PubMed Central

HHMI
HOWARD HUGHES MEDICAL INSTITUTE

AUTHOR MANUSCRIPT
Accepted for publication in a peer-reviewed journal

Published as: *Nat Struct Mol Biol*. 2012 August ; 19(8): 834–836.

Functional Characterization of an Active Rag-like Transposase

Cary G. Hencken[1,2], **Xianghong Li**[1], and **Nancy L. Craig, Ph.D.**[1]

[1]Howard Hughes Medical Institute and Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland USA.

Abstract

The formation of diverse immunoglobulin genes results in part from Rag protein-mediated DNA double strand breaks at the edges of immunoglobulin gene segments, followed by the combinatorial reassembly of these segments. We report that a *Transib* transposase from the insect *Helicoverpa zea* is active *in vitro* and that its breakage and joining activities mimic those of Rag, providing strong evidence that Rag and *Transib* transposases were derived from a common progenitor.

The assembly of an immunoglobulin gene mimics transposon excision, followed by joining of the flanking donor sites to form the coding gene segments[1] (Supplementary Figure 1). This excised intervening segment from an immunoglobulin gene can also mimic transposon insertion *in vitro* by inserting into a target DNA[2,3]. Sequence similarities between the Rag1 subunit of Rag protein that mediates immunoglobulin breakage and joining and the transposases of the fossil *Transib* transposon superfamily suggested that the core region of Rag1 was derived from a *Transib* transposon[4]. We have characterized an active *Transib* transposase and have found that the transposase from the *Transib* transposon called *Hztransib* of *Helicoverpa zea* [5], is active *in vitro*. Notably, the mechanism of its breakage and joining activities are like those of Rag, providing strong evidence that Rag1 and *Transibs* were derived from a common progenitor.

Hztransib was found in the insect cotton and corn bollworm pest *Helicoverpa zea*[5], by genome sequencing of the region around a Cytochrome P450 gene of interest. The fact that there is variation in its genomic position and copy number in different populations suggested that it might be a still active element[6]. To analyze its activity directly, we isolated the *Hztransib* transposase by tagging the ORF with a hexahistidine tag, expressing it in *E. coli* and affinity purifying the protein.

To evaluate *Hztransib*'s ability to excise the transposon end from flanking DNA and then join it to a target, we used a DNA fragment containing a 51 bp *Hztransib* end sequence and 37 bp of flanking DNA in which the interior 5′ end of the transposon was labeled as a substrate in reaction mixtures using a plasmid as target DNA and then displayed the products on a native agarose gel (Figure 1). Two products are visible in these reactions, one in which a single transposon end oligonucleotide joins to the plasmid target DNA, giving a nicked circle product (Single End Join = SEJ); alternatively, if two end oligonucleotides join

Correspondence should be addressed to Nancy L. Craig, Ph.D. Howard Hughes Medical Institute Department of Molecular Biology & Genetics 502 PCTB 725 N Wolfe St Johns Hopkins University School of Medicine Baltimore MD 21205 ncraig@jhmi.edu voice = 410-955-3933 fax = 443-287-7798.
[2]Present address: Department of Cell Biology, Harvard Medical School, Boston, Massachusetts, USA.

AUTHOR CONTRIBUTIONS
NLC conceived the project; CGH, XL and NLC designed the experiments; CGH and XL carried out the experiments; CGH, XL and NLC analyzed the data and wrote the paper.

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript HHMI Author Manuscript HHMI Author Manuscript

at the same target position, the plasmid is linearized (Double End Join = DEJ) (Figure 1, lane 2).

Mutational analysis of many transposases, integrases and the Rag1 subunit of Rag has identified three conserved acidic amino acids essential for catalytic activity that lie on an RNase H-like fold that closely juxtaposes them to bind essential divalent metal, usually Mg^{2+}, ions at the catalytic center[7,8]. These positions are conserved in *Transibs*[4]. *Hztransib* residues Asp126, Asp225, and Glu436 align with the cognate amino acids in the Rag1 subunit (Supplementary Figure 2) and mutation of these conserved amino acids in *Hztransib* significantly decreases the excision and target joining reaction (Figure 1, lanes 3-5). Such active site mutations also block Rag-mediated DNA end cleavage and target joining[8].

