MOLECULAR AND CELLULAR BIOLOGY, Nov. 1983, p. 1975-1984
0270-7306/83/111975-10$02.00/0

Vol. 3, No. 11

An Upstream Regulatory Domain of Avian Tumor Virus Long Terminal Repeat Is Required for the Expression of a Procaryotic Neomycin Gene in Eucaryotic Cells

S. ALEX MITSIALIS,† SHARI CAPLAN, AND RAMAREDDY V. GUNTAKA*‡

Department of Microbiology, Columbia University College of Physicians and Surgeons, New York, New York 10032

Received 4 March 1983/Accepted 29 August 1983

The regulatory elements present in the long terminal repeat (LTR) of avian sarcoma virus DNA were analyzed by recombinant DNA techniques coupled with DNA-mediated gene transfer in avian as well as mammalian cells. For this purpose, the neomycin resistance gene from transposon Tn*5* was inserted downstream from the avian sarcoma virus LTR, and the recombinant plasmid DNA was introduced into cells by the calcium phosphate technique. Cells resistant to the drug G-418 were selected. Analysis of the RNA transcripts made in vivo in these transformants indicated that initiation and termination of the transcripts occurred in the LTR sequences. Deletions were then introduced into the LTR, and their effect on transcription was also studied. These results allowed us to identify a strong regulatory sequence between nucleotides −299 and −114 in the LTR of avian sarcoma virus.

During infection by retroviruses, genomic RNA is converted into DNA by the virion-associated reverse transcriptase. In this process, two long terminal repeats (LTRs) containing nucleotide sequences derived from both 5′ (U5) and 3′ (U3) sequences of the template RNA are generated at both ends of the viral DNA (13, 24). This viral DNA is then integrated into host chromosomes in a unique arrangement so that the LTRs flank the structural viral genes of the provirus (25). The integrated provirus then serves as a template for viral RNA transcription, and it appears that the LTRs contain the required transcriptional elements, including signals implicated in the initiation and termination of transcription (25). In vitro transcription studies with cloned recombinant DNAs have revealed that the LTRs of murine (6) as well as avian (29) retroviruses contain sequences which can serve as efficient promoters for RNA polymerase II in cell-free lysates and that these sequences are localized between nucleotides −55 and +19 (with respect to the cap or transcription initiation site, which is +1) of the LTR (20). The LTR sequences have been shown to activate expression of heterologous coding sequences in tissue culture (7, 15, 17) as well as in tumors in which the LTRs activate cellular oncogenes (22, 23).

Recently, we have extended our in vitro studies (20) on the functional domains of the avian sarcoma virus (ASV) promoter region to in vivo systems by transferring the genes into eucaryotic cells using the neomycin resistance (Neo^r) gene as a selectable marker. The selection used was resistance to the aminoglycoside G-418, which has inhibitory activity against a wide range of procaryotic and eucaryotic organisms (14; P. J. L. Daniels, A. S. Yehaskel, and J. B. Morton, Program Abstr. Intersci. Conf. Antimicrob. Agents Chemother. 13th, Washington, D.C., abstr. no. 137, 1973). Colbere-Garapin et al. (4) have shown that a variety of tissue culture cell lines are sensitive to low concentrations of G-418. Placing the Neo^r gene from transposon Tn*5* under the control of the herpesvirus I thymidine kinase gene promoter, they were able to demonstrate expression of the Neo^r gene in transfected eucaryotic cells. In our studies, recombinant plasmid DNAs containing the structural gene for Neo^r and ASV promoter were introduced into tissue culture cells. Analyses of the Neo^r gene-specific transcripts in individual clones of G-418-resistant transformants permitted us to localize in vivo the transcriptional initiation site in the ATV LTR. In addition, our results indicated the presence of a second element in the ASV LTR, the 5′ boundary of which mapped upstream of position −114 in the LTR. This upstream element, although dispensable for efficient initiation of transcription in vitro, is indispensable for expression of genes

† Present address: Biological Laboratories, Harvard University, Cambridge, MA 02138.

‡ Current address: Department of Microbiology, School of Medicine, University of Missouri, Columbia, MO 65212.

under the control of the ASV LTR in vivo. This report will present these results.

MATERIALS AND METHODS

Restriction enzymes, *Escherichia coli* DNA polymerase I, T4 polynucleotide kinase, and bacterial alkaline phosphatase were obtained from Bethesda Research Laboratories, New England Biolabs, or Boehringer-Mannheim and were used as recommended by the manufacturers.

Radiolabeled reagents were obtained from New England Nuclear Corp., and nucleotide triphosphates were purchased from P-L Biochemicals. Glyoxal (Eastman Kodak) and formamide (Mallinckrodt) were deionized with mixed bed resin RG-501-X8, 20 to 50 mesh (Bio-Rad Laboratories) before use.

Plasmids. Construction and characterization of recombinant plasmids containing the avian tumor virus promoter and Neo genes have been described (21).

S1 nuclease mapping. The technique of Berk and Sharp (2) was used, with minor modifications, as described previously (20).

Transfections. The rat fibrosarcoma cell line XC and the quail line QT6 were used as recipient cells in most of these studies. The calcium phosphate coprecipitate method of Graham and van der Eb (8) with modifications (5, 28) was used. Two days before transfection, cultures were seeded at a density of 2×10^6 cells per 100-mm dish and were incubated in the normal growth media. Two hours before DNA addition, the cultures received fresh medium. In a typical experiment, 40 to 200 ng of plasmid DNA was mixed with 20 μg of carrier calf thymus DNA in 0.5 ml of HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid)-buffered saline (0.14 M NaCl, 5 mM KCl, 0.7 mM $Na_2HPO_2 \cdot 7H_2O$, 20 mM HEPES [pH 7.0], 0.1% glucose). $CaCl_2$ (50 μl of an 1.25 M solution) was added, and the fine precipitate was allowed to develop at room temperature for 20 min. The suspension was added to the culture, and the cells were incubated for 6 h under normal growth conditions. The media were then aspirated, and the monolayer in each original dish was trypsinized and seeded in the regular growth media into five new dishes. After 16 to 20 h, the cells were fed with fresh medium containing 0.22 mg of G-418 (a generous gift of P. L. Daniels, Schering Co.) per ml. From this time on, the cells were continuously exposed to the drug and were fed with fresh medium at intervals of 3 to 4 days. G-418-resistant colonies were scored 15 to 18 days after the onset of selection.

