Identification and Characterization of *vlf-1*, a Baculovirus Gene Involved in Very Late Gene Expression

JEANNE R. McLACHLIN[1] AND LOIS K. MILLER[1,2]*

Department of Entomology[1] and Department of Genetics,[2] University of Georgia, Athens, Georgia 30602

We have identified a gene required for strong expression of the polyhedrin gene by characterizing a mutant, *ts*B837, of the baculovirus *Autographa californica* nuclear polyhedrosis virus (Ac*M*NPV) which is temperature sensitive (*ts*) for occluded virus production at the nonpermissive temperature. Marker rescue experiments utilizing an overlapping set of Ac*M*NPV genomic clones revealed that the gene responsible for the *ts* mutant phenotype mapped to a region between 46 and 48 map units. Fragments (2.2 kb) from both wild-type Ac*M*NPV and *ts*B837 genomes spanning the mutated region were sequenced, and a single nucleotide difference was observed. This mutation is predicted to substitute a single amino acid within a 44.4-kDa polypeptide. Analysis of protein synthesis in wild-type- and mutant-infected cells at the nonpermissive temperature indicated that polyhedrin synthesis was dramatically reduced in the mutant. Northern (RNA) blot analysis revealed that the mutant had markedly reduced levels of polyhedrin transcripts. Transcripts of another very late gene, *p10*, were also reduced but to a lesser degree. The transcription of two late genes (603 ORF and *vp39*) was neither reduced nor temporally delayed. Thus, the gene encoding this very late expression factor, designated *vlf-1*, regulates the levels of very late gene transcripts, and the *ts*B837 mutation affects the levels of polyhedrin gene transcripts more strongly than those of *p10* transcripts. The product of the newly identified gene has a surprising but significant relationship to a family of integrases and resolvases.

The promoters of the genes encoding the polyhedrin and p10 proteins of *Autographa californica* nuclear polyhedrosis virus (Ac*M*NPV) are most frequently employed in baculovirus expression vector systems to drive heterologous gene transcription (reviewed in references 14 and 27). Both promoters are strongly activated during the final, or very late, stage of infection, which is characterized by the formation of polyhedral occlusion bodies in the nucleus. These large refractive occlusions, often referred to simply as polyhedra, are visible by light microscopy and are composed primarily of polyhedrin, a polypeptide which forms a paracrystalline matrix around multiple virions. Polyhedra are noninfectious in cell culture but serve in the horizontal spread of infection among larvae. Many baculovirus expression vector systems are designed to replace the polyhedrin gene (*polh*) with a heterologous gene; the absence of polyhedra then serves as a visual screen for recombinants. The product of the other major very late gene, p10, forms extensive fibrillar arrays in both the cytoplasm and nucleus, but p10 is not essential for either virus replication or polyhedron formation (44, 46, 49). Although the expression of other genes may also be activated during the very late phase of infection, *polh* and *p10* are unique with regard to the extreme abundance of their transcripts and gene products, accounting for the utility of their promoters in expression vector systems.

Unlike late baculovirus promoters, which are activated at between 6 and 12 h postinfection (p.i.), transcription from the *polh* and *p10* promoters is strongly activated at between 18 and 24 h p.i. However, both late and very late baculovirus genes are thought to be transcribed by the same α-amanitin-resistant RNA polymerase activity which is induced upon viral infection (8, 12, 13). In transient expression assays, *trans* activation of the *polh* promoter requires all the genes which are required to transactivate the promoter of *vp39*, a late gene encoding the major capsid protein (20, 21, 24, 29–33). However, the relatively low level of expression from the *polh* promoter in these assays suggests that an additional factor(s) is required to activate very late gene expression and remains to be identified.

cis-acting elements of representative late and very late promoters have been defined by mutational analyses (23, 26, 35, 36, 47). Late and very late transcripts generally initiate within the sequence TAAG. Linker scan analyses of both the *polh* and late *vp39* promoters confirm that this tetranucleotide motif is absolutely essential for promoter activity (23, 26, 38). In addition to the essential TAAG sequence, linker mutations have defined a role for the 8 bp upstream and 6 bp downstream of the TAAG sequence in the regulation of *vp39* (23). In contrast, linker mutations of the *polh* promoter show that the 50 bp from the TAAG transcriptional initiation site to the ATG translational initiation codon of the *polh* open reading frame (ORF) is the primary determinant of *polh* expression (26). On the basis of the effects of these mutations on the temporal regulation of *polh* transcription, the stability of *polh* RNAs, and the transcriptional initiation rate, Ooi et al. (26) concluded that the DNA sequences specifying the untranslated leader region of *polh* RNA are necessary for the "burst" of transcription observed during the very late occlusion phase. The phenotypic effects of the promoter mutations suggest that stimulation by a transcriptional activator, rather than inactivation of a repressor, is probably responsible for the very late burst of transcription. The *cis*-acting elements of the *p10* promoter have not been analyzed in as much detail, but the *p10* and *polh* promoter structures may be similar (36, 47). Because of subtle differences in timing of expression or competition between *p10* and *polh* promoters, however, it has been suggested that the regulation of *p10* may differ from that of *polh* (5, 40, 44).

In this study, we have characterized a temperature-sensitive (*ts*) mutant of Ac*M*NPV, *ts*B837, which produces normal levels of budded virus at the restrictive temperature (33°C) but fails

---

* Corresponding author. Mailing address: Department of Entomology, 413 Biological Sciences Bldg., The University of Georgia, Athens, GA 30602-2603. Phone: (706) 542-2294. Fax: (706) 542-2279. Electronic mail address: Miller@bscr.uga.edu.

to produce occlusion bodies (19). We have found that this mutant is defective in the expression of *polh* and, to a lesser extent, *p10* and that this defect is at the transcriptional level. We have traced the mutation to a novel gene which we have named very late expression factor 1 gene, *vlf-1*.

MATERIALS AND METHODS

Virus and cells. The *Spodoptera frugiperda* (fall armyworm) IPLB-Sf-21 (Sf-21) cell line (45) was adapted for growth at 23 and 33°C in TC-100 medium (GIBCO/BRL, Gaithersburg, Md.) with 10% fetal bovine serum and 0.26% tryptose broth (27). The *ts*B837 mutant of AcMNPV was originally isolated by treatment of wild-type (wt) AcMNPV (L1 variant) (18) with 5-bromodeoxyuridine, and its mutant phenotype was described and partially characterized (19). The *ts*B837 virus was plaque purified once more by us, and two well-separated isolates (C2 and C3) were propagated and titers were determined at the permissive temperature (23°C) for the present studies.

Recombinant plasmids. The plasmid pXma17 contains the *Sma*I-D genomic fragment of AcMNPV cloned into pA-CYC177 (6), corresponding to 39.6 to 49.6 map units (m.u.) (16). Restriction fragments of this plasmid were subcloned into Bluescript II KS+/− (Stratagene, La Jolla, Calif.) to generate the following plasmids: pSDPS5, containing a *Pst*I-*Sst*I fragment (41.2 to 44.9 m.u.) (20); pXB6, containing a *Sma*I-*Bgl*II fragment (39.5 to 44.5 m.u.); pXA7, containing a fragment from 45.2 m.u. to a *Sma*I site at 49.6 m.u.; pXK21, containing *Sst*I-G (45.1 to 47.2 m.u.); pXL8, containing *Sst*I-H (47.2 to 48.1 m.u.); pXBS2.1, containing a *Bst*EII-*Sma*I fragment (48.2 to 49.6 m.u.); pXSS1.3, containing *Sst*II-N (47.9 to 48.9 m.u.); pXABB4.5, containing a subclone of pXA7 from 45.2 m.u. to a *Bst*EII site at 48.2 m.u.; pXABgE, containing a subclone of pXA7 from 45.2 m.u. to an *Eco*RV site at 46.4 m.u.; and pXAEBs, containing an *Eco*RV-*Bst*EII fragment (46.4 to 48.2 m.u.) (Fig. 1).

Plasmid ptsES, which was used to sequence the region of the *ts*B837 genome determined by marker rescue to contain the *ts* mutation, was constructed by cloning the *Eco*RV-*Sma*I fragment (46.4 to 49.6 m.u.) from *ts*B837 into Bluescript II KS+/−. A hybrid fragment containing sequences from both wt and *ts*B837 DNAs was constructed by removing the *Eco*RV-*Bst*EII fragment (46.4 to 48.2 m.u.) from the wt AcMNPV pXA7 plasmid and replacing it with the comparable *Eco*RV-*Bst*EII fragment from the *ts* mutant plasmid, ptsES. The resulting plasmid, pXAtsEB, was sequenced across the *Eco*RV and *Bst*EII junctions, as well as in the internal region, confirming the presence of the *ts*B837 mutation. We used pXAtsEB in marker rescue experiments, and there was no evidence of positive rescue of *ts*B837 viral DNA (data not shown). We also cotransfected pXAtsEB with wt AcMNPV DNA to construct a *ts* virus, *ts*B837W, having the *ts*B837 mutation in a known wt background.

Marker rescue experiments. The *ts*B837 mutation was initially mapped by using a library of overlapping AcMNPV genomic clones that has been previously described (29), and it was further mapped by using subclones of pXma17 (described above). Marker rescue experiments were performed with Sf-21 cells adapted for growth at 33°C. Cells were cotransfected with *ts*B837 viral DNA (1 μg) and AcMNPV genomic clone DNA (1 to 6 μg) by calcium phosphate coprecipitation (27). Following transfection, the medium was replaced with complete TC-100, and the cells were incubated for 72 h at 33°C before being examined for the presence of occluded virus.

