JOURNAL OF VIROLOGY, July 1987, p. 2091–2099
0022-538X/87/072091-09$02.00/0

Vol. 61, No. 7

Nucleotide Sequence and Temporal Expression of a Baculovirus Regulatory Gene

LINDA A. GUARINO* AND MAX D. SUMMERS

Entomology Department, Texas A&M University, and Texas Agricultural Experiment Station, College Station, Texas 77843-2475

Received 2 February 1987/Accepted 31 March 1987

The nucleotide sequence of a *trans*-activating regulatory gene (IE-1) of the baculovirus *Autographa californica* nuclear polyhedrosis virus has been determined. This gene encodes a protein of 581 amino acids with a predicted molecular weight of 66,856. A DNA fragment containing the entire coding sequence of IE-1 was inserted downstream of an RNA promoter. Subsequent cell-free transcription and translation directed the synthesis of a single peptide with an apparent molecular weight of 70,000. Quantitative S1 nuclease analysis indicated that IE-1 was maximally synthesized during a 1-h virus adsorption period and that steady-state levels of IE-1 message were maintained during the first 24 h of infection. Northern blot hybridization indicated that several late transcripts which overlap the IE-1 gene were transcribed from both strands. The precise locations of the 5′ and 3′ ends of these overlapping transcripts were mapped using S1 nuclease. The overlapping transcripts were grouped in two transcriptional units. One unit was composed of IE-1 and overlapping γ transcripts which initiated upstream of IE-1 and terminated downstream of IE-1. The other unit, transcribed from the opposite strand, consisted of γ transcripts with coterminal 5′ ends and extended 3′ ends. The shorter, more abundant transcripts in this unit overlapped 30 to 40 bases of IE-1 at the 3′ end, while the longer transcripts overlapped the entire IE-1 gene. Transcription of several early *A. californica* nuclear polyhedrosis virus genes, in addition to 39K, was shown to be *trans*-activated by IE-1, indicating that IE-1 may have a central role in the regulation of β-gene expression.

The baculovirus *Autographa californica* nuclear polyhedrosis virus (AcNPV) has a double-stranded, covalently closed, circular DNA genome of 130 kilobases (kb). The expression of AcNPV genes in infected cells is coordinately regulated and sequentially ordered (6, 26, 40). Four general classes of genes are expressed from the AcNPV genome: α or immediate early, β or delayed early, γ or late, and δ or very late (14, 19, 28).

The regulation of baculovirus gene expression is an area of considerable interest. Several lines of evidence indicate that baculovirus gene expression is primarily regulated at the level of transcription. Hybridization mapping and cell-free translation of AcNPV transcripts purified at different times postinfection indicate that the composition of viral mRNA changes with time of infection (10, 11, 23, 38). Overlapping transcripts have been mapped in several regions of the AcNPV genome (12, 24, 29), although splicing of baculovirus transcripts does not appear to be common (23). Many of these overlapping units consist of transcripts of different temporal classes with coterminal 3′ ends and extended 5′ ends. It has been suggested that this type of overlapping transcriptional unit regulates temporal expression of baculovirus genes by sequential activation of upstream promoters and deactivation of downstream promoters (12). Some overlapping transcriptional units are composed of transcripts which share a common 5′ end and have extended 3′ ends. All units of this type reported to date are composed of either all early or all late transcripts which belong to a single temporal class (12, 29), indicating that 3′ extensions may not play a role in temporal regulation.

Little is known concerning the mechanisms which control temporal expression during baculovirus infection. Recently we described an α gene, IE-1, which *trans*-activates transcription of the β gene 39K (15). Expression of 39K was enhanced 1,000-fold when a baculovirus repeated element was *cis*-linked to the 39K gene (16, 17). The AcNPV enhancer was not active in the absence of the IE-1 gene, indicating that IE-1 may regulate the enhancer as well as the 39K promoter.

To further probe the function of IE-1, we determined the nucleotide sequence of the regulatory gene and compared the sequence of IE-1 with those of heterologous viral *trans*-activating genes. We also show that at least four other AcNPV early genes in addition to 39K require IE-1 for their expression in transient assay, suggesting that IE-1 has a central role in regulation of β-gene expression.

MATERIALS AND METHODS

Virus and cell culture. The conditions for cell culture and transfections have been described previously (15). *Spodoptera frugiperda* cells were infected with AcNPV E2 at a multiplicity of infection of 20. The virus was allowed to adsorb for 1 h at 24°C. The inoculum was removed, and the cells were incubated at 27°C for the indicated times. Zero time postinfection is defined as the end of the adsorption period.

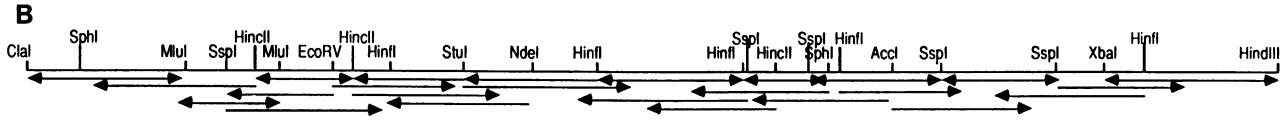
Cell-free transcription and translation. Sequences in the 5′-flanking region of IE-1 upstream of the initiating AUG were deleted with BAL 31, according to standard procedures (25). The plasmid p*Hin*dIII-G, which contains the complete IE-1 gene, was linearized with *Cla*I and digested with BAL 31 for various lengths of time. The DNA was then digested with *Sal*I and religated after repair of the ends with Klenow fragment. After an initial screening by agarose gel electrophoresis, several deletions were sequenced. One clone (pIE-33), in which the 5′-flanking sequences were deleted up to base 537 (Fig. 1), was digested with *Bam*HI, which cuts in

* Corresponding author.

