Tn*552* transposase catalyzes concerted strand transfer *in vitro*

ANDRES E. LESCHZINER, THOMAS J. GRIFFIN IV, AND NIGEL D. F. GRINDLEY*

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8114

ABSTRACT   The Tn*552* transposase, a member of the DDE superfamily of transposase and retroviral integrase proteins, has been expressed in soluble form. The purified protein performs concerted strand transfer *in vitro*, efficiently pairing two preprocessed transposon ends and inserting them into target DNA. For maximum efficiency, both participating DNA ends must contain the two adjacent transposase-binding sites that are the normal constituents of the Tn*552* termini. As is the case with transposition *in vivo*, the insertions recovered from the reaction *in vitro* are flanked by repeats of a short target sequence, most frequently 6 bp. The reaction has stringent requirements for a divalent metal ion. Concerted strand transfer is most efficient with Mg^{2+}. Although it stimulates strand transfer overall, Mn^{2+} promotes uncoupled, single-ended events at the expense of concerted insertions. The simplicity and efficiency of the Tn*552* transposition system make it an attractive subject for structural and biochemical studies and a potentially useful genetic tool.

The transposon Tn*552* is a 6.5-kb element isolated from *Staphylococcus aureus* responsible for resistance to β-lactam antibiotics (1, 2). It carries genes for three proteins likely to be involved in its transposition: a transposase (TnpA, previously called p480); a protein similar to the Mu B and Tn*7* TnsC proteins (TnpB, previously called p271), which, by analogy, may enhance transposition activity and play a role in target selection; and a resolvase (3). The presence of these three genes suggests that Tn*552* transposes by consecutive formation and resolution of a cointegrate between donor and target DNAs; however, this remains to be demonstrated.

The Tn*552* transposase belongs to the D,D(35)E superfamily of transposase proteins (3). Members of this superfamily include the integrase proteins of retroviruses and retrotransposons, the transposases of bacterial insertion sequences of the IS*3* family, bacteriophage Mu A protein, and TnsB of Tn*7*. The defining characteristic of this family of phosphoryl-transfer proteins is a homologous catalytic core domain containing a triad of invariant acidic residues, two aspartates and a glutamate, which are essential for catalytic activity and give the superfamily its name (3, 4). Despite their homology, transposases of the D,D(35)E superfamily form a rather divergent and heterogeneous group. They range in size from the small retroviral integrases (about 290 residues) to Mu A (663 residues) and Tn*7* TnsB (702 residues), and in many cases the region of homology does not appear to extend beyond the catalytic core (160–220 residues). Extensive biochemical studies of a few members of the family have provided a remarkably detailed picture of the transposition process including the substrates and proteins involved, the assembly of the transpositionally active protein–DNA complex, and the mechanism of strand transfer (5–7). However, complete understanding has been hampered by a lack of detailed structural information;

despite substantial efforts, three-dimensional structures have been limited to isolated transposase domains in the absence of substrates (8–10). We have turned to Tn*552* as a possible candidate for structural studies.

Of the well characterized members of the superfamily, Tn*552* transposase is most closely related to Tn*7* TnsB but differs significantly from it. At 480 residues, the Tn*552* transposase is much smaller, with the regions N- and C-terminal to the catalytic core about 100 residues shorter. Sequence alignments outside the 160-residue central catalytic region are tentative at best, bordering on the insignificant. The Tn*552* DNA substrate is also simpler. Each end of Tn*552* contains two adjacent 23-bp transposase-binding sites and, as is the case with other elements of this superfamily, terminates with the dinucleotide -$CA_{OH}3'$ (11). By contrast, the ends of Tn*7* (and also of Mu) are complex, containing three or more transposase-binding sites spread irregularly over 90–100 bp (5).

In this paper, we present an *in vitro* system for Tn*552* transposase-catalyzed concerted strand transfer. A previous attempt to obtain active Tn*552* transposase by Rowland and colleagues was only partially successful (11). Induced expression yielded a protein that could be solubilized only under strongly denaturing conditions. The purified and renatured protein bound specifically to the ends of Tn*552*. However, its catalytic activity was limited to uncoupled strand transfer of single ends in a reaction that required Mn^{2+} as the divalent cation. The protein was also unstable under the assay conditions and appeared to aggregate irreversibly. We have overcome these problems and produced a soluble protein proficient at synapsis and concerted strand transfer with Mg^{2+} as the metal ion.

MATERIALS AND METHODS

Expression and Purification of Tn*552* Transposase. The expression vector pKETH-8c carrying the Tn*552* transposase with an N-terminal His tag (11) and the *Escherichia coli* host strain BL21(DE3)[pLysS] used for the induction were kindly provided by S.-J. Rowland. Six to 10 liters of CSB1 medium (66 mM K_2HPO_4/33 mM NaH_2PO_4/15 mM $(NH_4)_2SO_4$/5 mM $MgCl_2$/5 mM sodium citrate/0.45 mM $FeCl_3$/44 μM $ZnCl_2$/30 μM $MnCl_2$/25 μM $CoCl_2$/25 μM Na_2MoO_4/24 μM H_3BO_3/20 μM $CaCl_2$/18 μM $CuCl_2$/0.6% glucose) supplemented with 10 μg/ml thiamin, 40 μg/ml kanamycin, and 20 μg/ml chloramphenicol were inoculated with a 1-liter overnight culture of the *E. coli* transposase expression strain grown in the same medium supplemented with 0.1% Casamino acids. Cells were grown at 30°C to an OD_{600} of 3–4, isopropyl β-D-thiogalactoside then was added to a final concentration of 0.4 mM, and growth was continued for an additional 2 hr. Cells (typically 11 g/liter) were harvested by centrifugation and the pellet was resuspended in 20 ml of 50 mM Tris·HCl (pH 7.5 at 4°C), 20% sucrose per liter of culture, and stored at −80°C. To lyse the cells, the suspension was thawed and brought to 1 M

*To whom reprint requests should be addressed at: Yale University, 266 Whitney Ave., P.O. Box 208114, New Haven, CT 06520-8114. e-mail: nigel.grindley@yale.edu.

