JOURNAL OF VIROLOGY, Dec. 1985, p. 779–789
0022-538X/85/120779-11$02.00/0

Vol. 56, No. 3

A New Pathway in the Generation of Defective Retrovirus DNA

JOHN C. OLSEN AND RONALD SWANSTROM*

Department of Biochemistry and Lineberger Cancer Research Center, University of North Carolina, Chapel Hill, North Carolina 27514

Received 23 May 1985/Accepted 30 August 1985

We used a retrovirus shuttle vector to make molecular clones of circular viral DNA from infected cells. One-third of the molecules examined had deletions that started within or near the U5 domain of the long terminal repeat (LTR) region and extended a variable distance toward the *gag* gene. We present evidence that some of these deletions arose by cleavage of a single LTR unit, in contrast to the cleavage of tandem LTR units associated with the integration reaction. These results suggest that in the formation of defective circular DNA, the U5 domain can be recognized and cleaved in the absence of an adjacent U3 domain. The cleavage of isolated U5 domains may represent an important mechanism responsible for the generation of certain forms of both defective circular DNA and defective integrated DNA.

Retroviruses go through an unusual replication cycle in which their single-stranded RNA genome is converted into double-stranded DNA in the infected cell. The synthesis of viral DNA is carried out in the cytoplasm by a virion-associated DNA polymerase, called reverse transcriptase. Viral DNA migrates to the nucleus, where it is first circularized and then integrated into the host chromosome to serve as template for the synthesis of new viral RNA (Fig. 1) (57).

The initial DNA product is a linear duplex molecule with a direct repeat at the ends that is generated during DNA synthesis (8, 15, 41). This repeat, called the long terminal repeat (LTR), is composed of sequence domains (U5 and U3) that are present at the ends of the RNA. In the nucleus some of the linear DNA is converted into several different forms of circular DNA (43), predominantly circles with either two copies of the LTR unit in tandem or circles with a single copy of the LTR unit (40, 41, 62). Panganiban and Temin (30) recently showed that the unique sequence domain formed at the circle junction site by joining the LTR units in tandem is active in the integration reaction. The ends of each LTR unit form a short, imperfect inverted repeat (56) representing a sequence domain that is required for integration (29). Only when the LTR units are joined in tandem in a circle are the sequences that form the inverted repeat juxtaposed. It is likely that the placement of these sequences next to each other at the circle junction site represents at least part of the signal that specifies the site of cleavage to open the circle during the integration reaction.

Covalently closed circles with two copies of the LTR unit represent a small minority of the retroviral circular DNA molecules found in infected cells, presumably because of rapid integration (30). Circles with one copy of the LTR unit predominate, possibly arising by homologous recombination between the LTR units at the ends of linear DNA (40, 41, 62). Even greater heterogeneity in the structure of circular DNA is detected when individual molecular clones are examined, revealing an array of deletions and inversions. The inversions appear to be the result of integration of circular DNA into itself, since the inverted region is bounded by LTR units which are each missing two terminal base pairs and there is a short direct repeat flanking the LTR units (44, 45, 51, 55), all hallmarks of integrated DNA (56). More common than the inversions are deletions in which one end of the deletion is within or near the LTR region (17, 18, 34, 45, 47); however, the relationship of these molecules to the integration reaction is unknown. It seems likely that insight into the process of defective circular DNA formation will help define the types of reactions that occur both in the generation of defective DNA and in the integration of viral DNA.

In an effort to understand the origin of defective circular viral DNAs we have used a retrovirus shuttle vector to make molecular clones of circular DNA molecules formed in vivo. One-third of the circular viral DNA molecules cloned had deletions associated with the LTR region. These deletions fell into two classes: small deletions at the circle junction site between tandem copies of the LTR unit and larger deletions with one endpoint at a variable location with respect to the U5 boundary of a single LTR unit. In all cases, the deletions started within or near an LTR unit and extended toward the *gag* gene. We interpret these results as evidence that circles with either one or two copies of the LTR unit can be cleaved within the LTR region as one step leading to the formation of deletions within circular DNA.

MATERIALS AND METHODS

Cell culture. Fertile chicken eggs (C/O, virus negative) were obtained from H and N Farms, Redmond, Wash. Primary cultures of chicken embryo fibroblasts (CEFs) were made from 12-day-old embryos. QT6 cells, from a continuous quail cell line (28), were obtained from S. Ortiz and H. E. Varmus. All cells were maintained at 37°C in medium 199 (GIBCO Laboratories, Grand Island, N.Y.) supplemented with 10% tryptose phosphate broth, 0.22% sodium bicarbonate, 4% fetal calf serum (GIBCO), and 1% heat-inactivated chicken serum (GIBCO). The medium for QT6 cells also contained 1% dimethyl sulfoxide.

Plasmids and bacteria. pAV-V1, an infectious plasmid clone of the Schmidt-Ruppin A (SR-A) strain of Rous sarcoma virus (RSV), was constructed from the SRA-2 clone (5) by P. Luciw. pRAV-10R, an infectious plasmid clone of the RAV-1 genome, was obtained from L. Sealy (39). To construct pOJ1, the 2.3-kilobase-pair (kbp) *Xba*I fragment spanning most of the *pol* and part of the *env* genes in pAV-V1 was removed by complete digestion with *Xba*I. The ends were filled in by a repair reaction with the large fragment (Klenow) of DNA polymerase I (New England BioLabs, Inc., Beverly, Mass.). The 2.3-kbp *Eco*RI-*Pvu*II

* Corresponding author.

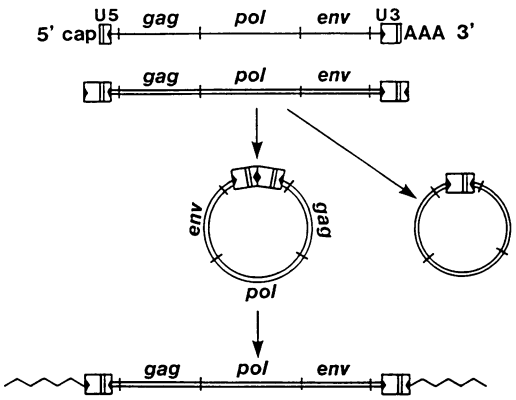

FIG. 1. Summary of retrovirus DNA synthesis and integration. Details are described in the text. At the top of the figure is a map of the retrovirus RNA genome. The *gag*, *pol*, and *env* genes are required for viral replication. The predominant forms of viral DNA are shown below the viral RNA. The large open box at each end of linear viral DNA represents an LTR unit. The arrowheads at the ends of each LTR unit represent the short inverted repeats at the termini of the LTR (15 bp for RSV). The wavy line represents flanking host cell chromosomal DNA.

fragment (after repair of the *Eco*RI end and phosphatase treatment) of pBR322 was joined to *Xba*I-cleaved pAV-V1 by blunt-end ligation with T4 DNA ligase (New England BioLabs). This fragment of pBR322 contains the ampicillin resistance (Amp^r) gene and ColE1 origin of replication. All plasmids were propagated in either HB101 or DH1, both of which are *recA*$^-$ strains of *Escherichia coli*. The genotype of strain DH1 was confirmed for the following markers: *endA1*, *hsdR17*, *thi-1*, and *gyrA96*.

DNA transfection and assays for virus production. Plasmid DNA was introduced into CEF cells by the calcium phosphate precipitate technique of Graham and Van der Eb (10) as described by DeLorbe et al. (5). Approximately 10 μg of plasmid DNA was added to a culture of 2×10^6 cells at 50% confluency. No carrier DNA was used in the transfection. Reverse transcriptase assays of media from infected cells were performed essentially as described previously (53) in 50-μl reaction mixtures containing 10 mM Tris hydrochloride (pH 7.4), 8 mM $MgCl_2$, 10 mM dithiothreitol, 0.5% (vol/vol) Triton X-100 (Research Products International Corp.), 2.4 μg of poly(A), 0.6 μg of oligo$(dT)_{12-18}$ (Collaborative Research, Inc., Waltham, Mass.), 5 μCi of [^{3}H]TTP (64 Ci/mmol; Chemical and Radioisotope Div., ICN Pharmaceuticals Inc., Irvine, Calif.), and 30 μl of culture medium (clarified at $2,000 \times g$ for 7 min). Reaction mixtures were incubated at 37°C for 2 h, and the incorporation of [^{3}H]TMP into DNA was assayed by the DEAE filter-binding technique (24) and liquid scintillation spectrophotometry.