A

```
1    ATCGATGTCTTTGTGATGCGCCGACATTTTTGTAGGTTATTGATAAAATGAACGGATACAGTTGCCCGACATTATCATTAAATCCTTGGCGTAGAATTTGTCGGGTCCATTGTCCGTGTG

121  CGCTAGCATGCCCGCTAACGGACCTCGTACTTTTGGCTTCAAAGGTTTTGCGCACAGACAAAATGTGCCACACTTGCAGCTCTGCATGTGTGCGCGTTACCACAAATCCCAACGGCGCAG

241  TGTACTTGTTGTATGCAAATAAATCTCGATAAAGGCGCGGCGCGCGAATGCAGCTGATCACGTACGCTCCTCGTGTTCCGTTCAAGGACGGTGTTATCGACCTCAGATTAATGTTTATCG

361  GCCGACTGTTTTCGTATCCGCTCACCAAACGCGTTTTTGCATTAACATTGTATGTCGGCGGATGTTCTATATCTAATTTGAATAAATAAACGATAACCGCGTTGGTTTTAGAGGGCATAA

481  TAAAAGAAATATTGTTATCGTGTTCGCCATTAGGGCAGTATAAATTGACGTTCATGTTGGATATTGTTTCAGTTGCAAGTTGACACTGGCGGCGACAAGATCGTGAACAACCAAGTGACT
                                                                                                    5'

601  ATGACGCAAATTAATTTTAACGCGTCGTACACCAGCGCTTCGACGCCGTCCCGAGCGTCGTTCGACAACAGCTATTCAGAGTTTTGTGATAAACAACCCAACGACTATTTAAGTTATTAT
     M  T  Q  I  N  F  N  A  S  Y  T  S  A  S  T  P  S  R  A  S  F  D  N  S  Y  S  E  F  C  D  K  Q  P  N  D  Y  L  S  Y  Y

721  AACCATCCCACCCCGGATGGAGCCGACACGGTGATATCTGACAGCGAGACTGCGGCACGTTCAAACTTTTTGGCAAGCGTCAACTCGTTAACTGATAATGATTTAGTGGAATGTTTGCTC
     N  H  P  T  P  D  G  A  D  T  V  I  S  D  S  E  T  A  A  R  S  N  F  L  A  S  V  N  S  L  T  D  N  D  L  V  E  C  L  L

841  AAGACCACTGATAATCTCGAAGAAGCAGTTAGTTCTGCTTATTATTCGGAATCCCTTGAGCAGCCTGTTGTGGAGCAACCATCGCCCAGTTCTGCTTATCATGCGGAATCTTTTGAGCAT
     K  T  T  D  N  L  E  E  A  V  S  S  A  Y  Y  S  E  S  L  E  Q  P  V  V  E  Q  P  S  P  S  S  A  Y  H  A  E  S  F  E  H

961  TCTGCTGGTGTGAACCAACCATCGGCAACTGGAACTAAACGGAAGCTGGACGAATACTTGGACAATTCACAAGGTGTGGTGGGCCAGTTTAACAAAATTAAATTGAGGCCTAAATACAAG
     S  A  G  V  N  Q  P  S  A  T  G  T  K  R  K  L  D  E  Y  L  D  N  S  Q  G  V  V  G  Q  F  N  K  I  K  L  R  P  K  Y  K

1081 AAAAGCACAATTCAAAGCTGTGCAACCCTTGAACAGACAATTAATCACAACACGAACATTTGCACGGTCGCTTCAACTCAAGAAATTACGCATTATTTTTACTAATGATTTTGCGCCGTAT
     K  S  T  I  Q  S  C  A  T  L  E  Q  T  I  N  H  N  T  N  I  C  T  V  A  S  T  Q  E  I  T  H  Y  F  T  N  D  F  A  P  Y

1201 TTAATGCGTTTCGACGACAACGACTACCAATTCCAACAGGTTCTCGACCATATGTCCGAAACTGGTTATTACATGTTTGTGGTTAAAAAAAGTGAAGTGAAGCCGTTTGAAATTATATTT
     L  M  R  F  D  D  N  D  Y  Q  F  Q  Q  V  L  D  H  M  S  E  T  G  Y  Y  M  F  V  V  K  K  S  E  V  K  P  F  E  I  I  F

1321 GCCAAGTACGTGAGCAATGTGGTTTACGAATATACAAACAATTATTACATGGTAGATAATCGCGTGTTTGTGGTAACTTTTGATAAAATTAGGTTTATGATTTCGTACAATTTGGTTAAA
     A  K  Y  V  S  N  V  V  Y  E  Y  T  N  N  Y  Y  M  V  D  N  R  V  F  V  V  T  F  D  K  I  R  F  M  I  S  Y  N  L  V  K

1441 GAAACCGGCATAGAAATTCCTCATTCTCAAGATGTGTGCAACGACGAGACGGCTGCACAAAATTGTAAAAAATGCCATTTCGTCGATGTGCACCACACGTTTAAAGCTGCTCTGACTTCA
     E  T  G  I  E  I  P  H  S  Q  D  V  C  N  D  E  T  A  A  Q  N  C  K  K  C  H  F  V  D  V  H  H  T  F  K  A  A  L  T  S

1561 TATTTTAATTTAGATATGTATTACGCGCAAACCACATTTGTGACTTTGTTACAATCGTTGGGCGAAAGAAAATGTGGGTTTCTTTTGAGCAAGTTGTACGAAATGTATCAAGATAAAAAT
     Y  F  N  L  D  M  Y  Y  A  Q  T  T  F  V  T  L  L  Q  S  L  G  E  R  K  C  G  F  L  L  S  K  L  Y  E  M  Y  Q  D  K  N

1681 TTATTTACTTTGCCTATTATGCTTAGTCGTAAAGAGAGTAATGAAATTGAGACTGCATCTAATAATTTCTTTGTATCGCCGTATGTGAGTCAAATATTAAAGTATTCGGAAAGTGTGCAG
     L  F  T  L  P  I  M  L  S  R  K  E  S  N  E  I  E  T  A  S  N  N  F  F  V  S  P  Y  V  S  Q  I  L  K  Y  S  E  S  V  Q

1801 TTTCCCGACAATCCCCCAAACAAATATGTGGTGGACAATTTAAATTTAATTGTTAACAAAAAAAGTACGCTCACGTACAAATACAGCAGCGTCGCTAATCTTTTGTTTAATAATTATAAA
     F  P  D  N  P  P  N  K  Y  V  V  D  N  L  N  L  I  V  N  K  K  S  T  L  T  Y  K  Y  S  S  V  A  N  L  L  F  N  N  Y  K

1921 TATCATGACAATATTGCGAGTAATAATAACGCAGAAAATTTAAAAAAGGTTAAGAAGGAGGACGGCAGCATGCACATTGTCGAACAGTATTTGACTCAGAATGTAGATAATGTAAAGGGT
     Y  H  D  N  I  A  S  N  N  N  A  E  N  L  K  K  V  K  K  E  D  G  S  M  H  I  V  E  Q  Y  L  T  Q  N  V  D  N  V  K  G

2041 CACAATTTTATAGTATTGTCTTTCAAAAACGAGGAGCGATTGACTATAGCTAAGAAAAACAAAGAGTTTTATTGGATTTCTGGCGAAATTAAAGATGTAGACGTTAGTCAAGTAATTCAA
     H  N  F  I  V  L  S  F  K  N  E  E  R  L  T  I  A  K  K  N  K  E  F  Y  W  I  S  G  E  I  K  D  V  D  V  S  Q  V  I  Q

2161 AAATATAATAGATTTAAGCATCACATGTTTGTAATCGGTAAAGTGAACCGAAGAGAGAGCACTACATTGCACAATAATTTGTTAAAATTGTTAGCTTTAATATTACAGGGTCTGGTTCCG
     K  Y  N  R  F  K  H  H  M  F  V  I  G  K  V  N  R  R  E  S  T  T  L  H  N  N  L  L  K  L  L  A  L  I  L  Q  G  L  V  P

2281 TTGTCCGACGCTATAACGTTTGCGGAACAAAAACTAAATTGTAAATATAAAAAATTCGAACTTTAATTAATTATACATATATTTTGAATTTAATTAATTATACATATATTTTATATTATT
     L  S  D  A  I  T  F  A  E  Q  K  L  N  C  K  Y  K  K  F  E  L  *                                                    3'

2401 TTTGTCTTTTATTATCGAGGGGCCGTTGTTGGTGTGGGGTTTTGCATAGAAATAACAATGGGAGTTGGCGACGTTGCTGCGCCAACACCACCTCCTTCCTCCTCCTTTCATCATGTATCT

2521 GTAGATAAAATAAAATATTAAACCTAAAAACAAGACCGCGCCTATCAACAAAATGATAGGCATTAACTTGCGTGACGCTGTCACTAACGTTGGACGATTTGCCGACTAAACCTTCATCGC

2641 CCAGTAACCAATCTAGACCCAAGTCGCCAACTAAATCACCAAACGAGTAAGGTTCGATGCACATGAGTGTTTGGCCCGCAGGAAGATCGCTAATATCTACGTATTGAGGCGAATCTGGGT

2761 CGGCGGACGGATCGCTGCCGCGACAAACTGTTTTTTCTACTTCATAGTTGAATCCTTGGCACATGTTGGTTAGTTCGGGCGGATTGTTAGGCAACAAGGGGTCGAATGGGCAAATGGTAA

2881 CATCCGACTGATTTAGATTGGGGTCTTGACGACAAGTGCGCTGCAATAACAAGCAGGCCTCGGCGATTTCTCCGGCGTCTTTACCTTGCACATAATAACTTCCGCCGGTGTTATTGATGG

3001 CGTTGATTATATCTTGTACTAGTGTGGCGGCGCTAAACAAGAAATAGCCCCGGTGGCCAAGAGTATGCCCGTTCCTCCTACTTTTAAGCTT 3091
```

