JOURNAL OF VIROLOGY, July 1987, p. 2264–2272
0022-538X/87/072264-09$02.00/0

Vol. 61, No. 7

Divergent Transcription of Early 35- and 94-Kilodalton Protein Genes Encoded by the *Hin*dIII K Genome Fragment of the Baculovirus *Autographa californica* Nuclear Polyhedrosis Virus†

PAUL D. FRIESEN* AND LOIS K. MILLER‡

Institute for Molecular and Agricultural Genetic Engineering, Department of Bacteriology and Biochemistry, University of Idaho, Moscow, Idaho 83843

Received 23 January 1987/Accepted 3 April 1987

The organization of viral genes within the 3.7-kilobase-pair *Hin*dIII-K/*Eco*RI-S region of the *Autographa californica* nuclear polyhedrosis virus genome (85 to 88 map units) was determined by using a combination of nucleotide sequencing, transcriptional mapping, and in vitro translation of hybrid selected RNA. Two nonoverlapping genes, extending in opposite directions and encoding polypeptides with molecular weights of 35,000 and 94,000 (35K and 94K polypeptides), were identified. Unspliced, messenger-active RNAs were transcribed from both genes early (2 h) after infection. Indicative of immediate-early genes, transcription of the divergent RNAs was unaffected by the protein synthesis inhibitor, cycloheximide. Late in infection, abundant RNAs were transcribed from promoters located at least 2.5 kilobase pairs upstream from the gene encoding the 35K polypeptide. These transcripts completely overlapped both the 35K and 94K polypeptide genes but apparently lacked protein-coding potential, suggesting that the transcripts may play a role in the suppression of early viral gene expression.

Autographa californica nuclear polyhedrosis virus (AcNPV), a member of the family *Baculoviridae*, possesses a double-stranded, circular DNA genome of approximately 128 kilobase pairs (kb). Viral replication occurs in the nucleus of the host insect cell and results in the production of two infectious forms, nonoccluded and occluded virus. Expression of the genome, including the synthesis of both viral forms, is temporally controlled through a cascade of regulatory events. On the basis of the requirement for de novo protein synthesis or viral DNA replication, expression can be divided into at least three phases: early (subdivided into immediate early and delayed early), late, and very late (occlusion specific) (15; for a review, see reference 7). An immediate-early gene (IE1) within the regulatory cascade has been identified; IE1 modulates the transcription of at least one later AcNPV gene via interaction with *cis*-acting enhancer elements (10, 11).

AcNPV transcription is characterized by the synthesis of a multitude of overlapping, polyadenylated RNAs. Transcripts with maximum synthesis during each of the temporal phases have been mapped throughout the viral genome (3, 6, 20, 30). Overlapping nests of transcripts composed of RNAs with identical 5′ or identical 3′ termini are common (6, 21, 28), yet splicing does not appear to be involved in the generation of these RNAs (20). The functional significance of the nests of overlapping RNAs is not clear.

The 3.7-kb *Hin*dIII-K/*Eco*RI-S region, located between 85 and 88 map units (m.u.) of the AcNPV genome, is transcribed into a complex set of overlapping RNAs with common 3′ ends (6). The transcriptional unit is bordered on one side (the *Hin*dIII B_2/*Eco*RI H fragments; 81 to 85 m.u.) by a different set of 3′ coterminal RNAs (21, 22) and on the other side, by one of the AcNPV enhancers, *hr*5 (11). Within the *Hin*dIII-K unit, separate promoters are involved in the transcription of early and late RNAs. The region is of additional interest since it is the site of integration (86.7 m.u.) of the retrotransposon TED, within the AcNPV mutant FP-D (25). TED transposes from the host insect cell (*Trichoplusia ni*) to the AcNPV genome during the infection process and promotes the transcription of new and abundant RNAs extending from the element into flanking *Hin*dIII-K sequences (8). Thus, analogous to the retroviruses and other transposable elements which act as insertion mutagens, TED integration alters the expression of nearby viral genes.

To analyze these alterations and to investigate the functional significance of the overlapping *Hin*dIII-K RNAs, we determined the organization of viral genes within this region. We report here that the *Hin*dIII K/*Eco*RI S genome fragments contained two nonoverlapping genes extending in opposite directions and encoding viral proteins of 35 and 94 kilodaltons (35K and 94K proteins). Transcription of each gene occurred as early as 2 h after infection and was unaffected during inhibition of protein synthesis, indicative of immediate-early genes. Integration of TED disrupted the coding region of the 94K protein gene.

MATERIALS AND METHODS

Cells and virus infection. *Spodoptera frigiperda* IPLB-SF-21 cells (34) were inoculated with extracellular AcNPV L-1 (19) at a multiplicity of 20 PFU per cell as described previously (6). To block protein synthesis, growth medium containing 100 μg of cycloheximide per ml was added to cell monolayers (10^7 cells per 100-mm plate) 30 min prior to infection. The viral inoculum, subsequent washes, and final growth medium also contained 100 μg of cycloheximide per ml.

Northern blot analysis. Poly(A)$^+$ RNA, isolated from the cytoplasmic fraction of infected cells, was subjected to electrophoresis, blotted to nitrocellulose, and hybridized to DNA probes as described previously (6). Glyoxalated DNA

* Corresponding author.

† Research paper no. 8751 of the Idaho Agricultural Experiment Station.

‡ Present address: Departments of Genetics and Entomology, University of Georgia, Athens, GA 30602.

