JOURNAL OF VIROLOGY, Dec. 1984, p. 750–760
0022-538X/84/120750-11$02.00/0

Vol. 52, No. 3

Multiple Binding Sites for Polyomavirus Large T Antigen Within Regulatory Sequences of Polyomavirus DNA

ALISON COWIE AND ROBERT KAMEN*
Genetics Institute, Boston, Massachusetts 02115

Received 6 June 1984/Accepted 24 August 1984

Polyomavirus large T antigen binds specifically to multiple sites within the regulatory region of the viral genome. Experiments done with crude extracts from wild-type virus-infected mouse cells and immunoprecipitation of protein-DNA complexes localized two high-affinity binding sites on the early region side of the DNA replication origin. Purification of the large T antigen by immunoaffinity chromatography made it possible to refine the analysis through application of DNase I footprinting. The high-affinity interactions were resolved into three closely spaced, but distinct, binding regions. These begin at a site only slightly overlapping the early boundary of the core replication origin, a location highly homologous to that of simian virus 40 large T antigen-binding site I, but then extend away from the origin toward the early coding sequence and thus span the early region transcriptional initiation sites. Each tight-binding region contains from two to four copies of the sequence 5′-(A=T)G(A>G)GGC-3′ repeated at 9- to 11-base-pair spacing. At high protein concentrations and at low ionic strength, additional sites within the core replication origin and in the enhancer region were protected from DNase I digestion. These minor binding sites also included repeats of sequences related to the consensus, but at different spacings. Our results suggest that, unlike simian virus 40 DNA, the polyomavirus genome may have distinct regions of interaction with its large T antigen which separately are involved in initiation of DNA replication and the regulation of viral transcription.

The large T antigen of polyomavirus is one of three T (for tumor) antigens encoded by the early region of the viral genome. Activities of this 100-kilodalton nuclear phosphoprotein have been implicated in several processes that occur after virus infection. These include the stimulation of host cell DNA synthesis, the initiation of viral DNA replication, the repression of early gene expression, the switch from early to late transcription, and the events that lead to integration, excision, and amplification of the virus DNA within the host chromosome (for a review, see reference 46). Expression of the large T antigen from DNA integrated into the cellular genome results in remarkable phenotypic alterations, but does not cause tumorigenic transformation. It can reduce the serum growth factor requirement of established rodent cell lines and complement the serum-dependent transformed state induced in such cells by the expression of the viral middle T antigen (42). Transfection of rodent embryo fibroblasts with DNA encoding the large T antigen, or encoding its amino-terminal 40 kilodaltons, results in the establishment of permanent cell lines without passage through crisis. The cell lines thus generated have the growth properties of nontransformed fibroblasts with relatively low serum requirements; they are dependent on continued expression of the large T antigen for propagation in culture (43). This establishment function of the large T antigen can complement the activity of cellular or viral oncogenes in the tumorgenic conversion of mortal rodent cells (29, 45). Polyomavirus large T antigen is thus one of an interesting class of oncogene products, also including adenovirus E1A and the myc protein, which reside in the nucleus, induce establishment (but not tumorgenic transformation) of cells, and function as transcriptional regulators (15, 21, 22, 29, 39, 45, 48).

Two biochemical activities of the polyomavirus large T antigen have been demonstrated in vitro: ATP hydrolysis and DNA binding. We have undertaken the detailed analysis of DNA binding with the view that this function is fundamental to the initiation of viral DNA replication, the repression of early transcription, and presumably the activation of late transcription. In addition, specific binding to cellular DNA sequences may be involved in cellular immortalization. Previous investigation has shown that large T antigen binds with high selectivity and affinity to sequences within the noncoding region of the polyomavirus genome (6, 13, 41). This portion of the viral DNA and its situation in a recombinant plasmid containing full-length strain A2 DNA is illustrated diagrammatically in Fig. 1. Within the 0.5 kilobase of DNA shown are a number of important sequences that regulate viral gene expression and DNA replication. The central feature is the origin of DNA replication (ORI), which has been mapped by deletion analysis (26, 32) to a minimal core of 65 to 70 base pairs (bp). The ORI comprises an AT-rich tract (15 of 16 base pairs) adjacent to a sequence with dyad symmetry followed by a 14-bp inverted repeat. The ORI sequences share approximately 80% homology with the corresponding region of simian virus 40 (SV40) DNA (53). To the late side of ORI (leftward in Fig. 1) lies the 244 bp demonstrated to contain the polyomavirus transcriptional enhancer (5, 31, 63). The precise positions of what appear to be several sequence elements within this region involved in enhancer function have yet to be well defined, but recent evidence (64) suggested that the principal element lies between a unique *Bcl*I site and the first *Pvu*II site (nucleotide [nt] 5128). Sequences within the enhancer are essential not only for viral transcription, but also in *cis* for viral DNA replication (32, 36, 63). Multiple initiation sites for late transcription occur within this region of the genome (3). On the early region side of ORI (rightward in Fig. 1) lie some 80 bp of DNA, spanning a unique *Bgl*I site, to which no essential function has yet been assigned. This is then followed by a TATA box and the cap sites of the early region transcripts (2). The pentanucleotide 5′-(G/T)(A/G)GGC-3′ occurs 15 times throughout this region: 9 of the copies on the early side of ORI, 4 copies within the ORI core, and 2 copies

* Corresponding author.

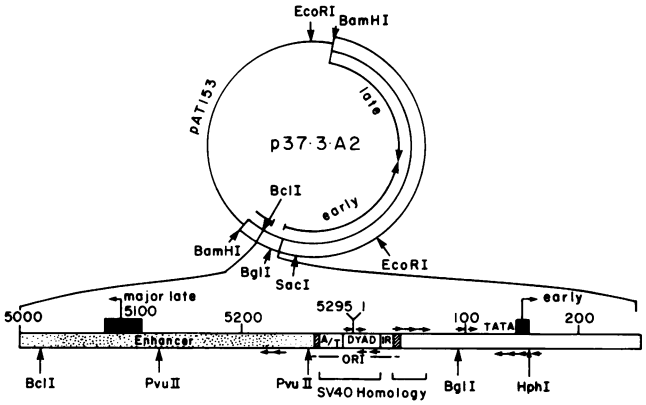

FIG. 1. Diagram of the regulatory region of polyomavirus. The regulatory region of the polyomavirus genome is shown with respect to its position in the recombinant plasmid p37.3.A2, which contains the entire viral DNA sequences cleaved with *Bam*HI and inserted into the homologous site of plasmid pAT153. The nucleotide numbering system is from Soeda et al. (54). The stippled area shows the enhancer sequences. The structural features of the ORI are indicated as follows: A/T, AT-rich region; DYAD, dyad symmetry; IR, inverted repeat. The crosshatched boxes represent the margin of error in the mapping of the minimal ORI core sequences sufficient for DNA replication (26, 32). The region of close homology with SV40 is shown (53). The early and late transcription initiation sites are marked by black boxes with arrows to show the direction of transcription. The small arrows above and below the diagram represent copies of the pentanucleotide 5'-(G/T)(A/G)GGC-3' and indicate on which strand of the DNA they occur. Certain relevant restriction endonuclease cleavage sites are indicated.

on the late side. This may be highly significant for polyomavirus large T antigen binding, because the analogous SV40 protein has been shown to interact directly with repeats of the same pentanucleotide sequence in SV40 DNA (4, 24, 55–59).

DNA binding studies on polyomavirus large T antigen have used, in almost all reports thus far, crude or slightly enriched preparations and indirect assay techniques such as DNA immunoprecipitation (41) or measurement of DNA-bound ATPase activity (13). By contrast, SV40 large T antigen has been extensively purified, and a rather comprehensive elucidation of its DNA interactions and of its role in transcriptional autoregulation (reviewed in references 59; 4, 17, 30, 37, 44, 55–58) has been achieved. The absence of genetically engineered recombinant viruses that overproduce the polyomavirus protein and of an efficient purification scheme has significantly impeded comparable definition of its DNA-binding properties.

Recently, Dilworth et al. (6) reported the purification of polyomavirus large T antigen utilizing a monoclonal antibody coupled to Sepharose. By a simple one-step method, the protein was purified to approximately 50% homogeneity from mouse 3T6 cells infected with wild-type virus. The large T antigen retained its specific DNA-binding activity, and preliminary experiments suggested that it could be used for DNase I footprinting. This work thus opened the way toward detailed characterization of polyomavirus large T antigen-DNA interaction by the application of methods possible only with purified protein.

In this paper we describe the interactions between polyomavirus large T antigen and regulatory sequences of the viral genome. We first used the indirect DNA immunoprecipitation approach to localize two high-affinity binding sites for the large T antigen within the noncoding region of viral DNA. These results were obtained with crude extracts derived from mouse cells infected with wild-type virus. Fractionation of the extracts by a modified version of the immunoaffinity purification described by Dilworth et al. (6) generated useful quantities of large T antigen which, as judged by the DNA immunoprecipitation method, retained the specific DNA-binding activity of the native molecule present in crude preparations. DNase I footprint experiments with the purified protein revealed multisite interactions between the T antigen and a series of closely situated binding regions in the viral DNA. Our results suggest that polyomavirus T antigen-binding regions functional in DNA replication may be separate from those involved in transcriptional regulation.

MATERIALS AND METHODS

Cells and viruses. Mouse 3T6 fibroblasts and monkey CV1 cells were routinely grown in Dulbecco modified Eagle medium supplemented with 5% fetal calf serum and antibiotics. Polyomavirus (A2 and A3 strain) was propagated in secondary mouse embryo cells at low multiplicity. SV40 (SV-S strain) was propagated in CV1 cells.