B



the multiple cloning sites, and *Xba*I, which cuts downstream of the translation termination codon. This 2.1-kilobase-pair (kbp) *Bam*HI-*Xba*I fragment was cloned into the *Bam*HI and *Xba*I sites of the vector Blue Scribe M13+ (Stratagene Cloning Systems, San Diego, Calif.). After linearization with *Xba*I, the plasmid pBS-33 was used for in vitro transcription with T7 polymerase according to the procedure recommended by the manufacturer, except that the DNase treatment was omitted. The RNA transcript was purified from the DNA template by gel electrophoresis in low-melting-temperature agarose.

After elution from agarose, the uncapped RNA transcript was translated in vitro with a rabbit reticulocyte lysate (Promega Biotech, Madison, Wis.) in the presence of [^{35}S]methionine. A 2-μl sample of the translation mix was diluted in loading buffer, and the proteins were separated on a 12% sodium dodecyl sulfate-polyacrylamide gel (21).

DNA sequence analysis. Appropriate restriction fragments were subcloned into bacteriophage M13 (27) and sequenced by the chain termination procedure (33). BAL 31 deletions of IE-1, constructed using standard protocols, were also sequenced. The sequences were compiled and analyzed by the programs of Devereaux et al. (9). The National Biomedical Research Foundation databank, release 9.0, was searched for homology to the IE-1 protein.

RNA extraction and gel electrophoresis. Whole-cell RNA was prepared by guanidinium isothiocyanate solubilization (7). Selection of poly(A)$^+$ RNA was by established techniques (25). Poly(A)$^+$ RNA was separated by electrophoresis through 1.5% agarose gel containing 2.2 M formaldehyde. The gels were transferred to nitrocellulose and hybridized with pIE-1 DNA uniformly labeled by nick translation or with strand-specific probes by standard procedures (25).

Nuclease S1 analysis was performed with 10 μg of whole-cell RNA from infected or transfected cells, as previously described (39). The hybridization temperatures were determined empirically for each probe and are indicated in the figure legends. The probes were labeled at the 5′ end, using T4 polynucleotide kinase, or at the 3′ end, using the Klenow fragment of *Escherichia coli* DNA polymerase or T4 DNA polymerase, following standard protocols (25). The S1-protected DNAs were analyzed in 4% polyacrylamide–7 M urea gels, using ϕX174-*Hae*III fragments as molecular standards.

RESULTS

Nucleotide sequence of IE-1. To sequence IE-1, appropriate restriction fragments were isolated, subcloned into M13 vectors, and sequenced by the chain termination method (27, 33). The sequencing strategy is indicated in Fig. 1B. The complete nucleotide sequence of pIE-1 and the predicted amino acid sequence of the longest open reading frame are presented in Fig. 1B. The polarity of the longest open reading frame is consistent with the direction of transcription previously determined for IE-1 (15).

The sequence indicates that the IE-1 gene product is composed of 581 amino acids with a molecular weight of 66,856. The protein is highly charged with the acidic residues aspartic acid, glutamic acid, asparagine, and glutamine, accounting for 25% of the total amino acids. The basic amino acids arginine, lysine, and histidine account for 13% of the residues, giving a net negative charge of 8. The hydrophobic residues account for only 6% of the total residues.

The 5′ end of the IE-1 mRNA was previously mapped to a position 341 bp upstream of a *Hin*fI site (15). Comparison of the sequence and mapping data indicates that the 5′ end is located approximately 50 bp upstream of the initiation codon. A putative TATA box (TATAAATT) is located 25 to 32 bp upstream of the cap site. A sequence similar to the consensus CAAT box (GGCATAAT) is located 69 to 77 bp upstream of the start site for transcription.

The 3′ end of IE-1 was mapped as described in the legend to Fig. 3B. The 3′ end is located 31 bp downstream of the translation termination codon. Although the 3′ end of IE-1 is contained within a 100-bp adenine-plus-thymine-rich region (92% A+T), there are no consensus polyadenylation signals on the IE-1 coding strand. This A+T-rich region also serves as the major transcriptional termination region for messages which are initiated downstream of and transcribed in the opposite orientation from IE-1 (see below and Fig. 6).

To gain insight into the possible function of IE-1, we compared the nucleotide sequence and the predicted amino acid sequence with the sequences of heterologous viral genes which are functionally analogous to IE-1. Comparison with adenovirus type 5 E1a protein (37), herpes simplex virus ICP4 (40), bovine papillomavirus E2 protein (6), simian virus 40 T antigen (31), Epstein-Barr virus nuclear antigen 1 (2), or the Tat protein of human T-cell lymphotrophic virus (34) exhibited no significant homology at either the nucleotide or amino acid level. In addition, we searched the National Biomedical Research Foundation databank for similar proteins. This search also revealed no significant homologies with other proteins.

Cell-free translation of IE-1 message. The plasmid pBS-33 was constructed as diagrammed in Fig. 2A. IE-1 message-sense RNA was generated in vitro from pBS-33 with the use of T7 polymerase. This uncapped RNA was used as a template for translation in a rabbit reticulocyte lysate. Cell-free translation yielded a polypeptide of apparent molecular weight 70,000 on sodium dodecyl sulfate-polyacrylamide gels (Fig. 2B). This empirical value is close to that predicted for the sequence (66,856). The synthesis of a protein of identical mobility was directed by poly(A)$^+$ RNA from AcNPV-infected cells but not from uninfected-cell RNA. The relative amount of the 70,000-molecular-weight protein directed by 6-h RNA was greater than that directed by 18-h RNA.

Time course of IE-1 transcription. To determine the peak time of IE-1 synthesis, quantitative S1 nuclease analysis was performed with an 892-nucleotide (nt) *Cla*I-*Hin*fI fragment specifically 5′ end labeled at the *Hin*fI site. Whole-cell RNA (10 μg) purified from *S. frugiperda* cells 0, 3, 6, 12, 18, and 24 h postinfection with AcNPV was hybridized with an excess of probe. After treatment with S1, the protected DNAs were analyzed on 4% polyacrylamide–7 M urea gels. The amount of IE-1-specific probe protected by infected-cell RNA (341 bp) was essentially constant during the early phase and declined slightly at later times (Fig. 3, left panel). With late RNA (12 to 24 h), the entire probe was also

FIG. 1. Nucleotide sequence and deduced amino acid sequence of AcNPV IE-1 (A). The restriction fragments indicated in panel B were subcloned into M13 and sequenced by the chain termination method (27, 33). The amino acid sequence of the longest open reading frame beginning with an initiation codon is indicated. The 5′ and 3′ ends of IE-1 are indicated, and the putative TATA and CAAT boxes are indicated.

