Promoter of the Adenovirus Polypeptide IX Gene: Similarity to E1B and Inactivation by Substitution of the Simian Virus 40 TATA Element

LEE E. BABISS¹ AND LYNNE D. VALES^{2*}

University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, New Jersey 08854-5635,² and The Rockefeller University, Box 279, 1230 York Avenue, New York, New York 10021¹

Received 28 August 1990/Accepted 26 October 1990

The promoter of the adenovirus polypeptide IX (pIX) gene consists of an SP1 binding site and a TATA box and is remarkably similar to the promoter of the E1B gene in which it is nested. Plasmid constructs containing the pIX gene with deletions in the SP1 or TATA sites were defective in pIX mRNA production in transient expression assays. These results were confirmed with analogous virus constructs. An oligonucleotide containing sequences within the pIX promoter region spanning the SP1 and TATA sites but not including the sequences downstream of the TATA box is sufficient to direct mRNA synthesis at +90 nucleotides within the pIX gene. While the simian virus 40 (SV40) early promoter is capable of directing pIX mRNA synthesis from the SV40 cap sites, substitution of the pIX TATA box with the SV40 TATA box results in barely detectable levels of pIX mRNA. These results will be discussed with respect to exchangeability of promoter elements and the possible role of the viral E1B 21-kDa protein in potentiating or stabilizing transcription factor TFIID binding to the pIX TATA element.

The adenovirus polypeptide IX (pIX) gene has been known to exhibit a strict regulatory pattern of expression during adenovirus infection. Only after adenovirus DNA replication occurs can pIX mRNA be observed (12). The position of the pIX gene, which is embedded within an upstream transcriptional unit, E1B (35), was found to be crucial to its regulation (29). The transcriptionally active E1B gene causes interference or occlusion of the pIX promoter at early times of infection. However, simple removal of the pIX gene from the E1B transcription unit was not sufficient to produce pIX mRNA. The E1B-encoded 21-kDa protein was necessary to obtain pIX mRNA production in the absence of DNA replication. DNA replication appeared to be required for pIX gene expression so that new templates, not engaged in transcription from the E1B gene, could support transcription of the downstream pIX gene. The finding that transcription across promoter regions can effectively silence downstream genes has been documented in several systems (1, 4, 11, 13). This phenomenon encompasses regulatory roles for transcription termination (10, 17, 22) or upstream promoter shutoff in order to activate downstream promoters that lie in the path of elongating transcription complexes (11).

Studies of adenovirus gene expression have provided tenable models for the description of cellular gene regulation and have focused on the identification and properties of cellular transcription factors that mediate promoter efficacy. The dissection of the promoter regions of several adenovirus genes has led to the identification of specific DNA elements that have been shown to or are presumed to interact with cellular transcription factors (for review, see reference 19). The basal levels of adenovirus gene expression, as well as the activation of several promoters mediated by the adenovirus E1A proteins, are being extensively analyzed by

several groups (for review, see reference 7). Interactions between the cellular transcription factors after binding to their cognate DNA sequence within the promoter region are believed to potentiate initiation of transcription by RNA polymerase II (for review, see reference 23). The early adenovirus transcription unit designated E1B has been shown to contain a relatively simple promoter region. Only the cellular transcription factors SP1 and the TATA box binding protein TFIID appear to be required to obtain optimal levels of E1B transcription (37). As well, extensive analyses of the E1B promoter indicated that the spacing between the two DNA elements that bind SP1 and TFIID proteins is crucial for E1B mRNA production (36). This finding suggested that the interaction between these two proteins may be constrained by the distance between their DNA binding sites and that this presumed interaction is important for initiation of E1B transcription.

In this report, we present an analysis of the pIX promoter region by transient expression assays and virus constructs to determine the DNA elements necessary for optimum pIX mRNA production. Our results indicate that the pIX promoter is remarkably similar to the E1B promoter, requiring consensus SP1 and TFIID binding sites. The promoter region of the pIX gene was found to be capable of directing sequences internal to the pIX gene. The distance between the consensus SP1 and TATA sites within the pIX gene is 20 nucleotides (nt), versus 11 nt within the E1B gene. Substitution of these extra 9 nt did not affect pIX mRNA production. The simian virus 40 (SV40) early promoter was found to support pIX mRNA synthesis with the SV40 early cap sites. However, while substitution of the E1B TATA box for the pIX TATA box led to comparable pIX mRNA levels, substitution of the SV40 early TATA box led to deficient production of pIX mRNA. These results suggest that the context of the SP1 and TATA sites within the pIX and SV40 promoters may be crucial to their respective promoter

* Corresponding author.

activities and that the pIX and E1B TATA boxes are functionally distinct from the SV40 TATA box.

MATERIALS AND METHODS

Cells, transfections, and infections. HeLa clone 2 cells and human 293 cells (constitutively express the adenovirus E1A and E1B genes) were obtained from the American Tissue Culture Collection and grown as monolayer cultures in Dulbecco's modified Eagle's medium supplemented with 10% defined calf serum (Hyclone). Transfection and infection experiments were performed as described previously (29).