The double strand breaks at the RSSs promoted by Rag involve both a hydrolysis reaction and a direct transesterification that occur by mechanisms used by other transposases and integrases[7].

Rag DNA breakage initiates with a nick that frees the 5′ end of the RSS from the flanking coding DNA segment[1]. The newly exposed coding segment 3′ OH then attacks its complementary strand, generating a hairpin on the coding DNA and introducing the double strand break that frees the RSS-bounded segment. Cellular enzymes open the hairpins and join the coding segments to form the immunoglobulin genes[1].

To analyze the mechanism by which *Hztransib* makes the double strand break at the transposon end to separate it from the flanking DNA, we used a DNA substrate fragment containing a 51 bp *Hztransib* end sequence and 37 bp of flanking DNA in which the interior 5′ end of the transposon and the flanking DNA was labeled and displayed the reaction products on a denaturing acrylamide gel (Figure 2a). The 37 nt species is the donor top strand fragment resulting from nicking at the junction of the flanking DNA with the top transposon strand (Figure 2b, lane 2). This nicking position is identical to that of Rag and distinct from that of *hAT* transposases that also excise via hairpin formation, which are initiated by a nick one nucleotide into the flanking donor DNA[9]. Present at lower amount is the 51 nt, 3′ OH transposon bottom strand fragment that results from the double strand break that separates the transposon end from the flanking donor DNA. We also assayed the *Hztransib* transposases with active site Asp126, Asp225, and Glu436 mutations, finding that nicking is not detectable (Figure 2b lanes 3,4) or is reduced (lane 5). We suspect that the residual nicking activity with the *Hztransib* mutant E436A results from the use of Mn^{2+} rather than Mg^{2+} as a co-factor[10](see also below). No recombination was observed using Mg^{2+} as a cofactor (data not shown).

With Rag, cleavage of the bottom strand occurs when the 3′ OH on the newly released donor top strand attacks its complementary strand, forming a hairpin on the flanking DNA. Such a hairpin is faintly visible in an incubation using WT protein and an intact oligonucleotide as a substrate (Figure 2b, lane 2). The hairpin is much more prominent, however, when a substrate “pre-nicked” at the 5′ end of the transposon is used as a substrate (Figure 2b, lane 7). A series of species is evident, the most slowly migrating of which co-migrates with an authentic 74 nt hairpin, indicating that the *Hztransib* transposase generates a double strand break via a hairpin. The ladder of species below the authentic hairpin may reflect degradation at the 3′ end of the hairpin or imprecise hairpinning; the use of the recessed 3′ OH in the 32 nt product may contribute to such imprecise hairpinning. Such species are not evident with a substrate that is pre-nicked to expose the 3′ OH end of the transposon (data not shown).

As is also true with Rag containing mutations in Rag1[7,8], reduced hairpinning is observed using the *Hztransib* transposase active site mutants D126A, D225A, and E436A (Figure 2b,

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

lanes 8-10). We suggest that the residual nicking activity with the *Hztransib* mutant E345A results from the use of Mn^{2+} as a co-factor[10].

We also find that *Hztransib* can promote target joining *in vitro* to form SEJs and DEJs using a "pre-cleaved" transposon end substrate in which the terminal 3′ OH is already exposed (Figure 3a). We observe about the same amount of SEJ and DEJ products using a 50 bp *Hztransib* end as with a 294 bp *Hztransib* ends (data not shown). As occurs with Rag1 mutants [8], less target joining is observed using the active site mutants D126A, D225A, and E436A *Hztransib* transposases (Figure 3b). We suggest that the residual target joining activity with the *Hztransib* mutants results from the use of high amounts of Mn^{2+} as a co-factor[10] (Supplemental Figure 3). When a reaction using a transposon fragment labeled at its interior 5′ end is displayed on a denaturing gel (Figure 3c), a product the length plasmid plus the transposon end is observed, consistent with joining of the 3′ OH end to the target DNA.

Finally, to further analyze the structure the products formed by the joining of *Hztransib* ends to a target DNA, we constructed a mini*Hztransib* transposon in which 50 bp *Hztransib* ends flanked a KanR gene fragment, incubated it with a target plasmid and transposase, and recovered insertion products by selection for KanR following transformation in *E. coli*. DNA sequence analysis of 44 insertions at different sites revealed that all had perfect 5 bp target site duplications that were enriched with 5′ CGnCG as occurs with RAG[1,3] (Supplementary Figure 4). Transposition is reduced more than 20-fold when the *Hztransib* transposases mutated at the conserved RNase H residues D126A, D225A, and E436A are used in the transformation transposition assay (Supplementary Figure 4).

