*lecular Biology of the Yeast* Saccharomyces*: Gene Expression*, E. W. Jones, J. R. Pringle, J. R. Broach, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1992), vol. 2, p. 583.
3. A. Johnson, in *Transcriptional Regulation*, S. L. McKnight and K. R. Yamamoto, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1992), p. 975.
4. B. M. Herschbach and A. D. Johnson, *Annu. Rev. Cell Biol.* **9**, 479 (1993).
5. C. A. Keleher, C. Goutte, A. D. Johnson, *Cell* **53**, 927 (1988); C. A. Keleher, S. Passmore, A. D. Johnson, *Mol. Cell Biol.* **9**, 5228 (1989); S. Passmore, R. Elble, B.-R. Tye, *Genes Dev.* **3**, 921 (1989).
6. A. D. Johnson and I. Herskowitz, *Cell* **42**, 237 (1985).
7. C. Goutte and A. D. Johnson, *ibid.* **52**, 875 (1988); A. Bender and G. F. Sprague, *ibid.* **50**, 681 (1987); A. M. Dranginis, *Nature* **347**, 682 (1990).
8. W. J. Gehring, M. Affolter, T. Burglin, *Annu. Rev. Biochem.* **63**, 487 (1994).
9. C. Wolberger, A. K. Vershon, B. Liu, A. D. Johnson, C. O. Pabo, *Cell* **67**, 517 (1991).
10. C. L. Phillips, A. K. Vershon, A. D. Johnson, F. W. Dahlquist, *Genes Dev.* **5**, 764 (1991).
11. R. T. Sauer, D. L. Smith, A. D. Johnson, *ibid.* **2**, 807 (1988); A. K. Vershon and A. D. Johnson, *Cell* **72**, 105 (1993).
12. A. Mak and A. D. Johnson, *Genes Dev.* **7**, 1862 (1993); M. R. Stark and A. D. Johnson, *Nature* **371**, 429 (1994); J. Strathern, B. Shafer, J. Hicks, C. McGill, *Genetics* **120**, 75 (1988); K. Tatchell, K. A. Nasmyth, B. D. Hall, *Cell* **27**, 25 (1981).
13. C. L. Phillips, M. R. Stark, A. D. Johnson, F. W. Dahlquist, *Biochemistry* **33**, 9294 (1994).
14. D. L. Smith, A. B. Desai, A. D. Johnson, *Nucleic Acids Res.* **23**, 1239 (1995).
15. A. K. Vershon, Y. Jin, A. D. Johnson, *Genes Dev.* **9**, 182 (1995).
16. C. Goutte and A. D. Johnson, *EMBO J.* **13**, 1434 (1994).
17. Y. Jin, J. Mead, C. Wolberger, A. K. Vershon, *Science* **270**, 290 (1995).
18. F. B. Johnson, E. Parker, M. Krasnow, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 739 (1995).

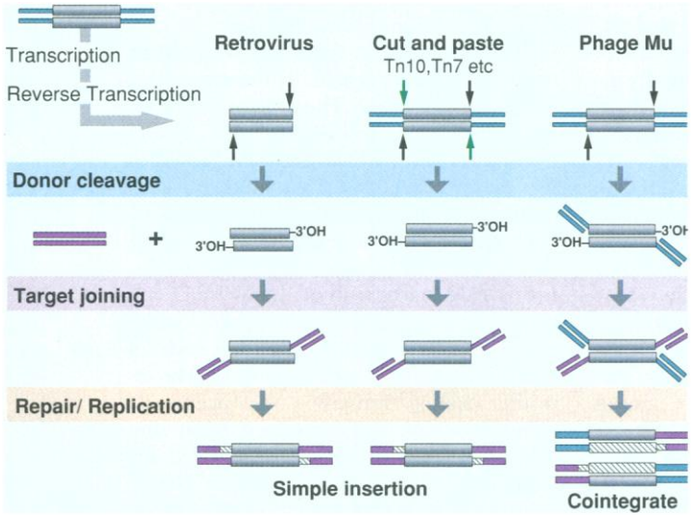
# Unity in Transposition Reactions

Nancy L. Craig

The process of transposition, in which pieces of DNA move around the genome, was made famous by Barbara McClintock's discovery of "jumping genes" and the surprising plasticity of DNA (*1*). Other biological transactions also make use of this process: the acquisition of bacterial genes for antibiotic resistance, the replication of certain bacteriophages, the integration of retroviruses, and the intracellular movement of retroviral-like elements. New evidence points to unexpected parallels among these many transposition events: They all occur by similar DNA breakage and joining reactions. Moreover, the structures of transposition proteins from very different biological sources have remarkable overall structural similarity, even though they lack extensive primary sequence homology.

