The Baculovirus-Integrated Retrotransposon TED Encodes *gag* and *pol* Proteins That Assemble into Viruslike Particles with Reverse Transcriptase

ROBERT A. LERCH AND PAUL D. FRIESEN*

Institute for Molecular Virology and Department of Biochemistry, Graduate School and College of Agricultural and Life Sciences, University of Wisconsin—Madison, Madison, Wisconsin 53706-1596

Received 4 November 1991/Accepted 15 December 1991

TED is a lepidopteran retrotransposon found inserted within the DNA genome of the *Autographa californica* nuclear polyhedrosis virus mutant, FP-D. To examine the proteins and functions encoded by this representative of the gypsy family of retrotransposons, the *gag*- and *pol*-like open reading frames (ORFs 1 and 2) were expressed in homologous lepidopteran cells by using recombinant baculovirus vectors. Expression of ORF 1 resulted in synthesis of an abundant TED-specific protein (Pr55^{gag}) that assembled into viruslike particles with a diameter of 55 to 60 nm. Expression of ORF 2, requiring a –1 translational frameshift, resulted in synthesis of a protease that mediated cleavage of Pr55^{gag} to generate p37, the major protein component of the resulting particles. Expression of ORF 2 also produced reverse transcriptase that associated with these particles. Both protease and reverse transcriptase activities mapped to domains within ORF 2 that contain sequence similarities with the corresponding functional domains of the *pol* gene of the vertebrate retroviruses. These results indicated that TED ORFs 1 and 2 functionally resemble the retrovirus *gag* and *pol* genes and demonstrated for the first time that an invertebrate member of the gypsy family of elements encodes active forms of the structural and enzymatic functions necessary for transposition via an RNA intermediate. TED integration within the baculovirus genome thus represents one of the first examples of transposon-mediated transfer of host-derived genes to an eukaryotic virus.

Of the known DNA animal viruses, the baculoviruses are unique in their ability to stably accommodate the spontaneous insertion of host-derived transposons during replication. Such events provide genetic diversity and may represent a significant force in virus evolution (for a recent review, see reference 5). Several different classes of mobile elements in lepidopterans (moths) (3, 8, 16, 45) have been identified as host insertions within the ~128-kb DNA genome of *Autographa californica* nuclear polyhedrosis virus (AcMNPV), the prototype of subgroup A baculoviruses. The largest of these is TED (transposable element D), a 7.5-kb host element that transposed from the genome of the nocturnal moth *Trichoplusia ni* (Lepidoptera; Noctuidae) to the AcMNPV genome during infection. TED is present as a single copy (TED^{FP-D}) within the AcMNPV mutant FP-D that was identified by its altered plaque morphology, distinguished by the presence of fewer virus polyhedra (31).

TED is a lepidopteran member of the gypsy family of retrotransposons (16), a distinct group of mobile elements that bear a striking resemblance to the retroviruses in structure and genetic organization. These transposons (including *Drosophila melanogaster* elements gypsy, 17.6, 297, and 412, and the more distantly related Ty3 element from *Saccharomyces cerevisiae*) possess long terminal repeats (LTRs) and internal open reading frames (ORFs) that are positioned in a manner analogous to that of the retroviral *gag* and *pol* genes (for reviews, see references 4, 6, and 12). The second or *pol*-like ORF, exemplified by TED (Fig. 1A), is the most highly conserved among gypsy members and includes regions that are markedly similar to the protease (PR), reverse transcriptase (RT), RNase H, and integrase domains

of the retrovirus *pol* gene (for retrovirus reviews, see references 9 and 43). Characteristic of the gypsy elements, the order of these *pol*-like domains is identical to that of the retroviruses. TED, 17.6, 297, and gypsy also possess a third ORF with a size and position analogous to those of the retroviral *env* gene.

Although the gypsy family represents one of the largest groups of invertebrate transposons, little is known about the proteins and functions encoded by these retroidlike elements. Of the retrotransposons, Ty1 from *S. cerevisiae* is best understood (for reviews, see references 6 and 24). Distinguished from the gypsy-related elements by a different order of *pol* domains and the lack of a third (*env*-like) ORF, Ty1 encodes proteins necessary for transposition via an RNA intermediate. The first ORF (*TYA*) encodes the polyprotein TYA that assembles into viruslike particles (VLPs) and is therefore analogous to the retroviral *gag* polyprotein (1). Each of the *pol*-like domains (PR, integrase, RT, and RNase H, respectively) within the second ORF (*TYB*) is required for Ty1 transposition (1, 13, 18, 48). The *TYB* gene products, including RT, are packaged into Ty-VLPs. Since Ty1-VLPs also contain Ty1 genomic RNA and various forms of Ty1 DNA, it is likely that formation of VLPs is required for transposition (7, 13, 14, 18, 30). The relative importance of *gag* proteins and VLPs for transposition is further suggested by the observation that no integration-competent retroid elements have been found without a *gag* gene (47).

The spontaneous integration of TED into the AcMNPV genome provides an opportunity to examine the gene functions of the gypsy family of retrotransposons and to determine whether TED insertion results in the acquisition of new host-derived genes by the baculovirus. In this report, we identify proteins and functions encoded by ORFs 1 and 2 of the virus-integrated element, TED^{FP-D}. To distinguish trans-

* Corresponding author.

poson-encoded proteins and facilitate high-level expression, we employed recombinant baculovirus vectors in which the 5' LTR of TED^{FP-D} was replaced by the strong promoter for the AcMNPV polyhedrin gene. TED proteins were subsequently identified after infection of cultured lepidopteran cells with specific virus vectors. By analyzing the expression of progressively longer sections of the internal portion of TED, proteins and their functions were mapped to their corresponding positions within the element. We report here that expression of TED ORFs 1 and 2 resulted in the production of VLPs that contain TED-specific RT. Thus, ORFs 1 and 2 (referred to hereafter as *gag* and *pol*, respectively) encode functions analogous to those of the retrovirus *gag* and *pol* genes. Our findings support a retroviruslike model for TED expression that includes the synthesis of essential functions for transposition. Therefore, by the criteria examined thus far, TED^{FP-D} may be competent for transposition from the AcMNPV genome.

MATERIALS AND METHODS

Cells and viruses. Established lepidopteran (moth) cell lines, *T. ni* TN368 (21) and *Spodoptera frugiperda* IPLB-SF21 (SF21) (44) were propagated in TC100 growth medium (GIBCO Laboratories) supplemented with 10% fetal bovine serum and 2.6 mg of tryptose broth per ml as previously described (15). Viruses used included the wild-type L-1 strain of AcMNPV (27), the AcMNPV few-polyhedra (FP) mutant FP-D that carries a single integrated copy of TED (31, 35), and an AcMNPV deletion mutant (Δ POLY) that lacks the polyhedrin gene and its promoter (kindly provided by Lois Miller [University of Georgia]).

