JOURNAL OF VIROLOGY, Sept. 1983, p. 600–610
0022-538X/83/090600-11$02.00/0

Vol. 47, No. 3

Polyomavirus Large T Antigen Binds Independently to Multiple, Unique Regions on the Viral Genome

BETSY J. POMERANTZ, CHRISTOPHER R. MUELLER, AND JOHN A. HASSELL*
Department of Microbiology and Immunology, McGill University, Montreal, Quebec H3A 2B4, Canada

Received 10 March 1983/Accepted 7 June 1983

To map the polyomavirus large T antigen binding sites on the viral genome we employed a quantitative immunoassay. Defined, radiolabeled fragments of the viral genome were reacted with crude nuclear extracts prepared from lytically infected mouse 3T6 cells, and the fragments bound by large T antigen were immunoprecipitated with anti-T serum and Formalin-fixed *Staphylococcus aureus*. The immunoprecipitated DNA was then analyzed by gel electrophoresis and autoradiography. By the use of a variety of restriction endonuclease-generated fragments of wild-type and mutant viral DNAs, the region of high-affinity binding was localized to a 153-base-pair stretch between nucleotides 5292 and 152. At least two independent binding sites lie within this region, one upstream and the other downstream of the *BglI* site at nucleotide 87. One of the binding sites is located within sequences required in *cis* for DNA replication; the other overlaps the TATA box and cap sites of the early transcription unit. The two sites share a common sequence, A/TGAGGC–N4/5–A/TGAGGC, which may serve as the recognition sequence for large T antigen.

Like other members of the papovaviruses, polyomavirus has served as a model to study gene structure and expression, DNA replication, and tumorigenesis. The entire sequence of the double-stranded, circular genome has been determined, and the boundaries of its two transcription units and the six viral proteins they encode have been largely elucidated (11). Accordingly, many recent studies have focused on the viral genetic control signals and the interaction of cellular and viral proteins with these regulatory regions. The large T antigen, encoded by the early transcription unit of polyomavirus, is an example of a regulatory protein that likely interacts with these genetic control signals. Polyomavirus large T antigen is comprised of 785 amino acids and has a calculated molecular weight of 87,991 (5, 33), but direct measurements of the molecular weight of the protein under denaturing conditions have led to estimates that range from 81,000 to 102,000 (38). Large T antigen is found almost exclusively in the nucleus of infected or transformed cells in close association with DNA (34). Consistent with this finding is the observation that the protein displays affinity for double-stranded DNA in vitro (6) and binds with specificity to a single site on the viral genome (7). An ATPase activity has been ascribed to large T antigen; however, the protein has not been purified sufficiently to assign this activity to it unequivocally (6).

Genetic analyses of the function of large T antigen have revealed that it is required for the initiation of each new round of viral DNA replication (4) and the repression of early transcription (2). The latter results in the reduced synthesis of the products of the early transcription unit, including large T antigen, and is an example of autogenous regulation (10). By analogy to simian virus 40 (SV40) (15, 27), it is thought that autogenous regulation results from the binding of polyomavirus large T antigen to promoter-proximal sequences, thereby blocking transcription initiation. Large T antigen must also interact with viral DNA to initiate DNA replication. This has been most elegantly demonstrated for SV40 by the use of mutants bearing *cis*-acting mutations that map within the origin of DNA replication (9, 31) and revertants of these origin-defective mutants that carry second-site mutations within large T antigen coding sequences (30), and by measurement of the physical interaction between SV40 large T antigen and viral DNA (36). Knowledge of the location of the SV40 T antigen binding sites on SV40 DNA, the limits of the origin of replication, and the position of the early promoter has permitted the formulation of models to explain autogenous regulation and replication initiation (15, 27, 37). By comparison, we know very little about the interaction of polyomavirus large T antigen with viral DNA or the boundaries of the replication origin and the early promoter. These studies

have lagged in part because abundant quantities of polyomavirus large T antigen are not available. Recently, however, McKay (24) described an immunoassay that permits the measurement of the direct interaction of SV40 large T antigen present in crude cell extracts with viral DNA. We have adopted this assay to map the polyomavirus large T antigen binding sites on the viral genome. Moreover, we and others have defined the boundaries of the polyomavirus early promoter (13; C. Mueller, A.-M. Mes, and J. A. Hassell, data not shown) and the sequences required in *cis* for polyomavirus DNA replication (26), thereby permitting us to examine the generality of the models proposed for autoregulation and replication initiation by SV40 large T antigen. Here we report on the spatial organization of the polyomavirus large T antigen binding sites, relative to the aforementioned genetic control signals, and discuss their arrangement in comparison to those found in SV40 DNA. Polyomavirus large T antigen binds independently to at least two closely spaced sites on the viral genome. One of the T antigen binding sites resides within a region of DNA required in *cis* for viral DNA replication and is likely the same site identified previously (7). The other site overlaps with sequences of the early promoter, including the TATA box and start sites for the early mRNAs.

MATERIALS AND METHODS

Cells and viruses. Cells were grown on plastic dishes in Dulbecco modified Eagle medium supplemented with 10% (vol/vol) fetal calf serum and antibiotics, and maintained at 37°C in a humidified CO_2 atmosphere (12). Polyomavirus stocks were prepared by infection of primary baby mouse kidney cultures at low multiplicities of infection (0.01 to 0.1 PFU per cell). The cells were harvested 10 to 14 days postinfection, and the virus was released by freezing and thawing the cells three times. Titers of about 10^9 PFU/ml were normally obtained.

Construction of recombinant plasmids. All plasmids were propagated in *Escherichia coli* HB101. Their structures are shown in the various figures. The construction of pPH1-8 (14) and pPBR2 (25) has been described previously. pPBR1 was constructed by ligation of the *Bam*HI (nucleotide 4632)-to-*Eco*RI (nucleotide 1560) fragment of polyomavirus DNA to *Bam*HI/*Eco*RI-cleaved pBR322 DNA. Note that pPBR2 and pPBR1 carry identical viral sequences but different plasmid sequences. The plasmid sequences within pPBR2 are derived from pMK16.1 (8, 19). The recombinant plasmid pPin7 is a derivative of pPBR2 in which the sole *Bgl*I site within the viral sequences has been converted to a *Hin*dIII site. This was accomplished by *Bgl*I cleavage of pPBR2 DNA, removal of the 3′ projections by incubation with the Klenow fragment of DNA polymerase I, and addition of a synthetic *Hin*dIII linker 10 base pairs (bp) long. pdPBHp was constructed by ligation of the *Bam*HI (nucleotide 4632)-to-*Hph*I (nucleotide 152) fragment of polyomavirus DNA to *Bam*HI/*Hin*dIII-cleaved pML-2 DNA (20). After ligation of the *Bam*HI ends of the vector and the viral insert, the remaining ends were treated with the Klenow fragment of DNA polymerase I to fill in the *Hin*dIII 5′ projection (on the vector) and to remove the *Hph*I 3′ projection on the viral sequences. Ligation across these blunt ends resulted in the construction of pdPBHp. pdPBBg was obtained by cloning the viral *Bam*HI-to-*Hin*dIII fragment of pPin7 within the large *Bam*HI-to-*Hin*dIII fragment of pML-2. The resulting recombinant contains polyomavirus sequences between nucleotides 4632 (*Bam*HI) and 90 (*Hin*dIII, formerly a *Bgl*I site). To confirm the structure of the recombinant plasmids described above, the virus-plasmid joints and the virus-linker junctions were sequenced by the chemical method of Maxam and Gilbert (22).