For mass propagation of individual transformants, well-isolated colonies were trypsinized in a 4-mm cloning cylinder attached to the surface of the dish with silicone grease (Dow-Corning). The cells were seeded in a 30-mm dish and propagated in the presence of 0.22 mg of G-418 per ml until confluent. Subsequently, they were transferred to 100-mm dishes for bulk culture.

Southern blot analysis of DNA. High-molecular-weight DNA was isolated, restricted with appropriate restriction endonucleases, and electrophoresed on 1% agarose slab gels as described before (12). The DNA was transferred to nitrocellulose filters (12). Hybridizations were carried out at 68°C for 16 to 20 h with appropriately labeled probes essentially as described (12).

RESULTS

Construction of deletion mutants and transfection of animal cells. Construction of recombinant plasmids bearing avian tumor virus DNA fragments has been described (11, 21). The plasmids pertinent to this study are shown schematically in Fig. 1. Plasmid pATV-6D3A (Fig. 1) contains two inserts of ASV DNA flanking the structural gene for Neo^r. The viral inserts included nucleotides between positions −299 and +44 (with respect to the ASV RNA cap site, which is defined as +1). Transcription originating in the upstream (left) insert should proceed into the Neo^r gene, and transcription originating in the downstream (right) insert should proceed into pBR322 sequences. Other recombinants of the pATV-6D3A prototype which were used in this study contain 3′ deletions in the upstream viral insert; the right endpoints of these deletions are +37 in pATV-6exo20A, +33 in pATV-6exo11A, +20 in pATV-6exo15A, and −23 in pATV-6exo28R (21). For example, pATV-6exo15A contains the sequence between −299 and +20, whereas pATV-6exo28R contains the sequence between −299 and −23. In the latter clone, the TATA sequence, but not the cap site, is present. pATV-6D3A*dl*1 was constructed by deleting the sequences between the *Eco*RI sites at positions −299 and −55 in the upstream viral insert of pATV-6D3A. pATV-6D3A*dl*2 was similarly constructed by deleting the sequences between the *Pvu*I site at position −114 in the upstream viral insert and the *Pvu*I site in the ampicillin resistance (Ap^r) gene of pBR322. pATV-6.49Nm1 was constructed by inserting the *Eco*RI fragment containing the Neo^r gene from pATV-6D3A (see Fig. 1) into the *Eco*RI site of pATV-6.49, the viral insert which contains U3 sequences between positions −86 and +44 (21). The above constructions have created plasmids in which the Neo^r gene is directly under the control of viral LTR sequences ranging from position +44 to positions −55 (pATV-6D3A *dl*1), −86 (pATV-6.49Nm1), or −114 (pATV-6D3A*dl*2) and position −299 (pATV-6D3A).

Plasmid DNAs from the above described recombinants were introduced into tissue culture cells by the calcium phosphate coprecipitation technique (8). G-418-resistant transformants were selected in media containing 0.22 mg of G-418 per ml. The rat fibrosarcoma line XC and the quail line QT6 (12) were the recipients in most of the experiments, but similar results were obtained with other cell lines such as mouse NIH 3T3, quail l6Q, human KB, mouse teratocarcinoma (C. Nguyen-Huu, unpublished results), and hamster H9 cells. At the concentration of G-418 used, cell growth was arrested within 2 to 3 days, and no surviving cells were present after 14 to 15 days of exposure. DNA from plasmid

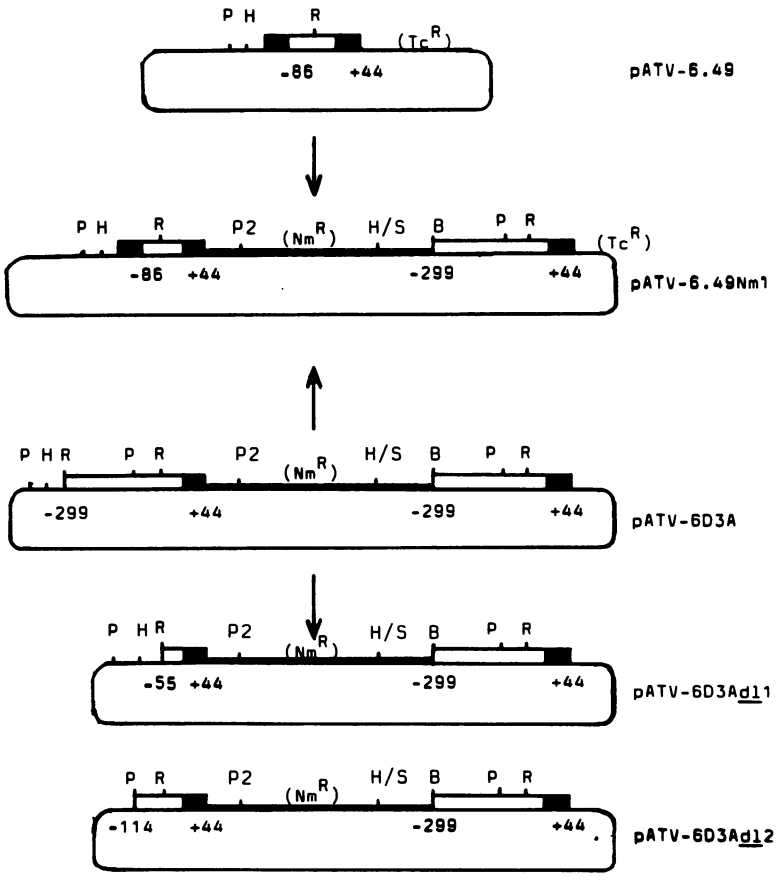

FIG. 1. Schematic representations of recombinant clones. Construction of pATV-6.49 Neo^r clone from pATV-6.49 was accomplished by inserting an *Eco*RI fragment containing part of LTRs from pATV-6D3A. pATV-6D3A*dl*1 was obtained by deleting the 245-base pair (bp) *Eco*RI fragment (−299 to −55) from pATV-6D3A. pATV-6D3*dl*2 was obtained by deleting a *Pvu*I frament (−114 in LTR to the *Pvu*I site in the pBR322 ampicillin gene). R, *Eco*RI; P, *Pvu*I; H, *Hin*cII. Thick line represents Neo sequences from Tn*5*; thin line indicates pBR322 sequences. Solid and open boxes represent U5 and U3 sequences, respectively, of LTR. It should be noted that the figures were drawn to indicate the position of the LTR sequences. pBR sequences were not drawn to indicate their sizes.

