) and the importance of particular sites in the IS10 outside end for HU-promoted transposition *in vitro* has not been investigated. The relative contributions of IHF and HU to outside-end transposition activity *in vivo* are discussed below.

The outside end also contains sequences important for several aspects of transposase gene expression: (i) the pIN promoter, at which transposase gene transcription initiates (Simons *et al.*, 1983), (ii) a GATC sequence within pIN through which DNA adenine methylation regulates transcription initiation (Roberts *et al.*, 1985) and (iii) sequences which, when transcribed into read-through transcripts by promoters outside of IS10, prevent translation of the transposase gene off of such transcripts (Davis *et al.*, 1985). The outside end also encodes determinants involved in IS10 anti-sense RNA control (Simons and Kleckner, 1983; Kleckner, 1988). However, all of the experiments reported here were performed under conditions where anti-sense RNA regulation did not occur: all outside end derivatives contained a mutation (R5) which virtually abolishes the activity of the anti-sense RNA (Simons and Kleckner, 1983).

We present below a detailed genetic analysis of bp 1–61 of the IS10 outside end. Mutations in this region were generated by physical methods, without respect to phenotype, and were subsequently analyzed *in vivo* for their effects on both outside-end transposition activity and transposase gene expression.

Results

Isolation of mutations

Most mutations were generated by the method of Myers and Maniatis (1985) in which single-stranded DNA is mutagenized *in vitro*, converted to double-stranded form, and fragments containing mutations are identified by an electrophoresis procedure which separates DNA fragments of identical length that differ in sequence by a single base pair. Mutant fragments were cloned and sequenced to identify mutations in the region of interest. Outside ends containing certain other mutations were generated using synthetic oligonucleotides (see Materials and methods).

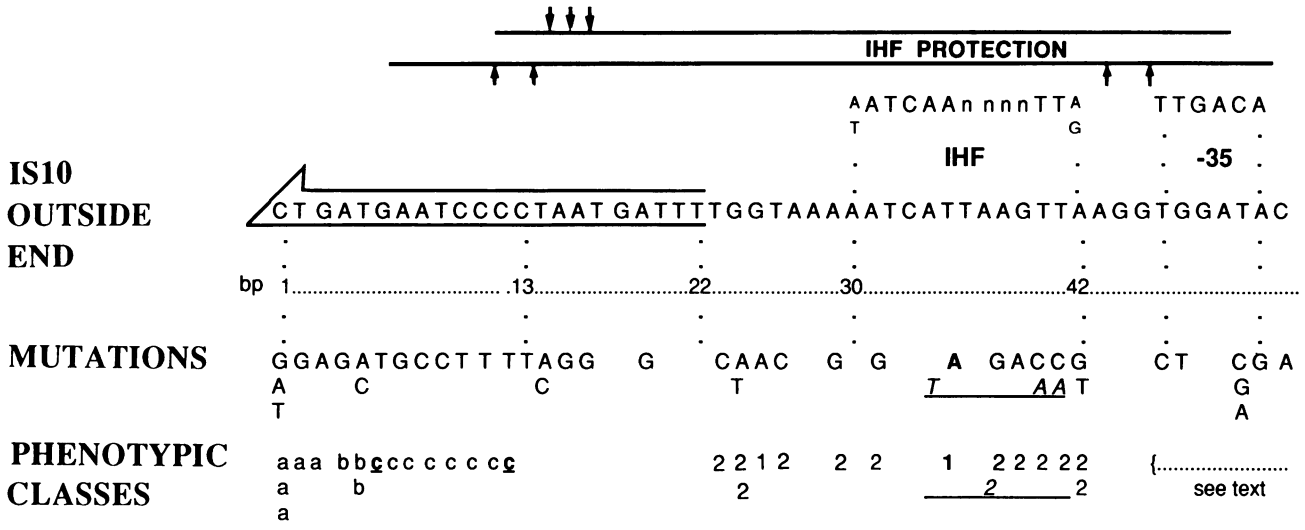

Fig. 1. The outside end of IS10, bp 1−52. Terminal inverted repeat sequence, written 5′ to 3′, is enclosed in large arrow; dots over bp 5, 6, 7, 15 and 18 indicate positions of differences with inside-end inverted repeat sequence. IHF consensus region and −35 region of transposase gene promoter (pIN) are indicated with the corresponding optimal consensus sequences shown. Region protected and positions enhanced by IHF for DNase I digestion are also indicated (D.Morisato and N.Kleckner, unpublished). Each mutation isolated and analyzed is indicated by the letter(s) corresponding to the mutant base(s); IHFup mutation (**35A**) is indicated in bold; triple mutation IHFdown (*34T*, *40A*, *41A*) is indicated in underlined italics. Phenotypic classes of mutations in bp 1−42 are indicated by symbols at corresponding positions in the line(s) below. Classes a, b, and c (including c) are from Table I; classes 1 and 2 are from Table II.

Strategy for genetic analysis

Each mutation was analyzed for its effects on outside-end activity (*in vivo*) in each of three different plasmid-borne mini-transposons (Figure 2): a construct with two outside ends carrying the same mutation (O* × O*), a construct with one mutant outside end and one wild-type outside end (O* × O^+), and a construct with one mutant outside end and one wild-type inside end (O* × I^+ or 'mini-IS10'). Transposition was measured in a mating-out assay (see Materials and methods) with transposase provided in *trans* from a separate, compatible plasmid.

For mutations in bp 1−42, the effects of each mutation in the three different mini-transposon constructs have been compared to one another in two ways. First, the effect of the mutation when present at both ends of a double outside-end construct (O* × O*) is compared to its effect when present at only one end of such a construct (O* × O^+). If the two ends function independently, the transposition defect of the double mutant, relative to its isogenic wild-type, should be the square of the transposition defect of the single mutant relative to its isogenic wild-type. This comparison is expressed in the parameter (B^2/A) in Tables I and II. For most but not all mutations, the observed (B^2/A) is close to 1.

