Determinants for hairpin formation in Tn10 transposition

John S. Allingham, Simon J. Wardle and David B. Haniford¹

Department of Biochemistry, University of Western Ontario, London, Ontario, Canada N6A 5B7

¹Corresponding author
e-mail: dhanifor@uwo.ca

Tn10 transposition involves the formation of a hairpin intermediate at the transposon termini. Here we show that hairpin formation exhibits more stringent DNA sequence requirements at the terminal two base pairs than either transpososome assembly or first strand nicking. We also observe a significant DNA distortion at the terminal base pairs upon transpososome assembly by chemical nuclease footprinting. Interestingly, mutations at these positions do not necessarily inhibit the formation of the distortion. However, it remains a possibility that the inhibitory effect of these mutations is due to a defect in protein–DNA interactions subsequent to this deformation. Terminal base pair mutations also inhibited strand transfer, providing evidence that transposase interactions with the terminal residues on both ‘transferred’ and ‘non-transferred’ strands are important for hairpin formation. We also demonstrate that mutation of a highly conserved tyrosine residue that is a component of the YREK motif, Y285, results in a phenotype comparable to that of the terminal base pair mutations. In contrast, a mutation at another conserved position, W265, is shown to relax the specificity of the hairpin formation reaction.

Keywords: DNA hairpin/DNA transposition/
protein–DNA interactions/Tn10/transposase

Introduction

The composite bacterial transposon Tn10 transposes by a non-replicative mechanism. The complete transposon, or one of its two flanking IS10 modules, is first excised by a pair of double strand breaks and then inserted into a new target site. These steps are carried out by the Tn10-encoded transposase protein, a 46 kDa protein that binds to specialized DNA sequences located at transposon termini or ‘ends’. Two transposase monomers bind per transposition module (Kennedy, 1999) and orchestrate the assembly of a higher order protein–DNA complex called a transpososome, in which the two transposon ends are held together. All of the chemical steps take place in the context of the transpososome (Kleckner *et al.*, 1996). A particularly fascinating aspect of the Tn10/IS10 transposition reaction is that the single active site in the transposase protein is used repeatedly to catalyze the four different chemical steps that occur at each transposon end

(Figure 1). In previous work, we have defined the orientation of different DNA strands in the single active site for each of the four chemical steps (Kennedy *et al.*, 2000). However, a more detailed understanding of each of these steps requires definition of the transposase–DNA interactions that govern them. This remains an important goal in this and other related transposition reactions.

As shown in Figure 1, Tn10 excision takes place through a hairpin mechanism (Kennedy *et al.*, 1998). The formation of a double strand break in DNA by a hairpin mechanism also occurs in V(D)J recombination and Tn5 transposition, and is likely for a number of plant transposons that utilize the cut-and-paste mode of DNA transposition (Coen *et al.*, 1989; Fedoroff, 1989; McBlane *et al.*, 1995; Colot *et al.*, 1998; Bhasin *et al.*, 1999). In the Tn10, Tn5 and V(D)J systems, hairpin formation involves attack of a 3' OH DNA strand terminus on a phosphate group situated directly across the double helix from the site of the initial nick, giving a perfectly self-complementary DNA hairpin (McBlane *et al.*, 1995; Kennedy *et al.*, 1998; Bhasin *et al.*, 1999). Also, in the V(D)J and Tn10 systems, it has been shown that this reaction takes place by a direct in-line attack mechanism wherein hairpin formation is contemporaneous with the complete release of transposon from flanking DNA sequences (van Gent *et al.*, 1996; Kennedy *et al.*, 2000).

Hairpin formation is considered to be a relatively difficult reaction because the functional groups involved in the reaction chemistry are separated by 16–18 Å in B-form DNA. It is therefore expected that this step would be associated with a significant distortion in the DNA structure at the site of the reaction chemistry. The first evidence of this came from studies in the V(D)J system where it was shown that the introduction of base pair mismatches immediately adjacent to the site of hairpin formation relieved the inhibitory effect of suboptimal flanking DNA sequences on hairpin formation (Cuomo *et al.*, 1996; Ramsden *et al.*, 1996). More recently, the structure of a Tn5 transpososome thought to represent a stage of the reaction following cleavage of the transposon from the donor DNA has been solved. In this complex, the terminal two base pairs of the transposon end are unpaired and base stacking interactions in this region are disrupted by the ‘flipping’ of one of the bases at the second position out of the double helix (Davies *et al.*, 2000). Details of the molecular events that govern hairpin formation have not been addressed previously in the Tn10 system and are the main focus of the current study.

Tn10 transposase interactions with its end sequences have been studied extensively by DNA footprinting and interference approaches as well as by functional studies on mutant forms of the end sequences (Huisman *et al.*, 1989; Haniford and Kleckner, 1994; Sakai *et al.*, 1995, 2000; Kleckner *et al.*, 1996; Sakai and Kleckner, 1996). The

Fig. 1. Structure of IS10 Right and transposase-mediated chemical steps in Tn10/IS10 transposition. IS10 Right (thick lines) encodes transposase protein and contains the binding determinants for transposase (filled rectangles) at its termini, the 'inside' (IE) and 'outside' (OE) end sequences. The inverted repeat sequence (half arrows) contains the primary specificity determinants for transposase binding. In addition, the OE also contains a binding site for *Escherichia coli* IHF (open rectangle) and a second 'distal' transposase-binding site. The four chemical steps that occur at each transposon end are indicated; thin lines indicate flanking donor DNA. The top strand shown is the 'non-transferred' strand (NTS) and the bottom strand is the 'transferred' strand (TS).

picture that emerges from these studies is that the primary determinants for end recognition are located in the interior portion of the inverted repeat sequence, including bp 6–13. DNA sequence information is also important at the tip of the inverted repeat as mutations at bp 1–3 strongly inhibit transposition. However, the effects of these mutations are manifested at a stage subsequent to the initial binding of transposase to the end sequence. Given that previous mutational studies were carried out prior to the identification of the hairpin intermediate in the Tn10 system, and considering the complexity of this reaction, the possibility arises that the terminal base pair mutations may interfere with hairpin formation.

A similar transposon end organization to that described above has also been demonstrated in the Tn5 and IS903 systems (Derbyshire *et al.*, 1987; Jilk *et al.*, 1996; Davies *et al.*, 2000). Interestingly, the terminal four base pairs of the outside ends of Tn10 and Tn5 are identical in sequence. Given that both systems employ the hairpin mechanism, this could reflect similar structural requirements for hairpin formation in the two systems. At the

protein level, the Tn10 and Tn5 transposases exhibit only 20% amino acid sequence identity. However, sequence similarity is extensive throughout the entire lengths of these two proteins, with the exception of a block of 44 amino acids present at the C-terminus of the Tn5 transposase. The most highly conserved block of amino acids in the two proteins is a region that contains the E of the DDE catalytic triad. This includes the Y-2-R-3-E-6-K signature, a hallmark of the IS4 family of transposons, of which both Tn10 and Tn5 are members (Mahillon *et al.*, 1985). The structure of the Tn5 synaptic complex has provided possible clues towards defining amino acid residues that play an important role in hairpin formation (Davies *et al.*, 2000). For example, W298 and Y237 stack against the flipped out base at position 2, and Y319 forms a hydrogen bond with a phosphate oxygen at the site of a sharp kink between residues 1 and 2. Such interactions are presumably important in positioning the 5' terminus of the non-transferred strand in close proximity to the 3' OH of the transferred strand. This information provides a framework for defining, through mutational analysis, amino acids in Tn10 transposase that may also function in hairpin formation. At the present time, only one residue in Tn10 transposase, P167, has been shown to play an important role in this step (Kennedy *et al.*, 1998). More generally, such studies will be important in addressing how closely related the structures of the two proteins are.