Recombinant DNAs. Plasmid p37.3.A2 (63) comprises *Bam*HI-cleaved polyomavirus DNA strain A2 cloned into the homologous site of plasmid pAT153 (62). Plasmids p43.25.67 and p43.34.70 are identical, except the *Pvu*II sites at nt 5128 and nt 5265, respectively (Fig. 1), have been replaced by an inserted *Xho*I linker (63). These plasmids were used to generate the pdl2000 series of deletion mutants from either *Pvu*II site (63). The pdl3000 plasmids are from a series of mutants with deletions from the single polyoma *Bgl*I site (nt 93; Fig. 1) that were described previously (23). Plasmids pΔ68 and pΔ81 also contain mutants deleted at the *Bgl*I site and were provided by M. Katinka (27). Plasmid p32.2 contains the fragment of polyomavirus DNA from the *Bcl*I site (nt 5021) to the *Bgl*I site (nt 93), spanning the origin of replication, cloned between the *Pvu*II and *Eco*RI site of plasmid pBR328 (52). Details of all recombinants used are presented in Table 1.

Preparation of radiolabeled DNA fragments. Polyomavirus and SV40 viral DNAs were isolated from 3T6 and CV1 cells, respectively, by selective sodium dodecyl sulfate extraction (20) and CsCl-ethidium bromide density gradient centrifugation. Plasmid DNAs were prepared from *Escherichia coli* strain HB101 by the alkaline sodium dodecyl sulfate method (1) and similar density gradient banding. All restriction endonucleases were purchased from New England Biolabs and used as recommended. Labeled DNA fragments for the immunoprecipitation assay were prepared by T4 DNA polymerase-catalyzed replacement synthesis (40). After digestion with the appropriate restriction endonucleases the digestion mixes were heated at 70°C for 5 to 15 min to inactivate the enzymes. The resultant DNA fragments were labeled by incubating 100 ng of a digest in 66 mM Tris acetate (pH 7.9)–33 mM potassium acetate–10 mM magnesium acetate)–1 mM dithiothreitol with 1 U of T4 DNA polymerase (PL Biochemicals) for 5 min at 37°C; 10 μCi of [α-^{32}P]dCTP (3,000 Ci/mmol; Amersham International) was then added, together with dATP, dGTP, and dTTP to 0.2 mM, and incubation was continued for 5 min at 37°C. Unlabeled dCTP was then added to 10 μM, and the incubation was continued for a further 10 min; the final reaction volume was 10 μl. The reaction was stopped with 90 μl of 10 mM Tris-hydrochloride (pH 7.5)–1 mM EDTA and heat inactivation of the enzyme at 70°C for 10 min.

For the DNase I footprint experiments, DNA fragments were labeled at the 5′ ends with [γ-^{32}P]ATP (3,000 Ci/mmol; New England Nuclear Corp.) and T4 polynucleotide kinase (PL Biochemicals) as described by Kamen et al. (25). For the DNA E-strand studies either strain A2 viral DNA or plasmid p37.3.A2 was cleaved with *Dde*I, labeled at the 5′ phosphate, and then cut with *Hin*fI to yield a fragment of 410 bp. For the DNA L-strand studies either p43.25.67 or p43.34.70 was cleaved with *Xho*I, labeled at the 5′ end, and then cut with *Hin*fI to give fragments of 556 and 413 bp, respectively. The resultant digests were resolved on 8% polyacrylamide gels in the presence of 8.3 M urea, but under nondenaturing conditions. The required fragments were eluted from the gel by the method of Maxam and Gilbert (33).

Preparation of crude extracts from infected cells. Mouse 3T6 cells or CV1 monkey cells were infected with polyomavirus A2 strain or SV40, respectively, at 50 to 100 PFU/cell. After 40 to 42 h at 37°C the cells were washed twice with Tris-buffered saline and lysed with 1 ml of lysis buffer (50 mM Tris-hydrochloride [pH 8.0], 150 mM NaCl, 2% Nonidet P-40, 1% Aprotinin [Sigma Chemical Co.]) per 90-mm culture dish. After 30 min on ice the lysates were centrifuged for 5 min in an Eppendorf microfuge. The supernatants were stored in liquid nitrogen.

Purification of polyomavirus large T antigen. We used the purification method described by Dilworth et al. (6), modified to reduce the level of contaminating actin in the final preparation. This proved necessary for the use of the partially purified protein in DNase I protection experiments because of the DNase-binding activity of actin. Typically, 1×10^9 to 2×10^9 mouse 3T6 fibroblasts were infected with wild-type polyomavirus strain A2 at 50 PFU/cell. The cells were grown at 32°C in Dulbecco modified Eagle medium containing 5% fetal calf serum and harvested at 40 h postinfection by lysis with 1 ml of 100 mM Tris-hydrochloride (pH 8.0)–100 mM NaCl–1% Nonidet P-40–1% Aprotinin per 2×10^7 cells on ice. After centrifugation at 5,000 rpm for 10 min, the supernatant was made 25 mM with potassium phosphate buffer (pH 7.2) and passed over a 10-g Bio-Gel HTP (Bio-Rad Laboratories) column. The HTP was washed with 10 volumes of 30 mM potassium phosphate buffer (pH 7.2) (actin was one of the proteins eluted at this phosphate concentration), and the remaining proteins, including the large T antigen, were then eluted with 5 volumes of 0.3 M potassium phosphate buffer (pH 7.2). To this fraction was added an equal volume of 1% Nonidet P-40, and it was then mixed with the αPyLT1 monoclonal antibody Sepharose as described by Dilworth et al. (6). After elution from the monoclonal antibody Sepharose with 3.5 M $MgCl_2$, the free antigen was dialyzed against 20 mM Tris-hydrochloride (pH 7.0)–0.1 M NaCl–1 mM EDTA–2 mM dithiothreitol–20% glycerol; bovine serum albumin was added to 0.01%, and the protein was stored in liquid nitrogen. The purity of each preparation was assayed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and silver staining before the addition of bovine serum albumin. The concentration of the large T antigen was approximated by visual comparison with known quantities of proteins run as markers on the gels; it was usually 10 to 50 μg/ml.

Immunoprecipitation of ^{32}P-labeled DNA fragment. The method used for immunoprecipitation of ^{32}P-labeled DNA was essentially a variation of the McKay procedure (35) as described by Hayday et al. (19), further modified as follows. ^{32}P-labeled DNA fragments (5 to 10 ng; ca. 10^{-3} pmol) were incubated with 100 μl of infected cell lysate in 1 ml of MB buffer (10 mM sodium phosphate [pH 7.0], 2 mM dithiothreitol, 0.01% [wt/vol] bovine serum albumin [Boehringer Mannheim], 1 mM EDTA, 0.05% Nonidet P-40, 3% dimethyl sulfoxide, 0.1 M NaCl, 5 μg of sheared salmon sperm DNA per ml) for 1 h at room temperature. Samples (5 μl each) of tissue culture supernatant, containing the monoclonal antibodies αPyLT1 and αPyLT4 (7) in the case of polyomavirus-infected cell lysate or 10 μl of tissue culture supernatant containing PabL416 (18) in the case of SV40-infected cell lysate, were then added for 1 h at room temperature followed by 20 μl of fixed *Staphylococcus aureus* cells to precipitate the DNA-protein complex. The immunosorbent was washed with 1 ml of MW buffer (20 mM Tris-hydrochloride [pH 8.0], 2 mM dithiothreitol, 0.01% bovine serum albumin, 0.5% Nonidet P-40, 1 mM EDTA, 0.1 M NaCl, 10 μg of sheared salmon sperm DNA per ml). The DNA was eluted from the immune complex with 1% sodium dodecyl sulfate–10 mM Tris-hydrochloride (pH 7.5)–1 mM EDTA, extracted twice with TE (10 mM Tris-hydrochloride [pH 7.5], 1 mM EDTA)-saturated phenol and precipitated with 2.5 volumes ethanol. The DNA fragments immunoprecipitated in this manner

TABLE 1. Structures of recombinant plasmids used in large T-Antigen binding studies

Plasmid	Vector	Polyomavirus DNA insert	Nucleotides deleted	Linker at site of deletion
p37.3.A2	pAT153	*Bam*HI linear	None	
p43.25.67	pAT153	*Bam*HI linear	None	*Xho*I
p43.34.70	pAT153	*Bam*HI linear	None	*Xho*I
pdl3027	pAT153	*Bam*HI linear	5291 to 103	*Xho*I
pdl3034	pAT153	*Bam*HI linear	77 to 127	*Xho*I
pdl3035	pAT153	*Bam*HI linear	86 to 133	*Xho*I
pdl3050	pAT153	*Bam*HI linear	5222 to 144	*Xho*I
pdl3041	pXf3	*Bam*HI linear	57 to 94	*Xho*I
pdl3044	pXf3	*Bam*HI linear	15 to 127	*Xho*I
pdl3045	pXf3	*Bam*HI linear	6 to 102	*Xho*I
pdl3048	pXf3	*Bam*HI linear	36 to 103	*Xho*I
pΔ81	pMK16*	*Bam*HI linear	43 to 147	*Sal*I
pΔ68	pMK16*	*Bam*HI linear	18 to 155	*Sal*I
p32.2	pBr328	*nt*5022-92	None	
pdl2015	pAT153	*Bam*HI linear	4654 to 68	*Xho*I
pdl2018	pAT153	*Bam*HI linear	4761 to 98	*Xho*I
pdl2019	pAT153	*Bam*HI linear	4831 to 36	*Xho*I
pdl2086	pAT153	*Bam*HI linear	4938 to 53	*Xho*I

were identified by agarose gel electrophoresis and quantitated by excision of the gel bands and Cerenkov counting. The radioactivity recovered in corresponding fragments excised from marker lanes served as normalization standards.

