

Structure of a cloned circular Moloney murine leukemia virus DNA molecule containing an inverted segment: Implications for retrovirus integration

(transposable DNA elements/long terminal repeats)

CHARLES SHOEMAKER, STEPHEN GOFF, ELI GILBOA, MICHAEL PASKIND, SUDHA W. MITRA, AND DAVID BALTIMORE

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Contributed by David Baltimore, April 14, 1980

ABSTRACT Closed circular Moloney murine leukemia virus (M-MuLV) DNA was prepared from recently infected cells and cloned in a λ vector. Four classes of cloned M-MuLV inserts were found: Class 1, full length 8.8-kilobase (kb) inserts with two tandem long terminal repeats (LTRs) of 600 base pairs; class 2, 8.2-kb inserts with a single copy of a LTR; class 3, M-MuLV DNA inserts with various portions deleted; and class 4, an 8.8-kb insert with an internal sequence inversion. Determination of nucleotide sequence at the junction between the two LTRs from a class 1 insert suggested that circularization occurred by blunt-end ligation of an 8.8-kb linear DNA. The class 4 molecule had an inversion that was flanked by inverted LTRs, each of which had lost two terminal base pairs at the inversion end points. Also, four base pairs that were present only once in standard M-MuLV DNA were duplicated at either end of the inversion. This molecule was interpreted as resulting from an integrative inversion in which M-MuLV DNA has integrated into itself. Its analysis thus provided explicit information concerning the mechanism by which retrovirus DNA integrates into host cell DNA. Models of retrovirus integration based on bacterial DNA transposition mechanisms are proposed.

Retroviruses are eukaryotic RNA viruses that, during their multiplication cycle, produce double-stranded DNA reverse transcripts capable of integrating into the host cell's genome (for review see ref. 1). The sequence of events involved in the integration of retrovirus DNA has yet to be demonstrated.

Reverse transcription of retroviral RNA produces linear double-stranded DNA in which both ends contain an identical sequence of 300-1200 nucleotides called a long terminal repeat (LTR) (2-4). A portion of the linear viral DNA molecules can enter the nucleus, whereupon some become circular (5) and contain either one or two copies of the LTR (2, 6). A mouse gene, *Fv-1*, appears to prevent both circularization and integration (7), which suggests that circularization may be a prerequisite to integration of certain retrovirus DNAs, although other explanations of those results are possible.

Integration of retrovirus DNA occurs at multiple chromosomal sites even within a single host cell (8-12); thus integration occurs either at random chromosomal sites or with low site specificity. The integrated proviral DNA has an orientation that is colinear with the unintegrated linear form and has LTR sequences at both ends (10-12), which suggests that retrovirus DNA recombines with host cell DNA at the termini of the LTRs.

Molecular cloning of retrovirus-related DNAs has provided a potent method for analyzing the detailed structure of individual molecules (13-15). Here we describe the cloning and

analysis of closed circular Moloney murine leukemia virus (M-MuLV) DNA that was obtained from recently infected cells. In addition to standard M-MuLV DNA that contained either one 600-base-pair LTR or two tandem LTRs, we found two additional classes of cloned insert. One had deletions of various lengths; another, represented by a single isolate, contained full-length M-MuLV but with an inverted segment.

Recombination occurring between two regions within a single circular DNA molecule can produce either two smaller circular molecules or an inversion of the DNA between the recombination sites. Analysis of the cloned M-MuLV DNA containing an inversion demonstrated that both inversion end points occurred at the termini of LTR sequences. This structure thus appeared likely to have resulted from a circular M-MuLV DNA molecule integrating within itself. Detailed analysis provided strong evidence that this molecule does indeed represent an integrative inversion and that the integration mechanism is analogous to that of prokaryotic transposable DNA elements (16, 17).

MATERIALS AND METHODS

Preparation of Circular Moloney Viral DNA. Clone-1 M-MuLV (18) was twice recloned by single-cell/single-virus techniques (19), and the resulting producer cell line (clone 3A) was used to prepare virus stocks. NIH/3T3 cells were infected with M-MuLV for 2 hr at a multiplicity of 1. After 24-hr incubation at 37°C, the cells were extracted for unintegrated viral DNA by the Hirt procedure (20). The DNA was banded on a CsCl/ethidium bromide gradient (21) and fractions were collected. Portions of each fraction were spotted onto nitrocellulose and analyzed for M-MuLV-specific DNA by hybridization with ³²P-labeled M-MuLV DNA (22). Fractions containing the lower band of circular viral DNA were pooled, extracted with isopropanol, dialyzed, and precipitated with ethanol.

