Overlapping TATA-Dependent and TATA-Independent Early Promoter Activities in the Baculovirus *gp64 envelope fusion protein* Gene

PHILIP H. KOGAN, XIAOYING CHEN, AND GARY W. BLISSARD*

Boyce Thompson Institute, Cornell University, Ithaca, New York 14853-1801

Received 6 September 1994/Accepted 8 November 1994

In previous studies to characterize basal and activated transcription from the early promoter of the *gp64 envelope fusion protein (efp)* gene of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus, the TATA box was identified as a functional element, essential for basal transcription from a minimal promoter construct. In the current study, we used discrete deletions and multiple point mutations that removed the functional TATA box from larger promoter constructs of the *gp64 efp* gene to reveal an overlapping TATA-independent transcriptional activity. TATA-independent transcriptional activity was inhibited *in vitro* by α -amanitin but not by tagetitoxin, demonstrating that like the TATA-dependent activity, the TATA-independent activity is mediated by RNA polymerase II. Using constructs in which the TATA box (TATATAA) was destroyed by substitution mutations, we identified four elements that are required for the TATA-independent activity. Two of the required elements, GATA (at –114) and CACGTG (at –104), were previously shown to specifically bind host transcription factors and activate transcription from the TATA-dependent wild-type *gp64 efp* promoter. The role of the early start site consensus CAGT sequence in TATA-independent transcription was also examined. Single-nucleotide substitution mutations in the CAGT sequence indicated that certain nucleotides within the CAGT start site were essential. In addition to the start site sequence and two upstream elements, a fourth essential element was identified in the 5' untranslated leader region (5'UTR). While the 5'UTR was not necessary for TATA-dependent transcription, deletion of a 10-bp 5'UTR sequence resulted in the loss of TATA-independent transcriptional activity. Although necessary, neither the GATA, CACGTG, start site region, nor 5'UTR element was alone sufficient for accurately initiated TATA-independent transcription from the consensus CAGT start site. Thus, the *gp64 efp* early promoter contains overlapping TATA-dependent and TATA-independent transcriptional activities. A number of common sequence elements (GATA, CACGTG, and start site CAGT) are involved in both of these activities, while one element (in the 5'UTR) is required only in the TATA-independent context.

In nuclear DNA viruses, transcription is often regulated by a combination of host and viral transcription factors. In the case of early genes that are transcribed by host RNA polymerases, the recognition of viral sequence elements and binding by host factors may be critical first steps in initiating infection. To examine these early recognition events in baculovirus infections, we have used the early promoter of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus (OpMNPV) *gp64 envelope fusion protein (efp)* gene to investigate host recognition and binding required for transcription by RNA polymerase II.

Baculoviruses are large, double-stranded DNA viruses that infect insect cells. After viral uncoating in the nucleus, early genes are transcribed by host RNA polymerase II. Evidence that RNA polymerase II and other host transcription factors are required for early gene expression includes α -amanitin sensitivity of early gene expression, the presence of common RNA polymerase II elements in many baculovirus early promoters, and accurate transcription from early gene promoters transfected into uninfected cells (3, 7, 14, 17, 18, 24, 27, 30, 40, 41). Several baculovirus early promoters have been examined in detail; these include the promoters of *39K*, *p35*, *IE-1*, *IE-2*

(also known as *IE-N*), *p143*, and *pe38* of *Autographa californica* multicapsid nuclear polyhedrosis virus (AcMNPV) as well as *IE-1*, *IE-2*, and *gp64 efp* of OpMNPV (5, 7, 8, 12, 14–16, 24, 26, 28, 40, 41). The OpMNPV *gp64 efp* gene is regulated by both early and late promoters. Studies of the OpMNPV *gp64 efp* early promoter have examined both minimal sequence requirements for accurate transcription initiation (basal promoter studies) as well as sequence elements required for activated transcription. A TATA-dependent basal promoter has been mapped to a 43-bp region of the *gp64 efp* promoter (3).

The 166-bp OpMNPV *gp64 efp* early promoter contains at least five to seven sequence regions involved in transcriptional regulation. In addition to the TATA box and start site CAGT sequence, two upstream elements that activate transcription above basal levels have been identified. We previously demonstrated that sequence-specific binding of host transcription factors occurs at both a GATA element and a CACGTG consensus site (19). These elements are located 40 and 30 bp upstream of the TATA box. Mutations in either element prevented specific host factor binding and reduced promoter activity in transient expression assays. Studies of elements required for basal transcription have focused on a TATA-dependent basal promoter defined as a 43-bp region that includes the TATA box and the conserved start site CAGT sequence (3). By using deletions, linker-scanning mutations, and single-nucleotide substitutions, the functional elements of the TATA-dependent basal promoter have been defined. In TATA-dependent basal promoter constructs (which contain

* Corresponding author. Mailing address: Boyce Thompson Institute, Cornell University, Tower Road, Ithaca, NY 14853-1801. Phone: (607) 254-1366. Fax: (607) 254-1242. Electronic mail address: gwb1@cornell.edu.

only the TATA box through the conserved start site CAGT sequences), the TATA box is required to position accurate transcription initiation and is an essential element. The conserved start site CAGT sequence is not essential but contributes to the level of basal transcription.

From these previous studies, we predicted that deletion of the TATA box from wild-type (wt) promoter constructs (containing the 166-bp upstream region) might inactivate this promoter. However, in the present study, we found that when the TATA box was removed (by deletion or substitution mutations) from promoter constructs containing 100 to 166 bp of upstream sequence, an accurately initiated TATA-independent transcriptional activity was detected. The characterization of TATA-independent promoters is of considerable interest since a number of baculovirus early genes do not appear to require a recognizable TATA box for transcription. In addition, the TATA-independent transcription observed when the *gp64* *efp* TATA box is destroyed appears to represent an example of basal promoter redundancy as has been observed in other baculovirus early genes (12, 14). In the current study, we examined the sequence elements necessary for the observed TATA-independent transcriptional activity of the OpMNPV *gp64* *efp* early promoter. To identify elements required for TATA-independent transcriptional activity, we first inactivated the TATA box with seven substitution mutations and used this TATA-less construct to generate additional deletion or substitution mutations. Using promoter constructs containing these mutations for transient expression and primer extension analysis, we identified three regions containing at least four sequence elements required for TATA-independent activity.

