Published in final edited form as:

Adv Drug Deliv Rev. 2010 September 30; 62(12): 1187–1195. doi:10.1016/j.addr.2010.06.006.

DDE Transposases: Structural Similarity and Diversity

Irina V. Nesmelova^{1,2} and Perry B. Hackett³

¹ Department of Physics and Optical Science, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223

² Department of Biochemistry, Molecular Biology & Biophysics, University of Minnesota, 321 Church Street, Minneapolis, MN 55455

³ Department of Genetics, Cell Biology & Development, University of Minnesota, 321 Church Street, Minneapolis, MN 55455

Abstract

DNA transposons are mobile DNA elements that can move from one DNA molecule to another and thereby deliver genetic information into human chromosomes in order to confer a new function or replace a defective gene. This process requires a transposase enzyme. During transposition DD[E/D]-transposases undergo a series of conformational changes. We summarize the structural features of DD[E/D]-transposases for which three-dimensional structures are available and that relate to transposases, which are being developed for use in mammalian cells. Similar to other members of the polynucleotidyl transferase family, the catalytic domains of DD[E/D]-transposases share a common feature: an RNase H-like fold that draws three catalytically active residues, the DDE motif, into close proximity. Beyond this fold, the structures of catalytic domains vary considerably, and the DD[E/D]-transposases display marked structural diversity within their DNA-binding domains. Yet despite such structural variability, essentially the same end result is achieved.

Keywords

DNA transposon; DD[E/D]-transposase; transposase; structure

The advances in the last two decades of the 20th century set the stage for a revolution in the delivery of genetic based therapeutics. Genetic medicines offer the possibilities of permanent solutions to chronic and acute medical problems such that the normal state of the individual is restored naturally rather than through odd therapeutic compounds that often are uneven in their activities and accompanied by undesirable side effects. Moreover, the levels of the natural biological product should be relatively constant at physiologically effective levels, rather than cycling through of high and low concentrations that result from taking of drugs or other therapeutics at periodic intervals. For the same reason, gene and cell-based therapies have the potential to provide a marked clinical and economic improvement over infused recombinant protein used in protein-replacement therapies. Thus, the delivery of these genetic medicines without adverse side effects is one of the most important problems facing gene and cell-based therapies. The essential goal of gene therapy is to provide what

Corresponding author: Nesmelova, I.V. (Irina.Nesmelova@uncc.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

all patients want, an improved quality of life. For these reasons gene therapy either by introduction of a therapeutic gene into cells of selected tissues or by engineering stem or pluripotent cells, which then are infused into the appropriate site, will become the treatment of choice for disorders such as hemophilia [1]. Gene therapy of tissues and/or cells is applicable to both genetic and acquired diseases such as cancer. This review discusses one promising vector system, DNA transposons, to achieve effective gene therapy without unwanted side effects and information that must be gained to improve this method of gene delivery to cells.

DNA transposons as integrating vectors

DNA transposons are mobile DNA elements that can move (transpose) from one location in a genome to another or from one DNA molecule to another (Fig. 1a). Autonomous DNA transposons encode a transposase enzyme needed for their activity; non-autonomous transposons have inactive transposase genes but can be mobilized by a transposase from an autonomous element. In mammalian cells non-autonomous transposons are in vast excess; indeed, in some genomes such as human none of the ca. 300,000 DNA transposons have a functional transposase gene. Transposons have three features that make them attractive as integrating vectors for functional genomics and genetic engineering/gene therapy. 1) Transposition is precise such that the genetic sequence that is mobilized for integration is absolutely defined. 2) Transposition directs the integration of single transposons. 3) Unlike viruses that often have highly immunogenic protein coats, transposons comprise only DNA and hence avoid immune and other defense mechanisms that cells employ to prevent integration of foreign DNA. Consequently, there is substantial interest in improving transpositional efficiency for use in targeting to specific sites in genomes [2–9].

Several DNA transposons are currently being developed for gene discovery and gene delivery to vertebrate organisms and their cells. These transposons include *Sleeping Beauty* (SB) [4,10], *Tol2* [2], *Frog Prince* [6], and *piggyBac* [3]. The SB transposon system in particular is based on a synthetic transposase that was resurrected from a *Tc1/mariner*-type transposon that was active in salmonid fish up until about 14 million years ago when it went extinct. Owing to its synthetic origins and further developments, the SB transposon is the most active of DNA transposons that are active in vertebrate animal cells [5] and because it was the first DNA transposon characterized in vertebrates, it has been characterized most extensively and consequently is the most widely used in research and proposed biomedical applications. In contrast to many of the other transposons used in vertebrates, SB appears to be the most random in preferences for integration sites, requiring only a TA dinucleotide basepair and largely independent of its transcriptional status [11]. Hence, owing the many uses to which vertebrate transposons are being put, high-resolution, structural information about transposase–transposon DNA complexes should facilitate the engineering of optimal structural components for high efficiency transposition.

The SB, Frog Prince and *piggyBac* transposase enzymes belong to the DD[E/D] family of transposases. DD[E/D]-transposases belong to a large superfamily of polynucleotidyl transferases, which includes RNase H, RuvC Holliday resolvase, RAG proteins, and retroviral integrases. In this review we compare the overall structures and discuss the conformational flexibility of transposases from the DD[E/D] family of transposons. Such information can be used to guide further development of transposons that are active in mammalian cells. The structures of catalytic domains of polynucleotidyl transferases and the mechanism of catalysis has been recently compared in detail [12] and they are mentioned only briefly here, as are relevant structural features of retroviral integrases that have been recently reviewed [13].

DD[E/D]-transposases

Transposase enzymes bind to the inverted terminal repeats (ITRs) of a transposon and catalyze the mobilization of the transposon to new location (Figure 1a). Transposase enzymes are divided into several families based on the mechanism utilized during transposition – DD[E/D]-transposases, rolling-circle or Y2-transposases, tyrosine (Y)-transposases, serine (S)-transposases, and a family, which encodes a combination of reverse transcriptase and endonuclease activities (RT/En) [14]. Members of the DD[E/D]-transposase family contain a characteristic motif of three acidic residues, two of which are aspartic acids and one is either glutamic or, in some cases, a third aspartic acid, a [D, D, (E or D)] motif. All DD[E/D]-transposases also contain a common structural motif that brings three catalytic residues into close proximity for catalysis, a RNase H-like fold [12–13, 15], a structural feature shared by the members of polynucleotidyl transferase superfamily. Based on functional analogy and supporting experimental evidence for at least some polynucleotidyl transferases [16–18], a two-metal catalytic mechanism patterned after that of the *Escherichia coli* DNA polymerase I [19] has been proposed. The three characteristic residues [D, D, (E or D)] are proposed to coordinate divalent metal ions required for the catalysis. The reaction begins with the hydrolysis of phosphodiester bonds at the termini of the ITRs of the transposon DNA, which leads to the formation of free 3'-OH groups. The 3'-OH groups are protected by the transposase until they are joined to the target DNA via a transesterification reaction, which is also catalyzed by the transposase [20–21].

