Nucleic Acids Research, 1998, Vol. 26, No. 8 **1927–1933**

Simple and efficient generation *in vitro* of nested deletions and inversions: Tn*5* intramolecular transposition

Dona York, Kelly Welch, Igor Yu. Goryshin and William S. Reznikoff*

Department of Biochemistry, 420 Henry Mall, University of Wisconsin-Madison, Madison, WI 53706, USA

Received January 14, 1998; Revised and Accepted February 23, 1998

ABSTRACT

We have exploited the intramolecular transposition preference of the Tn*5* *in vitro* transposition system to test its effectiveness as a tool for generation of nested families of deletions and inversions. A synthetic transposon was constructed containing an *ori*, an ampicillin resistance (Amp^r) gene, a multi-cloning site (MCS) and two hyperactive end sequences. The donor DNA that adjoins the transposon contains a kanamycin resistance (Kan^r) gene. Any Amp^r replicating plasmid that has undergone a transposition event (Kan^s) will be targeted primarily to any insert in the MCS. Two different size targets were tested in the *in vitro* system. Synthetic transposon plasmids containing either target were incubated in the presence of purified transposase (Tnp) protein and transformed. Transposition frequencies (Amp^r/Kan^s) for both targets were found to be 30–50%, of which >95% occur within the target sequence, in an apparently random manner. By a conservative estimate 10^5 or more deletions/inversions within a given segment of DNA can be expected from a single one-step 20 μl transposition reaction. These nested deletions can be used for structure–function analysis of proteins and for sequence analysis. The inversions provide nested sequencing templates of the opposite strand from the deletions.

INTRODUCTION

A number of methods for the generation of nested families of deletions have been developed for use as a diagnostic tool for protein structure–function analyses and/or as an aid in large scale genomic sequencing. Many nuclease-based *in vitro* methods for the generation of unidirectional deletions have been reported (1–7). All of these methods involve multiple enzymatic steps. In addition, PCR-based deletion studies have recently been reported. These studies involve PCR overlap extension (8), a 'megaprimer' procedure (9), a 'solid phase' method (10) or inverse PCR (11,12). These procedures, like the nuclease-based deletion methods, require a series of DNA manipulations.

A number of *in vivo* transposon-based methods have also been employed as a tool for functional and DNA sequence analyses. Transposon γδ, a member of the Tn*3* family, has been used both as a mobile binding site (intermolecular transposition) for universal primers (13–19) and as a method for generating intramolecular deletion and inversion events (20–22). At least two other transposon-mediated deletion generating systems have been described using Tn*9* (23) and Tn*5* (24). While all these transposon-based methods are successful in generating nested deletions, in general the frequency of transposition is low (<10^{-4}) and they require multiple genetic manipulations to find the desired products. Recently an *in vitro* system using Tn*3* transposase has been reported. However, the frequency of deletions in the target of interest is very low (25). In addition, this procedure requires an additional enzymatic reaction absolutely necessary to eliminate the high background noise of 'pseudo' deletion clones. We describe in this study a remarkably efficient one-step method for the generation of nested families of deletions and inversions using a Tn*5* *in vitro* transposition system.

Tn*5* is a composite transposon, composed of two terminal inverted insertion sequences (IS*50*) which flank three antibiotic resistance genes. Encoded within the IS*50* sequence is the 476 amino acid transposase protein (Tnp). Transposition requires both the Tnp protein and two 19 bp Tnp recognition sites at the ends of each IS*50* element. Tn*5* (24,26,27) [as well as Tn*7* (28) and Tn*10* (29,30)] undergoes a simple 'cut and paste' transposition process. After specific binding of Tnp to the two 19 bp inverted sequences a higher ordered protein–DNA complex (synapse) is formed by protein oligomerization. Double-stranded cleavage occurs yielding highly reactive 3′-OH groups at the ends. The excised transposon mediates a double-stranded staggered cleavage at a random target site (Goryshin *et al.*, submitted) and, in a strand exchange reaction, the 3′-OH of the transposon joins the 5′-ends of the target. There is a simple repair of the gap generated by the double-stranded cut, resulting in a short duplication (9 bp) of the target DNA surrounding the transposon (31–33).

A highly efficient *in vitro* transposition system for Tn*5* has recently been developed (27). It is important to note that there are only two macromolecular components needed in this system; a hyperactive version of Tn*5* transposase (EK54/MA56/LP372) and a substrate containing two inverted 19 bp end sequences (34).

* To whom correspondence should be addressed. Tel: +1 608 262 3608; Fax: +1 608 262 3453; Email: reznikoff@biochem.wisc.edu

Studies with this system have shown that both 3′ and 5′ cleavages occur precisely at the –1/+1 boundary of the 19 bp recognition site (27). In addition, target sites for subsequent strand exchange are fairly random (Goryshin *et al.*, submitted). Both intramolecular and intermolecular transposition can occur, however, at the low DNA concentrations used in the *in vitro* system, >95% of all transposition products are intramolecular events consisting of both deletions and inversions.