Recombinant plasmids. TED^{FP-D} from AcMNPV insertion mutant FP-D was cloned by inserting an 8.2-kb *XhoI*-*AsuII* fragment of viral DNA into the *XhoI* and *NotI* sites of the pBluescript (KS) vector (Stratagene) in which the *AsuII* and *NotI* ends were previously repaired with a Klenow fragment. The resulting plasmid, pTED/XA, contained the full-length TED element and 715 bp of viral DNA flanking the 5' LTR. To facilitate construction of transplacement plasmids for generating recombinant viruses, DNA fragments of TED that contained progressively longer sections of TED were first subcloned into pBluescript; each plasmid was designated according to the functional portion of TED it contained. Plasmid pGAG_{tr} contained all but 186 bp of the 3' end of the *gag* ORF and was constructed by adding an *XhoI* linker (5'-CCCTCGAGGG-3') to the *SmaI* end of the 1,151-bp *SmaI*-*SalI* fragment of TED (nucleotides [nt] 558 to 1706) and then inserting it into the *XhoI* and *SalI* sites of pBluescript. Plasmid pGAG contained the entire *gag* ORF (nt 558 to 1892) and was generated by inserting a *SalI*-*Sau3AI* fragment of TED (nt 1706 to 1944) into the *SalI* and *BamHI* sites of pGAG_{tr}. Similarly, plasmids pGAG/PR, pGAG/PR/RT, and pGAG/POL were generated by inserting progressively longer fragments of TED extending from the same *SalI* site (nt 1706) to the *SstI* (nt 2829), *XbaI* (nt 4707), or *HindIII* (nt 6075) site, respectively, into *SalI* and corresponding restriction sites of pGAG_{tr}. Transplacement plasmids in which TED sequences were fused to the polyhedrin promoter were constructed by inserting the *XhoI*-*SstII* fragments of the above plasmids into the corresponding *XhoI* and *SstII* sites of a modified form of the transplacement vector pEV55. pEV55 (32) was modified by removing *SmaI* and *PstI* sites and enlarging the polylinker to include unique restriction sites for *BglII*, *XhoI*, *EcoRI*, *PstI*, *SmaI*, *SpeI*, *XbaI*, *SstII*, and *KpnI*, respectively.

Construction of AcMNPV recombinants. Recombinant viruses were generated and propagated by established methods (reviewed by references 32 and 46). Briefly, 10 μ g of transplacement plasmid was mixed with 1 μ g of wild-type L-1 AcMNPV DNA and 30 μ g of Lipofectin (Bethesda Research Laboratories), and then added to SF21 cell monolayers. Recombinant viruses, harvested 4 days later, were identified by screening viral plaques for the occlusion-negative phenotype. After two additional rounds of plaque purification, viruses were analyzed for the insertion of TED-specific sequences into the polyhedrin locus by restriction mapping of isolated viral DNA (data not shown).

Radiolabeling of TED-specific proteins. SF21 or TN368 cell monolayers were inoculated with viruses, using a multiplicity of 10 PFU per cell, covered with growth medium, and incubated at 27°C. For pulse-label experiments, the medium was removed at 39 h postinfection and replaced with phosphate-buffered saline (PBS) (24) containing Trans³⁵S-Label (200 μ Ci; >1,000 Ci/mmol, methionine $\geq$ 70%, cysteine $\leq$ 15%; ICN Biomedicals, Incorp.) per ml. After a 1-h incubation at 27°C, the cells were dislodged, collected by centrifugation, suspended in lysis buffer (1% sodium dodecyl sulfate [SDS], 2.5% β -mercaptoethanol), and heated for 3 min at 100°C. The cell lysates were subjected to SDS-polyacrylamide gel electrophoresis (25) and autoradiography or fluorography with Fluoro-Hance (Research Products International Corp.). Molecular sizes of radiolabeled proteins were estimated by comparing electrophoretic mobilities with ¹⁴C-labeled molecular size standards. For pulse-chase experiments, infected SF21 cells were radiolabeled as described above, except that the labeling period was limited to 20 min, after which the radiolabel was replaced with growth medium supplemented with a 50-fold-higher-than-normal level of unlabeled methionine and cysteine. Cells were harvested at the indicated times after removal of the radiolabel, and SDS lysates were analyzed as described above. The proteins of TED-specific VLPs were radiolabeled in a similar manner except that at 36 h postinfection the medium was replaced with methionine-deficient Grace's medium (GIBCO Laboratories) that was supplemented with 2% fetal bovine serum and contained Trans³⁵S-Label (100 μ Ci/ml). VLPs were harvested 12 h later as described below.

Sedimentation and electron microscopy of TED VLPs. At 48 h postinfection, 10⁷ SF21 cells were dislodged and collected by low-speed centrifugation. The cells were suspended in PBS to give a density of 10⁶ cells per ml, Triton X-100 was added to 0.1%, and the mixture was incubated on ice for 15 min. The lysate was clarified by low-speed centrifugation, and VLPs were concentrated by sedimentation through a 30% (wt/vol) sucrose cushion in H β M buffer (50 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES] [pH 7.0], 0.1% β -mercaptoethanol) for 2.5 h at 285,000 $\times$ g (4°C) as described previously (11). The pellet was suspended in H β M buffer containing 0.1% Tween 20 and layered on a 20 to 60% (wt/wt) discontinuous sucrose gradient in the same buffer. After centrifugation for at least 16 h at 100,000 $\times$ g (4°C), the gradient was fractionated and monitored for A₂₅₄. For electron microscopy, VLPs were prepared as described above except that H β M buffer was replaced with TNE buffer (10 mM Tris [pH 7.5], 0.15 M NaCl, 1 mM EDTA). Particles were concentrated from gradient fractions containing the peak of TED protein by centrifugation, suspended in TNE buffer, and deposited on 400-mesh, carbon-coated Formvar grids. After the particles were stained with 1% (wt/vol) uranyl acetate, they were viewed by using a Philips 410 transmission electron microscope.

RT assays. Particulate material from infected-cell lysates was assayed for RT activity as described previously (20). In brief, samples (10 μ l) were mixed with an assay mixture consisting of 50 μ l of 50 mM Tris (pH 7.8), 75 mM KCl, 5 mM MgCl₂, 2.5 mM dithiothreitol, 0.05% Nonidet P-40, 15 μ g of poly(rA) · p(T)₁₂₋₁₈ per ml, [³H]dTTP (35 μ Ci/ml; 83 Ci/mmol; Amersham), and RNasin (150 U/ml; Promega) and incubated for 1 h at 37°C. One-third of the reaction mixture was spotted onto DE81 filters which were subsequently dried, washed four times with 0.5 M Na₂HPO₄, once with H₂O, and once with 95% ethanol. Incorporation of ³H was determined by liquid scintillation counting.

Isolation, cloning, and analysis of *T. ni* genomic copies of TED. Genomic DNA was isolated from TN368 cells using a standard protocol (2). In brief, cells were collected by low-speed centrifugation (1,500 $\times$ g for 10 min), washed with PBS, and suspended in digestion buffer (100 mM NaCl, 25 mM EDTA, 10 mM Tris [pH 8.0], 0.5% SDS, 0.1 mg of proteinase K per ml) at a density of 10⁸ cells per ml. After a 12- to 16-h incubation at 50°C, protein was removed by phenol extraction and the nucleic acid was recovered by ethanol precipitation. RNA was removed by digestion with 1 μ g of DNase-free RNase per ml in 0.1% SDS. To prepare genomic DNA fragments for cloning, 200 μ g of DNA was digested with *Xho*I and sedimented on a 10 to 40% (wt/vol) sucrose gradient in PBS for 21 h at 115,000 $\times$ g. DNA fragments with sizes from 5 to 20 kb were selected and cloned into the *Xho*I sites of the vector Lambda DASH (Stratagene). This strategy facilitated the isolation of full-length elements, since TED^{FP-D} lacks internal *Xho*I sites (16). DNA packaging, phage growth, and amplification were conducted as prescribed by the manufacturer. Clones containing TED sequences were identified by plaque hybridization (10), using a TED-specific probe consisting of the plasmid pTED/XA that was radiolabeled with [α -³²P]dATP (>5,000 Ci/mmol) by the random priming method (Amersham). For Southern blot analysis of TED-containing genomic clones, phage DNA was isolated (29) and simultaneously digested with restriction enzymes *Pst*I, *Bgl*II, and *Bam*HI. After agarose gel electrophoresis, the DNA fragments were transferred to Hybond-N (Amersham) and hybridized to plasmid pTED/XA DNA radiolabeled as described above.