In the current work, we have examined the effects of a selected group of transposon end mutations and amino acid substitutions in transposase on Tn10 excision. We define the specific stages at which these changes affect Tn10 excision and relate these changes to the interactions that take place between transposase and the tip of the outside end.

Results

Mutations at base pairs 1, 2 and 3 inhibit hairpin formation and resolution but not earlier steps

To learn more about the role of specific base pairs at the tip of the Tn10/IS10 outside end in transposon excision, we evaluated in detail the effects of mutations at bp 1–3 on individual steps in the excision reaction. Notably, mutations in this region do not affect transposase binding to the outside end (Huisman *et al.*, 1989; Haniford and Kleckner, 1994). For this investigation, we used the *in vitro* reaction system that employs a short linear outside end DNA fragment as the substrate (Sakai *et al.*, 1995). Duplex substrates formed by annealing 5'-³²P-end-labeled oligonucleotides (Figure 2A) were assembled into a transpososome by mixing with purified integration host factor (IHF) and transposase in a buffer system lacking a divalent metal ion. Mg²⁺ was then added to initiate reaction chemistry and aliquots were removed at various time points and treated with phenol to terminate the reaction. After purification, the DNA was subjected to electrophoresis on a high resolution denaturing gel where various species were separated and then visualized by phosphorimaging. Initial transpososome formation was evaluated by subjecting the reaction mixture to electrophoresis on a native polyacrylamide gel prior to the addition of the divalent metal ion. All possible mutations were made at positions 1–3 and at the first residue of the flanking donor DNA (–1).

Fig. 2. Effects of terminal base pair mutations on transposon excision. (A) An illustration of the linear outside end substrate used to investigate the effects of base pair mutations on transposon excision. Upper case letters indicate the terminal DNA sequence of the wild-type outside end and the lower case letter indicates the first residue in the flanking donor DNA. Numbers indicate base pair positions relative to the transposon–donor junction. Asterisks indicate the position of ^{32}P labels. Tn-B (or transposon bottom strand) is generated by first strand nicking and hairpin resolution. FD-T (or flanking donor top strand) is generated by hairpin formation. Sizes of Tn-B, HP (hairpin) and FD-T are indicated. (B) Wild-type and mutant substrates were assembled into transpososomes (see Materials and methods) and reaction chemistry was initiated by the addition of MgCl_2 . At the indicated time points, reactions were terminated by phenol extraction and the purified DNA was applied to an 8% denaturing polyacrylamide gel. URS is unreacted substrate and contains non-transferred and transferred strands, of 78 and 80 nt in length, respectively. The nomenclature for the mutants is as follows: the number indicates the position in the outside end of the mutation and the letters indicate the mutant bases; the non-transferred (top) strand base is indicated first, followed by the transferred (bottom) strand base. (C) The levels of ‘nicked’ and ‘resolved end’ species were determined from the data in (B) using the following equations: (i) resolved end (RE) = FD-T – HP; and (ii) nicked = Tn-B – RE. Values are expressed as a percentage of the amount of transpososome formed prior to initiation of reaction chemistry.

Two different mutations were made at position 4. In the experiments described below, we show that mutations specifically at bp 1 and 2 can have very strong inhibitory effects on hairpin formation and resolution without significantly affecting transpososome formation or ‘first strand nicking’.

Mutants were classified as either weak or strong based on the level of resolved end (see Figure 1) formed after 4 h in the presence of Mg^{2+} . Strong mutants, including 1A/1T, 1G/1C, 2C/2G and 2A/2T gave <5% of the wild-type level of resolved end (the nomenclature used to describe the end mutants is as follows: the number defines the base pair position relative to the transposon–donor junction, the

letter on the left is the base on the non-transferred and the letter on the right the base on the transferred strand). Weak mutants including 1T/1A, 2G/2C, 3C/3G, 3T/3A and 3A/3T gave resolved end levels ranging from 20 to 80% that of wild type. Base pair 4 and –1 mutants had no effect on resolved end formation. A representative set of time course experiments is presented in Figure 2B. In the wild-type time course, three products are visible, including ‘Tn-B’ (transposon bottom strand), ‘HP’ (hairpin) and ‘FD-T’ (flanking donor top strand). Tn-B is the most abundant product at the earliest time point. It accumulates throughout the time course, arising initially from first strand nicking and then later from hairpin resolution.

Fig. 3. Terminal base pair mutations do not inhibit first strand nicking. Levels of first strand nicking observed after 5 min in the experiment shown in Figure 2 are presented as a percentage of the initial transpososome formed for the indicated substrates. Values were calculated by summing Tn-B and HP (represents all molecules that have passed through the nicking stage) at 5 min and dividing this value by the amount of initial transpososome formed by the individual substrate. Error bars represent the standard deviation obtained from the variability in levels of first strand nicking between different experiments.

HP appears and then disappears, its level peaking at ~20 min, while FD-T first appears at the same time as HP and continues to accumulate over the course of the reaction. As previously described, appropriate treatment of these 'raw' data gives values for the amount of first strand nicking and resolved end formed at any time point (Figure 2C) (Kennedy *et al.*, 1998). In comparison with wild type, reactions with bp 1 and 2 mutants gave very low levels of hairpin and resolved end, even though high levels of first strand nicking were apparent. This latter point is indicated by the high levels of Tn-B formed (see also Figure 3). This is indicative of these mutations specifically interfering with steps after first strand nicking. The bp 3 mutant shown exhibited a less severe defect in formation of the hairpin and resolved end.

The extent of the defect in hairpin formation and resolution can be determined by calculating the average lifespan of the nicked and hairpin species, respectively. These values, which are equivalent to the inverse of the first order rate constant for the relevant steps, are calculated by dividing the area under the nicked curve for hairpin formation and the hairpin curve for hairpin resolution by the total number of molecules that have passed through these stages. For the mutant 1A/1T, the average lifespan for the nicked species was at least 24-fold longer than for wild type (i.e. ≥ 800 min versus 30 min). It was more difficult to obtain an accurate value for the lifespan of the hairpin itself for this mutant because of the low level of hairpin formed. Nevertheless, the absence of a noticeable decline in the hairpin level over a rather long period of time (from 80 to 480 min) together with only a small increase in hairpin formation (revealed by the increase in FD-T over the same time period) is indicative of a strong defect in hairpin resolution. These results are typical for 'strong' bp 1 and 2 mutants. In contrast, rates of

Table I. Effect of terminal base pair mutations on the kinetic parameters of transposon excision

Substrate	Nicked lifespan (min)	Hairpin lifespan (min)
Wild type	30	15
1G/1C	>800	ND
2A/2T	>800	ND
3T/3A	82	24
1C/1C	26	32
1G/1G	110	40
1C/1T	155	127
2A/2A	35	40
2T/2T	60	37
2T/2C	40	57

ND, not determined.

hairpin formation for bp 3 mutants were typically of the order of 3-fold slower than wild type, and inhibitory effects on hairpin resolution were relatively minor.

