Volume 14 Number 15 1986 Nucleic Acids Research

In vivo identification of sequence elements required for normal function of the adenovirus major late transcriptional control region

John Logan[1] and Thomas Shenk*

Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA

Received 9 April 1986; Revised and Accepted 30 June 1986

ABSTRACT

A series of adenovirus type 5 variants were constructed to identify the sequence elements which comprise the major late transcriptional control region in the context of the viral chromosome. The variant chromosomes carried a second copy of DNA sequence derived from the region surrounding the major late mRNA cap site. The reiterated segments replaced the normal transcriptional control region of the E1A gene. By monitoring the rate of E1A transcription subsequent to infection with the variants, it was possible to evaluate the capabilities of the substituted major late elements. A segment derived from -55 to +33 (relative to the major late cap site at +1) functioned for early transcription, in the presence of the E1A enhancer domain, but failed to direct enhanced levels of activity late after infection. A segment from -122 to +33 directed both early and enhanced late transcription. The rate of late E1A transcription directed by this element was about 40% of that displayed by the major late control region at its normal position. Inclusion of additional upstream sequences (to -565) did not increase late transcription rates. Thus, the segment between -122 and -52 contains a region required for normal function of the adenovirus major late control region.

INTRODUCTION

Late after adenovirus infection the preponderance of viral mRNAs are transcribed from the major late transcription unit. This unit has a control region at about 16.4 map units (1) and extends more than 30,000 base pairs to near 100 map units (2). The large primary transcript is processed to yield about 17 different mRNAs (3). These mRNAs all have the same capped 5' end as the primary transcript (4), but are differentiated by the use of a variety of 3' splice sites and poly A addition sites. Transcription of the major late unit occurs at a low level early after infection (5,6) and at a substantially increased level late after infection. The high level of transcription observed late after infection is probably due both to an increase in template number as a result of DNA replication and to an enhanced rate of initiation at the major late control region.

The DNA sequences required for initiation at the major late cap site have been investigated both in cell-free extracts and in HeLa cells transfected with recombinant plasmid DNAs. In both cases, two key elements have been identified. The first is a TATA homology (-25 to -30 relative to the cap

site) which appears to position the start site and influence the efficiency of transcription (7-11). The second is an element upstream of -50 which modulates transcriptional efficiency by a factor of 20- to 30-fold (8,11-13). Recently, a cellular factor has been identified which binds to the upstream element (-50 to -66) and enhances transcription in cell-free extracts (14-16).

Here we describe the domain required for normal function of the major late control element in the context of the adenovirus chromosome. A second copy of the region was substituted into the viral chromosome replacing the normal control region for an early transcription unit, E1A. Transcription of the modified E1A unit was then monitored and compared to the major late unit subsequent to infection. A segment mapping from -52 to +33 (relative to the major late cap site at +1) functioned for early transcription in the presence of the E1A enhancer domain, but failed to direct enhanced levels of activity late after infection. A segment derived from -122 to +33 contained elements necessary to direct low levels of transcription early and substantially increased levels late after infection. Thus, a critical element required for normal activity of this region lies between position -122 and -52.

MATERIALS AND METHODS

Plasmids, Viruses and Cells.

The recombinant plasmid pE1A-WT contains the left-end 3.8 map units (1339 base pairs) of the Ad5 genome (17). Viral sequences extend from 0 map unit at the pBR322 *Eco*RI site to 3.8 map units where an Ad5 *Xba*I site is fused to the plasmid *Pvu*II site. Adenovirus type 2 (Ad2) major late transcriptional control sequences were from pSmaF (18). Rearranged E1A genes carrying major late transcriptional control sequences were constructed by substituting portions of pSmaF into pE1A-WT.

