

## A Downstream-Element-Binding Factor Facilitates Assembly of a Functional Preinitiation Complex at the Simian Virus 40 Major Late Promoter

DONALD E. AYER† AND WILLIAM S. DYNAN\*

*Department of Chemistry and Biochemistry, University of Colorado, Campus Box 215, Boulder, Colorado 80309-0215*

Received 29 January 1990/Accepted 16 April 1990

**Recent work has shown that many promoters recognized by eucaryotic RNA polymerase II contain essential sequences located downstream of the transcriptional initiation site. We show here that the activity of a promoter element centered 28 base pairs downstream of the simian virus 40 major late initiation site appears to be mediated by a DNA-binding protein, which was isolated by affinity chromatography from HeLa cell nuclear extracts. In the absence of the other components of the transcriptional machinery, the protein bound specifically but weakly to its recognition sequence, with a  $K_d$  of approximately  $10^{-8}$  M. Analysis of kinetic data showed that mutation of the downstream element decreased the number of functional preinitiation complexes assembled at the promoter without significantly altering the time required for half the complexes to assemble. This suggests that in the absence of the downstream activating protein, preinitiation complexes are at least partially assembled but are not transcriptionally competent.**

Transcriptional initiation at eucaryotic mRNA promoters involves an ordered series of protein-DNA and protein-protein interactions. These interactions have been studied most extensively with promoters containing a TATA box homology. In the case of the adenovirus major late promoter, the general transcription factor TFIID binds the TATA box early in the reaction sequence and is later joined by other general transcription factors and RNA polymerase II (17, 23, 47, 58). Together these proteins form a preinitiation complex that can initiate transcription rapidly upon addition of nucleoside triphosphate substrates (33). The TFIID factor also plays a central role in transcription of the adenovirus E4 promoter; it is believed that transcriptional activator proteins induce a conformational change in TFIID, leading to enhanced promoter utilization (31, 35, 36).

In contrast to the extensive characterization of TATA box-containing promoters, much less is known about the mechanism of initiation at promoters that lack the TATA box homology. Such promoters are commonly associated with housekeeping and growth control genes (19, 26). In many cases, these promoters contain GC-rich sequences with potential binding sites for Sp1. Transcriptional initiation may be heterogeneous or may occur at a single site. To better understand the sequences required for initiation in the absence of a TATA box, we previously carried out an extensive linker scan mutational analysis of the simian virus 40 (SV40) major late promoter (3). This promoter, which lacks a recognizable TATA box homology, directs abundant synthesis of viral RNAs encoding late-gene products. Approximately 80 to 90% of the late RNAs have a 5' terminus at viral nucleotide 325, with the remainder distributed among various upstream start sites (28, 29, 45, 54).

Three *cis*-acting elements make up the core of the SV40 major late promoter (3). One element is centered 31 base pairs (bp) upstream of the initiation site in a position normally reserved for a TATA box. The second lies at the

transcription initiation site and perhaps functions analogously to the subsequently described initiator element of the TdT promoter (53). The third occupies a novel location centered 28 bp downstream of the transcriptional initiation site, within protein-coding sequence. The promoter structure determined by our linker scan analysis is consistent with the results of a number of earlier studies (4, 11, 45, 54). Although additional upstream sequences modulate the amount of late transcription, the region containing the three elements described here is sufficient to carry out the basic functions of a promoter, specifying both the position and start site for RNA synthesis (3).

The SV40 late promoter is one of several viral and cellular transcription units that have essential sequences downstream of the transcriptional initiation site. In principle, a downstream element may be recognized at either the RNA or DNA level and may influence RNA elongation, processing, and translation in addition to transcriptional initiation. The human immunodeficiency virus long terminal repeat promoter has a downstream element that appears to be recognized at the nascent RNA level (5; reviewed in reference 51). The bovine papillomavirus P1 and adenovirus major late promoters have downstream elements that interact with cellular DNA-binding proteins and so are presumably recognized at the DNA level (16, 40, 44, 55). Other promoters have downstream sequences required for expression, but in many cases the mechanism of action has not been well characterized (2, 7, 15, 22, 34, 42, 52, 56, 59).

In the present experiments, we have used DNA affinity chromatography to purify a protein from HeLa cell nuclear extracts that binds to the downstream element of the SV40 major late promoter. The isolated protein binds specifically but relatively weakly to its recognition site. In addition, the reaction conditions were manipulated to separate the transcription reaction at the late promoter into functional steps. These experiments showed that the downstream promoter element affects the number of functional preinitiation complexes without significantly altering the apparent first-order rate constant for complex assembly.

\* Corresponding author.

† Present address: Fred Hutchinson Cancer Research Center, 1124 Columbia St., Seattle, WA 98104.

## MATERIALS AND METHODS

**Plasmids.** The construction of pSVOL0 and pSVLS24 was described previously (3). These plasmids are isogenic except that pSVLS24 has a clustered point mutation in the downstream element of the SV40 major late promoter (see Fig. 3). The pSLS5/15 construct used in DNase footprinting experiments was constructed by recombining pSLS15 (3) and pSVLS5 (21) at the *Bgl*III site, which has the net effect of deleting SV40 nucleotides (nt) 106 to 273 and replacing them with a 10-bp *Bgl*III linker. pSLS5/15 × LS24 was made by substituting SV40 nt 294 to 768 of pSLS5/15 with the analogous fragment from pSVLS24.

**Cell culture and extracts.** HeLa cells were maintained and harvested as described before (3). Nuclear extracts were prepared for *in vitro* transcription as described by Dynan (20) except that the hypotonic lysis buffer and the extraction buffer contained 1 mg/ml each of the protease inhibitors leupeptin, aprotinin, soy bean trypsin inhibitor, pepstatin A, and phenylmethylsulfonyl fluoride. The final dialysis buffer had the same protease inhibitors at 0.1 mg/ml each.

**Protein purification.** Heparin-agarose was prepared as described before (18). A 16-ml column was washed with 1 column volume of TM buffer (50 mM Tris hydrochloride [Tris-HCl, pH 7.9], 12.5 mM MgCl<sub>2</sub>, 1 mM EDTA, 1 mM dithiothreitol, 20% glycerol) containing 1 M KCl and equilibrated with TM buffer containing 0.1 M KCl. Nuclear extract from a 16-liter culture of HeLa cells was adjusted to a conductivity equal to that of TM containing 0.1 M KCl and loaded onto the column. The column was washed with TM containing 0.2 M KCl until no protein was detected in the eluate. The column was eluted with TM containing 0.5 M KCl. Eluted fractions were assayed for protein by the method of Bradford (8). Fractions containing the bulk of the protein were pooled and dialyzed against TM containing 0.1 M KCl.

A DNA oligonucleotide affinity column was prepared by the method of Kadonaga and Tjian (38). Complementary oligonucleotides, 5'-GATCGCTGCGCCGGCTGTCACGC CAGGCCTC-3' and 5'-GATCGAGGCCTGGCGTGACAG CCGGCGCAGC-3', were synthesized and purified by preparative electrophoresis on 12% denaturing polyacrylamide gels, followed by gel filtration with Sephadex G-25. The concentration of each oligonucleotide was determined spectrophotometrically.

A 60-nmol amount of each oligonucleotide was mixed, allowed to anneal, and then phosphorylated and ligated essentially as described before (38). Ligated products were brought to 2 M ammonium acetate and ethanol precipitated two times. Two grams of CNBr-activated Sepharose-4B (Pharmacia) was prepared for coupling according to the manufacturer's instructions. Ligated DNA was covalently linked to the matrix as described before (38). The procedure yielded approximately 7 ml of packed resin containing 120 µg of DNA per ml.

A 4-ml DNA affinity column was washed with Z buffer (25 mM HEPES [N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid, pH 7.8], 12.5 mM MgCl<sub>2</sub>, 20% glycerol, 1 mM dithiothreitol, 0.1% Nonidet P-40 [NP-40]) containing 2 M KCl and equilibrated with Z buffer containing 0.1 M KCl. Pooled fractions from the 0.2 to 0.5 M heparin-agarose step fraction were dialyzed until the KCl concentration was 0.1 M, adjusted to 30 µg of poly(dI-dC) (alternating copolymer) per ml, incubated on ice for 20 min, and loaded onto the column at a flow rate of 0.15 ml/min. The column was washed with 8 ml of Z buffer containing 0.1 M KCl and

eluted with a 20-ml linear gradient of 0.1 to 1.25 M KCl in Z buffer, at a flow rate of 0.1 ml/min. Fractions (0.5 ml) were collected and assayed for specific DNA binding by electrophoretic mobility shift or DNase I footprinting. Active fractions were pooled and diluted with Z buffer to 0.1 M KCl. Fractions were adjusted to 10 mg of poly(dI-dC) per ml and subjected to a second round of DNA chromatography with a 2-ml column. Wash and elution volumes and flow rates were reduced twofold.