DNA sequencing. A set of nested deletion clones of pXAEBs was generated with exonuclease III and mung bean nuclease

(11). By utilizing double-stranded DNA templates, both strands were sequenced across the entire region by the dideoxynucleotide chain termination method (42). Sequencing of *ts*B837 DNA within ptsES was facilitated by oligonucleotide primers.

Metabolic labeling and protein analysis. Sf-21 cells adapted for growth at either 23 or 33°C were infected with wt or *ts*B837 virus at a multiplicity of infection equal to 20 PFU per cell at a density of 10^6 cells per 35-mm-diameter dish. After 1 h of adsorption, virus was removed and medium was replaced with complete TC-100 before the cells were returned to either 23 or 33°C. For analysis of total protein synthesis, complete TC-100 was replaced with TC-100 lacking both cysteine and methionine 1 h before and during the radiolabeling period. At each time point, the cells were labeled for 1 h in the presence of 25 μCi of Tran^{35}S-Label (ICN; ca. 1,000 Ci/mmol), which is a mixture of labeled cysteine and methionine. Samples were prepared and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on 12% gels (27). For the analysis of p10 production in cells infected with either wt or *ts*B837 virus, cells were treated in the same manner but were preincubated for 1 h in Sf-900 II SFM medium (GIBCO/BRL) lacking L-leucine before being labeled for 1 h with 20 μCi of L-[3,4,5-^{3}H]leucine (New England Nuclear; 168 Ci/mmol). Cells were harvested, and proteins were analyzed by SDS-PAGE with 15% polyacrylamide gels. Low-range-molecular-weight protein standards were purchased from GIBCO/BRL.

RNA isolation. Total cellular RNA was isolated from mock-infected, wt AcMNPV-infected, or *ts*B837-infected Sf-21 cells adapted for growth at 33°C. A total of 5×10^6 cells per 100-mm-diameter dish were infected at a multiplicity of infection of 20 PFU per cell. Viral adsorption was done for 1 h at room temperature prior to incubation at 33°C for the remainder of the infection period. Separate plates of cells were treated with cycloheximide (100 μg/ml) for 30 min before addition of the viral inoculum and then were maintained in cycloheximide during and after viral adsorption. In addition, some plates were treated with an inhibitor of DNA synthesis, aphidicolin, at 5 μg/ml after the 1-h viral adsorption period. The cycloheximide- and aphidicolin-treated cells were collected at 12 h p.i. Total cellular RNA was isolated by use of a guanidinium isothiocyanate procedure (7).

Northern (RNA) blot analysis. Total RNA (25 μg per lane) from infected cells was denatured by glyoxalation (22) and electrophoresed through a 1.2% agarose gel as described previously (27). The marker lane, which contained the 0.24- to 9.5-kb RNA ladder (GIBCO/BRL), was denatured with glyoxal before gel electrophoresis and stained separately with ethidium bromide. The remainder of the gel was transferred onto a nylon blotting membrane and hybridized at 42°C in the presence of 50% formamide. The probe used to detect *polh* transcripts was a gel-purified *Eco*RV-*Kpn*I fragment from pRIRV, a plasmid containing the right half of the AcMNPV *Eco*RI-I fragment extending from the *Eco*RV site to the *Eco*RI site. A probe specific for the 603 ORF was prepared by labeling pBSPRI-I(1), which contains the *Mlu*I-*Eco*RV fragment of AcMNPV *Eco*RI-I. The probe used to detect *p10* transcripts was generated by labeling plasmid pQP10, which contains all of AcMNPV *Hin*dIII-Q and *Hin*dIII-P. Transcripts of *vp39* were detected by using a *Nar*I-*Sca*I fragment probe made from the plasmid pSTSNM (43). A strand-specific probe complementary to *vlf-1* transcripts was prepared by using T3 RNA polymerase (GIBCO/BRL) and a deletion clone generated for sequencing which extended to approximately 700 bp upstream of the *Bst*EII site (48.2 m.u.) (arrows in Fig. 2).

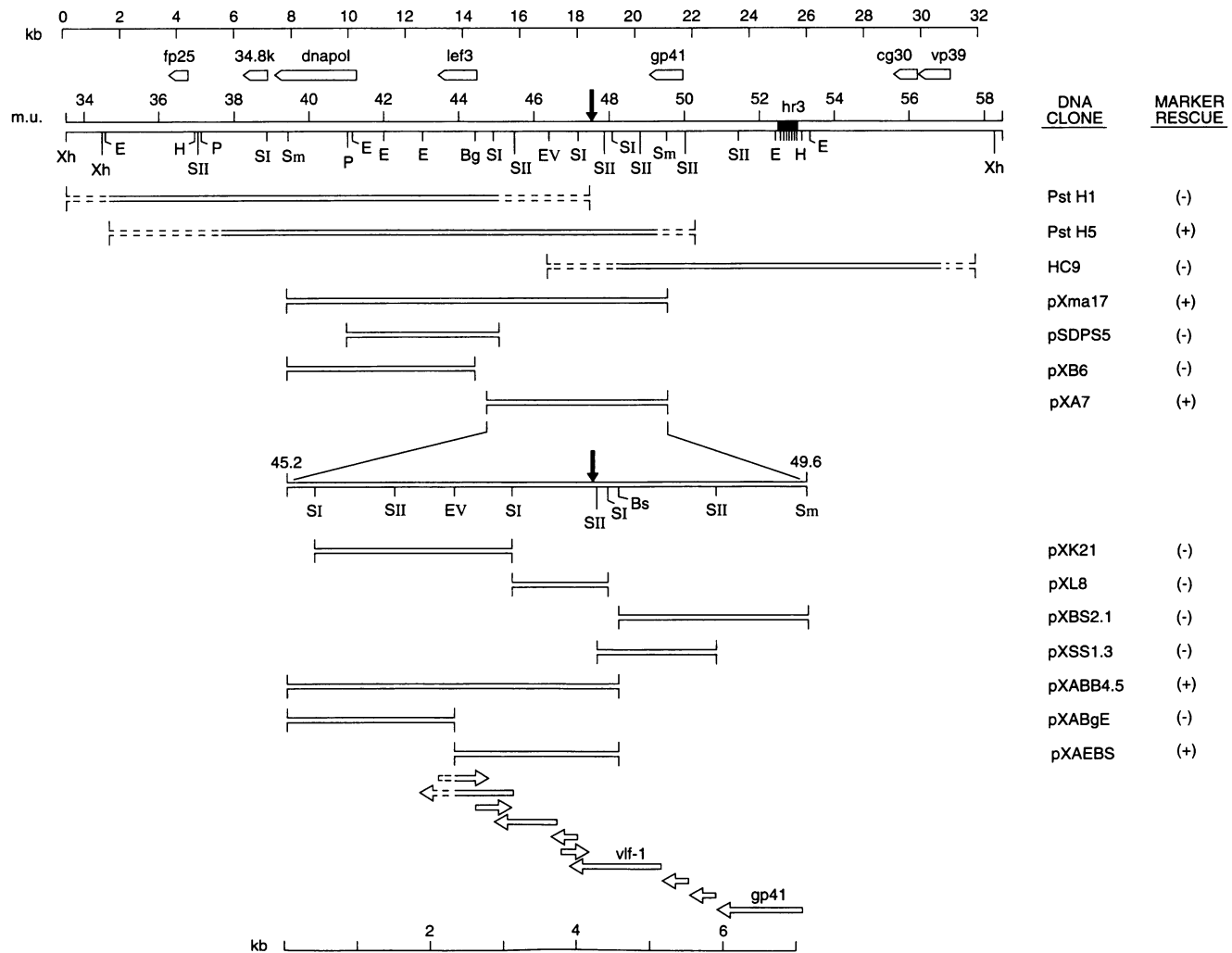**Multiple sequence alignment.** Protein sequence homology searches were performed by using the predicted amino acid

FIG. 1. Schematic diagram of clones of the AcMNPV genome from 34 to 58 m.u. and their abilities to marker rescue tsB837. Locations of previously identified and characterized genes in this region are indicated. PstH5, PstH1, and HC9 are overlapping lambda phage clones of AcMNPV DNA with approximate endpoints indicated by the broken lines. All other clones are plasmid subclones of the SmaI-D fragment of AcMNPV, pXma17. The results of marker rescue experiments with tsB837 appear in the right column. Clones that were able (+) or unable (−) to rescue tsB837 are indicated. The orientations and relative locations of predicted ORFs with sizes greater than 50 amino acids are indicated by open arrows on the lower section. The solid black arrow at approximately 47.5 m.u. shows the location of the single nucleotide difference detected in the EcoRV-BstEII fragment. The ORF affected by this mutation, vlf-1, is labeled, along with gp41, which was previously described by Whitford and Faulkner (48). E, EcoRI; Xh, XhoI; H, HindIII; SII, SstII; P, PstI; SI, SstI; Sm, SmaI; Bg, BglII; EV, EcoRV; Bs, BstEII.

sequence of VLF-1 (accession number S36692; AcMNPV hypothetical protein ORF 1137) and the basic local alignment search tool (BLAST) (2) to search the National Center for Biotechnology Information nonredundant peptide sequence database. Sequence alignments were conducted by using the Pileup multiple sequence alignment program of the Genetics Computer Group (Madison, Wis.) sequence analysis software package. Pileup initially calculates similarity scores by pairwise alignments of each possible pair of sequences. If necessary, the program inserts gaps to produce the alignment, inserting the gap at the same position in all the sequences for a particular cluster.

