

Tn5 *in Vitro* Transposition*

(Received for publication, September 29, 1997, and in revised form, December 21, 1997)

Igor Yu Goryshin and William S. Reznikoff‡

From the Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706

This communication reports the development of an efficient *in vitro* transposition system for Tn5. A key component of this system was the use of hyperactive mutant transposase. The inactivity of wild type transposase is likely to be related to the low frequency of *in vivo* transposition. The *in vitro* experiments demonstrate the following: the only required macromolecules for most of the steps in Tn5 transposition are the transposase, the specific 19-bp Tn5 end sequences, and target DNA; transposase may not be able to self-dissociate from product DNAs; Tn5 transposes by a conservative “cut and paste” mechanism; and Tn5 release from the donor backbone involves precise cleavage of both 3′ and 5′ strands at the ends of the specific end sequences.

Transposition is a process that gives rise to a variety of genomic rearrangements such as insertions, deletions, inversions, and chromosome fusions (1). Because transposition can play a profound role in genome evolution and in a variety of genetic diseases, how this process occurs and how its frequency is determined is of considerable interest. Bacterial transposons, such as Tn5, are convenient tools for studying these two questions.

The transposition process can be understood in the context of the following model (2, 3). The first step in transposition involves the sequence-specific binding of the transposon-encoded transposase (Tnp)¹ to the specific end sequences that define the ends of the transposable element. The second step involves formation of a synaptic complex in which the ends of the transposable element are brought together through Tnp oligomerization. In some well studied cases such as Tn10 (4, 5), synaptic complex formation is facilitated by one or more host proteins. The third step involves a Tnp-mediated nucleolytic attack on the phosphodiester bonds adjacent to the ends of the transposable element. The precise nature of these cleavages (whether just the 3′ ends of the transposon are liberated or both the 3′ and the 5′ ends and the exact location of the cleavage in the case of 5′ end release) is a special property of the system being studied. The fourth step involves Tnp binding to the target DNA sequence (target capture). Different transposons demonstrate various degrees of sequence specificity at this step. The fifth step involves a concerted nucleophilic attack by the transposon 3′-OH ends on phosphodiester bonds in both strands of

the target sequence leading to strand exchanges in which the transposon 3′-OH ends are covalently linked to the target 5′-PO₄ groups. The sixth step involves a release of Tnp from the complex (6). Finally, the transposition process is completed by a patch repair or a replicative resolution process.

It is clear that transposition is a complex multistep process and that Tnp participates in an intimate fashion in almost all of the steps. This means that Tnp is a multifunctional protein and that the activities of Tnp are likely to be manifested in a sequential fashion controlled by a variety of allosteric phenomena. Understanding the molecular details of the various steps is an important goal of current research efforts. To that end, a variety of systems have been studied *in vitro*. Some bacterial transposable elements (such as Mu, Tn7, and Tn10) have been particularly useful in this effort although, with the cases of Mu and Tn7 in particular, the systems are quite complex in regard to the protein and DNA participants (2, 7, 8). A simpler system involving essentially only one protein (Tnp) for most of the steps and simple DNA sequences might aid in the effort to relate structure to function. As we shall demonstrate in this communication, Tn5 is such a simple system.

Another critical feature of transposition is the frequency with which it occurs. For some transposable elements, a high frequency is allowable or desirable. Examples include the site-specific transposition process for Tn7 (9) and the lytic/replicative transposition process for Mu (2). However, many transposons have evolved mechanisms that facilitate the establishment of long term stable associations with the host genome. In these cases a very low frequency of transposition would be necessary. How can a low transposition frequency be ensured? In the case of Tn5, this low frequency appears to be the result of multiple features, two of which we have exploited in the development of an *in vitro* system. First, the low frequency of Tn5 transposition results from Tnp having suboptimal properties; it is a protein with a low level of activity. At least three classes of Tnp mutants have been isolated that enhance Tnp activity relative to different steps in the transposition process (10–12). Second, the activity of Tnp is further down-regulated by a trans-dominant negative variant of Tnp, the inhibitor (Inh) (13, 14). Inh synthesis can be blocked without impairing Tnp function by destroying the Inh initiator codon (10).

Tn5 is a composite transposon in which three genes encoding antibiotic resistance are flanked by two IS50 elements (see Fig. 1A for a schematic presentation). Tn5 has been described in detail in two reviews (3, 15). For the purpose of understanding the Tn5 transposition process, it is important to know the following. One of the IS50 elements (IS50R) encodes the cis-active 476-amino acid Tnp and the trans-dominant negative Inh. The inactive C-terminal truncated versions of Tnp and Inh encoded by IS50L and the antibiotic resistance functions in the central region play no role in the transposition process. The second critical component in transposition are the inverted 19-bp sequences that define the ends of Tn5, the so-called outside end (OE) sequences (Fig. 1C). There is a second type of

* This work was supported by National Institutes of Health Grant GM50692. 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.

‡ To whom correspondence should be addressed: Dept. of Biochemistry, University of Wisconsin, 420 Henry Mall, Madison, WI 53706.

¹ The abbreviations used are: Tnp, transposase; OE, outside end; IE, inside end; bp, base pair(s); Inh, inhibitor; CHAPS, 3-[(3-cholamidopropyl)dimethyl-ammonio]-1-propanesulfonate; PCR, polymerase chain reaction; DBB, donor backbone; Inter, intermolecular transposition product; ETF, excised transposon fragment.

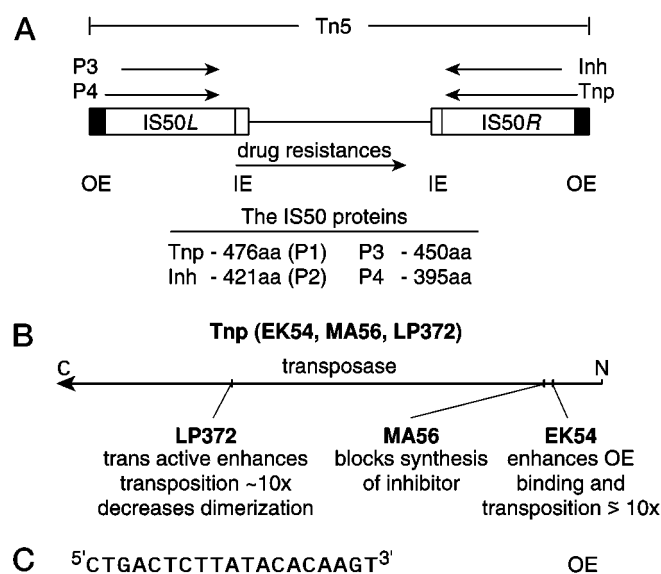


FIG. 1. Tn5. A, the features critical for Tn5 transposition include the transposase (*Tnp*), the inhibitor of transposition (*Inh*), and the 19-bp end sequences (the outside end, *OE*, and the inside end, *IE*). The P3 and P4 proteins are nonfunctional versions of *Tnp* and *Inh*. A more detailed description of Tn5 can be found in Refs. 3 and 15. B, hyperactive Tnp mutations used in this study. C, 19 bp outside end sequence.

19-bp sequence that can also be recognized by *Tnp*, the inside end (*IE*) sequence. *IE* is involved in *IS50* transposition (an *OE-IE* event) but not Tn5 transposition (an *OE-OE* event). Thus Tn5 transposition involves *Tnp* and the two appropriately spaced and oriented 19-bp *OE* sequences.