The nucleotide sequences of TED insertion sites within the *T. ni* genome were determined directly from phage DNA by using the dideoxy-chain termination method (37) and a modified form of T7 DNA polymerase (U.S. Biochemicals). Two synthetic oligonucleotide primers complementary to the LTRs (nt 73 to 56 and 7405 to 7420, respectively) were independently used to obtain the host DNA sequences flanking each element. Since both primers annealed to each LTR, superimposed sequencing ladders were generated. The host sequence was therefore obtained by direct comparison to the previously derived sequence of TED (16).

RESULTS

Identification of proteins encoded by the TED *gag* and *pol* ORFs. To examine the protein-coding potential of the *gag* and *pol* ORFs, these regions were subcloned from TED^{FP-D} and inserted into the AcMNPV genome under control of the strong promoter for the viral polyhedrin gene; this expedited high-level expression in lepidopteran cells. Our strategy involved the insertion of progressively longer portions of TED^{FP-D} (Fig. 1), beginning with the *gag* ORF and extending through the regions of the *pol* ORF that included the

conserved PR and RT domains. This preserved the overlapping organization of the *gag* and *pol* ORFs such that expression of the *pol* ORF required a frameshift in the -1 direction, as predicted by the nucleotide sequence of this element (16).

TED-encoded proteins were identified after recombinant virus infection of cultured cells from two different moth species, *S. frugiperda* (Fig. 2A) and *T. ni* (Fig. 2B) that either lack or possess resident copies of TED, respectively (28, 31). In both cell lines, recombinant virus vGAG_{tr}, containing a truncated version of the *gag* ORF (Fig. 1B), directed synthesis of a new, abundant 41-kDa polypeptide (p41) with a molecular mass comparable to that predicted from the nucleotide sequence of this portion of the *gag* ORF (40 kDa). Recombinant virus vGAG that contained the full-length *gag* ORF (Fig. 1B) produced a similarly abundant but larger 55-kDa protein (p55). While the apparent molecular mass of p55 was higher than predicted from the full-length *gag* ORF (46 kDa), this inconsistency could be due to an altered mobility caused by posttranslational modifications that are typical of retroviral *gag* proteins (9, 43). A less-abundant 77-kDa protein (p77) was also detected in vGAG-infected cells (see below). The smaller TED-specific proteins (see asterisks) most likely represented p55 degradation products, since their levels varied between repeated experiments.

Neither p41, p55, or p77 was detected in cells infected with wild-type AcMNPV or an AcMNPV deletion mutant (v Δ POLY) that lacked the polyhedrin gene (Fig. 2). This indicated that the overexpressed proteins were specifically encoded by the portion of the *gag* ORF inserted within the recombinant virus. This was supported by the observation that identical TED-specific proteins p41, p55, and p77 were synthesized by TED-containing viruses in the two different lepidopteran species (compare Fig. 2A and B) and further argued that they were not host proteins induced by AcMNPV infection. The possibility that the proteins resulted from the activation of cryptic (viruslike) promoters located within the internal body of TED was also ruled out by the absence of similarly abundant proteins in cells infected with virus FP-D that contains the single, spontaneously inserted copy of TED^{FP-D} (Fig. 2A and B, lane 3).

The conclusion that p55 is the full-length protein product of the TED *gag* ORF was supported by the observation that it was synthesized in cells infected by all recombinant viruses that contained the entire *gag* ORF (vGAG, vGAG/PR, vGAG/PR/RT, and vGAG/POL). Infections with recombinant viruses that contained the *gag* ORF and the 5' portion of the *pol* ORF (vGAG/PR and vGAG/PR/RT) or the entire *pol* ORF (vGAG/POL) (diagrammed in Fig. 1B) exhibited reduced levels of p55 and a new, 37-kDa protein (p37) in both cell lines (Fig. 2A and B, lanes 6 to 8). In addition, polypeptides larger than p55 (designated p182 and p195) were detected. The largest of these, p195, was detected only in cells infected with vGAG/POL that contained the entire *gag-pol* region. Its observed size was comparable to that predicted for the full-length TED *gag-pol* polyprotein (189 kDa).

The TED *gag* protein, p55, assembles into VLPs. To verify that p55 represented the TED *gag* gene product, we examined its potential to assemble into VLPs, a primary function of *gag* products of retroviruses and the retrotransposons Ty1 and copia (6, 9, 12, 43). To this end, the sedimentation properties of proteins encoded by the *gag* ORF expressed from vGAG-infected cells were examined on sucrose density gradients. Electrophoretic analysis of gradient fractions indicated that p55 was the predominant protein (Fig. 3A); it was most abundant in fractions with a density (1.22 to 1.25

FIG. 1. (A) Structural and genetic organization of the retrotransposon TED. The internal body of TED, flanked by 5' and 3' LTRs, contains three overlapping ORFs designated *gag*, *pol*, and ORF 3 (*env?*). The regions within the *pol* ORF that contain conserved PR, RT, and integrase (IN) domains are indicated. Restriction sites are based on the sequence of TED^{FP-D} (16) and are abbreviated as follows: B, *Bam*HI; Bg, *Bgl*II; H, *Hind*III; E, *Eco*RV; P, *Pst*I; S, *Sal*I; S3A, *Sau*3AI; Sm, *Sma*I; Ss, *Sst*I, and X, *Xba*I. The DNA fragments generated upon simultaneous digestion of plasmid pTED/XA (containing a full-length copy of TED^{FP-D}) with *Bam*HI, *Bgl*II, and *Pst*I are numbered by size: #1 (2,446 bp), #2 (2,085 bp), #3 (1,516 bp), #4 (967 bp), and #5 (214 bp); asterisks indicate these sites. In genomic clones, fragments X and Y contain flanking sequences and a portion of the 5' and 3' LTR, respectively; X and Y comprise a single fragment (X') within plasmid pTED/XA. (B) Portions of the TED *gag-pol* ORFs inserted into recombinant viruses. AcMNPV recombinants contained the indicated internal portions of TED fused to the polyhedrin promoter (striped box) and inserted at the polyhedrin locus. For each virus, the inserted TED sequences included 104 bp upstream from the predicted initiator codon (ATG) for the *gag* ORF and extended progressively downstream to the indicated restriction sites. Recombinant viruses are designated according to the portion of TED inserted. The misaligned open boxes depict the organization of TED ORFs that was preserved upon insertion into recombinant viruses; a frameshift in the -1 direction is required to align the *gag* and *pol* ORFs. An arrow marks the RNA start site within the polyhedrin promoter (118 bp upstream from the initiator codon) and the direction of transcription.

g/ml) comparable to that of VLPs from other retrotransposons (18, 20, 30, 36). Examination of the fractions containing the peak of p55 in the electron microscope revealed abundant VLPs (Fig. 3B). These particles have a spherical shape and a diameter of 55 to 60 nm, typical of other invertebrate retrotransposons. No such particles were detected in extracts of cells infected with recombinant viruses lacking the TED *gag* ORF (not shown). These results indicated that p55 was capable of assembling into VLPs in the absence of proteins encoded by other TED ORFs.