For analysis of virion RNA, clarified medium (30 ml) was centrifuged at $100,000 \times g$ for 75 min at 4°C to pellet virus particles. The particles were suspended in the same volume of fresh medium without serum and pelleted. Twice-pelleted virions were suspended in 200 μl of a solution containing 10 mM Tris hydrochloride (pH 7.4), 100 mM NaCl, 5 mM EDTA, 0.5% sodium dodecyl sulfate, 100 μg of yeast carrier RNA per ml, and 50 μg of proteinase K (Calbiochem-Behring, La Jolla, Calif.) per ml and then incubated for 20 min at 37°C. The solution containing viral RNA was extracted successively with equal volumes of phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform-isoamyl alcohol (24:1). The aqueous phase was made 0.4 M NaCl, and the RNA was precipitated by adding 3 volumes of ethanol. The precipitated RNA was collected by centrifugation and subjected to dot-blot hybridization analysis as described by Thomas (54), with nick-translated ^{32}P-pBR322 or ^{32}P-pAV-V1 DNA probes (35).

Rescue of unintegrated DNA into *E. coli*. Approximately 2×10^6 QT6 cells were infected with 2 ml of concentrated virus. The virus stock used in these experiments was from culture media pooled 11–24 days after transfection (see Fig. 3). The virus was pelleted from the clarified media as described above, suspended in 1/10 the original volume of media with serum, and frozen in small aliquots at −70°C prior to infection. At 24 h after infection, extrachromosomal DNA was separated from chromosomal DNA by the method of Hirt (14) and then extracted twice with phenol-chloroform-isoamyl alcohol (25:24:1). Nucleic acid in the aqueous phase was precipitated by the addition of 1.5 volumes of isopropanol. After 16 h at −20°C, the precipitate was collected by centrifugation ($8,000 \times g$ for 30 min), washed once with ice-cold 70% ethanol–40 mM NaCl, centrifuged as above for 10 min, and dried. The samples were dissolved in water and extracted by the acid-phenol method of Zasloff et al. (63) as described by DeLorbe et al. (5). This procedure removed most of the contaminating high-molecular-weight chromosomal DNA and improved the transformation efficiency. The partially purified circular DNA was dissolved in 10 μl of a solution containing 10 mM Tris hydrochloride (pH 8.0) and 0.1 mM EDTA and used to transform the DH1 strain of *E. coli* as described by Hanahan (13). Bacterial transformants were selected by growth on agar plates containing 50 μg of ampicillin per ml (Sigma Chemical Co., St. Louis, Mo.).

Analysis of rescued plasmid DNA. Plasmid DNAs from individual colonies were isolated from overnight cultures by the alkaline lysis method (24). Restriction endonucleases were from New England BioLabs or Bethesda Research Laboratories, Gaithersburg, Md., and used according to the recommendations of the manufacturers. Digested DNA was analyzed by electrophoresis in agarose (SeaKem ME; FMC Corp., Marine Colloids Div., Rockland, Maine). DNA fragments in the gel were transferred to nitrocellulose paper by the method of Southern (46) and hybridized with a ^{32}P-labeled probe specific for the 3′ end of the viral genome (50) or pBR322 labeled with ^{32}P by nick translation (35). Hybridization conditions were as described by Wahl et al. (60).

Nucleotide sequence analysis was done by the method of Maxam and Gilbert (25). DNA restriction fragments were end labeled with reverse transcriptase (Life Sciences, Inc.) in repair reactions containing the appropriate α-^{32}P-labeled deoxyribonucleoside triphosphates (3,000 Ci/mmol; New England Nuclear Corp., Boston, Mass.).

RESULTS

Construction of a retrovirus shuttle vector. We modified the genome of the avian retrovirus Rous sarcoma virus, Schmidt-Ruppin strain (SR-RSV), so that in its DNA form it would be capable of replication in bacteria. The starting material for the construction of this retrovirus shuttle vector was a plasmid (pAV-V1) containing the SRA-2 clone of RSV (5) that had been reorganized so that the viral transcription unit was intact. In this arrangement, the viral replicative genes were bounded by LTR units. We replaced the 2.3-kbp *Xba*I restriction endonuclease fragment from within the viral genome with a 2.3-kbp *Eco*RI-*Pvu*II fragment from the bacterial plasmid pBR322 (2) that contained the Amp^r gene and plasmid origin of replication (*ori*) (Fig. 2). The resulting

plasmid, pOJ1, retained *cis*-acting sequences required for replication as a retrovirus in avian cells (57) and contained sequences within the viral genome to permit propagation and selection in bacteria. Sequences required for expression of the viral subgenomic v-*src* mRNA, which encodes the viral transforming gene product pp60^{src}, were also present (48). Owing to the disruption of the viral *pol* and *env* genes, the viral genome in pOJ1 was defective for virus replication in the absence of complementation by *pol* and *env* gene products.

Introduction of the retrovirus shuttle vector into avian cells. Introduction of the parental SRA-2 clone of viral DNA into permissive avian cells by DNA-mediated transfection gives rise to a strongly transforming and replication-competent virus (5). Since pOJ1 was defective for replication, we introduced it into tertiary CEFs by cotransfection with a plasmid carrying the genome of an avian helper virus, RAV-1. The plasmids were cotransfected at a 1:1 molar ratio. RAV-1 contains intact *gag*, *pol*, and *env* genes necessary for replication but lacks sequences required for rapid morphological transformation. Virus production was monitored by the appearance of virion-associated reverse transcriptase in the culture medium. The peak of reverse transcriptase activity in the medium occurred 2 to 3 weeks after transfection (Fig. 3A). Concomitant with the rise in reverse transcriptase activity in the medium was the onset of morphological transformation; the cells had a transformed morphology indistinguishable from that observed after transfection with pAV-V1 (data not shown). This suggested that the defective pOJ1 sequences were replicated and that the infected cells were expressing the v-*src* gene product. Replication of sequences from pOJ1 was confirmed when RNA from virus particles was analyzed by the dot hybridization technique with radiolabeled DNA probes. RNA isolated from virus particles at the time of peak virus production hybridized with pBR322 sequences (Fig. 3B). Thus, we were able to prepare a virus stock that contained sequences present in the shuttle vector.

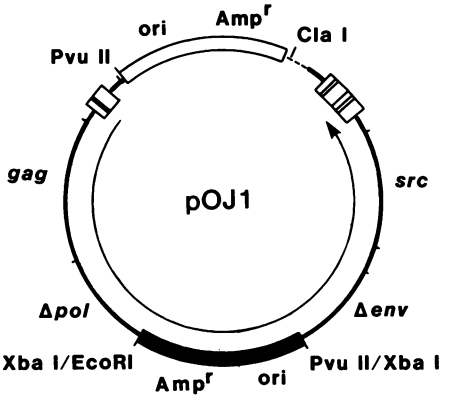

FIG. 2. Structure of the pOJ1 shuttle vector. Thin line, sequences derived from the SRA-2 clone of SR-RSV as placed in the plasmid pAV-V1. Solid box, the 2.3-kbp *Eco*RI-*Pvu*II Amp^r-*ori* fragment from pBR322. This fragment replaces a 2.3-kbp *Xba*I fragment containing part of the viral *pol* and *env* gene sequences. A single, complete LTR unit is present at the 5′ end of the RSV transcription unit. Tandem LTR units are present at the 3′ end. A small portion of RSV sequences are repeated outside the boundaries of the LTR region. About 260 bp of RSV sequences extend beyond the single LTR unit to the *Pvu*II site 31 bp upstream of the *src* stop codon. Also, 367 bp extend beyond the tandem LTR units towards *gag*. A fragment containing tandem LTR units was used in constructing pAV-V1; presumably only one of the LTR units that are in tandem is functional. Open box, 2.3-kbp *Eco*RI-*Pvu*II fragment containing the Amp^r-*ori* sequences from pBR322. Dashed line adjacent to the open box, *Eco*RI-*Bam*HI sequences derived from pBR322. The arrow shows the presumed transcription unit expected to be propagated as a retrovirus genome. The pBR322 sequences outside of the viral genome were used in the construction of pAV-V1, the progenitor of pOJ1.

