MECHANISM OF IS1 TRANSPOSITION IN E. COLI: CHOICE BETWEEN SIMPLE INSERTION AND COINTEGRATION

SUSAN W. BIEL AND DOUGLAS E. BERG

Departments of Microbiology and Immunology and of Genetics (Box 8093), Washington University Medical School, St. Louis, Missouri 63110

Manuscript received March 3, 1984
Revised copy accepted June 2, 1984

ABSTRACT

Insertion element IS1 and IS1-based transposon Tn9 generate cointegrates (containing vector and target DNAs joined by duplicate copies of IS1 or Tn9) and simple insertions (containing IS1 or Tn9 detached from vector sequences). Based on studies of transposon Tn5 we had proposed a conservative (non-replicative) model for simple insertion. Others had proposed that all transposition is replicative, occurring in a rolling circle structure, and that the way DNA strands are joined when replication terminates determines whether a simple insertion or a cointegrate is formed. We selected for the transposition of amp and cam resistance markers from pBR322::Tn9 plasmids to an F factor in recA- E. coli and identified products containing three and four copies of IS1, corresponding to true cointegrates (from monomeric plasmids), and simple insertions (from dimeric plasmids). The simple insertions with four copies of IS1 outnumbered those with three by a ratio of about 3:1, whereas true cointegrates containing three copies of IS1 were more numerous than those with four. A straightforward rolling circle model had predicted that the simple insertions containing three copies of IS1 should be more frequent than those with four. Because we obtained the opposite result we propose that simple insertions only arise when the element fails to replicate or if replication starts but then terminates prematurely. The two classes of products, simple insertions and cointegrates, reflect alternative conservative and replicative fates, respectively, of an early intermediate in transposition.

ANALYSES of transposable elements in prokaryotes have shown that their movement is independent of extensive DNA sequence homology and of the rec genes necessary for classical recombination. Transposition depends, instead, on element-specific proteins called transposases. These proteins may act by binding the ends of their cognate elements, introducing single- or double-strand breaks at the junctions of the element and its vector, cleaving the target DNA, joining element and target DNAs and, in certain cases, inducing the specific replication of the element (for reviews see IIDA, MAYER and ARBER 1983; BERG and BERG 1983; KLECKNER 1981).

Abbreviations used in text: amp, cam and tet, genes encoding resistances to ampicillin, chloramphenicol and tetracycline, respectively; ori, the origin of pBR322 replication.

1 Present address: United States Department of Agriculture, Southern Regional Research Center, 1100 Robert E. Lee Boulevard, New Orleans, Louisiana 70179.
Two types of transposition products have been described, cointegrates and simple insertions. Cointegrates consist of complete vector and target DNAs joined by duplicate copies of the transposable element. These can arise in \( \text{recA}^+ \) (recombination-proficient) cells after the completion of transposition: a reciprocal (homologous) crossover between different DNA molecules, each with a copy of the element, will join them together in a recombinational cointegrate. Cointegrates can also arise without recombination when only one DNA molecule contains the transposable element; in this case their formation requires replication during transposition (Shapiro 1979; Arthur and Sherratt 1979; Galas and Chandler 1981; Harshey and Bukhari 1981).

Simple insertions, in contrast, consist of the element moved to new genomic locations and free of vector sequences. Some arise secondarily from cointegrates by homologous recombination mediated by the \( \text{recA} \) protein or, in the case of transposon Tn3 and related elements, by a site-specific reaction termed resolution (see Heffron 1983). However, simple insertions are also found as the primary products of transposition, and there are two views of how these arise. Based on our analyses of Tn5, which only gives simple insertions, we proposed a model of conservative transposition involving double-strand cleavages to separate the element from its vector and the splicing of the element into a cleaved target DNA molecule. In this model DNA synthesis is limited to the repair of gaps at the element-target junctions and leads only to the characteristic target sequence duplication [9 base pairs (bp) for Tn5], not to duplication of the element (Berg 1977, 1983). Other models have been built assuming that simple insertion, like cointegration, involves replication (Galas and Chandler 1981; Harshey and Bukhari 1981; Hirschel, Galas and Chandler 1982; Hirschel et al. 1982; Kleckner 1981).

The insertion element IS1 (Galas and Chandler 1982) and bacteriophage Mu (Harshey and Bukhari 1981) are intriguing because they generate both cointegrates and simple insertions. To explain how IS1 might give rise to both types of products Galas and Chandler (1981) proposed that transposition is invariably linked to replication of the element in a rolling circle model (Figure 1); an equivalent model to explain the behavior of phage Mu was proposed independently by Harshey and Bukhari (1981). In their models transposition is initiated when the transposase protein joins a strand at one end of the element to the target DNA. The fork that is generated promotes replication of element and vector sequences, and only when the wave of replication is terminated by the action of transposase at another recognition site is the choice between cointegration or simple insertion made.