MATERIALS AND METHODS

Cells and TATA-less promoter constructs. Insect (*Spodoptera frugiperda*) Sf9 cells, a clonal derivative of the IPLB-Sf21 line, were cultured as previously described (3, 39). Plasmids used in this study were derived from a plasmid construct (p64CAT-166) containing 166 nucleotides (nt) of the upstream sequence (promoter and 5' untranslated region [5'UTR]) and the first seven codons of the OpMNPV *gp64* *efp* gene fused to the chloramphenicol acetyltransferase (CAT) open reading frame (ORF) by a 39-bp linker (7). Because the OpMNPV *gp64* *efp* gene contains multiple transcription start sites (one early and several late start sites), base positions are numbered relative to the common translation start site (+1). To examine the *gp64* *efp* early promoter in the absence of a TATA box, constructs containing deletions of the OpMNPV *gp64* *efp* TATA box were generated from plasmid p64CAT-166ΔTATA (kindly provided by Christian Gross). The p64CAT-166ΔTATA construct contains an *Xba*I site introduced into the TATA box by site-directed mutagenesis (GGGTA TATAA→GtctAgATAA). Constructs containing discrete 12-, 13-, and 21-bp deletions of the TATA box (-166Δ12TATA, -166Δ13TATA, and -166Δ21TATA) were generated from p64CAT-166ΔTATA by cleaving (within the mutated TATA box) with *Xba*I and digesting excessively with mung bean nuclease. To eliminate the possibility that the *gp64* *efp* ORF or bacterial linker sequences between the *gp64* *efp* ORF and CAT ORF fusion might cause uncertainty in the analysis of 3' deletions, we removed the *gp64* *efp* ORF and most sequence upstream of the CAT cassette to generate a construct in which the OpMNPV *gp64* *efp* 5'UTR was linked to the CAT ORF by only 6 nt of bacterial sequence. The sequence at the junction of the *gp64* *efp* promoter and the CAT ORF is 5'-CTACAAGgctaaatggagaaaaa-3' (baculovirus sequence is in uppercase and bacterial sequence is in lowercase; the CAT ORF is underlined). *gp64* *efp* protein-coding sequence and most of the bacterial linker sequence present in the p64CAT-166 construct were removed by a modification of the long primer-unique site elimination mutagenesis technique (36). The resulting plasmid was designated p64-166+TATA-1CAT to indicate that it contains the *gp64* *efp* promoter, a wt TATA box, and a CAT cassette linked to the baculovirus -1 position. For brevity, the p64-166+TATA-1CAT plasmid is referred to as -166+TATA. Sf9 cells transfected with either -166+TATA or the parent wt construct p64CAT-166 produced similar amounts of *gp64* *efp* RNA as determined by primer extension analysis. To destroy the TATA box and maintain wt spacing, we used the long primer-unique site elimination mutagenesis technique (36) to generate seven substitution mutations in the TATA box of construct -166+TATA. The 7 nt of the TATA box (TATATAA) were changed to GTC GACT (which generated a unique *Sal*I site in this plasmid), and the resulting construct was designated -166Δ7TATA.

5' deletions and point mutation constructs. To identify 5' sequences required

for TATA-independent transcriptional activity, we used the -166Δ7TATA construct to generate a series of 5' deletion constructs in the context of a TATA-minus promoter. The -166Δ7TATA plasmid was cleaved with restriction enzymes that cut at either -135 (*Bst*EII), -120 (*Rsa*I), -104 (*Pml*II), or -92 (*Mlu*I), made blunt with Klenow enzyme, and ligated with T4 DNA ligase. The resulting plasmids were designated -135Δ7TATA, -120Δ7TATA, -104Δ7TATA, and -92Δ7TATA. To examine the possible roles of two host transcription factor binding sites described previously (GATA and CACGTG), the same point mutations as used previously (19) were engineered into the -120Δ7TATA construct by standard cloning methods. To investigate the role of sequences upstream of the TATA box, we generated a construct containing only the -166 through -89 sequence of the *gp64* *efp* promoter linked to the CAT cassette. In this construct (pCAT-166/-89), 78 bp of baculovirus sequence were cloned upstream of the CAT ORF. The baculovirus sequence was excised from the p64CAT-166 construct with *Xho*I and *Mlu*I, made blunt with Klenow enzyme, ligated to *Hind*III linkers, and ligated into the *Hind*III site of the pCAT plasmid (29). In this construct (pCAT-166/-89), 59 bp of vector sequence separate the baculovirus sequence and the CAT ORF.

5'UTR and CAGT constructs. To study the importance of the 5'UTR region in TATA-independent transcription, we generated a 5'UTR deletion series by using the long primer-unique site elimination mutagenesis protocol (described above) to loop out sequences for deletion. Four oligonucleotide primers (-10 [5'-TTTCTCCATTTTAGCCTTGAGTGTTCTGTG-3'], -20 [5'-TTTCTCATTCTAGCTCCTGTGAAGAGCAC-3'], -30 [5'-TTTCTCCATTTTAGCA GCACAGTGAAC-3'], and -40 [5'-TTTCTCCATTTTAGCAACTGATT TACCAGAAC-3']) were synthesized to construct deletions of 9, 19, 29, and 39 bp in the 5'UTR. Each oligonucleotide contained the same 15-nt 5' sequence (5'-TTTCTCCATTTTAGC-3') located at the 3' end of the deletion, followed by 15 nt of sequence upstream of the deletion to ensure homologous binding during loop-out of the deleted regions. Deletions in the 5'UTR were generated in the context of the -166Δ7TATA promoter, and constructs were designated -166Δ7TATAΔ-10UTR, -166Δ7TATAΔ-20UTR, -166Δ7TATAΔ-30UTR, and -166Δ7TATAΔ-40UTR, where -10, -20, -30, and -40 indicate the last remaining 3' nt in the partially deleted UTR (see Fig. 6B). The largest deletion was also made in the context of a promoter that contains a wt TATA box (-166+TATA), and the resulting construct was designated -166+TATAΔ-40UTR. Mutations in the consensus CAGT start site were generated by using three oligonucleotides (5'-CACAGTGAACCTaATTACCAGA-3', 5'-CACAGTGAAGCTGATTTA CCAGA-3', and 5'-CACAGTGAACCaATTACCAGAA-3') for long primer-unique site elimination mutagenesis. All three of these mutations were made in the -166Δ7TATA context. The TATA-less constructs containing mutations at the CAGT start site were designated Δ7TATAΔtAGT, Δ7TATAΔCAGc, and Δ7TATAΔtgGT. Mutations and promoter regions of all constructs described above were confirmed by DNA sequencing.

Transfections, RNA preparation, and primer extension analysis. Transfections were performed in six-well plates as described previously (7), with minor modifications. Twenty micrograms of CsCl-purified supercoiled plasmid DNAs was transfected into 2.6×10^6 Sf9 cells per well. Forty-eight hours after transfection, total cellular RNA was isolated by a modification of the single-step acid guanidinium thiocyanate-phenol-chloroform method (10). After medium was removed from plates, cells were immediately lysed in 0.5 ml of solution D (4 M guanidine thiocyanate, 25 mM sodium citrate [pH 7], 0.5% sarcosyl, 0.1 M 2-mercaptoethanol) per well, and then 1/10 volume of 2 M sodium acetate (pH 4), was added. For some experiments, the lysed cells were stored at -80°C before extraction. Lysates were slowly pulled 10 times through 18-gauge needles to shear DNA, and then the pooled lysates (from six wells) were transferred to 15-ml polypropylene tubes containing 3 ml of phenol plus 0.6 ml of chloroform-isoamyl alcohol (24:1) on ice. After extraction, the RNAs were precipitated with ethanol and quantified by UV absorption. Isolated RNAs were annealed to a primer (CAT+40, 5'-TTCTTTACGATGCCATTGGG-3') that annealed 40 nt downstream of the CAT translation start site, and primer extension reactions were performed as described previously (3, 4) with avian myeloblastosis virus reverse transcriptase (Promega). In the primer extension analyses presented in Fig. 3, 5, and 6, 40 μg of total RNA was used per reaction. Poly(A) RNA was isolated from total RNA by oligo(dT)-cellulose column chromatography (1). The selected poly(A) RNA was used for the experiments presented in Fig. 1A [2.5 μg of poly(A) RNA per reaction] and Fig. 4 [25 μg of poly(A) RNA per reaction].