pKC56, which contains the structural gene for Neo^r without a promoter (21), produced no G-418-resistant colonies in dosages up to 2 μg of DNA per tissue culture dish (5×10^6 cells). In contrast, plasmid DNAs from pATV-6D3A and its 3′ deletion derivatives yielded, under the same conditions, a significant number of transformants (data not shown). The transformation efficiencies of all the 3′ deletions were comparable; it is of particular interest that plasmid pATV-6exo28R, which suffered a 3′ deletion downstream from −23 in the upstream viral insert, exhibited also a transformation efficiency comparable to that of the other plasmids. The efficiency of transfection by pATV-6exo28R DNA was compared with that of pATV-6exo11A DNA. This comparison was made because pATV-6exo11A DNA supports in vitro transcription at a rate at least 10- to 20-fold more than pATV-6exo28R DNA. The results (Fig. 2) indicate that the two recombinants exhibited, at the most, a twofold difference in the yield of G-418-resistant transformants.

Individual colonies were selected at random and propagated in mass culture as described above. The overwhelming majority of these clones were found to be stable during long-term propagation in the presence of 0.22 mg of G-418 per ml. The stability of the resistant phenotype in the absence of selection was examined for a number of clones. For this analysis, cells were propagated in media without G-418 for several passages, then an equal number of cells from each clone were plated in a number of dishes. Half of the dishes received regular medium, and the other half received medium containing G-418

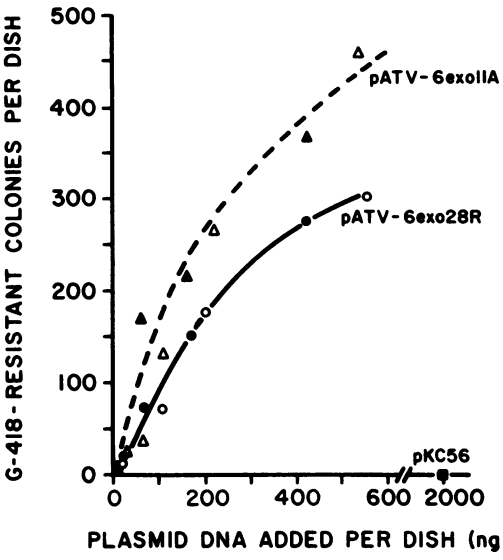

FIG. 2. Kinetics of transfection by pATV-6exo11A and pATV-6exo28R DNA. Tissue culture cell lines were seeded at a density of 2×10^6 cells per 100-mm dish and were incubated under regular growth conditions for 2 days; during this time, the number of cells increased to approximately 5×10^6 per dish. Supercoiled pATV-6exo11A DNA was then introduced as a coprecipitate with calcium phosphate, as detailed in the text. After 6 h of incubation, the monolayers were trypsinized, and the cells in each original dish were seeded to five dishes (diameter, 100 mm) and were incubated under normal growth conditions for 16 to 18 h. After that time, the cells were continuously exposed to media which contained 0.22 mg of G-418 per ml. Resistant colonies were counted 15 to 18 days after the onset of selection. The total number of G-418-resistant colonies arising in the quail line is plotted against the concentration of DNA added per original dish. Results are shown for QT6 (△) and XC (▲) cells receiving pATV-6exo11A DNA and QT6 (○) and XC (●) cells receiving pATV-6exo28R DNA.

at various concentrations. Surviving colonies were scored after 15 to 18 days. The results (Table 1) clearly show that the resistant phenotype was stable, indicating that selective pressure is not required to maintain the Neo^r gene.

Integration. The state of the plasmid DNA in the transformants was analyzed by Southern blot analysis of total genomic DNA isolated from individual G-418-resistant clones. The results obtained with a number of drug-resistant quail (QT6) and rat (XC) clones, which were transfected by *Sal*I-linearized plasmid DNA, indicated that as expected, the recombinational events involved in the integration of the linearized plasmid DNA had occurred at or close to the ends of the molecule (data not shown).

Representative data on the integration of plasmid DNA in cells transfected by supercoiled plasmid DNA are shown in Fig. 3. Total cellular DNA from G-418-resistant clones was digested with *Eco*RI (Fig. 3, lanes 2 through 10) or *Hin*cII (lanes 11 through 19) and electrophoresed in duplicate, as above. The blot of one set was hybridized with a Neo^r gene probe (Fig. 3A), and the blot of the other set was hybridized with a plasmid pBR322 probe (Fig. 3B). No specificity of integration on plasmid DNA could be detected. The 1.8-kilobase (kb) *Eco*RI fragment, which contains the Neo^r gene sequences, was found to be intact in the majority of the clones analyzed (Fig. 3A, lanes 2 through 10), indicating that integration of the transfecting plasmid DNA occurred in pBR322 sequences. In certain cases, as in the quail clones QD3C1 (Fig. 3, lanes 2) and Q15C3 (lanes 4), integration appears to have occurred within Tn*5* sequences. Multiple copies of transfecting plasmid were also evident in the rat clones X11C6, X11C7, and X15C2 (lanes 8 through 10 and 17 through 19), and in these clones certain plasmid DNA molecules appear to have integrated in pBR322 sequences and others in Tn*5* sequences, as evidenced by the appearance of different junction fragments (Fig. 3B).