Evidence that the terminal base pair mutations do not significantly affect earlier steps in the reaction comes from two sources. First, the amount of initial transpososome formed by the mutants did not differ by >5% relative to wild type (data not shown). Secondly, as shown in Figure 3, the absolute levels of first strand nicking for the mutants did not differ by >15% relative to wild type at an early time point in the reactions. This is shown by plotting the percentage transpososome converted to the nicked species at 5 min, a time point at which there is very little turnover of the nicked intermediate.

Mismatch base pairs at the terminal two residues of the outside end facilitate hairpin formation

It is likely that hairpin formation requires some disruption of base pairing and/or base stacking in the transposon-donor junction region. Presumably, such a distortion in the DNA structure would require the formation of transposase contacts with these residues. The results presented above show that hairpin formation is particularly sensitive to base pair changes at the terminal two residues of the outside end. This is probably a reflection of a loss of contacts and/or change in the potential of the terminal region to be deformed. We asked if the introduction of base pair mismatches at these positions would facilitate hairpin formation despite the loss of DNA sequence information content by increasing the potential of the DNA to be deformed. Time course experiments were performed on a set of six mismatch substrates including 1C/1C, 1C/1T, 1G/1G, 2T/2T, 2A/2A and 2T/2C. All of these mismatches increased the rate of hairpin formation relative to related non-mismatch mutants by at least 4-fold and maximally by 30-fold. These data are summarized in Table I. At bp 1, it appears that a mismatch which maintains the normal base on the non-transferred strand provides the best substrate for hairpin formation: compare 1C/1C versus 1G/1G, the relatively low activity of 1C/1T might reflect differences in the thermodynamic stabilities of DNA containing a C/C versus a C/T mismatch (Allawi and SantaLucia, 1998; Peyret *et al.*, 1999). Alternatively, keeping the normal base on the transferred strand at bp 2 maintains the best hairpin activity (compare 2A/2A versus

2T/2T). These results provide evidence for base-specific contacts at position 1 of the non-transferred strand and position 2 of the transferred strand, respectively. In contrast, base pair mismatches at position 3 and the first residue of the flanking donor DNA did not significantly influence the kinetics of hairpin formation (data not shown). Taken together, these results provide evidence that hairpin formation is dependent on base-specific contacts and a highly localized DNA distortion at the tip of the outside end. This is supported by the DNA footprinting experiments presented below.

Footprinting of wild-type and mutant transpososomes

Terminal base pair mutations that interfere with hairpin formation might inhibit the ability of the transposon end to adopt a structure suitable for hairpin formation. To test this possibility, we have looked at the structure of the DNA at the transposon–donor junction in transpososomes containing either wild-type ends or ends with terminal base pair mutations by performing 1,10-phenanthroline-copper (OP-Cu) footprinting. The OP-Cu reagent intercalates into the minor groove and initiates oxidative attack on the closest deoxyribose ring, ultimately leading to phosphodiester backbone cleavage both 3' and 5' of the ring (Sigman *et al.*, 1991). This nuclease activity is sequence dependent but not nucleotide specific, and is therefore sensitive to any conformational variability of B-form DNA (Schaeffer *et al.*, 1996). OP-Cu binding to the minor groove and therefore DNA cleavage can be blocked by protein binding and/or narrowing of the minor groove, while enhancement of OP-Cu cleavage can occur in response to widening of the minor groove. In fact, OP-Cu cleavage has proven to be a useful probe for identifying local conformational changes relating to unpaired regions and defects in helical stacking in open promoter complexes through hypersensitivity to the reagent (Spassky, 1992). Importantly, OP-Cu can be used *in situ*, and therefore its use is not limited by transpososome assembly conditions. In contrast, other potentially useful probes of DNA structure such as potassium permanganate and hydroxyl radicals turned out to be difficult to use in footprinting experiments because their reactivity was limited under the conditions of transpososome assembly. Transpososomes assembled with different outside end substrates and containing a 5'-³²P label on either the non-transferred or transferred strand were applied to a 5% native polyacrylamide gel to separate transpososomes from other species. The entire gel was then treated with the OP-Cu footprinting reagent. After isolation of 'bound' and 'free' forms, the DNA was subjected to electrophoresis on a high resolution denaturing gel alongside a G-specific Maxam and Gilbert DNA sequence ladder.

The results of OP-Cu footprinting experiments for wild-type, 1A/1T and 2C/2G 'pre-cleavage' transpososomes are presented in Figure 4A. A prominent feature of the wild-type footprint is the broad region of OP-Cu hypersensitivity on both strands in the transposon–donor junction region (summarized in Figure 4B). The hypersensitivity extends from bp 1 to 7 on the non-transferred strand and from bp -3 to 6 on the transferred strand. Within this

region, enhancements of DNA cleavage by the OP-Cu nuclease range from 1.4- to 5.4-fold relative to free DNA. The appearance of this hypersensitivity is not dependent on the presence of a divalent metal ion (data not shown), although Crellin and Chalmers have reported a hypersensitivity to hydroxyl radicals that is divalent metal ion dependent right at the transposon–donor junction (P.Crellin and R.Chalmers, personal communication). Other features of the non-transferred strand footprint closely resemble those previously reported for a DNase I footprint of the IHF-containing pre-cleavage transpososome (Sakai *et al.*, 2000). The pattern of protection and hypersensitivity observed for the terminal 20 bp shows a phasing of ~10 bp. This is indicative of transposase occupying a single face of the DNA (assuming that positions of hypersensitivity reflect sites of protein contact). Past this region, where transposase- and IHF-binding sites converge, the phasing breaks down and contacts are observed on both faces of the helix.

The broad region of OP-Cu hypersensitivity in the junction region provides an indication that the DNA in this part of the pre-cleavage transpososome adopts a highly distorted structure. Significantly, comparison of the other pre-cleavage transpososome footprints in Figure 4 shows little difference between the terminal base pair mutants and wild type. This could be an indication that these mutants inhibit hairpin formation by affecting a step in the reaction not directly involved in deforming the structure of the junction region. However, an alternative possibility that cannot be ruled out at this time is that the OP-Cu nuclease is not detecting subtle changes in DNA structure caused by the mutations.

Subsequently, we looked at another situation where the terminal base pair mutations might influence the ability of the junction region to be distorted. A nick was introduced at the site where transposase normally makes the first nick. The rationale here was that an effect of base pair mutations on DNA structure might be revealed only under conditions most closely representing the defective stage in the reaction, i.e. conversion of the nicked form of the transpososome to the hairpin intermediate. Two separate transferred strand oligonucleotides were annealed to a single long non-transferred strand oligonucleotide to produce a substrate where the internucleotide bond at the site of first strand nicking was missing. Due to the presence of the strand break in the transferred strand, it was only practical to carry out footprinting in the junction region on the non-transferred strand. Footprints of free and bound forms of the nicked substrate are indicated with an asterisk (F* and B* in Figure 4A). Nicked forms of wild-type and bp 2 mutations gave roughly the same pattern of OP-Cu hypersensitivity as the corresponding pre-cleavage transpososomes. Conversely, nicked transpososomes with bp 1 mutations exhibited a significantly reduced level of OP-Cu reactivity in the junction region as indicated by the bracket (Figure 4A and data not shown). Thus, we conclude that the combination of a nick and a position 1 mutation is sufficient to prevent the junction region from adopting a distorted structure that is detectable by OP-Cu. This could be related to the strong defect in hairpin formation exhibited by the bp 1 mutants. In addition, this could be an indication that bp 1 and 2 mutations interfere with hairpin formation in different ways.