The protein used in the electrophoretic mobility shift assays was further purified by Mono S cation-exchange chromatography. Active fractions from the second affinity column were pooled and diluted with Z buffer to 0.1 M KCl. The sample was applied at a flow rate of 0.15 ml/min to a 5/5 HR Mono S column previously equilibrated with Z buffer containing 0.1 M KCl. The column was washed with 5 column volumes of Z buffer containing 0.1 M KCl, and the proteins were eluted with a 20-ml linear gradient of 0.1 to 0.75 M KCl in Z buffer.

**DNase I footprinting.** pSLS5/15 and pSLS5/15 × LS24 were subcloned into polylinker-containing vectors. DNA was singly end-labeled with [ $\gamma$ -<sup>32</sup>P]ATP at an *Eco*RI site introduced just distal to the *Hinc*II site at SV40 nt 470. Labeled DNA was cut with *Hind*III at SV40 nt 5171, isolated on a 6% polyacrylamide gel, and recovered by electroelution. After ethanol precipitation, fragments were suspended in H<sub>2</sub>O to 10 nM concentration and stored at -20°C. This procedure generated a fragment that was labeled on the noncoding strand downstream of the nt 325 transcriptional initiation site. To make probes that were labeled on the coding strand, the *Eco*RI site was filled in with [ $\alpha$ -<sup>32</sup>P]dATP by using the Klenow fragment of *Escherichia coli* DNA polymerase I. The probe fragment was digested with *Hind*III and isolated as described above.

Binding reaction mixes for DNase I footprinting assays (20, 25) contained, in addition to the protein, 0.2 nM DNA probe, 10 µg of poly(dA-dT) (alternating copolymer) per ml, 25 mM Tris-HCl (pH 7.9), 6.25 mM MgCl<sub>2</sub>, 0.5 mM EDTA, 0.5 mM dithiothreitol, 10% glycerol, and approximately 80 mM KCl in a final volume of 50 µl. Samples were incubated on ice for 10 min, and 50 µl of 5 mM CaCl<sub>2</sub>-10 mM MgCl<sub>2</sub> was added. Samples were treated with dilute DNase I for 1 min and analyzed on 8% denaturing polyacrylamide gels as described before (20). Poly(dA-dT) carrier was used in place of the more usual poly(dI-dC) because the latter was shown to inhibit binding of the downstream activating protein (data not shown).

**Electrophoretic mobility shift assay.** To prepare labeled probes, pSVOL0 and pSVLS series mutants were digested with Asp 718 at SV40 nt 294 and end labeled with [ $\gamma$ -<sup>32</sup>P]ATP. After termination of the reaction, the DNA was digested with *Hinc*II at SV40 nt 470. The resulting 176-bp fragment was isolated on an 8% native polyacrylamide gel. After electroelution, the DNA was suspended in H<sub>2</sub>O to 10 nM concentration.

Binding reaction mixes containing 6.7 nM DNA, 6.25 mM MgCl<sub>2</sub>, 120 mM KCl, 25 mM HEPES (pH 7.9), 10% glycerol, and 50 µg of poly(dA-dT) per ml in a final volume of 15 µl were incubated on ice for 15 min. Loading dyes were added, and the samples were fractionated on a 5% native polyacrylamide gel (acrylamide-bisacrylamide, 49:1 [wt/wt]) containing 40 mM Tris, 36 mM glycine (pH 8.3), and 0.1% NP-40 (41) at 4°C. The gel was dried and exposed for autoradiography.

Quantitative binding experiments, which were performed to determine the protein-DNA binding constant, incorpo-

rated some modifications of the above procedure. An unlabeled DNA fragment was generated by digestion with Asp 718 (at SV40 nt 294) and *Bam*HI, which cuts at an introduced site immediately distal to SV40 nt 470. After preparative gel electrophoresis, DNA was allowed to elute by diffusion, centrifuged to remove polyacrylamide fragments, and incubated with 1 ml of TSK-Gel Toyopearl DEAE-650S (Supelco) for 2 h at room temperature. The matrix was poured into a column and washed with 10 ml of 10 mM Tris (pH 7.9)–1 mM EDTA–50 mM NaCl. DNA was eluted with the same buffer containing 500 mM NaCl. DNA-containing fractions were pooled, extracted with phenol-chloroform, and precipitated with ethanol. The pellet was suspended in 10 mM Tris-HCl (pH 7.9)–1 mM EDTA, and the DNA concentration was determined spectrophotometrically. The probe was radiolabeled with [ $\gamma$ - $^{32}$ P]ATP and polynucleotide kinase. The reaction was terminated by heating to 70°C for 10 min, and the DNA was diluted to the final working concentration without extraction or precipitation. Binding reaction mixes contained  $1.33 \times 10^{-10}$  M specific labeled DNA fragment, 6.6 mM Tris-HCl (pH 7.9), 0.13 mM EDTA, 9.2 mM HEPES (pH 7.8), 0.04% NP-40, 6.25 mM MgCl<sub>2</sub>, 10% glycerol, 1 mM dithiothreitol, 120 mM KCl, and 33  $\mu$ g of poly(dA-dT):poly(dA-dT) per ml in a final volume of 15  $\mu$ l. Reaction mixes were incubated for 20 min at 15°C, then 2  $\mu$ l of 0.8% NP-40–48% glycerol–0.02% bromphenol blue was added, and the samples were analyzed on a 5% polyacrylamide gel as described above (41). Binding was quantitated using an Automated Microbiology Systems (AMBIS) radioanalytic scanner. Values for concentrations of downstream activating protein and the equilibrium constant for protein-DNA binding were obtained by computer-assisted fitting of a theoretical binding curve, as described before (41).

**In vitro transcription.** In vitro transcription reactions were performed as described before (3) except that nuclear extract rather than whole-cell extract was used. Transcription reactions were divided into functional steps as described in the text and figure legends. Reaction mixes contained final concentrations of 20  $\mu$ g of DNA per ml, 250  $\mu$ M each nucleoside triphosphate, 25 mM Tris-HCl (pH 7.9), 50 mM KCl, 6.25 mM MgCl<sub>2</sub>, 0.5 mM dithiothreitol, 0.5 mM EDTA, and 10% glycerol. In a typical 50- $\mu$ l reaction mix, 15  $\mu$ l of nuclear extract was used. Following transcription, reaction mixes were extracted with phenol-chloroform (1:1, vol/vol) and ethanol precipitated as described previously (3). The pellets were suspended in 100  $\mu$ l of 2.5 mM CaCl<sub>2</sub>–5 mM MgCl<sub>2</sub> and incubated with 10 U of RNase-free DNase I for 10 min at 37°C. The reaction was terminated by the addition of 100  $\mu$ l of 200 mM NaCl–20 mM EDTA–1% sodium dodecyl sulfate. Reaction mixes were extracted with phenol-chloroform, precipitated with ethanol, and subjected to primer extension analysis as described before (3). Transcription was quantitated by AMBIS radioanalytic scanning. Curves in Fig. 6B and C were obtained by nonlinear least-squares fit with an equation of the form  $N = N_f(1 - e^{-kt})$ , where  $N$  is the number of preinitiation complexes at time  $t$ ,  $N_f$  is the limit approached by  $N$  as  $t$  goes to infinity, and  $k$  is the first-order rate constant that governs the reaction.

## RESULTS

**Purification of a factor that binds the downstream control element.** Nuclear extract from a 16-liter suspension culture of HeLa cells was fractionated on a heparin-agarose column and eluted with a 0.2 to 0.5 M step gradient of KCl. The bulk of the protein that eluted in this step fraction was pooled;

this represented about 10% of the protein present in the starting material. Although fractions prepared in this way were transcriptionally active, no specific binding of a factor to the downstream promoter element was detected by DNase I protection or electrophoretic mobility shift assays (data not shown).

The material was further fractionated by DNA affinity chromatography with sequences from the downstream promoter element. In previous studies, the downstream element was defined by linker scan mutants. Two mutants, LS23 and LS24, showed decreased transcription, whereas flanking mutants, LS22 and LS25, had wild-type activity (3). Thus, the entire element should be contained in the sequence between LS22 and LS25, which corresponds to SV40 nt 340 to 366. This sequence (5'-GCTGCCCGGGCTGTACGC CAGGCTC-3') and its complement were synthesized as oligonucleotides and prepared as an affinity matrix (38). The heparin-agarose 0.2 to 0.5 M KCl step fraction was applied to the column, which was then washed and eluted with a linear gradient of KCl.