In this communication we report the development and characterization of an *in vitro* transposition system for Tn5 using purified *Tnp*. We used hyperactive *Tnp* that contained three mutations (Fig. 1B) to accomplish this goal. EK54 enhances *OE* binding activity of *Tnp* (12). MA56 blocks the synthesis of *Inh* (10), thus removing its possible inhibitory activity from the reaction. LP372 enhances *Tnp* activity possibly by altering the dimerization potential for *Tnp* (11). The results of these studies show that most of the steps in the Tn5 transposition process involve three macromolecular components as follows: the 476-amino acid long Tn5 *Tnp*, the 19-bp *OE* sequences that define the ends of Tn5, and a target DNA sequence. In addition we demonstrate, among other features, that Tn5 transposition occurs by a conservative cut and paste mechanism in which the Tn5-donor backbone cleavage reaction involves a precise cleavage at the end of the *OE*.

It is quite likely that many other wild type transposase enzymes manifest a similar low level of activity. Thus the strategy that we have employed in developing the Tn5 *in vitro* transposition system should have general applicability. Finally, the high efficiency and simplicity of the Tn5 *in vitro* transposition plus the existence of many Tn5-related constructs (16) should make this system useful for many investigators.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Plasmids, and Reagents—*Escherichia coli* strains for overexpression of hyperactive transposase were BL21(DE3)pLysS (17) and BL21 *OmpT* (18). T7 RNA polymerase was encoded by λ DE3 phage (17) or by a plasmid, pT7pol26 (which provides tighter control of polymerase expression) (19). For all DNA manipulations and electroporation of reaction products, *E. coli* DH5 α (20) was used.

Plasmid pRZPET2, for overexpression of hyperactive mutant *Tnp*, was constructed as follows. The LP372 mutation from plasmid pRZ4870-LP372 (11) was introduced into pRZ5412-EK54 (12), by substituting the appropriate *NheI-BglIII* fragment within the *Tnp* gene. Then, the *BspHI-SalI* fragment of the resulting plasmid, carrying all three *Tnp* mutations (EK54, MA56, and LP372) in the transposase

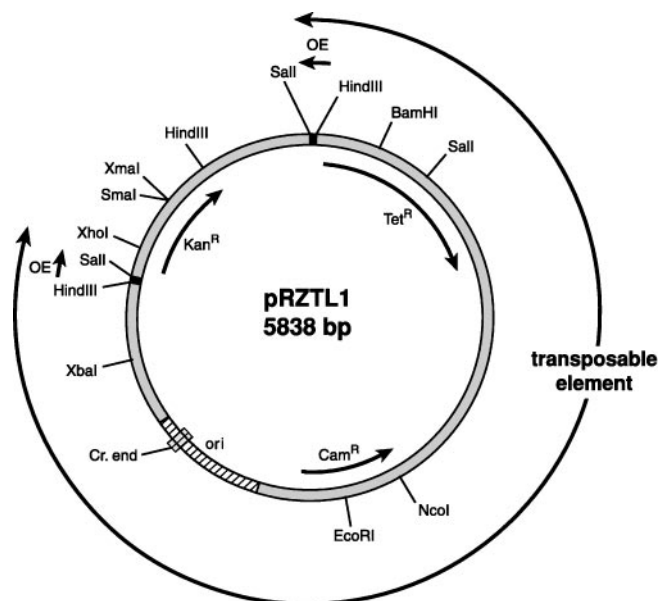


FIG. 2. Plasmid pRZTL1. pRZTL1 was used both as a substrate and a target DNA to study Tn5 transposition *in vitro*. Two black boxes are the *OE* sequences of Tn5 which define the ends of the transposable element. The shorter DNA sequence encoding kanamycin resistance is the donor backbone. Along with a p15 origin of replication and *cam*^R (chloramphenicol resistance) gene, the transposon contains a promoterless *tet*^R gene next to one *OE*. Tetracycline resistance can be activated after transposition into a unit of transcription in the proper orientation. A *lacZ'* fragment (also lacking its own promoter) is located next to the other *OE*, but this feature was not used in this study. The *HindIII* and *SalI* sites on both sides of the *OEs* were used to assemble the plasmid ("Experimental Procedures"). *NcoI* or *XhoI* sites were used to linearize the plasmid (Fig. 3). The cryptic end (*Cr.end*), recognized by *Tnp* under certain conditions (see text), is also shown. Its sequence is TTTTCA-GAGCAAGAGATTA with the last A as +1 of the end sequence. Underlined letters match the *IE* sequence (the match with *OE* is three letters weaker). Note, that position +1 is different from +1 of *OE* or *IE* but is still used for cleavage under relaxed conditions.

gene, was ligated with the large *NcoI-XhoI* fragment from pET21d (Novagen).

Plasmid pRZTL1 is represented in Fig. 2. pRZTL1 was constructed as follows. The appropriate *BamHI-SphI* fragment from pFMA50-1294 (21) was made blunt-ended with Mung bean nuclease and then was cloned into the *BsaAI* site of pTZ18U next to *lacZ'*. The *lacZ' XbaI-HindIII* fragment was then ligated to the large *XbaI-HindIII* fragment of pACYC184, creating pRZTL0. At this point both the *tet*^R gene and *lacZ'* have lost their transcription initiation signals. Finally the *kan*^R gene of Tn903 (on a *SalI-SalI* fragment) was cloned into the *HindIII* site with the assistance of a *SalI-HindIII* linker, encoding *OE*, in a four-fragment ligation reaction. The *SalI-OE-HindIII* linker was formed by annealing oligonucleotides S26 (5' TCGACTGACTCTTATACACAAGTA 3') and S27 (5' AGCTTACTTGTGTATAAGAGTCAG 3').

Plasmid pFMA1870 Δ used for the ligation-mediated PCR experiment is similar to pFMA187 (21) but with the *IE* of Tn5 substituted by *OE* (12) and the entire transposase gene deleted by cutting with *BspHI*, followed by ligation.

Bacteria were cultured in Luria broth (LB). Antibiotics were purchased from Sigma, and the final concentrations were 100 μ g/ml ampicillin, 15 μ g/ml tetracycline, 20 μ g/ml chloramphenicol, and 40 μ g/ml kanamycin. Isopropyl-1-thio- β -D-galactopyranoside was purchased from Diagnostic Chemical. Restriction enzymes were purchased from New England Biolabs and Promega and were used following the manufacturers' recommendations. Radionucleotides were purchased from Amersham Pharmacia Biotech. Oligonucleotides were purchased from Research Genetics.

DNA Preparation—pRZTL1 plasmid DNA for the experiment described in Fig. 3 was prepared as follows. After standard alkaline lysis purification (20), the DNA was run on a preparative 1% agarose gel, and supercoiled monomer DNA was cut out of the gel and was prepared with a GeneClean II kit (BIO 101). This protocol avoided the presence of multimeric forms of the plasmid and chromosomal DNA contaminants that would complicate the interpretation of the agarose gel patterns.

