Avian Sarcoma and Leukosis Virus *pol*-Endonuclease Recognition of the Tandem Long Terminal Repeat Junction: Minimum Site Required for Cleavage Is Also Required for Viral Growth

D. COBRINIK,¹ R. KATZ,² R. TERRY,² A. M. SKALKA,² AND J. LEIS^{1*}

Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106,¹ and Department of Molecular Oncology, Roche Institute of Molecular Biology, Hoffmann-La Roche Inc., Nutley, New Jersey 07110²

Received 18 November 1986/Accepted 22 February 1987

Integration of retroviral DNA is a site-specific reaction involving an endonuclease encoded by the viral *pol* gene (*pol*-endo). In vitro the *pol*-endo from avian sarcoma and leukosis viruses (ASLVs) cleaves both DNA strands near the U5-U3 junction of tandem long terminal repeats (LTR-LTR junction) in single-stranded and replicative form (RF)-I substrates. We have reported previously that the sequences that are required for cleavage of single-stranded substrates by the $\alpha\beta$ form of the *pol*-endo differ for the plus and minus strands (G. Duyk, M. Longiaru, D. Cobrinik, R. Kowal, P. deHaseth, A. M. Skalka, and J. Leis, J. Virol. 56:589–599, 1985). This is not the case with RF-I substrates, in which a maximum of 22 base pairs of U5 and 8 base pairs of U3 were required for $\alpha\beta$ *pol*-endo cleavage in each strand. Insertion of a palindromic octanucleotide (CATCGATG) at the LTR-LTR junction abolished cleavage in RF-I but not in single-stranded DNA substrates. Deletion of the four nucleotides (TTAA) at the junction prevented cleavage in the plus strand of RF-I DNA, but did not affect cleavage of single-stranded DNA. Furthermore, the $\alpha\beta$ form of ASLV *pol*-endo did not recognize heterologous LTR-LTR junction sequences from the reticuloendotheliosis virus or Moloney murine leukemia virus in either substrate form, despite their sequence and structural similarities to the ASLV junction. These results support a role for a sequence-specific interaction between the ASLV *pol*-endo and the LTR-LTR junction domains that are required for cleavage. By using the infectious Rous sarcoma virus clone pATV8-K, we introduced a set of deletions into the U5 region that would be incorporated into the LTR-LTR junction on viral replication. In the unintegrated provirus, the deletions started 43 base pairs from the LTR-LTR junction and extended various lengths toward the junction. Results of transfection studies with these clones indicated that the U5 sequences that are required for virus production in vivo correspond to those that are required for cleavage of RF-I DNA in vitro.

Reverse transcription of retroviral genomes leads to the production of circular DNA replication intermediates containing single or tandem long terminal repeat (LTR) elements (for a review, see reference 36). Circular DNAs containing tandem LTRs undergo an integration reaction (integrative recombination) in which short inverted repeat sequences at the LTR-LTR junction are recombined with host chromosomal DNA (29). The process requires an endonuclease activity that is encoded by the 3'-terminal one-third of the *pol* gene (the *pol*-endo) (9, 30, 33) and is characterized by the deletion of two nucleotides on the U5 and U3 sides of the virus integration site and by duplication of four to six nucleotides (depending on the virus) at the host integration site (38).

In the avian sarcoma and leukosis virus (ASLV) family of retroviruses, the major *pol* gene products are proteolytically processed from a common precursor to their mature forms by the viral protease p15 (1). The processed products include a 95-kilodalton β polypeptide, a 63-kilodalton α polypeptide with an amino acid sequence that is identical to the N-terminal two-thirds of β (12, 31), and a phosphorylated 32-kilodalton polypeptide (pp32) with a sequence that corresponds to that of the C-terminal one-third of β (16, 32). These proteins can be isolated from virus particles as an α monomer, a $\beta\beta$ dimer, an $\alpha\beta$ heterodimer, and a pp32 dimer

(18); the last three of which possess endonuclease activity (15, 18, 23). The ability of the *pol*-endo to cleave near the U5-U3 junction in both strands of replicative form (RF)-I and single-stranded DNA substrates (10, 17) and the requirement of the *pol*-endo coding domain for provirus integration in a variety of retrovirus systems (9, 30, 33) provide evidence for a role for this enzymatic activity in integration and support its postulated role in the production of defective proviruses (27). No other virus or host functions are as yet known to be involved in provirus integration.

We have been studying the interaction between the virus *pol*-endo and the LTR-LTR junction sequences to better understand its role in integrative recombination. By using single-stranded DNA substrates containing either the plus or minus strands of the LTR-LTR junction, the $\alpha\beta$, $\beta\beta$, and pp32 forms of *pol*-endo have been shown to specifically cleave three nucleotides 5' of the junction center (10, 11), whereas on RF-I substrates, cleavage has also been detected two nucleotides 5' of the junction (19; unpublished data). The sequences that are required for specific cleavage of single-stranded DNA substrates by the $\alpha\beta$ form of the *pol*-endo were found to differ for the plus and minus strands: (i) up to 44 nucleotides are required for minus-strand cleavage, while at most 22 are required for plus-strand cleavage; (ii) the cleavage site is near the 5' end of the required sequences in the minus strand and near the 3' end of the required sequences in the plus strand; and (iii) the required

* Corresponding author.

FIG. 1. Comparison of sequences at the junction of tandem LTRs. The sequence of the plus strand of RAV-2 is shown at the top. The conserved CA and TG dinucleotides that represent the 3' and 5' termini of the integrated provirus are shown in larger type. The arrowhead indicates the preferred $\alpha\beta$ *pol-endo* cleavage site in the plus strand. Below the RAV-2 sequences are the corresponding plus-strand sequences for AMV, endogenous virus ev-1, RAV-2, the TTA deletion construct, and the CLA insertion construct. Junction sequences for MuLV and REV-T are also shown. Dashes indicate nucleotides that are identical to those in the RAV-2 sequence, and the horizontal arrows indicate inverted repeats, with breaks at noncomplementary nucleotides. The vertical line indicates the position of the LTR-LTR junction.

sequences in each strand bear little resemblance to each other, despite their complementarity and their overlap with an imperfect inverted repeat (11). This ability to recognize specific sequences that bear little obvious relationship to one another suggests that the *pol-endo* may recognize structural features that are some distance from the cleavage site, as well as the specific sequences at the cleavage site. Furthermore, the ability of *pol-endo* to specifically cleave single-stranded DNA substrates raises the possibility that the two strands may be recognized independently in vivo, as well as in vitro (11).

In this study we determined the sequence requirements for specific cleavage of double-stranded RF-I DNA substrates by the $\alpha\beta$ form of *pol-endo*. We also studied the ability of the enzyme to cleave altered LTR-LTR junctions containing internal deletions and insertions and to cleave heterologous retrovirus junctions in single-stranded and RF-I DNA substrates (Fig. 1). In addition, the biological relevance of the sequence requirements for cleavage was evaluated by constructing infectious proviral genomes containing U5 deletions. Our results indicate that there are differences between the sequence requirements for $\alpha\beta$ *pol-endo* cleavage of single-stranded and RF-I DNA substrates, that cleavage of RF-I substrates requires sequence-specific interactions on both sides of the junction, and that the U5 sequences that are required for virus production correspond to those that are required for the cleavage of RF-I substrates.

