Structural and Functional Studies of Insertion Element IS200

Stephen Lam† and John R. Roth

Biology Department
University of Utah, Salt Lake City, UT 84112, U.S.A.

(Received 21 September 1983, and in revised form 15 August 1985)

The nucleotide sequence of the insertion element IS200 has been determined partially, including the junctions between the element and the host chromosome at the insertion site. At most, two bases (A-A) are found repeated at the junctions and could be duplications of host sequences generated by the insertion of the element. No obvious sequence repeats, either direct or inverted, have been detected between the sequences just within the two ends of the element. The element is an extremely strong block to host transcription across the insertion site. A sequence similar to known transcription termination signals was found just within the element near the right end. Removal of less than 50 base-pairs at the right end of the element abolishes the transcription block. The putative terminator sequence is located within this 50 base-pair region. Genetic studies suggest that the element contains a promoter located more than 93 base-pairs from its left end. The proposed promoter and terminator are in proper orientation to form an internal transcription unit.

1. Introduction

Insertion sequences (IS) belong to the class of genetic elements that can move from one location to another within the bacterial genome. These elements can also mediate various chromosomal rearrangements (including inversion, deletion, and fusion of DNA segments) and alter the expression of adjacent genes (for reviews, see Starlinger, 1989; Calos & Miller, 1980; Kleckner, 1981; Shapiro, 1983). Many IS elements have been studied in Escherichia coli. The complete nucleotide sequences of ISI, IS2, IS4 and IS5 have been determined (Ohtsubo & Ohtsubo, 1978; Johnsrud, 1979; Ghosal et al., 1979; Klaer et al., 1981; Engler & van Bree, 1981; Schoner & Kahn, 1981). All of the elements have terminal, inverted repeat sequences. The inserted elements are flanked by short duplications of host DNA at the sites of insertion. In all cases, nonsense codons are present in all three translation reading frames, which may partially account for the strong polar effects of these elements. In addition, sequences resembling known transcription termination signals (Rosenberg & Court, 1979) have been identified in IS2 and IS4.

Recently, a new IS element (IS200) was identified in Salmonella typhimurium. The element is specific to Salmonella and is present in almost all Salmonella species tested (Lam & Roth, 1983a). S. typhimurium strain LT2 has six copies of IS200, all of which have been mapped genetically (Lam & Roth, 1983b). The auxotrophic mutation hisD984 in the histidine operon was shown to be the result of insertion of IS200. Like most other insertion mutations, hisD984::IS200 is extremely polar, unlike other insertion mutations, however, it does not revert to prototrophy (less than 10^-11; see Lam & Roth, 1983a). Deletion derivatives that relieve the strong polar effect have been isolated (Hoppe & Roth, unpublished results). Genetic mapping studies show that some of these derivatives appear to have one endpoint at the site of the hisD984::IS200 mutation. These deletion derivatives have endpoints either within the 700 bpf of IS200 or at the junctions between IS200 and chromosomal material.

We have determined the nucleotide sequences at the junctions between the IS200 element and his sequences in the mutant hisD984::IS200, as well as the endpoints of several non-polar deletion derivatives. In this paper, we describe these results and those of other structural and functional studies, which demonstrate that IS200 contains transcription start and stop signals.

2. Materials and Methods

(a) Bacterial strains

All bacterial strains used in these studies are listed in Table 1. The nomenclature for Tn10 insertions has been

† Present address: CIBA-GEIGY, P.O. Box 12257, Research Triangle Park, NC 27709, U.S.A.