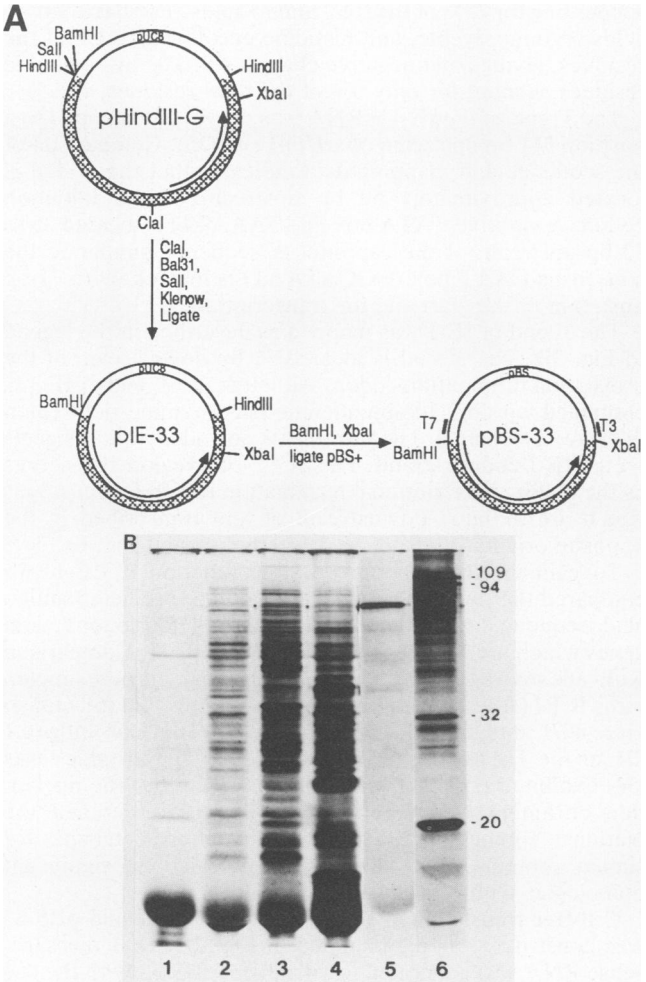


FIG. 2. Synthesis of IE-1 protein in cell-free translation system. (A) Construction of the plasmid pBS-33. Nucleotide sequences upstream of the IE-1 initiation codon were removed by BAL 31 nuclease. Then the coding sequence for IE-1 was cloned into the plasmid Blue Scribe M13+. Uncapped RNA transcribed in vitro from pBS-33 using T7 polymerase was translated in a rabbit reticulocyte lysate in the presence of [^{35}S]methionine (lane 5). The translation products were analyzed on a 12% sodium dodecyl sulfate-polyacrylamide gel (B). Lane 1, Endogenous synthesis (no RNA added); lane 2, translation products from uninfected-cell poly(A)$^+$ RNA; lane 3, translation products from 6-h-infected poly(A)$^+$ RNA; lane 4, translation products from 18-h-infected poly(A)$^+$ RNA. Lane 6, Translation products from brome mosaic virus RNA which were used for molecular size markers.

protected, indicating that at least one transcript which overlapped IE-1 was initiated upstream of the *Cla*I site.

The 3′ end of IE-1 was mapped using a 642-bp *Hin*fI-*Xba*I fragment exclusively labeled at the *Hin*fI site using the Klenow fragment of *E. coli* DNA polymerase. A single 363-nt, S1-resistant fragment was observed when the 3′ probe was hybridized to RNA isolated at 6 and 18 h postinfection. The locations of the 5′ and 3′ ends of IE-1 predict a size of 1,801 nt for IE-1 transcript, excluding the poly(A)$^+$ tail and assuming that the message is not spliced. With 18-h RNA, the entire probe was also protected, indicating that the transcripts which overlap the 5′ end of IE-1 are not 3′ coterminal with IE-1.

Identification of overlapping transcripts. To identify the transcripts which overlap IE-1, Northern blotting experiments were performed with polyadenylated RNA derived from *S. frugiperda* cells 6 and 18 h postinfection. Figure 4 shows the results obtained with pIE-1 DNA uniformly labeled by nick translation (lanes 1 through 3). During the early phase only one major transcript was evident. The size of this transcript (1.9 kb) is consistent with the size predicted for IE-1 message based on the 5′- and 3′-end mapping. Late in infection, several additional RNA transcripts were detected.

To determine whether these late transcripts were transcribed from the same or opposite strand as IE-1, poly(A)$^+$ RNA was hybridized with strand-specific probes. The plasmid pBS-33 was digested with *Bam*HI, and RNA was transcribed using T3 polymerase in the presence of [^{32}P]CTP to generate a probe which hybridized to IE-1 and to overlapping mRNAs transcribed from the same strand. Larger transcripts of 3.7, 3.5, and 3.1 kb overlap IE-1 on the same strand (Fig. 4, lanes 4 and 5). Digestion of pBS-33 with *Xba*I and transcription using T7 polymerase generated a probe to detect RNAs transcribed from the opposite strand to IE-1. The major 1.3-kb transcript was synthesized from the opposite strand as well as several less abundant, larger transcripts of 1.5, 2.1, 2.3, and 4.1 kb (Fig. 4, lanes 6 and 7).

The S1 nuclease data and Northern blot data presented above indicated that at least one transcript initiated upstream of the IE-1 promoter and terminated downstream of IE-1. To map this transcript(s), a 1.2-kbp *Bst*EII-*Cla*I fragment, specifically labeled at the *Cla*I site, was hybridized with 6- and 18-h RNA. After S1 treatment, the size of the protected fragments was determined by denaturing polyacrylamide gel electrophoresis (Fig. 5A). Two S1-resistant fragments (522 and 330 nt) were resolved when the probe was hybridized with 18-h RNA, but not with uninfected or

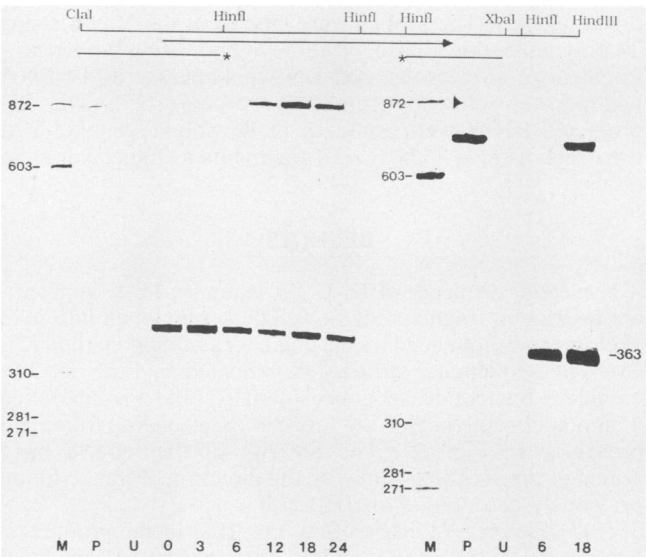


FIG. 3. Quantitative S1 nuclease analysis of the time course of transcription of the IE-1 gene. Whole-cell RNA (10 μg), uninfected (U) or at 0, 3, 6, 12, 18, and 24 h postinfection. Time zero refers to the end of a 1-h adsorption period. RNA was hybridized at 42°C with an IE-1-specific 5′-end probe, an 892-nt *Cla*I-*Hin*fI fragment specifically labeled at the *Hin*fI site (lanes 2 through 9), or at 49°C with an IE-1-specific 3′-end probe, a 642-nt *Hin*fI-*Xba*I fragment (lanes 10 through 12). ϕX174-*Hae*III molecular size markers are shown in lane 1.