MATERIALS AND METHODS

Reagents. [α - 32 P]dCTP (600 to 800 Ci/mmol) and [γ - 32 P]ATP (7,000 Ci/mmol) were purchased from New England Nuclear Corp., Boston, Mass. The Klenow fragment of DNA polymerase I, T4 DNA polymerase, T4 DNA ligase, and calf intestinal phosphatase were obtained from Boehringer Mannheim Biochemicals, Indianapolis, Ind. T4 polynucleotide kinase was obtained from Pharmacia Fine Chemicals, Piscataway, N.J. Restriction enzymes were from Boehringer Mannheim, except for *Hinf*I (U.S. Biochemicals) and *Mlu*I (New England BioLabs, Inc., Beverly, Mass.) All

reactions with commercial enzyme preparations were performed in accordance with the recommendations of the manufacturer. M13 reverse sequencing primers were from New England BioLabs (5'-AACAGCTATGACCATG3') and P-L Biochemicals, Inc. (Milwaukee, Wis.) (5'-CAG GAAACAGCTATGAC3'), and the pBR322 *Eco*RI site primer (clockwise; 5'-GTATCAGGAGGCCCT3') and *Hind*III site primer (counterclockwise; 5'-GCAATTTAACT GTGAT3') were from New England BioLabs.

Oligodeoxynucleotides. Oligodeoxynucleotides were synthesized on a solid support by the phosphoramidite method (5, 8) manually or by using a DNA synthesizer (350A; Applied Biosystems) and were purified as described previously (11). The M13 17-mer primer sequence was 5'-GTAAAACGACGGCCAGT3'. Oligodeoxynucleotides used in the construction of U5 deletions (see below) were GG (5'-ATTCGACTACGAGCACATG3'), NN (5'-CGGTC ACCAAATGAAGCCTTCTGC3'), Q (5'-TTTGGTGACCG AT3'), EE (5'-ATTGCATGAAGCAGAAGGCTTCA3'), S (5'-TTCATGTC3'), E29 (5'-ATTGCACATGCATGAAGCA GAAGGCTTCA3'), T (5'-TTCATGCATGTGC3'), R (5'-ATTGCAGAAGGCTTCATTGGTGACCGAT3'), E (5'-CATGAAGCAGAAGGCTTCA3'), and F (5'-TTCATG CATGTGCTCGTAGTCG3'). Oligodeoxynucleotides used to sequence M13 mp19ATV-SAC and pATV8-K U5 deletion mutants were N (5'-ATGAAGCCTTCTGC3') and minus primer (5'-TCACGTCGGGGTCCAC3'), respectively. Oligodeoxynucleotides used in the construction of M13 mp8 Δ TTAA(+) (plus orientation) (see below) were M (5'-TGTAAGTCTTATGC3'), G (5'-CGACTACGAGCACAT G3'), E, N, and F (as described above); and L (5'-CAA GAGTATTGCATAAGACTAC3'). Oligodeoxynucleotides used in the construction of clones containing the *Clal* insertion at the LTR-LTR junction were U (5'-CATCGA TG3') and V (5'-GGCATCGATGAATGTAGTCTTATGC3').

DNA clones. M13 clones mp18:5, 6, 7, 8, 9, and 10; mp2:SRBtd7 (SRBtd, Schmidt-Ruppin type B transformation defective); mp7:RAV 2-12 (RAV, Rous-associated virus); and mp7:SRBtd9 were as described previously (11). For this study, the inserts in the clones described above were in the plus orientation, except for clones td9 and 2-12. M13 mp18:REV-127(+) (REV-T, reticuloendotheliosis virus strain T) (plus orientation) and M13 mp19:REV-127(-) (minus orientation) were made by subcloning the *Xba*I-*Xho*I fragment of pAP127 (29) between the *Xba*I and *Sall* sites of mp18 and mp19, respectively. M13 mp19:MuLV(-) (MuLV: murine leukemia virus) was made by subcloning the *Kpn*I-*Pvu*II fragment of a plasmid containing the MuLV tandem LTR-LTR junction (34) between the *Hind*III and *Kpn*I sites of M13 mp19. M13 mp18:MuLV(+) was derived from the minus-orientation clone by inserting the *Eco*RI-*Hind*III fragment into M13 mp18.

Construction of clones containing altered LTR-LTR junctions. A palindromic octanucleotide containing a *Clal* restriction endonuclease site was added separately to the U3 and the U5 termini of two LTRs (Fig. 2). In step A, for the addition to the U5 terminus, linear double-stranded viral DNA was synthesized in melittin-permeabilized RAV type 2 (RAV-2) virions (4) in the presence of [α - 32 P]dATP (500 μ Ci/nmol). The product from a 500- μ l reaction mixture was extracted with sodium dodecyl sulfate and phenol, precipitated with ethanol, treated with RNase A (20 μ g/ml) and RNase T1 (100 U/ml) for 45 min at 37°C, and then treated with proteinase K (10 μ g/ml) in the presence of 0.2% sodium dodecyl sulfate for 20 min at 37°C. Following extraction with phenol, the DNA was passed through a Sephadex G100

column and precipitated with ethanol. The recovered DNA (200 ng) was ligated to oligodeoxynucleotide (U) which was phosphorylated at its 5' terminus (25). The *SalI*-*ClaI* fragment (Fig. 2) was inserted into the corresponding sites of pBR322 DNA by standard methods (25), and the desired clone was selected by using a probe containing homologous DNA that was obtained from a *KpnI* fragment of lambda Gd11 DNA (20). In step B, for the addition of a *ClaI* site to the U3 terminus of the LTR, a primer repair method was used (13). A 250-base-pair *MstII*-*EcoRI* fragment was obtained from a molecular clone of RAV-2 DNA (3) (Fig. 2). A 5' ³²P-labeled primer (V) which contained the *ClaI* sequence joined to a sequence homologous to the U3 terminus of the LTR plus strand was annealed to the denatured *MstII*-*EcoRI* fragment. It was then extended by the addition of the four deoxynucleotides and the Klenow fragment of *Escherichia coli* DNA polymerase (13). Labeled duplex was recovered from an agarose gel. The mismatch at the 3' end of the plus strand was removed by treatment with T4 DNA polymerase under conditions that favored its 3' exonuclease activity (25). The overhang containing the *ClaI* sequence was filled in by the polymerase activity after the addition of radioactive deoxynucleotide triphosphates. The labeled *ClaI*-*SphI* fragment (Fig. 2) was recovered from a polyacrylamide gel and inserted into the corresponding sites of pBR322 DNA. In step C, plasmid DNA from the clone obtained in step B was used as a vector for cloning a *ClaI*-*EcoRI* fragment from the U5 *ClaI* clone described in step A. The product (designated pBR:CLA) contained a reconstituted LTR-LTR junction with the *ClaI* oligodeoxynucleotide inserted between the two termini (Fig. 1). The *EcoRI*-*SphI* junction fragment was also inserted into M13 mp18 RF-I DNA to produce M13 mp18:CLA(+). A minus-orientation subclone was derived from M13 mp18:CLA(+) by releasing the insert with *EcoRI* and *HindIII* and reinserting it between the *EcoRI* and *HindIII* sites of M13 mp19, forming M13 mp19:CLA(-).