‡ Abbreviation used: bp, base-pair(s).
Table 1

Bacterial strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype†</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR3715</td>
<td>his-644</td>
<td>T. Kohno</td>
</tr>
<tr>
<td>TR3579</td>
<td>his-9569</td>
<td>I. Hoppe</td>
</tr>
<tr>
<td>TR3884</td>
<td>his-9574</td>
<td>I. Hoppe</td>
</tr>
<tr>
<td>TR3905</td>
<td>his-9580</td>
<td>I. Hoppe</td>
</tr>
<tr>
<td>TR397</td>
<td>his-9584</td>
<td>I. Hoppe</td>
</tr>
<tr>
<td>TR398</td>
<td>his-9588</td>
<td>P. E. Hartman</td>
</tr>
<tr>
<td>TR6282</td>
<td>his-2396</td>
<td></td>
</tr>
<tr>
<td>TR6287</td>
<td>his-5855</td>
<td></td>
</tr>
<tr>
<td>TR6288</td>
<td>his-5857</td>
<td></td>
</tr>
<tr>
<td>TR6289</td>
<td>his-5858</td>
<td></td>
</tr>
<tr>
<td>TT2144</td>
<td>zea-2::Tn10 his-9558</td>
<td>I. Hoppe</td>
</tr>
<tr>
<td>TT2146</td>
<td>zea-2::Tn10 his-9540</td>
<td>I. Hoppe</td>
</tr>
<tr>
<td>TT2148</td>
<td>zea-2::Tn10 his-9542</td>
<td>I. Hoppe</td>
</tr>
<tr>
<td>TT2151</td>
<td>zea-2::Tn10 his-9545</td>
<td>I. Hoppe</td>
</tr>
<tr>
<td>TT2152</td>
<td>zea-2::Tn10 his-9546</td>
<td>I. Hoppe</td>
</tr>
<tr>
<td>TT2153</td>
<td>hisG8543::Tn60 hisG9647::Tn5</td>
<td>D. Biek</td>
</tr>
<tr>
<td>TT6131</td>
<td>hisG9648::Tn3(tf-39) recA1</td>
<td></td>
</tr>
<tr>
<td>TT6783</td>
<td>metA22 trpB2 strA120 xyl-404 metE551 ile-1253 galE550 his-9539 zee-1::Tn10</td>
<td></td>
</tr>
<tr>
<td>TT6943</td>
<td>hisG9648::Tn3(tf-39) hisD9544::IS200</td>
<td></td>
</tr>
<tr>
<td>TT6944</td>
<td>hisG9648::Tn3(tf-39) hisD9547::IS200</td>
<td></td>
</tr>
<tr>
<td>TT6945</td>
<td>hisG9648::Tn3(tf-39) hisD9548::IS200</td>
<td></td>
</tr>
<tr>
<td>TT6946</td>
<td>hisG9648::Tn3(tf-39) hisD9549::IS200</td>
<td></td>
</tr>
<tr>
<td>TT6947</td>
<td>hisG9648::Tn3(tf-39) hisD9550::IS200</td>
<td></td>
</tr>
<tr>
<td>TT6948</td>
<td>hisG9648::Tn3(tf-39) hisD9551::IS200</td>
<td></td>
</tr>
<tr>
<td>TT6949</td>
<td>hisG9648::Tn3(tf-39) hisD9552::IS200</td>
<td></td>
</tr>
<tr>
<td>TT6950</td>
<td>hisG9648::Tn3(tf-39) hisD9553::IS200</td>
<td></td>
</tr>
<tr>
<td>TT6951</td>
<td>hisG9648::Tn3(tf-39) hisD9554::IS200</td>
<td></td>
</tr>
</tbody>
</table>

† Nomenclature for Tn10 insertions is as described by Chumley et al. (1979). Unless otherwise noted, all strains were constructed for this study.

described (Chumley et al., 1979). Strains TT6943 and TT6958 are derived from Bruce Stocker and C. Colson.