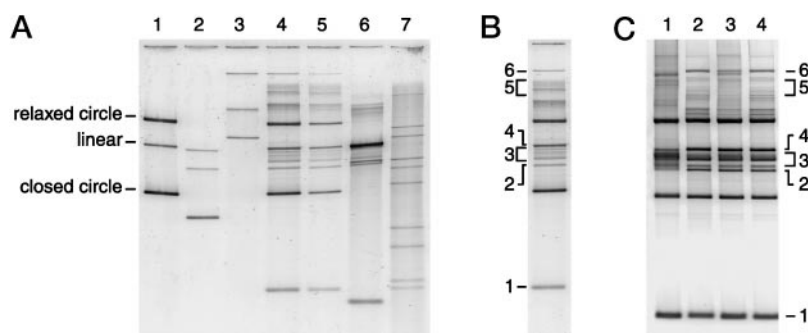


FIG. 3. Electrophoretic analysis of *in vitro* transposition products. A, data obtained using both circular and linear pRZTL1 substrates are shown. The bands were revealed on a 1% agarose gel following electrophoresis by staining with SYBR Green II (Molecular Probes) and scanned on a Fluorimager S1 (Molecular Dynamics). Lane 1 shows relaxed circular, linear, and supercoiled versions of pRZTL1. Lanes 2 and 3 show intramolecular and intermolecular transposition products recovered after electroporation. Lanes 4 and 5 present direct analyses following phenol extraction of two independent *in vitro* reactions using a mixture of closed and relaxed circular test plasmid substrates. In lane 6, linear pRZTL1 (*Xho*I-cut) was the reaction substrate. Lane 7, a *Bst*EII digest of lambda DNA. B, interpretation of reaction products. This figure reproduces lane 4 of A and shows an analysis of various products, based upon secondary restriction digest experiments, electroporation followed by DNA sequencing, and mobility of products side by side with appropriately sized standards (see text for details). Band 1, released DBB DNA. Band 2, excised Tn5 DNA (ETF). Bands in region 3, knotted and relaxed forms of single intramolecular inversion transposition events. Band 4, single end cleavage products. Bands in region 5, double transposition events (one intramolecular and one intermolecular). Band 6, intermolecular transposition products. C, a protein is associated with transposon containing transposition products. A typical reaction mixture was divided into 4 equal aliquots and then treated as follows prior to loading. Lane 1, sample had no further treatment except for a 2-fold dilution. Lane 2, sample was adjusted to 2% SDS, heated for 15 min at 68 °C, and then diluted 2-fold. Lane 3, sample was extracted 3 times with CHCl₃ and then diluted 2-fold. Lane 4, sample was extracted once with phenol, once with phenol/chloroform, and once with chloroform and then diluted 2-fold. The products in lane 4 are identified as described for B.

Control DNA plasmids were prepared in the same fashion but were subjected to sufficient vortexing during the GeneClean II procedure to produce linear and relaxed forms of the plasmids. For the rest of the experiments, pRZTL1 was prepared with the Qiagen Plasmid Maxi Kit, followed by an additional alkaline treatment, neutralization, and ethanol precipitation to enrich for supercoiled DNA. Minor amounts of multimers were present after this procedure.

Protein Purification—Hyperactive mutant Tn5 transposase was purified in general as described previously (22) with a few modifications. Cells for induction were grown starting from a single colony after transformation of *E. coli* BL21(DE3)pLysS with plasmid pRZPET2. The colony was chosen for the inoculation that gave the highest extent of killing on agar containing 0.1 mM isopropyl-1-thio-β-D-galactopyranoside. This test is based upon the lethal effect of transposase overproduction (23). One liter of cells was induced at $A_{600} = 0.6$ by the addition of 0.5 mM isopropyl-1-thio-β-D-galactopyranoside followed by incubation for 1.5 h at 37 °C. Cells were collected by centrifugation and washed once with TEGX (20 mM Tris-HCl, 1 mM EDTA, 10% glycerol, 0.1% Triton X-100, pH 7.5) buffer modified to contain 0.1 M NaCl and 0.1 mM phenylmethylsulfonyl fluoride and resuspended in 20 ml of the same buffer. The cells were lysed by two passages through a French pressure cell at 16,000 lb/in². The lysate was centrifuged for 20 min at 20,000 × *g* and then the supernatant was brought to 0.25% polyethyleneimine (Sigma), and the mixture was immediately centrifuged for 10 min at 10,000 × *g*. The supernatant was brought to 47% saturation with ammonium sulfate, and the pellet was collected by centrifugation at 10,000 × *g* for 15 min. The pellet was resuspended in 5 ml of 0.3 M NaCl/TEGX buffer containing 10% glycerol, dialyzed against this buffer, and applied to a 10-ml heparin-agarose (Sigma) column equilibrated in the same buffer. After washing with 3 column volumes of buffer, Tnp was eluted with a 100-ml linear (0.3–1.2 M) NaCl gradient. Tnp was eluted at approximately 0.6 M NaCl. Fractions with higher activity (earliest in the peak) were pooled and stored at –70 °C in the same buffer containing 10 mM CHAPS; Pierce). Tnp of this quality (86% pure, see Fig. 4, lane 2) was used in most experiments described in this paper. More highly purified Tnp was obtained, in addition, by applying a Reactive Yellow No. 3 (Sigma) column chromatography step² followed by chromatography on a second heparin-agarose column. Protein of this quality (96% pure, Fig. 4, lane 3; purified by D. York) gave essentially the same reaction efficiency. Estimation of the protein purity was accomplished by analyzing the Coomassie Blue-stained gel shown in Fig. 4 with a Personal Densitometer (Molecular Dynamics) set at the highest sensitivity.

***In Vitro* Reaction Conditions**—For a typical reaction, 2 μl (approximately 0.2 μg of protein/μl) of Tnp was added to 18 μl of pRZTL1 plasmid (approximately 1 μg of DNA) in the transposition reaction

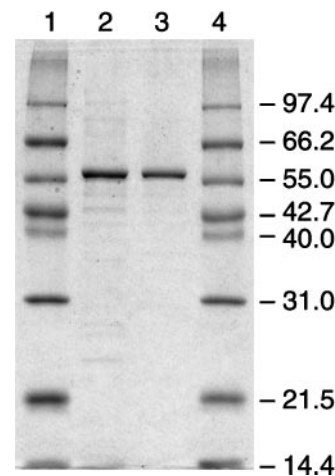


FIG. 4. 12% SDS-polyacrylamide gel electrophoresis analysis of purified hyperactive Tn5 transposase. Lanes 1 and 4 are molecular massweight markers (mid-range, Promega). Lane 2, standard Tnp preparation used for all experiments (86% purity). Lane 3, 96% pure Tnp. The 53-kDa Tnp is a very basic protein and known to run slower on this type of a gel than expected from its molecular mass. Note that both preparations are practically free from Tnp α and β degradation products usually found in wild type Tn5 Tnp preparations. This is probably an important feature of LP372 mutation (11).

buffer (0.1 M potassium glutamate, 25 mM Tris acetate, pH 7.5, 10 mM Mg²⁺ acetate, 50 μg/ml bovine serum albumin, 0.5 mM β-mercaptoethanol, 2 mM spermidine, 100 μg/ml tRNA; final concentrations). The reaction was incubated for 1 h at 20 °C and then diluted 2–3-fold in the same buffer and transferred to 37 °C. This procedure was performed to facilitate binding in the presence of CHAPS present in the reaction mixture as a component of the Tnp storage buffer (TEGX, 10 mM CHAPS). Dilution increased the cleavage reaction presumably due to dilution of CHAPS. CHAPS can be eliminated from the reaction (and storage buffer), and a simple incubation at 37 °C with no dilution is satisfactory if Qiagen-purified DNA is used. We nevertheless used the above two-step procedure for all experiments to synchronize the start of cleavage.