M13 mp8 ΔTTAA (plus orientation) was made by synthesizing and assembling a sequence comprised of the terminal 37 nucleotides of the RAV-2 U5 and the terminal 25 nucleotides of the RAV-2 U3, exclusive of the TTAA sequences at the LTR-LTR junction. All oligodeoxynucleotides were phosphorylated before use except for L and G. G was ligated to pE in the presence of F, pN to pF in the presence of E, GpE to pM in the presence of N, and L to pFpN in the presence of E. The assembled fragments (GpEpM and LpNpF) were annealed, blunt ends were made by filling in with the Klenow fragment of DNA polymerase I, and the product was cloned into the *HindII* site of M13 mp8. M13 mp9 ΔTTAA (minus orientation) was subcloned by inserting an *EcoRI*-*HindIII* fragment from M13 mp8 ΔTTAA(+) into the corresponding sites in M13 mp9.

Construction of retroviral genomes containing U5 deletions. The plasmid pATV8-K is a derivative of pATV8 (21). It contains a complete Rous sarcoma virus Prague strain (RSV Pr-C) provirus with a single LTR permuted around a unique *KpnI* site and cloned into the *KpnI* site of pMG4 (a pBR322 derivative which has a *KpnI*-*BamHI*-*XhoI* polylinker inserted between the pBR322 *EcoRI* and *SphI* sites; R. Katz, unpublished data). M13 mp19ATV-SAC was derived from pATV8-K by subcloning the 2.7-kilobase-pair *SacI* fragment, which contains a single LTR, into M13 mp19 in the plus orientation (same sense as the viral genome).

Deletions were introduced into the U5 portion of the single LTR as shown in Fig. 3. A *PvuI* fragment derived from mp19ATV-SAC was treated with calf intestinal phosphatase, followed by *HinfI*, and the indicated fragment (Fig. 3) was

FIG. 2. Construction of pBR:CLA. Details are provided in the text. Shaded areas indicate the U5 sequence; striped areas represent the *ClaI* sequence. Thick bars are viral sequences, and thin bars are the vector pBR322 sequence. Restriction endonuclease sites are abbreviated as follows: R, *EcoRI*; Sa, *SalI*; Sp, *SphI*; C, *ClaI*; M, *MstII*. dr is direct repeat. The asterisk denotes ³²P radioactivity. DNA sections in the figure are not drawn to scale.

isolated from an agarose gel. This fragment (0.125 pmol) was ligated to 2.5 pmol of labeled synthetic oligodeoxynucleotides (Δ6, Δ15, Δ21, and Δ28) which contain RAV-2 sequences stretching from the *BstEII* site within the RAV-2 primer binding site to various points within U5. The fragments contained *HinfI* sticky ends at the U5 terminus that were phosphorylated and *PvuI* sticky ends at the primer binding site terminus that were not (as described below). Ligation products, which were isolated from non-denaturing polyacrylamide gels (25), were ligated to the large *PvuI* fragment of mp19ATV-SAC, and the desired M13 recombinant phage were identified by plaque hybridization (2) by using a 5' ³²P-labeled oligodeoxynucleotide probe that was complementary to sequences at the U5 terminus (oligodeoxynucleotide NN). The *MluI*-*BstEII* fragment of each of the mp19ATV ΔU5 clones was inserted between the corresponding sites in pATV8-K DNA by standard methods (25).

The synthetic oligodeoxynucleotide fragments were assembled from the individual oligodeoxynucleotides listed above. All oligodeoxynucleotides were phosphorylated at their 5' termini by T4 polynucleotide kinase with [γ-³²P]ATP (50 Ci/mmol) before use, except for NN. Approximately 100 pmol of pEpQ, pEEpQ, and pE29pQ were made by ligation of the appropriate oligodeoxynucleotides in the presence of an equimolar amount of NN; and NNpF was made in the presence of E29, NNpE in the presence of E, NNpS in the presence of EE, and pGGpEpQ in the presence of F. After each phosphorylation and ligation step, oligodeoxy-

FIG. 3. Construction of U5 deletions in a cloned RSV provirus. Details are described in the text. Thick bars indicate sequences derived from pATV8-K, shaded areas indicate U5 sequences, thin bars indicate vector M13 mp19 sequences, and bent lines indicate deletions. CIP is calf intestinal phosphatase. Restriction sites are as follows: P, *PvuI*; S, *SacI*; H, *HinfI*; B, *BstEII*. DNA sections in the figure are not drawn to scale.

nucleotides were purified as described previously (11). The synthetic fragments that were used to construct each deletion were as follows: Δ6, pGGpEpQ:NNpF; Δ15, pE29pQ:NNpT; Δ21, pEEpQ:NNpS; and Δ28, pR:NN. The relevant portions of all of the constructs described above and those described in previous sections were verified by DNA sequencing analyses.

Primer extension mapping. RF-I DNA (1.5 μg), which was prepared as described by Zinder and Boeke (39), was incubated in the presence of 10 U of the avian myeloblastosis virus (AMV) αβ form of *pol-endo* (prepared as previously described [23]), 20 mM Tris hydrochloride (pH 8.4), 8 mM 2-mercaptoethanol, 60 mM NaCl, and 2 mM MnCl₂ (which was added after a 5-min preincubation at room temperature) in 18-μl reactions at 37°C. At the indicated times 1-μl fractions were removed for agarose gel analysis and 4.5-μl fractions were added to 4 μl of 2.5× HIN buffer (330 mM Tris hydrochloride [pH 7.5], 33 mM MgCl₂, 250 mM NaCl) and 0.5 to 1.0 pmol (7.5×10^6 to 15×10^6 cpm) of a 5' ³²P end-labeled oligodeoxynucleotide primer. The solution was heated to 95°C for 3 min and then cooled on ice. A total of 1 μl of a 0.5 mM solution of each of the four deoxynucleotide triphosphates and 1 U of the Klenow fragment of DNA polymerase I were added, and the reaction was incubated at 37°C for 20 min. The reaction was stopped by the addition of 10 μl of sequencing buffer and heated to 95°C for 2 min

before it was loaded onto 6% polyacrylamide sequencing gels (26).

