Molecular Model for the Transposition and Replication of Bacteriophage Mu and other Transposable Elements

James A. Shapiro

PNAS 1979;76:1933-1937
doi:10.1073/pnas.76.4.1933

This information is current as of October 2006.

This article has been cited by other articles:
www.pnas.org#otherarticles

E-mail Alerts
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

Rights & Permissions
To reproduce this article in part (figures, tables) or in entirety, see:
www.pnas.org/misc/rightperm.shtml

Reprints
To order reprints, see:
www.pnas.org/misc/reprints.shtml

Notes:
Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements

(DNA insertion elements/nonhomologous recombination/site-specific recombination/replicon fusion/topoisomerases)

JAMES A. SHAPIRO

Department of Microbiology, The University of Chicago, Chicago, Illinois 60637

Communicated by Hewson Swift, December 11, 1978

ABSTRACT A series of molecular events will explain how genetic elements can transpose from one DNA site to another, generate a short oligonucleotide duplication at both ends of the new insertion site, and replicate in the transposition process. These events include the formation of recombinant molecules which have been postulated to be intermediates in the transposition process. The model explains how the replication of bacteriophage Mu is obligatorily associated with movement to new genetic sites. It postulates that all transposable elements replicate in the transposition process so that they remain at their original site while moving to new sites. According to this model, the mechanism of transposition is very different from the insertion and excision of bacteriophage λ.

Recent research on transposable elements in bacteria has provided important insights into the role of nonhomologous recombination in genetic rearrangements (1–4). These elements include small insertion sequences (IS elements), transposable resistance determinants (Tn elements), and bacteriophage Mu (5). There are detailed differences in the genetic behavior of these various elements (such as differences in specificity of site selection for insertion), but there is a consensus that they all share underlying recombination mechanisms (3, 5). Although this consensus originally included the bacteriophage λ (cf. refs. 3 and 4), the considerations elaborated below argue that phage Mu and other transposable elements differ radically from the elegant and simple λ model (6, 7).

The mechanisms by which transposable elements move from one genetic site to another are still unknown. However, there has been a rapid accumulation of information about the structural consequences of transposition events, the genetic control of transposition, and the replication of bacteriophage Mu from work in many laboratories. While some of this information is not yet based on completely unambiguous data, the outlines appear sufficiently clear to propose a partially detailed molecular model to explain the transposition process. This model also explains how phage Mu replicates. Its justification draws examples from work on various elements (particularly Mu, IS1, Tn3, and Tn5) on the assumption that they are all mechanistically equivalent as far as the details of the present model go. This proposal does not address the question of site selection for transposition events and specifically deals only with events that occur after an initial donor–target complex has been formed.

OBSERVATIONS TO BE EXPLAINED

A model for the mechanism of transposition and replication of phage Mu and the insertion of other transposable elements should be able to explain the following observations.

(i) Insertion of a transposable element into a genetic site

results in the duplication of a short oligonucleotide sequence in the target. The duplicated sequence brackets the inserted element (8, 9). This characteristic of the insertion process has been detected in sequence data on insertions of IS1, Tn5, Tn9, and Tn10 [all having nine base pair (bp) duplications] and also insertions of IS2, Tn3, gamma-delta, and phage Mu (5 bp duplications) (summarized in ref. 10). Grindley and Sherratt have proposed a model to account for this phenomenon (10), and some features of their proposal are incorporated in the present model. Studies on the specificity of insertion and comparison of sibling Tn9 insertions indicate that there is no functional relationship between the sequences flanking the donor insertion and the repeat generated at the target site (11). A useful way to consider these observations is to think of insertion on both sides of the 5- or 9-bp sequence that is to be duplicated.