Fig. 4. OP-Cu DNA footprinting of pre-cleavage and nicked transpososomes. (A) Transpososomes were assembled with substrates as indicated in Figure 2 or using a substrate containing a nick at the transposon–donor junction of the transferred strand. A ^{32}P label was present at the 5' end on either the non-transferred or transferred strand. OP-Cu DNA footprinting was then carried out as described in Materials and methods. Equivalent amounts of bound (B) transpososome and free (F) DNA were then applied to an 8% denaturing polyacrylamide gel, alongside a G-cleavage lane prepared from the substrate DNA (not shown). Lanes representing footprinting reactions carried out on substrates containing a nick at the transposon–donor junction are labeled with an asterisk as F* or B* for free and bound, respectively. The numbered arrows indicate the base pair positions, with 1 being the first base pair of the transposon and –1 being the first base pair in the flanking donor DNA. (B) A summary of the footprint is presented for the wild-type pre-cleavage transpososome. Arrows indicate hypersensitive sites, with the length of the arrow being roughly proportional to the degree of hypersensitivity. Closed circles indicate positions of strong protection.

The effect of terminal base pair mutations on strand transfer

First strand nicking requires entry of only the transferred strand into the active site, whereas hairpin formation necessarily involves entry of both the transferred and non-transferred strands (Kennedy *et al.*, 2000). It follows that the terminal base pair mutations that strongly inhibit hairpin formation might influence primarily interactions between transposase and the non-transferred strand. If this were the case, these mutations would not be expected to influence the target strand transfer reaction (unless the

target and non-transferred strand must occupy the same space in the active site). We have addressed this issue by performing target strand transfer reactions with pre-cleaved substrates. It is necessary to use pre-cleaved substrates because target strand transfer requires complete removal of flanking DNA from both transposon ends.

^{32}P -labeled pre-cleaved substrates containing wild-type sequence or bp 1 or 2 mutations were assembled into 'double end break complexes' (DEBCs) and then mixed with an unlabeled 30 bp target DNA fragment containing the HisG1 hotspot for Tn10 insertion. To initiate

DEBc–target interactions, IHF was stripped from the transpososome with heparin and incubations were carried out for 1 h either in the presence of Mg^{2+} or in the absence of a divalent metal ion. The reactions were then applied to a 5% native polyacrylamide gel. If the DEBc is able to form a stable association with the target DNA (i.e. form a target capture complex, ‘TC-30’), a new species with a decreased mobility relative to the DEBc will be observed. We find in this assay that the amount of TC-30 formed after 1 h is roughly equivalent to the amount of strand transfer that has occurred if Mg^{2+} is present. This conclusion is based on the observation that the ratio of TC-30 to total DNA in the native gel is roughly equivalent to the ratio of strand transfer product to total DNA in a denaturing gel (Junop and Haniford, 1997 and data not shown). The results of the native gel analysis presented in Figure 5 demonstrate that all four of the mutants tested form much lower levels of TC-30, and thus of strand transfer product, than wild type (at least 5-fold lower). This supports the idea that the terminal base pair mutations interfere with hairpin formation by altering contacts between transposase and both the transferred and non-transferred strands.

Fig. 5. Terminal base pair mutations inhibit strand transfer. Double end break complexes (DEBcs) were assembled by mixing the indicated ^{32}P -labeled outside end substrates with IHF and transposase. After 2 h at 20°C, HisG1 target DNA was added to a final concentration of 37.5 pmol/μl and incubation was continued for an additional 20 min. Heparin was then added to remove IHF and, where indicated, $MgCl_2$ was added to promote strand transfer. Reactions were then applied to a native 5% polyacrylamide gel. In reactions where Mg^{2+} was included, the species labeled TC-30 contains at least 90% strand transfer products (data not shown). The positions of free outside end, IHF complex (IHFc), DEBc, TC-OE (a non-sequence-specific target capture complex in which the free outside end is the target DNA) and TC-30 (a sequence-specific target capture complex in which the HisG1 fragment is the target DNA) are shown.

Transposase mutants that specifically affect hairpin formation

As discussed in the Introduction, the structure of the Tn5 transpososome has revealed the identity of a number of amino acid residues that appear to be involved in distorting the DNA structure at the tip of the transposon DNA and are therefore good candidates for a role in hairpin formation. This includes Y319, W298 and Y237. Notably, Y319 is part of the YREK motif, a signature sequence of the IS4 family. We have asked if mutations at what are predicted to be the corresponding residues to Y319 and W298 in Tn10 transposase affect hairpin formation. A primary sequence alignment for part of the Tn5 and Tn10 transposase proteins is shown in Figure 6. From this alignment, we predict that Tn10 W265 is the counterpart of Tn5 W298. Tn10 Y285 was previously identified as a component of the YREK motif (Mahillon *et al.*, 1985). A Tn10 counterpart for Y237 is not obvious from the alignment (not shown). The properties of Tn10 Y285 and W265 mutants are described below.

Two mutations were made initially at the Y285 position, YS285 and YF285, to test the possibility that this highly conserved residue is involved in making a phosphotyrosine linkage during catalysis (Haniford *et al.*, 1989). This possibility was eliminated by the observation that while the YS285 mutation conferred a severe transposition defect, the YF285 mutation did not affect the transposition frequency. We have now purified the YS285 protein and tested it *in vitro* for transposon excision. The results shown in Figure 7A demonstrate that this mutant is defective for transposon excision and that, as for the terminal base pair mutations, this defect is manifested specifically at the hairpin formation stage.

We have also looked at the effect of two different position 265 substitutions, WV265 and WL265, on transposition *in vitro*. While both mutants form hairpins, it is evident, particularly from the WV265 mutant, that there is a relaxation in the specificity of the hairpin formation reaction. This is revealed most clearly from the appearance of FD-T*, a flanking top strand fragment that is one nucleotide shorter than FD-T. The evidence that FD-T* results from 3' OH attack at the 'incorrect' phosphate bond is that two forms of transposon end hairpin are detected, one of which is one nucleotide longer than the standard hairpin (Figure 7B). For WV265, ~40% of the hairpin formed is the longer hairpin species (HP*).

While we have not performed a detailed kinetic analysis on the WV265 and WL265 reactions, it is apparent that the kinetics of hairpin formation differ from wild type. Hairpin levels reach a maximum between 80 and 160 min and at 160 min for WL265 and WV265, respectively, instead of 20 min. The hairpin species also persists for

Fig. 6. A partial alignment of the amino acid sequences of the Tn5 and Tn10 transposase proteins. Amino acid identities and similarities are indicated by shaded and open boxes, respectively. The position of the Y-R-3-E-6-K motif, a 'signature' motif of the IS4 family of transposase proteins, is highlighted. The arrows indicate the Tn10 transposase residues mutated in this study.

