JOURNAL OF VIROLOGY, May 1985, p. 532–545
0022-538X/85/050532-14$02.00/0

Vol. 54, No. 2

Simian Virus 40 and Polyomavirus Large Tumor Antigens Have Different Requirements for High-Affinity Sequence-Specific DNA Binding

ARNO SCHELLER AND CAROL PRIVES*

Department of Biological Sciences, Columbia University, New York, New York 10027

Received 26 September 1984/Accepted 16 January 1985

By using a DNA fragment immunoassay, the binding of simian virus 40 (SV40) and polyomavirus (Py) large tumor (T) antigens to regulatory regions at both viral origins of replication was examined. Although both Py T antigen and SV40 T antigen bind to multiple discrete regions on their proper origins and the reciprocal origin, several striking differences were observed. Py T antigen bound efficiently to three regions on Py DNA centered around an *Mbo*II site at nucleotide 45 (region A), a *Bgl*I site at nucleotide 92 (region B), and another *Mbo*II site at nucleotide 132 (region C). Region A is adjacent to the viral replication origin, and region C coincides with the major early mRNA cap site. Weak binding by Py T antigen to the origin palindrome centered at nucleotide 3 also was observed. SV40 T antigen binds strongly to Py regions A and B but only weakly to region C. This weak binding on region C was surprising because this region contains four tandem repeats of GPuGGC, the canonical pentanucleotide sequence thought to be involved in specific binding by T antigens. On SV40 DNA, SV40 T antigen displayed its characteristic hierarchy of affinities, binding most efficiently to site 1 and less efficiently to site 2. Binding to site 3 was undetectable under these conditions. In contrast, Py T antigen, despite an overall relative reduction of affinity for SV40 DNA, binds equally to fragments containing each of the three SV40 binding sites. Py T antigen, but not SV40 T antigen, also bound specifically to a region of human *Alu* DNA which bears a remarkable homology to SV40 site 1. However, both tumor antigens fail to precipitate DNA from the same region which has two direct repeats of GAGGC. These results indicate that despite similarities in protein structure and DNA sequence, requirements of the two T antigens for pentanucleotide configuration and neighboring sequence environment are different.

Two well-characterized papovaviruses, simian virus 40 (SV40) and murine polyomavirus (Py), are sufficiently closely related that they share extensive structural as well as functional homology (6, 27, 31). In particular, the large tumor (T) antigens of these viruses have numerous similarities, although one of the notable differences between these two proteins is that the T antigen of SV40 is capable of most of the tumorigenic activities of that virus, whereas the T antigen of Py has few (24, 33). Despite this difference, the functions of the two proteins are virtually identical in direct control of viral replication and induction of the cellular processes required for viral replication. The structures of the T antigens from Py and SV40 are sufficiently similar that it is possible to match amino acid homologies with areas in the SV40 early gene products known to be involved with its functions in virus replication and stimulation of cellular DNA synthesis (6, 28, 30). Furthermore, the structures of the two T antigens are similar enough to permit on occasion cross-reactivity between high-titer anti-SV40 T-antigen antisera and the T antigen of Py (19).

In addition to the similarity of the two T antigens, the control regions of the two viruses have certain homology, especially in the regions bound by the two T antigens, as determined by DNase I and dimethyl sulfate protection by SV40 T antigen (2, 12, 32) and, recently, by protection from DNase I nicking by Py T antigen (3). The regions of greatest sequence homology include a palindrome at the site of initiation of viral DNA replication, with an adenine-thymine-rich region immediately adjacent to the palindrome. In addition, both regulatory regions contain tandem T-antigen binding sites adjacent to and including the origin palindrome. These sites contain the pentanucleotides GAGGC and GGGGC, which have been identified as the contact points for SV40 T-antigen binding (2, 12, 32). Although the distribution of these pentanucleotides is different in each of the two viruses, it has been reported that SV40 and Py T antigens bind to the regulatory region of the other virus (21). However, the locations of the binding regions on the other origin of replication for either T antigen were not identified, other than that the pentanucleotide GAGGC was involved.

To gain possible insight into both the similarities and differences of the two proteins in their roles in viral replication and interactions with DNA, we have examined and compared the specific binding of SV40 and Py T antigens to both viral DNAs. In addition, as there have been as yet few reports on the binding of these T antigens to specific nonviral DNA sequences, we also have examined the interaction of both proteins with DNA from the *Alu* family of repetitive DNA (11, 23) that bears homology to SV40 regulatory sequences. From the interactions of the two T antigens with these DNAs, we have drawn inferences concerning the distribution and the context of the pentanucleotides required to support high-affinity binding to DNA.

MATERIALS AND METHODS

Cells and viruses. Cos 7 cells are SV40-transformed monkey CV-1 cells (8). 3T6 cells were infected with Py wild-type strain LL 3. CV-1 cells were infected with SV40 wild-type strain 776 (33).

Preparation of cell extracts for DNA binding. Extracts from 10^7 cells were made as follows. A 150-mm plate of cells was

* Corresponding author.

washed four times with phosphate-buffered saline, drained well on ice, swelled with 2 ml of cold nuclear retention buffer {10 mM MES [2-(*N*-morpholino)ethanesulfonic acid], 4 mM $MgCl_2$, pH 6.0} for 10 min, and then scraped into a Dounce homogenizer. The cells were disrupted with 10 strokes of an A pestle and pelleted for 8 min at 8,000 rpm in an SS34 rotor. The pellet was suspended in 300 μl of high-salt HIP buffer (0.45 M NaCl, 0.02 M HEPES [*N*-2-hydroxyethylpiperazine-*N'*-2- ethanesulfonic acid], 1 mM EDTA, 1% Nonidet P-40, pH 8.5), incubated for 20 min, and then spun for 30 min in an SS34 rotor at 20,000 rpm. The supernatant could be stored frozen at −70°C, although with some loss (up to 50%) of specific DNA binding activity.

DNA fragment immunoassay. For binding, the supernatant was diluted with two volumes of HM dilution buffer (10 mM MES, 10 mM HEPES, 0.1% Nonidet P-40), and the pH was adjusted to the desired acidity (4°C), usually 6.50 in the case of Py extracts and 6.80 for SV40 extracts. As a rule, before adjusting the pH, the extract was preabsorbed with NHS (Capell Laboratories) and inactivated *Staphylococcus aureus* bacteria to reduce the background of nonspecifically bound labeled DNA restriction fragments. After pH adjustment, 5 ng of ^{32}P-labeled probe DNA, together with 1 μg of herring sperm DNA used as nonspecific carrier, was added into the extract. Generally, 1×10^6 cells of SV40-containing extract or 5×10^6 cells of Py-containing extract were used for each 5 ng of probe DNA.

After incubation on ice for 1 h, 10 μl of hamster antitumor antiserum was added for another 1 h, followed by 50 μl of 10% suspension of fixed *S. aureus* for 25 min. The protein DNA-*S. aureus* complex was isolated by centrifugation, and the pellet was washed three times by suspension in wash buffer. The DNA fragments were released by heating at 60°C for 30 min, and the supernatant was deproteinized and ethanol precipitated. The pellet was suspended in DNA sample buffer (buffer for running gels made 5% with Ficoll), and the DNA fragments were resolved by electrophoresis on either a 2% agarose or a 5% polyacrylamide gel, with 90 mM Tris-borate–2.5 mM EDTA. The gel was fixed in protein fixer (10% acetic acid, 10% trichloroacetic acid, 30% methanol), dried, and fluorographed with X-ray film with an intensifying screen at −70°C.

Modification of the binding assay. A number of the conditions used for SV40 T antigen, such as binding in 0.15 M NaCl and in the presence of carrier DNA (25), were usable for Py T antigen. However, unlike SV40 T antigen, Py T antigen exhibited a sharp pH optimum at 6.50, above which the signal decreased. Below pH 6.25 the ratio of origin-specific to nonspecific binding decreased. In trials with increased concentrations of Py T antigen, the pH optimum for specific origin binding broadened to include higher pH levels. Also, it was found that the Py origin binding activity in these extracts varied between batches but was routinely 5- to 10-fold less than that contained in SV40 T antigen extracts.

Since the efficiency of viral infection caused variation between batches of extract made from lytic infections and since the presence of viral DNA in such extracts might be a potent competitor for input DNA, a Py-transformed hamster cell line was used, as well as 3T6 cells lytically infected with Py. Comparison of the ratios of the separated binding regions derived after binding with hamster or mouse extract revealed no differences, and the two extracts were used interchangeably. Most variations in the level of specificity or background were due to the relative amount of specific binding activity in the extract rather than to the source of the cells used to make it. As a rule, for binding to ectopic DNAs, at least three times more extract was required for specific binding to be distinguishable over background.

Materials. Anti-Py antisera and anti-SV40 antisera were prepared from tumor-bearing hamsters. Anti-SV40 T-antigen monoclonal antibody was obtained from the tissue culture supernatant of pAB 416 mouse hybridoma cells (9). Formaldehyde-inactivated *S. aureus* bacteria were purchased from the Enzyme Center. Restriction endonucleases and DNA polymerase were purchased from New England BioLabs, Inc. DNAs used were p373A2, which is Py strain A2 inserted into the *Bam*HI site of pAT153 (24), and pSVR1, which is SV40 strain 776 inserted into the *Eco*RI site of pBR322. These plasmids were the generous gifts of R. Kamen and J. Manley, respectively. Also obtained from J. Manley was a *Bgl*II fragment, subcloned into the *Bam*HI site of pBR322, from the human β-globin locus (23) containing a representative of the *Alu* family. Nick translation to label DNA with [α-^{32}P]deoxynucleoside 5'-triphosphate was performed by the methods of Maniatis et al. (17) as described previously (25). Labeled DNAs were digested typically as 50 ng in 5 μl of the buffer recommended for each restriction endonuclease.

RESULTS

Py T antigen binds to discrete regions on Py DNA. Extracts of Py T-antigen-containing cells were used in the DNA fragment immunoassay to examine the specific binding of Py T antigen to Py DNA. Py T antigen is the viral gene product responsible for the high-affinity, sequence-specific DNA binding described below. This was first demonstrated in experiments with a Py T-antigen-specific monoclonal antibody (3), isolation of T-antigen-specific mutants of Py (10), and extracts from a rat cell line (18–37) expressing, of the three Py tumor antigens, only T polypeptide (C. Prives, K. Cary, B. Barnet, A. Scheller, K. Chowdhury, J. Bolen, and M. Israel, manuscript in preparation). When various restriction digests were used, the approximate limits and strengths of binding regions were estimated. These were found to extend the results of recently reported studies with the DNA fragment immunoassay (3, 7, 10, 22). In addition, the present study confirms the data from DNase I footprint techniques (3), which indicated that three specific binding regions are present on Py DNA for Py T antigen.