Preparation of DNA and its modification. Recombinant plasmid DNAs were isolated from bacteria and purified by CsCl density gradient centrifugation. Digestions of these DNAs with restriction endonucleases were performed according to the conditions specified by the manufacturers of the enzymes.

To end label the DNA fragments, 100 ng of restriction enzyme-cleaved DNA was incubated in a volume of 50 μl with 10 mM Tris-hydrochloride (pH 7.6), 5 mM $MgCl_2$, 1 mM dithiothreitol, 25 μCi of the appropriate α-^{32}P-deoxynucleotide triphosphate (3,000 Ci/mmol), and 5 U of the Klenow fragment of *E. coli* DNA polymerase I for 1 h at 15°C. The reaction was terminated by phenol extraction, and the aqueous phase was chromatographed on a Sephadex G-50 column to separate the unincorporated deoxynucleotide triphosphates from the labeled DNA fragments. Typically the specific activity of the labeled DNA fragments averaged 10^8 cpm/μg.

Preparation of crude nuclear extract. Nuclear extracts were prepared from 3T6 cells infected with polyomavirus essentially as described by McKay (24). From 0.5×10^9 to 1.0×10^9 3T6 cells growing on the surface of plastic petri dishes were infected with polyomavirus at a multiplicity of infection of about 10. At 40 to 46 h postinfection at 37°C, the cells were harvested by scraping and then washed twice with ice-cold phosphate-buffered saline solution. The cells were suspended in lysis buffer (10 mM Tris-hydrochloride [pH 7.4], 10 mM NaCl, 1.5 mM $MgCl_2$, 0.5 mM dithiothreitol, and 100 μg of phenylmethylsulfonyl fluoride per ml) at a density of 5×10^7 cells per ml. They were allowed to swell at 4°C for 10 min and then were lysed by 10 to 15 strokes with a type B pestle in a glass Dounce homogenizer. The nuclei were pelleted by centrifugation (2,000 rpm for 10 min at 4°C) and washed in lysis buffer. The nuclei were then suspended in nuclear extraction buffer (10 mM Tris-hydrochloride [pH 8.0], 300 mM NaCl, 0.5 mM dithiothreitol, and 100 μg of phenylmethylsulfonyl fluoride per ml) and incubated at 4°C for 5 to 15 min. The nuclei were pelleted by centrifugation, and the supernatant was used as a source of large T antigen. These preparations were stored at −70°C until used. We have tested various cell fractions for binding activity (cytoplasmic extract as well as 150, 300, and 600 mM NaCl nuclear extract). Some activity is eluted from the nuclei with 150 mM NaCl, and very little activity is found in the cytoplasmic fraction. The use of 600 mM NaCl leads to nuclear lysis and the release of proteins that bind

nonspecifically to the DNA. No specific binding activity was found in uninfected 3T6 cells (data not shown).

Immunoassay. The conditions of the immunoassay were essentially those described by McKay (24). Briefly, reactions were carried out in a 1-ml volume containing 20 mM $NaPO_4$ (pH 7.0), 2 mM dithiothreitol, 0.01% (wt/vol) bovine serum albumin, 0.1 mM EDTA, 0.05% (vol/vol) Nonidet P-40, 40 μg of phenylmethylsulfonyl fluoride per ml, and 3.0% (vol/vol) dimethyl sulfoxide. From 1 to 28 ng of labeled DNA fragments and various amounts of the crude T antigen extract were also included. After a 1-h incubation at 25°C, 2 μg of unlabeled rat cellular DNA was added, and incubation was continued for 20 min at 25°C. The inclusion of nonradioactive DNA reduces the nonspecific binding of large T antigen to the labeled DNA. The large T antigen-DNA complexes were immunoprecipitated by sequential incubation with antibodies directed against the polyomavirus T antigens and Formalin-fixed *Staphylococcus aureus*. Polyclonal antiserum directed against polyomavirus large T antigen was obtained from ascites fluid of brown Norwegian rats bearing tumors induced by polyomavirus-transformed rat cells. The antiserum was not purified before use. The latter reactions were performed at 25°C for 20 min in each instance. The immunocomplexes were then washed twice with 1 ml of 10 mM Tris-hydrochloride (pH 8.0)–150 mM NaCl–0.5% Nonidet P-40. Finally, the DNA was released from the immunocomplexes with 1% (wt/vol) sodium dodecyl sulfate–10 mM EDTA and deproteinized with phenol and chloroform-isoamyl alcohol (24:1, vol/vol), and a portion was electrophoresed through an agarose gel. Before autoradiography the gels were immersed in ethanol for 30 min and dried under vacuum with heating. Generally, the dried gels were exposed to Kodak XAR-5 film for 3 to 4 h in cassettes with Du Pont Lightning-Plus intensifying screens.

RESULTS

Specific binding of large T antigen to polyomavirus DNA. Because the amount of pure polyomavirus large T antigen obtainable from infected cells is insufficient to study its interaction with DNA by traditional binding assays (i.e., nitrocellulose filter binding), we used a technique that permits measurement of DNA-protein binding in complex cell extracts (24). This assay makes use of the fact that DNA bound to a particular protein can be separated from free DNA by immunoprecipitation with antibody specific to that protein. In brief, defined radiolabeled DNA fragments are reacted with crude nuclear preparations of large T antigen, and the fragments bound by T antigen are immunoprecipitated with an antiserum (isolated from rats bearing tumors induced by polyomavirus-transformed cells) and Formalin-fixed *S. aureus*. The immunocomplexes are washed, and the eluted DNA is analyzed by gel electrophoresis and autoradiography. By utilizing a variety of restriction endonuclease-generated fragments of viral DNA, whose sites of cleavage on the viral genome are known, a map can be established of the high-affinity large T antigen binding sites.