Figure 1 shows typical intramolecular transposition events and the two resulting products, inversions and deletions. Transposase binds to the ends and a synaptic complex is formed. Transposase cleaves precisely at the +1/–1 junctions (27), releasing the donor backbone (DBB). The excised transposon attacks 'itself' in one of two ways, as illustrated by the dotted lines from the left (L) and right (R) ends to the target. In a strand exchange reaction the 3′-OH of each end joins the 5′-end of the target. Resolution of the left intermediate pathway results in an inversion of the ends with respect to each other, shown on the bottom left. The size of this product (inversion) is equivalent to that of the excised transposon. Resolution of the target capture intermediate on the right pathway results in two separate deletion products, only one of which contains the *ori* and would therefore be viable when transformed.

In this study we have exploited the intramolecular transposition preference of the *in vitro* system to use as a tool for the generation of a nested family of deletions and inversions. Two different targets, a 1.5 and a 10.8 kb DNA segment, were tested in the *in vitro* system using a synthetic transposon. In the presence of transposase (2 h incubation) each of the target constructs were metabolized at a rate of 30–50%. For both targets there was a 2-fold higher frequency of deletions versus inversions. Moreover, the distribution of deletion and inversion end points were essentially random.

MATERIAL AND METHODS

Strains and plasmids

Escherichia coli strain DH5α [F⁻ϕ80d/*lac*ZΔM15, *end*A1, *rec*A1, *hsd*R17 (r_K^-, m_K^-), *sup*E44, *thi*-1, *gyr*A96, *rel*A1, Δ(*lac*ZYA-*arg*F), U169, λ⁻] was used for plasmid isolation. Plasmids pBR322 (35,36), pUC19 (37), pRZ7074 (38) and pRZTL4 (34) have been described previously.

PCR amplification

All PCR amplifications were performed in a PTC200 Peltier Thermal Cycler (MJ Research) using Taq DNA polymerase (Pharmacia). The standard PCR reaction (100 µl) contained the following: 10 mM Tris–HCl, 1.5 mM $MgCl_2$, 50 mM KCl, 0.5 µM each primer (or 1 0.2 µM single primer) and 1–10 ng template. Each reaction was subjected to a 'hot start' (97°C for 1 min) followed by incubation at 55°C. After a 5 min primer annealing step 40 µM dNTP mix (Pharmacia) and 5 U Taq polymerase were added. Twenty eight cycles were performed (denaturation at 95°C for 1 min, primer annealing at 55°C for 2 min and extension at 72°C for 1 min). After 28 cycles there was a final 5 min extension step at 72°C.

Plasmid pRZTL4, containing the kanamycin resistance gene (*Kan*ʳ) flanked by two 19 bp OE/IE mosaic sequences oriented in a head-to-head fashion, was used as a substrate for amplification using a single primer (purchased at Research Genetics Inc.), d(ACATGCATGC**TCA**C**TCA**C**TCA**AGATGTGTATAAGA-GACAGTCGACCTGCAGGG). The complementary portion of the OE/IE primer to the pRZTL4 substrate, shown above as

