MOLECULAR AND CELLULAR BIOLOGY, May 1986, p. 1599–1607
0270-7306/86/051599-09$02.00/0

Vol. 6, No. 5

Bidirectional Transcription from a Solo Long Terminal Repeat of the Retrotransposon TED: Symmetrical RNA Start Sites†

PAUL D. FRIESEN,* WILLIAM C. RICE, DAVID W. MILLER,‡ AND LOIS K. MILLER

Department of Bacteriology and Biochemistry, University of Idaho, and the Idaho Agriculture Experimental Station, Moscow, Idaho 83843

Received 15 November 1985/Accepted 27 January 1986

A single copy of the retrotransposon TED was found integrated within the DNA genome of the insect baculovirus, *Autographa californica* nuclear polyhedrosis virus. After excision of the element from the viral genome, a single long terminal repeat (LTR) remained behind. We have examined the effect of this solo TED LTR on the local pattern of viral transcription. Most prominent was the transcription of two sets of abundant RNAs; both originated within the LTR but extended in opposite directions into flanking viral genes. By promoting symmetric transcription of adjacent genes, the solo LTR has the capacity to activate or repress gene expression in two directions. Primer extension analysis demonstrated that the divergent LTR transcripts were initiated near the same point within a 22-base-pair sequence having hyphenated twofold symmetry. Analogous symmetries at the initiation sites of other retrotransposon LTRs, including copia and Ty, suggested that these sequences serve to establish the precise start for transcription.

The retrotransposons represent a group of eucaryotic transposable elements which bear a striking resemblance to the retroviral proviruses (for review, see reference 1). These elements, which include copia and copia-related transposons of *Drosophila melanogaster* and Ty of *Saccharomyces cerevisiae*, are flanked at both ends by long terminal repeats (LTRs). The LTRs carry signals necessary for initiation and termination of transcription. The transcribed internal portion of these elements contains several retroviruslike genes, including a gene (*pol*) with homology to reverse transcriptase (5, 9, 29, 38). In the case of Ty, the full-length RNA extending from LTR to LTR is copied into DNA by a process identical to that of the retroviruses, suggesting that transposition occurs through reverse transcription of an RNA intermediate (3).

By virtue of their ability to integrate randomly into the host genome and alter expression of nearby genes, the retroviruses represent an important class of insertion mutagens (for review, see reference 44). Similarly, the retrotransposons cause insertion mutations in *D. melanogaster* and *S. cerevisiae* (reviews in references 35 and 37). Integration within regulatory regions or introns of various genes has resulted in enhancement or inhibition of expression. These alterations occur at the level of transcription and are presumably due to insertion of new promoters, termination signals, or both, located within the retrotransposon. The mutagenic effects of inserted copia, gypsy, and Ty elements are suppressed by additional mutations at unlinked loci which restore transcription of the affected genes to wild-type levels (25, 32, 51). Mutation of the *S. cerevisiae SPT3* (suppressor of Ty) gene, for instance, abolishes normal Ty transcription, which suggests that the mutational effect of certain Ty insertions is due to transcription originating from the element (51). The exact mechanisms involved are still unclear.

TED is the first copia-related transposable element identified in a metazoan besides *D. melanogaster* (28). This mobile element is repeated in a dispersed fashion approximately 50 times in the chromosome of the moth *Trichoplusia ni* (a Lepidopteran noctuid). During an infection, TED was observed to transpose from the host cell to the circular DNA genome (~128 kilobases [kb]) of the baculovirus *Autographa californica* nuclear polyhedrosis virus (AcNPV) (28). This viral insertion mutant, FP-D, was identified by the reduced number of viral polyhedral occlusion bodies it produced in the nucleus of infected cells (for baculovirus reviews, see references 10 and 22 and L. K. Miller *in* R. R. Grandos and B. A. Federici, ed., *The Molecular Biology of the Baculoviruses*, in press).

We show here that TED has all the structural features of a retrotransposon. Integration of the single copy of TED into the AcNPV genome enabled us to study in detail the transcription and movement of a clonal isolate of a retrotransposon. In this report, we examined the effects on viral transcription of a single (solo) LTR left behind in the viral genome after excision of TED. These analyses were conducted in a permissive cell line (*Spodoptera frugiperda*) which lacks sequences homologous to TED, thereby eliminating potential influences of sister elements. The solo LTR was transcriptionally active and directed the synthesis of abundant RNAs extending in opposite directions into flanking viral genes. The divergent RNAs were initiated from a small 22-nucleotide transcriptional element within the LTR having partial twofold (dyad) symmetry. The role of DNA sequence symmetry in establishing the precise start for transcription from the LTR of TED and other retrotransposons is discussed.

MATERIALS AND METHODS

Virus and cell lines. AcNPV mutant FP-D (28, 34) was isolated after 25 serial passages of wild-type AcNPV (strain L-1; 24) in cultured cells (line TN-368; 16) of *T. ni* (the cabbage looper, Lepidoptera; Noctuidae). FP-D produces significantly reduced levels of polyhedral occlusion bodies in cell culture relative to wild-type AcNPV and is therefore referred to as a "few polyhedra" (FP) mutant. Plaque-purified FP-D

* Corresponding author.

† Research Paper no. 85510 of the Idaho Agriculture Experimental Station.

‡ Present address: Genetics Institute, Cambridge, MA 02115.

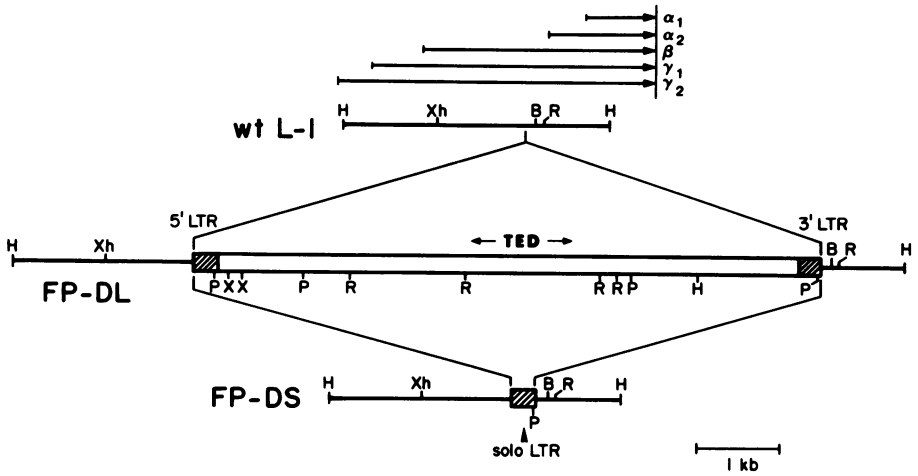

FIG. 1. Restriction maps of the *Hin*dIII-K fragment of wild-type (wt) L-1 AcNPV and the viral insertion mutants FP-DL and FP-DS. Mutant FP-DL carries the transposable element TED (7.3 kb) inserted 75 nucleotides to the left of the single *Bst*EII site within the *Hin*dIII-K restriction fragment of L-1 AcNPV (28). The *Hin*dIII-K region (85.1 to 87.5 map units) of the wild-type viral genome is transcribed from left to right into an overlapping group of five RNAs having a common 3′ end (top). TED is flanked by two LTRs designated here as 5′ and 3′ on the basis of structural homology with the retroviral proviruses (see text). Viral mutant FP-DS (bottom) carries a solo LTR (270 bp) of TED. H, *Hin*dIII; Xh, *Xho*I; P, *Pst*I; B, *Bst*EII; R, *Eco*RI; X, *Xma*I.

contains two virus populations, FP-DL and FP-DS, in an approximate ratio of 4 to 1 (Fig. 1). FP-DL generates a mixture of FP-DL and FP-DS during infection, whereas plaque-purified FP-DS remains genetically pure (28). Wild-type L-1 AcNPV, FP-D, and FP-DS viruses were routinely propagated in cultured cells (line IPLB-SF21; 46) of *S. frugiperda* (the fall armyworm, Lepidoptera; Noctuidae), using TC100 growth medium (GIBCO Laboratories, Grand Island, N.Y.) supplemented with 10% fetal bovine serum.