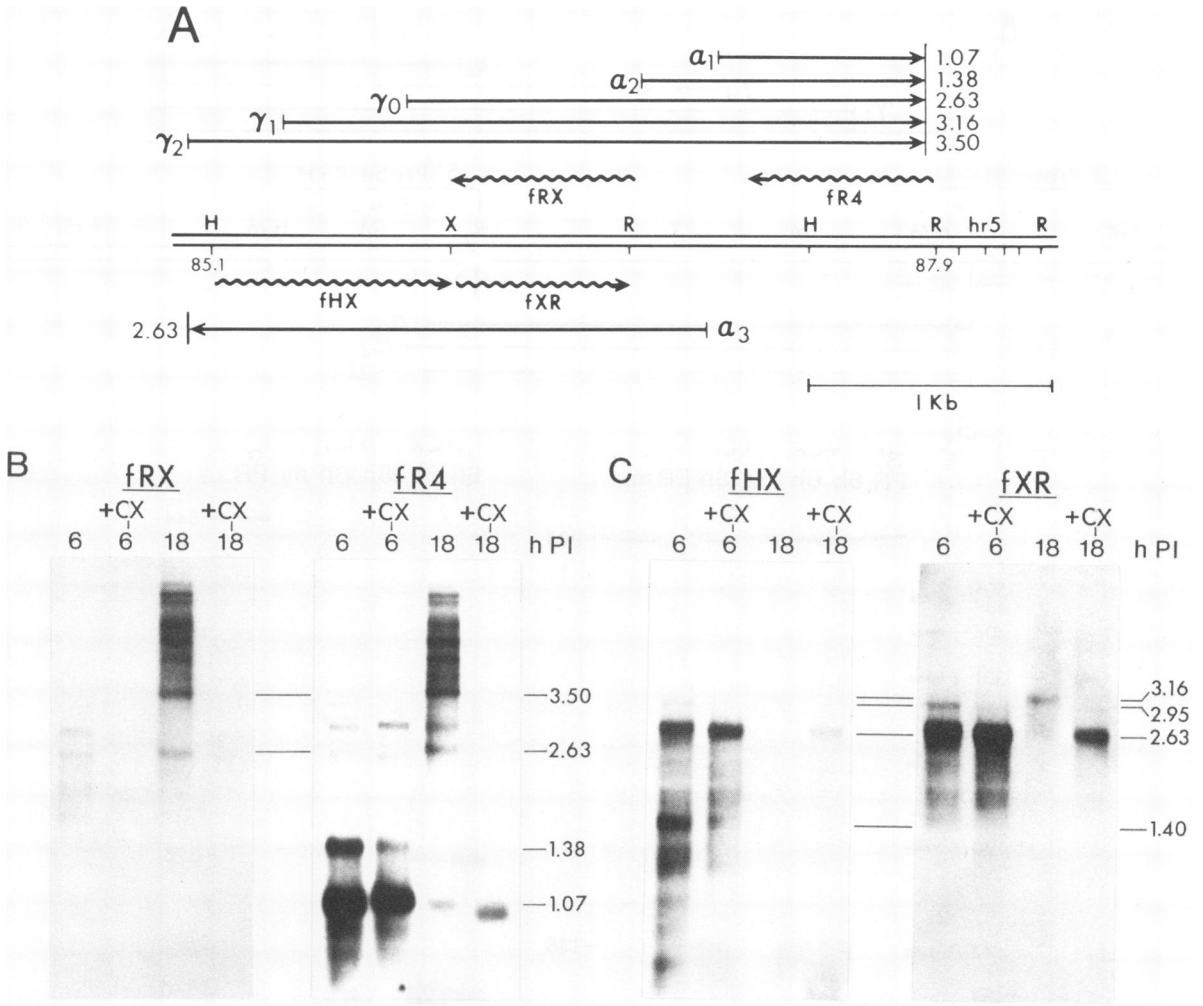

FIG. 1. Strand-specific Northern hybridization of *Hin*dIII-K AcNPV RNAs. Poly(A)$^+$ RNA, isolated 6 and 18 h after infection from cells incubated continuously in the presence (+CX) or absence of cycloheximide, was fractionated by electrophoresis and blotted to nitrocellulose. The filters were hybridized to the radiolabeled, single-stranded DNA probes shown in panel A (wavy lines); probes fRX (861-nucleotide *Eco*RI-*Xho*I insert) and fR4 (907-nucleotide insert proceeding leftward from the *Eco*RI site) extended right to left (5′ to 3′), while probes fHX (1,266-nucleotide *Hin*dIII-*Xho*I insert) and fXR (861-nucleotide *Xho*I-*Eco*RI insert) extended left to right (5′ to 3′). Previously mapped 3′ coterminal RNAs transcribed from left to right (clockwise) are shown (panel A, top); sizes are in kilobases. (B) Hybridization of probes fRX and fR4 (bottom DNA strand) to RNAs transcribed from left to right (clockwise). (C) Hybridization of probes fHX and fXR (top DNA strand) to RNAs transcribed from right to left (counterclockwise). Shown are autoradiograms of a 2-h exposure. H, *Hin*dIII; R, *Eco*RI; X, *Xho*.

fragments were used as size standards. Blots were subjected to autoradiography by using XAR5 film (Eastman Kodak Co., Rochester, N.Y.). To prepare strand-specific DNA probes, a 13-base oligonucleotide probe primer (Bethesda Research Laboratories, Gaithersburg, Md.) was annealed to single-stranded M13 bacteriophage DNA and extended with the Klenow fragment of *Escherichia coli* DNA polymerase I (Bethesda Research Laboratories) as described previously (24). By using a single radiolabeled nucleotide, [α-^{32}P]dCTP (3,000 Ci/mmol; New England Nuclear Corp., Boston, Mass.), specific activities were found to range from 0.5×10^8 to 1.0×10^8 cpm/μg. The cloned insert remained single stranded.

DNA cloning and sequencing. The restriction enzyme *Hin*dIII K, *Hin*dIII B_2, *Eco*RI S, and *Xho*I I fragments of the AcNPV L-1 genome were cloned into *E. coli* plasmids pBR322, pUC18, and pUC19 by using standard procedures (23). Various restriction fragments were further subcloned into the replicative forms of M13, mp18, and mp19. Overlapping deletions were constructed from M13 clones by using the exonuclease III target breakpoint method of Henikoff (12). The recombinant phage DNAs were sequenced by the dideoxy chain termination method of Sanger et al. (32) by using a 17-base universal primer (Bethesda Research Laboratories) and [^{35}S]dATP(α-S) (500 Ci/mmol; New England Nuclear). Nucleotide sequences shown were determined for both strands of DNA.

S1 nuclease analysis. The 5′ and 3′ ends of the 94K protein gene transcripts were determined by using the S1 nuclease procedure of Weaver and Weissmann (35). DNA probes were prepared by cleaving appropriate recombinant plasmids with a single restriction enzyme, radiolabeling the 5′ or 3′ ends with T4 polynucleotide kinase or T4 DNA polymerase (Bethesda Research Laboratories; 23), respectively, and cleaving with a second restriction enzyme. The desired DNA fragments, labeled exclusively at one end, were purified by agarose gel electrophoresis. Hybridization of DNA probes to poly(A)$^+$ RNA (10 μg) was conducted for 12 h at 42 or 47°C