In general, 25 and 90% of RF-I DNA was converted to RF-II DNA after 5 and 25 min of incubation, respectively, under these conditions, regardless of the insert DNA sequence. For clones with LTR-LTR junction sequences that were inserted into M13 in the plus orientation (M13 mp18:10, 5, 6, 7, 8, and 9, and M13 mp2:td7), the M13 17-mer primer was used to map plus-strand cleavage sites, and the M13 reverse sequencing primer was used to map minus-strand cleavage sites. For clones with the LTR-LTR junction in the minus orientation (M13 mp7:RAV 2-12 and M13 mp7:SRBtd9), the opposite was the case. Primers were end-labeled with T4 polynucleotide kinase and a molar excess of [γ -³²P]ATP and were purified by hydrophobic chromatography with NENsorb cartridges (New England Nuclear), as described by the manufacturer. Dideoxynucleotide sequencing reactions with the various supercoiled substrate DNAs and appropriate primers were performed as described previously (6) by using either end-labeled or nonradioactive primers. Primer extension mapping of single-stranded substrates was as described previously (11).

Transfection and virus assays. Chick embryo fibroblasts (C/O or line-O cells, Regional Poultry Laboratories, East Lansing, Mich.), which were grown in F-10 medium with 5% calf serum or in 199 medium with 4% fetal calf serum to 70% confluency in dishes (diameter, 60 mm), were transfected with 2 or 4 μg of *KpnI* digested pATV8-K and its U5 deletion derivatives by using the DEAE-dextran (150 μg/ml) procedure (24). Virus production was assayed by using a modification of the rapid reverse transcriptase assay (14) as follows: 5 μl of cell supernatants was assayed in a solution of 0.03% Nonidet P-40, 83 mM Tris hydrochloride (pH 8.2), 100 mM NaCl, 17 mM magnesium acetate, 33 mM dithioerythritol, 0.5 optical density unit of poly(G) per ml, and 0.25 optical density unit of oligo (dC)₁₂₋₁₈ per ml in a final volume of 30 μl. After 1 h of incubation at 37°C, 10-μl fractions were spotted onto DE-81 paper (Whatman, Inc., Clifton, N.J.), which was then washed three times for 15 min each with 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and rinsed twice with 95% ethanol. The paper was dried under a heat lamp and used for autoradiography, cut into squares for scintillation spectroscopy, or both. For infections, supernatants containing equal amounts of virus, as determined by reverse transcriptase activity, were added to uninfected cells; and virus production was assayed as described above.

RESULTS

Sequence requirements for cleavage by the αβ *pol-endo* near the LTR-LTR junction in RF-I DNA substrates. The αβ form of *pol-endo* cleaves preferentially within the CA dinucleotides 5' of the LTR-LTR junction, as shown in Fig. 1 for the plus strand of RAV-2. Using a primer extension assay, we found that the sequence requirements for this cleavage in single-stranded DNA substrates differ for the plus and minus strands (11). To determine if this is also the case for double-stranded supercoiled DNA, which is believed to be the substrate for integrative recombination, we repeated these analyses with RF-I substrates. The DNAs were digested for various lengths of time with the αβ *pol-endo* and then heated in the presence of an oligodeoxynucleotide primer that was complementary to sequences near the LTR-LTR junction. Annealed primers were extended by using the Klenow fragment of DNA polymerase I, and the cleavage

FIG. 4. Structure of clones that were used to determine the sequence requirements for *pol-endo* cleavage near the LTR-LTR junction of RF-I substrates. The number of base pairs contained on either side of the junction in each clone is indicated above the bars. The cleavage sites on the plus and minus strands are indicated by arrowheads above and below the bars, respectively. Cleavage by *pol-endo* is indicated by closed circles, and the absence of cleavage is indicated by open circles. Open circles without arrowheads indicate that cleavage at the same relative position was not observed. The + sign below clone M13 mp18:9 indicates the location of a minus-strand cleavage site that was enhanced in this clone. All clones contain LTR-LTR junction sequences in the plus orientation except td9, ΔTTAA, and 2-12. For clones 10, 5, 6, and 7, LTR sequences were inserted between the *Hind*II and *Sph*I sites of M13 mp18. For clones 8 and 9, LTR sequences were inserted into the *Hind*II site of M13 mp19. For td7, LTR sequences were inserted into the *Eco*RI site of M13 mp2, and for td9 and 2-12 they were inserted into the *Eco*RI site of M13 mp7. For pBRCLA, junction sequences were inserted between the *Eco*RI and *Sph*I sites of pBR322, and for ΔTTAA (minus orientation) they were inserted at the *Hind*II site of M13 mp9. The structure of an unintegrated RSV provirus containing tandem LTRs is shown at the top, where RI indicates *Eco*RI sites in U3, and the thin bars between U3 and U5 represent the R region.

sites were identified by comparing the resultant run off fragments to dideoxynucleotide sequencing ladders that were prepared with the same primer template. The relevant structure of the substrates that were tested is summarized in Fig. 4.

U5 requirements for RF-I cleavage. Clones 10, 5, and 6, which contain 36, 28, and 22 nucleotides of U5, respectively, were cleaved in the minus strand at the CA dinucleotide site three nucleotides 5' of the LTR-LTR junction; while clone 7, which contains only 17 nucleotides of U5, was not (Fig. 5A, heavy arrows). This result is in contrast with data that were obtained with single-stranded substrates, in which minus-strand cleavage at this site was not detected with clones 5 and 6 (11). Thus, 18 to 22 nucleotides of U5 are necessary for cleavage at the CA dinucleotide in the minus strand of RF-I DNA substrates, whereas 29 to 36 nucleotides are required in single-stranded substrates (see Fig. 8A and B). Clones 5 and 10, but not clones 6 and 7, were cleaved weakly two nucleotides 3' of the junction in the minus strand (Fig. 5A, light arrows). This site was observed previously in RF-I DNA (17).

U5 deletion clones 10, 5, and 6 were also cleaved at the CA dinucleotide in the plus strand, while clone 7 was not (data not shown). Thus, the U5 sequence requirements for

FIG. 5. Gel analysis of primer extended products from *pol-endo*-treated RF-I DNA substrates. For all panels, lanes 1, 2, and 3 indicate incubation with the *pol-endo* for 0, 5, and 25 min, respectively; and C, T, A, and G represent sequencing ladders corresponding to the indicated dideoxynucleotide chain terminators. Heavy arrows indicate the location of the preferred cleavage site within the CA dinucleotide 5' of the LTR-LTR junction, and light arrows indicate the location of a cleavage site two bases 3' of the junction in the minus strand (see Fig. 8B). (A) Mapping of minus-strand cleavage sites in clones M13 mp18:10, 5, 6, and 7, as described in the text. A partial nucleotide sequence is listed at the side of the figure, with a horizontal line denoting the LTR-LTR junction. The nucleotide sequence represents the complement of the DNA strand that is being cleaved by the *pol-endo*. End-labeled reverse sequencing primer (P-L Biochemicals) was used for the runoff and sequencing reactions. (B) Mapping of plus-strand cleavage sites in RF-I M13 mp8:9. An unlabeled 17-mer primer was used for dideoxynucleotide sequencing reactions, causing the sequencing products to run approximately one-half of a nucleotide slower than the corresponding runoff fragments. (C) Mapping of the minus-strand cleavage sites in RF-I M13 mp8:9. Reactions were performed as described for panel A.