(ii) Transposable elements can serve to fuse two replicons, and replicon fusion occurs when only one of the parental molecules originally contained a transposable element (12–14). In all cases in which there is data on the structure of the cointegrate product, it contains a directly repeated copy of the transposable element at each juncture of the two replicons (Fig. 1; refs. 12, 15, and 16). In the case of phage Mu, replicon fusion requires expression of the A gene but not of the B gene (12). In the case of Tn3, replicon fusion is observed only when a specific internal region of the transposon has been deleted (13, 15). Absence of this region causes a cis-dominant change in recombination behavior so that deleted transposons catalyze replicon fusion in the presence of a normal Tn3 (15). There is preliminary genetic evidence that derivatives of Tn5 can catalyze replicon fusion. These derivatives are generated by substitution of the internal Tn5 HindIII fragment by a trp fragment, and in this case the fusion phenotype is recessive to a wild-type Tn5 element (17). It has been suggested that the cointegrate fused replicons represent intermediates in the transposition process (15–17).

Abbreviations: IS, insertion sequence; bp, base pair.
The model involves four steps (which can be further subdivided) and is schematized in Fig. 3:

Step I: Four single-strand cleavages occur. These are probably not simultaneous and almost certainly not independent events. The donor molecule is cleaved on either strand at the extremity of the transposable element. The cleavage is drawn at the 3'-hydroxyl end on each strand. The endonuclease responsible for this cleavage is probably an element-specific enzyme (e.g., Mu A gene product). The target molecule is cleaved at sites chosen with greater or lesser specificity (depending on the element) to yield 5'- or 9-bp cohesive ends, as first postulated by Grindley and Sherratt (10). It is critical that the polarity of this cleavage be opposite that of the donor cleavage, so that this cleavage has been drawn at the 5'-phosphate ends of the oligonucleotide sequence on each strand. The endonuclease responsible for this cleavage is probably a cellular function that may be activated by the element-specific endonuclease, as first suggested by Grindley and Sherratt (10). Although the two pairs of cleavages...
must have opposite polarities, there is no compelling reason now for choosing the particular ones depicted in Fig. 3.

Step II: The donor and target strands are ligated to generate a y-shaped structure held together by the transposable element duplex as shown in Fig. 3. The ligation would almost certainly be catalyzed by a specific enzyme complex, perhaps including the same protein responsible for the original element-specific cleavages. Because both ends of the donor element have exposed 3'-hydroxyl groups and both ends of the target oligonucleotide have exposed 5'-phosphate groups, the joining can theoretically take place in either orientation. (The orientation may, however, be determined by the exact sequence of reaction.) Although there is probably some fraying at the ends of the transposable element duplex, this structure should be stable because the only regions that must be single-stranded are the 5- or 9-bp sequences from the target molecule.

I have described the reactions in steps I and II separately in order to clarify the sites of breaking and joining of DNA strands. This description is not meant to imply either a biochemical mechanism or a particular temporal sequence for these events. In order to guarantee maximal stability to the intermediate structures, for example, one could propose that there is initially only one cleavage of the target oligonucleotide followed by ligation to the donor element and then by a concerted cleavage–ligation reaction at the other target site. The types of biochemical reactions postulated are not unprecedented. T4 RNA ligase will join single DNA strands that are not held together by duplex regions (28). A number of topoisomerases will catalyze concerted cleavage–ligation (nicking–closing) reactions (29). In fact, the properties of one of these enzymes, DNA gyrase, are very similar to the postulated cellular function responsible for the staggered 5- or 9-bp cleavage of the target, for special conditions will induce gyrase to catalyze nonrandom DNA cleavages leaving a four-base cohesive end (ref. 30; A. Morrison and N. Cozzarelli, personal communication). The λ integrase complex will catalyze site-specific, concerted cleavage and ligation reactions involving different molecules to generate recombinant structures (7). And it has been proposed that the φX174 ctsA protein carries out a nonconcerted cleavage and ligation reaction that has a protein-stabilized nicked DNA intermediate (31).