A single 216-base-pair (bp) fragment (nucleotide [nt] 5265 to nt 187) was precipitated from a *Pvu*II-*Dde*I digest. (Fig. 1e). Further cleavage with *Bgl*I permitted fragments on either side of the *Bgl*I site at nt 90 to nt 92 to be precipitated (Fig. 1d). The use of *Mbo*II and *Hae*III restriction enzymes in conjunction with others led to the identification of three strong sites (regions A, B, and C) adjacent to the replication origin on the early side. The results of these experiments are presented.

(i) The middle site (B). After cleavage within the origin region with *Mbo*II, three fragments are precipitated: #12 (293 bp between nt 5045 and nt 45), #24 (86 bp between nt 46 and nt 132), and #19, (133 bp between nt 132 and nt 267) (Fig. 1a). Fragment #12 contains the origin palindrome, and fragment #19 includes the major early cap site (13) at nt 150. The intensity of #19 is weak, but data presented below indicate that cleavage at the *Mbo*II restriction site at nt 132 damages the binding region. Besides the three fragments identified above, the additional immune-precipitated fragments visible in the digest are the result of incomplete cleavage with *Mbo*II. These are the strong 222-bp fragment,

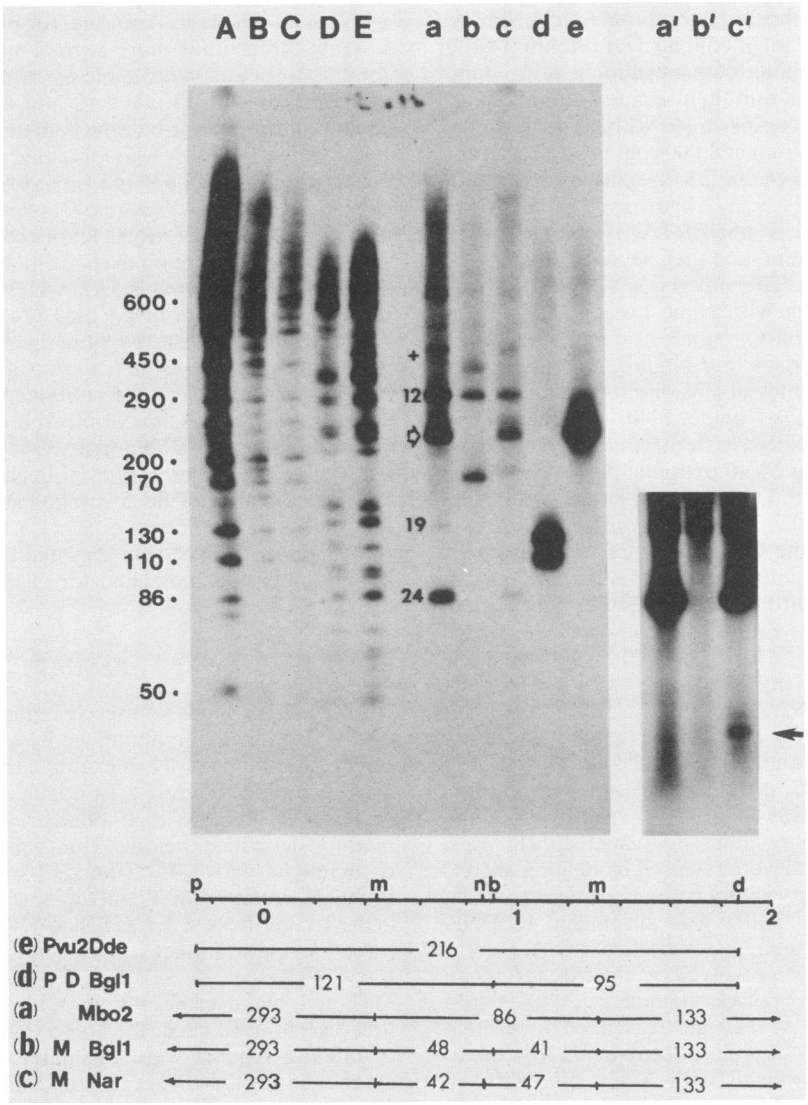

FIG. 1. Binding of Py T antigen to separated regions on Py DNA. Reactions were performed as described in the text. Extracts from 2 × 10^6 Py-transformed cells were used to bind 5 ng of labeled p373A2 DNA digested with *Mbo*II (a), *Mbo*II plus *Bgl*I (b), *Mbo*II plus *Nar*I (c), *Pvu*II plus *Dde*I plus *Bgl*I (d), or *Pvu*II plus *Dde*I (e). Ten-times longer exposures of lanes a to c are shown in lanes a', b', and c', respectively. Above the point of truncation, lanes a', b', and c' were very overexposed. Capital letters indicate the input digests. Numbers along the left side of the gel are sizes in bp. The expected fragments generated by *Mbo*II are identified on the autoradiogram as 12, 19, and 24. *Mbo*II, which is sensitive to methylation on adenosine and cytosine residues, could not be digested to completion. Bands representing partial digests by *Mbo*II appear, which include the 219-bp #24-19 (open arrow), the 380-bp #12-24 (+), and the 512-bp #12-24-19. *Nar*I, which is affected by nucleotides surrounding its recognition sequence, also was difficult to cut to completion. In lane c, less than half of the potential *Nar*I sites could be cleaved, as evinced from comparing the intensity of the band at 220 bp with its *Nar*I product at 170 bp. In lane c' a solid arrow points to the 47-bp *Nar*I-*Mbo*II fragment. Also, note that for lanes a and a', 10 ng of *Mbo*II was added instead of 5 ng. A map of the relevant cleavage sites within the *Pvu*II-*Dde*I fragment numbered every 100 bp appears at the bottom of the figure. Beneath this map are depicted each of the digests, with the sizes of the fragments expected in the binding region indicated. The order of these digests does not reflect the order of the lanes. At the left side of each digest, within parentheses, is the letter of the lane to which it refers. Symbols for the various restriction enzymes: a, *Alu*I; b, *Bgl*I; d, *Dde*I; m, *Mbo*II; n, *Nar*I; p, *Pvu*II.

which includes #24 and #19, and the weaker 380-bp fragment composed of #12 and #24. Additional cleavages in this region with other restriction endonucleases showed this to be the case.

Cleavage of *Mbo*II #24 with *Bgl*I prevented either of the resulting fragments from precipitating detectably (Fig. 1b). However, after digestion with *Nar*I (again a partial digest; see the legend to Fig. 1), which cuts at nt 84 to nt 86, a very small amount of a 47-bp fragment could be precipitated, visible with long exposure of the gel (Fig. 1c and c'). Since under even these conditions neither of the #24-*Bgl*I fragments is visible (Fig. 1b'), it was concluded that *Bgl*I effectively destroys this binding region. However, since the intensity of the precipitated #24-*Nar*I fragment is very much less than half that of #24, this indicates that *Nar*I still damages the binding region, probably cutting on the edge of it, but leaves enough of the region intact to permit some weak binding. It cannot be ignored that other factors can affect the intensity of precipitation of small DNA fragments. Local heterogeneities in the labeling of fragments by nick

translation can become pronounced in small fragments. Also, a minimum length may be required for T antigen to bind to a DNA fragment containing a specific binding region. Limits of the other side of the binding region were examined through digestion of *Mbo*II #24 with *Hae*III. A fragment of 62 bp was precipitated, which had an intensity little reduced proportionately from that of the 86-bp fragment (data not shown). This indicated that the sequences required for binding in this region lie slightly beyond the *Nar*I site at nt 84 but within the *Hae*III site at nt 106.

Although cleavage of *Mbo*II #24 with *Bgl*I prevented precipitation of the resulting halves (since fragments from either side of the *Bgl*I site at nt 92 extending past either *Mbo*II site at nt 45 or nt 132, as in the *Pvu*II-*Dde*I-*Bgl*I digest, were strongly precipitated), two strong binding regions ought to lie beyond the edges of #24. However, since neither #12 nor #19 was efficiently precipitated relative to the #24-#19 partial band, even after cleavage with *Bgl*I, each binding region probably was interrupted by *Mbo*II. Although it could be argued that the weakness of #19 would be attributable to very inefficient digestion by *Mbo*II at nt 132, two arguments can be brought against this. First, the ratio of the intensity of #19 to that of #24 remains low, which would be independent of the extent of digestion. Second, the proportions of partial products varied between batches of nick-translated DNA, and in a few cases the partial band at 220 bp was virtually undetectable.

(ii) The site adjacent to the replication origin (A). A strong binding region adjacent to the palindrome could be detected by treating the origin region with *Hae*III, which permitted two fragments to be precipitated: one, 94 bp long, between nt 12 and nt 106 (Fig. 2d), as well as one extending from nt 106 and beyond. With an *Hae*III-*Bgl*I digest, the 79-bp fragment (nt 12 to nt 91) also was precipitated efficiently (Fig. 2f). However, the intensity of the fragment was less than half that of the 94-bp fragment, as if the contribution of the binding region at the *Bgl*I site were substantial. Despite the certainty that the *Mbo*II site at nt 45 effectively damages the strong binding region, the endpoints of region A cannot be placed more definitely by this method than as being between the *Hae*III site at nt 12 and the *Bgl*I site at nt 92. However, the endpoints probably lie well within this region.

(iii) The site containing the major early mRNA caps (C). Evidence from this report and others (3, 13, 22) suggested that another strong binding region existed on the early side of the *Bgl*I site, and the *Mbo*II digest indicated that nt 132 was near the left boundary of this region. To keep sequences left of this site intact, a *Hae*III digest was used. Fragments rightward from nt 107, either the *Hae*III-*Dde*I 80-bp (nt 107 to nt 187) (Fig. 2d) or the *Hae*III-*Hph*I 48-bp (nt 107 to nt 155) (Fig. 2e) fragment, were precipitated. The ratio of the intensities between the 48-bp fragment and the 80-bp *Hae*III-*Dde*I fragment was less than that expected from a mere reduction in size. However, the ratio of intensities between the *Hae*III-*Hph*I 48-bp fragment and the *Hae*III-*Bgl*I 79-bp fragment (see above) is proportional to their sizes. This suggests that the binding region involves sequences extending past the *Hph*I site at nt 155. But, because of the weakness of *Mbo*II #19, another complete binding region cannot be involved, so *Hph*I may trim off DNA sequences not necessarily involved in specificity but still required for stable interaction with the protein. Therefore, the endpoints for this binding region would appear to extend from somewhat leftward of the *Mbo*II site at nt 132 up to around the *Hph*I site at nt 155.

To summarize the data on binding of Py T antigen to Py DNA by the DNA fragment immunoassay, available restriction sites led to the identification of three strong binding regions, A, B, and C. These happen to be interrupted by the *Mbo*II site at nt 45, the *Bgl*I site at nt 91, and the other *Mbo*II site at nt 132, respectively. In addition, a fourth region containing the origin palindrome also was bound (*Mbo*II-*Pvu*II 72-bp fragment; Fig. 2c), but considerably less efficiently than the others (see below). The regions are depicted as the upper band of each set of restriction digests (see Fig. 5).

