

# Do Variant SV40 Sequences Have a Role in the Maintenance of the Oncogenic Transformed Phenotype?

ROBERT POLLACK  
CAROL PRIVES  
JAMES MANLEY  
*Department of Biological Sciences  
Columbia University  
New York, New York*

## INTRODUCTION

SV40 is a small virus, but its approximately 5200 base pairs (bp) contains enough information either to productively grow in a permissive cell or to convert a normal cell into a transformed one (1). The conversion is usually genetically stable enough to generate clonable lines of transformed cells. Some of these lines lose so much of their normal growth control that they acquire the capacity to grow as tumors (2). Since the total DNA sequence of SV40 is known (1), it is somewhat surprising that the mechanism of action of SV40 in maintaining the various transformed states *in vitro* is not at all well understood. Also, and probably not coincidentally, precisely which SV40-encoded molecules are necessary for maintenance of any of these states also is an unsettled issue (1).

SV40 has been utilized as a model system both to study eukaryotic gene expression and regulation and to further understand viral-mediated neoplastic transformation. In this respect, the virus is ideal because it is small, coding for only six known polypeptides (VP-1, VP-2, VP-3, t and T antigens, and the agno protein); it has been entirely sequenced (1) and expresses its genes in a manner similar to eukaryotic chromosomal gene expression (Fig. 1). There are viral and cellular parallels for DNA replication, chromatin structure, RNA transcription,

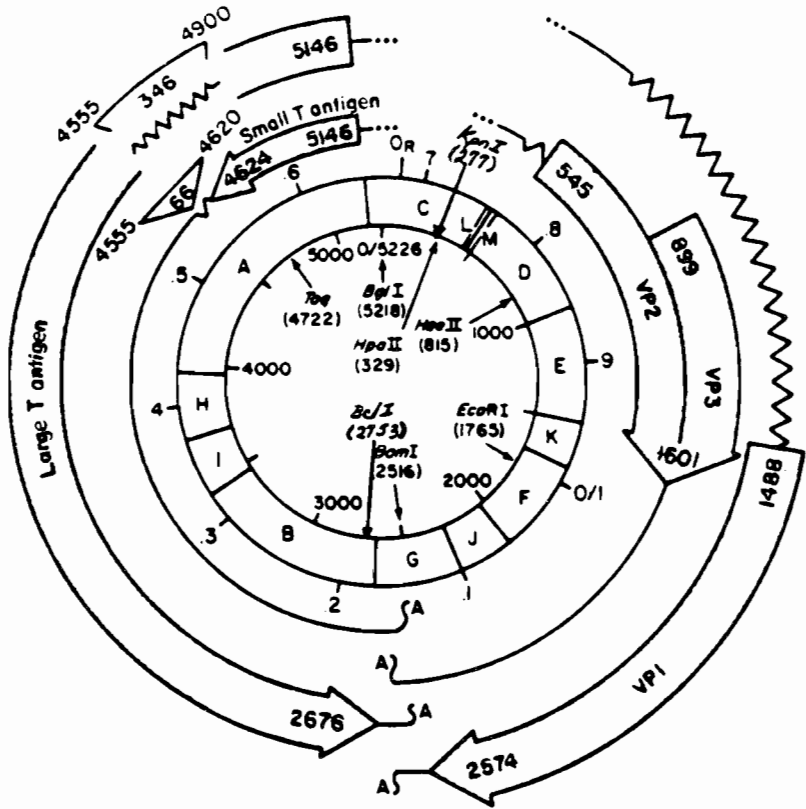


Fig. 1. Schematic representation of SV40 genome indicating locations of some principal features. The genome is represented as a circle with the origin of replication at the top. Outer boxed areas are *Hind*II and III fragments. Numbers outside the circle refer to locations in fractional genome length relative to the single *Eco*RI change site within SV40 DNA. Residue numbers and labels refer to SV40 sequence numbers (see refs, Appendix A). Sequence numbers for codons include termination codons. Shaded area with arrowheads indicate coding portions of mRNAs, with arrowheads pointing in 5'-3' direction. Introns are indicated by zig-zag lines. (This figure was taken, with minor modification, from ref. 1.)