These experiments have revealed that a *Transib* transposase uses the same mechanism for the same breakage and joining reactions as does the immunoglobulin gene recombinase Rag. The nick that initiates the double strand break occurs exactly at the junction of the transposon end and the flanking donor DNA, the subsequent double strand break occurs by hairpinning of the exposed 3′ OH donor end to its complementary strand, and the released 3′ OH end of the transposon directly attacks the target DNA. Moreover, mutation of a conserved catalytic triad of amino acids in both Rag1 and *Hztransib* transposase block breakage and joining. This functional analysis provides strong evidence that Rag and *Transib* transposases are descended from a common progenitor.

ONLINE METHODS

Cell Growth and Transposase Purification

An *E. coli* codon-optimized version of the Hztransib transposase gene was synthesized by DNA2.0 and cloned into the NcoI-EcoRI sites of the pBAD Myc/His (Invitrogen) vector and expressed in Top10 cells. One liter of cells was grown in LB + 100 ug/mL carbenecillin until about OD600 0.7, arabinose was added to 0.01% and cell growth continued overnight at 16°C. The cells were chilled, harvested by centrifugation and then resuspended in 10 ml Column Buffer (5 mM imidazole, 500 mM NaCl, 20 mM Tris (pH 7.8), 10 % glycerol). The cells were lysed in a French press, the lysate spun at 36,000 rpm for 70 min, the resulting supernatant filtered through a 0.45 um syringe filter and then diluted 1:1 with column buffer. The lysate was then loaded onto a 1 ml Nickel Sepharose, washed with 15 ml Column buffer, washed with Wash buffer (40 mM imidazole, 500 mM NaCl, 20 mM Tris (pH 7.8), 10 % glycerol), transposase eluted with Elution buffer (150 mM imidazole, 500 mM NaCl, 20 mM Tris (pH 7.8), 10 % glycerol) and dialyzed in Storage buffer (500 mM NaCl, 1 mM EDTA, 2 mM DTT, 20 mM Tris (pH 7.8), 25 % glycerol). The final protein concentration was 1 mg/ml and *Hztransib* transposase was about 5-10 % of the total protein.

HHMI Author Manuscript
HHMI Author Manuscript
HHMI Author Manuscript

Hztransib transposase mutants, D126A, D225A and E436A were generated using the Quick Change Site Directed Mutagenesis Kit (Stratagene) and the pBAD-*Hztransib* plasmid using the following oligonucleotides: Hz D1mut-Forward 5′CTAAATGGGGTTTCGCCGGCGCATCGAATCAG and Hz D1mut-Reverse 5′CTGATTCGATGCGCCGGCGAAACCCCATTTAG; Hz D2mut-For 5′ GATGACCATGATCGCCGGTAAGATCTGTACC and Hz D2mut-Rev 5′ GGTACAGATCTTACCGGCGATCATGGTCATC; and Hz Emut-For 5′GAAGAGGCGAGCGCGGCGCGCAACAAG and Hz Emut-Rev 5′CTTGTTGCGCGCCGCGCTCGCCTCTTC.

Transposon end fragments

DNA fragments and oligonucleotides were end-labeled using [γ-^{32}P] ATP and T4 polynucleotide kinase. A 294 bp *Hztransib* L end was synthesized by DNA2.0 and cloned into their pJ201 vector.

Short double-stranded DNA segments were assembled by incubating oligonucleotides together at 75°C, followed by cooling for 4 hours.

The flanking DNA + transposon end substrate used in Figure 1 was assembled using oligonucleotides CGH10 GCTAGTAACGGCCGCCAGTGTGCTAAGTGCAACGTCGCACGGTGGATCGAAAAT C GGCTCTAGAAGACATAGGATCTCGATGTCTCAA and 5′ end labeled CGH11 5′ TTGAGACATCGAGATCCTATGTCTTCTAGAGCCGATTTTCGATCCACCGTGCGA CGTTGCACTTAGCACACTGGCGGCCGTTACTAGC.