Recombinant plasmids. Restriction enzyme fragments *Hin*dIII-K and -Q of wild-type L-1 AcNPV and fragment *Hin*dIII-K of insertion mutants FP-D and FP-DS were cloned (28) and subcloned using plasmids pBR322 and pUC8 (47). Recombinant plasmids were propagated in *Escherichia coli* JM83, and plasmid DNA was prepared as described by Holmes and Quigley (17).

DNA sequencing. Sequence analysis was conducted by subcloning the FP-DS solo LTR and the 5′ and 3′ LTRs of TED into the replicative form of M13, mp18 or mp19. Dideoxy-chain termination sequencing as described by Sanger et al. (40) was performed on the single-stranded M13 recombinant phage DNAs using a 15-base oligonucleotide primer obtained from Bethesda Research Laboratories (Bethesda, Md.). The nucleotide sequences shown were determined for both DNA strands.

Northern blot analysis of RNA. Monolayer IPLB-SF21 cells, infected at a multiplicity of infection of 20 with wild-type or FP-DS AcNPV, were harvested at the indicated times, and RNA was extracted from the cytoplasmic fraction with phenol as described (14). Poly(A)$^+$ RNA was purified by oligo(dT)-cellulose chromatography (18). The poly(A)$^+$ RNA (2 to 10 μg) was denatured by glyoxalation (27), fractionated by electrophoresis on agarose gels, and transferred to nitrocellulose. Recombinant plasmids, as indicated in the text, were radiolabeled by nick translation (5×10^8 cpm/μg) and hybridized to RNA blots using conditions described by Thomas (43). After washing, the blots were subjected to autoradiography using Kodak XAR5 film.

S1 nuclease mapping of RNA transcripts. The locations of the 5′ and 3′ ends of individual transcripts were determined using the S1 nuclease procedure of Weaver and Weissman (49). DNA probes were prepared by cleaving the appropriate recombinant plasmids with a single restriction enzyme. The 5′ and 3′ ends were radiolabeled with T4 polynucleotide kinase or T4 DNA polymerase, respectively (26), then cleaved with a second enzyme to yield fragments labeled exclusively at one end. The desired fragments were purified by agarose or acrylamide gel electrophoresis. RNA-DNA hybridizations were conducted at 50°C in 80% formamide–40 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] (pH 6.4)–0.4 M NaCl–1 mM EDTA. S1-resistant fragments were denatured, fractionated by electrophoresis on 4% polyacrylamide–7 M urea–TBE (89 mM Tris borate, pH 8.3, 2 mM EDTA) gels, and subjected to autoradiography. Size standards were derived from *Taq*I or *Msp*I digests of plasmid pUC19 (53).

Primer extension analysis. For analysis of the 5′ ends of RNA transcripts initiating within the solo LTR, two DNA primers were prepared as described above and purified by polyacrylamide gel electrophoresis. These primers included the 34-base-pair (bp) *Eco*RV-*Alu*I fragment (5′ end labeled at the *Eco*RV site), extending from position −178 to −145 within the TED LTR (RNA cap site +1), and the 74-bp *Taq*I-*Bst*EII fragment (5′ end labeled at the *Bst*EII site), extending from position +80 to +153 within the flanking viral region. Both primers were annealed separately to poly(A)$^+$ RNA (2 μg) from FP-DS-infected cells and extended with reverse transcriptase as described (14). To obtain a sequencing pattern of the region, the above primers were treated briefly with exonuclease III, annealed to appropriate single-strand M13, mp18 or mp19 subclones, and extended using the dideoxy-chain termination procedure of Sanger et al. (40). Extended primers were precipitated with ethanol, denatured, and fractionated by electrophoresis on 43-cm-long 6% polyacrylamide–8 M urea–TBE gels.

RESULTS

Effects of the insertion of the solo TED LTR on viral RNA transcription. FP-DL is an FP mutant of L-1 AcNPV which produces reduced levels of polyhedral occlusion bodies in cell culture. The mutant carries the host (*T. ni*)-derived

transposable element TED (7.3 kb) inserted within the AcNPV restriction fragment *Hin*dIII-K (Fig. 1). Characteristic of a retrotransposon, TED is flanked at both ends by LTRs. Upon excision of the element from the viral genome (possibly via homologous recombination between the LTRs), a solo LTR (0.27 kb) is left behind at the same site (Fig. 1, bottom). The resulting mutant, FP-DS, also displays the FP phenotype.

We explored the effect of the solo LTR on transcription of viral genes at the site of TED insertion. This region of the wild-type genome, encompassing the *Hin*dIII-K restriction fragment, is transcribed into an overlapping group of RNAs that terminate at a common 3′ site (Fig. 1, top). Each RNA, representing three temporal classes (α, β, and γ), is transcribed from a unique promoter (14).

Northern analysis of RNA from wild-type and FP-DS infections revealed that the solo LTR had a dramatic effect on transcription within the viral *Hin*dIII-K region (Fig. 2). Most prominent was the synthesis of four new RNAs (1.05, 1.68, 2.38, and 3.50 kb) in FP-DS-infected cells; these RNAs appeared at 6 h and accumulated through 24 h after infection (Fig. 2, right panel). The unusually high levels of these RNAs were demonstrated by comparing the Northern blots of Fig. 2, which depict RNA isolated from equivalent numbers of cells during wild-type (left) and FP-DS (right) infections. Each of the three temporal classes of viral RNA was detected at the normal time after FP-DS infection, but their levels differed markedly from those observed during wild-type infection. Most affected were the immediate early α transcripts located downstream from the inserted LTR; the major α_1 (1.07-kb) RNA was reduced approximately 10-fold, and the α_2 (1.38-kb) RNA was reduced to levels below detection (the α_1 RNA and the new FP-DS RNA [1.05 kb] are different transcripts; see below). In contrast, the early β RNA accumulated to levels three- to fourfold higher than normal at 6 h. After the appearance of the new FP-DS RNAs, both the β and the γ RNAs were reduced relative to their wild-type levels.