Besides p55, several proteins cosedimented with TED VLPs; the largest and most abundant of these was p77 (Fig. 3A). While the association of p77 with VLPs suggested that p77 also possessed *gag*-like functions, its origin is not yet clear. There are various possibilities, including that p77 is a C-terminal fusion protein resulting from the infrequent translational frameshift at the junction of the *gag* and *pol* ORFs or a posttranslational modification of p55 that results in an altered electrophoretic mobility.

The PR domain within the TED *pol* ORF encodes a protease that cleaves *gag*-containing polyproteins. The synthesis of putative *gag-pol* polyproteins in cells infected with recombinant viruses that contained the TED *pol* ORF suggested that the retroviruslike PR and RT domains were expressed. To determine whether these domains were functional, we examined infected cells for the presence of TED-specific PR and RT activities, respectively. The presence of a *pol*-encoded PR was suggested by pulse-label analyses (Fig. 2)

which showed reduced levels of p55 and a concomitant appearance of p37 in cells infected exclusively with recombinant viruses that contained the putative PR domain. To further examine the presence of a *gag*-specific protease, we examined the precursor-product relationship of p55 and p37 by using pulse-chase experiments (Fig. 4).

When cells were infected with recombinant vGAG (containing only the *gag* ORF), the level of p55 (and associated p77) was unaltered throughout the 3-h chase period (Fig. 4, lanes 14 to 18). However, in cells infected with recombinant viruses that included, in addition to the *gag* ORF, the PR domain alone (vGAG/PR [lanes 4 to 8]) or the entire *pol* ORF (vGAG/POL [lanes 9 to 13]), p55 disappeared during the chase, exhibiting a half-life of 15 to 30 min. The putative *gag-pol* fusion protein p195 synthesized in vGAG/POL-infected cells (lanes 9 to 13) also disappeared rapidly during the chase. The rate of p55 disappearance paralleled the appearance and steady increase in p37 suggesting that p55 was cleaved to generate p37. The finding that p37 was the major protein component of the *gag-pol*-specific VLPs (see below) supported this precursor-product relationship. If p55 were processed into a minimum of two proteins, one of which was p37, then the size of the companion cleavage product would be 10 to 15 kDa. Thus far, we have not identified such a protein by these methods, because all other proteins appearing during the chase were accounted for by AcMNPV-specific proteins (compare with Δ POLY [lanes 1 to 3]). This may be due in part to an aberrant mobility of this

FIG. 2. Identification of TED-specific proteins in recombinant virus-infected cells. *S. frugiperda* (SF21) cells (A) and *T. ni* (TN368) cells (B), infected 39 h previously, were radiolabeled for 1 h with [35 S]methionine-cysteine, then lysed, and subjected to SDS-polyacrylamide gel electrophoresis. Autoradiograms are shown with the positions of molecular size standards (in kilodaltons) shown on the left of each panel. Cells were infected with the indicated viruses: wild-type (WT) AcMNPV (lane 1), AcMNPV polyhedrin-deletion mutant Δ POLY (lane 2), TED insertion mutant FP-D (lane 3), and recombinant viruses vGAG_{tr} (lane 4), vGAG (lane 5), vGAG/PR (lane 6), vGAG/PR/RT (lane 7), and vGAG/POL (lane 8); the internal portion of TED contained by each recombinant virus is depicted in Fig. 1B. TED-specific proteins are indicated by size on the right, while the AcMNPV proteins polyhedrin (poly) and gp64 are indicated on the left.

protein or inefficient radiolabeling with 35 S because of the uneven distribution of methionine and cysteine residues within the *gag* ORF.

The TED *pol* ORF encodes an active RT. To extend our analysis to the RT domain of TED, we tested infected cells for the presence of RT. The association of this activity with retrotransposon VLPs (13, 18, 30, 38) suggested that inclusion of the TED *pol* ORF in recombinant viruses would result in production of *pol*-specific proteins, including RT, that associated with TED VLPs. This provided a means by which RT could be concentrated (via ultracentrifugation), thereby facilitating its characterization. To distinguish RT from other DNA polymerase activities, we used an assay that measured synthesis of complementary DNA in response to a single-stranded RNA template [poly(rA)] and an oligo(dT) primer.

When the portion of the *pol* ORF containing sequence similarities with the conserved retrovirus domains for RT was included in recombinant viruses (vGAG/PR/RT and vGAG/POL), a 50- to 60-fold increase in RT activity was detected in the particulate material from infected cells compared with cells infected with viruses lacking the *pol* ORF (vGAG and Δ POLY) (Fig. 5). Cells infected with the latter viruses exhibited activity nearly identical to that of uninfected *S. frugiperda* cells, indicating that AcMNPV infection alone failed to induce this activity. Preliminary characterization indicated that the TED-specific RT had a broad temperature range (from 25 to 37°C), with maximum activity

at 25°C and pH 7.8; moreover, Mn^{2+} functionally substituted for Mg^{2+} without loss of activity (data not shown). Cells infected with virus FP-D, that contains a full-length copy of TED, produced twice the RT activity of vGAG- or Δ POLY-infected cells, suggesting that the TED LTR was also capable of directing expression of the *pol* ORF in the intact element. The higher levels of activity synthesized by recombinant viruses vGAG/PR/RT and vGAG/POL compared with that of FP-D may be a result of the relative strength of the polyhedrin promoter in directing expression of TED-specific proteins.

The *gag-pol* ORFs direct assembly of TED-VLPs that contain RT and cleaved *gag* protein, p37. The ability to concentrate intracellular RT by ultracentrifugation indicated that this activity was associated with particulate material, possibly VLPs. To characterize such particles, intracellular material from vGAG/POL-infected cells was subjected to sedimentation on sucrose density gradients. Analysis of the gradients for RT revealed a prominent peak of activity at a density of 1.16 g/ml (Fig. 6A); this activity cosedimented with the major peak of absorbance (A_{254}) within these gradients (see below). No activity was detected in gradients of intracellular material from cells infected with viruses that lacked the TED *pol* ORF (data not shown). To examine the TED-specific proteins present, radiolabeled particulate material from vGAG/POL-infected cells was sedimented under identical conditions. Electrophoretic analysis of the gradient fractions indicated that p37 was the predominant protein

FIG. 3. Sedimentation and electron microscopy of TED VLPs from vGAG-infected cells. (A) Sucrose gradient sedimentation. SF21 cells were radiolabeled with [35 S]methionine-cysteine from 36 to 48 h after infection with recombinant virus vGAG. Intracellular particulate material was collected by ultracentrifugation, sedimented on a discontinuous 20 to 60% sucrose gradient, and analyzed by SDS-polyacrylamide gel electrophoresis. An autoradiogram is shown in which fractions are indicated from top to bottom of the gradient (left to right, respectively); the topmost fractions were omitted because of the absence of detectable TED-specific proteins. Also shown are the intracellular proteins (see the legend to Fig. 2) from infections with AcMNPV polyhedrin-deletion mutant vΔPOLY (lane 1) and recombinant viruses vGAG/POL (lane 2) and vGAG (lane 3). The position of TED-specific proteins (p55 and p77) and molecular size standards (in kilodaltons) are indicated on the right. (B) Electron micrographs of gradient-purified VLPs. Unlabeled particles collected from gradient fractions containing the peak of p55 were negatively stained and visualized without prior fixation. Bar, 100 nm.