To test the applicability of this method for measuring the binding of polyomavirus large T antigen to specific viral DNA fragments, we incubated increasing amounts of nuclear extract from lytically infected mouse cells with a fixed quantity of end-labeled DNA fragments obtained by *Hin*fI cleavage of a recombinant plasmid that carries the *Hin*dIII-1 fragment of the viral genome. This segment of DNA contains those viral sequences between nucleotides 3918 and 1656. We have used the nucleotide numbering scheme suggested by Soeda et al. (33) and report nucleotide numbers in the clockwise direction on the conventional polyomavirus physical map. The results are shown in Fig. 1. The antiserum selectively precipitates the *Hin*fI-4 fragment of pPH1-8 DNA, which is solely composed of polyomavirus sequences (from nucleotide 5073 to 385). No other fragments are selectively precipitated, and increasing the extract concentration within the reaction mixture results in increased precipitation of the *Hin*fI-4 fragment (Fig. 1). In the absence of nuclear extract a fraction of all the fragments is retained in the immunocomplexes (Fig. 1). This is due to the presence of DNA-binding proteins in the antiserum preparations (data not shown). Moreover, large T antigen binds nonspecifically to DNA fragments, and this leads to their immunoprecipitation according to size. The largest fragments are preferentially precipitated under these conditions (Fig. 1). Two independent lines of evidence suggest that it is polyomavirus large T antigen, and not middle or small T antigen, that binds specifically to the viral DNA. First, immunoprecipitation of [^{35}S]methionine-labeled proteins from the nuclear extract used as a source of large T antigen revealed the presence of large T antigen and small quantities of small T antigen, but no middle T antigen (data not shown). Middle T antigen is not solubilized from the membrane fraction of infected cells with the purification procedure used (17) and hence does not appear in either the cytoplasmic or nuclear fraction. It is unlikely that small T antigen binds specifically to the viral genome because the cytoplasmic fraction from polyomavirus-infected 3T6 cells, which is enriched for small T antigen by comparison to large T antigen, is incapable of binding to viral DNA with specificity (data not shown). Second, monoclonal antibodies that recognize antigenic determinants unique to large T antigen immunoprecipitate the same DNA fragments that are precipitated with conventional antiserum (data not shown). These data suggest that large T antigen mediates the immunoprecipitation of specific viral DNA fragments.

To map the T antigen binding sites on the viral genome more precisely and to determine whether binding to DNA occurs from other areas of the genome, we performed the binding assay with polyomavirus DNA isolated from infected cells and subgenomic, cloned viral DNA fragments cleaved with a variety of restriction endonucleases. *Hpa*II cleavage of polyomavirus DNA yields eight fragments whose locations are shown in Fig. 2A. In the presence of nuclear extract from infected cells the antiserum selectively precipitates the *Hpa*II-5 or -6 fragment of viral DNA, or both (Fig. 2A). These fragments comigrated in this gel. Analysis of the immunoprecipitated fragments on polyacrylamide gels clearly revealed that large T antigen binds to the *Hpa*II-5 but not to the *Hpa*II-6 fragment of polyomavirus DNA (data not shown). Additional evidence that large T antigen binds to the *Hpa*II-5 fragment is shown in Fig. 2B. In this set of reactions, end-labeled, *Hpa*II-cleaved pPH1-8 DNA was used as the substrate in the binding reactions. The *Hpa*II-5 fragment of polyomavirus DNA (the fifth largest fragment in the *Hpa*II digest of pPH1-8 DNA) is specifically retained in immunocomplexes with antiserum and *S. aureus*. This fragment is located between nucleotides 5291 and 399 (Fig. 2). To more narrowly define the large T antigen binding region we employed the recombinant plasmid pPBR1 and the restriction endonuclease *Hph*I. pPBR1 is composed of the small *Bam*HI-to-*Eco*RI fragment of polyomavirus DNA (nucleotides 4632 to 1560) and the large *Bam*HI-to-*Eco*RI fragment of pBR322 DNA (Fig. 2). Reaction of the crude preparation of polyomavirus large T antigen with the end-labeled fragments obtained by *Hph*I cleavage of pPBR1 DNA yielded one fragment after immunoprecipitation (Fig. 2C). This DNA fragment contains those polyomavirus sequences that lie between the *Bam*HI cleavage site at nucleotide 4632 and the *Hph*I cleavage site at nucleotide 153 (Fig. 2C). The data obtained by performing the binding assay with the *Hin*fI (Fig. 1), *Hpa*II, and *Hph*I cleavage products of the various recombinant plasmids (summarized in Fig. 2D) allow us to map a high-affinity, large T antigen binding site between nucleotides 5292 and 153 on the polyomavirus genome. This 154-bp fragment spans the sites where DNA replication and early transcription initiate.

Multiple large T antigen binding sites within a restricted area of the viral genome. In SV40 DNA there are three closely spaced large T antigen binding sites that together span a region of about 120 bp in length (29, 35, 36). SV40 large T antigen binds to these sites with different affinities (23, 27, 28). Thus, large T antigen binding to one site (binding site I) precedes and may facilitate the subsequent binding of more protein to sites II and III. Moreover, large T antigen binds very poorly or not at all to binding sites II and III when binding site I has been altered by mutation (23, 27). To ascertain whether the large T antigen binding region in polyomavirus DNA was organized in a related way, we measured the binding of polyomavirus large T antigen to DNA fragments containing subsets of the sequences known to encompass the high-affinity large T antigen binding region. This was accomplished by the use in the DNA binding immunoassays of DNA substrates that carried only part of the large T antigen binding region. To perform this

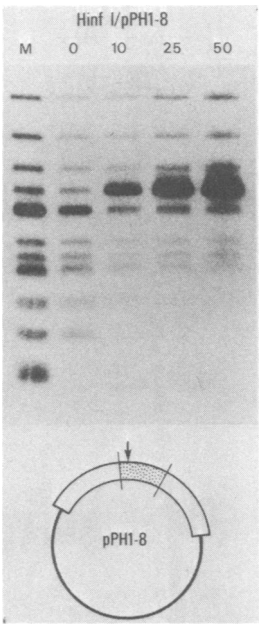

FIG. 1. Immunoprecipitation of the *Hin*fI restriction fragments of pPH1-8 after reaction with increasing quantities of nuclear extract from polyomavirus-infected cells. pPH1-8 DNA was cleaved to completion with *Hin*fI, and the fragments were end labeled and reacted with increasing quantities of nuclear extract from polyomavirus-infected 3T6 cells. In these reactions 10 ng of DNA fragments per ml was incubated with 0, 10, 25 and 50 μl of extract and processed as described in the text. M indicates the marker lane, which shows the *Hin*fI fragments of pPH1-8 DNA used as substrate. Several fragments comigrate in this gel; they include the *Hin*fI-5, -6, and -7 fragments and the *Hin*fI-10 and -11 fragments of pPH1-8 DNA. The circular diagram below the autoradiogram depicts the recombinant plasmid pPH1-8. Polyomavirus sequences are represented by the boxed area, and plasmid sequences are indicated by the thick line. Lines intersecting the circle represent those *Hin*fI cleavage sites that bracket the *Hin*fI-4 fragment, and stippled areas represent fragments which are specifically immunoprecipitated. The arrow indicates the junction between polyomavirus nucleotides 5292 and 1.