modification, splicing, and translation (1). Depending on the host cell, SV40 induces either a lytic or nonlytic response. The lytic response makes more complete use of the viral genetic information. The kidney cells of the African green monkey (and derived cell lines) are the permissive host for the lytic response to SV40. In these cells, the infectious cycle can be divided into early and late phases. In the early phase E strand-specific RNA is expressed, presumably on a single

unspliced transcript which is then subjected to two alternate splicing pathways. This generates mRNAs for the large T and small t antigen. After T antigen is synthesized, there is a burst of viral and cellular DNA synthesis, the viral DNA being replicated on a circular minichromosome. The role of the large T antigen in these processes has been demonstrated by genetic and biochemical analyses. The T antigen binds with high affinity to sequences at the viral origin of replication *ori*. The promoter for early strand RNA synthesis is within the *ori* region, and T antigen, in binding to this region, brings about considerable down regulation of the early mRNAs. After viral DNA replication begins, L strand specific transcription initiates, and after the later 19 S and 16 S viral mRNAs are processed, their translation producing the capsid proteins. VP-1 (abundant) and VP-2 + VP-3 (minor) are synthesized. Considerably more L strand than E strand mRNA, and mRNA translation products are synthesized during the late phase. The simplest explanation for this is that E strand mRNAs are autogeneously regulated by T antigen, and L strand mRNAs levels are increased due to an increase in DNA template number brought about by the combination of T antigen function, but this has not yet been proved. Electron microscopic and restriction enzyme analysis have demonstrated that although the SV40 minichromosome resembles cellular chromatin in the classical histone-DNA "bead" structure, there is a nucleosome-free gap in the noncoding regulatory sequences which include the 300 bp *ori* region and sequences extending in the late direction upward until approximately 0.72  $\mu\text{m}$ .

In nonpermissive cells, there is little or no viral DNA replication, presumably due to the lack of the viral DNA replication-specific factors present in AGMK cells. The remainder of the genome, the early region, is expressed in transformed cells.

## SV40 VIRAL GENES AND TRANSFORMATION

The SV40 early region encodes two well-characterized proteins, the large (94K) and small (17K) T antigens. Although the large T is found in SV40 transformed cell lines and although DNA encoding the early region of SV40 is necessary for transformation, the actual mechanism(s) by which SV40 T antigens maintained the transformed state remain poorly understood (1). The problem of mechanism is especially thorny because normal fibroblasts display distinct serum- and anchorage-dependent growth controls and because both of these controls can be independently disrupted by SV40 transformation (3).

**TABLE I**  
**Assays for Selecting and Characterizing SV40 Transformed Cells**

Assay	Change in transformed cell
Physiological	
Serum	Grows in (<1%) serum
Anchorage	Grows in agar as a physical colony
Focus	Grows to high cell density
Nonselective	All possible stable changes
Tumorigenicity	Grows in animal directly as a tumor
Chemical	
Insulin	Grows in defined medium lacking insulin
Calcium	Grows in chelex-serum plus calcium-free medium

In molecular terms, transformation must be the consequence of an interaction between SV40-encoded molecules and molecules of the transformed cell. Further complexity is therefore introduced into the role of SV40 in transformation by the choices of which cell to transform and which selective assay to use in generating transformed clones (Table I).

Given all these possible variations on the theme of the capability of SV40 to transform and to maintain transformation, it is perhaps inevitable that the virus, although small, makes a great deal of use of each base pair. SV40 encodes a number of mRNAs and proteins (Fig. 1) even when integrated into the genome of the nonpermissive cell. One of these proteins, the large T antigen, is itself a molecule with many functions and domains (Table II). Thus, a major unanswered question at this time is whether the phenotypic events in transformation and lytic expression are linked to different domains of T antigen.

Simple genetics has not been able to complete this task, because only mutations in SV40 sequences needed for virus growth can be detected by classic selection techniques. A set of recent technical developments has made it possible now to map SV40 domains in a new and complete way. The construction of SV40 mutants using restriction enzymes and recombinant technology has allowed the analysis of SV40s which are totally defective for growth (e.g., origin minus) but which can transform cells quite well at least in certain assays. At the same time, the major transforming gene product of WT SV40, the large T antigen, has been purified and has been found to display many different activities in cell-free preparations. Many monoclonal antibodies to

**TABLE II**  
**Properties of Partially and Highly Purified**  
**SV40 T Antigens**

---

Activities
Binds to double- and single-stranded DNA
Binds to chromatin
Binds specifically to sites at the viral origin of replication
ATPase Activity [stimulated by poly(dT)]
Specific SV40 DNA Unwinding Activity
Binds to and stabilizes the 54K host protein
Exists as 5-7 S, 18 S and 22 S oligomers
Stimulates cell DNA synthesis (after microinjection)
Induces host tumor rejection response
Posttranslational modifications
N-terminal acetylated
Phosphorylated ( <i>ser</i> and <i>thr</i> ; not <i>tyr</i> )
ADP ribosylated

---

WT T antigen have been prepared; these recognize different specific regions of the T antigen protein (4).