Analysis of Reaction Products—Reaction products were analyzed by agarose gel electrophoresis followed by staining with SYBR Green II (Molecular Probes) following the manufacturer's recommendations, and pictures were taken with a Fluorimager S1 (Molecular Dynamics) and analyzed with software provided by the same company.

DNA Sequencing—Sequencing of transposition products and PCR

² L. Braam, personal communication.

fragments was accomplished with a modified dideoxy chain termination procedure with use of 10% dimethyl sulfoxide (Sigma), boiling, and snap-cooling (24), Sequenase 2.0 (U. S. Biochemical Corp.), terminating mixtures (U. S. Biochemical Corp.), and as a sequencing buffer, KGB (20).

Ligation-mediated PCR—Ligation-mediated PCR was used to define the 5' end exposed after transposon excision during the *in vitro* reaction. Ligation-mediated PCR was performed essentially as described (25, 26). After a standard reaction with plasmid pFMA18700Δ, 2 μl of the reaction mixture was added to 20 μl of a primer extension mixture (1× KGB, 1 mM primer S30: 5'ACCTCGGTTCAAAGAGTTGG 3'). After heating in a boiling bath for 3 min and annealing at 37 °C for 15 min, 1 mM dNTPs and 2 units of Sequenase 2.0 (U. S. Biochemical Corp.) were added, and the reaction was incubated for 20 min at 37 °C. Then Sequenase was heat-inactivated, and after a 2-fold dilution in water, ATP (10 mM final), 1 unit of ligase (Promega), and the anchor linker (5 μM final) were added. The anchor linker was assembled from the following oligonucleotides: S28, 5' GGCTCGGGACCGTGGCTAGCATTAGT 3'; and S29, 5' ACTAATGCTTAG 3' (26). After overnight ligation at 16 °C, chloroform extraction, ethanol precipitation, and resuspension in 20 μl of H₂O, 5 μl was added to 100 μl of a standard PCR reaction with 0.2 μM primers S28 and S31 (S31, 5' GGTAGCTCAGAGAACCCTCG 3'). After 30 cycles (93 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min) products of the reaction were run on a 1.5% agarose gel, and after staining with ethidium bromide, the DNA band corresponding to expected fragment (250 bp) was recovered from the agarose using the GeneClean procedure and directly sequenced as described above. The primer for sequencing was S32, 5' CGAAAACCGCCCTTGAAGG 3'.

RESULTS

Experimental System—The four critical components of an *in vitro* transposition system are the DNA substrate, the DNA target, the protein component(s), and the reaction solution conditions. These are discussed below in the context of our studies on Tn5.

Extensive *in vivo* experimentation has shown that two properly oriented 19-bp OE sequences (or for IS50 transposition, one OE sequence, and one 19-bp IE sequence) are the only substrate DNA sequences required for Tn5-specific transposition (27–30). Thus we constructed as a substrate DNA molecule plasmid pRZTL1 in which the only Tn5-related sequences are two inverted 19-bp OEs (Fig. 2). These two inverted 19-bp OE sequences flank an origin of replication, a chloramphenicol resistance gene, and a “silent” tetracycline resistance gene that can only be expressed following a DNA rearrangement placing this gene downstream of an active promoter. In the experiments to be reported in this communication, pRZTL1 was also used as a target molecule. Transposition was assayed biologically by electroporating the plasmid DNA preparation following the indicated treatment into *E. coli* DH5α and quantifying the number of tetracycline-resistant colonies. The structures of the resulting plasmid DNAs in representative tetracycline-resistant cells were analyzed. In addition, the reaction products were directly analyzed following phenol extraction by agarose gel electrophoresis (see Fig. 3, A and B).

The only protein component added to the Tn5 transposition reactions was purified, hyperactive mutant Tnp. Three purification procedures for preparing Tnp were utilized with equivalent results. In most experiments we utilized Tnp purified by a procedure similar to that described previously (22). These preparations were approximately 86% pure as judged by Coomassie stain analysis of SDS-polyacrylamide electrophoresis gels (see Fig. 4, lane 2). Two additional purification steps were imposed on some preparations (reactive dye Yellow No. 3 chromatography and a second heparin agarose chromatography) to yield a preparation that was over 95% pure (see Fig. 4, lane 3). Finally, one preparation of a GST-Tnp fusion protein that had been purified by affinity chromatography was used. Since all preparations gave essentially identical results (data not shown), these experiments suggest that Tnp is the only protein required for Tn5 transposition; however, we cannot rule out

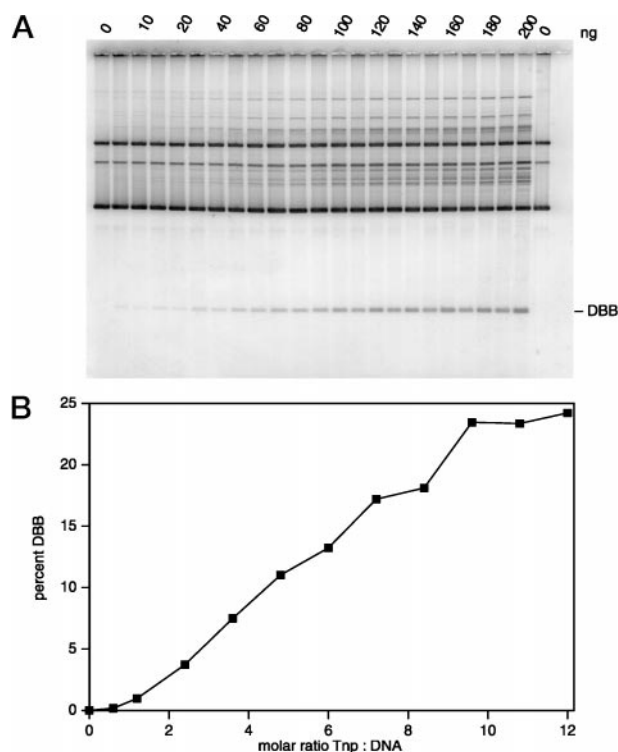


FIG. 5. Reaction dependence on Tnp concentration. Transposase was diluted in storage buffer, and 2 μl of each dilution was added to 18 μl of reaction buffer containing 0.26 pmol of pRZTL1 DNA. The amount of transposase added varied from 0 to 3.8 pmol (200 ng). The reactions were incubated for 4 h at 37 °C and then analyzed by 1% agarose gel electrophoresis. *A*, a 1% agarose gel analysis of reaction products stained with SYBR Green II. Labeled lanes contain reaction products generated by the indicated amounts of Tnp. Unlabeled lanes contained a dilution series of the 200-ng reaction products mixed with 0-ng reaction products. These were included to ensure detection linearity. *B*, yield of DBB as a function of the Tnp:substrate DNA molar ratio in different reactions. 100% DBB would be equal to the amount of DBB generated from all of the added substrate molecules.

that a low level host protein contaminant is also required for the reaction. Moreover, the electrophoretic mobility of many of the reaction products is dependent on a prior phenol extraction (see below) implying that Tnp release from these products has not been effected in our defined system.

