Cell, Vol. 19, 795–805, March 1980, Copyright © 1980 by MIT

The Inverted Repeats of Tn*5* Are Functionally Different

Steven J. Rothstein, Richard A. Jorgensen,* Kathleen Postle† and William S. Reznikoff
Department of Biochemistry
College of Agricultural and Life Sciences
University of Wisconsin
Madison, Wisconsin 53706

Summary

The inverted repeats of Tn*5*, which have identical restriction endonuclease cleavage patterns, have different functional properties. They differ with respect to RNA polymerase binding, full promotion of neomycin resistance, the polypeptides coded for by the repeats and their function in the transposition process. There is a weak RNA polymerase binding site present in one repeat and not in the other which seems to be important for neomycin resistance. The two inverted repeats code for polypeptides of different molecular weights, with each repeat appearing to encode two polypeptides. The polypeptides from only one of the repeats of Tn*5* appear to be absolutely required for Tn*5* transposition.

Introduction

The discovery in recent years of a variety of transposable elements has been helpful in elucidating some of the dynamic processes involved in shaping the organization of procaryotic genomes. The ability of transposons to "move" from one replicon to another allows the rapid proliferation of an environmentally selected marker within bacterial populations. In retrospect, it is reasonable that the DNA encoding resistance to a number of different antibiotics, conferring strong selective advantage in the presence of drug, has often been found to be localized on transposable elements. A variety of other genetic markers ranging from the lactose operon (Cornelis, Ghosal and Saedler, 1978) to enterotoxin genes (So, Heffron and McCarthy, 1979) have also been found to reside on transposons in naturally occurring isolates. It seems possible that virtually any gene could exist in this state.

Many of the transposable drug-resistant elements, when analyzed by electron microscopy, have been found to contain repeated DNA sequences at the ends of the element (Kleckner, 1977). It has been suggested that these repeated sequences were originally derived from two homologous IS insertion sequences, with the IS elements providing the genes necessary for the transposition process (Kleckner, 1977; MacHattie and Jackowski, 1977). If this hypothesis is correct, then in those transposable elements such as Tn*5*, where the inverted repeats are still substantial in length, one would expect that the transposition functions would be encoded primarily by the repeated DNA. The resistance gene would then be localized in the unique central region of the transposon. One of the purposes of this study was to attempt to analyze the importance of the inverted repeats of Tn*5* in the transposition process.

Transposition involves the "movement" of a distinct physical segment of DNA from one site in a cell's DNA to another, in the absence of general recombination systems (Kleckner, 1977). The exact mechanism by which this process takes place is still poorly understood. Deletion analysis of Tn*3*, which encodes resistance to ampicillin, has demonstrated that removal of its 39 bp inverted repeat resulted in a transposition-deficient mutant. No complementation by wild-type DNA was possible. A deletion in the central region, however, which also led to a loss in the ability to transpose ampicillin resistance, was complementable by wild-type Tn*3* DNA (Heffron et al., 1977). Thus in Tn*3* there appears to be a cis-dominant structural feature of the DNA at the ends of the element, as well as a trans-acting factor, presumably a protein encoded by the transposon, needed for full transposition activity.

Tn*5* is a 5400 bp piece of DNA which encodes resistance to some aminoglycoside antibiotics including kanamycin and neomycin. It has 1450 bp inverted repeats when measured by electron microscopy (Berg et al., 1975). The repeats flank a central region 2500 bp in length (Berg et al., 1975). Of the more than fifteen restriction sites already mapped in the inverted repeats, each is in an identical location in the two repeats (Jorgensen, Rothstein and Reznikoff, 1979; and our unpublished results). We have found, however, that the two inverted repeats in Tn*5* have different functional properties.

We have constructed a variety of deletion, insertion and substitution mutants of Tn*5*. Analysis of these mutants has allowed us to study some of the transcription, translation and transposition properties of Tn*5*. The two repeats have different properties with respect to RNA polymerase binding. Only one of the inverted repeats is capable of promoting full neomycin resistance. Furthermore, the two repeats encode polypeptides of different molecular weights, and it appears as though only one of the repeats encodes a protein absolutely required for the transposition process.

Results

Construction of Mutants of Tn*5*

The restriction endonuclease map shown in Figure 1 was generated previously by an analysis of pRZ102 and pRZ112 plasmid DNAs (Jorgensen et al., 1979).

* Present address: Carnegie Institution of Washington, 290 Panama Street, Stanford, California 94305.
† Present address: Department of Microbiology, University of California College of Medicine, Irvine, California 92717.

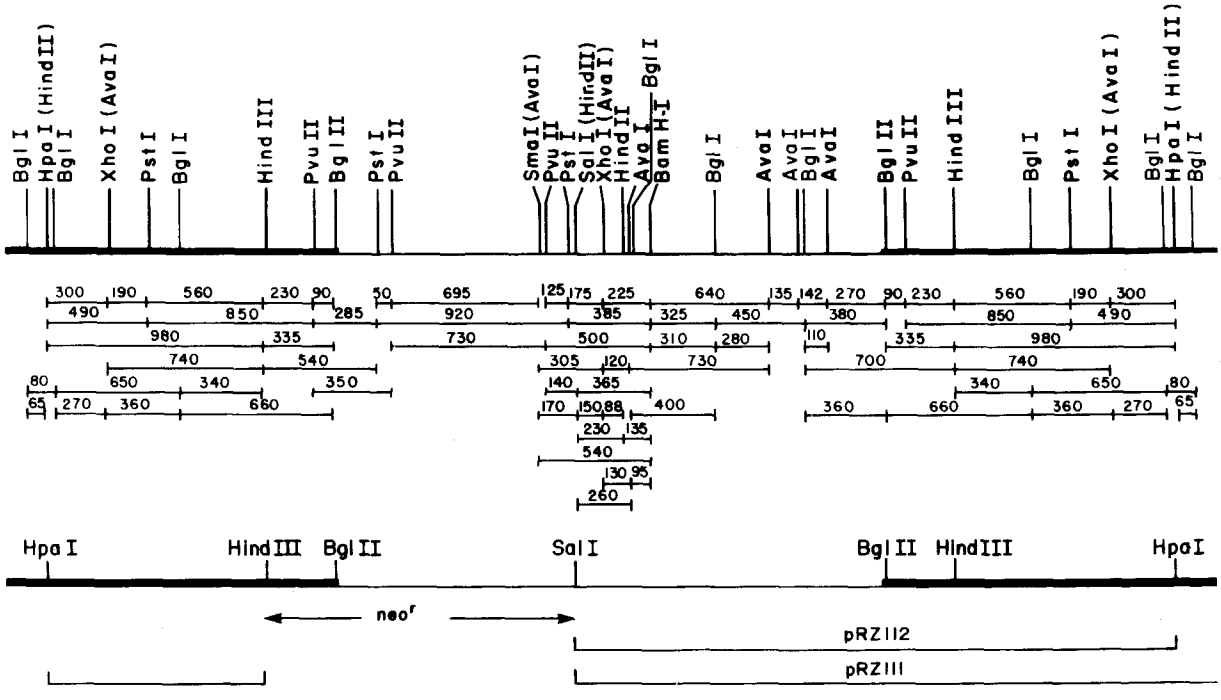


Figure 1. Restriction Map of Tn5

This restriction map of Tn5 was described earlier (Jorgensen et al., 1979). The DNA region necessary for neomycin resistance is shown, as is the structure of pRZ112.

The region essential for neomycin resistance lies between the Hind III site in the left inverted repeat and the Sal I site in the central region of Tn5. The restriction endonuclease map reveals a number of cleavage sites of interest for the generation of mutants. All of the mutants described below were constructed starting with either pRZ102 or λ:Tn5 DNA. The detailed methods by which the mutations were constructed and analyzed are described in Experimental Procedures. Their structures are shown in Figure 2.