Fig. 7. Effect on transposon excision of amino acid substitutions at conserved aromatic amino acids. **(A)** Kinetic analysis of transposon excision for wild type and three mutant forms of transposase was carried out as described in Figure 2. The novel species FD-T* is detected only in reactions with position 265 mutants and is a result of the formation of a hairpin species that is one nucleotide longer than the 'standard' hairpin (see text). **(B)** Aliquots from the 20 and 160 min time points obtained in (A) were run for an additional hour on an 8% denaturing gel. A portion of the gel containing the standard 128 nt hairpin (HP) species is shown for the indicated transposase mutants. The 129 nt hairpin (HP*) is readily visible for the WV265 mutant. **(C)** Transposon excision was performed with a 3'-end-labeled substrate. A portion of the denaturing gel is shown that displays the non-transferred strand cleavage products, Tn-T and Tn-T*. The latter arises from resolution of the alternative hairpin species (HP*). This experiment provides a direct measure of the amount of resolved end formed, i.e. resolved end = Tn-T + Tn-T*.

longer periods of time. A direct measure of resolved end formation was also made by 3' end labeling the substrate as shown in Figure 7C. Here, the sum of Tn-T and Tn-T* (the longer form of Tn-T produced by cleaving the top strand one nucleotide into the flank) represents the total resolved end formed. At 12 h, resolved end values for WL265 and WV265 were ~90 and 30% that of wild type, respectively. The greater abundance of Tn-T relative to Tn-T* could mean either that the longer hairpin is resolved less efficiently relative to the standard hairpin or that at a

relatively high frequency this hairpin is resolved off-center to generate Tn-T.

Discussion

The DNA sequence at the tip of the transposon dictates the efficiency of hairpin formation

We have shown that base pair mutations at positions 1–3 inhibit hairpin formation to varying degrees without

significantly affecting initial transpososome assembly or the introduction of the first nick at the transposon–donor junction. Mutations at bp 1 and 2 displayed the strongest inhibition, increasing the average lifespan of the nicked species at least 24-fold, while mutations at bp 3 had much less severe effects. Conversely, mutations on either side of the terminal three base pairs had no effect on hairpin formation. It follows that DNA sequence information in only a highly localized region at the tip of the transposon end appears to be important for hairpin formation. In contrast, the DNA sequence requirements in the same region are fairly relaxed for transpososome assembly and first strand nicking. Thus, the substrate-binding pocket of the active site must be able to accommodate any of the four nucleotides at the terminal three residues at early stages of the reaction. However, it is unlikely that this condition applies to later stages such as hairpin formation and resolution.

The DNA in the transposon–donor junction region adopts a distorted structure

Our DNA footprinting of transpososomes with OP-Cu, a DNA cleavage reagent whose binding to DNA and therefore cleavage activity is extremely sensitive to the width of the minor groove, has provided evidence for a distorted DNA structure in the transposon–donor junction region of the transpososome. We have shown that there is a broad region of hypersensitivity to OP-Cu spanning bp 7 to –3. This pattern of OP-Cu hypersensitivity is reminiscent of that observed for the open complex of the *Escherichia coli lacUV5* promoter prior to elongation (Spassky, 1992). The junction region of transpososomes in other transposition systems, including Mu and Tn5, is also hypersensitive to cleavage by chemical nucleases, indicating the presence of an altered DNA structure in this location (Lavoie *et al.*, 1991; Savilahti *et al.*, 1995; Wang *et al.*, 1996; Bhasin *et al.*, 2000). In Tn5, this distortion in DNA structure coincides with a bend in the DNA backbone close to the transposon–donor junction (York and Reznikoff, 1997). In addition, a very similar pattern of OP-Cu hypersensitivity to that reported here has been observed for the signal end complex in V(D)J recombination (Akamatsu and Oettinger, 1998). There is also strong evidence for a distorted DNA structure being formed at the terminus of human immunodeficiency virus (HIV) DNA prior to cleavage of the viral ends by integrase (Scottoline *et al.*, 1997).

In the Tn10 system, we have shown that terminal base pair mutations do not prevent the DNA distortion in the junction region from taking place (except if a nick and a bp 1 mutation are present). There are several possible interpretations of this result with regard to hairpin formation. (i) The observed DNA distortion may not have anything to do with hairpin formation. (ii) The DNA distortion may be linked to hairpin formation and its occurrence dependent on base-specific contacts with the terminal residues; however, the OP-Cu reagent might not be able to detect subtle structural differences conferred by the terminal base pair mutations. (iii) Again the DNA distortion is linked to hairpin formation but the distortion itself does not require base-specific contacts with the terminal residues. In this case, sequence-specific contacts with the terminal residues would probably be confined to a

step subsequent to the DNA distortion, perhaps the movement of the non-transferred strand into the active site prior to hairpin formation. In this scenario, the effect seen by the bp 1 mutations in the presence of the nick may suggest that base-specific contacts, although initially not involved in generating the distortion, are formed after first strand nicking and are required to maintain the distortion, providing a suitable substrate for hairpin formation. However, it is also possible that first strand nicking as well as hairpin formation require the DNA distortion. This is unlikely for the second scenario because nicking does not exhibit sequence specificity.

One way in which the DNA distortion could occur in a sequence-independent manner is through a ‘protein wedge’ mechanism. There are now several examples where a DNA-binding protein inserts a wedge of hydrophobic amino acids into a DNA helix (reviewed in Werner *et al.*, 1996). Insertion of the wedge typically results in the flipping of a single base out of the helix as the wedge replaces this base, and this is accompanied by a substantial distortion in the local DNA structure. The protein wedge interacts with adjacent bases through non-specific stacking interactions, although the positioning of the wedge is typically directed by sequence-specific interactions outside the zone of insertion. Interestingly, in the Tn5 synaptic complex, amino acid intercalation is observed and this coincides with base flipping and the introduction of a sharp kink near the 5′ terminus of the non-transferred strand. This kink results in the transferred and non-transferred strands being positioned closely together, as would be required for hairpin formation (Davies *et al.*, 2000).

Hairpin formation requires sequence-specific contacts at the termini of both transferred and non-transferred strands

One or both DNA strands at the transposon terminus might provide contact points necessary for hairpin formation. A way of distinguishing between these possibilities is to ask if the terminal base pair mutations interfere with strand transfer when present on pre-cleaved ends. The logic here is that if terminal base pair mutations inhibit hairpin formation by interfering only with contacts on the non-transferred strand, then there is no a priori reason for strand transfer to be blocked as this step may be dependent exclusively on contacts made with the transferred strand. We found that four of four terminal base pair mutations tested in the context of a double end break transpososome showed strong defects in strand transfer. These results therefore support the idea that hairpin formation is dependent on transposase making terminal base pair contacts with both transferred and non-transferred strands. This is supported further by the results of our base pair mismatch analysis. However, an important caveat to the strand transfer experiment is that it is not completely clear if the above mutants interfere with sequence-specific target capture. The mutants all form the non-sequence-specific target capture complex (TC-OE) but we do not have direct evidence that they are able to support the transition to the sequence-specific target capture complex (TC-30), a prerequisite for strand transfer, i.e. there was a failure to detect the TC-30 complex in the presence of Mg²⁺ for these mutants. While this may appear to provide

strong evidence for a defect in sequence-specific target capture, it is possible that it is the failure to carry out strand transfer efficiently that significantly reduces the amount of TC-30 that can be detected.

The role of terminal base pairs in related DNA transposition systems

There is indirect evidence that terminal base pair mutations in the IS903 system have effects similar to those described here for the Tn10 system. IS903 differs from Tn10 in that it is able to transpose by both non-replicative and replicative transposition, although the majority of events normally take place by the former mechanism (Weinert *et al.*, 1984). It has been shown that specific mutations at bp 1 dramatically increase the frequency of co-integrate formation, the product of replicative transposition. It has been suggested that this could be the result of the mutation blocking full transposon excision by interfering with hairpin formation, in which case a nicked intermediate could go on to carry out strand transfer prior to cleavage of the second DNA strand (Tavakoli and Derbyshire, 1999).