Fractions were assayed for specific binding by DNase I footprinting. Fractions that eluted between 0.35 and 0.5 M KCl induced DNase I-hypersensitive sites at several positions between nt 340 and nt 365 in the wild-type probe but not in the LS24 mutant (data not shown). The induction of specific hypersensitive sites without clear regions of DNase I protection is often indicative of subsaturating levels of protein binding. To further enrich for binding activity, the active fractions were pooled, diluted, subjected to a second round of affinity chromatography with the same matrix, and assayed by DNase I footprinting. A distinct peak of protecting activity eluted between 0.32 and 0.56 M KCl (fractions 20 to 24) (Fig. 1). In addition, flanking DNase I-hypersensitive sites were induced (large arrowheads) at the same positions as with the first column.

A separate activity eluted from the column at approximately 0.25 to 0.3 M KCl and protected sequences between the Sp1-binding sites and the A/T-rich region of the SV40 early promoter (Fig. 1, fractions 16 to 18, near 309-bp marker). This activity bound equally well to the wild-type and the mutant probe (data not shown), suggesting that its enrichment with the affinity column was fortuitous.

To further characterize the binding to the downstream promoter element, protein titrations were performed with wild-type and LS24 mutant probes labeled on either the coding or noncoding strand (Fig. 2A and B). The DNase I-protected region extended between nt 346 and nt 365 on the noncoding strand, as determined by alignment with Maxam-Gilbert sequence markers (44a). DNase I-hypersensitive sites were induced at nt 340, nt 343, nt 344, nt 366, and nt 367. There was no protection of LS24 mutant DNA, although slight DNase I hypersensitivity was induced at nt 366 and 367 with the highest amount of protein tested (Fig. 2A). Based on the amount of protein required to give equivalent hypersensitivity, it appeared that the protein bound the mutant probe 5- to 10-fold less tightly than the wild-type probe. On the coding strand, protection extended approximately from nt 346 to nt 363, with flanking hypersensitivity at nt 343, nt 344, nt 367, and nt 368. There was no protection of LS24 mutant DNA. A summary of the protected nucleotides and the DNase I-hypersensitive sites on both the coding and the noncoding strands is given in Fig. 3.

**Correlation between binding and transcription.** Radiolabeled probes were prepared from the wild-type promoter and from several linker scan mutants encompassing the downstream element (3). Protein that had been purified over two

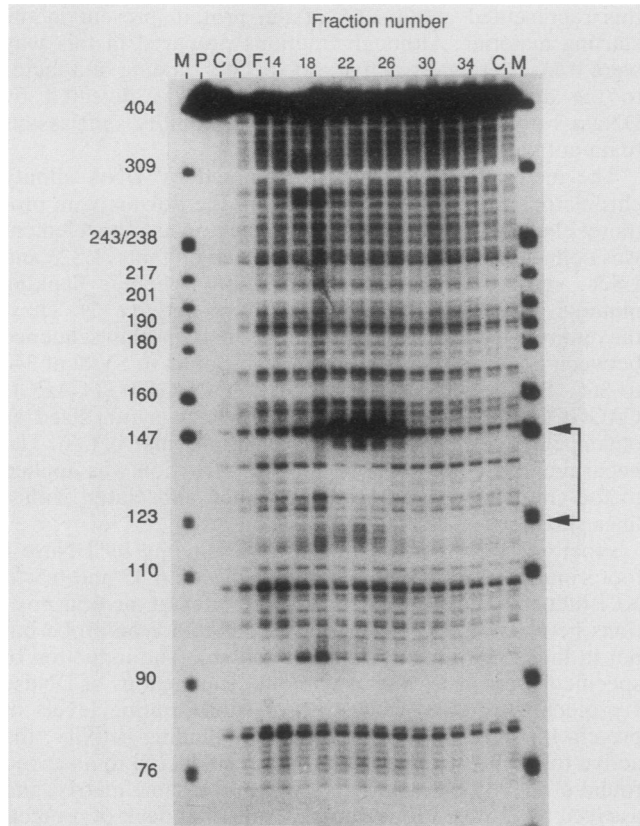


FIG. 1. DNase I footprinting assay of fractions from DNA affinity column. The pSLS5/15 noncoding-strand probe, which has a wild-type downstream element, was prepared and used in DNase I footprinting assays as described in Materials and Methods. Lanes M, pBR322 *Hpa*II-cut DNA size markers, with sizes indicated in base pairs. Lane P, Probe DNA uncut by DNase I. Lanes C, Control binding reaction mixes with no protein. Lane O, Assay of DNA affinity column output. Lane F, Assay of column flowthrough. Remaining lanes, Assays of even-numbered column fractions as indicated. Arrows mark DNase I-hypersensitive nucleotides; bracket marks protected nucleotides.

affinity columns and a Mono S cation-exchange column (data not shown) was added to each of the probes, and the resulting protein-DNA complexes were resolved by an electrophoretic mobility shift assay (Fig. 4) (24, 27). The uppermost complex (arrow) was seen with the wild type, LS21, LS22, and LS25 mutant probes, which are transcriptionally active (3), but not with the LS23 and LS24 mutant probes. Thus, the presence of this complex correlated precisely with late promoter transcriptional activity. This correlation suggests that the protein we have isolated is a transcriptional activator, and we shall refer to it subsequently as the downstream activating protein (DAP). The amount of the lower complexes varied between experiments. Because these complexes were not affected by promoter mutations, they may reflect binding by DAP to another site or binding by other proteins which may be present as contaminants in the preparation.

Genetic analysis suggested that the start site element and the downstream control element interact with each other (3). If the interaction between the two control elements were mediated by DAP itself, one would expect to see altered binding with the LS21 probe, which has a mutated start site element. However, this was not seen, suggesting that any

interaction between the start site element and the downstream element is mediated indirectly, perhaps by RNA polymerase II or other proteins present in the transcription complex.

**Determination of the equilibrium constant for DAP binding.** Quantitative binding studies were carried out with the electrophoretic mobility shift assay (Fig. 5). To obtain an estimate of DAP concentration, the amount of protein was held constant and the total DNA concentration was varied by adding nonradioactive but otherwise identical competitor DNA to the reaction mix. The fraction of the DNA bound by protein decreased as competitor was added. Radioactivity was quantitated by AMBIS scanning. As indicated in the figure, several closely spaced protein-DNA complexes were grouped together in the analysis. Computer-assisted least-squares fitting of a theoretical binding curve (Fig. 5C) showed the concentration of active protein to be  $(2.4 \pm 0.3) \times 10^{-9}$  M. The method of data analysis has been described in detail elsewhere (41). The standard deviation given here was determined by standard linear approximation (6) and may be an underestimate (41).

As determined by  $R_f$  values, the most slowly migrating band in the group in Fig. 5A is a species not present in Fig. 4. The new band was the first to disappear on the DNA titration and may therefore represent DNA molecules with more than one protein bound. Because this band was seen only at lower DNA concentrations (lanes 2 to 5), it is expected to have little effect on the empirically determined value for the concentration of DAP. The most rapidly migrating band in the group apparently corresponded to one of the lower complexes in Fig. 4 and probably represents binding by DAP or another protein to a site outside the promoter region. Because this complex disappeared at the same time as the promoter-bound complex, there was again no effect on the determination of DAP concentration.

To obtain an estimate of the equilibrium constant for DAP binding, another experiment was performed in which protein concentration was varied while the amount of DNA was held constant (Fig. 5B). A theoretical curve was fit to the data by using values for active DAP concentration determined from the previous experiment and adjusting the parameter that represents the dissociation constant (Fig. 5D). In this experiment, the apparent  $K_d$  was  $(6.8 \pm 0.2) \times 10^{-9}$  M. Because the lower, non-promoter-bound complex represents one-third to one-half of the total binding, however, this value should be adjusted correspondingly, giving a  $K_d$  for binding to the promoter site in the range of  $10^{-8}$  M.

**The downstream element affects assembly of preinitiation complexes.** The SV40 late promoter functions efficiently in an *in vitro* transcription system containing HeLa whole-cell or nuclear extract. The downstream element is required for wild-type levels of RNA synthesis at the nt 325 start site (3). To further characterize the role of this element in transcription, the transcription reaction was divided into individual functional steps.

In the absence of nucleoside triphosphates, components of the transcriptional machinery associate with eucaryotic promoters to form preinitiation complexes that are able to initiate transcription rapidly upon addition of nucleoside triphosphates. Preinitiation complex formation has been studied most extensively with the adenovirus major late promoter. These complexes are stable to washing (1), protect defined sequences of the DNA from nuclease attack (13, 58), and migrate as discrete species in an electrophoretic mobility shift assay (12).