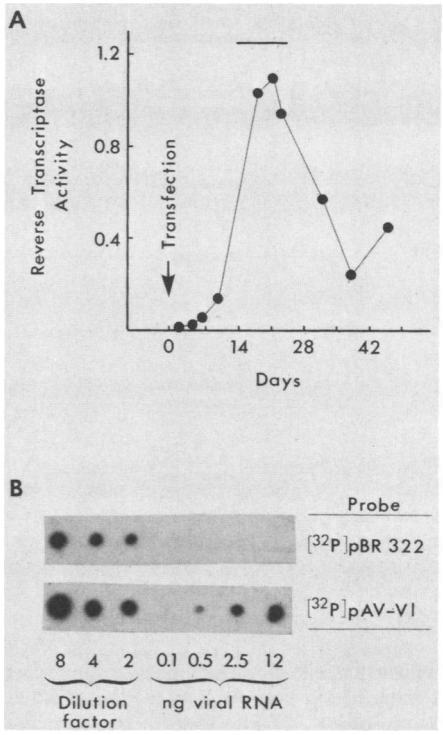

FIG. 3. Propagation of pOJ1 and RAV-1 genomes in CEFs. (A) After transfection, virus production was monitored by assaying for reverse transcriptase activity in the media. Reverse transcriptase activity is plotted relative to the activity obtained with media from CEFs chronically infected with SR-A RSV. On days of peak virus production (indicated by the bar) the media were pooled and stored at −70°C for subsequent analysis or infection. (B) Analysis of viral RNA by the dot-blot hybridization technique. RNA isolated from the equivalent of 8, 4, or 2 ml of pooled media was spotted on the left-hand portion of each filter. On the right-hand portion of each filter the indicated amounts of purified RSV 70S RNA were spotted. The radioactive DNA probes used were as indicated. Preparation of the filter and hybridization conditions are described in Materials and Methods.

Conversion of the shuttle vector from viral RNA to circular DNA and rescue in bacteria. The virus stock containing RAV-1 and shuttle vector sequences was used to infect QT6 cells, a continuous quail cell line (28). The QT6 cell line was chosen because of its susceptibility to infection by subgroup A avian retroviruses (including the RAV-1 helper virus) and because these cells accumulate circular viral DNA (12). At 24 h after infection, extra-chromosomal DNA was isolated by the Hirt fractionation procedure (14) followed by acid-phenol extraction to enrich for covalently closed circular DNA (63). This DNA was used to transform *E. coli* DH1 by the transformation procedure of Hanahan (13). Bacteria containing the shuttle vector sequences were selected by their resistance to ampicillin. In three separate experiments, we obtained 25 ampicillin-resistant bacterial colonies when

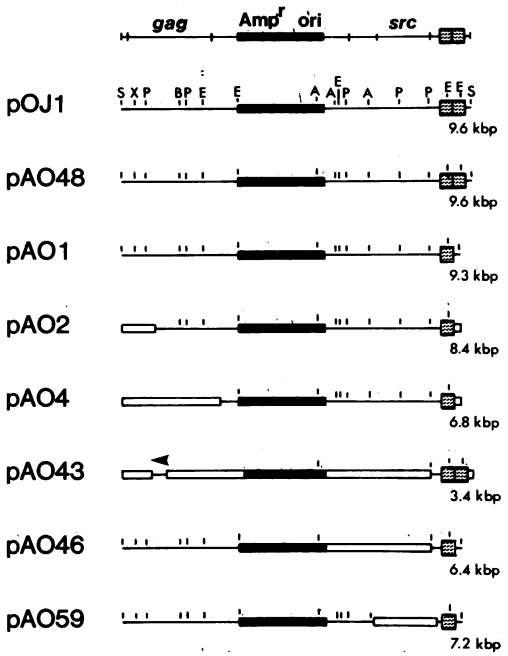

FIG. 4. Physical maps of selected rescued shuttle vector clones. Each DNA molecule is presented in the permuted form that is obtained by *Sac*I digestion. The map for the large *Sac*I fragment from the pOJ1 parent is shown for reference along with a genomic map (top line). The other maps were derived from individual plasmids rescued in the shuttle experiments. The thin line represents viral sequences. The solid box is the Amp^r-*ori* insert. Each hatched box represents a single LTR unit. Deleted sequences are indicated by an open box. The arrow above the thin line of pAO43 indicates the position of an inversion present in this isolate. The boundaries of all of the deletions shown in this figure were determined by nucleotide sequence analysis (Fig. 6), except for the deletion in pAO59 and the right-hand boundary of the inverted sequence in pAO43. Restriction endonuclease sites: S, *Sac*I; X, *Xho*I; P, *Pvu*II; B, *Bgl*II; E, *Eco*RI; A, *Acc*I. There were no *Cla*I cleavage sites in any of the molecules.

unintegrated DNA from about 10^7 infected QT6 cells was used to transform *E. coli*. Preparations of extrachromosomal DNA from uninfected cells did not give rise to ampicillin-resistant colonies.

Physical mapping of the rescued plasmids. To determine the structures of the rescued viral genomes, plasmid DNA was isolated from individual ampicillin-resistant bacterial colonies, and a physical map of restriction endonuclease sites was constructed by using *Sac*I, *Eco*RI, *Pvu*II, *Xho*I, *Bgl*II, *Cla*I, *Sph*I, and *Acc*I. Examples of the structural features of 25 rescued plasmids are shown in Fig. 4 and can be summarized as follows.

(i) All of the rescued plasmids contained sequences present in the shuttle vector plasmid used in the initial transfection; however, they differed from the parental plasmid in a way that is consistent with prior replication as a retroviral genome. Only sequences between the LTR units should be replicated as a retrovirus genome. With the exception of one class of recombinant molecules (see below), only the sequences in the parental plasmid within the LTR units that flanked the nearly intact viral genome were present in the rescued plasmids.

(ii) Six of the rescued plasmids had two copies of the LTR unit in tandem (e.g., pAO48) although several of these appeared to have deletions within the LTR region. The rest had one copy of the LTR unit (e.g., pAO1).

(iii) The sizes of the rescued plasmids ranged from 3.4 to 9.6 kbp; the latter is the size expected for recovery of a full-length genome. Eight of the rescued plasmids appeared to have the structure expected for an intact genome (e.g., pAO48). Seventeen of the rescued plasmids had deletions ranging in size from about 0.6 to 3.2 kbp. Four plasmids had more than one deletion.

(iv) Six of the rescued plasmids had deletions in the 5′ portion of the viral genome (e.g., pAO2 and pAO4). In each case, the 5′ boundary of the deletion mapped very near or within an LTR unit, whereas the 3′ boundary of the deletion extended for a variable distance (900 to 3,200 bp) into the *gag* or *pol* regions. Deletions extending significantly farther into the *pol* region would not be expected because of the requirement to maintain the Amp^r region of the vector for selection in bacteria.

(v) Ten of the rescued genomes had a similar deletion of about 3 kbp that included most of the *env* and *src* genes (e.g., pAO43 and pAO46).

(vi) Four of the rescued plasmids (e.g., pAO59) had various deletions that, by restriction endonuclease mapping, were restricted to part or all of the *src* gene. The deleted sequences of one isolate, pAO59, appeared to include all of the *src* sequences. This rescued viral genome may be analogous to transformation-defective RSV genomes previously described (58), or may be a recombinant with the RAV-1 helper virus.

A comparison of the physical maps of all 25 molecules revealed similarities in structure that suggested that a limited number of mechanisms were involved in generating the

TABLE 1. Summary of the structures of the rescued plasmids[a]

Rescued plasmid	Size (kbp)	Intact	Δ*gag*[b]	Δ*src*	3 kb Δ*env-src*	Other	Recombinant with RAV-1	1 LTR	2 LTR
pAO1	9.3	+						+	
pAO2	8.4		+					+	
pAO4	6.8		+					+	
pAO5	6.6				+				+[c,d]
pAO6	9.3	+						+	
pAO7	9.3	+						+	
pAO8	9.3	+						+	
pAO15	9.3	+						+	
pAO43	3.4		+		+				+[c]
pAO46	6.4				+			+	
pAO47	9.3	+						+	
pAO48	9.6	+							+[d]
pAO49	9.3	+						+	
pAO50	6.4				+			+	
pAO51	6.4				+			+	
pAO52	6.7				+				+
pAO53	5.7		+	+				+	
pAO54	6.4				+			+	
pAO55	6.7				+				+
pAO56	6.4				+			+	
pAO57	5.1		+	+				+	
pAO58	8.4					Δ*env*		+	
pAO59	7.2			+			?	+	
pAO60	6.7				+				+
pAO61	4.8		+	+			+	+	

[a] Percentage in each class: intact, 32; Δ*gag*, 24; Δ*src*, 16; Δ*env-src*, 40; other, 4; recombinant, 8; 1 LTR, 76; 2 LTR, 24.
[b] Δ refers to a deletion in the specified region of the genome.
[c] Tandem LTR units with an insertion at the circle junction site.
[d] Tandem LTR units with a deletion at the circle junction site.

deletions (Table 1). Nucleotide sequence analysis was carried out on the relevant regions of selected plasmids to determine the exact nature of the deletions.