The ways in which different DD[E/D]-transposases generate a DNA substrate for strand transfer vary [14]. Some transposases (e.g., MuA, Tn3) cleave only one strand at each transposon end. The generated branched intermediate with the 3'-transposon ends joined to the target and the 5'-ends attached to the donor flanks is resolved by DNA replication (replicative transposition). Other transposases (e.g. Tn5, Tn7, Tn10) cleave both DNA strands generating an excised transposon ready for insertion (cut-and-paste transposition). This process can include the formation of a hairpin linking two strands of the excised transposon (Tn5, Tn10) [22–23] or of the flanking sequences (*Hermes*) [24].

Folded structures of DD[D/E]-transposases

DD[E/D]-transposases have multiple functional domains (Figure 1b). Thus far, several structures of either full-length (e.g., Tn5 and Mos1), truncated DD[E/D]-transposases or their functional domains have been determined [17–18, 25–36] (Table 1).

The catalytic domain

The catalytic domains of different DD[E/D]-transposases have very little sequence similarity and may differ in size significantly (Figure 2a). Nevertheless, as mentioned above, all DD[E/D]-transposases (and polynucleotidyl transferases) contain a region with a similar, RNase H-like, α - β - α fold [12–13, 15] that was first observed in RNase H [37–38]; a beta-sheet with five strands (the first three strands are anti-parallel, and the last two strands are parallel) flanked by alpha-helices (Figures 2b–e). However, beyond the RNase H-like fold, catalytic domains are structurally diverse. Moreover, in some transposases the RNase H-like fold can be interrupted by the insertion sequence of variable length and structure (the grey boxes in Figure 2a and the grey lines in Figure 2b–2e). The RNase H-like fold of Tn5 transposase is interrupted by an insertion of 90 amino acids that includes two long beta-strands that participate in DNA-binding with the N-terminal DNA-binding domain [25] (Figure 2b). Likewise, the catalytic domain of *Hermes* transposase is disrupted by almost 300 residues alpha-helical insertion [30] (Figure 2c). The inserts in the two transposases are not structurally related. Unlike the insertion domain of Tn5 transposase, the insertion

domain of *Hermes* transposase also contributes to the formation of transposase oligomer [30].

The RNase H-like fold of some transposases remains uninterrupted, but an additional amino acid sequence can be appended, e.g., two additional, N-terminal and C-terminal alpha-helices (Figure 2e) that cap the RNase H-like core [18] of the Mos1 catalytic domain. These helices occupy a position similar to the insertion sequence in the Tn5 catalytic domain [25]. The catalytic domain of MuA transposase consists of the N-terminal subdomain (residues 258–490) with the uninterrupted RNase H-like fold and the C-terminal subdomain (residues 491–560) with six anti-parallel beta-strands forming a β -barrel [31] (Figure 2d). The C-terminal subdomain is thought to participate in nonspecific protein-DNA interactions due to overall positive electrostatic potential [31] and also is important for the synaptic assembly as revealed by point mutations [39].

The DNA-binding domain

While the catalytic domain of DD[E/D]-transposases possesses some DNA-binding activity that is usually non-specific and/or weak, the DNA-binding domain of transposases is responsible for the sequence-specific recognition of the ITRs of the cognate transposon. Usually, it is near the N-terminus of the transposase. As with the catalytic domains, the DNA-binding domains of different DD[E/D]-transposases differ significantly in sequence and length. They do not have a unified structural organization although they all contain a distinct combination of helices and turns (Figure 3). Transposases Tc3 and MuA possess a helix-turn-helix (HTH) DNA-binding motif, one of the most common motifs that proteins use to bind to DNA (Figure 3). The HTH motif is also predicted for the DNA-binding domain of Tc1 and *mariner* transposases [40]. In this regard, DNA-binding motifs in some DD[E/D]-transposases resemble those of many other types of DNA-binding proteins.

The DNA-binding domains of DD[E/D]-transposases can have a single DNA-recognition motif (Tn5, *Hermes*) or a modular structure (MuA, Tc3) (Figure 3). In the latter case, each of the subdomains recognizes different parts of the transposon DNA and can also participate in protein–protein interactions that occur during transposition. The DNA-binding domain of Tn5 transposase [25] is composed of four compacted helices connected by flexible linkages and turns (Figure 3a). Although the second and fourth helices form contacts with the transposon DNA, most contacts are with the fourth helix. The site-specific, DNA-binding domain of *Hermes* transposase in a monomeric form is composed of three helices (Figure 3b), however it exists as a highly intertwined dimer with a tightly packed hydrophobic core [30]. A similar domain has been also observed in another member of polynucleotidyl transferases family, RAG1 [41].

The DNA-binding domain of MuA transposase (Figure 3c) is further subdivided into I α -, I β -, and I γ -domains [42]. Structurally, I α belongs to the family of winged HTH DNA-binding proteins [34, 43]. In the model of MuA–DNA complex [34], the recognition helix is the second helix of the HTH motif. It fits into the major groove and the loop connecting the second and third beta-strands also contacts the DNA. The I β subdomain is composed of five helices connected by loops of various lengths [35]. The I γ subdomain comprises a four-helix bundle tightly packed around a hydrophobic core [36]. Despite the absence of significant amino acid sequence identity, the folding topologies of the second, third and fourth helices of the I β subdomain are very similar to those of the first three helices of the I γ subdomain, as well as to the fold found in members of the homeodomain family of HTH DNA-binding proteins [35].

The DNA-binding domains of Tc3 and Mos1 transposases have similar bipartite structures [29,33] (Figures 3d and 3e) that can be superimposed over backbone C α atoms within 3.1 Å.