SV40 sequences integrated in transformed cells have been mapped, cloned, and sequenced (5). These rescued sequences have also been used to transform cells, thus constructing a novel genetic cycle from eukaryotic host to plasmid vector and back. If specific regions of SV40 T carry out specific functions, then this cycle is a powerful way to apply the knowledge acquired from analysis of a new functional domain.

SV40 T antigen regulates both SV40 DNA replication and SV40 transcription. As SV40 DNA has become so malleable, its replication and transcription also have therefore become subject to experimental manipulation, for example, by directed methylation of specific restriction-enzyme sites or by deletion of sequences required for transcription (6).

## INITIAL HYPOTHESIS

The SV40 DNA sequence has a second open coding frame in the distal portion of the early transcript (7). If permitted, a splice into this frame would generate a new sort of protein, with a very hydrophobic C-terminus. This postulated hydrophobic T might be related to the 100K T we have observed in SV101 and its anchorage-independent

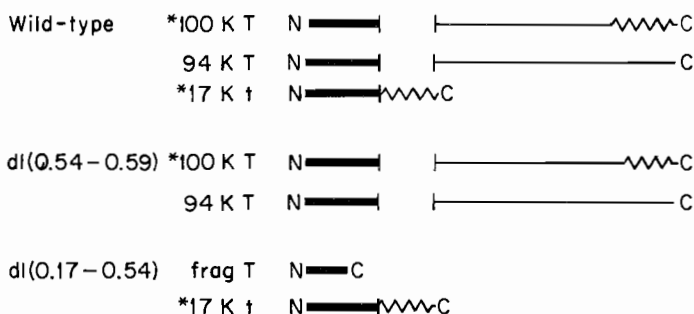


Fig. 2. Working model of SV40 gene-products in anchorage-transformation. Hypothesis: (1) To maintain anchorage independence, a cell must contain at least one molecule of the class N —hydrophobic—C, in order to bring the N-terminus common domain to its functional site. (2) To first efficiently initiate and then to maintain anchorage independence, a cell needs both a molecule of the class N —hydrophobic—C and a wild-type 94K T.

\* A molecule of the class: N —hydrophobic—C.

derivatives (8). How might such a T antigen be generated? So far, the search in wt-infected lytic cells for an mRNA with this splice has not been productive. However, although mutant viruses and virus sequences in transformed cells have not been examined, there is no need to necessarily impose a second splice on the transformed cells in order to generate a 100K T. Alternative mutational methods exist. These form the basis of our hypothesis. We hypothesize first that the 100K T antigen in SV101 is the product of a mutated SV40 within SV101 and that this mutation is a deletion (as small as a base pair) that brings the mRNA into the second reading frame for the C-terminal, resulting in hydrophobic C-terminus for large T (Fig. 2).

At first sight, a difficulty with this hypothesis appears to arise from the fact that extensive C-terminal deletion mutants can transform in the anchorage assay. However, small t connects also the common N-terminus to a hydrophobic C-terminus. We propose that both 17K and 100K have the same function in transformation. Recently, the laboratories of K. Danna (personal communication) and T. Shenck (Cambridge T.V. meeting, Abstract 1981) have used transfection of PBR-cloned short SV40 sequences to show that the stretch of SV40 DNA from the origin to 0.54 is sufficient to transform in the stringent anchorage assay, albeit with a lower limitation frequency ( $\sim 1\%$ ) compared to one full wild-type sequence. When proteins in such transformed cells are examined by immunoprecipitation, small t and a

fragment of large T are both detected, but the C-terminal of large T has been deleted.

A second set of data that must be included in any hypothesis of anchorage transformation comes from deletions in 0.54–0.59 map units. These remove the coding sequence for the C-terminal of small t. In some deletions a splice function as well is removed, eliminating even N-terminal t fragments. Anchorage-transformed mouse lines generated by the 54/59 mutant dl 884 have 100K T and 94K T (8).