The following list includes the accession numbers for the sequences used in the multiple sequence alignments; the sequences were derived from either the GenBank or Swiss-Prot database: VLF1 AcMNPV (S36692); Strep int, *Streptomyces ambofaciens* Intr_stram (P15435); Tn2603, *Escherichia coli* Rgnlacb_1 (J02967); L54, Vint_Bpl54 (P20709); Cre, Phage MyP1cre_2 (X03453); P4, Phage MyP4cg_4 (X51522); HK22, Phage STHK022_1 (X51962); Lambda, Phage Lamcg_29 (J02459); Tn554A, *Staphylococcus aureus* ISTN554_1 (X03216); Tn554B, *S. aureus* ISTN554_2 (X03216); P22, Phage Vint_ Bpp22 (P04890); HP1, Phage Vint_Bphp1 (P21442); pColBM, *E. coli* Ecoorid_2 (J03312); Rci, *E. coli* Cibrci_7 (D90039); Mini F, *E. coli* Protein D Fplmcg_4 (M12987); P2, Phage PP2Int_1 (M27836); Phi186, Phage P0186CON_4 (X04449); Phi80, Phage Vint_BPPH8 (P06155); FimB, *E. coli* EcFimbe_1 (X03923); FimE, *E. coli* Ecfimbe_2 (X03923); Tn4430, *Bacillus thuringiensis* Btpgi2xx_1 (X13481); Dlp12, *E. coli* Intd_Ecoli (P24218); pSDL2, *Salmonella dublin* (A38114); Xprb_Ec, *E. coli* Xprb_Ecoli (P21891); Tn2426, *Shigella sonnei* TrnTn21 cas_1 (M86913); Phi42, Phage U01872_2 (U01872); Tn916, *Enterococcus faecalis* Trn916ent_10 (L15633); Flp_Yeast, Fl-p_Yeast (P03870); pSM1, Yeast Ysolsm1_4 (M18275); pSB2,

Yeast Ysosb2_1 (M18274); pSB3, Yeast Zppsb3_1 (X02608); pYSR1, Yeast Zrpsr1_2 (X02398); and pKD1, Yeast Flp_Kudr (P13783). Two protein sequences not shown have sequences in region A and region B identical to those of Tn2603: *Pseudomonas aeruginosa* 38.4K protein (JQ0300) and *Klebsiella pneumoniae* integrase (S32184).

A second group of proteins used for a multiple sequence alignment included Human eIF-4A (P04765), *Drosophila* eIF-4A (Q02748), Yeast eIF-4A (P10081), Dros ME31B (P23128), Human P54 (P26196), *Drosophila* Deadbxa (Z23266), Tobacco NeIF4A2 (X61205), Tobacco NeIF4A3 (X61206), and *Arabidopsis* eIF4A1 (X65052); not shown are Mouse eIF-4A (P10630) and Rabbit eIF-4A (P29562).

Nucleotide sequence accession number. The sequence data from this paper have been deposited with the EMBL/GenBank libraries under accession number L35905.

RESULTS

Mapping of the *ts*B837 mutation. Marker rescue experiments were performed to determine the general location of the mutation in *ts*B837. Viral DNA from *ts*B837 was cotransfected individually with a set of overlapping clones (29) which collectively span the entire AcMNPV genome. Rescue of the occlusion-negative phenotype to the wt phenotype was scored by the presence of polyhedral inclusion bodies in transfected Sf-21 cells held at the nonpermissive temperature (33°C). The mutation in *ts*B837 did not map to *polh*. Instead, the mutant was rescued by the PstH5 (34.8 to 50.3 m.u.) and pXma17 (39.5 to 49.6 m.u.) clones but not by two other overlapping clones in this region, PstH1 and HC9 (Fig. 1). Subclones spanning pXma17 were then tested for their ability to rescue. Plasmids pSDPS5 and pXB6, covering the left half of *Sma*I-D, were negative, while pXA7, covering the right half, was positive. Thus, the mutation was localized to the right side of *Sma*I-D between 45.2 and 49.6 m.u. Subclones of pXA7, indicated in the lower half of Fig. 1, were then tested for their abilities to rescue *ts*B837. Two of these clones, pXABB4.5 (45.2 to 48.2 m.u.) and pXAEBS (46.6 to 48.2 m.u.), were positive. As the size of the fragments decreased, the level of rescue observed decreased. The smallest subclone which provided reliable rescue data was the 2.2-kb pXAEBS DNA; rescue was not observed with a smaller plasmid, pXL8, probably because of the small size of the fragment. The dependence of marker rescue on fragment size was noted previously (39) and is probably due to the decreased frequency of homologous recombination with smaller fragments.

Nucleotide sequence of the 46.4- to 48.5-m.u. regions of AcMNPV and *ts*B837. The wt AcMNPV insert within pXAEBs (46.4 to 48.2 m.u.), which rescued *ts*B837, was sequenced (Fig. 2). A portion of this sequence (nucleotides 756 to 2635) corresponds with but is not identical to a sequence published recently (15). A genomic clone of *ts*B837, corresponding to m.u. 46.4 to 48.5, was also sequenced (nucleotides 1 to 2635). The wt and *ts*B837 sequences differed only at a single nucleotide, which was changed from C to T. This mutation is located at nucleotide 1915, which was within a 1.1-kb ORF oriented in the direction opposite to that of the standard AcMNPV map. The mutation from G to A on the coding strand would result in a predicted codon change from a cysteine to a tyrosine residue. Our sequence differed from that of Kool et al. (15) at four positions, which are indicated in boldface in Fig. 2 (nucleotides 1145, 1154, 1183, and 1869). These differences were observed by us in both the wt AcMNPV and *ts*B837. Only one of those differences (nucleotide 1869) was in the same ORF as that affected by the *ts*B837 mutation, designated ORF

1137 by Kool et al. (15), and it did not change the predicted amino acid sequence of the gene product. The full sequence of AcMNPV became available during the review of this paper (accession no. L22858). Comparison of the four identified differences between our sequence and that which had been previously published (15) with the full sequence shows that our sequence agrees with that of Ayres et al. (4) at positions 1145, 1154, and 1183, leaving only the difference noted in our sequence at position 1869.

Protein synthesis patterns in wt AcMNPV- and *ts*B837-infected cells. At the permissive temperature of 23°C, the kinetics of protein synthesis in Sf-21 cells infected with either wt or *ts*B837 mutant virus were virtually identical, and by 48 h p.i., normal levels of polyhedrin production were observed with all viruses (Fig. 3A). In contrast, polyhedrin was virtually absent in the *ts*B837-infected cells incubated at 33°C (Fig. 3B) compared with the abundant levels observed at 33°C in the wt-infected cells. At 33°C, the infection by both wt and *ts*B837 proceeded at a more rapid rate than at 23°C. By 24 h at 33°C, abundant polyhedrin synthesis occurred in wt-infected cells, but synthesis was greatly reduced in *ts*B837-infected cells at the nonpermissive temperature. The weak band at the polyhedrin position in the *ts*B837 patterns likely corresponds to another protein known to migrate at this position (34). Synthesis of a smaller, ca. 20-kDa polypeptide which is regulated like polyhedrin and may be polyhedrin related was also reduced in *ts*B837-infected cells held at 33°C. Several polypeptides (e.g., polypeptides of 25 and 50 kDa) were synthesized in greater abundance at 24 h p.i. in *ts*B837-infected cells than in wt-infected cells, suggesting that the shutoff of synthesis of some polypeptides may not occur normally at this time p.i. At earlier times p.i. (i.e., 6 and 12 h p.i.), however, the patterns are temporally and quantitatively similar, including the transition between the early and late phases and the shutoff of host protein synthesis.

We then specifically examined the effect of the *ts*B837 mutation on the synthesis of p10, which lacks both cysteine and internal methionine residues. Cells were radiolabeled with [³H]leucine, and total cellular protein from infected cells at 33°C was analyzed on a high-percent-acrylamide gel. Normal levels of p10 were observed in wt-infected cells by 24 h p.i.; however, the cells infected with *ts*B837 showed a significant reduction in p10 production (Fig. 4), although some p10 was synthesized.

Northern blot analysis. Northern blot analyses were performed to further define the effect of the *ts*B837 mutation on the transcription of specific late and very late genes. RNA was isolated at selected times from infected cells grown at 33°C, and *polh* transcripts were identified with a probe extending from the *Eco*RV site at position −93 to the *Kpn*I site at position +637. Since the probe was double stranded, it hybridized both to *polh* sense transcripts (1.3, 3.4, and 4.9 kb) and to a late 3.2-kb antisense RNA (Fig. 5, upper left panel; 3.2- and 3.4-kb transcripts comigrate in this gel) known to extend from the 1629 ORF downstream of *polh* through the 603 ORF upstream of *polh* (25). The levels of *polh* sense transcripts were dramatically reduced in the *ts* mutant-infected cells compared with the abundant *polh* sense transcripts observed for wt-infected cells by 24 and 48 h p.i. Levels of the antisense RNA (3.2 kb) were comparatively constant from 12 to 48 h p.i.