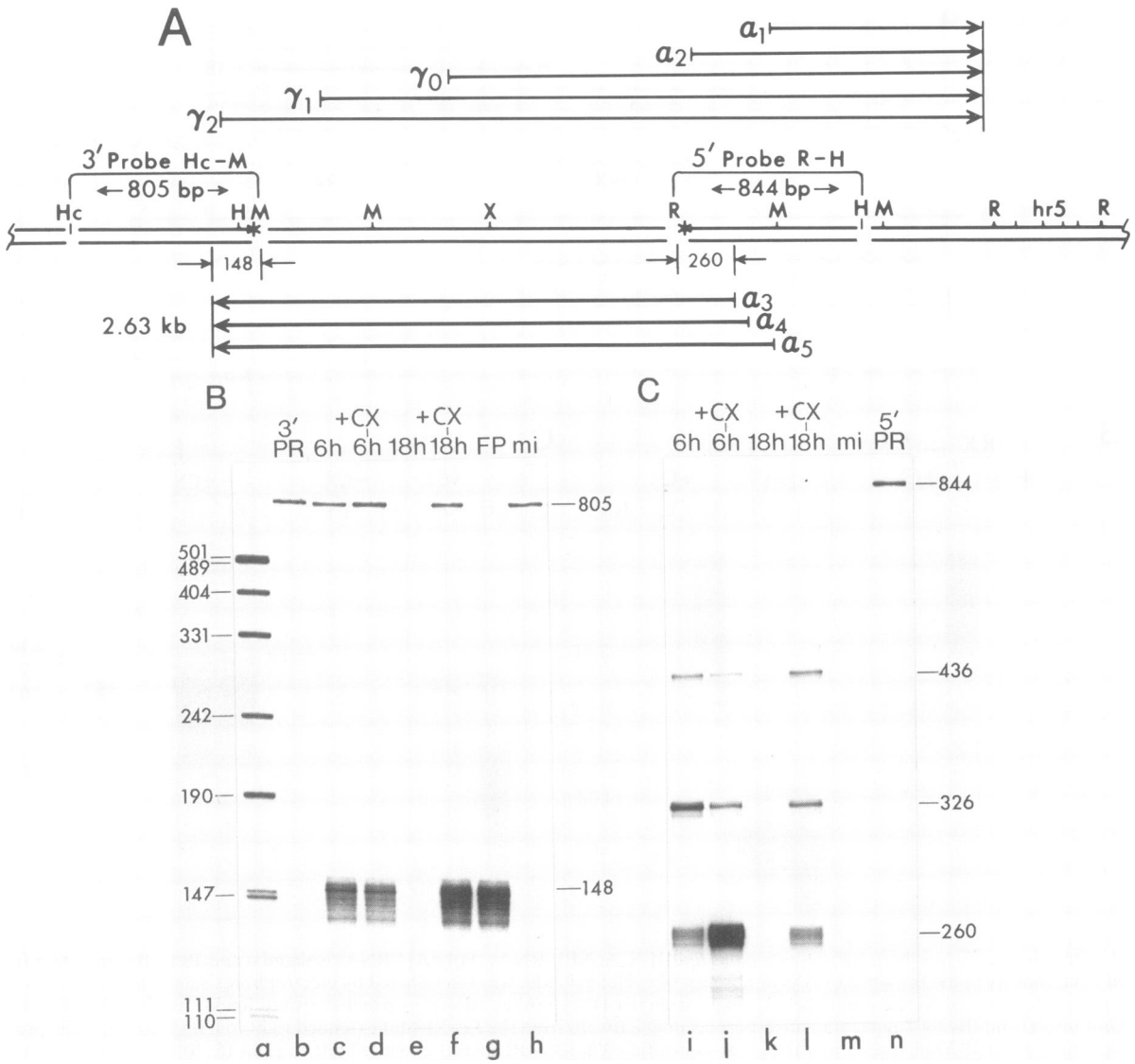

FIG. 2. S1 nuclease mapping of the 5′ and 3′ ends of the 2.63-kb α_3 RNA. Poly(A)$^+$ RNA, isolated 6 and 18 h after infection from an equal number of cells treated with cycloheximide (+CX) or without, was hybridized at 42°C to a 3′-end-labeled probe (805-bp *Hin*cII-*Msp*I fragment [3′ Hc-M]) or at 47°C to a 5′-end-labeled probe (844-bp *Eco*RI-*Hin*dIII fragment [5′ R-H]) as diagrammed in panel A; the * denotes the position of the end label. After treatment with S1 nuclease, the resistant 3′-end-labeled fragments (B) and the resistant 5′-end-labeled fragments (C) were subjected to polyacrylamide gel electrophoresis and autoradiography. Each panel includes the S1-resistant fragments generated after hybridization to mock-infected (mi) RNA (lanes h and m) and end-labeled DNA probes (PR) prior to S1 treatment (lanes b and n). Sizes of molecular weight standards (lane a) are indicated in nucleotides.

as previously described (6). S1-resistant fragments were denatured and subjected to electrophoresis on 43-cm-long 6% polyacrylamide–8 M urea–TBE (100 mM Tris borate [pH 8.3], 2 mM EDTA) gels. Size standards consisted of 3′-end-labeled *Msp*I-digested plasmid pUC19 and dideoxy chain termination sequencing ladders.

Hybrid selection and in vitro translation. Viral RNA was hybrid selected as described by Esche and Siegmann (4), except that strand-specific DNA was used. In brief, 10 to 20 μg of single-stranded M13 recombinant phage DNA was denatured and added dropwise to nitrocellulose filters (1 cm^2; Schleicher & Schuell, Inc., Keene, N.H.). The filters were washed with 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 trisodium citrate [pH 7.0]) and then baked for 2 h at 80°C in a vacuum. Hybridization was conducted for 12 h at 42°C in 50% formamide–10 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] (pH 6.4)–0.4 M NaCl–1 mM EDTA with total cytoplasmic RNA (1 to 3 mg) isolated from cells 6 h after infection. The filters were washed with 1× SSC–0.1% sodium dodecyl sulfate–2 mM EDTA, and the bound RNA was eluted by boiling for 1 min in 1 mM EDTA containing 10 μg of calf liver tRNA (Boehringer Mannheim Biochemicals, Indianapolis, Ind.). In vitro translation in rabbit reticulocyte lysates (Promega Biotec, Madison, Wis.), made mRNA dependent by the method of Pelham and Jackson (26), was conducted at 30°C for 1 h by using [^{35}S]methionine (1,200 Ci/mmol; New England Nuclear) and conditions described previously (33).