Plasmid derivations. Plasmid constructs containing only the pIX gene with various amounts of sequence upstream of the pIX cap sites were generated by sequential *Bal31* digestions of pMK2004 DNA containing adenovirus type 5 (Ad5) nt 1 to 5780 (3) linearized at nt 3328 with *Bgl*II. The termini of the digested candidates were then made blunt ended with the Klenow fragment of DNA polymerase I and ligated to *Bam*HI linker DNA. The DNA fragments containing portions of the pIX gene were separated and isolated from low-melting-point agarose after digestion with *Bam*HI and *Xho*I (nt 5780). DNA fragments were then ligated to pMK2004 DNA containing *Bam*HI and *Xho*I termini to generate plasmid constructs containing only the pIX gene. DNA fragments were also ligated to the *Bgl*II and *Xho*I termini of linearized pMK2004 DNA carrying Ad5 nt 1 to 3328 (*Bgl*II) to generate plasmid constructs carrying Ad5 nt 1 to 5780 with deletions within the E1B gene that extended from nt 3328 to the 5' endpoint of the deleted fragment of the pIX gene. The DNA sequence of the region encompassing the pIX cap sites was determined in all cases. The 5' deletion endpoint of one plasmid construct containing only the pIX gene was found to be +90 nt (nt 3670) relative to the pIX cap sites. The DNA fragment contained within nt +90 (*Bam*HI) and nt 5640 (*Sst*I) of this construct was ligated to pGem-1 linearized with *Hind*III and *Sst*I in the presence of the oligonucleotide carrying the pIX promoter (−56 [*Hind*III] to −20 [*Bam*HI]). The final derivative was pGem-1 carrying the pIX promoter fused to +90 nt of the pIX gene. The 5' endpoint of another pMK2004 construct containing only the pIX gene was found to be nt −20 relative to the pIX cap sites. The DNA fragment contained within nt 3560 (*Bam*HI) and nt 5780 (*Xho*I) of this −20 construct was ligated to pGem-1 DNA linearized with *Hind*III and *Sal*I in the presence of an oligonucleotide containing the pIX promoter (−56 nt [*Hind*III] to −20 nt [*Bam*HI]) with either the pIX TATA, SV40 TATA, E1B TATA, sub9, or del9 mutations. The resulting derivatives were pGem-1 DNAs containing only the pIX gene from nt 3524 to 5780 with various mutations within the pIX promoter. pMK2004(2x21) was used to isolate the SV40 early promoter region (enhancer-minus) contained within *Bam*HI-linkered DNA (6). This fragment was ligated to the *Bam*HI-linearized −20 construct (pMK2004), and the orientation of the insert was determined. The product was pMK2004 DNA containing the SV40 early region fused to the pIX gene at nt 3560 (nt −20).

Virus constructions. To facilitate the localization of the pIX mutations (−56 and −38) on an intact adenovirus genome, the pMK2004 plasmids containing these mutated viral DNA sequences were linearized with *Xho*I at nt 5780. A DNA fragment extending from nt 5780 to 8250 was isolated from the Ad5 genome by *Xho*I digestion and, following low-melting-point agarose gel purification, ligated to the pMK plasmid sequences. The orientation of the viral se-

quences was confirmed by restriction enzyme analysis and resulted in contiguous viral DNA sequences on the plasmid genome extending from nt 1 (*Bam*HI) to 8250 (*Xho*I). Initial attempts to introduce these mutated viral sequences into the wild-type virus *dl*309 genome (20) by both in vivo overlap recombination (31) and in vivo ligation failed, most likely owing to the instability of these viral sequences in the viral capsid (data not shown) (15). To correct for the proper-sized recombinant viral genome, we used the LLX1 virus (8) as the parent, since its genome size would permit the pIX sequences to be localized and packaged efficiently. Like *dl*309, the LLX1 virus is phenotypically wild type (8). Plasmid DNA was linearized at nt 8250 with *Xho*I and cotransfected into human 293 cells with LLX1 virus DNA-protein complex (9) that was digested with *Cl*aI (cuts only at nt 918) and *Pae*R71 (cuts at nt 5780 and 8250). In vivo ligation at nt 8250 will yield a full-length viral genome and give rise to plaques on the transfected plates. The presence of the mutations in the resulting viruses was confirmed by restriction enzyme analysis of viral DNA. Mutant viruses were plaque purified once before further analysis, and virus stocks were prepared and counted on human 293 cells as described previously (3).

RNA preparation and analysis. Cells were harvested and lysed in 0.5% Nonidet P-40 (25), and oligo(dT)-cellulose selection of cytoplasmic RNA was done as described previously (29). The SP6 probe (Ad5 nt 3328 [*Bgl*II] to 3787 [*Pst*I]) for E1B and pIX mRNA analyses and the SP6 probe (Ad5 nt 1334 [*Xba*I] to 1774 [*Sst*I]) for E1A mRNA analyses have been described previously (29). All E1B mRNAs utilize the same acceptor splice site at position Ad5 nt 3595 (30), which lies 13 nt downstream of the pIX cap sites. pIX mRNA is not spliced (2). pIX mRNA is therefore expected to protect a fragment of SP6 RNA that is 13 nt longer than that protected by the second exon of the major E1B mRNAs. The first exon of the 22S E1B mRNA utilizes a donor splice site at Ad5 nt 3511 and is expected to protect a fragment of 183 nt of the same SP6 probe. Conditions of probe hybridization and analysis of RNase T₂-protected fragments were performed as described previously (29).

RESULTS

pIX gene expression from plasmids containing deletions upstream of the pIX cap sites. Our previous analyses have shown that expression of the pIX gene can be disassociated from DNA replication when two conditions exist: E1B transcription across the pIX promoter is absent and the 21-kDa E1B protein is present (29). These results were obtained by using plasmid and viral constructs containing the pIX gene with 250 nt upstream of the pIX cap sites. E1B-encoded protein was provided during cotransfection experiments with a plasmid construct, *dl*30E1A−, that contained the E1B gene with a relatively small deletion of 300 nt that removed the pIX promoter. These plasmids are shown in Fig. 1. To define the regions of the pIX gene that are required for optimal levels of pIX mRNA synthesis in the absence of overriding E1B transcription, several plasmid constructs containing only the pIX gene with various amounts of DNA upstream of the pIX cap sites were isolated. These constructs were analyzed for pIX mRNA synthesis in transient expression assays performed in HeLa cells (see Materials and Methods). Cytoplasmic, polyadenylated RNA was isolated, and pIX mRNA was scored with an antisense SP6 probe that spanned the pIX cap sites (see Materials and Methods). Plasmids containing various