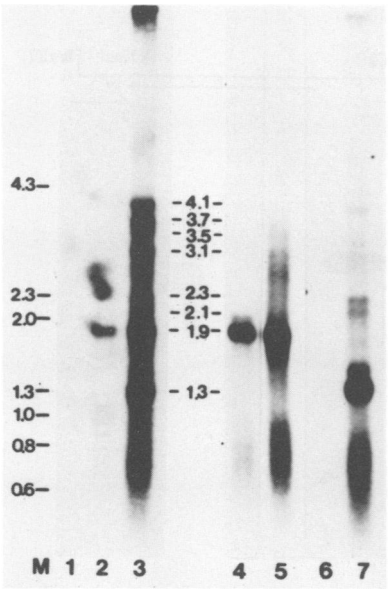


FIG. 4. RNA blot hybridization and S1 analysis of transcripts which map to the IE-1 region of the AcNPV genome. (A) Poly(A)$^+$ RNA was purified from uninfected (lane 1) *S. frugiperda* cells or AcNPV-infected cells 6 h (lanes 2, 4, and 6) or 18 h (lanes 3, 5, and 7) postinfection. RNA samples were separated on a 1.5% agarose–2.2 M formaldehyde gel, transferred to nitrocellulose, and hybridized with uniformly labeled pIE-1 DNA (lanes 1 through 3) or strand-specific probes (lanes 4 through 7). The molecular size markers are indicated on the left.

6-h RNA. The S1-protected fragment at 300 nt was not consistently observed.

The 3′ ends of the overlapping transcripts were mapped using a 431-bp *Xba*I-*Hin*dIII probe, exclusively 3′ end labeled at the *Xba*I site (Fig. 5B). After hybridization and nuclease S1 treatment, a protected band of 423 nt was observed with 18-h infected-cell RNA. No S1-resistant bands were detected with uninfected or 6-h RNA. Based on the location of the 5′ and 3′ ends, these transcripts presumably corresponded to the 3.7- and 3.5-kb transcripts observed in Fig. 4. The origin of the 3.1-kb transcript detected in Fig. 4 is not clear from this analysis. An explanation consistent with the size of the 3.1-kb trancripts is that they arose from initiation at the IE-1 start site and termination at the late transcriptional termination site.

To map the approximate location of the transcripts which hybridize to the opposite strand, Northern blot experiments were performed with subclones of pIE-1. The results indicated that all transcripts hybridized to the right-most sequences of IE-1, while only the longer transcripts hybridized to the left-most sequences (data not shown). This result suggested that the transcripts formed a nested set of mRNAs with extended 3′ ends and a common 5′ end which maps downstream of IE-1.

The precise location of the common 5′ end of the nested transcripts was determined by nuclease S1 mapping. A 713-bp *Hin*dIII-*Eco*RV fragment, specifically 5′ end labeled at the *Hin*dIII site (Fig. 5C), was hybridized with total cell RNA and treated with S1 nuclease, and the protected fragments were sized on acrylamide gels. Late RNA protected 434 nt of probe, whereas no S1-resistant fragments were detected with uninfected or early RNA.

The 3′ end of the major 1.3-kb transcript was mapped using a 667-bp *Xba*I-*Sph*I fragment specifically labeled at the *Xba*I site (Fig. 5D). S1 analysis with late RNA revealed a band with some heterogeneity at 312 nt and a single band at 324 nt. On a longer exposure of this gel a larger, less abundant fragment at 512 nt was observed (data not shown). Based on the location of the 5′ and 3′ ends, the transcripts which terminate 312 and 324 nt left of the *Xba*I site would both comprise the 1.3-kb transcript. The longer transcript presumably accounts for the 1.5-kb transcript. The smaller transcripts terminate in the A+T rich region to the left of the IE-1 termination site. Therefore, IE-1 and the 1.3-kb transcript share 30 to 40 nt of complementary sequence. The larger transcript contains 230 bp of sequence complementary to IE-1. Analysis of the nucleotide sequence in this region indicated that there were consensus polyadenylation signals (ATTAAA) located 20 to 30 bp upstream of the 3′ ends of the transcripts.

The entire probe was also protected, indicating that the longer transcripts terminate to the left of the *Sph*I site. These were mapped using an *Acc*I-*Cla*I fragment specifically 3′ end labeled at the *Acc*I site (Fig. 5D). Hybridization and S1 nuclease analysis indicated that transcripts terminate 430 and 600 nt to the left of the *Acc*I site. This accounts for the 2.1- and 2.3-kb transcripts. The largest transcript (4.1 kb) which maps to the opposite strand apparently terminates to the left of the *Cla*I site, and the exact map location of this transcript was not determined.

Figure 6A summarizes the mapping of the overlapping transcripts, and Fig. 6B shows the location of open reading frames in the region of IE-1. All of the transcripts are synthesized late, with the exception of IE-1. These transcripts represent two separate transcriptional units, all of which overlap each other to some degree. The unit containing IE-1 consists of transcripts with 5′ and 3′ ends both extended. The cluster containing the major 1.3-kb transcript consists of transcripts with common 5′ ends and extended 3′ ends. The transcripts in this unit are potentially protein coding (Fig. 6B and unpublished data). There is also an open reading frame opposite and left of IE-1. Although transcripts which map to this region were not detected by Northern blot analysis (Fig. 4), S1 analysis indicated the presence of a transcript which would encode this open reading frame (data not shown).

Additional genes *trans*-activated by IE-1. To determine whether IE-1 *trans*-activated transcription of other viral genes in addition to 39K, we analyzed transient expression of several early genes of AcNPV. The plasmid p*Hin*dIII-Q contains the promoter for two early transcripts which share a common 5′ end. These transcripts are maximally synthesized at 12 h postinfection. *S. frugiperda* cells were cotransfected with p*Hin*dIII-Q in the presence and absence of IE-1. RNA was purified 24 h posttransfection and hybridized with an 890-bp *Xho*I-*Eco*RV probe, specifically 5′ end labeled at the *Xho*I site (Fig. 7A). A single 600-nt fragment was protected with RNA purified from cells cotransfected with both fragments. This is in good agreement with the previously reported 588 nt from the *Xho* site to the 5′ end for these transcripts (30). No protected bands were observed with RNA from cells transfected with p*Hin*dIII-Q alone.

Several transcripts have been mapped to the *Hin*dIII-K fragment of AcNPV DNA (12, 13). Four of these have been classified as α transcripts; two are maximally synthesized at 2 h postinfection and two are synthesized at 6 h postinfection. *S. frugiperda* cells were transfected with *Bgl*II-D, which contains all of *Hin*dIII-K, in the presence and absence of pIE-1. RNA was isolated 24 h posttransfection and hybridized with specific 5′-end-labeled probes (Fig. 7B). A