Cloning of Circular M-MuLV DNA. Charon 21A phage (23) DNA was treated with *Hind*III (New England BioLabs) and calf intestinal phosphatase (Boehringer Mannheim). The DNA fraction prepared from infected cells was cleaved with *Hind*III, extracted with phenol, and reprecipitated with 1 μ g of vector DNA. The mixture was ligated with T4 DNA ligase (New England BioLabs) in 5 μ l of buffer (50 mM Tris, pH 7.5/15 mM dithiothreitol/10 mM MgCl₂/1 mM ATP) for 20 hr at 14°C. The product DNA was then packaged into phage (24), and the resulting phage stock was plated on LE392 cells. Plaques were screened by using the procedure of Benton and Davis (25). Approximately one M-MuLV recombinant plaque was ob-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: M-MuLV, Moloney murine leukemia virus; LTR, long terminal repeat; kb, kilobase; IS, insertion sequence(s).

tained per 10³ infected NIH/3T3 cells. Phage containing M-MuLV sequences was purified by repeated cycles of screening and plating.

Preparation and Analysis of Cloned M-MuLV DNA. Large-scale phage lysates (26) or small plate stocks (27) were prepared, and DNA was isolated from the phage. Restriction endonuclease (New England BioLabs) digestions were performed under the conditions recommended by the supplier. DNA labeling and sequence determination were performed by the method of either Maxam and Gilbert (28) or Maat and Smith (29).

Transfer of recombinant inserts from Charon 21A to the plasmid pBR322 was performed as follows. Purified pBR322 closed circular DNA was digested with *Hind*III and treated with calf intestinal phosphatase. This DNA was ligated to the *Hind*III-digested recombinant Charon 21A DNA and then used to transform χ 1776 *Escherichia coli* (30) by the method of Enea *et al.* (31). Plasmid DNA from the transformed ampicillin-resistant, tetracycline-sensitive bacterial colonies was purified as described by Katz *et al.* (32).

Recombinant DNA procedures were carried out according to the National Institutes of Health guidelines in force at the time the work was performed.

RESULTS

Molecular Cloning of M-MuLV DNA. Unintegrated, closed circular M-MuLV DNA was prepared from infected NIH/3T3 cells and cleaved into full-length linear molecules at its single *Hind*III site (3). These molecules were ligated to *Hind*III-digested Charon 21A DNA and then packaged *in vitro* into λ phage particles. Recombinant plaques containing M-MuLV DNA (about 2% of the total) were detected by plaque-filter hybridization, and 20 positive clones were picked at random for purification.

Four Classes of M-MuLV Inserts. The recombinant inserts were cleaved from the phage DNA with *Hind*III and further digested with *Xho* I (which cuts M-MuLV DNA once; see map in Fig. 2), and the DNA fragments were resolved by electrophoresis through 0.7% agarose gels (Fig. 1). By this analysis we could classify the 20 recombinant phages into four classes. Class 1 recombinant phage contained an 8.8-kilobase (kb) insert (seen as 5.5- and 3.3-kb fragments) but with some 8.2-kb insert also present (seen as 4.9- and 3.3-kb fragments). Because both sizes of DNA arose from a single plaque isolate, some of the 8.8-kb insert can apparently lose 600 base pairs during passage in *E. coli*. Inserts of both sizes were transferred to the plasmid pBR322, where they were stable and could be purified separately. Restriction enzyme mapping showed that the 8.8-kb insert consisted of a full-length copy of M-MuLV DNA containing two tandem copies of the LTR (Fig. 2). The map presented in Fig. 2 is identical (at all sites examined) to the map of infectious, linear M-MuLV DNA (3), although the molecule is permuted because the *Hind*III site determines its ends. The 8.2-kb DNA was identical to the 8.8-kb insert except for the presence of only a single LTR. Apparently, homologous recombination between the tandem, directly repeated LTRs occurs during passage of the phage through *E. coli* (13-15). The frequency of this event is probably increased by the fact that an 8.8-kb insert in Charon 21A DNA is approaching the upper size limit for packaging into phage coats, thus creating a selective advantage for those DNA molecules containing an 8.2-kb insert.