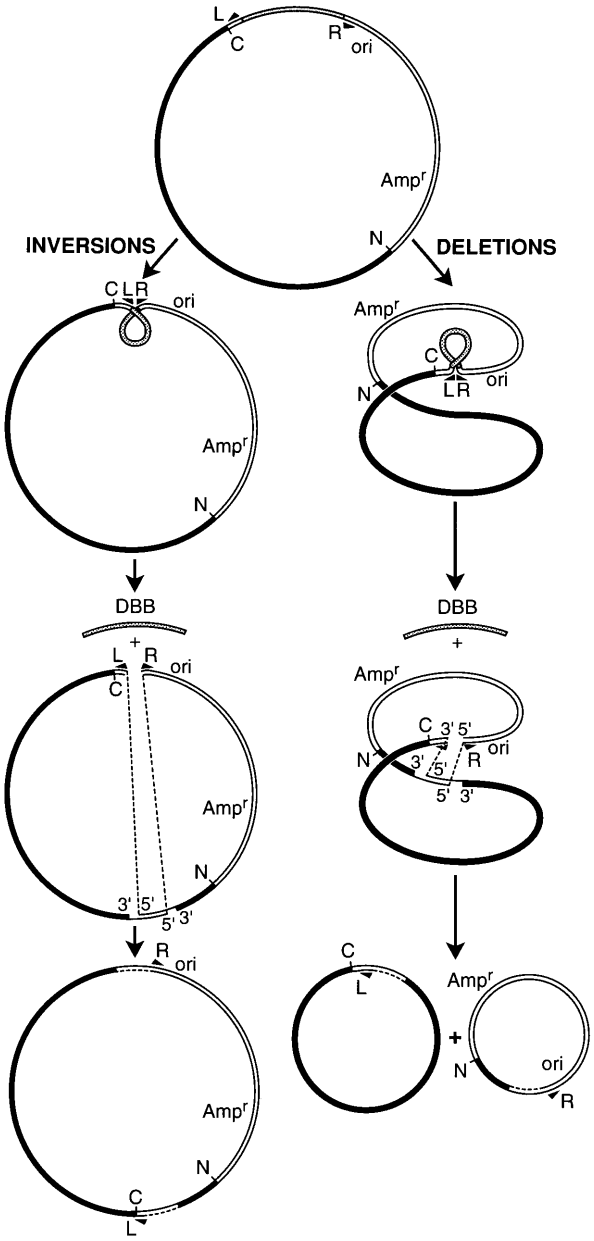

Figure 1. Model for Tn*5* transposase-mediated intramolecular transposition into a hypothetical gene (N indicates N-terminus and C indicates C-terminus) in deletion vector plasmid pRZ9075. Inversion and deletion formation pathways are shown to the left and right of the figure respectively. The positions of the ampicillin resistance gene (*AMP*) and the Col E1 origin of replication (*ori*) are shown. The two inverted 19 bp mosaic end sequences are indicated by the symbols L (left end) and R (right end). After synapse formation, as shown by the first intermediate of each pathway, double-ended transposase-mediated cleavage occurs, resulting in release of the donor backbone (DBB). Subsequent staggered cleavage of the target by 3′-OH of the L and R ends and joining of the ends to the target is shown by dotted lines. Resolution of the intermediates in the pathway results in formation of an inversion (lower left) or two deletion products (lower right). Only the deletion product which retains the *ori* will survive after transformation. Staggered cleavage of the target site results in a 9 bp duplication of the target DNA adjacent to each end (indicated by a dotted line in the final products).

double underlined, contains a 19 bp mosaic OE/IE sequence (read +19 to +1 in the 5′→3′ direction). Upstream of the complementary portion the primer also includes three nonsense codons (in bold) and an *Sph*I site (single underlined). Amplification of pRZTL4 with the OE/IE mosaic primer was performed as described above. Reactions were phenol extracted, ethanol precipitated and digested overnight with *Sph*I. PCR products (1325 bp) were visualized on a 1% agarose gel and isolated using Qiaex II (Qiagen).

Deletion vector construction

Plasmid pBR322 (35,36) was digested with *Afl*III and *Eco*RV, the sticky ends filled-in with Klenow and religated (pRZ9073). The multi-cloning site (MCS) of pUC19 (37) was isolated as an *Eco*RI–*Hin*dIII fragment (54 bp) and cloned into the *Eco*RI and *Hin*dIII sites of pRZ9073, resulting in pRZ9074. The isolated *Sph*I-digested PCR product containing the *Kan*^{r} gene flanked by two inverted 19 bp OE/IE mosaic sequences, as described above, was cloned into the *Sph*I site of the MCS of pRZ9074. This plasmid was designated pRZ9075.

Plasmid pRZ7074 (38), a pET21d derivative containing the entire *Tnp* gene (with the MA56 mutation that eliminates Inh protein production), was digested with *Bgl*II. The 1530 bp fragment containing the *Tnp* gene under control of the T7 promoter was isolated and cloned into the *Bam*HI site of pRZ9075. This plasmid was designated pRZ9076. Plasmid pRZ9077 was constructed by inserting a 10.8 kb segment of the *Escherichia coli* chromosome (77.6 min), isolated as an *Eco*RI fragment, into the *Eco*RI site of pRZ7074.

Protein purification

EK54/MA56/LP372 Tnp protein was purified as previously described (27). The homogeneity, as determined by densitometric scan of a Coomassie stained SDS–PAGE gel, was found to be 96%.

Generation of nested deletions and inversions

Plasmid pRZ9075 containing either the *Tnp* gene under control of the T7 promoter or the 10.8 kb segment of the *E.coli* chromosome was used in the *in vitro* transposition reaction (as described in 27). A 20 µl reaction volume containing 0.1 M potassium glutamate, 25 mM Tris–acetate, pH 7.5, 10 mM Mg^{2+} acetate, 50 µg/ml BSA, 0.5 mM β-mercaptoethanol, 2 mM spermidine, 100 µg/ml tRNA, 0.037–0.12 pmol plasmid and 5.6 pmol transposase protein was incubated for 2 h at 37°C. Reactions were phenol extracted and ethanol precipitated. Half of each reaction was transformed into electrocompetent DH5α cells. Aliquots of 100 µl of a 1:100 dilution of each transformation were plated onto Amp^{100} plates. A number of Amp^r colonies were replica-plated onto Amp^{100} and Kan^{20} plates. Amp^r/Kan^s colonies were inoculated into LB containing Amp^{100} and grown overnight. Isolated plasmids (prepared with the Wizard SV miniprep kit from Promega) from each culture were screened for size by electrophoresis on a 1% agarose gel. The end–target junctions were determined by sequencing using the US Biochemical T7 Sequenase v.2.0 kit according to the protocols provided. A universal primer, complementary to just upstream of the *Eco*RV site within the pBR322 vector, was used for all sequencing of nested deletions and inversions. The sequence of the primer is d(CGCAAGAGGCCCGGCAGTAC).

RESULTS AND DISCUSSION

The *in vitro* Tn*5* transposition system, as developed by Goryshin and Reznikoff (27), is a highly efficient one-step reaction. It contains two macromolecular components, a hyperactive form of Tn*5* transposase and DNA containing two inverted 19 bp Tnp recognition sites. The hyperactive form of Tn*5* Tnp (EK54/MA56/LP372) dramatically increases the transposition frequency (>10^3 fold) *in vitro* (27). The hyperactive end sequences used were chosen based on an *in vivo* mutagenesis study by Zhou *et al.* (34) and subsequently tested in the *in vitro* system (Goryshin and Reznikoff, personal communication). Transposition frequencies using these mosaic end sequences increased substantially both *in vivo* and *in vitro*. Together with the hyperactive form of Tnp, the efficiency of the *in vitro* system is maximized. At the low DNA concentration used intramolecular transpositional events comprise >95% of the products (both deletions and inversions) generated. We have exploited the intramolecular preference of the *in vitro* system to test its effectiveness as a tool to generate nested families of deletions and inversions.