RF-I plus-strand cleavage is identical to that defined for the RF-I minus strand (Fig. 4).

U3 requirements for RF-I cleavage. We also analyzed LTR-LTR junctions of U3 deletion mutants (Fig. 4). With

RF-I M13 mp2:SRBtd7 DNA, which has an extensive U3 deletion starting with the third nucleotide from the junction (35), we did not observe cleavage near the LTR-LTR junction in either strand. There was also no detectable cleavage in M13 mp2:SRBtd9 and M13 mp2:RAV 2-12 RF-I DNAs, both of which have extensive U3 deletions starting with the first nucleotide of U3 (data not shown; Fig. 4).

The minimum number of U3 nucleotides that are required for cleavage at the plus-strand CA dinucleotide in RF-I substrates is not clear, because this site was not cleaved in M13 mp2:SRBtd7, whereas in M13 mp8:9 it was (Fig. 4 and 5B). Both of these clones have two nucleotides of U3, but the flanking sequences differ. Thus, more than two U3 nucleotides may have a role in plus-strand cleavage, but they appear not to be specifically recognized. Because the plus-strand CA dinucleotide was cleaved in M13 mp8:8 RF-I DNA (data not shown; Fig. 4), a maximum of eight nucleotides of U3 may be required for plus-strand cleavage in RF-I substrates. In contrast, U3 sequences are not required for plus-strand cleavage in single-stranded substrates (11) (see Fig. 8A and B).

The U3 requirements for specific minus-strand cleavage were the same for RF-I and single-stranded substrates. M13 mp8:8 RF-I DNA was cleaved at the CA dinucleotide in the minus strand (data not shown), while M13 mp8:9 RF-I DNA was not (Fig. 5C). Interestingly, cleavage at the minus-strand site to the left of the junction mentioned above (Fig. 5A, light arrow) was reproducibly and strongly enhanced in the latter substrate (Fig. 5C). Taken together, the results presented above indicate that up to eight nucleotides of U3 play a role in cleavage of both DNA strands in RF-I substrates.

$\alpha\beta$ *pol-endo* cleavage of RF-I and single-stranded substrates with deletions and insertions at the LTR-LTR junction. In most retroviruses, a four-base TTAA sequence separated the conserved CA and TG dinucleotides at the LTR-LTR junction (Fig. 1). Deletion of the terminal T nucleotide, however, does not seem to affect integration in MuLV mutants (7). To determine whether the spacing across the junction is critical for cleavage of the RAV-2 substrates by the $\alpha\beta$ *pol-endo*, deletion and insertion constructs were prepared (see above), and the cleavage sites were mapped.

The insertion construct contained a palindromic octadeoxynucleotide sequence CATCGATG, which included a *Cla*I site, inserted into the RAV-2 LTR-LTR junction (Fig. 1). As with wild-type substrates, single-stranded *CLA* insertion DNAs were cleaved in each strand in the conserved CA sequences (Fig. 6A and B). When the *CLA* insertion junction was present in an RF-I substrate, it was not cleaved at the CA dinucleotide in either strand but was weakly cleaved at the minus-strand site to the left of the LTR-LTR junction described above (Fig. 6C). Detection of this site indicates that cleavage and runoff conditions were sufficient to detect cleavage at the usual CA dinucleotide site, had it occurred. Also, control runoff reactions with pBR:CLA DNA that was cut near the LTR-LTR junction with a restriction enzyme (*Hinf*I) demonstrated successful runoff analysis on both strands (data not shown). Results with the *Cla*I insertion construct were similar to those obtained with the plus strand of substrates with U3 deletions in that single-stranded, but not RF-I, DNAs were cleaved at the CA dinucleotide.

The deletion construct contained 35 base pairs of the U5 side and 23 base pairs of the U3 side of the RAV-2 LTR-LTR junction, but not the TTAA sequence located at the junction center (Fig. 1). The modified junction was cleaved in both

strands of single-stranded substrates at the CA cleavage site (Fig. 6D). The RF-I DNA that contained this junction was also cleaved in the minus strand, but cleavage at the CA site was not detected in the plus strand (data not shown). The results with the insertion and deletion constructs suggest that an intact LTR-LTR junction is important for plus-strand cleavage in RF-I but not for single-stranded substrates. However, it appears that an intact junction is not necessarily required for minus-strand cleavage of RF-I substrates; the octanucleotide insertion prevented cleavage, while the TTAA deletion did not. The effects of the insertions and deletions on $\alpha\beta$ *pol-endo* cleavage of RF-I substrates are summarized in Fig. 4.

$\alpha\beta$ *pol-endo* does not cleave at the LTR-LTR junction of unrelated retroviruses. Although the inverted repeat and CA and TG dinucleotides are conserved at the LTR termini of retrotransposons, sequences outside of these are essentially nonhomologous (36, 38). Because the $\alpha\beta$ form of *pol-endo* commonly cleaves weakly at secondary sites that are nonhomologous to the preferred site at the LTR-LTR junction (data not shown), it is possible that the enzyme recognizes structural features that are common to all cleavage sites. To test whether such structural features are recognized at the LTR-LTR junctions of other retroviruses, the avian myeloblastosis virus (AMV) $\alpha\beta$ *pol-endo* was incubated with single-stranded and RF-I DNAs that contain the LTR-LTR junctions of REV-T or MuLV (Fig. 1). Using standard assay conditions, we could not detect specific cleavage in either strand of the REV-T and the MuLV LTR-LTR junctions in single-stranded or RF-I substrates (negative data not shown). These results suggest that the AMV $\alpha\beta$ *pol-endo* specifically recognizes the ASLV LTR-LTR junction sequences.

Relationship between the sequence requirements for *pol-endo* cleavage and retrovirus replication. The studies described above define a domain consisting of 18 to 22 base pairs of U5 and 3 to 8 base pairs of U3 which is necessary to cleave both strands of RF-I substrates in vitro. Only the eight terminal base pairs of U3 are conserved among members of the ASLV retrovirus family (Fig. 1, see RSV, endogenous virus ev-1, and AMV sequences). Thus, it seems likely that this domain may be required for cleavage in vivo as well as in vitro (11). While the U5 sequences of ASLVs are also conserved, homology extends beyond the region that is required for specific cleavage in either single-stranded or RF-I DNA substrates (3). To evaluate the significance of the U5 domain required for *pol-endo* cleavage in vitro, a series of deletions was introduced into this region in a cloned infectious RSV provirus, pATV8-K. The U5 terminal sequences were replaced with various lengths of the U5 sequences from the related virus RAV-2 (Fig. 3). The resultant deletions started 43 base pairs from the junction and extended to positions 38, 29, 23, and 16 (clones pATV Δ 6, Δ 15, Δ 21, Δ 28, respectively). Thus, they represent a nested set of deletions that reproduce the sequences and the borders that were used to define the U5 requirements for in vitro cleavage in single-stranded and RF-I DNA substrates (cf. Fig. 4). Provirus containing these deletions was introduced into endogenous genome-free (line-O) primary chick embryo fibroblasts by using the DEAE-dextran transfection procedure, and the resultant virus production and spread was monitored by assaying for reverse transcriptase activity in cell supernatants.