NaCl and passed twice through the French press at 1,300–1,500 psi. The lysate was clarified by centrifugation (130,000 × g for 60 min) and then diluted with 50 mM Tris·HCl (pH 7.5 at 4°C), 20% sucrose, and 3.8 M $(NH_4)_2SO_4$ in 50 mM Tris·HCl (pH 7.5 at 4°C) to lower the NaCl concentration to 0.5 M and bring the final solution to 40% saturation in $(NH_4)_2SO_4$. The precipitated protein was recovered by centrifugation (35,000 × g for 30 min), resuspended in about 1/6 of the original volume of 50 mM Tris·HCl (pH 7.5 at 4°C)/0.5 M NaCl/20% glycerol (buffer A) and dialyzed against 40 vol of the same buffer for 2–3 hr. Imidazole was added to a final concentration of 12.5 mM, and the sample was loaded onto a 10-ml Ni^{2+}-NTA column (Qiagen) and preequilibrated in buffer A plus 12.5 mM imidazole. The column was washed with 5–10 vol of the same buffer and then 10–15 vol of buffer A plus 25 mM imidazole. Protein was eluted from the column with 10 vol of buffer A plus 200 mM imidazole and 20 mM EDTA. Fractions containing transposase were identified by SDS/PAGE, pooled, and dialyzed overnight against 50 vol of buffer A plus 750 mM $(NH_4)_2SO_4$/1 mM DTT/1 mM EDTA. The dialyzed sample was loaded onto an 8-ml phenyl Superose HR10/10 column (Pharmacia) equilibrated in the same buffer. The column was washed with about 10 vol of the loading buffer followed by about 3 vol of buffer A plus 375 mM $(NH_4)_2SO_4$. The column was then eluted with a 20-vol linear gradient of 375 mM–0 mM $(NH_4)_2SO_4$ in buffer A plus 1 mM DTT/1 mM EDTA. Fractions containing transposase were identified by SDS/PAGE, pooled, diluted 3-fold with 50 mM Tris·HCl (pH 7.5 at 4°C)/20% glycerol/1 mM DTT/1 mM EDTA (buffer B) plus 200 mM NaCl and dialyzed overnight against about 70 vol of the same buffer. One half of the dialyzed sample was loaded onto a Mono S HR5/5 column (Pharmacia) equilibrated in buffer B plus 200 mM NaCl. The column was washed with 10 vol of the loading buffer and then with a 5-vol gradient of 200–350 mM NaCl in buffer B. The column was then eluted with a 40-vol gradient of 350–600 mM NaCl in buffer B. The procedure was repeated with the second half of the sample. Fractions containing transposase were pooled, the protein was precipitated by addition of 3.8 M $(NH_4)_2SO_4$/50 mM Tris·HCl (pH 7.5 at 4°C) to 40% saturation in $(NH_4)_2SO_4$, recovered by centrifugation (17,000 × g for 10 min), and resuspended to the desired final concentration in 50 mM Tris·HCl (pH 7.5 at 4°C)/2.0 M NaCl/50% glycerol/1 mM DTT/1 mM EDTA. Protein concentration was determined from the A_{280}; the theoretical extinction coefficient (ε_{280} = 68,674 M^{-1}·cm^{-1}) was calculated from the amino acid sequence (12). The sample was then dialyzed to lower the NaCl concentration to 0.5 M. Final concentrations were confirmed from the A_{280} reading. Typical yields were around 36 mg of transposase per 100 g of cells. The transposase preparation used throughout retains the N-terminal His tag.

Preparation of DNA Substrates. A plasmid carrying seven tandemly repeated copies of a DNA fragment containing the two binding sites for TnpA from the right end of Tn*552*, the CA 3′, and two additional base pairs at the "inside" end was digested with *Xmn* I and *Tth111* I to release RI&II, a 50-bp substrate with the sequence shown in Fig. 5*A*. The fragment was purified on an 8% native polyacrylamide gel in TBE. RI was made by annealing AL80 (bottom strand in Fig. 5*A*) and AL88R (top strand in Fig. 5*A*), which were obtained from the Keck Biotechnology Resources Facility, Yale University Medical School. DNA concentrations were determined by measuring A_{260}.