RNA Polymerase Binding to Tn*5*

The binding of RNA polymerase to DNA can be used to a first approximation for the localization of promoters. RNA polymerase binding to a variety of restriction digests of Tn5 DNA revealed three sites of stable complex formation. The results of one set of binding experiments are shown in Figure 3.

An Eco RI + Hind II + III digest of pRZ112 produces four fragments. Only the largest band, which contains virtually the entire Col E1 plasmid, and the smallest, which is the Eco RI-Hind II fragment containing the left end of Tn5, bind polymerase strongly. The 80 bp of Col E1 DNA present in this fragment are almost certainly not the source of the interaction between RNA polymerase and the fragment, since the 95 bp Eco RI-Pst I fragment which contains these sequences does not bind RNA polymerase (our unpublished results).

The binding of RNA polymerase to a Pst I + Hind II digest of pRZ112 shows that the Pst I-Hind II fragment at the other end of Tn5 also binds polymerase. Since only ~10 bp of Col E1 are present in the fragment, the binding site must be within Tn5.

A third, weaker RNA polymerase binding site in Tn5 can also be detected. When Eco RI + Hind II + III-cleaved DNA is used in the binding experiment, the Hind III-Hind II (also Sal I) 1450 bp fragment also binds RNA polymerase weakly (Figure 3a). When Pst I + Hind II-cleaved DNA is bound to RNA polymerase, a Tn5 Pst I-Pst I 1085 bp fragment also binds weakly (Figure 3b). Further analysis of RNA polymerase binding to pRZ112 DNA, shown in Figure 3c, demonstrated that the Hind III-Bgl II fragment contained this third, weaker binding site. The Hae III fragment in pRZ102 encompassing this DNA also binds RNA polymerase weakly. However, the Hae III fragment that contains the Hind III-Bgl II fragment in the other inverted repeat shows no binding activity. This was the first indication that the two inverted repeats of Tn5 are not identical. The locations of the RNA polymerase binding sites in Tn5 are shown in Figure 3d.

Inversion of the Central Region of Tn*5* Can Affect Drug Resistance

The insertion of DNA at the Bgl II site in pRZ112 has previously been shown to reduce the level of neomycin resistance drastically (Jorgensen et al., 1979). The expression of the drug resistance is therefore dependent upon the adjacent inverted repeat for full activity. The RNA polymerase binding experiments discussed

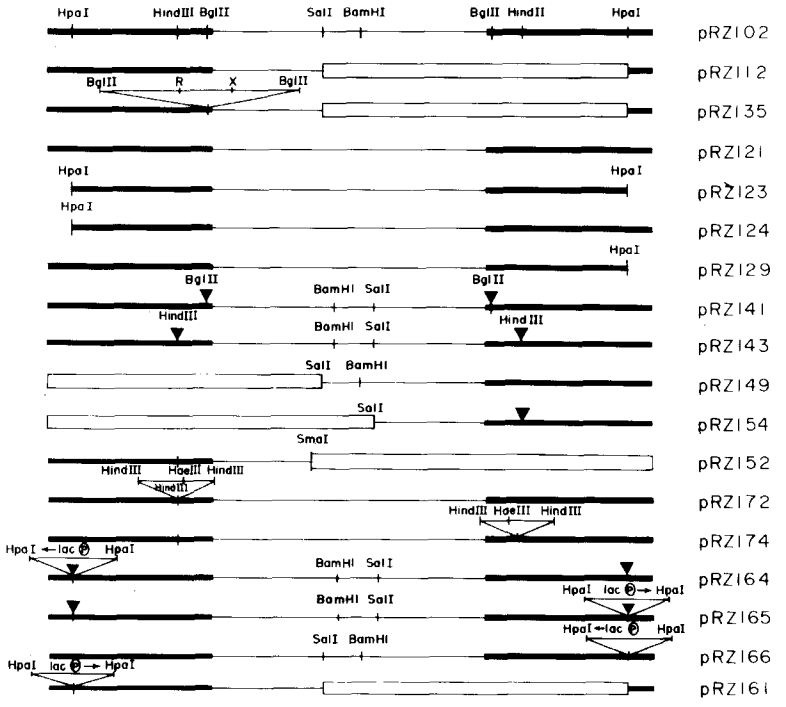


Figure 2. Structures of Tn5 Mutant Plasmids

The construction and analysis of the Tn5 mutants are described in Experimental Procedures. pRZ112 and pRZ135 have been described (Jorgensen et al., 1979). (▼) denotes the presence of an inversion of Tn5, with its locations delineating the inversion's endpoints. The Bam HI and Sal I sites are in reverse order compared with wild-type, where an inversion is present. pRZ154 was made from pRZ143 (the Hind III inversion strain) and thus still codes for neomycin resistance despite the absence of the left half of Tn5. Both pRZ149 and pRZ154 are actually substitution strains rather than deletion strains, since the Hae III 203 bp promoter-operator fragment is substituted for the Tn5 DNA removed. There is a single Hae III cut in the 515 bp fragment inserted in pRZ172 and pRZ174. Its location denotes the orientation of the insert with respect to the inverted repeats. The direction of transcription of the lactose promoter is designated by the arrows in pRZ161, pRZ165, pRZ166 and pRZ161.

above demonstrated that the two inverted repeats have different binding properties in the region between the Hind III and Bgl II sites. It seemed possible that an inversion of the unique region of Tn5, bringing the neomycin resistance gene adjacent to the other inverted repeat, might lead to an alteration in the level of drug resistance.

Two inversions of Tn5 shown in Figure 2 can be used to test this prediction. As shown in Table 1, pRZ143 (an inversion at the Hind III sites) codes for a level of resistance close to that coded for by wild-type Tn5, pRZ102. pRZ141, however, which inverts the central region at the Bgl II sites, reduces the level of resistance dramatically. The deletion, pRZ112, also has wild-type resistance levels. As noted above, when DNA is inserted into the Bgl II site of pRZ112, as in pRZ135, there is a drastically reduced level of resistance.

These results illustrate that not only is some function of the repeat adjacent to the neomycin resistance gene necessary for expression of full drug resistance, but that the other repeat is incapable of supplying all the necessary sequences. This difference between the two inverted repeats must lie between their Bgl II and Hind III sites.

Transcription

The RNA polymerase binding sites near the ends of the inverted repeats in Tn5 are potential promoters for the transcription of the repeats. The in vitro transcription of purified restriction fragments has been useful in delineating the site of transcription initiation in several systems, including the lactose promoter (for example see Maquat and Reznikoff, 1978). Furthermore, the direction of transcription can be tested by transcribing restriction fragments having different endpoints and analyzing the size of the transcripts which run off the ends of the fragment. The RNA polymerase binding site near the end of the left inverted repeat was analyzed in this fashion.

Restriction fragments were isolated which all contained the Col E1 Eco RI site on one end and had varying endpoints in Tn5. The Eco RI-Hpa I fragment is 260 bp, the Eco RI-Hae II fragment 210 bp and the Eco RI-Hae III fragment 160 bp long. In vitro transcription was carried out under the conditions described in Experimental Procedures, and the RNA products were run on a polyacrylamide-urea gel and analyzed by autoradiography. The results are shown in Figure 4. Discrete transcripts are made from each of these fragments, and their sizes decrease in amounts corresponding to the decrease in the size of the restriction fragment. In vitro RNA synthesis from a site near the outside end of this repeat therefore transcribes the inverted repeat toward the unique region of the transposon.

Two discrete transcripts, separated by approximately 10 bases, are synthesized utilizing each of the restriction fragments as templates. The size difference between the two must occur at the 5′ end of the transcripts, since a change at the 3′ endpoints does not affect their synthesis. In the case of the Eco RI-Hae III fragment, very little of the larger RNA is produced. The significance (if any) of having two in vitro transcription initiation sites ten bases apart is unknown.