Class 2 recombinant phage contained only inserts of 8.2-kb, which were indistinguishable from the 8.2-kb inserts present in the class 1 population (Fig. 1). Although molecules of this type have been observed *in vivo* in recently infected cells (6),

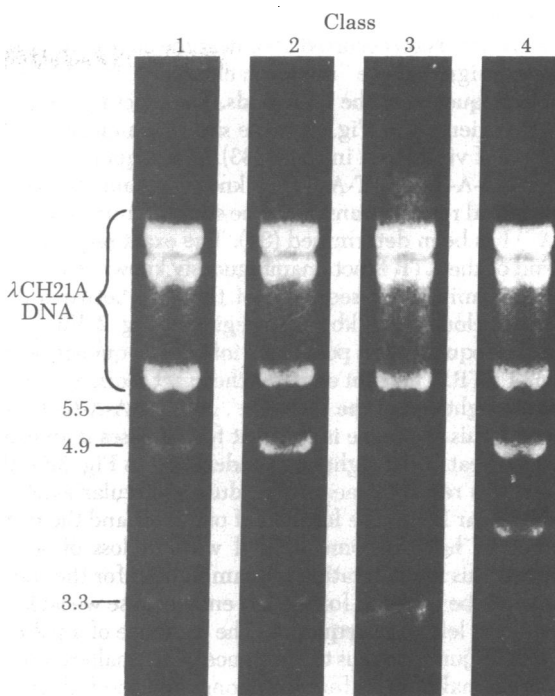


FIG. 1. Four classes of cloned M-MuLV DNA inserts. Charon 21A phage containing M-MuLV DNA inserts were prepared and plaque-purified. DNA was prepared from the individual isolates, digested with *Hind*III and *Xho* I, resolved by electrophoresis on 0.7% agarose gels, and stained with ethidium bromide. Lanes providing representative examples of the four distinct classes of recombinant inserts are displayed. Of 20 clones that were analyzed, 3 belonged to class 1, 11 to class 2, 5 to class 3, and 1 to class 4. Sizes are shown in kb.

the predominance of this class of recombinant phage among those analyzed (11 of 20) was probably enhanced as a consequence of the selective pressure favoring loss of a LTR through homologous recombination during plaque purification.

Several recombinant phage isolates (class 3) contained inserts of various sizes <8.2 kb. One clone of this class with an insert of 6.5 kb was analyzed by restriction enzyme mapping. It consisted of apparently normal M-MuLV DNA with a single deletion that excised the LTRs and adjacent sequences on both sides.

Class 4 recombinant phage was represented by only a single isolate. It contained an insert of 8.8 kb that did not produce 8.2-kb variants during passage. Detailed restriction enzyme mapping, some of which is displayed in Fig. 2, indicated that, relative to a standard 8.8-kb insert, this insert contained an in-

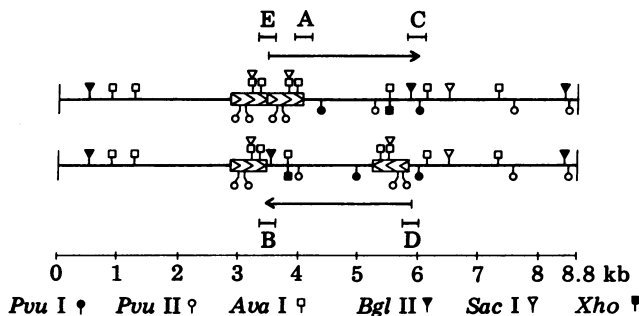


FIG. 2. Restriction enzyme maps of a class 1 cloned M-MuLV DNA insert (Upper) and the class 4 insert (Lower). Regions designated A-E identify the location of sequences presented in Fig. 3. The arrows indicate the region of DNA that is inverted in the class 4 insert relative to the class 1 insert.

ternal inversion of 2.4 kb of the M-MuLV DNA sequence. Each end of the inversion occurred at or near the edge of a LTR. This clone is designated the "inversion clone."