Various analogues of Tn*552* encoding kanamycin resistance were made by PCR, by using as template a linear DNA fragment containing the Tn*903 kan* gene. The primers used for the left end contained LI+LII (AL25; 5′-cttggcgcgccatactagT-GTTACTTTACTTGATATATGAGAATGATTTTACCT-AGTAAATGAGAAcgtcaagtcagcgtaatgctc), LI (AL27; 5′-cttggcgcgccatactagTGTTACTTTACTTGATATATGAGA-

Acgtcaagtcagcgtaatgctc), or no repeats (AL23; 5′-cttggcgcgccatactagTGagcaatgtcctatattaaagtac). Those for the right end contained RI+RII (AL26; 5′-cttccatggactagTGTC-ACTTTGCTTGATATATGAGAATTATTTAACCTTAT-AAATGAGAAgttgtgtctcaaaatctctgatg), RI (AL28; 5′-cttccatggactagTGTCACTTTGCTTGATATATGAGAA-gttgtgtctcaaaatctctgatg), or no repeats (AL24; 5′-cttccatggact-agTGaaagcaacgcactttaaataag) (Tn*552* sequences shown in uppercase). All primers were designed to contain the "preprocessing" *Spe*I site. PCR products were digested with *Spe*I and cloned into the *Spe*I site of LITMUS 28. The desired transformants were identified by restriction analysis, and the transposon-LITMUS 28 junctions—including the TnpA-binding sites—were confirmed by sequencing. The plasmid pAL101 carries the Tn*552kan* substrate, with the *kan* gene flanked by the outer 48 bp from each end of Tn*552*. The Tn*552kan* elements were obtained by digesting the appropriate plasmid with *Spe*I and purifying the fragments from low-melting-point agarose gels (SeaPlaque GTG, American Bioanalytical, Natick, MA) using either β-agarase (New England Biolabs) followed by ethanol precipitation and resuspension in TE or with the Geneclean II kit (Bio 101) followed by resuspension in TE.

Oligonucleotide substrates and the various Tn*552kan* DNAs were 5′ labeled by using [γ-^{32}P]ATP and T4 polynucleotide kinase (New England Biolabs). Labeled fragments were separated from unincorporated nucleotides with G-25 Quick Spin columns (Boehringer Mannheim). RI&II was labeled in its double-stranded form. AL88R was hybridized to the labeled AL80 in equimolar amounts.

Strand Transfer Assays. Typically, reaction mixtures (30 μl) contained 120 ng of transposase, 180–270 ng of pUC19; 50 fmol–3 pmol of DNA substrate; 10 mM Mg^{2+} (unless otherwise indicated), 133 mM NaCl, 2 mM EDTA, 50% (wt/vol) glycerol; and 50 mM Tris·HCl (pH 7.5 at 4°C). Reactions were incubated at 37°C for 1–3 hr and subsequently the products were precipitated with ethanol. Pellets were resuspended in 10 mM Tris·HCl, pH 7.5/0.1 mM EDTA, digested with *Pst*I when appropriate, then quenched by addition of 1/4 volume of 5× Ficoll dye mix (2.5% SDS/75 mM EDTA/50 mM Tris·HCl, pH 7.5/25% Ficoll/bromophenol blue/xylene cyanol) and proteinase K (0.5–1.0 mg/ml final concentration). After at least 30 min at 37°C, samples were analyzed by 0.8% agarose (SeaKem GTG, American Bioanalytical) gel electrophoresis in TBE buffer, and products were detected by ethidium bromide staining and autoradiography.

Cloning and Sequencing of Transposition Products. Competent JM109 cells (Promega) were transformed with the gel-purified concerted transposition product; alternatively, JM109 cells were electroporated directly with the outcome of a standard strand transfer reaction (resuspended in H_2O). Transformants were selected with kanamycin. Plasmid DNAs were prepared by using Qiagen kits. Transposon–target junction sequences were determined using the primers 5′-GTTA-ATTGGTTGTAACACTGGC (AL32) and 5′-TGATGATA-TATTTTTATCTTGTGC (AL31) for left and right junctions, respectively. Both primers anneal within the *kan* gene, close to the Tn*552* ends.

RESULTS

Believing that the limited activity observed earlier (11) might have resulted from the necessity to denature the initially insoluble protein, we set out to purify Tn*552* transposase in a soluble form. Using the same combination of expression plasmid and host strain as Rowland *et al.* (11), we explored a variety of growth and induction parameters, including culture medium, growth temperature, cell density at induction, length of induction, and concentration of inducer. We found that using a defined medium and growing the cultures at 30°C were

critical to our success. A pool of soluble His-tagged transposase was obtained and purified using standard methods (see *Materials and Methods*) to better than 95% purity.

In Vitro Concerted Strand Transfer with Tn*552kan.* To test for activity, we incubated our soluble transposase with a labeled, preprocessed (i.e., with exposed $CA_{OH}3'$ ends) substrate, Tn*552kan*, which consists of a kanamycin-resistance gene flanked by the terminal 48 bp from the transposon's ends (see Fig. 1*A*). Concerted strand transfer of both ends of Tn*552kan* into target plasmid (pUC19) should result in a relaxed circle whose length will be that of pUC19 (2.7 kb) plus that of the Tn*552kan* (1.1 kb). Digestion with *Pst*I (which does not cut Tn*552kan*) should give a 3.8-kb linear species. Insertions of a single end should produce nicked circles containing a 1.1-kb "tail." Digestion of these products with *Pst*I should result in Y-shaped molecules with two branches of variable length (assuming random insertion) that will not migrate as a discrete species on a gel.