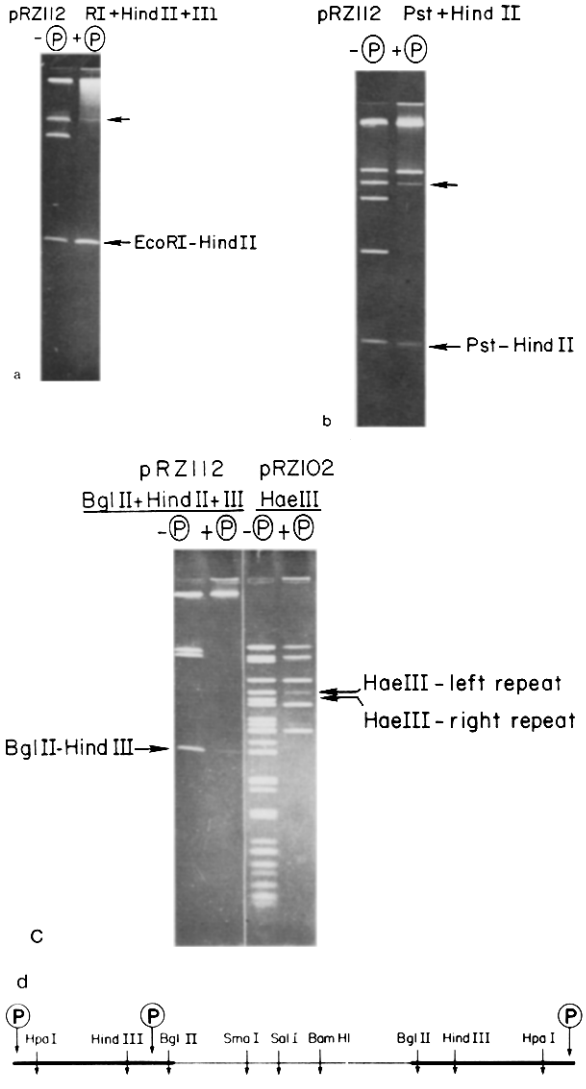


Figure 3. RNA Polymerase Binding to Tn5 DNA

(a) The RNA polymerase binding fragments of pRZ112 DNA digested with Hind II + III + Eco RI (+Ⓟ) are electrophoresed in a lane adjacent to a total Hind II + III + Eco RI digest of pRZ112 DNA (−Ⓟ) on a 5% polyacrylamide gel. The isolation of the RNA polymerase binding fragments is described in Experimental Procedures. The strong binding fragment noted is the Eco RI-Hind II fragment from the left end of Tn5. The weak binding fragment noted is the Hind III-Hind II (Sal I) fragment. (The Hind II site is the one also cleaved by Sal I in the unique region of Tn5.)

(b) The RNA polymerase binding fragments of pRZ112 DNA digested with Hind II + Pst I (+Ⓟ) electrophoresed in a lane adjacent to a total Hind II + Pst I digest of pRZ112 (−Ⓟ) on a 5% polyacrylamide gel. The strong binding fragment is the Pst I-Hind II fragment containing the right end of Tn5. The weak binding fragment noted is a Pst I-Pst I fragment which contains the Hind III-Bgl II fragment in the left repeat of Tn5.

(c) The RNA polymerase binding fragments of pRZ112 DNA digested with Hind II + III + Bgl II (+Ⓟ) electrophoresed in a lane adjacent to a total Hind II + III + Bgl II digest of pRZ112 (−Ⓟ) on a 5% polyacrylamide gel. The figure also shows the RNA polymerase binding fragments of pRZ102 digested with Hae III (+Ⓟ) electrophoresed next to a total Hae III digest of pRZ102 (−Ⓟ). The Hae III fragment which contains the sequences between the Bgl II and Hind III sites in the left inverted repeat of Tn5 binds RNA polymerase weakly, while the restriction fragment immediately below it, which contains the Bgl II-Hind III region, does not bind at all (see Jorgensen et al., 1979, for the location of the Hae III sites in Tn5.)

(d) The location of the RNA polymerase binding sites in Tn5.

The RNA polymerase binding site near the end of the right inverted repeat is also able to initiate transcription from a site similar in location to that described above for the left repeat (our unpublished results), although no detailed analysis concerning the direction of transcription was carried out.

Tn5-Encoded Protein Synthesis

The synthesis of proteins encoded by extrachromosomal elements can be analyzed through the use of minicell-producing bacterial strains. Plasmid DNA present in the cell, but not chromosomal DNA, segregates into the minicells, and only transcription and translation stemming from the plasmid should occur.

pRZ102 and pRZ104 both contain wild-type Tn5 transposed into different sites in the colicin gene of Col E1. A comparison of polypeptide synthesis in minicells directed by pRZ102 and pRZ104 with that by Col E1 reveals the synthesis of five new polypeptides that can be ascribed to Tn5 (Figure 5). One further difference between pRZ102, pRZ104 and Col E1-directed polypeptide synthesis involves the size of the colicin protein synthesized. Both Tn5 insertions produce a protein of smaller molecular weight, as would be expected.

The approximate molecular weights of the five Tn5 polypeptides are 58K, 54K, 53K, 49K and 26K. The 26K protein almost certainly must be neomycin phosphotransferase II (NPTII), which confers drug resistance upon the cell. Its molecular weight has been reported to be 25K, very close to that measured here (Matsuhashi et al., 1976). The smallest of the other four polypeptides (protein 4) was found to be encoded by the left inverted repeat of Tn5 (Postle and Reznikoff, 1979). The combined molecular weights of these polypeptides would appear to require a coding capacity larger than that of Tn5.

To understand the Tn5-encoded polypeptide synthesis more thoroughly, the Tn5 mutants were analyzed in minicells. pRZ172 has an insertion at the Hind III site in the left inverted repeat, while pRZ174 has an insertion in the Hind III site of the right repeat. The polypeptides encoded by these plasmids as well as pRZ152 and pRZ149, which delete the right or left halves of the transposon, respectively, are also shown in Figure 5.

The Hind III insertion into the right inverted repeat leads to a failure to synthesize polypeptides 1 and 2, while an insertion in the opposite repeat leads to an absence of 3 and 4. Each of the insertion strains synthesize two fusion proteins, f1 and f2. The fusion proteins in the two insertion strains are very similar in

Table 1. Neomycin Resistance of Strains Containing Tn5 Mutant Plasmids

Plasmid Present	EOP_{50} (μg/ml neo)
pRZ102	90
pRZ112	75
pRZ135	2
pRZ143	90
pRZ141	10

Analysis of neomycin levels required to kill 50% of the cells (EOP_{50}s) of strains containing Tn5 wild-type or mutant plasmids. The strains were tested as described in Experimental Procedures. The structures of the plasmids are shown in Figure 2.

molecular weight. The difference in size between f1 and f2 (42K and 38K) is similar to the corresponding differences between polypeptides 1 and 2 and polypeptides 3 and 4.

Assuming that transcription initiates near the outside edge of the transposon, the fact that there are two distinct fusion proteins from each repeat strongly supports the notion that the difference in molecular weight between the polypeptides from each repeat occurs at their N terminal rather than their C terminal ends. Furthermore, the similarity in size of the fusion proteins from the two different repeats implies that the difference between the two inverted repeats of Tn5 with respect to protein synthesis is most profound near the inside edge of the inverted repeat. Thus the difference in the polypeptides coded for by the two repeats with respect to size occurs at the C terminal end.