DNase I footprinting. Protection from partial DNase I digestion of labeled DNA fragments bound by polyomavirus large T antigen was examined by a modification of the method of Galas and Schmitz (12). Known amounts of 5′ end-labeled DNA fragment (1×10^{-3} to 2×10^{-3} pmol) were incubated with between 1 and 20 pmol (based on the monomeric molecular weight) of partially purified large T antigen in a final volume of 100 μl with DB buffer (20 mM Tris-hydrochloride [pH 7.0], 0.1 M NaCl, 2 mM dithiothreitol, 1 mM EDTA, 0.01% [wt/vol] bovine serum albumin) at 20°C per 30 min. Partial DNase I digestion was done by the addition of $MgCl_2$ to 5 mM, $CaCl_2$ to 1 mM , and DNase I (Miles-Yeda) to a final concentration of 5 ng/ml and incubation for a further 10 min at 20°C. The reaction was terminated by the addition of an equal volume of 2 M ammonium acetate–100 mM EDTA–100 μg of sheared salmon sperm DNA per ml. The samples were extracted with TE-saturated phenol and precipitated from ethanol. They were then dissolved in formamide dye mix (33) and run on 8% polyacrylamide-urea denaturing gels. The gels were dried before autoradiography.

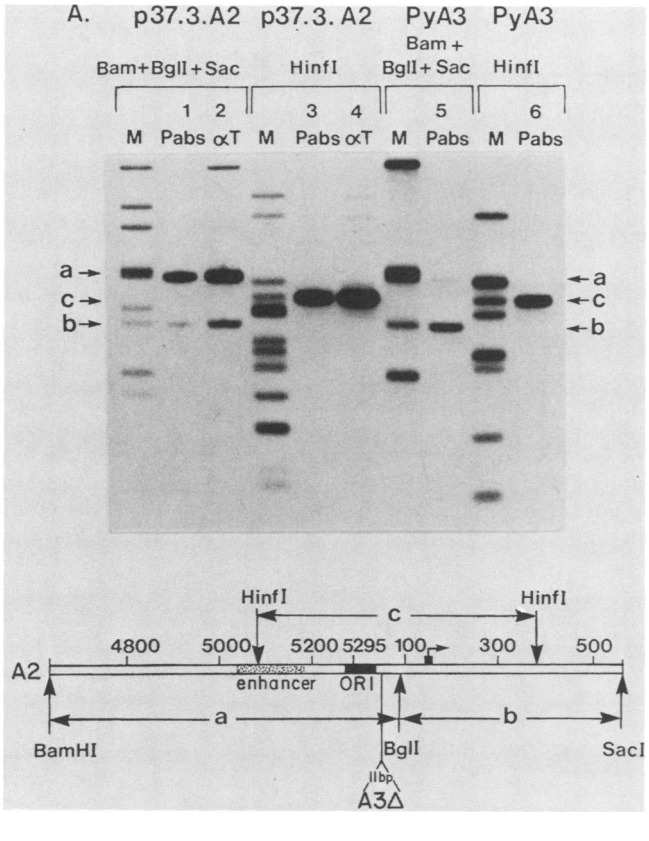

FIG. 2. Specific immunoprecipitation of polyomavirus ORI region DNA. ^{32}P-labeled restriction endonuclease digests (5 to 10 ng, ca 10^{-3} pmol) were incubated with cell extract from 2×10^6 infected 3T6 cells as described in the text. Lanes 1 and 2 show immunoprecipitation from a combined *Bam*HI-*Bgl*I-*Sac*I digest of the plasmid p37.3.A2 with either αPyLT1 plus αPyLT4 monoclonal antibodies or with anti-tumor serum. Lane 5 shows the same digest with strain A3 DNA after immunoprecipitation with αPyLT1 and αPyLT4. Lanes 3, 4, and 6 show the results of immunoprecipitation from a *Hin*fI digest of either p37.3.A2 or A3 DNA, respectively, with either αPyLT1 plus αPyLT4 or antitumor serum. The diagram indicates the positions of the *Hin*fI fragment c and the *Bam*HI-*Bgl*I-*Sac*I fragments a and b with respect to each other and the polyomarvirus regulatory region. The position of the 11-bp deletion in A3 strain DNA is marked (53). The lanes labeled M comprise 1/50 of the input DNA used as size markers and to aid quantitation. Note that the fragment a band in the marker lanes of the *Bam*HI-*Bgl*I-*Sac*I digest comigrates with another fragment and therefore appears to be disproportionately intense.

RESULTS

Specific immunoprecipitation of polyomavirus large T antigen bound to regulatory region DNA demonstrates two tight binding sites. Our initial experiments used as a source of the viral large T antigen extracts prepared from 3T6 mouse cells infected with wild-type virus. These studies demonstrated highly specific DNA interactions, approximated the locations of two high-affinity binding sites, and provided an assay to assess the activity of the protein through the purification used for subsequent experimentation.

The immunoprecipitation method developed by McKay (35) allowed us to examine the DNA-binding activity in unfractionated extracts. In our application of this procedure mixtures of radiolabeled DNA fragments were incubated with crude lysates of infected cells. The assay was made highly selective for specific DNA-protein interactions by the addition of nonradioactive carrier DNA (5 μg of salmon sperm DNA per 5 to 10 ng of radiolabeled restriction fragments) and 100 mM NaCl to the binding buffer used by McKay (35). Those fragments bound to the large T antigen were immunoprecipitated by reaction with either polyclonal antitumor serum or a mixture of two different rat anti-large T antigen monoclonal antibodies (7), followed by adsorption to fixed *S. aureus* bacteria. The labeled DNA eluted from the washed immunocomplexes was identified by agarose gel electrophoresis and autoradiography.

Figure 2 shows the results obtained with two different restriction endonuclease digests of recombinant plasmid p37.3.A2 (Fig. 1), which contains full length viral A2 strain DNA cloned at the *Bam*HI site into plasmid pAT153. The unique DNA fragment immunoprecipitated from the *Hin*fI digestion (band c in Fig. 2) by either antitumor serum (lane 4) or monoclonal antibodies (lane 3) spans the enhancer region, ORI, and the early region promoter (Fig. 2). No other DNA fragments were detectably immunoprecipitated (the adjacent lane M shows the input DNA digest). We therefore concluded that the large T antigen binds with high specificity to one or more sequences within the regulatory region of the polyomavirus genome. The single high-affinity binding site localized by Gaudray et al. (13), with a different procedure, is included within DNA fragment c.

To test the hypothesis proposed by Gaudray et al. (13) and Kamen et al. (25) that polyomavirus large T antigen binds separately to sequences at the transcriptional initiation sites as well as to those in the ORI region, we digested plasmid p37.3.A2 with the endonuclease *Bgl*I, which cleaves polyomavirus DNA at a single site between the ORI and the early region cap sites, plus *Bam*HI and *Sac*I. Two fragments were immunoprecipitated from this digest (Fig. 2, lanes 1 and 2): fragment a, which includes ORI, and fragment b, which spans the TATA box and the cap sites. Fragment a was more efficiently immunoprecipitated than fragment b. By excision of the gel bands and Cerenkov counting the molar ratio between fragments a and b was determined to be approximately 4:1. The regulatory region thus contains at least two

TABLE 2. Mapping of the large T antigen high affinity binding regions with deletion mutants and endonuclease cleavage

Plasmid	Nucleotides present within fragment: a	Nucleotides present within fragment: b	Endonuclease digest[a]	Relative immunoprecipitation (%)
p37.3.A2	4632 to 93		1	100
p32.2	5022 to 93		2	100
pdl2019	37 to 93		1	78
pdl2086	54 to 93		1	48
pdl2015	69 to 93		3	0
p32.2	5022 to 46		4	0
p32.2	47 to 93		4	0
pdl3035	4632 to 85		5	100
pdl3034	4632 to 76		5	100
pdl3041	4632 to 56		5	0
pΔ81	4632 to 42		6	0
pdl3048	4632 to 35		5	0
p37.3.A2		94 to 569	1	100
pdl2018		99 to 484	7	76
pdl3027		104 to 569	5	75
pdl3034		128 to 569	5	25
pdl3035		134 to 569	5	10
pdl3050		145 to 385	8	0
pdl3027		104 to 185	9	100
pdl3027		104 to 153	10	11
pdl3027		154 to 469	10	0
pdl3027		104 to 135	4	0
pdl3027		136 to 263	4	0

[a] Key to the endonuclease digestions. 1, *Bam*HI 1-*Bgl*I-*Sac*I; 2, *Rsa*I-*Eco*RI; 3, *Bgl*I-*Pst*I; 4, *Mbo*II; 5, *Bam*HI-*Xho*I-*Sac*I; 6, *Bam*HI-*Sal*I-*Sac*I; 7, *Xho*I-*Pst*I; 8, *Hin*I; 9, *Dde*I; 10, *Hph*I.

independent large T antigen-binding sites, with apparently different affinities, located on either side of the *Bgl*I site. From the experiment shown in Fig. 2 the binding region within fragment b, on the early side of the *Bgl*I site, maps between this position (nt 93) and the *Hin*fI site at nt 385. Further analysis with a number of different restriction enzymes and deletion mutants (Tables 1 and 2) demonstrated that all of the sequences important for binding within fragment b are located between nt 103 and nt 188. This DNA spans the early region TATA box and cap sites (2, 25). Truncation of the DNA fragment at the principal cap site (nt 153) by cleavage with endonuclease *Hph*I reduced the binding by 10-fold (Table 2).

