

Identification and Characterization of a Protein Kinase Gene in the *Lymantria dispar* Multinucleocapsid Nuclear Polyhedrosis Virus

DAVID S. BISCHOFF AND JAMES M. SLAVICEK*

Forestry Sciences Laboratory, Northeastern Forest Experimental Station,
USDA Forest Service, Delaware, Ohio 43015

Received 23 August 1993/Accepted 1 December 1993

The *Lymantria dispar* multinucleocapsid nuclear polyhedrosis virus (LdMNPV) gene encoding vPK has been cloned and sequenced. LdMNPV vPK shows a 24% amino acid identity to the catalytic domains of the eucaryotic protein kinases nPKC from rabbits, HSPKCE from humans, APLPKCB from *Aplysia californica*, and dPKC98F from *Drosophila melanogaster*, and homology to several other protein kinases from yeasts, mice, and bovines. The homology suggests that vPK is a serine/threonine protein kinase as defined by Hanks et al. (S. K. Hanks, A. M. Quinn, and T. Hunter, *Science* 241:42–52, 1988). Temporal expression studies indicate that vPK is expressed throughout the infection cycle beginning at 4 h postinfection, first as a delayed-early gene and subsequently as a late gene. Sequence analysis and primer extension reactions confirm the presence of distinct early and late transcription initiation regions. Expression of vPK with a rabbit reticulocyte system generated a 31-kDa protein, which is in close agreement with the predicted size of 32 kDa from the amino acid sequence. Phosphorylation activity of in vitro-expressed vPK was demonstrated by using calf thymus histones.

The *Lymantria dispar* multinucleocapsid nuclear polyhedrosis virus (LdMNPV) is a double-stranded DNA virus that is pathogenic to the gypsy moth (*L. dispar*). The gypsy moth, an insect nonindigenous to the United States, feeds on over 300 species of trees and shrubs and is a serious defoliating pest in northeastern forests. The life cycle of LdMNPV is similar to that of other baculoviruses in that it has two morphologically distinct forms: a nonoccluded budding virus that infects cells within the same larva and an occluded form or polyhedron that enables transmission from one larva to another (for a review, see reference 4).

Baculovirus genes are temporally regulated and can be divided into four categories based on the time of expression (for a review, see reference 10). Early genes are those that precede viral DNA replication and are further divided into two groups: immediate early and delayed early. Both immediate- and delayed-early genes are expressed from host RNA polymerase II and have promoters that resemble typical eucaryotic promoters (11, 18), although delayed-early genes may require other viral proteins as enhancers of transcription. Recently, it has been suggested that there should be only one category of early genes, as *Autographa californica* multinucleocapsid nuclear polyhedrosis virus (AcMNPV) immediate- and delayed-early gene transcription can be efficiently initiated with noninfected *Spodoptera frugiperda* (Sf9) nuclear extracts (12). The levels of transcription are increased when AcMNPV-infected Sf9 nuclear extracts are used. Therefore, the previous distinction between immediate- and delayed-early gene transcription may be due to different levels of promoter activity and transactivation by viral proteins. Late genes are transcribed by a virus-induced α -amanitin-resistant polymerase (11, 59) which recognizes the specific late promoter sequence NTAAG,

where N is usually an A (15, 25, 41). Late and very late gene expression occurs concomitant with DNA replication (21) and includes viral genes required for nucleocapsid and polyhedron formation.

Reversible phosphorylation of proteins is a common mechanism in eucaryotes for regulation of cellular processes, including gene transcription and cell division (for a review, see reference 19). Protein kinases have been identified in vaccinia virus (24), Sendai virus (42), Semliki Forest virus (52), Sindbis virus (52), and RNA tumor viruses (17). Many of these enzymes are virally encoded, including protein kinases from herpes simplex virus (8), cytomegalovirus (51), and African swine fever virus (ASFV) (3). Several host and viral proteins are phosphorylated following infection with baculoviruses (22, 30, 60). Protein kinase activity has also been found to be associated with occluded and extracellular forms of AcMNPV (31) and with purified capsids of the *Plodia interpunctella* granulosis virus (57). In *P. interpunctella* granulosis virus, the protein kinase is thought to phosphorylate nucleocapsid proteins, enabling the release of viral DNA from the capsids (58). A novel protein kinase that is induced in *Bombyx mori* cells after infection with *B. mori* multinucleocapsid nuclear polyhedrosis virus has recently been identified (61). It is not known whether the gene for this protein kinase is located on the viral or host genome.

In this paper, we report the cloning and characterization of the protein kinase gene, vPK, of LdMNPV. vPK exhibits approximately a 24% amino acid identity to the catalytic domains of the eucaryotic protein kinases nPKC from rabbits (33), HSPKCE from humans (2), APLPKCB from *Aplysia californica* (32), and dPKC98F from *Drosophila melanogaster* (46). Expression of the vPK gene in vitro generated a 31-kDa protein, which was in close agreement with the 32-kDa protein predicted from the amino acid sequence. The in vitro-generated vPK exhibited protein kinase activity with histones serving as a substrate. Northern (RNA) analysis reveals that vPK is

* Corresponding author. Mailing address: USDA Forest Service, Northeastern Forest Experimental Station, Forestry Sciences Laboratory, 359 Main Rd., Delaware, OH 43015. Phone: (614) 369-4476. Fax: (614) 363-1437.

initially expressed as a delayed-early gene but late in the infection cycle is expressed as a late gene.

MATERIALS AND METHODS

Cells and virus. *L. dispar* 652Y cells (13) were grown as monolayers in Goodwin's IPL-52B medium supplemented with 6.25 mM glutamine and 10% fetal bovine serum. Cell cultures were inoculated with LdMNPV clonal isolate 5-6 (49) at a multiplicity of infection of 10. Virus was removed after a 1-h adsorption period and replaced with fresh medium.

RNA isolation and Northern blot analysis. Infected *L. dispar* 652Y cells were harvested at various times postinfection (p.i.). Cytoplasmic RNA was isolated as described by Friesen and Miller (9). For inhibitor studies, cells were treated with cycloheximide at 100 µg/ml for 0.5 h prior to virus adsorption or aphidicolin at 5 µg/ml after adsorption. Inhibitors were maintained in the medium throughout the time course. RNA was separated on 1.2% agarose-formaldehyde gels and transferred to nitrocellulose. Northern blots were performed as described by Mahmoudi and Lin (28). Probes were radiolabelled with a nick translation kit from Bethesda Research Laboratories and [α - 32 P]dCTP from NEN.

Construction of a λ gt11 cDNA library. A cDNA library was constructed by using the Riboclone cDNA Synthesis System from Promega. Poly(A) RNA was isolated and purified from *L. dispar* cells infected with LdMNPV at 7 h p.i. First-strand synthesis was accomplished with avian myeloblastosis virus reverse transcriptase by using a poly(T) primer with an adapter containing a unique *Not*I site. After second-strand synthesis of the cDNA, *Eco*RI linkers were ligated to the ends. The cDNAs were then cloned into λ gt11 after digestion with *Eco*RI-*Not*I.