FIG. 1. Schematic representation of the left-end portion of the Ad5 genome is shown at the top of the diagram. The arrows indicate the direction of transcription of the E1A, E1B, and pIX primary transcripts. The sites of transcription initiation are nt 498 for E1A; nt 1701 and 1703 for E1B; and nt 3579 and 3581 for pIX. All E1B mRNAs utilize the same splice acceptor site located 13 nt downstream of the pIX cap sites at nt 3595. pIX mRNA is not spliced. The two plasmid constructs shown are derivatives of pMK2004 DNA (diagonal lines) containing portions of the left end of the Ad5 genome. *dl30E1A⁻* contains Ad5 nt 1334 to 5780 with a deletion within the E1B gene extending across the pIX promoter from Ad5 nt 3328 to 3640. This construct contains only the E1B transcription unit. The -250 construct contains Ad5 nt 3328 to 5780, which encompasses the pIX gene with 250 nt upstream of the pIX cap sites. This construct contains only the pIX transcription unit.

amounts of DNA upstream of the pIX cap sites were not capable of producing detectable levels of pIX mRNA alone in HeLa cells (data not shown); however, all the constructs retaining at least 56 nt of upstream DNA sequences (-56) produced comparable amounts of pIX mRNA when cotransfected with *dl30E1A⁻* (Fig. 2). The -56 construct contained a consensus sequence for binding of the cellular transcription factor SP1 as well as a TATA motif (see Fig. 5). The -38 nt construct contained a deletion of this SP1 consensus sequence and resulted in reduced levels of pIX mRNA, while the -20 construct, devoid of the TATA motif, was

FIG. 2. Deletion analysis of the pIX gene. Denaturing polyacrylamide gel analysis of RNase T₂-resistant products obtained after hybridization of antisense SP6 probe with cytoplasmic polyadenylated RNA isolated from HeLa cells cotransfected with *dl30E1A⁻* and the plasmid constructs containing only the pIX gene indicated. The sequence remaining upstream of the pIX cap sites is shown. pIX mRNA is expected to protect 205 or 207 nt and E1B mRNA derived from *dl30E1A⁻* is expected to protect 147 nt (29). Low but detectable levels of a 227-nt protected fragment were observed in analyses of RNA from transfections with the -20 construct. The size of this protected fragment is consistent with unspliced transcripts that initiated within vector sequences and were transcribed across the pIX sequences remaining in this plasmid. These transcripts would be expected to protect nt 3560 to 3787 of the SP6 RNA probe. Lane M shows radiolabeled *Hpa*II-digested pBR322 DNA as markers.

FIG. 3. DNA elements required for pIX transcription. (A) Schematic representation of pMK2004 DNA containing E1A, E1B, and pIX genes with deletions within the E1B gene that extend from nt 3328 (nt -250 pIX caps) to nt 3525 (nt -56) or 3541 (nt -38). The DNA sequence downstream of the 3' deletion endpoint of these constructs is shown, and the SP1 and TATA sites are underlined. The directions of transcription of E1A, E1B, and pIX mRNAs are indicated with arrows. (B) Analysis of RNA isolated from HeLa cells infected for 16 h with virus derivatives of plasmid constructs shown in panel A. Aliquots of RNA, isolated from cells infected with the virus indicated at the top of the lanes, were probed for E1B and pIX mRNA and for E1A mRNA with SP6-E1A (see Materials and Methods). E1A mRNA is expected to protect a fragment of 296 nt. Lane M shows markers. wt, wild type.

incapable of supporting pIX mRNA synthesis (Fig. 2). These results indicated that sequences upstream of the SP1 consensus sequence were not important to pIX mRNA synthesis. We also noted, as we did previously (29), the presence of some transcription across the pIX promoter even in the absence of the E1B promoter, as indicated by the presence of a protected fragment of 195 nt that migrated below the protected fragments corresponding to pIX mRNA. The 195-nt fragment arises from transcription across the pIX promoter with subsequent usage of the E1B splice acceptor site that lies downstream of the pIX cap sites (Ad5 nt 3595). The promoter directing this transcription is believed to reside within the vector sequence and has been detected by others (14).

Several of the plasmid constructs containing deletions upstream of the pIX cap sites were analyzed for pIX mRNA synthesis when the E1A gene and most of the E1B gene were restored upstream of the pIX gene. The constructs contained the same 5' deletion endpoint at -250 nt relative to the pIX cap sites but differed in their 3' endpoints (for example, see constructs in Fig. 3A). As expected, in the presence of overriding E1B transcription, none of these plasmids gave rise to detectable pIX mRNA synthesis in transfected cells (data not shown). This result is in keeping with our previous findings that E1B transcription across the pIX promoter is inhibitory to pIX gene expression.