Results were also obtained with polyomavirus strain A3 DNA. This strain differs from strain A2 DNA by an 11-bp deletion (nt 44 through 55) (53), removing one of the three AGAGGC repeats adjacent to the ORI (Fig. 1 and 2). *Hin*fI fragment c was efficiently immunoprecipitated from the digest of strain A3 DNA (Fig. 2, lane 6), but in the *Bgl*I-*Bam*HI-*Sac*I digest, fragment a binding was only 8% efficient relative to the strain A2 fragment. Binding to A3 strain fragment b was at the same efficiency as observed for A2 DNA (Fig. 2, lane 5). The difference between A2 and A3 DNAs indicated to us that the binding region within fragment a involves the deleted sequences and suggested that the number of AGAGGC hexanucleotide sequences within the binding region (two in strain A3 DNA versus three in A2) influences the binding affinity when measured by the selective immunoprecipitation assay. Our data further implied that the strong binding site located within fragment a was not important for viability because strains A2 and A3 are both wild type and grow equally well in culture.

Origin-proximal binding site maps outside the ORI region. The data presented in Fig. 2 suggest that the high-affinity binding site within fragment a might not include the ORI sequence. We define the ORI as one of the two elements necessary in *cis* for viral DNA replication (32, 36, 63). It has been mapped by deletion analysis to a minimal core sequence of ca. 70 bp between a late boundary at nt 5265 through 5269 (32) and an early boundary at nt 36 through 42 (26). The DNA sequence from nt 42 to 64 plays a quantitative role (its removal reduces replicatory efficiency to 10 to 20% of wild-type level), but it is not essential for DNA replication (26; our unpublished observations). The second element essential for DNA replication is noncontiguous and occurs within the enhancer element (32, 36, 63).

To better localize the binding region within fragment a and to ask more explicity whether large T antigen binds to the ORI, we tested a series of restriction endonuclease fragments extending from the *Bam*HI site (nt 4632) to various endpoints within fragment a. These were derived from a set of deletion mutants (Table 1) with *Xho*I or *Sal*I linkers at the site of deletion. The results (Fig. 3) demonstrate the absence of detectable binding until the fragment extends well past the ORI to nt 76. The large T antigen did not bind to the DNA fragments, which include competent replication origins such as those with endpoints at nt 42 or nt 56. This effect is not an artifact caused by the use of restriction endonuclease fragments with ends near the essential region. A deletion-insertion mutant in which polyomavirus DNA from nt 43 to 147 was replaced with a portion of the mouse metallothionine gene (G. M. Veldman, unpublished results) did not contain any large T antigen-binding sites detectable by immunoprecipitation (data not shown), yet this DNA replicated at 10 to 20% of the wild-type level after transfection into permissive mouse cells.

The experiments to locate the binding region within fragment a were extended to include a number of additional restriction fragments derived from wild-type or deletion mutant DNAs (Table 2). We concluded that there are no sequences essential or sufficient for high-affinity DNA binding before nt 36 or after nt 76. This would imply that the

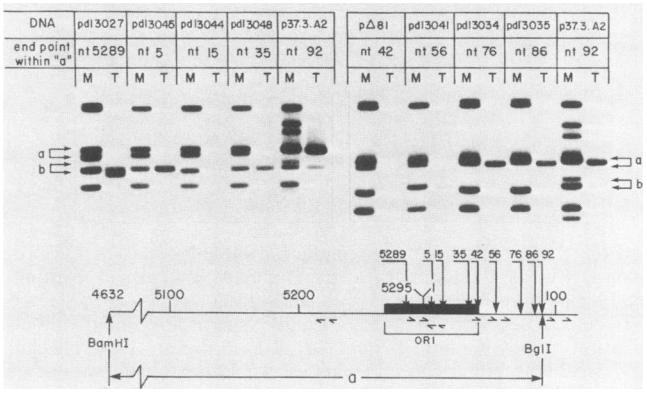

FIG. 3. Mapping the region of fragment a to which the large T antigen binds. Deletion mutants with endpoints within fragment a (Table 1) were cleaved with *Bam*HI, *Sac*I, and either *Bgl*I in the case of p37.3.A2, *Xho*I for pdl3027, pdl3045, pdl3044, pdl3048, pdl3041, pdl3034, and pdl3035, or *Sal*I for pΔ81 to cleave the linkers at the deletion sites. After incubation with infected cell extract and immunoprecipitation with polyclonal antibodies αPyLT1 and αPyLT4, the resultant DNA fragments were separated on 2% agarose gels with 1/50 of the radiolabeled DNA digests as marker tracks. The diagram shows the deletion endpoints in fragment a with respect to ORI. The repeated pentanucleotide sequence (G/T)(A/G)GGC is indicated by the half-arrows.

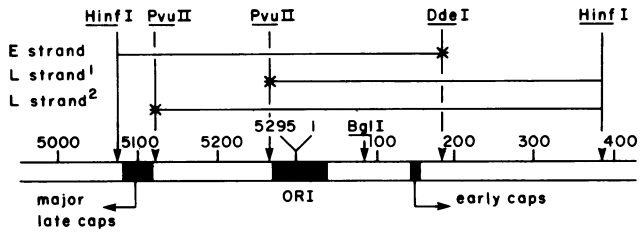

FIG. 4. DNA fragments used in the DNase I footprinting studies. The DNA fragments used in Fig. 5, 6, and 7 are shown aligned with various restriction endonuclease sites and the following structural features of the viral DNA: ORI, early RNA cap sites, and late RNA cap sites. Viral DNA or p37.3.A2 recombinant DNA was used for the E-strand fragment. A recombinant plasmid (p43.34.70) that has an *Xho*I linker inserted into the *Pvu*II site at nt 5265 was used for the L-strand[1] fragment. A similar plasmid (p43.25.67) has the *Pvu*II site at nt 5128 replaced with an *Xho*I linker and was used for the L-strand[2] fragment. The position of the 5′ ^{32}P label is marked by an asterisk.

binding region is contained within these limits and therefore overlaps the early boundary of the ORI core by no more than a few base pairs.

Purification of polyomavirus large T antigen. DNA fragment immunoprecipitation is a sensitive technique for study of a DNA-binding protein present in crude cell extracts. When combined with a sufficiently extensive set of DNA deletion mutants, it can be used, as illustrated above, to approximate the positions of specific, independent DNA-binding regions. We considered these assignments as preliminary, however, because the method does not directly identify the DNA sequences contacted by the protein in wild-type DNA. We were also unable to study potentially cooperative binding that could extend into the ORI. A satisfactory understanding of the large T antigen-DNA interaction would only be possible if the protein were purified. It was our good fortune at this time to enter into a collaboration with S. Dilworth to evaluate an immunoaffinity purification procedure he had developed by using monoclonal antibodies against polyomavirus large T antigen (7). We demonstrate elsewhere (6) that his single-step immunoaffinity purification yields, from wild-type virus-infected 3T6 cells, protein 30 to 50% pure which retains its DNA-binding specificity. We have now modified the procedure to reduce frequent contamination with a protein, presumed to be actin, that interfered with the DNase I footprint analysis. Details of the purification steps are described above.

Detection of three high-affinity binding regions by DNase I footprinting. With the availability of purified protein, albeit in small quantities, we were able to apply the footprinting procedure of Galas and Schmitz (12) to study the binding interactions more definitively. The ^{32}P-labeled DNA fragments used for the footprint studies are illustrated in Fig. 4. We incubated the DNA with increasing amounts of purified large T antigen in a fixed final reaction volume. A typical saturation pattern is shown in Fig. 5. Three distinct regions of the DNA, designated A, B, and C, were protected from DNase I digestion with as little as 1 pmol (5 μl of a solution of ca. 20 μg/ml) of large T antigen monomer.

Alignment of the protected regions with the marker lanes and the A and G sequencing track allowed precise localization with respect to the DNA sequence (see Fig. 8). Region A is adjacent to the replication origin, overlapping slightly the ORI early boundary (Fig. 1). Region B spans the single *Bgl*I site at nt 93 and ends just before the TATA box, where there are a few base pairs unprotected. Region C extends from the TATA sequence to halfway between the early cap sites (nt 148 through 153 [32]) and the ATG of the coding region (nt 173).

The positions of binding regions A and C agree very well with the conclusions derived from DNA immunoprecipitation mapping experiment. Binding of large T antigen to region B, however, could not be demonstrated by the indirect assay method. In repeated efforts, we were unable to immunoprecipitate an endonuclease *Mbo*II fragment (nt 48 through 133), which includes all of region B. The significance of this discrepancy between footprint and immunoprecipitation results is considered below.