Differences in the binding of SV40 T antigen to multiple sites on Py DNA. The ability of the T antigen encoded by the related papovavirus SV40 to bind to Py DNA was examined. Similar restriction digests were used to compare the relative efficiency of binding to the different regions. As was the case with Py T antigen, the *Pvu*II-*Dde*I 216-bp fragment was the major fragment precipitated by SV40 T antigen (Fig. 3a). However, a striking difference between the response of the two T antigens was revealed after digestion of this fragment with *Bgl*I. The 125-bp fragment containing the origin and region A was precipitated efficiently, but the 95-bp fragment containing region C was bound only weakly (Fig. 3b). Trials with the *Mbo*II digest revealed that fragments #12 and #24 were precipitable but #19 was not (Fig. 3c and e). For contrast, compare the Py T-antigen precipitates in Fig. 2. The partial digest of #24-#19 was precipitable, but after digestion with *Bgl*I the resulting 170-bp fragment no longer was strongly precipitable (Fig. 3d). The failure of SV40 T antigen to bind efficiently to region C was confirmed by the *Hae*III digests. Compared with the 95-bp *Hae*III fragment containing regions A and B, the 80-bp *Hae*III-*Dde*I fragment containing only region C was rather weak (Fig. 3f). Consequently, the *Hae*III-*Hph*I 48-bp fragment (Fig. 3g and h) was detectable only after much longer exposures than those shown. Thus, although SV40 T antigen does bind efficiently to regions A and B, demonstrating the same sensitivity to *Mbo*II and *Bgl*I at these sites as does Py T antigen, it binds only very weakly to region C. As this latter region contains four tandem repeats of GPuGGC (three of them GAGGC), which ought to ensure efficient binding to this region, this is a somewhat surprising observation. A summary of the binding of SV40 T antigen to discrete regions on Py DNA is depicted as the lower band of each restriction digest (see Fig. 5).

Evidence that Py T antigen and SV40 T antigen bind weakly to sequences at the Py origin of replication. In addition to identifying three discrete strong binding regions adjacent to the viral replication origin, several experiments indicated that Py T antigen binds weakly but specifically to sequences containing the origin palindrome. First, a *Pvu*II-*Mbo*II digest yielded a 72-bp fragment (nt 5265 to nt 45) which was of an intensity considerably less than that of *Mbo*II #24 (Fig. 4a). Second, an alternative digestion of *Mbo*II #12 with *Hpa*II, which cuts at nt 1, resulted in the weak binding of a 47-bp fragment mapping between nt 1 and nt 45 (Fig. 4b). The intensity of the 47-bp *Hpa*II-*Mbo*II fragment was greatly reduced compared with that of the 72-bp *Pvu*II-*Mbo*II fragment. Although the possibility remained that sequences outside of the origin palindrome adjacent to nt 45 played a role in precipitating *Mbo*II #12, a third digest with *Bst*NI-*Hph*I indicated that this was not the case. Digestion of the origin with *Bst*NI permitted the 79-bp fragment (nt 5227 to nt 10) to be precipitated very weakly (Fig. 4c), thereby indicating that sequences within the origin palindrome contribute to the precipitation of *Mbo*II #12. From the reduction in efficiency of binding relative to the 72-bp *Pvu*II-*Mbo*II

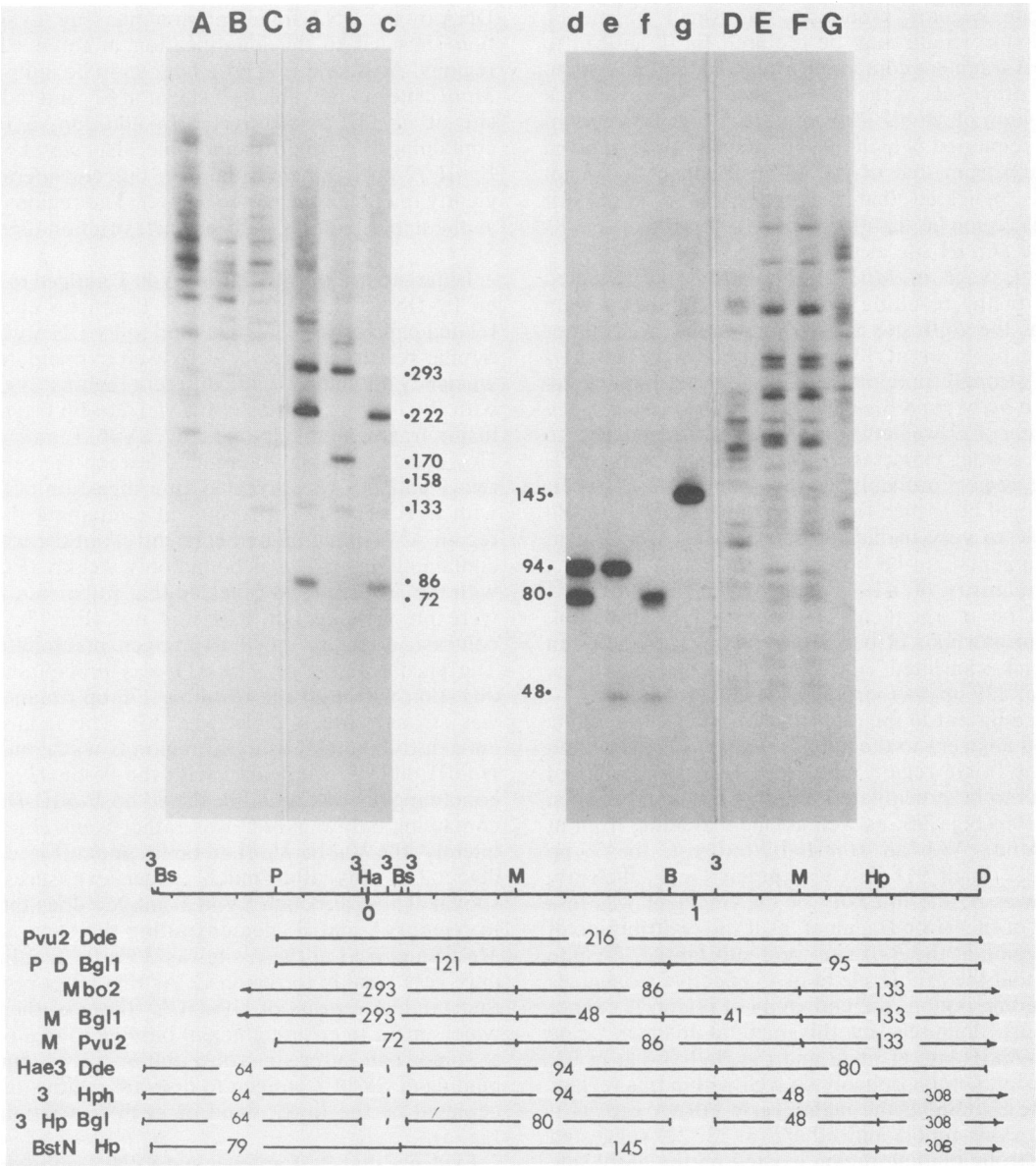

FIG. 2. Binding of Py T antigen to separated regions on Py DNA. Reactions were performed as described in the legend to Fig. 1. The enzymes used were *Mbo*II (a), *Mbo*II plus *Bgl*I (b), *Mbo*II plus *Pvu*II (c), *Hae*III plus *Dde*I (d), *Hae*III plus *Hph*I (e), *Hae*III plus *Bgl*I plus *Hph*I (f), or *Bst*NI plus *Hph*I (g). Capital letters indicate the input digests. Numbers refer to the size in bp. At the bottom of the figure is a map of the relevant cleavage sites in *Pvu*II-*Dde*I, which is numbered every 20 bp according to Soeda et al. (27). Maps of each of the restriction digests, with sizes of the fragments expected spanning the binding regions indicated, appear below. Symbols for the various restriction enzymes: B, *Bgl*I; Bs, *Bst*NI; D, *Dde*I; Ha, *Hpa*II; Hp, *Hph*I; M, *Mbo*II; P, *Pvu*II; 3, *Hae*III. In lane c, approximately half of the sample was lost. The 72-bp band from the *Mbo*II-plus-*Pvu*II digest (lane c) is barely visible in this figure but can be more clearly distinguished in a four-times longer exposure of the same lane appearing in Fig. 4a below. Also, note that a 158-bp band appears, which should extend from nt 5265 to nt 133, and that the bands larger than 300 bp have been cleaved by *Pvu*II.

fragment, it is evident that *Hpa*II and *Bst*NI (as well as *Hae*III) impinge upon the origin binding region.

Binding to the Py origin palindrome by SV40 T antigen differed. Although precipitation of the 72-bp *Pvu*II-*Mbo*II fragment was close to the efficiency observed for Py T antigen (Fig. 4a, cf. p and s), cleaving within the origin palindrome with either *Hpa*II or *Bst*NI prevented SV40 T antigen from precipitating the origin palindrome containing DNA. Unlike Py T antigen, neither the 47-bp *Hpa*II-*Mbo*II fragment (Fig. 4b) nor the 79-bp *Bst*NI fragment (Fig. 4c) was precipitated by SV40 T antigen. Thus, the distribution of SV40 T antigen on the origin palindrome differs from that of Py T antigen. Furthermore, there is less of a possibility that the remainder of region A left in *Mbo*II #12 contributes to the precipitation of that fragment. The difference in binding to the Py origin palindrome is included in the summary presented in Fig. 5.