Second, the effect of the mutation in the two singly mutant constructs was compared to determine whether the effect of the mutation differed depending upon whether the second end in the reaction was an inside end (O* × I^+) or an outside end (O* × O^+), parameter (C/B) in Tables I and II. Most but not all mutations have the same effect in both situations; the observed (C/B) is close to 1.

All mutations in bp 1−61 were analyzed in an intact IS10 element marked with a Kan^r determinant; this element is isogenic to the mini-IS10 (O* × I^+) construct (Figure 2). The effect of each mutation on transposase gene expression was determined by measuring the ability of the mutant marked IS10 to provide transposase in *trans* to a wild-type outside end mini-transposon construct (O^+ × O^+), carried on a separate plasmid. Also, marked IS10 elements were examined directly for their own level of 'self-driven' transposition. In most but not all cases, the effect of a mutation on self-driven transposition was simply the product of its separately measured effects on outside-end activity in the mini-IS10 construct and on transposase expression. Marked IS10 and mini-IS10 elements were constructed both with and without a set of *rrnB* transcription terminators upstream of IS10 bp 1 in order to both evaluate and/or eliminate effects of read-through transcription initiated outside the element.

Mutations in the inverted repeat (bp 1−22)

Mutations in the inverted repeat fall into distinct classes which correspond to the map positions of the mutations (Table I, Figure 1).

Class (a) consists of all five mutations in bp 1, 2 and 3, including all three possible changes at bp 1. The severity of the mutant defects among members of this class varies over a 1000-fold range, but none is among the strongest mutations (see below). The characteristic feature of this class is that all of these mutations, and only these mutations, have a high (B^2/A) ratio (5−9), in contrast to all other inverted repeat mutations, which have a (B^2/A) ratio of less than 3 and usually close to 1.

Class (b) consists of the three mutations at bp 4 and 5. The characteristic feature of this class is that these mutations all have extremely weak effects, the very weakest of any in the outer 13 bp. None of these mutations affects transposition by more than ±30%.

Class (c) consists of all eight mutations between bp 6 and 13, one mutation at each base pair. The characteristic feature of this class is that these mutations all have extremely strong effects. Seven of the eight have the strongest phenotypes of any mutations identified in this analysis; the eighth mutation is also very strong; only two other mutations (1A and 3A) are equally strong.

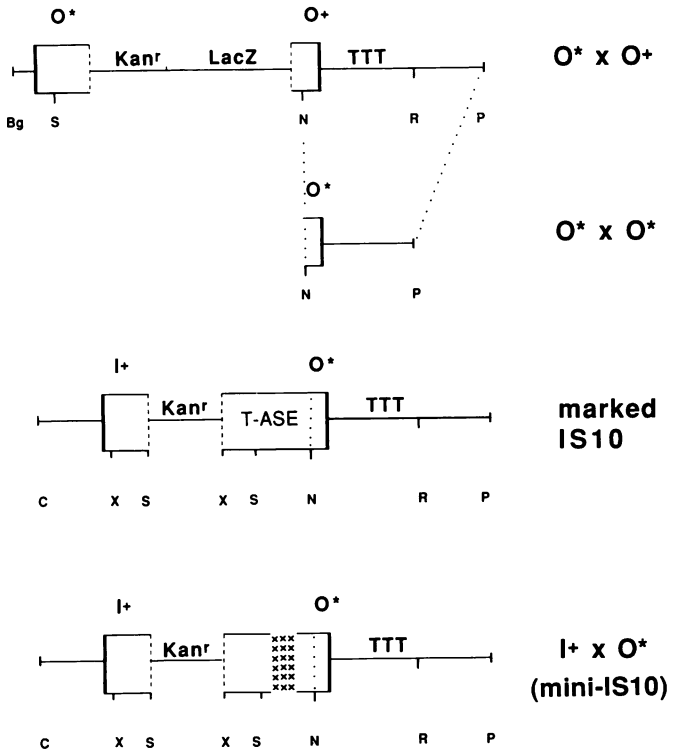

Fig. 2. Mutations were subcloned from the pGC plasmids on which they were isolated into the (O* × O$^+$) construct by substitution of a *Bgl*II (Bg)–*Sal*I (S) fragment. Constructs containing a given mutation at both ends (O* × O*), were constructed from the (O* × O$^+$) plasmids by substitution of a mutant *Nde*I (N)–*Pst*I (P) fragment, also from the original pGC plasmid; this step unavoidably required removal of a set of transcription terminators located upstream of the affected end of the element. By similar strategies, the mutation was moved from the pGC plasmid into a marked IS10 construction. Mini-IS10 derivatives (I$^+$ × O*) were then made by in-frame deletion of transposase gene sequences (indicated by vertical rows of x's). The marked IS10 and (I$^+$ × O*) derivatives shown here contain transcription terminators upstream of the mutant terminus, and are identical to the (O* × O$^+$) construct between the *Nde*I and *Pst*I sites. Derivatives of these two constructs lacking the terminators were also constructed, in essence by substitution of the same *Nde*I–*Pst*I fragment used to create the (O* × O*) construct. The actual lengths of the transposable elements in these four constructs are about 5 kb for the (O × O) constructs, 3.0 kb for the marked IS10, and 2.7 kd for the (I$^+$ × O*) mini-IS10 element. Additional details are provided in Materials and methods.

The two mutations which define the ends of the Class (c) region, 6T and 13T, have an additional unique property: they are the only mutations which exhibit substantially different activities in combination with inside and outside ends. Both mutations confer a less severe defect in combination with an inside end (C/B = 5–10). The basis for this differential behavior remains to be investigated.

Mutations in bp 14–19 all have weak phenotypes. The strongest (16G) confers a 20-fold defect when present at both ends, while others confer defects of 6-fold or less. Four of these mutations occur at bp 14, 15 and 16 and thus provide sharp delineation of this region from that defined by mutations in Class (c). However, too few mutations in this region have been analyzed to be certain that they comprise a single homogeneous group.

None of the mutations in the inverted repeat affect expression of transposase significantly (by more than ±30%; data not shown).