Of particular importance are the results of *Hin*cII digestion of genomic DNA from transformed cells (Fig. 3A, lanes 11 through 19). This enzyme generates a fragment of 1.9 kb from plasmid DNA, which includes the upstream LTR (see Fig. 1). Integration at or near the LTR

TABLE 1. Stability of G-418-resistant phenotype of representative clones[a]

Clone[b]	Propagation without selection (wk)	Colonies surviving in G-418 at concn (mg/ml):		
		0	0.22	0.66
QT6		356	0	
Q11.2	9	550	670	ND[c]
Q11C5	6	152	158	ND
X11C6	4	410	294	335
X28C2	7	398	426	289

[a] Individual G-418-resistant colonies were cloned, and the cells were propagated for two or three passages in mass culture in media containing 0.22 mg of G-418 per ml. The cultures were then grown in regular media, without the drug, for the times indicated. The cultures were routinely passaged every 4 to 5 days. To test the stability of the phenotype, equal numbers of cells from each clone were seeded in several dishes and incubated in media containing G-418 at the concentrations indicated. G-418-resistant colonies were counted 18 days after the onset of selection.

[b] QT6, Parental quail cell line; Q11.2 and Q11C5, QT6 derivatives transformed with pATV-6exo11A DNA; X28C2, XC line derivative transformed with pATV-6exo28R DNA.

[c] ND, Not done.

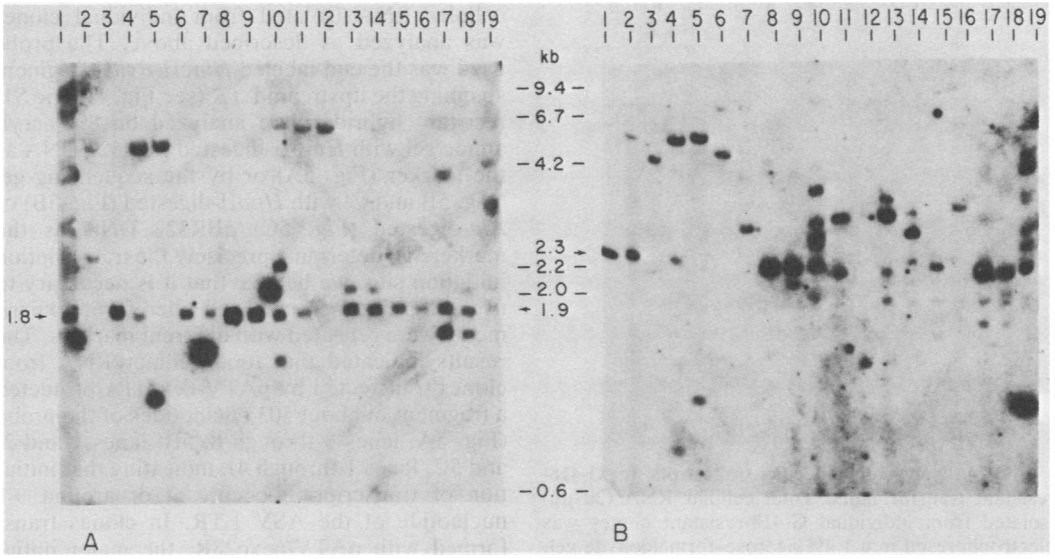

FIG. 3. Analysis of plasmid sequences in the genomes of clones transformed with supercoiled plasmid DNA. Total cellular DNA from individual G-418-resistant clones was analyzed by Southern blotting as described in the text. Two identical blots were hybridized with a nick-translated *Bgl*II-*Bam*HI fragment of the Neo^r gene (A) or with nick-translated pBR322 DNA (B) DNA from each clone (20 μg) was digested with *Eco*RI (lanes 2 through 10); an additional 20-μg DNA sample from the same clones was digested with *Hin*cII (lanes 11 through 19). Lanes 2 through 6 and 11 through 15: DNA from clones derived from the QT6 cell line by transformation with supercoiled DNA from pATV-6D3A (clone QD3C1, lanes 2 and 11), pATV-6exo11A (clone Q11C7, lanes 3 and 12), pATV-6exo15A (clones Q15C3 and Q15C4, lanes 4 and 13 and 5 and 14, respectively), and pATV-6exo20A (clone Q20C1, lanes 7 and 16). Lanes 7 through 10 and 15 through 19: DNA from clones derived from the XC cell line by transformation with supercoiled DNA from pATV-6D3A (clone XD3C1, lanes 7 and 16), pATV-6exo11A (clones X11C6 and X11C7, lanes 8 and 17 and 9 and 18, respectively), and pATV-6exo15A (clone X15C2, lanes 10 and 19). Lane 1: 50 pg of pATV-6exo11A plasmid DNA digested with *Eco*RI.

should result in the elimination of this fragment from genomic DNA digests. As can be seen in Fig. 3A, lanes 11 through 19, a 1.9-kg fragment containing Neo^r gene sequences was present in all of the clones except one (lane 12). Since the 1.8-kb *Eco*RI fragment and 1.9-kb *Hin*cII fragment are intact in the integrated molecules, it is very likely that integration did not occur in the upstream LTR. However, the data do not rule out the involvement of the −55 to +44 sequence of the LTR, downstream to the Neo gene fragment, in integration.

Transcription of Neo^r gene-specific RNA. To study the nature of Neo^r gene-specific transcripts in the transformants, total cellular RNA was purified from individual clones and was analyzed by agarose gel electrophoresis followed by blotting and hybridization with Neo^r gene-specific probe, as described above. The results (Fig. 4) indicate that in a majority of the clones, a transcript of 2.0 kb was present. If transcription on the integrated plasmid DNA is under the control of the ASV promoter initiated in the upper viral insert and terminated in the downstream viral insert, a Neo^r gene-containing transcript of 1.8 kb is expected. The presence of a transcript of this size strongly suggests that both initiation and termination of transcription occurred within the viral inserts. The length of slightly over 1.8 kb would occur if the polyadenylation signal present in the LTR at positions −2 to −7 (see Fig. 1) is efficiently recognized by the appropriate enzymatic complexes, resulting in the addition of about 0.2 kb of polyadenylic acid at the 3′ terminus of the transcript. It should also be pointed out that in all clones examined, a minor transcript of about 0.9 kb was also observed. This, we believe, is due to termination in the Neo gene sequences, perhaps the termination site used in bacteria. In addition to the 2.0 kb transcript in some clones, other, longer, transcripts were also present. These could be due to termination in the adjacent cell sequences.