In vitro transcription. In vitro transcription reactions were performed as described previously (3, 17, 19) except that some reaction mixtures included α-amanitin to inhibit RNA polymerase II activity or tagetitoxin to inhibit RNA polymerase III activity. Briefly, 30-μl reaction mixtures containing 0.3 μg of linearized plasmid and 44 μg of Sf9 nuclear extract were incubated for 40 min in the presence or absence of inhibitor. To discriminate between RNA polymerase II and RNA polymerase III transcription, the inhibitors α-amanitin and tagetitoxin were added to the reaction mixtures. α-amanitin was added to a final concentration of 1 μg/ml, and tagetitoxin (Tagetin; Epicentre Technologies) was added to a final concentration of 10 U per reaction. RNAs were isolated and used in primer extension reactions as described above.

FIG. 1. Analysis of TATA-independent transcription from the *gp64 efp* early promoter. (A) Primer extension analysis of *gp64 efp* early promoter constructs containing 166 nt of upstream sequence with mutations that delete or replace the TATA box. The diagram on the left shows the locations of the TATA box and the conserved CAGT start site sequence relative to the ORF in the p64CAT-166 promoter construct (wt TATA). The wt sequence of the TATA box region and sequences from two constructs that lack the TATA box are shown below the diagram. The -166Δ-12TATA construct contains a deletion of the 7-nt TATA box and 5 upstream nt. In the -166Δ7TATA construct, seven single-nucleotide substitutions replace the TATA box with the underlined sequence gtcgact. Shown on the right are the results of primer extension analysis of construct p64CAT-166, -166Δ-12TATA, or -166Δ7TATA, containing either the wt TATA box (wt TATA), the TATA box deletion (Δ-12TATA), or the TATA box substitution (Δ7TATA), respectively. The location of the primer extension product corresponding to initiation at the A of the CAGT consensus is indicated on the left (arrowhead), and two lanes of markers from a DNA sequencing reaction are shown on the right. For primer extension reactions, RNA was isolated from Sf9 cells transfected with constructs containing either wt or mutant TATA box regions. RNA from mock-transfected cells was included as a control (mock) in primer extension experiments. (B) Effects of specific RNA polymerase inhibitors on TATA-independent transcription. The TATA-less promoter construct -166Δ7TATA was used for in vitro transcription analysis in the presence of either α-amanitin (1 μg/ml), tagetitoxin (10 U), or no inhibitor. RNAs from in vitro transcription reactions were analyzed by primer extension analysis and electrophoresed on a sequencing gel. Primer extension products corresponding to transcription from the A of the CAGT consensus are identified by an arrowhead on the right. The A lane from a sequencing reaction used as a marker is shown on the left.

RESULTS

TATA-independent transcription from the *gp64 efp* early promoter. In this study, we examined the role of the TATA box in the context of the larger activated promoter. To determine if the TATA box is required in this context, we generated promoter constructs containing 166 bp upstream of the translation start site and introduced discrete deletions or point mutations (substitutions) in the TATA box. For deletion mutations, we generated 12-, 13-, and 21-bp deletions that removed the 7-bp TATA box and a few base pairs of adjacent sequence. To remove the TATA box and retain wt spacing within the promoter region, we generated a construct containing seven substitutions that replaced the 7 bp of the TATA box (TATATAA) with the sequence GTCGACT. In constructs containing the entire 166-bp promoter region (Fig. 1A), removal of the TATA box sequence reduced but did not eliminate accurately initiated transcription from the early start site. Figure 1A shows the results of primer extension analysis of

RNA isolated from Sf9 cells transfected with plasmids containing the wt 166-bp promoter (wt TATA), a construct containing a 12-bp deletion of the TATA box region (-166Δ-12TATA), and a construct containing seven substitution mutations within the TATA box (-166Δ7TATA). In constructs that removed or replaced the TATA box (Fig. 1A, -166Δ-12TATA and -166Δ7TATA), accurately initiated transcriptional activity was retained, but at levels substantially lower than that of the wt promoter (Fig. 1A, lanes 1 to 3, arrowhead). Quantitation of the bands shown in Fig. 1A by PhosphorImager analysis indicated that the Δ-12TATA and Δ7TATA constructs produced only approximately 5 and 10%, respectively, of the RNA detected from the wt promoter containing an intact TATA box. Similar results were observed from additional deletion constructs containing either 13- or 21-bp deletions of the TATA box region (not shown). Thus, removal of the TATA box by deletion or substitution mutations demonstrates that a TATA-independent promoter activity is present, and transcription is accurately initiated from the wt transcription start site.

FIG. 2. Summary of 5' deletion analysis data from the current and a previous study (7) of upstream sequences required for accurately initiated transcription. Horizontal bars represent constructs containing various amounts of the *gp64 efp* early promoter upstream of the *gp64 efp* ORF. Negative numbers indicate the number of nucleotides remaining upstream of the *gp64 efp* ORF. Construct pCAT-166/-89 (bottom) contains *gp64 efp* upstream sequences from -166 to -89 linked to the CAT ORF by 59 nt of vector sequence. This construct was examined for discrete transcription initiation either within the baculovirus sequence or at the appropriate distance downstream of the cloned baculovirus sequence. On the right of each construct, a + or - indicates the ability of each construct to mediate accurate transcription initiation either in transient expression assays or by in vitro transcription. Accurate transcription initiation (at the A of the start site CAGT; arrow) was determined by primer extension analysis as described in Materials and Methods.

To determine if the TATA-independent promoter activity was mediated by RNA polymerase II or III, we examined the TATA-independent promoter activity in the presence of either α -amanitin, which inhibits RNA polymerase II (9), or tagetitoxin, which inhibits RNA polymerase III (38). In vitro transcription of a TATA-less promoter construct (-166 Δ 7TATA) in the presence of these inhibitors of RNA polymerases revealed that the TATA-independent promoter activity is mediated by RNA polymerase II (Fig. 1B). When linearized plasmid containing the -166 Δ 7TATA promoter was incubated without inhibitor in an Sf9 nuclear extract, accurate transcript initiation was observed (Fig. 1B, lane 1). Identical reactions were performed in the presence of either α -amanitin (lane 2) or tagetitoxin (lane 3). In this experiment, only α -amanitin was inhibitory, indicating that the observed TATA-independent transcription is RNA polymerase II dependent.