A Tn*5* transposition deletion vector, pRZ9075 (Fig. 2A), was constructed as described in Materials and Methods. The transposon portion of the vector (2163 bp) consists of an origin of replication, the *Amp*^{r} gene, an MCS and the two hyperactive inverted 19 bp end sequences. Immediately adjacent to each end, as shown in Figure 2A, are three stop codons in all three reading frames (designated NS). The DBB portion (1255 bp) of the deletion vector (see Fig. 1) flanked by the two ends encodes the *Kan*^{r} gene. The *Kan*^{r} gene provides a convenient screen for detecting transposition events. The MCS allows for insertion of any DNA sequence for deletion analysis. Construction of the vector is such that any Amp^r replicating plasmid that has undergone a transpositional event (Kan^s) will be targeted primarily to any insert in the MCS.

As an initial test to evaluate the efficiency of the system, a 1.5 kb *Bgl*II fragment containing the *Tnp* gene under control of a T7 promoter was cloned into the *Bam*HI site of pRZ9075. The 5′→3′ orientation of the gene (pRZ9076; see Fig. 2B) allows for C-terminal deletions of the *Tnp* gene to be generated in the *in vitro* system. The positioning of stop codons in all three reading frames adjacent to each end in addition to a regulatable promoter provided with the gene of interest in the MCS allows for overexpression of C-terminal deletion proteins. As described above, intramolecular events predominate in the Tn*5* *in vitro* transposition system, resulting in both deletion and inversion products. Figure 2C illustrates the two predicted intramolecular events featuring the *Tnp* gene as the target (see also Fig. 1). The solid lines show a typical end attack that results in deletions. Upon transformation only the deletion product which retains the origin (thick line) will survive, resulting in a C-terminal deletion of the Tnp *gene*. As shown in Figure 2C, there is a 9 bp duplication of the target DNA after strand exchange and repair of a gap due to staggered cleavage of the target site by the Tnp–DNA complex. The dotted lines show the results of an end attack that results in an inversion. Also indicated is the single primer used to evaluate both deletion and inversion products.

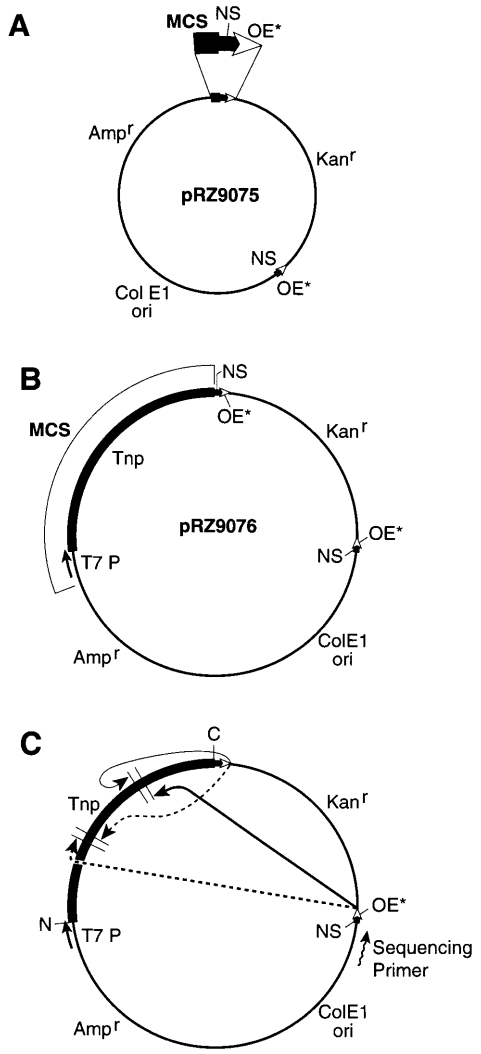

Figure 2. Deletion vectors and possible intramolecular products. (**A**) The synthetic Tn*5* transposon, pRZ9075, consists of an ampicillin resistance gene (*Amp*^{r}), the Col E1 origin, a multi-cloning site (MCS), two hyperactive mosaic end 19 bp end sequences (OE*) and three stop codons (NS). A blow-up of the region containing the MCS and one set of NS and OE* is shown. The donor backbone encodes the kanamycin resistance gene (*Kan*^{r}). (**B**) Plasmid pRZ9076, a derivative of pRZ9075, contains the 1.5 kb *Tnp* gene under control of the T7 promoter in the MCS. (**C**) pRZ9076 and the two predicted intramolecular products. Solid lines illustrate a typical end attack that results in deletions. The thicker solid line indicates the deletion product which would survive after transformation (retains the origin). The dotted lines demonstrate a typical end attack that would result in an inversion. The paired vertical lines between the arrows (which represents end attacks) indicate the 9 bp duplications that are the result of staggered cleavage at the target sites. The location of the sequencing primer used to analyze the end–target junctions is also indicated.