Analysis of circle junction regions with tandem LTR units. We used the following criteria to distinguish between circle junction regions with one or two copies of the LTR unit: (i) cleavage within the LTR unit by either *Eco*RI or *Sph*I generates a 330-bp fragment for molecules with intact tandem copies of the LTR unit, and (ii) cleavage with *Pvu*II, which has sites that flank the LTR unit, gives fragments from this region of about 1,330 or 1,660 bp from plasmids with one or two intact copies of the LTR unit, respectively. On the basis of these criteria, 4 of 25 rescued plasmids tested had circle junctions with tandem copies of the LTR unit. Two other plasmids, pAO5 and pAO48, had physical maps consistent with the presence of deletions of about 140 and 10 bp, respectively, within tandem LTR units. The remaining 19 plasmids all had physical maps that were consistent with circles containing one copy of the LTR unit (14 plasmids) or indicated the presence of large deletions that extended up to or into a single copy of the LTR unit (5 plasmids).

We determined the nucleotide sequence at the LTR-LTR junction site for pAO5 and pAO48, the molecules with deletions in the tandem LTR units, and for pAO52 and pAO43, two of the isolates that appeared by restriction endonuclease mapping to have intact tandem LTR units. The results are summarized in Fig. 5. pAO52 had a circle junction region with two complete copies of the LTR unit in tandem, identical in sequence to the circle junction region of the SRA-2 parental clone (47). The sequence of pAO43 was identical in this region except that an extra A · T base pair was present precisely at the circle junction site. Sequence analysis of pAO5 showed a deletion of 141 bp that was located entirely within the U3 region adjacent to the circle junction site. In addition, an extra T · A base pair was present at the circle junction site (Fig. 5). pAO48 had an 11-bp deletion at the site of joining between the LTR units. An ambiguity in the position of the deletion arises because the dinucleotide T-T flanks the deleted sequence (Fig. 5). All 11 bp may have been lost from the U3 region or, alternatively, 2 of the 11 bp may have been lost instead from the U5 region.

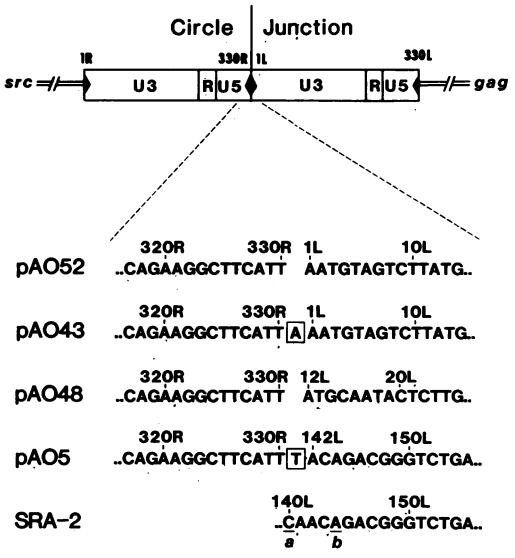

FIG. 5. Nucleotide sequence analysis at the circle junction region from selected rescued shuttle vectors. Nucleotide sequences are presented for the (+) strand. The numbering system used is that of Swanstrom et al. (47), with the separate LTR units designated L (left end of the genome next to *gag*) and R (right end of the genome next to v-*src*). The sequence of pAO52 at the circle junction site is identical to that of SRA-2 (47). The boxed nucleotides shown in the sequences of pAO43 and pAO5 represent insertions. The *a* and *b* below the sequence of SRA-2 denotes deletion endpoints for molecules described by Ju and Skalka (17) and Katz et al. (18), respectively. The arrowheads show the positions of the inverted repeats that bound each LTR unit.

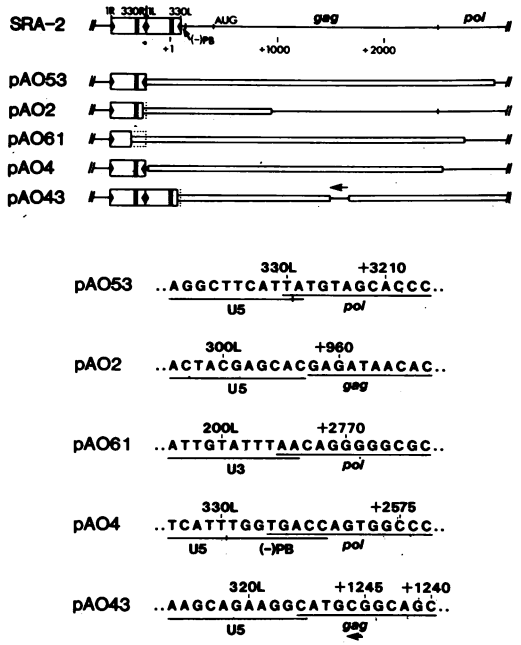

FIG. 6. Nucleotide sequence analysis at the sites of large deletions extending from the LTR region. The line drawing of SRA-2 is shown at the top for reference. (−)PB denotes the binding site for the $tRNA^{Trp}$ primer used to initiate the first (minus) strand of viral DNA. The position of the translation start codon of the *gag* gene is indicated by AUG. The locations of deleted sequences are shown by open boxes. The locations of deleted LTR sequences are emphasized by the dotted boxes. The arrowheads in the large boxes of the U3 and U5 domains of the LTR units denote the positions of the short, imperfect inverted repeat sequences present at the boundaries of each LTR unit. The arrow above the thin line of pAO43 indicates the position of the inversion present in this isolate (see text). Nucleotide sequences at the site of the deletion are presented for the plus strand. The numbering system used for the LTR sequences is that of Swanstrom et al. (47), with the LTR units distinguished by an L (left, next to *gag*) and R (right, next to v-*src*) designation. The numbering system used for non-LTR regions and sequence alignment is from Schwartz et al. (37) and Swanstrom et al. (49), where numbering starts from the 5′ end of the RNA. In most cases one or more common nucleotides were present at each boundary of the deletion. These common nucleotides are indicated by the overlapping lines below the sequence.

Large deletions extending from the LTR region. As indicated by the mapping experiments described above, a significant percentage (24%) of the rescued plasmids had large deletions in the 5′ end of the viral genome, with one end of the deletion associated with a single LTR unit. To understand the mechanism of this type of deletion formation, we determined the nucleotide sequence at the sites of deletion

from five selected plasmids. The results are shown in Fig. 6. Like the deletions associated with tandem LTR units, one end of each deletion was close to a U5 boundary and all the deletions extended toward the *gag* gene. However, both the positions of the deletion endpoints relative to the U5 boundary and the sizes of the deletions were much more variable than in the previous molecules. The deletions in pAO2 and pAO61 started within the LTR at 24 and 124 to 126 bp upstream of the U5 boundary, respectively (Fig. 6). (The uncertainty of the precise location of the deletion boundary in pAO61 is due to a common dinucleotide that flanks the deleted region.) The deletion in pAO53 started precisely at the U5 boundary or 1 bp into the $tRNA^{Trp}$ binding site [(−)PB] adjacent to the U5 domain. The uncertainty in this deletion boundary is extended 1 bp to the right if the deletion occurred between tandem LTR units. The deletion in pAO2 ended at position 957 from the RNA cap site, in the middle of the *gag* gene. The deletions of pAO61 and pAO53 extended well into the *pol* gene. Because of the large deletions, it is not clear whether these molecules arose as a result of deletion events associated with the circle junction site of tandem LTR units, like those described above for pAO5 and pAO48, or whether they were derived from circles with a single copy of the LTR unit.

Two other molecules, pAO4 and pAO43, shared the common features of a large deletion in the 5′ region of the viral genome and one end of the deletion at a variable point relative to the U5 boundary. However, for these two molecules the deletion events were clearly associated with a U5 domain that was not adjacent to a U3 domain in tandem LTR units. In pAO4 the deletion started 4 to 8 bp to the right of the U5 boundary within the adjacent domain of the $tRNA^{Trp}$ primer-binding site. In pAO43 the deletion started 6 to 7 nucleotides within the U5 domain, but this site was at the outer boundary of tandem copies of the LTR unit.

In several experiments we tested the possibility that the deletions had either occurred or were selected for in bacteria. In one experiment, 50 pg of pOJ1 was mixed with lysed QT6 cells. Circular DNA was then isolated by the same procedure used to purify DNA in the shuttle experiment. The DNA was used to transform *E. coli* DH1. One-hundred Amp^r colonies were picked and grown up to small cultures. Plasmid DNA was isolated from each culture and analyzed by agarose gel electrophoresis after digestion with *Eco*RI or *Sac*I. We found no variant fragment sizes in either the large fragments internal to the genome or the small fragments associated with the LTR region. Although this test does not rule out deletion formation in bacteria at some low level, it does exclude the possibility that the high frequency of deleted molecules that we observed occurred during the cloning step.