(b) Recombinant phages and plasmids

The recombinant phage M13Ho176 (Barnes, 1979) and plasmid pWB91 (Barnes, 1977) were kindly provided by W. Barnes. The phage M13Ho176 carries a 3300 bp insert in the M13 phage intergenic region; this insert contains the control region (hisO) and the first 2 structural genes (hisG and hisD) of the wild-type S. typhimurium his operon. This same portion of the his operon is present in the plasmid pWB91, a ColEl derivative. The mutation hisG9648::Tn5(tnf-39) is a transposition-defective derivative of hisG9648::Tn5 (Biek & Roth, 1981). Strains carrying this mutant insertion element in the hisG gene retain hisD gene expression, and the resulting strain is phenotypically HisD+ (can grow on histidinol as the source of histidine). The double mutant hisG9648::Tn5(tf-39) hisD984::IS200 was constructed by transduction. P22 phage were grown on a strain (TT6131) containing the mutation hisG9648::Tn5(tf-39) and used as donor to transduce a strain containing the mutation hisD984::IS200 to kanamycin resistance. Kanamycin-resistant transductants that are HisD- contain both mutations. Such a transductant (TT6943) was then used as donor to transduce to kanamycin resistance a strain (TT6783) carrying a non-transducible his deletion in the chromosome and the plasmid pWB91, which includes the promoter-proximal portion of the wild-type his operon (including the his control region.

d (e) DNA sequencing

DNA sequencing was carried out according to the methods of Sanger et al. (1977), Barnes (1978b) and Maxam & Gilbert (1980).

(f) Southern hybridization methods

All materials and methods for Southern (1975) hybridization experiments have been described (Lam & Roth, 1983a,b).

g (g) Construction of pWB91-hisG::Tn5 hisD984::IS200

The mutation hisG9648::Tn5(tf-39) is a transposition-defective derivative of hisG9648::Tn5 (Biek & Roth, 1981). Strains carrying this mutant insertion element in the hisG gene retain hisD gene expression, and the resulting strain is phenotypically HisD+ (can grow on histidinol as the source of histidine). The double mutant hisG9648::Tn5(tf-39) hisD984::IS200 was constructed by transduction. P22 phage were grown on a strain (TT6131) containing the mutation hisG9648::Tn5(tf-39) and used as donor to transduce a strain containing the mutation hisD984::IS200 to kanamycin resistance. Kanamycin-resistant transductants that are HisD- contain both mutations. Such a transductant (TT6943) was then used as donor to transduce to kanamycin resistance a strain (TT6783) carrying a non-transducible his deletion in the chromosome and the plasmid pWB91, which includes the promoter-proximal portion of the wild-type his operon (including the his control region.

(c) Media, plaque growth and transductional methods

All media, plaque growth and transductional methods have been described (Lam & Roth, 1983b).

(d) Preparation of DNA

Single-stranded and double-stranded DNA of various phages and plasmids were prepared according to Barnes (1977a,b). Restriction fragments were isolated and recovered from gels by the method described by Maxam & Gilbert (1980).

(e) DNA sequencing

DNA sequencing was carried out according to the methods of Sanger et al. (1977), Barnes (1978b) and Maxam & Gilbert (1980).

(f) Southern hybridization methods

All materials and methods for Southern (1975) hybridization experiments have been described (Lam & Roth, 1983a,b).

(g) Construction of pWB91-hisG::Tn5 hisD984::IS200

The mutation hisG9648::Tn5(tf-39) is a transposition-defective derivative of hisG9648::Tn5 (Biek & Roth, 1981). Strains carrying this mutant insertion element in the hisG gene retain hisD gene expression, and the resulting strain is phenotypically HisD+ (can grow on histidinol as the source of histidine). The double mutant hisG9648::Tn5(tf-39) hisD984::IS200 was constructed by transduction. P22 phage were grown on a strain (TT6131) containing the mutation hisG9648::Tn5(tf-39) and used as donor to transduce a strain containing the mutation hisD984::IS200 to kanamycin resistance. Kanamycin-resistant transductants that are HisD- contain both mutations. Such a transductant (TT6943) was then used as donor to transduce to kanamycin resistance a strain (TT6783) carrying a non-transducible his deletion in the chromosome and the plasmid pWB91, which includes the promoter-proximal portion of the wild-type his operon (including the his control region.
hisD, hisD and part of hisC. Since hisD984::Tn5(raf-39) is defective in transposition, kanamycin-resistant trans-
ductants arise only when the Tn5 mutation recombines
into the plasmid pWB91. Kanamycin-resistant trans-
ductants were grown in rich medium containing
kanamycin to allow for segregation of non-transduced
plasmids, and then scored for their HisD- phenotype.
Transductants that were phenotypically HisD had
inherited IS200 with the donor Tn5 insertion.