Adenovirus constructs containing deletions within E1B up to

FIG. 4. Regulatory regions of the pIX gene. (A) Diagram of pMK2004 DNA (diagonal lines) containing the SV40 promoter region without the SV40 enhancer (bold line) fused to sequences containing the pIX gene downstream of the pIX TATA box (Ad5 nt 3560 to 5780). The endpoints of adenovirus DNA are indicated. The expected direction of transcription initiating from the SV40 early promoter is shown with an arrow. pMK2004 DNA carrying the pIX promoter and pIX gene is also shown. The pIX promoter (nt -56 to -20, bold line) is fused to a portion of the pIX coding region between nt 3670 and 5780. The expected direction of transcription directed by the pIX promoter is indicated with an arrow. (B) Analysis of RNA isolated from HeLa cells transfected with the -56 construct alone (lane 1) or with *dl30E1A*⁻ (lane 2) and with SV40/pIX alone (lane 3) or with *dl30E1A*⁻ (lane 4). Lane M shows markers. (C) Similar analysis of RNA isolated from HeLa cells transfected with the -56 construct alone (lane 1) or with *dl30E1A*⁻ (lane 2) and with pIX/+90 pIX alone (lane 3) or with *dl30E1A*⁻ (lane 4). Lane M shows markers.

or including the pIX SP1 site. Several of these plasmids were then used to derive adenovirus constructs (see Materials and Methods) in order to confirm the minimal upstream sequence elements required to observe wild-type levels of pIX mRNA during infection. The amount of pIX mRNA synthesized from virus derivatives of these constructs was analyzed at early and late times of infection of HeLa cells (see Materials and Methods). As expected, none of the viruses produced pIX mRNA early in infection due to interference mediated by E1B transcription across the pIX promoter (data not shown). Virus constructs containing deletions between -250 and -150 or -100 nt produced wild-type levels of pIX mRNA in infected cells at late times (data not shown). The deletion mutant retaining 56 nt upstream of the pIX cap sites (nt -250 to -56, SP1⁺) produced wild-type levels of pIX mRNA at late times of infection (Fig. 3B). However, the deletion mutant that lacked the SP1 consensus site (-250 to -38, SP1⁻) was defective in pIX mRNA production at late times of infection (Fig. 3B). In fact, pIX mRNA was not observed at all from this virus, in contrast to the reduced but detectable levels of pIX mRNA obtained from the -38 construct with E1B in *trans* (Fig. 2). Both virus mutants produced normal levels of E1A mRNA (Fig. 3B). In order to ensure that the defective production of pIX mRNA from the SP1⁻ virus was not a consequence of insufficient production of E1B protein from this mutant virus, a coinfection experiment was performed with SP1⁻ virus and a virus construct of *dl30*. pIX mRNA was not observed at early or late times of coinfection into HeLa cells (data not shown). These results confirmed our transfection analyses regarding the minimal upstream sequence requirements for optimum pIX gene expression.

Sequences downstream of the pIX TATA box do not affect transcription from a heterologous promoter. These initial observations suggested a similarity between the pIX and the reported E1B promoter (37) elements with respect to the SP1 and TATA site requirements. There are also, however, marked differences in the expression of these two promoters. While E1A has been shown to activate adenoviral early-gene expression, including E1B gene expression (for

review, see reference 7), pIX gene expression does not appear to respond to E1A in transfection or infection experiments (29). Furthermore, the basal level of E1B gene expression in the absence of E1A is detectable in transient expression assays (29, 32) as well as in viral constructs (37); this is not the case for the pIX gene. Even in the absence of overriding E1B transcription, whether in plasmid or viral constructs, pIX gene expression is not detectable in the absence of E1B protein. In light of the similarity between the upstream sequence elements necessary to obtain optimal levels of E1B and pIX mRNAs, we prepared several other plasmid constructs in order to determine whether the sequence extending from -56 to -20 nt upstream of the pIX cap sites constituted the complete promoter and regulatory region of the pIX gene. In all of the following constructs, the upstream E1A gene and E1B promoters were absent so that E1B-mediated promoter interference was not a factor in examining the activity of the pIX promoter.

First, the SV40 early promoter was positioned upstream of the -20 construct, which lacked the upstream pIX promoter elements. This construct is diagrammed in Fig. 4A and was used in transient expression assays to determine whether any regions within or downstream of the pIX gene negatively affect the transcription of the heterologous SV40 early promoter. As shown in Fig. 4B, while pIX mRNA was only observed from the -56 construct when HeLa cells were cotransfected with *dl30E1A*⁻ DNA, steady-state levels of pIX mRNA directed by the SV40 early promoter were clearly obtained from the SV40-pIX fusion plasmid alone. The SV40-pIX mRNA protected a fragment of the SP6 probe extending from nt 3560 (-20 nt pIX cap sites) to nt 3775 (3' end of the SP6 antisense RNA) and probably initiated at the cap sites of the SV40 early promoter. This result suggested that sequences within and downstream of the pIX gene did not negatively interfere with the functioning of the SV40 early promoter.

DNA sequence spanning only the pIX SP1 and TATA sites is sufficient to initiate transcription at two cap sites internal to the pIX gene. Second, we tested the DNA sequence encompassing the delineated pIX promoter for its ability to direct a

heterologous mRNA species. In this case, the pIX promoter sequences extending from nt -56 to +57 were joined to the β -globin gene devoid of regulatory and promoter elements. β -globin mRNA was not detectable in cells transfected with the plasmid bearing the pIX- β -globin fusion (data not shown). However, in the presence of the 21-kDa E1B protein produced by *dl30E1A*⁻ plasmid DNA, β -globin mRNA was observed, as was utilization of the normal pIX cap sites (data not shown). This result suggested that the pIX sequences extending from -56 to +57 nt were sufficient to direct β -globin mRNA in an E1B-dependent manner.