Plasmid pRZ9076 (Fig. 2B) was used in the *in vitro* transposition system. Purified pRZ9076 (0.061 pmol) plasmid (Qiagen Maxi kit) was incubated in the presence of purified hyperactive Tnp protein (5.6 pmol) at 37°C. After 2 h the reaction was phenol extracted and precipitated. Half of the reaction was transformed into electrocompetent DH5α cells. One hundred microliters of a 1/100 dilution of the transformation mix was plated onto Amp^{100} plates and incubated overnight, yielding ~100 colonies per plate. Two hundred Amp^r colonies were subsequently replica-plated onto both Amp^{100} and Kan^{20} plates. Of the 200 Amp^r colonies tested 72 were found to be Kan^s; a transposition frequency of 36%. Plasmid analysis was performed on the 72 independent Amp^r/Kan^s isolates. Size determination of the purified plasmids was performed by electrophoresis on 1% agarose gels to determine the approximate size of the deletions in the 1.5 kb *Tnp* gene generated in the transposition assay (data not shown). As controls upper and lower size limit plasmids were used in the first and last lanes of each gel. As the upper size control an inversion product of pRZ9076 was used. This plasmid is the same size (3693 bp) as the excised transposon (DBB released) and still contains the entire 1.5 kb *Tnp* gene insert. The lower size limit is a plasmid that has undergone a transposition event resulting in release of the DBB and deletion of the entire *Tnp* gene and T7 promoter region (~2100 bp). As a result of plasmid size determination 41 plasmids were chosen for sequencing. Thirty eight of the plasmids sequenced had deletions within the *Tnp* gene, one had a deletion at the 3′-end of the *Amp*^{r} gene in the pBR322 vector and two were the result of inversions within the *Tnp* gene.

To determine the relative ratio of deletions versus inversions in a random sampling of Amp^r/Kan^s colonies 24 additional isolates were analyzed. In Figure 3A the 24 purified plasmids were run on a 1% agarose gel using the upper size control (U) and lower size plasmid (L), as described above. Plasmids that migrate at a position similar to that of U would be expected to be the result of an inversion, while those plasmids that migrate between U and L would be expected to be the result of a deletion event. The transposition junction of all 24 samples were sequenced. All plasmids shown in Figure 3A whose mobility fell between the two control plasmids were found to be deletions. The plasmid in lane 4 was the result of an intermolecular event. Of the 24 Amp^r/Kan^s isolates 58% were the result of deletions, 38% were the result of inversions and 4% represented an intermolecular event. Of the 58% deletions found only one was in the vector region. Most importantly, as diagramed in Figure 3B, both deletions (indicated by vertical lines above the *Tnp* gene line representing 3′ end points) and inversions (vertical lines below *Tnp* gene line representing 5′ end points) generated from all the studies with the *Tnp* gene are randomly distributed within the 1.5 kb sequence.

The positioning of three stop codons in all three reading frames immediately adjacent to the mosaic end sequences (see Fig. 2A) should terminate Tnp N-terminal peptides expressed by the T7 promoter (provided with the *Tnp* gene in the MCS). Proteins from various deletions were overexpressed and were found to be of the expected molecular weight, as determined by SDS–PAGE (data not shown). Thus positioning of the stop codons adjacent to the end sequences allows structure–function deletion analysis of any protein.

A number of interesting observations were made as a result of this study with the *Tnp* gene. We observed a 30–50% transposition frequency (Kan^s/Amp^r), of which 95% resulted in randomly distributed deletions and inversions within the *Tnp* gene. Therefore, from a single one-step *in vitro* transposition assay 10^5 or more deletions/inversions can be expected within the target. The unequal distribution, 58 versus 38%, of deletions and inversions was an unexpected result. Theoretically both should

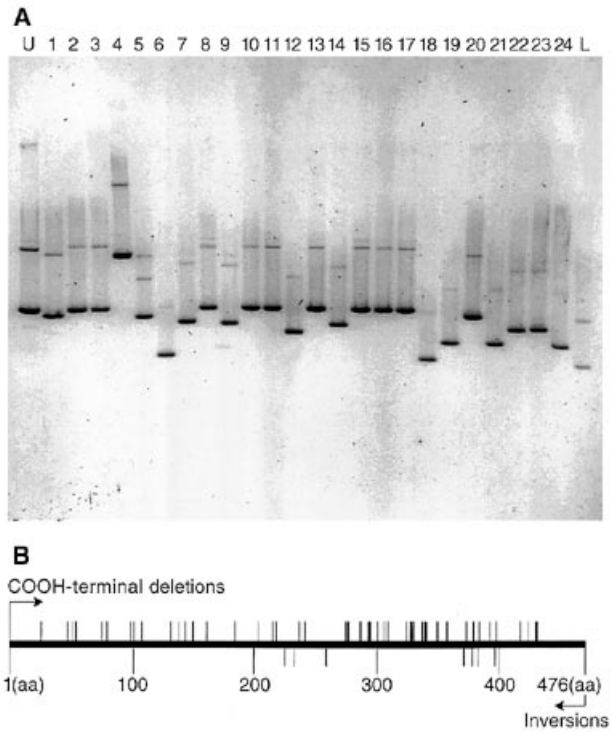

Figure 3. Size determination and sequence analysis of transposition products using pRZ9076. (**A**) Following a transposition reaction using pRZ9076 (described in Materials and Methods), plasmids from 24 Amp^r/Kan^s colonies were isolated. Plasmids were screened for size on a 1% agarose gel (85 V for 5 h at 4°C). The upper size control plasmid (U) represents an inversion product of pRZ9076. The lower size control plasmid (L) represents a deletion event resulting in release of the donor backbone (DBB) and deletion of the entire 1.5 kb *Tnp* and the T7 promoter region. (**B**) Distribution of deletions and inversions. The horizontal bar represents the 476 amino acids of Tn*5 Tnp*. The vertical lines above the bar indicate the 3′ end points of C-terminal deletions and the vertical lines below the bar indicate the 5′ end points of inversions. Thicker vertical lines indicate positions where two or more independent deletions or inversions were found.