Mutations in bp 23–42

Mutations in bp 23–42 fall into two classes on the basis of their outside end transposition phenotypes, Classes 1 and 2 of Table II. Included in this analysis are two oligonucleotide-generated mutations in the IHF consensus region: a single base pair change (IHFup) and a triple base pair change (IHFdown) (Figure 1). These mutations take the IHF sequence closer to and farther from consensus, respectively. Correspondingly, they strongly increase or decrease IHF binding in both DNase protection and gel-retardation assays, and IHF activation of outside end activity *in vitro* (D.Morisato and N.Kleckner, unpublished). Class 1 mutations include IHFup and a second mutation with a similar set of phenotypes; Class 2 mutations include IHFdown and other mutations with similar, but generally weaker, phenotypes.

The two classes of mutations are defined by the mini-IS10 (O* × I$^+$) construct: both Class 1 mutations slightly but reproducibly increase transposition of this element while all Class 2 mutations decrease its transposition. The strongest Class 2 mutations occur in the IHF consensus sequence, while mutations mapping in the 'interstitial' region between the inverted repeat and the consensus sequence have weaker effects. Although the effects of these mutations are relatively small, they are significant; the same pattern of phenotypes has been observed in several independent experiments.

Many Class 1 and Class 2 mutations also have significant effects when present in doubly mutant outside-end constructs (O* × O*). The IHFdown mutation has the strongest effect, reducing transposition to 7% of the wild-type level, and mutation 41C is also very strong; the effects of other Class 2 mutations are weaker, or, in a few cases, negligible. Paradoxically, the two IHFup mutations reproducibly decrease (O* × O*) transposition as well.

The effects of Class 1 and Class 2 mutations in singly mutant outside-end constructs (O* × O$^+$) are so small as to be essentially undetectable. The only number in Column B of Table II which is reproducibly different from 1 is the defect of the IHFdown mutations (0.7). These results were expected from the previous finding that even a complete deletion of this region results in only a 75% reduction in transposition activity (Way and Kleckner, 1984; see Discussion).

Comparisons of mutant effects in different constructs reflect the above patterns. The ratio (B^2/A) is high (5 and 7) for the two strongest Class 2 mutations, for which B is close to 1 and A is low. A deletion of the region beyond bp 27 of IS10 (Way and Kleckner, 1984), behaves analogously: (OΔ × OΔ) = 0.004, (OΔ × O$^+$) = 0.25 and (B^2/A) = 16. Correspondingly, for weaker Class 2 mutations, (B^2/A) is slightly greater than or equal to 1. The ratio (C/B) essentially reflects mutant effects on the (O* × I$^+$) element (C), since the singly mutant outside end constructs exhibit no defect (B = 1).

Some mutations in bp 23–42 also affect the level of complementing transposase made by the marked IS10 element. For several mutations, the effect on transposase expression is opposite to its effect on transposition: both Class 1 mutations decrease the level of transposase (7- to 12-fold), while the IHFdown mutation and a second Class 2 mutation (39A) increase transposase levels (3- to 10-fold); several other Class 2 mutations also give slight increases. However, the effects of a few Class 2 mutations on

Table I. Mutations in bp 1–22

Class	Mutation	(WT base)	Relative activity (A) O* × O*	(B) O* × O$^+$	(C) O* × I$^+$ (+T)	(B^2/A)	(C/B)
	None		≡1	≡1	≡1		
a	1G	(C)	0.07	0.7	0.77	**7**	1.1
a	1A	(C)	0.0009	0.09	0.06	**9**	0.7
a	1T	(C)	0.2	1	0.72	**5**	0.7
a	2G	(T)	0.02	0.35	0.27	**6**	0.7
a	3A	(G)	0.001	0.1	0.15	**10**	1.5
b	4G	(A)	**1**	1	0.77	1	0.8
b	5A	(T)	**0.7**	0.9	0.95	1.2	1.0
b	5C	(T)	**1.3**	1.3	0.79	1.3	0.7
c	6T	(G)	*0.0003*	0.01	0.08	0.33	**8, 10**
c	7G	(A)		*0.003+*			
c	8C	(A)	0.003	0.1	0.2	3.3	2
c	9C	(T)	*0.00004*	0.006	0.006	0.9	1
c	10T	(C)		*0.002*	*0.003*		
c	11T	(C)	*<0.00001*	0.004	0.002	≥1.6	0.5
c	12T	(C)	*0.00002*	0.004	0.002	0.8	0.5
c	13T	(C)	*0.0002*	0.01	0.03	0.5	**3, 10**
	14A	(T)	0.35	0.7	0.75	1.4	1
	14C	(T)	0.15	0.5	0.35	1.7	0.7
	15G	(A)	1.1	1	0.85	0.9	0.8
	16G	(A)	0.05	0.3	0.18	1.8	0.6
	19G	(A)	0.3	0.5	0.32	0.8	0.6

Relative transposition frequencies in mating-out assays (see Materials and methods) for three mini-transposon constructs (Figure 2) carrying mutations in bp 1–22. The last two columns compare data from other columns (also see text). With mutations 6T and 13T, (C/B) for (O* × I$^+$) constructs both with and without terminators is shown to emphasize the significance of the high numbers; with other mutations, (C/B) for (O* × I$^+$) constructs lacking terminators was always about 1 (data not shown). The transposition levels for the three wild-type constructs were 2.5×10^5 (A and B) and 2.5×10^4 (C) Kan^r exconjugants per ml of mating mixture. Background in the absence of transposase is ≤3 per ml.
(+T) indicates the (O* × I$^+$) element carrying *rrnB* terminators (Figure 2). (+) = mutation 7G creates a GATC site whose As at the bp 8 and 9 of IS10 are methylated by Dam methylase; analysis of this mutation in a Dam$^-$ host suggests that it is one or both of these methylations which causes the defective phenotype. In fact, this mutation creates a GATC at precisely the site where one occurs naturally at the inside end of IS10, where it is only methylation of bp 9 that is inhibitory.