Upstream elements required for TATA-independent transcription. To identify elements required for the observed TATA-independent activity, we examined three regions: sequence upstream (5') of the substituted TATA box, the conserved start site CAGT, and sequence in the 5' untranslated leader region of the early mRNA. Previous studies of wt promoter constructs (containing a TATA box) showed that all transcriptional activity was lost when the TATA box region was sequentially deleted from the upstream region (7). These results were confirmed both in vitro (data not shown) and in vivo. The results of 5' deletions on accurate transcription initiation are summarized in Fig. 2. Although 5' deletions down to -77 bp (upstream of the translation start site) retained activity (Fig. 2, -166 through -77), deletion of the 15 bp containing the TATA box resulted in the loss of detectable transcription initiation (Fig. 2, -77 versus -62). These results confirmed our previous observations (7) and showed that sequences downstream of the TATA box were not sufficient for the TATA-independent *gp64 efp* promoter activity. This finding suggested that sequences upstream of the TATA box were essential.

To map the upstream sequences required for TATA-independent transcription, we used the -166 Δ 7TATA construct

(which contains seven substitution mutations in the TATA box) to generate a series of 5' deletion constructs in a TATA-minus context. These constructs contain the previously described seven nucleotide substitutions in the TATA box (Fig. 1A, -166 Δ 7TATA) and successive deletions of the upstream sequence (Fig. 3). As observed earlier, we detected an approximately 90% decrease in transcription when the TATA box was removed by substitution mutations (Fig. 3; compare lanes 1 and 2). A further substantial decrease in detectable RNA was observed when sequence between -135 and -120 was removed but accurate initiation was still detected (Fig. 3, lanes 3 and 4). However, when the region between -120 and -104 was deleted, all detectable activity was lost (Fig. 3, lanes 4 and 5, arrowhead). This result indicates that an essential component of the TATA-independent promoter is present in the region between -120 and -104. The -120 to -104 region contains the GATA and CACGTG elements that were previously shown to bind host transcription factors and activate transcription in the wt promoter. In addition, we previously demonstrated that binding of host transcription factors to these two elements is independent of the presence of a TATA box (19).

GATA and CACGTG elements. To determine if these GATA and CACGTG elements were involved in accurately initiated TATA-independent transcription, we generated a series of constructs containing substitution mutations in the GATA and CACGTG motifs, all within the context of a -120 promoter lacking the TATA box (Fig. 4, -120 Δ 7TATA). The point mutations used in this analysis were shown in a previous study to prevent binding of host transcription factors and to reduce transcription from the *gp64 efp* early promoter (19). Substitution mutations that prevented host factor binding to either the CACGTG site (atCGTG or CACGac) or the GATA site (ctcgAA) or both were examined for effects on TATA-independent transcription (Fig. 4). In primer extension experiments using 25 μ g of poly(A) mRNA, constructs containing mutations in either the GATA or the CACGTG element did not produce detectable levels of accurately initiated RNA transcripts (Fig. 4; compare lane 1 with lanes 2 to 5). These data

FIG. 3. Mapping upstream elements required for TATA-independent transcription. To map upstream elements required for TATA-independent transcription, a series of 5' deletions was constructed in the context of a TATA-less promoter construct (-166Δ7TATA). Deletion constructs are indicated by horizontal lines on the left. For each 5' deletion construct, the amount of sequence remaining upstream of the translation start site is indicated. Each deletion construct was used to transfect Sf9 cells. The results of primer extension analysis using RNAs isolated from transfected cells are shown on the right. A primer extension gel is shown, with promoter constructs indicated above the lanes. A plasmid containing a wt TATA box (-166+TATA) was included as a control (lane 1). Transcription initiation at the A of the start site CAGT is indicated by an arrowhead on the right. Marker lanes (G and A sequencing reactions from a defined template) are shown on the left. The results of primer extension analyses are summarized (+ or -) on the right of each construct diagram.

indicate that both GATA and CACGTG elements were necessary for TATA-independent transcription.

Because deletions and point mutations within the upstream sequence resulted in the loss of TATA-independent transcriptional activity, it was possible that the requirement for this activity lay solely within the region upstream of the TATA box (containing the GATA and CACGTG elements) were alone sufficient for the TATA-independent activity, we generated a construct containing 78 bp from the upstream region (-166 to -89) linked to a CAT cassette and examined the upstream sequence alone in transient expression assays (Fig. 2, pCAT-166/-89). Primer extension analysis (using a CAT-specific

primer) of transfected cell RNAs indicated that the -166 to -89 region was not sufficient for accurately initiated transcription at the appropriate location downstream of the cloned baculovirus sequence and revealed no discrete initiation from within the baculovirus sequence. Therefore, while the -120 to -89 upstream region is necessary for the observed TATA-independent transcriptional activity, these data indicate that this region is not alone sufficient for accurately initiated transcription.

Consensus CAGT start site. Because many baculovirus early promoters and insect RNA polymerase II genes contain a conserved start site CAGT sequence, and because it has previously been shown that this sequence plays a role in TATA-dependent transcription (3, 14), we also investigated the pos-

FIG. 4. Identification of required upstream elements by point mutations and transient expression analysis. The diagram on the left represents a TATA-less promoter construct (-120Δ7TATA) that contains sequences from nt -120 to -1 relative to the *gp64 efp* translation start site and seven point mutations that destroy the TATA box (see Materials and Methods). The locations of GATA and CACGTG binding sites are overlined on the wt upstream sequence below the diagram. Point mutations in GATA and CACGTG binding sites (-120 and -102 region) are indicated in lowercase and underlined. Results from primer extension of RNAs isolated from transfected Sf9 cells are shown on the right. Lanes 1 to 5 show primer extension analysis of constructs shown on the left. Lane 6 shows results from cells transfected with a negative control plasmid (pCAT). Marker lanes from sequencing reactions (G and C) are included on the left and right. The position of the band corresponding to the accurate transcription initiation site is indicated (arrowhead).

FIG. 5. Effects of point mutations in the conserved start site CAGT sequence on TATA-independent transcription. Point mutations were generated in a TATA-less promoter construct containing sequences from nt -166 to -1 relative to the *gp64 efp* translation start site. In the parental construct, seven point mutations (substitutions) were inserted to destroy the TATA box (-166 Δ 7TATA; see Materials and Methods). Single- and double-nucleotide substitutions in the start site CAGT sequence are indicated in lowercase and underlined. For primer extension analysis, RNAs were isolated from transfected Sf9 cells. Primer extension analysis of promoter constructs containing either a wt promoter (+TATAwtCAGT; lane 1), a TATA-less promoter (Δ 7TATAwtCAGT; lane 2), or TATA-less promoters with single (Δ 7TATA Δ tAGT and Δ 7TATA Δ CAGc; lanes 3 and 4) or double (Δ 7TATA Δ tgGT; lane 5) point mutations in the start site CAGT sequence are shown on the right. For markers, a sequencing reaction is included on the left (G, A, T, C). The band corresponding to the accurate transcription initiation site is indicated (arrowhead). As a negative control, a plasmid containing no promoter (pCAT) was also included in transfections (lane 6).