The same RNA blot was also used to study transcription of the 603 ORF specifically (Fig. 5, lower left panel). As noted with the *polh* probe, the levels of the 3.2-kb RNA from the 603 ORF were similar at 12 h p.i. for both wt- and *ts*B837 mutant-infected cells; levels of the 3.2-kb RNA declined thereafter. Since this same blot was stripped before it was

```
       EcoRV          .          .          .          .          .          .          .          .          .
   1   GATATCTTTGAACGTGCACGTTTTCAATTTGTCCTTATCGGCCAACGCAAGTTTCAATTCGCTCTGTAAAGTTTCTAAAATTTTGTCTTTATTGTTGTCA    100

 101   AATTCGTGCGTGTTGCGTTCCAACCACAATTTGAACGGCTCGTCGACAAAAATCGTGCGCAACACCTCGTACAACTGTCTGCCTAACGTGTACACTTGCT    200

 201   CGTATTCTTTCATGCTGACCTCTTTGCTAACGTACATTACTAAAAAATCTACAAGTATTTTCAAACATTTGTAATAGGCGACGTATTTTGATTTAAGTTT    300

 301   TAAACCGTCCACCGTGTATTCGTCCACGTTCGCATCGACCACTTTTCGATTATTATCGCCGCTTGTTGCCGGCGCGTCGGCCTGTTCGGTTTTAACTATA    400

 401   TCCGGTTCAATATTTAAAGTTTCAAAAGATTTAATGGCATTCATAAAATCATCTTTTTGCTTTGGCGTGGTCAATGGTAAATCTATCGAGGAGTTGTCGT    500

 501   CCGTGTGCTCTTCGGGCACGCTGTTCAGACGTAACGTAATCTTTTTGGGATCGTCTTCATCGGGTATCAAATCGGCTTTAATTTTATTAGAATTGAGCAA    600

 601   CGACATGGTGGTCGCTTGTAAATTTAATAAATTAATTAAAGACTGAAATTGTATATTGCACAAATTTATTTTCATTTTTATTGATCTTACTATTAATACG    700

                                                                      SstI
 701   ATGGCAGTTGGTATGCTTCATCCATTTTTGTGACTAGAAAATTTGCTAAAAAACTGAGCTCGTCCTGTGTTAAAACGTTGTCGTCCACGAATCTATGCAA    800

 801   TGTAAATGTTACACTGACATTGTTTAACAATGCATGTATTAAAAAAATCAACCTGTCGCCTACTGAGTTTATTAGAAGAGTCGACCGTTTCTACTAGTTTG    900

 901   TAGATTTTGTTATTTTCAATTTCATTGTTTAAAAACATGTTAACTACTCGTTTGAGTTTAAGCGAAAAATCCTTGTCCGGATAGACTTGTTCGCACAGCC   1000

1001   AATTGCTAAGAGTGGTTTTGACCACGGACACCTTGGTGGTGAACGTCGTCGATTTGACCAGTTCGGTGAAAAAGTTTTTCATTAAATTGGACATTTTAAC   1100

1101   AAACACTTATCAATCTATTGAGCTGGTATTTTTGTTTAGAATCGCATCAAGCGCTTGCTCGATCTCCAATTTTTTTCGGACGCTCTTAGCTTTATGACTC   1200

1201   GGTATGTCTTCTACGGTAGACTCGGTGTTCTTACTTATAATGGCCGGGCTGACGATAATAAACACGAGAAACAATATGAGCAGATACAAAAAGATGCTGT   1300

1301   TTTCCTTTTTGTCATACACTAGGCTAAATATGGCCAGTGCGCCCAACAACAAATATAAATTCATTTTTATTCCCTTACTCTATTCGTTGCGATAGTACAA   1400
 379                                                                                    *  E  N  R  Y  Y  L    374
                                                                       r-
1401   CAACGATTCTCCCGACGAACCGGACGAATTGCGATTATGCTGCGCGTCGTCGTCGTCGTTGTTGTTCTCCTCTTCGCTGCTCGTTTCGTCTAAACCTATA   1500
 373    L  S  E  G  S  S  G  S  S  N  R  N  H  Q  A  D  D  D  D  N  N  N  E  E  E  S  S  T  E  D  L  G  I     341

1501   TTGTATTTGTTCAAGTAATGTTTGGTGCTTGCGGAGGATTCGTGGTTCATTAATTTGGCCACTTTTTGTAAAGGCACGCCGCTATTGTATAGGTTACTGC   1600
 340    N  Y  K  N  L  Y  H  K  T  S  A  S  S  E  H  N  M  L  K  A  V  K  Q  L  P  V  G  S  N  Y  L  N  S     308

1601   TCAAATAATGTCTTATCATGTTGCTGCGCGGCCGTTCCATCTCGACGCCCGACTCTTCAAGGAGTCGCCTGAAATCTTTGAAGGGCGTCGAGGTGTTTTT   1700
 307    S  L  Y  H  R  I  M  N  S  R  P  R  E  M  E  V  G  S  E  E  L  L  R  R  F  D  K  F  P  T  S  T  N  K   274

1701   AGATATTTGCAAAATGGTCGGGTTTCGTGAATAAATCTCGCGTGCCAATTCCAACGGTTTCATTTTGATGTTGTTGAGTGTGTTATTACGACTGCGTTTT   1800
 273     S  I  Q  L  I  T  P  N  R  S  Y  I  E  R  A  L  E  L  P  K  M  K  I  N  N  L  T  N  N  R  S  R  K    241

1801   CGCTTTAAATTAATCGTGTCGCTGTGCAGTTTTCCTCTTTTAATTAGCACGTTGAGATCGTCCACGCTAAGTTGGCGCGCTTCGTTGATTCGCATACCCG   1900
 240    R  K  L  N  I  T  D  S  H  L  K  G  R  K  I  L  V  N  L  D  D  V  S  L  Q  R  A  E  N  I  R  M  G     208
            T              .           SstII  .
1901   TCCCTAACATGATGCAAAACACTATCGCGCCCCTAATTAGACCGCGGTCGTGAACATAATCGCTGTTGAGCATTTTAATTTTATCATTAATAAAATTTAA   2000
 207    T  G  L  M  I  C  F  V  I  A  G  R  I  L  G  R  D  H  V  Y  D  S  N  L  M  K  I  K  D  N  I  F  N  L   174

                          SstI
2001   TATGGTATCTATTACGTTTTTAAGCATTAAATTCTTTTCCTTTTCCCTGATATTTTTGAGCTCCTTGTCGCGCGGCAGCATAACCATGCGGGGAATTTTG   2100
 173     I  T  D  I  V  N  K  L  M  L  N  K  E  K  E  R  I  N  K  L  E  K  D  R  P  L  M  V  M  R  P  I  K    141
                                                                 <-|
                                                               BstEII         .
2101   TATTCGGGCAAGTTCATCATGTTGGTGTAAAAGTTTATAGTCAACTGTAGTGTTTCTTTGGTGACCGAGCGAAGTTCGAGCATGCGCCTGCACAGTTCTT   2200
 140    Y  E  P  L  N  M  M  N  T  Y  F  N  I  T  L  Q  L  T  E  K  T  V  S  R  L  E  L  M  R  R  C  L  E     108

2201   GGGGATCAATGAGAAGTGTTTGGTTTTCTATCGAGTCAAACTCCTTGTCCAACGAGTACGACATGTCTTCCAGGTGAACATCGTCTACCGAGCAGTACAC   2300
 107    Q  P  D  I  L  L  T  Q  N  E  I  S  D  F  E  K  D  L  S  Y  S  M  D  E  L  H  V  D  D  V  S  C  Y  V    74

2301   AATTTTAATGAATCGAGACTTGTAACTTTTTAAAGTGGTGGGCGCAAACGGTTTGGGGAACATGTACTTGCTCCACAGACTGTTGTTTTTCACCTCGTCG   2400
  73     I  K  I  F  R  S  K  Y  S  K  L  T  T  P  A  F  P  K  P  F  M  Y  K  S  W  L  S  N  N  K  V  E  D    41

2401   GGCGTGCATCGTTGCCGATCGGTGGCCAAATCGAACACGGACTCGAACCGGGGAGCGGATTGAATTTTTATTTTCCAAGAATTAAAATTGTTTTCGTTGC   2500
  40    P  T  C  R  Q  R  D  T  A  L  D  F  V  S  E  F  R  P  A  S  Q  I  K  I  K  W  S  N  F  N  N  E  N      8

2501   GAACATTAAAACCGTTCATTGTGGTTAATCAAATTTATTAAAAACAAAAGGAGAATCGGTGTCAATACTATCCGAATATTGTTGTTGTTCTCTTAATATT   2600
   7    R  V  N  F  G  N  M                                                                                   1

2601   ACGAAATAATATATTACATACAGCAGTAAGAATAA   2635
```