(h) Selection of non-polar derivatives of hisD984::IS200

Since hisD984::IS200 is absolutely polar, it is
phenotypically HisD- (by direct damage) and HisC-
because it blocks transcription. Genes located promoter-
distal to hisC- function because of a promoter located
between the hisC and hisB genes (Ciampi et al., 1982).
When hisD984::IS200 acquires an F'hisD+ hisC-
plasmid, it becomes phenotypically HisD+ (can grow on
histidinol as a histidine source). The diploid remains
hisD- and will not grow on minimal medium. Therefore,
selection can be made for HisC revertants. Among these
are the chromosomal deletion mutants that affect IS200
and are described in this paper.

(i) Construction of strains containing hisG::Tn10 and non-
polar deletion derivatives of hisD984::IS200

Non-polar deletion derivatives of hisD984::IS200 have
restored transcription of the adjacent downstream hisC
gene by removing the transcription stop signal at the
right end of hisD984::IS200. Since the deletions were
selected for hisC gene expression, all must retain the hisC
gene. Two exceptions are deletions his-9543 and 9545 (see
the strains described in the last paragraph to
hid- and will not grow on minimal medium. Therefore.
on the plasmid.

If transcription resulting strains were examined for their HisC phenotype.
transductants were grown in rich medium containing
is defective in transposition. kanamycin-resistant trans-
ductants arise only when the Tn5 mutation recombines

The approximate location of the mutation
hisD984::IS200 within the hisD gene was
determined from the following observations.
Genetic mapping studies (Hoppe & Roth, personal
communication) indicated that hisD984 is located
between the mutations hisD6404 and hisD2578.
The mutation hisD6404 has been shown by Bossi &
Roth (1980) to be a C to T base substitution at the
49th nucleotide from the beginning of the hisD
gene. The mutation hisD2578 has been shown by
Isoso & Yourno (1974) to affect a run of C residues,
which we now know starts at the 149th nucleotide
from the beginning of hisD. Thus, the site of
insertion of the mutation hisD984::IS200 is located
between the 49th and 149th nucleotides from the
beginning of hisD.

Figure 1 shows a restriction map of the region,
including the insertion. The nucleotide sequences of
the left and right junctions between the insertion
mutation and adjacent his material were
determined by the chain-termination method of
Sanger et al. (1977) and Barnes (1978b). The
template used was single-stranded DNA isolated
from M13H0176-hisD984::IS200. The primer used
for obtaining the left junction sequences was
pWB91 DNA digested with the enzyme PvuII. The
PvuII site in hisD (see Fig. 1) is the only PvuII site
on pWB91 (W. Barnes, personal communication).
The primer used for obtaining the right junction
sequences was the EcoRI-HindIII fragment
internal to IS200 (see Fig.1), isolated from
replicative form DNA of M13H0176-
hisD984::IS200. Autoradiograms showing the
junction sequences are presented in Figure 2.

Additional IS200 sequence near the right
junction was determined by the chemical cleavage
method of Maxam & Gilbert (1980), using the
HindIII fragment that contains the right junction.
Figure 1. Restriction map of the cloned region of the *S. typhimurium* his operon. The top line shows a genetic map of the entire his operon. The direction of transcription is from left to right. The bottom line shows the region of the operon that is present in M13Ho6-hisD984::IS200, with the locations of relevant restriction sites indicated above the line.

The following structural features are apparent: (1) no his sequence is missing; (2) there is a possible duplication of, at most a 2 bp sequence (A-A) at the corresponding wild-type his sequences (W. Barnes, personal communication), are shown in Figure 3.

The fragment was isolated from pWB91-hisG::Tn5 hisD984::IS200 plasmid DNA. Combined sequence data of the two junctions, together with the corresponding wild-type his sequences (W. Barnes, personal communication), are shown in Figure 3.