The flanking DNA + transposon end substrate used in Figure 2 lanes 1-5 was assembled using 5′ end labeled oligonucleotides CGH10 GCTAGTAACGGCCGCCAGTGTGCTAAGTGCAACGTCGCACGGTGGATCGAAAAT C GGCTCTAGAAGACATAGGATCTCGATGTCTCAA and CGH11 5′ TTGAGACATCGAGATCCTATGTCTTCTAGAGCCGATTTTCGATCCACCGTGCGA CGTTGCACTTAGCACACTGGCGGCCGTTACTAGC. The pre-nicked flanking DNA + transposon end substrate used in Figure 2 lanes 6-10 was assembled from oligonucleotides 5′ end-labeled CGH12 5′ GCTAGTAACGGCCGCCAGTGTGCTAAGTGCAACGTCG, CGH2 5′ CACGGTGGATCGAAAATCGGCTCTAGAAGACATAGGATCTCGATGTCTCAA and CGH11 5′ TTGAGACATCGAGATCCTATGTCTTCTAGAGCCGATTTTCGATCCACCGTGCGA CGTTGCACTTAGCACACTGGCGGCCGTTACTAGC.

The 294 bp *Hztransib* L end fragment used on Figure 3a and b was generated by PCR from the *Hztransib* L end in pJ201 by PCR amplification with the 5′ end labeled primers (*Hztransib* = **bold**) 5′ **CACGGTGGATCGAAAATCGGCT** and 5′ **CGCAAGTGATGCCGCAATATG**. The pre-cleaved 50 bp *Hztransib* end fragment used in Figure 3d was assembled from oligonucleotides 5′ CACGGTGGATCGAAAATCGGCTCTAGAAGACATAGGATCTCGATGTCTCAA and 5′ TTGAGACATCGAGATCCTATGTCTTCTAGAGCCGATTTTCGATCCACCGTG.

The *Hztransib* L50-KanR-R50 fragment used in Supplementary Figure 4 was generated by PCR using a Kan gene as template with the primers 5′ CACGGTGGATCGAAAATCGGCTCTAGAAGACATAGGATCTCGATGTCTCAACA CTAGTATTACCCTGTTATCC and 5′CACGGTGGATCGAAAATCGGCTGTAGAAGACATAGGATCTCGATGTCTTACA ACAAAGCCGCCGTCCCGTC.

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

In vitro transposition reactions

Reaction mixtures were 20 μl and included 0.2 pmol transposon end DNA, 0.3 pmol pUC19 as target DNA, and 3 μl of *Hztransib* protein fraction in Storage Buffer. They contained 8 mM $MnCl_2$, 2 mM DTT, 0.01% BSA and 25 mM HEPES (pH 8.0) and (from Storage Buffer) 75 mM NaCl, 0.15 mM EDTA, 0.3 mM DTT, 3.0 mM Tris (pH 7.8), 3.7 % glycerol. Reactions were incubated at 30°C for 1 hr (or as indicated) and then stopped by the addition of STOP buffer (10 mM EDTA, 10% SDS, 0.15% Bromophenol Blue, 0.25% Xylene Cyanol).

Electrophoresis

Native samples were displayed on 1% agarose gels in 1 × TAE. For alkaline agarose gels, *in vitro* transposition reactions were ethanol precipitated and resuspended in 50 mM NaOH and 1 mM EDTA following the addition of 6x loading dye (300 mM NaOH, 6 mM EDTA, 18% Ficoll, 0.15% Bromophenol Blue, 0.25% Xylene Cyanol). Denatured samples were displayed on a 1% agarose gel in 50 mM NaOH and 1 mM EDTA. For denaturing acrylamide gels, samples were phenol extracted, ethanol precipitated and resuspended in 12 μl 95% formaldehyde, 20 mM EDTA, 0.05% Bromophenol Blue, 0.05% Xylene Cyanol, boiled and then placed on ice. Samples were displayed on denaturing acylamide gels containing 8% EXPLORER acrylamide and 7M urea in 1 × TBE.

Transformation Assay for *Hztransib* integration

300 ng *Hztransib* L50-KanR-R50 fragment and 500 ng pUC19 were used as substrates in *in vitro* transposition reactions. The reaction mixtures were phenol extracted, ethanol precipitated and resuspended in 10 μl H_2O.5 μl was transformed into 25 μl Mach1 cells (Invitrogen), LB added to 500 μl and transposition products recovered by selection for kanamycin resistance.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the other members of the Craig lab for fruitful discussions, Helen McComas for her assistance with the figures and text, and Steve Desiderio for his comments on the manuscript. NLC is an Investigator of the Howard Hughes Medical Institute.