Table II. Mutations in bp 23–42

Class	Mutation	Location	Relative activity (A) O* × O*	(B) O* × O$^+$	(C) O* × I$^+$ (+T)	(B^2/A)	(C/B)	(D) Tase (+T)
	Wild-type		≡1	≡1	≡1			≡1
1	IHFup	C	0.7	0.8	1.1	0.9	1.2	0.08
	25A	I	0.8	1	1.3	1.3	1.3	0.14
2	IHFdown	C	0.07	0.7	0.2	7	0.3	9.6
	23C	I	0.9	1.1	0.85	1	0.9	
	24A	I	0.9	1	0.65	1	0.7	1.6
	24T	I	0.8	1	0.65	1.3	0.7	1.3
	26C	I	0.85	1.2	0.65	1.2	0.7	1.2
	29G	I	0.8	0.9	0.55	1	0.6	1.1
	31G	C	0.65	0.9	0.32	1.5	0.3	1.2
	38G	NC	1	1.1	0.75	1	0.8	0.37
	39A	NC	0.6	0.9	0.35	1.7	0.4	3.2
	40C	C	0.8	1	0.25	1.3	0.3	1.1
	41C	C	0.2	0.8	0.12	5	0.1	1.2
	42G	C	0.8	0.9	0.55	1	0.6	1.5

Relative transposition frequencies for three mutant mini-transposon constructs (A), (B) and (C), and for a wild-type mini-transposon complemented by a mutant marked IS10 element (D). Comparisons, (+T), transposition levels for (A), (B) and (C), and background as in Table I; for (D), transposition level is 1×10^4 Tet^r exconjugants per ml of mating mixture. Tase = transposase

transposase expression do not correlate with their effects on outside end transposition activity: the mutation with the strongest effect on transposition (41C) has almost no effect on transposase expression and one mutation (38G) weakly decreases transposition but also decreases transposase expression 3-fold.

Table III. Mutations in bp 46–61

Mutation	Predicted phenotype	Relative activity (D) Tase (+T)	(E) Tase (−T)	(E/*D*)	(F) IS10 (−T)	(*F*/*E*)
None		≡1	≡1		≡1	
46C	−35dn	**0.006**	0.09	15*	0.05	0.55
47T	−35up	**16**	12	0.75	10	0.83
50C	−35up	**24**	20	0.83	15	0.75
50A	−35nc	**0.83**	0.69	0.83	0.46	0.67
50G	−35nc	**34**	15	0.44	15	1.0
51G	−35dn	**0.23**	0.24	1.0		
52A		*6.0*	2.7	0.45	3.9	1.4
53G		0.65	1.4	2.1	0.73	0.52
55G		0.93	0.83	0.9	0.95	1.1
57T	Pot⁻	**0.78**	4.7	***6.0***	1.3	0.27
58C	Pot⁻	**2.1**	17	***8.1***	6.7	0.39
60T	Pot⁻	**2.5**	22	***8.5***	6.5	0.29
61C	Pot⁻	**0.43**	3.6	***8.4***	1.1	0.30

Relative transposition frequencies for a wild-type mini-transposon complemented by mutant marked IS10 elements with (D) and without (E) *rrnB* terminators upstream of the IS10 outside end, and the ratio of the two values (third column). Also: relative self-driven transposition frequencies of mutant marked IS10 elements lacking terminators (F). Rightmost column (F/*E*) compares the relative self-driven transposition frequency (F) with relative level of transposase complementation (E) for constructs lacking transcription terminators. Mutations in the −35 region of pIN either take the sequence closer to consensus (−35up), away from consensus (−35dn) or change one non-consensus base to another non-consensus base (−35nc). Transposition levels for wild type constructs were 2×10^4 (B), 7×10^3 (C) and 5×10^3 (D) drug resistant exconjugant colonies per ml of mating mixture; levels for (A) and background are as in previous tables.

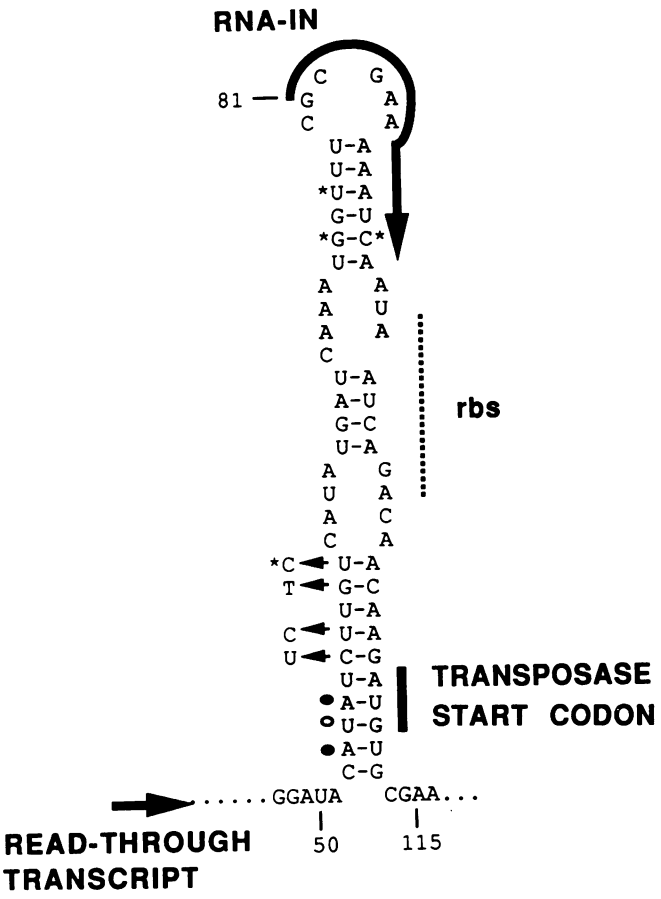

Fig. 3. Effects of mutations on IS10 protection structure formed by read-through transcripts entering the transposase gene from upstream (Davis *et al.*, 1985). Closed arrows = mutations from this analysis; * = mutations previously isolated by R.W.Simons and colleagues (Kleckner, 1988); all of these mutations are predicted to weaken the affected stem regions. Filled circles = sites of two mutations from this analysis which do not affect protection; open circle = site of R5 mutation present in all constructs analyzed here to eliminate anti-sense control, which changes a C at this position to the indicated U. Failure of these latter three mutations to affect the Pot phenotype suggests that integrity of the lower part of the stem is not important. Numbers indicate the positions of the corresponding template base pairs in the sequence of IS10. RNA-IN, the normal transposase mRNA, initiates at bp 81.