The Tnp concentration dependence of the transposition reaction has been examined in the experiment shown in Fig. 5. Various amounts of transposase, from 0 to 3.8 pmol, were added to a standard 20-μl reaction containing 0.26 pmol of pRZTL1 and incubated for 4 h at 37 °C. The products were then analyzed by agarose gel electrophoresis. The yield of donor backbone (DBB) product, which is easiest to quantify, is shown in Fig. 5*B*. It is obvious that the reaction is dependent upon the amount of Tnp added and that this dependence is slightly sigmoidal.

Two reaction buffers have been used in our studies. The first buffer is similar to that reported for the analysis of Tn10 transposition (31). By using the hyperactive EK54, MA56, LP372 Tn5 Tnp, these conditions were found to yield both bona fide Tn5 transposition events and additional events that involved cutting at a cryptic sequence similar to IE located as a direct repeat of its partner OE sequence (data not shown; the location of the cryptic sequence is indicated in Fig. 2). In addition, one of the transposition events generated (out of six studied) under these conditions had a 10-bp target site duplication as opposed to the expected 9-bp duplication (data not shown). Apparently, these conditions gave rise to transposition type reactions with relaxed specificity for Tn5. A similar observation

was made for Tn10 transposition reactions using this type of buffer (31). A second set of conditions (using the potassium glutamate-based buffer described under "Experimental Procedures") gave rise almost exclusively to events that involved the correct OE sequences (however, when substrate/target DNA preparations generated by the alkaline lysis procedure were used, a low frequency of cryptic end cutting was apparent) and generated only the expected 9-bp target site duplications in the transposition products. This buffer was examined in more detail by studying the dependence on spermidine concentration (optimal transposition occurred at 2 mM) and the dependence on the presence of Mg²⁺ (as expected, no transposition occurred in the absence of Mg²⁺) (data not shown).

The transposition reaction has been shown to be dependent upon the presence of two OE sequences. First, *in vitro* transposition product molecules either involve precise joining of the OE sequences to new sequences or, for intermediates, precise cleavages at the OE-DBB boundaries (see analyses below). Second, Tnp fails to metabolize plasmid DNA (such as pRZTL0 from which pRZTL1 was derived, see "Experimental Procedures") that lacks OE sequences (data not shown). In separate studies we have examined the effect on the transposition reaction of single base pair changes in just one of the two OE sequences. Jilk *et al.* (32) had previously demonstrated that some single base pair changes (for instance at position 6) in one OE completely blocked *in vivo* transposition but instead led to the formation of rare adjacent deletions next to the wild type OE sequence. This same experiment has been repeated for the *in vitro* system with identical results.³ Thus we conclude that not only does the *in vitro* system depend upon the presence of OE sequences but also that a single defect in one can effectively block the reaction.

Analysis of Reaction Products—Efficient *in vitro* transposition was shown to have occurred both by transformation and by direct electrophoretic analysis of phenol-extracted products. As noted above, tetracycline resistance can only arise in this system if the transposable element is transposed to a site that is downstream of an active promoter. In a typical experiment utilizing purified hyperactive Tnp (containing mutations EK54, MA56, and LP372), the frequency of transformed cells that were tetracycline-resistant (the indicator product of transposition events) was 0.5% of the chloramphenicol-resistant cells (having received pRZTL1 DNA). However, this number underestimates the total transposition frequency because the detection system limits the target to 1/18 of the plasmid DNA (*i.e.* the chloramphenicol resistance gene was the only available target for intramolecular events which represented at least 95% of the products in our standard system (see below)), and this gene is 660 bp long relative to the plasmid length of 5838 bp. In addition only inserts of one orientation will yield transcription of the tetracycline resistance gene.

Plasmid DNA was analyzed from over 300 independent tetracycline-resistant transformants.⁴ Two types of final products representing intramolecular events and intermolecular events were recovered. An analysis of representative DNAs is shown in lanes 2 and 3 of Fig. 3A. As expected, the relative abundance of intermolecular *versus* intramolecular events is related to the DNA concentration in the transposition reaction. For instance, a reaction performed at 0.05 μg of DNA/ml yielded 5% intermolecular and 95% intramolecular products, whereas a reaction performed at 0.005 μg of DNA/μl yielded only intramolecular products. The transposon-target junctions on these

plasmids have been sequenced and proven to be the products of true Tn5 transposition events with 9-bp repeats flanking the OEs.⁴ When the same reaction was performed with purified MA56 Tnp, no tetracycline-resistant transformants were found (in contrast to over a thousand transformants generated in a standard reaction using EK54, MA56, LP372 Tnp) suggesting that the EK54 + LP372 changes have increased the *in vitro* activity of Tnp by at least 10³-fold.

Phenol-extracted products of two independent reactions utilizing supercoiled substrate DNA and hyperactive Tnp were directly analyzed by agarose gel electrophoresis (Fig. 3A, lanes 4 and 5). The DNA was recovered from different positions in the gel and studied by restriction digestion analysis and by transformation into competent cells, followed by restriction digestion and sequence analysis of the resulting plasmids. These analyses led to the following interpretations (see Fig. 3B).

The DNA in band 1 has an apparent molecular weight (as judged by its electrophoretic mobility) and a restriction digest pattern consistent with these molecules being released donor backbone DNA. This conclusion is verified by the results shown in Fig. 3A, lane 6, in which the reaction substrate was pRZTL1 linearized at the *Xho*I site within the donor DNA backbone. In this case the resulting product is the expected donor backbone subfragment.

Band 2 is excised Tn5 DNA. This was determined from its apparent molecular weight (notice that its mobility is identical to that of the linearized intramolecular transposition product in lane 2 that should be the same length as the excised Tn5 DNA) and its restriction digestion pattern (data not shown) (see also the experiment described in Fig. 7B).

The bands in region 3 contain knotted and relaxed forms of single intramolecular inversion transposition events. These structures were determined by analyzing the effect(s) of digesting these species with *Nco*I that cuts once in the transposon sequence. This digestion gave rise to a single molecular weight species whose electrophoretic mobility was that of a linear transposon sequence. Intramolecular transposition events of supercoiled DNA is expected to yield knotted transposon length circular products if Tnp is keeping the transposon topologically closed after cleavage but prior to strand transfer.

Region 3 may also contain catenated deletion circles (alternative products of intramolecular transposition events), but these would not have yielded uniformly sized products upon *Nco*I digestion.

Band 4 contains the products of single cleavage events (next to one or the other of the OE sequences). This was determined from its apparent molecular weight and its restriction digestion pattern (data not shown).

The bands in region 5 contain products of double transposition events (one intramolecular and one intermolecular). These products were identified by analyzing the plasmids found following transformation. The presence of multiple bands is presumably due to the presence of various numbers of knots in the products.

The DNA in band 6 contains intermolecular transposition products (single events) as determined by analyzing product DNAs following transformation.

The efficient production of the 1256-bp released donor backbone DNA fragment, the appearance of the excised Tn5 DNA fragment, the production of transposition products in the apparent absence of replication machinery, and the absence of DNA products that migrate more slowly than intermolecular transposition products lead us to the conclusion that Tn5 transposes via a "cut and paste" conservative mechanism. This conclusion is supported by the kinetic analysis of the transposition reaction.

³ I. Y. Goryshin, J. A. Miller, and W. S. Reznikoff, unpublished observations.