The large deletions of Tn5, pRZ152 and pRZ149, effect protein synthesis in the expected fashion. The deletion strain which does not contain the right inverted repeat of Tn5 (pRZ152) does not synthesize polypeptides 1 and 2, while the left half deletion strain (pRZ149) fails to produce 3 and 4. The 26K NPTII protein is still coded for by pRZ152 but not by pRZ149, which does not confer neomycin resistance.

pRZ124 and pRZ129 each delete Tn5 DNA from the Hpa I site outward, into the right or left inverted repeat, respectively. pRZ123 removes both outside edges of the transposon. Since the RNA polymerase binding sites and the transcription initiation site described earlier are in the material deleted by these strains, one might expect that these mutations would have a strong effect on the polypeptides synthesized by the inverted repeats. This turns out to be the case. Figure 6 is an autoradiogram of the polypeptides synthesized in minicells containing pRZ121, pRZ123, pRZ124 and pRZ129. pRZ121 contains a wild-type

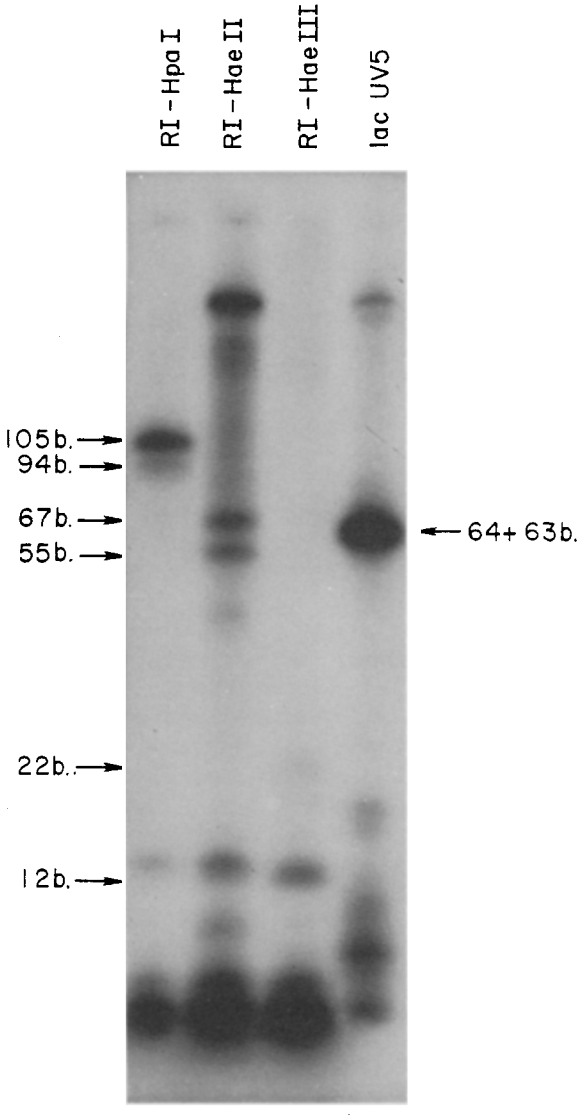


Figure 4. Transcription from the RNA Polymerase Binding Site near the Left End of Tn5

The transcription utilizing isolated restriction fragments as templates is described in Experimental Procedures. The RNA synthesized was electrophoresed on a 10% polyacrylamide/7 M urea gel and the resulting pattern was determined by autoradiography. Both the Eco RI-Hpa I fragment and the Eco RI-Hae II fragment give some small transcripts when used as a template. In the case of the *lac* UV5 transcription also shown here (using a Hae 203 lactose promoter fragment having the UV5 mutation), these small transcripts have proved to be abortive starts (Maquat and Reznikoff, 1978). This is probably also the case for the Tn5 transcription. The sizes of the transcripts produced when the Eco RI-Hae III fragment was used as a template are consistent with transcription initiating at the same site as with the Eco RI-Hpa I and Eco RI-Hae II fragments. Since the smaller transcript is not very different in size from the presumed "abortive" transcripts, however, we cannot rule out the possibility that it results from some other process. When the Eco RI-Hae II fragment is utilized as a template, a large transcript is also synthesized. This is presumably a result of having two "sticky" ends, possibly allowing concatemer formation.

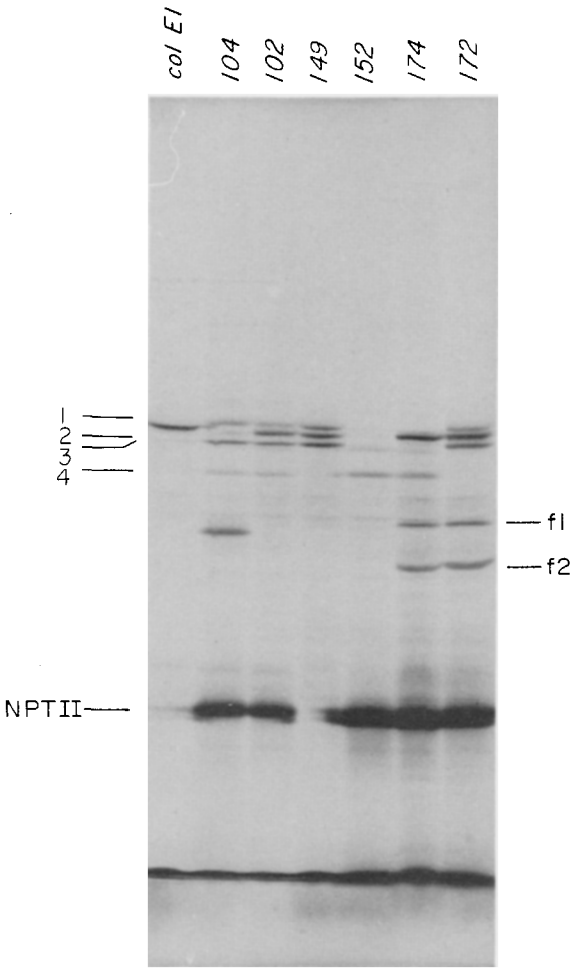


Figure 5. Localization of the Coding Regions of the Tn5 Polypeptides

Minicells containing Col E1, pRZ104, pRZ102, pRZ149, pRZ152, pRZ174 or pRZ172 were labeled with ^{35}S-methionine. The proteins were electrophoresed on a 12.5% SDS-polyacrylamide gel, and the resulting autoradiogram is shown. Polypeptide 3 was very faint in pRZ102- and pRZ104-containing strains. The fusion proteins synthesized in pRZ172 and pRZ174 are noted as f1 and f2. The colicin protein in Col E1 is abbreviated in the Col E1::Tn5 strains since the inversions are in the colicin gene. In pRZ102, the colicin protein electrophoreses between polypeptides 1 and 2, whereas in pRZ104 it is considerably smaller (~37K). It should be noted that polypeptide 3 is very faint and at times difficult to see.

Tn5 inserted into pVH51 DNA (along with some flanking λ sequences) and codes for the same five Tn5 polypeptides seen earlier for pRZ102 and pRZ104. Minicells containing pRZ123 do not make polypeptide 1 and 3 and have dramatically reduced synthesis of 2 and 4. Minicells containing pRZ124, which is a deletion of the end of the right inverted repeat, no longer make polypeptide 1 and have a decreased synthesis of 2, whereas pRZ129 with its deletion in the left repeat does not code for polypeptide 3 and has a decreased synthesis of polypeptide 4. In all cases, the NPTII protein is still produced.

Several conclusions can be drawn from these results. First, as was observed with the insertion strains above, different polypeptides are coded by the two inverted repeats. Second, the results are consistent with the presence of a promoter in each repeat between the Hpa I site and the end of the transposon. The weaker synthesis of polypeptides 2 and 4 in pRZ124 and pRZ129, respectively, would presumably be a result of read-through transcription from other promoters present on the plasmid. Finally, each inverted repeat does appear to encode two polypeptides.