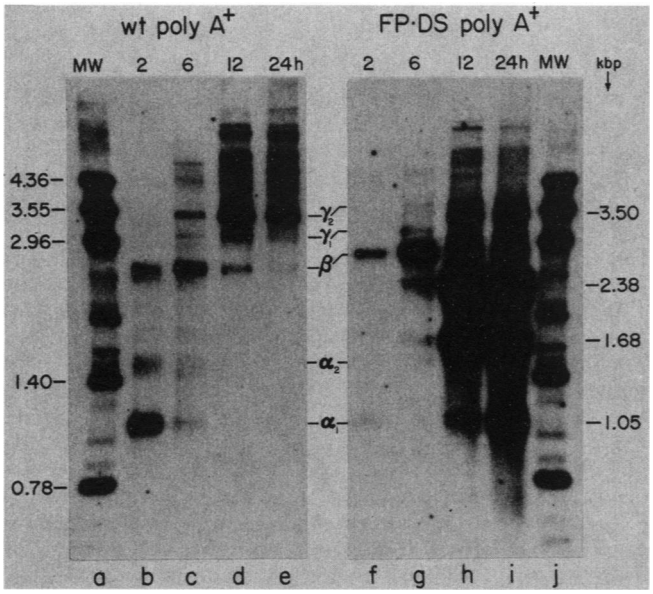

FIG. 2. Northern blot analysis of RNA transcribed within the *Hin*dIII-K region of wild-type (wt) and FP-DS AcNPV. Poly(A)$^+$ RNA was isolated from parallel cultures of *S. frugiperda* cells (line IPLB-SF21) 2, 6, 12, and 24 h after infection with wild-type (left) or FP-DS (right) AcNPV. After denaturation with glyoxal and fractionation on a single 1.7% agarose gel, the wild-type and FP-DS RNAs were transferred to nitrocellulose and hybridized to the wild-type *Hin*dIII-K fragment cloned into pBR322 and radiolabeled by nick translation. Sizes of the glyoxalated molecular weight standards (lanes a and j) in kilobase pairs (kbp) are indicated on the left, and sizes (kbp) of the FP-DS induced RNAs are indicated on the right. Equivalent exposures (4 h) of both autoradiograms are shown to illustrate the relative abundance of wild-type and FP-DS RNAs.

The β and γ RNAs, but not the α RNAs, from FP-DS-infected cells were approximately 0.3 kb larger than their wild-type counterparts (Fig. 2). This was consistent with the notion that these later transcripts extended uninterrupted through the inserted LTR (0.27 kb) and terminated at their normal position downstream. Additional Northern hybridization, using a probe composed entirely of the LTR and lacking viral sequences, demonstrated that the FP-DS β, γ_1, and γ_2 RNAs were complementary to the solo TED LTR and confirmed that these RNAs spanned the LTR (data not shown).

The FP-DS-induced RNAs originate within the solo TED LTR. To determine the location and direction of transcription of the four new RNAs within the *Hin*dIII-K region of mutant FP-DS, we employed the S1 nuclease mapping procedure of Weaver and Weissman (49). Preliminary Northern blot analysis indicated that the 1.05- and 2.38-kb RNAs were homologous to the solo LTR and to viral sequences located to the left of the LTR. The 5′ ends of these leftmost RNAs were therefore mapped with a 980-bp *Xho*I-*Pst*I probe (5′ end labeled at *Xho*I) containing the solo LTR (Fig. 3A). A single 920-bp fragment was protected from S1 nuclease by FP-DS RNA (Fig. 3A, lane d). These data and subsequent mapping of the 3′ termini (see below) indicated that the 2.38- and 1.05-kb RNAs were transcribed from right to left with their 5′ ends located within the solo LTR (270 bp) approximately 210 bp from the left edge. The 720-bp fragment (Fig. 3A, lane e) protected by wild-type viral RNA represented a minor viral transcript(s) extending leftward from the edge of the LTR through the viral *Xho*I site.

Northern analyses indicated that the two other major FP-DS RNAs (1.68 and 3.50 kb) were complementary to DNA probes derived from viral sequences to the right of the LTR but not to a probe from the LTR itself (data not shown). To map these RNAs, a 1,025-bp *Pst*I-*Hin*dIII probe (5′ end labeled at *Hin*dIII) extending from the LTR to the rightmost *Hin*dIII site was hybridized to RNA from FP-DS-infected cells and treated with S1 nuclease (Fig. 3B). Surprisingly, FP-DS RNA protected the full-length (1,025-bp) probe (Fig. 3B, lane c) and indicated that the 1.68- and 3.50-kb RNAs were transcribed from left to right, opposite that of the 1.05- and 2.38-kb RNAs above. Moreover, these RNAs originated within the LTR from a position left of the *Pst*I site. Primer extension analysis confirmed this and mapped the 5′ ends of these RNAs to the nearest nucleotide (see below). That the two transcripts were not detected in Northern analyses using the LTR alone as a probe was attributed to the unusually high A-T content (75%) at the right end of the LTR from which these RNAs originated (see also below). As expected, wild-type RNA protected a 1,010-bp fragment (Fig. 3B, lane e) which represented the *Hin*dIII-K β and γ transcripts spanning the probe (excluding the LTR) from left to right. No detectable fragments were protected by RNA from mock-infected cells (Fig. 3B, lane d) or by FP-DS RNA when hybridized to a *Pst*I-*Hin*dIII probe 3′ end labeled at the *Hin*dIII site (data not shown).

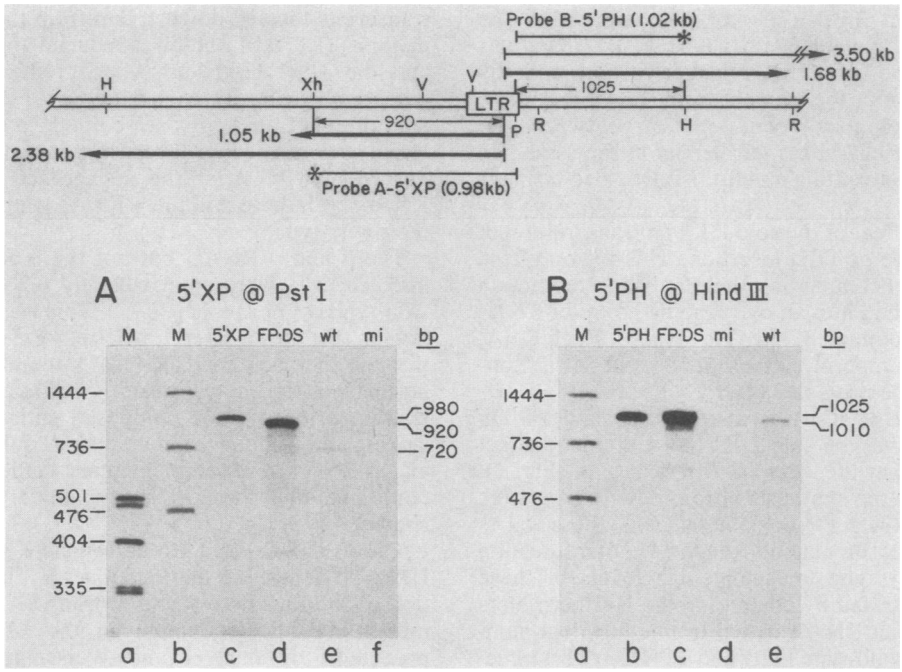

FIG. 3. S1 nuclease mapping of the 5′ ends of the FP-DS-induced RNA transcripts. Poly(A)$^+$ RNA from cells 12 h after infection with wild-type (wt) or FP-DS AcNPV or RNA from mock-infected (mi) cells (0.5 μg each) was hybridized to the 0.98-kb *Xho*I-*Pst*I fragment 5′ end labeled exclusively at the *Xho*I site (designated probe A; see panel A) or to the 1.02-kb *Pst*I-*Hin*dIII fragment 5′ end labeled at the *Hin*dIII site (designated probe B; see panel B). The locations of probes A and B are illustrated in the top diagram; the * denotes position of the 5′ end label. After treatment with S1 nuclease, the resistant fragments were denatured and resolved on a 4% acrylamide–7 M urea–TBE gel. Autoradiograms of 8-h (A) and 12-h (B) exposures are shown. Molecular weight standards (M) in bp and 5′ end labeled DNA probes (5′XP and 5′PH) are indicated.