The rate of preinitiation complex formation at the SV40

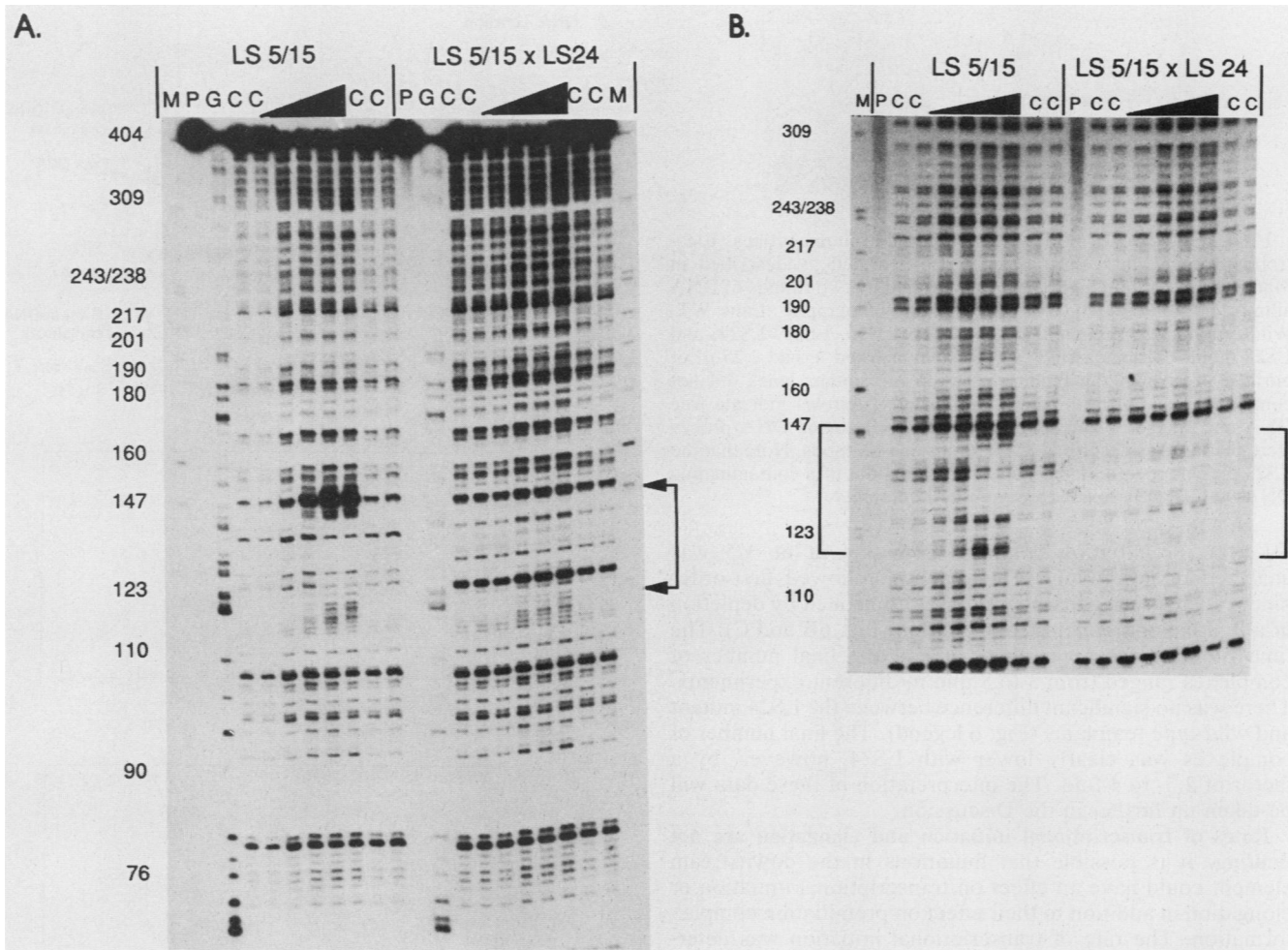


FIG. 2. Titration of protein binding to the downstream element. DNase I footprinting assay was done with noncoding-strand probes (A) and coding-strand probes (B). Probes were pSLS5/15 (wild-type downstream element) and pSLS5/15 × LS24 (mutant downstream element) as indicated. The protein was the peak fraction from the second affinity column. Lanes M, P, and C are as in Fig. 1. Lane G, Maxam-Gilbert guanosine-specific sequencing ladder. Dark wedge indicates increasing volume of protein: 5, 10, 15, and 25  $\mu$ l, respectively. Arrows and brackets are as in Fig. 1.

major late promoter was measured by incubating template DNAs for a variable time with HeLa cell nuclear extract. Nucleoside triphosphates were added to initiate RNA synthesis (Fig. 6). After 90 s, further initiation was blocked by the addition of Sarkosyl to 0.2% final concentration, and elongation was allowed to proceed for 30 min. Control experiments showed that initiation, but not elongation, of SV40 late RNA was inhibited by  $\geq 0.1\%$  Sarkosyl (data not

shown). The amount of RNA synthesized under these conditions reflects the number of functional preinitiation complexes present at the time that nucleoside triphosphates were added.

Two prominent RNA molecules were synthesized in in vitro transcription reactions (Fig. 6A). An initiation site at nt 170 was used in vitro but not in vivo; sequences required for initiation at this site have been described in detail elsewhere

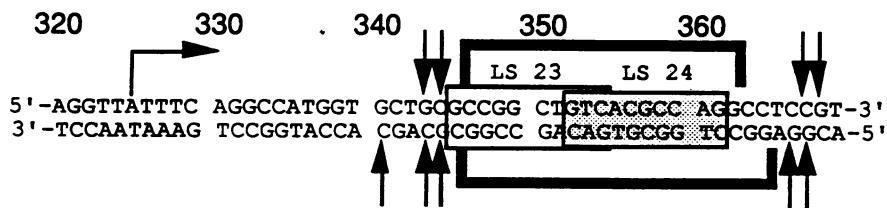


FIG. 3. Nucleotides protected by protein binding to downstream element. The sequences shown include the nt 325 major late cap site of SV40 (rightward arrow). Numbers at top indicate SV40 nucleotide positions. Boxes indicate sequences replaced by 10-bp linker (CGAGATCTCG) in the LS23 and LS24 mutants, respectively. Vertical arrows indicate bases that become hypersensitive to DNase I in the presence of bound protein. Brackets show regions on either strand that are protected from DNase I digestion.

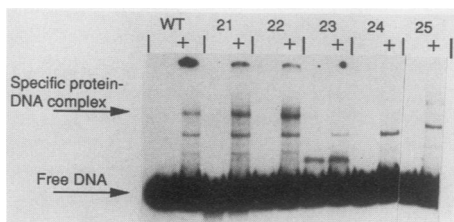


FIG. 4. Binding of DAP to wild-type and mutant probes. Electrophoretic mobility shift assay was performed as described in Materials and Methods. Protein was purified by two rounds of DNA affinity chromatography and Mono S chromatography. Lane WT, Wild-type probe; remaining lanes, LS21, LS22, LS23, LS24, and LS25 mutant probes as indicated. Lanes marked + had 2.25  $\mu$ l of purified protein in the binding reaction mix; other lanes did not contain protein in the binding reaction mix. Arrows indicate free DNA and specific protein-DNA complexes. Labeled DNA probes were prepared as described in Materials and Methods. Note that the LS23 probe appeared to contain a small amount of a contaminating DNA with slightly lower electrophoretic mobility.

(21, 32). Preinitiation complex formation at nt 325 was analyzed quantitatively. The reaction followed first-order kinetics, suggesting that the rate was controlled by depletion of a limiting component from solution (Fig. 6B and C). The time required for appearance of half the final number of complexes ranged from 3 to 8 min in different experiments. There was no significant difference between the LS24 mutant and wild-type templates (Fig. 6 legend). The final number of complexes was clearly lower with LS24, however, by a factor of 2.5- to 4-fold. The interpretation of these data will be taken up further in the Discussion.

**Rates of transcriptional initiation and elongation are not limiting.** It is possible that mutations in the downstream element could have an effect on transcriptional initiation or elongation in addition to their effect on preinitiation complex formation. The rate of transcriptional initiation was determined by measuring the time required for preinitiation complexes to be converted to Sarkosyl-resistant elongation complexes (Fig. 7A and B). Preinitiation complexes were formed for 60 min at 30°C, with either wild-type or LS24 mutant template. The incubation temperature was decreased to 20°C so as to slow the initiation reaction enough to permit accurate rate measurements. Nucleoside triphosphates were added, and at various times Sarkosyl was added to a final concentration of 0.2% to block further initiation. Elongation was allowed to proceed for 30 min at 30°C.