Viral DNA isolation and Southern blot analysis. Nonoccluded virus from plaque-purified LdMNPV isolate 5-6 was isolated from infected 652Y cells and used as a source of genomic DNA for Southern blot analysis. Medium was decanted from the cells, and cellular debris was removed by centrifugation at $550 \times g$ for 10 min. Virus was pelleted by centrifugation at $104,000 \times g$ for 45 min at 4°C. The pellet was resuspended in $0.1 \times \text{TE}$ (1 mM Tris, 0.1 mM EDTA [pH 8.0]) at 4°C overnight. One volume of $2 \times$ DNA extraction buffer (20 mM Tris [pH 7.5], 120 mM NaCl, 20 mM EDTA [pH 8.0], 2% sodium dodecyl sulfate [SDS], and 40 µg of proteinase K per ml) was added, and the solution was incubated for 1 h at 50°C. The solution was adjusted to 1% Sarkosyl and incubated for an additional hour at 50°C. Viral DNA was extracted with 1 volume of Tris-buffered phenol and 2 volumes of chloroform-isoamyl alcohol (24:1) and then precipitated with the addition of 2 volumes of ethanol. Viral DNA was digested with restriction endonucleases and fractionated on Tris-borate-EDTA-0.8% agarose gels. Southern blot analysis was performed on nitrocellulose with probes generated as described above.

Sequencing. The *vPK* sequence was obtained on both strands by the dideoxynucleotide method of Sanger et al. (44). M13 vectors were used to generate single-stranded DNA templates and then sequenced with the Sequenase kit from U.S. Biochemicals. Specific primers for sequencing were synthesized on an Applied Biosystems 381A DNA Synthesizer. Sequence from plasmids was obtained with the fmole DNA Sequencing System from Promega by using protocols supplied with the kit. [α - 35 S]dATP was supplied from NEN. Sequence analysis was done by using MacVector from IBI.

Primer extension mapping of early and late transcripts. Primer extension reactions were performed by the method of Crawford and Miller (6). RNA was isolated from *L. dispar* cells

infected with LdMNPV clonal isolate 5-6 at 6 and 48 h p.i. An 18-base primer that is complementary to the sequence shown in Fig. 3 (positions 320 to 339) was used in the reactions after being end labelled with [γ - 32 P]dATP from NEN. The primer was extended by using Moloney murine leukemia virus reverse transcriptase. Primer extension products were fractionated on 6% polyacrylamide-8 M urea gels and visualized by autoradiography.

In vitro transcription and translation of *vPK*. The 1.2-kb *Pst*I-*Hind*III fragment was subcloned into plasmid pT7/T3 α 18 from Bethesda Research Laboratories. The resulting plasmid, pPH1.2, has *vPK* under control of the T7 promoter. *vPK* was expressed from pPH1.2 by using the T_NT Coupled Reticulocyte Lysate System and T7 RNA polymerase from Promega by using directions provided with the kit. The expressed protein was labelled by the addition of [35 S]Met obtained from NEN. Reaction products were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) and autoradiography.

Determination of protein kinase activity of *vPK*. Kinase activity was determined by using T_NT-expressed *vPK* without the addition of any labelled amino acid. Phosphorylation assays were conducted as described previously (3) with only minor changes as detailed below. Briefly, the kinase reactions were performed in a 20-µl volume of a mixture of 50 mM Tris (pH 7.5), 5 mM dithiothreitol, 5 mM dATP, 1 µCi of [γ - 32 P] dATP, and 10 µg of protein substrate. The reactions were initiated by the addition of 10 µl of T_NT extract, and the mixtures were incubated at 30°C. The reactions were stopped after 10 and 30 min by addition of 20 µl of $2 \times$ SDS sample application buffer (125 mM Tris [pH 6.8], 4% SDS, 40% glycerol, 1 M β -mercaptoethanol, 0.002% bromophenol blue) and analyzed by SDS-PAGE and autoradiography.

Nucleotide sequence accession number. The nucleotide sequence of the LdMNPV *vPK* gene has been deposited in GenBank under accession no. M04322.

RESULTS

Temporal analysis of *vPK*. A λ gt11 cDNA library was constructed by using RNA isolated at 7 h p.i. from *L. dispar* 652Y cells infected with clonal isolate 5-6. Since early viral transcripts had previously been mapped between 11 and 18 kb (6.7 to 11 map units) on the viral genome (49), the 5.1-kb *Eco*RV-I fragment was used as a probe to screen the 7-h cDNA library. Two λ -cDNA clones (λ 5-2 and λ 5-6) which hybridized to this probe were identified and purified. The lengths of the cDNA inserts were determined by digestion of the phage DNA with *Not*I-*Eco*RI. Phage λ 5-2 contained an insert approximately 880 bp in length, and λ 5-6 contained an insert that was approximately 750 bp (data not shown). The cDNA inserts were gel purified from both phages and used as probes on Northern blots to identify the sizes of the transcripts being expressed. Both cDNAs hybridized to the same 950-bp transcript (Fig. 1). This gene was later designated *vPK* on the basis of homology to other protein kinases (see below). *vPK* is first expressed at 4 h p.i. and continues to be expressed throughout infection. The *vPK* cDNA clones also hybridized to larger transcripts (1.6, 2.1, and 3.3 kb) expressed at 48 and 72 h p.i., indicating that at least part of the *vPK* gene is present on overlapping transcripts. These types of transcripts are commonly found in other baculoviruses (26). It has been suggested that overlapping transcripts can be used as a means of viral gene regulation (29).

Genomic mapping and sequencing of *vPK*. *vPK* was localized to the 4.1-kb *Pst*I fragment located at approximately 9.8 to 13.9 kb (6.0 to 8.5 map units) on the viral genome by Southern blot

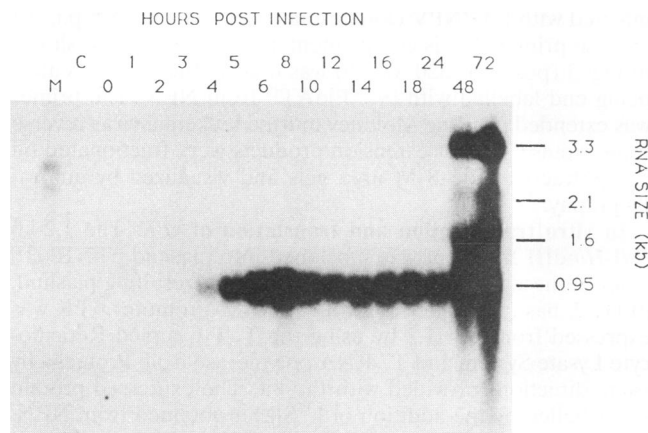


FIG. 1. Expression analysis of the LdMNPV *vPK* gene. *L. dispar* 652Y cells were infected with 10 50% tissue culture infective dose units of LdMNPV isolate 5-6 per cell. At the times indicated above the lanes the cells were harvested, and cytoplasmic RNA was isolated. The RNA was separated by agarose-formaldehyde gel electrophoresis, blotted, and probed with a 32 P-labelled *vPK* cDNA clone. Uninfected cells were used as a control (lane C). Lane M contains size standards. The lengths of the *vPK* gene transcripts are indicated at the right.

analysis using the *vPK* cDNA insert from λ 5-6 (data not shown). Further mapping determined that the entire gene was present on a 1.2-kb *Pst*I-*Hind*III fragment (Fig. 2A) covering the region from 9.8 to 11.0 kb (6.0 to 6.7 map units) on the genome. This fragment was subcloned into plasmid pT7/T3 α 18 from Bethesda Research Laboratories to generate pPH1.2, which has *vPK* under control of the T7 promoter. The nearly intact *vPK* cDNA (from λ 5-2) and the partial cDNA clone lacking the 5' region of the *vPK* gene (from λ 5-6) were used as probes for Southern blotting to determine the direction of transcription of *vPK* (data not shown). These data indicated that *vPK* is translated from left to right with respect to the viral genome.