The divergent FP-DS RNAs terminate within flanking viral sequences. The above data indicated that the new FP-DS RNAs extended in opposite directions from the LTR into flanking viral genes. To map the 3′ termini of the two RNAs extending leftward, FP-DS RNA was hybridized to a 1,740-bp *Hin*dIII-*Eco*RV probe (3′ end labeled at *Eco*RV) (Fig. 4A). Two major S1-resistant fragments of 595 and 1,740 bp were detected (Fig. 4A, lane c). The 595-bp fragment corresponded to the 3′ end of the 1.05-kb RNA. The 2.38-kb RNA protected the full-length probe (1,740 bp), and its size suggested that it terminated approximately 100 bp to the left of the *Hin*dIII site (Fig. 4, top diagram). RNA from wild-type-infected cells also generated 595- and 1,740-bp fragments as well as a new 395-bp fragment (Fig. 4A, lane e). These fragments, however, were present in amounts approximately 50-fold lower than those generated by equivalent amounts of FP-DS RNA. The 1.05-kb FP-DS RNA therefore terminated left of the viral *Xho*I site, apparently along with a minor wild-type RNA(s) also transcribed from right to left.

To map the 3′ termini of RNAs extending rightward from the LTR, FP-DS RNA was hybridized to a 1,590-bp *Hin*dIII-*Xho*I probe (3′ end labeled at *Hin*dIII) derived from the adjacent *Hin*dIII-Q restriction fragment (87.5 to 88.7 map units). This resulted in protection of a 540-bp fragment (Fig. 4B, lane e) identical in size to that protected by wild-type RNA (Fig. 4B, lane d) but in amounts approximately 10-fold higher. Thus the major 1.68-kb FP-DS RNA extending rightward from the LTR terminated along with the overlapping α, β, and γ viral RNAs. Longer exposures (not shown) revealed protection of a full-length 1,590-bp fragment corresponding to the minor 3.50-kb FP-DS transcript which terminated approximately 1,800 bp further downstream. Both the 1.68- and 3.50-kb FP-DS RNAs were detected by hybridization to the *Hin*dIII-Q restriction fragment in Northern blot analyses (data not shown).

Figure 5 summarizes the above data and illustrates the transcriptional pattern of the *Hin*dIII-K region of both wild-type (top) and FP-DS (bottom) viruses. Insertion of the solo TED LTR resulted in transcription of two new sets of abundant RNAs. These RNAs originated within the LTR but extended in opposite directions, two leftward (1.05L and 2.38L) and two rightward (1.68R and 3.50R). At least two RNAs (1.05L and 1.68R) terminated at poly(A) sites used by viral RNAs within the region. The exact correlation between the sizes of the S1-resistant hybrids and the corresponding RNAs [excluding the 100 or so bases comprising the 3′ poly(A) tracts] also indicated that little or no RNA splicing occurred.

The solo TED LTR directs divergent transcription from a single 22-bp sequence with dyad symmetry. Both sets of divergent FP-DS RNAs were initiated within a narrow region of the TED LTR. To determine the 5′ start sites to the nearest nucleotide, we employed the method of primer extension. Figure 6 illustrates the sequence of the solo TED LTR as inserted within the viral *Hin*dIII-K restriction fragment. The 5′ ends of the RNAs extending leftward were mapped by annealing FP-DS RNA to an *Eco*RV-*Alu*I primer (position −178 to −145 within the LTR; Fig. 6) 5′ end labeled at the *Eco*RV site. Extension with reverse transcriptase (Fig. 7A) yielded a single 176-base product (lane 5) which terminated with the thymidine at position −3. No extension products were detected when reverse transcriptase was incubated with tRNA alone (Fig. 7A, lane 6). The

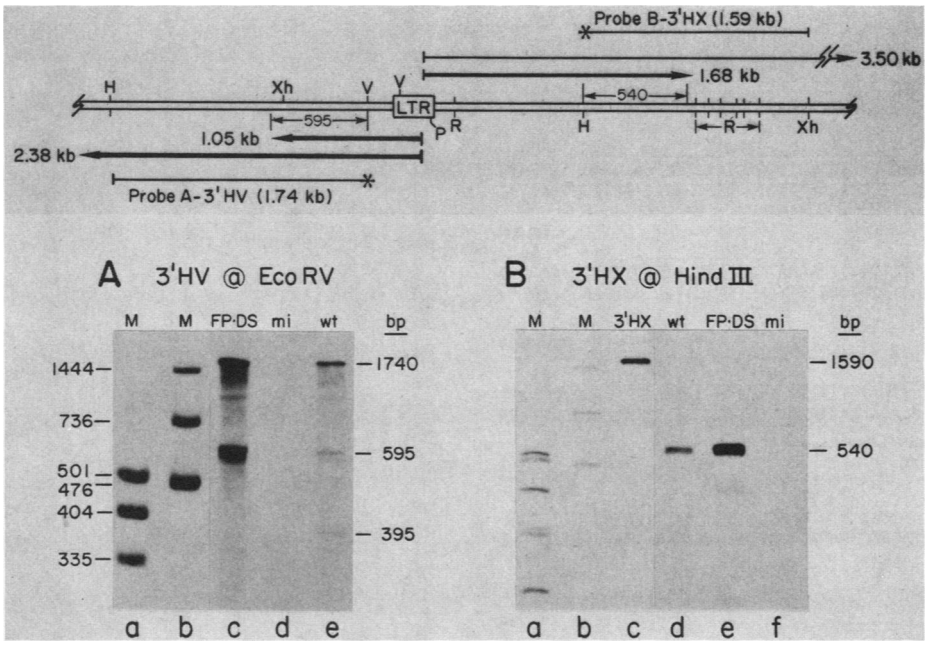

FIG. 4. S1 nuclease mapping of the 3′ ends of the FP-DS-induced RNA transcripts. Poly(A)-containing RNA from cells 12 h after infection with wild-type (wt) or FP-DS AcNPV (1.0 and 0.4 μg, respectively) and RNA (0.4 μg) from mock-infected (mi) cells were hybridized to the 1.74-kb *Hin*dIII-*Eco*RV fragment 3′ end labeled exclusively at the *Eco*RV (V) site (probe A; see panel A) or the 1.59-kb *Hin*dIII-*Xho*I fragment 3′ end labeled exclusively at the *Hin*dIII site (probe B; see panel B). The * denotes positions of the 3′ end label for each probe (top). S1-resistant fragments were resolved on polyacrylamide gels as described in the legend of Fig. 3. Autoradiograms of 24-h (A) and 12-h (B) exposures are shown. Molecular weight standards (lanes a and b) are indicated in bp.

generation of a single extension product with FP-DS RNA indicated that the 5′ ends of the leftward 1.05L and 2.38L RNAs were coterminal beginning with the adenosine at position −3 (Fig. 7C).