In a second control experiment we tested the possibility that molecules with deletions in the 5′ region of the genome were more efficiently cloned than molecules with these sequences intact. Such a circumstance would arise if there were a "poison" sequence present near the 5′ end of the RSV genome and would result in an overrepresentation of rescued molecules lacking these sequences. Majors and Varmus (23) reported such a sequence in the 5′ region of the mouse mammary tumor virus genome. To determine whether an analogous sequence was present in RSV DNA we reintroduced several rescued plasmids with and without 5′ deletions back into *E. coli* cells and determined their relative transformation efficiency. We found that all of the rescued plasmids tested, including pAO52, pAO46, pAO1, pAO2, and pAO4, were capable of transforming *E. coli* DH1 to ampicillin resistance with an efficiency comparable to that of pBR322 (13). There was no significant difference in transformation efficiency between plasmids containing or lacking 5′ sequences (data not shown). We conclude that the RSV genome does not contain a poison sequence in the 5′ region of the genome that would inhibit its cloning or replication as a plasmid. We cannot exclude, however, the possibility that unintegrated DNA isolated from quail cells contains some unknown secondary modification (e.g., methylation) that would inhibit its replication as a plasmid.

A specific deletion in *env* and *src*. We repeatedly rescued a class of molecules that had a 3-kbp deletion that mapped in the *env* and *src* regions. We determined the nucleotide sequence at the boundaries of this deletion for two rescued plasmids. Both pAO5 and pAO46 had the same sequence at the site of the deletion (not shown). One of the borders is precisely at the right-hand boundary of the Amp^r-*ori* insert, and the other is part of a *Pvu*II recognition site 31 bp upstream of the termination codon of the v-*src* gene. The two deleted molecules that were sequenced contained the *Pvu*II recognition sequence at the deletion site. All of the rescued plasmids that had this type of deletion, as indicated by restriction endonuclease mapping, also had a *Pvu*II site that mapped to this position. This same positioning of the Amp^r-*ori* sequences and v-*src* sequences occurs in pOJ1 outside of the retrovirus genomic transcription unit (Fig. 2). Thus, the most likely explanation for the appearance of this deletion at high frequency is homologous recombination between the two Amp^r-*ori* sequences present in the pOJ1 plasmid during the transfection step.

We detected only one of the products of the homologous recombination event as a rescued plasmid. We did not expect to see the other product, since it would have contained the part of the parental plasmid lacking sequences needed for packaging viral RNA (33). Thus, the other recombination product would have been lost during virus propagation.

Screen for recombination between shuttle vector and helper virus genomes. There were two points in the shuttle protocol where intermolecular recombination was likely to occur. Genetic recombination between avian retrovirus genomes is extraordinarily high; previous studies have indicated that 10 to 40% of the progeny genomes isolated after a mixed infection have undergone at least one recombination event (19, 59). In addition, intermolecular recombination between cotransfected DNA molecules also occurs at a high frequency (20, 26). We searched for the product of a recombination event by using restriction endonuclease site polymorphisms between the helper virus and the shuttle vector.

In an initial test we screened for the acquisition by the rescued shuttle vector of an *Xho*I restriction endonuclease site in the *pol* gene of the RAV-1 genome (see reference 16 for a restriction endonuclease map). Rescued plasmids were digested with both *Sac*I and *Xho*I and then subjected to electrophoresis through an agarose gel. Digestion of a nondeleted, nonrecombined rescued plasmid should yield two DNA species, of 9.2 and 0.4 kb. The appearance of a 2.0-kbp band with a corresponding decrease in the size of the 9.2-kbp band would be indicative of the acquisition of the new *Xho*I site. In a screen of 22 rescued plasmids that had maintained this region of the genome, we failed to find any cases in which this *Xho*I site was acquired from the RAV-1 helper virus (data not shown).

In a second test, we screened for the loss of an *Sph*I restriction endonuclease site present in the U3 region of the

LTR unit of the shuttle vector but absent in the U3 region of the RAV-1 helper virus genome (Fig. 7). Rescued plasmids were digested with *Xho*I and *Sph*I, and the products were fractionated through an agarose gel. The shuttle vector contained *Sph*I sites in the U3 region of the LTR unit and within the v-*src* gene (Fig. 7), whereas the RAV-1 helper virus genome lacked both of these *Sph*I sites. Both genomes had an *Xho*I site near the 5′ end of *gag*. Digestion of a nonrecombined rescued plasmid with both enzymes should yield low-molecular-weight DNA species of 454 and 759 bp. An additional low-molecular-weight band of 330 bp would be expected with plasmids containing tandem LTR units (Fig. 7). If recombination had resulted in a rescued shuttle vector containing sequences derived from a RAV-1 LTR unit, then all three low-molecular-weight bands would be missing. In 2 of the 24 rescued plasmids tested, the expected low-molecular-weight bands were absent, suggesting the acquisition of RAV-1 LTR sequences in these instances. In one case (pAO61) we subjected the U3 region to nucleotide sequence analysis and compared the resulting sequence with that of the U3 region of RAV-1 (C. Bova and R. Swanstrom, unpublished data) and to the U3 region of the parent SRA-2 sequence in pOJ1. Our results indicate that the U3 region was derived from RAV-1 (Fig. 7). One boundary of recombination was obscured by a deletion that removed all of U5 and *gag* and a portion of *pol* (Fig. 6). We did not determine the other boundary of recombination, although it extended at least 50 nucleotides beyond the U3 domain.

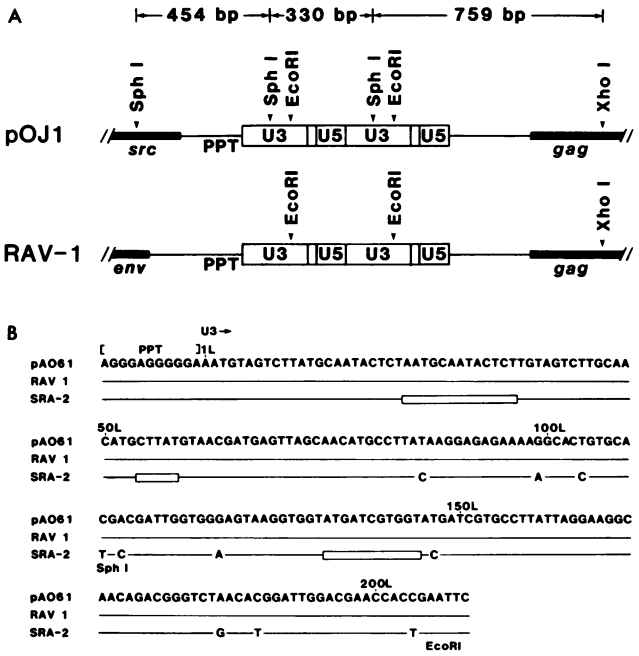

FIG. 7. Screen for recombination between shuttle vector and helper virus genomes. (A) A physical map of *Eco*RI and *Sph*I restriction endonuclease sites in the pOJ1 shuttle vector and the RAV-1 helper virus genome. PPT, Polypurine tract. (B) DNA sequence near the *Sph*I restriction endonuclease polymorphic site in the U3 region. The nucleotide sequence for pAO61 from the polypurine tract to the *Eco*RI site in the LTR sequence is shown in the top line. The RAV-1 sequence was obtained by C. Bova and R. Swanstrom (unpublished data). The pOJ1 sequence is that of the SRA-2 clone (47). Nucleotide gaps generated during the sequence alignment are indicated by open boxes. Nucleotide changes are indicated. The continuous line represents sequence identity.

DISCUSSION

We have used an avian retrovirus shuttle vector to clone circular retroviral DNA directly into bacteria. We examined in detail the structures of 25 viral genomes cloned as plasmids. Approximately one-third of the molecules had deletions associated with the LTR region. Many of the molecules with deletions were similar in structure to molecules described previously. However, several molecules had structures not observed previously in viral DNA. The structures of these molecules offer new insight into the origin of many forms of defective viral DNA.

Features of the shuttle vector system. The only requirement for cloning the vector in its viral genome form was the maintenance of the ampicillin resistance gene and the plasmid origin of replication. All circular forms of viral DNA cloned with equal efficiency regardless of the number of LTR units in the molecule or the presence or absence of specific regions of the viral genome. Furthermore, we could detect no rearrangements or deletions attributable to the cloning step. Thus, the cloned molecules were representative of the different forms of viral DNA that existed in vivo.