The major product of Tn*552kan* transposition formed in the presence of Mg^{2+} comigrated with a marker for the 3.8-kb nicked circle (Fig. 1*B*, lane 2, and data not shown) and is readily visible in the ethidium-stained gel (lane 9). The identity of this species as the product of concerted strand transfer was confirmed by its conversion, upon *Pst*I digestion, to a 3.8-kb linear species (Fig. 1*B*, lane 3). Other products visible in lane 3 included (*i*) a band of about 2 kb that remained unchanged

by *Pst*I digestion and was also formed in reactions lacking pUC19 target DNA (data not shown)—we infer this results from concerted insertion of one molecule of Tn*552kan* into another; (*ii*) a minor band that migrated slightly faster than the 3.8-kb nicked circles and disappeared upon *Pst*I digestion—we infer this to be the nicked form of pUC19 with a 1.1-kb tail that results from insertion of a single transposon end; (*iii*) a smear of radioactivity below the substrate that likely contains the products of intramolecular events; and (*iv*) a complex mixture of minor bands that migrated slower than the 3.8-kb nicked circles—we have not attempted to identify these but assume they result from multiple insertions into the same target and insertions into multimers of pUC19.

A divalent metal ion is essential for strand transfer activity (lane 1), and the use of alternative divalent metal ions changes the strand transfer activity of the Tn*552* transposase. In the presence of Mn^{2+}, we saw an increase in both transposition activity and product complexity (Fig. 1*B*, lanes 4 and 5). However, very little concerted transposition was observed; rather, the major product appeared to be the single-ended insertion. An additional band migrating just ahead of these products can be observed. Although we have not determined its identity we believe it might contain θ-shaped molecules resulting from insertion of both ends of a Tn*552kan* element in two independent, single-ended events. With Ca^{2+}, transposition activity was very low and the only products detected resulted from single-end strand transfer (data not shown).

The outcome of the reactions was independent of whether the target DNA was supercoiled or linear. When we performed reactions with Mg^{2+} or Mn^{2+} using *Pst*I-linearized pUC19 as the target (Fig. 1*B*, lanes 6 and 7), the patterns of products were virtually identical to those obtained after *Pst*I digestion of the reactions with supercoiled target DNA (compare lanes 3 and 6, and 5 and 7 in Fig. 1*B*).

Tn*552kan* Insertions Are Flanked by Direct Repeats of a Target Sequence. One of the hallmarks of transposition is the presence, flanking the inserted element, of copies of a short, duplicated target sequence. To determine whether insertion of Tn*552kan in vitro* was associated with target duplication, we performed a standard reaction with supercoiled pUC19 as target and Mg^{2+} as the divalent cation. Total DNA from this reaction, or DNA purified from an ethidium-stained gel slice containing the band predicted to be the concerted strand transfer product was used to transform *E. coli*. We analyzed plasmid DNAs from 14 colonies resistant to kanamycin. All contained single insertions of Tn*552kan* with the transposon
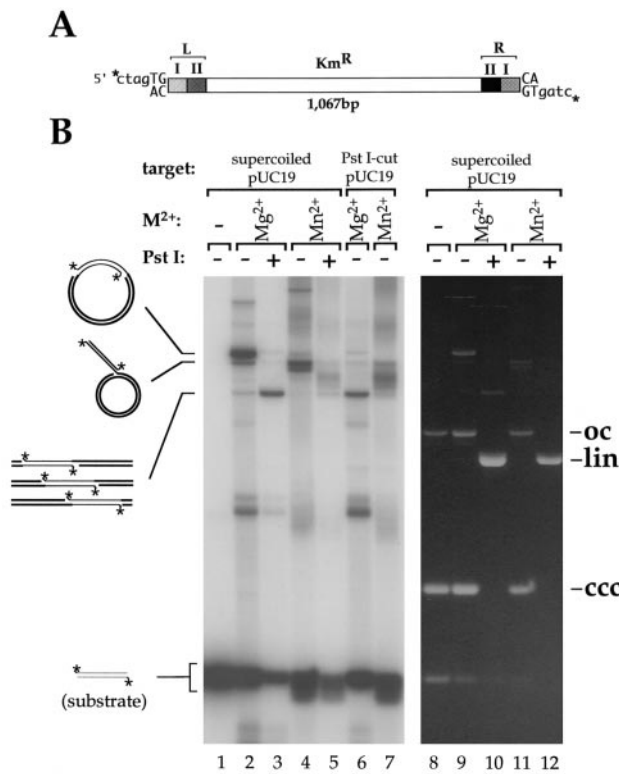

Fig. 1. Strand transfer reactions with Tn*552kan*. (*A*) Tn*552kan*: the shaded boxes represent the 23-bp binding sites at the transposon ends. The asterisks indicate the position of the ^{32}P label. (*B*) Transposition reactions (3 hr) contained 270 ng of supercoiled or *Pst*I-linearized pUC19 as target, about 50 fmol of a mixture of unlabeled and ^{32}P-labeled Tn*552kan* as the substrate, 120 ng transposase, and 10 mM MgCl$_2$ or MnCl$_2$ as indicated. Reactions run in lanes 3, 5, 10, and 12 were digested with *Pst*I. Reaction products were resolved by agarose gel electrophoresis and detected by both ethidium bromide staining (*Right*) and autoradiography (*Left*). The predicted structures of the reaction products are shown; thin lines and thick lines represent the substrate and target DNAs, respectively. oc, open circle form of pUC19; lin, linear pUC19; ccc, covalently closed circular form of pUC19.