Py T antigen binds to discrete regions on SV40 DNA. Considering the similarities and differences of binding of SV40 T antigen on the Py origin, it was conceivable that Py T antigen might differ from SV40 T antigen in its distribution on and relative affinity for the SV40 origin. The binding of SV40 T antigen to its own origin has been characterized extensively and serves as a context for the efficiency of

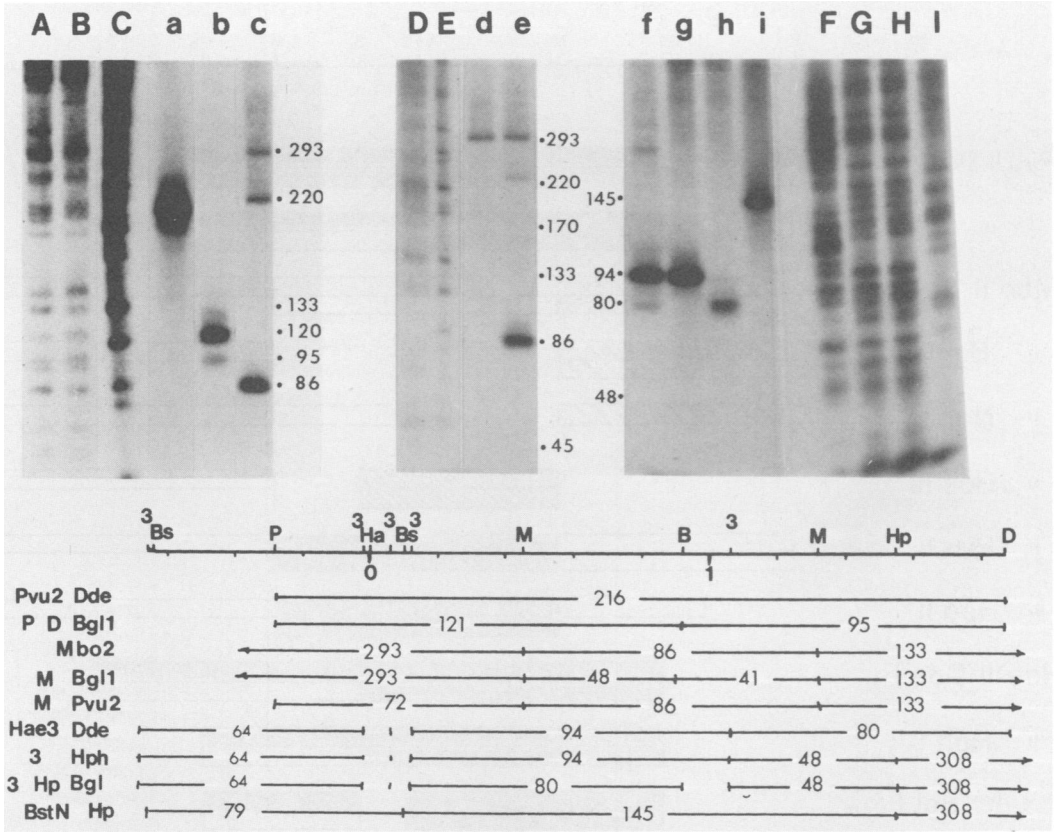

FIG. 3. Binding of SV40 T antigen to separated regions on Py DNA. Reactions were performed as described in the legend to Fig. 1. Before the addition of DNA, the extracts were adjusted to pH 6.75. The following digests were used: *Pvu*II plus *Dde*I (a), *Pvu*II plus *Dde*I plus *Bgl*I (b), *Mbo*II (c), *Mbo*II plus *Bgl*I (d), *Mbo*II (e), *Hae*III plus *Dde*I (f), *Hae*III plus *Hph*I (g), *Hae*III plus *Bgl*I plus *Hph*I (h), and *Bst*NI plus *Hph*I (i). Capital letters indicate the input digests. Lanes a to c, d and e, and f to i are from three separate trials. The map below is the same as that described in the legend for Fig. 2.

precipitation, as well as a measure of the validity and applicability of the assay.

All of the SV40 DNAs (Fig. 6) were digested with *Bst*NI, which gives endpoints at nt 5092 and nt 160. To simplify identification, the digests will be referred to by the other enzymes used. As shown previously, the 232-bp *Hin*dIII fragment (nt 5171 to nt 160) was precipitated efficiently by SV40 T antigen (Fig. 6a, under SV40 T). After digestion of

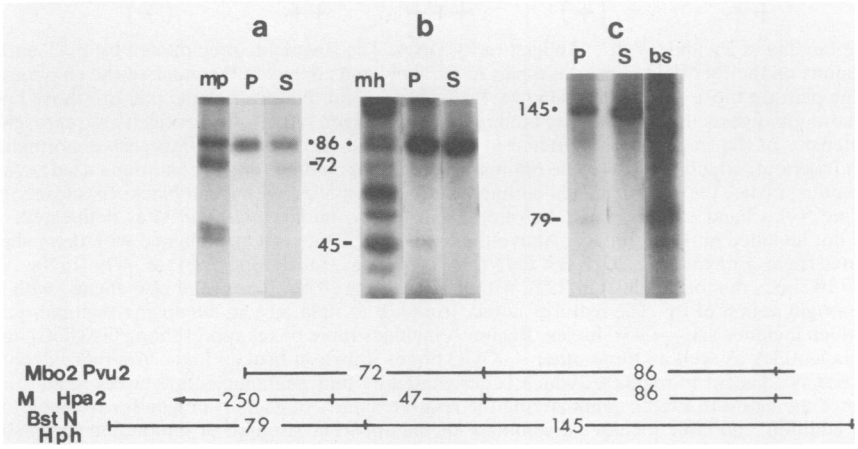

FIG. 4. Binding of Py and SV40 T antigen to the Py origin of replication. Reactions were performed as described in the legend to Fig. 1. The following digests of Py DNA were used: *Mbo*II plus *Pvu*II (a), *Mbo*II plus *Hpa*II (b), and *Bst*NI plus *Hph*I (c). The lanes with input digests are indicated by mp (a), mh (b), and bs (c). (See the map in fig. 5 for reference.) In each set of digests, Py extract was used for the left lane (P) and SV40 extract was used for the right lane (S). Compared with the *Mbo*II-*Pvu*II trial from Fig. 2c, lane a is exposed four times longer. The lanes are truncated for comparison, since lanes b and c are very overexposed. The faint fragments of interest are at 72 bp (a), 45 bp (b), and 79 bp (c). Each of the restriction digests which span the origin region is indicated below, with the sizes of the fragments expected.

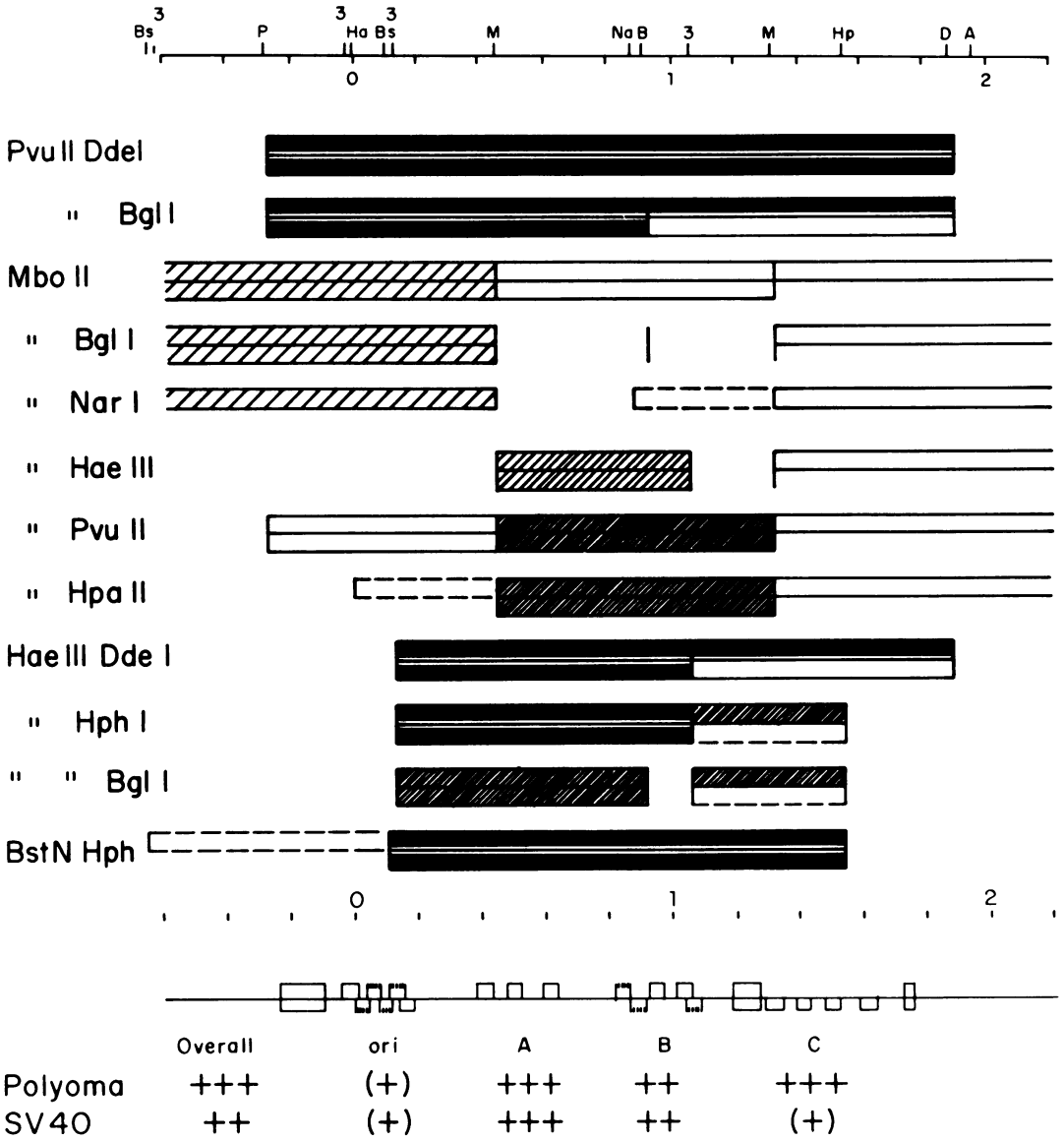

FIG. 5. Summary of the binding of Py and SV40 T Antigen to Py DNA. The fragments precipitated by Py T antigen and SV40 T antigen used to map the binding regions on the Py origin appear as a pair of parallel bands opposite the name of the enzymes used for the digest. For each digest, the upper of the pair are those fragments bound by Py T antigen and the lower of the pair are those bound by SV40 T antigen. The fragments from the restriction digests used to map the binding of Py T antigen and SV40 T antigen are represented as bands of differing darkness relating to the intensity of the fragment precipitated in the DNA immunoassay. The darkness is normalized to binding, for each protein, to the intact origin fragment, which is taken to be optimal efficiency of binding for the conditions used, even if the absolute binding is relatively weak. In descending order, the scale is bands outlined in solid lines filled with (i) all black, (ii) closely hatched, and (iii) loosely hatched, (iv) an open outline, (v) a band outlined in broken lines, and (vi) no outline, with cut sites delineated, indicating no detectable binding. Partial digests are not included in the summary. Above the bound fragments is a map aligned with them showing the location of the restriction sites that generated them. Enzymes: B, *Bgl*I; Bs, *Bst*NI; D, *Dde*I; Ha, *Hpa*II; Hp, *Hph*I; M, *Mbo*II; Na, *Nar*I; P, *Pvu*II; 3, *Hae*III. The numbering noted every 100 bp is that of Soeda et al. (27), with a mark every 20 bp. Below and also aligned with the restriction fragments is a topological map of the origin region of Py. The features noted, from left to right, are an adenosine-thymine-rich stretch and a box next to the origin palindrome, which includes six – –GGC boxes. Region A includes three boxes representing GAGGC, and region B has two solid boxes as canonical pentanucleotides as well as three other – –GGC boxes drawn in broken lines. Another adenosine-thymine-rich region which includes the TATA box is adjacent to region C, which is designated by four pentanucleotide boxes in an orientation opposite to that of region A. At the bottom of the figure is a table summarizing the relative affinity of each T antigen for the intact regulatory region, under the heading of Overall. In addition, the table includes a summary of the apparent strength of interaction for each of the binding regions, including the origin palindrome (ori).

the *Bst*NI fragment with *Dde*I, SV40 T antigen precipitated both the 137-bp fragment containing site 1 (nt 5092 to nt 5131) and the 173-bp fragment containing sites 2 and 3 (nt 5229 to nt 160) (Fig. 6b). As expected, the fragment containing site 1 was precipitated 5- to 10-fold more efficiently than was that containing sites 2 and 3. Cleaving the 173-bp fragment with *Bgl*I dramatically reduced the efficiency of binding to the resulting 167-bp fragment (Fig. 6c and e').