The restriction map of the 1.2-kb *Pst*I-*Hind*III fragment is shown in Fig. 2A. Specific subclones of this fragment were generated in M13 vectors and then used to determine the nucleotide sequence of the fragment. The sequencing strategy used is outlined in Fig. 2B. Computer analysis of the 1.2-kb *Pst*I-*Hind*III fragment revealed several open reading frames (ORFs) that may encode proteins, with the largest ORF (867 bp) being in frame 2 (Fig. 2C).

Characteristics of the nucleotide sequence. The nucleotide sequence of the 1,247-bp *Pst*I-*Hind*III fragment and the predicted amino acid sequence of *vPK* are presented in Fig. 3. *vPK* begins at position 287 and ends at position 1109. This gene would encode a 274-amino-acid protein with a predicted molecular mass of 32 kDa.

Approximately 183 bp upstream of the *vPK* start codon is a region with the sequence AACGTGAC (position 104). This sequence is similar to the CGT motif, with the consensus sequence A(AT)CGTGTR (where R is a purine residue), that has been identified upstream of many early genes in AcMNPV (7). In addition, there are two identical 13-bp direct repeats upstream of *vPK* beginning at nucleotide positions 39 and 171. These repeats are composed entirely of GC residues with the core sequence GCGCG. GC-rich repeats with this core sequence were identified upstream of early genes in AcMNPV and may function as enhancers of viral transcription (7). Although there is no consensus early promoter sequence

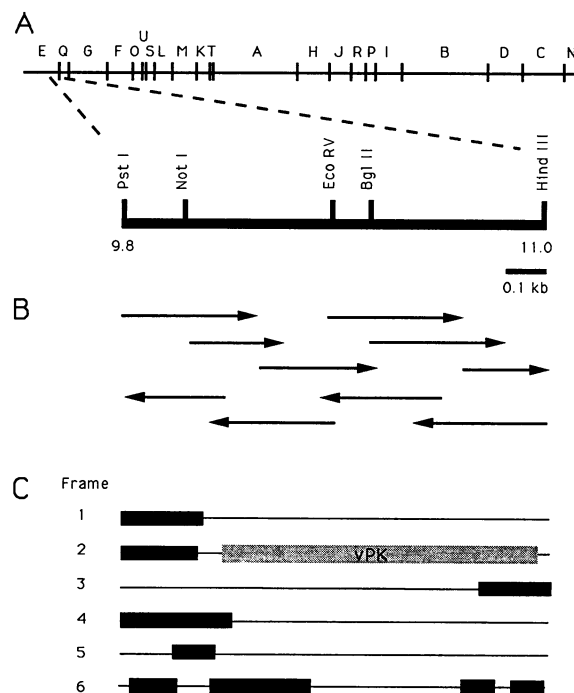


FIG. 2. Genomic location of LdMNPV *vPK*, sequencing strategy, and ORF analysis. (A) *Bgl*II restriction map of the LdMNPV viral genome (39). The enlarged map indicates the 1.2-kb *Pst*I-*Hind*III fragment that encodes *vPK* at 9.8 to 11.0 kb (6.0 to 6.7 map units) on the genome. The relevant restriction sites used in subcloning and sequencing are indicated. (B) Sequencing strategy. The arrows below the map indicate sequences obtained from specific subclones and from oligonucleotide primers. (C) ORF analysis of this fragment in all six reading frames: 1, 2, and 3, from the coding strand of DNA; 4, 5, and 6, from the noncoding strand. The black boxes indicate ORFs that are at least 25 amino acids in length and which begin with an ATG start codon as determined by using the MacVector program from IBI. The shaded box indicates the ORF encoding LdMNPV *vPK* (frame 2).

(TATA box) that can be identified upstream of *vPK*, the AT-rich regions at positions 140 (ATTATC) and 241 (AAATAA) might serve as an early promoter and drive expression of *vPK* at early stages of viral infection. The ATTATC sequence at position 140 is similar to the sequence (ATTATG) that has been identified as a possible promoter sequence in the *CG30* gene of AcMNPV (53). The AT-rich region at position 241 also has the consensus late promoter sequence ATAAG (41) which may drive expression of *vPK* during the late phase of viral replication. It is possible that this late promoter is also acting as a TATA box because of the presence of the other early gene elements not present in AcMNPV. Alternatively, expression of *vPK* may not require a TATA-like promoter. Close examination of the early start site sequence (CTCATTGC) reveals similarity to transcriptional start sites in *D. melanogaster*, with the consensus sequence NTCAGTYN, that do not appear to require TATA boxes (20). Downstream of *vPK* at position 1110 is a consensus polyadenylation site, AATAAA.

Characteristics of the protein sequence. The predicted amino acid sequence of *vPK* was compared with other sequences in GenBank at the National Center for Biotechnology Information by using the BLAST network service (1). *vPK* shows homology to the catalytic domains of protein kinases

Pst I
 1 CTGC AGT CGC CGT ATC CTC GGC GAG TCG TCG CTA CGC ACG GGC GCG
 47 GCG GCG GCG CAC ACC CGA TTG AAA AAT CTC AAC AGG TCG ACG GTT
 92 TCG CGC AGA TCG AAC GTG AGC ACG CCG GGC GAC TCG ACC GCC GAG
 137 GTC ATT ATC TTG TCT ATC AGC CAT TGC GCG CTC GCG GGC GCG GCG
 182 GCG GCG ACC TCC GAC ACG AGC CGC AGC GCC TCG TCG CGG GGA GGG
 227 TAG GCT TCG TTG TGA AAT AAG TTA CTT TCG GGC TCA TTG CCC GCG
 272 TAC GAC TTG AAC ATA ATG GAC GCG CTG ATC GGG GAC TTT GCG GAT
 317 TTT CAC AAA GAG TGC AGC GCG CGC ACC GCG CTC CAC CTC GTC AAC
 362 GGC AAG TTC GGC AAG GTG TCC GTG TGG AAA CAC GGA CCG ACT CAA
 407 AAG TCC TTC TAC AAG CGA ATC GAG CAC AAA CAC TTC AAT GCC
 452 ATC GAG CCG TTC TGA CAC CAC CTG ATG AAG TTC AAC AAA TAT TTT
 497 CTC AGA CTC TTC TAC TCG TTG CAT TCG CTG CGC GAA CAC TTA CTA
 542 GTA ATG GAC TAC ATT CCC GAC GGG GAC CTG TTC GAC CTG ATG CAA
 587 ACA GAG CCC CGG CTG CGG GAG CCA GAG ATA TCC TTG ATC GCC TAC
 632 CAG CTC ATA GAC GCC CTG CAA GCC CTG CAC AAA CAC AAC GTG GTG
 677 CAC AAC GAC ATC AAG CTG GAG AAC GTG CTG TAC AGG CGG TTC GAG
 722 CAG ATC TAC GTG TGC GAC TAC GGG CTG TGC AAG ATC GCC GGT TCG
 767 CCC TCC ACG TTC GAG GGC ACG GTG GAC TAC TTC TCG CCG GAG AAG
 812 ATA AAC AAA CAC GCC GCG GCG GAC TCC GAC TGG TGG GCC GTG
 857 GGC GTG CTG CTC TAC GAA ATA TCC ACC GGC AAA CAC CCC TTC AAG
 902 CTC GAC CAG GAC AGC AGC CTG GAC GTG GAG ACC CTG CAC AAG CGG
 947 CAG ATA CAG CTC GAC GTG ACC TTT CCC GCC GAC TTT GAC AAC CCC
 992 TTC CTA GAA GAG TTT ATT TGT TTT CTG TTG GGC TAT TGC TAC GAG
 1037 TAC AGA GCC CAC AGT TAT GAG GTC ATT CAA AAG AAC ACC TAC TGG
 1082 AAG AGT ATT GTT CAT TGG AAG CAA CGA TAA TAA ATA TTT AAA AAT
 1127 ATA TAT AGT GTT TTA TTA TTT AAA TAA CAA CCA AGG GGT GAG GTT
 1172 AAA AAA AAT AAC AAC CAA CGG TGA GGT TAA AAA AAT ATA TTT ATT
 Hind III
 1217 GCA TGC GGT TAA CAA TAG ATT CAT ACA AGC TT

FIG. 3. Nucleotide sequence of the 1.2-kb *Pst*I-*Hind*III fragment and the predicted amino acid sequence of vPK. Early gene elements that have been identified in AcMNPV are indicated with a single (GC motif) or double (CGT motif) underline. The late promoter sequence at position 241 and the polyadenylation site at position 1110 are boxed. Transcriptional start sites are indicated with single arrows for early start sites or double arrows for late start sites. The position of the poly(A) tail is indicated with an asterisk.