Fig. 2. Tn*552kan* insertion sites in pUC19. Sequences of the flanking repeats. The positions in pUC19, shown on the right, indicate the first nucleotide of the target repeat.
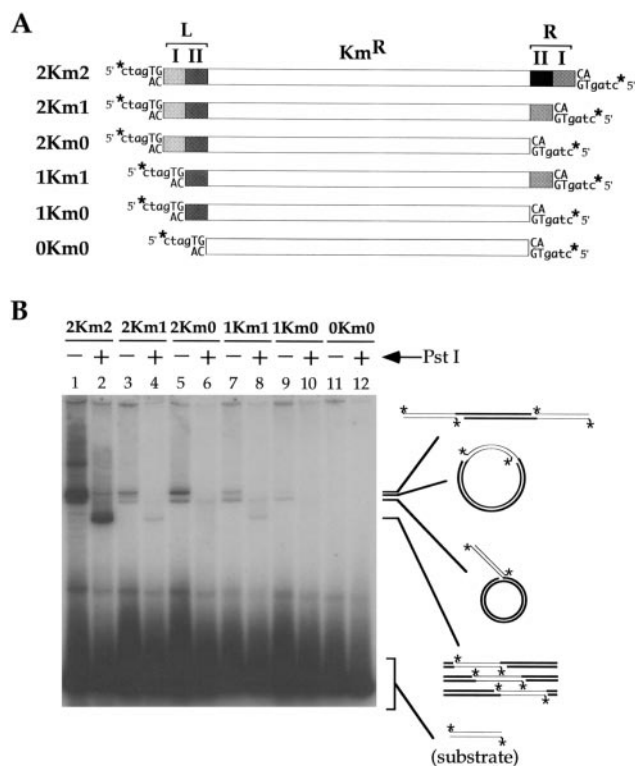
FIG. 3. Requirement for the transposase binding sites in transposition. (*A*) Structure of the 2Km2–0Km0 series of Tn*552kan* elements (details are identical to those in Fig. 2). (*B*) Transposition reactions (2 hr) contained 10 mM Mg^{2+}, 225 ng of supercoiled pUC19, and about 18 fmol of ^{32}P-labeled substrate as indicated. Samples in even-numbered lanes were digested with *Pst*I. Product analysis was as before.

ends flanked by direct repeats of pUC19 target sequences (Fig. 2). Interestingly, the size of the target duplication was variable, ranging between 6 and 9 bp (Fig. 2). The analyzed insertions did not appear to be clustered in any particular region of pUC19 (Fig. 2).

All Four 23-bp Repeats Are Required for Efficient Concerted Strand Transfer. To examine the role of the two 23-bp repeats located at each Tn*552* end, we constructed derivatives of Tn*552kan* in which these repeats were progressively removed from one or both ends (Fig. 3*A*). Deletion of a single (inner) 23-bp repeat from just one end severely impaired the ability of transposase to use both transposon ends in concerted strand transfer; however, such concerted strand transfer remained detectable not only with 2Km1 but also with 1Km1 (Fig. 3*B*; compare the yields in lanes 2, 4, and 8 of linear products obtained after digestion with *Pst*I). For those transposons that have just one complete end (2Km1 or 2Km0), the most frequent event is coupled insertion of two different donor molecules yielding the 4.7-kb linear product (the upper band of the triplet; Fig. 3*B*, lanes 3 and 5). For those substrates in which the least defective end contains just a single (outer) 23-bp repeat (1Km1 or 1Km0), the preferred (though very inefficient) event is the uncoupled insertion of a single end (Fig. 3*B*, lanes 7 and 9).

Product Yield and Distribution Are Sensitive to Reaction Conditions. The activities of transposase are affected by glycerol and NaCl concentration (Fig. 4). Strand transfer activity under otherwise standard conditions (i.e., at 150 mM NaCl) required a high level of glycerol. Strand transfer products were undetectable at 25% glycerol or less but were obtained at high efficiency over a broad range of concentrations from 50 to 75% (Fig. 4, lanes 1–5). At a high glycerol concentration, concerted insertions were stimulated by NaCl, achieving a maximum efficiency at 100–200 mM (Fig. 4, lanes

6–11). By contrast, the level of single-ended insertions (and of most higher-molecular-weight products) decreased dramatically as NaCl concentration was increased from 25 to 150 mM.

***In Vitro* Strand Transfer with Oligonucleotide Substrates.** A 50-bp "preprocessed" oligonucleotide substrate, RI&II, with the tandem transposase-binding sites is an efficient minimal substrate for concerted strand transfer. Coupled insertion of two oligonucleotide substrates at a single site should linearize the target plasmid. As shown in Fig. 5, a reaction with the labeled RI&II substrate in the presence of Mg^{2+} and high NaCl concentration gave high yields of a single product that migrates just behind linears of the pUC19 target DNA; this product also is visible in the ethidium-stained gel. We infer this to be the linear product of concerted insertion. Digestion with *Pst*I (which cuts the target once) resulted in a heterogeneous smear of labeled DNA running ahead of linear pUC19, both confirming our conclusion and indicating an absence of any strongly preferred target sites. [Similar results were obtained using a left-end substrate (LI&II) containing both 23-bp repeats (data not shown), indicating that left and right ends are equally reactive.] The high yield of concerted insertions required an elevated NaCl concentration (compare lanes 3 and 7) and the presence of both transposase-binding sites; a 25-bp substrate, RI, with just a single transposase-binding site (and a preprocessed end) gave no products under standard conditions (lanes 1 and 2).

Replacing Mg^{2+} with Mn^{2+} had two main effects. As expected from the results with Tn*552kan*, it reduced the coupling of strand transfer events. Open circular products, indicative of single-end insertions, were obtained in significant amounts, even with the RI&II substrate at a high NaCl concentration (Fig. 5, compare lanes 3 and 12). Mn^{2+} also made the transposase permissive for the suboptimal 25-bp RI substrate, particularly at low NaCl concentration, resulting in the formation of substantial amounts of single-end insertions (Fig. 5, lanes 10, 11, 14, and 15).

