Characterization of an Immortalizing N-Terminal Domain of Polyomavirus Large T Antigen

PATRICIA S. HOLMAN, OLE V. GJOERUP, TANJA DAVIN, AND BRIAN S. SCHAFFHAUSEN*

Department of Biochemistry, Tufts University School of Medicine, Boston, Massachusetts 02111

Received 9 August 1993/Accepted 29 October 1993

Polyomavirus large T antigen has an N-terminal domain of approximately 260 amino acids which can immortalize primary cells but lacks sequences known to be required for DNA binding and replication. Treatment of full-length large T with either V8 protease or chymotrypsin yields an N-terminal fragment of 36 to 40 kDa and a C-terminal fragment of ~60 kDa. This finding suggests a division of the protein into two domains. Proteolysis experiments show that the N-terminal domain does not have strong physical association with the rest of the protein. It also does not self-associate. A construct expressing only the N-terminal 259 amino acids is sufficient for immortalization. The independently expressed N-terminal domain is multiply phosphorylated, although at a lower level than the same region in full-length large T. The 259-residue protein binds to both pRb and p107 with somewhat lower efficiency than the full-length protein.

Polyomavirus large T antigen is a protein with a variety of functions. In productive infection, large T initiates viral DNA replication (16), and in transformation, it is responsible for integration (10) and excision (3) of the viral genome. It represses early message production (8, 14) but activates cellular transcription (22). Analysis of transformation of both established and primary cells suggests additional functions. While middle T antigen may be the most obvious transforming protein of polyomavirus, large T also plays a role. Transformation of primary embryo fibroblasts is a multistep process that requires both middle T and large T (24). One function of large T is immortalization, which allows primary cells to grow in long-term culture without entering crisis (31). Although most cells transformed by temperature-sensitive large T mutants remain transformed when shifted to the nonpermissive temperature (15), some established cell lines retain a continuing requirement for large T or a fragment of it (32, 41). In FR3T3 cells, large T can activate middle T for transformation (6). Additional insight into effects of large T on growth control comes from analysis of myogenesis. Large T prevents differentiation of myoblasts and the induction of muscle-specific genes (28).

Large T antigen of polyomavirus, like that of simian virus 40 (SV40), appears to be a modular protein. For example, large T is a DNA-binding protein (9, 39). Genetic analysis has identified a region from residues 282 to approximately 400 that is critical for this activity (46, 47). A functional domain of large T involved in transformation has been identified by cleavage of polyomavirus DNA. The *Bam*HI-*Eco*RI DNA fragment that encompasses the origin and proximal portion of the early region is capable of immortalizing cells (31). This fragment and a similar fragment extending to the *Hind*III site (2) are nearly as efficient as the intact protein. The fragment that ends at the *Eco*RI site can cooperate with middle T (31) or *ras* (24, 29) for transformation of primary cells. The N-terminal ~40% of large T alone is therefore capable of immortalizing primary cells and complementing middle T in transformation. Despite its important biological activities, this fragment has been subjected to

very limited biochemical analysis. Proteins of various sizes (40, 37, and 35 kDa) have been reported in independently isolated cell lines (19, 31). The construct used in these experiments did not contain a termination codon, however, making it difficult to determine exactly what protein species were present.

We are interested in dissecting the structure and function of polyomavirus large T. We have defined an N-terminal domain that immortalizes and binds pRb but is clearly distinct from known DNA-binding elements and other sequences identified as being required for replication. This domain of 259 amino acids is smaller than that suggested by previous genetic data. There appears to be little interaction between it and the C-terminal portion of the protein.

MATERIALS AND METHODS

DNA constructs. Plasmid LTR880 consists of large T cDNA (nucleotides 173 to 4658) ligated to a *Cla*I-*Bam*HI fragment containing the Harvey murine sarcoma virus long terminal repeat (LTR). The resulting fragment was inserted into pBR322 between the *Nar*I and *Bam*HI sites. pNT expresses large T residues 1 to 336 and was made by inserting the SV40 *Bcl*I-*Bam*HI fragment (nucleotides 2770 to 2533) encoding the early polyadenylation signal into LTR880. This fragment was inserted by using an *Eco*RI linker which also encodes two termination codons. pP6 encodes amino acids 1 to 259 and was made by changing codon 260 to a termination codon, using site-directed mutagenesis (48) of plasmid pPRB. pPRB expresses full-length large T under control of the Harvey murine sarcoma virus LTR (33b).

To produce a recombinant baculovirus which expresses full-length large T, pBLT was made by insertion of a 4,078-bp *Bam*HI-*Bcl*I fragment containing large T cDNA into *Bam*HI-cleaved baculovirus transfer vector pVL941 (26). A baculovirus transfer vector encoding residues 1 to 259, pBNT, was made by insertion of a PCR fragment from wild-type large T sequence, using a 3' primer that included a termination codon in place of bases encoding residue 260 into pBlueBac (Invitrogen).