Each Tc3 and Mos1 subdomain contains three helices and has a characteristic HTH DNA-binding motif [29,32–33]. The overall conformations of the two subdomains are almost the same - their structures are superimposable within 2 Å for the C α atoms. Both subdomains of the Tc3 and Mos1 transposase form contacts with DNA; however, the N-terminal subdomain binds with a higher specificity than the C-terminal subdomain [44]. A long linker that is stretched along the minor groove of the DNA fragment connects the two subdomains. The Tc3 and Mos1 bipartite DNA-binding domains are structurally similar to the *paired* DNA-binding domain of Pax transcription factors [45]. The highest degree of similarity is observed for the N-terminal subdomain. A bipartite DNA-binding domain, similar to that of Tc3 and Mos1 transposases, has been predicted for their close relative, SB [4,46] transposase (Figure 3f) as well as for other Tc1-related transposases [40].

Synaptic complex organization

Transposition happens in the context of a protein–DNA complex called the synaptic complex (or transpososome) (Figure 1a), which contains at least two transposase enzymes (Tn5, Tn10, Mos1) and the ends of the transposon DNA [25, 29, 47–48]. The number of transposase monomers that enter synaptic complex is not conserved between different DD[E/D]-transposases. MuA [49–51] or HIV-1 integrase [52] are shown to act as tetramers. Tetrameric assembly has been also proposed for mariner transposases Tc3, SB, and Himar 1 [33, 46, 53–54]. Apparently, synaptic complexes may include even higher order transposase multimers as suggested by a crystal structure of *Hermes* transposase hexamer, which, however, has been resolved for a truncated version of the protein in the absence of DNA molecules [30]. One of the reasons for such divergence in structures of synaptic complexes may arise from the spatial constraints due to the variety of structures assumed by the inserted motifs in the catalytic domains of some DD[E/D]-transposases and the structural variability of their DNA-binding domains. The dependence of synaptic complex formation on the properties of the structural organization of the DNA-binding domain has been reviewed for transposases in the Tc1/*mariner* superfamily [55]. In addition, other cellular proteins can be involved in the synaptic assembly at different stages of transposition (e.g., MuA, Tn7, Tn10 transposition).

For a long time, the available high-resolution structural information for the synaptic complexes of the full-length DD[E/D]-transposases, as well as for polynucleotidyl transferases, was limited to the structure of Tn5 transposase–DNA complex representing a synaptic complex at the stage following cleavage from donor DNA [25]. Recently, the crystal structure of the paired-end complex formed by the DNA sequences from inverted terminal repeats of transposon ends and the full-length Mos1 transposase have been resolved [29]. Additionally, a 3D reconstructed image at a 34 Å resolution from electron micrographs of the MuA complex consisting of a tetramer of MuA transposase and two identical 50-bp DNA segments. The path of the DNA through the complex, determined using electron spectroscopic imaging of the DNA-phosphorus, provides a structural model for the Mu-cleaved donor complex [51].

Crystal structures of the Tn5 and Mos1 (Figure 4a) transposase–DNA complexes [25, 29] provide direct experimental evidence of strikingly different overall architectures of synaptic complexes formed by DD[E/D]-transposases. The Tn5 forms a compact, approximately globular dimeric assembly, in which a transposase dimer binds to two *OE* DNA sequences of the transposon ends arranged in an anti-parallel fashion [25]. The Mos1 transposase–DNA complex contains a dimer of transposase and four DNA sequences mimicking the cleaved right transposon end [29]. Two of four DNA sequences (*IR_A* and *IR_B*) are bound by the DNA-binding domains of transposase monomers and, in contrast to the Tn5 synaptic complex, approach the catalytic sites in parallel (Figure 4a). Two additional DNA sequences

are bound only by the transposase catalytic domains and could represent likely binding sites for DNA flanking the transposon [29]. In this arrangement, two transposase monomers crisscross each other at the position of linkers connecting the DNA-binding and catalytic domains (Figure 4b). The crystal structure of a full-length Mos1 transposase in complex with DNA [29] resolved a controversy about the oligomeric state of Mos1 transposase [56–57] by showing that it forms a dimer. The remarkable difference of synaptic complex structures of Tn5 and Mos1 transposases, which serve essentially the same purpose, raises a question related to the engineering of highly active transposons for genetic applications: what is the minimal structure of DD[E/D]-transposases required to carry out the reaction of transposition?

Despite the different overall architectures, the Tn5 and Mos1 transposase–DNA complexes share two prominent common features. First, each monomer of the transposase forms contacts with both DNA fragments (termed *cis/trans* contacts): the DNA-binding domain binds to the DNA sequence at one end of the transposon whereas the catalytic domain of the same monomer binds to (and cleaves) the other transposon end [25,29,58] (Figure 4). Second, protein–DNA contacts are extensive and play a major role in maintaining the structure of the synaptic complex [25,29,58].

The cleavage of the transposon at one end by the transposase subunit that is bound to the other end, *i.e.*, in *trans*, was initially established for MuA transposase based on biochemical data [59–60] and later confirmed by a 3D model of MuA transposase–DNA complex [51]. A *trans*-cleavage was also found for another member of polynucleotidyl transferase family, RAG-1 protein [61]. A member of the *mariner* family of transposases, the synthetic *Sleeping Beauty* transposase, has been proposed to operate via *trans*-cleavage [54] (Figure 4b), although other geometries are not excluded [46]. The fact that the *trans*-cleavage is observed for the members of polynucleotidyl transferase family that are not closely related, suggests that it can be a common theme in DNA transposition.

Extensive protein–DNA contacts provide most of the interactions that stabilize the complex of MuA transposase [51]. These interactions might be responsible for the increasing stability of the synaptic complex as the transposition progresses [62]. Similarly, Tn10 transposase forms extensive contacts with the DNA at the onset of transposition although the contact map changes as reaction proceeds [63].

Both features, the *trans*-cleavage and extensive protein–DNA contacts, are likely necessary to hold the DNA ends together while the multi-step reaction of transposition is carried out. Accordingly, *trans* cleavage may be expected for other DD[E/D]-transposases as well. More structural data will help to decipher the organization of synaptic complexes of different transposases to clarify their mechanism of concerted cleavage and integration.