DISCUSSION

We have developed a protocol for the expression of Tn*552* transposase and its purification from the soluble fraction,
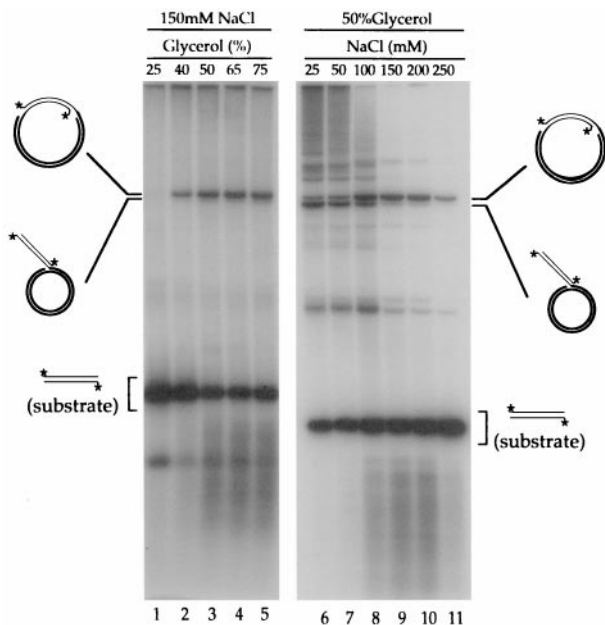

FIG. 4. Effects of glycerol and NaCl on strand transfer activity. Reactions (1 hr) contained 180 ng (*Left*) or 270 ng (*Right*) of supercoiled pUC19 as target, about 50 fmol of ^{32}P-labeled 2Km2 as substrate, and 10 mM Mg^{2+}. Concentrations of NaCl and glycerol were as indicated. Product analysis was as before.
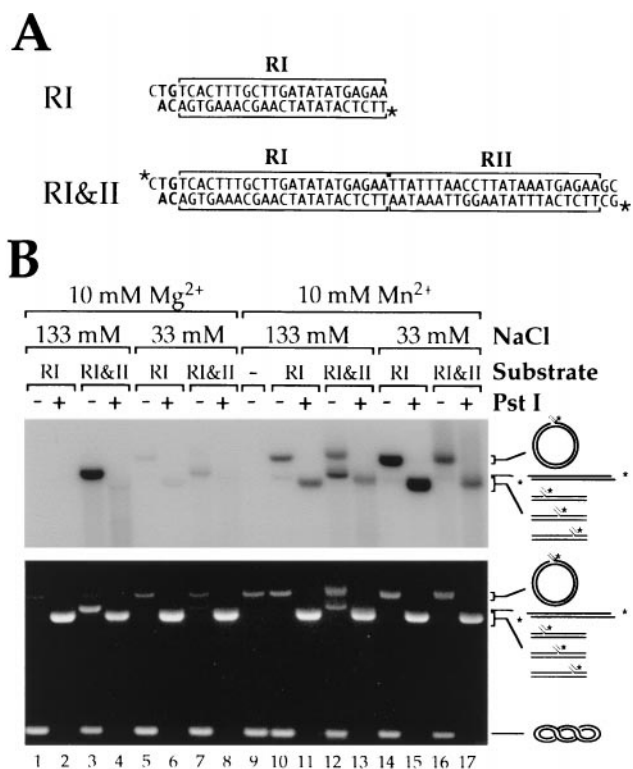
A

B

FIG. 5.   Strand transfer reactions with oligonucleotide substrates. (*A*) The RI and RI&II substrates. The 23-bp transposase-binding sites are indicated with horizontal square brackets. Asterisks indicate the position of the ^{32}P label. (*B*) Strand transfer reactions contained 270 ng of supercoiled pUC19 as target, about 2 pmol of a mixture of unlabeled and ^{32}P-labeled RI or RI&II, 120 ng of TnpA, either 133 mM or 33 mM NaCl (lanes 5–8 and 14–17), and 10 mM MgCl$_2$ or MnCl$_2$. Reaction products were resolved by agarose gel electrophoresis and detected by both ethidium bromide staining (*Lower*) and autoradiography (*Upper*).

avoiding the use of any denaturing steps. The transposase purified in this way efficiently catalyzes concerted insertions into an added target DNA, using substrate DNA with preprocessed ends, indicating that the protein is competent at both synapsis and strand transfer. The reaction requires only substrate and target DNAs, transposase protein, buffer, Mg^{2+}, and glycerol. The substrate DNA may be a transposon analogue with two complete 48-bp Tn552 ends or a short oligonucleotide with a single 48-bp end; in either case, coupled insertions of two ends are obtained. The observed activity differs significantly from that of a previous preparation of the Tn552 transposase that was limited to single-end insertions in a Mn^{2+}-dependent reaction (11). Similarly efficient synapsis and strand transfer reactions with short oligonucleotide substrates have been observed with Mu and Tn10 transposases (13, 14).