Cells and viruses. NIH 3T3 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% calf serum. Cell lines expressing full-length, N-terminal 259-amino-acid or 336-amino-acid large T sequences were made by

* Corresponding author. Mailing address: Department of Biochemistry, Tufts University School of Medicine, 136 Harrison Ave., Boston, MA 02111. Phone: (617) 956-6876. Fax: (617) 956-6409.

cotransfection of large T constructs along with plasmid pRS-Vneo, using calcium phosphate precipitation (17). 3T3 cell lines expressing large T sequences were identified by immunofluorescence and were maintained in 200 μ g of Geneticin (Sigma) per ml.

Rat embryo fibroblasts were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. Primary embryo fibroblasts were prepared from 12- to 15-day Fischer rat embryos (Charles River Laboratories) (1). Cells at the first plating or after one passage were transfected by CaPO_4 precipitation with 6 μ g of test plasmid and 2 μ g of pRSVneo per 60-mm-diameter plate seeded with 3×10^5 cells. DNA precipitate was left on cells for 18 to 20 h. Cells were then replated at 30% confluence and after 24 h placed into selective medium containing 1 mg of Geneticin per ml. Selection medium was changed every 4 days. Individual colonies were picked after 10 to 14 days and placed into 96-well Linbro plates. Cells were passaged at a 1:5 dilution into 24-well plates as they approached confluence. Lines were considered immortal after reaching the fifth passage.

Spodoptera frugiperda Sf9 cells were maintained at 27°C in Grace's insect cell culture medium (GIBCO) supplemented with 10% fetal calf serum, Yeastolate, and lactalbumin hydrolysate. Recombinant viruses were made by transfection of pBLT and pBNT into Sf9 cells along with wild-type *Autographa californica* nuclear polyhedrosis virus DNA by CaPO_4 precipitation followed by plaque purification. Infected Sf9 cells were used at approximately 40 h postinfection.

Antisera. A rabbit antiserum prepared against polyomavirus small T (30) was used for immunoprecipitation. Monoclonal antibody PN116 recognizes an epitope in the first 79 amino acids of large T. Rabbit antiserum against a maltose-binding protein fusion containing large T residues 269 to 785 was used to identify the C-terminal domain. The C-terminal domain constructs and antibodies will be described elsewhere (16a).

Protein analysis. Procedures for T-antigen extraction and immunoprecipitation have been published elsewhere (34, 36). [^{35}S]methionine labelling was carried out in Hanks' salts (Flow Laboratories) and [^{32}P] labelling was carried out in phosphate-free medium for 2 h. T antigens were usually extracted in TEB (pH 9) (0.137 M NaCl, 0.01 M Tris-HCl [pH 9.0], 0.001 M MgCl_2 , 0.001 M CaCl_2 , 10% [vol/vol] glycerol, 1% [vol/vol] Nonidet P-40) for 20 min on ice and then subjected to clearing in a microcentrifuge for 10 min. T antigens were immunoprecipitated by using antiserum and protein A-Sepharose (Pharmacia) or monoclonal antibodies and anti-mouse immunoglobulin G-agarose (Sigma). Immune complexes were washed twice with phosphate-buffered saline, twice with 0.5 M LiCl-0.1 M Tris-HCl (pH 8.0), and once with deionized water. A second immunoprecipitation was sometimes performed. Primary precipitates were boiled for 2 min in 0.4% sodium dodecyl sulfate (SDS)-50 mM triethanolamine-Cl (pH 7.4)-100 mM NaCl-2 mM EDTA (pH 7.4)-2 mM 2-mercaptoethanol. Iodoacetamide to 10 mM and 1/4 volume of 10% Triton X-100 were added to the supernatant. After immunoprecipitation, the samples were analyzed on discontinuous buffer SDS-gels (23).

Procedures for mapping of large T with hydroxylamine or *Staphylococcus aureus* V8 protease have been described previously (4, 7). Hydroxylamine cleavage was performed by first equilibrating the beads in 0.2 M K_2CO_3 (pH 9.0) and then incubating them in 0.2 M K_2CO_3 (pH 9.0) containing 2 M hydroxylamine (Kodak) for 90 min at 45°C. V8 digestion was carried out by using 100 μ g of enzyme per ml. For enzymatic digestion of baculovirus-expressed large T, infected cells were extracted in TEB (pH 9). Aliquots were digested with V8 or

chymotrypsin for 30 min at 4 or 37°C and then boiled in SDS sample buffer.

For gradient analysis, extracts (20 mM N-2-hydroxyethyl-piperazine-N'-2-ethanesulfonic acid [HEPES; pH 7.9], 150 mM NaCl, 5% glycerol, 1% Nonidet P-40) of 3T3 cells expressing pP6 were layered on a 5 to 20% sucrose gradient and run for 5 h at 55,000 rpm. Fractions were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE).