Conformational flexibility of DD[E/D]-transposases

Transposases must be sufficiently flexible to allow conformational rearrangements of their domains to bind the transposon DNA and to supply a catalytic site during each step of transposition. Thus, the MuA transposase undergoes several conformational transitions: upon binding to the transposon it acquires a conformation needed for tetramer formation, then it rearranges to perform a joining reaction of the Mu transposon to the target DNA, and lastly it undergoes a conformational transition to allow the disassembly of synaptic complex and release from the DNA (see [64–66] and references therein). The Tn5 transposase probably undergoes an allosteric structural transition prior to the formation of a synaptic complex to remove inhibitory interactions between the N-terminal and C-terminal domains [25,67].

In part, the conformational flexibility of the DD[E/D]-transposases is due to the presence of flexible linkers connecting the polypeptide domains. However, some regions within domains may also exhibit significant flexibility as indicated poorly defined electron density in crystal structures. In particular, crystallographic [68–72] and molecular dynamics simulation studies [73] of HIV-1 integrase have demonstrated that the surface loop overhanging the active site (residues 141–148) has a high degree of flexibility, which apparently could be correlated to integrase activity [70,72]. In contrast, the equivalent loop in the crystal structure of MuA catalytic domain is ordered [31]. However, the third catalytic residue E392, positioned on this loop, is too far from the other two catalytic residues to be able to participate in coordinating a metal atom [31]. Although the loop is ordered, there are several glycine residues in the vicinity that may allow a conformational rearrangement of the loop to bring all three catalytic residues close together in active conformation [31].

Binding to DNA alters the conformational flexibility of a transposase either by inducing the structure in some regions of the molecule or by shifting the equilibrium towards one, preferable conformational state. The recent crystal structure of the Mos1 paired-end complex [29] revealed that the segment formed by residues 162–189 (clamp loop-linker) becomes ordered in the presence of DNA and appears to play a crucial role in transposase dimerization, making protein–protein interactions across the dimer interface as well as important protein–DNA contacts. Likewise, NMR relaxation studies of the I α subdomain of the MuA DNA-binding domain in the DNA-unbound form showed that residues 38 to 46 in the disordered loop (wing) between the second and third beta-strands are highly mobile on a pico- to nanosecond time scale [43]. The same amino acid residues in the wing of the homologous DNA-binding domain of Mu repressor protein remain mobile in the absence of DNA but form a well-defined conformation that is immobilized within a minor groove of the DNA upon binding [74]. The wing mobility enables a rapid sampling of many possible conformations and facilitates transient recognition of the cognate DNA during transposition [43].

On the other hand, single molecule fluorescent studies of one of the subdomains from the DNA-binding domain of Tc3 transposase revealed that it fluctuates between two or more conformations in the absence of DNA. These conformational changes continue upon binding to DNA. However, the equilibrium between conformational states shifts and a particular conformation of the transposase persists for longer time when bound to DNA compared to the DNA-unbound form [75]. Thus the DNA-binding domain of Tc3 transposase retains its conformational flexibility even after binding to DNA, but the binding to DNA enhances the probability of formation of a preferred DNA-bound conformation.

Although here we have focused on several specific transposase enzymes, extensive studies of Tn10 transposon suggest that successful transposition might require significant conformational changes not only in the transposase, but in all components of the synaptic assembly, including the DNA [63,76–77].

Conclusions

All DD[E/D]-transposases catalyze mobilization of DNA segments using mechanisms that share two reaction steps: (1) hydrolysis of the phosphodiester bonds at each end of a transposon DNA to generate free 3'-OH groups and (2) joining of these groups to a target DNA sequence through a single-step transesterification. They all possess an RNase H-like fold containing the motif of three catalytically active residues, DD[E/D] motif. However, beyond this the sequences and structures of DD[E/D]-transposases show amazing variability, which is reflected in different ways by which DD[E/D]-transposases recognize transposon DNA, assemble into a synaptic complex, and complete the cleavages and

ligations that result in transposition. The comparison of the crystal structure of the Tn5 complex with the recent structure of Mos1 transposase–DNA complex demonstrates how dissimilar the overall architectures of synaptic assembly can be. Moreover, during transposition DD[E/D]-transposases undergo a series of conformational changes that adjust their binding competence and alter their organization.

Thus far the development of DNA transposons for genetic applications relied on the phylogenetic and biochemical approaches to direct important improvements of transposon activity. More information about the molecular details of the transposase–transposon DNA interactions, such as packing, hydrogen-bonding and electrostatic interactions responsible for nucleotide-specific recognition, would provide a foundation for the rational design of improved transposons and transposases. Such information can be derived from the three-dimensional structures of transposase–DNA complexes. Given the limited number, yet wide variation, of available three-dimensional structures of DD[E/D]-transposases and the structural variability, more structural studies of transposases complexed with transposon DNA segments that reflect different steps in the transposition reaction are necessary. This is especially true for transposons such as the *Sleeping Beauty* transposon system that are under study as vectors for human gene therapy and tools for functional genomics.

Acknowledgments

This work was supported by the faculty seed grant from the Academic Health Center of the University of Minnesota to IVN and PBH.

Abbreviations used

ITR	inverted terminal repeat
SB	Sleeping Beauty transposase
OE	outside end
bp	base pairs
HTH	helix-turn-helix
NLS	nuclear localization signal
ISD	integration/insertion site duplication
PDB	protein databank

References

1. Mannucci PM, Tuddenham EG. The hemophilias--from royal genes to gene therapy. *N Engl J Med*. 2001; 344:1773–1779. [PubMed: 11396445]
2. Balciunas D, Wangenstein KJ, Wilber A, Bell J, Geurts A, Sivasubbu S, Wang X, Hackett PB, Largaespada DA, McIvor RS, Ekker SC. Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. *PLoS Genet*. 2006; 2:e169. [PubMed: 17096595]
3. Hackett PB, Ekker SC, Largaespada DA, McIvor RS. Sleeping beauty transposon-mediated gene therapy for prolonged expression. *Adv Genet*. 2005; 54:189–232. [PubMed: 16096013]
4. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell*. 1997; 91:501–510. [PubMed: 9390559]
5. Mates L, Chuah MK, Belay E, Jerchow B, Manoj N, Acosta-Sanchez A, Grzela DP, Schmitt A, Becker K, Matrai J, Ma L, Samara-Kuko E, Gysemans C, Pryputniewicz D, Miskey C, Fletcher B, Vandendriessche T, Ivics Z, Izsvak Z. Molecular evolution of a novel hyperactive Sleeping Beauty