sible role of the CAGT start site sequence in the TATA-independent context. In another study, using saturating point mutations to study the basal TATA-dependent promoter, we found that the C and A at positions 1 and 2 of the start site CAGT were especially sensitive to mutagenesis, while mutations at other positions (nt 3 and 4) were better tolerated (unpublished data). Two strategies were used to disrupt the conserved start site CAGT in TATA-independent constructs. First, single-nucleotide substitutions were made at positions 1 and 4 of the CAGT, leaving the A at the start site intact. CAGT was changed to either tAGT or to CAGc (Fig. 5, tAGT and CAGc). Second, we generated a double-substitution mutation that destroyed the C and A at the start site. CAGT was changed to tgGT (Fig. 5, tgGT). Using primer extension analysis, we found that the tAGT mutation had little or no effect on accurate initiation (Fig. 5; compare lanes 2 and 3). However, the CAGc and tgGT mutations eliminated accurate transcription initiation (Fig. 5, lanes 4 and 5). Thus, while two mutations of the CAGT start site eliminated detectable transcripts (tgGT and CAGc), a third mutation (tgGT) did not have a significant effect. This finding suggests that while nucleotides within the conserved 4-nt CAGT start site sequence appear to play an essential role in TATA-independent transcription, a strict conservation of the 4-nt CAGT sequence is not required.

5'UTR element. Results of this study indicated that elements upstream of the TATA box were not independently capable of initiating TATA-independent activity. In addition, results of this and previous studies (3, 7) indicated that sequences at the conserved start site were not sufficient for accurate transcript initiation. Next we examined the role of sequences in the 5'UTR. To determine if sequences in the 5'UTR were required, we removed 39 bp from the 43-bp 5'UTR region of a TATA-less promoter construct (Fig. 6A, -166 Δ 7TATA Δ -40UTR). As a control, the same deletion was generated in the wt promoter containing a TATA box (Fig. 6A, -166+TATA Δ -40UTR). Although removing the 39 bp of 5'UTR sequence had no significant effect on transcription from the wt promoter, which contains an intact TATA box (Fig. 6A; compare lanes 1

and 3), all accurate transcription initiation was lost when 39 bp of untranslated leader region was removed from the TATA-less promoter (Fig. 6A; compare lanes 2 and 4). Thus, the 5'UTR contains an element required for the TATA-independent promoter activity. To map this element more precisely, we used a site-directed mutagenesis protocol to generate a series of sequential deletions in the 5'UTR region of a TATA-less promoter (Fig. 6B). The four deletion constructs shown in Fig. 6B (-166 Δ 7TATA Δ -10UTR, -166 Δ 7TATA Δ -20UTR, -166 Δ 7TATA Δ -30UTR, and -166 Δ 7TATA Δ -40UTR) are identical except that each has a decreasing amount of the 5'UTR from -1 to -40 (negative numbers preceding "UTR" indicate the last remaining 3' nt in the 5'UTR of each construct). To examine the effects of these deletions, Sf9 cells were transfected with each construct and RNAs were isolated and examined by primer extension (Fig. 6B), using a primer located 40 nt downstream of the CAT translation start codon. Because each successive 10-bp deletion results in an mRNA that is 10 nt shorter, the predicted primer extension products form a ladder in 10-nt increments (Fig. 6B). As shown in Fig. 6B the -166 Δ 7TATA, -166 Δ 7TATA Δ -10UTR, and -166 Δ 7TATA Δ -20UTR constructs all produced RNAs that were initiated accurately (Fig. 6B, lanes 2 to 4, small circles). However, neither the -166 Δ 7TATA Δ -30UTR nor the -166 Δ 7TATA Δ -40UTR construct produced accurately initiated transcripts (Fig. 6B, lanes 5 and 6, larger circles). Thus, accurate initiation was lost when 5'UTR sequences between -20 and -30 were removed. These data indicate that sequences essential for TATA-independent activity are located between -20 and -30 bp upstream of the *gp64 efp* translation start site. This corresponds to sequences between 14 and 24 bp downstream of the transcription start site (Fig. 6B).

DISCUSSION

In previous studies (3, 7), we demonstrated that the OpM-NPV *gp64 efp* early promoter contains a functional TATA box that plays an important role in early transcription. In the cur-

A

B

C

rent study, we found that while the TATA box is important for high levels of transcription, discretely removing the TATA box revealed an overlapping basal promoter activity that is TATA independent. To map elements required for this TATA-independent activity, we first generated a TATA-less promoter that was otherwise wt in sequence and maintained wt nucleotide spacing in the 166-nt promoter region. The TATA-less promoter was used to examine three regions by mutagenesis: (i) sequence upstream of the (deleted) TATA box, (ii) sequence at the conserved start site CAGT, and (iii) sequence in the 5'UTR region between the transcription and translation start sites. In these studies, we show that the TATA-independent activity required at least four elements within a 100-bp region (-120 to -20): GATA and CACGTG elements in the -120 to -104 region, sequences at the transcription start site, and an element in the 5'UTR.

GATA and CACGTG elements. Two of the elements required for the TATA-independent transcriptional activity, the GATA and CACGTG motifs, were previously identified as transcription factor binding sites that served as activation elements in the wt (TATA-containing) promoter (19). In the TATA-containing promoter, we showed that each of these upstream elements (GATA and CACGTG) specifically binds host transcription factors and that the TATA box is not required for binding to either element (19). Furthermore, we showed that each of these elements independently contributes to the functional activation of the wt TATA-dependent promoter. However, in the TATA-independent context, both elements appear to be required for transcriptional activity. Point mutations in either or both elements in -120 Δ 7TATA constructs (Fig. 4) eliminated all detectable TATA-independent transcription. It is possible that the two elements have differential contributions. However, because transcriptional activity from the -120 Δ 7TATA construct is very low, such differences may not be detectable by the techniques used in this study.

CAGT start site sequence. We also examined the role of the conserved start site CAGT sequence in TATA-independent promoter activity. Although one substitution mutation (CAGT changed to Δ AGT) had little or no effect, two mutations (CAGT changed to CAG Δ or Δ gGT) eliminated transcription, indicating that a portion of this conserved 4-nt sequence functions as part of a required element for the TATA-independent promoter activity. The observation that the Δ AGT mutation had little or no effect on TATA-independent transcription

indicates that a perfectly conserved 4-nt CAGT sequence at the start site is not required for the TATA-independent promoter activity. Thus, while essential sequences were identified within the start site CAGT, perfect conservation of the CAGT sequence was not required. This is perhaps not surprising since the conserved CAGT start site sequences are most commonly found in association with an upstream TATA box in insect and baculovirus genes transcribed by RNA polymerase II. Indeed, the promoters of three early genes that lack TATA boxes (DNA polymerase gene, CG30 gene, and *lef-1*) do not contain the CAGT start site consensus (31, 42, 45). In *Drosophila melanogaster*, it has been demonstrated that in both TATA-containing and TATA-less promoters, the transcription factor IID (TFIID) complex can recognize sequences at an initiator element (34, 35). While there is similarity between the conserved CAGT start site of baculovirus genes and this *Drosophila* initiator element, the CAGT baculovirus consensus is not independently active as an initiator in the *gp64* *efp* early promoter (3, 6). Interestingly, the start site mutation that retained transcriptional activity in this study (Δ AGT) is consistent with the consensus [(G/A/T)(T/C)A(G/T)TG] found when the *Drosophila hsp70* promoter was used for TFIID selection of preferred initiator binding sequences (34). The mutations that lacked transcriptional activity (Fig. 5, CAG Δ and Δ gGT) do not match that consensus. Thus, we might speculate that the essential sequences at the transcription start site are involved in TFIID binding or contact and that these sequences are especially important when the TATA box is absent.