The transcriptional initiation reaction was essentially complete after 20 s at 20°C (Fig. 7A). The data were normalized to determine the fraction of the preinitiation complexes that were converted to elongation complexes after various times (Fig. 7B). There was no significant difference in the rate of conversion when wild-type and LS24 mutant templates were compared. Furthermore, the rate of transcriptional initiation was far too rapid to be limiting in the overall transcription reaction.

The rate of transcriptional elongation was estimated by measuring the time required for an initiated RNA polymerase II to synthesize an RNA long enough to be assayed by primer extension analysis. Preinitiation complexes were formed by using either the wild-type or LS24 mutant template. Nucleoside triphosphates were added, and after 2 min Sarkosyl was added to a final concentration of 0.2% to block further initiation. Reactions were terminated at various times, and RNA synthesis was measured by the primer extension assay (Fig. 7C). The amount of RNA synthesis

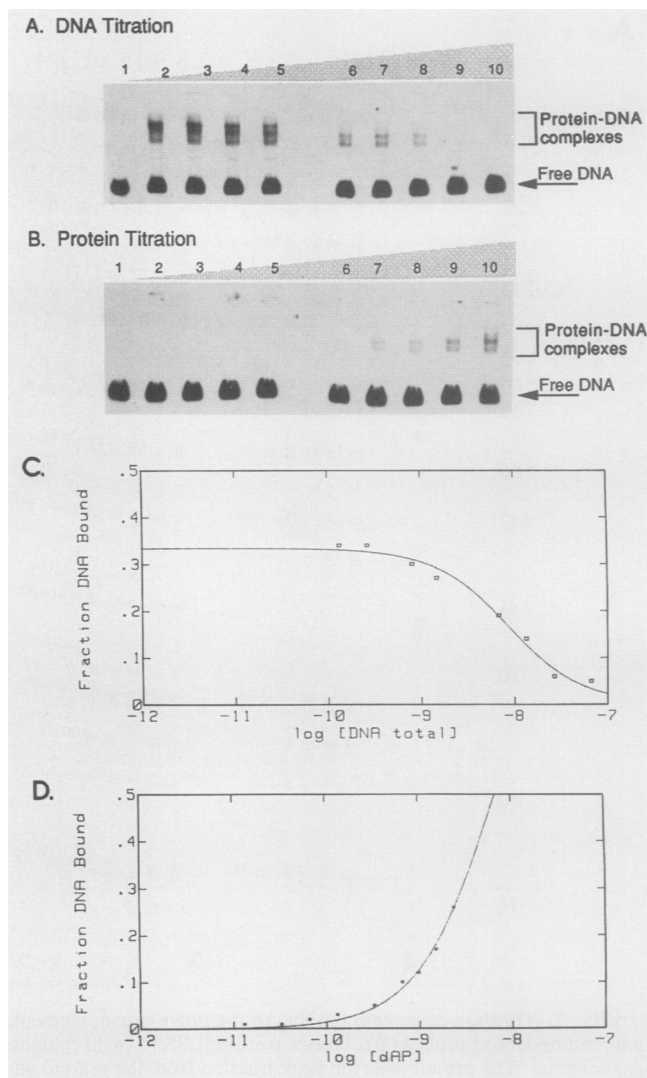
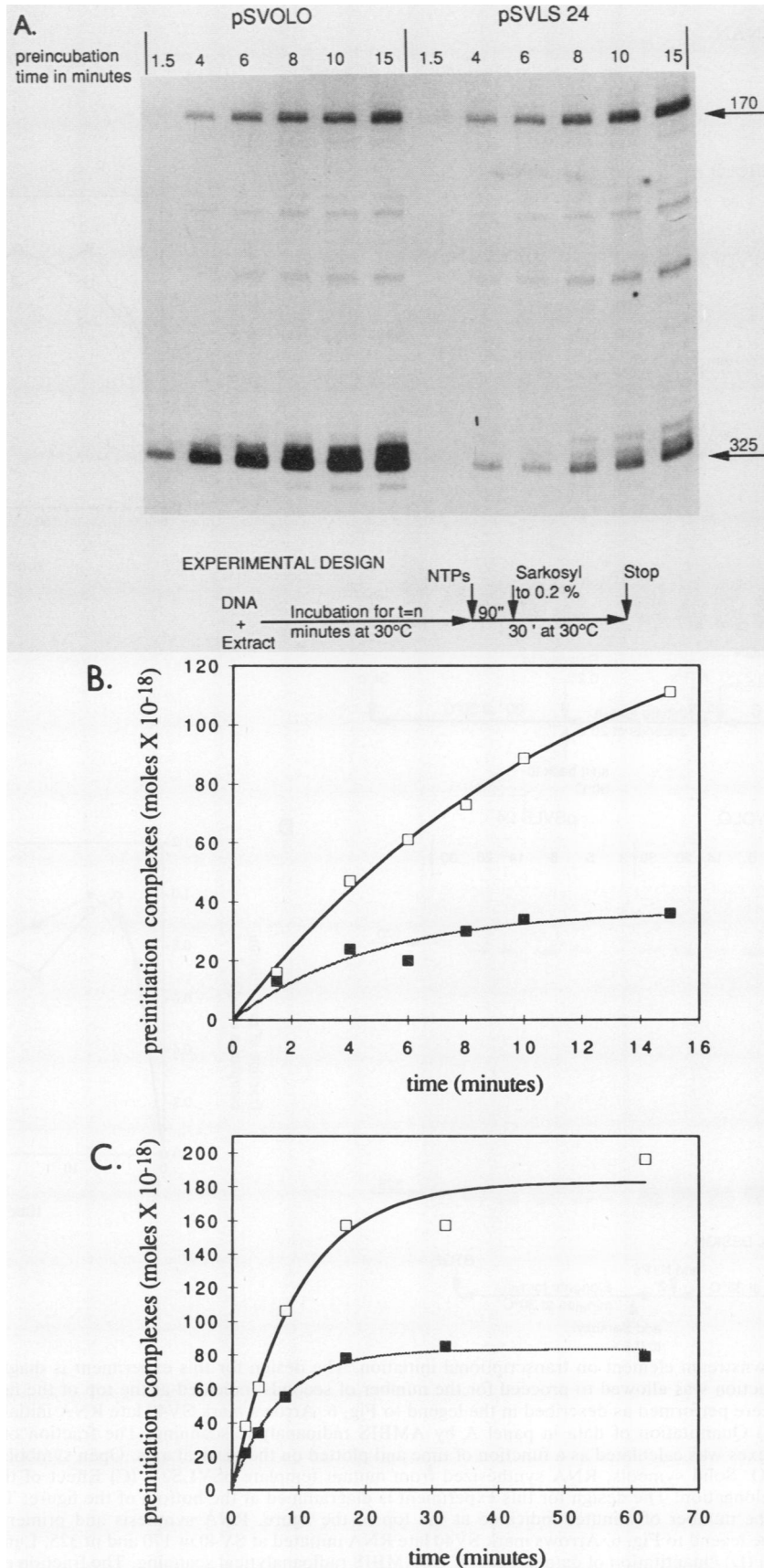
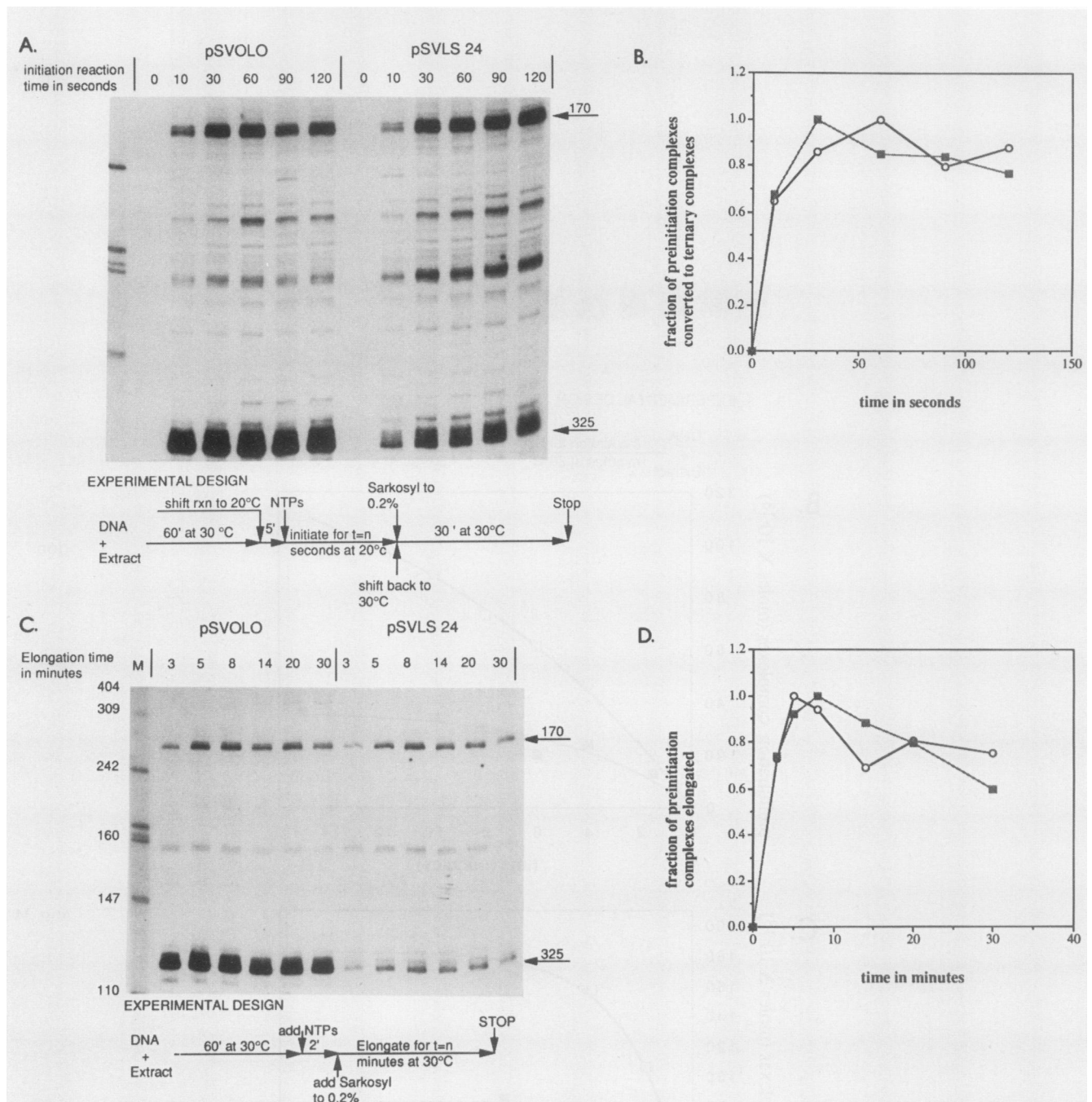