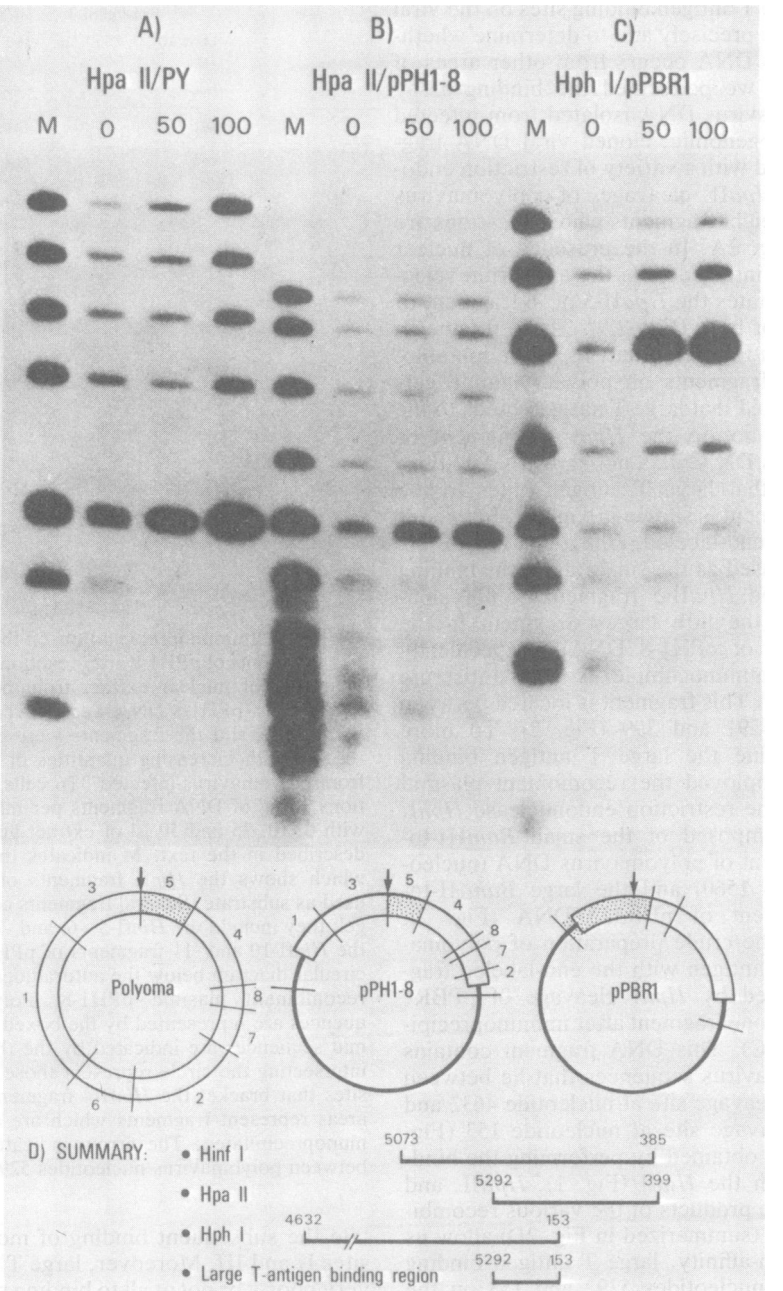

FIG. 2. Mapping the polyomavirus DNA sequences to which polyomavirus large T antigen binds. Immunoprecipitation of the various DNA fragments was performed as described for Fig. 1. Different quantities of nuclear extract from polyomavirus-infected cells were reacted with 28 ng of the *Hpa*II fragments of polyomavirus DNA per ml (A), 10 ng of the *Hpa*II fragments of pPH1-8 DNA per ml (B), and 25 ng of the *Hph*I fragments of pPBRI DNA per ml (C). In (D) the polyomavirus DNA fragments which were immunoprecipitated in the assays shown in Fig. 1 and 2 are displayed to define those sequences which they commonly possess. The symbols and markings are the same as those shown in Fig. 1. Note that the numbering of the *Hpa*II fragments of pPH1-8 DNA (B) corresponds to the polyomavirus numbering scheme (A) and not to their position in the gel shown above the figure.

analysis, we used a recombinant plasmid called pPin7. pPin7 is a derivative of pPBR2 (25) that carries a *Hin*dIII linker at the *Bgl*I site (beginning at nucleotide 87) in polyomavirus DNA. A 3-bp stretch of viral DNA (nucleotides 91, 92, and 93) have been removed and replaced with 10 bp of linker DNA. Cleavage of pPin7 DNA simultaneously with *Hin*dIII, *Bam*HI, and *Eco*RI yields four fragments of which two (fragments 2 and 4; Fig. 3A) are composed only of viral sequences. Employment of these DNA fragments, which had been radiolabeled at their termini, in the binding assay revealed the presence of a large T antigen binding site on both fragments 2 and 4 (Fig. 3A). Therefore, there are at least two large T antigen binding sites in pPin7 DNA, one upstream and one downstream of the inserted *Hin*dIII linker. To insure that both of these sites were located entirely within the previously mapped T antigen binding region we utilized defined DNA fragments prepared from several recombinant plasmids that carried one or the other T antigen binding site alone in the binding assay.

pdPBHp is a recombinant plasmid that carries polyomavirus sequences between nucleotide 4632 (the *Bam*HI site) and nucleotide 152 (near an *Hph*I cleavage site). These viral sequences were cloned between the *Hin*dIII and *Bam*HI sites of the vector pML-2, a deletion derivative of pBR322 (20). Cleavage of pdPBHp DNA simultaneously with *Bam*HI and *Bgl*I yields four fragments whose locations are shown in Fig. 3B. When these fragments were end labeled and reacted with a crude preparation of large T antigen, two fragments were specifically immunoprecipitated (fragments 2 and 3; Fig. 3B). One of the fragments (number 2) carries only those viral sequences between nucleotides 94 and 152. These results confirm those shown in Fig. 3A and reveal that one of the large T antigen binding sites (between nucleotides 94 and 152) resides entirely within the region previously shown to harbor sites of binding of large T antigen (Fig. 2D).