The rest of the mutants delineated in Figure 2 were analyzed in the minicell system. The results are summarized in Table 2. If the two repeats synthesize similar polypeptides from the outside edge of the repeat up to the Hind III sites, then the Hind III inversion strain (pRZ143) should synthesize a set of fusion proteins having the same molecular weight as those made from pRZ102. The set of larger polypeptides, however, will now be encoded by the left inverted repeat. In addition, a deletion of this repeat (as in pRZ154) should not synthesize polypeptides 1 and 2, while a deletion of the wild-type left repeat would not synthesize polypeptides 3 and 4, as shown above.

The results of a minicell experiment using these strains as well as the Bgl II inversion pRZ141 confirm these predictions, as shown in Table 2. As predicted, pRZ143 codes for a set of polypeptides of the same molecular weight as does pRZ102. Minicells containing pRZ154, again as expected, fail to produce any polypeptides having the same molecular weight as 1 and 2. Finally, pRZ141 synthesizes all four inverted repeat polypeptides. NPTII production encoded by pRZ141 is reduced as expected, since this plasmid does not endow a strain with full neomycin resistance.

Another set of Tn5 mutants shown in Figure 2 involved the insertion of the lactose, 789 bp Hind II fragment into the Hpa I sites of pRZ112 and pRZ102. Their minicell polypeptide synthesis is also described in Table 2. pRZ164 has the central Hpa I fragment of Tn5 inverted. The *lac* promoter is inserted in the left repeat and transcribes away from the central region of Tn5. Since the differences in the two repeats with respect to protein synthesis probably occur near the central region, one would expect that an inversion at the Hpa I site would lead to a concomitant inversion of the proteins encoded by the two repeats. pRZ164 should therefore not make polypeptides 1 and 2 (or at least very little 2), and should synthesize polypeptides 3 and 4 even though its insertion is in the left inverted repeat. Furthermore, since *lac* transcription is not directed toward the central region, no fusion protein should be made. These expectations are confirmed by our results. pRZ165 is similar to pRZ164 in that the central region of Tn5 is inverted and the *lac* promoter transcribes away from the rest of the transposon. In this case, however, the inserted fragment is in the right inverted repeat. As expected, this insertion

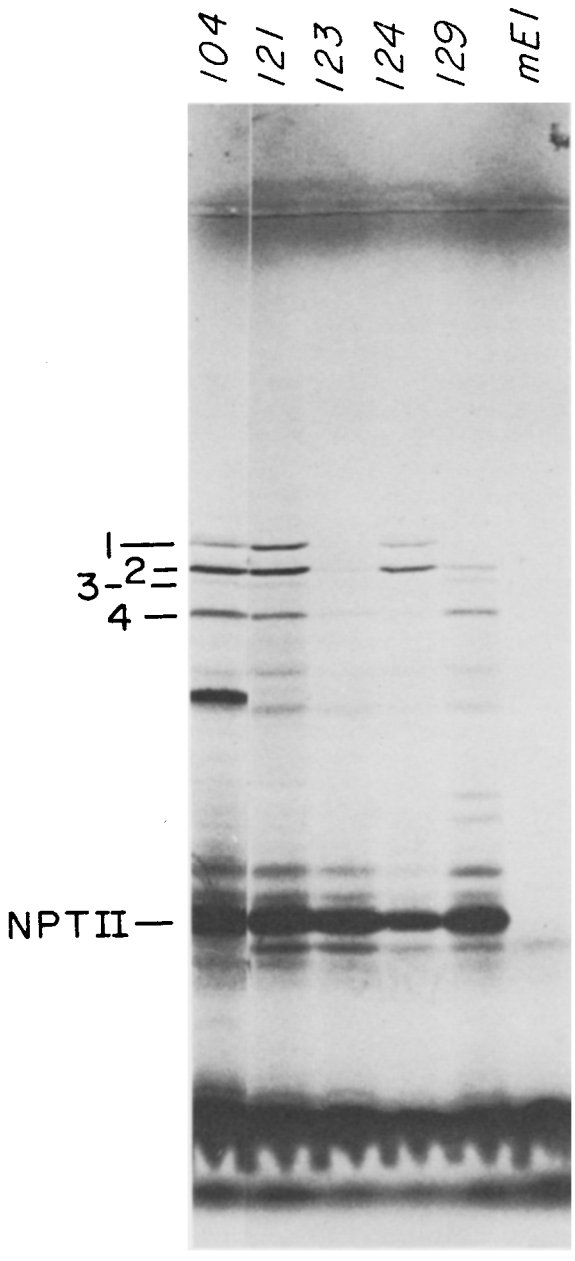


Figure 6. Proteins Encoded by Tn5 Deletion Strains

Minicells containing pRZ104, pRZ121, pRZ123, pRZ124, pRZ129 or pVH51 were labeled with ^{35}S-methionine. The proteins were electrophoresed on a 12.5% SDS-polyacrylamide gel, and the resulting autoradiogram is shown.

prevents synthesis of polypeptide 3 and synthesizes little if any of 4.

pRZ166 has the *lac* promoter inserted into the Hpa I site in the right repeat of pRZ102. The *lac* promoter transcribes toward the central region of the transposon. Polypeptide 1 is not synthesized, and polypeptide 2 synthesis is greatly reduced in this strain. However, a large fusion protein is made. This is presumably a fusion product of the N terminus of *lac*Z and one of the polypeptides made from that inverted repeat. (The first 145 amino acids of *lac*Z are coded for by the Hind II 789 bp fragment.)

The final *lac* insertion studied, pRZ161, has the *lac* promoter inserted in the Hpa I site of pRZ112. The promoter transcribes toward the central region of the transposon. Neither polypeptide 3 or 4 is synthesized, as expected, and a large fusion protein is produced. (The coding region of polypeptides 1 and 2 is eliminated in the deletion.) The fusion protein is smaller than that found in pRZ166. Since an identical piece of *lac*Z is fused to the polypeptides in these two strains, this result is consistent with the idea that the primary difference in the polypeptides coded for by the two inverted repeats occurs at their C terminal ends.

When the *lac* promoter is in the correct orientation for transcription to read through the inverted repeat, fusion proteins that apparently fuse β-galactosidase to an inverted repeat polypeptide are made. These fusion proteins have been shown to be under the control of the lactose promoter, since their production in minicells is stimulated by IPTG (our unpublished results).

Two polypeptides of different molecular weights appear to be encoded by each inverted repeat. It seemed possible that the explanation for having two polypeptides coded for by the same region might be trivial. A pulse-chase experiment was attempted to determine whether the larger polypeptide would be cleaved into the smaller. No noticeable difference could be observed in the polypeptide pattern. This does not eliminate processing as the source of the two polypeptides.

Transposition Tests on Tn*5* Mutants

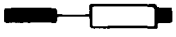
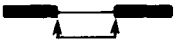
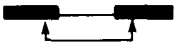
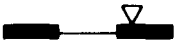
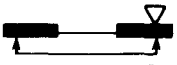
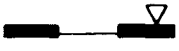
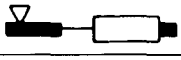
Since the insertion mutants in the inverted repeats of Tn*5* effect different polypeptides depending upon which repeat is mutated, it was of interest to determine whether transposition was also differentially affected.

To test the ability of the mutants to transpose, the plasmid DNAs were transformed into a W3110 nal$^R\lambda_{bb\,nin}$ lysogen and transposition onto $\lambda_{bb\,nin}$ was analyzed. Transducing particles were tested by lysogenizing into hfl$^-$, a strain which forms λ lysogens at high frequency. Kan^R cells were selected as described in Experimental Procedures, and the number of transductants was compared to the number of plaque-forming units.

The results are shown in Table 2. There is a large difference between the transposition frequencies of insertions into the left inverted repeat compared with those into the right repeat. The left repeat insertions transpose at a frequency that is only 4–5 fold less than wild-type Tn5. Right arm insertions transpose at a frequency at least two orders of magnitude less than either wild-type or left repeat insertions.