- transposase enables robust stable gene transfer in vertebrates. *Nat Genet.* 2009; 41:753–761. [PubMed: 19412179]
6. Miskey C, Izsvak Z, Plasterk RH, Ivics Z. The Frog Prince: a reconstructed transposon from *Rana pipiens* with high transpositional activity in vertebrate cells. *Nucleic Acids Res.* 2003; 31:6873–6881. [PubMed: 14627820]
 7. VandenDriessche T, Ivics Z, Izsvak Z, Chuah MK. Emerging potential of transposons for gene therapy and generation of induced pluripotent stem cells. *Blood.* 2009; 114:1461–1468. [PubMed: 19471016]
 8. Yant SR, Park J, Huang Y, Mikkelsen JG, Kay MA. Mutational analysis of the N-terminal DNA-binding domain of sleeping beauty transposase: critical residues for DNA binding and hyperactivity in mammalian cells. *Mol Cell Biol.* 2004; 24:9239–9247. [PubMed: 15456893]
 9. Zayed H, Izsvak Z, Walisko O, Ivics Z. Development of hyperactive sleeping beauty transposon vectors by mutational analysis. *Mol Ther.* 2004; 9:292–304. [PubMed: 14759813]
 10. Hackett PB, Largaespada DA, Cooper LJ. A transposon and transposase system for human application. *Mol Ther.* 2010; 18:674–683. [PubMed: 20104209]
 11. Berry C, Hannenhalli S, Leipzig J, Bushman FD. Selection of target sites for mobile DNA integration in the human genome. *PLoS Comput Biol.* 2006; 2:e157. [PubMed: 17166054]
 12. Nowotny M. Retroviral integrase superfamily: the structural perspective. *EMBO Rep.* 2009; 10:144–151. [PubMed: 19165139]
 13. Jaskolski M, Alexandratos JN, Bujacz G, Wlodawer A. Piecing together the structure of retroviral integrase, an important target in AIDS therapy. *Febs J.* 2009; 276:2926–2946. [PubMed: 19490099]
 14. Curcio MJ, Derbyshire KM. The outs and ins of transposition: from mu to kangaroo. *Nat Rev Mol Cell Biol.* 2003; 4:865–877. [PubMed: 14682279]
 15. Rice PA, Baker TA. Comparative architecture of transposase and integrase complexes. *Nat Struct Biol.* 2001; 8:302–307.
 16. Nowotny M, Gaidamakov SA, Crouch RJ, Yang W. Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell.* 2005; 121:1005–1016. [PubMed: 15989951]
 17. Lovell S, Goryshin IY, Reznikoff WR, Rayment I. Two-metal active site binding of a Tn5 transposase synaptic complex. *Nat Struct Biol.* 2002; 9:278–281. [PubMed: 11896402]
 18. Richardson JM, Dawson A, O'Hagan N, Taylor P, Finnegan DJ, Walkinshaw MD. Mechanism of Mos1 transposition: insights from structural analysis. *Embo J.* 2006; 25:1324–1334. [PubMed: 16511570]
 19. Beese LS, Steitz TA. Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *Embo J.* 1991; 10:25–33. [PubMed: 1989886]
 20. Craig NL. Unity in transposition reactions. *Science.* 1995; 270:253–254. [PubMed: 7569973]
 21. Mizuuchi K. Transpositional recombination: mechanistic insights from studies of mu and other elements. *Annu Rev Biochem.* 1992; 61:1011–1051. [PubMed: 1323232]
 22. Bhasin A, Goryshin IY, Reznikoff WS. Hairpin formation in Tn5 transposition. *J Biol Chem.* 1999; 274:37021–37029. [PubMed: 10601258]
 23. Kennedy AK, Guhathakurta A, Kleckner N, Haniford DB. Tn10 transposition via a DNA hairpin intermediate. *Cell.* 1998; 95:125–134. [PubMed: 9778253]
 24. Zhou L, Mitra R, Atkinson PW, Hickman AB, Dyda F, Craig NL. Transposition of hAT elements links transposable elements and V(D)J recombination. *Nature.* 2004; 432:995–1001. [PubMed: 15616554]
 25. Davies DR, Goryshin IY, Reznikoff WS, Rayment I. Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science.* 2000; 289:77–85. [PubMed: 10884228]
 26. Steiniger-White M, Bhasin A, Lovell S, Rayment I, Reznikoff WS. Evidence for “unseen” transposase–DNA contacts. *J Mol Biol.* 2002; 322:971–982. [PubMed: 12367522]
 27. Davies DR, Mahnke Braam L, Reznikoff WS, Rayment I. The three-dimensional structure of a Tn5 transposase-related protein determined to 2.9-Å resolution. *J Biol Chem.* 1999; 274:11904–11913. [PubMed: 10207011]