We report here that among simple insertions from dimeric forms of pBR322::Tn9 plasmids those with four copies of IS1 outnumber those with three copies of IS1. This outcome suggests to us that simple insertions do not arise by a rolling circle mechanism. We propose, therefore, that IS1 transposition can, but need not be, replicative. Both products may arise from a single intermediate: cointegrates whenever the element replicates and simple insertions if the element does not.
FIGURE 1.—Model of transposition associated with unidirectional (rolling circle) replication (adapted from Harshey and Bukhari 1981; Galas and Chandler 1981). Thick lines correspond to strands of the transposable element. The top circle represents a monomeric pBR322::Tn9 plasmid in which direct repeats of ISI (thick lines) separate the cam gene (of Tn9) and the sequences of pBR322. Sequences replicated during transposition are indicated by dashed lines. Other models for transposition have been presented by Berg (1977, 1983), Grindley and Sherratt (1978), Arthur and Sherratt (1979), Shapiro (1979) and Ohtsubo et al. (1980).

MATERIALS AND METHODS

Bacterial plasmids: The F factor pOX38 (Guyer et al. 1980) used as a target for transposition is about 52 kb long and contains the genes of F needed for autonomous replication and conjugal transfer but lacks the insertion elements present in wild-type F. The locations of restriction endonuclease cleavage sites in plasmid pBR322 (Stukcliffe 1978) and in transposon Tn9 which contains direct terminal repeats of ISI (called IS1A and IS1B) and encodes resistance to chloramphenicol (Cam') (Alton and Vapnek 1979; Ohtsubo and Ohtsubo 1978; Machida et al. 1982) are known. The pBR322::Tn9 plasmids pBRG44 and pBRG45 were constructed by transposition of Tn9 from an E. coli chromosomal site to a derivative of plasmid pBR322 mutated in its EcoRI site [by EcoRI cleavage, repair with DNA polymerase I (Klenow fragment) and blunt end ligation]. Our selection for transposition of Tn9 to pBR322 has been described (Biel, Adelt and Berg 1984). Restriction endonuclease mapping showed that in pBRG44 the Tn9 element is bracketed by the SauIHA sites at pBR322 positions 3225 and 3330. It is oriented such that IS1A is closest to amp and IS1B is closest to ori (Figure 2). In pBRG45, Tn9 is located in the promoter region of pBR322's tet gene just upstream from the HindIII site and oriented such that IS1A is
Figure 2.—Monomeric (left) and dimeric (right) forms of the pBR322::Tn9 plasmid pBRG44.
Abbreviations: cam, amp and tet, genes encoding resistances to chloramphenicol, ampicillin and tetracycline, respectively; ori, origin of replication; R and H, sites cleaved by restriction endonucleases EcoRI and HindIII, respectively. The EcoRI site normally present in pBR322 has been mutated so that the only site in pBRG44 cleaved by EcoRI is in the cam gene. The thick lines represent IS1 sequences, and their orientations are indicated by arrows.

nearest the HindIII site. pBRG44 encodes resistance to chloramphenicol, ampicillin and tetracycline; pBRG45 encodes resistance to chloramphenicol and ampicillin but not to tetracycline because the Tn9 insertion inactivated the tet gene promoter.

Dimeric forms of pBRG44 (Figure 2) and pBRG45 were generated by homologous recombination in vivo and maintained in recA+ E. coli (see BERG 1983).

Bacterial strains: All bacterial strains were derived from E. coli K-12. DB1504 (BERG, WEISS and CROSSLAND 1980) is F− recA+ ΔtrpE5 Δ(proB-lac) and was used as the host to generate dimeric forms of pBR322::Tn9 plasmids. DB1986 (BERG 1983) is ΔtrpE5 Δlac Δ(srl::Tn10-recA306) and harbors the F factor pOX38. Derivatives of DB1986 carrying pBR322::Tn9 plasmids were used as donor strains in bacterial matings as described (BIEL, ADELT and BERG 1984). DB1648 (EGNER and BERG 1981) is F− recA1 ΔtrpE5 Δ(proB-lac) rpsL and was used as the recipient in conjugation.