⁴ I. Y. Goryshin, J. A. Miller, Y. V. Kil, V. A. Lanzov, and W. S. Reznikoff, submitted for publication.

Transposition Products (but Not Released Donor Backbone DNA) Have Associated Proteins—One of the final steps in the transposition process is release of the transposase from the rearranged DNA molecules. We have examined whether this process occurs in our system by studying the agarose gel mobility of reaction products either directly from the reaction mixture, after chloroform extraction, after SDS treatment, or after phenol extraction. An equivalent amount of released donor backbone DNA is apparent in all four preparations (see Fig. 3C), but many other product DNAs are not apparent in the untreated preparation as compared with the phenol-extracted or SDS-treated preparations. CHCl_3 treatment appears to give a mixture of product mobilities. Since Tnp and bovine serum albumin are the only proteins present in stoichiometric quantities in these reactions, these results suggest that Tnp is tightly associated with Tn5 containing products after the transposition reaction.

Transposition Reaction Time Course—Considerable information regarding the transposition reaction mechanism can be deduced by studying the time course for the generation of various key reaction products. Preliminary experiments indicated that the hyperactive Tn5 Tnp will bind to OE sequences (and presumably form synaptic complexes) but will not catalyze transposition when incubated with pRZTL1 DNA at 20 °C (data not shown). We used this observation to perform a synchronized transposition reaction. A standard reaction system was incubated at 20 °C for 1 h, and then the temperature was brought to 37 °C. Samples were removed at various times between 0 and 9 h, phenol-extracted, and analyzed by agarose gel electrophoresis.

The results are shown in Fig. 6, A–C. In Fig. 6A we have marked the three product molecules that can be unambiguously quantitated, the excised transposon fragment (ETF) which should be an intermediate in a cut and paste transposition event, the released donor backbone (DBB), which should be a dead-end product, and the intermolecular transposition product (Inter). These products were quantitated using a fluorimeter, and the appearance of products with time is shown in Fig. 6B. The ratio of the abundance of ETF/DBB and ETF/Inter is presented in Fig. 6C. All products appear early, but the ETF appears to be chased into other products as would be expected for a cut and paste transposition reaction.

We have also analyzed the kinetics of intermolecular transposition product formation. The intermolecular transposition product appearance resembles that for DBB except that after 6 h there is a reduction in their abundance, perhaps suggesting that they are being used as transposition substrates.

Linear DNA Acts as an Efficient Substrate—Fig. 3A, lane 6, displays the result of an experiment in which pRZTL1 DNA linearized in the donor backbone was used as a substrate. This result clearly demonstrates that linear DNA containing two inverted OE sequences can undergo Tnp-mediated cleavage and strand transfer. We have also performed reactions with pRZTL1 linearized by cleavage within the transposable element and reactions with pRZTL1 cut within both the donor backbone and the transposable element (thus breaking any covalent linkage between the two OE sequences). In all cases Tnp mediates DNA rearrangements (data not shown).

Tn5 Tnp Produces a Precise Cleavage of the OE Donor Backbone Junction—We have used ligation-mediated PCR analysis of a typical reaction mixture to define the 5' cleavage product generated by Tnp at one end of the Tn5 substrate. After a 3-h transposition reaction the DNA was phenol-extracted, denatured, and hybridized to a primer located 250 bp away from and facing one OE sequence. The primer was extended using Sequenase. A pre-existing nick in the template strand will give

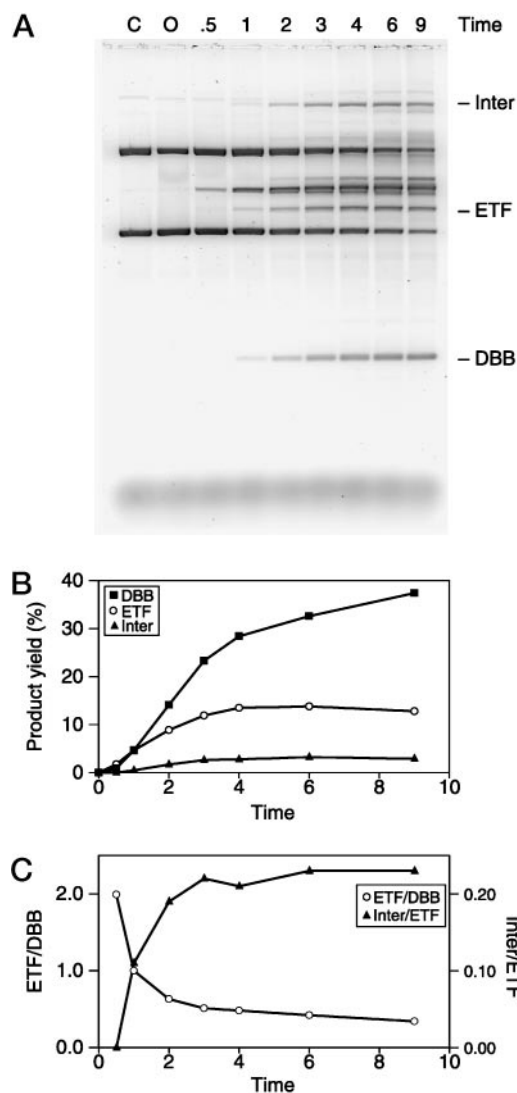


FIG. 6. Time course of transposition reaction. A, an agarose gel analysis of phenol-extracted products found after 0–9 h incubation at 37 °C (all samples had been preincubated for 1 h at 20 °C except for that in lane C). Three types of products have been quantitated as follows: excised transposon fragment (ETF), donor backbone fragment (DBB), and intermolecular transposition product (Inter). B, time course of DBB, ETF, and Inter appearance. Since the sizes of three molecular types are different, the yields were normalized to the size of the initial substrate. The normalization coefficients are 4.65, 1.27, and 0.56 for DBB, ETF, and Inter, respectively. Using these corrections, each product is shown as a percent of its maximum possible yield. Note that intermolecular transposition products (with a final yield in this experiment of about 3%) represent only a small portion of the intramolecular + intermolecular transposition product total (about 5% in a typical experiment as judged from transformation tests), so the total abundance of final transposition products must be higher than the abundance of ETF but cannot be precisely estimated from Fig. 3A. C, ratio of ETF to DBB, and Inter to ETF. The first ratio decreases and the second one increases with time indicating that ETF is an intermediate in the reaction pathway.

rise to a blunt end to which a double-stranded linker was ligated. A PCR reaction was then performed utilizing the appropriate two primers (one from the linker and the other adjacent to the original primer), and the resulting molecule was sequenced. As shown in Fig. 7A, there is one ligation-mediated PCR product generated in the vicinity of the OE, and this corresponds to a cleavage event precisely between the +1 position of the OE and the –1 position of the DBB. This defines the 5' end of the Tnp-cleaved molecules. If we assume that the 3' end that is transferred in the overall reaction represents the position of the 3' cleavage reaction (which is true for other