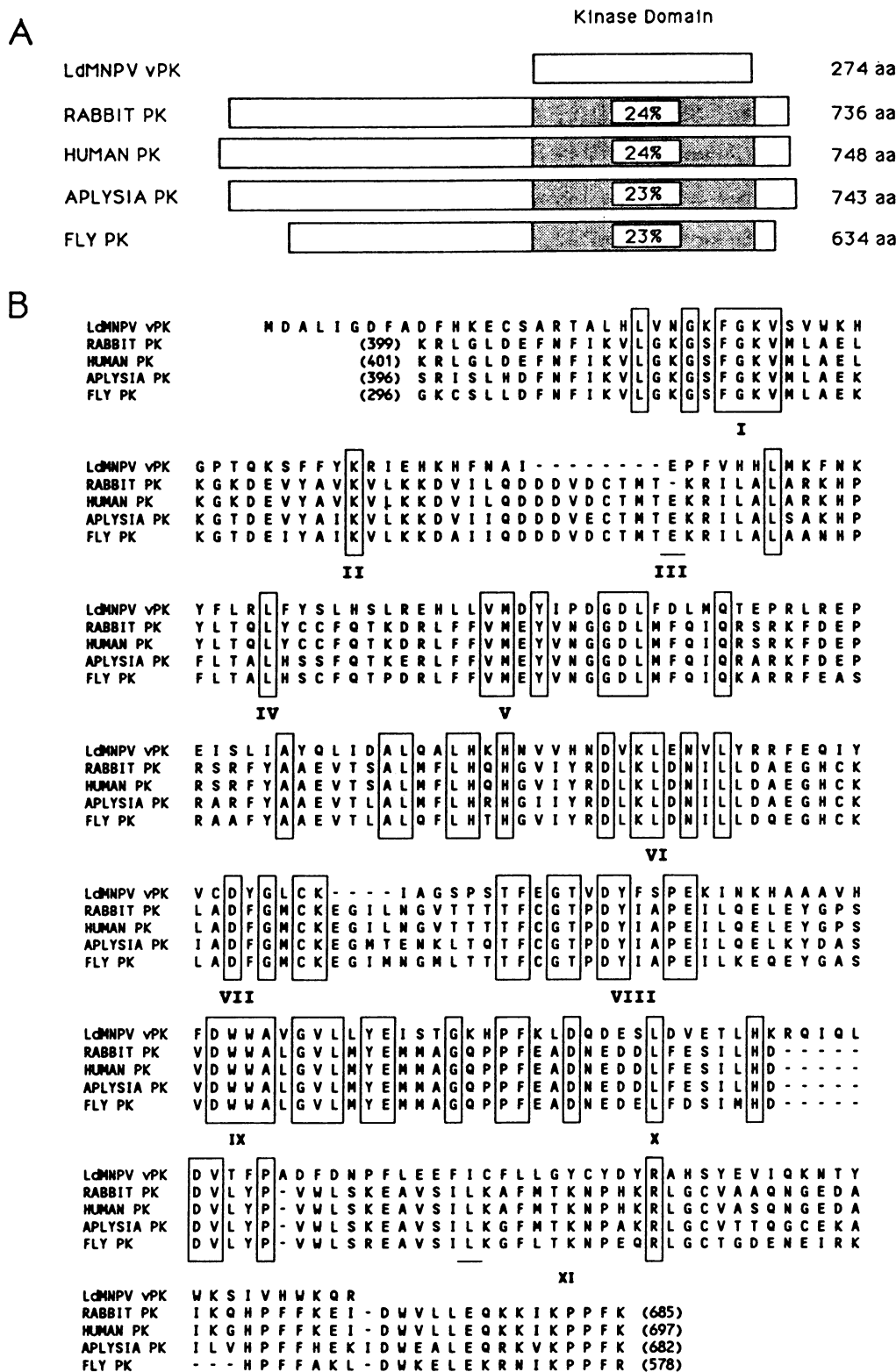
from a variety of organisms, including nPKC from rabbits (33), HSPKCE from humans (2), APLPKCB from *Aplysia californica* (32), and dPKC98F from *D. melanogaster* (46). vPK and the protein kinases shown are approximately 24% identical over the entire length of vPK (Fig. 4). There is approximately 44% amino acid similarity between vPK and the protein kinases (nPKC, 45%; HSPKCE, 45%; APLPKCB, 43%; and dPKC98F, 43%) when conservative residue changes are taken into consideration. The vPK homology extends only over the catalytic domains of the eucaryotic protein kinases. The other protein kinases are much longer than vPK, since they also have an N-terminal regulatory domain that is absent in vPK.

Homologous subdomains found in all protein kinase cata-

lytic domains (16) are indicated in Fig. 4, including the consensus ATP-binding site (G-X-G-X-G-X-V-X₁₄-K, where X is any amino acid) found in region I (56). The first glycine in the LdMNPV vPK ATP-binding site is replaced with a valine residue. Most protein kinases are highly conserved in this region, although there are exceptions. In casein kinase II (45) and Nim-1 (43), the third glycine residue is replaced by a serine. Another protein kinase, Mik1 (27), has the first, second, and third glycine residues replaced with histidine, serine, and serine, respectively. A protein kinase that is encoded by ASFV has recently been identified. Although the ATP-binding site in the ASFV protein kinase is missing two of the conserved glycine residues (the first glycine is replaced by glutamate, and the third is replaced with asparagine), the enzyme could still phosphorylate calf thymus histone protein in vitro (3).