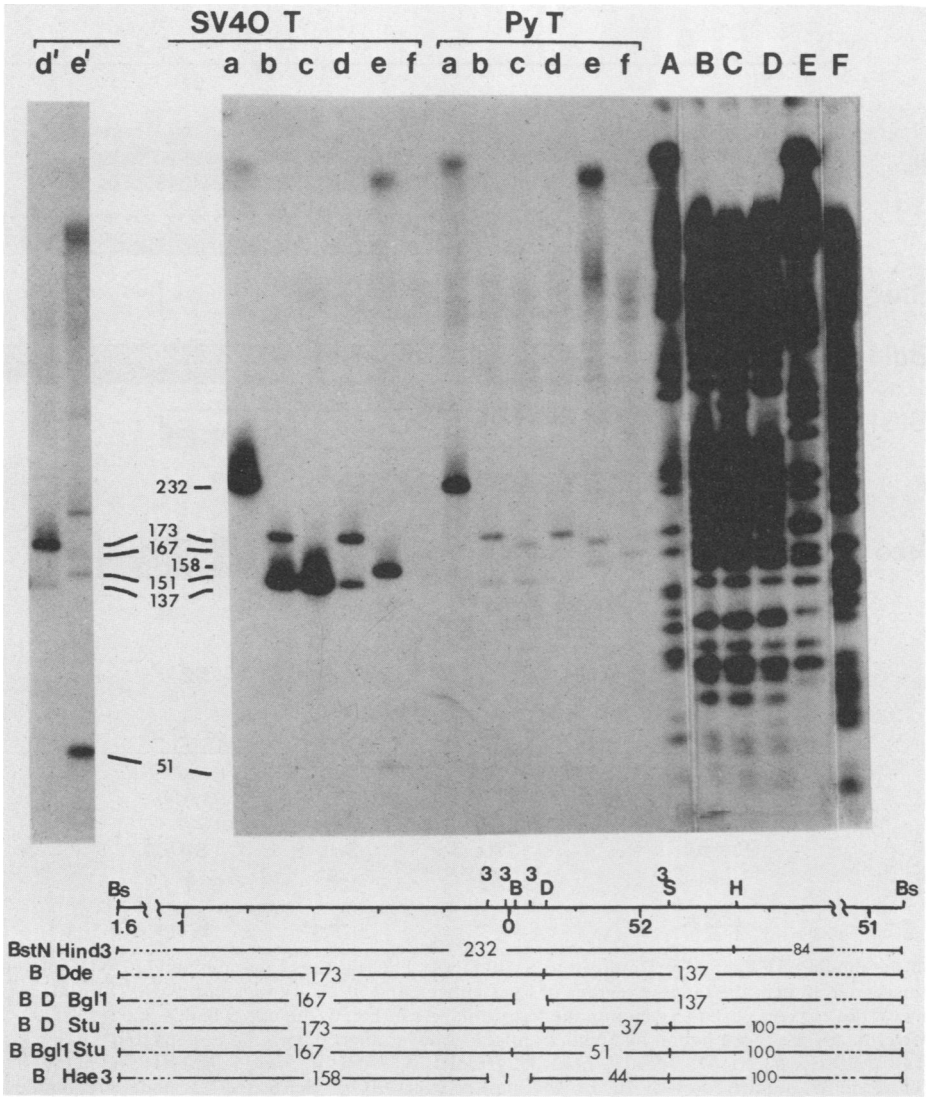

FIG. 6. Binding of SV40 T antigen and Py T antigen to SV40 DNA. Binding was performed as described in the text. Either extract from 5×10^5 Cos 7 cells was adjusted to pH 6.75 or extract from 5×10^6 Py-transformed cells was adjusted to pH 6.50 before the addition of DNA. All digests include *Bst*NI, together with *Hin*dIII (a), *Dde*I (b), *Dde*I plus *Bgl*I (c), *Dde*I plus *Stu*I (d), *Bgl*I plus *Stu*I (e), or *Hae*III (f). The *Stu*I digests were occasionally incomplete, so (d') and (e') were included to show Cos 7 extract in a different trial with more complete *Stu*I digestion, to allow visualization of bands at 167 bp, and to compare the 51-bp band in (e') with the 173-bp band in (d'). The sets including either SV40 extract or Py extract are indicated. Capital letters indicate the input digests. A map below, using the Buchman et al. numbering of SV40 as described in reference 33 (in U of 100 bp), shows the location of the enzyme sites used. The sizes of fragments for each of the input digests are indicated. These include Bs, *Bst*NI; D, *Dde*I; H, *Hin*dIII; S, *Stu*I; and 3, *Hae*III.

Also, cleaving the 137-bp fragment with *Stu*I at nt 5190 resulted in a 37-bp fragment (nt 5190 to nt 5228) to which binding was not detected (Fig. 6d), thereby indicating that *Stu*I disrupts site 1. Any partial digests by *Stu*I were most effectively exploited by SV40 T antigen (e.g., the 137-bp fragment in lane d and the 151-bp fragment in lane e [Fig. 6]). Thus, it appeared that the portions of site 1 required for efficient binding by SV40 T antigen border the *Stu*I site and that the solitary GAGGC on the late side of the *Stu*I site does not act efficiently by itself. In addition, the portion of site 2 required for strong binding includes the sequences between the *Dde*I site at nt 5228 and the *Bgl*I site at nt 5239 (12, 15). This region contains the first pentanucleotide of site 2. Cutting the 311-bp *Bst*NI fragment with *Stu*I and *Bgl*I resulted in efficient precipitation of the 51-bp fragment (nt 5190 to nt 5240) as well as the weak 167-bp fragment (same as above) (Fig. 6e and e'). The dramatic increase in efficiency of precipitation occurs after increasing the length of the 37-bp *Stu*I-*Dde*I fragment (or the 44-bp *Hae*III fragment from between nt 5291 and nt 5234) to include the first pentanucleotide of site 2 so as to isolate a palindrome bracketed by facing pentanucleotides. The efficient precipitation of the 51-bp *Stu*I-*Bgl*I fragment may indicate that some synergy occurs between polypeptides bound to damaged site 1 and damaged site 2. To detect binding at site 3, the 311-bp *Bst*NI fragment was digested with *Hae*III. However, the 158-bp fragment containing site 3 (nt 6 to nt 160) was undetectable (Fig. 6f). From this series of digests, SV40 T antigen binds most strongly to site 1 and less efficiently to site 2, whereas binding to site 3 was undetectable under the conditions used in this assay. The region around the *Stu*I site at site 1 and the first pentanucleotide in site 2 appear to be

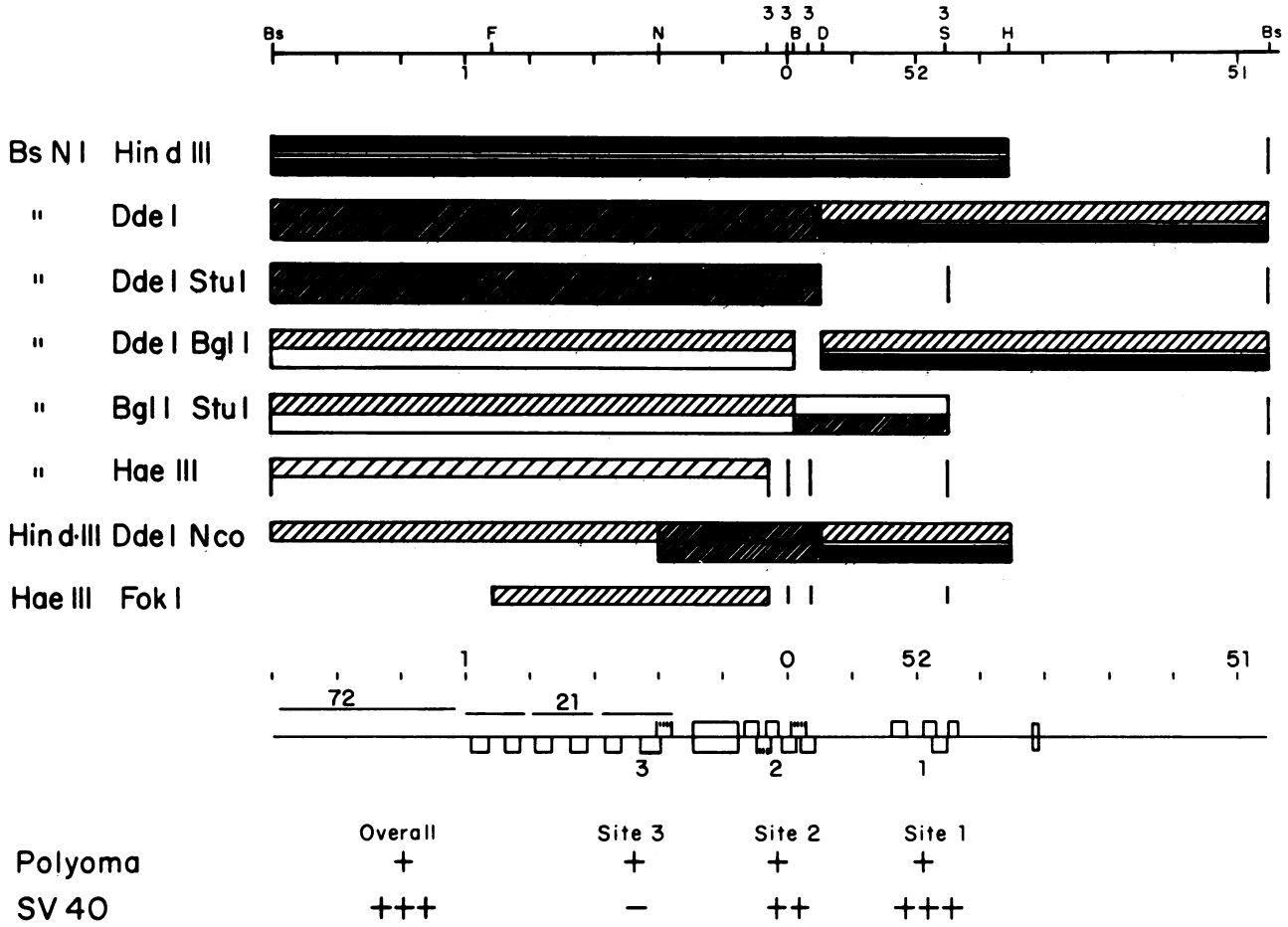

FIG. 7. Summary of the binding SV40 and Py T antigen to SV40 DNA. The fragments precipitated by Py T antigen and SV40 T antigen from the restriction digests used to map the binding regions of Py T antigen and SV40 T antigen on SV40 DNA are represented by a pair of bands opposite the name of the enzymes used to generate the fragments. As in Fig. 5, the upper bands of each set are those bands bound by Py T antigen and the lower bands are those bound by SV40 T antigen. Above, aligned with the precipitated bands, is a restriction map with the following symbols: B, *Bgl*I; Bs, *Bst*NI; D, *Dde*I; F, *Fok*I; H, *Hin*dIII; N, *Nco*I; S, *Stu*I; 3, *Hae*III. The numbering is that of Buchman et al. as described in reference 33, numbered every 100 bp and marked every 20 bp. Below the fragments precipitated is a topological map of the origin of SV40. From left to right are indicated the enhancer 72, the promoter as three 21-bp repeats with boxes indicating sequences it contains as GPuGGC, and a – –GGC box at the edge of the promoter in the opposite orientation. An adenosine-thymine-rich stretch which includes the TATA box is adjacent to the origin palindrome, which contains two pairs of GAGGC residues (solid lines) in opposite orientation. Another cluster of pentanucleotide boxes in both orientations appears to the left of the ATG start codon. Under the map, binding sites 3, 2, and 1 are indicated. As in Fig. 5, a table summarizing the interaction appears at the bottom of this figure.

the sequences required for strong binding to SV40 DNA. As this hierarchy of affinities is consistent with previously published work (2, 30, 32), it appears that the DNA immunoassay gives some accurate measure of the extent of binding regions through the use of restriction fragments.