In the Tn5 system, mutational studies have shown that base pair changes at the tip of the outside end strongly inhibit DNA transposition, although it has not yet been established what step(s) in transposition is/are affected (Makris *et al.*, 1988). However, based on the structure of the Tn5 synaptic complex, it is likely that mutations in this region will affect hairpin formation (Davies *et al.*, 2000). This structure has also revealed the presence of base-specific transposase contacts at positions 2 and 3 of the non-transferred strand and position 3 of the transferred strand.

In V(D)J recombination, binding sites for the RAG recombinase are thought to be located primarily in the signal end, although the DNA hairpin is formed at the coding end (Fugmann *et al.*, 2000). Similarly to results reported here, mutations at bp 1 and 2 of the heptamer sequence of the signal end (analogous to bp 1 and 2 in the Tn10 end) have much stronger effects on hairpin formation than on first strand nicking (Cuomo *et al.*, 1996; Ramsden *et al.*, 1996; Li *et al.*, 1997). In contrast, base pair mismatches at a greater number of positions, both in the signal end and into the coding flank, were found to facilitate hairpin formation in the V(D)J system (Cuomo *et al.*, 1996; Ramsden *et al.*, 1996). This could be an indication that a more substantial DNA distortion is required for hairpin formation in V(D)J recombination relative to Tn10 transposition.

The effects of terminal base pair mutations have also been studied extensively in the retroviral integration systems. Strong inhibition of viral cleavage has been reported for any base pair change at the invariant 'CA' dinucleotide in the HIV system in the presence of Mg²⁺, the physiological divalent metal ion for this reaction (Esposito and Craigie, 1998). This is strong evidence for base-specific contacts being made at these positions. Direct evidence for a base-specific contact at position 1 of the transferred strand has been generated through chemical cross-linking; K159 was shown to cross-link strongly to the A at this position (Jenkins *et al.*, 1997). Interestingly, this contact was shown to be dispensable when a mismatch is present at this position (Esposito and

Craigie, 1998). This suggests that the contact is normally important for distorting the DNA structure at the end of the retrovirus.

Conserved aromatic amino acids play an important role in hairpin formation

As discussed above, a number of amino acid residues stand out as good candidates for being involved in hairpin formation in the Tn5 system. Using the information from the Tn5 system as a guide, we tested two positions in Tn10 transposase for a role at this step. Both Y285 and W265 are shown to function in hairpin formation. The phenotype of the YS285 mutant closely resembles that of the terminal base pair mutations in that hairpin formation is strongly inhibited but transpososome assembly and first strand nicking occur normally. In contrast, WV265 has a novel phenotype in that it relaxes the specificity of the hairpin formation reaction. We think the most straightforward way of accounting for this is that the valine substitution affects the conformation of the non-transferred strand so that there is an increased probability of the 'incorrect' phosphate group entering the active site subsequent to first strand nicking. The defect in hairpin formation conferred by the YS285 mutation may also result from inappropriate positioning of the non-transferred strand in the active site. The equivalent residue in Tn5 transposase contacts the non-transferred strand at the internucleotide bond between bp 1 and 2 (Davies *et al.*, 2000). In HIV integrase, the equivalent α -helix to that containing the YREK motif, $\alpha 4$ (Dyda *et al.*, 1994), contains K159, the residue shown to contact bp 1 of the transferred strand (Heuer and Brown, 1997; Jenkins *et al.*, 1997). Taken together, these results are suggestive of functional conservation between HIV integrase, Tn5 and Tn10 transposase to an extent beyond that already documented (Mizuuchi, 1997; Davies *et al.*, 2000).

In addition to providing further insight into structure-function relationships in the Tn10 system, the above results are significant for some additional reasons. First, prior to this study, there were no functional data to support the possibility that the tyrosine of the YREK motif actually plays an important role in hairpin formation. Given the specialized role of this residue and the conservation of this motif throughout the IS4 family, it may be that transposition through a hairpin mechanism is another unifying feature of this family of transposable elements. Secondly, the above results provide an early indication that the active sites of the Tn10 and Tn5 transposase proteins have a similar organization despite the low degree of amino acid sequence identity. Finally, identifying mutants in the Tn10 system that are very likely to be involved in positioning the non-transferred strand in the active site provides us with new tools to investigate the position target DNA occupies in the active site. At this point, it is not clear if the non-transferred strand occupies the same position in the active site as the target DNA (Kennedy *et al.*, 2000).

Materials and methods

DNA substrates

Outside end substrates used in the assembly of the pre-cleavage transpososomes were generated by annealing a 78 nucleotide (nt) 'non-transferred strand' oligonucleotide (78-nts) and an 80 nt 'transferred

strand' oligonucleotide (80-ts). Only the DNA sequences of the wild-type versions of these oligonucleotides are given here: 78-nts WT, 5'-TAC-ACACACTCAGCGCTGATGAATCCCCTAATGATTTTGGTAAAA-ATCATTAAGTTAAGGTGGATACACATCTTGTC-3' and 80-ts WT, 5'-TATGACAAGATGTGTATCCACCTTAACCTTAATGATTTTACCA-AAAATCATTAGGGGATTCATCAGCGCTGAGTGTGTGTA-3'.

Outside end substrates used in the assembly of the nicked transpososomes were generated by annealing three oligonucleotides, including 78-nts (above) and two bottom strand oligonucleotides, a 15 nt 'donor' oligonucleotide (15-ts-d) and a 65 nt 'transposon' oligonucleotide (65-ts-t). The DNA sequences of the wild-type forms of these oligonucleotides are: 15-ts-d WT, 5'-CGCTGAGTGTGTGTA-3' and 65-ts-t WT, 5'-TATGACAAGATGTGTATCCACCTTAACCTTAATGATTTTACCAAAATCATTAGGGGATTCATCAG-3'.

Pre-cleaved outside end substrates were generated by annealing a 63 nt non-transferred strand transposon oligonucleotide (63-nts-t) and a 65 nt transferred strand transposon oligonucleotide (65-ts-t). The DNA sequence of the wild-type form of the 63-nts-t oligonucleotide is: 5'-CTGATGAATCCCCTAATGATTTTGGTAAAAATCATTAAAGTTAAGGTGGATACACATCTTGTC-3'.

The 30 bp HisG1 target substrate was assembled by annealing two 30 nt oligonucleotides containing an off-center HisG1 consensus site. The sequences of these strands were: HisG1-30mer-top, 5'-TAATTTACACACTCAGCGCTGATTGCGAT-3'; and HisG1-30mer-bottom, 5'-CATCGCAATCAGGCGCTGAGTGTCTAAATT-3'. Oligonucleotides were purchased from Perkin Elmer Applied Biosystems and were gel purified prior to use. Yields of full-length deoxyoligonucleotides were quantified by measuring the A₂₆₀ where an absorbance of 1 corresponds to 35 µg/ml.