To map the other large T antigen binding site we constructed another recombinant plasmid, pdPBBg, and cleaved it with the restriction endonuclease *Hpa*II. pdPBBg is composed of pML-2 DNA and the *Bam*HI (nucleotide 4632)-to-*Bgl*I (nucleotide 90) fragment of polyomavirus DNA (Fig. 3C). The viral sequences were cloned as a *Bam*HI-to-*Hin*dIII fragment within the large *Bam*HI-to-*Hin*dIII fragment of pML-2 DNA. Cleavage of pdPBBg DNA with *Hpa*II yields 19 fragments, and 2 of these (fragments 1 and 2) carry viral sequences. Employment of the DNA binding assay with these end-labeled fragments revealed that one of the fragments was specifically immunoprecipitated (fragment 2; Fig. 3C). The viral sequences borne by this fragment span the region from the *Hpa*II site at nucleotide 5292 to the *Bgl*I site at nucleotide 90. These results demonstrate that large T antigen can bind independently to at least two sites within the 153-bp region that includes nucleotides 5292 and 152. One of the large T antigen binding sites is located between nucleotides 5292 and 90, whereas the other maps between nucleotides 94 and 152. We refer to the latter as binding site I and to the former as binding site II throughout the remaining text.

While our research was under way, Gaudray et al. (7) reported the existence of a large T antigen binding site on polyomavirus DNA that overlaps with binding site II. However, they were unable to find evidence for any other binding sites on the viral genome. These investigators employed an indirect assay that measures the release by viral DNA fragments of ATPase activity from infected-cell extracts bound to calf thymus DNA-cellulose columns. The ATPase activity is inhibited by antibodies directed against the polyomavirus T antigens and is only released by specific viral DNA fragments (7). One explanation for the discrepancy between the data of Gaudray et al. and ours is that these investigators employed 300 mM KCl in the large T antigen (ATPase) release assay, whereas we used 20 mM NaCl. If the affinity of large T antigen for site I is reduced at high salt concentrations, then this would provide an explanation for the difference between their data and ours. To test this hypothesis we measured the affinity of large T antigen for the specific viral sequences at several different NaCl concentrations. First, the end-labeled fragments generated by *Hpa*II cleavage of pPH1-8 DNA (see Fig. 2B for a map) were used as substrate in the binding assay, and the reactions were carried out at 20, 50, 100, 150, and 300 mM NaCl. The polyomavirus *Hpa*II-5 fragment, which carries sites I and II, was specifically immunoprecipitated at all salt concentrations below 150 mM NaCl, but very little binding could be demonstrated at salt concentrations greater than 150 mM NaCl (Fig. 4). Second, the affinity of large T antigen for the separated binding sites was assayed at various NaCl concentrations. Employment of the end-labeled fragments generated by simultaneous cleavage of pPin7 DNA with *Bam*HI, *Hin*dIII, and *Eco*RI (see Fig. 3A for a map) in the binding assay at various salt concentrations revealed that the affinity of large T antigen for site I was less affected at high salt concentration than was its affinity for site II (Fig. 4). At 20 mM NaCl the fragments bearing site I (the second largest fragment) and site II (the smallest fragment) were specifically immunoprecipitated. However, at 100 mM NaCl the fragment bearing site II

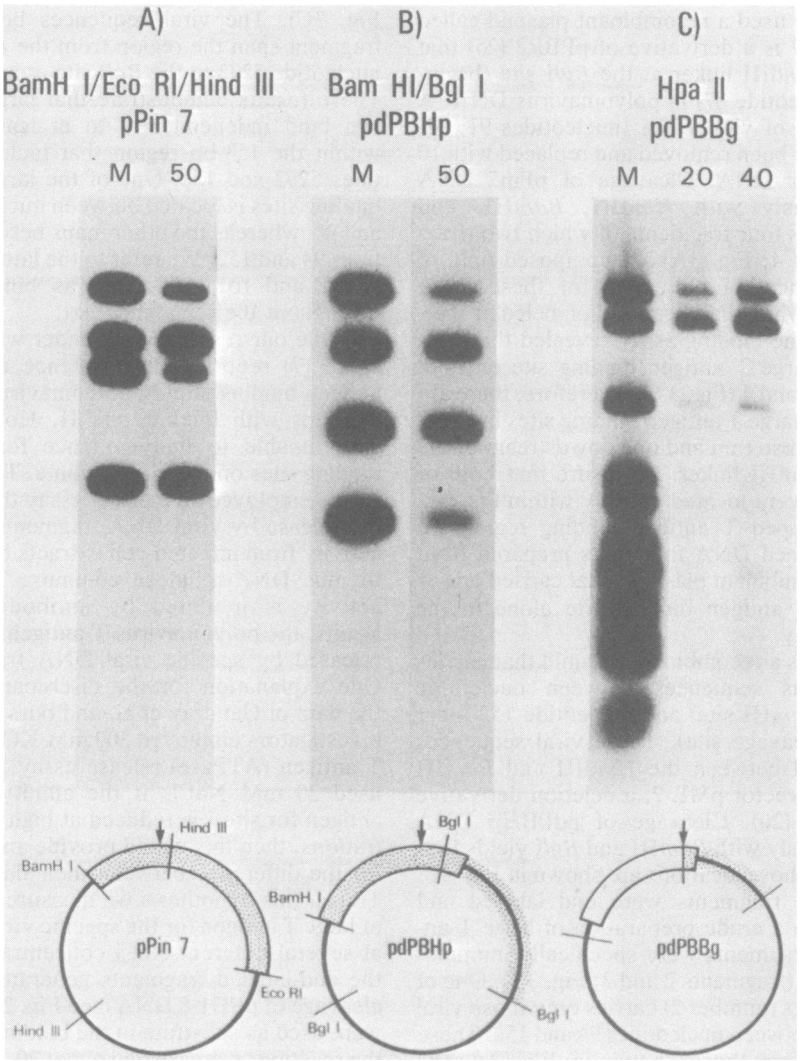

FIG. 3. Mapping of two large T antigen binding sites on the polyomavirus genome. End-labeled DNA fragments that contain only a portion of the previously identified large T antigen binding region were employed in the immunoassay. (A) End-labeled fragments obtained after the simultaneous digestion of pPin7 DNA with *Hin*dIII, *Eco*RI, and *Bam*HI were reacted at 10 ng/ml with 50 μl of nuclear extract and processed as described in the text. Similarly, 10 ng of pdPBHp DNA per ml, cleaved with *Bam*HI and *Bgl*I and end labeled, was reacted with 50 μl of nuclear extract (B). Note that the ends of DNA fragments generated by *Bgl*I cleavage (which produces 3′ protrusion) are not as efficiently labeled by the Klenow fragment of DNA polymerase I as are termini with 5′ protrusions (i.e., *Bam*HI, *Hin*dIII; lane M in B). This is more evident in shorter exposures of the autoradiogram, shown in (B), which show that the two smallest fragments that bear *Bam*HI-*Bgl*I ends are more efficiently labeled than the two largest that bear only *Bgl*I ends. (C) End-labeled fragments of pdPBBg DNA obtained by *Hpa*II digestions were reacted at 10 ng/ml with 20 and 40 μl of nuclear extract.