FIG. 4. Pulse-chase analysis of proteins synthesized in recombinant virus-infected cells. SF21 cells were radiolabeled for 20 min with [35 S]methionine-cysteine 39 h after infection with the polyhedrin-deletion mutant v Δ POLY, and the TED-containing recombinant viruses vGAG/PR, vGAG/POL, and vGAG, respectively. The radiolabel was removed and replaced with medium containing excess unlabeled methionine and cysteine to initiate the chase. Cells were lysed at the indicated times (0 min, 15 min, 30 min, 1 h, and 3 h) thereafter, and radiolabeled proteins were analyzed by SDS-polyacrylamide gel electrophoresis. A fluorogram is shown with 14 C-labeled molecular size standards (in kilodaltons) indicated on the right along with TED-specific proteins p195, p77, p55, and p37.

(Fig. 6B); p55 was also present but in reduced amounts. The viruslike sedimentation properties of p37 resembled that of p55 alone (Fig. 3) and suggested that p37 represented the major *gag* protein component of vGAG/POL-derived VLPs.

In contrast to the peak of RT activity, the gradient distribution of radiolabeled p37 was quite broad (Fig. 6, compare panels A and B). To determine whether this was the result of an association (or aggregation) of p37 with cellular components such as polysomes, we examined the effect of removing RNA from intracellular particulate material prior to centrifugation. Treatment with RNase A eliminated the predominant gradient peak of A_{254} and thus demonstrated that a major component of this material was composed of ribonuclease-sensitive RNA (data not shown). RNase treatment also altered the gradient distribution of RT (Fig. 6A). The predominant peak of activity at 1.16 g/ml (dashed line) was replaced by a new peak (solid line) with a higher density (1.23 g/ml); the additional activity near the top of the gradient most likely represented free (unassociated) protein. Examination of the gradient distribution of p37 after RNase treatment (Fig. 6C) indicated a parallel change to a higher and more uniform density of 1.23 g/ml. Separate analyses confirmed that this peak of p37 cosedimented with RT activity (data not shown). Thus, a significant proportion of the intracellular RT was associated with p37-containing

FIG. 5. Relative levels of RT synthesized in infected *S. frugiperda* (SF21) cells. Extracts of uninfected cells or cells infected 48 h previously with the AcMNPV deletion mutant v Δ POLY, AcMNPV insertion mutant FP-D, and recombinant viruses vGAG, vGAG/PR/RT, and vGAG/POL, respectively, were subjected to ultracentrifugation ($285,000 \times g$) to concentrate particulate material. Samples of such material from the equivalent of 5×10^5 cells were assayed for RT by measuring DNA synthesis using a poly(rA) · p(T)₁₂₋₁₈ template-primer mix and [3 H]dTTP (17). RT activity is reported as counts per minute (cpm) of [3 H]dTTP incorporated above background, defined as the counts per minute incorporated in the absence of cell extract. Virus designations are as shown in Fig. 1B.

vGAG/POL-derived VLPs. Since RNase had no effect on the gradient distribution (including the average density of 1.23 g/ml) of vGAG-derived VLPs (data not shown), these results suggested that vGAG/POL-derived VLPs differentially associated with RNase-sensitive components that affected particle density. Further analyses are required to determine the basis for this association.

The AcMNPV copy TED^{FP-D} structurally resembles TED elements in the host *T. ni* genome. As a first step in determining whether expression of TED elements residing in the host genome is identical to that of TED^{FP-D} present in the baculovirus genome, we cloned full-length copies of TED from *T. ni*. The structural organization of 18 genomic clones (G1 through G18) was then compared directly to TED^{FP-D} by restriction mapping and Southern blot hybridization (Fig. 7). Simultaneous digestion with enzymes *Bam*HI, *Bgl*II, and *Pst*I generated five DNA fragments (fragments 1 to 5) that were diagnostic of the LTRs and internal portions of TED^{FP-D} (Fig. 1A). The elements within 10 of the 18 clones (G2, G5, G7, G8, G9, G11, G13, G14, G17, and G18) were indistinguishable from TED^{FP-D}. This was confirmed by other analyses that employed a different set of enzymes (*Eco*RV, *Hind*III, and *Sst*I) (data not shown). Since fragment X of these genomic clones is composed of sequences extending from the *Pst*I site within the 5' LTR to the next restriction site (*Bam*HI, *Bgl*II, or *Pst*I) within the DNA that flanks the LTR (Fig. 1A), it is potentially diagnostic of the position of TED elements within the host genome. Thus, the observed differences in the size of fragment X for the 10 genomic clones (Fig. 7, see arrows on gels) suggested that each of these TED elements was derived from a different location within the *T. ni* genome.

Determination of the nucleotide sequence at the site of

FIG. 6. Sedimentation analysis of recombinant virus vGAG/POL-derived VLPs. (A) Sucrose gradient distribution of RT activity. Intracellular particles were collected from SF21 cells 48 h after infection with recombinant virus vGAG/POL by ultracentrifugation. One-half of the preparation was incubated with RNase A (60 μ g/ml) for 30 min at 37°C, and then both halves were sedimented on separate 20 to 60% discontinuous sucrose gradients. Fractions were collected and assayed for RT activity which is reported as counts per minute (cpm) of [3 H]dTTP incorporated as described in the legend to Fig. 5. The RT profiles for untreated particles (—RNase) (---) and RNase-treated particles (+RNase) (—) are shown from top to bottom of each gradient (left to right, respectively). (B) Gradient distribution of TED-specific proteins without RNase treatment. Intracellular particles from vGAG/POL-infected cells, radiolabeled with [35 S]methionine-cysteine from 36 to 48 h after infection, were collected and sedimented on a sucrose gradient as described above for panel A. The indicated fractions, numbered from top to bottom of the gradient (left to right), were subjected to SDS-polyacrylamide gel electrophoresis. An autoradiogram is shown with the positions of molecular size standards (in kilodaltons) indicated on the left. The topmost fractions were omitted because of their lack of detectable protein. (C) Gradient distribution of TED-

TED integration confirmed that of 10 genomic clones examined, at least 8 were distinct copies, since the flanking host sequences differed (Fig. 8). Although the flanking sequences were identical for clones G3, G4, and G5, the identical restriction patterns of these three clones (Fig. 7) suggested that they represented identical copies of the same element; the alternative possibility that each clone represented TED and host DNA that was amplified within the *T. ni* genome through a mechanism other than transposition (e.g., unequal cross-over) is not ruled out. The terminal nucleotides of the 5' and 3' LTRs (TGTTA and TAATT, respectively) of TED were identical for all 10 genomic clones. Moreover, within each clone, the 4 nt flanking the 5' and 3' LTR, respectively, were identical. Thus, like the integration of TED^{FP-D} within the AcMNPV genome, integration of host elements apparently resulted in a duplication of 4 host nt. Although limited, this analysis also suggested that TED has no apparent sequence specificity for the site of integration, which distinguishes it from other members of the gypsy family that exhibit a high degree of specificity (12).