The results indicate that virus was produced after transfection with Δ 6, Δ 15, and Δ 21, while none was detected after transfection with Δ 28 (Fig. 7A), even after passage of

FIG. 6. Gel analyses of primer extended products from $\alpha\beta$ *pol-endo*-treated substrates that contain mutations at the LTR-LTR junction. Cleavage sites are mapped in the plus (A) and minus (B) strands of single-stranded M13 mp18:CLA (+) and M13 mp19:CLA (-), respectively. Single-stranded DNAs containing the octanucleotide *Clal* insertion between the junction of two LTRs was incubated with the $\alpha\beta$ form of *pol-endo* for 0 (lanes 1) or 20 (lanes 2) min, and primer extension analysis was carried out as described previously (11). Arrows indicate the preferred cleavage site. C, T, A, and G are as described in the legend to Fig. 5. A partial nucleotide sequence, which represents the complement of the DNA strand being cleaved by the *pol-endo*, is presented at the side of the figure. The braces denote the octanucleotide insertion at the LTR-LTR junction. (C) Mapping of cleavage sites in the minus strand of pBR:CLA RF-I DNA incubated in the presence of the $\alpha\beta$ form of *pol-endo* for 0 (lane 1) or 20 (lane 2) min. Reactions were performed as described in the legend to Fig. 5, except that an end-labeled *EcoRI* site primer was used for runoff analysis and sequencing. The heavy arrow indicates the absence of cleavage at the site observed in wild-type substrates, and the light arrow indicates cleavage at a site two nucleotides to the left of the junction. (D) Mapping of cleavage sites in single-stranded M13 mp8 Δ TTAA plus-strand and M13 mp7 Δ TTAA minus-strand DNAs. Single-stranded circular DNAs in which the TTAA was missing from the junction were analyzed as described in panel A. The arrowhead indicates site of TTAA deletion.

transfectants for as long as 5 weeks. Therefore, the U5 terminal sequences that are required for virus production correspond to those that are required for specific $\alpha\beta$ *pol-endo* cleavage of both strands of RF-I substrates. It thus appears likely that replication of mutant Δ 28 is blocked due to the loss of sequences required for *pol-endo* cleavage.

Although virus production after transfection was not abolished by deletion of U5 sequences that were more than 21

base pairs from the junction (i.e., in Δ 6, Δ 15, and Δ 21), the appearance of virus was delayed when compared with that of a wild-type provirus clone. A similar delay was noted when line-O chick embryo fibroblasts were infected with equal amounts of virus (as estimated by reverse transcriptase units) from wild-type RSV and Δ 15 and Δ 21 mutants (Fig. 7B), as well as for Δ 6 (data not shown). In this case, progeny virus was first detected within 7 days for the wild type and

only after 2 weeks for $\Delta 15$ and $\Delta 21$ mutants. The delay may be related to the effect of the deletion (as few as six base pairs in $\Delta 6$) or, less likely, to the effect of the substitution of RAV-2 sequences for RSV sequences, which differ at five positions in $\Delta 6$.

DISCUSSION

The ASLV *pol-endo* cleaves preferentially at the plus- and minus-strand CA dinucleotides near the LTR-LTR junction in single-stranded as well as in double-stranded RF-I substrates (10, 11). Results of this study indicate that the sequence requirements for this cleavage differ with the two substrates (Fig. 8A and B). Cleavage of the plus strand requires 18 to 22 nucleotides of U5 in both single-stranded and RF-I DNAs. However, plus-strand cleavage occurs in the absence of U3 sequences in single-stranded DNAs (11) but not in RF-I DNAs (td9 and 2-12; see Fig. 4). Moreover, one RF-I DNA that contained two U3 nucleotides (td7) was not cleaved in the plus strand, while another RF-I DNA that contained two U3 nucleotides (M13 mp8:9) was (Fig. 4). Taken together, these results suggest that specific recognition of U3 sequences may not be required for plus-strand cleavage in RF-I DNA. Rather, up to eight U3 nucleotides may fulfill a structural requirement for cleavage, and these may be substituted by other sequences, as in mp8:9.

For *pol-endo* cleavage of the minus strand at the LTR-LTR junction, three to eight nucleotides of U3 are required in both single-stranded and RF-I substrates (Fig. 8A and B).

FIG. 7. (A) Reverse transcription assay for virus production 14 days after transfection of line-O chick embryo fibroblasts with pATV-K and U5 deletion mutants. The designations are as follows: WT, pATV8-K wild type; CON, mock transfection. Reverse transcriptase activity was measured as described in the text, using poly(G)-oligo(dC)₁₂₋₁₈ as template primer, and an autoradiogram of the dCMP incorporation is shown. (B) Kinetics of appearance of virus after infection of cells with equal amounts of RSV (Pr-C) and virus mutants containing U5 deletions $\Delta 15$ and $\Delta 21$. Viral titers were estimated by assaying reverse transcriptase activity as described for panel A, except that dCMP incorporation was quantitated by scintillation spectroscopy.

FIG. 8. Schematic representation of the sequence requirements for $\alpha\beta$ *pol-endo* cleavage and virus production. (A) Sequence requirements for cleavage of the plus (top) and the minus (bottom) strands of the RAV-2 LTR-LTR junction in single-stranded DNA substrates (11). Solid areas indicate the maximum required sequences; numbers indicate the distance (in nucleotide base pairs) from the LTR-LTR junction; arrows indicate the preferred cleavage sites in the CA dinucleotides three nucleotides 5' of the junction (Fig. 1); dots indicate nucleotides that may or may not be required for cleavage. (B) Sequence requirements for cleavage of the plus (top) and minus (bottom) strands of RF-I substrates. The dashed arrow indicates cleavage at a site on the minus strand two nucleotides 3' of the junction. Other notations are as described for panel A. (C) Sequence requirements for virus production following transfection of line-O chick embryo fibroblasts with provirus mutants $\Delta 6$, $\Delta 15$, $\Delta 21$, and $\Delta 28$. The numbers indicate the number of base pairs that were deleted. The expected structure of the junction of tandem LTRs in unintegrated circular viral DNA for each U5 deletion mutant is as shown. The U5 sequences between the deletions and the LTR-LTR junction are from RAV-2; all other sequences are from RSV Pr-C. Virus production from cells that were transfected with each mutant provirus (as detected by reverse transcriptase activity) is indicated at the right.

These correspond to the sequences that are highly conserved among members of the ASLV family and include the cleavage site itself. On the U5 side, an additional 9 to 15 nucleotides are required for cleavage of minus-strand, single-stranded substrates, as compared with RF-I substrates (Fig. 8A and B). The reason for this is unclear.