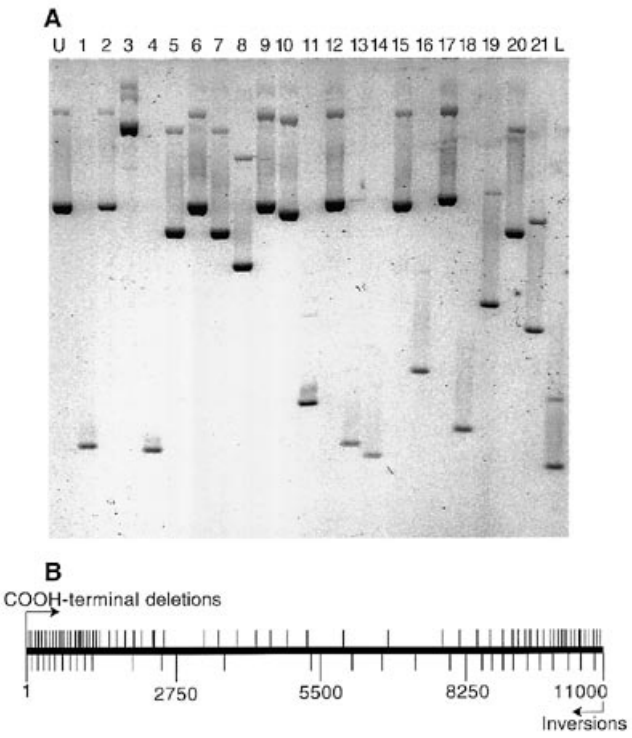

Figure 4. Size determination and sequence analysis of transposition products using pRZ9077. (**A**) Twenty one isolates from a random sampling of 100 Amp^r/Kan^s colonies, the result of a transposition reaction using pRZ9077, were screened for size on a 1% agarose gel (40 V for 18 h at 4°C). The upper size control plasmid (U) represents an inversion product of pRZ9077. The lower size control plasmid (L) is a result of a deletion event resulting in release of the DBB and deletion of the entire 10.8 kb DNA segment. (**B**) Distribution of deletions and inversions within the 10.8 kb DNA segment. The horizontal bar represents 11 kb of DNA. The vertical lines above the bar indicate the 3′ end points of C-terminal deletions. The vertical lines below the bar indicate the 5′ end points of inversions. Thicker vertical lines indicate positions where two or more independent deletions were found.

occur with equal probability. We discuss possible explanations for this outcome later.

Also, from this study we observed an ~180 bp gap between the closest end sequence and the farthest 3′ deletion end point within the C-terminus of the *Tnp* gene (Fig. 3B). This spacer region suggests that the DNA persistence length constraint on transposition is in the same range as that of synapse formation, as described by Goryshin *et al.* (39). In their report changing the length of the donor DNA (from 200 to 64 bp) between the end sequences had dramatic effects on the frequency of transposition *in vivo*. When the end sequences are in proximity transposition is inhibited. Although transposition is restored with longer donor DNA lengths (between 66 and 174 bp), a definite periodic relationship between DNA length and transposition exists. This implies that the overall architecture of the synaptic complex is stringent. In this study the intramolecular nature of the *in vitro* system likely imposes additional constraints on transposition. Therefore, DNA looping constraints probably affect both synapse formation and subsequent intramolecular target capture (see Fig. 1).

To test this system as a tool for sequencing large segments of DNA a 10.8 kb fragment from the *E.coli* chromosome was cloned into the MCS of the deletion vector. This construct, pRZ9077, was used in the *in vitro* transposition system as described above. As with pRZ7076, 30–50% of all Amp^r colonies had undergone a true transposition event (Kan^s). Electrophoretic mobility of plasmids from 100 Amp^r/Kan^s isolates was analyzed on a 1% agarose gel. An example of one of the gels is shown in Figure 4A. The upper size limit control plasmid in the first lane represents an inversion product of pRZ9077, while the lower size limit, shown in the last lane, represents a plasmid in which the DBB as well as the entire 10.8 kb segment has been removed. Dideoxy sequencing of the transposition junction of all 100 Amp^r/Kan^s isolates revealed the following: 68% of the isolates were found to be the result of deletions, 28% were the result of inversions, 3% represented one-sided transposition events (40) and 1% resulted from an intermolecular event. A one-sided transposition event is the result of cleavage and subsequent strand transfer at only one end.

As shown in Figure 4B, the distribution of deletions (3′ end points represented by vertical lines above the DNA bar) and inversions (5′ end points represented by vertical lines below the DNA bar) differs somewhat from that of the *Tnp* gene (see Fig. 3B). The clustering of deletions/inversions near the N- and

C-termini is probably a reflection of the fact that there is a decrease in the local concentration of synaptic complex relative to the target DNA with increasing DNA length. However, in general, deletions and inversions are found throughout the entire region.