This analysis indicated that while half of the elements were similar to TED^{FP-D}, considerable heterogeneity among other genomic copies of TED exists. Studies of gypsy elements within the *Drosophila* genome have indicated a similar heterogeneity of individual elements (26, 34). The most common alteration among the host TED clones was the absence of internal DNA fragments (Fig. 7). For example, clones G3, G4, G6, G10, G12, and G16 each lacked DNA fragment 2 corresponding to TED *ORF3*. Since clones G3, G4, and G6 possessed a 3' LTR (Fig. 8), loss of fragment 2 was due to an internal deletion. At least two clones (G1 and G15) exhibited patterns indicative of elements with truncated 5' ends. Besides deletions, other alterations could be explained by the presence of *Xho*I sites in the host copy of TED that resulted in the cloning of less than full-length elements. Genomic clones G19 and G20, selected for analysis because of their comparatively weak hybridization to the TED probe used, exhibited only one or two TED-specific fragments (Fig. 7). Such patterns were compatible with the presence of single (solo) LTRs or other *T. ni* sequences minimally related to TED.

DISCUSSION

Expression strategy and function of TED *gag* and *pol* ORFs.

By employing the baculovirus expression system in which progressively longer sections of TED were inserted into recombinant viruses under control of the late polyhedrin promoter, we have identified TED structural proteins and several enzymatic activities encoded by the *gag* and *pol* ORFs. Our current model for expression of TED, based on the determined relationships of various proteins and their apparent function, is presented in Fig. 9. Since a majority of the TED elements cloned directly from the *T. ni* genome structurally resembled the viral isolate TED^{FP-D} used in this study (Fig. 7), the model is expected to be representative of full-length elements residing within the host genome. Moreover, since TED proteins were indistinguishable when over-

specific proteins after RNase treatment. An equivalent portion of radiolabeled intracellular particles prepared in panel B was incubated with RNase A and analyzed by sedimentation on an identical sucrose gradient and gel electrophoresis. Radiolabeled proteins (Fig. 2) from vGAG/POL-infected cells were included in panels B and C (lanes 1 and 19, respectively).

FIG. 7. Southern blot analysis of TED-containing clones derived from the *T. ni* genome. Plasmid pTED/XA (containing the full-length copy of TED^{FP-D}) and phage DNA from the *T. ni* genomic clones (G1 to G20) were digested simultaneously with restriction enzymes *Bam*HI, *Bgl*II, and *Pst*I, subjected to agarose gel electrophoresis, blotted, and hybridized to a radiolabeled probe consisting of pTED/XA. An autoradiogram is shown with DNA size markers (in kilobase pairs) on the left. The five restriction fragments (#1 to #5) of pTED/XA (lane 22) that contain TED sequences are labeled on the right; the position of each fragment within TED is depicted in Fig. 1A. Fragment X' corresponds to a *Pst*I-*Pst*I fragment of pTED/XA that contains the 5' LTR and plasmid sequences. The analogous fragment (X) from the *T. ni* genomic clones (see arrows on gel) contains the 5' LTR and flanking host sequences and is therefore unique for each independent copy of TED. The corresponding Y fragments of these clones contain the host sequences flanking the 3' LTR but only 16 nt of the LTR (Fig. 1A) and were not detected by the hybridization methods used.

produced in cultured *T. ni* and *S. frugiperda* cells (Fig. 2), expression is similar in lepidopteran hosts with or without resident copies of TED.

The primary translation product of the TED *gag* ORF is p55. When expressed alone, p55 assembled into VLPs with a morphology and density similar to that of VLPs from copia and Ty1 (1, 18, 20, 30, 36, 38). Upon coexpression of the TED *pol* ORF, p55 was cleaved to produce p37 (Fig. 4). The finding that p37 was the major protein component of vGAG/POL-derived VLPs (Fig. 6) indicated that it also exhibited *gag*-like functions and further supported the precursor-product relationship of p55 and p37. On the basis of these results, p55 represents the TED *gag* precursor (hereafter designated as Pr55^{gag}). While alternative methods (e.g., immunoprecipitation) will be required to identify the second cleavage product of Pr55^{gag}, a coding location at the 3' terminus of the *gag* ORF is reasonable. From its nucleotide sequence (16), the acidic and tyrosine-rich composition of this portion of the *gag* ORF distinguishes it from the adjacent region that is highly basic and proline-rich (predicted to comprise the nucleocapsid domain) (Fig. 9); analogous

acidic domains exist at the 3' end of the *gag* ORF of retrotransposons 17.6 and 297 (28). Proteolytic separation of this 10- to 15-kDa C-terminal domain from Pr55^{gag} would produce a protein comparable in size to that of p37. Moreover, the cleavage would position the putative nucleocapsid region at the C terminus of p37 such that the resulting NH₂-capsid-nucleocapsid-COOH arrangement would be similar to that of the retroviral *gag* polyproteins (9, 43). In the absence of further processing, a single protein (p37) may therefore provide both capsid and nucleocapsid functions.

As predicted by the model (Fig. 9), p195 represents the full-length *gag-pol* polyprotein. This was supported by several findings. First, the size of p195 was comparable to that expected for a TED *gag-pol* fusion protein (189 kDa); it was the largest of the TED-specific proteins identified and was only observed in cells infected with recombinant vGAG/POL that contained the entire TED *gag-pol* region. Second, the level of p195 was significantly lower than that of the *gag* proteins Pr55^{gag} and p37 (Fig. 2). This was consistent with a frameshift event that is required for expression of *pol*-encoded products; translational frameshifts typically decrease the yield of the fusion protein relative to that of the nonfused protein (43). Ribosomal slippage is predicted to occur at the UUU codon of the heptanucleotide G GAU UUU (TED nt 1856 to 1862) that is located within the overlap region of the *gag* and *pol* ORFs (Fig. 9) and closely matches known -1 frameshift sequences of vertebrate retroviruses (22, 41). Lastly, p195 exhibited a short half-life, suggesting that it was rapidly cleaved after synthesis, a process that mimics the processing of retrovirus *gag-pol* polyproteins (9, 43). Indeed, p195 processing was accompanied by the appearance of PR and RT activities in recombinant virus-infected cells. The PR activity that mediated cleavage of Pr55^{gag} (and possibly p195) mapped to the 5' portion of the *pol* ORF (nt 1706 to 2829) that contains a domain similar to the highly conserved domain for the active site of the retrovirus aspartyl proteases (23, 42). Likewise, RT was encoded by the portion of the *pol* ORF (nt 2829 and 4707) that contains sequences similar to the conserved retrovirus domains for RT. Studies are under way to identify the mature polypeptides for these activities, as well as those potentially encoded by the adjacent RNase H and integrase domains, each of which is essential for retrotransposition (6).

Potential role of VLPs in TED transposition. The finding that TED directs assembly of VLPs demonstrated for the first time that an invertebrate member of the gypsy family encodes a functionally active *gag* gene product. As suggested by studies of Ty1, an apparent prerequisite for retrotransposition is the formation of VLPs that contain RT and full-length RNA (7, 13, 14, 18, 30). Indeed, TED-encoded RT was associated with vGAG/POL-derived particles, although demonstration of this required the removal of ribonuclease-sensitive material prior to the sedimentation of intracellular particles (Fig. 6). The association was confirmed by the cosedimentation of RT and VLPs when extracellular particulate material was examined in the absence of ribonuclease (28). Whether these extracellular particles were generated by budding from the plasma membrane or baculovirus-mediated cell lysis remains to be determined. Direct demonstration of TED's transpositional competency will depend on the packaging of full-length RNA into newly assembled VLPs and the subsequent integration of reverse transcribed copies within a recipient DNA. Nevertheless, our preliminary findings suggest that despite sequence and structural differences, the major invertebrate groups of retrotransposons (gypsy, copia,