Figure 2. Autoradiograms of sequencing gels showing the junctions between hisD984::IS200 and his sequences. The left panel shows the left junction sequence, the middle panel shows wild-type his sequence, and the right panel shows the right junction sequence. Sequencing was done according to the methods of Sanger et al. (1977) and Barnes (1979).
Figure 3. Nucleotide sequences at the junctions of the insertion mutation hisD984 : : IS200. Only one strand of the DNA is shown. The top and bottom lines show wild-type sequences. The middle 2 lines show the sequences at the 2 junctions. Numbering of his sequence starts at the beginning of the hisD gene. Numbering of IS200 sequence starts at the left end of hisD984 : : IS200, and uses numbers followed by primes (1', 10', etc.). For the purpose of numbering, the length of the inserted sequence is arbitrarily taken to be 700 bp. The arrows indicate sequences that can potentially form a structure resembling known transcription termination signals (see Fig. 7). WT, wild-type. Sequence hyphens are omitted for clarity.

(a) A possible terminator near the right end of hisD984 : : IS200

Insertion mutant hisD984 : : IS200 blocks transcription of the adjacent promoter-distal hisC gene. Therefore, IS200 must include a transcription termination signal. Several spontaneous deletion derivatives of IS200 have been selected that have regained hisC expression (see Fig. 4). This selection method is described in Materials and Methods. Each of these deletions must have removed the region(s) of IS200 that is responsible for the polar effect. By analyzing the extent of such deletions, it was possible to infer the approximate location of the transcription termination site. The site removed by all non-polar deletion mutants lies within 50 bp of the right end of IS200. Examination of the DNA sequence in this region revealed a likely candidate for the termination signal.

The IS200 element contains a HindIII site near its right end. Restriction fragments ending at this site had been used to obtain DNA sequences at the right end of the element. Although the exact DNA sequence near the site was not determined, the approximate distance between the site and the right end of the element was estimated, from the spacing on the gels to be less than 50 bp. Several non-polar deletion derivatives entering IS200 from the right (see Fig. 4) were examined by Southern hybridization experiments, to see if they have removed this HindIII site. The results are shown in Figure 5. In these experiments, the DNA was digested with HindIII, and the probe used was M13Hol6. Lane 1 contains DNA from the wild-type strain LT2, lanes 2 to 8 contain DNA from the deletion derivatives, and lane 9 contains DNA from the parent strain carrying the mutation hisD984 : : IS200 (TR6238). HindIII does not cut within the cloned portion of hisC.
the wild-type his operon present in M13H0176. The wild-type strain LT2 therefore yielded only one hybridized fragment. The strain containing hisD984::IS200 (TR6238) yielded two hybridized fragments, because of the HindIII site within IS200. The larger fragment contains the his control region hisO, hisG and part of hisD; the smaller fragment (which runs much faster and is barely visible in Fig. 5) contains part of hisD and part of hisC (see Fig. 5). Each of the strains containing, respectively, the deletions his-9569 (lane 2), his-9585 (lane 3), his-9546 (lane 4), his-9540 (lane 5), and his-9574 (lane 6) yielded one hybridized fragment with a molecular weight differing from that of the larger fragment in the parent strain carrying hisD984::IS200 (TR6238; lane 9), suggesting that these deletions have removed the HindIII site in IS200 and have fused the two fragments in the parent strain (TR6238). The deletions his-9542 (lane 7) and his-9545 (lane 8) yielded one hybridized fragment with the same molecular weight as the large fragment in the parent strain (TR6238), suggesting that these two deletions, which enter the element from the right have not removed the HindIII site but have removed essentially all of the second fragment. An alternative possibility is that these deletions have removed the HindIII site, but the fusion fragment is fortuitously of the same size as the larger fragment in the parent strain (TR6238). This possibility is eliminated by the Southern blot experiments presented below.