vPK was found to contain protein kinase subdomains II through XI as defined by Hanks et al. (16). Region II contained the invariant lysine residue (at position 46) which is essential for protein kinase activity (5). Regions III and IV are defined by an invariant glutamic acid and an isoleucine, leucine, or valine, respectively. vPK contains a glutamic acid at position 57 and a leucine residue at position 73. The area of vPK between residues 86 and 95 shows homology to region V, which exhibits heterogeneity among kinase proteins. vPK contains a valine, methionine, tyrosine, and aspartic acid at positions 86, 87, 89, and 95, respectively, that are present in numerous protein kinases. Regions VI and VII are highly conserved among all protein kinases. The consensus sequences for regions VI and VII are D-V/I/L-X-X-X-N-V/I/L and V/I/L-X-D-F-G, respectively. Serine/threonine protein kinases can be identified with the consensus sequence D-L-K-P-E-N in region VI. vPK contains residues D-V-K-L-E-N-V at positions 133 to 140 in region VI, indicating that it is a serine/threonine kinase. vPK region VII has the consensus sequence with the exception of a tyrosine in place of the phenylalanine at position 152 (V-X-D-Y-G). The amino acid sequence G-T-X-X-Y-X-S-P-E at positions 166 to 174 in vPK is in close agreement with the region VIII consensus sequence, G-T/S/P-X-X-Y-X-A/P/S-P/L-E, which also indicates a serine/threonine protein kinase. The region IX consensus sequence of D-X-W/Y/F-A/S-X-G-V/I/L is well conserved at positions 186 to 192 (D-X-W-A-X-G-V) in vPK. Region X, a heterogeneous area of approximately 20 residues, generally begins approximately 15 residues from the highly conserved V/I/L residue in region IX and often with an aspartic acid residue. An aspartic acid is located 15 residues from the conserved valine in vPK and is followed by the sequence Q-X-E-X-L-X-X-X-X-L, which is present in several protein kinases. Typically, region XI is located approximately 20 to 52 residues from region X and contains an I/V/L/N/C residue followed by a highly conserved arginine 11 positions down. vPK contains an isoleucine and arginine at positions 241 and 252, respectively, which matches the placement of these residues in the region XI consensus sequence (Fig. 4). In addition, vPK contains all six highly conserved regions identified in herpesvirus protein kinases (5), which correspond to regions I, II, III, VII, VIII, and IX defined by Hanks et al. (16).

Mapping of 5' and 3' ends of the vPK transcript. Temporal analysis indicated that vPK was expressed throughout the infection cycle of the virus. To determine whether transcription of vPK was initiated at the same regions at both early and late times, primer extension reactions were carried out with RNA from either 6 or 48 h p.i. to map the 5' ends of the vPK transcripts. The 6-h primer extension reaction demonstrated that there are two start sites for early gene expression 25 and 26 bp upstream of the vPK start codon (Fig. 5A). At a later



time (48 h), there are multiple transcriptional start sites (40, 42, 43, 45, and 48 bp upstream of the *vPK* start codon) detected within the ATAAG late promoter sequence (Fig. 5A).

The 3' end of the *vPK* transcripts was determined by sequencing the λ 5-6 cDNA insert in pBluescript by using M13 reverse primer (Fig. 5B). The poly(A) tail begins at nucleotide position 1129 (Fig. 3), 13 bp downstream of the AATAAA polyadenylation sequence. The sizes of the transcripts predicted from the DNA sequence by using the early and late transcription start sites [not including the poly(A) tail] are approximately 876 bp (early) and 898 bp (late). Since poly(A) tails can be up to approximately 60 bp in length, the predicted size of the *vPK* transcripts is in close agreement with the 950-bp *vPK* transcript seen in the temporal expression experiments.

Temporal expression of *vPK* in the presence of DNA and protein synthesis inhibitors. Since *vPK* was expressed throughout the entire infective cycle of the virus, starting at 4 h p.i. throughout 72 h p.i., it was necessary to determine whether the *vPK* gene is an immediate-early, delayed-early, or late gene. RNA was isolated from 652Y cells infected with clonal isolate 5-6 in the presence or absence of DNA or protein synthesis inhibitors (37). Aphidicolin inhibits host and viral DNA replication and can therefore be used to distinguish between early and late genes. Late gene expression initiates after the onset of DNA replication, which occurs at approximately 20 h p.i. in LdMNPV (40). Cycloheximide is an inhibitor of cytoplasmic protein synthesis that can be used to determine whether a gene should be classified as a delayed-early or immediate-early gene, since delayed-early genes require other viral proteins to be expressed at high levels. The 0.95-kb *vPK* transcript was detected at 12 and 24 h p.i. from cells that were infected in the presence of aphidicolin, implying that initially *vPK* is expressed as an early gene (Fig. 6). At later times (>48 h p.i.), there is no transcript detected in the presence of aphidicolin, indicating that *vPK* is now being expressed from a late promoter, which is not active when aphidicolin is present. The lack of the *vPK*

transcript at 12 h p.i. in the presence of cycloheximide indicates that *vPK* is a delayed-early gene.

In vitro transcription and translation of *vPK*. To determine whether the ORF designated *vPK* encoded a protein, pPH1.2 was used to express the gene by using a rabbit reticulocyte transcription and translation kit from Bethesda Research Laboratories. A band with an apparent molecular mass of 31 kDa is seen after SDS-PAGE and autoradiography (Fig. 7). The size of the *vPK* protein is predicted to be 32 kDa from the nucleotide sequence. No radiolabelled band is detected from parent plasmid pT7/T3 α 18.

Protein kinase activity of *vPK*. To determine whether *vPK* encoded a functional protein kinase, extracts from the rabbit reticulocyte transcription and translation reactions were tested for the ability to phosphorylate substrates in an in vitro protein kinase assay (Fig. 8). The extracts expressing *vPK* from plasmid pPH1.2 could specifically phosphorylate histone H2A (lanes 4 and 10), which is not phosphorylated in extracts containing control plasmid pT7/T3 α 18 (lanes 1 and 7). Extracts containing pT7/T3 α 18 could phosphorylate both histone H2B (lanes 1 and 7) and glycogen phosphorylase *b* (lanes 3 and 9) but could not phosphorylate reduced carboxamidomethylated and maleylated lysozyme (lanes 2 and 8) or the other histone proteins (lanes 1 and 7). This phosphorylation is due to endogenous protein kinase activity within the rabbit reticulocyte lysate. This kinase activity has been observed when other proteins, such as the adenovirus E1A protein, have been expressed in similar systems (38).

DISCUSSION

We report in this paper the cloning and characterization of the LdMNPV gene *vPK*, the first baculovirus-encoded protein kinase gene to be identified. LdMNPV *vPK* exhibits approximately 24% amino acid identity to the catalytic subunits of the protein kinase C gene nPKC of rabbits (33), the protein kinase C gene dPKC98F of *D. melanogaster* (46), the *Homo sapiens*

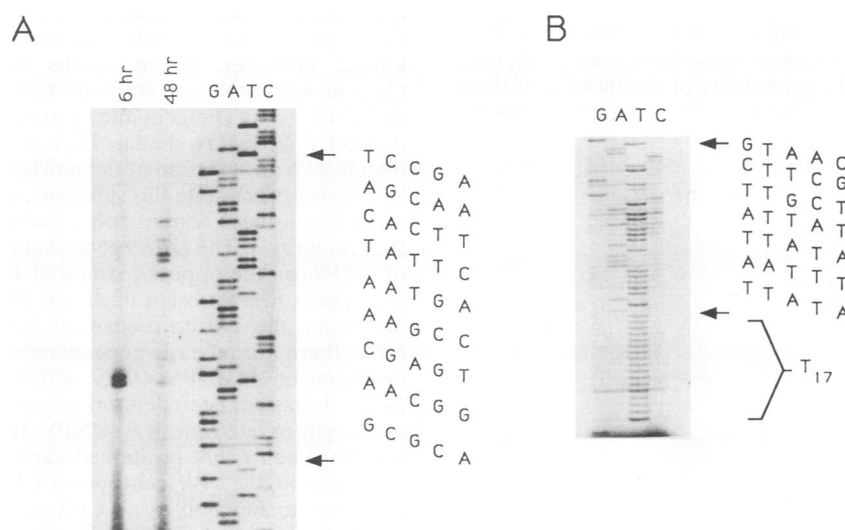


FIG. 5. Determination of 5' and 3' ends of the *vPK* transcripts. (A) Primer extension analysis of *vPK* RNAs isolated at 6 and 48 h p.i. Total RNA (40 μ g) was incubated with an end-labelled 18-bp primer that is complementary to the *vPK* nucleotide sequence at positions 320 to 339 (Fig. 3). After extension of the primer with Moloney murine leukemia virus reverse transcriptase, the extension products were fractionated by PAGE and visualized by autoradiography. The sequencing ladder was generated with the same primer. (B) Sequencing analysis of the 3' end of the *vPK* cDNA from λ 5-6. The cDNA was cloned into pBluescript after digestion with *NotI*-*EcoRI* and sequenced by using M13 reverse primer.