TED Clone #	5' flanking sequence	5' LTR	3' LTR	3' flanking sequence
FP-D	TAAAACA <u>AA</u> TG	TGTTA--	--TAATT	<u>AA</u> TGCGCAAGC
G2	AACTATGCAG <u>TT</u>	TGTTA--	--TAATT	<u>AG</u> TTGTG
G5	TTTAAGTTGAC <u>CA</u>	TGTTA--	--TAATT	<u>ACC</u> ATGAGT
G7	ATTTTATGTCC <u>CT</u>	TGTTA--	--TAATT	<u>CCCT</u> TGTAAAC
G8	CTGTTTGG <u>CA</u>	TGTTA--	--TAATT	<u>GGC</u> AGA
G11	CACAAGATAA <u>CT</u>	TGTTA--	--TAATT	<u>AACT</u> TTCCA
G18	CTTGTTGCAG <u>CG</u>	TGTTA--	--TAATT	<u>AGCG</u> TGATGG
G3, 4 & 6	TCTTGAAGCTG <u>CT</u>	TGTTA--	--TAATT	<u>TGCT</u> TTTATAG

FIG. 8. Comparison of the nucleotide sequence at various integration sites of TED within the *T. ni* genome. The nucleotides flanking the 5' and 3' LTR are shown along with the 5 terminal nt of each LTR, respectively. The insertion site of TED within the genome of AcMNPV mutant FP-D is also shown. The 4 viral nt duplicated upon TED integration are underlined. The 4 nt of host *T. ni* DNA that flank each LTR are also underlined.

and Ty) use similar strategies for movement, including the production of VLPs with RT. As has already been shown for several vertebrate retroviruses (11, 19, 33), baculovirus vectors should prove useful in further defining the role of VLPs and *gag-pol* polyproteins in transposition of TED and other invertebrate elements.

Evolutionary relationship between TED and the retroviral proviruses. Comparison of nucleotide sequences and gene organization of retroviral elements has led to the suggestion that the retroviruses originated from *gag-pol*-containing retrotransposons that subsequently acquired a third (*env*) gene making extracellular cell-to-cell movement possible (40, 47). Because of their striking resemblance to the retroviruses, including the presence of a third ORF analogous to the *env* gene, members of the gypsy family of retrotransposons may provide clues in retrovirus evolution. Our studies here extend this resemblance to include functions and mecha-

nisms of expression of the first two ORFs (*gag* and *pol*) on the basis of the finding that TED produces VLPs with enzymatic activities (RT and a *gag-pol* PR) normally associated with retrovirus particles. Of key interest is the role of the third TED ORF in element movement and potential infectivity of TED VLPs. Preliminary experiments using recombinant baculoviruses indicated that ORF3 encodes an *env*-like glycoprotein (39). Characterization of this gene product may be useful in ascertaining whether TED represents a retrotransposon that has acquired *env*-like functions or whether it is an endogenous retrovirus that because of undefined mutations has lost the capacity for extracellular movement.

Potential advantages of host transposon and virus associations. The insertion of host transposons has the potential to incorporate new regulatory sequences, genes, or both into the baculovirus genome and thereby increase genetic diver-

FIG. 9. Model for the expression of TED-encoded proteins and their function. The retrotransposon TED contains three ORFs (overlapping open boxes) designated *GAG*, *POL*, and *ORF3*, respectively, that are flanked by identical LTRs (hatched). TED encodes polyproteins Pr55^{gag} and p195 (solid bars) which by analogy to the retroviruses, are predicted to have a common N terminus; the relative abundance of TED proteins is indicated by the thickness of each bar. Synthesis of p195, the putative *gag-pol* polyprotein, requires a frameshift in the -1 direction that is predicted to occur at the sequence UUUU (underlined) within the overlap region of the *gag* and *pol* ORFs. Expression of the *pol* ORF produces PR that mediates cleavage of the *gag* precursor Pr55^{gag} and generates p37, the major protein component of TED VLPs. The predicted capsid (CA) domain and the basic, proline-rich nucleocapsid (NC) domain of Pr55^{gag} are indicated along with the comparatively acidic (Ac) domain at the C terminus. The Pr55^{gag} cleavage site and the smaller companion cleavage product to p37 remain to be identified. The *pol* ORF also encodes an active RT that is associated with TED-VLPs. The putative *pol* domains for RNase H (RH) and integrase (IN) are indicated. Full-length RNA transcripts initiate (see arrow) and terminate within the 5' and 3' LTRs, respectively.

sity of the virus (reviewed by reference 5). Representing the first example of the spontaneous insertion of host-derived genes within an eukaryotic DNA virus, TED^{FP-D} integration provided AcMNPV with several new genes, including those for *gag* proteins, a PR, RT, and others not examined here. Through mechanisms analogous to retrovirus transduction (43), TED may also introduce host-derived genes that could confer selectable advantages to the virus and thereby contribute to baculovirus evolution. By insertion within the AcMNPV genome, TED also acquired new functions, including those encoded by the virus for extracellular transmission. Through baculovirus integration, TED therefore gains the potential to escape the confines of the cell in which it resides and spread through the virus population, a feature not unexpected of selfish DNA. Thus, the association between host transposon and virus may be mutually beneficial.

A major question to be addressed is whether baculoviruses promote interspecies movement of transposons by acting as a shuttle vehicle (8, 16, 31). A necessary step in this process is successful transposition of the element after integration within the virus genome. Our findings indicate that TED^{FP-D} retains the function of several of the gene products required for retroviral transposition. In addition, the U3-R-U5 organization of the TED LTRs directs the synthesis of a full-length genomic RNA with redundant termini from the AcMNPV genome (16). As such, these RNAs represent potential substrates for reverse transcription and intermediates in the transposition process. To date therefore, our evidence suggests that TED is capable of transposition after integration within AcMNPV. Studies to test this possibility are under way.

ACKNOWLEDGMENTS

We thank John Fleming for assistance in the construction of several TED-containing plasmids, Lois Miller (University of Georgia) for providing the AcMNPV polyhedrin-deletion mutant, C. Yong Kang (University of Ottawa, Canada) for unpublished VLP protocols, and Renate Bromberg for assistance in electron microscopy.

This work was supported in part by Public Health Service grant AI25557 from the National Institute of Allergy and Infectious Diseases. R.A.L. was supported by Viral Oncology Training grant CA09075 from the National Institutes of Health.