The level of Neo^r gene-specific transcripts seemed to vary between individual clones. No correlation seemed to be possible between the amount of Neo^r gene-specific transcripts and the permissiveness or nonpermissiveness to ASV of the recipient cells (in Fig. 4, lanes 1 through 5 represent permissive quail cells and lanes 6 through 9 represent nonpermissive rat cells). In

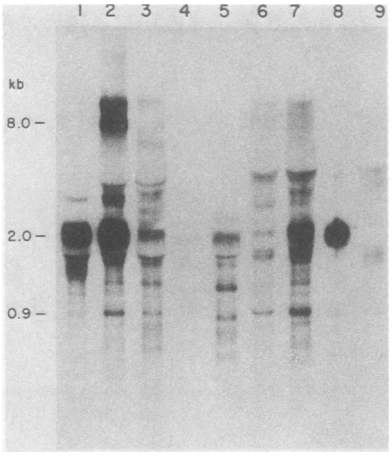

FIG. 4. Neo^r gene-specific transcripts in G-418-resistant transformants. Total cellular RNA (20 μg) isolated from individual G-418-resistant clones was electrophoresed in a 1.4% agarose–formaldehyde gel. Gels were routinely stained with ethidium bromide and photographed under UV light to visualize the rRNA bands. Equal intensity of the rRNA bands among samples was presumed to show that an equal amount of RNA was loaded in each lane and that no gross degradation had occurred. *E. coli* RNA was also run in parallel; the 16S and 23S procaryotic rRNA species, together with the 18S and 28S eucaryotic rRNA species, served as size markers. The gel was blotted onto nitrocellulose paper and was hybridized with the nick-translated *Bam*HI-*Bgl*II Neo^r gene fragment. Lanes 1 through 5: RNA from clones derived from the QT6 cell line by transformation with pATV-6exo11A DNA (lanes 1 through 3, clones Q11.2, Q11C5, and Q11C6, respectively) or with pATV-6exo28R DNA (lanes 4 and 5, clones Q28C1 and Q28C2). Lanes 6 through 9: RNA from clones derived from the XC cell line by transformation with pATV-6exo11A DNA (lanes 6, 7, and 8, clones X11C6, X11C7, and X11.2, respectively) or with pATV-6exo28R DNA (lane 9, clone X28C2).

the clones transformed by pATV-6exo28R, a 3′ deletion mutant of the upstream viral insert, the levels of the Neo^r transcripts were found to be in general less than those in the wild type (pATV-6D3A), as evidenced by the presence of only a faint band (Fig. 4, lanes 4 and 5). That this was a real transcript was indicated by the size of the S1-protected fragment (Fig. 5A, lanes 7 and 8). However, these levels appear to be sufficient to maintain the drug-resistant phenotype (Fig. 4, lanes 4 and 5, and pATV-6exo28R). Since some of the differences in the amount of Neo^r gene transcripts may have been copy number effects, no quantitative statement can about the levels of Neo^r-specific RNA be made in these studies.

S1 mapping of RNA. More specific analysis of the initiation site of the Neo^r gene transcript, which is under the control of the ASV promoter, involves S1 nuclease protection studies. Total cellular RNA isolated from individual clones was analyzed as described above. The probe used was the end-labeled *Hin*cII-*Pvu*II fragment spanning the upstream LTR (see Fig. 1). The S1-resistant hybrids were analyzed on 8% acrylamide gel with *Hac*III-digested pBR322 DNA as the marker (Fig. 5A) or by the sequencing gel (Fig. 5B and C) with *Hpa*II-digested (Fig. 5B) or *Alu*-digested (Fig. 5C) pBR322 DNA as the marker. To determine precisely the transcription initiation site, we believe that it is necessary to use different markers, and therefore, experiments were repeated with different markers. The results indicated that total cellular RNA from clones transfected by pATV-6exo11A protected a fragment of about 303 nucleotides of the probe (Fig. 5A, lanes 3 through 6, 5B, lanes 1 and 2, and 5C, lanes 1 through 4), indicating that initiation of transcription occurs at or around +1 nucleotide of the ASV LTR. In clones transformed with pATV-6exo28R, the major initiation point was about 20 nucleotides downstream from the TATA element of the ATV promoter Neo^r gene, as indicated by the protection of a 255-nucleotide fragment (Fig. 5A, lanes 7 through 9, 5B, lane 3, and 5C, lanes 4 through 6). Since the ^{32}P label was introduced at the *Pvu*II site in the Neo^r gene, which is located about 270 nucleotides downstream, the results indicate that transcription initiation occurs in the LTR sequences about 33 nucleotides upstream from the Neo^r insert, which is the +1 nucleotide of the ASV.

Requirement of upstream sequences. In vitro studies had suggested that LTR sequences upstream of position −55 had no effect on the activity of the ASV promoter in HeLa lysates (20). To investigate whether a similar conclusion can be drawn on the activity of the promoter in vivo, recombinant plasmids were constructed which carried the Neo^r gene under the control of LTR derivatives containing 5′ deletions (see Fig. 1), with the endpoint at position −55 (see Fig. 1). Transfection by these plasmid DNAs produced no G-418-resistant colonies (Table 2). These results were highly reproducible and were the same whether supercoiled DNA or linearized plasmid DNA was used. To localize the upstream sequence required for successful expression of the Neo gene, other 5′ deletions were constructed. The plasmid containing additional sequence between −114 and −55 also gave negative results (Table 2). These results strongly suggest the existence of an element of the ASV LTR, other than the "TATA" element, which is dispensable for efficient initiation of transcription in vitro but is nevertheless essential for in vivo expression of genes under the control of this promoter. This element appears to have a 5′ boundary upstream of position