FIG. 5. Determination of the dissociation constant for the DAP-DNA complex. DNA probe was prepared from wild-type SV40 DNA, and electrophoretic mobility shift assay was performed as described in Materials and Methods. The DAP protein was purified by two rounds of DNA affinity chromatography and by Mono S chromatography. (A) Titration with increasing amounts of DNA. Binding reaction mix in lane 1 contained no protein. Other binding reaction mixes contained 5.5  $\mu$ l of purified DAP. All reaction mixes contained  $1.33 \times 10^{-10}$  M radiolabeled DNA and were supplemented with various amounts of the identical nonradiolabeled fragment. Shaded wedge indicates increasing total concentration of DNA in the binding reaction mixes:  $1.33 \times 10^{-10}$  M,  $2.67 \times 10^{-10}$  M,  $8.0 \times 10^{-10}$  M,  $1.47 \times 10^{-9}$  M,  $2.80 \times 10^{-9}$  M,  $6.80 \times 10^{-9}$  M,  $1.35 \times 10^{-8}$  M,  $2.68 \times 10^{-8}$  M, and  $6.68 \times 10^{-8}$  M DNA for lanes 2 to 10, respectively. (B) Titration with increasing amounts of purified DAP. Each binding reaction mix contained  $1.33 \times 10^{-10}$  M radiolabeled probe. Shaded wedge indicates increasing total concentration of DAP in the binding reaction mixes: no DAP,  $1.31 \times 10^{-11}$  M,  $3.27 \times 10^{-11}$  M,  $6.53 \times 10^{-11}$  M,  $1.31 \times 10^{-10}$  M,  $6.53 \times 10^{-10}$  M,  $9.8 \times 10^{-10}$  M,  $1.52 \times 10^{-9}$  M, and  $2.39 \times 10^{-9}$  M DAP for lanes 1 to 10, respectively. (C) Quantitation of data in panel A by AMBIS radioanalytic scanning. Solid line indicates theoretical binding curve (41). Best fit was obtained with a DAP concentration of  $(2.4 \pm 0.3) \times 10^{-9}$  M and a  $K_d$  of  $(4.8 \pm 0.7) \times 10^{-9}$  M. (D) Quantitation of data in panel B. Solid line indicates theoretical binding curve (41). Best fit was obtained with a  $K_d$  of  $(6.8 \pm 0.2) \times 10^{-9}$  M.



**FIG. 6. (A)** Effect of the downstream promoter element on preinitiation complex assembly. The design of the experiment is diagrammed at bottom. Preinitiation complexes were formed on pSVOLO and pSVLS24 DNA for the number of minutes indicated at the top of the figure. In vitro RNA synthesis and primer extension analysis was performed as described in Materials and Methods. Arrows mark SV40 late RNA initiated at SV40 nt 170 and nt 325. 90", 90 s; 30', 30 min. **(B)** Quantitation of data in panel A by AMBIS radioanalytic scanning. Open symbols, RNA synthesized from the wild-type template (pSVOLO). Solid symbols, RNA synthesized from the mutant template (pSVLS24). **(C)** Independent experiment similar to experiment in panel B except that the time course was extended to 60 min. In this experiment, the time required for half the complexes to assemble was 6.6 min for the wild-type template and 4.9 min for LS24.



**FIG. 7. (A)** Effect of downstream element on transcriptional initiation. The design for this experiment is diagrammed at the bottom of the figure. The initiation reaction was allowed to proceed for the number of seconds indicated at the top of the figure. RNA synthesis and primer extension analysis were performed as described in the legend to Fig. 6. Arrows mark SV40 late RNA initiated at SV40 nt 170 and nt 325. rxn, Reaction mix. **(B)** Quantitation of data in panel A by AMBIS radioanalytic scanning. The fraction of preinitiation complexes converted to ternary complexes was calculated as a function of time and plotted on the vertical axis. Open symbols, RNA synthesized from wild-type template pSVOLO. Solid symbols, RNA synthesized from mutant template pSVLS24. **(C)** Effect of the downstream promoter element on transcriptional elongation. The design for this experiment is diagrammed at the bottom of the figure. Transcriptional elongation was allowed to occur for the number of minutes indicated at the top of the figure. RNA synthesis and primer extension analysis were performed as described in the legend to Fig. 6. Arrows mark SV40 late RNA initiated at SV40 nt 170 and nt 325. Lane M, DNA markers, with sizes indicated in base pairs. **(D)** Quantitation of data in panel C by AMBIS radioanalytical scanning. The fraction of preinitiation complexes elongated to the point where product RNA could be assayed by primer extension analysis was calculated as a function of time and plotted on the vertical axis. Open symbols, RNA synthesized from wild-type template pSVOLO. Solid symbols, RNA synthesized from mutant template pSVLS24.



was normalized to the number of functional preinitiation complexes present on each template (Fig. 7D). Elongation was largely complete by 5 to 7 min after the addition of nucleoside triphosphates, long before the end of the standard 30-min incubation period used in other experiments. The difference between the wild-type and the LS24 mutant templates appears to be within experimental error. As the period of elongation increased past 10 min, there was a slight reduction in the amount of detectable RNA, perhaps due to endogenous RNases present in the HeLa cell nuclear extract.

RNA polymerase II must transcribe at least 125 nucleotides to synthesize an RNA that is long enough to be assayed under the conditions used here. A substantial amount of this RNA appeared within 5 min, implying that the rate of transcriptional elongation was at least 0.5 nucleotide per second. A rate of 7 nucleotides per second was reported in a purified system containing calf thymus RNA polymerase II (37).

### DISCUSSION

A number of eucaryotic transcription units have recently been shown to contain control elements downstream of the RNA synthesis initiation site (2, 3, 7, 15, 16, 22, 34, 40, 42, 44, 52, 56, 59). In the experiments presented here, we have shown that HeLa cell nuclei contain a protein, DAP, that binds to the downstream element of the SV40 major late promoter. The protein does not bind to promoter mutants lacking the downstream element. The isolation of a DNA-binding protein specific for the downstream element argues that the element is recognized at the DNA level. Recognition of the element at the DNA level is consistent with the observed effect of the element on preinitiation complex formation and with previous experiments ruling out mechanisms related to changes in RNA stability and processing (3).