FIG. 2. Nucleotide sequence from 46.4 to 48.5 m.u. The predicted amino acid sequence of *vlf-1* (designated ORF 1137 by Kool et al. [15]) is shown running in the opposite direction with respect to the standard Ac*M*NPV map. The single nucleotide difference in the *ts*B837 DNA is indicated by a C-to-T change at nucleotide 1915. This would represent a change from cysteine to tyrosine at residue 202. The differences between our wt Ac*M*NPV sequence and that reported by Kool et al. (15) are indicated in boldface at nucleotides 1145, 1154, 1183, and 1869. The region labeled for a strand-specific probe complementary to *vlf-1* transcripts is indicated (arrows).
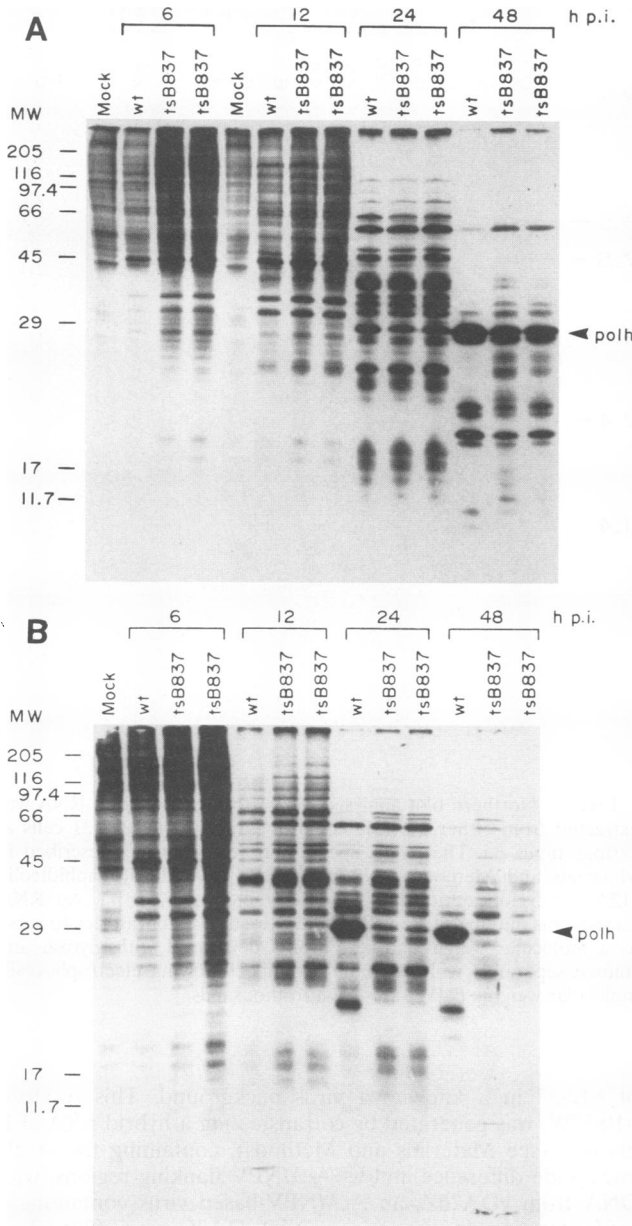
FIG. 3. Time course of protein expression in wt AcMNPV- and tsB837-infected Sf-21 cells at 23°C (A) and 33°C (B). The time p.i. is indicated at the top of each lane. At each time point, cells were pulse-labeled with a mixture of [35S]methionine and cysteine, 1 h prior to collection of the cells. Total cellular protein was analyzed by SDS-PAGE on 12% gels, and the labeled proteins were displayed by fluorography. The positions of protein size standards are indicated on the left, with molecular weights (MW) given in thousands.

FIG. 4. Analysis of p10 protein production in wt AcMNPV- and tsB837-infected cells at 33°C. The time p.i. is indicated at the top of each lane. Cells were labeled for 1 h prior to each time point with [3,4,5-3H]leucine to visualize p10 protein. Total cellular proteins were analyzed by SDS-PAGE on 15% gels, and the labeled proteins were displayed by fluorography. The positions of protein size standards are indicated on the left, with molecular weights (MW) given in thousands.

hybridized to the polh-specific probe (Fig. 5, upper left panel), the differences in the levels of polh RNAs observed between wt and tsB837 lanes of this blot were not due to unequal loading of RNA on the gels.

Northern blot analyses of p10 and vp39 transcription were also performed by using a common blot. This blot was first probed for p10 (Fig. 5, upper right panel) and then stripped and reprobed with a vp39-specific probe (lower right panel). The probe used to examine p10 transcription encompasses both the HindIII-P and -Q fragments. Lower levels of the major p10-specific transcripts (0.7 and ca. 2.6 kb) were observed in cells infected with the mutant virus than in cells infected with the wt virus. However, the difference between the levels of p10 transcripts observed in ts and wt virus-infected cells was not as great as the difference observed for polh transcripts. The probe used to detect p10 transcripts also hybridizes to a 1.5-kb transcript specific for p26 (37), a gene which is transcribed at both early and late times in infection (12). This 1.5-kb RNA was present at similar levels in both wt- and tsB837-infected cells and at relatively uniform levels from 6 through 48 h p.i. The expression of the late gene vp39 was also not affected by the tsB837 mutation, as indicated by the levels of the major 2.2-kb transcript (43) expressed maximally beginning in the late phase of infection (e.g., 12 h p.i.) in both wt- and tsB837-infected cells.

Characterization of VLF-1 transcription. We performed a Northern blot analysis to examine the temporal regulation of vlf-1 transcription. Total cellular RNA was isolated from both wt- and tsB837-infected cells grown at 33°C. Separate samples were treated with either an inhibitor of DNA synthesis, aphidicolin, or an inhibitor of protein synthesis, cycloheximide, and the RNA was extracted at 12 h p.i. A strand-specific RNA probe which would be specific for vlf-1 transcripts was generated from an AcMNPV clone spanning the region of approximately 47.7 to 48.2 m.u., as indicated by the arrows in Fig. 2.
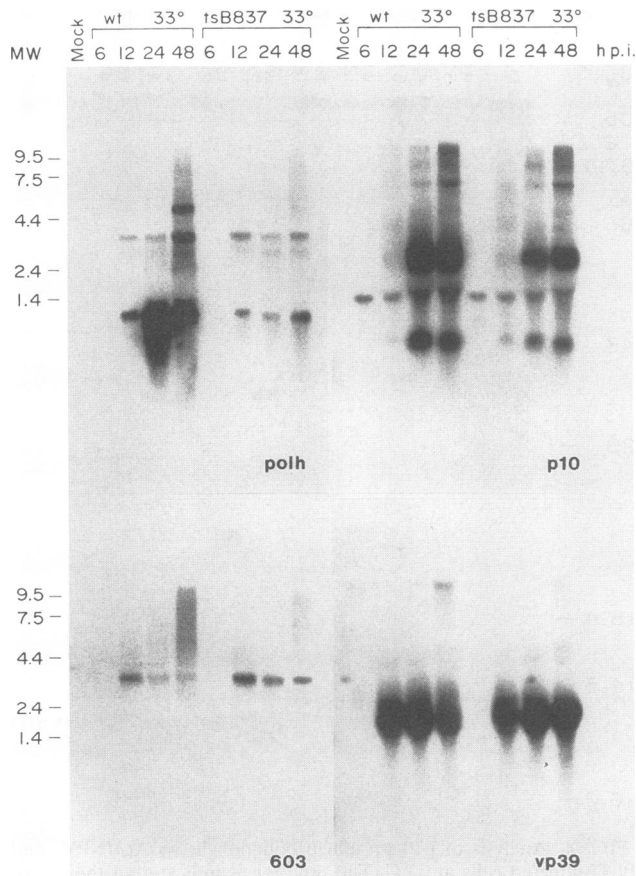
FIG. 5. Northern blot analysis of Ac*M*NPV transcripts from wt Ac*M*NPV- and *ts*B837-infected cells at 33°C. Total cellular RNA was collected from cells at various times p.i. as indicated at the top of each lane. The sizes of RNA standards (RNA Ladder; GIBCO/BRL) were estimated from an ethidium bromide stain of an outside lane removed prior to blotting; molecular weights (MW) are in thousands. Individual probes specific for transcripts from *polh, p10,* 603 ORF, and *vp39* were used, as indicated on each panel, and are described in detail in Materials and Methods.
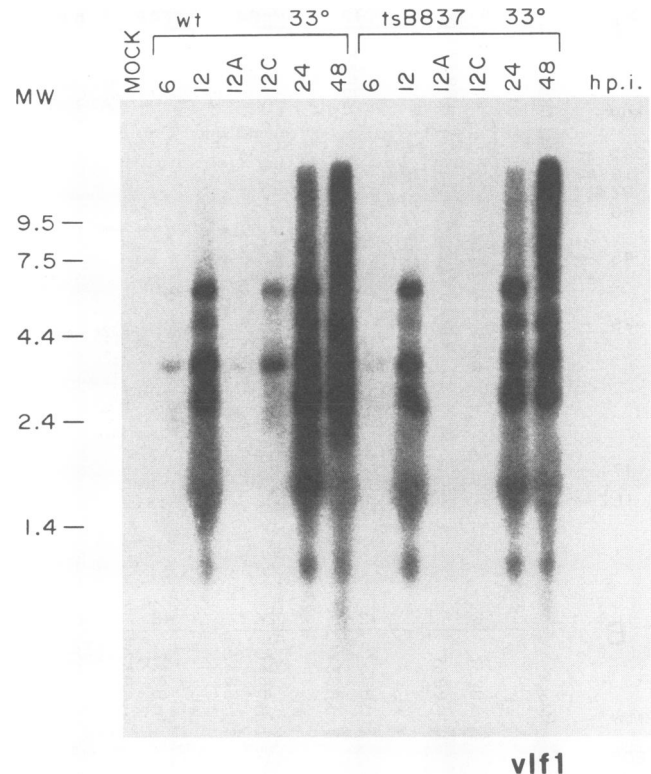

FIG. 6. Northern blot analysis of *vlf-1* transcripts. Total RNA was extracted from either mock-infected or virus-infected SF-21 cells at various times p.i. The strand-specific probe for *vlf-1* is described in Materials and Methods. RNA from cells treated with aphidicolin (12A) or cycloheximide (12C) was extracted at 12 h p.i. An RNA Ladder was purchased from Bethesda Research Laboratories to serve as a molecular weight marker. It was denatured with glyoxal and stained separately with ethidium bromide following electrophoresis; molecular weights (MW) are given in thousands.