REFERENCES

1. Matthews AG, Oettinger MA. Nat Immunol. 2009; 10:817. [PubMed: 19621044]
2. Schatz DG, Swanson PC. Annu Rev Genet. 2011; 45:167. [PubMed: 21854230]
2. Hiom K, Melek M, Gellert M. Cell. 1998; 94:463. [PubMed: 9727489]
3. Agrawal A, Eastman QM, Schatz DG. Nature. 1998; 394:744. [PubMed: 9723614]
4. Kapitonov VV, Jurka J. PLoS Biol. 2005; 3:e181. [PubMed: 15898832]
5. Chen S, Li X. Gene. 2008; 408:51. [PubMed: 18031956]
6. Du E, Ni X, Zhao H, Li X. Insect Mol Biol. 2011; 20:291. [PubMed: 21166910]
7. Hickman AB, Chandler M, Dyda F. Crit Rev Biochem Mol Biol. 2010; 45:50. [PubMed: 20067338]
8. Kim DR, et al. Genes Dev. 1999; 13:3070. [PubMed: 10601033]
10. Landree MA, Wibbenmeyer JA, Roth DB. Genes Dev. 1999; 13:3059. [PubMed: 10601032]
11. Fugmann SD, Villey IJ, Ptaszek LM, Schatz DG. Mol Cell. 2000; 5:97. [PubMed: 10678172]
9. Zhou L, et al. Nature. 2004; 432:995. [PubMed: 15616554]
10. Yang W, Lee JY, Nowotny M. Mol Cell. 2006; 22:5. [PubMed: 16600865]

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

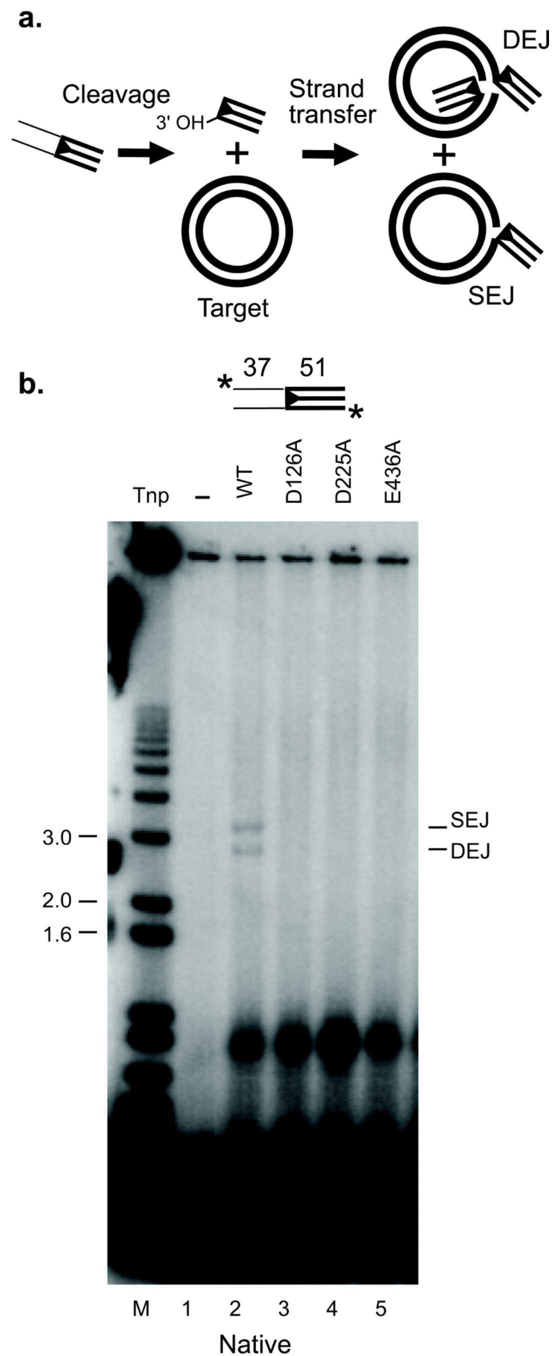

Figure 1. *Hztransib* transposase can promote coupled cleavage and strand transfer

a. The pathway of transposition. The transposase introduces a double strand break between a transposon end and the flanking donor DNA to expose the 3′ OH transposon end. The joining of one 3′ OH to one target strand a plasmid target generates a nicked circle product (Single End Join; SEJ). The joining of two 3′ OH ends to two target strands at complementary positions in a single target plasmid yields a linearized version of the plasmid attached to the two end oligonucleotides (Double End Join; DEJ).