Assembly of substrates and transpososomes

Oligonucleotides were ³²P labeled at their 5' or 3' ends with polynucleotide kinase (NEB) and [γ-³²P]ATP (NEN) or terminal transferase (MBI Fermentas) and [α-³²P]ddATP (Amersham Pharmacia Biotech), respectively, and annealed to form double-stranded substrates. Typically, 20 pmol of the transferred strand and non-transferred strand oligonucleotides were mixed in 1× DNA buffer, heated to 85°C and then cooled to 25°C over a period of several hours; for assembly of the nicked substrate, a 4-fold excess of 15-ts-d was used. All substrates were stored in 1× DNA buffer in a volume of 100 µl.

Transpososomes were formed by mixing 0.2 pmol of DNA substrate, 0.35 pmol of IHF and 0.3 pmol of transposase in a final volume of 20 µl. IHF was added prior to the addition of transposase and the assembly reaction was incubated at 20°C for 2 h. Buffer conditions were as described in Sakai *et al.* (1995). For kinetic, strand transfer and footprinting analyses, transpososome assembly reactions were carried out on a 10×, 3× and a 50× scale with respect to the reaction volume, respectively. Transposase and IHF proteins were purified as previously reported (Chalmers and Kleckner, 1994).

Kinetic analyses of donor cleavage

A 10 µl aliquot of each assembly reaction was analyzed on a native 5% polyacrylamide gel prior to the addition of MgCl₂ in order to determine the efficiency of transpososome formation. Gels were analyzed by phosphorimaging. Assembly efficiency (percentage of total input substrate converted to transpososome) was calculated by dividing the peak area of transpososome by the total input substrate (calculated from the peak areas of transpososome, IHF complex, free DNA and material trapped in wells). DNA was then eluted from the transpososome-containing gel slice and applied to an 8% denaturing polyacrylamide gel where transferred and non-transferred strand oligonucleotides were separated. From the corresponding peak areas, labeling efficiencies of the two oligonucleotides were determined. MgCl₂ was added to the remaining portion of each assembly reaction to a final concentration of 5 mM and, at the indicated time points, 10 µl aliquots were removed and phenol extracted to terminate the reaction. DNA was recovered by ethanol precipitation, resuspended in denaturing gel load mix [0.3% (w/v) bromophenol blue, 0.3% xylene cyanol, 10 mM EDTA in 97.5% (v/v) deionized formamide] and applied to an 8% denaturing polyacrylamide gel. Gels were analyzed by phosphorimaging, the levels of individual cleavage products were calculated initially as a percentage of total substrate, values were corrected for differences in loading efficiency and labeling efficiency and then converted to a percentage of total transpososome formed.

Strand transfer assay

Target capture complexes were generated by mixing 20 µl of the DEBcs assembled from pre-cleaved transposon ends with 37.5 pmol of 30 bp target substrate containing the HisG1 hotspot. Reactions were incubated for 20 min at 20°C. A 1 µl aliquot of heparin (0.005 mg/ml) and then, where indicated, 1 µl of 0.1 mM MgCl₂ was added to the reactions to initiate strand transfer. Incubation was continued for an additional hour at 20°C. Reactions were then analyzed on a 5% native polyacrylamide gel.

OP-Cu footprinting

Pre-cleavage or nicked transpososomes were assembled in a 50× reaction using substrates that were 5' end-labeled on either the transferred or non-transferred strand. Following a 2 h incubation at 20°C, CaCl₂ was added to a final concentration of 5 mM and incubation was continued for an additional 2 h. The entire sample was then loaded onto a 5% native polyacrylamide gel where transpososomes were separated from free DNA and IHF complex. *In situ* OP-Cu footprinting was then performed as described by Sigman *et al.* (1991). Following chemical treatment, gels were subjected to autoradiography and DNA was extracted from the relevant gel slices into 1× TE buffer by the crush and soak method. DNA was purified further by spin column chromatography using a Biospin-6 column (Bio-Rad), and then concentrated using a Millipore-10K NMWL ultrafree filter. Finally, the DNA was precipitated with ethanol, resuspended in denaturing loading dye, heated to 90°C for 2 min and loaded onto an 8% denaturing polyacrylamide gel alongside a Maxam-Gilbert G-reaction ladder generated from the same end-labeled substrate DNA (Maxam and Gilbert, 1980). Approximately 100 000 c.p.m. was loaded per lane. Band intensities were determined by phosphorimaging.

Acknowledgements

We are grateful to G.Chaconas, C.Brandl, P.Pribil and C.Coros for comments on the manuscript, A.Dar and B.Jardine for their contributions to closely related projects, and M.Junop for providing us with an alignment between Tn10 and Tn5 transposase proteins. This research was funded by a grant to D.B.H. from the Canadian Institutes of Health Research. J.S.A. was supported by an MRC Studentship.

References

- Akamatsu,Y. and Oettinger,M.A. (1998) Distinct roles of RAG1 and RAG2 in binding the V(D)J recombination signal sequences. *Mol. Cell. Biol.*, **18**, 4670–4678.
- Allawi,H.T. and SantaLucia,J.,Jr (1998) Thermodynamics of internal C:T mismatches in DNA. *Nucleic Acids Res.*, **26**, 2694–2701.
- Bhasin,A., Goryshin,I.Y. and Reznikoff,W.S. (1999) Hairpin formation in Tn5 transposition. *J. Biol. Chem.*, **274**, 37021–37029.
- Bhasin,A., Goryshin,I.Y., Steiniger-White,M., York,D. and Reznikoff,W.S. (2000) Characterization of a Tn5 pre-cleavage synaptic complex. *J. Mol. Biol.*, **302**, 49–63.
- Chalmers,R.M. and Kleckner,N. (1994) Tn10/IS10 transposase purification, activation, and *in vitro* reaction. *J. Biol. Chem.*, **269**, 8029–8035.
- Coen,E., Robbins,T., Almeida,J., Hudson,A. and Carpenter,R. (1989) Consequences and mechanisms of transposition in *Antirrhinum majus*. In Berg,D.E. and Howe,M.M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 413–436.
- Colot,V., Haedens,V. and Rossignol,J.L. (1998) Extensive, nonrandom diversity of excision footprints generated by Ds-like transposon Ascot-1 suggests new parallels with V(D)J recombination. *Mol. Cell. Biol.*, **18**, 4337–4346.
- Cuomo,C.A., Mundy,C.L. and Oettinger,M.A. (1996) DNA sequence and structure requirements for cleavage of V(D)J recombination signal sequences. *Mol. Cell. Biol.*, **16**, 5683–5690.
- Davies,D.R., Goryshin,I.Y., Reznikoff,W.S. and Rayment,I. (2000) Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science*, **289**, 77–85.
- Derbyshire,K.M., Hwang,L. and Grindley,N.D.F. (1987) Genetic analysis of the interaction of the insertion sequence IS903 transposase with its terminal inverted repeats. *Proc. Natl Acad. Sci. USA*, **84**, 8049–8053.
- Dyda,F., Hickman,A.B., Jenkins,T.M., Engelman,A., Craigie,R. and Davies,D.R. (1994) Crystal structure of the catalytic domain of HIV-1