For immunoblotting, proteins were transferred to nitrocellulose by electroblotting (49). In the case of full-length T and N-terminal domain blotting, monoclonal antibody PN116 was used at a 1:50 dilution, followed by sheep anti-mouse immunoglobulin G conjugated to horseradish peroxidase. For C-terminal domain detection, a 1:10,000 dilution of serum was used, followed by goat anti-rabbit immunoglobulin G conjugated to horseradish peroxidase. The enhanced chemiluminescence method (Amersham) was used for detection. Radioactivity and chemiluminescence were detected with XAR5 film (Kodak).

Binding of T antigens to pRb and p107. Bacterially expressed pRb- and p107-glutathione S-transferase (GST) fusion proteins were bound to glutathione-agarose beads as described by Kaelin et al. (21). Plasmids used were pGEX-2T, its derivatives pGT-RB(379-928) and pGT-RB(379-928, 706 C>F), p107 (252-816), and p107 (252-816, 713 C>F) (12). TEB (pH 9) extracts of 3T3 cell lines expressing different T constructs were bound to beads for 1 h at 4°C. Complexes were then washed four times with NETN (20 mM Tris [pH 8], 100 mM NaCl, 1 mM EDTA, 0.5% Nonidet P-40) and analyzed by Western immunoblot.

RESULTS

The N-terminal (336-amino-acid) construct used to establish function yields shortened protein comprising approximately 270 amino acids. Our aim was to define biochemically the residues of the N-terminal immortalizing domain inferred from early genetic experiments. That work on biological function used a large T fragment defined by a convenient *Eco*RI site. Our characterization started with this fragment as well. We prepared a polyomavirus cDNA which encoded this sequence followed by an added termination codon and the SV40 early polyadenylation signal. This cDNA was then introduced into a vector containing the Harvey murine sarcoma virus LTR. The resulting construct would encode the first 336 amino acids of large T. The construct, pNT, was used to establish a stable line in NIH 3T3 cells. This cell line was labeled with [^{35}S]methionine and analyzed for large T sequences by immunoprecipitation (Fig. 1). A minor band migrating at a size consistent with the full-length 336-residue protein (arrow) was observed. However, the major band observed corresponded to a smaller polypeptide with an apparent molecular mass of 38 kDa. The sizes appear larger than those calculated because large T is known to behave anomalously on polyacrylamide gels (36) and because the high level of phosphorylation in this region affects gel mobility (5). From this gel, it appeared that proteolysis of the 336-amino-acid fragment was occurring.

To determine which end of the protein was being truncated, proteolytic mapping was performed. Hydroxylamine cleavage of full-length large T releases an N-terminal 210-amino-acid fragment (4). An identical fragment was produced when the major protein from the pNT immunoprecipitate was digested (not shown). Therefore, the truncation occurred at the C terminus, leaving the N terminus intact. To obtain a more precise size estimate, the 38-kDa protein was compared with known N-terminal formic acid fragments of full-length large T

FIG. 1. Expression of the N-terminal 336-amino-acid and full-length large T proteins. NIH 3T3 cells expressing protein from LTR constructs encoding large T amino acids 1 to 336 (lane 1) or full-length large T (lane 2) were labeled with [^{35} S]methionine. Labeled proteins were immunoprecipitated twice and run on a 10% polyacrylamide gel. The 336-amino-acid protein is indicated by the arrow. The major ~270-amino-acid polypeptide as well as full-length large T are indicated by arrowheads. Positions of prestained molecular weight markers are indicated at the right in kilodaltons.

(4). It migrated on gels to a position intermediate between large T peptides of 255 and 285 amino acids (not shown). This finding suggested that the 336-residue protein was being proteolytically processed into a domain of approximately 270 amino acids.

A similar N-terminal fragment, as well as a complementary C-terminal fragment, can be derived from proteolysis of intact large T. Protein domains can often be separated by mild proteolysis of susceptible hinge regions (51). One explanation for the results with the *EcoRI* fragment is that the 336-amino-acid polypeptide includes one domain, as well as additional sequence that is insufficient to fold into a tight, protease-resistant structure. If this is the case, then intact large T should be cleaved by exogenous proteases to yield fragments similar to the 270-amino-acid polypeptide seen with pNT.