DNA Sequence at the LTR Ends. The exact right end of the LTR (as oriented in Fig. 2) is the site at which reverse transcription of viral RNA initiates (33). The sequence at this site (3' . . . A-G-A-A-A-G-T-A-A 5') is known because the sequence of the initial reverse transcript (the start of "minus strong stop DNA") has been determined (34). The exact sequence of the left end of the LTR is not unambiguously known. However, we have determined the sequence of the joint between the two LTRs in a clone of 8.8-kb DNA (region E, Fig. 2; Fig. 3E), and from this sequence it is possible to infer the sequence at the left end of a LTR. The joint contains the exact sequence known to be at the right end of the LTR (3' . . . A-G-A-A-A-G-T-A-A 5'). Abutting this sequence is one that for 13 bases is an exact inverted repeat of the right end (underlined in Fig. 3E). We interpret this result to mean that, during circularization of an 8.8-kb linear DNA, the left end of one LTR and the right end of another have become ligated without loss of sequence. Whereas this interpretation is unambiguous for the right end, it can only be inferred for the left end because we lack knowledge of the left end's sequence. The existence of a palindrome at the LTR junction plus the presence of an unaltered right end sequence make this inference strong, and the following data will be interpreted in light of this assumption.

DNA Sequence at the Inversion Sites. The approximate locations of the inversion end points in the inversion clone are evident from the mapped restriction sites (Fig. 2). The nucleotide sequence of these regions from a standard 8.8-kb clone are presented in Fig. 3 A and C.

To place the exact sites of inversion and to examine its consequences, the sequences at either end of the inverted segment were determined (Fig. 2, regions B and D; Fig. 3 B and D). At the left end, region B, the LTR right end sequence was evident, but it lacked the final two A-T base pairs. This sequence was followed by one that matched a sequence in region C, starting with the four base pairs $\begin{smallmatrix} C-T-G-A \\ C-A-C-T \end{smallmatrix}$. At the other end of the inverted segment, region D, the left end of an LTR sequence was evident, but again it lacked two A-T base pairs. Where the LTR

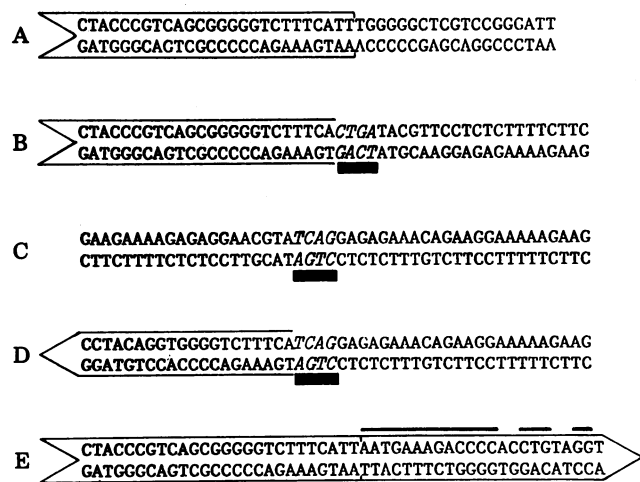


FIG. 3. Nucleotide sequence of relevant regions from a class 1 cloned M-MuLV DNA insert and the class 4 insert. Sequences A-E are obtained from regions of DNA identified by the same letter in Fig. 2. An enclosed sequence is from a LTR. Directionality of the enclosed sequences is from left to right, as oriented in Fig. 2. Heavy underlining and italics identify the four base pairs that are repeated at the inversion end points. Base pairs which form a palindrome around the LTR junction in E are identified with underlining or overlining.

sequence ended, the sequence from region C began but included the same four base pairs (albeit inverted) that were found at the other inversion site.

We conclude that during the inversion process two base pairs were lost from each end of the LTR sequence involved in the inversion and furthermore four bases at the inversion site were reiterated and appeared at both ends of the inverted segment.

DISCUSSION

Closed circular M-MuLV DNA has been extracted from recently infected cells and cloned in a λ vector. Four different classes of insert have been detected. Two of the classes represent standard 8.2- or 8.8-kb M-MuLV DNA containing, respectively, one or two LTRs. A third class contained inserts with deletions of various lengths, whereas a fourth class contained an insert with an inversion. One of four of the clones with 8.2-kb inserts and two of two of the 8.8-kb cloned inserts were infectious when excised with *Hind*III, polymerized by ligation, and transfected into NIH/3T3 cells (unpublished results). Thus, these cloned inserts contain all of the necessary functions for viral replication, virion maturation, and infection. It is not known what defect exists in the uninfected cloned DNA inserts.