Intracellular viral proteins were radiolabeled by incubating cell monolayers (10^6 cells per 35-mm plate), infected

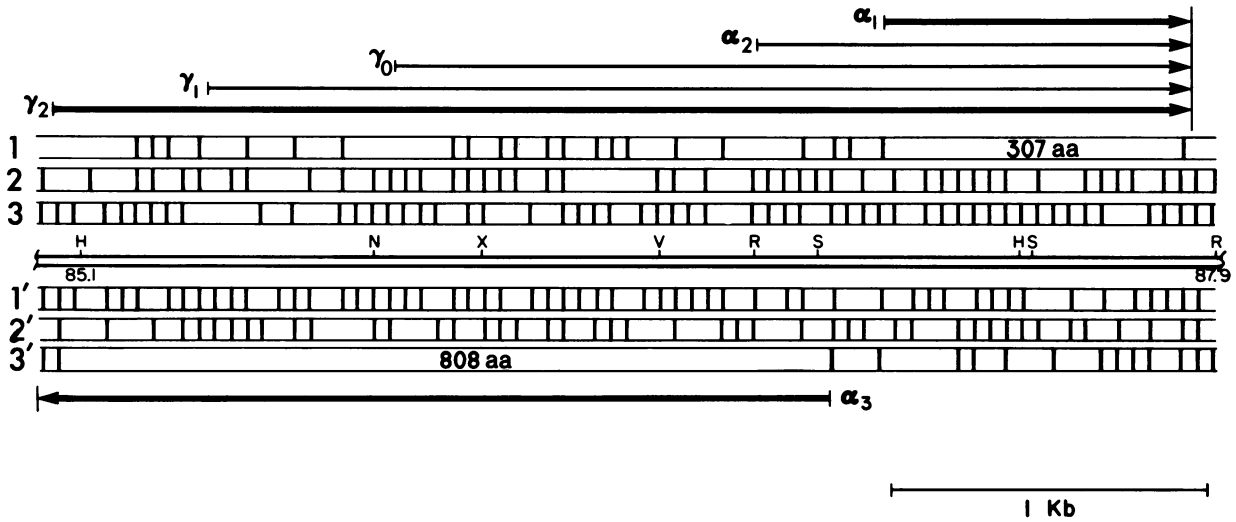

FIG. 3. Distribution of ORFs within the overlapping group of *Hin*dIII-K RNA transcripts. Translational termination codons TAG, TGA, and TAA (vertical bars) between 85 and 88 m.u. are depicted for all six reading frames (1, 2, and 3, top DNA strand; and 1', 2', and 3', bottom DNA strand). The location of RNA transcripts and their relative abundance are indicated; RNAs α_1, α_3, and γ_2 represent the most abundant species. ORFs were identified by using Cornell DNA sequence analysis programs (9). H, *Hin*dIII; N, *Nsi*I; R, *Eco*RI; S, *Sal*I; V, *Eco*RV; X, *Xho*I.

previously with 20 PFU per cell, in methionine-deficient growth medium containing 250 μCi of [^{35}S]methionine per ml. After 1-h incubations at 6, 12, and 24 h after infection, cells were collected, washed with phosphate-buffered saline (19), and lysed by the addition of sodium dodecyl sulfate and β-mercaptoethanol to 1%. Intracellular proteins and in vitro translation products were subjected to electrophoresis on 10% polyacrylamide gels (18), followed by autoradiography using Kodak XAR5 film.

RESULTS

Transcription of AcNPV early and late *Hin*dIII-K RNAs. During infection, the *Hin*dIII-K/*Eco*RI-S region is transcribed from left to right (clockwise on the circular AcNPV map) into an overlapping group of RNAs with common 3' ends (Fig. 1A). The smallest RNAs are transcribed earliest (2 to 6 h) and are replaced in time with larger RNAs initiated from promoters located further upstream (6). Previous studies indicated that the early RNAs are transcribed in the presence of the protein synthesis inhibitor, cycloheximide, and the DNA synthesis inhibitor, aphidicolin, while the later RNAs are not (29). This suggests that the *Hin*dIII-K region is transcribed into both immediate-early (α) and late (γ) RNAs.

To map these RNAs more precisely, Northern blots of poly(A)$^+$ RNA were hybridized to strand-specific DNA probes of the *Hin*dIII K fragment (Fig. 1A). Probe fR4 hybridized to the two smallest RNAs (1.07 and 1.38 kb) early (6 h) after infection (Fig. 1B, right panel). Early accumulation of the 1.07-kb α_1 RNA was unaffected by cycloheximide, while that of the 1.38-kb α_2 RNA was reduced. By late times (18 h), the α RNAs were replaced by a heterogeneous group of longer RNAs (2.63 and 3.50 kb and larger) also transcribed from left to right. These late γ RNAs were not detected in cells treated with cycloheximide; instead, small amounts of the α_1 RNA remained. The location of these RNAs was confirmed by an additional probe, fRX, located entirely upstream from the α_1 and α_2 transcripts. fRX hybridized to the late γ RNAs (18 h) but only from cells not treated with cycloheximide (Fig. 1B, left panel).

Probes fHX and fXR, derived from the opposite (top) strand of the *Hin*dIII K fragment, detected several RNAs transcribed from right to left (Fig. 1C). The most abundant transcript, α_3 (2.63 kb), hybridizing to both probes, was synthesized early (6 h) but disappeared later (18 h). Early accumulation of RNA α_3 was unaffected by cycloheximide. By 18 h, only a single minor RNA (3.16 kb), sensitive to cycloheximide, was detected; longer exposures indicated that this RNA hybridized to both fHX and fXR probes. Probe fXH, but not fXR, also hybridized to a smaller 1.40-kb early RNA (Fig. 1C, left panel).

Mapping of 5' ends of the divergent α RNAs to a common 210-bp region. The 5' and 3' ends of the early α_1 and α_2 RNAs (1.07 and 1.38 kb), transcribed left to right (clockwise), were mapped previously (6). The counterclockwise 2.63-kb α_3 transcript, on the other hand, represented a cycloheximide-insensitive RNA which previously escaped detection because of its similar electrophoretic mobility with another RNA, γ_0 (2.63 kb), transcribed in the clockwise direction (Fig. 1A). To map the 5' end of the α_3 RNA, poly(A)$^+$ RNA was hybridized to an 844-base-pair (bp) *Eco*RI-*Hin*dIII probe, 5' end labeled exclusively at the *Eco*RI site (Fig. 2A). Early RNA (6 h) protected three fragments of 260, 326, and 436 bp from S1 nuclease (Fig. 2C, lane i). All three fragments were also protected by RNA isolated both early (6 h) and late (18 h) from cells treated with cycloheximide (lanes j and l); 6 h after infection, the 260-bp fragment was enhanced relative to the other two. Consistent with the early window for synthesis of the α_3 RNA, none of the three 5' ends were detected 18 h after infection (lane k). No fragments were protected by RNA from mock-infected cells (lane n). These and additional data (below) indicate that the 2.63-kb RNA possessed at least three 5' ends; the most frequently used RNA start site (α_3) mapped approximately 160 bp away from the divergent α_1 start site (Fig. 2A).

The 3' end of the 2.63-kb RNA was mapped by using a 805-bp *Hin*cII-*Msp*I probe, 3' end labeled exclusively at the *Msp*I site (Fig. 2A). Early (6-h) RNA protected a tightly clustered group of fragments ranging in size from 144 to 149 bp, with the 148-bp fragment predominating (Fig. 2B). The

same 3′ fragments were protected by RNA from cycloheximide-treated cells both early (6 h) and late (18 h) but not by RNA from untreated cells late in infection. Identical fragments were also protected by very late RNA (24 h) isolated from cells infected with the AcNPV mutant (FP-DS) (lane g). Thus, the α_3 RNA and the abundant RNAs extending leftward from the TED long terminal repeat (LTR), inserted (87.6 m.u.) within the *Hin*dIII K fragment (8), terminated at a common 3′ site.