In addition, we also questioned whether the presence of an acceptor splice site (nt 3595) used in the processing of the E1B primary transcript (24, 30), which lies downstream of the pIX cap sites (nt 3580 and 3582), affects pIX mRNA production. pIX mRNA is not spliced (2). Therefore, a third construct was derived so that the pIX promoter in the form of an oligonucleotide (see Materials and Methods) was positioned at nt +90 (nt 3670) within the pIX gene (Fig. 4A). This construct, pIX/+90 pIX, is devoid of the E1B acceptor splice site. Cells harboring plasmid DNA containing either the -56 wild-type construct or the pIX/+90 pIX fusion did not produce detectable levels of mRNA directed by the pIX promoter (Fig. 4C, lanes 1 and 3, respectively). However, cells cotransfected with *dl30E1A*⁻ gave rise to mRNA directed by the pIX promoter in both cases (lane 2, -56, and lane 4, pIX/+90 pIX). In the case of the pIX/+90 pIX fusion, protected fragments of the expected size for mRNA initiated approximately 24 nt downstream of the pIX TATA box contained within the oligonucleotide. Interestingly, the size of the protected fragments indicated that initiation of transcription occurred at two sites that appeared to be 2 bp apart (probably at nt 3688 and 3690), analogous to wild-type pIX mRNA. This result indicated that the presence of the E1B acceptor splice site downstream of the pIX mRNA initiation sites was inconsequential to pIX mRNA production and furthermore that the oligonucleotide containing only the pIX promoter region, without the pIX mRNA initiation sites, was capable of driving transcription in an E1B-dependent manner.

Similarity of E1B and pIX promoters. This last series of experiments suggested that the sequences upstream of the pIX cap sites contained within nt -56 to -20 constituted the pIX promoter and regulatory region for pIX gene expression. Comparison of the DNA sequence upstream of the E1B and pIX cap sites revealed a remarkable similarity in the two promoters (Fig. 5). While there are 11 and 20 nt between the SP1 and TATA sites in the E1B and pIX promoters, respectively, there is considerable homology between these spacer sequences; the notable difference is the extra 9 nt within the pIX promoter. Previous studies indicated that the spacing between the SP1 and TATA sites of the E1B promoter was optimal at 11 nt and that increasing the distance between these two sites was detrimental to E1B mRNA production (36).

In order to examine the effect of the extra 9 nt on basal pIX mRNA production, two constructs containing perturbations of this region were derived. Oligonucleotides containing the wild-type pIX promoter (-56 to -20) or a deletion of the extra 9 nt (del9) or a substitution of the extra 9 nt with *Bam*HI linker DNA (sub9) were placed upstream of the -20 construct (Fig. 5). The del9 construct was derived in order to determine whether placing the SP1 and TATA sites closer together, as they occur in the E1B promoter, would affect basal levels of pIX mRNA synthesis, while the sub9 construct was designed to examine the sequence specificity of

```

E1B   GGGCGGGGCTTAAAGGGTATATAAT
      SP1                TATA

pIX   GGGCGTGGCTTAAGGGTGGGAAAGAATATATAA
      SP1                TATA

del(9) GGGCGTGGCTTAAGGGATATATAA
      SP1                TATA

sub(9) GGGCGTGGCTTAAGGGCCGGATCCCATATATAA
      SP1                TATA

pIX/E1B GGGCGTGGCTTAAGGGTGGGAAAGAATATATAAT
      SP1                TATA

pIX/SV GGGCGTGGCTTAAGGGTGGGAAAGAATATTTAT
      SP1                TATA

```

FIG. 5. DNA sequence of wild-type and mutant pIX promoters and E1B promoter. The Ad5 DNA sequence contained within the pIX and E1B promoters is shown. Bold underlines delineate the SP1 and TATA consensus sequences. The sequence homology between spacer DNA in the pIX and E1B promoters is underlined. The DNA sequences of synthetic oligonucleotides del9 and sub9 are shown. del9 contains a deletion of the 9 nt present in the pIX promoter but absent in the E1B promoter, and sub9 contains a substitution of these 9 nt with *Bam*HI linker DNA. Two oligonucleotides containing the pIX promoter with a substitution of the pIX TATA box for either the E1B TATA box (pIX/E1B) or the SV40 TATA box (pIX/SV) are also shown.

this region. These constructs were transfected into HeLa cells in the absence and presence of E1B protein produced from *dl30E1A*⁻ DNA (Fig. 6). Neither the del9 nor sub9 construct gave different basal levels of pIX gene expression alone. In the presence of E1B, sub9 gave levels of pIX mRNA similar to those in the control, indicating that the exact sequence of the 9 nt was inconsequential to pIX mRNA synthesis. However, the del9 construct gave reduced levels of pIX mRNA relative to the control and sub9 cases. This result was unexpected, since del9 is nearly identical to E1B with respect to the promoter region. While this result indicated that the sequence of the extra 9 nt within the pIX promoter was not important, the optimal spacing between the SP1 and TATA sites appeared to include these 9 nt.

Decreased pIX gene expression when the SV40 TATA box is substituted for the pIX TATA box. The pIX promoter appeared to be similar to both the E1B and SV40 early promoters with respect to the requirements for SP1 and TATA sites. We therefore compared the pIX, E1B, and SV40 early TATA sites for their ability to function in the context of the pIX promoter in the absence and presence of E1B protein. Plasmid constructs containing the pIX gene and a substitution of the pIX TATA box with either the E1B or SV40 early TATA box (Fig. 5) were made (see Materials and Methods) and analyzed for pIX mRNA synthesis in

FIG. 6. Synthesis of pIX mRNA from the del9 and sub9 mutant pIX promoters. SP6 analysis of HeLa cells transfected with wild-type (wt), del9, or sub9 in the absence and presence of *dl30E1A*⁻ DNA, as indicated. Lane M shows markers.