The same approach was used to map the 5′ start sites of the RNAs (1.68R and 3.50R) extending rightward from the LTR. A *Taq*I-*Bst*EII primer located to the immediate right of the LTR (position +80 to +153) 5′ end labeled at the *Bst*EII site yielded a single product (153 nucleotides) after hybridization to FP-DS RNA and extension with reverse transcriptase (Fig. 7B, lane 5). The terminus of this extended primer corresponded to the thymidine of position +1. Thus both rightward RNAs (1.68R and 3.50R) were coterminal and began with the adenosine at position +1.

These data demonstrated that both sets of TED LTR RNAs initiated within two nucleotides of one another but extended in opposite directions (Fig. 7C). Examination of the nucleotide sequence of this novel transcriptional arrangement revealed that the 5′ start sites (+1 and −3) were symmetrically positioned within a 22-bp element (see box) having partial (hyphenated) twofold symmetry. The element, including the polyadenylation signal (AATAAA), possesses

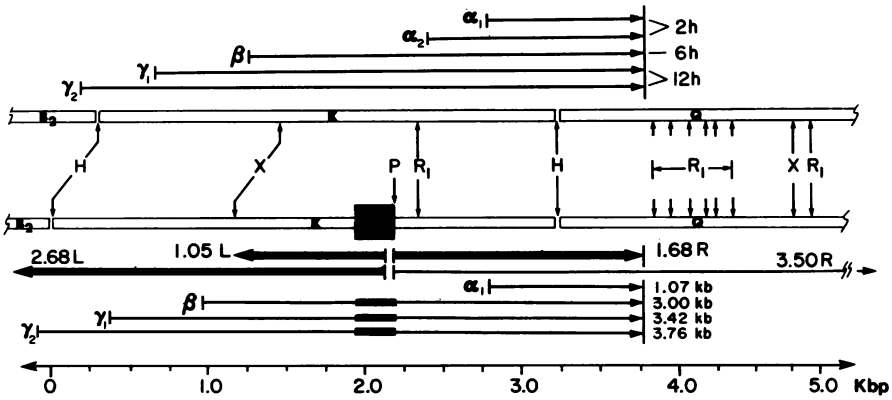

FIG. 5. Transcription of the *Hin*dIII-K region of wild-type AcNPV (top) and insertion mutant FP-DS (bottom). The wild-type *Hin*dIII-K fragment is transcribed into five overlapping RNAs (α_1, α_2 β, γ_1, and γ_2), all terminated at a common 3′ site located within the adjacent *Hin*dIII-Q fragment as shown. Each RNA is transcribed from a unique promoter (14). The presence of the solo TED LTR (270 bp, shaded box) in mutant FP-DS resulted in transcription of four abundant RNAs (bold arrows), two extending to the left (1.05L and 2.38L) and two to the right (1.68R and 3.50R) from the LTR. FP-DS viral RNAs β (3.00 kb), γ_1 (3.42 kb), and γ_2 (3.76 kb) spanned the LTR and were 270 nucleotides larger than their wild-type counterparts (top). Scale at bottom indicates distance in kilobase pairs (kbp). H, *Hin*dIII; X, *Xho*I; P, *Pst*I; R_1, *Eco*RI.

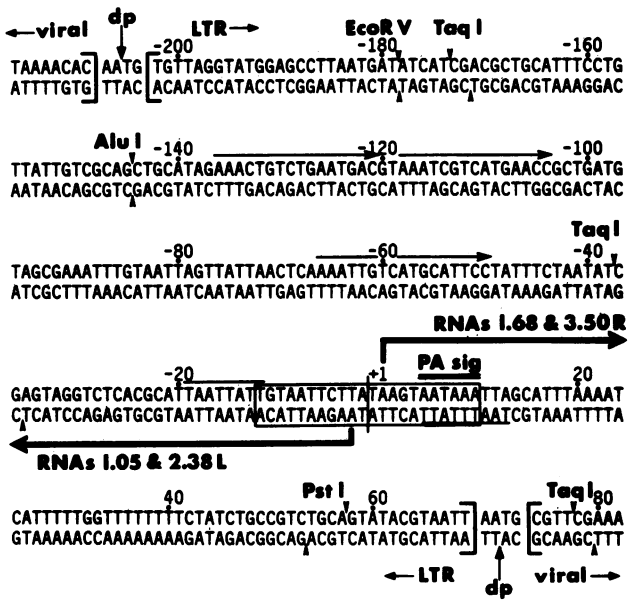

FIG. 6. Nucleotide sequence of the solo TED LTR of FP-DS. The sequence (270 nucleotides) of both DNA strands of the solo TED LTR, along with a small portion of flanking *Hin*dIII-K viral sequences, is shown. Brackets designate the ends of the LTR and illustrate the four viral nucleotides duplicated upon element insertion. The sequence is numbered from the transcriptional start site (+1) of the RNAs extending rightward (1.68R and 3.50R) as determined by primer extension analysis (see Fig. 7). These RNAs, as well as the major RNAs extending leftward (1.05L and 2.68L), originate within a 22-bp region (see box, positions −12 through +10) having partial twofold (dyad) symmetry. Sequences of potential TATA boxes (overlined) and the poly(A) signal (double line, position +5) are indicated. Thin arrows designate 16-bp tandem repeats (imperfect) which bear partial homology with the core sequence of other viral enhancers (50). Nucleotide sequences, including those of the wild-type *Hin*dIII-K fragment at the site of integration, were determined for both DNA strands using the dideoxy-chain termination method of Sanger et al. (40).