General methods: The DNAs of small (pBR322 and pBR322::Tn9) and also of the larger (pOX38-pBR322::Tn9) chimeric plasmids were extracted by an alkaline-SDS lysis procedure (BIRNBOIM and DOLY 1979). Recipes for media, protocols for bacterial growth and matings and electrophoresis in agarose gels have been described (BERG 1983; BERG, SCHMANDT and LOWE 1983; BIEL, ADELT and BERG 1984; MANIATIS, FRITCHE and SAMBROOK 1982). Restriction endonucleases were used according to the suppliers’ (New England Biolabs and Bethesda Research Laboratories) instructions.

Rationale

Definitions: Chimera is used as a general term to indicate any product of transposition between replicons. Simple insertions consist of a fragment of the donor molecule moved to a new location and separated from previously adjacent vector segments. By contrast, cointegrates consist of vector and target sequences joined by duplicate copies of the mobile element.

Experimental approach: The transposable element Tn9 contains direct terminal repeats of the smaller mobile element IS1 bracketing a region encoding chloramphenicol resistance (see Figure 2). The IS1 elements encode the proteins and provide the sites necessary for Tn9 movement, and probably any short DNA segment becomes transposable when bracketed by a pair of IS1
elements (Machida et al. 1982). Chandler, Clerget and Galas (1982) studied the movement of IS1-based transposons from sites in a repressed prophage to a plasmid and found the probability of transposition to be a simple inverse function of DNA length. They interpreted this outcome as support for their rolling circle model, a reflection they believed of a gradual decay of a specialized replication fork. We have tested this interpretation using dimeric and monomeric forms of pBR322::Tn9 plasmids (Figure 2).

Selection for the joint transposition of cam and amp determinants from dimeric pBR322::Tn9 plasmids permits the recovery of several classes of simple insertions (Figure 3). One class contains three copies of IS1 and single copies of the cam and amp segments; two other classes contain four copies of IS1, one with duplicate copies of the cam segment and a single copy of the amp segment, and the other with a single copy of the cam segment and duplicate copies of the amp segment. In each class only part of the dimeric donor plasmid had transposed and, thus, each is a simple insertion, not a cointegrate. According to the Galas-Chandler rolling circle model, simple insertions containing three copies of IS1 should outnumber those containing four copies of IS1 because the replication fork need travel less distance to generate the former class.

pBR322::Tn9 plasmids are nonconjugative, and transposition of their resistance genes to the F factor pOX38 can be selected in bacterial matings (Berg et al. 1980; Galas and Chandler 1982). Derivatives of strain DB1986 harboring the pBR322::Tn9 plasmids pBRG44 and pBRG45 were crossed with recA- recipient strain DB1648, and Cam' Amp' exconjugants were selected. Electrophoresis of plasmid DNA extracted from exconjugants generally revealed a single DNA species somewhat larger (about 8-20 kb) than pOX38 (Galas and Chandler 1982; Biel, Aedelt and Berg 1984). However, a small fraction (about 15%) contained additional prominent bands of smaller DNAs, generally the size of pBR322::IS1 or pBR322::Tn9. Our tests (Biel, Aedelt and Berg 1984) indicated that this heterogeneity reflects an instability of certain chimeras unrelated to the mechanism of transposition. The breakdown continued after transfer to new hosts, and cells harboring the unstable chimeras tended to make small colonies with fast growing sectors. We attributed this instability to a combination of factors including (1) rare recA-independent homologous recombination (which occurs even in plasmids that contain no IS elements; Doherty, Morrison and Kolodner 1983), (2) deleterious effects of certain insertions in pOX38 when maintained at high copy number and (3) growth rate selection of clones in which smaller plasmids had outcompeted these chimeras in replication (Biel, Aedelt and Berg 1984). Because unstable chimeras were infrequent, instability was apparently unrelated to transposition, and characterization of chimeric DNA structures is easiest with homogeneous plasmid preparations, the unstable chimeras were not studied further.

The structures and numbers of IS1 elements present in stable pOX38-pBR322::Tn9 can be inferred from the patterns of DNA fragments produced by digestion with appropriate restriction endonucleases. The pBR322::Tn9 plasmids used here contain a single HindIII site in the pBR322 component
RESULTS AND DISCUSSION

We analyzed the structures of simple insertions formed by the transposition of the cam and amp determinants of dimeric pBR322::Tn9 plasmids to the F factor pOX38 in recA^-E. coli. The data in Table 1 show that simple insertions containing four IS1 elements outnumber those containing only three by about 3:1. In addition, among simple insertions with four IS1 elements, those with two copies of the pBR322 segment were about eightfold more frequent than those with two copies of the cam segment (designated type II and type I, respectively, in Table 1).