H5*d*l309 or H5*d*l340 serve as the wild-type Ad5 parents in these studies. These viruses were generated from H5*wt*300 which is a plaque purified derivative of a virus stock obtained from H. Ginsberg. H5*d*l309 was selected as a variant that contains only one *Xba*I cleavage site, located at 3.8 map units (1339 base pairs from the left end of the chromosome), and it displays a wild-type phenotype (19). H5*in*340 (17) was derived from H5*d*l309 and carries a second copy of the left-end Ad5 E1A enhancer and *cis*-acting packaging sequences at the right end of the viral chromosome, i.e. the left-end 0-355 base-pair segment has replaced the normal right-end 107 base pairs. Substitutions constructed in pE1A-WT were rebuilt into intact viral chromosomes by ligation of *Xba*I-cleaved plasmid DNA to the 3.8-100 map unit *Xba*I-generated fragment from *d*l309 or *in*340 (20). The recombined molecules are then used to transfect permissive cells. The resulting viruses contain a normal left end, having lost pBR322 sequences presumably by copying the right-end terminal repeat during replication or through recombination. The substituted viruses (*sub*360-*sub*365) are diagramed in Figure 1. *sub*361 and 363 lack left-end packaging sequences and have been constructed in an *in*340 background while the remainder utilized *d*l309 as parent.

The 293 cell line (21) was maintained in medium containing 10% calf serum. Spinner cultures of HeLa cells (obtained from the American Type Culture Collection) were grown in medium containing 7% calf serum.

RNA Preparation and Analysis

To analyze steady state RNAs, cytoplasmic, polyadenylated RNA was isolated from HeLa cells (17) either early (5 hr) or late (20 hr) after infection at a multiplicity of 25 plaque-forming units per cell. RNAs were quantitated by blot analysis using an E1A-specific probe (Ad5 sequence position 1-1572). Primer extension analysis of the 5' ends of E1A mRNAs (22) utilized a single-stranded, 5' end-labeled oligonucleotide (Ad5 sequence position 574-589).

Transcription rates were measured in isolated nuclei essentially as described by Groudine et al. (23) and Hofer and Darnell (24). Nuclei prepared from infected cells were incubated for 15 min at 30°C in the presence of ^{32}P-UTP (750uCi/ml; 410Ci/mmole), nuclear RNA was isolated, degraded by treatment with 0.2N NaOH for 10 min at 4°C, and hybridized to single-stranded DNA probes bound to nitrocellulose (100ug genome equivalents) by the method of McKnight and Palmiter (25). M13 DNAs carrying region specific sequences served as probe DNAs: E1A, 1-4.5; E1B, 5.5-9.8; IVa2, 12-15.7; L1, 31.5-37.5; E2A, 63.6-68 map units. After the first round of hybridization a second DNA-containing filter was added to each mix to ensure quantitative results.

RESULTS

Construction and Propagation of Variants

It was not practical to mutate the major late transcriptional control region at its normal position. Mutations could not only potentially compromise function of the entire transcription unit, but could also inactivate the viral DNA polymerase coded within region E2B on the opposite DNA strand. Therefore, we chose to reiterate and substitute portions of the major late control region for the E1A control region. If the resulting alterations generated a defective virus, it could be propagated on 293 cells, a transformed human line that contains and expresses the Ad5 E1A and E1B genes (21). Function of the major late segments could then be assayed by monitoring E1A expression subsequent to infection of HeLa cells with the substituted viruses.

Rearranged E1A genes were initially prepared in recombinant plasmids, and then rebuilt into intact viral chromosomes substituting for the normal E1A gene. The E1A regions of the resulting variants are diagrammed in figure 1. The E1A insertions include 565 (sub362 and 363), 245 (sub360 and 361), 122 (sub364) or 52 base pairs (sub365) of sequence normally located immediately upstream of the major late cap site. These inserts have been placed into two backgrounds. In one case, all E1A transcriptional control elements known to function within infected cells (17) have been removed (sub361 and 363). In