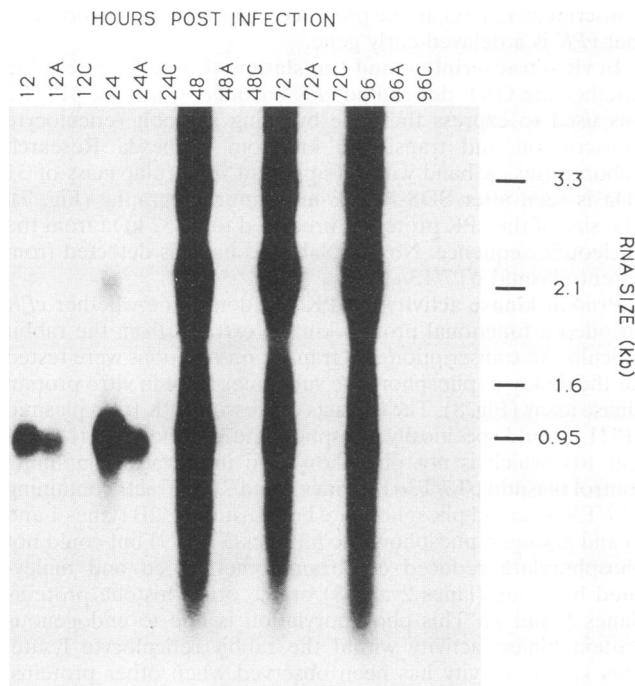


FIG. 6. Temporal analysis of *vPK* transcripts in the presence of inhibitors of DNA and protein synthesis. *L. dispar* 652Y cells were infected with isolate 5-6, and total RNA was isolated at the times indicated. As a control, the first lane for each time point contains RNA in the absence of any inhibitors. Aphidicolin, an inhibitor of host and viral DNA synthesis, was added after the 1-h adsorption period in lanes A at a final concentration of 5 μ g/ml. Cycloheximide, an inhibitor of cytoplasmic protein synthesis, was added 0.5 h prior to the virus adsorption period in lanes C at a final concentration of 100 μ g/ml. The inhibitors were maintained in the medium throughout the time course. The length of the *vPK* transcripts is indicated at the right.

protein kinase C epsilon gene (2), and the sea slug protein kinase gene APLPKCB from *Aplysia californica* (32) (Fig. 4). In addition, *vPK* exhibited significant homology to a large number of protein kinases such as those from yeasts (54), mice (55), and bovines (48) (the probability of significance of these matches ranged from $3.5e^{-25}$ to $1.6e^{-15}$, as determined by the

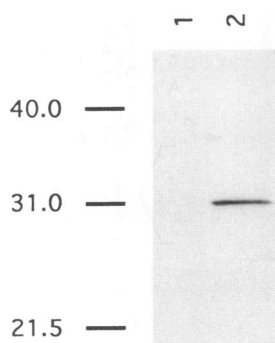


FIG. 7. SDS-PAGE analysis after induction of *vPK* under control of the T7 promoter. Autoradiogram of *vPK* expressed in a rabbit reticulocyte transcription and translation system and labelled with [35 S]Met. Lane 1, control plasmid pT7/T3 α 18; lane 2, pPH1.2 expressing *vPK*. Molecular mass standards are indicated to the left in kilodaltons.

BLAST program [1] from the National Center for Biotechnology Information). Hanks et al. have identified 11 regions that are conserved among many protein kinases (16). Sequence analysis of *vPK* revealed the presence of all 11 regions (Fig. 4).

Although *vPK* has been shown to phosphorylate histone H2A (Fig. 8) in an in vitro phosphorylation assay, the role in the virus life cycle has not been determined. Protein phosphorylation is a common means of regulation of many cellular processes, including gene transcription, control of metabolic pathways, cell division, and membrane transport (for a review, see reference 19). Protein kinase activity has been found to be associated with the virions of many RNA and DNA viruses of mammals, and several, such as herpes simplex virus (8), cytomegalovirus (51), and ASFV (3), have been found to be virally encoded. Recently, Zemskov et al. identified a novel protein kinase that was induced upon infection of *B. mori* with *B. mori* multinucleocapsid polyhedrosis virus, although it has not been determined whether this protein kinase is encoded by that virus (61). Several recent reports have implied that protein phosphorylation may also play a role in baculovirus infection or replication (23, 47). Many proteins are phosphorylated after infection with baculoviruses (22, 30, 60). Protein kinase activity is also associated with the occluded and extracellular forms of AcMNPV (31) and may be involved in the release of the DNA from the nucleocapsid in *P. interpunctella* granulosis virus (58).

Protein kinases are typically composed of regulatory and catalytic subunits which can be encoded by the same or different genes (for a review, see reference 50). Examples of protein kinases in which the catalytic and regulatory subunits are encoded by different genes can be found in mice (55), bovines (48), yeasts (54), and ASFV (3). Binding of a signal to the regulatory subunit activates the protein kinase. Without the regulatory subunit, the protein kinase is always activated. Since *vPK* has only the catalytic subunit, it is possible that there is a virus-encoded regulatory subunit to moderate its activity. Alternatively, if *vPK* is involved in phosphorylating nucleocapsid proteins to release DNA, there may not be any need for regulation, as the DNA may be released immediately upon entering the cell.

A computer search of known protein sequences in GenBank also revealed that *vPK* shows 40% identity to the AcMNPV ORF9 protein (35), which appears to be a truncated protein kinase. However, recent results reported by Reilly and Guarino suggest that an entire protein kinase gene is present in AcMNPV (36). The genomic locations of LdMNPV *vPK* and AcMNPV ORF9 are similar. Both genes are located approximately 2 kb downstream of the polyhedrin genes in both of the baculoviruses. Despite this similarity, examination of the DNA sequence in these regions reveals several differences between the two viruses. The ORF8 gene that is immediately upstream of ORF9 (on the opposite strand of DNA) in AcMNPV does not appear to be present upstream of LdMNPV *vPK*, at least not within the 290 bp sequenced. Furthermore, unlike LdMNPV, there are no early gene elements that can be identified upstream of AcMNPV ORF9, although there is a late promoter. It has not been determined whether ORF9 is expressed as an early or late gene in AcMNPV. It is possible that genomic rearrangements that positioned early gene elements in close proximity to the *vPK* gene occurred, thus enabling the late promoter to function as an early promoter in LdMNPV (or alternatively removing the early gene elements in AcMNPV). Sequence examination of the region downstream of AcMNPV ORF9 (from 190 to 271 bp) revealed a sequence exhibiting a nucleotide identity ranging from 67 to 85% with the AcMNPV homologous repeats. This homologous repeat region is not present in the DNA that was sequenced downstream of the