At least six transcripts with estimated sizes of 1.2, 1.6, 2.8, 3.5, 4.6, and 6.0 kb were observed. Two transcripts (3.5 and 6.0 kb) were observed at 6 h p.i. and were also visible in both the aphidicolin- and cycloheximide-treated cells following infection with wt virus (Fig. 6). These bands were also visible in the *ts*B837-infected cells with a longer exposure of the autoradiogram (additional exposure not shown) but appeared to be less abundant. The presence of these bands in cycloheximide- and aphidicolin-treated cells indicates a low level of early gene transcription occurring in the absence of DNA or protein synthesis. The difference in the levels of these early transcripts in wt- and *ts*B837-infected cells may suggest that VLF-1 is a component of the virus particle or is exerting some effect early in infection. These two early transcripts persisted through 48 h p.i. The majority of transcripts (1.2, 1.6, 2.8, and 4.6 kb) were first observed by 12 h p.i., were sensitive to aphidicolin and cycloheximide, and were more abundant at 24 and 48 h p.i., properties that are characteristic of late gene expression.

Generation of a VLF-1 *ts* virus. Because *ts*B837 was originally generated by treatment of infected cells with bromodeoxyuridine, we constructed a mutant virus containing the single nucleotide change (nucleotide 1915, affecting amino acid 202)

of *ts*B837 in a known wt virus background. This *ts* virus, *ts*B837W, was generated by cotransfecting a hybrid pXAtsEB plasmid (see Materials and Methods), containing the single nucleotide flanking regions, with wt Ac*M*NPV flanking regions, with DNA from vDA26Z, an Ac*M*NPV-based virus containing a *lacZ* insertion in the nonessential DA26 gene (28). Two recombinants were identified and plaque purified three times. The *ts* phenotype of these recombinants for occlusion body production at 23 but not 33°C was confirmed, thus verifying that this single nucleotide difference confers the *ts* phenotype to *ts*B837.

Sequence homology of VLF-1. We performed a sequence comparison based on the predicted amino acid sequence of VLF-1 by using the program "basic local alignment search tool" or BLAST (2). The baculovirus VLF-1 protein appears to have striking homology with a family of previously identified integrase proteins (1, 3). The results of a multiple sequence alignment performed by using Pileup in the Genetics Computer Group sequence analysis software is presented in Table 1. Residues conserved between VLF-1 and the integrase proteins are indicated in boldface, and the highly conserved arginine residue in region A and conserved arginine, glycine, histidine, and tyrosine residues in region B are marked with asterisks. Table 1 includes several proteins of bacterial origin (i.e., Strep int, pSDL2, Xprb_Ec, and Tn2426) that did not

TABLE 1. Sequence identities in two regions of VLF-1 and a family of integrase proteins[a]

Protein[b]	Position 1[c]	Region A	No. of intervening amino acids[d]	Region B
		*		* ** * * *
VLF-1	203	IMLGTGMRINEARQLSVDDLN	70	ERPRSNMIRHYLSSNLYNSG.VPLQKVAKLMNHES.SASTKHYLN
Strep int				DEWTPRELRHSFVSLLSDRG.VPLEEISRLVGHSGTAVTEEVYRK
Tn2603	139	LLYGTGMRISEGLQLRVKDLD	112	KPATPHTLRHSFATALLRSG.YDIRTVQDLLGHSDV.STTMIYTH
L54	175	VQALTGMRIGELLALQVKDVD	98	KKVTTHTLRHTHISLLAEMN.ISLKAIMKRVGHRDEKTTIKVYTH
Cre	166	IAYNTLLRIAEIARIRVKDIS	98	LAWSGHSARVGAARDMARAG.VSIPEIMQAGGWTNVNIVMNYIRN
P4	237	LSLLTFVRSSELRFARWDEFD	85	TEVCGHGFRTMARGALGESGLWSDDAIERQLSHSERNNVRAAYIH
HK22	205	LAVVTGQRVGDLCRMKWSDIN	78	NPPTFHELR.SLSARLYRNQ.IGDKFAQRLLGHKS.DSMAARYRD
Lambda	205	LAVVTGQRVGDLCEMKWSDIV	78	DPPTFHELR.SLSARLYEKQ.ISDKFAQHLLGHKS.DTMASQYRD
Tn554A	191	LMYEGGLRIGEVLSLRLEDIV	86	IEFTSHMLRHTHATQLIREG.WDVAFVQKRLGHAHVQTTLNTYVH
Tn554B	356	IVQECGMRISELCTLKKGCLL	84	YRFHAHAFRHTVGTRMINNG.MPQHIVQKFLGHES.PEMTSRYAH
P22	208	FALSTGLRRSNIINLEWQQID	81	EDFRFHDLRHTWASWLVQAG.VPISVLQEMGGWESI.EMVRRYAH
HP1	200	ICLATGARWSEAETLTQSQVM	55	KGQLTHVLRHTFASHFMMNG.GNILVLKEILGHSTI.EMTMRYAH
pColBM	66	TLWNTGARINEALALTRGDFS	102	VPVTPHTFRHSYAMHMLYAG.IPLKVLQSLMGHKSI.SSTEVYTK
Rci	148	LALETAMRQGEILALRWEHID	63	EDLHFHDLRHEAISRFFELGSLNVMEIAAISGHRSM.NMLKRYTH
Mini F	63	TLWNTGARINEALALTRGDFS	102	VPVTPHTFRHSYAMHMLYAG.IPLKVLQSLMGHKSI.SSTEVYTK
P2	187	LCLSTGARWGEAARLKAENII	57	MGQATHALRHSFATHFMING.GSIITLQRILGHTRI.EQTMVYAH
Phi186	196	ICLATGARWSEAESLRKSQLA	56	AGQLTHVLRHTFASHFMMNG.GNILVLQRVLGHTDI.KMTMRYAH
Phi80	240	GCRTAEIRLSERSWFRLDDNE	73	AEWSLHDMRRTIATNLSELG.CPPHVIEKLLGHQMVGVMAHYNLH
FimB	40	LCFIHGFRASEICRLRISDID	76	LEIHPHMLRHSCGFALANMG.IDTRLIQDYLGHRNI.RHTVWYTA
FimE	34	LAYRHGMRISELLDLHYQDLD	77	TQTHPHMLRHACGYELAERG.ADTRLIQDYLGHRNI.RHTVRYTA
Tn4430	138	LLAYTGVRISEALSIKMNDFN	71	NVITPHQLRHFFCTNAIEKG.FSIHEVANQAGHSNI.HTTLLYTN
Dlp12	208	FALATGLRKSNIINLEWQQID	81	EDFRFHDLRHTWASWLIQSG.VPLSVLQEMGGWESI.EMVRRYAH
pSDL2				VPVTPHTFRHSYAMHMLYAG.IPLKVLQALMGHKSV.SSTEVYTK
Xprb_Ec				EKLSPHVLRHAFATHLLNHG.ADLRVVQMLLGHSDL.STTQIYTH
Tn2426				KPATPHTLRHSFATALLRSG.YDIRTVQDLLGHSDV.STTMIYTH
Phi42	190	IMICTGCRVSGARNIKIEHIN	71	PIITSHALRHTHCSYLLAKG.VSIHYISKRLGHKNIAITTSVYSH
Tn916	137	ILLKTGLRISEFGGLTLPDLD	99	PHITPHSLRHTFCTNYANAG.MNPKALQYIMGHANIAMTLNYYAH
Flp_Yeast	184	ATFINCGRFSDIKNVDPKSFK	96	NGPKSHIGRHLMTSFLSMKGLTELTNVVGNWSDKRA...SAVARTTYTH
pSM1	133	TSFFNCSRHSDLKNADPTKFE	92	HGPKSHLGRHLMSSYLSHTNHGQWVSPFGNWSAGKDTVESNVARAKYVH
pSB2	183	LTFMNCCRANDLKNADPSTFQ	92	NGPKAHLGRHLMASYLGNNSLKSEATLYGNWSVERQEGVSKMADSRYMH
pSB3	179	MTFFNCCRNNDLMNVDPSTFK	93	HGPKSHLGRHLMNSFLHKNELDSWANSLGNWSSSQNQRESG.ARLGYTH
pYSR1	199	ATFMNCCRADDLKNTDIKTFE	92	NGPKAHLGRHVTASYLSNNEMDKEATLYGNWSAAREEGVSRVAKARYMH
pKD1	178	GAVGNCCRYSDLKNLDPRTFE	93	NGPKSHFGRHLMATFLSREKGKYVSSLGNWAGDRE.IQSAVARSHYSH

[a] Residues conserved between VLF-1 and the integrase proteins are in boldface, and the highly conserved arginine residue in region A and conserved arginine, glycine, histidine, and tyrosine residues in region B are marked with asterisks.
[b] See Materials and Methods for database accession numbers.
[c] Position of the first amino acid listed relative to the N terminus of the predicted protein.
[d] Number of amino acid residues between the end of region A and the beginning of region B.

have any sequence homology with the other proteins within region A but did have the highly conserved residues in region B with the exception of the first histidine. The beginning of region A relative to the N terminus of the predicted protein is indicated, as is the number of residues separating the more N-terminal region A from region B.