The prominent regions of DNase I protection (A through C) manifest in Fig. 5 confirmed that the high-affinity binding sites for polyomavirus large T antigen occur outside of the ORI core sequence. Nevertheless, inspection of data such as that presented for the DNA E strand in Fig. 5 suggested that the T antigen was also perturbing the DNase I digestion pattern in the ORI region, but that this required higher

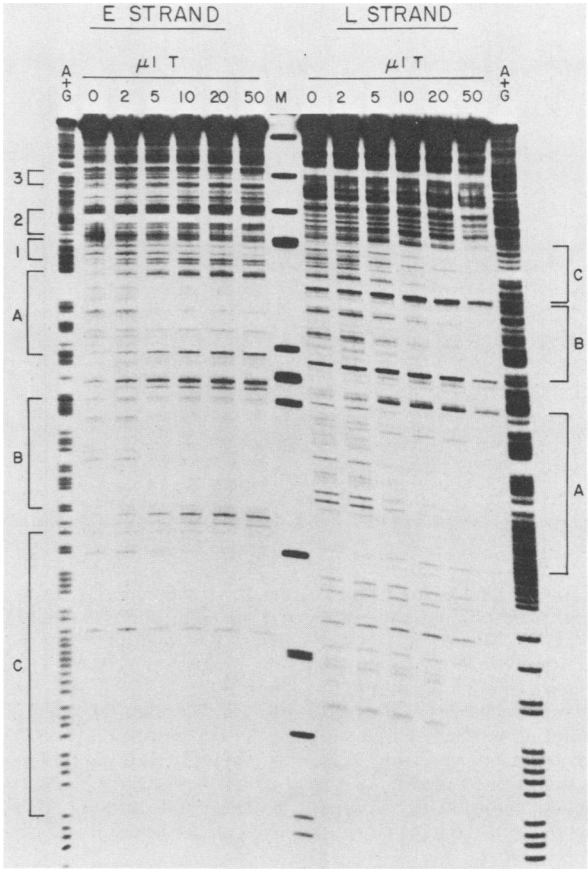

FIG. 5. Binding of polyomavirus large T antigen to the noncoding region of the genome. To localize the binding sites of large T antigen to both E- and L-DNA strands, approximately 10^{-3} pmol of 5′-^{32}P-labeled restriction endonuclease fragments (E strand and L strand[1] from Fig. 4) were incubated with increasing amounts of partially purified large T antigen in a total volume of 100 μl of DB buffer as described in the text. After DNase I digestion the samples were subjected to denaturing gel electrophoresis in an 8% polyacrylamide-urea sequencing gel. The A+G sequencing reaction (33) of each fragment was used as a marker track together with a size marker.

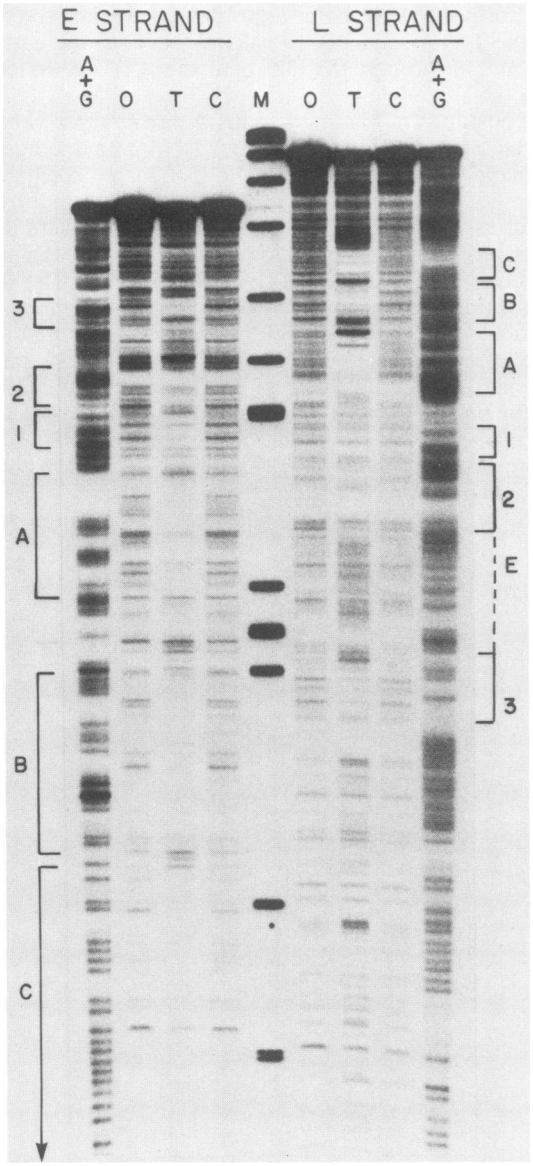

FIG. 6. Binding of large T antigen at high concentration. To localize the binding sites of large T antigen, approximately 10^{-3} pmol of 5'-^{32}P-labeled DNA fragments (E strand and L strand[2] in Fig. 4) were incubated with 100 μl (ca. 20 pmol of monomer) large T antigen at room temperature for 30 min. After DNase I digestion the resultant fragments were resolved on an 8% urea-polyacrylamide denaturing gel. Size markers and the A+G cleavage reaction (33) of each fragment were run as marker lanes. Lane O represents the reaction in the presence of 100 μl of DB buffer, lane T shows the reaction with 100 μl of large T antigen, and lane C shows the reaction with 100 μl of large T storage buffer (see the text for buffer compositions).

concentrations of protein. The putative additional binding sites are indicated as 1, 2, and 3 in Fig. 5 for consistency with conclusions derived from the subsequent experiments described below.

Detection of additional binding regions with ORI and the enhancer sequences by DNase I footprinting. To further investigate the interaction between large T antigen and the ORI region, we increased the amount of protein in the binding reaction to the maximum level technically feasible (Fig. 6). We also used the longer 5' ^{32}P-labeled fragment for the DNA L-strand probe (Fig. 4). To show that the protection was specific for the presence of T antigen, control DNase I digestions were done in parallel in the binding buffer and the large T antigen storage buffer (see above for details). As a result, T antigen-specific effects, in addition to protection of regions A, B, and C, were revealed (Fig. 6). Considering first the E-strand DNA, we note above region A an enhanced cleavage site and then a region of protection. As this protected region is interrupted by a site of enhanced digestion, we tentatively define two adjacent binding sites, 1 and 2. Above region 2, some 50 bp removed, we note a further region of DNase I protection, indicated in Fig. 6 as region 3. Protection of regions 1 and 2 was less obvious on the DNA L strand, but region 3 was apparent, as was a marked region of enhanced DNase I digestion between regions 2 and 3 (denoted by E in Fig. 6). The interactions at regions 1 and 2 map around nt 5295/1, within the ORI core sequence (cf. Fig. 1 and 8). Region 3 is on the late side of the ORI (around nt 5220; see Fig. 8).

We next tried varying the NaCl concentration in the binding reaction. All previous experiments had been carried out at 0.1 M NaCl. Figure 7 shows the results obtained by using the same amount of large T antigen in each reaction, but increasing the NaCl concentration from 33 mM to 0.2 M. At 33 and 50 mM NaCl the protection at sites 1, 2, and 3 was more clearly visible on both DNA strands than at 100 mM, as was the region of enhanced digestion (E) between 2 and 3. At 150 mM NaCl no protection of sites 1, 2, and 3 was detected, but regions A, B and C were unchanged. The regions protected at low ionic strength map to the same positions at those shown at high protein concentration, i.e., on either side of the nt 5295/1 and around nt 5220. The dependence of the protection of sites 1, 2, and 3 on high protein concentration or low ionic strength may indicate that the large T antigen has a significantly lower affinity for these sites than for regions A, B, and C, which could explain why they were not detected in the immunoprecipitation assay. At present, however, we cannot rule out the alternative explanation that a minor subset of T-antigen molecules binds to regions 1, 2, and 3.

DISCUSSION

We have presented results that define multiple interactions between polyomavirus large T antigen and the viral genome. The protein is known to be involved in the initiation of each successive round of viral DNA replication and in the repression of early transcription (46) and, by analogy to SV40 large T antigen, could also play a direct role in the augmentation of late transcription (28). Specific DNA binding perhaps mediates each of these activities. Our aim was to identify the DNA sequences recognized by the protein with the anticipation that localization of binding sites with respect to other regulatory signals would suggest their function.

Indirect mapping experiments with crude protein preparations. The study of polyomavirus large T antigen has been impeded by technical difficulties encountered in its purification. It was nevertheless possible to examine its DNA-binding properties by the application of two different indirect approaches. The earliest report (13) used a novel method in which the T antigen present in rather crude preparations was absorbed to double-stranded calf thymus DNA-cellulose. Polyomavirus T antigen ATPase activity could be specifically eluted by incubation of the DNA-cellulose with plasmids