28. Klenchin VA, Czyz A, Goryshin IY, Gradman R, Lovell S, Rayment I, Reznikoff WS. Phosphate coordination and movement of DNA in the Tn5 synaptic complex: role of the (R)YREK motif. *Nucleic Acids Res.* 2008; 36:5855–5862. [PubMed: 18790806]
29. Richardson JM, Colloms SD, Finnegan DJ, Walkinshaw MD. Molecular architecture of the Mos1 paired-end complex: the structural basis of DNA transposition in a eukaryote. *Cell.* 2009; 138:1096–1108. [PubMed: 19766564]
30. Hickman AB, Perez ZN, Zhou L, Musingarimi P, Ghirlando R, Hinshaw JE, Craig NL, Dyda F. Molecular architecture of a eukaryotic DNA transposase. *Nat Struct Mol Biol.* 2005; 12:715–721. [PubMed: 16041385]
31. Rice P, Mizuuchi K. Structure of the bacteriophage Mu transposase core: a common structural motif for DNA transposition and retroviral integration. *Cell.* 1995; 82:209–220. [PubMed: 7628012]
32. van Pouderoyen G, Ketting RF, Perrakis A, Plasterk RH, Sixma TK. Crystal structure of the specific DNA-binding domain of Tc3 transposase of *C.elegans* in complex with transposon DNA. *Embo J.* 1997; 16:6044–6054. [PubMed: 9312061]
33. Watkins S, van Pouderoyen G, Sixma TK. Structural analysis of the bipartite DNA-binding domain of Tc3 transposase bound to transposon DNA. *Nucleic Acids Res.* 2004; 32:4306–4312. [PubMed: 15304566]
34. Clubb RT, Omichinski JG, Savilahti H, Mizuuchi K, Gronenborn AM, Clore GM. A novel class of winged helix-turn-helix protein: the DNA-binding domain of Mu transposase. *Structure.* 1994; 2:1041–1048. [PubMed: 7881904]
35. Schumacher S, Clubb RT, Cai M, Mizuuchi K, Clore GM, Gronenborn AM. Solution structure of the Mu end DNA-binding domain of phage Mu transposase: modular DNA recognition by two tethered domains. *Embo J.* 1997; 16:7532–7541. [PubMed: 9405381]
36. Clubb RT, Schumacher S, Mizuuchi K, Gronenborn AM, Clore GM. Solution structure of the I gamma subdomain of the Mu end DNA-binding domain of phage Mu transposase. *J Mol Biol.* 1997; 273:19–25. [PubMed: 9367742]
37. Yang W, Hendrickson WA, Crouch RJ, Satow Y. Structure of ribonuclease H phased at 2 Å resolution by MAD analysis of the selenomethionyl protein. *Science.* 1990; 249:1398–1405. [PubMed: 2169648]
38. Katayanagi K, Miyagawa M, Matsushima M, Ishikawa M, Kanaya S, Ikehara M, Matsuzaki T, Morikawa K. Three-dimensional structure of ribonuclease H from *E. coli*. *Nature.* 1990; 347:306–309. [PubMed: 1698262]
39. Namgoong SY, Kim K, Saxena P, Yang JY, Jayaram M, Giedroc DP, Harshey RM. Mutational analysis of domain II beta of bacteriophage Mu transposase: domains II alpha and II beta belong to different catalytic complementation groups. *J Mol Biol.* 1998; 275:221–232. [PubMed: 9466905]
40. Pietrokovski S, Henikoff S. A helix-turn-helix DNA-binding motif predicted for transposases of DNA transposons. *Mol Gen Genet.* 1997; 254:689–695. [PubMed: 9202385]
41. Yin FF, Bailey S, Innis CA, Ciubotaru M, Kamtekar S, Steitz TA, Schatz DG. Structure of the RAG1 nonamer binding domain with DNA reveals a dimer that mediates DNA synapsis. *Nat Struct Mol Biol.* 2009; 16:499–508. [PubMed: 19396172]
42. Nakayama C, Teplow DB, Harshey RM. Structural domains in phage Mu transposase: identification of the site-specific DNA-binding domain. *Proc Natl Acad Sci U S A.* 1987; 84:1809–1813. [PubMed: 3031651]
43. Clubb RT, Mizuuchi M, Huth JR, Omichinski JG, Savilahti H, Mizuuchi K, Clore GM, Gronenborn AM. The wing of the enhancer-binding domain of Mu phage transposase is flexible and is essential for efficient transposition. *Proc Natl Acad Sci U S A.* 1996; 93:1146–1150. [PubMed: 8577730]
44. Colloms SD, van Luenen HG, Plasterk RH. DNA binding activities of the *Caenorhabditis elegans* Tc3 transposase. *Nucleic Acids Res.* 1994; 22:5548–5554. [PubMed: 7838706]
45. Czerny T, Schaffner G, Busslinger M. DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev.* 1993; 7:2048–2061. [PubMed: 8406007]

46. Izsvak Z, Khare D, Behlke J, Heinemann U, Plasterk RH, Ivics Z. Involvement of a bifunctional, paired-like DNA-binding domain and a transpositional enhancer in Sleeping Beauty transposition. *J Biol Chem.* 2002; 277:34581–34588. [PubMed: 12082109]
47. Bolland S, Kleckner N. The three chemical steps of Tn10/IS10 transposition involve repeated utilization of a single active site. *Cell.* 1996; 84:223–233. [PubMed: 8565068]
48. Sakai J, Chalmers RM, Kleckner N. Identification and characterization of a pre-cleavage synaptic complex that is an early intermediate in Tn10 transposition. *Embo J.* 1995; 14:4374–4383. [PubMed: 7556079]
49. Lavoie BD, Chan BS, Allison RG, Chaconas G. Structural aspects of a higher order nucleoprotein complex: induction of an altered DNA structure at the Mu-host junction of the Mu type 1 transpososome. *Embo J.* 1991; 10:3051–3059. [PubMed: 1655409]
50. Mizuuchi M, Baker TA, Mizuuchi K. Assembly of the active form of the transposase-Mu DNA complex: a critical control point in Mu transposition. *Cell.* 1992; 70:303–311. [PubMed: 1322248]
51. Yuan JF, Beniac DR, Chaconas G, Ottensmeyer FP. 3D reconstruction of the Mu transposase and the Type 1 transpososome: a structural framework for Mu DNA transposition. *Genes Dev.* 2005; 19:840–852. [PubMed: 15774720]
52. Guiot E, Carayon K, Delelis O, Simon F, Tauc P, Zubin E, Gottikh M, Mouscadet JF, Brochon JC, Deprez E. Relationship between the oligomeric status of HIV-1 integrase on DNA and enzymatic activity. *J Biol Chem.* 2006; 281:22707–22719. [PubMed: 16774912]
53. Lipkow K, Buisine N, Lampe DJ, Chalmers R. Early intermediates of mariner transposition: catalysis without synapsis of the transposon ends suggests a novel architecture of the synaptic complex. *Mol Cell Biol.* 2004; 24:8301–8311. [PubMed: 15340089]
54. Cui Z, Geurts AM, Liu G, Kaufman CD, Hackett PB. Structure-function analysis of the inverted terminal repeats of the sleeping beauty transposon. *J Mol Biol.* 2002; 318:1221–1235. [PubMed: 12083513]
55. Brillet B, Bigot Y, Auge-Gouillou C. Assembly of the Tc1 and mariner transposition initiation complexes depends on the origins of their transposase DNA binding domains. *Genetica.* 2007; 130:105–120. [PubMed: 16912840]
56. Auge-Gouillou C, Brillet B, Hamelin MH, Bigot Y. Assembly of the mariner Mos1 synaptic complex. *Mol Cell Biol.* 2005; 25:2861–2870. [PubMed: 15767689]
57. Richardson JM, Finnegan DJ, Walkinshaw MD. Crystallization of a Mos1 transposase-inverted-repeat DNA complex: biochemical and preliminary crystallographic analyses. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2007; 63:434–437.
58. Steiniger-White M, Rayment I, Reznikoff WS. Structure/function insights into Tn5 transposition. *Curr Opin Struct Biol.* 2004; 14:50–57. [PubMed: 15102449]
59. Aldaz H, Schuster E, Baker TA. The interwoven architecture of the Mu transposase couples DNA synapsis to catalysis. *Cell.* 1996; 85:257–269. [PubMed: 8612278]
60. Savilahti H, Mizuuchi K. Mu transpositional recombination: donor DNA cleavage and strand transfer in trans by the Mu transposase. *Cell.* 1996; 85:271–280. [PubMed: 8612279]
61. Swanson PC. The DDE motif in RAG-1 is contributed in trans to a single active site that catalyzes the nicking and transesterification steps of V(D)J recombination. *Mol Cell Biol.* 2001; 21:449–458. [PubMed: 11134333]
62. Surette MG, Buch SJ, Chaconas G. Transpososomes: stable protein-DNA complexes involved in the in vitro transposition of bacteriophage Mu DNA. *Cell.* 1987; 49:253–262. [PubMed: 3032448]
63. Crellin P, Chalmers R. Protein-DNA contacts and conformational changes in the Tn10 transpososome during assembly and activation for cleavage. *Embo J.* 2001; 20:3882–3891. [PubMed: 11447129]
64. Chaconas G. Studies on a “jumping gene machine”: higher-order nucleoprotein complexes in Mu DNA transposition. *Biochem Cell Biol.* 1999; 77:487–491. [PubMed: 10668626]
65. Harshey RM, Jayaram M. The mu transpososome through a topological lens. *Crit Rev Biochem Mol Biol.* 2006; 41:387–405. [PubMed: 17092824]
66. Williams TL, Jackson EL, Carritte A, Baker TA. Organization and dynamics of the Mu transpososome: recombination by communication between two active sites. *Genes Dev.* 1999; 13:2725–2737. [PubMed: 10541558]