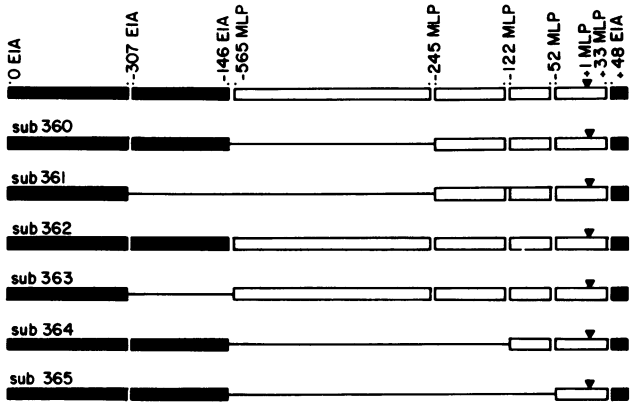

Figure 1. Diagram of Ad5 variants which carry segments of the major late control region in place of E1A transcriptional control elements. Solid blocks designate segments from the E1A region; open blocks are from the major late region. Single lines indicate that the corresponding segments are not present in the construction. Numbering is relative to the E1A or major late cap sites at +1. Thus, sub360 carries an upstream, E1A-specific segment (-499 to -146 relative to the E1A cap site), it lacks the -145 to +47 E1A segment and carries in its place a major late control region segment (-245 to +33 relative to the major late cap site).

the other, an E1A enhancer domain (17) remains upstream of the inserted major late sequences (sub360, 363, 364, 365). Initially, the variants were propagated on 293 cells to complement potential E1A defects, but all of the viruses turned out to grow as well as wild-type virus on HeLa cells.

The Reiterated Late Control Elements Are Active Early After Infection

As mentioned above, the major late transcription unit is active at a low level early after infection. To determine whether the reiterated major late control elements function early, E1A transcription rates were measured 4 hr after infection with the variants (Table 1). All of the rearranged genes were transcriptionally active at this time. When the E1A enhancer domain (-146 to -307 relative to the E1A start site at +1) was present upstream of the late control region, it increased the level of transcription directed by the late element about two-fold (compare sub360 which contains the enhancer to sub361 which lacks it, Table 1). The rate of E1A transcription directed by transplanted late control elements in the absence of an E1A enhancer was identical to that for the L1 mRNA family whose transcription is directed by the late control region at its normal position on the viral chromosome (compare E1A to L1 rates in sub361-infected cells, Table 1). Thus, the transplanted control segments appear to function normally at their new location early after infection.

The variant which includes the E1A enhancer but carries the smallest segment of the late control region (sub365, -52 to +33 relative to the late

Table 1. Rate of E1A and L1-specific transcription in wild-type and mutant-infected HeLa cells.

Virus	4 hr after infection ^{32}P-CPM x 10^{-2}		20 hr after infection ^{32}P-CPM x 10^{-3}		E1A/L1
	E1A	L1	E1A	L1	
d1309	3.2	1.2	1.50	23.6	0.06
sub360	2.8	1.4	10.5	21.4	0.49
sub361	1.2	1.2	9.4	20.3	0.46
sub362	2.5	--	11.3	24.2	0.47
sub363	1.1	--	8.8	20.7	0.43
sub364	2.2	--	7.8	19.5	0.40
sub365	3.3	--	3.0	22.4	0.13

Nuclei were prepared from infected cells at the indicated times after infection, incubated for 15 min at 30°C in the presence of ^{32}P-UTP, and nuclear RNA was prepared. RNA was hybridized to single-stranded probe DNAs carried in M13 immobilized on nitrocellulose filters, and radioactivity was quantitated. Results have been normalized to actin-specific CPM and are presented as CPM per 1000 nucleotides of probe sequence. Dashes indicate the measurement was not made.

cap site) is as active for early transcription as the wild-type E1A domain carried in d1309 (Table 1). This suggests that a TATA box and 5' start site are sufficient to fully activate an E1A gene containing its enhancer domain.

RNA blot analysis of early E1A-specific mRNAs present in cells infected with the substitution mutants confirmed the rate experiment (Figure 1). Steady state mRNA levels were highest in cells infected with variants which retained the E1A enhancer domain. The major late control region is clearly active early after infection when transposed to a new position on the viral chromosome.