Base pairs 43–61

None of the mutations beyond bp 43 has any significant effect on outside-end transposition activity, consistent with previous deletion analysis (Way and Kleckner, 1984; Morisato and Kleckner, 1985). However, nine mutations in this region do affect transposase expression (Table III). All nine do so by altering previously identified features of IS10.

Six mutations affect the activity of the transposase gene promoter, pIN as revealed by the fact that they increase or decrease transposase expression by more than 2-fold in constructs where expression is protected from upstream promoters by transcription terminators (Table III, Column D). Five of these mutations (46C, 47T, 50C, 50G, 51G) alter the −35 region of the pIN promoter; the four which take the sequence towards or away from the −35 consensus sequence have the predicted effects on expression (Table III, Figure 1). The effects of these mutations on transposase expression are probably a consequence of direct effects on RNA polymerase interaction at the pIN promoter. The basis for the effect of mutation 52A is not clear. None of these mutations alters outside-end transposition activity (data not shown). Furthermore, the effects of these mutations on self-driven transposition of the marked IS10 element are the same as their effects on transposase expression (*F*/*E*) = 0.5–1.4.

Four other mutations, 57T, 58C, 60T and 61C, exhibit normal or nearly normal levels of expression in constructs containing transcription terminators but exhibit 6- to 15-fold higher levels of transposase expression when terminators are absent than when they are present [Table III, (*E*/*D*)]. This 'protection'-defective phenotype (Pot⁻) is diagnostic of mutations which abolish IS10's ability to prevent expression of transposase off of read-through transcripts initiated upstream of pIN. IS10 sequences present in read-through transcripts form an extensive secondary structure, not found in pIN transcripts, which sequesters the transposase gene start codon and thus blocks translation (Davis *et al.*, 1985; J.Rayner, T.Sonnebend and R.W.Simons in preparation). All four Pot⁻ mutations identified here are predicted to destabilize this secondary structure (Figure 3). None of these mutations affect ouside-end transposition activity in mini-transposon constructs (data not shown).

However, the four Pot⁻ mutations differ from the pIN mutations in one important way: in a marked IS10 element lacking transcription terminators, these mutations increase self-driven transposition only one third as much as they increase the level of complementing transposition functions made by this element [(*F*/*E*) = 0.27–0.39].

Discussion

Organization of the inverted repeat

Genetic analysis suggests that the inverted repeat can be divided into four sub-regions on the basis of mutational phenotypes: bp 1–3, 4–5, 6–13 and 14–22. It seems likely, at least for the first three regions, that differences in mutant phenotype correspond to differences in function for the affected base pairs. This result is particularly striking, since there is no reason *a priori* that different functional domains should necessarily correspond to discrete base pair segments.

Mutations in bp 6–13 have the strongest phenotypes. The simplest interpretation is that transposase makes its strongest sequence-specific contacts with DNA at these positions. However, a role of IHF in this region is not excluded, since IHF protection extends through bp 7 (Figure 1 and below). The only detailed information about transposase interactions in these base pairs is that N6 adenine methylation at bp 9 of the inside end strongly inhibits inside-end activity (Roberts *et al.*, 1985), suggestive of major groove contact(s) at this position.

The region from bp 1 to 3 is also important for transposition, since three of the mutations in this segment confer defects of 50-fold or more in the doubly mutant construct. Interestingly, at bp 1, which is the site at which transposase separates transposon sequences from flanking DNA and rejoins them to target DNA sequences, two of the three base pair changes at this position (CG to GC or TA) confer mild defects, 5- and 14-fold, while the third mutation (CG to AT) confers a very strong defect. An interesting possibility would be that these base pairs provide primarily structural information, and that a mutant defect reflects creation of an inappropriate structure rather than absence of particular sequence specific protein–DNA contacts.

Mutations in bp 1–3 have high values of the parameter (B^2/A): each mutation has a less severe effect in combination with a wild-type outside end than in combination with an outside end bearing the same mutation. The formal interpretation of this result is that a wild-type outside end can compensate for the particular defect conferred by a change in this region.

The five mutations in bp 14–19 have much weaker phenotypes than mutations in bp 6–13 and 1–3. These mutations are probably affecting transposase interactions rather than IHF interactions. Unlike similarly weak Class 2 mutations in bp 23–42, which probably do affect IHF interactions (see below), mutations in bp 14–19 confer stronger defects when present at both ends (O* × O*) than when present in combination with an inside end (O* × I^+) and do confer significant effects even in the singly mutant outside-end construct (O* × O^+).

On the basis of mutations analyzed thus far, the regions at bp 4–5 appear to contain almost no sequence-specific information.

Comparison with other transposon inverted repeats

Comparison of the results presented above with results from detailed analyses of the IS903 and IS50 inverted repeat sequences suggests that the inverted repeats of all three elements could share a very similar functional organization.

When we analyze data for 36 mutations in the inverted repeat of IS903 (Derbyshire *et al.*, 1987) by the same approaches used above for IS10 mutations (see Materials and methods), two striking points of similarity are revealed. First, the nine strongest IS903 mutations occur in the central portion of the inverted repeat, between bp 6 and 15, analogous to the bp 6–13 region of IS10. Second, of the 13 mutations analyzed in singly and doubly mutant constructs analogous to (O* × O^+) and (O* × O*) above, each of three mutations at bp 1, 2 and 3 exhibits a very high ratio of the parameter (B^2/A), with values of 17–130, while mutations at bp 4–15 exhibit much lower ratios, 3–9.

In fact, these two regions of IS903 have already been proposed to constitute functionally distinct domains on completely different grounds. The IS903 mutations were analyzed in mini-transposon constructs which carry the IS903 transposase gene in *cis* immediately adjacent to the pair of transposon ends. In such constructs, mutations at positions 6–15 confer very different phenotypes depending upon whether the mutation is in the end proximal or distal to the transposase gene, while mutations at positions 1–3 confer similar phenotypes in both cases. These and other observations have led Derbyshire *et al.* (1987 and personal communication) to propose that the bp 6–15 region is involved in an initial transposase binding step while the bp 1–3 region is involved in a later step.