The U5 and U3 sequence requirements for the cleavage of RF-I DNA are essentially the same in both strands, suggesting that an interaction of the *pol-endo* with sequences on both sides of the LTR-LTR junction may be involved in RF-I cleavage, as was proposed previously from genetic studies of murine retroviruses (7). Our observations suggest that the ASLV *pol-endo* specifically recognizes duplex DNA extending 18 to 22 base pairs into U5 and recognizes somewhat less specifically U3 sequences extending 3 to 8 base pairs from the LTR-LTR junction.

A *pol-endo* interaction across the LTR-LTR junction provides an explanation for our results that were obtained with the CLA insertion and the TTAA deletion substrates. Each of these constructs was cleaved in both strands in single-stranded DNA, indicating that disruption of the wild-type sequence across the junction does not impair the interactions that are necessary for single-strand cleavage. In RF-I substrates, however, the TTAA deletion was cleaved only in the minus strand, while the CLA insertion DNA was not cleaved in either strand. It should be noted that the CLA

insertion construct maintains the palindromic nature of the tandem LTR junction and has sequences at the center of the inverted repeat that resemble the wild-type junction, as shown below:

Normal junction: TTCATT AATGTA
 · · · · · X X · · · X ·
CLA insertion junction: TTCATTCATG CATGAATGTA

Thus, the inability of the *pol-endo* to cleave the CLA insertion and the TTAA deletion constructs in RF-I DNA may reflect a requirement for interactions between the U5 and U3 domains in which spacing, the highly conserved T-A base pairs, or both are important.

It has been shown previously (19) that insertion of a single A nucleotide into a Schmidt-Ruppin LTR-LTR junction does not prevent specific plus- or minus-strand cleavage of RF-I substrates by the pp32 form of *pol-endo*. Also, deletion of two, but not one, terminal U5 nucleotides at the MuLV LTR-LTR junction prevents integration (7). These results suggest that *pol-endo* cleavage is not blocked by a single nucleotide insertion, but may be blocked with deletions of two or more nucleotides. Our in vitro results show that removal of four nucleotides from the junction blocks cleavage of the plus strand, whereas insertion of eight nucleotides blocks cleavage of both strands.

The failure of the AMV $\alpha\beta$ *pol-endo* to recognize MuLV and REV-T junctions in single-stranded and RF-I substrates indicates that substrate recognition by *pol-endo* is species specific and that common features such as inverted repeats or conserved terminal nucleotides are insufficient for recognition. This conclusion is supported by the different U5 sequence requirements for spleen necrosis virus (SNV) and RSV virus production (28; see Fig. 8C) and by the antigenic divergence of the MuLV *pol-endo* from those of SNV and ASLV (37). Our in vitro results suggest that the major portion of the sequences recognized by the ASLV *pol-endo* are in U5.

Under conditions in which they are most active, both the $\alpha\beta$ and pp32 forms of the ASLV *pol-endo* cleave at a site one nucleotide removed from that expected for integration. Cleavage at the expected site is observed only under limiting reaction conditions with pp32 (17; M. Katzman and D. Cobrinik, unpublished observations). However, both of these *pol-endo* forms appear to recognize the same features of the LTR-LTR junction, since they require the same minimum sequences for cleavage of single-stranded DNA substrates (unpublished observations). Furthermore, the correlation between the sequences required for cleavage by the $\alpha\beta$ *pol-endo* and for viral growth (see below) suggests that the interactions studied here are biologically important.

Results of transfections with proviruses that have deletions in the U5 domain show that the region which is required for $\alpha\beta$ cleavage of RF-I substrates is also required for viral growth. Cells transfected with pATV Δ 28, in which only 15 of the 18 to 22 nucleotides required for cleavage were present, did not produce detectable levels of virus, whereas transfection of provirus with 21 of the nucleotides did (Fig. 8B and C). Since cleavage at the LTR-LTR junction of the circular provirus occurs during retrovirus DNA integration (29), it seems likely that mutant Δ 28 virus production is blocked at the integration step. Although integration may not be an absolute requirement for virus production (22, 28), the inability of SNV and MuLV mutants which contain alterations in the junction region (7, 28) or in the *pol-endo* domain (9, 30, 33) to integrate is correlated with 10^3 - to 10^6 -fold decreases in virus titer. A similar impairment in Δ 28 repli-

cation efficiency after DEAE-dextran transfection would render virus production undetectable by the reverse transcriptase assay used here.

Mutants that had deletions adjacent to the U5 domain that was required for cleavage ($\Delta 21$, $\Delta 15$, and $\Delta 6$) produced virus. However, the appearance of virus was delayed compared with that of the wild type (Fig. 7B). Because line-Q cells do not have endogenous virus sequences with strong homology to the ASLVs and because the delay in virus spread is reproducible, it is unlikely that this poor growth reflects the necessity to repair the defect via recombination with endogenous viruses or through pseudoreversion. The growth defect could be caused by a partial defect in integration. If so, these mutants identify additional sequences which are important to this event. However, it is also possible that steps other than integration are affected. This highly conserved region (3) encodes other important functions such as templating reverse transcription of strong-stop DNA, forming the 5' end of the genomic and subgenomic viral mRNAs, and possibly packaging. The RSV sequences that were deleted, replaced, or both by RAV-2 sequences in pATV $\Delta 6$ could define a new *cis*-acting function of the LTR which is not required for provirus integration but which affects the efficiency of virus replication.

ACKNOWLEDGMENTS

We are grateful to A. Panganabin and H. Temin for providing pAP127 and to S. Goff for providing a subclone containing the MuLV tandem LTR junction. Geoffry Duyk constructed a M13ΔTTAA and a MuLV subclone. Fern Alexander contributed to construction of the *Clal* insertion clones, and Laverne Soskey provided excellent technical assistance.

This investigation was supported in part by Public Health Service grant CA38046 from the National Cancer Institute to J.L. D.C. is supported in part by National Institute of Health training grant T32-GM-07250 from the National Institute of General Medical Science.