Inherent in the rolling circle model of Galas and Chandler (1981) and of Harshey and Bukhari (1981) was a prediction that the smallest of the simple
MECHANISM OF IS1 TRANSPOSITION

I. MECHANISM OF IS1 TRANSPOSITION

insertions would be most frequent because their formation would have entailed the least DNA synthesis: expected to predominate was the class with three rather than four IS1 elements and, among those with four IS1 elements, the type with two copies of the cam segment and one of the amp segment (the 1102-bp cam segment is about one-fourth the length of the 4363-bp pBR322 plasmid). Our experiments (Table 1) gave just the opposite result.

Among true cointegrates (derived from monomeric pBR322::Tn9 plasmids), those with three IS1 elements were severalfold more frequent than those with four. This outcome is predicted by several models of transposition (including the rolling circle model) and thus serves as a valuable control. It is because the distribution of simple insertions from dimeric plasmids is opposite to that predicted by the rolling circle model of GALAS and CHANDLER (1981) and HARSHLEY and BUKHARI (1981) that it became necessary to imagine how else transposition might occur.

Elements related to Tn3 must move by a replicative process, since they invariably generate cointegrates (GILL et al. 1978; ARTHUR and SHERRATT 1979; HEFFRON 1983). However, the diversity evident in comparisons of DNA sequences, functional organization, regulation and transposition product struc-

Figure 4.—Representative digests of pOX38-pBR322::Tn9 chimeras. In each pair of lanes, the left contains the EcoRI digest, and the right contains the HindIII digest. The numbers 4, 5, 4, 6, etc. above each pair of lanes indicate the number of IS1 elements that we interpret these chimeras to contain. The designations (aca), (cacac) and (cac) indicate the arrangement of pBR322 (a) and chloramphenicol resistance (c) segments deduced for these chimeras as shown in Figure 3.
TABLE 1
Size of DNA segments transposed

<table>
<thead>
<tr>
<th>Donor</th>
<th>3 copies IS1</th>
<th>4 copies IS1</th>
<th>5 copies IS1</th>
<th>&gt;5 copies IS1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimeric pBRG44</td>
<td>12</td>
<td>4</td>
<td>32</td>
<td>27</td>
</tr>
<tr>
<td>Dimeric pBRG45</td>
<td>12</td>
<td>5</td>
<td>45</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>86</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>Monomeric pBRG44</td>
<td>72</td>
<td>11</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Monomeric pBRG45</td>
<td>46</td>
<td>1</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>118</td>
<td>40</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

* Each pOX38-pBR322::Tn9 chimera was selected to contain both the *cam* and the *amp* determinants of the indicated pBR322::Tn9 plasmid and was characterized by restriction endonuclease digestion (Figures 3 and 4). Each chimera is independent, having been derived from a mating of a separate donor subclone of strain DB1986 harboring the indicated pBR322::Tn9 plasmid with recipient strain DB1648.

* Type I and type II indicate the arrangement of the segments encoding chloramphenicol resistance (of Tn9) and ampicillin resistance (of pBR322) in chimeras containing four copies of IS1 (see Figure 3). Type I contains two copies of *cam* and one copy of *amp* (cac subclass), and type II contains two copies of *amp* and one copy of *cam* (aca subclass) (see Figure 3).

atures suggests that the current test of prokaryotic transposable elements may be derived from several different ancestral sources and that there may be more than one mechanism of transposition. Thus, we have suggested that transposon Tn5, unlike Tn3, may move by a conservative process since its transposition does not result in cointegrates (BERG 1977, 1983).

We propose that IS1-mediated transposition leading to simple insertion is conservative, whereas leading to cointegrates is replicative (Figure 5). In our model, transposition is initiated by the binding of the transposase protein to both ends of the element, the cutting of one or both DNA strands at each of the element-vector junctions and the joining of the element’s free ends to the cleaved target DNA. Single-strand breaks at the element-vector junctions lead to a pair of replication forks and to cointegrates. Double-strand breaks, on the other hand, separate vector and element sequences, do not permit the formation of replication forks and hence result in simple insertions. Thus, in our model, the structure of the transposition product is determined at the initiation step, not at termination.

The concentration of transposase may limit the rate of transposition, and transposase may be synthesized only sporadically from any single copy of IS1. Studies with mutated IS1 elements had shown that the IS1-encoded transposase operates effectively only in cis, and that it is most active on the element that had encoded it (MACHIDA et al. 1982). This suggests that transposase may bind preferentially to an end of the element from which it was synthesized and then slide along the donor molecule in search of a second recognition site.