Comparison of the wild-type IS50 and IS10 inverted repeat sequences also reveals two features of interest (Figure 4). First, the sequence relationships between the two sequences suggest that the central portions of the two inverted repeats, bp 6–12, are responsible for the specificity of the two transposases for their particular ends. [IS10 transposase does not promote transposition of a mini-IS50 construct (P.R. Errada, A.Delong and N.Kleckner, unpublished).] Base pairs 1–5 are not providing specificity: bp 1–4 are identical for both elements, and the IS10 mutation 5C changes the IS10 sequence to the corresponding IS50 base pair without significant phenotypic effect. The significance of this sequence relationship, which is not shared with IS903 or other known IS elements, remains to be determined. In contrast, IS10 outside end function is totally abolished by a mutation which places the corresponding IS50 base pair at position 6 of IS10 (mutation 6T). Finally, IS10 and IS50 either are or can be made identical at bp 13, 14 and 16 without significant effect, and mutations in this region of IS10 do not have strong phenotypes in any case (above).

Second, at the outside end of IS50, the binding site for the DnaA protein, which is an essential host factor for IS50 transposition (Yin and Reznikoff, 1987), occurs at bp 8–16 of the inverted repeat, directly overlapping the presumptive critical region. In IS10, the region of IHF protection in IS10 extends across the analogous region, through bp 7, and some specific points of enhancement occur between bp 11 and 18 (Figure 1). It is intriguing to speculate that interaction of host factors with transposase and/or DNA within this critical region might be a general feature of IS element inverted repeats. Insertion sequence IS1 may provide a third such case. IHF binding sites overlap the terminal inverted repeats at both ends of this element; however, the role of IHF in IS1 transposition is not yet understood (Gamas *et al.*, 1987).

The inverted repeats of IS50 have also been subjected to mutational analysis (Phadnis and Berg, 1987; Makris *et al.*, 1988). One clear conclusion of this work is that bp 4 at the

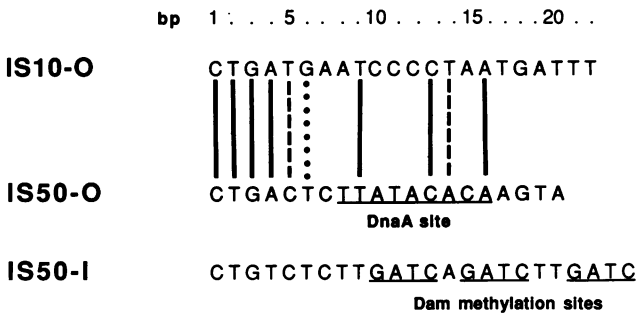

Fig. 4. Comparison of the IS10 outside end with IS50 inside and outside ends (Makris *et al.*, 1988; Phadnis and Berg, 1987). Straight vertical lines indicate positions of identity; dashed lines indicate positions at which IS10 mutation to the IS50 base-pair has a weak or no phenotype (5C and 14A); closed circles indicate a position at which an IS10 mutation to the corresponding IS50 base pair (6C) has a strong deleterious effect. Sites of DnaA binding and *dam* methylation at the ends of IS50 are indicated (Yin and Reznikoff, 1987, 1988).

end of IS50 contains little or no sequence-specific information, a finding analogous to that obtained above for IS10 bp 4 and 5.

The Tn*3* group of transposons, which are very different from the IS elements, also have terminal inverted repeats. Comparison among related elements suggests that among this group as well, transposase/terminus specificity resides within a small region of the inverted repeat that is located several base pairs from the end of the element (D.Sherratt, 1989).

Base pairs 23–42

IHF as a host factor in the transposition reaction. The IHFdown mutation provides a very useful tool for specifically probing the role of IHF in the transposition reaction *in vivo*. *In vitro*, this mutation abolishes IHF-promoted transposition, reducing it at least 50-fold, and does not affect HU-promoted transposition (D.Morisato and N.Kleckner, unpublished).

IHFdown confers a 15-fold decrease in transposition when present at both ends of an outside end construct (O* × O* = 0.07). We interpret this result to mean that IHF contributes at least 95% of auxiliary host factor activity *in vivo*. This interpretation is supported by the close correspondence between the phenotypes of the IHFdown mutation and other mutations in Class 2 of Table II; if a second host factor were contributing significantly, more complex phenotypic patterns might have been anticipated. The low level transposition observed with the IHFdown mutant could reflect the action of HU protein, residual activity of IHF and/or the presence of a third host factor.

The IHFdown mutation confers a much smaller decrease when present at only one end in combination with a wild-type outside end than when present at both ends (B^2/A = 10). Since a singly-mutant outside end construct bearing this mutation transposes at nearly the same frequency as wild-type (≥50%), we interpret these results as an extreme case of 'compensation': transposition of a Tn*10*-like construct, with two outside ends, requires that IHF be present at only one end of the transposon. This is a more specific version of the proposal of Way and Kleckner (1984).

The effects of the IHFup mutation on outside end activity are complex and more work is required to explain them fully. The small decrease in transposition of the doubly mutant (O* × O*) IHFup element may mean that IHF activity is nearly optimal for transposition, and that too much IHF activity at the outside end can be inhibitory. *In vitro* experiments support this conclusion (P.R.Errada and N.Kleckner, unpublished).

IHF protects a large segment of the IS10 outside end, from bp 7 to 50. The IHF consensus sequence includes nine of the 13 bp between bp 30 and 42 of IS10, and the stronger mutations (Class 2) all occur within this region, though not all at consensus base pairs. In addition, a series of somewhat weaker mutations occur in the interstitial region located between the consensus sequence and the inverted repeat, suggesting that this region is also involved in the IHF interaction. Other analyses have also suggested that information outside of the IHF consensus sequence can affect IHF binding (Gamas *et al.*, 1987). The generally weak phenotypes of single base changes throughout the IHF binding region is consistent with previous indications that IHF binding is the sum of a large number of relatively weak interactions (Nash *et al.*, 1987 and personal communication).