Extract from Sf9 cells infected with baculovirus expressing full-length large T was treated with *S. aureus* V8 protease. The resulting extract was analyzed by SDS-PAGE. Figure 2A shows a blot of such a gel probed with a monoclonal antibody that recognizes large T sequences in the N terminus. A single major band with an apparent molecular mass of 36 kDa was observed (lane 4). By contrast, probing an identical blot with antibody against C-terminal sequences (Fig. 2B, lane 4) revealed a major fragment of approximately 60 kDa. This was not detected by the N-terminal antibody and represented the C-terminal remainder of the protein. The simplest interpretation of this result is that the molecule has been split into two pieces. Treatment of large T extract with chymotrypsin also suggested that large T was sensitive to proteolysis in the same region as found with V8. Lane 5 of Fig. 2B shows that chymotryptic cleavage yields a C-terminal fragment similar to that obtained with V8. An N-terminal fragment slightly larger than that obtained with V8 was also seen in the chymotryptic digest when it was probed with the N-terminal antibody (Fig. 2A, lane 5). This fragment might be as large as 280 amino acids.

The N-terminal domain does not associate strongly with the C-terminal domain or with itself. The possibility of interactions between the N- and C-terminal domains of large T was investigated in two ways. First, Sf9 cells were coinfecting with viruses encoding large T residues 1 to 259 and 264 to 785. Neither immunoprecipitation nor sucrose gradient analysis

FIG. 2. Large T protein cleaved by proteases and probed with specific antibodies. (A and B) Sf9 cells infected with a baculovirus expressing full-length large T were extracted with T-antigen extraction buffer and subjected to proteolysis. The resulting samples were run on a 10% polyacrylamide gel and transferred to nitrocellulose. (A) A blot probed with a monoclonal antibody to the N terminus of large T; (B) an identical blot probed with rabbit antiserum directed against the C terminus of large T. Lane 1 contains protein from Sf9 cells infected with a virus encoding amino acids 1 to 259 of large T. Full-length large T was treated with no protease (lane 2), no protease with 30-min incubation at 37°C (lane 3), 1 μ g of V8 protease (lane 4), or 0.1 μ g of chymotrypsin (lane 5). (C) Cleavage of large T by V8 releases two independent fragments. Sf9 extract with full-length large T was treated with V8 protease and then immunoprecipitated with an antibody which recognizes only the C terminus of large T. Total extract before immunoprecipitation (lane 1) was compared with a washed immunoprecipitate (lane 2). After transfer, the blot was probed with monoclonal antibody PN116, which recognizes the N terminus of large T. Sizes are indicated in kilodaltons.

suggested substantial association of the two domains. A concern with these experiments is that the two domains needed to come together in solution. A similar experiment was therefore carried out with proteolyzed large T. We prepared extracts in which large T was cleaved with V8 as described above and then used a C-terminal specific antibody in an immunoprecipitation. Both total extract and the immunoprecipitate were blotted and probed with antibody which recognizes N-terminal determinants (Fig. 2C). The immunoprecipitate (lane 2) was quite depleted of the N-terminal fragment compared with total extract (lane 1). In other words, there was not an association between the N-terminal domain and the rest of the molecule strong enough to permit its efficient immunoprecipitation.

Large T is known to self-associate (50). Evidence to be presented elsewhere indicates that sequences in the C terminus are required for self-association (33b). To determine whether the N-terminal domain alone can oligomerize, extracts from mouse cells expressing the N-terminal 259-amino-acid construct were run on sucrose gradients. Western blots of fractions from this gradient (Fig. 3) demonstrated that the 259-amino-acid large T fragment sedimented close to a 29-kDa carbonic anhydrase marker and distinctly more slowly than a 67-kDa bovine serum albumin (BSA) marker. This behavior implies that the protein was present as a monomer in cells.

The N-terminal domain is highly phosphorylated. Large T is a highly phosphorylated protein (5, 20, 34–36). The phosphorylation occurs in two principal regions around residues 70 to 187 and residues 255 to 285 (4). Isoelectric focusing indicates a lower stoichiometry of phosphorylation for the fragment when expressed alone than for the same region in the full-length molecule (not shown). Where most large T molecules have three to five phosphates in their N-terminal domains, the isolated domains generally have one to two.

FIG. 3. Sucrose gradient of the N-terminal domain. Extract from P6 cells expressing amino acids 1 to 259 was sedimented on a 5 to 20% sucrose gradient for 5 h at 55,000 rpm. Fractions were analyzed by SDS-PAGE. The gel was transferred to nitrocellulose and probed with a mouse antibody which can recognize the first 79 amino acids of large T. The bottom of the gradient is at the left. Positions in the gradient of the 29-kDa carbonic anhydrase (CA) marker as well as the 67-kDa BSA marker are indicated.

V8 mapping was used to compare phosphopeptides from full-length and N-terminal domain proteins (Fig. 4A). Phosphopeptide maps of large T are complex, and the number of sites and their exact positions have not yet been fully determined. Previous work has established that V8 phosphopeptides 1 to 4, 8, and 9 come from the N-terminal domain (4). These peptides were present in the fragment, although the amount of phosphopeptide 3 was reduced. Since V8 achieves only a partial resolution of the phosphorylation sites, it remains possible that some specific sites within these peptides are affected. Phosphopeptides 5 and 7 contain the more C-terminal phosphorylation cluster around residues 255 to 285. These were missing from the domain construct. A spot was observed near the position of peptide 7. Analysis of the hydroxylamine fragment from amino acids 1 to 210 (Fig. 4B) showed that this new phosphopeptide arose from the N-terminal region (arrowhead). It therefore represents a peptide for which phosphorylation was either only allowed or enhanced when the domain was separated from the rest of the molecule.