At present the shuttle vector system is rather inefficient, yielding approximately eight clones per 4×10^6 infected cells. In these experiments, no attempt was made to select for the presence of the shuttle vector during virus growth. One approach that has been used by others to increase the number of copies of viral DNA is amplification of circular DNA by using a simian virus 40 or polyoma virus origin of replication (1, 3). However, amplification may give rise to molecules that are not necessarily representative of the retroviral DNA population.

Two types of recombination events were detected in the cloned shuttle vector molecules. The products of apparent intramolecular recombination were observed in approximately 40% of the rescued plasmids. These molecules resulted from a recombination event between the two copies of Amp^r-*ori* sequences that were present in the original plasmid (Fig. 2). The intramolecular recombination event must have occurred during the DNA transfection step, since one of the regions involved probably could not be propagated as a viral genome to permit recombination at a later point. There were five examples of internal deletions within the *env* and *src* regions of the viral genome that were not characterized in detail.

The products of intermolecular recombination were also detected in the shuttle vector plasmids. Using a restriction endonuclease site polymorphism, we detected a recombinant between the shuttle vector and the RAV-1 helper virus used to complement the vector growth in avian cells. There are two points in the shuttle procedure at which recombination is expected to occur at a high frequency. During DNA transfection recombination is enhanced (20, 26), and during virus propagation retroviral genomes recombine at a high frequency (19, 59). At present we cannot determine at which point intermolecular recombination occurs.

Two classes of deletions in circular DNA involving the LTR region. Deletions associated with the LTR region fell into two classes. Molecules in the first class had small deletions within tandem copies of the LTR unit. In the two molecules we examined (pAO5 and pAO48; Fig. 5), one end of the deletion was either at the circle junction site or 2 bp to the left in U5 (there is a 2-bp ambiguity in the placement of the deletion in pAO48). The deletions extended a variable but short distance into the adjacent U3 domain. These molecules are similar to five molecules described previously, in which

one end of the deletion was either at or within 2 bp of the circle junction site with the deletions extending between 61 and 157 bp into U3 (17, 18, 47).

The second class of molecules had larger deletions (0.9 to 3 kbp) and had one endpoint of the deletion within or near a single LTR unit or at the outer boundary of tandem LTR units (pAO2, pAO4, pAO43, pAO53, pAO57, and pAO61; Fig. 6 and Table 1). All six of the molecules in this class had deletions that started in the LTR region and extended toward the *gag* gene through sequences required for the initiation of viral DNA synthesis (52) and a region believed to be required for packaging viral RNA into virions (21, 33, 42). Therefore, these deletions must have arisen during or after the final round of viral DNA synthesis, since RNAs with these deletions would not have been previously packaged into virions to serve as template. Except for pAO4, there was no significant homology (direct or inverted repeats) at the borders of the deleted sequences. Thus, we conclude that the mechanism(s) leading to this class of deletions is not dependent on sequence homology. pAO43 also had a small inversion of viral sequences associated with the deletion of *gag* and *pol* (Fig. 6). The exact size of the inversion was not determined. In addition to the large size of the deletions, another feature distinguishes most of the members of this class of molecules from those of the first class. While all the deletion endpoints lay near the U5 boundary of the LTR unit as in the first class, the positions were much more heterogeneous with respect to that boundary than the rather precise position of the deletion endpoint associated with deletions in tandem LTR units (with the exception of pAO53).

Cleavage signals within the LTR unit. In the integration reaction, circular DNA is cleaved at the circle junction site between tandem LTR units and inserted into the host chromosome with the loss of 2 bp from both the U5 and U3 domains (30, 56). Evidence from several sources suggests that the precise cleavage at the circle junction site during integration is carried out by a virus-encoded endonuclease. Such an activity is encoded within the *pol* gene and is found in virus particles. In avian leukosis virus the activity is associated with the DNA polymerase and also as a carboxy-terminal fragment of the polymerase that has only the endonuclease activity (9, 11, 36). This fragment, called pp32, binds in vitro with some specificity to the LTR region (27). Both the pp32 protein and the endonuclease activity associated with the polymerase cleave cloned viral DNA within the circle junction region (7; D. Grandgenett, personal communication). The endonuclease activities from murine leukemia virus and reticuloendotheliosis virus are less well characterized, but there is genetic evidence indicating a direct role for an analogous activity in the integration reaction. Virus isolates with mutations near the 3′ end of the *pol* reading frame, mutations which do not affect the polymerase activity, are unable to integrate viral DNA in newly infected cells (6, 31, 38). Thus it appears that a viral protein is responsible for recognizing and cleaving the circle junction site of viral DNA during integration.

One model that can explain the origin of circles with deletions that have one endpoint in the LTR region is aberrant cleavage of circular DNAs by the viral endonuclease. In this scheme, the initial cleavage occurs at or near the U5 boundary in the LTR unit and the second cleavage takes place at a variable distance to the right of the first cleavage, toward the *gag* gene. In the molecules with deletions within tandem LTR units, the deletions start either precisely at the circle junction boundary or 2 bp away (Fig. 5; 17, 18, 47). It is possible that these deletions arose as the result of cleavage at or within 2 bp of the circle junction site followed by either misplacement of the second cleavage beyond the appropriate site in U3 or removal of bases from the U3 domain by an exonuclease prior to ligation. At present it is not possible to distinguish between deletion formation involving aberrant cleavage of the circle junction site of tandem LTR units and deletions associated with circle closure, exonucleolytic removal of bases from the ends of linear DNA prior to circle formation, or incomplete synthesis of viral DNA followed by circle closure. As noted previously (17), inappropriate initiation of the second strand of viral DNA can result in a truncated linear molecule that, if circularized, would have a deletion at the circle junction site.

Larger deletions extending from single LTR units could also result from aberrant cleavage by the viral endonuclease. This model appears to be especially useful in explaining the structure of two of the molecules we analyzed (pAO4 and pAO43; Fig. 6). pAO4 had a large deletion, very similar to the other members of the second class of deleted molecules except that the end associated with the LTR unit was adjacent to the U5 domain, approximately 5 bp into the $tRNA^{Trp}$ primer binding site. There was a single intact LTR unit in this molecule, with the adjacent deletion extending for about 2,560 bp and terminating in the *pol* gene. If the precursor to this molecule was an intact circle, then the cleavage associated with the LTR unit must have occurred near a U5 domain in the absence of an adjacent U3 domain. Alternatively, since there is a 5-bp sequence homology at the site of the deleted sequences, the deletion in pAO4 could be explained by invoking a mechanism involving limited sequence homology such as slippage during DNA synthesis. In pAO43, a U5 domain was cleaved in the absence of an adjacent U3 domain since the U5 domain was at the outer boundary of tandem LTR units. The U5 domain next to U3 at the circle junction site was unaltered in this molecule, although there was an extra A · T base pair at the circle junction (Fig. 5). If the deletions in pAO4 and pAO43 were initiated by an aberrant cleavage, the exact site of the cleavage is difficult to assign owing to the possibility of exonucleolytic attack prior to ligation to form a circle. In pAO4, this cleavage would have had to be imprecise with respect to the U5 boundary since the deletion endpoint lies just outside the U5 domain.

pAO43 has tandem LTR units with a deletion extending from the outer U5 domain toward *gag*. pAO53, pAO2, and pAO61 (Fig. 6) have similar structures, except that in these molecules the deletions extend from a single LTR unit. A single mechanism involving the cleavage of an isolated U5 domain in circles with either one or two copies of the LTR unit can account for the structure of all of these molecules. However, for molecules with deletions that extend from within a single LTR unit mechanisms of deletion formation involving aberrant circle closure of linear DNA, exonuclease activity on linear DNA prior to circularization, and incomplete viral DNA synthesis are also consistent with the observed structures.

Although we only observed deletions that extended from or near the U5 domain toward *gag*, Shoemaker et al. (45) observed some molecules with deletions that begin at the U3 boundary of an LTR unit and extend toward the *env* gene in clones of circular murine leukemia virus DNA. Also, Duyk et al. (7) showed that the endonuclease from RSV can cleave separated strands of DNA representing the LTR region, suggesting that either the U5 or U3 domain can be cleaved first. One possible explanation for the absence of molecules

with deletions that extend from the U3 domain in the circles we analyzed might be because of our small sample size. This would be especially significant if aberrant cleavage events at the boundaries of LTR units occurred more frequently in the U5 domain than in the U3 domain.