FIG. 7. Synthesis of pIX mRNA when the TATA box is exchanged. SP6 analysis of HeLa cells cotransfected with *dI30E1A*⁻ and the -20 construct (TATA⁻) or plasmids containing the pIX gene with either the pIX TATA, E1B TATA, or SV40 early TATA box as indicated. Lane M shows markers.

transient expression assays. None of the constructs yielded pIX mRNA in the absence of E1B protein (data not shown). The pIX gene with the E1B TATA box produced amounts of pIX mRNA similar to those of the wild-type construct in the presence of E1B protein in *trans* (Fig. 7). In contrast, however, pIX mRNA was barely detectable for the construct carrying the SV40 early TATA box (Fig. 7). The small amount of pIX mRNA observed was accurately initiated at the pIX cap sites, and aberrant initiation sites were not detected. These results indicated that switching to the SV40 early TATA box led to a marked decrease in the function of the pIX promoter and that the typical stimulation observed with E1B protein could not be attained. The functional TATA elements of the pIX and SV40 early promoters were therefore not exchangeable in spite of the apparent similarity of these two transcription units.

DISCUSSION

An interesting facet of the pIX gene is that its promoter is very similar in structure to the promoter of the E1B gene, in which it is embedded. Our studies with transfection and infection assays indicated that pIX mRNA synthesis is dependent on SP1 and TATA sites within the pIX promoter. In fact, deletion of the SP1 consensus sequence has more drastic effects on pIX mRNA synthesis from the virus than from the plasmid constructs. Perhaps the formation of initiation complexes at the pIX promoter in the absence of SP1 is less stable in the virus than in plasmid DNA. We have not detected any requirements for DNA sequences upstream of the pIX SP1 site or downstream of the pIX TATA box that are important to pIX gene expression. Our studies support those of Matsui (21), which showed that sequences between nt -70 and -45 as well as the TATA box were important to pIX gene expression in transient expression assays, although the identity of the SP1 site was not recognized in that previous work. However, we have not found any sequences internal to the pIX gene that affect pIX gene expression *in vivo*. In studies performed by Matsui, sequences from +33 to +122 appeared to inhibit pIX gene expression *in vitro*, but these sequences alone did not appear to affect pIX gene expression *in vivo*. In our case, deletion of sequences between nt -20 and +90, which overlap the sequences deleted in constructs used by Matsui (+33 to +122), did not give rise to increased mRNA production in HeLa cells with the pIX/+90 pIX fusion construct. In addition, our studies *in vivo* did not detect any inhibitory sequences within the pIX

gene that affect the function of the SV40 early promoter in directing pIX mRNA synthesis. Matsui concluded that the effect of these downstream sequences could be observed *in vivo* when the pIX promoter consisted only of the TATA box (that is, sequences upstream of the TATA box were also deleted). However, the transcripts produced *in vivo* were analyzed with a 5'-end-labeled probe derived from the wild-type pIX gene; the start sites of transcription for the deleted mRNA observed were therefore not scored. Therefore, it is possible that these transcripts may have originated from vector sequences or from the SV40 enhancer present on the plasmid constructs used.

Our results do indicate a marked similarity between the pIX and E1B genes with respect to their promoter regions. The major difference in these two promoters is the additional 9 nt between the SP1 and TATA sites of the pIX promoter relative to the E1B promoter. Our studies indicate that deletion of these 9 nt results in lower levels of pIX mRNA synthesis, but substitution of these 9 nt with *Bam*HI linker DNA does not affect pIX mRNA levels. These results suggest that the spacing provided by these 9 nt rather than the sequence is important for optimal pIX gene expression. In contrast, the spacing between the SP1 and TATA sites of the E1B promoter is optimal without these 9 nt, although Wu and Berk (36) found that the addition of 10 nt between the SP1 and TATA sites did not significantly decrease E1B mRNA levels *in vivo*. While optimal levels of E1B mRNA were obtained with constructs having the wild-type spacing of 11 nt and slightly decreased levels were observed with a spacing of 21 nt, the opposite was true for the pIX gene. In the case of the pIX gene, optimum levels of pIX mRNA were observed with the wild-type spacing of 20 nt and declined when only 11 nt were present.

In attempting to compare the pIX, E1B, and SV40 early promoters, all of which consist of SP1 and TATA sites, we compared pIX mRNA levels when the pIX promoter contained either the E1B or SV40 early TATA box in place of the pIX TATA box. Studies by Wu et al. (37) have shown that correctly initiated E1B mRNA is formed even in the absence of the E1B TATA box with virus constructs containing linker-scanning mutations. We did not detect pIX mRNA or aberrantly initiated pIX mRNA with virus constructs containing deletions that extend across the SP1 and TATA sites of the pIX promoter, but we did not analyze mutants of the pIX promoter that retain only the SP1 site. Previous studies by Benoist and Chambon (5) have shown that mutations of the SV40 early TATA box do not lead to decreased levels of SV40 early mRNA but instead result in additional sites of transcription initiation. The SV40 early TATA box (TATTTAT), which is very similar to the E1A TATA box (TATTTATA), does not contain the recognized consensus TATA sequence; however, studies by Hahn et al. (16) have shown that the yeast transcription factor TFIID does bind to these sites. When the pIX TATA box was exchanged for that of E1B, no remarkable effects on pIX mRNA levels were observed. However, a substantial decline in pIX mRNA levels was observed with the SV40 early TATA box in constructs containing the pIX gene. Nonetheless, the SV40 TATA box in the context of the SV40 promoter was capable of directing pIX mRNA synthesis. These results strongly suggest that different TATA sequences contribute to the levels of promoter activity and that TATA sites are not necessarily interchangeable.

The sequence of the TATA site within the context of a given promoter appears to participate in the extent to which that promoter functions. Perhaps the presence of multiple

SP1 sites within the SV40 early promoter are necessary to facilitate transcription because the SV40 TATA site is inefficiently used *in vivo*. In contrast, the pIX promoter contains only one SP1 site, which, in conjunction with the SV40 TATA box, is inadequate to support transcription initiation. Recent studies by Wefald et al. (33) showed that replacement of the myoglobin TATA box (TATAAA) with the SV40 TATA box (TATTTAT) abolished myoglobin enhancer activity of the chloramphenicol acetyltransferase (*cat*) reporter gene in transient expression assays. This same construct retained the response to another muscle-specific enhancer, creatine kinase enhancer. This work also showed that the SV40 early promoter directing *cat* mRNA could be made responsive to the myoglobin enhancer when the SV40 TATA box was replaced with TATAAA. The sequence of the TATA box itself, therefore, was shown to contribute directly to responsiveness to specific enhancer elements.