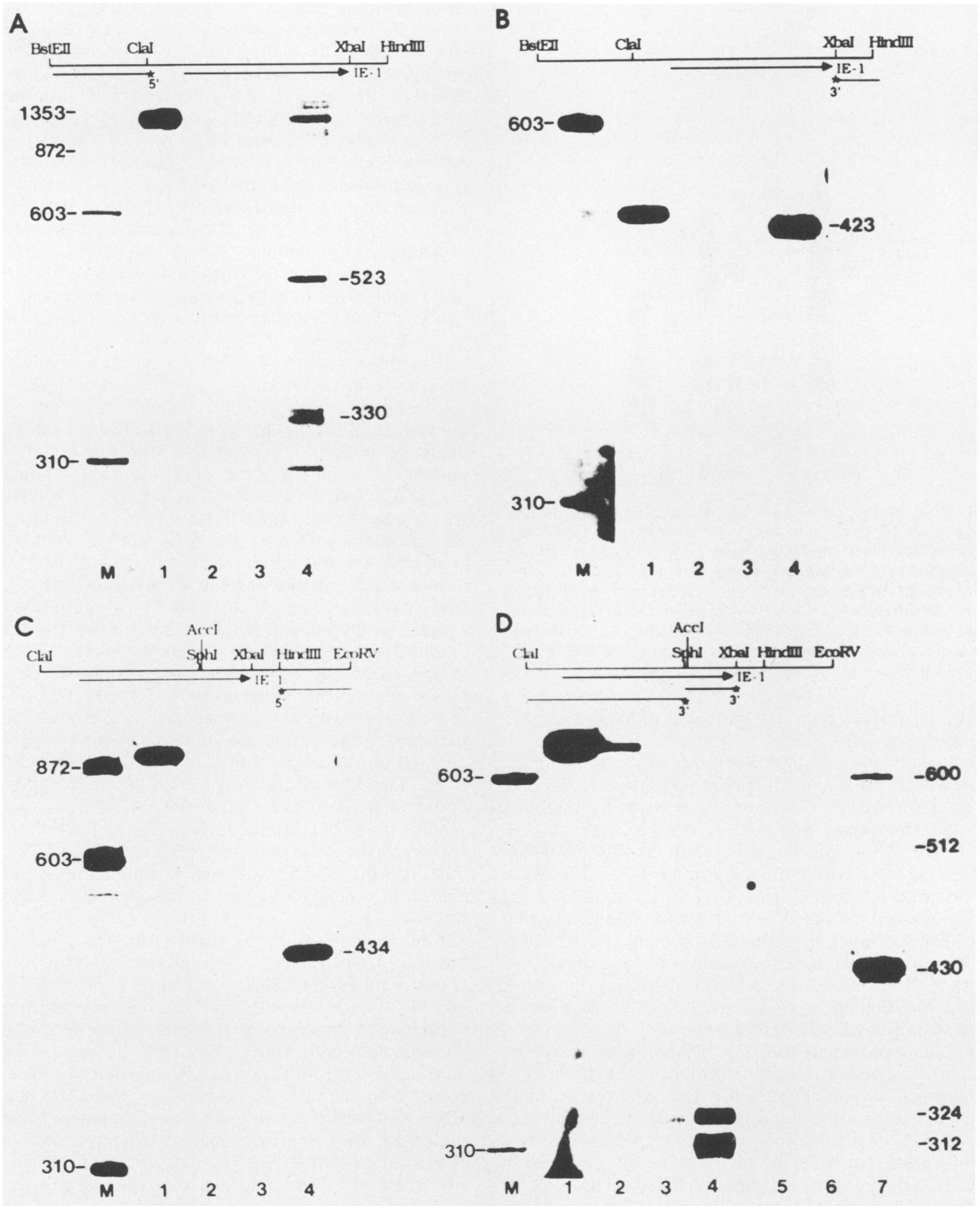


FIG. 5. S1 nuclease mapping of the transcripts which overlap IE-1. The locations of the probes relative to IE-1 are indicated in the diagram at the top of each panel. In all panels the migration of the relevant ϕX174 molecular markers is indicated in lane M. The migration of the undigested probes is shown in lane 1; the size in nucleotides of the S1-resistant fragments after hybridization with uninfected cell RNA (lane 2), 6-h-infected RNA (lane 3), or 18-h-infected RNA (lane 4) is indicated on the right. In panel D only, lanes 5, 6, and 7 show the S1-resistant fragments after hybridization with uninfected or 6- or 18-h RNA, respectively. The probe for panel A was a 1.2-kb *Bst*EII-*Cla*I fragment specifically 5′ end labeled at the *Cla*I site. The hybridization temperature was 49°C. The probe used in panel B was a 431-bp *Xba*I-*Hin*dIII fragment specifically 3′ end labeled at the *Xba*I site. The hybridization temperature was 45°C. The probe used in panel C was a 713-bp *Xba*I-*Hin*dIII fragment 5′ end labeled at the *Xba*I site. The hybridization temperature was 42°C. The probes used in panel D were a 667-bp *Xba*I-*Sph*I fragment (lanes 1 through 4) or a 2,140-bp *Acc*I-*Cla*I fragment (lanes 5 through 7), both 3′ end labeled at the *Xba*I or *Acc*I site. The hybridization temperatures were 49 and 42°C, respectively.

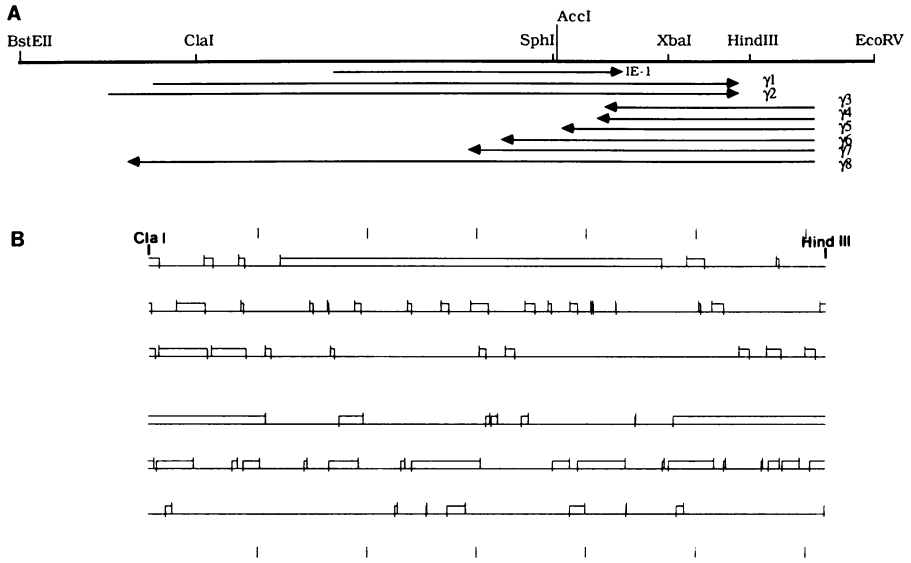


FIG. 6. Location of the transcripts which map to the region of the IE-1 gene (A) and the location of open reading frames (B).