The concerted insertions obtained *in vitro* show all the hallmarks of normal transposition. When recovered from transformed *E. coli* cells and characterized, all insertions are accompanied by short target duplications, indicating that the target attack *in vitro* occurs with the expected staggered breaks. These target duplications are somewhat variable in size, ranging from 6 to 9 bp. However, variability may also be a feature of Tn552 transposition *in vivo*; of just three transpositions characterized, two were flanked by 6-bp repeats (1, 15) whereas the third had a 7-bp duplication (16). Within the relatively small sample analyzed, we saw no site specificity or clustering of insertion sites, suggesting that concerted strand transfer is rather nonspecific regarding target sequences. In fact, analysis of whole reactions at the molecular level has

shown strand transfer to occur at virtually every position in the target DNA (unpublished data). These characteristics, coupled with the efficiency with which the reaction is catalyzed, make Tn552 transposase a potentially useful biotechnological tool.

Strand transfer activity of our transposase preparation is supported by either Mg^{2+} or Mn^{2+} but the outcome is influenced by the choice of metal. Under otherwise standard conditions, Mg^{2+}, which is expected to be the relevant cation *in vivo*, leads to a much cleaner reaction profile with a higher yield of concerted strand transfer products and an overall reduction in single-ended events. In the presence of Mn^{2+}, the total strand transfer is increased (see, for example, Fig. 5) but most products appear to result from uncoupled single-ended events and the yield of coupled events drops. Increased permissiveness in the presence of Mn^{2+} has been observed in several other systems, for example, retroviral integrases (17–19), Mu transposase (20), Tn10 transposase (21), and DNA polymerases (22, 23)—but the exact nature of this effect is not understood.

The concentrations of both glycerol and NaCl have a pronounced effect on strand transfer activity. Our protein requires unusually high levels of glycerol: more than 25% was necessary to detect reaction products at all, and their accumulation peaked only when the reactions were carried out in the presence of 50–75% glycerol (see Fig. 4). Glycerol stabilizes protein structure through the "solvophobic" effect; its favorable interaction with the bulk water solvent makes the contact between hydrophobic residues on the protein and water less unfavorable than those with the glycerol solution, effectively driving glycerol away from the protein surface. This preferential hydration of the protein surface is a driving force for minimizing its area (24). The extremely high concentrations of glycerol required in our reactions might be stabilizing the structure, preventing partial denaturation and aggregation, or inducing a conformational change that would normally be driven by other factor(s) *in vivo*.

The most interesting effect of NaCl is the selective and increasing inhibition of single-ended events and concomitant stimulation of concerted insertions as the concentration is raised from 50 to 150 mM. Possible explanations can be divided into two general classes. Full, concerted transposition normally would occur within a synaptic complex consisting of a transpososome (the two transposon ends bound and brought together by transposase—most likely four transposase protomers in the case of Tn552) plus target DNA. Within this complete complex, strand transfer events generally are concerted either because they are kinetically coupled (i.e., the rates of catalysis are faster than the rate of dissociation) or because intersubunit communication ensures that both catalytic events are primed to occur virtually simultaneously. The high incidence of single-ended events at low NaCl concentration could result either from uncoupling of the catalytic events within complete synaptic complexes or from dispensing with the requirement for a complete synaptic complex and allowing aberrant transposase-end–target complexes (lacking an end or some subunits, or incompletely engaging the target) to initiate strand transfer. High NaCl concentration may both inactivate or dissociate aberrant complexes and promote the formation of transpososomes or transpososome–target complexes.

Efficient concerted strand transfer requires both transposase-binding sites contained within the 48-bp terminal DNA segment. The reactions performed with oligonucleotide substrates (Fig. 5 and data not shown) suggest that under our standard conditions, the right and left ends are fully interchangeable. This is in contrast to what was observed in the analogous reaction with Mu transposase where two right ends proved to be far more efficient substrates (13). The requirement for a pair of binding sites at each end suggests that a

Tn*552* transpososome contains four protomers of the transposase (assuming each site binds one transposase monomer).

We have been unable, so far, to detect 3′ end processing activity by Tn*552* transposase under our current assay conditions. This was somewhat unexpected given the efficiency with which the protein catalyzes strand transfer with preprocessed substrates. One possible explanation is that TnpB is required for this activity. Although MuB is not required for formation of the Mu transpososome and cleavage at the Mu 3′ ends (25), cleavage at the ends of Tn7 requires the TnsC protein (26). In the absence of the terminal cleavages, TnpB may be required to stabilize the transpososome or to activate the endonuclease function of the transposase.

Tn*552* can now be added to a short list of transposable elements for which efficient concerted strand transfer is achievable *in vitro* (6, 14, 26–30). Two of the best-characterized systems are for other, distantly related members of the D,D(35)E superfamily—Mu and Tn7 (for reviews see refs. 5–7). For both of these systems the protein and substrate DNA requirements are complex. Transposition involves both the transposase (MuA or the heteromeric TnsAB of Tn7) and an ATP-dependent, nonspecific DNA-binding protein (MuB or Tn7 TnsC) that activates the transposase and plays a role in target presentation. In addition, Tn7 transposition requires one of two other proteins, either TnsD or TnsE, that determine alternative pathways of target site selection. The ends of Mu each contain three binding sites for its transposase; these are distributed irregularly over 175 bp at the left end and 80 bp at the right. In addition, Mu contains an internal sequence, the IAS, which is bound by the extreme N-terminal domain of MuA and is important for assembly of the transpososome but not for its subsequent activity. Tn7 has seven binding sites for TnsB: three, spanning 150 bp at the left end, and four, covering 90 bp at the right; again, the spacing between sites varies. Tn*552* appears to provide a simpler transposition system. Not only is its transposase considerably smaller but also its natural substrate is less complex; the ends are of identical structure, containing just the two abutting transposase-binding sites, and the transposon presumably has no equivalent to the IAS.