Sequence Requirements for Normal Late Function of the Reiterated Late Control Elements

E1A and L1 transcription rates were measured 20hr after infection of HeLa cells. Normally, E1A is transcribed at a much lower rate than L1 at this late time (d1309, E1A to L1 ratio is 0.06, Table 1). E1A expression is increased to levels that approach late unit transcripts when controlled by a transposed late element (e.g. sub360, E1A to L1 ratio is 0.49, Table 1). Thus, the control sequences positioned upstream of the E1A coding region can function to direct transcription late after infection nearly as well as the late control region at its normal position. In contrast to early transcription the E1A enhancer domain has little effect on late activity (e.g. compare sub360 which contains the enhancer to sub361 which lacks it, Table 1).

Similar levels of late transcriptional activity were observed for inserts comprising as much as 565 bp (sub362) or as little as 122 bp (sub364) upstream of the late initiation site. A segment extending from -52 to +33 relative to the late start site was considerably less active (sub365, E1A to

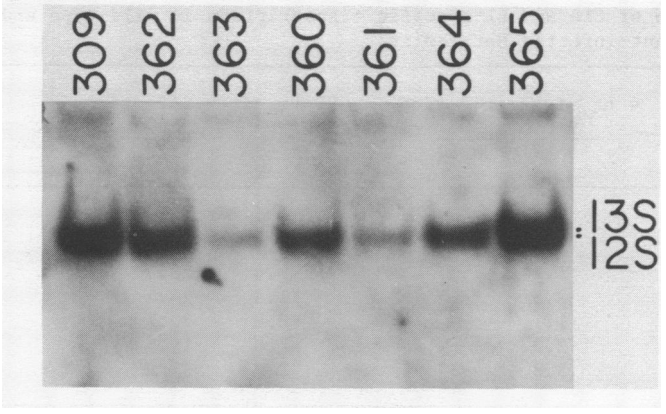

Figure 2. Blot analysis of viral E1A-specific mRNAs present in HeLa cells 5 hr after infection with mutant or wild-type viruses. Cytoplasmic, polyadenylated RNA was prepared, subjected to electrophoresis, transferred to a nitrocellulose filter, and probed using a ^{32}P-labeled, E1A-specific DNA segment. 13S and 12S mRNA species have not separated in this analysis.

L1 ratio is 0.13, Table 1). A key sequence element required for late transcription is located at least in part between -52 and -122 relative to the major late transcriptional start site.

RNA blot analysis of late E1A-specific mRNAs again confirmed the rate experiment (Figure 3A). In addition to the 13S and 12S mRNAs observed early (Figure 2), a late-specific 9S species predominates late after infection. As

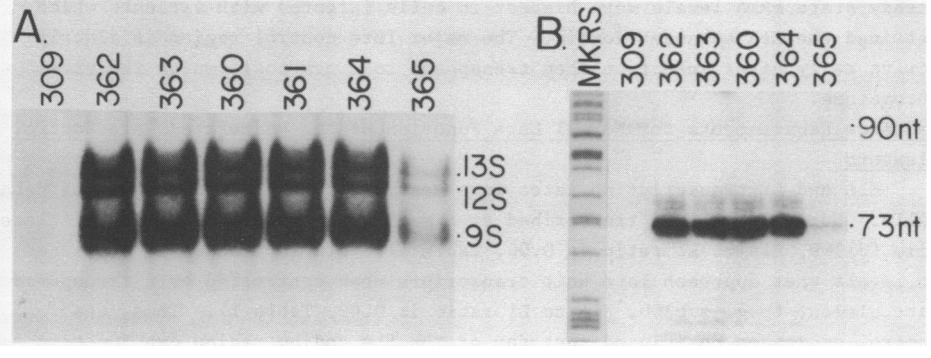

Figure 3. E1A-specific mRNAs present in HeLa cells 20 hr after infection with mutant or wild-type viruses. (A) RNA blot analysis. Cytoplasmic, polyadenylated RNA was prepared, subjected to electrophoresis, transferred to a nitrocellulose filter, and probed using a ^{32}P-labeled E1A-specific DNA segment. Bands corresponding to 13S, 12S and 9S mRNA species are labeled. (B) Primer extension analysis. A single-stranded, 5'-end-labeled oligonucleotide (nucleotide 574-589) was hybridized to polyadenylated RNA and extended using reverse transcriptase. Products of the reaction were resolved in an 8% polyacrylamide gel containing 8M urea. 90 and 73 nucleotide extension products (size determined relative to markers) are designated.