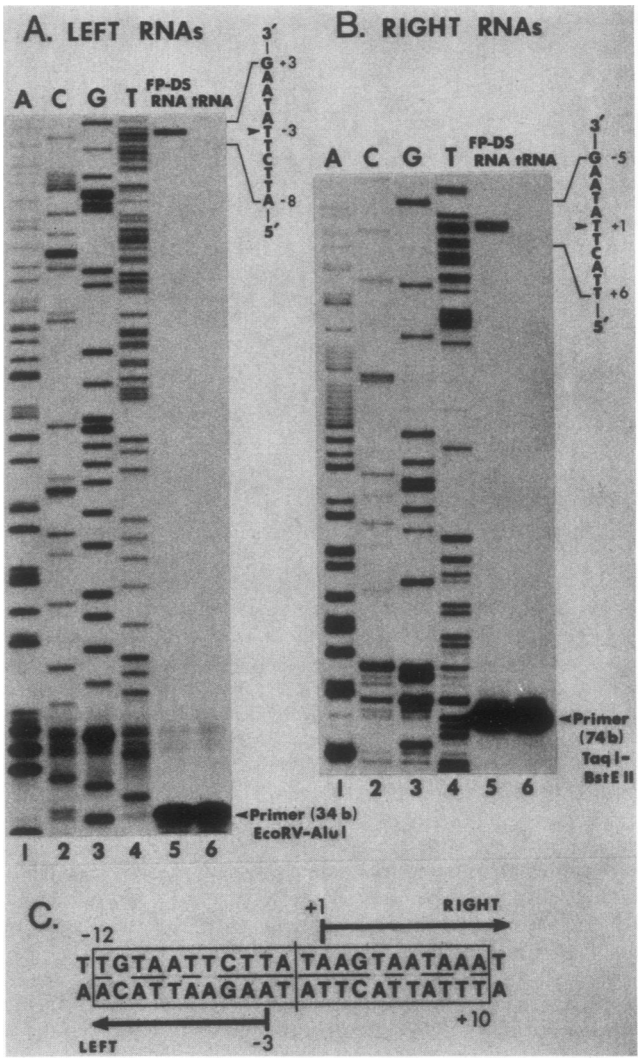

FIG. 7. Primer extension mapping of the 5′ ends of RNAs extending leftward (A) or rightward (B) from the solo LTR of FP-DS. A 34-bp *Eco*RV-*Alu*I primer (position −178 to −145, see Fig. 6) 5′ end labeled exclusively at the *Eco*RV site (A), or a 74-bp *Taq*I-*Bst*EII primer (position +80 to +153) 5′ end labeled at the *Bst*EII site (B), was annealed to FP-DS RNA. The primers were extended with reverse transcriptase, denatured, and fractionated on 6% polyacrylamide–8 M urea gels. The locations of the 5′ ends of the leftward and rightward RNAs were determined by comparing the positions of the extended primers with the pattern generated by annealing the same primers to appropriate M13 clones of the LTR, followed by extension using the dideoxy-chain termination procedure (see Materials and Methods). Lanes 5 and 6, respectively, show primers extended in the presence (FP-DS) or absence (tRNA) of FP-DS RNA; lanes 1 through 4 indicate A, C, G, and T sequencing reactions, respectively. (C) Positions of the symmetrical 5′ termini of both sets of divergent RNAs (−3 leftward, +1 rightward) within the 22-nucleotide sequence with hyphenated dyad symmetry (see box). Symmetrical nucleotides are underlined.

8 of 11 nucleotides on each side of the central axis which are symmetrical with respect to distance and identity. The TED LTR represents the first example of a eucaryotic element capable of directing transcription of divergent RNAs from a single initiation site.

DISCUSSION

TED is a Lepidopteran retrotransposon. The TED element, repeated about 50 times in the Lepidopteran (*T. ni*) genome, has many features diagnostic of a eucaryotic retrotransposon. It is flanked by two LTRs (270 bp) which contain imperfect inverted repeats (7 of 10 nucleotides) at their termini (Fig. 8). Each LTR starts with the trinucleotide TGT, as do the LTRs of copia, Ty, and the integrated proviral forms of many retroviruses (44), but ends with ATT similar only to the *Drosophila* elements 17.6 and gypsy (Fig. 8). Analogous to the latter elements, TED transposition resulted in duplication of four nucleotides (AATG) at the site of insertion; the presence of small direct repeats (three to six nucleotides) adjacent to the ends of transposable elements is characteristic of the insertional event (4).

TED also possesses a putative tRNA binding site for the initiation of reverse transcription (Fig. 8). The 18-nucleotide sequence is identical to the primer binding site of element 297, which is in turn complementary to the 3′ end of *Drosophila* $tRNA^{Ser}$ (38). Unlike the retroproviral primer binding site which lies immediately adjacent to the 5′ LTR, the putative primer binding sites of TED, 297, 17.6, and gypsy overlap the LTR by one nucleotide (Fig. 8). TED also contains a purine-rich region immediately adjacent to the rightmost LTR (data not shown) which may act to initiate second-strand DNA synthesis. Morever, a full-length RNA (7 kb) was transcribed from one LTR to the other when TED

ELEMENT	LTR TERMINI 5' END	3' END	5' END of INTERNAL DOMAIN (tRNA Binding Site)
TED	TGTTAGGTAT.....ATACGTAATT		T GGCGCAGTCGGTAGGAT....
297	GTGACGTATT.....CAATTTTACT		T GGCGCAGTCGGTAGGAT....
17.6	AGTGACATAT.....ATTTGCAATT		T GGCGCAGTCGATGTGAT....
gypsy	AGTTAACAAC.....TTATATAATT		T GGCGCCCAACCAACAAT....
copia	TGTTGGAATA.....AATTACAACA		A GGTTATGGGCCCAGTCCA...
Ty912	TGTTGGAATA.....CCATTTCTCA		A TGGTAGCGCCTGTGCTTC...
	→LTR	LTR←	5'LTR←

FIG. 8. Comparison of the nucleotide sequence of the TED LTR with corresponding sequences of other retrotransposons. The 5' and 3' termini of the TED LTR with indicated boundaries are listed along with those of five retrotransposons from *Drosophila* or *S. cerevisiae*. The putative primer (tRNA) binding site of TED, located immediately adjacent to the leftmost LTR, is also compared. The sequence TGG complementary to the 3' end of most tRNAs (CCA) is underlined. Nucleotide sequences shown are of the copia-like elements 297 (19), 17.6 (38), and gypsy bx^{34e} (13) (which is identical to mdg4 [2]) and copia (9, 29) as well as the *S. cerevisiae* Ty element Ty912 (5).

was integrated within the host chromosome or the FP-DL viral genome (unpublished data). Structural analogy with the retroproviruses predicts that transcription is from the left (5') to the right (3') LTR.

Pending identification of a reverse transcriptase-like (*pol*) gene, we have concluded that TED is a Lepidopteran (moth) retrotransposon whose size (7.3 kb) and structural features place it within the subfamily of retrotransposons comprised of *Drosophila* elements 17.6 (7.43 kb) and 297 (6.5 kb), among others. As such, TED represents the first retrotransposon identified in a metazoan besides *Drosophila* (retroproviruses not included).

Effects of the solo TED LTR on AcNPV RNA transcription. To begin examination of the transcription of TED as integrated within the AcNPV genome, we first determined the effects of a single LTR-excision derivative on viral transcription. Figure 5 illustrates our data on the transcription occurring at the site of insertion of the solo LTR. Most prominent was the transcription of two sets of new RNA; both were initiated within the LTR but extended in opposite directions. The two RNAs extending rightward (1.68R and 3.50R) overlapped the immediate early viral RNAs, α_1 and α_2. Should these new RNAs represent functional mRNAs for α polypeptides, then the LTR promoted reactivation of the immediate early genes. The two RNAs extending leftward (1.05L and 2.38L) were complementary to the viral β and γ RNAs (Fig. 5). Consequently, these LTR RNAs possess the capacity to hybridize to viral mRNA and thereby block expression at the level of translation in a manner similar to that proposed for antisense RNA (20, 33). By directing symmetric transcription of adjacent viral genes, the solo TED LTR represents a mobile promoter element with the potential to augment or repress gene expression irrespective of its orientation.

Although inserted in the middle of a group of overlapping viral RNAs (Fig. 5), the solo LTR had little if any effect on the temporal specificity of viral transcription. Each temporal class of viral RNA (α, β, and γ) was transcribed during the normal wild-type periods. The LTR did, however, alter the levels of these RNAs (Fig. 2). The immediate early α_1 and α_2 RNAs were dramatically reduced relative to the later β and γ RNAs. This reduction occurred before the transcription of the LTR-induced RNAs, suggesting that the LTR (inserted 150 nucleotides upstream from the 5' cap of the α_2 RNA) disrupted sequences, chromatin structure, or both, regulating the efficiency of transcription from the immediate early promoters.

The LTR sequences were not spliced out of the poly(A)$^+$ β and γ RNAs which spanned the LTR and terminated at their normal position (Fig. 5). Thus, the LTR insertion may have disrupted coding regions, thereby inactivating viral polypeptides. Interestingly, no new RNAs were detected which terminated within the LTR. This suggested that the poly(A) signal located at position +5 and followed shortly by several CA dinucleotides (11) was not utilized; we cannot rule out, however, the possibility that LTR-terminated RNAs were unstable and therefore not detected here. In a similar situation, insertion of a copia element within an intron of the white gene of *Drosophila* resulted in termination of white RNAs within or near the 3' LTR but not the 5' LTR or a solo LTR of the element (25, 54). A solo LTR and the 5' LTR from Moloney murine leukemia virus also failed to direct termination or polyadenylation of transcripts when inserted into a provirus of Rous sarcoma virus (45). It was suggested that the local context strongly influences LTR termination behavior.