It thus appears that the two repeats have functional

Table 2. The Proteins Synthesized by the Tn5 Mutants and Their Levels of Transposition

Plasmid present	Structure	Tn5 proteins synthesized					Fusion proteins synthesized	% Transposition of w.t.
		1	2	3	4	NPTII		
102		+	+	+	+	+		100%
112		−	−	+	+	+		<0.5%
141		+	+	+	+	<		n.d.
143		+	+	+	+	+		100%
154		−	−	+	+	+		n.d.
172		+	+	−	−	+	38K + 34K	27%
174		−	−	+	+	+	38K + 34K	<0.5%
164		−	<	+	+	+		<0.5%
165		+	+	−	−	+		21%
166		−	<	+	+	+	~79K	<0.5%
161		−	−	−	−	+	~73K	n.d.

Some of the mutants synthesize fusion proteins of different molecular weight than the wild-type Tn5 polypeptide. This was shown above for pRZ172 and pRZ174. pRZ161 and pRZ166 each fuse the first 145 amino acids of β-galactosidase to one of the inverted repeat polypeptides. Since the fusion in pRZ161 is between β-galactosidase and the left repeat polypeptide, the fusion protein is somewhat smaller than that coded for by pRZ166, where the fusion is to the right repeat polypeptide. The amount of transposition onto $\lambda_{bb\,nin}$ of some of the Tn5 inversion and insertion strains is also shown. The values given are the number of kanamycin-transducing particles per plaque-forming unit and are an average of three different experiments. The various types of mutants are indicated by (□) for a deletion, (▽) for an insertion, (⊥↑) for an inversion.

differences with respect to transposition. An insertion at identical locations in the two repeats has dramatically different effects on the observed level of transposition.

The final point that can be made about the transposition test concerns pRZ143, the strain containing an inversion at the Hind III site. It transposes at a frequency close to that of wild-type. Assuming that the larger polypeptides synthesized (now off the left repeat) are still of primary importance for transposition, it must be assumed that fusion of the left repeat proteins up to the Hind III site, to the right repeat proteins beyond that site, still allows for complete pRZ143 function. This emphasizes further that the main difference between the two repeats lies between the Hind III sites and the central region of the transposon.

Discussion

The inverted repeats of Tn5 are clearly different. They have different RNA polymerase binding patterns; only one of them promotes full neomycin resistance; they code for polypeptides of different molecular weights; and they function differently in the transposition process. This was somewhat surprising at first, given their identical restriction endonuclease cleavage maps, which ruled out any gross deletions or insertions. Further analysis of the difference between the two repeats revealed that the meaningful functional dissimilarities appear to occur only near the central region of transposon.

A model consistent with the evidence described here for the transcription and translation of Tn5 is shown in Figure 7. Five polypeptides have been found to be coded for by Tn5, two being translated from each inverted repeat. The fifth is the neomycin phosphotransferase II protein. The details of this model are discussed below. The inverted repeats between the outside edge and the Hind III sites appear to be functionally identical. The difference in size of the two polypeptides synthesized off either one of the repeats does occur in this region, however, as indicated by the fact that the Hind III insertion strains still synthesize

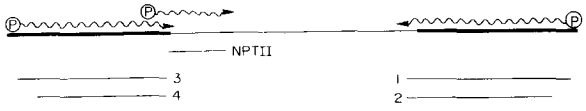


Figure 7. Model of Tn5 Transcription and Translation

Proteins 1 and 2 differ in length primarily at the N terminal end of the polypeptides, as do proteins 3 and 4. The right inverted repeat polypeptides differ in size from the left inverted repeat polypeptides at the C terminal end. If this model is correct, the coding region of polypeptides 3 and 4 overlaps the control region of the neomycin resistance protein.

two polypeptides from each repeat. The two fusion proteins have a difference in molecular weight close to that between proteins 1 and 2 that between proteins 3 and 4 from the right and left inverted repeats, respectively.

The region between the Hind III and Bgl II sites differs in the two inverted repeats with respect to RNA polymerase binding, as well as the ability to promote full drug resistance when adjacent to the neomycin resistance gene. An inversion of the central region of Tn5 at the Bgl II sites lowers the level of neomycin resistance drastically. The right inverted repeat must lack some DNA sequences present in the left repeat necessary for high levels of resistance. As mentioned above, an inversion at the Hind III sites does not alter neomycin resistance, indicating that this difference between the two repeats must lie between the Hind III and Bgl II sites. It is still not known whether there are differences in the polypeptides encoded by this region in the two repeats.

The polypeptides encoded by the right inverted repeat are approximately 6000 daltons larger than those encoded by the left repeat. This size difference almost certainly must be localized primarily at the C terminal ends of the polypeptides, near the unique region of the transposon. It seems probable that the larger polypeptides of the right repeat are actually encoded in part by the central region of Tn5 (our unpublished results).

The detailed mechanism of transposition has not yet been elucidated. It seems probable, however, that in general a protein capable of recognizing and acting upon the ends of the transposon is required. In Tn5, we have found that insertions in the right repeat have a far greater effect on the level of transposition than insertions in the left repeat. While it is theoretically possible that some sort of structural feature present only in the right repeat is affected by these insertions, it seems far more plausible that the important factor is their effect on the polypeptides synthesized by this repeat. Hence it would not be totally surprising if one (or both) of the polypeptides coded for by the right inverted repeat turned out to act as some sort of transposase enzyme. It has been reported that there is a trans-acting factor coded for by Tn5 that is necessary for transposition. It was also suggested that the probable coding region for this factor would be in the right inverted repeat of Tn5, which correlates well with the results presented here (Meyer, Boch and Shapiro, 1979). The function of the left inverted repeat-encoded polypeptides is still not known. Insertions in this repeat do lower the level of transposition 4–5 fold compared with wild-type.

The proposed location of three Tn5 promoters and their direction of transcription are also shown in Figure 7. All three are found in regions which bind RNA polymerase. The binding site near the end of the left inverted repeat has been found to initiate transcription in vitro, and its transcription is toward the central region of the transposon, as expected. The binding site near the end of the right inverted repeat also initiates transcription in vitro, but the direction of transcription has not been determined.

The fusion proteins made in the insertion strains strongly support the notion that the inverted repeats are transcribed starting from the ends of the transposon. The Hind III insertion strains synthesize fusion proteins of the predicted size. It is hard to imagine how proteins of this size could be made if transcription were in the opposite direction. Most importantly, insertion of the lactose promoter-operator fragment leads to production of a larger fusion protein only when lactose transcription is directed toward the central region of the transposon. Induction of the lactose promoter by IPTG demonstrates that the synthesis of this fusion protein is now under *lac*I gene control, and its transcription and translation must be from near the outside edge of the transposon toward the unique region.

The region between the Hind III and Bgl II sites has an important role for the synthesis of NPTII since an insertion or inversion at the Bgl II site affects the level of resistance. A protein of molecular weight similar to that of NPTII is synthesized, although in smaller amounts in both the insertion or inversion strains (see above; our unpublished results). It therefore seems that the region between the Bgl II and Hind III sites is likely to contain the promoter for the NPTII gene. If this is the case, then the promoter for NPTII overlaps with the structural gene for the left inverted repeat polypeptides.

Assuming that the inverted repeats of Tn5 were derived originally from a pair of homologous insertion sequences and that the gene organization model presented in Figure 7 is correct, then the current structure of the transposon may tell us something about its evolution. The left inverted repeat would have to have undergone at least two changes. The first would presumably have involved mutation(s) creating a promoter-like activity between the Hind III and Bgl II sites. A subsequent deletion might have removed the DNA between this promoter and the structural gene for NPTII, placing neomycin resistance under the control of the "new" promoter and in the process deleting part of the original IS sequence as well as the original

NPTII promoter. This series of events would firmly link the "left" inverted repeat and the resistance functions of Tn*5*.