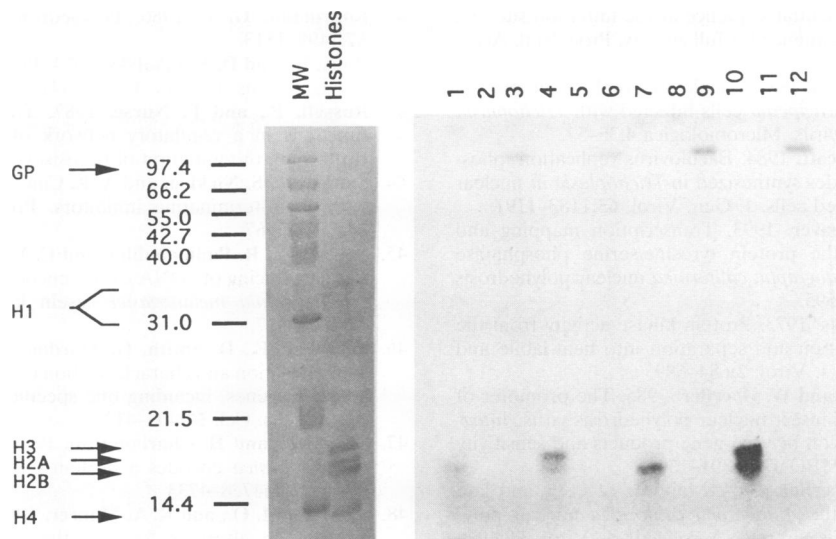


FIG. 8. In vitro phosphorylation of protein kinase substrates. Protein substrates (10 μ g) were incubated with rabbit reticulocyte extracts containing plasmid pPH1.2 expressing ν PK (lanes 4 to 6 and 10 to 12) or control plasmid pT7/T3 α 18 (lanes 1 to 3 and 7 to 9). The reactions were carried out at 30°C for 10 or 30 min (lanes 1 to 6 and 7 to 12, respectively), and the products were analyzed by SDS-PAGE. Molecular weight (MW) standards and histones were run on the same gel and stained with Coomassie brilliant blue. The locations of the standards in kilodaltons and the locations of the histones (H1, H3, H2A, H2B, and H4) and glycogen phosphorylase *b* (GP) are indicated to the left. Protein substrates tested were calf thymus histones (lanes 1, 4, 7, and 10), reduced carboxamidomethylated and maleylated lysozyme (lanes 2, 5, 8, and 11), and glycogen phosphorylase *b* (lanes 3, 6, 9, and 12).

LdMNPV ν PK. The homologous repeat regions function both as transcriptional enhancers (14) and as DNA replication origins in AcMNPV (34).

The role of the protein kinase encoded by the ν PK gene is not presently understood. Earlier studies have identified protein kinase activity associated with nonoccluded and occluded virions of AcMNPV (31) and with viral capsids of *P. interpunctella* granulosis virus (57). In addition, a protein tyrosine/serine-threonine phosphatase gene has been identified in AcMNPV (23, 47). These results suggest a regulatory and/or functional role for ν PK during viral replication. Studies are in progress to ascertain the functional role of ν PK.

REFERENCES

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment tool. *J. Mol. Biol.* **215**:403–410.
- Basta, P., M. B. Strickland, W. Holmes, C. R. Loomis, and D. J. Burns. 1992. Sequence and expression of human protein kinase C-epsilon. *Biochim. Biophys. Acta* **1132**:154–160.
- Baylis, S. A., A. H. Banham, S. Vydellingum, L. K. Dixon, and G. L. Smith. 1993. African swine fever virus encodes a serine protein kinase which is packaged into virions. *J. Virol.* **67**:4549–4556.
- Blissard, G. W., and G. F. Rohrmann. 1990. Baculovirus diversity and molecular biology. *Annu. Rev. Entomol.* **35**:127–155.
- Chen, W. S., C. S. Lazar, M. Poenie, R. Y. Tsien, G. N. Gill, and M. G. Rosenfeld. 1987. Requirement for intrinsic protein tyrosine kinase in the immediate and late actions of the EGF receptor. *Nature (London)* **328**:820–823.
- Crawford, A. M., and L. K. Miller. 1988. Characterization of an early gene accelerating expression of late genes of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **62**:2773–2781.
- Dickson, J. A., and P. D. Friesen. 1991. Identification of upstream promoter elements mediating early transcription from the 35,000-molecular-weight protein gene of *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **65**:4006–4016.
- Frame, M. C., F. C. Purves, D. J. McGeoch, H. S. Marsden, and D. P. Leader. 1987. Identification of the herpes simplex virus protein kinase as the product of the viral gene US3. *J. Gen. Virol.* **68**:2699–2704.
- Friesen, P. D., and L. K. Miller. 1985. Temporal regulation of baculovirus RNA: overlapping early and late transcripts. *J. Virol.* **54**:394–400.
- Friesen, P. D., and L. K. Miller. 1986. The regulation of baculovirus gene expression, p. 31–49. *In* W. Doerfler and P. Bohm (ed.), *The molecular biology of baculoviruses*. Springer-Verlag KG, Berlin.
- Fuchs, L. Y., M. S. Woods, and R. F. Weaver. 1983. Viral transcription during *Autographa californica* nuclear polyhedrosis virus infection: a novel RNA polymerase induced in infected *Spodoptera frugiperda* cells. *J. Virol.* **48**:641–646.
- Glocker, B., R. R. Hoopes, Jr., and G. F. Rohrmann. 1992. In vitro transactivation of baculovirus early genes by nuclear extracts from *Autographa californica* nuclear polyhedrosis virus-infected *Spodoptera frugiperda* cells. *J. Virol.* **66**:3476–3484.
- Goodwin, R. H., G. J. Tomkins, and P. McCrawley. 1978. Gypsy moth cell lines divergent in viral susceptibility. *In Vitro (Rockville)* **14**:485–494.
- Guarino, L. A., M. A. Gonzalez, and M. D. Summers. 1986. Complete sequence and enhancer function of the homologous DNA regions of *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **60**:224–229.
- Guarino, L. A., and M. D. Summers. 1988. Functional mapping of *Autographa californica* nuclear polyhedrosis virus genes required for late gene expression. *J. Virol.* **62**:463–471.
- Hanks, S. K., A. M. Quinn, and T. Hunter. 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**:42–52.
- Hatanaka, M., E. Twiddy, and W. V. Gilden. 1972. Protein kinase associated with RNA tumor viruses and other budding RNA viruses. *Virology* **47**:536–538.
- 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**:195–201.
- Hunter, T. 1987. A thousand and one protein kinases. *Cell* **50**:823–829.
- Jarrell, K. A., and M. Meselson. 1991. *Drosophila* retrotransposon