Although isolated DAP binds specifically to its recognition site, the interaction is relatively weak. The binding activity was not detected in crude fractions, perhaps because of interference from nonspecific DNA-binding proteins. Quantitative binding experiments with isolated protein suggested that DAP bound to the downstream element with a  $K_d$  of approximately  $10^{-8}$  M. For comparison, other promoter-selective transcription factors bind from 10 to 1,000 times more tightly under similar conditions. Sp1 binds to a consensus recognition site with a  $K_d$  of  $5 \times 10^{-10}$  M (41), major late transcription factor binds to a site in the adenovirus major late promoter with a  $K_d$  of  $10^{-10}$  to  $10^{-12}$  M (14), NF-1 binds to a site in adenovirus with a  $K_d$  of  $2.1 \times 10^{-11}$  M (49), and the *Drosophila* heat shock transcription factor binds to a site in a heat shock promoter with a  $K_d$  of  $4 \times 10^{-12}$  M (60).

The weak affinity of DAP for its cognate site is reminiscent of the interaction between the general transcription factor TFIID and TATA box-containing promoters. Like DAP, binding of TFIID to its recognition site is not detected until the protein has been partially purified, and yeast TFIID has recently been shown to bind the adenovirus major late promoter with a  $K_d$  of  $2 \times 10^{-9}$  M (30). Binding of TFIID appears to be stabilized by other components of the transcription complex (50). The same may be true of DAP. We were unable to construct fully functional promoters where the downstream element was inverted or shifted from its normal location (data not shown), suggesting that DAP may occupy a fixed position in the preinitiation complex. Further purification of the other proteins required for initiation at the SV40 major late promoter will be required in order to test the possibility that these proteins stabilize DAP binding.

Attempts to demonstrate a transcription activation function for DAP directly were not successful. The primary difficulty was in obtaining an extract that was selectively depleted of DAP activity. Oligodeoxynucleotides that represent the DAP-binding site did not inhibit transcription when added as a competitor (data not shown). It is possible that the binding of DAP to its cognate site is too weak or that the protein is too abundant to allow quantitative depletion. Addition of purified DAP to nondepleted extracts had no effect on transcription (data not shown).

The primary effect of mutations in the downstream element was on preinitiation complex assembly. The final number of preinitiation complexes was lower with the LS24 mutant than with wild-type templates. The difference does not appear to be attributable to a slower intrinsic rate of complex formation on the mutant template. The time required for half the final number of complexes to form was not significantly different for wild-type and mutant DNAs (Fig. 6). Neither does the difference appear to be attributable to lower stability of the complexes formed on the mutant template, as competition experiments suggested that complex formation was essentially irreversible on both templates under the conditions used (data not shown). A plausible way to reconcile these observations is to suppose that mutation of the downstream element increases the fraction of preinitiation complexes that are either partially assembled or misassembled and thus defective for transcription. Such complexes either fail to initiate at all or produce abortive transcripts too short to be detected.

Although abortive initiation has been most thoroughly studied with procaryotic RNA polymerases, this reaction has been shown to occur with eucaryotic RNA polymerase II as well (43). For example, a recent study showed that the incidence of abortive initiation in vitro at the adenovirus major late promoter may be as high as 50 to 80%. It is interesting to speculate that DAP might influence the balance between abortive and productive initiation at the SV40 late promoter. In our previous study, we failed to detect an increase in short RNA with templates carrying mutations in the downstream element. However, these experiments might not have detected RNAs that were only a few bases long.

At present, it is not known whether binding sites for DAP occur in promoters for cellular genes. Farnham and Means (22) have identified several sequences in cellular promoters that are similar to the downstream element of SV40. One of these, CCCGCTGCCA (matches to SV40 underlined), which lies shortly downstream of the mouse dihydrofolate reductase gene initiation site, is in a region required for promoter activity and has been shown to bind a cellular protein (22). Another sequence, GTGGCTGCCA, which lies downstream of the mouse rpL32 gene initiation site, is also in a region required for promoter activity and also has been shown to bind a cellular protein (2). Both of these sequences are very strong candidates for interaction with DAP. Another possible candidate for interaction with DAP is the sequence TGTCAGTCCTG, which occurs downstream of the mouse primase gene initiation site (46). It has not been reported whether this sequence is required for transcription.

It is possible that the relatively weak interaction between DAP and the downstream element of the SV40 major late promoter contributes to physiological regulation of viral transcription. It is believed that template amplification plays a role in the increase in late transcription that occurs after replication, although other factors, such as T antigen transactivation, are also involved (9, 10, 39, 48, 57). One may

imagine a role for DAP as follows. Early in infection, there are few template molecules. Because of the weak binding affinity, only a fraction of the template molecules are bound by DAP, restricting the level of late-gene expression. Assuming that free DAP is not exhausted, as DNA is replicated, the fraction of the template bound by DAP remains constant, but there is an increase in the number of templates with DAP bound and a proportionate increase in the level of total late transcription. This predicted *in vivo* behavior is similar to the behavior seen in the *in vitro* DNA titration experiment shown in Fig. 5A. Provided that factors such as T antigen autoregulation limit the increase in early transcription, the model predicts that a switch from predominantly early-gene to predominantly late-gene expression will occur as a direct consequence of template amplification.

#### ACKNOWLEDGMENTS

We thank Kip Murphy and James McSwiggen for assistance in fitting theoretical curves to experimental data, P. Farnham for access to unpublished data, and R.-B. Markowitz for extensive comments on the manuscript.

This work was supported by Public Health Service grant GM-35866 from the National Institutes of Health.