Structure of the Inverted Clone. To interpret the structure of the inversion clone it was necessary to know the exact ends of the LTR sequence. This sequence was inferred from the sequence at the joint of two LTRs in an 8.8-kb cloned DNA. The sequence abutting the known sequence from the right LTR end (34) is an exact inverted repeat of the right end for 13 bases, with 18 of the first 22 bases appearing in a palindromic (inverted repeat) structure. We have interpreted these data to mean that the left and right ends of a linear 8.8-kb DNA are blunt-end ligated to form the 8.8-kb circular DNA, thus generating the palindrome. It could be that the left end has more (or even less) sequence than is evident in the circular form, but until such sequence is shown, we will assume that the left end sequence is an inverted repeat of the right end sequence.

The inversion clone contains two LTR sequences in inverted orientation separated by 2400 bases of inverted M-MuLV DNA (Fig. 2). Thus, as a consequence of the inversion of unique M-MuLV sequences, one end of each LTR abuts a novel sequence. The DNA sequence at each of these joints has been determined (Fig. 3 B and D), and the same four-base sequence was found at each end. Furthermore, two bases were absent from the LTR sequence at these joints.

Models of Integration. We interpret these sequence results as supporting the relevance of bacterial DNA transposition mechanisms (16, 17) to the integration of M-MuLV DNA into host cell DNA. In the present case, "host cell" is the M-MuLV DNA itself and the result is integrative inversion. Rather than presenting models to explain the integrative inversion, we have chosen to show in Fig. 4 two models of the mechanism for integration of M-MuLV DNA into host cell DNA because it represents the presumably more important circumstance and is simpler to depict.

The model for retrovirus integration that most closely approximates the Shapiro model (17) for transposition is model I of Fig. 4. This model assumes that the integrating form of M-MuLV DNA is an 8.2-kb circle with one copy of the LTR sequence. Integration would occur as a consequence of staggered cuts in the host and viral DNA, joining of ends, and replication of reiterated sequence. Specifically, an enzyme must recognize the ends of the LTR and make single-stranded nicks on opposite strands two bases into the LTR sequence (cuts 3 and 4). An enzyme must also cut the host cell DNA (presumably at a random sequence or one of low specificity), leaving single-