5'UTR element. Using discrete deletions of 5'UTR sequences in a TATA-less promoter construct, we show that the 5'UTR contains an element that is required for TATA-independent transcription (Fig. 6A and B). Interestingly, sequences in the 5'UTR do not appear to be required for either activation of the *gp64* *efp* TATA-dependent promoter or stability of the early transcript, since removal of the 5'UTR had no effect on steady-state levels of mRNA from TATA-dependent transcription (Fig. 6A). Because the 5'UTR is not required for mRNA stability, yet removal of sequences between -30 and -20 in the 5'UTR resulted in loss of TATA-independent transcription (Fig. 6B), we can conclude that the 5'UTR contains an element essential for transcription in the TATA-minus context. When the 5'UTR from the OpMNPV *gp64* *efp* early transcript was compared with the predicted 5'UTRs from the *Choristoneura fumiferana* multicapsid nuclear polyhedrosis virus (CfM-

FIG. 6. Analysis of the role of *gp64* *efp* 5'UTR sequences in TATA-dependent and TATA-independent transcription initiation. (A) Constructs containing either the wt 5'UTR (-166+TATA and -166 Δ 7TATA) or a 39-nt deletion of the majority of the 5'UTR sequence (-166+TATA Δ -40UTR and -166 Δ 7TATA Δ -40UTR) were transfected into Sf9 cells, and RNAs were examined by primer extension analysis with a -40 CAT primer (see Materials and Methods). The panels on the right show the results of deleting the 5'UTR from a promoter containing an intact TATA box (lanes 1 and 3) and from a TATA-less promoter (lanes 2 and 4). Marker lanes (not shown) used to identify transcripts shown in lanes 1 to 4 consisted of sequencing ladders from characterized DNA templates. Locations of the primer extension products corresponding to the initiation site within the start site CAGT are indicated (arrowheads). (B) For fine mapping of 5'UTR sequences required for TATA-independent transcription, a series of deletions was generated in the 5'UTR of a TATA-less promoter construct. On the left, TATA-less promoter constructs containing either a wt 5'UTR (-166 Δ 7TATA) or successive deletions in the 5'UTR (-166 Δ 7TATA Δ -10UTR, -166 Δ 7TATA Δ -20UTR, -166 Δ 7TATA Δ -30UTR, and -166 Δ 7TATA Δ -40UTR) are represented by horizontal bars. Deleted sequences are indicated by thin lines. Nucleotides deleted in constructs identified as -10, -20, -30, and -40 consist of nucleotides -1 through -9, -1 through -19, -1 through -29, and -1 through -39, respectively. Results of primer extension analysis of RNAs isolated from transfected cells are shown in the panel on the right, with constructs indicated at the top of the gel. Primer extension products from a wt TATA-containing construct (-166+TATA) and mock-transfected cells (mock) are shown in lanes 1 and 7, respectively. Predicted primer extension products are separated by 10 nt as a result of the increasing size of deletions in the 5'UTR. Expected positions for products from constructs -166 Δ 7TATA Δ -30UTR and -166 Δ 7TATA Δ -40UTR are indicated by large closed circles on the gel. The sequence of the wt 5'UTR sequence is shown at the top of the diagram. Sequences conserved between the OpMNPV and AcMNPV *gp64* *efp* 5'UTR sequences are indicated by dots below the nucleotide sequence. A repeated sequence that is partially conserved between OpMNPV and AcMNPV is indicated by dashed lines with an arrowhead above the sequence. (C) Alignments of baculovirus *gp64* *efp* 5'UTR sequences with the 5'UTR of the *Drosophila hsp70* gene. The top three sequences show an alignment of the 5'UTR sequences from the early transcript of the OpMNPV *gp64* *efp* gene with the homologous sequences predicted from the *gp64* *efp* genes of AcMNPV and CfMNPV. The bottom sequence shows the 5'UTR sequence of the *Drosophila hsp70* gene. Arrows show the sites of transcription initiation and negative numbers show the locations of the OpMNPV *gp64* *efp* 5'UTR nucleotides (relative to the translation start site at +1). Sequence identity among all three baculovirus sequences is indicated by grey boxes, and a 14-nt region of conserved baculovirus sequence is indicated by a dashed line above the sequence. Identity between the baculovirus consensus and the *Drosophila hsp70* 5'UTR is indicated by vertical lines. The thick black bar above the sequence (-30 to -20) indicates the mapped location of baculovirus 5'UTR sequences required for TATA-independent transcription from the *gp64* *efp* gene. Circles below the sequence indicate TFIID contact points with the *Drosophila hsp70* 5'UTR, as determined by hydroxyl radical interference (35).

NPV) and AcMNPV *gp64 efp* genes, a moderate degree of sequence identity was observed (Fig. 6C). This sequence identity includes a 14-nt region (Fig. 6C, -10 to -23, dashed line) that is perfectly conserved at 12 of 14 nt in all three baculovirus genes. Although deletion of 10 of the 14 nt in this region (-10 to -20) did not affect TATA-independent transcription (Fig. 6B, -166Δ7TATA-20), deletion of an additional 10 nt (-10 to -30, which removed the remaining 4 nt of the conserved 14-nt region and 6 nt of additional upstream sequence) eliminated all detectable TATA-independent transcriptional activity (Fig. 6B, -166Δ7TATA-30). It is possible that sequences immediately upstream of the conserved 14-nt region play the more critical role in TATA-independent transcription. Recent work has shown that the TFIID transcription complex makes contact with sequences downstream of the initiation site in the *Drosophila hsp70* promoter and that transcription complex formation becomes more dependent on these downstream sequences when either the TATA box or the start site is removed from this promoter (35). In those studies, hydroxyl radical interference was used to identify important contacts in the +25 region (relative to transcription initiation) of the *hsp70* promoter (35). A comparison of the OpMNPV and predicted AcMNPV and CfMNPV *gp64 efp* 5'UTRs with the 5'UTR of the *hsp70* promoter is shown in Fig. 6C. It is noteworthy that there are 6 nt of sequence identity between the *hsp70* 5'UTR and the conserved 14-nt region described above for baculovirus *gp64 efp* gene 5'UTRs. More interesting, though, is the observation that the locations of 5'UTR sequences required for TATA-independent transcription in the baculovirus *gp64 efp* gene (Fig. 6C, -20 to -30, black bar) correlate well with the locations of nucleotides identified as TFIID contact points by hydroxyl radical interference studies of the *hsp70* gene (35) (Fig. 6C, circles). The locations of the baculovirus 5'UTR element and the *hsp70* physical contact points are similar in spacing from the transcription start site and from the 6 nt conserved in the 5'UTR region. Further studies will be necessary to determine if similar nucleotide positions in the *gp64 efp* promoter are involved in direct contacts with the TFIID transcription complex during TATA-independent transcription. In summary, the data from these studies of the OpMNPV *gp64 efp* 5'UTR show that the 5'UTR contains a transcriptional regulatory element that is not required for early transcription when a functional TATA box is present but becomes essential when the TATA box is inactivated or removed.