67. Mahnke Braam LA, Goryshin IY, Reznikoff WS. A mechanism for Tn5 inhibition. carboxyl-terminal dimerization. *J Biol Chem.* 1999; 274:86–92. [PubMed: 9867814]
68. Dyda F, Hickman AB, Jenkins TM, Engelman A, Craigie R, Davies DR. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science.* 1994; 266:1981–1986. [PubMed: 7801124]
69. Bujacz G, Alexandratos J, Qing ZL, Clement-Mella C, Wlodawer A. The catalytic domain of human immunodeficiency virus integrase: ordered active site in the F185H mutant. *FEBS Lett.* 1996; 398:175–178. [PubMed: 8977101]
70. Greenwald J, Le V, Butler SL, Bushman FD, Choe S. The mobility of an HIV-1 integrase active site loop is correlated with catalytic activity. *Biochemistry.* 1999; 38:8892–8898. [PubMed: 10413462]
71. Goldgur Y, Dyda F, Hickman AB, Jenkins TM, Craigie R, Davies DR. Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. *Proc Natl Acad Sci U S A.* 1998; 95:9150–9154. [PubMed: 9689049]
72. Lubkowski J, Yang F, Alexandratos J, Wlodawer A, Zhao H, Burke TR Jr, Neamati N, Pommier Y, Merkel G, Skalka AM. Structure of the catalytic domain of avian sarcoma virus integrase with a bound HIV-1 integrase-targeted inhibitor. *Proc Natl Acad Sci U S A.* 1998; 95:4831–4836. [PubMed: 9560188]
73. Lee MC, Deng J, Briggs JM, Duan Y. Large-scale conformational dynamics of the HIV-1 integrase core domain and its catalytic loop mutants. *Biophys J.* 2005; 88:3133–3146. [PubMed: 15731379]
74. Wojciak JM, Iwahara J, Clubb RT. The Mu repressor-DNA complex contains an immobilized ‘wing’ within the minor groove. *Nat Struct Biol.* 2001; 8:84–90. [PubMed: 11135677]
75. Daniel DC, Thompson M, Woodbury NW. DNA-binding interactions and conformational fluctuations of Tc3 transposase DNA binding domain examined with single molecule fluorescence spectroscopy. *Biophys J.* 2002; 82:1654–1666. [PubMed: 11867477]
76. Sakai JS, Kleckner N, Yang X, Guhathakurta A. Tn10 transpososome assembly involves a folded intermediate that must be unfolded for target capture and strand transfer. *Embo J.* 2000; 19:776–785. [PubMed: 10675347]
77. Haniford DB. Transpososome dynamics and regulation in Tn10 transposition. *Crit Rev Biochem Mol Biol.* 2006; 41:407–424. [PubMed: 17092825]
78. Lohe AR, De Aguiar D, Hartl DL. Mutations in the mariner transposase: the D,D(35)E consensus sequence is nonfunctional. *Proc Natl Acad Sci U S A.* 1997; 94:1293–1297. [PubMed: 9037046]

Figure 1. DNA transposons: the mechanism of transposition and molecular characteristics of the transposases featured in this review

(a) The *cut-and-paste* mechanism of transposition for a DNA transposon (purple, with the Inverted Terminal Repeats (ITRs, filled arrows) flanking a gene (X)) mediated by transposase (blue spheres). The reaction occurs in four steps: (i) binding of the transposase enzyme to the inverted terminal repeats, ITRs, of the transposon DNA; (ii) formation of synaptic complex, in which two ITRs of the transposon DNA are brought together and held by the transposase enzymes; (iii) excision of the transposon DNA from its original location, DNA-1; and (iv) integration of excised transposon DNA into the target DNA, DNA-2. The transposition process generates a duplication of the integration site (**ISD**), which can vary from 2 to 16 bp, depending on the transposon/transposase. Cellular DNA repair enzymes rejoin the cleaved source of the transposon (DNA-1) often leaving a footprint of extra basepairs (DNA-1 with black slash lines). The figure shows a single transposase binding to each ITR although, as described in the text, often multiple transposases bind at each end. (b) Schematic structure of some DD[E/D]-transposases. The sizes of the transposases are given in amino acids. Multidomain structure of the transposases is shown with structurally independent domains indicated by different shapes. All transposases contain the catalytic domain shown by pink rectangles and the transposon-binding domain shown by grey ovals. In addition to these domains some transposases have C-terminal (Tn5, MuA) or N-terminal (*Hermes*) domains. The C-terminal domain of Tn5 transposase, also referred to as the dimerization domain, is responsible for protein–protein interactions during the formation of the synaptic complex [25]. The N-terminal domain of *Hermes* transposase contains a nuclear localization signal and is proposed to bind zinc and to participate in nonspecific DNA binding. The C-terminal domain of MuA transposase, or domain III, is further subdivided into III α and III β domains. Domain III α is required for the active protein–DNA complex assembly, binds DNA, and possesses endonuclease activity by itself. Domain III β interacts with MuB protein, which allosterically activates MuA transposase for the transesterification step (see [64] and references therein). The HIV-1 integrase consists of three domains: the N-terminal zinc-binding domain that has been implicated in protein-protein interactions and multimerization, the catalytic domain, and the C-terminal nonspecific DNA-binding domain. The catalytic domain of integrase shares common structural features with DD[E/D]-transposases. The N-terminal domain contains a highly-conserved between integrases motif that is composed of two histidine and two cysteine amino acids (the HHCC motif) and binds zinc. In the presence of zinc, the N-terminal domain of integrase forms a four-helical structure. The C-terminal domain structurally resembles Src homology domains (for review see [13]).