We have yet to determine which, if any, of the above mechanisms for altered gene expression is responsible for the observed reduction in the number of viral occlusions produced by insertion mutants FP-D and FP-DS. Occlusion body production is affected by mutations at several locations within the baculovirus genome (12, 34; Miller, in press).

Possible role of DNA symmetry in transcriptional initiation. A surprising finding in our analyses was that the solo TED LTR directed transcription of divergent RNAs originating within a 22-bp sequence with dyad symmetry (Fig. 7). While there exist several eucaryotic genes which are divergently transcribed from a common promoter region, including mouse dihydrofolate reducatase (6), *S. cerevisiae* galactose (21), and simian virus 40 early and late genes (15), the TED LTR contains one of the first examples of a promoter capable of directing divergent transcription from a single point.

The symmetrical arrangement of the 5' ends of the divergent RNAs within the 22-bp TED element (+1 and −3, respectively) suggested that one function of the dyad symmetry is to establish the precise start site for transcription. Various transcription regulatory factors have been shown to interact at symmetrical DNA sequences in procaryotes, adenovirus, and *D. melanogaster* (31, 41, 52). Moreover, several retrotransposon LTRs possess symmetrical sequences at the site of transcription initiation (Fig. 9). The major RNA transcript of Ty912, for example, initiates within a 22-bp sequence with hyphenated dyad symmetry (centered 232 bp from the left edge of the LTR). The copia LTR possesses a similar element (centered 113 bases from the left edge); the lower degree of symmetry of the copia sequence might explain the observed heterogeneity in RNA start sites (TTC) located within and slightly downstream from the element. Interestingly, the LTR of the retrovirus human T-lymphotropic virus type III/lymphadenopathy virus contains several symmetrical elements, one positioned at the RNA start site (Fig. 9). Viral transcription initiates from this same site despite deletion of nearby TATA sequences (position −22), suggesting that proper initiation is more dependent on sequences at the start site than on the TATA box (36).

A major question posed by the observed bidirectional transcription of the solo TED LTR is whether the same

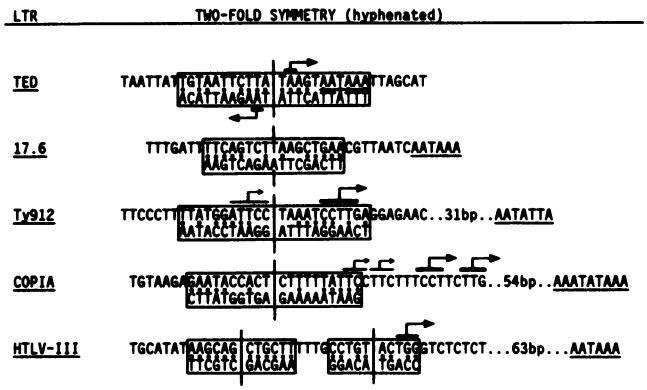

FIG. 9. Regions of twofold symmetry at the transcriptional initiation sites of several retrotransposon LTRs. The rectangular boxes surrounding each vertical axis designate nucleotide sequences with hyphenated twofold symmetry; symmetrical nucleotide pairs are indicated. Arrows depict the location of the 5′ start site and the direction of transcription of each RNA. Poly(A) signals are underlined. The exact start site of the *D. melanogaster* 17.6 RNA has not been determined; however, sequence homology with the retrovirus avian leukemia-sarcoma virus suggests that RNA initiates within the region shown (23). LTR sequences are those of the *Drosophila* retrotransposons 17.6 (23, 38) and copia (9, 29), the *S. cerevisiae* retrotransposon Ty912 (5, 8), and the retrovirus human T-lymphotropic virus type III/lymphadenopathy virus (30, 39, 42, 48).

promoter(s) is used by the intact transposon for directing symmetrical transcription from its 5′ and 3′ LTRs. The 3′ LTR, for instance, should have the capacity to direct transcription of an RNA extending in the direction opposite that of the major (genomic) RNA originating within the 5′ LTR. Transcriptional overlap interference, occurring when transcription of full-length RNA from one LTR inhibits initiation from the other (7), may play a role in regulating such antisense transcription. Northern blot analysis has indicated that, besides the full-length TED RNA, smaller RNAs complementary to the 3′ internal portion of TED are transcribed (unpublished data). We are in the process of determining whether these RNAs are initiated from the 3′ LTR or are derived via splicing of a precursor transcribed from the 5′ LTR.

ACKNOWLEDGMENTS

We thank Martin Hodge for excellent technical assistance.

This work was supported in part by Public Health Service grant AI17338 from the National Institute of Allergy and Infectious Disease and by a grant from the University of Idaho Biomedical Research Support Grant Program (BRSG S07 RR07170).