Summary. To explore the relationships between virus and host in the initiation of early baculovirus transcription, we have used the early promoter of the baculovirus *gp64 efp* gene as a model of early recognition and regulation. A central focus of these studies is the identification and mapping of virus-encoded sequence elements that mediate the virus-host interactions. We report here on the discovery of a TATA-independent transcriptional activity in the *gp64 efp* early promoter. This activity requires two elements that we previously identified (GATA and CACGTG), as well as sequences within the start site CAGT, and an element present in the 5'UTR. The identification and characterization of a TATA-independent early promoter activity from the *gp64 efp* gene is important in two contexts. First, there appear to be a number of TATA-less early promoters in the baculovirus genome, and very little is known about their transcriptional regulation. Second, this TATA-independent promoter activity is an example of redundancy in baculovirus regulatory elements. Redundancy in early and late baculovirus regulation appears to be a recurring theme, especially in the context of essential functions (such as replication) or expression of essential gene products (see below).

A number of baculovirus early genes do not have clearly identified or defined TATA boxes located appropriately (approximately 30 bp upstream) relative to the mapped transcription start sites. Examples include the DNA polymerase gene (45), *lef-1* (31), and the CG30 gene (42). Therefore, there is no absolute requirement for a TATA box for early baculovirus transcription and TATA-independent transcription is a general feature of a subclass of baculovirus early promoters. In the AcMNPV genome, many large ORFs do not have either consensus late promoter motifs or readily identifiable TATA boxes located upstream (2). Because these are unlikely to be late genes, we would speculate that many of these genes may be regulated by additional TATA-independent early promoters. Until transcription has been mapped and characterized for a great number of baculovirus early genes, it will be difficult to determine the relative contributions of TATA-independent transcription in baculovirus early gene expression.

Several examples of promoter redundancy in baculovirus gene regulation are notable. Like the OpMNPV *gp64 efp* early promoter, the *p35* gene of AcMNPV encodes a redundant TATA-independent transcriptional activity. Although the *p35* gene contains a TATA box, previous studies showed that deletion of this TATA box did not reduce transcriptional activity to zero (12). However, similar to the results reported here, substitution mutations in the *p35* gene TATA box reduced transcriptional activity to less than 10% of that observed from the wt construct (12). In contrast to the TATA-independent activity described for *gp64 efp*, sequences downstream of the *p35* TATA box appear to be sufficient to mediate TATA-independent transcriptional activity. Another example of early promoter redundancy is the *39K* (or *pp31*) gene of AcMNPV, which contains dual TATA boxes within 17 nt (14). Linker-scanning mutations through these two TATA boxes indicated that each is functional, directing transcription from different start sites. Removal of either of the TATA boxes does not eliminate transcriptional activity (14). It is not known whether a TATA-independent transcriptional activity may also be associated with this promoter. A number of baculovirus late genes also appear to encode redundant elements in their promoters. Examples include the multiple functional TAAG motifs observed in the late promoters from the *gp64 efp*, *gp37*, *gp41*, *lef-6*, *p39* (*capsid*), *sod*, and *v-ubi* genes (4, 5, 13, 32, 43, 44, 46–48).

The incorporation of multiple independent (redundant) transcriptional regulatory elements may be a common feature of important baculovirus genes. Indeed, where regulatory sequences have been studied or are easily recognizable, this phenomenon has been observed in both early and late genes. In addition to redundancy in transcriptional regulation, multiple and apparently independent origins of replication appear to be reiterated at intervals around the circular AcMNPV genome (11, 20–23, 33), even though all *hr* elements are not required (37) and the reiteration of such elements appears to be inconsistent with preliminary data suggesting a rolling-circle model for baculovirus genomic replication (25). It is tempting to speculate on possible biological roles for redundancy in baculovirus regulatory elements. Redundant regulatory elements might show differential activities in different host species and may thereby influence the host range of a virus. Alternatively, redundant regulatory elements may simply provide a measure of insurance against the prolonged mutagenic stresses that these viruses encounter in the environment. Thus, while the role of redundancy in baculovirus regulatory elements remains unclear, the observation that this phenomenon is apparently widespread suggests that it may play an important role in the life cycle of the virus.

ACKNOWLEDGMENTS

We thank Christian Gross for providing clone p64CAT-166ΔTAGA and Rosalind Wei for excellent technical assistance. We also thank David Garrity for valuable discussions and helpful comments on the manuscript.

This study was supported by grant AI31130 from the National Institutes of Health and by Boyce Thompson Institute project 1250-01.