b. Reactions contained wildtype or mutant *Hztransib* transposase as indicated. We suspect the lower transposase-dependent band likely results from end binding.

HHMI Author Manuscript
HHMI Author Manuscript
HHMI Author Manuscript

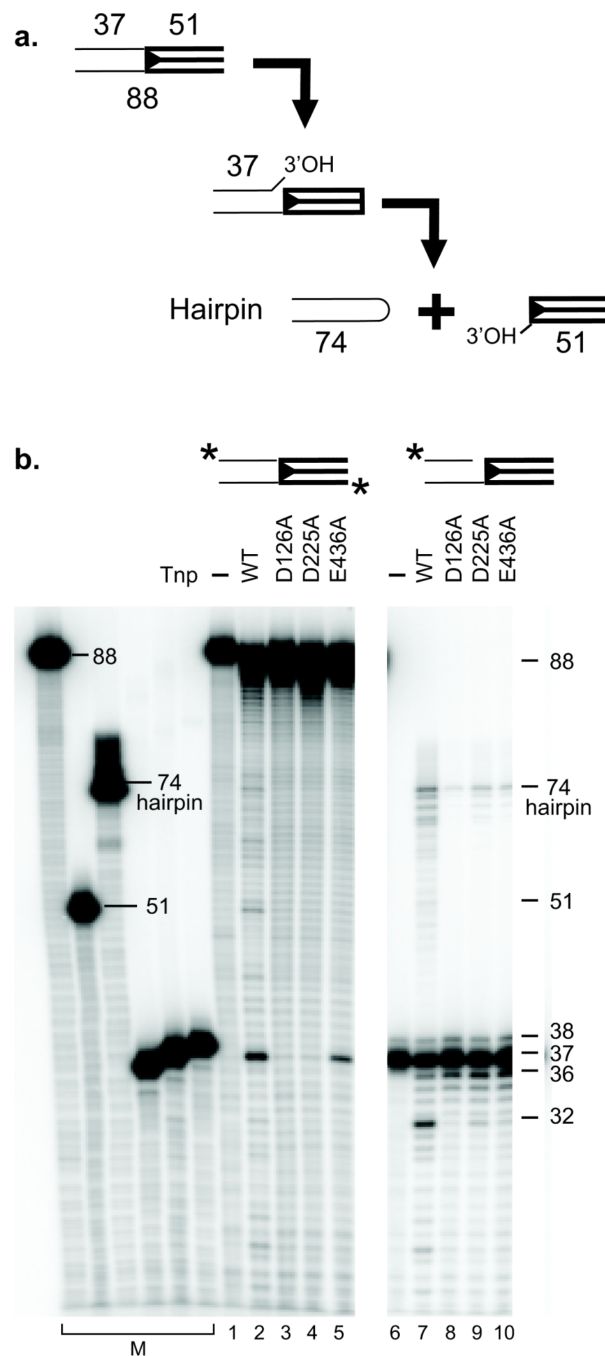

Figure 2. Analysis of the *Hztransib* DNA cleavage mechanism

a. The mechanism of transposition. Transposition begins with a nick at the 5′ end of the transposon end. The newly exposed donor 3′ OH then attacks the complementary strand to form a hairpin and release the transposon end.

b. *In vitro* transposition reactions were carried out using the transposon end fragments labeled as indicated, transposases as indicated and displayed on a denaturing acrylamide gel. The image is of a single gel with a masked lane.