LITERATURE CITED

1. **Baltimore, D.** 1985. Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome. Cell **40**:481–482.
2. **Bayev, A. A., N. V. Lyubomirskaya, E. B. Dzhumagaliev, E. V. Ananiev, I. G. Amiantova, and Y. V. Ilyin.** 1984. Structural organization of transposable element mdg4 from *Drosophila melanogaster* and a nucleotide sequence of its long terminal repeats. Nucleic Acids Res. **12**:3707–3723.
3. **Boeke, J. D., D. J. Garfinkel, C. A. Styles, and G. R. Fink.** 1985. Ty elements transpose through an RNA intermediate. Cell **40**:491–500.
4. **Calos, M. P., and J. H. Miller.** 1980. Transposable elements. Cell **20**:579–595.
5. **Clare, J., and P. Farabaugh.** 1985. Nucleotide sequence of a yeast Ty element: evidence for an unusual mechanism of gene expression. Proc. Natl. Acad. Sci. USA **82**:2829–2833.
6. **Crouse, G. F., E. J. Leys, R. N. McEwan, E. G. Frayne, and R. E. Kellens.** 1985. Analysis of the mouse dhfr promoter region: existence of a divergently transcribed gene. Mol. Cell. Biol. **5**:1847–1858.
7. **Cullen, B. R., P. T. Lomedico, and G. Ju.** 1984. Transcriptional interference in avian retroviruses: implications for the promoter insertion model of leukaemogenesis. Nature (London) **307**: 241–245.
8. **Elder, R. T., E. Y. Loh, and R. W. Davis.** 1983. RNA from the yeast transposable element Tyl has both ends in the direct repeats, a structure similar to retrovirus RNA. Proc. Natl. Acad. Sci. USA **80**:2432–2436.
9. **Emori, Y., T. Shiba, S. Kanaya, S. Inouye, S. Yuki, and K. Saigo.** 1985. The nucleotide sequences of copia and copia-related RNA in *Drosophila* virus-like particles. Nature (London) **315**:773–776.
10. **Faulkner, P.** 1981. Baculovirus, p. 3–37. *In* E. W. Davidson (ed.), Pathogenesis of invertebrate microbial diseases. Allanheld Osmun Publ., Totowa, N.J.
11. **Fitzgerald, M., and T. Shenk.** 1981. The sequence 5′-AAUAAA-3′ forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell **24**:251–260.
12. **Fraser, M. J., G. E. Smith, and M. D. Summers.** 1983. Acquisition of host cell DNA sequences by baculoviruses: relationship between host DNA insertions and FP mutants of *Autographa californica* and *Galleria mellonella* nuclear polyhedrosis viruses. J. Virol. **47**:287–300.
13. **Freund, R., and M. Meselson.** 1984. Long terminal repeat nucleotide sequence and specific insertion of the gypsy transposon. Proc. Natl. Acad. Sci. USA **81**:4462–4464.
14. **Friesen, P. D., and L. K. Miller.** 1985. Temporal regulation of baculovirus RNA: overlapping early and late transcripts. J. Virol. **54**:392–400.
15. **Gidoni, D., J. T. Kadonaga, H. Barrera-Saldaña, K. Takahashi, P. Chambon, and R. Tjian.** 1985. Bidirectional SV40 transcription mediated by tandem Spl binding interactions. Science **230**:511–517.
16. **Hink, W. F .** 1970. Established insect cell line from the cabbage looper, *Trichoplusia ni*. Nature (London) **225**:466–467.
17. **Holmes, D. S., and M. Quigley.** 1981. A rapid method for the preparation of bacterial plasmids. Anal. Biochem. **114**:193–197.
18. **Hruby, D. E., and W. K. Roberts.** 1976. Encephalomyocarditis virus RNA: variations in polyadenylic acid content and biological activity. J. Virol. **19**:325–330.
19. **Ikenaga, H., and K. Saigo.** 1982. Insertion of a movable genetic element, 297, into the T-A-T-A box for the H3 histone gene in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **79**: 4143–4147.
20. **Izant, J. G., and H. Weintraub.** 1985. Constitutive and conditional suppression of exogenous and endogenous genes by anti-sense RNA. Science **229**:345–352.
21. **Johnston, M., and R. W. Davis.** 1984. Sequences that regulate the divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:1440–1448.
22. **Kelly, D. C.** 1982. Baculovirus replication. J. Gen. Virol. **63**:1–13.
23. **Kugimiya, W., H. Ikenaga, and K. Saigo.** 1983. Close relationship between the long terminal repeats of avian leukosis-sarcoma virus and copia-like movable genetic elements of *Drosophila*. Proc. Natl. Acad. Sci. USA **80**:3193–3197.
24. **Lee, H. H., and L. K. Miller.** 1978. Isolation of genotypic variants of *Autographa californica* nuclear polyhedrosis virus. J. Virol. **27**:754–767.
25. **Levis, R., K. O'Hare, and G. M. Rubin.** 1984. Effects of transposable element insertions on RNA encoded by the white gene of *Drosophila*. Cell **38**:471–481.
26. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **McMaster, G. K., and G. G. Carmichael.** 1977. Analysis of single- and double-stranded nucleic acids on polyacrylamide

and agarose gels by using glyoxal and acridine orange. Proc. Natl. Acad. Sci. USA 74:4835–4838.
28. **Miller, D. W., and L. K. Miller.** 1982. A virus mutant with an insertion of a copia-like transposable element. Nature (London) 299:562–564.
29. **Mount, S. M., and G. M. Rubin.** 1985. Complete nucleotide sequence of the *Drosophila* transposable element copia: homology between copia and retroviral proteins. Mol Cell Biol. 5:1630–1638.
30. **Muesing, M. A., D. H. Smith, C. D. Cabradilla, C. V. Benton, L. A. Lasky, and D. J. Capon.** 1985. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. Nature (London) 313:450–457.
31. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. 53:293–321.
32. **Parkhurst, S. M., and V. G. Corces.** 1985. *Forked*, gypsys and suppressors in *Drosophila*. Cell 41:429–437.
33. **Pestka, S., B. L. Daugherty, V. Jung, K. Hotta, and R. K. Pestka.** 1984. Anti-mRNA: specific inhibition of translation of single mRNA molecules. Proc. Natl. Acad. Sci. USA 81: 7525–7528.
34. **Potter, K. N., and L. K. Miller.** 1980. Correlating genetic mutations of a baculovirus with the physical map of the DNA genome, p. 71–80. *In* B. N. Fields, R. Jarnisch, and C. F. Fox (ed.), Animal virus genetics. Academic Press, Inc., New York.
35. **Roeder, G. S., and G. R. Fink.** 1983. Transposable elements in yeast, p. 299–328. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York.
36. **Rosen, C. A., J. S. Sodroski, and W. A. Haseltine.** 1985. The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. Cell 41:813–823.
37. **Rubin, G. M.** 1983. Dispersed repetitive DNA in *Drosophila*, p. 329–361. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York
38. **Saigo, K., W. Kugimiya, Y. Mutasuo, S. Inouye, K. Yoshioka, and S. Yuki.** 1984. Identification of the coding sequence for a reverse transcriptase-like enzyme in a transposable genetic element in *Drosophila melanogaster*. Nature (London) 312: 659–661.
39. **Sanchez-Pescador, R., M. D. Power, P. J. Barr, K. S. Steimer, M. M. Stempien, S. L. Brown-Shimer, W. W. Gee, A. Renard, A. Randolf, J. A. Levy, D. Dina, and P. A. Luciuw.** 1985. Nucleotide sequence and expression of the AIDS-associated retrovirus (ARV-2). Science 227:484–492.
40. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.
41. **Sawadogo, M., and R. G. Roeder.** 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell 43:164–175.
42. **Starcich, B., L. Ratner, S. F. Joseph, T. Okamoto, R. C. Gallo, and F. Wong-Staal.** 1985. Characterization of the long terminal repeat sequences of HTLV-III. Science 227:538–540.
43. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA 77:5201–5205.
44. **Varmus, H. E.** 1983. Retroviruses, p. 411–503. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York.
45. **Varmus, H. E., N. Quintrell, and S. Ortiz.** 1981. Retroviruses as mutagens: insertion and excision of a nontransforming provirus alter expression of a resident transforming provirus. Cell 25:23–26.
46. **Vaughn, J. L., R. H. Goodwin, G. L. Thompkins, and P. McCawley.** 1977. Establishment of two insect cell lines from the insect *Spodoptera frugiperda* (Lepidoptera: Noctuidae). In Vitro 13:213–217.
47. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with universal primers. Gene 19:259–268.
48. **Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Allzon.** 1985. Nucleotide sequence of the AIDS virus, LAV. Cell 40:9–17.
49. **Weaver, R. F., and C. Weissman.** 1979. Mapping of RNA by a modifiction of the Berk-Sharp procedure: the 5′ terminus of 15S β-globin mRNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucleic Acids Res. 7:1175–1193.
50. **Weiher, H., M. König, and P. Gruss.** 1983. Multiple point mutations affecting the simian virus 40 enhancer. Science 219:626–631.
51. **Winston, F., K. J. Durbin, and G. R. Fink.** 1984. The SPT3 gene is required for normal transcription of Ty elements in *S. cerevisiae*. Cell 39:675–682.
52. **Wu, C.** 1984. Activating protein factor binds in vitro to upstream control sequences in heat shock gene chromatin. Nature (London) 311:81–84.
53. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119.
54. **Zachar, Z., D. Davison, D. Garza, and P. M. Bingham.** 1985. A detailed developmental and structural study of the transcriptional effects of the insertion of the *copia* transposon into the white locus of *Drosophila melanogaster*. Genetics 111:495–515.