Mizuuchi and colleagues (13, 31) have found reaction conditions that simplify dramatically the overall requirements for Mu transposition. In the presence of DMSO (as well as glycerol), dependence on MuB, the IAS, substrate superhelicity, and even the first 76 residues of MuA (the IAS-binding domain) is lost. Under the modified conditions, a DNA substrate with just the terminal 50 bp of the Mu right end and a few flanking base pairs is assembled efficiently into a transpososome, cleaved at the transposon end, and inserted into target DNA. It may well be significant that the active minimal substrate is organized similarly to the Tn*552* substrate, with a pair of adjacent transposase-binding sites. This arrangement, found at only one end of the wild-type Mu and Tn7 transposons, probably provides an optimal scaffold for assembly of the Mu transpososome, which is known to contain a tetramer of the transposase (32). By adopting this arrangement at both ends, Tn*552* may have obviated the necessity for the structural complexities found in Mu and Tn7. The reaction we have described here with Tn*552* transposase is very similar to this simplified Mu reaction; indeed its lack of dependence on TnpB (the MuB analogue) may be attributable to the somewhat unusual reaction conditions or to the use of pre-processed transposon ends. Its overall simplicity (particularly its independence of TnpB) makes Tn*552* transposase attractive for structural studies.

1. Rowland, S. J. & Dyke, K. G. (1989) *EMBO J.* **8,** 2761–2773.
2. Murphy, E. & Novick, R. P. (1979) *Mol. Gen. Genet.* **175,** 19–30.
3. Rowland, S. J. & Dyke, K. G. (1990) *Mol. Microbiol.* **4,** 961–975.
4. Polard, P. & Chandler, M. (1995) *Mol. Microbiol.* **15,** 13–23.
5. Craig, N. L. (1996) *Curr. Top. Microbiol. Immunol.* **204,** 27–48.
6. Mizuuchi, K. (1992) *Annu. Rev. Biochem.* **61,** 1011–1051.
7. Lavoie, B. D. & Chaconas, G. (1996) *Curr. Top. Microbiol. Immunol.* **204,** 83–102.
8. Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. & Davies, D. R. (1994) *Science* **266,** 1981–1986.
9. Rice, P. & Mizuuchi, K. (1995) *Cell* **82,** 209–220.
10. Bujacz, G., Jaskolski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R.A. & Skalka, A. M. (1995) *J. Mol. Biol.* **253,** 333–346.
11. Rowland, S. J., Sherratt, D. J., Stark, W. M. & Boocock, M. R. (1995) *EMBO J.* **14,** 196–205.
12. Gill, S. C. & von Hippel, P. H. (1989) *Anal. Biochem.* **182,** 319–326.
13. Savilahti, H., Rice, P. A. & Mizuuchi, K. (1995) *EMBO J.* **14,** 4893–4903.
14. Sakai, J., Chalmers, R. M. & Kleckner, N. (1995) *EMBO J.* **14,** 4374–4383.
15. Rice, L. B., Carias, L. L., Marshall, S. H. & Bonafede, M. E. (1996) *Plasmid* **35,** 81–90.
16. Paulsen, I. T., Gillespie, M. T., Littlejohn, T. G., Hanvivatvong, O., Rowland, S. J., Dyke, K. G. & Skurray, R. A. (1994) *Gene* **141,** 109–114.
17. Kulkosky, J., Jones, K. S., Katz, R. A., Mack, J. P. & Skalka, A. M. (1992) *Mol. Cell. Biol.* **12,** 2331–2338.
18. Drelich, M., Wilhelm, R. & Mous, J. (1992) *Virology* **188,** 459–468.
19. Katzman, M., Katz, R. A., Skalka, A. M. & Leis, J. (1989) *J. Virol.* **63,** 5319–5327.
20. Kim, K., Namgoong, S. Y., Jayaram, M. & Harshey, R. M. (1995) *J. Biol. Chem.* **270,** 1472–1479.
21. Junop, M. S. & Haniford, D. B. (1996) *EMBO J.* **15,** 2547–2555.
22. Beckman, R. A., Mildvan, A. S. & Loeb, L. A. (1985) *Biochemistry* **24,** 5810–5817.
23. Van de Sande, J. H., Loewen, P. C. & Khorana, H. G. (1972) *J. Biol. Chem.* **247,** 6140–6148.
24. Timasheff, S. N. (1993) *Annu. Rev. Biophys. Biomol. Struct.* **22,** 67–97.
25. Surette, M. G., Buch, S. J. & Chaconas, G. (1987) *Cell* **49,** 253–262.
26. Bainton, R., Gamas, P. & Craig, N. L. (1991) *Cell* **65,** 805–816.
27. Chalmers, R. M. & Kleckner, N. (1994) *J. Biol. Chem.* **269,** 8029–8035.
28. Lampe, D. J., Churchill, M. E. & Robertson, H. M. (1996) *EMBO J.* **15,** 5470–5479.
29. Goryshin, I. Y. & Reznikoff, W. S. (1998) *J. Biol. Chem.* **273,** 7367–7374.
30. Ton-Hoang, B., Polard, P. & Chandler, M. (1998) *EMBO J.* **17,** 1169–1181.
31. Craigie, R. & Mizuuchi, K. (1987) *Cell* **51,** 493–501.
32. Lavoie, B. D., Chan, B. S., Allison, R. G. & Chaconas, G. (1991) *EMBO J.* **10,** 3051–3059.