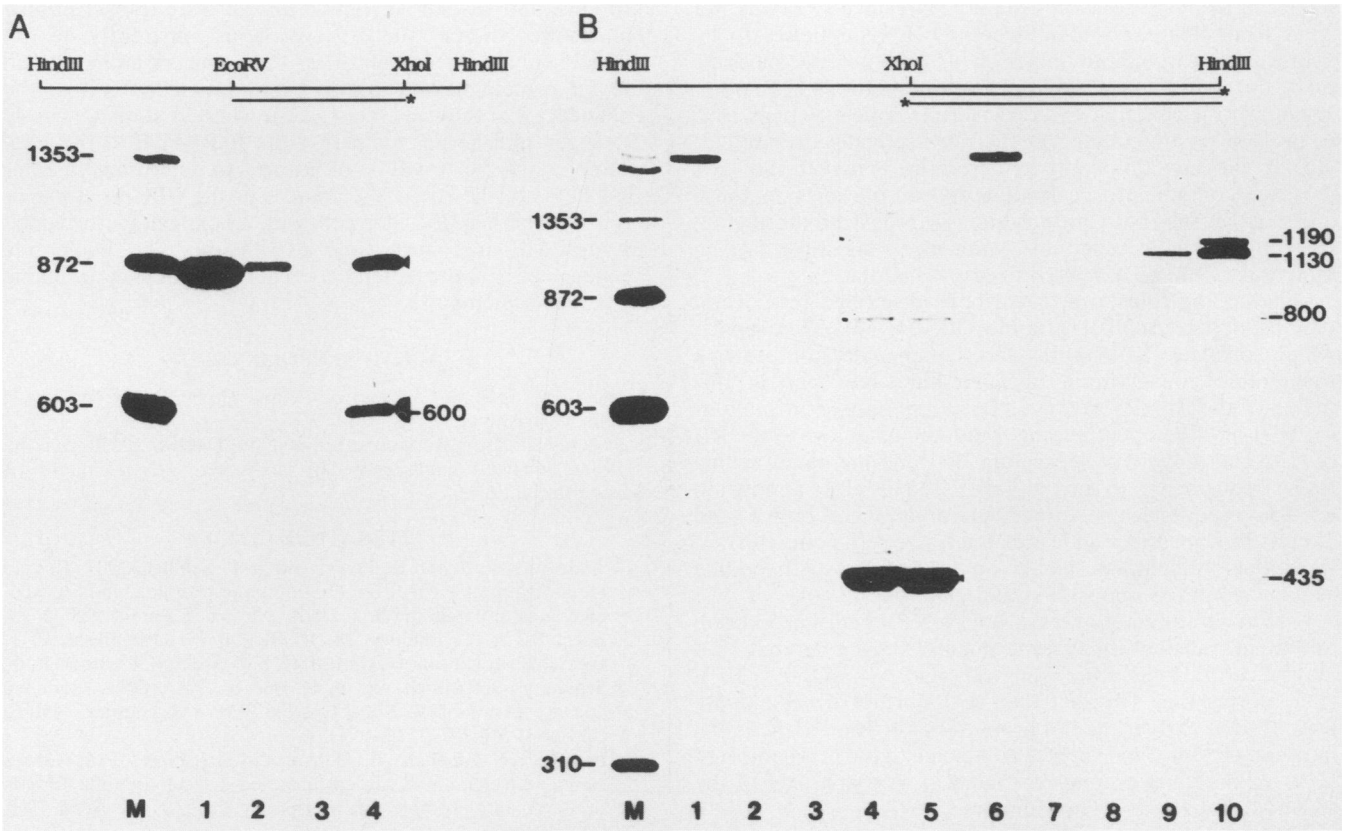


FIG. 7. S1 analysis of AcNPV delayed-early genes *trans*-activated by IE-1. (A) S1 nuclease analysis of total cell RNA purified from untransfected cells (lane 2), cell transfected with p*Hin*dIII-Q alone (lane 3), or cells transfected with p*Hin*dIII-Q and pIE-1 (lane 4) and hybridized with an 890-bp *Xho*I-*Eco*RV probe 5′ end labeled at the *Xho*I site. The undigested probe is shown in lane 1. (B) S1 nuclease analysis of total cell RNA purified from untransfected cells (lanes 2 and 7), cells transfected with *Bgl*II-E (lanes 3 and 8) or p*Bgl*II-E and pIE-1 (lanes 4 and 9), or RNA purified from cells 6 h postinfection (lanes 5 and 10). The probe used in the hybridizations shown in lanes 2 through 5 was a 1.7-kbp *Hin*dIII-*Xho*I fragment specifically 5′ end labeled at the *Hin*dIII site. The undigested probe is shown in lane 1. The probe used in the hybridizations in lanes 7 through 10 was the same 1.7-kbp *Hin*dIII-*Xho*I probe, but specifically 5′ end labeled at the *Xho*I site. The undigested probe is shown in lane 6.

1.7-kbp *Hin*dIII-*Xho*I probe labeled at the *Hin*dIII site was used to detect the transcripts which extend 440 bases and 830 bases 5′ from the *Hin*dIII site (12). In this analysis, 435- and 800-nt fragments were protected after hybridization with RNA from cells cotransfected with both plasmids and subsequent S1 nuclease treatment. No protected fragments were observed with untransfected cell RNA or with RNA from cells transfected with *Bgl*II-D alone.

The same 1.7-kbp *Hin*dIII-*Xho*I probe was specifically 5′ end labeled at the *Xho*I site to detect the two early RNAs transcribed from the opposite strand (13). Only one of the S1-resistant fragments which were protected with 6-h RNA was observed with RNA from cells cotransfected with IE-1 and p*Bgl*II-D. No protected bands were detected with RNA purified from cells transfected with p*Bgl*II-D alone or with untransfected cell RNA.

DISCUSSION

The AcNPV α regulatory gene IE-1 *trans*-activates expression of the β gene 39K (15). To learn more about this regulatory gene, we determined the nucleotide sequence of the gene, studied its temporal expression, and tested whether IE-1 *trans*-activated other viral genes.

Transcription of IE-1 reached its maximal level by the end of a 1-h adsorption period (time zero), and essentially constant levels of IE-1 were maintained for 24 h. Because S1 nuclease analysis measures steady-state levels, it was not clear from this experiment whether IE-1 continues to be synthesized throughout infection or whether the message made during the adsorption period persists for 24 h. In other regions of the genome, early transcripts appear to peak at 2, 6, or 12 h postinfection and then dramatically decline (12, 30). It will be of interest to determine whether the IE-1 message continues to be synthesized during the late phase and whether the IE-1 gene product is synthesized throughout infection. The presence of "antisense" message late in infection would be expected to arrest translation.

Overlapping transcripts have been described for several regions of the AcNPV genome (12, 24, 30). This report, which correlates S1 mapping and sequence data, indicates a remarkable conservation of nucleotide sequence in this region of the AcNPV genome. There are only 65 nt separating IE-1 and the open reading frame opposite and to the left of IE-1 and only 50 nt separating IE-1 and the open reading frame opposite and to the right of IE-1. This close apposition of coding sequences necessitates overlapping of both 5′ and 3′ ends of transcripts and, presumably, overlapping of transcriptional regulatory signals on one strand with coding sequences on the opposite strand.

IE-1 *trans*-activates at least four AcNPV genes in addition to 39K in transient assay experiments, suggesting that IE-1 plays a central role in the temporal regulation of the AcNPV gene expression. The fact that IE-1 did not *trans*-activate one of the early tandem promoters of *Hin*dIII-K is an intriguing result. This suggests tht a clue to the mechanism of IE-1 *trans*-activation may be found in a comparison of the two adjacent *Hin*dIII-K promoters.

The results reported here indicate that three of the four early transcripts which map to the *Hin*dIII-K fragment should be classified as β genes based upon their requirement for IE-1 in transient assay experiments. Two of the *Hin*dIII-K transcripts which are activated by IE-1 have previously been categorized as α genes based upon their (2 h) time of maximum synthesis (12). It is noteworthy that IE-1 is maximally synthesized well before this time. The other two *Hin*dIII-K transcripts activated by IE-1 were originally classified as β transcripts based upon their time of maximal synthesis (6 h). However, this classification was subsequently changed to α based upon the fact that these transcripts were synthesized in the presence of cycloheximide (13). It has been well established that cycloheximide-blocking experiments in adenovirus-infected cells can produce artifactual results (22, 29). Early genes which are *trans*-activated by the adenovirus E1a gene product are actually synthesized at higher levels in the presence of cycloheximide than in its absence. Presumably this is due to the fact that leakage through the cycloheximide block can produce sufficient levels of E1a protein to *trans*-activate the other early genes. We have previously shown that even low levels of IE-1 can efficiently *trans*-activate the 39K promoter in the presence of the AcNPV enhancer elements.

trans-Activating regulatory proteins have been identified in many viruses. These include the simian virus 40 large T antigen (20), the adenovirus E1a protein (4, 18), the immediate-early herpesvirus proteins (8), the E2 gene product of bovine papillomavirus (36), the nuclear antigen 1 of Epstein-Barr virus (30), and the Tat proteins of human T-cell leukemia virus types 1 and 2 (34). Although the mechanism of action of these regulatory proteins is unknown, they can be arranged into two groups based on their target sequences. Some of these proteins, namely, Epstein-Barr virus nuclear antigen 1, the E2 protein of bovine papillomavirus, the Tat proteins, and IE-1 of AcNPV, *trans*-activate transcriptional enhancers, whereas the other proteins apparently do not regulate enhancers. Epstein-Barr virus nuclear antigen 1 and the E2 protein have recently been shown to bind to the enhancers that they activate (1, 32), whereas simian virus 40 T antigen and herpes simplex virus ICP4 bind to the viral promoters (3, 5). It will be of interest to determine whether IE-1 binds to the AcNPV enhancer or the 39K promoter or both, or whether IE-1 interacts with a host cell factor which in turn activates viral gene expression. The nucleotide sequence of IE-1 presented here will be valuable in designing these experiments.