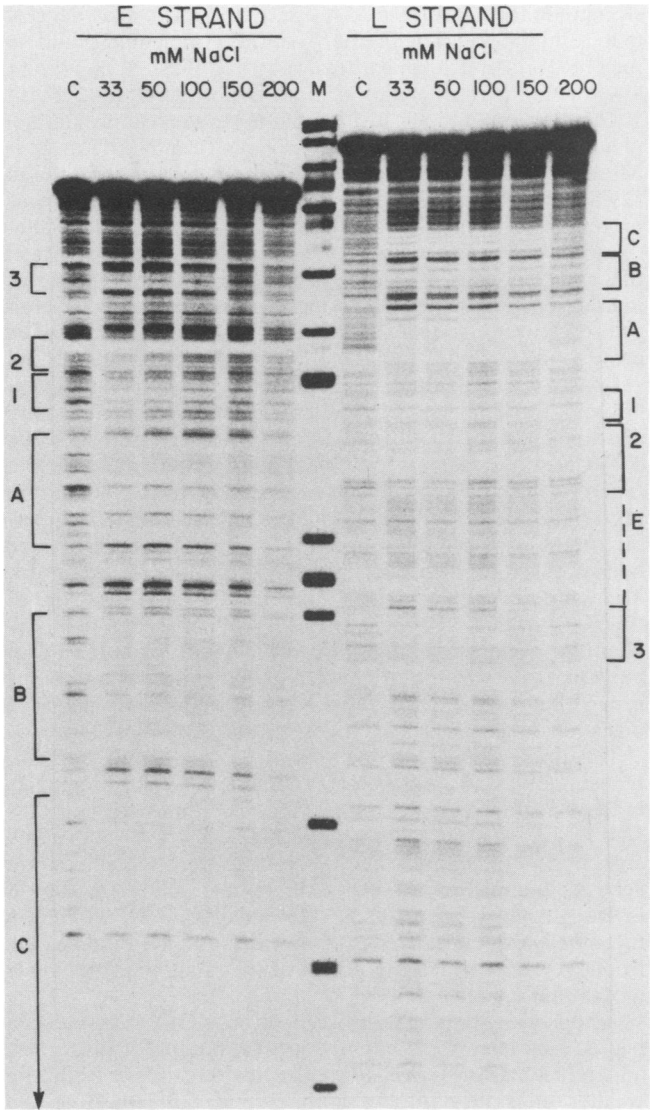

FIG. 7. Effect of NaCl concentration on large T antigen DNA binding. The effect of NaCl concentration on large T antigen binding was determined by incubating 1×10^{-3} to 2×10^{-3} pmol of 5'-^{32}P-labeled DNA fragments (Fig. 4) with 40 μl (ca. 8 pmol of monomer) of partially purified large T antigen in a total volume of 100 μl at different NaCl concentrations as shown. After DNase I digestion the samples were run on an 8% polyacrylamide-urea denaturing gel. The control lanes, marked C, were carried out at 100 mM NaCl in the absence of large T antigen. Because the change in NaCl concentration also affected the rate of DNase I digestion, the amount of enzyme was varied between 2.5 ng/ml and 33 mM and 10 ng/ml at 200 mM to obtain equivalent degress of digestion.

containing the viral regulatory region. Guadray et al. (13) were able to measure the affinity of the protein for its binding site and to approximate the genomic location of this site by use of several different restriction endonucleases and deletion mutants. In this way they identified a single high-affinity binding region immediately adjacent to the core replication origin on the early side. They suggested that two of the three copies of the hexanucleotide AGAGGC repeated within the binding region may be required for high-affinity interaction. No further independent binding sites were detected by this method, but it was found that sequences on both sides of the high-affinity site contributed to the binding strength (P. Gaudray and F. Cuzin, personal communication).

Several laboratories subsequently used the McKay (35) DNA immunoprecipitation technique to study the DNA-binding properties of large T antigen in unfractionated extracts. Our results, reported above, localized two independent binding sites for the large T antigen (Table 2; Fig. 2 and 3). A higher-affinity site in strain A2 DNA was included between nt 37 and nt 76. It did not overlap the replication origin by more than a few nucleotides. A lower-affinity site, located between nt 104 and nt 188 (Fig. 1 and 8), spanned the TATA box and the initiation sites for early region transcription and contained three TGAGGC hexanucleotides and one AGAGGC hexanucleotide. This site is an obvious candidate

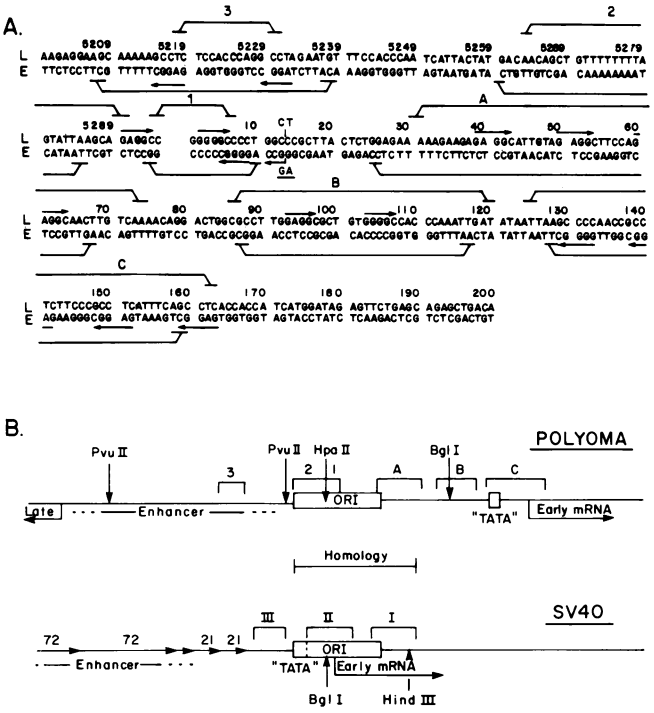

FIG. 8. Diagrammatic representation of the polyomavirus large T antigen-binding sites. (A) The DNA sequence of the polyomavirus regulatory region from nt 5200 to 200 (numbering system of Soeda et al. [54] as modified by Tyndall et al. [63]). The published sequence of strain A2 polyomavirus DNA (53) contains three errors in the ORI region. The A at nt 5 is actually a G, and two base pairs (CT) must be inserted between nt 13 and nt 14. These changes eliminate the differences between the A2 and A3 ORI sequence (9). To avoid confusion between the numbering system employed in previous papers and the work presented here, we have refrained from adjusting the numbering of the sequence to include the additional dinucleotide. Instead we represent the additional bases above and below the sequence shown in this figure. The areas protected from DNase I digestion by the large T antigen are bracketed. The pentanucleotide repeats 5'-(G/T)(A/G)GGC-3' are marked with arrows. The ORI core extends from nt 5269 to between nt 36 and nt 42 (26, 32). (B) The top line is a diagram of the binding sites of polyomavirus large T antigen shown with respect to the functional elements of the regulatory region: the early promoter. The ORI core sequences and the enhancer. The bottom line is a similar diagram of SV40 aligned with polyomavirus at the ORI homology. The position of the early RNA start sites and the enhancer sequences are shown, with arrows indicating the 72-bp and 21-bp repeats (taken from reference 37).

for the interaction mediating repression of transcription from the major early cap sites.

In strain A2 DNA, the higher-affinity site was the origin-proximal site. It was immunoprecipitated about four times more efficiently than the origin-distal site. Strain A3, a wild-type variant of polyomavirus has an 11-bp deletion within the origin-proximal T antigen-binding site which removes one of the three AGAGGC sequences (53). This reduced the apparent binding affinity by more than 10-fold and led us to speculate that at least three repeats must be present in a binding site for efficient immunoprecipitation under the highly selective conditions used.

The DNA immunprecipitation results we describe confirm and extend related reports (19, 41). Pomerantz et al. (41), using an A3-type strain, found two binding sites and reached similar conclusions for the location of the origin-distal site, although they failed to note the 10-fold reduction in binding efficiency caused by *Hph*I cleavage at nt 153. They assumed that the origin-proximal site included the replication origin, but did not present supporting evidence. It is clear from Fig. 3 and Table 2 that the ORI core sequence neither has high affinity for polyomavirus large T nor contributes significantly to the binding strength we measure for the adjacent site. One could postulate that binding of large T antigen to the origin is dependent on simultaneous occupation of the origin-proximal high-affinity site, but because recombinant plasmids lacking both of these T antigen binding sites replicate (26; our unpublished results), such cooperative binding, if it occurs, cannot be essential.

DNase I footprinting experiments. The availability of monoclonal antibodies directed against polyomavirus large T antigen allowed us to purify enough of the protein from wild-type virus-infected 3T6 cells for direct analysis of DNA binding by DNase I footprinting. The antibody used for immunoaffinity fractionation, αPyLT1, reacts with all of the large T antigen detectable in polyomavirus-infected cells (7). It is not known whether polyomavirus large T antigen oligomerizes in vivo or whether the forms of the protein that have been previously identified (7, 61) have different DNA binding activities. By using the αPyLT1 monoclonal antibody, the risk of enriching for particular subspecies was minimized.

Our initial footprint analyses (Fig. 5) demonstrated three closely spaced, but noncontiguous, regions of DNase I protection (designated A, B, and C in Fig. 8). The lengths of the protected regions were approximately 43, 35, and 38 bp, respectively. The locations of regions A and C agree with the deductions made from the DNA immunoprecipitation experiments. Region B, which spans the *Bgl*I site at nt 93, was not detectable in the indirect assay. We also note that footprint saturation experiments such as that shown in Fig. 5 indicated nearly equivalent protection of all three sites, whereas immunoprecipitation data implied a much greater affinity of T antigen for region A than for region C. These discrepancies suggest that the footprint experiments were relatively insensitive to differences in binding affinity because of the high reactant concentrations employed.

In addition we detected protection of three other regions, identified as 1, 2, and 3 in Fig. 8. Regions 1 and 2 together span most of the ORI core sequence. These studies have thus afforded the first evidence for an interaction between T antigen and the sequence necessary for DNA replication. Region 3 is near sequences known to have enhancer function. Approximately 10 times more T antigen was required to protect regions 1, 2, and 3 than to saturate regions A, B, and C (Fig. 5), and this protection was highly sensitive to ionic strength in the 33 to 150 mM range (Fig. 7). This may suggest that a minor fraction of the T antigen molecules bind to regions 1, 2, and 3, that the affinity for these sites is very low, or that a combination of both hypotheses pretains.