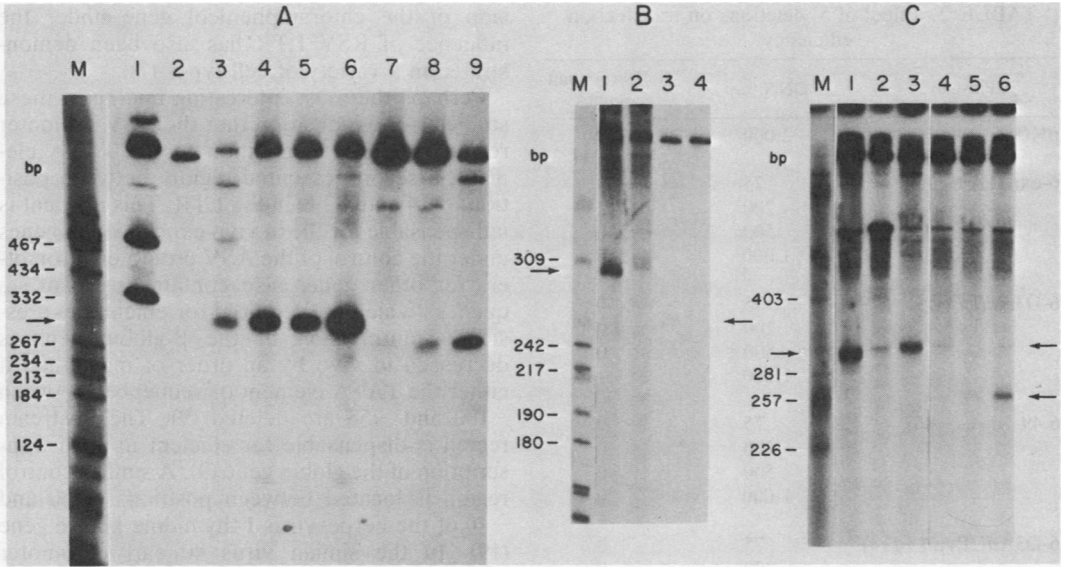

FIG. 5. (A) S1 nuclease mapping of the 5′ ends of Neo^r gene transcripts in G-418-resistant transformants. Total cellular RNA (20 μg) isolated from individual G-418-resistant transformants was analyzed under conditions described previously (20). Briefly, the *Hin*cII- *Pvu*II fragment was end labeled, and about 10^5 cpm was incubated with RNA in 15 μl of hybridization buffer (80% formamide, 0.4 M NaCl, 40 mM PIPES [piperazine-*N*, *N*′-bis(2-ethanesulfonic acid)], pH 7.0) and hybridized at 53 to 54°C for 2 h. Then, 0.25 ml of S1 buffer (0.03 M acetate buffer [pH 4.5], 0.3 M NaCl, 1 mM $ZnCl_2$, 5% glycerol) was added and incubated with 200 U of S1 nuclease for 45 min at 42°C. The hybrids were precipitated and analyzed on 8% acrylamide gel. The gel was exposed to X-ray film. Lane 1, *Hin*cII-*Pvu*II probe digested with *Eco*RI; lane 2, RNA from control QT6 cells; lanes 3 through 6, RNA from clones derived by transfection with pATV-6exo11A DNA of QT6 cells (lanes 3 and 4, clones Q11C6 and Q11C7) or XC cells (lanes 5 and 6, clones X11C6 and X11C7); lanes 7 through 9, RNA from clones derived by transfection with pATV-6exo28R DNA of QT6 cells (lanes 7 and 8, clones Q28C1 and Q28C2) or XC cells (lane 9, clone X28C2). (B) and (C) Sizes of S1-protected fragments as determined on 7.5% acrylamide–urea sequencing gels. RNAs from Q11C5 (lanes 1), Q11C6 (lanes 2), and Q28X1 (lanes 3 in B and 5 in C) were used to hybridize with the end-labeled *Hin*cII-*Pvu*II fragment as probe. Lane 4 in B contains *E. coli* RNA. Lanes 3, 4, and 6 in C contain S1-resistant hybrids from clones Q11C7, XC111, and Q28C2, respectively. Lanes M in B and C contain *Hpa*II- and *Alu*I-digested fragments, respectively, as markers.

−114, the left endpoint of the smallest LTR deletion used here, and −234, the 5′ limit of LTR sequences. Thus, at least part of the second regulatory domain is located between −234 and −114. Note that the recombinant plasmids used in this study have a sequence up to −299. However, in all the integrated proviruses, the left end boundary of the upstream LTR includes sequences up to −234 only (25). Therefore, the second upstream regulatory element may be located between −234 and −114 instead of between −299 and −114, although the influence of the sequence between −299 and −234 on transcription cannot be ruled out. To identify the exact upstream sequence required for gene expression, deletions from the 5′ end of the LTR, i.e., from −234 downstream, are necessary. These experiments are in progress.

To strengthen the argument that the effect of upstream deletion is on transcription, XC cells were treated with supercoiled DNA from the wild type (pATV-6D3A) or from a deletion mutant (pATV-6D3A*dl*1) for 36 h. RNA was isolated and analyzed for neomycin-specific transcripts by Northern blotting. The results (Fig. 6) indicate that under transient expression conditions there is no evidence for the synthesis of the 2.0-kb Neo-specific RNA in the deletion mutant (Fig. 6, lane 2). Two transcripts of 2.0 and 4.2 kb (lane 1) were observed in pATV-6D3A-transfected cells. The longer transcript might be due to the absence of termination in the LTR. In such case the entire plasmid DNA (4.2 kb) could be transcribed to give rise to this product. This view is partly supported by the observation that in the deletion mutant, initiation of transcription from supercoiled DNA can take place in the downstream LTR to give rise to the 4.2 kb transcript (lane 2). Since the 2.0-kb Neo-specific transcript is absent from the mutant, we can conclude that deletion of the upstream sequence affects initiation of transcription, further supporting our contention that the upstream sequence is required for gene expression.

TABLE 2. Effect of 5′ deletions on transfection efficiency[a]

Plasmid[b]	DNA (ng)	G-418-resistant colonies
pKC56	2,000	0
6-exo11A (−299)	75	63
	200	145
	500	250
	1,000	555
6-D3A*dl*1 (−55)	75	0
	200	0
	500	0
	1,000	1
6.49-Nm1 (−86)	75	0
	200	0
	500	0
	1,000	0
6-D3A*dl*2Pvu (−114)	75	0
	200	0
	500	0
	1,000	1

[a] XC cells were exposed to supercoiled plasmid DNAs at the concentrations indicated, and transformants were counted after 16 days.