Location of ORFs within the *Hin*dIII K/*Eco*RI S fragments. To facilitate determination of the organization of viral genes, the AcNPV genome was sequenced from 85 to 88 m.u. and analyzed for possible protein-coding regions, as evidenced by the presence of long open reading frames (ORFs; Fig. 3). The largest ORF encoded by the top DNA strand was 0.92 kb long, corresponding to 307 codons, and mapped to a location most directly transcribed by the 1.07-kb α_1 RNA. While the longer RNAs (α_2, γ_0, γ_1, and γ_2) also contained this ORF, numerous translational termination codons were located at the 5′ portion of each. The largest ORF encoded by the bottom strand was 2.42 kb long, corresponding to 808 codons, and was completely transcribed by the 2.63-kb α_3 RNA (Fig. 3).

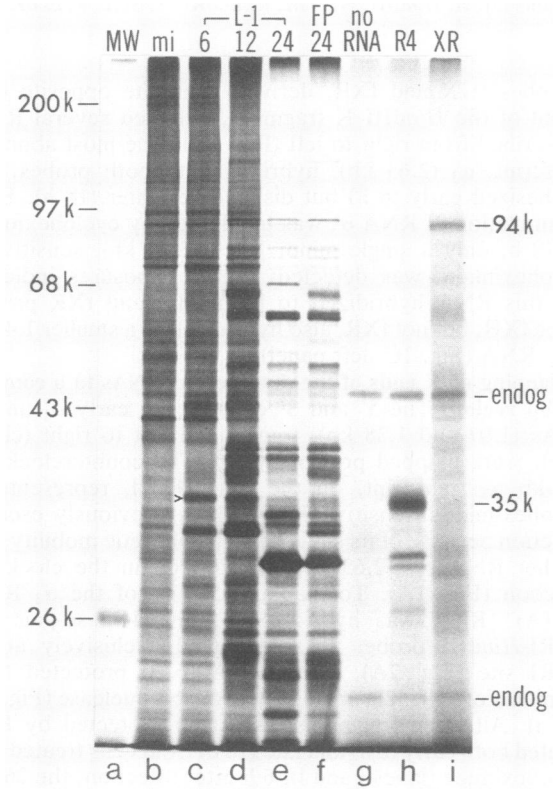

FIG. 4. In vitro translation products of the early α_1 and α_3 RNAs. Specific RNA, selected from total cytoplasmic RNA (6 h after infection) by hybridization to single-stranded M13 clones of AcNPV DNA, was translated in mRNA-dependent extracts of rabbit reticulocytes. The radiolabeled polypeptides synthesized in response to no RNA (lane g), fR4-selected RNA (lane h), and fXR-selected RNA (lane i) were subjected to electrophoresis on 10% polyacrylamide gels along with [^{35}S]methionine-labeled intracellular proteins of mock-infected (mi, lane b) and L-1 AcNPV-infected cells 6, 12, and 24 h after infection (lanes c to e). Also included were pulse-labeled proteins of cells infected with the AcNPV mutant, FP-DS, 24 h after infection (lane f). The positions of molecular weight standards (lane a) and endogenous (endog) reticulocyte proteins are indicated. A 3-day autoradiogram is shown.

Identification of viral proteins encoded by the early α_1 and α_3 RNAs. The presence of two long ORFs (307 and 808 amino acids) suggests that the *Hin*dIII-K region encodes at least two viral proteins and that the early RNAs α_1 and α_3 represent messengers for the two viral genes. To identify the viral gene products, these RNAs were independently hybrid selected from early (6-h) RNA and translated in protein synthesis cell extracts (Fig. 4).

Probe fR4, complementary to RNA α_1 (Fig. 1), selected RNA which directed the synthesis of a predominant 35K polypeptide and several minor polypeptides ranging from 26 to 30 kilodaltons (Fig. 4, lane h). The size of the 35K polypeptide was in good agreement with the molecular weight (34,800) of the α_1 gene product predicted from its nucleotide sequence (Fig. 5). Electrophoretic examination of intracellular proteins radiolabeled at intervals after infection (Fig. 4, lanes c to e) revealed an abundant virus-induced polypeptide which comigrated with the 35K protein in vitro (arrow). Its early period of intracellular synthesis, from 2 to 12 h after infection (data not shown), paralleled the presence of the α_1 RNA in infected cells (6), suggesting that this protein is a likely intracellular candidate for the α_1 gene product.

Probe fXR, complementary to the α_3 RNA, hybrid selected RNA which directed the synthesis of a predominant 94K polypeptide (Fig. 4, lane i). The size of this protein was in close agreement with the molecular weight (94,400) predicted for the α_3 gene product (Fig. 6). In vitro translation of fXR-selected RNA exhibited a higher background of lower-molecular-weight products than did that selected by fR4. These heterogeneous products did not correspond to those translated in response to total cytoplasmic RNA (6 h after infection) and most likely represented premature termination products of the 94K protein, rather than products of RNAs nonspecifically hybridized to fXR DNA.

Unlike the 35K protein, an intracellular candidate for the 94K protein was not detected in pulse-labeled infected cell lysates (Fig. 4). In an attempt to facilitate such identification, the proteins of cells infected with wild-type AcNPV and mutant FP-DS were compared. Since the insertion of the 270-bp LTR in virus FP-DS (8) disrupted the 94K ORF, the normal 94K protein should be absent in FP-DS-infected cells. While there were several differences between the polyacrylamide gel profiles (Fig. 4, lanes e and f) of wild-type and FP-DS-infected cells, no differences were observed in the 80- to 100-kilodalton protein range. Thus, intracellular synthesis of the 94K protein was either obscured by the synthesis of similar-sized host proteins early in infection or it was processed into a protein(s) with an altered mobility.

DISCUSSION

Gene organization of the *Hin*dIII K/*Eco*RI S genome fragments. We used a combination of transcriptional mapping, nucleotide sequencing, and in vitro translation of hybrid-selected RNA to determine the organization and regulation of genes from 85 to 88 m.u. on the circular AcNPV genome. Our findings are summarized in Fig. 7.