was not specifically immunoprecipitated, whereas the fragment carrying site I was (Fig. 4). These observations suggest that the affinity of large T antigen for site I is greater than is its affinity for site II, and they directly contradict the explanation provided earlier to account for the difference between the observations of Gaudray et al. and ours. An alternative hypothesis is that there are several different large T antigen species within infected cells that interact at different sites on the viral genome. The species with ATPase activity may bind only to site II and not to site I. Whatever the explanation, it is clear that the two large T antigen binding sites or the large T antigen species that interact with these sites are not equivalent.

DISCUSSION

We have identified a 153-bp segment of DNA within the polyomavirus genome to which large T antigen binds with high affinity. This region is located outside of structural gene sequences and can be subdivided into at least two large T antigen binding sites. Binding of large T antigen to each site occurs independently of binding to the other site, and under our standard conditions of assay, the affinity of large T antigen for each site appears to be nearly the same (Fig. 3 and 4). We have numbered these sequences binding sites I and II, with binding site I lying closest to the position where the 5′ ends of the most abundant early mRNAs map (Fig. 5).

Binding site I lies within a 58-bp sequence bracketed by nucleotides 94 and 152. Located within this area of the viral genome are the early Goldberg-Hogness TATA box (polyomavirus nucleotides 120 through 130) and a 5-bp stretch of DNA that displays homology to the consensus cap site, PyCATTCPu (polyoma virus nucleotides 149 through 155). Both of these elements are required for accurate and efficient initiation of transcription in vitro (18; Mueller et al., data not shown). However, neither of these sequences is absolutely necessary for early gene expression in vivo. Viable deletion mutants of polyomavirus have been isolated that remove all the sequences between nucleotides 94 and 152 (1), and mutant, recombinant viral DNAs have been constructed that also remove the aforementioned sequences with little consequence on early gene expression (Mueller et al., data not shown). That these sequences play a role at all in early gene transcription has been determined by the analyses of the phenotype of mutants of the viral genome cloned in plasmids. These studies revealed that deletion of the early TATA box and cap sites from DNA molecules that had no other upstream promoter elements resulted in a 100-fold reduction in the specific transforming activity of the DNA (Mueller et al., data not shown). Therefore, these sequences are components of the polyomavirus early promoter that function in conjunction with other upstream elements. The location of T antigen binding site I, relative to the early promoter elements and the position from which transcription initiates, strongly suggests an involvement of this site in autogenous gene regulation.

Binding site II is located within a 90-bp tract of DNA encompassed between nucleotides 5292 and 90 (Fig. 5). The initiation of viral DNA replication is known to occur near this region of the genome (3), and this binding site overlaps with sequences that are required in *cis* for viral DNA replication (21, 26, 38). However, part of binding site II is dispensable for lytic growth of the virus for constructed mutants (1, 21), and

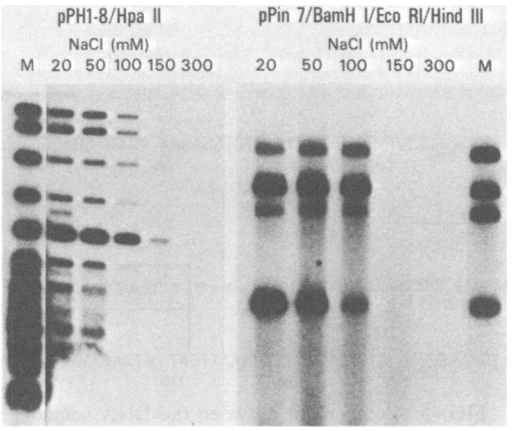

FIG. 4. Reactions of nuclear extracts from polyomavirus-infected 3T6 cells with DNA substrates at a variety of NaCl concentrations. End-labeled *Hpa*II DNA fragments of pPH1-8 DNA were reacted at 8 ng/ml with 50 μl of crude nuclear extract at the NaCl concentrations shown, and the samples were processed as described in the text (left-hand panel). The *Bam*HI, *Eco*RI, and *Hin*dIII fragments of pPin7 DNA were reacted at 10 ng/ml with 100 μl of crude nuclear extract at the NaCl concentrations shown, and the samples were processed as described in the text (right-hand panel).

variant strains of polyomavirus have been isolated that lack sequences contained within this site (5, 32). It is not known whether large T antigen binds to the DNA of the aforementioned mutants, but the protein does bind to the DNA of the A2 and A3 strains (7) and to our strain of polyomavirus. The A3 strain has 11 bp deleted between nucleotides 43 and 55 (nucleotides 44 through 54 are deleted) as compared to the A2 strain (5, 32; Fig. 5). Moreover, the strain of polyomavirus used in our studies has 10 bp deleted between nucleotides 44 and 55 as compared to the A2 strain (Fig. 5). Therefore, binding site II is composed of no more than 79 bp of DNA. It is very likely that at least part of large T antigen binding site II is required for viral DNA replication because mutant viral DNA molecules that lack binding site II fail to replicate when provided with large T antigen in *trans* (26).

Comparison of the DNA sequences in large T antigen binding sites I and II reveals that the sequence A/TGAGGC, spaced by 4 or 5 bp, is repeated three times within each site (7, 32; Fig. 5). These sequences occur on the opposite DNA strand and in the opposite orientation when the two binding sites are compared (Fig. 5). However, only two of the repeats in each site are required for large T antigen binding in vitro. Deletion of the central AGAGGC within binding site II, as occurs in the A3 and our strain of polyomavirus, does not impair the binding of