Expression of the N-terminal domain (amino acids 1 to 259)

FIG. 4. V8 analysis of phosphorylated large T and N-terminal domain. (A) Full-length large T (LT)- or pNT (NT)-expressing cells were labeled with $^{32}\text{PO}_4$ for 2 h. After immunoprecipitation, the proteins were resolved on cylindrical SDS-polyacrylamide gels (34). In the second dimension, the cylinders were overlaid with 2 ml of 100- $\mu\text{g}/\text{ml}$ V8 protease and the peptides were resolved on 12.5% polyacrylamide gels. Positions of individual V8 peptides are indicated by arrowheads. (B) Immunoprecipitates were treated with hydroxylamine as described previously (4). After treatment, the N-terminal 210-amino-acid fragments were resolved on cylindrical gels and digested with V8 protease as for panel A.

TABLE 1. Immortalization assay

Plasmid	Large T sequence	Immortalization efficiency ^a (%)
pIBI24	None	0/87 (0), 0/62 (0)
pPRB	1-785	11/93 (11), 15/82 (18)
pP6	1-259	8/96 (8), 8/81 (10)

^a Rat embryo fibroblasts were transfected with the indicated plasmids along with pRSVneo. G418-resistant colonies were picked 3 to 4 weeks after transfection and passaged at a 1:5 dilution each time they reached confluence. Cell lines were considered immortal after five passages.

is sufficient for immortalization. Earlier experiments showed that the *Eco*RI construct was sufficient for immortalization. Since both the full-length 336-amino-acid as well as the truncated ~270-amino-acid forms would be expected to be present in cells, it was not possible to determine whether the smaller protein was capable of immortalizing. Since the most N-terminal residue known to be involved in DNA binding is residue 286 (47), the larger, but not the smaller, polypeptide might have DNA-binding capacity.

Primary Fischer rat embryo fibroblasts were transfected with large T constructs. Table 1 shows that the 259-amino-acid construct is sufficient for immortalization. The surviving colonies were expanded and tested for the presence of T sequences by Western analysis. All cell lines expressed proteins of the expected size (not shown). This seems to rule out a requirement for DNA binding in immortalization as well as reducing the size of the immortalizing domain.

The N-terminal domain binds both pRb and p107. Within the N-terminal domain of large T are sequences known to be required for binding to the product of the Rb gene as well as to p107 (11). Residues in large T required for binding have been identified around amino acid 140. We tested the ability of the 259-amino-acid protein to bind to pRb and p107, using GST fusion proteins. GST was fused to wild-type pRb or p107 as well as to mutant forms containing point mutations which are known to ablate binding to large T (12, 21). Glutathione-agarose beads with pRb- or p107-GST fusion proteins bound were incubated with cell extracts containing large T. Both the full-length and 259-amino-acid proteins bound to the wild-type proteins but not to the mutants (Fig. 5). A comparison of the relative intensities of the bands from cell extracts and bound protein suggests that the 259-amino-acid polypeptide bound with lower efficiency than full-length large T to both cellular proteins. Large T bound approximately four times better than the N terminus alone to pRb and 10 times better to p107.

DISCUSSION

It is known that polyomavirus large T antigen is a multifunctional protein. Although domains have been identified at the genetic level, much less is known about the protein structure and how it relates to function. Treatment with proteases, as well as extended incubation of the extract at 37°C (not shown), yielded N-terminal fragments of similar sizes. Fragments representing the N-terminal ~260 to 280 amino acids and C-terminal ~500 amino acids are released by such treatment. It therefore appears that the region from residues 260 to 280 reflects a protease-sensitive hinge. This separation of the protein into two domains fits with the genetic data that place elements directly involved in DNA and ATP binding, as well as other functions needed for replication, from amino acids 284 to the end of molecule and elements required for immortalization in the N terminus. Previous constructs used for immor-

FIG. 5. Binding of large T to pRb and p107. Full-length large T (LT; lanes 1, 3, 5, and 7) as well as the N-terminal 259-residue protein (NT; lanes 2, 4, 6, and 8) from NIH 3T3 cells were bound to pRb-GST (A) or p107-GST (B) fusion proteins bound to glutathione-agarose beads. Lanes 1 and 2 show starting material and represent approximately one-fourth the amount of protein used in binding experiments. Fusion proteins were wild type (lanes 3 and 4), large T-binding mutants (lanes 5 and 6), and pGEX parent vector (lanes 7 and 8).

talization contained some of the sequences involved in DNA binding. Our results showed that independent expression of the first 259 amino acids of large T, a region which includes none of these sequences, was sufficient for immortalization.