As with the 10.8 kb DNA segment, distribution of end points of larger DNAs would be subject to probability density (referred to as the Jacobson–Stockmayer factor; 41). Jacobson and Stockmayer examined the probability of cyclization of linear random coiled polymers in which the probability density is the effective concentration of one end in the vicinity of the other end. When comparing the cyclization of two DNA molecules the Jacoson–Stockmayer factor predicts that the ratio of the lengths of two molecules raised to the power 3/2 is equal to the ratios of their probability densities. In the current study the probability density is the effective concentration of the synaptic complex to any given target along the DNA length. Based on the results with the 10.8 kb segment in these studies we can approximate the relative frequency of deletions/inversions end points for any size DNAs. For example, the farthest target site relative to an end in the 10.8 kb segment would be in the center, or 5 kb away. The farthest target of a DNA segment 20 kb in length would be 10 kb. Therefore, the probability of a deletion/inversion event near the center of a 20 kb DNA would be 35% of that of the 10.8 kb segment near its center [1/2 (ratio of the two lengths)$^{3/2}$].

Two additional observations from the *in vitro* studies with the 10.8 kb segment are consistent with the *Tnp* construct results. There was a gap (~150 bp) from the closest mosaic end sequence and the farthest 3′ deletion end point. In addition, there was a 2- to 3-fold higher occurrence of deletions within the 10.8 kb DNA segment than inversions.

Three adjacent deletions (deletion rearrangements apparently resulting from transposase action at only one end) were isolated in the 10.8 kb insert study. Based on *in vivo* studies by Jilk *et al.* (40) we believe that these events may also have occurred as a result of a failure to form a synaptic complex. One-ended –1/+1 junction transposition events were found to occur *in vivo* when the transposable element contained one wild-type end sequence and a deletion (or a binding site mutant). These one-ended transposition products were proposed to be the result of a failure to form a synaptic complex followed by cleavage occurring only at the –1/+1 junction of the wild-type end. In this study we believe that the one-sided transposition events observed in the 10.8 kb segment are the result of binding of Tnp to only one end and subsequent cleavage and strand transfer of that end into the DBB. This results in a viable deletion product in which both ends are present on the same molecule.

In this study we observed an unequal distribution of deletion and inversion products. We believe that a portion of the deletions are the result of non-concerted end usage in our *in vitro* system despite the formation of synaptic complexes. A typical transposition reaction is thought to occur as follows: Tnp protein binds to the end sequences, a synaptic complex forms through protein oligomerization, cleavage occurs at the –1/+1 junction of both ends simultaneously, the target is captured and strand transfer occurs. The additional deletions that we isolated probably result from a failure to precisely coordinate dual end –1/+1 cleavage following synaptic complex formation. This results in a lag time between cleavages that is sufficient for target capture and strand transfer of the first end cleaved by itself. We believe that synapse formation occurs because of the observed ‘spacer’ region between an end and the closest deletion end point. The apparent lack of coordination in dual end –1/+1 cleavage may be an artifact of our *in vitro* conditions.

CONCLUSIONS

The Tn*5 in vitro* transposition system provides an efficient one step procedure for generating nested families of deletions and inversions. Nested families of deletions can easily be used in protein studies. Positioning of stop codons adjacent to the end sequences leads to formation of truncated proteins for structure–function analyses. Generation of both deletions and inversions from a single one-step reaction provides the necessary substrates (from both strands) for overlapping sequence determination.

Two new deletion vectors have recently been completed. One construct is identical to pRZ9075, however, the kanamycin gene was replaced with DNA encoding the α subunit of β-galactosidase. This allows for an immediate screen for transposition events (white colonies on Amp^{100}/X-Gal/IPTG plates) after transformation. As a test of this new construct the *Tnp* gene under control of the T7 promoter was cloned in. Transposition frequencies with this construct are comparable with those with pRZ9076 (data not shown).

For generation of N-terminal deletions for protein analysis a second deletion vector was constructed. Immediately adjacent to the mosaic end sequence closest to the origin of replication (in place of the stop codons in the original deletion vector) was placed a T7 promoter region with a His tag and kinase region. However, there is one drawback of this construct. Of all deletions generated in the gene of interest using this construct only one third will be in the correct reading frame. However, the high efficiency of the reaction should guarantee more than enough ‘correct’ deletions for subsequent study. Tests of this vector with the *Tnp* gene in the MCS are currently under investigation. Initial results are promising.

ACKNOWLEDGEMENTS

This paper is dedicated to the memory of the late Claire M.Berg, who helped introduce W.S.R. to the excitement of *E.coli* molecular genetics during the summer of 1964 and who pioneered the use of transposons in genomic research. This work was supported in part by NIH grant GM50692. D.Y. was supported by the University–Industry Relations Program of the University of Wisconsin. K.W. was a participant in the University of Wisconsin Summer Research Program (NSF REU Site Grant BIR-9424074). W.S.R. is the Evelyn Mercer Professor of Biochemistry and Molecular Biology.

REFERENCES

1 Guo,L.-H., Yang,R.C.A. and Wu,R. (1983) *Nucleic Acids Res.*, **11**, 5521–5540.
2 Hattori,M., Tsukahara,F., Furuhata,Y., Tanahashi,H., Hirose,M., Saito,M., Tsukuni,S. and Sakaki,Y. (1997) *Nucleic Acids Res.*, **25**, 1802–1808.
3 Henikoff,S. (1984) *Gene*, **28**, 351–359.
4 McCombie,W.R., Kirkness,E., Fleming,J.T., Kerlavage,A.R., Iovannisci,D.M. and Martin-Gallardo,A. (1991) *Methods (Orlando)*, **3**, 33–40.
5 Okita,T.W. (1985) *Anal. Biochem.* **144**, 201–211.
6 Ozkayanak,E. and Putney,S.O. (1987) *BioTechniques*, **5**, 770–773.
7 Zhu,K.H. and Marshall,J.M. (1995) *BioTechniques*, **18**, 222–224.
8 Horton,R.M. and Pease,L.R. (1991) In McPherson,M.J. (ed.), *Directed Mutagenesis: A Practical Approach*. IRL Press, Oxford, UK, pp. 217–247.
9 Ogel,Z.B. and McPherson,M.J. (1992) *Protein Engng*, **5**, 467–468.