A BLAST search was performed by using the sequence contained only in region B of VLF-1. A high degree of similarity was observed in this region with regard to the spacing of several highly conserved residues found in a family of RNA helicases (Table 2).

DISCUSSION

We have identified a gene, vlf-1, which is involved in the regulation of polyhedrin synthesis at the transcriptional level. After confirming that the temperature-sensitive mutant tsB837 presents a wt phenotype at the permissive temperature but forms plaques lacking occlusion bodies at the high temperature (19), we mapped the ts mutation by marker rescue to between 46.6 and 48.2 m.u. of the AcMNPV genome. Comparison of the sequence of wt AcMNPV and tsB837 in this region revealed an alteration of a single nucleotide which is predicted to change amino acid 202 of a 44.4-kDa polypeptide from a cysteine to a tyrosine.

The factor affected by the tsB837 mutation, VLF-1, appears to be specific for regulating the levels of very late transcripts. The kinetics of protein synthesis in wt- or tsB837-infected cells were similar during the early and late phases of infection at the nonpermissive temperature. However, little or no polyhedrin was synthesized in tsB837-infected cells at the nonpermissive temperature, even by 48 h. Northern blot analysis revealed that the levels of both polh and p10 RNAs were reduced. Transcription of several late genes was comparatively unaffected by the tsB837 mutation, either temporally or quantitatively, at the times p.i. examined, so the mutational effect is being exerted primarily at the very late transcriptional level. Although we have not yet determined whether VLF-1 is exerting its effect at the level of transcriptional initiation or RNA stability, it is likely that the effect is at transcriptional initiation, since previous analysis of polyhedrin promoter mutants indicated that transcriptional initiation rather than RNA stability played the major role in regulating the levels of polh RNA. Although the burst of polh expression did not occur in tsB837-infected cells, low levels of polh transcripts were detected in the absence

TABLE 2. Comparison of region B sequence of VLF-1 with RNA helicases

Source and protein[a]	Residue no.[b]	Sequence in region B[c]
AcMNPV VLF-1	295	**ERPRSMNIRHYLSSNLYN** SGVPLQKVAK **LMNHESSAST KHYLNKYNIG**
Human eIF-4A[d]	316	REF**RS**GSSRVLITTDLLA RGIDVQQVSL VI**N**YDLPTNR EN**Y**IHR..**IG**
Drosophila eIF-4A	312	KQF**RS**GSSRVLITTDLLA RGIDVQQVSL VI**N**YDLPSNR EN**Y**IHR..**IG**
Yeast eIF-4A	304	KEF**RS**GSSRILISTDLLA RGIDVQQVSL VI**N**YDLPANK EN**Y**IHR..**IG**
Drosophila ME31B	339	HDFRQGLCRNLVCSDLFT RGIDVQAVNV VI**N**FDFPRMA ET**Y**LHR..**IG**
Human P54	378	HDFRNGLCRNLVCTDLFT RGIDIQAVNV VI**N**FDFPKLA ET**Y**LHR..**IG**
Drosophila Deadbxa	295	SRFKSNQIRTLIATDVAA RGLDIPSVEL V**MN**HILPRTP KE**Y**IHR..V**G**
Tobacco NeIF4A2	323	REF**RS**GSSRVLITTDLLA RGIDVQQVSL VI**N**YDLPTQP EN**Y**LHR..**IG**
Tobacco NeIF4A3	301	AEF**R**GGTTRVLITTDVWA RGLDVQQVSL VI**N**YDLPNNR EL**Y**IHR..**IG**
Arabidopsis eIF4A1	322	REF**RS**GSSRVLITTDLLA RGIDVQQVSL VI**N**FDLPTQP EN**Y**LHR..**IG**

[a] See Materials and Methods for database accession numbers.

[b] Position of the first amino acid listed relative to the N terminus of the predicted protein.

[c] Residues conserved between VLF-1 and the other proteins are in boldface.

[d] Two other mammalian eIF-4 alpha kinases, rabbit and mouse, are also found in existing databases and have sequences in this region identical to that of the human gene shown. Database accession numbers are listed in Materials and Methods.

of functional VLF-1, presumably from the low-affinity TAAG sequence (23) or some leakiness of the ts mutation.

Levels of p10 transcripts were reduced in tsB837-infected cells, but the reduction was not as dramatic as that observed for polh RNA levels (Fig. 5). A difference in effects on p10 and polh expression was also noted at the protein level. If VLF-1 is acting as a transcriptional initiation factor, it is possible that the ts lesion in VLF-1 results in a conformational change. This conformational change may differentially affect the interaction of VLF-1 with the polh and p10 promoters. It is also possible that the ts mutation affects the interaction of VLF-1 with other proteins (e.g., another factor or the polymerase itself); if this is the case, then the transcriptional regulation of p10 and polh differs with regard to the level of dependence on VLF-1.

A differential effect of VLF-1 on polh and p10 transcription is consistent with prior reports. A carefully controlled study noted subtle differences in the temporal regulation of a reporter gene driven by either the polh or p10 promoter; the reporter gene under p10 promoter control was found to be activated slightly earlier but expressed to lower levels later in infection relative to what was observed for the polh promoter (40). Other studies suggest that the two promoters differ with regard to their abilities to compete for limiting factors in infected cells (5, 44). Thus, although both genes are strongly activated during the very late phase of infection, they are likely to be regulated somewhat differently, although our results show that both genes require VLF-1 to achieve maximal RNA levels.

Like that of late genes, transcription of polh and p10 initiates from an essential conserved sequence, TAAG (23, 26, 35, 38, 47). A few additional nucleotides immediately flanking the TAAG sequence are conserved in polh promoters of all baculoviruses studied as well as in the p10 promoter (41). These sequences may provide specific recognition signals for VLF-1. However, it is more likely that VLF-1 interacts directly or indirectly with the transcriptional burst sequence of the polh promoter located in the 50 bp downstream of the TAAG sequence (23, 26). Whether there is a similarly positioned burst sequence in the p10 promoter and whether it provides as strong an impetus to p10 transcription remain to be determined.

The vlf-1 gene is the first gene implicated specifically in very late gene expression. We are currently refining a transient expression assay used to define late expression factor genes (lefs) to be responsive to vlf-1 in order to isolate other genes involved in the regulation of polh and p10 expression. Although transactivation of the polh promoter in transient ex-

pression assays requires all the lefs required by the late vp39 promoter for transactivation, the very late promoters clearly require additional factors for maximum expression.

The predicted sequence of the 44.4-kDa VLF-1 shows striking similarity to the sequences of a large class of resolvases and integrases found in Saccharomyces cerevisiae, prokaryotes, and phages (Table 1). This class of genes includes the integrase genes of temperate phages of E. coli (e.g., lambda and P22), the cre gene of phage P1, the integrases and resolvases of a variety of prokaryotic transposons (e.g., Tn554 and Tn4430), the FLP gene of S. cerevisiae, and the bacterial fimB and fimE genes involved in gene switching. VLF-1 has significant sequence similarity with the two extended regions of sequence similarity which are found in most members of this group of integrases. Region A, which contains a rigorously conserved arginine residue, is located 55 to 112 residues N terminal to region B, which contains, in most cases, the motif $HXXRHX_9$ $GX_{11-12}HX_{8-9}Y$ (1). With the exception of the first H residue of the region B motif, VLF-1 has all the sequences usually conserved in this family of integrases and resolvases. Combined with the remaining sequence similarity in the two regions, the sequence relationship is striking and easily sufficient to assign VLF-1 as a member of this group. However, it is far from clear at this time why a gene which presents the phenotype of a transcriptional activator should be related to a family of integrases and resolvases. At least one of the genes in this class, the gene encoding D protein of mini-F (17), encodes a transcriptional repressor which also functions as a site-specific resolvase.

While examining the sequence relationships between VLF-1 and resolvases, we also noted another interesting sequence relationship. Region B of VLF-1 and this family of integrases and resolvases also has sequence similarity to a region of several RNA helicases from plants, insects, and mammals (Table 2). This subset of a much larger family of RNA helicases contain a motif, $RX_4RX_6LX_3GX_3QXVX_4NX_9Y$, in which the second arginine (R), glycine (G), and tyrosine (Y) residues are spatially and rigorously conserved, as in the integrase-resolvase family. This region, however, is not one of the seven regions of homology conserved in RNA helicases (9, 10). This region of VLF-1, which lies downstream from (C terminal to) the tsB837 mutation, may be involved in nucleic acid interactions.

Viruses with mutations in vlf-1, such as tsB837, are expected to be of considerable utility in both baculovirus expression vector systems and baculovirus pesticide production. Mutants with a temperature-sensitive VLF-1 should facilitate the isola-

tion of recombinant viruses carrying cytotoxic genes at the restrictive temperature yet permit subsequent analysis of the properties and effects of the cytotoxic product at the permissive temperature. It should also be possible to develop transformed cell lines expressing vlf-1, thus allowing the expression of foreign genes carried by vlf-1 deletion viruses only in that cell line(s) expressing vlf-1. Similarly, occluded viral pesticides that are defective in occlusion by virtue of a vlf-1 defect may be produced in vlf-1-expressing cell lines but be limited in their ability to be transmitted in the environment. Other applications of vlf-1 mutants are also envisioned.