Effects of bp 23–42 mutations on transposase expression. The effects of some mutations in bp 23–42 are consistent with the possibility that IHF directly inhibits transcription initiation from the transposase gene promoter, pIN: mutations IHFup and 25A decrease expression while mutations IHFdown and 39A increase expression. Such an effect is mechanistically reasonable, since the sequences protected by IHF extend across the −35 region of pIN. However, more direct experiments are required to establish IHF as a negative regulator. The effects of some Class 2 mutations on transposase expression do not correlate with their effects on outside end transposition activity (above), and some or all of the mutational effects on transcription could be due to secondary effects. In particular, *in vitro* transcription analysis suggests the existence of an additional promoter within this region which competes with pIN (J.Campbell and N.Kleckner, unpublished); mutations in such a competing promoter would alter pIN expression.

The unusual phenotype conferred by Pot mutations. All four Pot⁻ mutations identified in this analysis increase the ability of a marked IS10 element to provide complementing transposition functions more than they increase the level of self-driven transposition of that element. Since none of these mutations affects outside end activity, this difference implies that the mutations increase the level of complementing functions more than they increase the level of transposase available for self-driven transposition. This differential effect can be explained as follows.

IS10 transposition is regulated by Dam methylation (Roberts *et al.*, 1985). A fully methylated IS10 element is inactive, because the pIN promoter and the activity of the inside end of the element are each individually inhibited by methyl group(s) at a specific GATC site in each of the two determinants. Immediately after an IS10 element is replicated, it becomes temporarily hemi-methylated, until such time as the relevant GATC sites are acted upon by Dam methylase. For one of the hemi-methylated IS10 elements, both pIN and the inside end are activated; transposase is made and acts successfully at the activated end to promote transposition.

For an IS10 element bearing a Pot$^-$ mutation, transposase is still made from pIN only during the period of hemi-methylation. However, transposase is also made from read-through transcripts, and read-through transcription, which is not Dam-regulated, occurs throughout the cell cycle. The complementation assay for transposase expression measures all of this transposase, while the self-driven transposition assay measures only the transposase made during the period of hemi-methylation, when the element's inside end is activated to receive it. Since the contribution of read-through transcription to this transposase is less than its contribution to total transposase, Pot$^-$ mutations have a smaller effect on self-driven transposition than on the level of complementing transposition functions.

Materials and methods

Bacterial strains and media

All transposition assays were carried out using NK7378 (*recA56 ΔlacproXIII arg*$^-$*/F'pOX38*) as a donor strain and NK6641 (*recA56 ΔlacproXIII arg*$^-$ *strA*) as a recipient. NK7378 is NK5830 with the *pOX38* plasmid (Guyer *et al.*, 1980) substituted for *F'lacpro*; NK5830 and NK6641 are described in Foster *et al.* (1981). Multi-copy plasmids were introduced into donor strains by transformation; for strains containing two compatible plasmids, the plasmids were introduced separately in two successive steps. All such plasmids were present as monomers. LB media is as described by Miller (1972); antibiotics were obtained from Sigma.

Plasmids

Mutations were isolated in pOH56 and pOH59. These two plasmids were constructed by inserting into the polylinker regions of pGC1 and pGC2 (Myers and Maniatis, 1985), a *Bam*HI to *Sal*I segment containing bp 1–80 of IS10 from wild-type Tn*10*. This segment extends from an *Nru*I site in IS10 (converted to a *Sal*I site) through 37 bp of flanking *hisG* material, up to a *Hae*III site (converted to a *Bam*HI site) (Foster *et al.*, 1981; Halling *et al.*, 1982). The (O* × O$^+$) constructs (Figure 2) were generated by subcloning the mutant *Bam*HI–*Sal*I fragment from a wild-type pGC derivative and from each mutant pGC derivative into pNK1214, which is described in Huisman and Kleckner (1987). Essentially, the *Bgl*II–*Pst*I fragment indicated in Figure 2 formed half of a pBR-based Amp^r plasmid, the other half of which contains a portion of *Salmonella* DNA normally located upstream of the histidine operon plus half of the *amp* gene and origin from pBR322, and is the *Pst*I–*Nco*I segment of pNK75 (Foster *et al.*, 1981). pNK1215 (mutation *del* of Table I) is a deletion of pNK1214 from the *Bgl*II site to a *Bcl*I site located at bp 70 of the adjacent IS10 outside end (Huisman and Kleckner, 1987). (O* × O*) constructs were generated from the (O* × O$^+$) constructs by substitution of an *Nde*I–*Pst*I piece from the mutant pGC plasmids as shown in Figure 2; the *Nde*I site is at ~bp 63 of IS10. Marked IS10 and mini-IS10 (I$^+$ × O*) elements were constructed in a separate series based on pNK1256, which is a derivative of the marked IS10 plasmid pNK1166 (Roberts *et al.*, 1985) having a deletion between two *Hpa*I sites in Tn*10 tet* gene material located adjacent to the inside end. Three manipulations were carried out on this plasmid, in somewhat different orders for different mutations. (i) A mutant IS10 outside end was introduced on a *Bgl*II–*Nde*I fragment from the original pGC derivative. (ii) A series of *rrnB* transcription terminators was introduced on a *Bgl*II–*Pst*I fragment from pNK2333, which is a derivative of pRS391 (Simons *et al.*, 1987) carrying a *Bgl*II linker at its *Bam*HI site. (iii) The transposase gene was inactivated by deletion of material between the two (unique) *Nco*I sites in IS10. pNK1152 is a *Bam*HI–*Sal*I Tet^s deletion derivative of pACYC which contains at its unique *Hin*dIII site a mini-Tn*10*-Tet^r element composed of two appropriately oriented IS10 outside ends flanking the *tet* gene fragment from Tn*10*; details of construction are available upon request. pNK629 is a derivative of pACYC184 containing at its unique *Eco*RI site the *Eco*RI fragment of pNK474 (Morisato *et al.*, 1983) which carries the IS10 transposase gene driven by a *ptac* promoter.