Table 2. Rate of transcription at 20 hr after infection of HeLa cells with wild-type or mutant virus.

Probe	E1A-Sense Strand ^{32}P-CPM x 10^{-3}		IV2-Sense Strand ^{32}P-CPM x 10^{-2}	
	dl309	sub362	dl309	sub362
E1A	1.5	11.3	--	--
E1B	2.2	9.9	2.5	1.9
IVa2	1.1	7.8	12.5	13.6
L1	23.6	24.2	--	--
E2A	--	--	10.8	11.1

Results from nuclear run-off analyses have been normalized to actin-specific CPM and are presented as CPM per 1000 nucleotides of probe sequence. Dashes indicate the measurement was not made.

predicted by transcription rates, cells infected with sub365 contained reduced steady state levels of E1A mRNAs (Figure 3A). The structure of E1A mRNA 5' ends was established by primer extension analysis (Figure 3B). All of the substitution mutants initiated transcription at the same position, which corresponded to the normal initiation site for the major late transcription unit.

The Reiterated Late Element Stimulates Enhanced Transcription Downstream of the E1A Unit

To measure effects of enhanced E1A transcription on neighboring units, transcription rates for the E1B and IVa2 units were assayed. The E1B unit resides immediately downstream of the E1A unit, and it utilizes the same coding strand. The IVa2 unit is located just downstream of the E1B unit, but its mRNAs are coded from the opposite DNA strand. The rate of E1B transcription was enhanced about four-fold in sub362- as compared to dl309-infected cells (Table 2). This increase is due to transcriptional readthrough from the E1A unit since the steady state level of bona fide E1B transcripts is unchanged (data not shown). The IVa2 anti-sense strand was also transcribed at an abnormally high rate in sub362-infected cells, but the rate of L1 transcription was not changed (Table 2). Polymerase which initiates at the E1A cap site can read through the E1A, E1B and IVa2 regions without encountering a termination site. Since L1 transcription was not influenced in sub362, it appears that the polymerase does not cross the major late transcriptional control region.

Transcription from the IVa2-sense strand was also monitored (Table 2). Enhanced transcription on the E1A-sense strand did not influence transcription rates on the IVa2 strand.

DISCUSSION

We (22) and others (26-29) have demonstrated previously that the adenovirus major late transcriptional control region can function efficiently

when transposed to the left end of the viral chromosome. Here, we define the extent of sequences which must be transposed to maintain normal activity of the control region. A segment extending from +33 to -122 relative to the major late start site (sub364, Table 1 and Figure 3) signaled late transcription at 40% the level of the control region at its normal location. Activity was increased only slightly when additional upstream sequences (to -565, sub363, Table 1 and Figure 3) were included in the transposed segment. Transcription was not enhanced late after infection when the inserted segment extended from +33 to -52 (sub365, Table 1 and Figure 3). Thus a key sequence required for late activity within infected cells must reside at least in part between -52 and -122 relative to the major late start site.

This observation fits well with previous reports of a requirement for sequences upstream of -50 for optimal activity of this control region in both short-term transfection assays and in cell-free extracts (8,11-13). It seems likely that the key upstream element is located at -50 to -66, the binding site for a cellular factor which has recently been identified (14-16). Our data suggest that this factor is responsible, at least in part, for the enhanced activity of the control region late after infection.

Although the E1A enhancer element augmented activity of the +33 to -122 major late control element early after infection, it had no effect at late times (compare sub360 to sub361, Table 1). Perhaps the mechanism of enhancement by the E1A element is somehow incompatible with late activity of the +33 to -122 element. Alternatively, the control region might be maximally active late after infection and unable to respond to additional stimulation provided by a juxtaposed enhancer element.

Even though it did not display enhanced late activity, the +33 to -52 segment directed nearly normal early function (sub365, Table 1 and Figure 2). Apparently, the cellular factor which binds to the -50 to -66 domain does not contribute to the early activity observed in sub365-infected cells. It is possible and perhaps even likely, that the E1A enhancer domain present in sub365 functionally substitutes for the normal major late upstream domain. Thus, we can not yet rule out a role for upstream sequences and the cellular factor in early activity of the major late control region.