REFERENCES

- Adams, S. E., J. Mellor, K. Gull, R. B. Sim, M. F. Tuite, S. M. Kingsman, and A. J. Kingsman. 1987. The functions and relationships of Ty-VLP proteins in yeast reflect those of mammalian retroviral proteins. *Cell* 49:111-119.
- Ausubel, F. M., R. Brent, E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1989. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
- Beames, B., and M. D. Summers. 1990. Sequence comparison of cellular and viral copies of host cell DNA insertions found in *Autographa californica* nuclear polyhedrosis virus. *Virology* 174:354-363.
- Bingham, P. M., and Z. Zachar. 1989. Retrotransposons and the FB transposon from *Drosophila melanogaster*, p. 485-502. In D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
- Blissard, G. W., and G. F. Rohrmann. 1990. Baculovirus diversity and molecular biology. *Annu. Rev. Entomol.* 35:127-155.
- Boeke, J. D., and V. G. Corces. 1989. Transcription and reverse transcription of retrotransposons. *Annu. Rev. Microbiol.* 43: 403-434.
- 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.
- Cary, L. C., M. Goebel, B. G. Corsaro, H. Wang, E. Rosen, and M. J. Fraser. 1989. Transposon mutagenesis of baculoviruses: analysis of *Trichoplusia ni* transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. *Virology* 172:156-169.
- Coffin, J. M. 1990. Retroviridae and their replication, p. 1434-1500. In B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), *Fields virology*, 2nd ed. Raven Press, Ltd., New York.
- Davis, L. G., M. D. Digner, and J. F. Battey. 1986. Basic methods in molecular biology. Elsevier Science Publishing, Inc., New York.
- Delchambre, M., D. Gheysen, D. Thines, C. Thiriart, E. Jacobs, E. Verdin, M. Horth, A. Burny, and F. Bex. 1989. The GAG precursor of simian immunodeficiency virus assembles into virus-like particles. *EMBO J.* 8:2653-2660.
- Echallier, G. 1989. *Drosophila* retrotransposons: interactions with genome. *Adv. Virus Res.* 36:33-105.
- Eichinger, D. J., and J. D. Boeke. 1988. The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: cell-free Ty1 transposition. *Cell* 54:955-966.
- Fink, G. R., J. D. Boeke, and D. J. Garfinkel. 1986. The mechanism and consequences of retrotransposition. *Trends Genet.* 2:118-123.
- Friesen, P. D., and L. K. Miller. 1985. Temporal regulation of baculovirus RNA: overlapping early and late transcripts. *J. Virol.* 54:392-400.
- Friesen, P. D., and M. S. Nissen. 1990. Gene organization and transcription of TED, a lepidopteran retrotransposon integrated within the baculovirus genome. *Mol. Cell. Biol.* 10:3067-3077.
- Friesen, P. D., W. C. Rice, D. W. Miller, and L. K. Miller. 1986. Bidirectional transcription from a solo long terminal repeat of the retrotransposon TED: symmetrical RNA start sites. *Mol. Cell. Biol.* 6:1599-1607.
- Garfinkel, D. J., J. D. Boeke, and G. R. Fink. 1985. Ty element transposition: reverse transcriptase and virus-like particles. *Cell* 42:506-517.
- Gheysen, D., E. Jacobs, F. de Foresta, C. Thiriart, M. Francotte, D. Thines, and M. De Wilde. 1989. Assembly and release of HIV-1 precursor Pr55^{gag} virus-like particles from recombinant baculovirus-infected insect cells. *Cell* 59:103-112.
- Heine, C. W., D. C. Kelly, and R. J. Avery. 1980. The detection of intracellular retrovirus-like entities in *Drosophila melanogaster* cell cultures. *J. Gen. Virol.* 49:385-395.
- Hink, W. F. 1970. Established insect cell line from cabbage looper, *Trichoplusia ni*. *Nature (London)* 225:466-467.
- Jacks, T., H. D. Madhani, F. R. Masiarz, and H. E. Varmus. 1988. Signals for ribosomal frameshifting in the rous sarcoma virus *gag-pol* region. *Cell* 55:447-458.
- Katoh, I., T. Yasunaga, Y. Ikawa, and Y. Yoshinaka. 1987. Inhibition of retroviral protease activity by an aspartyl proteinase inhibitor. *Nature (London)* 329:654-656.
- Kingsman, A. J., and S. M. Kingsman. 1988. Ty: a retroelement moving forward. *Cell* 53:333-335.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* 227:680-685.
- Lambertsson, A., S. Andersson, and T. Johansson. 1989. Cloning and characterization of variable-sized gypsy mobile elements in *Drosophila melanogaster*. *Plasmid* 22:22-31.
- Lee, H. H., and L. K. Miller. 1978. Isolation of genotypic variants of *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* 27:754-767.
- Lerch, R. A., and P. D. Friesen. Unpublished data.
- Malik, A. N., P. M. McLean, A. Roberts, P. S. Barnett, A. G. Demaine, J. P. Banga, and A. M. McGregor. 1990. A simple high yield method for the preparation of lambda gt10 DNA suitable for subcloning, amplification and direct sequencing. *Nucleic Acids Res.* 18:4031-4032.
- Mellor, J., M. H. Malim, K. Gull, M. F. Tuite, S. McCready, T. Dibbayawan, S. M. Kingsman, and A. J. Kingsman. 1985. Reverse transcriptase activity and Ty RNA are associated with virus-like particles in yeast. *Nature (London)* 318:583-586.
- 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.
32. Miller, L. K. 1988. Baculoviruses as gene expression vectors. *Annu. Rev. Microbiol.* **42**:177–199.
 33. Overton, H. A., Y. Fuji, I. R. Price, and I. M. Jones. 1989. The protease and *gag* gene products of the human immunodeficiency virus: authentic cleavage and post-translational modification in an insect cell expression system. *Virology* **170**:107–116.
 34. Pfeifer, M., and W. Bender. 1988. Sequences of the gypsy transposon of *Drosophila* necessary for its effects on adjacent genes. *Proc. Natl. Acad. Sci. USA* **85**:9650–9654.
 35. 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. Cox (ed.), *Animal virus genetics*. Academic Press, Inc., New York.
 36. Saigo, K., T. Shiba, and T. Miyake. 1980. Virus-like particles of *Drosophila melanogaster* containing t-RNA and 5S ribosomal RNA. I. Isolation and purification from cultured cells and detection of low molecular weight RNAs in the particles, p. 411–424. In E. Kurstak, K. Maramorosch, and A. Dübenborfer (ed.), *Invertebrate systems in vitro*. Elsevier/North-Holland Biomedical Press, Amsterdam.
 37. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 38. Shiba, T., and K. Saigo. 1983. Retrovirus-like particles containing RNA homologous to the transposable element *copia* in *Drosophila melanogaster*. *Nature* (London) **302**:119–124.
 39. Szatkowski, M. T., and P. D. Friesen. Unpublished data.
 40. Temin, H. M. 1980. Origin of retroviruses from cellular moveable genetic elements. *Cell* **21**:599–600.
 41. Ten Dam, E. B., C. W. A. Pleij, and L. Bosch. 1990. RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. *Virus Genes* **4**:121–136.
 42. Toh, H., R. Kibuno, M. Hayashida, T. Miyata, W. Kugimiya, S. Inouye, S. Yuki, and K. Saigo. 1985. Close structural resemblance between putative polymerase of a *Drosophila* transposable genetic element 17.6 and the *pol* gene product of Moloney murine leukaemia virus. *EMBO J.* **4**:1267–1272.
 43. Varmus, H., and P. Brown. 1989. Retroviruses, p. 53–108. In D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
 44. Vaughn, J. L., R. H. Goodwin, G. L. Thompson, and P. McCawley. 1977. Establishment of two insect cell lines from the insect *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *In Vitro* (Rockville) **13**:213–217.
 45. Wang, H. H., M. J. Fraser, and L. C. Cary. 1989. Transposon mutagenesis of baculoviruses: analysis of TFP3 lepidopteran transposon insertions at the *FP* locus of nuclear polyhedrosis viruses. *Gene* **81**:97–108.
 46. Webb, N. R., and M. D. Summers. 1990. Expression of proteins using recombinant baculoviruses. *Technique* **2**:173–188.
 47. Xiong, Y., and T. H. Eickbush. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* **9**:3353–3362.
 48. Youngren, S. D., J. D. Boeke, N. J. Sanders, and D. J. Garfinkel. 1988. Functional organization of the retrotransposon Ty from *Saccharomyces cerevisiae*: Ty protease is required for transposition. *Mol. Cell. Biol.* **8**:1421–1431.