The response of Py T antigen to SV40 DNA was strikingly different. The intact origin within the 232-bp *Hin*dIII fragment was precipitated (Fig. 6a, under Py T), although with reduced efficiency relative to how it precipitates the Py regulatory region. However, unlike SV40 T antigen, after cleavage of the *Bst*NI fragment with *Dde*I, the 173-bp fragment (sites 2 and 3) was brought down at least two times more efficiently than the 137-bp fragment (site 1) (Fig. 6b). Furthermore, cleavage of the 173-bp fragment with *Bgl*I reduced the intensity of the resulting 167-bp fragment to equal that of the 137-bp fragment (Fig. 6c). From the intensity of the bands in the 167-bp *Bgl*I fragment (nt 5240 to nt 160), it appears unlikely that the remainder of site 2 contributes strongly to the binding of this fragment to any degree greater than how SV40 T antigen binds to this remnant. This was confirmed by digestion of the *Bst*NI fragment with *Hae*III. The intensity of the resulting 158-bp band (Fig. 6f) was not much different from the 137-bp *Dde*I band or the 167-bp *Bgl*I fragment. This indicates that Py T antigen binds to site 3 with an apparent affinity similar to that for either site 1 or site 2. Py T antigen also bound to the site 3 fragment with ends at nt 6 and at nt 97 as cut by *Hae*III and *Fok*I (data not shown). Therefore, although Py T antigen binds to each of the three binding regions on SV40 DNA, unlike SV40 T antigen, Py T antigen does not demonstrate relative differences in efficiency of binding these sites on the SV40 origin. The results from comparing the binding of SV40 T antigen and Py T antigen to SV40 DNA is summarized (Fig. 7).

Another example illustrates the difference in binding of SV40 and Py T antigens to the two viral DNAs (Fig. 8). The binding of SV40 T antigen appears to be efficient on both DNAs, approaching, on the 94-bp *Hae*III Py fragment, an

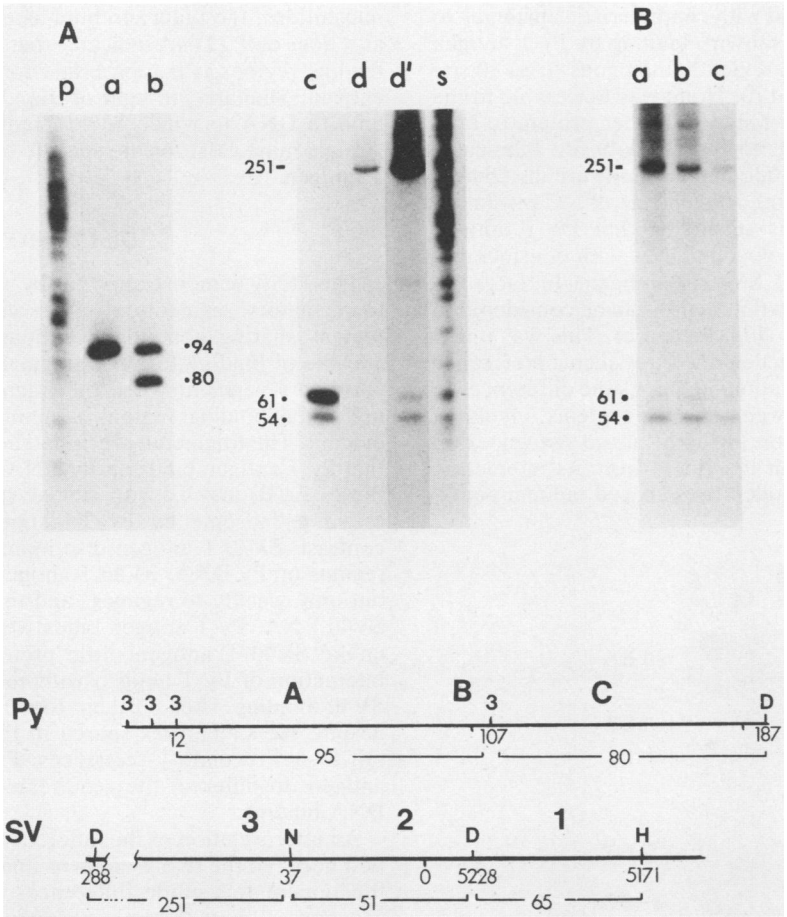

FIG. 8. Efficiency of binding of SV40 and Py T antigens to separated binding regions from SV40 and Py DNA. (A) With the *Hae*III-plus-*Dde*I digest of p373A2 DNA (p, a, and b), extract from 2×10^6 Cos 7 cells (a) or 2×10^6 Py-transformed cells (b) was used. For binding to the pSVRI *Hin*dIII-plus-*Dde*I-plus-*Nco*I digest (c, d, and s), extract from 5×10^5 Cos 7 cells (c) or 5×10^6 Py-transformed cells (d) was used. A 10-fold longer exposure of lane d appears in lane d'. Input digests are in lanes p and s. A map of the restriction enzyme sites is shown below. The location of the fragments in the SV40 digest is as follows. The 61-bp fragment (site 1) extends from the *Hin*dIII site at nt 5171 to the *Dde*I site at nt 5228. The 54-bp fragment (site 2) lies between the *Dde*I site and the *Nco*I site at nt 37. The 251-bp fragment (site 3) extends from the *Nco*I site to the next *Dde*I site at nt 288. Unlike previous trials, the extracts were bound through antiserum to *S. aureus* and washed before the binding of DNA. (B) Response of Py T antigen to the same *Hin*dIII-*Dde*I-*Nco*I digest of pSVRI as in (A) but performed in binding buffer containing 0.15 (a), 0.18 (b), or 0.22 (c) M NaCl. Washes of the DNA in the immune pellet were performed in 0.15 M NaCl wash buffer. The binding reaction was performed as described in the legend to Fig. 6. The same three fragments as those bound in part A are indicated.

intensity similar to that of Py T antigen on that fragment (Fig. 8A, a and b). However, Py T antigen binds weakly to SV40 DNA, with site 1 and site 2 only weakly apparent, even after a 10-times longer exposure (Fig. 8d and d'). It is noteworthy that with the SV40 origin separated into these smaller fragments, the binding of Py T antigen to site 2 appears to be about three times greater than that to site 1. Again, despite this weak overall binding, the affinity of Py T antigen for SV40 site 3 is much greater than that of SV40 T antigen for this site. SV40 T antigen binding to site 3 is not evident even after much longer exposures. Although Py T antigen does bind to SV40 site 3, binding to this region is more sensitive to increased salt concentration and shows a differential reduction relative to its binding to separated sites 1 and 2. Between the salt range of 0.15 and 0.22 M NaCl, the efficiency of precipitation of isolated site 3 dropped to 13% its initial value, whereas those of sites 1 and 2 were reduced only to 44% of their initial value (Fig. 8B). This greater sensitivity indicates that affinity of Py T antigen for site 3 depends upon interactions which are less optimal than for the other binding regions, possibly being more sensitive to ionic interactions with the DNA backbone than the other binding regions.

Binding to *Alu* family DNA. When considering the ability of the two T antigens to bind similar viral DNA sequences, the question was raised as to whether there were any nonviral DNA sequences to which the two proteins would bind. A striking homology has been noted between half of the origin of replication of the papovaviruses and the designated consensus sequence of the *Alu* family middle repetitive DNA (11). Examinations were made of the binding of Py T antigen and SV40 T antigen to a representative of *Alu* family DNA located 3' to the human beta-globin locus (23). This DNA included sequences homologous to half of SV40 site 1, as well as to half of SV40 site 2. The sequences of the two homologous regions are depicted, together with a comparison to SV40 sequences (Fig. 9). Note that a single base change interrupts the homology in each case. The two regions are separable through restriction digests. The site 1 homology appears in a 120-bp *Hin*cII-*Bst*NI fragment or in a 165-bp *Hin*cII-*Dde*I fragment. The region homologous to site 2 appears in a 65-bp *Bst*NI fragment. Under conditions in

which both T antigens bound with characteristic efficiency to both viral DNAs (data not shown), binding by Py T antigen to fragments containing the region homologous to SV40 site 1 was observed (Fig. 9b and d). None was detectable to the 65-bp *Bst*NI fragment. The failure of either protein to bind efficiently to the region homologous to SV40 site 2 indicates that two direct pentanucleotide repeats alone are insufficient for tight binding. Furthermore, the change of a C residue to T may be disruptive. It is surprising that Py T antigen precipitated the site 1 homologous DNA with considerably greater efficiency than did SV40 T antigen. In fact, the binding of SV40 T antigen to that region can be considered to be only marginally specific (Fig. 9a and c). This was unexpected since SV40 T antigen demonstrates such a preference for SV40 site 1. This observation highlights the differences in sequence requirements between the two proteins, involving the content or the orientation, or both, of the sequences on DNA. However, although it is evident from its interaction with Py DNA that Py T antigen uses spaced tandem pentanucleotides, the failure to bind as efficiently on SV40 DNA as it does on Py DNA indicates that some aspect of the SV40 binding regions is inappropriate for optimal usage by Py T antigen. Similarly, in light of the ability of Py T antigen to bind to DNA to which SV40 T antigen binds poorly, constraints must exist for the specific binding of DNA by SV40 T antigen.