Two genes encoding 35K and 94K viral polypeptides were identified within the *Hin*dIII-K/*Eco*RI-S region. The coding frames extended uninterrupted in opposite directions and were separated from one another by 210 bp. Transcription of the poly(A)$^+$ mRNAs for the 35K and 94K protein genes, α_1 (1.07 kb) and α_3 (2.63 kb), respectively, occurred as early as

FIG. 5. Nucleotide sequence and predicted amino acid sequence of the early gene encoding the 35K polypeptide. The gene extends left to right (5′ and 3′, clockwise) from 87.1 to 87.8 m.u. on the AcNPV genome. The DNA strand having the polarity of the major α_1 RNA (1.07 kb) is shown and numbered from the intitator codon (ATG, +1). The 5′ start site (position −34) and the 3′ termination site (position +908) of the α_1 RNA are underlined. Potential regulatory signals for transcription, including the TATA box and polyadenylation signal, are indicated along with the start of the divergent 94K ORF located in the opposite DNA strand.

2 h after infection. Intracellular transcription of both RNAs, originating from within the intergenic region, was unaffected by the protein synthesis inhibitor cycloheximide (Fig. 1). Thus, the apparent lack of a requirement for de novo viral protein synthesis and the very early transcription of these RNAs classified the 35K and 94K protein genes as immediate-early genes. Conversely, transcription of the γ RNAs (2.63 kb and longer), initiated from promoters upstream from the 35K protein gene late in infection (Fig. 7), was not detected in cells arrested for protein synthesis. That the DNA synthesis inhibitor, aphidicolin, also blocked the transcription of these RNAs (29) demonstrated the additional requirement for viral DNA synthesis and classified these transcripts as late.

Nucleotide sequence of the 94K protein gene. The protein-coding region of the 94K protein gene extended uninterrupted for 2,409 bp (803 amino acids; Fig. 6). Three different RNA start sites, α_3, α_4, and α_5 (−13, −79, and −194) were mapped by S1 nuclease analysis (Fig. 2C). While the 5′ leader of the α_3 RNA was unusually short (13 nucleotides), the sequences surrounding the first encountered initiator codon, ATG (+1), conformed to the consensus for eucaryotic mRNAs, including the adenosine residue at −3 (16, 17); this was also the case for the translation initiator of the 35K protein gene (Fig. 5). The sequence at the most frequently used RNA start sites, α_3 and α_4 (Fig. 6), closely matched a well-conserved heptanucleotide, ATCA(G,T)T(C,T) found at the 5′ end of many insect genes (13, 16). An AT-rich sequence containing a consensus TATA box was positioned 25 bp upstream from the α_3 start but downstream from the α_4 start. Interestingly, transcription from the α_3 start was stimulated in cells treated with cycloheximide, while that from α_4 and α_5 was reduced slightly (Fig. 2C). Termination of 94K protein gene transcription occurred within an AT-rich sequence (+2449 to 2455) located 10 bp downstream from the second of two polyadenylation signals, AATAAA (5). Thus, the expected size of an unspliced 94K protein gene transcript (2.47 to 2.65 kb) corresponded well to the size of the major 2.63-kb α_3 RNA (Fig. 1), which included a 130- to 150-nucleotide poly(A) tract.

Another early, but less abundant, RNA (1.4 kb) was homologous to the C-terminal end of the 94K protein gene (Fig. 1C). Preceded by a consensus TATA box, a potential start site for this RNA was located (+1186) which resembled the α_3 and α_4 sites (Fig. 5). The messenger activity, if any, of this RNA has yet to be determined. Only 50 bp further upstream, a rarely used transcriptional termination site (+1140) was previously mapped (8). This AT-rich sequence is directly preceded by a polyadenylation signal (+1126) and raises the interesting question of why a majority of the 94K protein gene transcripts ignored this termination site for the preferred site 1.3 kb downstream (+2453). The relative frequency of termination at the upstream signal (+1140) increased dramatically after integration of the solo TED LTR within the 94K protein gene (+384, Fig. 6); the LTR directed transcription of abundant RNAs extending in both directions into the 94K and 35K protein genes (8).

Nucleotide sequence of the 35K protein gene. The coding region of the 35K protein gene extended for 897 bp (299 amino acids) in the direction opposite that of the adjacent 94K protein gene (Fig. 5). Only a single RNA start site, α_1 (−35), was mapped immediately upstream by S1 nuclease and primer extension analyses (6). Transcriptional termination occurred in an AT-rich sequence (+908) located 10 bp downstream from a single polyadenylation signal; thus, the size of the expected 35K protein gene transcript corre-