Further Southern hybridization analyses of these deletion derivatives indicate that the deletion his-9542 does not remove the HindIII site. The results of these analyses are shown in Figure 6. The top half of Figure 6 shows an autoradiogram from Southern hybridization experiments designed to define the left endpoints of the two deletions his-9542 and his-9545. The bottom half shows the interpretation of the data. Four different strains were examined in two sets of experiments: the
Studies of IS200

Figure 6. Southern blot analysis of deletion derivatives of hisD984::IS200. Top: autoradiogram of Southern blot. The probe used was M13H016, which includes material indicated by the unbroken horizontal line in the map, but without IS200. Two sets of double digest experiments are shown: in the first set (lanes 4 to 7), DNA was digested with HincII and HindIII; in the second set (lanes 8 to 11), DNA was digested with HincII and EcoRI. Four strains were analyzed in each set of experiments: the parent strain containing hisD984::IS200 (TR6238; lanes 5 and 9) and 3 deletion derivatives: his-9540 (TT2146; lanes 4 and 8), his-9542 (TT2148; lanes 6 and 10) and his-9545 (TT2151; lanes 7 and 11). Lanes 2 and 3 contain DNA from the parent strain (TR6238) digested with EcoRI and HindIII, respectively. Lanes 1 and 12 contain the molecular weight standards: DNA from the plasmid pWB91, digested with HincII. This plasmid contains the promoter-proximal portion of the his operon. Bottom: restriction map of the region and interpretation of the hybridization data.

Parent strain (containing hisD984::IS200) and three deletion derivatives, his-9540, his-9542 and his-9545. The deletion his-9540 has been sequenced partially (discussed below). Although the exact junction has not been determined, it is known that the deletion enters IS200 from the right leaving approximately 70 bp at the left end of hisD984::IS200 and has thus removed both the EcoRI and the HindIII sites internal to the IS200 element.

In the first set of experiments (Fig. 6, lanes 4 to 7), DNA from each strain was digested by the two enzymes HincII and HindIII. The larger fragment in each lane (4 to 7) corresponds to the HincII-HindIII fragment, including part of hisG and most of the IS element. It can be seen that for his-9540 (lane 4), this fragment size is different from that of the other strains (lanes 5, 6 and 7). This is expected, since the deletion has removed the HindIII site, thus joining the fragments on either side of it. The other two deletion strains (lanes 6 and 7) yielded fragments with the same size as the parent (lane 5). The simple interpretation is that the deletions end to the right of the HindIII site, thus leaving the
HincII-HindIII fragment intact. The second set of experiments (Fig. 6, lanes 8 to 11) demonstrates that the EcoRI site of IS200 also is removed by deletion his-9540 and retained by the other deletions. This set of experiments is identical to the first, except that the DNA was digested with the enzymes HincII and EcoRI. The larger fragment in each lane (8 to 11) corresponds to the HincII-EcoRI fragment spanning hisG and part of the IS element. Deletion mutant his-9540 (lane 8) yielded a large fragment, different in size from that of the other strains, while the other two deletion strains (lanes 10 and 11) yielded fragments of identical size to those of the parent (lane 9). Thus deletions his-9542 and his-9545 leave both the EcoRI and the HindIII sites of IS200. In both of the above digests, extra bands are predicted that are not apparent in the Figure. Several of these fragments comigrate as the smaller band. (Note the differences in intensity of the band in various deletions.) The other fragments are much smaller and blot so poorly they are not apparent in the photograph.

The above conclusion is strengthened because one can predict that a larger HincII-HindIII fragment would have been detectable had these deletions removed the HindIII site. Genetic mapping data (Hoppe & Roth, personal communication) indicate that the right endpoint of his-9542 is within hisD and, therefore, retains the EcoRI site in hisC. Deletion his-9545, though entering the hisC gene, expresses hisC function. It must have removed only a small, apparently dispensable, portion of the hisC gene and is also likely to have retained the EcoRI site in hisC. The distance from EcoRI to HindIII in hisC is detectably larger than the distance from EcoRI to HindIII in IS200. (This is demonstrated by the difference in the smallest bands in lanes 2 and 3 in Fig. 6.) Knowing the above facts, one can predict that if deletions 9542 or 9545 ended between the EcoRI and the HindIII sites of IS200, the HincII-HindIII fragment generated would have been considerably larger than the parent fragment. This was not the case (lanes 5 to 7 in Fig. 6).