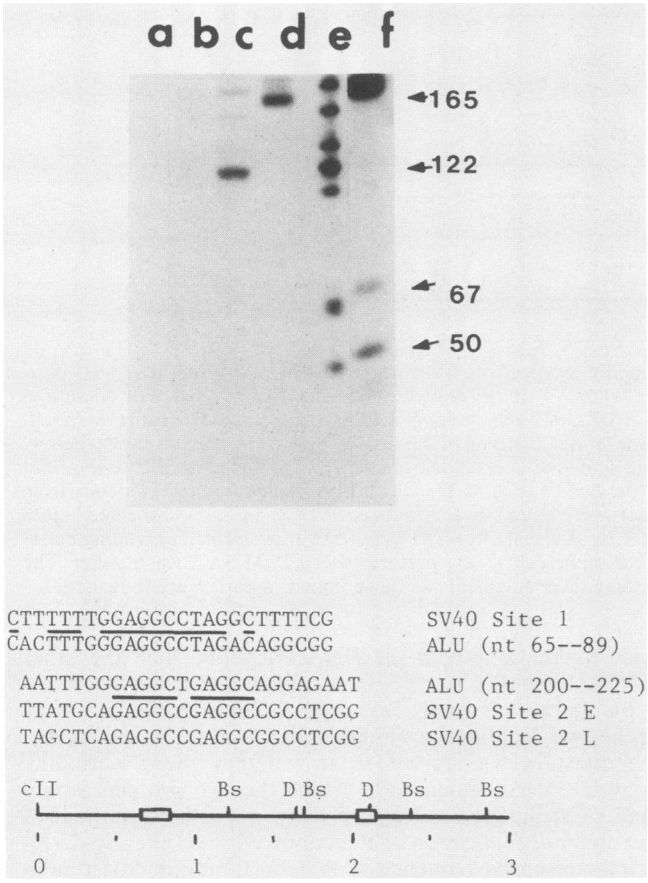

FIG. 9. Binding of SV40 and Py T antigen to *Alu* DNA. Extracts from 2×10^6 Cos 7 cells (a and b) or from 5×10^6 Py-transformed cells (c and d) were attached to *S. aureus* through antiserum and washed before DNA binding. The complexes were bound to *Alu* DNA (23) digested with either *Hin*cII plus *Bst*NI (a and c) or *Hin*cII plus *Dde*I (b and d). Lanes e and f are the respective input DNAs. Below are the pertinent sequences compared with their homologous regions in SV40 with the homologies underlined. Note the homology between the two *Alu* DNA sequences. At the bottom is a rough map showing the location of the enzymes used: Bs, *Bst*NI; D, *Dde*I; cII, *Hin*cII. On the map, nt 1 corresponds to nt 15984 as sequenced by Poncz et al. (23). The homology regions on the map are indicated by boxes. Numbering is every 100 bp.

DISCUSSION

This study demonstrates that Py and SV40 T antigens bind to regulatory segments of either virus at multiple discrete regions sharing homology. Although the T antigens were capable of binding to DNA fragments containing the intact regulatory segment when the fragment was separated into individual binding regions a number of differences were evident. The fragments precipitated by Py T antigen indicate that Py T antigen binds on Py DNA at four regions, three of these, A, B, and C, with similar strengths. Binding to the origin palindrome by Py T antigen is much weaker. In contrast, SV40 T antigen discriminates among the different regions on Py DNA: SV40 T antigen binds regions A and B but only weakly to region C and the origin palindrome. On SV40 DNA, Py T antigen binds weakly, overall. However, unlike SV40 T antigen, little preference is evident in the interaction of Py T antigen with each of the three different SV40 binding sites. Taken together, this indicates that despite the similarities shared in the ability to bind to the other viral regulatory sequences, Py T antigen and SV40 T antigen do differ in the requirements of each for specific DNA binding.

An interpretation of the differences in the specific interaction between the two T antigens and DNA can be made as a function of the subtle differences between the nucleotide sequences. From this it is possible to derive guidelines for the arrangement of the pentanucleotides required to support tight binding by each T antigen. In all binding regions, the sequence GPuGGC is arranged in groups which must be maintained intact to support strong specific binding (2; this study). However, the failure of the T antigens to bind all sets of pentanucleotides with equal strength implies that constraints exist for optimal binding. Characteristics which might affect specific binding include the sequences within each pentanucleotide (composition of the – – GGC box), the sequences surrounding the pentanucleotide (context), and the orientation of the pentanucleotides with respect to one another, together with the distance in nt's between them (orientation and spacing). The nt sequences of the regulatory regions of the two viruses which contain the pentanucleotides in the binding regions are reproduced (Fig. 10).

Composition of the other two nucleotides in the – – GGC boxes on SV40 has been considered by DeLucia et al. (2) to influence the strength of SV40 T-antigen DNA binding. A similar effect is evident with the binding of Py T antigen to DNA. Both region B and region C on Py include GGGGC, but these were not detected by Gaudray et al. (7) and were reported by Dilworth et al. to be less strongly bound than region A (3). Furthermore, the origin palindrome, which is much more weakly bound by Py T antigen, has only two classical GAGGC, between which are interposed four other – – GGC stretches (6, 14, 34). Similarly, the binding of Py T antigen to SV40 site 3, which is unlike SV40 sites 1 and 2 in that it lacks any pentanucleotides of the composition GAGGC, is differentially sensitive to increases in salt concentration. However, it is difficult to determine from our studies whether SV40 T antigen and Py T antigen differ in

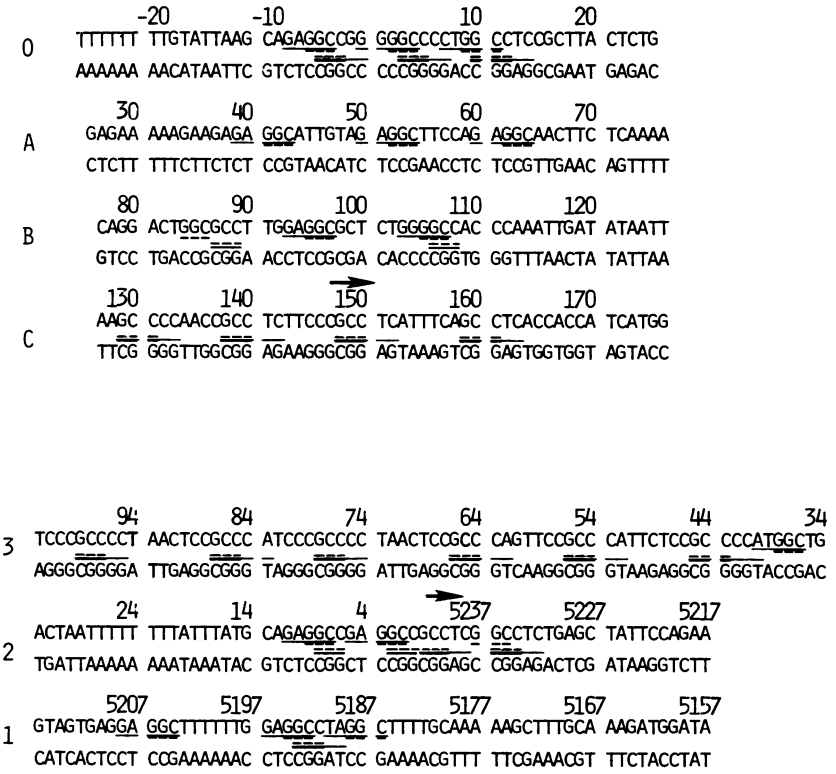

FIG. 10. Nucleotide sequence of Py and SV40. Nucleotide sequence of Py (6, 27, 34) and SV40 (BBB in reference 33), separated into each of the segments which include each of the binding regions. For Py, these binding regions are O (origin palindrome), A, B, and C. On SV40, these are 3, 2, and 1. Pentanucleotide boxes are indicated in the nucleotide sequence by NNGGC. Also, the major early transcription start site is indicated by an arrow over nt 150 on Py and over nt 5240 on SV40.

their tolerance for other deviations from the canonical pentanucleotide.

The nucleotide context of the pentanucleotide can influence how it is bound by the T antigen molecule. Examination of Py region C and SV40 site 3, each of which is bound poorly by SV40 T antigen, reveals that both contain the motif GPuGGC, followed immediately 3′ by the dinucleotide GG to generate GPuGGCGG. In site 3, which is undetectable in this assay, every pentanucleotide has this form. However, in region C, which is weakly bound but still visible, only the two internal pentanucleotides of four have this context. Pentanucleotide context may help to explain the overall weak binding of Py T antigen on SV40 DNA. A survey of the Py regulatory region reveals that the dinucleotide CC is conspicuous in its absence 5′ to the pentanucleotides. It appears two of six times on the origin palindrome directly 5′ to a – – GGC box but nowhere in the strong binding regions. In contrast, a comparison of the pentanucleotides on the SV40 regulatory region reveals that CC frequently is directly 5′ to a – – GGC box: two of four times in site 1, four of six times in site 2, and one of seven times in site 3. However, this context effect may be less pronounced for the binding of Py T antigen to DNA than for that observed for SV40 T antigen.

Examination of the sequence of the DNA fragments bound by the T antigens reveals a preference for binding to groups of pentanucleotides which are spaced such that the GGC residues in each set are on the same side of the DNA helix (12, 32). A review of the spacing within each set indicates that pentanucleotides in a tail-to-head orientation within the binding regions on SV40 DNA are separated by 7 bp (±1 bp), whereas on those on Py are spaced usually by 5 bp. Within the origin palindromes, pairs of pentanucleotide repeats in the tail-to-tail orientation on SV40 are spaced by 7 bp, whereas those on Py are spaced by 5 bp. It is likely that the greater spacing of the pentanucleotides on SV40 hinders their effective utilization by Py T antigen, conceivably by the prevention of mutual stabilization of polypeptides on adjacent pentanucleotides. However, the closer spacing of the pentanucleotides on Py DNA appears to be better tolerated by SV40 T antigen. Both proteins are reported to bind to poly-*Xho*I linkers (21), in which the pentanucleotides in the tail-to-head orientation are spaced by 3 bp but in the tail-to-tail orientation are spaced by 6 bp.

Alternative spacial arrangements of – – GGC boxes may contribute to specific DNA binding. The binding of Py to *Alu* DNA homologous to SV40 site 1 raised the possibility that Py T antigen binds more efficiently than SV40 T antigen to pentanucleotides in a diagonally opposed configuration. This would place the T antigen molecules on opposite sides of the DNA helix, in contradiction of the model for SV40 T antigen binding presented by Jones and Tjian (12), which proposed that T antigen is arranged to be on only one side of the DNA helix. A similar diagonal opposition appears in SV40 site 2, with the pentanucleotide TCGGC on one strand opposed to GAGGCCGAGGC on the other. These TCGGC strings include the dimethyl sulfate-protected guanine residues at nt 5234 and nt 7 (2, 32), which appear between the more ideal pentanucleotides. Even so, instead of denoting another binding region, these residues may reflect merely the difference in the way that SV40 T antigen binds to pentanucleotides in the tail-to-tail orientation as opposed to those

arranged tail to head. However, another grouping includes this TCGGC, in the form CCTCGGC, 5 bp 3′ to a GAGGC on the same strand, and this tandem repeat may be the source of binding by Py T antigen to SV40 site 2. Likewise, on the *Alu* DNA fragment bound well by Py T antigen, but weakly by SV40 T antigen, there appears GAGGC 5 bp 5′ to CAGGCGG. Therefore, it is possible to correlate the binding behavior of each T antigen with its response to the context as well as the spacing of the pentanucleotides.