The N-terminal domain appears to be relatively independent. Pulse-chase experiments showed little difference in stability between the full-length large T and the 259-amino-acid fragment (not shown). It does not self-associate significantly, even when expressed in the baculovirus system. It also does not appear to have strong contact with the rest of the molecule. Coexpression of the two separate domains in Sf9 cells shows very little complex formation. In addition, cleavage within the hinge region which tethers the two domains largely severs any association between them. However, these results do not rule out weaker interactions. The extent of direct contact between the two domains in the intact molecule remains to be determined. One interpretation of the enhanced phosphopeptide seen in the hydroxylamine fragment of pNT is that a surface that is unavailable in the intact molecule is exposed.

The N-terminal domain is clearly sufficient to bind both pRb and p107. This is expected since the ability of large T to bind pRb has been shown to be required for immortalization (25). There is, however, a quantitative decrease in binding of the independent domain compared with full-length large T. This could be related to the reduced stoichiometry of phosphorylation of the N-terminal protein. Mapping to be presented elsewhere (19a) shows that association of large T with pRb is favored when particular large T phosphopeptides are observed. Alternatively, there may be additional binding elements in the C-terminal domain of large T that could contribute to its affinity for pRb or p107. One example of this type of behavior at the protein level is seen with polymerase α , which binds to two separate regions of SV40 large T (13).

Finally, it is interesting to compare these results with those for SV40. SV40 large T also shows an N-terminal domain (40, 42, 52). Cleavage with trypsin, for example, cuts between residues 130 and 131 (40). *dl1137*, which encodes 121 N-terminal amino acids, is capable of transforming C3H 10T1/2

cells (45). One group has reported that a slightly longer construct encoding 147 amino acids transforms primary rat embryo fibroblasts or REF52 cells (43, 44), although others (45, 53) have found a requirement for additional large T sequences. The SV40 fragment could be seen as similar to that of polyomavirus. Polyomavirus has two nuclear localization signals (33); the second one, starting at residue 280, is homologous to the SV40 signal around residue 128. In each protein, therefore, an N-terminal domain that extends to this nuclear localization signal can be derived. There is also a cluster of phosphorylation sites near or just after the end of the domain: positions 255 to 285 for polyomavirus (4) and positions 106, 112, 120, 123, and 124 for SV40 (38). However, there are also obvious differences. The polyomavirus fragment is much larger. The additional sequence is apparently important since mutations there, such as in *dl8* (18) and *dl1015* (27), affect replication. It is also much more highly phosphorylated; in SV40, only serines 106 and 112 are modified in the N-terminal domain. Finally, there appears to be more interaction between the domains in the case of SV40. Schwyzer et al. (40) showed originally that the 130-residue tryptic fragment remained with the immune complex after cleavage, and Scheidtmann and colleagues (37) inferred interactions from the effect of mutation of residue 677 on the phosphorylation of residue 112.

ACKNOWLEDGMENTS

We thank Bill Kaelin and Mark Ewen for gifts of GST fusion constructs. We gratefully acknowledge Jennifer Feldmann and Jim DeCaprio for help with generation of monoclonal antibodies. We thank Tom Roberts and Josh Niclas for critical reading of the manuscript.

This work was supported by NIH grant PO-CA50661.

REFERENCES

- Asselin, C., and M. Bastin. 1985. Sequences from polyomavirus and simian virus 40 large T genes capable of immortalizing primary rat embryo fibroblasts. *J. Virol.* **56**:958-968.
- Asselin, C., J. Vass-Marengo, and M. Bastin. 1986. Mutation in the polyomavirus genome that activates the properties of large T associated with neoplastic transformation. *J. Virol.* **57**:165-172.
- Basilico, C., D. Zouzas, G. Della-Valle, S. Gattoni, V. Colantoni, R. Fenton, and L. Dailey. 1980. Integration and excision of polyoma virus genomes. *Cold Spring Harbor Symp. Quant. Biol.* **44**:611-620.
- Bockus, B. J., and B. Schaffhausen. 1987. Localization of the phosphorylations of polyomavirus large T antigen. *J. Virol.* **61**:1155-1163.
- Bockus, B. J., and B. Schaffhausen. 1987. Phosphorylation of polyomavirus large T antigen: effects of viral mutations and cell growth state. *J. Virol.* **61**:1147-1154.
- Bouchard, L., F. Mathieu, and M. Bastin. 1988. Polyoma large T can activate middle T expression by a hit-and-run mechanism. *Oncogene* **2**:379-386.
- Carmichael, G., B. S. Schaffhausen, G. Mandel, T. J. Liang, and T. L. Benjamin. 1984. Transformation by polyoma virus is drastically reduced by substitution of phenylalanine for tyrosine at residue 315 of middle-sized tumor antigen. *Proc. Natl. Acad. Sci. USA* **81**:679-683.
- Cogen, B. 1978. Virus-specific early RNA in 3T6 cells infected by a tsA mutant of polyoma virus. *Virology* **85**:222-230.
- Cowie, A., and R. Kamen. 1984. Multiple binding sites for polyomavirus large T antigen within regulatory sequences of polyomavirus DNA. *J. Virol.* **52**:750-760.
- Della Valle, G., R. G. Fenton, and C. Basilico. 1981. Polyoma large T antigen regulates the integration of viral DNA sequences into the genome of transformed cells. *Cell* **23**:347-355.
- Dyson, N., R. Bernards, S. H. Friend, L. R. Gooding, J. A. Hassell, E. O. Major, J. M. Pipas, T. Vandyke, and E. Harlow. 1990. Large T antigens of many polyomaviruses are able to form complexes