We may now ask, what parts of the SV40 genome, if any, do all three kinds of anchorage-transformed cells share? Providing our initial hypothesis about the 100K T is correct, then the answer is: *All anchorage transformed lines have the N-terminal common domain of the early region connected to a hydrophobic C-terminal domain* (Fig. 2). For WT transformants, two molecular forms, 100K T and 17K T, meet this requirement. For dl 884 transformants, 100K T meets it, and for transformants made from infection with extensive C-terminal-T deletion mutants, the 17K small t meets it. If this hypothesis is correct, we may then say:

1. The 94K T may not be the molecular species responsible for anchorage transformation, because it lacks the hydrophobic C-terminus and because we have seen no mouse anchorage transformants without such a link of N-terminus to hydrophobic C-terminus.

2. The 94K T is necessary for *efficient initiation* of anchorage transformation. We hypothesize that this is so because it enables SV40 to replicate during initiation. Mutation to second frame in one copy of a tandem array would not cause total loss of 94K T function.

3. In species other than mouse, the hydrophobic C-terminus might be unnecessary due to different and more efficient transforming interactions of these host molecules (such as 54K) with their 94K T. This would explain why WT transformed rat cells, for instance, both seem to lack 100K T and to be very efficiently transformed by dl 54/59 mutants (1).

If our frameshift theory is correct, we would expect the 100K protein to have a new, hydrophobic C-terminal domain in addition to the amino acid sequence present in the 94K T antigen. If this is what we find, we can ask: Does such a domain cause the protein to localize in the plasma membrane, where it might operate like transforming proteins from RNA oncoviruses? We can test this hypothesis in several ways. If there is a convenient enzymatic site, the domain of interest could be cleaved off the protein and isolated on a polyacrylamide-SDS gel. Antisera, monoclonal or conventional, could be raised

against this domain alone. The 100K protein could then be traced in SV101 and its sublines by immunofluorescent second antibodies. Alternatively, the protein itself—or the new domain alone—could be microinjected into the nuclei of mouse embryo fibroblasts and followed by anti-T sera. Or the protein could be traced, with the same anti-sera and fluorescent second antibodies, in mouse embryo fibroblasts transformed with 100K T coding sequences.

Such a new domain could also be studied using recombinant DNA techniques. For instance, most of the sequences between the promoter and the new domain could be deleted from the pBR322 subclone, and the new construct used in an attempt to produce copies of the hydrophobic peptide. Such a construct, lacking an intron, might or might not be expressed in mouse cells. To test if the membrane localization is due to the new domain, another well-characterized, intron-containing coding sequence (such as the gene for  $\beta$ -globin) could be spliced in to replace the N-terminal T antigen sequence. It could then be observed if the new hybrid molecule accumulated in the plasma membrane.

Since the SV101 sublines which no longer grow in agar still produce 94K protein—though not 100K protein—it might be argued that a gene dosage—or domain dosage—effect is at work (8). It is possible that we will find, not a frameshift mutation producing a new domain, but rather a duplication of a domain already present in large T antigen. If so, we could use binding and protection studies to see if that domain binds either DNA or 54K nonviral T antigen with high affinity.

Cellular studies involving the function of the 100K protein as a whole would involve either microinjection or calcium-phosphate mediated transfer of cloned sequences into mouse embryo fibroblasts. It could then be determined if 100K protein without 94 or 17K protein can transform mouse cells in a quantitative manner. It would be especially intriguing to subculture lines transformed by injected 100K coding sequences to see if viral DNA rearrangements occur with high frequency.

#### REFERENCES

1. Tooze, J. (1981) "Molecular Biology of Tumor Viruses," 2nd ed., Chapters 1–4 and Appendix. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
2. Freedman, V., and Shin, S. (1974) *Cell* 3, 355.
3. Risser, R., and Pollack, R. (1974) *Virology* 59, 477.
4. Scheller, A., Covi, L., Barnet, B., and Prives, C. (1982) *Cell* 29, 375.

5. Botchan, M. *et al.* (1980) *Cell* **20**, 143; Gluzman, Y. (1981) *ibid.* **23**, 175.
6. Fradin, A., Manley, J., and Prives, C. (1982) *P.N.A.S.* **79**, 5142.
7. Mark, D., and Berg, P. (1979) *Cold Spring Harbor Symp. Quant. Biol.* **44**, 55.
8. Chen, S., Verderame, M., Lo, A., and Pollack, R. (1981) *Mol. Cell Biol.* **1**, 994.