We have not detected a class of molecules containing an inversion flanked by LTR units as first described by Shoemaker et al. (44, 45). These molecules, presumably the result of autointegration events, were detected at very low levels (3%) in the clones of murine leukemia virus DNA examined (45). Our failure to observe these types of molecules may be due to the small sample size. In theory, autointegration events can give rise to circles with inversions flanked by LTR units or to circles with a single LTR unit adjacent to a deletion. We do not believe that the molecules with large deletions we have characterized are the result of a precise autointegration event for two reasons: (i) two of the molecules (pAO4 and pAO43) have deletions that were not initiated from the circle junction site between tandem LTR units; and (ii) as a group, the large deletions in the 5′ region of the genome start at variable distances from the U5 boundary, in contrast to the precise cleavage of U5 associated with the normal integration reaction (56) or the autointegration reaction that results in an inversion (45, 51).

In total, one-third of the molecules we examined had deletions associated with the LTR region that arose during a single round of virus replication. At present we do not know what role cellular enzymes play in cleavage and ligation events. It has been reported that circular DNA molecules are present at higher levels after infection with RSV in the quail cell line QT6 than in other avian cells (12). It may be that host factors function inefficiently in these cells, promoting the formation of deleted molecules.

Implications for the origin of defective integrated DNA. One type of defective provirus can be accounted for directly by a deletion event in circular DNA prior to integration. In pAO43 a deletion event occurred at the outside U5 boundary of tandem LTR units. If a circular DNA molecule like pAO43 had integrated, the resulting provirus would have been flanked by LTR units but would have had an internal deletion starting at or near the upstream LTR unit and extending into *gag*. One such provirus has been detected near the c-*myc* gene in an avian leukosis virus-induced bursal lymphoma (32). The deletion starts 16 to 17 bp to the right of U5 in the $tRNA^{Trp}$ primer binding site and extends about 1,100 bp into *gag* (61). Another class of mutant proviruses may also be explained by this type of aberrant cleavage event. If the size of these deletions were smaller in some molecules, proviruses could be generated that would express viral genes but fail to package viral RNA. Two such mutants have been described (21, 22, 42), and one of these mutants efficiently packages cellular RNA in place of viral RNA (4, 22).

In Fig. 8 is a summary of several different types of viral DNA structures seen in infected cells. Linear DNA is the precursor to circular DNA with either one or two copies of the LTR unit. We propose that deleted circular DNA molecules can arise by an inappropriate cleavage signaled by a U5 domain in circular DNA. Cleavage at the circle junction site of tandem LTR units can lead to a small deletion extending into the adjacent U3 domain. Cleavage at a U5 domain not positioned next to a U3 domain can result in a molecule that has a large deletion extending into the *gag* gene. If the latter deletion takes place at the outside boundary of tandem LTR units, integration may still occur, resulting in a defective provirus with an internal deletion in the 5′ region of the genome. The generation of all of these forms of defective viral DNA could share a common feature in being initiated by an aberrant cleavage in the U5 domain.

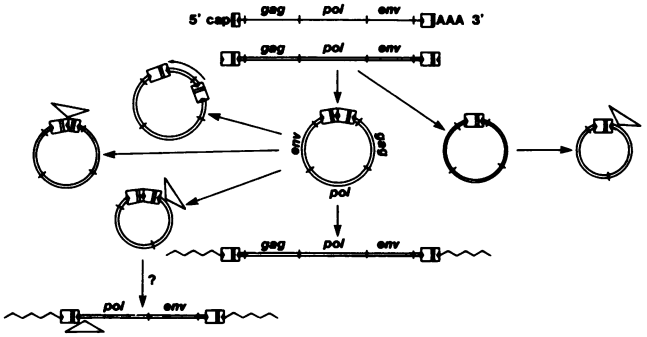

FIG. 8. Proposed pathways for the generation of several forms of viral DNA in infected cells.

ACKNOWLEDGMENTS

We thank Bill DeLorbe and Richard Parker for helpful discussion and Barbara Leonard for preparation of the manuscript. We are grateful to Bill Farmerie, John Manfredi, and Ron Shehee for their assistance in carrying out portions of this work.

This investigation was supported by Public Health Service grants R01 CA33147 and 5T32 CA09156 awarded by the National Cancer Institute, and by the North Carolina Biotechnology Center. J.O. is a fellow of the Leukemia Society of America.

ADDENDUM IN PROOF

Duyk et al. (J. Virol. **56**:589–599, 1985) have shown that the polymerase-associated endonuclease will cleave within the U5 domain in the absence of an intact adjacent U3 domain. This observation is consistent with our results which indicate that deletions formed in vivo can extend from the boundary of a U5 domain.

LITERATURE CITED

1. **Berger, S. A., and A. Bernstein.** 1985. Characterization of a retrovirus shuttle vector capable of either proviral integration or extrachromosomal replication in mouse cells. Mol. Cell. Biol. **5**:305–312.
2. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heynecker, H. W. Boyer, J. H. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
3. **Cepko, C. L., B. E. Roberts, and R. C. Mulligan.** 1984. Construction and application of a highly transmissible murine retrovirus shuttle vector. Cell **37**:1053–1062.
4. **Chen, P.-J., A. Cywinski, and J. M. Taylor.** 1985. Reverse transcription of 7S L RNA by an avian retrovirus. J. Virol. **54**:278–284.
5. **DeLorbe, W. J., P. A. Luciw, H. M. Goodman, H. E. Varmus, and J. M. Bishop.** 1980. Molecular cloning and characterization of avian sarcoma virus circular DNA molecules. J. Virol. **36**:50–61.
6. **Donehower, L. A., and H. E. Varmus.** 1984. A mutant murine leukemia virus with a single missense codon in *pol* is defective in a function affecting integration. Proc. Natl. Acad. Sci. USA **81**:6461–6465.
7. **Duyk, G., J. Leis, M. Longiaru, and A. M. Skalka.** 1983. Selective cleavage in the avian retroviral long terminal repeat sequence by the endonuclease associated with the αβ form of avian reverse transcriptase. Proc. Natl. Acad. Sci. USA **80**:6745–6749.
8. **Gilboa, E., S. W. Mitra, S. Goff, and D. Baltimore.** 1979. A