Isolation of mutations

Mutations were isolated from pOH56 and pOH59 according to the method of Myers and Maniatis (1985). Briefly, single-stranded DNA was mutagenized *in vitro* with each of several different mutagens, converted to double-stranded form with reverse transcriptase and an appropriate oligonucleotide primer, digested with *Bgl*II and *Sal*I to release a fragment of ~130 bp containing the end of IS10, and the DNA electrophoresed in denaturing gradient gels which can separate molecules of identical length that differ in composition by a single base pair. Fragments migrating faster or slower than the wild-type fragment were eluted from the gel, cloned back into pGC1, and sequenced over the entire length of the mutagenized segment. Plasmids containing mutations in the region of interest were kept, and fragments subcloned into other plasmids for analysis (above). Several mutations were generated by synthesizing pairs of complementary oligonucleotides which were inserted into an appropriate plasmid backbone as double-stranded linkers to generate ends which were perfectly isogenic to those used for random mutagenesis: mutations 1A, 1G, 1T, 2G, 3A, 5A, 8C, 12T, IHFup and IHFdown.

Mating out transposition assays

Fresh transformants were picked with a sterile capillary into 1 ml of LB broth containing appropriate antibiotics and grown overnight to saturation at 37°C. Overnights were then diluted 1/20 into 1 ml of fresh broth lacking antibiotics, grown at 37°C on a roller drum with rapid rotation until cells reached a concentration of ~0.5×10^8 per ml, and then grown for an additional 90 min with very slow rotation. At this point, 2.5 ml of recipient cells (grown to 5×10^8 in LB) were added and incubation continued with slow rotation for ~90 min. Tubes were then placed on ice and plated on selective media, LB plates containing 200 μg/ml streptomycin sulfate (Sigma) and either kanamycin (50 μg/ml) or tetracycline (50 μg/ml). Previous control experiments have shown that matings performed in parallel differ in the concentration of total exconjugants by no more than ±20%. Thus, relative transposition frequencies are expressed as the relative concentrations of drug resistant exconjugants per ml of mating mixture plated. The absolute frequencies of transposition of the several wild-type constructs differ because of differences in the length, the nature (inside or outside) of the ends, the immediate context of the transposing segment, and the level of transposase provided either in *trans* from a *ptac* fusion or in *cis* from pIN (see table legends).

Analysis of IS903 mutations

Calculation of (B^2/A) for IS903 is more complicated than for IS10 because the IS903 mutations were analyzed in constructs having the transposase gene in *cis* to the pair of transposon ends, and the phenotypes of some mutations vary dramatically depending upon whether they are located proximal or distal to the transposase gene. For IS903, the parameter analogous to (B^2/A) is the product of the defects in the two singly mutant constructs, having the mutation either proximal or distal, divided by the defect in the doubly mutant construct. The fact that even mutations in the critical region exhibit (B^2/A) ratios significantly greater than 1 may be significant; perhaps mutations in the critical region can also be compensated by a wild-type end in some situations. The difference with IS10 could result from differences in behavior in the two types of tester constructs used, or intrinsic differences between the two elements.

Acknowledgements

This research was funded in part by a grant to N.K. from the National Science Foundation and in part by a grant to O.H. from the CNRS of France (730038) and from the Association de la Recherche sur le Cancer (6617). O.H. and L.S. are staff members of the Institut National de la Recherche Agronomique. We are grateful to A.Toussaint for advice concerning this manuscript.

References

Berthold,V. and Geider,K. (1976) *Eur. J. Biochem.*, **71**, 443–449.
Craig,N. and Nash,H. (1984) *Cell*, **39**, 707–716.
Davis,M.A., Simons,R.W. and Kleckner,N. (1985) *Cell*, **43**, 379–387.
Derbyshire,K.M., Hwang,L. and Grindley,N.D.F. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8049–8053.
Foster,T., Davis,M.A., Takeshita,K., Roberts,D.E. and Kleckner,N. (1981) *Cell*, **23**, 201–213.
Gamas,P., Chandler,M.G., Prentki,P. and Galas,D.J. (1987) *J. Mol. Biol.*, **195**, 261–272.
Guyer,M.S., Reed,R.R., Steitz,J.A. and Low,K.B. (1980) *Cold Spring Harbor Symp. Quant. Biol.*, **45**, 135–140.
Halling,S.M., Simons,R.W., Way,J.C., Walsh,R.B. and Kleckner,N. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2608–2612.
Huisman,O. and Kleckner,N. (1987) *Genetics*, **116**, 185–189.
Kleckner,N. (1989) In Berg,D.E. and Howe,M.M. (eds). *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 225–267.

Makris,J.C., Nordmann,P.L. and Reznikoff,W.S. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2224–2228.
Myers,R.M. and Maniatis,T. (1985) *Science*, **229**, 242–245.
Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Morisato,D. and Kleckner,N. (1985) *Cell*, **39**, 181–190.
Morisato,D. and Kleckner,N. (1987) *Cell*, **51**, 101–111.
Morisato,D., Way,J.C,. Kim,H.-J. and Kleckner,N. (1983) *Cell*, **32**, 799–807.
Nash,H.A., Bauer,C.E. and Gardner,J.F. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4049–4053.
Phadnis,S.H. and Berg,D.E. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 9118–9122.
Roberts,D.E., Hoopes,B.C., McClure,W.R. and Kleckner,N. (1985) *Cell*, **43**, 117–130.
Sherratt,D.J. (1989) In Berg,D.E. and Howe,M.M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 163–184.
Simons,R.W. and Kleckner,N. (1983) *Cell*, **34**, 683–691.
Simons,R.W., Hoopes,B., McClure,W. and Kleckner,N. (1983) *Cell*, **34**, 673–682.
Simons,R.W., Houman,F. and Kleckner,N. (1987) *Gene*, **53**, 85–96.
Way,J.C. and Kleckner,N. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3452–3456.
Yin,J.C.P. and Reznikoff,W.S. (1987) *J. Bacteriol.*, **169**, 4637–4645.
Yin,J.C.P. and Reznikoff,W.S. (1988) *J. Mol. Biol.*, **199**, 35–45.

Received on February 28, 1989