While this manuscript was in preparation, Cowie and Kamen (1), by using DNase I footprinting under conditions of low binding stringency, also identified weak binding by Py T antigen both to the Py origin palindrome and to other weak regions in the enhancer. The region bound in the enhancer has a grouping of three – –GGC boxes, two of which are separated by 8 intervening bp. This interval can be considered to be too great to permit efficient interaction between bound Py T antigen monomers. As binding was evident only under conditions of reduced ionic strength, it is not surprising that we failed to detect binding to this region. However, in some trials with SV40 T antigen, we were able to observe the specific precipitation of a *Pvu*II fragment harboring only this binding region to an extent greater than that observable in Fig. 4 (A. Scheller and C. Prives, unpublished data). In a recent study, Wright et al. (36) examined the binding of SV40 T antigen to nonviral DNA containing single pentanucleotides either with compositions of considerable variety or arrayed in groups not in optimal orientation. Binding to these sequences, as observed through DNase I protection, was evident only under conditions of very low stringency. These observations, along with ours, indicate a requirement for a critical spacing of the pentanucleotides to promote specific DNA binding by T antigen. Conceivably, this may ensure the interaction of T antigen with viral DNA functioning over the background of cellular chromatin containing sequences partially homologous to the viral origin. Furthermore, since regulatory regions at the 5′ end of genes are guanosine-cytosine rich, we suggest that the inhibitory contexts imposed by dicytosine or diguanosine residues adjacent to a – –GGC box in cellular DNA are significant. These constraints can minimize binding by T antigen to areas of the genome in which pentanucleotides are most likely to appear.

The question of whether either of the two T antigens would complement the other for viral functions involving DNA binding was not resolved. Although it appears that binding of both T antigens to the viral origin palindrome differs, the possibility remains that weak binding is sufficient to align cellular components and stimulate initiation of viral DNA replication or modify transcriptional activity at low levels. Despite the requirement for specificity in DNA binding, considerable plasticity appears in the system. Recently a number of mutations and small deletions within the origin palindrome of Py and SV40 have been isolated (4, 14, 16, 26, 34). A wide spectrum of effects was observed upon DNA replication, including little or no change, outright abolition, or dependence upon a second-site pseudoreversion within the protein coding sequence (18, 26). As Tegtmeyer et al. have noted (29), the function of T antigen bound to the origin must allow DNA and RNA multienzyme complexes to pass the binding regions, which implies that T antigen either should remain bound to single-stranded DNA or should dissociate from the DNA, in which case weak binding may be a desirable feature. Perhaps nonideal pentanucleotides like TCGGC in site 2 are involved. A point mutation within this sequence was described by Shortle and Nathans as forming small sharp plaques but overproducing both viral DNA and T antigen protein (26). In this case, weak binding by SV40 T antigen might be responsible.

Interactions between molecules bound to adjacent regions have been observed through both biochemical and genetic manipulations (4, 5, 20). Moreover, the structure of the T polypeptide must include in its complexity the coordination of the interactions between individual units, as when the protein is bound to pentanucleotides in different orientations (35). However, the configuration of the pentanucleotides as well as the number in each set indicates that the two proteins should display significant differences in shape, probably to interact with cellular components which differ slightly between each normal permissive host.

ACKNOWLEDGMENTS

We thank B. Kamen, A. Cowie, P. Gaudray, and J. Hassell for helpful discussions and communication of unpublished data. Initial experiments were done with the help of M. Israel and K. Chowdhury.

This work was supported by Public Health Service grant Ca-26905 to C.P. from the National Institutes of Health.

LITERATURE CITED

1. **Cowie, A., and R. Kamen.** 1984. Multiple binding sites for polyomavirus large T antigen within regulatory sequences of polyoma DNA. J. Virol. **52**:750–760.
2. **DeLucia, A. L., B. A. Lewton, R. Tjian, and P. Tegtmeyer.** 1983. Topography of simian virus 40 A protein-DNA complexes: arrangement of pentanucleotides interaction sites at the origin of replication. J. Virol. **46**:143–150.
3. **Dilworth, S., A. Cowie, R. Kamen, and B. Griffin.** 1984. DNA binding activity of polyoma virus large tumor antigen. Proc. Natl. Acad. Sci. U.S.A. **81**:1941–1945.
4. **DiMaio, D., and D. Nathans.** 1980. Cold-sensitive regulatory mutants of simian virus 40. J. Mol. Biol. **140**:129–142.
5. **DiMaio, D., and D. Nathans.** 1982. Regulatory mutants of simian virus 40: effect of mutations at a T antigen binding site on DNA replication and expression of viral genes. J. Mol. Biol. **156**:531–558.
6. **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 acid homology with SV40. Cell **17**:715–724.
7. **Gaudray, P., C. Tyndall, R. Kamen, and R. Cuzin.** 1981. The high affinity binding site on polyoma virus DNA for the large-T protein. Nucleic Acids Res. **9**:5697–5710.
8. **Gluzman, Y.** 1981. SV40 transformed simian cells support the replication of early SV40 mutants. Cell **23**:175–182.
9. **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.
10. **Hayday, A., 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 which has extensive amino acid homology with simian virus 40 large T-antigen. J. Virol. **45**:693–699.
11. **Jellinek, W., T. Toomey, L. Leinwand, C. Duncan, P. Biro, P. Chaudry, S. Weissman, C. Rubin, C. Houck, P. Deininger, and C. Schmid.** Ubiquitous, interspersed repeats in mammalian genomes. Proc. Natl. Acad. Sci. U.S.A. **77**:1398–1402.
12. **Jones, K., and R. Tjian.** 1984. Essential contact residues within SV40 large T antigen binding sites I and II identified by alkylation interference. Cell **36**:155–162.
13. **Kamen, R., P. Jat, R. Treisman, J. Favaloro, and W. Folk.** 1982. 5′ termini of polyoma virus early region transcripts synthesized in vivo by wild-type virus and viable deletion mutants. J. Mol. Biol. **159**:189–224.
14. **Katinka, M., and M. Yaniv.** 1983. DNA replication origin of polyoma virus: early proximal boundary. J. Virol. **47**:244–248.
15. **Lewton, B. A., A. L. DeLucia, and P. Tegtmeyer.** 1984. Binding of simian virus 40 A protein to DNA with deletions at the origin of replication. J. Virol. **49**:9–13.
16. **Luthman, H., M.-G. Nilsson, and G. Magnussen.** 1982. Non-con-

tiguous segments of the polyoma genome required in cis for DNA replication. J. Mol. Biol. **161**:533–550.
17. **Maniatis, T., A. Jeffrey, and D. Kleid.** 1975. Nucleotide sequence of rightward operator of phage lambda. Proc. Natl. Acad. Sci. U.S.A. **72**:1184–1188.
18. **Margolskee, R., and D. Nathans.** 1984. Simian virus 40 mutant T antigens with relaxed specificity for the nucleotide sequence at the viral DNA origin of replication. J. Virol. **49**:386–393.
19. **McCormick, F., D. Lane, and S. Dilworth.** 1982. Immunological cross-reaction between large T-antigen of SV40 and polyomavirus. Virology **116**:382–387.
20. **Myers, T., 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.
21. **Pomerantz, B. J., and J. A. Hassell.** 1984. Polyomavirus and simian virus 40 large T antigens bind to common DNA sequences. J. Virol. **49**:925–937.
22. **Pomerantz, C. J., C. R. Mueller, and J. A. Hassell.** 1983. Polyomavirus large T antigen binds independently to multiple unique regions on the viral genome. J. Virol. **47**:600–610.
23. **Poncz, M., E. Schwartz, M. Balantine, and S. Surrey.** 1983. Nucleotide sequence analysis of the δ/β globin gene regions in humans. J. Biol. Chem. **258**:11599–11609.
24. **Rasoulzadegan, M., A. Cowie, A. Carr, N. Glaichenhaus, R. Kamen, and F. Cuzin.** 1982. The roles of individual polyoma virus early proteins in oncogenic transformation. Nature (London) **300**:713–718.
25. **Scheller, A., L. Covey, B. Barnet, and C. Prives.** 1982. A small subclass of SV40 T antigen binds to the viral origin of replication. Cell **29**:375–383.
26. **Shortle, D., and D. Nathans.** 1978. Regulatory mutants of simian virus 40 constructed with base substitutions in origin of DNA replication. J. Mol. Biol. **131**:801–817.
27. **Soeda, E., J. Arrand, N. Smolar, J. Walsh, and B. Griffin.** 1980. Coding potential and regulatory signals of the polyoma virus genome. Nature (London) **283**:445–453.
28. **Soprano, K. J., N. Galanti, G. J. Jonak, S. McKerchner, J. M. Pipas, K. W. C. Peden, and R. Baserga.** 1983. Mutational analysis of simian virus 40 T antigen: stimulation of cellular DNA synthesis and activation of silent rRNA genes by mutants with deletions in the T-antigen gene. Mol. Cell Biol. **3**:214–219.
29. **Tegtmeyer, P., B. A. Lewton, A. L. DeLucia, V. G. Wilson, and K. Ryder.** 1983. Topography of simian virus 40 A protein-DNA complexes: arrangement of protein bound to the origin of replication. J. Virol. **46**:151–161.
30. **Tenen, D., T. Taylor, L. Haines, M. Bradley, R. Martin, and D. Livingston.** 1983. Binding of simian virus 40 large T antigen from virus-infected cells to wild-type and mutant viral replication origins. J. Mol. Biol. **168**:791–807.
31. **Thomas, T., P. Vollmer, and W. R. Folk.** 1981. Nucleotide sequence changes in polyoma virus A gene mutants. J. Virol. **37**:1094–1098.
32. **Tjian, R.** 1978. Protein-DNA interactions at the origin of simian virus 40 DNA replication. Cold Spring Harbor Symp. Quant. Biol. **43**:655–662.
33. **Tooze, J. (ed.).** 1981. Molecular biology of tumor viruses, part 2: DNA tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Treizenberg, S. J., and W. R. Folk.** 1984. Essential nucleotides in the polyomavirus origin region. J. Virol. **51**:437–444.
35. **Wilson, V. G., M. J. Tevethia, B. A. Lewton, and P. Tegtmeyer.** 1982. DNA binding properties of simian virus 40 temperature-sensitive A proteins. J. Virol. **44**:458–466.
36. **Wright, P. J., A. L. DeLucia, and P. Tegtmeyer.** 1984. Sequence-specific binding of simian virus 40 A protein to nonorigin and cellular DNA. Mol. Cell Biol. **4**:2631–2638.