- integrase: similarity to other polynucleotidyl transferases. *Science*, **266**, 1981–1985.
- Esposito,D. and Craigie,R. (1998) Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein–DNA interaction. *EMBO J.*, **17**, 5832–5843.
- Fedoroff,N.V. (1989) Maize transposable elements. In Berg,D.E. and Howe,M.M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 375–411.
- Fugmann,S.D., Lee,A.I., Shockett,P.E., Villey,I.J. and Schatz,D.G. (2000) The RAG proteins and V(D)J recombination: complexes, ends and transposition. *Annu. Rev. Immunol.*, **18**, 495–527.
- Haniford,D.B. and Kleckner,N. (1994) Tn10 transposition *in vivo*: temporal separation of cleavages at the two transposon ends and roles of terminal basepairs subsequent to interaction of ends. *EMBO J.*, **13**, 3401–3411.
- Haniford,D.B., Chelouche,A.R. and Kleckner,N. (1989) A specific class of IS10 transposase mutants are blocked for target site interactions and promote formation of an excised transposon fragment. *Cell*, **59**, 385–394.
- Heuer,T.S. and Brown,P.O. (1997) Mapping features of HIV-1 integrase near selected sites of viral and target DNA molecules in an active enzyme–DNA complex by photo-cross-linking. *Biochemistry*, **36**, 10655–10665.
- Huisman,O., Errada,P.R., Signon,L. and Kleckner,N. (1989) Mutational analysis of IS10's outside end. *EMBO J.*, **8**, 2101–2109.
- Jenkins,T.M., Esposito,D., Engelman,A. and Craigie,R. (1997) Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking. *EMBO J.*, **16**, 6849–6859.
- Jilk,R.A., York,D. and Reznikoff,W.S. (1996) The organization of the outside end of transposon Tn5. *J. Bacteriol.*, **178**, 1671–1679.
- Junop,M.S. and Haniford,D.B. (1997) Factors responsible for target site selection in Tn10 transposition: a role for the DDE motif in target DNA capture. *EMBO J.*, **16**, 2646–2655.
- Kennedy,A.K. (1999) Mechanistic aspects of Tn10 transposition. PhD thesis, University of Western Ontario, London, Canada.
- Kennedy,A.K., Guhathakurta,A., Kleckner,N. and Haniford,D.B. (1998) Tn10 transposition via a DNA hairpin intermediate. *Cell*, **95**, 125–134.
- Kennedy,A.K., Haniford,D.B. and Mizuuchi,K. (2000) Single active site catalysis of the successive phosphoryl transfer steps by DNA transposases: insights from phosphorothioate stereoselectivity. *Cell*, **101**, 295–305.
- Kleckner,N., Chalmers,R.M., Kwon,D., Sakai,J. and Bolland,S. (1996) *Tn10 and IS10 Transposition and Chromosome Rearrangements: Mechanism and Regulation In Vivo and In Vitro*. Springer, Berlin, Germany.
- Lavoie,B.D., Chan,B.S., Allison,R.G. and Chaconas,G. (1991) Structural aspects of a higher order nucleoprotein complex: induction of an altered DNA structure at the Mu–host junction of the Mu type I transpososome. *EMBO J.*, **10**, 3051–3059.
- Li,W., Swanson,P. and Desiderio,S. (1997) RAG-1- and RAG-2-dependent assembly of functional complexes with V(D)J recombination substrates in solution. *Mol. Cell. Biol.*, **17**, 6932–6939.
- Mahillon,J., Seurinck,J., Van Rompuy,L., Delcour,J. and Zabeau,M. (1985) Nucleotide sequence and structure organization of an insertion sequence element (IS231) from *Bacillus thuringiensis* strain berliner 1715. *EMBO J.*, **4**, 3895–3899.
- Makris,J.C., Nordmann,P.L. and Reznikoff,W.S. (1988) Mutational analysis of insertion sequence 50 (IS50) and transposon 5 (Tn5) ends. *Proc. Natl Acad. Sci. USA*, **85**, 2224–2228.
- Maxam,A.M. and Gilbert,W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.*, **65**, 499–560.
- McBlane,J.F., van Gent,D.C., Ramsden,D.A., Romeo,C., Cuomo,C.A., Gellert,M. and Oettinger,M.A. (1995) Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell*, **83**, 387–395.
- Mizuuchi,K. (1997) Polynucleotidyl transfer reactions in site-specific DNA recombination. *Genes Cells*, **2**, 1–12.
- Peyret,N., Senevieratne,P.A., Allawi,H.T. and SantaLucia,J., Jr (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A:A, C:C, G:G and T:T mismatches. *Biochemistry*, **38**, 3468–3477.
- Ramsden,D.A., McBlane,J.F., van Gent,D.C. and Gellert,M. (1996) Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. *EMBO J.*, **15**, 3197–3206.
- Sakai,J. and Kleckner,N. (1996) Two classes of Tn10 transposase mutants that suppress mutations in the Tn10 terminal inverted repeat. *Genetics*, **144**, 861–870.
- Sakai,J., Chalmers,R.M. and Kleckner,N. (1995) Identification and characterization of a pre-cleavage synaptic complex that is an early intermediate in Tn10 transposition. *EMBO J.*, **14**, 4374–4383.
- Sakai,J.S., Kleckner,N., Yang,X. and Guhathakurta,A. (2000) Tn10 transpososome assembly involves a folded intermediate that must be unfolded for target capture and strand transfer. *EMBO J.*, **19**, 776–785.
- Savilahti,H., Rice,P.A. and Mizuuchi,K. (1995) The phage Mu transpososome core: DNA requirements for assembly and function. *EMBO J.*, **14**, 4893–4903.
- Schaeffer,F., Rimsky,S. and Spassky,A. (1996) DNA-stacking interactions determine the sequence specificity of the deoxyribonuclease activity of 1,10-phenanthroline-copper ion. *J. Mol. Biol.*, **260**, 523–539.
- Scottoline,B.P., Chow,S., Ellison,V. and Brown,P.O. (1997) Disruption of the terminal basepairs of retroviral DNA during integration. *Genes Dev.*, **11**, 371–382.
- Sigman,D.S., Kuwabara,M.D., Chen,C.-H.B. and Bruce,T.W. (1991) Nuclease activity of 1,10-phenanthroline-copper in study of protein–DNA interaction. *Methods Enzymol.*, **208**, 414–433.
- Spassky,A. (1992) Passage of RNA polymerase from open complex to elongation mode at the *Escherichia coli lacUV5* promoter: nucleolytic hypersensitivity as a probe for complex conformational changes. *Biochemistry*, **31**, 10502–10509.
- Tavakoli,N.P. and Derbyshire,K.M. (1999) IS903 transposase mutants that suppress defective inverted repeats. *Mol. Microbiol.*, **31**, 1183–1195.
- van Gent,D.C., Mizuuchi,K. and Gellert,M. (1996) Similarities between initiation of V(D)J recombination and retroviral integration. *Science*, **271**, 1592–1594.
- Wang,Z., Namgoong,S.-Y., Zhang,X. and Harshey,R.M. (1996) Kinetic and structural probing of the precleavage synaptic complex (type 0) formed during phage Mu transposition. *J. Biol. Chem.*, **271**, 9619–9626.
- Weinert,T.W., Derbyshire,K., Highson,F.M. and Grindley,N.D.F. (1984) Replicative and conservative transpositional recombination of insertion sequences. *Cold Spring Harb. Symp. Quant. Biol.*, **49**, 251–260.
- Werner,M.H., Gronenborn,A.M. and Clore,G.M. (1996) Intercalation, DNA kinking, and the control of transcription. *Science*, **271**, 778–784.
- York,D. and Reznikoff,W.S. (1997) DNA binding and phasing analyses of Tn5 transposase and a monomeric variant. *Nucleic Acids Res.*, **25**, 2153–2160.

Received February 19, 2001; revised and accepted April 12, 2001