[b] Numbers in parentheses define the 5′ limit of the U3 sequences in the upper viral insert.

DISCUSSION

The results presented here clearly demonstrate that the avian tumor virus LTR can efficiently drive expression of heterologous genes in a variety of eucaryotic cells. Introduction of recombinant plasmid DNAs containing ATV LTR and the Neo^r gene from transposon Tn*5* into tissue culture cells results in the appearance of G-418-resistant colonies. In these transformants, an RNA transcript of about 2.0 kb is found. S1 nuclease mapping of the RNA has indicated, as expected, that the Neo^r transcript is initiated in the LTR approximately at nucleotide +1. The presence of the 2.0-kb transcript in these transformants also suggests that the termination signals, which are postulated to be in the LTR, are also efficiently recognized by the transcriptional apparatus of the host cells.

In the majority of cells transformed by plasmid pATV-6D3A or pATV-6exo11A, the levels of Neo^r-specific transcripts are usually high, reaching about 500 copies per cell as determined by comparing known amounts of denatured neomycin DNA fragment, which was run in parallel. In pATV-6exo28R, which suffered a deletion downstream from the TATA box, the levels were in general low but sufficient to render the cells resistant to G-418. High levels of expression of the chloramphenicol gene under the influence of RSV LTR has also been demonstrated in a variety of cell types (7).

Perhaps the most interesting finding in these studies is the indication that the ASV promoter region includes, in addition to the TATA element, a second essential region, between position −114 and −299 in the LTR. This element is indispensable for the in vivo expression of genes under the control of the ASV promoter. Promoters for other genes also contain upstream sequences which are required for efficient expression. Transcription of the β-globin gene is decreased in vivo by an order of magnitude if either the TATA element or sequences between −100 and −58 are deleted (9). The upstream region is dispensable for efficient in vitro transcription of the globin gene (9). A similar control region is located between positions −100 and −40 of the herpesvirus I thymidine kinase gene (19). In the simian virus 40 early promoter system, sequences upstream of position −160 are dispensable for efficient in vitro transcription but indispensable for transient in vivo expression of the T antigen or for long-term transformation of recipient cells (1).

In simian virus 40, a 72-bp repeat appears to be required for the expression of early genes (1, 27). Analogous elements, which are referred to as enhancers, have been shown to exist in other viral systems (murine sarcoma virus, Rous sar-

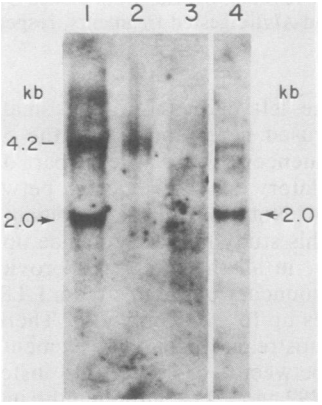

FIG. 6. Further evidence for the requirement of upstream sequence for transcription. XC cells were treated with supercoiled DNA for 36 h. Cells were harvested, and RNA was isolated and analyzed for neomycin-specific sequences as described in the legend to Fig. 4. RNAs isolated from the wild type (pATV-6D3A) and the deletion mutant (pATV-6D3A) were run in lanes 1 and 2, respectively. Lane 3, Control from cells which did not receive DNA. Lane 4, RNA from a stable transformant (X11C7; Fig. 4, lane 7) was used as a positive control.

coma virus) and function in *cis* even when positioned farther away from the coding sequences (17a, 27). Since the Neo^r gene is flanked by the LTR sequence, we would anticipate expression of the Neo gene if this sequence acts as an enhancer because the upper regulatory sequence in one LTR was intact in all the clones. Since no transformants were found when the sequence from the positions −299 to −114 in the upstream (relative to the Neo^r gene) LTR was deleted, we can conclude that this sequence does not activate gene transcription when inserted away from the structural gene. The discrepancy between our results and the results obtained by Luciw et al. (17a) cannot be reconciled at this time. It is possible that different host cell types used in these studies might account for these differences. Kriegler and Botchan (16) recently showed that the enhancing effect of the simian virus 40 and murine sarcoma virus LTRs is host specific. Nevertheless, recent data indicate that the enhancer works best when it is inserted 5′ to the structural gene (26). The other possibility is that the adjacent pBR322 sequences might in some way suppress the enhancing effect of the downstream LTR.

The sequence 5′-GCAATAT-3′ is located at positions −175 to −169 of the ASV LTR. This sequence exhibits homology with the consensus sequence that has been postulated to represent the core of the upstream control elements of many RNA polymerase II promoters (3). The main function of this element may be to provide an entry site for RNA polymerase in the promoter region, thus determining the rate of transcription in certain promoters, whereas the TATA element defines both the rate and the topography of transcriptional initiation. In cell-free lysates, entry of RNA polymerase on the DNA template seems to occur mainly through the ends of the molecule (J. L. Manley, personal communication). Thus, the upstream control element of the promoter is expected to play little or no role on the rate of transcriptional initiation in vitro. However, in vivo, this element may regulate entry of the transcriptional apparatus in the promoter region either by interacting with proteins which are involved in positive control of transcription or by participating in the generation of transcriptionally active open chromatin structure (18). In this connection, it should be pointed out that Groudine et al. (10) have observed a highly DNase-sensitive region in the LTR which can be correlated with expression of viral genes. They proposed that this region is probably involved in phasing the chromatin structure which regulates gene transcription. Removal of this upstream region may in some way perturb the chromatin formation, which in turn prevents expression of the Neo^r gene.

ACKNOWLEDGMENTS

We thank H. E. Varmus, C. Nguyen Huu, D. Figurski, and C. Milcarek for suggestions on the manuscript, S. Malik for constructing clone 6.49Nm1, and S. Gowda for S1 mapping experiment. Our sincere thanks to P. Daniels, Shering Corp., for generously donating G-418.