- with the retinoblastoma protein. *J. Virol.* **64**:1353–1356.
12. Ewen, M. E., B. Faha, E. Harlow, and D. M. Livingston. 1992. Interaction of p107 with cyclin A independent of complex formation with viral oncoproteins. *Science* **255**:85–87.
  13. Fanning, E., and R. Knippers. 1992. Structure and function of simian virus 40 large tumor antigen. *Annu. Rev. Biochem.* **61**:55–85.
  14. Farmerie, W. G., and W. R. Folk. 1984. Regulation of polyomavirus transcription by large tumor antigen. *Proc. Natl. Acad. Sci. USA* **81**:6919–6923.
  15. Fluck, M. M., and T. L. Benjamin. 1979. Comparison of two early gene functions essential for transformation in polyoma virus and SV40. *Virology* **96**:205–228.
  16. Francke, B., and W. Eckhart. 1973. Polyoma gene function required for viral DNA synthesis. *Virology* **55**:127–135.
  - 16a. Gjoerup, O. V., et al. Unpublished data.
  17. Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**:1044–1051.
  18. Griffin, B. E., and C. Maddock. 1979. New classes of viable deletion mutants in the early region of polyoma virus. *J. Virol.* **31**:645–656.
  19. Guizani, I., P. Clertant, and F. Cuzin. 1987. Biochemical properties associated with the immortalizing domain of the large T protein of polyoma virus. *Biochem. Biophys. Res. Commun.* **144**:973–979.
  - 19a. Holman, P. S., et al. Unpublished data.
  20. Hassauer, M., K. H. Scheidtmann, and G. Walter. 1986. Mapping of phosphorylation sites in polyomavirus large T antigen. *J. Virol.* **58**:805–816.
  21. Kaelin, W. G. J., D. C. Pallas, J. A. DeCaprio, F. J. Kaye, and D. M. Livingston. 1991. Identification of cellular proteins that can interact specifically with the T/E1A-binding region of the retinoblastoma gene product. *Cell* **64**:521–532.
  22. Kingston, R. E., A. Cowie, R. I. Morimoto, and K. A. Gwinn. 1986. Binding of polyomavirus large T antigen to the human hsp70 promoter is not required for *trans* activation. *Mol. Cell. Biol.* **6**:3180–3190.
  23. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
  24. Land, H., L. F. Parada, and R. A. Weinberg. 1983. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. *Nature (London)* **304**:596–602.
  25. Larose, A., N. Dyson, M. Sullivan, E. Harlow, and M. Bastin. 1991. Polyomavirus large T mutants affected in retinoblastoma protein binding are defective in immortalization. *J. Virol.* **65**:2308–2313.
  26. Luckow, V. A., and M. D. Summers. 1989. High level expression of non-fused foreign genes with A. californica nuclear polyhedrosis virus expression vectors. *Virology* **170**:31–39.
  27. Magnusson, G., M. G. Nilsson, S. M. Dilworth, and N. Smolar. 1981. Characterization of polyoma mutants with altered middle and large T antigens. *J. Virol.* **39**:673–683.
  28. Maione, R., G. M. Fimia, and P. Amati. 1992. Inhibition of *in vitro* myogenic differentiation by a polyomavirus early function. *Oncogene* **7**:85–93.
  29. Mougneau, E., L. Leimieux, M. Rassoulzadegan, and F. Cuzin. 1984. Biological activities of v-myc and rearranged c-myc oncogenes in rat fibroblast cells in culture. *Proc. Natl. Acad. Sci. USA* **81**:5758–5762.
  30. Pallas, D. C., C. Schley, M. Mahoney, E. Harlow, B. S. Schaffhausen, and T. M. Roberts. 1986. Polyomavirus small t antigen: overproduction in bacteria, purification, and utilization for monoclonal and polyclonal antibody production. *J. Virol.* **60**:1075–1084.
  31. Rassoulzadegan, M., Z. Naghashfar, A. Cowie, A. Carr, M. Grisoni, R. Kamen, and F. Cuzin. 1983. Expression of the large T protein of polyoma virus promotes the establishment in culture of “normal” rodent fibroblast cell lines. *Proc. Natl. Acad. Sci. USA* **80**:4354–4358.
  32. Rassoulzadegan, M., R. Seif, and F. Cuzin. 1978. Conditions leading to the establishment of the N (a gene dependent) and A (a gene independent) transformed states after polyoma virus infection of rat fibroblasts. *J. Virol.* **28**:421–426.