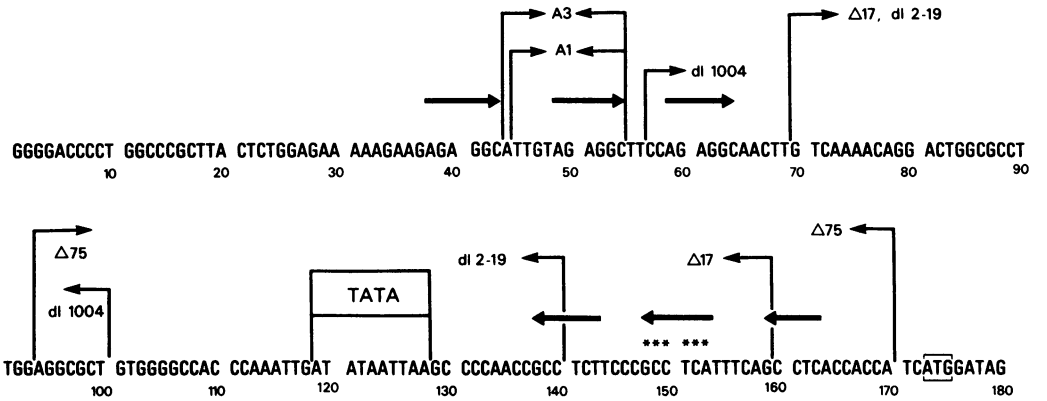

FIG. 5. Comparison between the DNA sequences in large T antigen binding sites I and II. Nucleotides 1 through 180, reading left to right, have the same polarity and sense as the early mRNAs of the A2 strain of polyomavirus. The sequences deleted in the various strains and viable mutants are indicated by the vertical lines bearing arrows. A1 refers to our strain of polyomavirus. The TATA consensus sequence and the putative initiation codon for translation of the early mRNAs are boxed. The asterisks denote the 5′ termini of the early mRNAs (7). Repeats of the A/TGAGGC sequence are marked by the horizontal arrows. Note that the repeats within binding site I occur on the opposite DNA strand from the one shown. Large T antigen binding site I has been mapped between nucleotides 94 and 152 (bottom); site II is located between nucleotides 5292 and 90 (top).

large T antigen to this site. Similarly, deletion of the terminal TGAGGC that straddles nucleotide 160 does not impair the binding of large T antigen to site I. Large T antigen binds efficiently to DNA fragments from pdPBHp DNA that contain viral sequences between nucleotides 90 and 152 (site I; Fig. 3B). In this context it is noteworthy that all of the viable deletion mutants isolated to date retain at least two copies of the hexanucleotide repeat, spaced by 5 bp, in binding site II (Fig. 5). Taken together, these observations suggest that a tandem repeat of the hexanucleotide sequence A/TGAGGC, separated by 4 or 5 bp, serves as a recognition site for large T antigen binding.

Although there is a great deal of sequence conservation between the two large T antigen binding sites, there are nonetheless important differences between them. First, the two sites appear to play distinct roles during the virus life cycle. Binding site I is likely the locus with which large T antigen interacts to repress early transcription, and site II probably serves as part of the replication origin of polyomavirus. Besides their functional distinctiveness, the two binding sites interact differently with large T antigen. The affinity of large T antigen for site I is greater than its affinity for site II at high salt concentrations. This was indicated by the observation that the binding of large T antigen to site II is reduced at high salt concentrations (100 mM NaCl), whereas binding to site I is hardly affected (Fig. 4). This could be due to the presence of three repeats of the AGAGGC sequence in the DNA fragment bearing site I and only two repeats of this sequence in the DNA fragment containing site II (Fig. 4 and 5). However, we have also observed that some crude nuclear preparations of large T antigen contain more activity for binding site II than for site I (data not shown). These apparently contradictory results can be explained if we postulate that there are at least two forms of large T antigen within infected cells, with different affinities for the two binding sites, and that the ratio of the two forms varies during the course of infection. A further embellishment of this hypothesis is that the large T antigen species that interacts with site II possesses an associated or intrinsic ATPase activity. This would explain the results of Gaudray et al. (7).

Comparison of the large T antigen binding sites in SV40 DNA and polyomavirus DNA relative to the genetic control regions of the viruses reveals a remarkable degree of similarity. Both viral DNAs contain multiple binding sites, and each of these can be assigned independent functions. Binding sites I of polyomavirus and SV40 both straddle sequences important for accurate initiation of early transcription and are likely the sites with which large T antigen interacts to repress early mRNA synthesis. Similarly, binding sites II of the viruses are probably involved in the initiation of viral DNA synthesis, for they form part of the origins of viral DNA replication. The similarity between the T antigen binding sites in SV40 and polyomavirus DNA is not limited to their position relative to viral control signals. Inspection of the sequences within the T antigen binding sites of SV40 and polyomavirus DNA reveals the occurrence of homologous sequences. The sequence GAGGC

is found repeated twice in SV40 binding site I, spaced by a run of 6 bp, and this repeat occurs again in binding site II, spaced by a single base pair. The pentamer is also found three times in binding site III, but it occurs on the opposite strand and the sequences are widely separated. The related sequence A/TGAGGC is repeated and spaced by 4 to 5 bp in both binding sites in polyomavirus DNA. It is possible that the hexamer repeat and the related pentamer repeat serve as recognition sequences for polyomavirus and SV40 large T antigen. Because these repeats are homologous, one would predict that the SV40 and polyomavirus T antigens could interact with each other's binding sites. This prediction is supported by the observation that SV40 large T antigen binds with specificity to sequences that encompass both polyomavirus large T antigen binding sites, and polyomavirus large T antigen specifically binds to SV40 DNA fragments containing the three SV40 large T antigen binding sites (B. J. Pomerantz and J. A. Hassell, data not shown). Whether the pentameric and hexameric repeats serve as T antigen recognition sites and whether they are sufficient for specific binding to occur remain to be determined. Since this research was completed Hayday et al. (16) obtained independent evidence that at least two large T antigen binding sites exist in polyomavirus DNA.

ACKNOWLEDGMENTS

We thank Monica Naujokas and Cathy Collins-Maynard for technical assistance. We also thank William Muller for pdPBBg and pPin7 DNA, Nicholas Acheson for helpful discussions and a sample of polyomavirus, Steve Dilworth for samples of large T antigen monoclonal antibodies, and Ron McKay for introducing us to the DNA binding-immunoprecipitation technique.

This research was supported by the Medical Research Council and the National Cancer Institute of Canada. C.R.M. is supported by a research studentship from the Natural Sciences and Engineering Research Council of Canada, and J.A.H. is a research scholar of the National Cancer Institute.