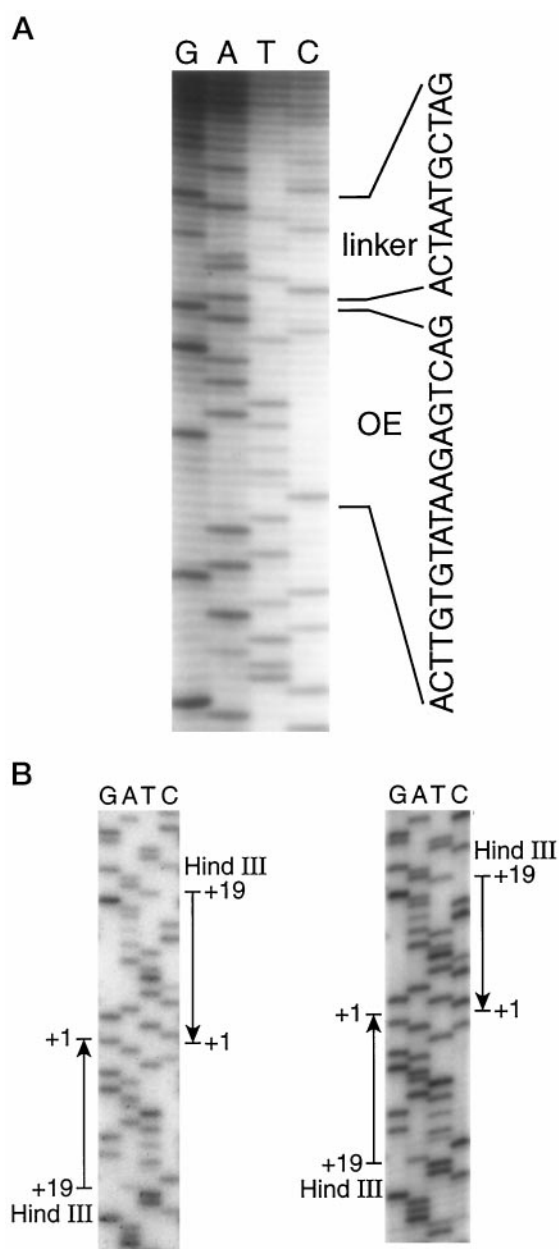


FIG. 7. Analysis of 5' and 3' transposable element ends exposed during Tn5 *in vitro* transposition. *A*, ligation mediated PCR analysis of 5' end at the OE of Tn5. Exposed 5' end at the OE (nick or cleaved) was converted to a blunt end by primer extension. Then a linker 5' ACTAATGCTAG 3' and 3' TGATTACGATCGGTGCCAG-GCTCGG 5' was ligated to the blunt ends, defining a joint between the linker and the exact position of the exposed 5' end. After PCR and recovery of a fragment of predicted length from an agarose gel, DNA sequence analysis was performed. As seen in the figure, the text of a linker follows position +1 of OE. The same approach has been applied to the *in vivo* system with the same results. *B*, sequence analysis of the two OE of ETF after recovery from a gel, ligation, and electroporation. As can be seen on the gel text, one OE, after position +1, is followed in inverted orientation by position +1 of the opposite OE. Two HindIII sites (see Fig. 2) are also shown to stress symmetry of the pattern. This junction can only arise if both OE sequences were cut by transposase as blunt ends at position +1.

transposition systems) then we conclude that Tn5 Tnp generates a double strand cleavage precisely between the OE +1 position and the -1 position of the donor DNA backbone. This conclusion is supported by the analysis of the excised transposon fragment described below.

The precise nature of the Tn5 Tnp cleavage products was

also studied by sequence analysis of artificially circularized Tn5 excised transposon fragments. As shown in Fig. 3, treatment of pRZTL1 with purified hyperactive Tn5 Tnp releases the OE-defined sequence (band 2, Fig. 3B) from the donor backbone. DNA was extracted from the band 2 region of a typical gel-fractionated reaction mixture. This DNA was treated with T4 DNA ligase and transformed into cells selecting for chloramphenicol resistance (the pRZTL1 Tn5 construct encodes a p15 origin of replication and a chloramphenicol resistance gene, see Fig. 2). Five of 12 chloramphenicol-resistant colonies contained plasmids of the expected molecular weight (the other 7 contained smaller molecules). The sequence of the possible OE-OE junction was determined for all 5. An example sequencing gel is presented in Fig. 7B. All five DNAs had the same precise inverted OE-OE junction sequence with no extra or missing nucleotide pairs. This structure is most easily explained by the Tn5 excised transposon fragment having been generated by precise 3'-OH and 5'-PO₄ cleavages at the precise ends of the OE sequences. The shorter plasmids are most likely deletion derivatives resulting from the presence of the inverted repeat structure at the junction. The same blunt end cleavage of transposon ends was reported for Tn10 (33).

DISCUSSION

A critical requirement for studying transposition is the availability of a defined *in vitro* transposition system. In this communication we describe such a system for the bacterial transposon Tn5, and we use this system to address several key mechanistic questions. Initially we attempted to develop such an *in vitro* system utilizing purified wild type Tn5 Tnp. The resulting experiments demonstrated occasional, barely detectable levels of transposition as measured by a transformation assay (data not shown). We reasoned that perhaps the wild type Tnp was functionally inefficient because it had evolved to be largely inactive as a mechanism for ensuring that transposition occurred at a very low frequency *in vivo*. Based upon this assumption, we turned to hyperactive mutants of Tnp. Purified triple EK54/MA56/LP372 mutant Tnp proved to catalyze efficient *in vitro* Tn5 transposition that could be assayed either biologically or by direct gel analysis. Furthermore, the products of these transposition events were similar to those catalyzed by wild type Tnp *in vivo* in two key respects as follows: the transposition events utilized the 19-bp OE sequences, and the transposition events caused 9-bp target site duplications. In a separate study, we demonstrate that the target site preferences are the same for the *in vitro* as for the *in vivo* transposition.⁴ Thus the hyperactive EK54/MA56/LP372 Tnp system is a reasonable facsimile of Tn5 transposition.

Macromolecules Involved in Tn5 Transposition, the 19-bp End Sequences and Tnp—Transposition is a complex multistep process, and thus it is not surprising that many transposition systems are very elaborate involving a complex series of DNA sequences and several different polypeptides. *In vivo* studies suggested that this might not be the case for Tn5. If this were true, an *in vitro* Tn5 transposition system might offer a unique opportunity to study how a single polypeptide and single type of DNA sequence had evolved to perform multiple sequential steps. *In vivo* genetic studies have suggested that the only Tn5 related sequences required for transposition are two inverted 19-bp OE sequences (27–30). This situation is in clear distinction with that found for Mu (in which four 22-bp att sites and an internal enhancer sequence participate) (2) and Tn7 (in which approximately 150 bp are required on each end) (8). Even in the case of Tn10, an IHF binding sequence adjoining the 23-bp end sequence is intimately involved in the reaction resulting in the preferred end sequences being 70 bp long (4). The results reported in this communication are consistent with two inverted

and appropriately spaced 19-bp OE sequences being the only specifically required cis-active sequence for Tn5 transposition.