10 Yohda,M., Kato,N. and Endo,I. (1995) *DNA Res.*, **2**, 175–181.
11 Imai,Y., Matsushima,Y., Sugimura,T. and Terada,M. (1991) *Nucleic Acids Res.*, **19**, 2785–2785.
12 Pues,H., Holtz,B. and Weinhold,E. (1997) *Nucleic Acids Res.*, **25**, 1303–1304.
13 Berg,C.M., Berg,D.E. and Groisman,E.A. (1989) In Berg,D.E. and Howe,H.H. (eds), *Mobile DNA*. ASM Press, Washington, DC, pp. 879–925.
14 Berg,C.M., Vartak,N.B., Wang,G., Xu,X., Liu,L., MacNeil,D.J., Gewain,K.M., Wiater,L.A. and Berg,D.E. (1992) *Gene*, **113**, 9–16.
15 Berg,C.M., Wang,G., Strausbaugh,L.D. and Berg,D.E. (1993) *Methods Enzymol.*, **218**, 279–306.
16 Liu,L., Whalen,W., Das,A. and Berg,C.M. (1987) *Nucleic Acids Res.*, **15**, 9461–9469.
17 Martin,C.H., Mayeda,C.A., Davis,C.A., Ericsson,C.L., Knafels,J.D., Mathog,D.R., Celniker,S.E., Lewis,E.B. and Palazzolo,M.J. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 8398–8402.
18 Strathmann,M., Hamilton,B.A., Mayeda,C.A., Simon,M.I., Meyerowitz,E.M. and Palazzolo,M.J. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1247–1250.
19 Strausbaugh,L.D., Bourke,M.T., Sommer,M.T., Coon,M.E. and Berg,C.M. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 6213–6217.
20 Krishnan,B.R., Jamry,I., Berg,D.E., Berg,C.M. and Chaplin,D.D. (1995) *Nucleic Acids Res.*, **23**, 117–122.
21 Wang,G., Blakesley,R.W., Berg,D.E. and Berg,C.M. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 7874–7878.
22 Wang,G., Xu,X., Chen,J.-M., Berg,D.E. and Berg,C.M. (1994) *J. Bacteriol.*, **176**, 1332–1338.
23 Ahmed,A. (1984) *J. Mol. Biol.*, **178**, 941–948.
24 Tomcsanyi,T., Berg,C.M., Phadnis,S.H. and Berg,D.E. (1990) *J. Bacteriol.*, **172**, 6348–6354.
25 Morita,M., Umemoto,A., Li,Z.-X., Nakazono,N. and Sugino,Y. (1996) *DNA Res.*, **3**, 431–433.
26 Berg,D.E. (1989) In Berg,D.E. and Howe,M.M. (eds), *Mobile DNA*. ASM Press, Washington, DC, pp. 184–210.
27 Goryshin,I.G. and Reznikoff,W.S. (1998) *J. Biol. Chem.*, in press.
28 Bainton,R., Gamas,P. and Craig,N.L. (1991) *Cell*, **65**, 805–816.
29 Benjamin,H.W. and Kleckner,N. (1989) *Cell*, **59**, 373–383.
30 Morisato,D. and Kleckner,N. (1984) *Cell*, **39**, 181–190.
31 Berg,D.E., Johnsrud,L., McDivitt,L., Ramabhadran,R. and Hirschel,B.J. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2632–2635.
32 Nag,D.K., DasGupta,U., Adelt,G. and Berg,D.E. (1985) *Gene*, **34**, 17–26.
33 Schaller,H. (1979) *Cold Spring Harbor Symp. Quant. Biol.*,**43**, 401–408**.**
34 Zhou,M., Bhasin,A. and Reznikoff,W.S. (1998) *J. Mol. Biol.*, **276**, 913–925..
35 Bolivar,F., Rodriguez,R.L., Greene,P.J., Betlach,M.C., Heyneker,H.L. and Boyer,H.W. (1977) *Gene*, **2**, 95–113.
36 Sutcliffe,J.G. (1979) *Cold Spring Harbor Symp. Quant. Biol.*, **2**, 77–90.
37 Yanisch-Perron,C., Viera,J. and Messing,J. (1985) *Gene*, **33**, 103–119.
38 Wiegand,T.W. and Reznikoff,W.S. (1994) *J. Mol. Biol.*, **235**, 486–495.
39 Goryshin,I., Kil,Y.V. and Reznikoff,W.S. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 10834–10838.
40 Jilk,R.A., Makris,J.C., Borchardt,L. and Reznikoff,W.S. (1993) *J. Bacteriol.*, **175**, 1264–1271.
41 Jacobson,H. and Stockmayer,W.H. (1950) *J. Chem. Phys.*, **18**, 1600.