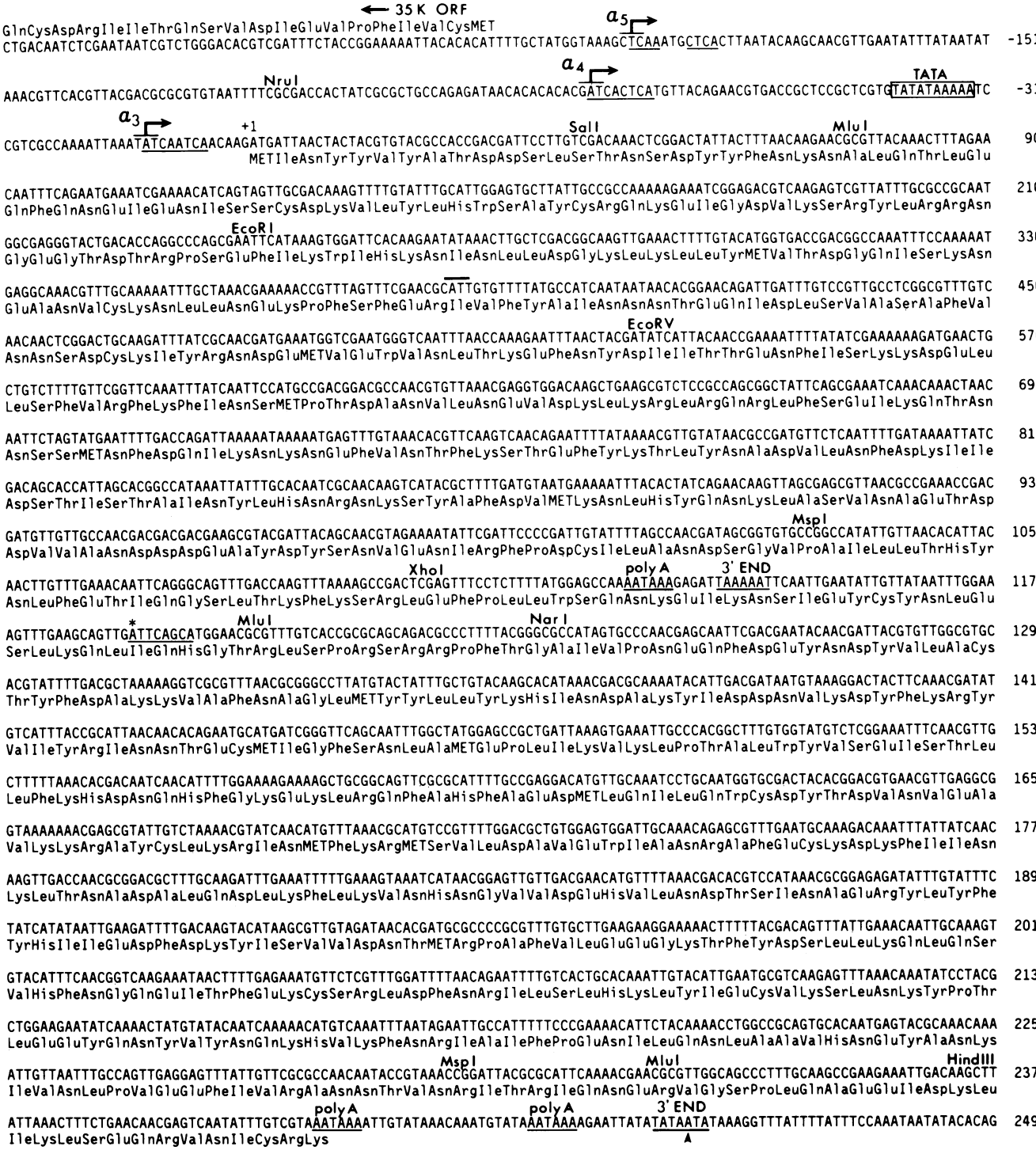

FIG. 6. Nucleotide sequence and predicted amino acid sequence of the early gene encoding the 94K polypeptide. The gene extends right to left (5′ to 3′, counterclockwise) from 86.8 to 85.0 m.u. on the AcNPV genome. The DNA strand having the polarity of the major α_3 RNA (2.63 kb) is shown and is numbered from the initiator codon (ATG, +1). The 5′ start sites α_3, α_4, and α_5 (positions −13, −79, and −194, respectively) and the 3′ termination site (+2453) are underlined. The location of potential regulatory signals for transcription are shown. The four bases, CATT (+384), duplicated upon integration of the retrotransposon TED (8) are overlined. The start of the coding region of the divergent gene encoding the 35K protein is located 210 bp upstream from the 94K protein gene in the opposite DNA strand.

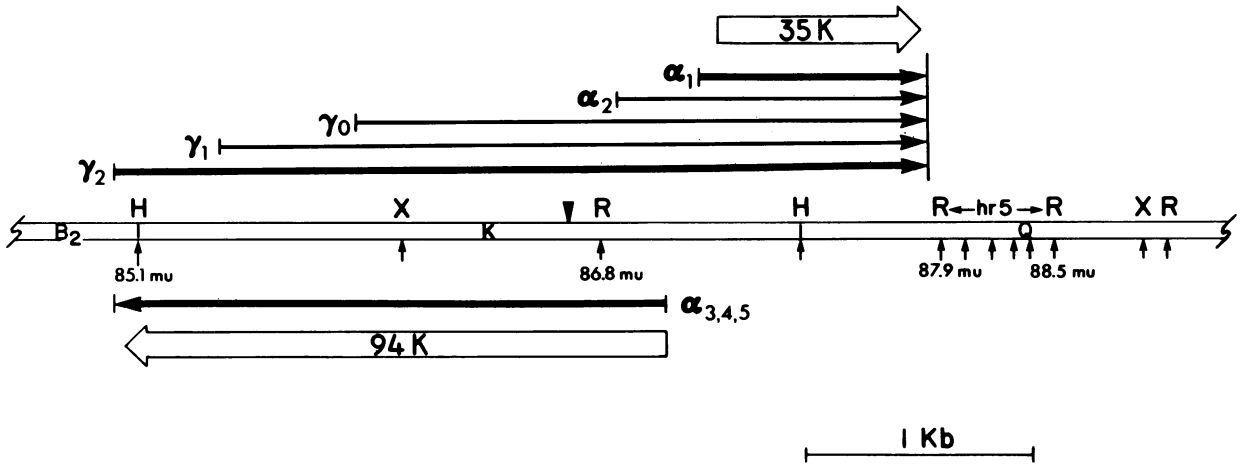

FIG. 7. Genetic and transcriptional organization of the AcNPV genome from 85 to 88 m.u. The 3.7-kb region encompassed by the *Hin*dIII K (85.1 to 87.5 m.u.) and *Eco*RI S (86.8 to 87.9 m.u.) fragments contains two divergent genes encoding 35K and 94K proteins. The coding region of each gene is depicted by open arrows. Both genes are transcribed by multiple, overlapping RNAs. The locations of the most abundant RNAs, α_1, α_3, and γ_2, are shown. The site of integration (▼) of the retrotransposon TED of AcNPV mutant FP-D is within the gene encoding the 94K polypeptide. The *hr*5 region, containing multiple *Eco*RI sites (87.9 to 88.5 m.u.) and involved in the enhancement of AcNPV gene transcription (11), is located immediately downstream from the gene encoding the 35K polypeptide. The *Hin*dIII K and Q fragments are marked. H, *Hin*dIII; R, *Eco*RI; X, *Xho*I.

sponded well to the 1.07-kb α_1 RNA identified by Northern blots (Fig. 1B). The 5′ end of the minor 1.38-kb α_2 RNA (Fig. 7) mapped approximately 450 bp upstream from the 35K protein gene initiator codon. The presence of numerous 5′ termination codons suggests that, if messenger-active, this RNA also encodes the 35K protein. Like the nearby 94K protein gene promoter region, an AT-rich sequence containing a consensus TATA box was located 20 bp upstream from the α_1 start site. While sequences at or near the α_1 start site bore partial homology to the 94K protein gene α_3 and α_4 starts, the α_1 start also resembled a portion of the highly conserved 12-mer (AATAAGTATTTT, matches underlined) located at or near the transcriptional start of baculovirus hyperexpressed genes (31). The significance of this sequence in viral gene expression is currently unclear.

The function of the 35K and 94K proteins remains to be determined. The early window (2 to 12 h) for the synthesis of their mRNAs suggests an early role in viral replicative processes. The isolation of AcNPV mutants, FP-D and FP-DS, which carry insertions of the retrotransposon TED (7.3 kb) or a single TED LTR (0.27 kb), respectively, suggests that the 94K gene is nonessential for viral replication. Insertion of the TED sequences, containing numerous translational termination codons, disrupted the 94K protein gene ORF (8). Not ruled out is the possibility that a truncated 94K protein or proteins translated from the new RNAs originating from within the TED LTR provide essential functions normally provided by the intact 94K protein gene.