ACKNOWLEDGMENTS

REFERENCES

1. Abremski, K. E., and R. H. Hoess. 1992. Evidence for a second conserved arginine residue in the integrase family of recombination proteins. Protein Eng. 5:87–91.
2. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.
3. Argos, P., A. Landy, K. Abremski, J. B. Egan, E. Haggard-Ljungquist, R. H. Hoess, M. L. Kahn, B. Kalionis, S. V. L. Narayana, L. S. Pierson III, N. Sternberg, and J. M. Leong. 1986. The integrase family of site-specific recombinases: regional similarities and global diversity. EMBO J. 5:433–440.
4. Ayres, M. D., S. C. Howard, J. Kuzio, M. Lopez-Ferber, and R. D. Possee. 1994. The complete DNA sequence of Autographa californica nuclear polyhedrosis virus. Virology 202:586–605.
5. Chaabihi, H., M. H. Ogliastro, M. Martin, C. Giraud, G. Devauchelle, and M. Cerutti. 1993. Competition between baculovirus polyhedrin and p10 gene expression during infection of insect cells. J. Virol. 67:2664–2671.
6. Chang, A. C. Y., and S. N. Cohen. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic miniplasmid. J. Bacteriol. 134:1141–1156.
7. Chirgwin, J. M., A. E. Przybyia, R. J. McDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294–5299.
8. Glocker, B., R. R. Hoopes, L. Hodges, and G. F. Rohrmann. 1993. In vitro transcription from baculovirus late gene promoters: accurate mRNA initiation by nuclear extracts prepared from infected Spodoptera frugiperda cells. J. Virol. 67:3771–3776.
9. Gorbalenya, A. E., and E. V. Koonin. 1988. One more conserved sequence motif in helicases. Nucleic Acids Res. 16:7734.
10. Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov. 1988. A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. FEBS Lett. 235:16–24.
11. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28:351–359.
12. Huh, N. E., and R. F. Weaver. 1990. Categorizing some early and late transcripts directed by the Autographa californica nuclear polyhedrosis virus. J. Gen. Virol. 71:2195–2200.
13. Huh, N. E., and R. F. Weaver. 1990. Identifying the RNA polymerases that synthesize specific transcripts of the Autographa californica nuclear polyhedrosis virus. J. Gen. Virol. 71:195–202.
14. King, L. A., and R. D. Possee. 1992. The baculovirus expression system. A laboratory guide. Chapman & Hall, London.
15. Kool, M., R. Broer, D. Zuidema, R. W. Goldbach, and J. M. Vlak. 1994. Nucleotide sequence and genetic organization of a 7.3 kb region (map unit 47 to 52.5) of Autographa californica nuclear polyhedrosis virus fragment EcoRI-C. J. Gen. Virol. 75:487–494.
16. Kool, M., and J. M. Vlak. 1993. The structural and functional

17. Lane, D., R. de Feyter, M. Kennedy, S.-H. Phua, and D. Semon. 1986. D protein of miniF plasmid acts as a repressor of transcription and as a site-specific resolvase. Nucleic Acids Res. 14:9713–9728.
18. Lee, H. H., and L. K. Miller. 1978. Isolation of genotypic variants of Autographa californica nuclear polyhedrosis virus. J. Virol. 27:754–767.
19. Lee, H. H., and L. K. Miller. 1979. Isolation, complementation, and initial characterization of temperature-sensitive mutants of the baculovirus Autographa californica nuclear polyhedrosis virus. J. Virol. 31:240–252.
20. Li, Y., A. L. Passarelli, and L. K. Miller. 1993. Identification, sequence, and transcriptional mapping of lef-3, a baculovirus gene involved in late and very late gene expression. J. Virol. 67:5260–5268.
21. Lu, A., and L. K. Miller. Identification of three late gene expression factors within the 33.8 to 45.4 m.u. region of Autographa californica nuclear polyhedrosis virus. Submitted for publication.
22. 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.
23. Morris, T. D., and L. K. Miller. 1994. Mutational analysis of a baculovirus major late promoter. Gene 140:147–153.
24. Morris, T. D., J. W. Todd, B. Fisher, and L. K. Miller. 1994. Identification of lef-7: a baculovirus gene affecting late gene expression. Virology 200:360–369.
25. Ooi, B. G., and L. K. Miller. 1990. Transcription of the baculovirus polyhedrin gene reduces the levels of an antisense transcript initiated downstream. J. Virol. 64:3126–3129.
26. Ooi, B. G., C. Rankin, and L. K. Miller. 1989. Downstream sequences augment transcription from the essential initiation site of a baculovirus polyhedrin gene. J. Mol. Biol. 210:721–736.
27. O'Reilly, D. R., L. K. Miller, and V. A. Luckow. 1992. Baculovirus expression vectors: a laboratory manual. W. H. Freeman & Co., New York.
28. O'Reilly, D. R., A. L. Passarelli, I. F. Goldman, and L. K. Miller. 1990. Characterization of the DA26 gene in a hypervariable region of the Autographa californica nuclear polyhedrosis virus genome. J. Gen. Virol. 71:1029–1037.
29. Passarelli, A. L., and L. K. Miller. 1993. Three baculovirus genes involved in late and very late gene expression: ie-1, ie-n, and lef-2. J. Virol. 67:2149–2158.
30. Passarelli, A. L., and L. K. Miller. 1993. Identification and characterization of lef-1, a baculovirus gene involved in late and very late gene expression. J. Virol. 67:3481–3488.
31. Passarelli, A. L., and L. K. Miller. 1993. Identification of genes encoding late expression factors located between 56.0 and 65.4 map units of the Autographa californica nuclear polyhedrosis virus genome. Virology 197:704–714.
32. Passarelli, A. L., and L. K. Miller. Identification and transcriptional regulation of the baculovirus lef-6 gene. Submitted for publication.
33. Passarelli, A. L., J. W. Todd, and L. K. Miller. A baculovirus gene involved in late gene expression predicts a large polypeptide with a conserved motif of RNA polymerases. Submitted for publication.
34. Pennock, G. D., C. Shoemaker, and L. K. Miller. 1984. Strong and regulated expression of Escherichia coli β-galactosidase in insect cells with a baculovirus vector. Mol. Cell. Biol. 4:399–406.
35. Possee, R. D., and S. C. Howard. 1987. Analysis of the polyhedrin gene promoter of the Autographa californica nuclear polyhedrosis virus. Nucleic Acids Res. 15:10233–10248.
36. Qin, J., A. Liu, and R. F. Weaver. 1989. Studies on the control region of the p10 gene of the Autographa californica nuclear polyhedrosis virus. J. Gen. Virol. 70:1273–1279.
37. Rankin, C., B. F. Ladin, and R. F. Weaver. 1986. Physical mapping of temporally regulated, overlapping transcripts in the region of the 10K protein gene in Autographa californica nuclear polyhedrosis virus. J. Virol. 57:18–27.
38. Rankin, C., B. G. Ooi, and L. K. Miller. 1988. Eight base pairs

encompassing the transcriptional start point are the major determinant for baculovirus polyhedrin gene expression. Gene **70**:39–49.

39. **Ribeiro, B. R., K. Hutchinson, and L. K. Miller.** 1994. A mutant baculovirus with a temperature-sensitive IE-1 transregulatory protein. J. Virol. **68**:1075–1084.

40. **Roelvink, P. W., M. M. M. van Meer, C. A. D. de Kort, R. D. Possee, B. D. Hammock, and J. M. Vlak.** 1992. Dissimilar expression of *Autographa californica* multiple nucleocapsid nuclear polyhedrosis virus polyhedrin and p10 genes. J. Gen. Virol. **73**:1481–1489.

41. **Rohrmann, G. F.** 1986. Polyhedrin structure. J. Gen. Virol. **67**:1499–1514.

42. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

43. **Thiem, S. M., and L. K. Miller.** 1989. Identification, sequence, and transcriptional mapping of the major capsid protein gene of the baculovirus *Autographa californica* nuclear polyhedrosis virus. J. Virol. **63**:2008–2018.

44. **van Oers, M. M., D. Malarme, J. M. P. Jore, and J. M. Vlak.** 1992. Expression of the *Autographa californica* nuclear polyhedrosis virus p10 gene: effect of polyhedrin gene expression. Arch. Virol. **123**:1–11.

45. **Vaughn, J. L., R. H. Goodwin, G. J. Tompkins, and P. McCawley.** 1977. The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera: Noctuidae). In Vitro **13**:213–217.

46. **Vlak, J. M., F. A. Klinkenberg, K. J. M. Zaal, M. Usmany, E. C. Klinge-Roode, J. B. F. Geervliet, J. Roosien, and J. W. M. Van Lent.** 1988. Functional studies on the p10 gene of *Autographa californica* nuclear polyhedrosis virus using a recombinant expressing a p10-beta-galactosidase fusion gene. J. Gen. Virol. **69**:756–776.

47. **Weyer, U., and R. D. Possee.** 1989. Analysis of the promoter of the *Autographa californica* nuclear polyhedrosis virus p10 gene. J. Gen. Virol. **70**:203–208.

48. **Whitford, M., and P. Faulkner.** 1992. Nucleotide sequence and transcriptional analysis of a gene encoding gp41, a structural glycoprotein of the baculovirus *Autographa californica* nuclear polyhedrosis virus. J. Virol. **66**:4763–4768.

49. **Williams, G. V., D. Z. Rohel, J. Kuzio, and P. Faulkner.** 1989. A cytopathological investigation of *Autographa californica* nuclear polyhedrosis virus p10 gene function using insertion/deletion mutants. J. Gen. Virol. **70**:187–202.