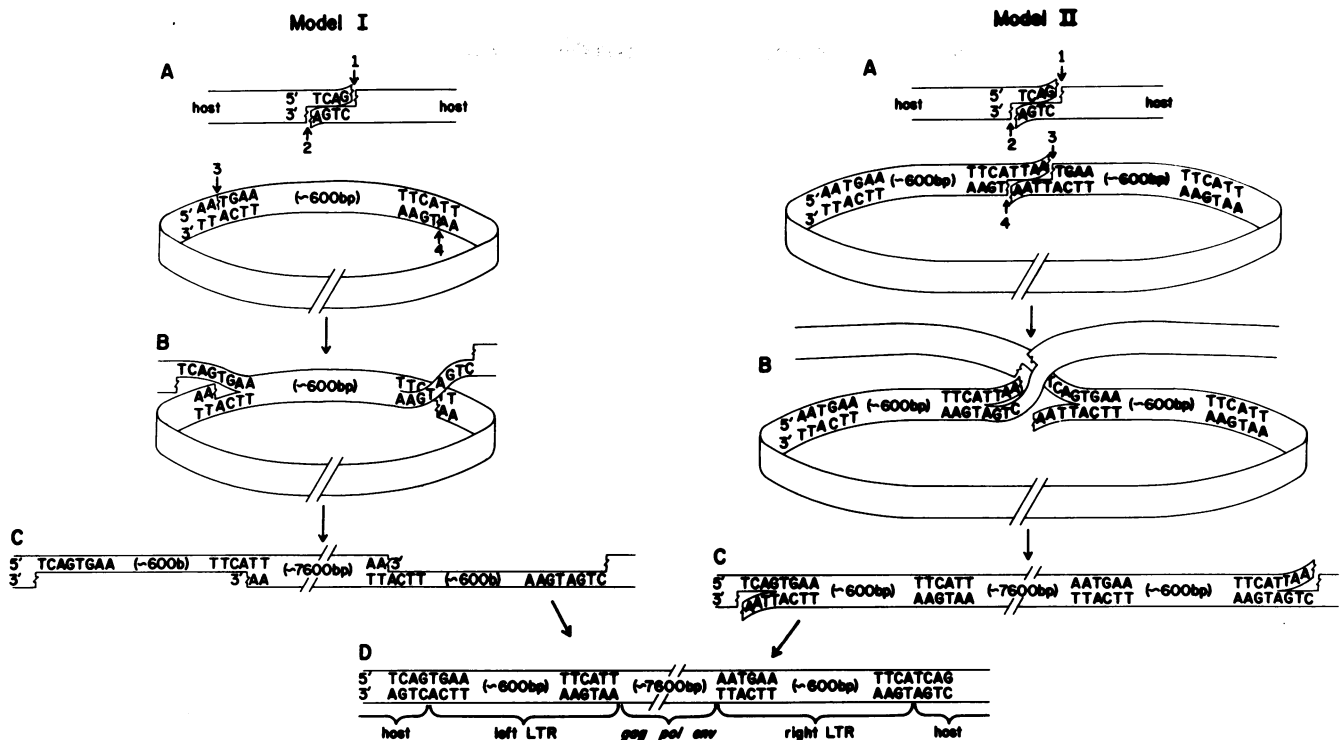


FIG. 4. Two models of retrovirus integration into host cell DNA. *Model I.* A circular M-MuLV DNA with one LTR is the integrating form. (A) Host DNA is nicked at sites 1 and 2 (the four-base-pair sequence inside of the nicks is the target site for retrovirus integration found in the inversion clone but is presented here for illustrative purposes and is not meant to imply sequence specificity of the integration site). Circular M-MuLV DNA containing a single LTR is nicked at sites 3 and 4. (B) The 3' ends of nicks 1 and 2 ligate to the 5' ends of nicks 3 and 4, respectively (the nicks could have occurred so that 5' ends on host DNA and 3' ends on viral DNA were used). (C) The strands separate through the 600-base-pair LTR. (D) The resultant gaps are filled, producing the integrated form of M-MuLV proviral DNA.

Model II. A circular M-MuLV DNA molecule containing two tandem LTRs is the preintegrative form. (A) Host DNA is nicked as in model I, whereas viral DNA is nicked at sites 3 and 4 so that a four-base-pair staggered cut (3' overhang) is produced, surrounding the LTR junction. (B) The 3' ends of nicks 1 and 2 then ligate to the 5' ends of nicks 3 and 4, respectively, forming a four-base-pair mismatch shown uncoiled in (C). Upon repair of the mismatch, the same end product is formed as for model I.

stranded nicks on opposite strands four bases apart (cuts 1 and 2). One enzyme must leave 5'-end overlapping sequences and the other must leave 3'-end overlapping sequences. By joining the overlapping sequences to each other forming 3'-5' joints, an integration intermediate structure can be formed (Fig. 4, model I B). If the base pairs holding together the LTR are then separated, a linear structure (Fig. 4, model I C) is formed. (It is unlikely that this actually exists as an intermediate, but it is shown here because it demonstrates graphically that integration has taken place). After joining, the two free 3' ends can initiate replication of the LTR and the four-base repeats. The resulting double-stranded DNA (Fig. 4D) will have the viral genes flanked by LTRs that lack two bases; this integrated viral DNA will be flanked by four-base direct repeats of host cell DNA.

An alternative model, that yields the same integrated structure, can be derived by postulating that an 8.8-kb circular viral DNA with two tandem LTRs is the integrating form (Fig. 4, model II). Although host DNA is cut in the same way as in model I, cuts 3 and 4 in the viral DNA would have to be two bases to either side of the joining point of the two LTRs. The linkage to host DNA would thus leave four unpaired bases at either end (Fig. 4, model II B), and the final replication step would be preceded by removal of the mismatched sequence. This model has the attraction that the site of nicking on the viral DNA is unique to the 8.8-kb circular form and is not present following integration.

Other models can be proposed based on these principles. For instance, *linear* DNA with terminal LTRs could integrate just as shown in Fig. 4, model II, except that the integrating viral

DNA would have free ends. Whereas the circular form has these ends joined and thus places all of the needed DNA ends close to one another, a linear form could be imagined to be held in a similar configuration by a protein.