LITERATURE CITED

1. **Bendig, M. M., T. Thomas, and W. R. Folk.** 1980. Regulatory mutants of polyoma virus defective in DNA replication and the synthesis of early proteins. Cell **20**:401–409.
2. **Cogen, B.** 1978. Virus-specific early RNA in 3T6 cells infected by a tsA mutant of polyoma virus. Virology **85**:222–230.
3. **Crawford, L. V., A. K. Robbins, and P. M. Nicklin.** 1974. Location of the origin and terminus of replication in polyoma virus DNA. J. Gen. Virol. **25**:133–138.
4. **Francke, B., and W. Eckhart.** 1973. Polyoma gene function required for viral DNA synthesis. Virology **55**:127–135.
5. **Friedmann, T., A. Esty, P. LaPorte, and P. Deininger.** 1979. The nucleotide sequence and genome organization of the polyoma early region: extensive nucleotide and amino acid homology with SV40. Cell **17**:715–724.
6. **Gaudray, P., P. Clertant, and F. Cuzin.** 1980. ATP phosphohydrolase (ATP ase) activity of a polyoma virus T-antigen. Eur. J. Biochem. **109**:553–560.
7. **Gaudray, P., C. Tyndall, R. Kamen, and F. Cuzin.** 1981. The high affinity binding site on polyoma virus DNA for the viral large T protein. Nucleic Acids Res. **9**:5697–5710.
8. **Gluzman, Y., R. J. Frisque, and J. Sambrook.** 1979. Origin-defective mutants of SV40. Cold Spring Harbor Symp. Quant. Biol. **44**:293–300.
9. **Gluzman, Y., J. F. Sambrook, and R. J. Frisque.** 1980. Expression of early genes of origin-defective mutants of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **77**:3898–3902.
10. **Goldberger, R. F.** 1974. Autogenous regulation of gene expression. Science **183**:810–816.
11. **Griffin, B. E.** 1982. Structure and genomic organization of SV40 and polyoma virus, p. 61–123. *In* J. Tooze (ed.), DNA tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. **Hassell, J., E. Lukanidin, G. Fey, and J. Sambrook.** 1978. The structure and expression of two defective adenovirus 2/simian virus 40 hybrids. J. Mol. Biol. **120**:209–247.
13. **Hassell, J. A., C. Mueller, A.-M. Mes, M. Featherstone, M. Naujokas, B. Pomerantz, and W. Muller.** 1982. The construction of polyoma virus vectors: functions required for gene expression, p. 71–77. *In* Y. Gluzman (ed.), Eukaryotic viral vectors. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **Hassell, J. A., W. C. Topp, D. B. Rifkin, and P. E. Moreau.** 1980. Transformation of rat embryo fibroblasts by cloned polyoma virus DNA fragments containing only part of the early region. Proc. Natl. Acad. Sci. U.S.A. **77**:3978–3982.
15. **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.
16. **Hayday, A. C., F. Chaudry, and M. Fried.** 1983. Loss of polyoma virus infectivity as a result of a single amino acid change in a region of polyoma virus large T-antigen which has extensive amino acid homology with simian virus 40 large T-antigen. J. Virol. **45**:693–699.
17. **Ito, Y., J. R. Brocklehurst, and R. Dulbecco.** 1977. Virus-specific proteins in the plasma membrane of cells lytically infected or transformed by polyoma virus. Proc. Natl. Acad. Sci. U.S.A. **74**:4666–4670.
18. **Jat, P., J. W. Roberts, A. Cowie, and R. Kamen.** 1982. Comparison of the polyoma virus early and late promoters by transcription *in vitro*. Nucleic Acids Res. **10**:871–887.
19. **Kahn, M., R. Kolter, C. Thomas, D. Figurski, R. Meyer, E. Remaut, and D. R. Helinski.** 1979. Plasmid cloning vehicles derived from plasmids Col E1, F, R6K, and RK2. Methods Enzymol. **68**:268–280.
20. **Lusky, M., and M. Botchan.** 1981. Inhibition of SV40 replication in simian cells by specific pBR322 DNA sequences. Nature (London) **293**:79–81.
21. **Luthman, H., M. G. Nilsson, and G. Magnusson.** 1982. Noncontiguous segments of the polyoma genome required in *cis* for DNA replication. J. Mol. Biol. **161**:533–550.
22. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labelled DNA with base specific chemical cleavages. Methods Enzymol. **65**:499–560.
23. **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.
24. **McKay, R. D. G.** 1981. Binding of a simian virus 40 T-antigen related protein to DNA. J. Mol. Biol. **145**:471–488.
25. **Mes, A.-M., and J. A. Hassell.** 1982. Polyoma viral middle T-antigen is required for transformation. J. Virol. **42**:621–629.
26. **Muller, W. J., C. R. Mueller, A.-M. Mes, and J. A. Hassell.** 1983. Polyomavirus origin for DNA replication comprises multiple genetic elements. J. Virol. **47**:586–599.
27. **Myers, R. M., D. Rio, A. Robbins, and R. Tjian.** 1981. SV40 gene expression is modulated by the cooperative binding of T-antigen to DNA. Cell **25**:373–384.
28. **Myers, R. M., and R. Tjian.** 1980. Construction and analysis of simian virus 40 origins defective in tumor antigen binding and DNA replication. Proc. Natl. Acad. Sci. U.S.A. **77**:6491–6495.

29. **Shalloway, D., T. Kleinberger, and D. M. Livingston.** 1980. Mapping of SV40 DNA replication origin region binding sites for the SV40 T-antigen by protection against exonuclease III digestion. Cell **20**:411–422.
30. **Shortle, D., R. F. Margolskee, and D. Nathans.** 1979. Mutational analysis of the SV40 replicon: pseudorevertants of mutants with a defective replication origin. Proc. Natl. Acad. Sci. U.S.A. **76**:6128–6131.
31. **Shortle, D., and D. Nathans.** 1979. Regulatory mutants of SV40: constructed mutants with base substitutions at the origin of DNA replication. J. Mol. Biol. **131**:801–817.
32. **Soeda, E., J. R. Arrand, N. Smolar, and B. E. Griffin.** 1979. Sequences from the early region of polyoma virus DNA containing the viral replication origin and encoding small, middle, and (part of) large T-antigen. Cell **17**:357–370.
33. **Soeda, E., J. R. Arrand, N. Smolar, J. E. Walsh, and B. E. Griffin.** 1980. Coding potential and regulatory signals of the polyoma virus genome. Nature (London) **283**:445–453.
34. **Takemoto, K. K., R. A. Malmgren, and K. Habel.** 1966. Immunofluorescent demonstration of polyoma tumor antigen in lytic infection of mouse embryo cells. Virology **28**:485–488.
35. **Tegtmeyer, P., B. Andersen, S. B. Shaw, and V. G. Wilson.** 1981. Alternative interactions of the SV40 A protein with DNA. Virology **115**:75–87.
36. **Tjian, R.** 1978. The binding site on SV40 DNA for a T-antigen related protein. Cell **13**:165–179.
37. **Tjian, R.** 1981. T-antigen binding and the control of SV40 gene expression. Cell **26**:1–2.
38. **Turler, H.** 1980. The tumor antigens and the early functions of polyoma virus. Mol. Cell. Biochem. **32**:63–93.