#### LITERATURE CITED

- Arias, J. A., and W. S. Dynan. 1989. Promoter-dependent transcription by RNA polymerase II using immobilized enzyme complexes. *J. Biol. Chem.* **264**:3223-3229.
- Atchison, M. L., O. Meyuhos, and R. B. Perry. 1989. Localization of transcriptional regulatory elements and nuclear factor-binding sites in mouse ribosomal protein gene rpl32. *Mol. Cell. Biol.* **9**:2067-2074.
- Ayer, D. E., and W. S. Dynan. 1988. Simian virus 40 major late promoter: a novel tripartite structure that includes intragenic sequences. *Mol. Cell. Biol.* **8**:2021-2033.
- Barkan, A., and J. E. Mertz. 1981. DNA sequence analysis of simian virus 40 mutants with deletions mapping in the leader region of late viral mRNAs: mutants with deletions similar in size and position exhibit varied phenotypes. *J. Virol.* **37**:730-737.
- Berkhout, B., R. H. Silverman, and K.-T. Jeang. 1989. Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell* **59**:273-282.
- Bevington, P. R. 1969. Data reduction and error analysis for the physical sciences. McGraw-Hill, New York.
- Bourachot, B., M. Yaniv, and P. Herbomel. 1989. Control elements situated downstream of the major transcriptional start site are sufficient for highly efficient polyomavirus late transcription. *J. Virol.* **63**:2567-2577.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**:248-254.
- Brady, J., J. B. Bolen, M. Radonovich, N. Salzman, and G. Khoury. 1984. Stimulation of simian virus 40 late gene expression by simian virus 40 tumor antigen. *Proc. Natl. Acad. Sci. USA* **81**:2040-2044.
- Brady, J., and G. Khoury. 1985. *trans*-Activation of the simian virus 40 late transcription unit by T-antigen. *Mol. Cell. Biol.* **5**:1391-1399.
- Brady, J., M. Radonovich, M. Vodkin, V. Natarajan, M. Thoren, G. Das, J. Janik, and N. P. Salzman. 1982. Site-specific base substitutions and deletions that enhance or suppress transcription of the SV40 major late RNA. *Cell* **31**:625-633.
- Buratowski, S., S. Hahn, L. Guarente, and P. A. Sharp. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**:549-561.
- Cai, A., and D. S. Luse. 1987. Variations in template protection by the RNA polymerase II transcription complex during the initiation process. *Mol. Cell. Biol.* **7**:3371-3379.
- Chodosh, C. A., R. W. Carthew, and P. A. Sharp. 1986. A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol. Cell. Biol.* **6**:4723-4733.
- Coen, D. M., S. P. Weinheimer, and S. L. McKnight. 1986. A genetic approach to promoter recognition during *trans*-induction of viral gene expression. *Science* **234**:53-59.
- Cohen, R. B., L. Yang, J. A. Thompson, and B. Safer. 1988. Identification of a downstream sequence and binding protein that regulate adenovirus major late promoter transcription *in vitro*. *J. Biol. Chem.* **263**:10377-10385.
- Davison, B. L., J.-M. Egly, E. R. Mulvihill, and P. Chambon. 1983. Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. *Nature (London)* **301**:680-686.
- Davison, B. L., T. Leighton, and J. C. Rabinowitz. 1979. Purification of *Bacillus subtilis* RNA polymerase with heparin-agarose. *J. Biol. Chem.* **254**:9220-9226.
- Dynan, W. S. 1986. Promoters for housekeeping genes. *Trends Genet.* **2**:196-197.
- Dynan, W. S. 1987. DNase I footprinting as an assay for transcriptional regulatory proteins, p. 75-87. *In* J. K. Setlow and A. Hollaender (ed.), *Genetic engineering: principles and methods*, vol. 9. Plenum Publishing Corp., New York.
- Dynan, W. S., and S. A. Chervitz. 1989. Characterization of minimal simian virus 40 promoter: enhancer elements in the 72-base-pair repeats not required. *J. Virol.* **63**:1420-1427.
- Farnham, P. J., and A. L. Means. 1990. Sequences downstream of the transcription initiation site modulate the activity of the murine dihydrofolate reductase promoter. *Mol. Cell. Biol.* **10**:1390-1398.
- Fire, A., M. Samuels, and P. A. Sharp. 1984. Interactions between RNA polymerase II, factors, and templates leading to accurate transcription. *J. Biol. Chem.* **259**:2509-2516.
- Fried, M., and D. M. Crothers. 1981. Equilibrium and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**:6505-6525.
- Galas, D. J., and A. Schmitz. 1978. DNase footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* **5**:3157-3170.
- Gardiner-Garden, M., and M. Frommer. 1987. CpG islands in vertebrate genomes. *J. Mol. Biol.* **196**:261-282.
- Garner, M. M., and A. Revzin. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* **9**:3047-3060.
- Ghosh, P. K., V. B. Reddy, J. Swinscoe, P. V. Choudary, P. Lebowitz, and S. M. Weissman. 1978. The 5'-terminal leader sequence of 16S mRNA from cells infected with simian virus 40. *J. Biol. Chem.* **253**:3643-3647.
- Ghosh, P. K., V. B. Reddy, J. Swinscoe, P. Lebowitz, and S. M. Weissman. 1978. Heterogeneity and 5'-terminal structures of the late RNAs of simian virus 40. *J. Mol. Biol.* **126**:813-846.
- 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.
- Hai, T., M. Horikoshi, R. G. Roeder, and M. R. Green. 1988. Analysis of the role of the transcription factor ATF in the assembly of a functional preinitiation complex. *Cell* **54**:1042-1051.
- Hansen, U., and P. A. Sharp. 1983. Sequences controlling *in vitro* transcription of SV40 promoters. *EMBO J.* **2**:2293-2303.
- Hawley, D. K., and R. G. Roeder. 1985. Separation and partial characterization of three functional steps in transcription initiation by human RNA polymerase II. *J. Biol. Chem.* **260**:8163-8172.
- Homa, F. L., T. M. Otal, J. C. Glorioso, and M. Levine. 1986. Transcriptional signals of a herpes simplex virus type 1 late ( $\gamma$ 2) gene lie within bases -34 to +124 relative to the 5' terminus of the mRNA. *Mol. Cell. Biol.* **6**:3562-3666.
- Horikoshi, M., M. F. Carey, H. Kakidani, and R. G. Roeder. 1988. Mechanism of activation of a yeast activator: direct effect

- of GAL4 derivatives on mammalian TFIID-promoter interactions. *Cell* **54**:665-669.
36. Horikoshi, M., T. Hai, Y. Lin, M. R. Green, and R. G. Roeder. 1988. Transcription factor ATF interacts with the TATA factor to facilitate the establishment of a preinitiation complex. *Cell* **54**:1033-1042.
  37. Kadesch, T. R., and M. J. Chamberlin. 1982. Studies of in vitro transcription by calf thymus RNA polymerase II using a novel duplex DNA template. *J. Biol. Chem.* **257**:5286-5295.
  38. Kadonaga, J. T., and R. Tjian. 1986. Affinity purification of sequence-specific DNA-binding proteins. *Proc. Natl. Acad. Sci. USA* **84**:3653-3657.
  39. Keller, J. M., and J. C. Alwine. 1984. Activation of the SV40 late promoter: direct effects of T antigen in the absence of viral replication. *Cell* **36**:381-389.
  40. Leong, K., W. Lee, and A. J. Berk. 1990. High-level transcription from the adenovirus major late promoter requires downstream binding sites for late-phase-specific factors. *J. Virol.* **64**:51-60.
  41. Letovsky, J., and W. S. Dynan. 1989. Measurement of the binding of the transcription factor Sp1 to a single GC box recognition sequence. *Nucleic Acids Res.* **17**:2639-2653.
  42. Lewis, E. D., and J. L. Manley. 1985. Control of adenovirus late promoter expression in two human cell lines. *Mol. Cell. Biol.* **5**:2433-2442.
  43. Luse, D. S., and G. A. Jacob. 1987. Abortive initiation by RNA polymerase II in vitro at the adenovirus 2 major late promoter. *J. Biol. Chem.* **262**:14990-14997.
  44. Mansour, S. L., T. Grodziker, and R. Tjian. 1986. Downstream sequences affect transcriptional initiation from the adenovirus major late promoter. *Mol. Cell. Biol.* **6**:2684-2694.
  - 44a. Maxam, A., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499-560.
  45. Piatak, M., P. K. Ghosh, L. C. Norkin, and S. M. Weissman. 1983. Sequences locating the 5' ends of the major simian virus 40 mRNA late forms. *J. Virol.* **48**:503-520.
  46. Prussak, C. E., M. T. Almazan, and B. Y. Tseng. 1989. Mouse primase p49 subunit molecular cloning indicates conserved and divergent regions. *J. Biol. Chem.* **264**:4957-4963.
  47. Reinberg, D., M. Horikoshi, and R. G. Roeder. 1987. Factors involved in specific transcription by mammalian RNA polymerase II: functional analysis of initiation factors IIA and IID and identification of a new factor operating at sequences downstream of the initiation site. *J. Biol. Chem.* **262**:3322-3330.
  48. Rio, D., A. Robbins, R. Myers, and R. Tjian. 1980. Regulation of simian virus 40 early transcription in vitro by a purified tumor antigen. *Proc. Natl. Acad. Sci. USA* **77**:5706-5710.
  49. Rosenfeld, P. J., and T. S. Kelly. 1986. Purification of nuclear factor 1 by DNA recognition site affinity chromatography. *J. Biol. Chem.* **261**:1398-1408.
  50. Sawadogo, M., and R. G. Roeder. 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* **43**:165-175.
  51. Sharp, P. A., and R. A. Marciniak. 1989. HIV TAR: an RNA enhancer. *Cell* **59**:229-230.
  52. Sinclair, J. H., J. F. Burke, D. Ish-Horowitz, and J. H. Sang. 1986. Functional analysis of the transcriptional control regions of the *copia* transposable element. *EMBO J.* **5**:2349-2354.
  53. Smale, S. T., and D. Baltimore. 1989. The initiator as a transcription control element. *Cell* **57**:103-113.
  54. Somasekhar, M. B., and J. E. Mertz. 1985. Sequences involved in determining the locations of the 5' ends of the late RNAs of simian virus 40. *J. Virol.* **56**:1002-1013.
  55. Stenlund, A., G. Bream, and M. R. Botchan. 1987. A promoter with an internal regulatory domain is part of the origin of replication in BPV-1. *Science* **236**:1666-1671.
  56. Thummel, C. S. 1989. The *Drosophila E74* promoter contains essential sequences downstream from the start site of transcription. *Genes Dev.* **3**:782-792.
  57. Tjian, R. 1981. T antigen binding and the control of SV40 gene expression. *Cell* **26**:1-2.
  58. Van Dyke, M. W., R. G. Roeder, and M. Sawadogo. 1988. Physical analysis of transcription complex assembly on class II gene promoters. *Science* **241**:1335-1338.
  59. Wright, S., A. Rosenthal, R. Flavell, and F. Grosfeld. 1984. DNA sequences required for expression of  $\beta$ -globin genes in murine erythroleukemia cells. *Cell* **38**:265-273.
  60. Wu, C., S. Wilson, B. Walker, I. David, T. Paisley, V. Zimarino, and H. Ueda. 1987. Purification and properties of *Drosophila* heat shock protein. *Science* **238**:1247-1253.