The E1B TATA box (TATATAAT) was completely functional in the context of the pIX promoter, as expected given its close sequence homology with the pIX TATA box (TATATAA). Studies performed by Wu et al. (37) have shown that E1A-mediated activation of the E1B promoter involves the E1B TATA box. Studies performed by Simon et al. (27) have also shown that E1A activation of the heat shock promoter occurs via the TATA box. We have not been able to observe E1A activation of the pIX gene in cotransfection or coinfection experiments (29). In addition, E1A activation of the pIX gene containing the E1B TATA box was not detected in cotransfection experiments with pE1A (data not shown). It is possible that the TATA box is necessary but not sufficient to mediate E1A activation in this case.

Recent studies by Herrman and Mathews (18) have shown that the 21-kDa E1B protein nonspecifically stabilizes transfected DNA, leading to higher levels of DNA and higher steady-state levels of mRNA. In addition, mutations in the 21-kDa E1B protein result in degradation of cellular and viral DNA during infection (26, 28, 34). The virus construct Ad Bgl6, containing a deletion of the E1A and most of the E1B genes, was incapable of producing pIX mRNA alone during infection of HeLa cells (29). Over 16 h of infection, we did not observe degradation of the viral DNA (Southern analysis of DNA extracted from infected-cell nuclei by the Hirt protocol; data not shown). During coinfection of Ad Bgl6 with the viral equivalent of *dl30* (E1A⁺, E1B⁺, pIX⁻), pIX mRNA was clearly observed from Ad Bgl6 in the absence of DNA replication (29). Our results cannot, at this time, completely rule out the possibility that the 21-kDa E1B protein leads to stabilization of plasmid constructs carrying the pIX gene in cotransfection experiments.

Another possibility to consider is that E1B protein stabilizes pIX mRNA. We believe this is unlikely because all E1B mRNAs identified to date contain pIX mRNA sequences from nt +13 to the 3' terminus. E1B protein would therefore be stabilizing pIX mRNA differentially based upon the pIX mRNA sequence from the cap sites up to nt +13. However, the pIX/+90 pIX construct used in this study indicates that the sequences between nt -21 and +89 are dispensable for synthesis of doubly capped mRNA directed by the pIX promoter in the presence of E1B protein.

We are actively investigating the role of E1B protein in regulating pIX gene expression. Our results have shown, however, that even in the presence of E1B protein, pIX mRNA is only inefficiently produced from the pIX gene carrying the SV40 TATA box. Therefore, in this case, E1B protein was incapable of restoring optimal levels of pIX

mRNA. E1B protein may differentially alter the amount or conformation of transcription factor TFIID bound to the pIX promoter, dependent on the TATA box sequence. The possibility exists that E1B protein may stabilize DNA and may also activate transcription in some cases.

ACKNOWLEDGMENTS

We thank Hamish Young for helpful suggestions concerning the isolation of the recombinant viruses. We are very grateful to James E. Darnell, Jr., for consistent and generous support and advice. We thank Danny Reinberg for many helpful discussions and for critical review of this manuscript.

This work was aided by a grant from the American Cancer Society (JFRA-235) and by Public Health Service grant CA48707-10 from the National Institutes of Health to L.E.B. and by National Institutes of Health grant GM08670 to L.D.V.

REFERENCES

- Adhya, S., and M. Gottesman. 1982. Promoter occlusion: transcription through a promoter may inhibit its activity. *Cell* 29:939-944.
- Alestrom, P., G. Akusjarvi, M. Perricaudet, M. B. Mathews, D. F. Klessig, and U. Pettersson. 1980. The gene for polypeptide IX of adenovirus type 2 and its unspliced messenger RNA. *Cell* 19:671-681.
- Babiss, L. E., and H. S. Ginsberg. 1984. Adenovirus type 5 early region 1b gene product is required for efficient shutoff of host protein synthesis. *J. Virol.* 50:202-212.
- Bateman, E., and M. R. Paule. 1988. Promoter occlusion during ribosomal RNA transcription. *Cell* 54:985-992.
- Benoist, C., and P. Chambon. 1981. *In vivo* sequence requirements of the SV40 early promoter region. *Nature (London)* 290:304-310.
- Bergsma, D. J., D. M. Olive, S. W. Hartzell, and K. N. Subramanian. 1982. Territorial limits and functional anatomy of the simian virus 40 replication origin. *Proc. Natl. Acad. Sci. USA* 79:381-385.
- Berk, A. J. 1986. Adenovirus promoters and E1A transactivation. *Annu. Rev. Genet.* 20:45-79.
- Brunet, L. J., L. E. Babiss, C. S. H. Young, and D. R. Mills. 1987. Mutations in the adenovirus major late promoter: effects upon viability and transcription during infection. *Mol. Cell. Biol.* 7:1091-1100.
- Chinnadurai, G., S. Chinnadurai, and M. Green. 1978. Enhanced infectivity of adenovirus type 2 DNA and a DNA-protein complex. *J. Virol.* 26:195-199.
- Connelly, S., and J. L. Manley. 1989. A CCAAT box sequence in the adenovirus major late promoter functions as part of an RNA polymerase II termination signal. *Cell* 57:561-571.
- Corbin, V., and T. Maniatis. 1989. Role of transcriptional interference in the *Drosophila melanogaster* Adh promoter switch. *Nature (London)* 337:279-282.
- Crossland, L. D., and H. J. Raskas. 1983. Identification of adenovirus genes that require template replication for expression. *J. Virol.* 46:737-748.
- Cullen, B. R., P. T. Lomedico, and G. Ju. 1984. Transcriptional interference in avian retroviruses—implications for the promoter insertion model of leukaemogenesis. *Nature (London)* 307:241-245.
- Fuhrman, S. A., C. V. Beveren, and I. M. Verma. 1981. Identification of a RNA polymerase II initiation site in the long terminal repeat of Moloney murine leukemia viral DNA. *Proc. Natl. Acad. Sci. USA* 78:5411-5415.
- Ghosh-Choudhury, G., Y. Haj-Ahmad, and F. L. Graham. 1987. Protein IX, a minor component of the human adenovirus capsid, is essential for the packaging of full length genomes. *EMBO J.* 6:1733-1739.
- Hahn, S., S. Buratowski, P. A. Sharp, and L. Guarente. 1989. Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences. *Proc. Natl. Acad. Sci. USA* 86:5718-5722.
- Henderson, S. L., K. Ryan, and B. Sollner-Webb. 1989. The