Inversion Did Not Occur in *E. coli*. By interpreting the integrative inversion events as an effective integration of viral DNA into chromosomal DNA, we are making the tacit assumption that the event occurred in the newly infected mouse cell, not while the recombinant λ phage was growing in its bacterial host. This assumption seems valid for four reasons. First, it is unlikely that bacteria could carry out such a precise and unique process on a DNA that evolved in mammalian host cells. Second, in an accompanying paper, Dhar *et al.* (35) report evidence of a four-base repeat surrounding integrated Moloney murine sarcoma virus DNA. Third, we have passed unarranged clones containing 8.8-kb M-MuLV DNA in λ phage through multiple cycles, plaque-purified them, and found no inversion (or deletion) structures in 32 clones examined. Finally, we can detect the presence of the inversion clone (as well as deletion clones) by filter hybridization analysis (36) of phage DNA that has undergone only a single cycle of plaque purification.

Analogy to Other Systems. DNA structures like those of retrovirus provirus have been reported in other systems. In bacteria, complex transposable elements are flanked by inverted or directly repeated IS (insertion sequence) elements that have within themselves terminal inverted repeats (reviewed in ref. 37). When they transpose, a small number of base pairs at the recipient site are reiterated at either end of the transposon (38,

39). In several instances, bacterial IS sequences surround antibiotic resistance genes that are carried with the elements during transposition in a manner analogous to the LTRs carrying the M-MuLV genes (37). In one recent study (40), a plasmid with a single IS sequence was found to have integrated into another plasmid, generating an insert flanked by IS sequences; this structure is exactly that predicted by model I in Fig. 4 and strengthens the analogy of bacterial and retrovirus systems.

In *Drosophila*, a number of different moderately repeated and transposable DNA segments—called “copia”, 412 and 297—have been characterized that represent a unique sequence flanked by a direct repeat sequence (41). Determination of the nucleotide sequence of the terminal regions from several of these elements has demonstrated that transposition of these elements creates a five-base-pair duplication at the site of insertion. (P. Dunsmuir, W. J. Brorein, M. Simon, and G. Rubin, personal communication).

Can an Integrated Provirus Transpose? The analogy to transposons raises the question of whether an integrated provirus of a retrovirus can transpose. The answer to this question is not known; no case of transposition has been reported but it might be difficult to recognize without a serious search. We find it significant that, when M-MuLV DNA integrates, two bases of the LTR are lost. If the two bases are crucial to recognition, this may protect the integrated DNA from nicking by the enzyme involved in the integration and might therefore inhibit transposition. As noted above, model II (Fig. 4) naturally incorporates protection against nicking after integration.

Site for Integration. A study of cells infected by M-MuLV has shown multiple sites of integration (9). Thus, the four-base staggered nicks presumably can be made at many sites in the host cell genome. Comparing our four-base sequence with those of Dhar *et al.* (35), we find no apparent common features. More regularity may become apparent when more site sequences have been determined.

This work was supported by Grants CA-14051 and CA-26717 from the National Cancer Institute and Contract N01-CP-53562 from the Division of Cancer Cause and Prevention, National Cancer Institute. S.G. is a Postdoctoral Fellow of the Jane Coffin Childs Memorial Fund for Medical Research. S.W.M. is a Postdoctoral Fellow of the American Cancer Society, Massachusetts Division. D.B. is an American Cancer Society Research Professor.