The division of the protected region between approximately nt 5260 and nt 10 into two subregions (1 and 2 in Fig. 8) is based on the hypersensitivity of the sequence around the nt 5295/1 junction to DNase I when the rest of the region was protected (Fig. 6, E-DNA strand). Region 1 is far smaller than the other protected regions and therefore we suspect that region 1 plus 2 together comprise a single domain of interaction. Further work, using more concentrated T antigen preparations and methods such as dimethyl sulfate protection for footprinting should resolve this issue.

Consensus recognition sequences. Repeated pentanucleotides of the consensus family 5′-(G>T)(A>G)GGC-3′ are distributed in the three regions of SV40 DNA contacted by SV40 large T antigen (4, 56–58) and are thought to represent recognition sequences. Soeda et al. (53) originally noted that the related hexanucleotide, AGAGGC, is repeated numerous times near the polyomavirus origin. These have since been implicated in polyomavirus large T antigen binding (13, 41). As noted in Fig. 1 and shown more clearly in Fig. 8, there are in fact 15 copies of the SV40 consensus pentanucleotide sequence arranged in five clusters of two to four tandem repeats within the 250 bp spanning the polyomavirus replication origin. The regions protected from DNase I digestion by large T antigen coincide with these clusters.

The three strong binding regions A, B, and C contain, respectively, 3, 2, and 4 pentanucleotides. The homologies in these regions extend to six bases, the consensus being 5′-(A=T)G(A>G)GGC. The repeat intervals are all between 9 and 11 bp, indicating that the recognition sequences are accessible on the same side of the double helix. Although it is likely that sequences outside the putative recognition hexanucleotide are also important in T-antigen binding, the frequency with which the repeat occurs at similar spacing is most remarkable.

A consensus polyomavirus repeat of 5-′G(A>G)GGC can be deduced from the comparison of all binding regions. It is apparent that the strong binding regions differ from the weaker ones in both the number and the length of the conserved sequence and in the spacing of the repeats. All of the results we have obtained by DNA immunoprecipitation can be understood under the hypothesis that a threefold repeat spaced at one turn of the helix is optimal for high-affinity binding. The different arrangement of the recognition sequences within regions 1 plus 2 and in region 3 may imply a different mode of interaction between the protein and the DNA.

Implications for papovavirus large T antigen function. One of the several significant differences between the genome organization of murine polyomaviruses and their simian counterpart, SV40, is the displacement of the polyomavirus early region promoter from the core ORI (Fig. 8). The early region cap sites and TATA box are included within the SV40 ORI sequence (14), and DNA essential for effective in vitro transcription lies immediately to the late side of the ORI (16, 37, 44). By contrast, the principal polyomavirus early region TATA box and cap sites are some 80 bp on the early side of the ORI, within DNA sequence not homologously represented in the SV40 genome. The sequences at the polyomavirus ORI, although very similar to the SV40 sequence, are not transcriptionally active in vitro, and indeed no specific sequence upstream of the TATA box is important for polyomavirus early transcription in vitro (23).

The large T antigens of polyomavirus and SV40 both repress early region transcription and activate DNA replication. SV40 large T antigen recognizes three adjacent regions spanning its ORI, called I, II, and III (Fig. 8), which bind noncooperatively and with different affinities in the order of I > II > III (55, 57, 58, 60). In vitro experiments have suggested that SV40 large T antigen-binding regions I and possibly II are involved in the repression of early transcription, which initiates within region II (17, 44). Considerable genetic evidence demonstrates that the DNA within SV40 region II is essential for viral DNA replication, whereas sequences within region I has a quantitative, if nonessential, role (8, 34, 35, 38, 49–51).

Coincident with the displacement of the polyomavirus early promoter from the ORI is a major shift in the location of high-affinity large T antigen-binding regions. Region A of polyomavirus occurs at a position precisely analogous to SV40 region I and also appears to be the highest-affinity site, at least in strain A2 DNA. Sequences within polyomavirus region A, like that comprising SV40 region I, influence the quantitative efficiency of DNA replication, but are not essential. Regions B and C, unlike SV40 II and III, extend toward and across the early transcriptional regulatory sequences rather than in the replication ORI direction (Fig. 8). The location of the major binding regions strongly suggests that these are primarily involved in transcriptional regulation.

We have detected polyomavirus T antigen binding to the ORI (region 1 and 2), at a position very similar to that of SV40 region II, but this interaction is either much weaker than that to regions A, B, and C or is the property of a small subset of large T-antigen molecules. Of the polyomavirus large T antigen-binding regions, only 1 and 2 occur within the sequence essential for DNA replication. We therefore suspect that the minor interaction we have detected is that which is principally involved in the replication initiation function of the protein. This topological segregation of functional binding sites distinguishes polyomavirus from SV40 and should prove useful in the further study of the multifunctional interaction between the viral protein and its genome.

The binding at region 3 occurs outside the ORI core within the viral enhancer region, but within a sequence that can be deleted with no phenotypic consequence (32). Region 3, however, includes the sequence mutated and duplicated in polyomavirus variants able to grow in the embryonal carcinoma cell line F9 (10, 27, 47). Furthermore this sequence has been shown to be necessary in cis for both viral DNA replication and early gene expression in embryonal carcinoma cells (11, 31). Future experiments will test whether large T antigen binding at region 3 is involved in enhancer function, and whether region 3 plays a role in the activation of later gene expression.

ACKNOWLEDGMENTS

This work was initiated in the Transcription Laboratory of the Imperial Cancer Research Fund, London.

We are grateful to Steve Dilworth and Furzana Chaudry for their advice and to Steve Dilworth for the monoclonal antibodies used. Steve Lupton, Parmjit Jat, and Chiara Tyndall kindly provided deletion mutants. We would also like to thank Peter Tegtmeyer and Angelo DeLucia, whose interest and comments stimulated our search for an ORI interaction.

LITERATURE CITED

1. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
2. **Cowie, A., P. Jat, and R. Kamen.** 1982. Determination of sequences at the capped 5′ ends of polyomavirus early region transcripts synthesized *in vivo* and *in vitro* demonstrates an unusual microheterogeneity. J. Mol. Biol. **159**:225–255.
3. **Cowie, A., C. Tyndall, and R. Kamen.** 1981. Sequences at the capped 5′ ends of polyomavirus late region mRNAs: an example of extreme terminal heterogeneity. Nucleic Acids Res. **9**:6305–6322.
4. **DeLucia, A. L., B. A. Lewton, R. Tjian, and P. Tegtmeyer.** 1983. Topography of simian virus 40 A protein-DNA complexes: arrangement of pentanucleotide interaction sites at the origin of replication. J. Virol. **46**:143–150.
5. **deVillers, J., and W. Schaffner.** 1981. A small segment of polyomavirus DNA enhances the expression of the clone β-globin gene over a distance of 1400 base pairs. Nucleic Acids Res. **9**:6251–6264.
6. **Dilworth, S. M., A. Cowie, R. Kamen, and B. E. Griffin.** 1984. DNA binding activity of polyomavirus large T antigen. Proc. Natl. Acad. Sci. U.S.A. **81**:1941–1945.
7. **Dilworth, S. M., and B. E. Griffin.** 1982. Monoclonal antibodies against polyomavirus tumor antigens Proc. Natl. Acad. Sci. U.S.A. **79**:1959–1963.
8. **DiMaio, D., and D. Nathans.** 1982. Regulatory mutants of SV40. Effect of mutations at a T antigen binding site on DNA replication and expression of viral genes. J. Mol. Biol. **156**:531–548.
9. **Friedman, T., A. Esty, P. LaPorte, and D. Deininger.** 1979. The nucleotide sequence and genome organization of the polyoma early region: extensive nucleotide and amino homology with SV40. Cell **17**:715–724.
10. **Fujimura, F. K., P. L. Deininger, T. Friedmann, and E. Linney.** 1981. Mutation near the polyoma DNA replication origin permits productive infection of F9 embryonal carcinoma cells. Cell **23**:809–814.
11. **Fujimura, F. K., and E. Linney.** 1982. Polyoma mutants that productively infect F9 embryonal carcinoma cells do not rescue wild type polyoma in F9 cells. Proc. Natl. Acad. Sci. U.S.A. **79**:1479–1483.
12. **Galas, D. J., and A. Schmitz.** 1978. DNase footprinting: a simple method for detection of protein-DNA binding specificity. Nucleic Acids Res. **5**:3157–3710.
13. **Gaudray, P., C. Tyndall, R. Kamen, and F. Cuzin.** 1981. The high affinity binding site on polyomavirus DNA for the viral large T antigen. Nucleic Acids Res. **21**:5697–5710.
14. **Ghosh, P. K., and P. Lebowitz.** 1981. Simian virus 40 early mRNAs contain multiple 5′ termini upstream and downstream from a Hogness-Goldberg sequence; a shift in 5′ termini during the lytic cycle is mediated by large T antigen. J. Virol. **40**:224–240.
15. **Green, M. R., R. Treisman, and T. Maniatis.** 1983. Transcriptional activation of cloned human β-globin genes by viral immediate-early gene products. Cell **35**:137–148.
16. **Hansen, U., and P. Sharp.** 1983. Sequences controlling *in vitro* transcription of SV40 promoters. EMBO J. **2**:2293–2303.
17. **Hansen, U., D. G. Tenen, D. M. Livingston, and P. A. Sharp.** 1981. T antigen repression of SV40 early transcription from two promoters. Cell **27**:603–612.
18. **Harlow, E., L. V. Crawford, D. C. Pim, and N. M. Williamson.** 1981. Monoclonal antibodies specific for simian virus 40 tumor antigens. J. Virol. **39**:861–869.
19. **Hayday, A. C., F. Chaudry, and M. Fried.** 1983. Loss of polyomavirus infectivity as a result of a single amino acid change in a region of polyomavirus large T antigen which has extensive amino acid homology with simian virus 40 large T antigen. J. Virol. **45**:693–699.
20. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26**:365–369.
21. **Houweling, A., P. J. van den Elsen, and A. J. van der Eb.** 1980. Partial transformation of primary rat cells by the leftmost 4.5% fragment of adenovirus 5 DNA. Virology **105**:537–550.
22. **Imperiale, M. J., L. T. Feldman, and J. R. Nevins.** 1983. Activation of gene expression by adenovirus and herpes virus regulatory genes acting in *trans* and by a *cis*-acting adenovirus enhancer element. Cell **35**:127–136.