This work was supported by Public Health Service grant CA28990 from the National Cancer Institute. R.V.G. thanks the National Cancer Institute for a research career development award.

LITERATURE CITED

1. **Benoist, C., and P. Chambon.** 1981. *In vivo* sequence requirements of the SV40 early promoter region. Nature (London) **290**:304–309.
2. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
3. **Breathnach, R., and P. Chambon.** 1981. Organization and expansion of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
4. **Colbere-Garapin, F., F. Horodniceau, P. Kourilsky, and A. C. Garapin.** 1981. A new dominant hybrid selectable marker for higher eukaryotic cells. J. Mol. Biol. **150**:1–14.
5. **Cooper, G. M., and H. M. Temin.** 1976. Lack of infectivity of the endogenous avian leukosis virus-related genes in the DNA of uninfected chicken cells. J. Virol. **17**:422–430.
6. **Fuhrman, L. A., C. Van Beverin, and I. 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. U.S.A. **78**:5411–5415.
7. **Gorman, C. M., G. T. Merlino, M. C. Willingham, I. Pastan, and B. Howard.** 1982. The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc. Natl. Acad. Sci. U.S.A. **79**:6777–6781.
8. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus DNA. Virology **52**:456–467.
9. **Groseveld, G. C., E. de Boer, C. K. Shewmaker, and R. A. Flavell.** 1982. DNA sequences necessary for transcription of the rabbit β-globin gene *in vivo*. Nature (London) **295**:120–126.
10. **Groudine, M., R. Eisenman, and H. Weintraub.** 1981. Chromatin structure of endogenous retroviral genomes and activation by an inhibitor of DNA methylation. Nature (London) **292**:311–317.
11. **Guntaka, R. V., and S. A. Mitsialis.** 1980. Cloning of avian tumor virus DNA fragments in plasmid pBR322: evidence for efficient transcription in *E. coli* from a virus-coded promoter. Gene **12**:113–121.
12. **Guntaka, R. V., P. Y. Rao, S. A. Mitsialis, and R. Katz.** 1980. Modification of avian sarcoma proviral DNA sequences in nonpermissive XC cells but not in permissive chicken cells. J. Virol. **34**:569–572.
13. **Hsu, T. W., J. L. Sabran, G. E. Mark, R. V. Guntaka, and J. Taylor.** 1978. Analysis of unintegrated avian RNA tumor virus double-stranded DNA intermediates. J. Virol **28**:810–818.
14. **Jimenez, A., and J. Davies.** 1980. Expression of a transposable antibiotic resistance element in Saccharomyces. Nature (London) **287**:869–871.
15. **Joyner, A., Y. Yamamoto, and A. Bernstein.** 1982. Retrovirus long terminal repeats activate expression of coding sequences for the herpes simplex virus thymidine kinase gene. Proc. Natl. Acad. Sci. U.S.A. **79**:1573–1577.
16. **Kriegler, M., and M. Botchan.** 1983. Enhanced transformation by a simian virus 40 recombinant virus containing a Harvey murine sarcoma virus long terminal repeat. Mol. Cell. Biol. **3**:325–339.
17. **Lee, F., R. Mulligan, P. Berg, and G. Ringold.** 1981. Glucocorticoids regulate expression of dihydrofolate reductase cDNA in mouse mammary tumor virus chimaeric plasmids. Nature (London) **294**:228–232.

17a. **Luciw, P., J. M. Bishop, H. E. Varmus, and M. R. Capecchi.** 1983. Location and function of retroviral and SV40 sequences that enhance biochemical transformation after microinjection of DNA. Cell **33**:705–716.

18. **Mathis, D. J., and P. Chambon.** 1981. The SV40 early region TATA box is required for accurate *in vitro* initiation of transcription. Nature (London) **290**:310–315.

19. **McKnight, S. L., E. R. Gavis, R. Kingsbury, and R. Axel.** 1981. Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell **25**:385–398.

20. **Mitsialis, S. A., J. L. Manley, and R. V. Guntaka.** 1983. Localization of active promoters for eucaryotic RNA polymerase II in the long terminal repeat of avian sarcoma virus DNA. Mol. Cell. Biol. **3**:811–818.

21. **Mitsialis, S. A., F. Young, P. Palese, and R. V. Guntaka.** 1981. An avian tumor virus promoter directs expression of plasmid genes in E. coli. Gene **16**:217–225.

22. **Neel, B. G., W. S. Hayward, H. L. Robinson, J. Fang, and S. M. Astrin.** 1981. Avian leukosis virus-induced tumors have common provirus integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. Cell **23**:323–334.

23. **Payne, G., S. Courtnidge, L. Crittenden, A. Fadly, J. M. Bishop, and H. E. Varmus.** 1981. Analysis of avian leukosis virus DNA and RNA in bursal tumors: viral gene expression is not required for maintenance of the tumor state. Cell **23**:311–322.

24. **Shank, P. R., S. H. Hughes, H. J. Kung, J. E. Majors, N. Quintrell, R. V. Guntaka, J. M. Bishop, and H. E. Varmus.** 1978. Mapping unintegrated avian sarcoma virus DNA: termini of linear DNA bear 300 nucleotides present once or twice in two species of circular DNA. Cell **15**:1383–1395.

25. **Varmus, H. E.** 1982. Form and function of retroviral proviruses. Science **216**:812–820.

26. **Wasylyk, B., C. Wasylyk, P. Augeneau, and P. Chambon.** 1983. The SV40 72bp repeat preferentially potentiates transcription starting from proximal natural or substitute promoter elements. Cell **32**:503–512.

27. **Weiher, H., M. Konig, and P. Gruss.** 1983. Multiple point mutations affecting the simian virus 40 enhancer. Science **219**:626–631.

28. **Wigler, M., S. Silverstein, L. S. Lee, A. Pellicer, Y. C. Cheng, and R. Axel.** 1977. Transfer of purified herpesvirus thymidine kinase gene to cultured mouse cells. Cell **11**: 223–232.

29. **Yamamoto, T., B. deCrombrugghe, and I. Pastan.** 1980. Identification of a functional promoter in the long terminal repeat of Rous sarcoma virus. Cell **22**:787–797.