The *in vivo* analyses of protein participants in Tn5 transposition are not as clear since several host mutants do affect the frequency of Tn5 transposition (reviewed in Refs. 3 and 15). The most obvious candidate for a required host protein is DnaA, because DnaA binds specifically to the OE sequence (34) and because the frequency of Tn5 transposition is reduced approximately 10-fold in a *dnaA* host (35). Our *in vitro* studies suggest that no host protein including DnaA is required for Tn5 transposition at least through the strand transfer step. We have used three different types of purified Tnp. The standard preparation is 86% pure, and two additional chromatographic steps result in 96% pure Tnp (see Fig. 4). Both preparations are active in mediating efficient *in vitro* transposition. Furthermore, a GST-Tnp fusion purified by affinity chromatography is also active in mediating efficient transposition (data not shown). It is possible, of course, that a required host protein has copurified with Tnp in all of these preparations. In fact topoisomerase I is known to be present in the 86% pure preparation (but not in the 96% pure transposase preparation), although topoisomerase I is not active in our standard transposition reaction conditions.⁵ However, it is highly unlikely that any contaminating protein is present in the stoichiometric amounts usually required for host proteins involved in transposition.

There are two caveats to our proposal that Tnp is the only required protein for transposition. First, we have been using hyperactive mutant Tnp. It is possible that one or more of the mutant changes bypasses the need for a required host protein. A more interesting possibility is suggested by the results shown in Fig. 3C. Some protein(s) appears to be tightly associated with many of the Tn5 transposition products. Presumably this protein is Tnp. Thus release of Tnp may require a host factor. It is of interest that release of Mu Tnp from its products requires the presence of ClpX (6). We are currently testing whether DnaA is this required protein.

We have also studied the Tnp concentration dependence for the reactions leading to DBB release (Fig. 5). These reactions demonstrate a modest sigmoidal dependence upon Tnp abundance suggesting that very few Tnp protomers are needed for this partial transposition process. Since Tnp has gel exclusion properties of a monomeric peptide,⁶ we propose that the reaction is initiated by the binding of single Tnp molecules to each of the two ends, and no other Tnp molecules are needed at least until the transposon is cleaved free of the DBB.

Cut and Paste Mechanism—It was first hypothesized 20 years ago that Tn5 transposes through a conservative cut and paste mechanism (36). This is in contrast to the replicative transposition mechanism used by Mu, Tn3, and related transposable elements (1). The data presented in this communication confirm this hypothesis. A cut and paste transposition mechanism is strongly supported by the time course analysis presented in Fig. 6, in which the excised transposon band is generated first and then is reduced in abundance in parallel with the appearance of final transposition products. It is also

important to note that certain DNA structures that would be produced by a replicative process, in particular cointegrate type molecules that would travel more slowly than relaxed intramolecular products, are not observed in our experiments (see Figs. 3, 5, and 6). Finally, our analyses of the Tn5 donor backbone cleavage products strongly support the proposed cut and paste mechanism and specifically indicate that both strands are cleaved at the precise OE donor backbone boundary (Fig. 7).

Usefulness of the Tn5 System—Other investigators studying transposition systems may find the isolation and use of hyperactive mutants a useful approach. In addition, the Tn5 *in vitro* system is defined, efficient, and simple to use, thus it should find applications in the *in vitro* manipulation of DNA.

Acknowledgments—We thank Lisa Braam and Dona York for help in the preparation of Tnp; Maggie Zhou for providing the EK54 mutation; Warren Johnson for comments on the manuscript; Barb Shriver for preparing bacterial media and other solutions; and Michael Cox for helpful discussions.

REFERENCES

- Berg, D. E., and Howe, M. M. (1989) *Mobile DNA*, American Society for Microbiology, Washington, D. C.
- Mizuuchi, K. (1992) *Annu. Rev. Biochem.* **61**, 1011–1051
- Reznikoff, W. S. (1993) *Annu. Rev. Microbiol.* **47**, 945–963
- Sakai, J., Chalmers, R. M., and Kleckner, N. (1995) *EMBO J.* **14**, 4374–4383
- Bolland, S., and Kleckner, N. (1996) *Cell* **86**, 223–233
- Levchenko, I., Luo, L., and Baker, T. A. (1995) *Genes Dev.* **9**, 2399–2408
- Haniford, D. B., and Chaconas, G. (1992) *Curr. Opin. Genet. & Dev.* **2**, 698–704
- Gary, P. A., Biery, M. C., Bainton, R. J., and Craig, N. L. (1996) *J. Mol. Biol.* **257**, 301–316
- Craig, N. L. (1991) *Mol. Microbiol.* **5**, 2569–2574
- Wiegand, T. W., and Reznikoff, W. S. (1992) *J. Bacteriol.* **174**, 1229–1239
- Weinreich, M. D., Gasch, A., and Reznikoff, W. S. (1994) *Genes Dev.* **8**, 2363–2374
- Zhou, M., and Reznikoff, W. S. (1997) *J. Mol. Biol.* **271**, 362–373
- Bieck, D., and Roth, J. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 6047–6051
- Johnson, R. C., Yin, J. C.-P., and Reznikoff, W. S. (1982) *Cell* **30**, 873–882
- Berg, D. E. (1989) in *Mobile DNA* (Berg, D. E., and Howe, M. M., eds) pp. 185–210, American Society for Microbiology, Washington, D. C.
- Berg, C. M., Berg, D. E., and Groisman, E. A. (1989) in *Mobile DNA* (Berg, D. E., and Howe, M. M., eds) pp. 879–925, American Society for Microbiology, Washington, D. C.
- Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol.* **187**, 60–89
- Grodberg, J., and Dunn, J. J. (1988) *J. Bacteriol.* **170**, 1245–1253
- Mertens, N., Remaut, E., and Walter, F. (1995) *Bio/Technology* **13**, 175–179
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Goryshin, I. Y., Kil, Y. V., and Reznikoff, W. S. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10834–10838
- de la Cruz, N. B., Weinreich, M. D., Wiegand, T. W., Krebs, M. P., and Reznikoff, W. S. (1993) *J. Bacteriol.* **175**, 6932–6938
- Weinreich, M. D., Yigit, H., and Reznikoff, W. S. (1994) *J. Bacteriol.* **176**, 5494–5504
- Winship, P. R. (1989) *Nucleic Acids Res.* **17**, 1266
- Mueller, P. R., and Wold, B. (1989) *Science* **246**, 780–786
- Saveliev, S., and Cox, M. (1995) *Genes Dev.* **9**, 248–255
- Johnson, R. C., and Reznikoff, W. S. (1983) *Nature* **204**, 280–282
- Sasakawa, C., and Berg, D. E. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 7293–7297
- Phadnis, S. H., Huang, H. V., and Berg, D. E. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 5908–5912
- Weinreich, M. D., Mahnke-Braam, L., and Reznikoff, W. S. (1994) *J. Mol. Biol.* **241**, 166–177
- Chalmers, R. M., and Kleckner, N. (1994) *J. Biol. Chem.* **269**, 8029–8035
- Jilk, R. A., Makris, J. C., Borhardt, L., and Reznikoff, W. S. (1993) *J. Bacteriol.* **175**, 1264–1271
- Benjamin, H. W., and Kleckner, N. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4648–4652
- Fuller, R. S., Funnell, B. E., and Kornberg, A. (1984) *Cell* **38**, 889–900
- Yin, J. C.-P., and Reznikoff, W. S. (1987) *J. Bacteriol.* **169**, 4637–4645
- Berg, D. E. (1977) in *DNA Insertion Elements, Plasmids, and Episomes* (Bukhari, A. I., Shapiro, J. A., and Adhya, S. L., eds) pp. 205–212, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

⁵ H. Yigit and W. S. Reznikoff, manuscript in preparation.

⁶ L. Braam and W. S. Reznikoff, unpublished observations.