Experimental Procedures

The media used for growing phage, isolating plasmids and minicell growth and labeling are the same as those described by Jorgensen and Reznikoff (1979) and Postle and Reznikoff (1979).

RNA polymerase was purified as described by Burgess and Jendrisak (1975). The following enzymes were purchased from New England Biolabs: Bam HI, Bgl II, Pst I, Sal I, Hpa I, Hind II, Sma I and T4 DNA ligase. Hind III and Eco RI were gifts from J. Gardner. M. luteus DNA polymerase was a gift from E. Selsing and R. D. Wells. α-^{32}P-UTP was purchased from Amersham-Searle (250 Ci/mmole), as was ^{35}S-methionine (600–800 Ci/mmole).

Bacterial Strains

λ:Tn*5* (Berg et al., 1975) or pRZ102 (Jorgensen and Reznikoff, 1979) were used as the source of Tn*5* for all these studies. C600 SF8 (Struhl, Cameron and Davis, 1976) was used as the original recipient in transformations. All plasmids were grown in either C600 SF8 or MO (F$^-$ *str*A$^-$ isogenic for HFR Hayes). χ984 (F$^-$*min*A1, *tsx*63, *pur*E41, *sup*E41, *sup*E42, *pdx*C3 *min*B2, *his*53, *met*C65, *rps*L97, T3^R, *xyl*-14, *ilv*277, *cyc*B2, *cyc*A1, from R. Curtiss) and DS410 (*min*A, *min*B, *ara*, *lac*Y, *str*^{R}, *mal*A, *xyl*, *azi*^{R}, *ton*^{A}, λ$^-$, *mtl*, *thi*, from D. Sherratt) were the minicell-producing strains used. W3110 (F$^-$ *trp*A 33) lysogenic for $\lambda_{bb\,nin}$ was transformed with the various plasmids containing Tn*5* mutants. These strains were then used for growing phage for the transposition tests. hfl-1 (Belfort and Wulff, 1971) was used as a recipient for λ:Tn*5* transducing particles.

Selection of Transformants after Ligation

The basic ligation, transformation and selection procedures have been described elsewhere (for example see Jorgensen et al., 1979).

The selected transformants were tested for plasmid size and presence by a modification of the procedure described by Barnes (1977). When no size difference was expected or to test for the regeneration of restriction cleavage sites, small crude preparations of plasmid DNA capable of being cleaved by restriction enzymes were prepared using the technique of Cameron, Phillippsen and Davis (1977).

Construction of Mutants of Tn*5*

pRZ121, pRZ123, pRZ124 and pRZ129 all resulted from ligating a partial Hpa I digest of λ:Tn*5* into Hind II-cleaved pVH51. Neomycin-resistant colonies were selected after transformation and the plasmids present in the transformants were screened for plasmid size. Several different clones containing plasmids of different molecular weights were selected for further analysis. Restriction analysis of purified plasmid DNAs with Hpa I and Hind II was used to demonstrate which Tn*5* and which λ fragments were still present in the various transformants.

The inversions of Tn5 were constructed by digesting pRZ102 DNA with either Hind III or Bgl II and ligating the two resulting fragments together. Colicin-resistant colonies were selected and the resulting transformants were screened for neomycin resistance and plasmid size. The DNA cleaved with Bgl II gave several transformants that were considerably less neomycin-resistant than wild-type pRZ102 and these, along with several others, were selected for further analysis. One of the less resistant colonies contained a plasmid the same size as pRZ102, while the others were larger. None of the Hind III-cleaved DNA gave transformants less resistant than wild-type Tn*5*.

Small crude preparations of DNA were prepared from a number of the transformants as described above. The Bgl II inversion candidate that was the same size as pRZ102 and less resistant, as well as several of the Hind III transformants, had an Eco RI-Sal I fragment larger than that of pRZ102. This is indicative of an inversion. Purified DNA was then isolated and its structure was tested further by repeating the Eco RI-Sal I digest, as well as by digesting with several other restriction enzymes including an Eco RI + Bam HI double digest (which gives a smaller Tn*5* fragment than in pRZ102) and Hae III (which gives a pattern identical to that of the wild-type). It should be noted that the Hind III inversion with its identical phenotype and internally identical restriction pattern cannot be differentiated in any simple way from wild-type Tn*5* if it transposes onto a different vector.

The substitution of the lactose Hae III 203 promoter-operator fragment containing the UV5 mutation for the Eco RI-Sal I fragment of pRZ102 and pRZ145 (the Hind III inversion) was achieved by cleaving the plasmid DNA with Eco RI and Sal I, filling in the "sticky ends" with M. luteus DNA polymerase (Backman, Ptashne and Gilbert, 1976) and ligating in the Hae 203 fragment (a gift from L. Maquat). The transformed cells were selected for colicin resistance and screened on XG plates for constitutive β-galactosidase production. The size of the plasmids present in the transformants was analyzed and minipreparations of plasmid DNA were cleaved with Eco RI + Hind III and Eco RI + Sal I to ensure that the Eco RI and Sal I cleavage sites were restored.

Insertion of the lactose Hind II 789 wild-type promoter-operator fragments into pRZ102 and pRZ112 was performed by cleaving the vector with Hpa I, adding purified Hind II 789 fragment (a gift from S. Hardies) and incubating the DNA with T4 DNA ligase. Blue Neo^r colonies on XG-neo plates were selected and plasmid DNA was isolated. pRZ112 Hpa I inserts were tested for regeneration of the Hpa I site by Hpa I digests and the orientation of the lactose promoter by a Hae II digest.

The insertion of the Hind II 789 fragment into pRZ102 was tested for site of insertion by Pst I and Pst I + Eco RI digests; orientation of the central Hpa I fragment of Tn*5* was tested by comparing Eco RI-Sal I digests with Eco RI-Bam HI digests and the orientation of the lactose promoter was tested by Hae II digests. Only those plasmids in which the Hpa I sites were regenerated were kept for further analyses.

The insertion of the 515 bp Hind III fragment from the central region of Tn*10* into Tn*5* was accomplished by first cloning it into pRZ112 and pRZ141, which each contain only one of the inverted repeats of Tn*5* and consequently have a single Hind III site. A Hind III digestion of pRT3 DNA, which contains most of Tn*10* (Jorgensen and Reznikoff, 1979), gives three fragments, a large one including all of Col E1, a 2500 bp fragment which contains the tetracycline resistance gene and the 515 bp fragment wanted. This mixture was ligated into Hind III-cleaved pRZ112 or pRZ141 with neomycin or colicin resistance selected, respectively. The transformants were screened for tetracycline resistance and plasmid size, and tet^S plasmids larger than the original vector were isolated. The presence of the Hind III fragment was verified by digesting the resulting plasmid with Hind III.

To test the Hind III insertions for transposition, the missing half of Tn*5* had to be regenerated. The Hind III insertion into pRZ141 (which contains only the right inverted repeat) was cleaved with Sal I + Eco RI and ligated to Eco RI + Sal I-cut pRZ102, and neomycin-resistant transformants were selected. Plasmids larger than pRZ102 were selected and tested for the presence of the Hind III insertion.

pRZ112 (which retains the entire left inverted repeat) containing the Hind III insertion was cleaved with Eco RI + Bgl II (it does not contain a Sal I site) and ligated to an Eco RI + Bgl II digest of pRZ102. Since it was known that inversion of the Bgl II fragment in Tn*5* leads to lower neomycin resistance, transformants were selected at high levels of neomycin to eliminate this possibility. A plasmid larger than pRZ102 was screened for further analysis and found to have the desired structure.