detailed model of reverse transcription and tests of crucial aspects. Cell **18**:93–100.
9. **Golomb, M., D. P. Grandgenett, and W. Mason.** 1981. Virus-coded DNA endonuclease from avian retrovirus. J. Virol. **38**:548–555.
10. **Graham, F. L., and A. J. Van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
11. **Grandgenett, D. P., A. C. Vora, and R. D. Schiff.** 1978. A 32,000-dalton nucleic acid-binding protein from avian retrovirus cores possesses DNA endonuclease activity. Virology **89**:119–132.
12. **Guntaka, R. V., O. C. Richards, P. R. Shank H.-J. Kung, N. Davidson, E. Fritsch, J. M. Bishop, and H. E. Varmus.** 1976. Covalently closed circular DNA of avian sarcoma virus: purification from nuclei of infected quail tumor cells and measurement by electron microscopy and gel electrophoresis. J. Mol. Biol. **106**:337–357.
13. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.
14. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26**:365–369.
15. **Hsu, T. W., J. L. Sabran, G. E. Mark, R. V. Guntaka, and J. M. Taylor.** 1978. Analysis of unintegrated avian RNA tumor virus double-stranded DNA intermediates. J. Virol. **28**:810–818.
16. **Hughes, S.** 1982. Appendix B: restriction maps of retroviral proviruses and corresponding cellular oncogenes, p. 1293–1315. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses: molecular biology of tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Ju, G., and A. M. Skalka.** 1980. Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. Cell **22**:379–386.
18. **Katz, R. A., C. A. Omer, J. H. Weis, S. A. Mitsialis, A. J. Faras, and R. V. Guntaka.** 1982. Restriction endonuclease and nucleotide sequence analyses of molecularly cloned unintegrated avian tumor virus DNA: structure of large terminal repeats in circle junctions. J. Virol. **42**:346–351.
19. **Kawai, S., and H. Hanafusa.** 1972. Genetic recombination with avian tumor viruses. Virology **49**:37–44.
20. **Kopchick, J. J., and D. W. Stacey.** 1984. Differences in intracellular DNA ligation after microinjection and transfection. Mol. Cell. Biol. **4**:240–246.
21. **Koyama, T., F. Harada, and S. Kawai.** 1984. Characterization of a Rous sarcoma virus mutant defective in packaging its own genomic RNA: biochemical properties of mutant TK15 and mutant-induced transformants. J. Virol. **51**:154–162.
22. **Linial, M., E. Medeiros, and W. S. Hayward.** 1978. An avian oncovirus mutant (SE21Q1b) deficient in genomic RNA: biological and biochemical characterization. Cell **15**:1371–1381.
23. **Majors, J. E., and H. E. Varmus.** 1981. Nucleotide sequence of host-proviral junctions for mouse mammary tumor virus. Nature (London) **289**:253–258.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
26. **Miller, C. K., and H. M. Temin.** 1983. High-efficiency ligation and recombination of DNA fragments by vertebrate cells. Science **220**:606–609.
27. **Misra, T. K., D. P. Grandgenett, and J. T. Parsons.** 1982. Avian retrovirus pp32 DNA-binding protein. I. Recognition of specific sequences on retrovirus DNA terminal repeats. J. Virol. **44**:330–343.
28. **Moscovici, C., M. G. Moscovici, H. Jimenez, M. M. C. Lai, M. J. Hayman, and P. K. Vogt.** 1977. Continuous tissue culture cell lines derived from chemically induced tumors of Japanese quail. Cell **11**:95–103.
29. **Panganiban, A. T., and H. M. Temin.** 1983. The terminal nucleotides of retrovirus DNA are required for integration but not for virus production. Nature (London) **306**:155–160.
30. **Panganiban, A. T., and H. M. Temin.** 1984. Circles with two tandem LTRs are precursors to integrated retrovirus DNA. Cell **36**:673–679.
31. **Panganiban, A. T., and H. M. Temin.** 1984. The retrovirus *pol* gene encodes a product required for DNA integration: identification of a retrovirus *int* locus. Proc. Natl. Acad. Sci. USA **81**:7885–7889.
32. **Payne, G. S., S. A. Courtneidge, L. B. Crittendon, A. M. Fadly, J. M. Bishop, and H. E. Varmus.** 1981. Analysis of avian leukosis virus DNA and RNA in bursal tumors: viral gene expression is not required for maintenance of the tumor state. Cell **23**:311–322.
33. **Pugatsch, T., and D. W. Stacey.** 1983. Identification of a sequence likely to be required for avian retroviral packaging. Virology **128**:505–511.
34. **Rassart, E., L. DesGroseillers, and P. Jolicoeur.** 1981. Molecular cloning of B- and N-tropic endogenous BALB/c murine leukemia virus circular DNA intermediates: isolation and characterization of infectious recombinant clones. J. Virol. **39**:162–171.
35. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase. J. Mol. Biol. **113**:237–251.
36. **Samuel, K. P., T. S. Papis, and J. G. Chirikjian.** 1979. DNA endonucleases associated with the avian myeloblastosis virus DNA polymerase. Proc. Natl. Acad. Sci. USA **76**:2659–2663.
37. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32**:853–869.
38. **Schwartzberg, P., J. Colicelli, and S. P. Goff.** 1984. Construction and analysis of deletion mutants in the *pol* gene of Moloney murine leukemia virus: a new viral function required for productive infection. Cell **37**:1043–1052.
39. **Sealy, L., M. L. Privalsky, G. Moscovici, C. Moscovici, and J. M. Bishop.** 1983. Site specific mutagenesis of avian erythroblastosis virus: erb-B is required for oncogenicity. Virology **130**:155–178.
40. **Shank, P. R., J. C. Cohen, H. E. Varmus, K. R. Yamamoto, and G. M. Ringold.** 1978. Mapping of linear and circular forms of mouse mammary tumor virus DNA with restriction endonucleases: Evidence for a large specific deletion occurring at high frequency during circularization. Proc. Natl. Acad. Sci. USA **75**:2112–2116.
41. **Shank, P. R., S. H. Hughes, H.-J. Kung, J. E. Majors, N. Quintrell, R. V. Guntaka, J. M. Bishop, and H. E. Varmus.** 1978. Mapping unintegrated avian sarcoma virus DNA: Termini of linear DNA bear 300 nucleotides present once or twice in two species of circular DNA. Cell **15**:1383–1395.
42. **Shank, P. R., and M. Linial.** 1980. Avian oncovirus mutant (SE21Q1b) deficient in genomic RNA: Characterization of a deletion in the provirus. J. Virol. **36**:450–456.
43. **Shank, P. R., and H. E. Varmus.** 1978. Virus-specific DNA in the cytoplasm of avian sarcoma virus-infected cells is a precursor to covalently closed viral DNA in the nucleus. J. Virol. **25**:104–114.
44. **Shoemaker, C., S. Goff, E. Gilboa, M. Paskind, S. W. Mitra, and D. Baltimore.** 1980. Structure of a cloned circular Moloney murine leukemia virus DNA molecule containing an inverted segment: implications for retrovirus integration. Proc. Natl. Acad. Sci. USA **77**:3932–3936.
45. **Shoemaker, C., J. Hoffmann, S. P. Goff, and D. Baltimore.** 1981. Intramolecular integration within Moloney murine leukemia virus DNA. J. Virol. **40**:164–172.
46. **Southern, E.** 1980. Gel electrophoresis of restriction fragments. Methods Enzymol. **68**:152–176.
47. **Swanstrom, R., W. J. DeLorbe, J. M. Bishop, and H. E. Varmus.** 1981. Nucleotide sequence of cloned unintegrated avian sarcoma virus DNA: viral DNA contains direct and inverted repeats similar to those in transposable elements. Proc. Natl. Acad. Sci. USA **78**:124–128.
48. **Swanstrom, R., R. C. Parker, H. E. Varmus, and J. M. Bishop.** 1983. Transduction of a cellular oncogene: the genesis of Rous sarcoma virus. Proc. Natl. Acad. Sci. USA **80**:2519–2523.
49. **Swanstrom, R., H. E. Varmus, and J. M. Bishop.** 1982. Nucle-

otide sequence of the 5′ noncoding region and part of the *gag* gene of Rous sarcoma virus. J. Virol. **41**:535–541.
50. **Tal, J., H.-J. Kung, H. E. Varmus, and J. M. Bishop.** 1977. Characterization of DNA complementary to nucleotide sequences adjacent to poly(A) at the 3′ terminus of the avian sarcoma virus genome. Virology **79**:183–197.
51. **Tamura, T., and T. Takano.** 1982. Long terminal repeat (LTR)-derived recombination of retroviral DNA: sequence analysis of an aberrant clone of baboon endogenous virus DNA which carries an inversion from the LTR to the *gag* region. Nucleic Acids Res. **10**:5333–5343.
52. **Taylor, J. M.** 1977. An analysis of the role of tRNA species as primers for the transcription into DNA of RNA tumor virus genomes. Biochim. Biophys. Acta **473**:57–71.
53. **Tereba, A., and K. G. Murti.** 1977. A very sensitive biochemical assay for detecting and quantitating avian oncornaviruses. Virology **80**:166–176.
54. **Thomas, P.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
55. **Van Beveren, C., E. Rands, S. K. Chattopadhyay, D. R. Lowy, and I. M. Verma.** 1982. Long terminal repeat of murine retroviral DNAs: sequence analysis, host-proviral junctions, and preintegration site. J. Virol. **41**:542–556.
56. **Varmus, H. E.** 1982. Form and function of retroviral proviruses. Science **216**:812–820.
57. **Varmus, H. E., and R. Swanstrom.** 1982. Replication of retroviruses, p. 369–512. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses: molecular biology of tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
58. **Vogt, P. K.** 1971. Spontaneous segregation of nontransforming viruses from cloned sarcoma viruses. Virology **46**:939–946.
59. **Vogt, P. K.** 1971. Genetically stable reassortment of markers during mixed infection with avian tumor viruses. Virology **46**:947–952.
60. **Wahl, G. M., M. Stern, and G. R. Stark.** 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl paper and rapid hybridization using dextran sulphate. Proc. Natl. Acad. Sci. USA **76**:3683–3687.
61. **Westaway, D., G. Payne, and H. E. Varmus.** 1984. Proviral deletions and oncogene base-substitutions in insertionally mutagenized c-*myc* alleles may contribute to the progression of avian bursal tumors. Proc. Natl. Acad. Sci. USA **81**:843–847.
62. **Yoshimura, F. K., and R. A. Weinberg.** 1979. Restriction endonuclease cleavage of linear and closed circular murine leukemia viral DNAs: discovery of a smaller circular form. Cell **16**:323–332.
63. **Zasloff, M., G. D. Ginder, and G. Felsenfeld.** 1978. A new method for the purification and identification of covalently closed circular DNA molecules. Nucleic Acids Res. **5**:1139–1152.