Thus, the deletion his-9542, which abolished the strong polar effect of the mutation hisD984::IS200, has removed less than 50 bp of the right end of the element. The conclusion is, therefore, that a strong polar signal is encoded within 50 bp of the right end of hisD984::IS200. This sequence was scanned for potential termination sequences. The most promising candidate is presented in Figure 7. All deletions that destroy the transcription blocking ability of IS200 remove this sequence. The sequence has a strong resemblance to known transcription termination sites (Rosenberg & Court, 1979; Platt & Bear, 1983). Therefore, we feel that the sequence is likely to serve as a transcription block within IS200.

The DNA sequences of the deletion derivatives his-9540, his-9557, his-9582 and his-9587 have been determined. The sequence data obtained support the above conclusion. The deletions were crossed into M13H176 by homologous recombination (Lam & Roth, 1983a). Single-stranded DNA was isolated from the resulting phages and sequences were obtained by the chain-termination methods of Sanger et al. (1977) and Barnes (1978b). The primer used in these experiments was pWB91 DNA digested with the enzyme PvuII. The extents of the deletions are shown in Figure 8. All of the deletions end within the IS200 element, leaving behind different amounts of the element. The nucleotide sequences at the joint points of the deletions are shown in Figure 9. The DNA sequences of deletion his-9540 has been determined only partially. The exact sequence of an approximately 10 bp region surrounding the joint point was not determined. This uncertainty is indicated by brackets in Figure 9. It can be seen that the two deletions entering the IS200 element from the left, his-9582 and his-9557, have removed practically all of the element, leaving only 5 bp and 11 bp, respectively, at the right end, whereas the two deletions entering the element from the right, his-9587 and his-9540, have left 93 bp and approximately 70 bp, respectively, at the left end of the element. (We

---

**Figure 7.** Potential transcription terminator structure near right end of hisD984::IS200. This structure was removed by all deletions analyzed that removed the transcription blocking ability of IS200. Sequence hyphens have been omitted but base-pairing is indicated.

---

**Figure 8.** Extent and general location of 4 non-polar deletion derivatives of hisD984::IS200, his-9582, his-9587 and his-9540.
Figure 9. Nucleotide sequences at the joint points of 4 non-polar deletion derivatives of hisD984::IS200: his-9582, his-9557, his-9587 and his-9540. Only one strand of the DNA is shown. The numbering system is as described in the legend to Fig. 3. The vertical arrows indicate the junction between his and IS200 sequences. In deletion his-9587, sequence A-T on both sides of the joint point makes it impossible to identify the actual site of the deletion: only one possibility is shown. The DNA sequence of his-9540 has been determined only partially. The exact sequence of an approximately 10 bp region surrounding the joint point was not determined. This uncertainty is indicated by brackets in the Figure. Sequence hyphens are omitted for clarity.

Studies of IS200

(c) IS200 includes a promoter function

The deletion derivatives of hisD984::IS200 described have apparently restored hisC gene expression by removing a transcription stop signal at the right end of hisD984::IS200. This transcript could have started at the his promoter and extended across the remaining portion of hisD984::IS200 and into hisC. Alternatively, the transcript could start within the IS200 element and extend across the right end of the element into hisC. To distinguish between the two possibilities, a hisG::Tn10 mutation was introduced into a number of strains carrying different non-polar derivatives of hisD984::IS200 (see Materials and Methods). These Tn10 insertions in hisG are absolutely polar (Ciampi et al., 1982). The hisG::Tn10 mutation prevents any transcription initiated at the his promoter from reaching the hisD and hisC genes. Strains containing both the hisG::Tn10 insert and the non-polar derivative of IS200 were tested for their ability to express the hisC gene. (If the strains are HisC+ they will complement an F' hisC+ mutation.) If transcription originates only at the his promoter, the strains should become HisC-; if transcription is initiated within hisD984::IS200, the strains should remain HisC-.