Step III: There is filling in of the unpaired oligonucleotide sequences at the ends and semiconservative replication of the transposable element. The ligation step actually results in formation of a replication fork at each end of the transposable element. Displacement synthesis from the unligated 3'-hydroxyl groups will result in the formation of large single-stranded regions adjacent to an unligated 5' end, and so seems unlikely to proceed very far without discontinuous replication on the opposite strand. (This situation will not be altered by reversing the polarities of the original cleavages.) The replication may proceed from both forks to the interior of the element, from one fork completely across the element, or (conceivably but not likely) bidirectionally from an internal site. The first two possibilities provide a very attractive explanation of why Mu must transpose in order to replicate. Polymerization and completion of this replication require no new biochemical activities, but in some cases specific proteins may be involved (e.g., Mu B gene product during vegetative phage growth). The result of this replication will be two recombinant duplexes each containing a semiconserved transposable element adjacent to the target 5- or 9-bp sequence. This oligonucleotide sequence has been duplicated in the process. Note that if the donor and target molecules are both intact replicons, the a and b duplex arms will be connected, and the c and d arms will also be connected. In other words, the recombinant products in Fig. 3 will constitute the join regions of a fused replicon (Fig. 1).

Step IV: Site-specific reciprocal recombination takes place between the two transposable elements to generate both the original donor molecule (with some newly synthesized material in the transposable element) and the target molecule containing an inserted element flanked by a 5- or 9-bp repeat. Although homologous recombination systems can catalyze this step, there must be some recA-independent process because transposition occurs normally in recA strains. This recombination probably involves specific proteins (e.g., the product lacking in Tn5-‐trp hybrids; ref. 17) acting on specific sites (e.g., the site deleted in certain Tn3 mutants that generate replicon fusions; ref. 13). It is possible, however, that there is a general cellular recombination function that recognizes particular sites when the DNA is replicating (cf. ref. 16). In this case, I would suggest that steps III and IV overlap in time, so that the reciprocal recombination event occurs between duplicated segments of the replicating element before the replication process is finished. This last step is not essential for Mu replication, and its efficiency may vary from element to element. In the case of λ::Tn9 infection of rec- Escherichia coli cells, cointegrate formation (steps I–III) occurs approximately as often as transposition (steps I–IV) (16).

DISCUSSION

This model satisfactorily explains the six observations listed above and is not inconsistent with any of the other genetic phenomena [such as adjacent deletion formation (32)] that appear to be common to transposable elements. In fact, it is particularly instructive to consider the consequences of one simple alteration in the process just outlined: a failure to complete step IV. (A simple genetic formalism facilitates this kind of analysis because the net result of steps I–III is to generate two complete recombinant duplexes: a-element–oligonucleotide-d and c-oligonucleotide–element-b. Thus, it is not necessary to go through all the details of cleavage, ligation, and replication to work out recombinant structures.) If donor and target molecules are both intact circular replicons, the product will be a cointegrate replicon fusion with directly repeated elements, and it is certainly suggestive that Tn3 and Tn5 appear to generate fuses only when part of the transposon has been deleted (13, 15, 17). If the donor and target sites are both on the same replicon, the recombinant products of the replication step will depend on the orientation of the ligation step. If the duplex arms bracketing the transposable element (marked a and b in Fig. 3) and those bracketing the target site (c and d) are arranged as shown in the upper panel of Fig. 4, the products will be two circles. (a, b, c, and d in Fig. 3 are shown as A, B, C, and D in Fig. 4.) The larger one could represent an adjacent deletion mutant of the parental replicon. Both could represent the heterogeneous Mu-containing circles observed during Mu replication (19, 33). Thus, the model readily accounts for two observations that did not enter into the original formulation. If the duplex arms of the donor element and target site are arranged as depicted in the lower panel of Fig. 4, then the products will form an inversion bracketed by inverted repeats of the transposable element. It is tempting to speculate that such events might have given rise to structures like those of Tn5 and Tn10 (3). These hypothetical adjacent inversion events do not yield structures identical to those already observed by Kleckner and Ros (34). Although it is possible to generate their structures in the same four steps outlined above, I think the discrepancy provides an important test of the model. The model predicts that inversions such as those schematized in the lower panel of Fig. 4 will be the most common kind of adjacent inversion when there has been no selection for genetic fusion or loss of a particular transposon function. Although the data have not yet been published, Faelen and Toussaint refer to Mu-mediated inversions of the structure depicted in Fig. 4 (35).
A particularly appealing feature of the model is the intimate relationship it postulates between transposition and replication. This relationship has been proposed for phage Mu (19), and the model provides an explanation of many unusual facts of Mu replication. Furthermore, it makes clear predictions about the structure involved in the initiation of Mu-specific DNA synthesis.