Figure 6 diagrams how sliding of transposase could account for the relative abundance of simple insertions containing four copies of IS1 elements and also
MECHANISM OF IS1 TRANSPOSITION

For the excess of the \textit{aea} subclass among simple insertions with four \textit{IS1} elements (Table 1). Transposase synthesized from a single \textit{IS1} element (arbitrarily, \textit{IS1} number 1), binds to the "I" end of that element. It migrates until a second recognition site is bound, most frequently the I site of \textit{IS1} number 4, less frequently the O end of \textit{IS1} number 3, and results in simple insertions with four and three copies of \textit{IS1}, respectively. Alternative binding to the O ends of \textit{IS1} numbers 1 and 2 would generate the minority (\textit{cac}) subclass of simple insertions containing four \textit{IS1} elements.

Since true cointegrates from monomeric pBR322::Tn9 containing three copies of \textit{IS1} outnumber those with four copies of \textit{IS1}, the \textit{cac} subclass of cointegrates might, similarly, have been expected to outnumber the \textit{aea} subclass since less DNA need be replicated. That the expected excess was not observed

**Figure 5.**—Proposed formation of simple insertions and cointegrates from a single species of transposition intermediate. Represented here is transposon Tn9 which contains direct terminal repeats of the insertion element \textit{IS1} (thick lines). We propose that a key intermediate in transposition is generated by the joining of single strands at each end of the element to the target DNA. The choice between conservative (left branch) and replicative (right branch) fates depends on the relative probabilities of double- vs. single-strand breakage of the element vector-junctions prior to the onset of replication. Alternatively, breakage of arms of replication forks anytime before they meet would also generate simple insertions and would involve a limited amount of DNA synthesis instead of being completely conservative. The drawing depicts the transposition of the entire Tn9 element. The transposition of a single \textit{IS1} element would be initiated by breaks at its two ends.
FIGURE 6.—Proposed tracking patterns of transposase synthesized sporadically from an IS1 element. Solid boxes, IS1 elements; filled circles, transposase; dashed and dotted lines, possible paths taken by transposase synthesized from a single IS1 element; stippled boxes, segments transposed when transposase acts at the indicated ends of IS1 elements numbers 1 and 3 or 1 and 4. O and I correspond to the outside and inside ends of IS1 elements, respectively, as arranged in transposon Tn9. Even though the O and I ends of a single IS1 element differ in DNA sequence, because IS1 is repeated directly in Tn9, the O and I ends of adjacent IS1 elements are identical in sequence. Even though the DNA is depicted here in extended form, we imagine that when transposase tracks in search of a second recognition site it may remain complexed with the first binding site. The two ends of the element would, accordingly, be brought together by the transposase protein complex (see BERG 1977).

(Table 1) also suggests to us fundamental differences between simple insertion and cointegration early in the transposition pathway. Transcription impinging on the end of an IS element decreases its ability to participate in transposition (SASAKAWA et al. 1982; MACHIDA et al. 1983; CHANDLER, CLERGET and GALAS 1982; BIEL, ADELT and BERG 1984). The unexpected abundance of the aca subclass among cointegrates may reflect this positional effect on those transposition events that are replicative.

We conclude that the striking inverse dependence of transposition frequency on DNA length during simple insertion events (CHANDLER, CLERGET and GALAS 1982; this work) may reflect only the probability that transposase bound to one recognition site will encounter a second site and thus be able to initiate transposition. It does not require that transposition occur by a replicative mechanism.

Because IS1, unlike Tn5, generates cointegrates as well as simple insertions, it seems possible that simple IS1 insertion might also result from an initially replicative event interrupted by chance cleavages of the replication forks. This modification envisions a variable amount of DNA synthesis associated with such insertions, as has been found during the transposition of bacteriophage Mu in vitro (MIZUUCHI 1983). However, even this proposal, simple insertion due to an interruption of replication fork movement, differs fundamentally from roll-
ing circle models which require a full round of DNA synthesis during every successful transposition event and commitment to simple insertion or to coinTEGRATION only when the replication phase is complete.

We are grateful to G. ADELT for skillful technical assistance and to C. M. BERG for critical reading of this manuscript. This work was supported by United States Public Health Research grants 5 RO1 AI 14267 and 1 RO1 AI 18980 to D. E. B. S. W. B. was supported by Postdoctoral Training grant T32 AI-07172.

LITERATURE CITED


Corresponding editor: G. MOSIG