Central to all transposition reactions are breakage events that precisely expose the 3′ tips of the transposable element; these exposed 3′ tips are then joined to the target DNA (*2*). Sometimes DNA cleavage reactions also occur at the 5′ tips of the element; whether this occurs has a profound influence on the structure of the transposition products. Thus, although all transposition reactions involve DNA breakage and joining, several different types of recombination products can emerge, depending on which DNA strands are broken and joined (see figure).

Retroviral integration (*1*–*3*) and the replication of phage Mu (*1*, *2*, *4*) exhibit breakage and joining at only the 3′ ends of mobile elements. In retroviral integration, the viral RNA genome is converted to double-stranded DNA by reverse transcription. This DNA product is then cleaved by the retrovirally encoded integrase to expose 3′-OH ends at the embedded tips of the actual retroviral DNA. Strand transfer reactions then join these exposed 3′ ends to staggered positions on the target DNA, one transposon end joining to one target strand and the other end joining to a displaced position on the other target strand. As a result, the transposon is covalently joined to the target DNA but is flanked by short gaps that reflect the staggered positions of target joining; host DNA repair functions then repair these flanking gaps. The resulting product, in which the retroviral DNA is covalently joined to the target DNA, is called a simple insertion (see figure).




**The DNA-processing reactions that underlie the translocation of three mobile elements.** Gray boxes, mobile elements; blue lines, flanking donor DNA; black arrow, cleavage at the 3′-OH ends; green arrow, cleavage at the 5′ ends; red line, target DNA; cross-hatching, DNA replication.

In the case of bacteriophage Mu replication, the phage DNA is embedded in host chromosomal DNA. DNA cleavage reactions executed by MuA protein, the Mu-encoded transposase, introduce single-strand nicks at both tips of the element. These cleavages cleanly expose the 3′-OH ends of the transposon, separating them from flanking bacterial DNA but leaving the transposon covalently linked at its uncleaved 5′ ends to flanking DNA; the exposed 3′ ends of the transposon are then joined by strand transfer reactions to staggered positions on the target DNA. This transposition product is then replicated by host DNA replication to generate a product called a cointegrate in which the donor backbone, target, and two transposon copies are linked (see figure). In some cases, such as Tn3, another element-encoded recombination system further processes the cointegrate to generate a target molecule containing a simple insertion and regenerate the donor (*5*).

Identical chemical steps executed by element-encoded transposition proteins underlie retroviral integration and Mu replication: DNA cleavage reactions expose the 3′-OH ends of the elements, followed by strand transfer reactions that covalently join these 3′ element ends to the target DNA. Although the products of retroviral integration and Mu transposition are distinct, their differences arise not from the transposition reactions per se but rather from the state of the DNA substrate, in particular at the 5′ ends. Mu remains linked to the donor site while also inserting into the target site, because no cleavage occurs at the 5′ ends of the element; by contrast, in the retrovirus the substrate DNA contains only the transposable DNA segment.

The author is in the Howard Hughes Medical Institute, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. Email: nancy_craig@qmail.bs.jhu.edu

The DNA cleavage and strand transfer reactions at the 3′ ends of retroviruses and phage Mu proceed by similar mechanisms. Both are $Mg^{2+}$-dependent reactions and likely result from direct, one-step transesterifications without covalent, protein-DNA intermediates (6). In DNA cleavage, $H_2O$ is the nucleophile that attacks a phosphodiester bond at the transposon tip, resulting in phosphodiester bond hydrolysis and exposure of the 3′-OH transposon end. In strand transfer, the newly exposed 3′-OH transposon tip is the attacking nucleophile; attack by the 3′-OH transposon end on the target DNA results in the covalent linkage of the transposon end to a target strand. The chemical similarity of these reactions and the fact that mutation of the proteins that execute these reactions [retroviral integrases (7) and the MuA transposase (8)] block both the 3′ breakage and joining steps led to the suggestion that the same (or at least closely related) active sites execute both of these reactions.