The model makes several other genetic and biochemical predictions. Among these are the following:

(i) Transposition of an element to a new site does not remove the element from its original site. Thus, cells will accumulate multiple elements with time unless transposition is repressed. The discovery of a transposition repressor for Tn3 is particularly interesting in regard to this problem (13, 15).

(ii) Transposable element systems will code for three kinds of functions: a strand cleavage function for step I, a ligation function for step II, and a specific recombination function for step IV. As suggested above, one protein may participate in both element-specific cleavage and ligation reactions. The recombination function may be either a specific enzyme–site system (analogous to the integrase–att system of λ) or simply a site which makes use of some cellular rec-independent system for illegitimate recombination. Transposable elements that replicate, such as phage Mu, may also have specific proteins for step III.

(iii) The recombinant structures formed after step III are obligatory intermediates in transposition events. For two circular molecules this means a replication structure (Fig. 1). For genetically linear molecules (such as λ phages) this means that a transposable element will serve as a site for reciprocal recombination when the full process is aborted after step III. In such crosses, only one parental molecule will carry the transposable element, and recombination events will occur at various sites on the other parental molecule. Both recombinant products will carry the transposable element. These recombination events will not require DNA sequence homology nor homologous recombination systems and will differ from the postulated step IV reaction, which will occur at defined sites on both parental molecules.

(iv) The donor molecule in a transposition event will be regenerated with two segments of newly synthesized DNA on opposite strands, and the newly inserted transposon will contain parental DNA from the donor molecule in two segments on opposite strands. The transposable element in both donor and target molecules will consist of only hybrid (new-old) DNA.

There are four additional aspects of this model that merit emphasis. First, there is no autonomous existence of the transposable element. This is one of the striking features of phage Mu, in which even the virus particle contains an inserted prophage rather than an autonomous viral genome free of host DNA (19). Second, all reactions are limited to the local regions of the donor element and the target site. Thus, there is no requirement for physical circularity of either molecule. Third, the donor molecule is regenerated intact and does not have to commit suicide in order to donate its transposable element to a new site. Thus, it is possible to have transposition between sites in the same replicon. Fourth, the initial interactions between donor and target duplexes involve end-to-end joining of single DNA strands. Thus, there is no role for base-pairing (i.e., genetic homology) in the recombination process at the site of insertion. This feature distinguishes the present proposal from mechanisms based on the activity of restriction enzymes.

This model owes a great deal to the earlier proposal of Grindley and Sherratt (10). The ideas of staggered cleavages at the target site, of replicating the transposable element, and of regenerating the donor molecule were first explicitly stated by them. Thus, prediction i above is the same for both models and distinguishes them from models based on λ excision and insertion. There are, however, important differences between the two proposals. The model described above postulates that cleavages and ligations at both ends of the transposable element occur prior to DNA replication. The Grindley–Sherratt model postulates displacement synthesis across the transposable element between the first ligation at one of its ends and the subsequent ligation at the other end. Thus, the structure of transposition intermediates, the nature of cleavage and ligation reactions, and the functions of the element-specific enzymes are different. In particular, prediction iii of this model includes two potentially stable recombinant duplexes as intermediates. By postulating a strand switch during displacement synthesis across the transposable element, it is possible to generate recombinant duplex structures from the Grindley–Sherratt model. However, these are not intermediates for transposition, so that the presence of a rec-independent reciprocal recombination system for sibling elements is not a feature of their proposal. The nature of element-specific replication is also very different in the two models. According to Grindley and Sherratt, the transposable element is duplicated by two successive displacement reactions from free 3′-hydroxyl ends formed by separate cleavage events. These reactions can result in conservative replication of at least part of the transposable element. The model proposed here postulates strictly semiconservative replication of the transposable element. Thus, prediction i above the presence of only hybrid DNA in transposition products does not apply to the Grindley–Sherratt model. An additional (and more intriguing) difference is the possible role of a DNA replication termination function in transposition. To complete replication of the transposable element by the present model poses the same topological problems as finishing replication of any circular molecule that passes through a theta structure (36). Thus, the total process almost certainly requires the biochemical activities needed for normal termination and separation of daughter molecules. Without these, the two recombinant duplex products of step III would not be formed. Such an activity is not required by the Grindley and Sherratt replication scheme. The tests for
many of the alternative predictions of these two models await development of in vitro transposition systems. Both models indicate that transposition is a more complex process than integrative Λ recombination (7).