Nitrocellulose Filter Binding Assay

The RNA polymerase binding experiments described here were performed similarly to those described by Jones et al. (1977). 4 μg of restricted plasmid DNA were mixed with 6 μg of RNA polymerase in 10 mM Tris–HCl (pH 7.9), 100 mM KCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM DTT and 40 μg BSA in a 100 μl reaction volume. After a 30 min incubation at 37°C, 100 μg of heparin were added to remove nonspecifically bound RNA polymerase and the mixture was incubated at 37°C for 3 min.

Stable complexes were then isolated by running the mixture through nitrocellulose filters and removing bound complexes from the filter by mixing with a solution of 10 mM Tris and 0.1% SDS. After ethanol precipitation, the DNA fragments that were bound by the RNA polymerase were electrophoresed on a 5% polyacrylamide gel and compared to an aliquot of the restriction digest.

In Vitro Transcription Assay

In vitro transcription was carried out as described by Maquat and Reznikoff (1978). Restriction fragments isolated from 5% polyacrylamide were incubated with purified RNA polymerase for 30 min in 30 mM Tris–HCl (pH 7.9), 0.1 mM EDTA, 10 mM $MgCl_2$, 100 mM KCl, 0.1 mM dithiothreitol and 0.5 mg bovine serum albumin per ml. 2 μg of heparin were then added. Finally, transcription was initiated by adding nucleotide triphosphates to a final concentration of 200 mM ATP and GTP, 10 mM CTP and 10 mM α-^{32}P-UTP (Amersham-Searle, 250 Ci/mmole). Synthesis was terminated after 10 min by the addition of cold transcription buffer and extracted with phenol. After removing the aqueous layer, 100 μg of carrier tRNA and Na acetate to 0.3 M were added, and the RNA was precipitated with 2.5 vol of ethanol. The precipitate was resuspended and electrophoresed on a 10% polyacrylamide/7 M urea gel.

Measurement of Neomycin Resistance Levels

The level of neomycin resistance conferred by the Tn*5* plasmids in C600 SF8 was measured by a modification of the procedure of Tait, Rodriguez and Boyer (1977). An identical number of cells (between 100 and 400) was plated on nutrient plates containing 0, 2, 5, 10, 20, 30, 40, 55, 70, 85 and 100 mg/ml neomycin. The number of surviving colonies was counted and the concentration of neomycin necessary to kill 50% of the cells was determined graphically, this value being the EOP_{50} of the strain.

Minicells

Minicells were purified as described by Roozen, Fenwick and Curtiss (1971). Minicells obtained from a 350 ml culture were washed and resuspended in 2 ml of ¼ Met assay medium (Difco) in M9 salts, 0.5% glucose, 2 μg pyridoxine per ml, 40 μg adenine per ml. They were then labeled with 50 mCi ^{35}S-methionine (spec. act. 600–800 Ci/mmole; Amersham-Searle) for 30 min, pelleted and frozen. The pellets were resuspended in 100 ml sample buffer (Laemmli, 1970) and heated to 90°C for 2–5 min. 20 μl were then electrophoresed on a 12.5% SDS-polyacrylamide gel (Laemmli, 1970).

For inducing the lactose promoter with IPTG, the minicells were divided into two aliquots; one was incubated with 1 mM IPTG during labeling while the other was not.

The pulse-chase experiments involved two types of procedures. The first was simply to add an excess of cold methionine (to a final concentration of 200 g/ml) and incubate for an additional 30 min before pelleting the minicells. Alternatively, the minicells were labeled for 30 min, collected on a Millipore filter, washed with methionine assay medium and resuspended in methionine assay medium supplemented with 100 μg/ml methionine. The minicells were then incubated for an additional 30 min prior to pelleting.

Molecular weight standards were electrophoresed in an adjacent lane to the minicell proteins. The gel was stained with Coomassie blue and destained, and the mobility of the standards was compared to that of the labeled protein on the autoradiograph.

Measurements of the Level of Transposition

The various Tn*5*-containing plasmids to be tested for transposition were transformed into $\lambda_{bb\,nin}$ lysogens. They were then induced in 3 ml cultures and the resulting phage were titered. The phage from each Tn*5*-containing strain were analyzed for their ability to transduce kanamycin resistance by incubating an approximately equal number of phage with an aliquot of hfl-1 and plating on nutrient plates containing 30 μg/ml kanamycin. The resulting colonies were then tested on 100 μg/ml streptomycin plates to ensure that they were not surviving W3110nal^R cells.

Acknowledgments

We are grateful for support from the NIH to W.S.R. W.S.R. was also supported in part by a Career Development Award from the NIH and was a recipient of the Harry and Evelyn Steenbock Career Development Award.

The authors would like to thank J. Davies, J. Yin, R. Johnson, L. Wray, J. Shapiro and D. Berg for helpful discussions. Special thanks are also due to S. Kagan for helping with the manuscript.

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

Received October 4, 1979; revised December 19, 1979

References

Backman, K., Ptashne, M. and Gilbert, W. (1976). Proc. Nat. Acad. Sci. USA *73*, 4174–4178.

Barnes, W. M. (1977). Science *195*, 393–394.

Belfort, M. and Wulff, S. (1971). In The Bacteriophage Lambda, A. D. Hershey, ed. (New York: Cold Spring Harbor Laboratory), pp. 739–742.

Berg, D. E., Davies, J., Allet, B. and Rochaix, J. (1975). Proc. Nat. Acad. Sci. USA *72*, 3628–3632.

Burgess, R. R. and Jendrisak, J. J. (1975). Biochemistry *14*, 4634–4638.

Cameron, J. R., Phillippsen, P. and Davis, R. W. (1977). Nucl. Acids Res. *4*, 1429–1448.

Cornelis, G., Ghosal, D. and Saedler, H. (1978). Mol. Gen. Genet. *160*, 215–224.

Heffron, F., Bedinger, P., Champoux, J. J. and Falkow, S. (1977). Proc. Nat. Acad. Sci. USA *74*, 6012–6016.

Jones, B. B., Chan, H., Rothstein, S., Wells, R. D. and Reznikoff, W. S. (1977). Proc. Nat. Acad. Sci. USA *74*, 4914–4918.

Jorgensen, R. A. and Reznikoff, W. S. (1979). J. Bacteriol. *138*, 705–714.

Jorgensen, R. A., Rothstein, S. and Reznikoff, W. S. (1979). Mol. Gen. Genet. *177*, 65–72.

Kleckner, N. (1977). Cell *11*, 11–23.

Laemmli, U. (1970). Nature *227*, 680–685.

MacHattie, L. and Jackowski, J. B. (1977). In DNA Insertion Elements, Plasmids, and Episomes, A. I. Bukhari, J. A. Shapiro and A. L. Adhya, eds. (New York: Cold Spring Harbor Laboratory), pp. 219–228.

Maquat, L. and Reznikoff, W. S. (1978). J. Mol. Biol. *125*, 467–490.

Matsuhashi, Y., Sawa, T., Takeuchi, T. and Umezawa, H. (1976). J. Antibiotics *29*, 204–207.

Meyer, R., Boch, G. and Shapiro, J. (1979). Mol. Gen. Genet. *171*, 7–13.

Postle, K. and Reznikoff, W. S. (1979). J. Mol. Biol. *131*, 619–636.

Roozen, D., Fenwick, R. and Curtiss, R. (1971). J. Bacteriol. *107*, 21–33.

So, M., Heffron, F. and McCarthy, B. J. (1979). Nature *277*, 453–456.

Struhl, K., Cameron, J. R. and Davis, R. D. (1976). Proc. Nat. Acad. Sci. USA *73*, 1471–1475.

Tait, R., Rodriquez, R. and Boyer, H. (1977). Mol. Gen. Genet. *151*, 327–331.