REFERENCES

- Aviv, H., and P. Leder. 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci. USA* **69**:1408-1412.
- Ayres, M. D., S. C. Howard, J. Kuzio, M. Lopez-Ferber, and R. D. Possee. 1994. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* **202**:586-605.
- Blissard, G. W., P. H. Kogan, R. Wei, and G. F. Rohrmann. 1992. A synthetic early promoter from a baculovirus: roles of the TATA box and conserved start site CAGT sequence in basal levels of transcription. *Virology* **190**:783-793.
- Blissard, G. W., R. L. Quant-Russell, G. F. Rohrmann, and G. S. Beaudreau. 1989. Nucleotide sequence, transcriptional mapping, and temporal expression of the gene encoding P39, a major structural protein of the multicapsid nuclear polyhedrosis virus of *Orgyia pseudotsugata*. *Virology* **168**:354-362.
- Blissard, G. W., and G. F. Rohrmann. 1989. Location, sequence, transcriptional mapping, and temporal expression of the gp64 envelope glycoprotein gene of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* **170**:537-555.
- Blissard, G. W., and G. F. Rohrmann. 1990. Baculovirus diversity and molecular biology. *Annu. Rev. Entomol.* **35**:127-155.
- Blissard, G. W., and G. F. Rohrmann. 1991. Baculovirus gp64 gene expression: analysis of sequences modulating early transcription and transactivation by IE1. *J. Virol.* **65**:5820-5827.
- Carson, D. D., M. D. Summers, and L. A. Guarino. 1991. Molecular analysis of a baculovirus regulatory gene. *Virology* **182**:279-286.
- Chambon, P. 1974. Eucaryotic RNA polymerases. *Enzymes* **10**:261-331.
- Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**:156-159.
- Cochran, M. A., and P. Faulkner. 1983. Location of homologous DNA sequences interspersed at five regions in the baculovirus AcMNPV genome. *J. Virol.* **45**:961-970.
- 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.
- Guarino, L. A. 1990. Identification of a viral gene encoding a ubiquitin-like protein. *Proc. Natl. Acad. Sci. USA* **87**:409-413.
- Guarino, L. A., and M. Smith. 1992. Regulation of delayed-early gene transcription by dual TATA boxes. *J. Virol.* **66**:3733-3739.
- Guarino, L. A., and M. W. Smith. 1990. Nucleotide sequence and characterization of the 39k gene region of *Autographa californica* nuclear polyhedrosis virus. *Virology* **179**:1-8.
- Guarino, L. A., and M. D. Summers. 1987. Nucleotide sequence and temporal expression of a baculovirus regulatory gene. *J. Virol.* **61**:2091-2099.
- Hoopes, R. R., and G. F. Rohrmann. 1991. *In vitro* transcription of baculovirus immediate early genes: accurate mRNA initiation by nuclear extracts from both insect and human cells. *Proc. Natl. Acad. Sci. USA* **88**:4513-4517.
- Huh, N. E., and R. F. Weaver. 1990. Identifying the RNA polymerases that synthesize specific transcripts of the *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virol.* **71**:195-202.
- Kogan, P. H., and G. W. Blissard. 1994. A baculovirus gp64 early promoter is activated by host transcription factor binding to CACGTG and GATA elements. *J. Virol.* **68**:813-822.
- Kool, M., P. M. M. van den Berg, J. Tramper, R. W. Goldbach, and J. M. Vlak. 1993. Location of two putative origins of DNA replication of *Autographa californica* nuclear polyhedrosis virus. *Virology* **192**:94-101.
- Kool, M., and J. M. Vlak. 1993. The structural and functional organization of the *Autographa californica* nuclear polyhedrosis virus genome. *Arch. Virol.* **130**:1-16.
- Kool, M., J. T. M. Voeten, R. W. Goldbach, J. Tramper, and J. M. Vlak. 1993. Identification of seven putative origins of *Autographa californica* multiple nucleocapsid nuclear polyhedrosis virus DNA replication. *J. Gen. Virol.* **74**:2661-2668.
- Kool, M., J. T. M. Voeten, R. W. Goldbach, and J. M. Vlak. 1994. Functional mapping of regions of the *Autographa californica* nuclear polyhedrosis virus genome required for DNA replication. *Virology* **198**:680-689.
- Krappa, R., K. A. Behn, F. Jahnel, W. Doerfler, and M. D. Knebel. 1992. Differential factor binding at the promoter of early baculovirus gene *pe38* during viral infection: GATA motif is recognized by an insect protein. *J. Virol.* **66**:3494-3503.
- Leisy, D. J., and G. F. Rohrmann. 1993. Characterization of the replication of plasmids containing *hr* sequences in baculovirus-infected *Spodoptera frugiperda* cells. *Virology* **196**:722-730.
- Lu, A., and E. B. Carstens. 1991. Nucleotide sequence of a gene essential for viral DNA replication in the baculovirus *Autographa californica* nuclear polyhedrosis virus. *Virology* **181**:336-347.
- Lu, A., and E. B. Carstens. 1992. Transcription analysis of the *EcoRI* D region of the baculovirus *Autographa californica* nuclear polyhedrosis virus identifies an early 4-kilobase RNA encoding the essential *p143* gene. *J. Virol.* **66**:655-663.
- Lu, A., and E. B. Carstens. 1993. Immediate-early baculovirus genes transactivate the *p143* gene promoter of *Autographa californica* nuclear polyhedrosis virus. *Virology* **195**:710-718.
- Mackett, M., G. L. Smith, and B. Moss. 1984. General method for production and selection of infectious vaccinia virus recombinants expressing foreign genes. *J. Virol.* **49**:857-864.
- Nissen, M. S., and P. D. Friesen. 1989. Molecular analysis of the transcriptional regulatory region of an early baculovirus gene. *J. Virol.* **63**:493-503.
- Passarelli, A. L., and L. K. Miller. 1993. Identification and characterization of *lef-1*, a baculovirus gene involved in late and very late gene expression. *J. Virol.* **67**:3481-3488.
- Passarelli, A. L., and L. K. Miller. 1994. Identification and transcriptional regulation of the baculovirus *lef-6* gene. *J. Virol.* **68**:4458-4467.
- Pearson, M., R. Bjornson, G. Pearson, and G. Rohrmann. 1992. The *Autographa californica* baculovirus genome: evidence for multiple replication origins. *Science (Washington, D.C.)* **257**:1382-1384.
- Purnell, B. A., P. A. Emanuel, and D. S. Gilmour. 1994. TFIID sequence recognition of the initiator and sequences farther downstream in *Drosophila* class II genes. *Genes Dev.* **8**:830-842.
- Purnell, B. A., and D. S. Gilmour. 1993. Contribution of sequences downstream of the TATA element to a protein-DNA complex containing the TATA-binding protein. *Mol. Cell. Biol.* **13**:2593-2603.
- Ray, F. A., and J. A. Nickoloff. 1992. Site-specific mutagenesis of almost any plasmid using a PCR-based version of unique site elimination. *BioTechniques* **13**:342-346.
- Rodems, S. M., and P. D. Friesen. 1993. The *hr5* transcriptional enhancer stimulates early expression from the *Autographa californica* nuclear polyhedrosis virus genome but is not required for virus replication. *J. Virol.* **67**:5776-5785.
- Steinberg, T. H., D. E. Mathews, R. D. Durbin, and R. R. Burgess. 1990. Tagetitoxin: a new inhibitor of eukaryotic transcription by RNA polymerase III. *J. Biol. Chem.* **265**:499-505.
- Summers, M. D., and G. E. Smith. 1987. A manual of methods for baculovirus vectors and insect cell culture procedures. Texas Agricultural Experiment Station bulletin no. 1555.
- Theilmann, D. A., and S. Stewart. 1991. Identification and characterization of the IE-1 gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* **180**:492-508.
- Theilmann, D. A., and S. Stewart. 1992. Molecular analysis of the transactivating *ie-2* gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* **187**:84-96.
- 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.
- Thiem, S. M., and L. K. Miller. 1989. Identification, sequence, and transcriptional mapping of the major capsid protein gene of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **63**:2008-2018.
- Tomalski, M. D., R. Eldridge, and L. K. Miller. 1991. A baculovirus homolog of a copper zinc superoxide dismutase gene. *Virology* **184**:149-161.
- Tomalski, M. D., J. Wu, and L. K. Miller. 1988. The location, sequence, transcription, and regulation of a baculovirus DNA polymerase gene. *Virology* **167**:591-600.
- Whitford, M., and P. Faulkner. 1992. Nucleotide sequence and transcriptional analysis of a gene encoding gp41 a structural glycoprotein of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **66**:4763-4768.
- Whitford, M., S. Stewart, J. Kuzio, and P. Faulkner. 1989. Identification and sequence analysis of a gene encoding gp67, an abundant envelope glycoprotein of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **63**:1393-1399.
- Wu, J., and L. K. Miller. 1989. Sequence, transcription, and translation of a late gene of the *Autographa californica* nuclear polyhedrosis virus encoding a 34.8k polypeptide. *J. Gen. Virol.* **70**:2449-2460.