- promoter proximal rDNA terminator augments initiation by preventing disruption of the stable transcription complex caused by polymerase read-in. *Genes Dev.* 3:212-223.
18. Herrmann, C. H., and M. B. Mathews. 1989. The adenovirus E1B 19-kilodalton protein stimulates gene expression by increasing DNA levels. *Mol. Cell. Biol.* 9:5412-5423.
 19. Jones, N. C., P. W. J. Rigby, and E. B. Ziff. 1988. Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. *Genes Dev.* 2:267-281.
 20. Jones, N. C., and T. Shenk. 1979. Isolation of adenovirus type 5 host range deletion mutants defective for transformation in rat embryo cells. *Cell* 17:683-689.
 21. Matsui, T. 1989. Novel regulation of transcription initiation of the peptide IX gene of adenovirus type 2. *Mol. Cell. Biol.* 9:4265-4271.
 22. McStay, B., and R. H. Reeder. 1990. An RNA polymerase I termination site can stimulate the adjacent ribosomal gene promoter by two distinct mechanisms in *Xenopus laevis*. *Genes Dev.* 4:1240-1252.
 23. Mermelstein, F. H., and D. Reinberg. 1989. Initiation of transcription by RNA polymerase II. *Biochim. Biophys. Acta* 1009:1-10.
 24. Montell, C., E. F. Fisher, M. H. Caruthers, and A. J. Berk. 1984. Control of adenovirus E1B mRNA synthesis by a shift in the activities of RNA splice sites. *Mol. Cell. Biol.* 4:966-972.
 25. Nevins, J. R. 1980. Definition and mapping of adenovirus-2 nuclear transcription. *Methods Enzymol.* 65:768-785.
 26. Pilder, S., J. Logan, and T. Shenk. 1984. Deletion of the gene encoding the adenovirus type 5 early region 1B 21,000-molecular-weight polypeptide leads to degradation of viral and host cell DNA. *J. Virol.* 52:664-671.
 27. Simon, M. C., T. M. Fisch, B. J. Benecke, J. R. Nevins, and N. Heintz. 1988. Definition of multiple, functionally distinct TATA elements, one of which is a target in the *hsp70* promoter for E1A regulation. *Cell* 52:723-729.
 28. Subramanian, T., M. Kuppuswamy, J. Gysbers, S. Mak, and G. Chinnadurai. 1984. 19-kD tumor antigen coded by early region 1B of adenovirus 2 is required for efficient synthesis and for protection of viral DNA. *J. Biol. Chem.* 259:11777-11783.
 29. Vales, L. D., and J. E. Darnell, Jr. 1989. Promoter occlusion prevents transcription of the adenovirus polypeptide IX gene until after DNA replication. *Genes Dev.* 3:49-59.
 30. Virtanen, A., and U. Pettersson. 1985. Organization of early region 1B of human adenovirus type 2: identification of four differently spliced mRNAs. *J. Virol.* 62:435-443.
 31. Volkert, F. C., and C. S. H. Young. 1983. The genetic analysis of recombination using adenovirus overlapping terminal DNA fragments. *Virology* 125:175-193.
 32. Yoshida, K., L. Venkatesh, M. Kuppuswamy, and G. Chinnadurai. 1987. Adenovirus transforming 19-kD antigen has an enhancer-dependent *trans*-activation function and relieves enhancer repression mediated by viral and cellular genes. *Genes Dev.* 1:645-658.
 33. Wefald, F. C., B. H. Devlin, and R. S. Williams. 1990. Functional heterogeneity of mammalian-TATA box sequences revealed by interaction with a cell-specific enhancer. *Nature (London)* 344:260-262.
 34. White, E., T. Grodzicker, and B. W. Stillman. 1984. Mutations in the gene encoding the adenovirus early region 1B 19,000-molecular-weight tumor antigen cause degradation of chromosomal DNA. *J. Virol.* 61:410-419.
 35. Wilson, M. C., N. W. Fraser, and J. E. Darnell, Jr. 1979. Mapping of RNA initiation sites by high doses of UV irradiation: evidence for three independent promoters within the left 11% of the ad-2 genome. *Virology* 94:175-184.
 36. Wu, L., and A. Berk. 1988. Constraints on spacing between transcription factor binding sites in a simple adenovirus promoter. *Genes Dev.* 2:403-411.
 37. Wu, L., D. S. E. Rosser, M. C. Schmidt, and A. Berk. 1987. A TATA box implicated in E1A transcriptional activation of a simple adenovirus 2 promoter. *Nature (London)* 326:512-515.