That a retroviral integrase and a bacterial transposase are indeed fundamentally related has now been spectacularly demonstrated by the structures of the catalytic domains of these proteins at the atomic level. The overall topology of the HIV integrase (9) and MuA transposase (10) structures is similar, although there is little extended primary sequence homology between them. That these proteins might be structurally related was not entirely unanticipated; these and other recombinases contain a signature array of conserved acidic amino acids, the D,D(35)E motif, so-called because of the usually 35–amino acid spacing between the last two residues (7, 11) These conserved amino acids are critical for the 3′ end processing reactions (7, 8), suggesting that these amino acids are part of (or at least closely related to) the active sites of the enzymes. Moreover, in both the integrase and MuA transposase structures, these amino acids are close together (9, 10), forming a plausible binding site for a metal ion cofactor essential to recombination.

What was unanticipated was that the structures of the catalytic domains of these recombinases would also be so profoundly related to the structures of two other nucleic acid–processing enzymes, ribonuclease H (12) and the Holliday junction–resolving enzyme RuvC (13). All four proteins display the same central overall protein topology with clustering of acidic residues (9, 10). Thus, all of these proteins—of disparate biological function—are members of a superfamily that executes metal ion–dependent polynucleotidyl transferase reactions.

How are the DNA-processing reactions that underlie recombination coordinated so as to result in recombination rather than breakage? In Mu, a tetramer of transposase (that is, four metal-dependent phosphoryl transfer centers) actually executes recombination (14); the precise arrangement of these active sites and how they may communicate with each other is not yet known. Two of the active sites may mediate the strand cleavage steps and the other two the strand transfer steps. The central regulatory step in Mu recombination is the transition of MuA from the monomer to the active tetramer form (15), a step that depends on critical features of the substrate DNA and other protein cofactors. The retroviral integration system may, like Mu, use a tetramer of active sites, but the architecture of the retroviral system has not yet been directly established (16).

Many mobile elements translocate by a cut-and-paste mechanism that is seemingly distinct from that of retroviral integration and Mu replication—the introduction of double-strand breaks at the ends of an element to generate an excised transposon that is then inserted into a target site. This paradigm was first established for the bacterial transposons Tn10 and Tn7 (1, 2, 17, 18), but it is now clear that numerous elements, including the P element of *Drosophila* (19), the Tc elements of *Caenorhabditis elegans* and other organisms (20), and likely many other transposition systems (1, 21), use this mechanism.

The fundamental chemistry at the heart of these reactions is actually the same as in Mu and retroviruses. The critical act in recombination is to expose the 3′-OH transposon ends, which are then joined to the target DNA. But with these cut-and-paste elements, recombination initiates by double-strand breaks at the transposon ends that separate the element from flanking donor DNA, rather than with Mu- and retrovirus-like single-strand nicks. Double-strand breaks result in cleavage ends of these at both the 3′ and 5′ elements; the resulting excised transposon is conceptually equivalent to the processed retroviral DNA. Recombination is completed by strand transfer reactions that join the 3′ transposon ends to the target DNA.

As these double-strand break elements are processed by the same chemistry as the retroviruses, it is not surprising that double-strand break transposases can have amino acid similarities to the retroviral recombinases. Essential D,D(35)E motifs are present in recombinases that execute cut-and-paste reactions (20–22), and critical acidic amino acids that may function like the D,D(35)E motif have also been identified in other transposases (17). Indeed, clusters of acidic amino acids that can interact with metal ions may be present in many recombinases.

How do the double-strand break transposases execute cleavages at both 3′ and 5′ tips of mobile elements? In the case of Tn10, a single gene product, the Tn10 transposase, executes all the DNA processing and transfer reactions (17). Tn7, however, executes the initiating double-strand breaks through a collaboration of two Tn7-encoded D,D(35)E proteins (22). In an alternative pathway to replication, phage Mu displays a distinctive mechanism for element excision via double-strand breaks; following MuA-mediated cleavage and strand transfer at the 3′ ends of Mu DNA, an as yet unidentified nuclease distinct from transposase itself cuts the connections between the 5′ ends of Mu and the flanking donor DNA, resulting in a simple insertion (2, 4).