Figure 2. Schematic presentation and three-dimensional structures of the catalytic domains of DD[E/D]-transposases

(a) Schematic structures of the catalytic domains of Tn5, *Hermes*, MuA, and Mos1 transposases. Blue boxes show the catalytic residues. Numbers correspond to the amino acids in the protein sequence. (b) Structure of the catalytic domain of Tn5 transposase (PDB code 1muh). (c) Structure of the catalytic domain of *Hermes* transposase (PDB code 2bw3). (d) Structure of the catalytic domain of MuA transposase (PDB code 1bco). (e) Structure of the catalytic domain of Mos1 transposase (PDB code 2f7t). Secondary structure elements of the RNase H-like fold are shown in ribbon representation. Grey lines show structural elements of the catalytic domains that are not part of the RNase H-like fold and correspond to grey boxes in panel A. Catalytic residues are shown in a ball-and-stick representation and metal ions are indicated by purple spheres (only seen in panels *b* and *e*). The presence of D284 in place of glutamic acid in Mos1 and other *mariner* transposases is essential; mutant D284E is inactive, most likely because the longer side chain of E284 occupies the second metal-binding site [78].

Figure 3. Three-dimensional structures of the DNA-binding domains of DD[E/D]-transposases DNA-binding domains are shown schematically as grey ovals. Numbers correspond to the amino acids in the protein sequence. Secondary structural elements are shown in ribbon representation. The structure of the DNA-binding domain from Tc3 and Mos1 transposases are shown in a complex with the cognate ITR DNA sequences (green lines). The structures are taken from Protein Data Bank. (a) Structure of the DNA-binding domain of Tn5 transposase (PDB code 1muh) [25]. (b) Structure of the DNA-binding domain of *Hermes* transposase [30] (PDB code 2bw3). (c) Modular structure of the DNA-binding domain of MuA transposase [34–36] (PDB codes 1tns, 2ezh, 2ezk). (d) Structure of the bipartite DNA-binding domain of Tc3 transposase [33] (PDB code 1u78). (e) Structure of the bipartite DNA-binding domain of Mos1 transposase [29] (PDB code 3hos). (f) Schematic structure of the bipartite DNA-binding domain of SB transposase. Based on sequence similarities, the secondary structure of the SB DNA-binding domain is predicted to be similar to the structure of the DNA-binding domains of Tc3 and Mos1 transposases [4,46].

Figure 4. Structures of the synaptic complexes of DD[E/D]-transposases

(a) Crystal structures of Tn5 and Mos1 synaptic complexes [25,29]. Arrows indicate the anti-parallel arrangement of the transposon ends in Tn5 complex and the parallel arrangement of the transposon ends (IR_A and IR_B) in Mos1 complex. Two additional DNA sequences (FL_A and FL_B) that are not bound by the transposase DNA-binding domains may represent binding sites for DNA flanking the transposon. (b) Transposase molecules are shown as ovals that bind to the ITRs (broad rectangles) of their respective transposons (narrow grey lines). Tn5, Mos1 and MuA transposases binding to one end of a transposon *trans*-operate on the other end, as depicted by the solid arrows. Mu transposition proceeds through a series of higher-order nucleoprotein complexes [64,66]. We show a stable synaptic complex formed by a tetramer of MuA in which two DNA ends are paired [50]. Only two monomers in the MuA tetramer are active for cleavage. Each of the ITRs for an SB transposon has two nearly identical binding sites (DRs) for SB transposase, with tighter binding to an internal DR_i compared to the binding at an outer DR_o . The data support a model wherein the SB transposases also operate in *trans* [54], as depicted by either the dashed set of arrows, the dotted set of arrows, or both. The dashed set of arrows indicates the criss-cross inter-monomer contacts.

Table 1Protein Data Bank (**PDB**) access codes for the structures of DD[E/D]-transposases.

Transposase	Organism	PDB code	References
Tn5	<i>E. coli</i>	1muh (E54K, M56A, L372P, 20bp OE DNA)	[25]
		1mur (E54K, M56A, L372P, 20bp OE DNA, 2 Mn ²⁺)	[17]
		1mus (E54K, M56A, D119K, K120A, E345K, L372P, 20bp OE DNA, 2 Mn ²⁺) URL: http://www.rcsb.org/pdb	
		1mm8 (ME DNA)	[26]
		1b7e (transposase inhibitor)	[27]
		3ecp (E54K, M56A, L372P, G477, 20bp OE 5'-phosphorylated DNA)	[28]
Mos1	<i>Drosophila mauritiana</i>	3hos (T216A, 28/25bp DNA, Mg ²⁺)	[29]
		3hot (T216A, 28/25bp DNA, Mn ²⁺)	
<i>Hermes</i>	<i>Musca domestica</i>	2bw3 (residues 79–612: DNA-binding, catalytic, and “inserted” domains)	[30]
<i>Catalytic domain</i>			
MuA	Bacteriophage Mu	1bco, 1bcm	[31]
Mos1	<i>Drosophila mauritiana</i>	2f7t (T216A, Mg ²⁺)	[18]
<i>DNA-binding domain</i>			
Tc3	<i>Caenorhabditis elegans</i>	1tc3 (residues 1–65, 20/21bp DNA)	[32]
		1u78 (bipartite domain, 26bp DNA)	[33]
I α domain of MuA	Bacteriophage Mu	1tns, 1tnt	[34]
I β subdomain of MuA	Bacteriophage Mu	2ezk, 2ezl	[35]
I γ subdomain of MuA	Bacteriophage Mu	2ezh, 2ezi	[36]