  33. Richardson, W. D., B. L. Roberts, and A. E. Smith. 1986. Nuclear location signals in polyoma virus large-T. *Cell* **44**:77–85.
  - 33a. Rose, P., et al. Unpublished data.
  - 33b. Rose, P., and B. S. Schaffhausen. Unpublished data.
  34. Schaffhausen, B., and T. L. Benjamin. 1981. Comparison of phosphorylation of two polyoma virus middle T antigens *in vivo* and *in vitro*. *J. Virol.* **40**:184–196.
  35. Schaffhausen, B. S., and T. L. Benjamin. 1979. Phosphorylation of polyoma T antigens. *Cell* **18**:935–946.
  36. Schaffhausen, B. S., J. E. Silver, and T. L. Benjamin. 1978. Tumor antigen(s) in cells productively infected by wild-type polyoma virus and mutant NG18. *Proc. Natl. Acad. Sci. USA* **75**:79–83.
  37. Scheidtmann, K. H., M. Buck, J. Schneider, D. Kalderon, E. Fanning, and A. E. Smith. 1991. Biochemical characterization of phosphorylation site mutants of simian virus 40 large T antigen: evidence for interaction between amino- and carboxy-terminal domains. *J. Virol.* **65**:1479–1490.
  38. Scheidtmann, K. H., B. Echle, and G. Walter. 1982. Simian virus 40 large T antigen is phosphorylated at multiple sites clustered in two separate regions. *J. Virol.* **44**:116–133.
  39. Scheller, A., and C. Prives. 1985. Simian virus 40 and polyomavirus large tumor antigens have different requirements for high-affinity sequence-specific DNA binding. *J. Virol.* **54**:532–545.
  40. Schwyzer, M., R. Weil, G. Frank, and H. Zuber. 1980. Amino acid sequence analysis of fragments generated by partial proteolysis from large simian virus 40 tumor antigen. *J. Biol. Chem.* **255**:5627–5634.
  41. Seif, R., and F. Cuzin. 1977. Temperature-sensitive growth regulation in one type of transformed rat cells induced by the ts-a mutant of polyoma virus. *J. Virol.* **24**:721–728.
  42. Simmons, D. T. 1988. Geometry of the simian virus 40 large tumor antigen-DNA complex as probed by protease digestion. *Proc. Natl. Acad. Sci. USA* **85**:2086–2090.
  43. Sompayrac, L., and K. J. Danna. 1991. The amino-terminal 147 amino acids of SV40 large T antigen transform secondary rat embryo fibroblasts. *Virology* **181**:412–415.
  44. Sompayrac, L., and K. J. Danna. 1992. An amino-terminal fragment of SV40 T antigen transforms REF52 cells. *Virology* **191**:439–442.
  45. Srinivasan, A., K. W. Peden, and J. M. Pipas. 1989. The large tumor antigen of simian virus 40 encodes at least two distinct transforming functions. *J. Virol.* **63**:5459–5463.
  46. Sunstrom, N.-A., N. H. Acheson, and J. A. Hassell. 1991. Determination of the origin-specific DNA-binding domain of polyomavirus large T antigen. *J. Virol.* **65**:6998–7003.
  47. Tang, W. J., and W. R. Folk. 1989. Asp-286→Asn-286 in polyomavirus large T antigen relaxes the specificity of binding to the polyomavirus origin. *J. Virol.* **63**:242–249.
  48. Taylor, J. W., J. Ott, and F. Eckstein. 1985. The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA. *Nucleic Acids Res.* **13**:8764–8785.
  49. Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**:4350–4354.
  50. Wang, E. H., and C. Prives. 1991. ATP induces the assembly of polyoma large tumor antigen into hexamers. *Virology* **184**:399–403.
  51. Wetlaufer, D. B. 1973. Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci. USA* **70**:697–701.
  52. Wun-Kim, K., and D. T. Simmons. 1990. Mapping of helicase and helicase substrate-binding domains on simian virus 40 large T antigen. *J. Virol.* **64**:2014–2020.
  53. Zhu, J., P. W. Rice, L. Gorsch, M. Abate, and C. N. Cole. 1992. Transformation of a continuous rat embryo fibroblast cell line requires three separate domains of simian virus 40 large T antigen. *J. Virol.* **66**:2780–2791.