ACKNOWLEDGMENTS

This work was supported by a grant from the American Cancer Society (MV-45). John Logan received fellowship support from the Otsuka Pharmaceutical Company. Thomas Shenk is an American Cancer Society Research Professor.

[1]Present address: Agricultural Research Division, American Cyanamid Company, PO Box 400, Princeton, NJ 08540, USA

*To whom correspondence should be addressed

REFERENCES

1. Evans, R.M., Fraser, N., Ziff, E., Weber, J., Wilson, M. and Darnell, J.E. (1977) Cell 12, 733-739.
2. Weber, J., Jelinek, W. and Darnell, J.E., Jr. (1977) Cell 10, 611-616.
3. Chow, L.T., Roberts, J.M., Lewis, J.B. and Broker, T.R. (1977) Cell 11, 819-836.
4. Gelinas, R.E. and Roberts, R.J. (1977) Cell 11, 533-544.
5. Chow, L.T., Broker, T. and Lewis, J.B. (1979) J. Mol. Biol. 134, 265-303.
6. Shaw, A.B. and Ziff, E.B. (1980) Cell 22, 905-916.
7. Hu, S-L, and Manley, J.L. (1981) Proc. Natl. Acad. Sci. USA 78, 820-824.
8. Hen, R., Sassone-Corsi, P., Corden, J., Gaub, M.P. and Chambon, P. (1982) Proc. Natl. Acad. Sci. USA 79, 7132-7136.
9. Concino, M., Goldman, R.A., Caruthers, M.H. and Weinmann, R. (1983) J. Biol. Chem. 258, 8493-8496.
10. Concino, M.F., Lee, R.F., Merryweather, J.P. and Weinmann, R. (1984) Nucl. Acids Res. 12, 7423-7433.
11. Miyamoto, N.G., Moncollin, V., Wintzerith, M., Hen, R., Egly, J.M. and Chambon, P. (1984) Nucl. Acids Res. 12, 8779-8799.
12. Jove, R. and Manley, J.L. (1984) J. Biol. Chem. 259, 8513-8521.
13. Yu, Y-T. and Manley, J.L. (1984). Nucl. Acids Res. 12, 9309-9321.
14. Carthew, R.W., Chodosh, L.A. and Sharp, P.A. (1985) Cell 43, 439-448.
15. Miyamoto, N.G., Moncollin, V., Egly, J.M. and Chambon, P. (1985) EMBO J. 4, 3563-3570.
16. Sawadogo, M. and Roeder, R.G. (1985) Cell 43, 165-175.
17. Hearing, P. and Shenk, T. (1983) Cell 33, 695-703.
18. Weil, P.A., Luse, D.S., Segall, J. and Roeder, R.G. (1979) Cell 18, 469-484.
19. Jones, N. and Shenk, T. (1979) Cell 17, 683-689.
20. Stow, N.D. (1981) J. Virol. 37, 171-180.
21. Graham, F.L., Smiley, J., Russell, W.C. and Nairu, R. (1977) J. Gen. Virol. 36, 59-72.
22. Logan, J. and Shenk, T. (1984) Proc. Natl. Acad. Sci. USA 81, 3655-3659.
23. Groudine, M., Peretz, M. and Weintraub, H. (1981) Mol. Cell. Biol. 1, 281-288.
24. Hofer, E. and Darnell, J.E. (1981) Cell 23, 585-593.
25. McKnight, G.S. and Palmiter, R.D. (1979) J. Biol. Chem. 254, 9050-9058.
26. Solnick, D. (1983) The EMBO Journal 2, 845-851.
27. Berkner, K.L. and Sharp, P. (1984) Nucl. Acids Res. 12, 1925-1941.
28. Berkner, K.L. and Sharp, P. (1985) Nucl. Acids Res. 13, 841-857.
29. David, A.R., Kostek, B., Mason, B.B., Hsiao, C.L., Morin, J., Dheer, S.K. and Hung, P.P. (1985) Proc. Natl. Acad. Sci. USA 82, 7560-7564.