23. **Jat, P., U. Novak, A. Cowie, C. Tyndall, and R. Kamen.** 1982. DNA sequences required for specific and efficient initiation of transcription at the polyomavirus early promoter. Mol. Cell. Biol. **2**:737–751.
24. **Jones, K. A., and T. Tjian.** 1984. Essential contact residues within SV40 large T antigen binding sites I and II identified by alkylation-interference. Cell **36**:155–162.
25. **Kamen, R., P. Jat, R. Treisman, J. Favaloro, and W. R. Folk.** 1982. 5′ termini of polyomavirus early region transcripts synthesized *in vivo* by wild type and viable deletion mutants. J. Mol. Biol. **159**:189–224.
26. **Katinka, M., and M. Yaniv.** 1983. DNA replication origin of polyomavirus: early proximal boundary. J. Virol. **47**:244–248.
27. **Katinka, M., M. Yaniv, M. Vasseur, and D. Biangy.** 1980. Expression of polyoma early functions in mouse embryonal carcinoma cells depends upon sequence rearrangements in the beginning of the late region. Cell **20**:393–399.
28. **Keller, J. M., and J. C. Alwine.** 1984. Activation of the simian virus 40 late promoter: direct effect of T antigen in the absence of viral DNA replication. Cell **36**:381–389.
29. **Land, H., L. F. Parada, and R. A. Weinberg.** 1983. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. Nature (London) **304**:596–602.
30. **Lewton, B. A., A. L. DeLucia, and P. Tegtmeyer.** 1984. Binding of simian virus A protein to DNA with deletions at the origin of replication. J. Virol. **49**:9–13.
31. **Linney, E., and S. Donerly.** 1983. DNA fragments from F9 PyEC mutants increase expression of heterologous genes in transfected F9 cells. Cell **35**:693–699.
32. **Luthman, H., M. G. Nilsson, and G. Magnusson.** 1982. Noncontinuous segments of the polyoma genome required in *cis* for DNA replication. J. Mol. Biol. **161**:533–550.
33. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labelled DNA with base specific chemical cleavages. Methods Enzymol. **65**:499–560.
34. **McKay, R., and D. DiMaio.** 1981. Binding of an SV40 T antigen-related protein to the DNA of SV40 regulatory mutants. Nature (London) **289**:810–813.
35. **McKay, R. D. G.** 1981. Binding of an SV40 T antigen related protein to DNA. J. Mol. Biol. **145**:471–488.
36. **Muller, W. J., C. R. Mueller, A.-M. Mes, and J. A. Hassel.** 1983. Polyoma virus origin for DNA replication comprises multiple genetic elements. J. Virol. **47**:586–599.
37. **Myers, R. M., D. C. Rio, A. K. Robbins, and R. Tjian.** 1981. SV40 gene expression is modulated by the cooperative binding of T-antigen to DNA. Cell **25**:373–384.
38. **Myers, R. M., and R. Tjian.** 1980. Construction and analysis of SV40 origins defective in tumor binding and DNA replication. Proc. Natl. Acad. Sci. U.S.A. **77**:6491–6495.
39. **Nevins, J. R.** 1981. Mechanism of activation of early viral transcription by the adenovirus E1A gene product. Cell **26**:213–220.
40. **O'Farrell, P.** 1981. Replacement synthesis method of labelling DNA fragments. Bethesda Research Laboratories Focus **3**:1.
41. **Pomerantz, B. J., C. R. Mueller, and J. A. Hassel.** 1983. Polyoma virus large T antigen binds independently to multiple, unique regions on the viral genome. J. Virol. **47**:600–610.
42. **Rassoulzadegan, M., A. Cowie, A. Carr, N. Glaichenhaus, R. Kamen, and F. Cuzin.** 1982. The roles of individual polyomavirus early proteins in oncogenic transformation. Nature (London) **300**:713–718.
43. **Rassoulzadegan, M., Z. Naghashfar, A. Cowie, M. Grisoni, R. Kamen, and F. Cuzin.** 1983. Expression of the large T antigen of polyomavirus promotes the establishment in culture of "normal" rodent fibroblast cell lines. Proc. Natl. Acad. Sci. U.S.A. **80**:4354–4358.
44. **Rio, D. C., and R. Tjian.** 1983. SV40 T antigen binding site mutations that affect autoregulation. Cell **32**:1227–1240.
45. **Ruley, H. E.** 1983. Adenovirus early region 1A enables viral and cellular transforming genes to transform primary cells in culture. Nature (London) **304**:602–606.
46. **Schaffhausen, B.** 1982. Transforming genes and gene products of polyoma and SV40. Crit. Rev. Biochem. **13**:215–286.
47. **Sekikawa, K., and A. J. Levine.** 1981. Isolation and characterization of polyoma host range mutants that replicate in nullipotential embryonal carcinoma cells. Proc. Natl. Acad. Sci. U.S.A. **78**:1100–1104.
48. **Shiroki, K., H. Handa, H. Shimojo, S. Yano, S. Ojima, and K. Fukinaga.** 1977. Establishment and characterization of rat cell lines transformed by restriction endonuclease fragments of adenovirus 12 DNA. Virology **82**:462–471.
49. **Shortle, D., and D. Nathans.** 1978. Local mutagenesis; a method for generating viral mutants with base substitutions in preselected regions of the viral genome. Proc. Natl. Acad. Sci. U.S.A. **75**:2170–2174.
50. **Shortle, D., and D. Nathans.** 1979. Regulatory mutants of simian virus 40 constructed mutants with base substitutions at the origin of DNA replication. J. Mol. Biol. **131**:801–817.
51. **Shortle, D., and D. Nathans.** 1979. Mutants of SV40 with base substitutions at the origin of DNA replication. Cold Spring Harbor Symp. Quant. Biol. **43**:663–668.
52. **Soberon, X., L. Covarrubias, and F. Bolivar.** 1980. Construction and characterization of new cloning vehicles. Gene **9**:287–305.
53. **Soeda, E., J. R. Arrand, N. Smolar, and B. E. Griffin.** 1979. Sequence from early region of polyomavirus DNA containing viral replication origin and encoding small, middle and (part of) large T antigens. Cell **17**:357–370.
54. **Soeda, E., J. R. Arrand, N. Smolar, J. E. Walsh, and B. E. Griffin.** 1980. Coding potential and regulatory signals of the polyomavirus genome. Nature (London) **283**:445–453.
55. **Tegtmeyer, P., B. Anderson, S. B. Shaw, and V. G. Wilson.** 1981. Alternative interactions of the SV40 A protein with DNA. Virology **115**:75–87.
56. **Tegtmeyer, P., B. A. Lewton, A. L. DeLucia, V. G. Wilson, and K. Ryder.** 1983. Topography of SV40 A protein-DNA complexes: arrangement of protein bound to the origin of replication. J. Virol. **46**:151–161.
57. **Tenen, D. G., L. L. Haines, and D. M. Livingston.** 1982. Binding of an analogue of the simian virus 40 T antigen to wild-type and mutant viral replication origins. J. Mol. Biol. **157**:473–492.
58. **Tenen, D. G., T. S. Taylor, L. L. Haines, M. K. Bradley, R. G. Martin, and D. M. Livingston.** 1983. Binding of simian virus 40 large T antigen from virus infected monkey cells to wild-type and mutant viral replication origins. J. Mol. Biol. **168**:791–808.
59. **Tjian, R.** 1978. The binding site of SV40 DNA for a T antigen-related protein. Cell **13**:165–179.
60. **Tjian, R.** 1981. Regulation of viral transcription and DNA replication by the SV40 large T antigen. Curr. Top. Microbiol. Immunol. **93**:5–24.
61. **Turler, H., and C. Solomon.** 1977. Characterization of polyoma T antigens. INSERM **69**:131–144.
62. **Twigg, A. J., and D. Sherrat.** 1980. Trans-complementable copy-number mutants of plasmid ColE1. Nature (London) **283**:216–218.
63. **Tyndall, C., G. Lamantia, C. M. Thacker, J. Favaloro, and R. Kamen.** 1981. A region of the polyomavirus genome between the replication origin and late protein coding sequences is required in *cis* for both early gene expression and viral DNA replication. Nucleic Acids Res. **9**:6231–6250.
64. **Weber, F., J. deVilliers, and W. Schaffner.** 1984. An SV40 "enhancer trap" incorporates exogenous enhancers or generates enhancers from its own sequences. Cell **36**:983–992.