- promoter includes an essential sequence at the initiation site and requires a downstream sequence for full activity. *Proc. Natl. Acad. Sci. USA* **88**:102–104.
21. Kelly, D. C., and T. Lescott. 1981. Baculovirus replication: protein synthesis in *Spodoptera frugiperda* cells infected with *Trichoplusia ni* nuclear polyhedrosis virus. *Microbiologica* **4**:35–57.
 22. Kelly, D. C., and T. Lescott. 1984. Baculovirus replication: phosphorylation of polypeptides synthesized in *Trichoplusia ni* nuclear polyhedrosis virus-infected cells. *J. Gen. Virol.* **65**:1183–1191.
 23. Kim, D., and R. F. Weaver. 1993. Transcription mapping and functional analysis of the protein tyrosine/serine phosphatase (PTPase) gene of the *Autographa californica* nuclear polyhedrosis virus. *Virology* **195**:587–595.
 24. Kleiman, J., and B. Moss. 1973. Protein kinase activity from the vaccinia virus: solubilization and separation into heat-labile and heat-stable components. *J. Virol.* **2**:684–689.
 25. Knebel, D., H. Lubbert, and W. Doerfler. 1985. The promoter of the late p10 gene in the insect nuclear polyhedrosis virus *Autographa californica*: activation by viral gene products and sensitivity to DNA methylation. *EMBO J.* **4**:1301–1306.
 26. Lubbert, H., and W. Doerfler. 1984. Mapping of early and late transcripts encoded by the *Autographa californica* nuclear polyhedrosis virus genome: is viral RNA spliced? *J. Virol.* **50**:497–506.
 27. Lungdren, K., N. Walworth, R. Booher, M. Dembski, M. Kirschner, and D. Beach. 1991. mik1 and weel cooperate in inhibitory tyrosine phosphorylation of cdc2. *Cell* **64**:1111–1122.
 28. Mahmoudi, M., and V. K. Lin. 1989. Comparison of two different hybridization systems in Northern transfer analysis. *Bio/Technology* **7**:331–333.
 29. Mainprize, T. H., K. Lee, and L. K. Miller. 1986. Variation in the temporal expression of overlapping baculovirus transcripts. *Virus Res.* **6**:85–99.
 30. Maruniak, J. E., and M. D. Summers. 1981. *Autographa californica* nuclear polyhedrosis virus phosphoproteins and synthesis of intracellular proteins after virus infection. *Virology* **109**:25–34.
 31. Miller, L. K., M. J. Adang, and D. Browne. 1983. Protein kinase activity associated with the extracellular and occluded forms of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **46**:275–278.
 32. National Center for Biotechnology Information. Nucleotide sequence accession number M94884.
 33. Ohno, S., Y. Akita, Y. Konno, S. Imajoh, and K. Suzuki. 1988. A novel phobol ester receptor/protein kinase, nPKC, distantly related to the protein kinase C family. *Cell* **53**:731–741.
 34. Pearson, M., R. Bjornson, G. Pearson, and G. Rohrman. 1992. The *Autographa californica* baculovirus genome: evidence for multiple replication origins. *Science* **257**:1382–1384.
 35. Possee, R. D., T. Sun, S. C. Howard, M. D. Ayres, M. Hill-Perkins, and K. L. Gearing. 1991. Nucleotide sequence of the *Autographa californica* nuclear polyhedrosis 9.4 kbp *EcoRI*-I and -R (polyhedron gene) region. *Virology* **185**:1229–1241.
 36. Reilly, L. M., and L. A. Guarino. 1993. The *EcoRI* fragment of AcMNPV contains a protein kinase gene, p. A24. Abstr. 12th Am. Soc. Virol. Meet.
 37. Rice, W. C., and L. K. Miller. 1986. Baculovirus transcription in the presence of inhibitors and in nonpermissive *Drosophila* cells. *Virus Res.* **6**:155–172.
 38. Richter, J. D., J. M. Slavicek, J. F. Schneider, and N. C. Jones. 1988. Heterogeneity of adenovirus type 5 E1A proteins: multiple serine phosphorylations induce slow-migrating electrophoretic variants but do not affect E1A-induced transcriptional activation or transformation. *J. Virol.* **62**:1948–1955.
 39. Riegel, C. I., C. Lanner-Herrera, and J. M. Slavicek. *J. Gen. Virol.*, in press.
 40. Riegel, C. I., and J. M. Slavicek. Unpublished data.
 41. Rohrmann, G. F. 1986. Polyhedron structure. *J. Gen. Virol.* **67**:1499–1513.
 42. Roux, L., and D. Kolakofsky. 1974. Protein kinase associated with Sendai virions. *J. Virol.* **13**:545–547.
 43. Russell, P., and P. Nurse. 1987. The mitotic inducer *nim-1*⁺ functions in a regulatory network of protein kinase homologs controlling the initiation of mitosis. *Cell* **49**:569–576.
 44. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 45. Saxena, A., R. Padmanabha, and C. V. C. Glover. 1987. Isolation and sequencing of cDNA clones encoding alpha and beta subunits of *Drosophila melanogaster* casein kinase II. *Mol. Cell. Biol.* **7**:3409–3417.
 46. Schaeffer, E., D. Smith, G. Mardon, W. Quinn, and C. Zuker. 1989. Isolation and characterization of two new *Drosophila* protein kinase C genes, including one specifically expressed in photoreceptor cells. *Cell* **57**:403–412.
 47. Sheng, Z., and H. Charbonneau. 1993. The baculovirus *Autographa californica* encodes a protein tyrosine phosphatase. *J. Biol. Chem.* **268**:4728–4733.
 48. Showers, M. O., and R. A. Maurer. 1986. A cloned bovine cDNA encodes an alternate form of the catalytic subunit of cAMP-dependent protein kinase. *J. Biol. Chem.* **261**:16288–16291.
 49. Slavicek, J. M. 1991. Temporal analysis and spatial mapping of *Lymantria dispar* nuclear polyhedrosis virus transcripts and in vitro translation polypeptides. *Virus Res.* **20**:223–236.
 50. Soderling, T. R. 1990. Protein kinases: regulation by autoinhibitory domains. *J. Biol. Chem.* **265**:1823–1826.
 51. Somogyi, T., S. Michelson, and M.-J. O. Masse. 1990. Genomic location of a human cytomegalovirus protein with protein kinase activity (PK68). *Virology* **174**:276–285.
 52. Tan, K., and F. Sokol. 1974. Virion-bound protein kinase in Semliki Forest and Sindbis viruses. *J. Virol.* **13**:1245–1253.
 53. Thiem, S. M., and L. K. Miller. 1989. A baculovirus gene with a novel transcription pattern encodes a polypeptide with a zinc finger and a leucine zipper. *J. Virol.* **63**:4489–4497.
 54. Toda, T., S. Cameron, P. Sass, M. Zoller, and M. Wigler. 1987. Three different genes in *S. cerevisiae* encode the catalytic subunits of the cAMP-dependent protein kinase. *Cell* **50**:277–287.
 55. Uhler, M. D., J. C. Chrivia, and G. S. McKnight. 1986. Evidence for a second isoform of the catalytic subunit of cAMP-dependent protein kinase. *J. Biol. Chem.* **33**:15360–15363.
 56. Wierenga, R. K., and W. G. J. Hol. 1983. Predicted nucleotide-binding properties of p21 protein and its cancer-associated variant. *Nature (London)* **302**:842–844.
 57. Wilson, M. E., and R. A. Consigli. 1985. Characterization of a protein kinase activity associated with purified capsids of the granulosis virus infecting *Plodia interpunctella*. *Virology* **143**:516–525.
 58. Wilson, M. E., and R. A. Consigli. 1985. Function of a protein kinase activity associated with purified capsids of the granulosis virus infecting *Plodia interpunctella*. *Virology* **143**:526–535.
 59. Yang, C. L., D. A. Stetler, and R. F. Weaver. 1991. Structure comparison of the *Autographa californica* nuclear polyhedrosis virus-induced RNA polymerase and the three nuclear RNA polymerases from the host, *Spodoptera frugiperda*. *Virus Res.* **20**:251–264.
 60. Zemskov, E. A., E. B. Abramova, and V. S. Mikhailov. 1991. Virus-induced phosphoproteins in silkworm pupae infected with nuclear polyhedrosis virus. *Biokhimiya* **56**:2207–2217. (In Russian.)
 61. Zemskov, E. A., E. B. Abramova, and V. S. Mikhailov. 1992. Induction of a novel protein kinase in pupae of the silkworm *Bombyx mori* after infection with nuclear polyhedrosis virus. *J. Gen. Virol.* **73**:3231–3234.