1. Bishop, J. M. (1978) *Annu. Rev. Biochem.* **47**, 35–88.
2. Shank, P. R., Hughes, S. H., Kung, H., Majors, J. E., Quintrell, N., Guntaka, R. V., Bishop, J. M. & Varmus, H. E. (1978) *Cell* **15**, 1383–1395.
3. Gilboa, E., Goff, S., Shields, A., Yoshimura, F., Mitra, S. & Baltimore, D. (1979) *Cell* **16**, 863–874.
4. Shank, P. R., Cohen, J. C., Varmus, H. E., Yamamoto, K. R. & Ringold, G. M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2112–2116.
5. Shank, P. R. & Varmus, H. E. (1978) *J. Virol.* **25**, 104–114.
6. Yoshimura, F. K. & Weinberg, R. A. (1979) *Cell* **16**, 323–332.
7. Jolicœur, P. & Rassart, E. (1980) *J. Virol.* **33**, 183–195.
8. Keshet, E. & Temin, H. M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3372–3376.
9. Steffen, D. & Weinberg, R. A. (1978) *Cell* **15**, 1003–1010.
10. Cohen, J. C., Shank, P. R., Morris, V. L., Cardiff, R. & Varmus, H. E. (1979) *Cell* **16**, 333–345.
11. Hughes, S. H., Shank, P. R., Spector, D. H., Kung, H., Bishop, J. M., Varmus, H. E., Vogt, P. K. & Breitman, M. L. (1978) *Cell* **15**, 1397–1410.
12. Sabran, T. L., Hsu, T. E., Yeater, C., Kaji, A., Mason, W. S. & Taylor, J. M. (1979) *J. Virol.* **29**, 170–178.
13. Hager, G. L., Chang, E. H., Chan, H. W., Garon, C. F., Israel, M. A., Martin, M. A., Scolnick, E. M. & Lowy, D. R. (1979) *J. Virol.* **31**, 795–809.
14. Tronick, S. R., Robbins, K. C., Canaani, E., Devore, S. G., Andersen, P. R. & Aaronson, S. A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 6314–6318.
15. Oliff, A. I., Hager, G. L., Chang, E. H., Scolnick, E. M., Chan, H. W. & Lowy, D. R. (1980) *J. Virol.* **33**, 475–486.
16. Grindley, N. D. & Sherratt, D. (1979) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 1257–1261.
17. Shapiro, J. A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1933–1937.
18. Fan, H. & Paskind, M. (1974) *J. Virol.* **14**, 421–429.
19. Shields, A., Witte, O. N., Rothenberg, E. & Baltimore, D. (1978) *Cell* **14**, 601–609.
20. Hirt, B. (1967) *J. Mol. Biol.* **26**, 365–371.
21. Radloff, R., Bauer, W. & Vinograd, J. (1967) *Proc. Natl. Acad. Sci. USA* **57**, 1514–1521.
22. Goff, S. (1978) Dissertation (Stanford Univ., Stanford, CA).
23. Blattner, F. R., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Richards, J. E., Slighton, J. L., Tucker, P. W. & Smithies, O. (1978) *Science* **202**, 1279–1284.
24. Sternberg, N., Tiemeier, D. & Enquist, L. (1977) *Gene* **1**, 255–280.
25. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–181.
26. Enquist, L., Tiemeier, D., Leder, P., Weisberg, R. & Sternberg, N. (1976) *Nature (London)* **259**, 596–598.
27. Cameron, J. (1976) Dissertation (Stanford Univ., Stanford, CA).
28. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
29. Maat, J. & Smith, A. J. H. (1978) *Nucleic Acids Res.* **5**, 4537–4545.
30. Curtiss, R., III, Inoue, M., Pereira, D., Shu, J. C., Alexander, L. & Rock, L. (1977) *Miami Winter Symp.* **13**, 99–111.
31. Enea, V., Vovis, G. F. & Zinder, N. D. (1975) *J. Mol. Biol.* **96**, 495–509.
32. Katz, L., Kingsbury, D. T. & Helinski, D. R. (1973) *J. Bacteriol.* **114**, 577–585.
33. Gilboa, E., Mitra, S. W., Goff, S. & Baltimore, D. (1979) *Cell* **18**, 93–100.
34. Haseltine, W. A., Kleid, D. C., Panet, A., Rothenberg, E. & Baltimore, D. (1976) *J. Mol. Biol.* **106**, 109–131.
35. Dhar, R., McClements, W. L., Enquist, L. W. & Vande Woude, G. F. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3937–3941.
36. Southern, J. (1975) *J. Mol. Biol.* **38**, 503–517.
37. Kleckner, N. (1977) *Cell* **11**, 11–23.
38. Reed, R. R., Young, R. A., Steitz, J. A., Grindley, N. D. F. & Guyer, M. S. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4882–4886.
39. Calos, M. P., Johnsrud, L. & Miller, J. H. (1978) *Cell* **13**, 411–418.
40. Ohtsubo, E., Zenilman, M. & Ohtsubo, H. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 750–754.
41. Potter, S. S., Brorein, W. J., Dunsmuir, P. & Rubin, G. M. (1979) *Cell* **17**, 415–427.