There are clearly many details to fill in. The mechanisms governing the formation of the initial donor–target complex remain to be elucidated. This is a particularly interesting question because both genetic and sequence data indicate that there are differences between the ways in which Tn9 and Tn10 choose insertion sites (11, 33). Steps III and IV are vague because the natures of both replication initiation and site-specific recombination events need to be specified. Nonetheless, this model has some original features, provides new explanations for previous observations, makes several testable predictions, and hopefully will aid in the design of both genetic and biochemical experiments for elucidation of the behavior of transposable elements.

Note Added in Proof. I thank Arianne Toussaint and Nigel Grindley for bringing some additional points about Mu replication and Mu-mediated chromosome rearrangements to my attention: (i) Mu DNA appears to replicate unidirectionally starting at the immunity (c) terminus (C. Wiffelman and P. van de Putte (1977), in ref. 3, pp. 329–333). (ii) When a Mu prophage mediates deletion formation at a distant region of the genome, the resulting mutant cell contains two copies of Mu—one linked to the deletion and one at the original site (35). Thus, in this case transposition has occurred without loss of the donor prophage. (iii) Both ends of the Mu prophage are required for the formation of an adjacent deletion (35). This is consistent with the model for adjacent deletion formation depicted in Fig. 4. (iv) Mu-mediated recombination events that require only one cycle of transposition and replication (e.g., adjacent deletion formation and integration of Agal into the chromosome of a Mu lysogen) occur at the same frequencies in the presence or absence of Mu B function (12, 35). On the other hand, Mu-mediated recombination events that require two or more cycles of transposition and replication (e.g., deletion formation at sites distant from a prophage) occur at the frequency of a single event in presence of Mu B function but at the frequency of two successive independent events in the absence of Mu B function (35). This observation suggests a function for the Mu B gene product: the presence of Mu B product permits the Mu A function to catalyze repeated rounds of transposition (such as needed to initiate successive rounds of vegetative replication). In the absence of Mu B product, the Mu A function can catalyze only a single round of transposition and replication. This hypothesis predicts that density-labeled Mu A + B phage will form hybrid density prophages after infection of sensitive bacteria.

I thank Lorne MacHattie for introducing me to the mysteries of Tn9, for long hours of invaluable discussions, and for making these figures understandable. I am grateful to Nicholas Cozzarelli, Alan Morrison, Pat Brown, Pat Higgins, and Ken Kreuzer for communicating unpublished results and helping me ensure the biochemical plausibility of the model; to Ahmad Bukhari, Martha Howe, and Arianne Toussaint for many helpful lessons on phage Mu biology; to Ron Gill and Jeff Miller for communicating unpublished results on replicon fusion experiments; and to Nigel Grindley for going over our respective models to clarify their similarities and differences. Research in my laboratory is supported by grants from the National Institute of General Medical Sciences (GM 24060), the University of Chicago Cancer Research Center (CA 14598-05-307), and the National Science Foundation (PCM 77-08591). I am the recipient of a Research Career Development Award (1-K04-AI00118-01) from the National Institute of Allergy and Infectious Diseases.