The biological and conceptual connections between many different mobile elements have long been recognized (1). It is now clear that biologically diverse systems also share fundamental biochemical and structural similarities.

## References


1. D. E. Berg and M. M. Howe, *Mobile DNA* (American Society for Microbiology, Washington, DC, 1989).
2. K. Mizuuchi, *Annu. Rev. Biochem.* **61**, 1011 (1992).
3. S. Goff, *Annu. Rev. Genet.* **26**, 527 (1992).
4. B. D. Lavoie and G. Chaconas, *Curr. Top. Microbiol. Immunol.* **204**, 83 (1995).
5. D. Sherratt, in (*1*), pp. 163–184.
6. K. Mizuuchi, *J. Biol. Chem.* **267**, 21273 (1992).
7. A. Engelman and R. Craigie, *J. Virol.* **66**, 6361 (1992); J. Kulkosky *et al.*, *Mol. Cell Biol.* **12**, 2331 (1992); D. C. van Gent, A. A. M. Oude Groeneger, R. H. A. Plasterk, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9598 (1992).
8. T. A. Baker and L. Luo, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6654 (1994); K. Kim, S.-Y. Namgoong, M. Jayaram, R. M. Harshey, *J. Biol. Chem.* **270**, 1472 (1995).
9. F. Dyda *et al.*, *Science* **266**, 1981 (1994).
10. P. Rice and K. Mizuuchi, *Cell* **82**, 209 (1995).
11. O. Fayet *et al.*, *Mol. Microbiol.* **4**, 1771 (1990); P. Radstrom *et al.*, *J. Bacteriol.* **176**, 3257 (1994); R. Rezsohazy, B. Hallet, J. Delcour, J. Mahillon, *Mol. Microbiol.* **9**, 1283 (1993); T. G. Doak, F. P. Doerder, C. L. Jahn, G. Herrick, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 942 (1994).
12. W. Yang, W. A. Hendrickson, R. J. Crouch, Y. Satow, *Science* **249,** 1398 (1990); K. Katayanagi *et al.*, *Nature* **347**, 306 (1990); J. F. Davies II, Z. Hostomska, Z. Hostomky, S. R. Jordan, D. A. Matthews, *Science* **252**, 88 (1991).
13. M. Ariyoshi *et al.*, *Cell* **78**, 1063 (1994).
14. T. A. Baker, E. Kremenstova, L. Luo, *Genes Dev.* **8**, 2416 (1994); B. D. Lavori, B. S. Chan, R. G. Allison, G. Chaconas, *EMBO J.* **10**, 3051 (1991).
15. M. Mizuuchi, T. A. Baker, K. Mizuuchi, *Cell* **70**, 303 (1992); T. A. Baker and K. Mizuuchi, *Genes Dev.* **6**, 2221 (1992).
16. A. B. Hickman *et al.*, *J. Biol. Chem.* **269**, 29279 (1994).
17. N. Kleckner *et al.*, *Curr. Top. Microbiol. Immunol.* **204**, 49 (1995); R. M. Chalmers and N. Kleckner, *J. Biol. Chem.* **269**, 8029 (1994).
18. N. L. Craig, *Mol. Microbiol.* **5**, 2579 (1991); *Curr. Top. Microbiol. Immunol.* **204**, 27 (1995).
19. P. D. Kaufman and D. C. Rio, *Cell* **69**, 27 (1992).
20. H. G. van Luenen, S. D. Colloms, R. H. A. Plasterk, *ibid.* **79**, 293 (1994); J. C. Vos and R. H. Plasterk, *EMBO J.* **13**, 6125 (1994).
21. B. R. Calvi, T. J. Hong, S. D. Findley, W. M. Gelbart, *Cell* **66**, 465 (1991).
22. E. May and N. L. Craig, unpublished data; R. Sarnovsky and N. L. Craig, unpublished data.