HHMI Author Manuscript

HHMI Author Manuscript

HHMI Author Manuscript

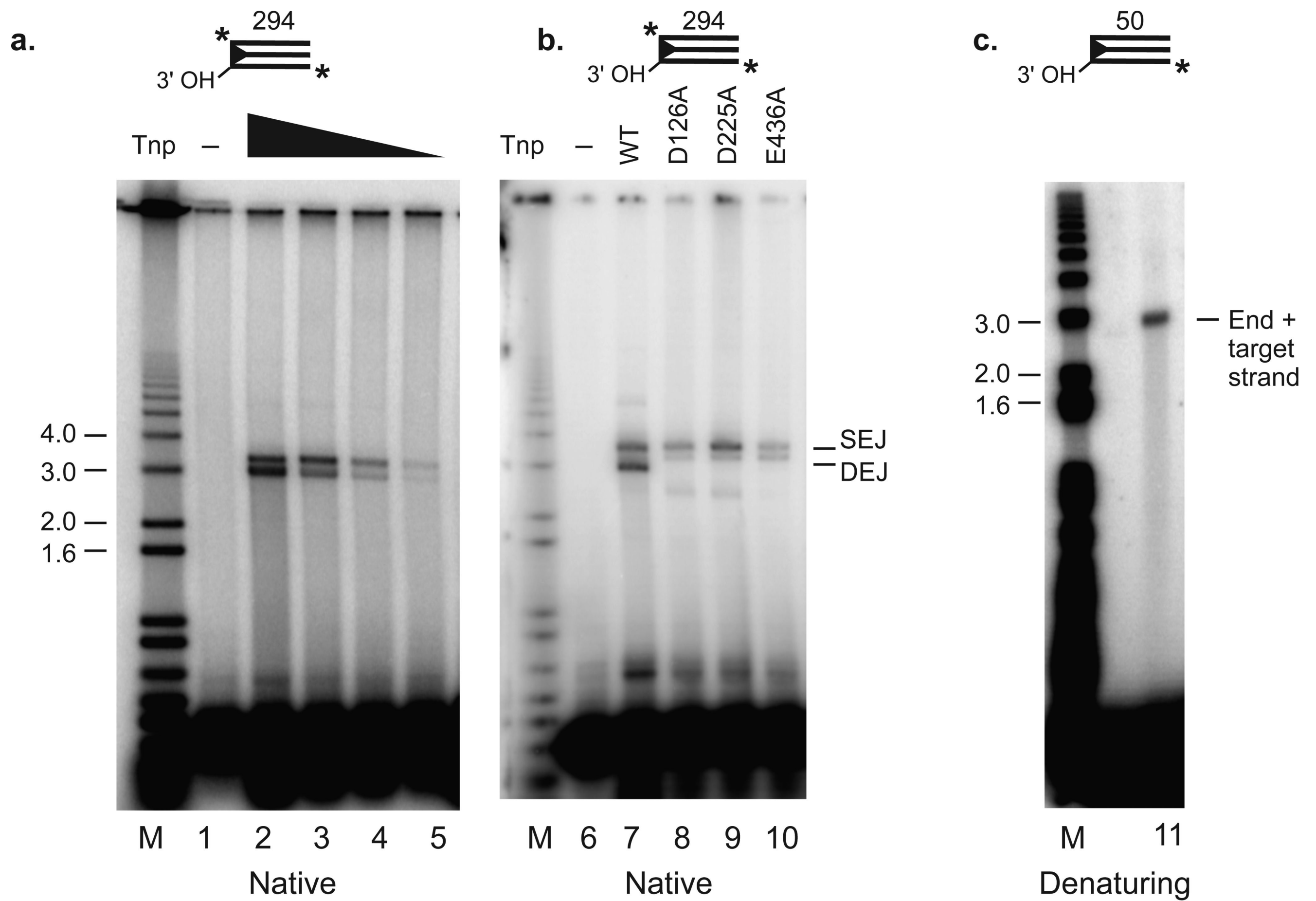

Figure 3. *Hztransib* Transposition *in vitro* using pre-cleaved transposon ends

a. Target joining by decreasing amounts of wildtype transposase. The reactions contained 1×, 1/3×, 1/9× and 1/27× transposase.

b. Target joining by transposase mutants. The transposon end substrate was a 294 bp *Hztransib* left end labeled at both 5′ ends (asterisk) and reactions contained wildtype or mutant *Hztransib* transposase as indicated. The reaction products were displayed on a native agarose gel.

c. The 3′ OH transposon end joins to target DNA. The transposon end substrate was a 50 bp *Hztransib* left end labeled only at the internal 5′ end* (asterisk) of the strand that joins covalently to target DNA.