Potential regulatory functions for the overlapping *Hin*dIII-K RNAs. Late in infection (12 to 24 h), the *Hin*dIII-K region was transcribed into heterogeneous RNAs whose 5′ ends lay 2 kb or more upstream from the 35K protein gene (Fig. 7). The 3.50-kb γ_2 RNA is an abundant representative of this group. The numerous translational termination codons found in all three reading frames of the 5′ half of the γ_2 RNA (Fig. 3) suggest that this RNA encodes a disproportionately small protein, if any. We found no evidence for frequent splicing of this RNA, although the presence of spliced RNAs in low abundance was not ruled out. The abundance of the γ RNAs and the observation that they overlap both the 35K and 94K protein genes suggest an alternative role for these late transcripts, involving the regulation of the early genes. Since the γ RNAs are complementary to the overlapping 94K protein gene (Fig. 7), they possess the capacity to hybridize to the 94K protein gene α_3 transcripts and thereby block translation. Such an antisense mechanism (14, 27) would facilitate the turnoff of 94K protein synthesis no longer required late in infection. Additionally, transcriptional interference (promoter occlusion) caused by transcription complexes initiating from the upstream γ promoters and blocking transcriptional initiation at downstream promoters (1, 2, 6) would further down-regulate the expression of the 94K and 35K protein genes. Both *cis*-acting mechanisms may operate in concert with other diffusable viral gene products to turn off early α gene expression.

ACKNOWLEDGMENTS

We thank Karen Adams for technical assistance and Kathy Hoefer for typing the manuscript.

This work was supported in part by Public Health Service grant AI23719 from the National Institute of Allergy and Infectious Disease and by a grant from the University of Idaho Research Council.

LITERATURE CITED

1. **Adhya, S., and M. Gottesman.** 1982. Promoter occulsion: transcription through a promoter may inhibit its activity. Cell **29**:939–944.
2. **Cullen, B. R., P. T. Lomedico, and G. Ju.** 1984. Transcriptional interference in avian retroviruses: implications for the promoter insertion model of leukamogenesis. Nature (London) **307**: 241–245.
3. **Erlandson, M. A., and E. B. Carstens.** 1983. Mapping early transcription products of *Autographa californica* nuclear polyhedrosis virus. Virology **126**:398–402.
4. **Esche, H., and B. Siegmann.** 1982. Expression of early viral gene products in adenovirus type 12-infected and -transformed cells. J. Gen. Virol. **60**:99–113.
5. **Fitzgerald, M., and T. Shenk.** 1981. The sequence 5′-AAUAAA-

3′ forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell **24**:251–260.
6. **Friesen, P. D., and L. K. Miller.** 1985. Temporal regulation of baculovirus RNA: overlapping early and late transcripts. J. Virol. **54**:392–400.
7. **Friesen, P. D., and L. K. Miller.** 1986. Regulation of baculovirus gene expression. Curr. Top. Microbiol. Immunol. **131**:31–49.
8. **Friesen, P. D., W. C. Rice, D. W. Miller, and L. K. Miller.** 1986. Bidirectional transcription from a solo long terminal repeat of the retrotransposon TED: symmetrical RNA start sites. Mol. Cell. Biol. **6**:1599–1607.
9. **Fristensky, B., J. Lis, and R. Wu.** 1982. Portable microcomputer software for nucleotide sequence analysis. Nucleic Acids Res. **10**:6451–6463.
10. **Guarino, L. A., and M. D. Summers.** 1986. Functional mapping of a *trans*-activating gene required for expression of a baculovirus delayed-early gene. J. Virol. **57**:563–571.
11. **Guarino, L. A., and M. D. Summers.** 1986. Interspersed homologous DNA of *Autographa californica* nuclear polyhedrosis virus enhances delayed-early gene expression. J. Virol. **60**:215–223.
12. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351–359.
13. **Hultmark, D., R. Klemenz, and W. J. Gehring.** 1986. Translational and transcriptional control elements in the untranslated leader of the heat-shock gene hsp22. Cell **44**:429–438.
14. **Izant, J. G., and H. Weintraub.** 1985. Constitutive and conditional suppression of exogenous and endogenous genes by anti-sense RNA. Science **229**:345–352.
15. **Kelly, D. C., and T. Lescott.** 1981. Baculovirus replication: protein synthesis in *Spodoptera frugiperda* cells infected with *Trichoplusia ni* nuclear polyhedrosis virus. Microbiologica (Bologna) **4**:35–57.
16. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12**:857–872.
17. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44**:283–292.
18. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
19. **Lee, H. H., and L. K. Miller.** 1978. Isolation of genotypic variants of *Autographa californica* nuclear polyhedrosis virus. J. Virol. **27**:754–767.
20. **Lübbert, 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.
21. **Lübbert, H., and W. Doerfler.** 1984. Transcription of overlapping sets of RNAs from the genome of *Autographa californica* nuclear polyhedrosis virus: a novel method for mapping RNAs. J. Virol. **52**:255–265.
22. **Mainprize, T. M., K.-J. Lee, and L. K. Miller.** 1986. Variation in the temporal expression of overlapping baculovirus transcripts. Virus Res. **6**:85–99.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Meinkoth, J., and G. Wahl.** 1984. Hybridization of nucleic acids immobilized on solid supports. Anal. Biochem. **138**:267–284.
25. **Miller, D. W., and L. K. Miller.** 1982. A virus mutant with an insertion of a *copia*-like transposable element. Nature (London) **299**:562–564.
26. **Pelham, H. R. B., and R. J. Jackson.** 1976. Efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. **67**:247–256.
27. **Pestka, S., B. L. Daugherty, V. Jung, K. Hotta, and R. K. Pestka.** 1984. Anti-mRNA: specific inhibition of translation of single mRNA molecules. Proc. Natl. Acad. Sci. USA **81**:7525–7528.
28. **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.
29. **Rice, W. C., and L. K. Miller.** 1986. Baculovirus transcription in the presence of inhibitors and in nonpermissive *Drosophilia* cells. Virus Res. **6**:155–172.
30. **Rohel, D. Z., and P. Faulkner.** 1984. Time course and mapping of *Autographa californica* nuclear polyhedrosis virus transcripts. J. Virol. **50**:739–747.
31. **Rohrmann, G. F.** 1986. Polyhedrin structure. J. Gen. Virol. **67**:1499–1513.
32. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
33. **Shih, D. S., C. T. Shih, D. Zimmern, R. R. Rueckert, and P. Kaesberg.** 1979. Translation of encephalomyocarditis virus RNA in reticulocyte lysates: kinetic analysis of the formation of virion proteins and a protein required for processing. J. Virol. **30**:472–480.
34. **Vaughn, J. L., R. H. Goodwin, G. L. Thompkins, and P. McCawley.** 1977. Establishment of two insect cell lines from the insect *Spodoptera frugiperda* (Lepidoptera: Noctuidae). In Vitro (Rockville) **13**:213–217.
35. **Weaver, R. F., and C. Weissman.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S β-globin mRNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucleic Acids Res. **7**:1175–1193.