LITERATURE CITED

1. Alexander, F., J. Leis, D. Soltis, R. Crowl, W. Danho, M. Poonian, Y. Pan, and A. M. Skalka. 1986. Proteolytic processing of ASLV *pol-endo* recombinant proteins reveals another *pol* gene domain. *J. Virol.* **61**:534-542.
2. Benton, W., and R. Davis. 1977. Screening lambda gt recombinant clones by hybridization to single plaques *in situ*. *Science* **196**:180-182.
3. Bizub, D., R. Katz, and A. M. Skalka. 1984. Nucleotide sequence of noncoding regions in Rous-associated virus-2. Comparisons delineate conserved regions important in replication and oncogenesis. *J. Virol.* **49**:557-565.
4. Boone, L., and A. M. Skalka. 1980. Two species of full length cDNA are synthesized in high yield by melittin treated avian retrovirus particles. *Proc. Natl. Acad. Sci. USA* **77**:847-851.
5. Caruthers, M., S. Beaucage, Y. Efcavitch, E. Fisher, R. Goldman, P. deHaseth, W. Mandecki, M. Matteucci, M. Rosendahl, and Y. Stabinsky. 1983. Chemical synthesis and biological studies on mutated gene-control regions. *Cold Spring Harbor Symp. Quant. Biol.* **47**:411-418.
6. Chen, E., and P. Seeburg. 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. *DNA* **4**:165-170.
7. Colicelli, J., and S. Goff. 1985. Mutants and pseudorevertants of Moloney murine leukemia virus with alterations at the integration site. *Cell* **42**:573-580.
8. deHaseth, P., R. Goldman, C. Cech, and M. Caruthers. 1983. Chemical synthesis and biochemical reactivity of bacteriophage lambda PR promoter. *Nucleic Acids Res.* **11**:3773-787.
9. Donehower, L., and H. 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.
10. 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 $\alpha\beta$ form of reverse transcriptase. *Proc. Natl. Acad. Sci. USA* **80**:6745–6749.
 11. Duyk, G., M. Longiaru, D. Cobrinik, R. Kowal, P. deHaseth, A. M. Skalka, and J. Leis. 1985. Circles with two tandem long terminal repeats are specifically cleaved by *pol* gene-associated endonuclease from avian sarcoma and leukemia viruses: nucleotide sequences required for site-specific cleavage. *J. Virol.* **56**:589–599.
 12. Gibson, W., and I. Verma. 1974. Studies on the reverse transcriptase of RNA tumor viruses: structural relatedness of two subunits of avian RNA tumor viruses. *Proc. Natl. Acad. Sci. USA* **71**:4991–4994.
 13. Goeddel, D., H. Shepard, E. Yelverton, D. Leung, R. Crea, A. Slona, and S. Pestka. 1980. Synthesis of human fibroblast interferon by *E. coli*. *Nucleic Acids Res.* **8**:4057–4073.
 14. Goff, S., P. Traktman, and D. Baltimore. 1981. Isolation and properties of Moloney murine leukemia virus mutants: use of a rapid assay for release of virion reverse transcriptase. *J. Virol.* **38**:239–248.
 15. Golomb, M., and D. Grandgenett. 1979. Endonuclease activity of purified RNA-directed DNA polymerase from avian myeloblastosis virus. *J. Biol. Chem.* **254**:1606–1613.
 16. Grandgenett, D., R. Knaus, and P. Hippenmeyer. 1983. Antibodies against a synthetic peptide of the avian retrovirus pp32 protein and a α DNA polymerase subunit. *Virology* **130**: 257–262.
 17. Grandgenett, D., and A. Vora. 1985. Site-specific nicking at the avian retrovirus LTR circle junction by the viral pp32 endonuclease. *Nucleic Acids Res.* **13**:6205–6221.
 18. Grandgenett, D., A. Vora, and R. Schiff. 1978. A 32,000-dalton nucleic acid-binding protein from avian retrovirus cores possesses DNA endonuclease activity. *Virology* **89**:119–132.
 19. Grandgenett, D., A. Vora, R. Swanstrom, and J. Olsen. 1986. Nuclease mechanism of the avian retrovirus pp32 endonuclease. *J. Virol.* **58**:970–974.
 20. Hishinuma, F., P. DeBona, S. Astrin, and A. M. Skalka. 1981. Nucleotide sequence of acceptor site and termini of integrated avian endogenous provirus ev-1: integration creates a 6 base pair repeat of host DNA. *Cell* **23**:155–164.
 21. Katz, R., C. Omer, J. Weis, S. Mitsialis, A. Faras, and R. Guntaka. 1982. Restriction endonuclease and nucleotide sequence of molecularly cloned unintegrated avian tumor virus DNA: structure of large terminal repeats in circle junctions. *J. Virol.* **42**:346–351.
 22. Kopchick, J., G. Ju, A. M. Skalka, and D. Stacey. 1981. Biological activity of cloned retroviral DNA in microinjected cells. *Proc. Natl. Acad. Sci. USA* **78**:4383–4387.
 23. Leis, J., G. Duyk, S. Johnson, M. Longiaru, and A. M. Skalka. 1983. Mechanism of action of the endonuclease associated with the $\alpha\beta$ and $\beta\beta$ forms of avian RNA tumor virus reverse transcriptase. *J. Virol.* **45**:727–739.
 24. Lopata, M., D. Cleveland, and B. Sollner-Webb. 1984. High level transient expression of a chloramphenicol acetyl transferase gene by DEAE-dextran mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment. *Nucleic Acids Res.* **12**:5707–5717.
 25. Maniatis, T., E. Fritsch, and J. Sambrook. 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 26. Maxam, A., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**:560–564.
 27. Olsen, J., and R. Swanstrom. 1985. A new pathway in the generation of defective retrovirus DNA. *J. Virol.* **56**:779–789.
 28. Panganabin, A., and H. Temin. 1983. The terminal nucleotides of retrovirus DNA are required for integration but not for virus production. *Nature (London)* **306**:155–160.
 29. Panganabin, A., and H. Temin. 1984. Circles with two tandem LTRs are precursors to integrated retrovirus DNA. *Cell* **36**: 673–679.
 30. Panganabin, A., and H. 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.
 31. Rho, H., D. Grandgenett, and M. Green. 1975. Sequence relatedness between the subunits of AMV reverse transcriptase. *J. Biol. Chem.* **250**:5278–5280.
 32. Schiff, R., and D. Grandgenett. 1980. Partial phosphorylation in vivo of the avian retrovirus pp32 DNA endonuclease. *J. Virol.* **36**:889–893.
 33. Schwartzberg, P., J. Colicelli, and S. Goff. 1984. Construction and analysis of deletion mutations in the *pol* gene of Moloney murine leukemia virus: a new function required for establishment of the integrated provirus. *Cell* **37**:1043–1052.
 34. Shoemaker, C., S. Goff, E. Gilboa, M. Paskink, S. 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.
 35. Skalka, A. M., G. Duyk, M. Longiaru, P. deHaseth, R. Terry, and J. Leis. 1984. A role for the avian retroviral reverse transcriptase in the integration of retroviral DNA. *Cold Spring Harbor Symp. Quant. Biol.* **49**:651–659.
 36. Skalka, A. M., and J. Leis. 1984. Retroviral DNA integration. *BioEssays* **1**:206–210.
 37. Tanese, N., M. Roth, and S. Goff. 1986. Analysis of retroviral *pol* gene products with antisera raised against fusion proteins produced in *Escherichia coli*. *J. Virol.* **59**:328–340.
 38. Varmus, H. 1983. Retroviruses. In J. Shapiro (ed.), *Mobile genetic elements*, p. 411–503. Academic Press, Inc., New York.
 39. Zinder, N., and J. Boeke. 1982. The filamentous phage (F ϕ) as vectors for recombinant DNA—a review. *Gene* **19**:1–12.