Five derivatives were tested. The deletions his-9585 (TR5395), his-9587 (TR5397), and his-9542 (TT2148) have removed his sequences as well as IS200 sequences (see Fig. 4). For the other two derivatives, his-9588 (TR5388) and his-9538 (TT2144), no deletions of his material was detected in genetic mapping crosses (Hoppe & Roth, personal communication); presumably, these derivatives harbor changes internal to the IS element that affect the terminator structure. Introduction of the hisG::Tn10 mutation into the strains containing his-9585 and his-9587 rendered these strains HisC-, indicating that transcription was initiated at the his promoter only; presumably these deletions remove the promoter within IS200. The other three strains (9542, 9588 and 9538) remained HisC-, indicating that transcription was initiated within hisD984::IS200 sequences.
4. Discussion

The Salmonella-specific transposable element, IS200, shows several unusual features.

(1) Although insertion of the element did not remove any his sequences, the insertion mutation does not revert by precise excision. Most transposable elements studied do excise precisely at detectable frequencies (for a review, see Calos & Miller, 1980). The only other exceptions are Tn1 (Weinstock & Botstein, 1979; Hernalsteens et al., 1977) and Mu (Bukhari et al., 1971). The failure of the IS200 insertion to revert may be a consequence of the lack of repeated sequences within the element or due to the minimal duplication of host sequences flanking the insertion (Albertini et al., 1982). In other elements, ability to excise precisely has been attributed to both of these structures (for a review, see Kleckner, 1981).

(2) At most, 2 bp of his sequences may have been duplicated at the ends of the element. Since this is the only insertion site studied, the actual ends of the element cannot be determined. All other transposable elements studied have been found to generate more extensive duplication of host sequences at the insertion sites (see Calos & Miller, 1980). The extent of this duplication differs among elements, but the majority generate either a 9 bp or a 5 bp duplication. There does not seem to be any correlation between the length of duplication and excision frequency. The element IS1, which generates an 11 bp or 12 bp repeat (Haberman et al., 1979), excises precisely at a frequency of 1 per 10^6 cells (Pfeifer, cited by Starling, 1980), whereas IS2, which generates a 9 bp repeat, reverts at a frequency as high as 10^-6 per cell plated (Jordan et al., 1968).

(3) Comparison of the sequences at the ends of IS200 revealed no obvious relationship between them. There is no hint of an inverted or a direct repeat, which is a common feature among transposable elements. It is interesting to note that phage Mu, which also does not revert spontaneously, is the only other transposable element without obvious inverted repeats.

Southern hybridization studies and sequence data suggest that a transcription termination signal is located at the right end of the IS200 element. The element is strongly polar. A non-polar deletion derivative that enters the element from the right was shown to have removed less than 50 bp at the right end of the element. The two derivatives tested that enter the element from the left were both shown to have removed IS200 material from this 50 bp region. Examination of the DNA sequence in this region revealed a sequence similar to known transcription termination signals (Rosenberg & Court, 1979). It seems likely that this structure is responsible for the transcription block caused by the insertion.

The DNA sequence of four deletion derivatives that end within the IS200 element have been determined. All have very different endpoints, suggesting that they are spontaneous deletions that were probably not formed by events catalyzed by the transposition activity of the element.

Genetic studies of the derivatives of hisD984::IS200 that express the downstream hisC gene indicate that, in some of these derivatives, transcription initiates within IS200 sequences. Of the three mutants that show this property, his-9538 and his-9588 have not been studied physically. Genetic mapping studies indicate that the mutations involved have not removed any his sequences adjoining the insertion (Hoppe & Roth, personal communication). The changes in hisD984::IS200 sequences in these mutations may be localized near the region of the terminator structure. The third mutation showing transcription from within IS200, his-9542, has been shown by Southern hybridization studies to have removed less than 50 bp at the right end of hisD984::IS200. Thus, most of the element remained intact in the strain carrying this mutation. The mutation his-9587, which showed no transcription initiated from within hisD984::IS200 sequences, has removed all but 83 bp of hisD984::IS200 sequences at the left end. These data are consistent with the hypothesis that hisD984::IS200 contains a promoter located more than 93 bp from its left end. The proposed promoter and terminator are in proper orientation to form a transcription unit within IS200.

References


*Edited by I. Herskowitz*