ACKNOWLEDGMENTS

We thank Melinda Worley Smith and David Baker for expert technical assistance.

This investigation was supported by grant DMB-8510270 from the National Science Foundation and by Texas Agricultural Experiment Station Project 6316.

LITERATURE CITED

1. **Androphy, E. J., D. R. Lowy, and J. T. Schiller.** 1987. Bovine papillomavirus E2 *trans*-activating gene product binds to specific sites in papillomavirus DNA. Nature (London) **325**:70–73.
2. **Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deiniger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchell, C. Seguin, P. S. Tuffnell, and B. G. Barrell.** 1984. DNA sequence and expression of the B95-8 Epstein Barr virus genome. Nature (London) **310**:207–211.
3. **Beard, P., S. Faber, K. W. Wilcox, and L. I. Pizer.** 1986. Herpes simplex virus immediate early infected-cell polypeptide 4 binds to DNA and promotes transcription. Proc. Natl. Acad. Sci. USA **83**:4016–4020.
4. **Berk, A. J., F. Lee, T. Harrison, J. Williams, and P. Sharp.** 1979. Pre-early adenovirus 5 gene product regulates synthesis of early viral RNA's. Cell **17**:935–944.
5. **Brady, J., J. Bolen, M. Radonovich, N. Salzman, and G. Koury.** 1985. Stimulation of SV40 late gene expression by SV40 tumor antigen. Proc. Natl. Acad. Sci. USA **81**:2040–2044.
6. **Chen, E. Y., P. M. Howley, A. D. Levinson, and P. H. Seeburg.** 1982. The primary structure and genomic organization of the

bovine papillomavirus type 1 genome. Nature (London) **299**:529–534.
7. **Chirgwin, J. M., A. E. Przybla, R. J. McDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid for sources enriched in ribonuclease. Biochemistry **18**:5294–5297.
8. **Dixon, R. A. F., and P. A. Schaffer.** 1980. Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. J. Virol. **36**:189–203.
9. **Devereaux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–396.
10. **Erlandson, M. A., J. Gorden, and E. B. Carstens.** 1985. Size and map locations of early transcriptional products on the AcNPV genome. Virology **142**:12–23.
11. **Esche, H., H. Lubbert, B. Siegmann, and W. Doerfler.** 1982. The translation map of the *Autographa californica* nuclear polyhedrosis virus AcNPV genome. EMBO J. **1**:1629–1633.
12. **Friesen, P. D., and L. K. Miller.** 1985. Temporal regulation of baculovirus RNA: overlapping early and late transcripts. J. Virol. **54**:392–400.
13. **Friesen, P. D., and L. K. Miller.** 1986. The regulation of baculovirus gene expression. Curr. Top. Microbiol. Immunol. **131**:31–49.
14. **Gordon, J. D., and E. B. Carstens.** 1984. Phenotypic characterization and physical mapping of a termperature-sensitive mutant of *Autographa californica* nuclear polyhedrosis virus defective in DNA synthesis. Virology **138**:69–81.
15. **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.
16. **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.
17. **Guarino, L. A., M. A. Gonzalez, and M. D. Summers.** 1986. The sequence and enhancer function of the homologous DNA regions of *Autographa californica* nuclear polyhedrosis virus. J. Virol. **60**:224–229.
18. **Jones, N., and T. Shenk.** 1979. An adenovirus 5 early product function which regulates expression of other genes. Proc. Natl. Acad. Sci. USA **76**:3665–3669.
19. **Kelly, D. C., and T. Lescott.** 1981. Baculovirus replication: protein synthesis in *Spodoptera frugiperda* cells infected with *Trichoplusia ni* nuclear polyhedrosis virus. Microbiologica **4**:35–47.
20. **Keller, J. M., and J. C. Alwine.** 1984. Activation of the SV40 late promoter: direct effects of T-antigen in the absence of DNA replication. Cell **36**:381–389.
21. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
22. **Lewis, J. B., and M. B. Matthews.** 1980. Control of adenovirus early gene expression: a class of immediate early gene products. Cell **21**:303–313.
23. **Lubbert, H., and W. Doerfler.** 1984. Mapping of early and late transcripts encoded by the *Autographa californica* nuclear polyhedrosis virus: is viral RNA spliced? J. Virol. **50**:497–506.
24. **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.
25. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **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.
27. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
28. **Miller, L. K., R. E. Trimarchi, D. Browne, and G. D. Pennock.** 1983. A temperature-sensitive mutant of the baculovirus *Autographa californica* nuclear polyhedrosis virus defective in an early function required for further gene expression. Virology **126**:376–380.
29. **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.
30. **Rawlins, D. R., G. Milma, S. D. Hayward, and G. S. Hayward.** 1985. Sequence-specific DNA binding of Epstein-Barr virus nuclear antigen (EBNA-1) to clustered sites in the plasmid maintenance region. Cell **42**:859–868.
31. **Reddy, V. B., B. Thimappaya, R. Dhar, K. N. Subramanian, B. S. Zain, J. Pan, P. K. Ghosh, M. L. Celma, and S. M. Weissman.** 1978. The genome of simian virus 40. Science **200**:494–502.
32. **Reisman, D., and B. Sugden.** 1986. *trans*-Activation of an Epstein-Barr viral transcriptional enhancer by the Epstein-Barr viral nuclear antigen 1. Mol. Cell. Biol. **6**:3838–3846.
33. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5468.
34. **Sodroski, J. G., C. A. Rosen, and W. A. Haseltine.** 1984. *trans*-Acting transcriptional activator of the long terminal repeat of human T lymphotrophic viruses in infected cells. Science **225**:381–385.
35. **Southern, E.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–513.
36. **Spalholz, B. A., Y. C. Yang, and P. M. Howley.** 1985. *trans*-Activation of a bovine papilloma virus transcriptional regulatory element by the E2 gene product. Cell **42**:183–191.
37. **Van Ormont, H., J. Maat, A. De Waard, and A. J. Van der Eb.** 1978. The nucleotide sequence of the transforming HpaI-E fragment of adenovirus type 5 DNA. Gene **4**:309–328.
38. **Vlak, J. M., G. E. Smith, and M. D. Summers.** 1981. Hybridization and in vitro translation of *Autographa californica* nuclear polyhedrosis virus mRNA. J. Virol. **40**:762–771.
39. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of an RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S β-globin mRNA precursor and mature β-globin mRNA have identical map coordinates. Nucleic Acids Res. **7**:1175–1193.
40. **Whitton, J. L., and J. B. Clements.** 1984. Replication origins and a sequence involved in coordinate induction of the immediate-early gene family are conserved in an intergenic region of herpes simplex virus. Nucleic Acids Res. **12**:1730–1739.