JOURNAL OF VIROLOGY, June 1981, p. 958-967
0022-538X/81/060958-10$02.00/0

Vol. 38, No. 3

DNA Sequences of Polyoma Virus Early Deletion Mutants

NINA SMOLAR† AND BEVERLY E. GRIFFIN*

Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2, England

Received 26 November 1980/Accepted 4 March 1981

The DNA sequences of four "early" viable deletion mutants of polyoma virus have been determined. Two of these (*dl*-8 and *dl*-23) are mutants with deletions in the region of the genome that codes for parts of both large and middle T-antigens, and two (*dl*-6 and *dl*-28) are mutants with deletions around the viral origin of replication. The former mutants have altered transformation properties relative to wild-type virus, and *dl*-8 appears to be replication deficient (B. E. Griffin and C. Maddock, J. Virol. **31**:645-656, 1979). Sequences are discussed in terms of the altered phenotypes observed for the various mutants, the DNA structures and protein sequences that are affected by the deletions, and how these might affect the biological properties of the mutants.

Recent studies on polyoma virus have shown that cellular transformation by this virus does not require the entire viral genome (18, 21, 30). Thus, in transfection experiments with viral DNA, a fragment that contains only about half of the "early region" (the early region being defined as the part of the genome, about 50%, transcribed before the onset of DNA replication) was found to be fully competent to transform rat cells in culture, and these cells produced tumors on injection into rats (30). The fragment employed in this study was derived from the part of the DNA (between 65.4 and 0/100 map units [14]) which has especially complicated coding properties. Throughout the coding sequence (from 73.8 to 0/100 map units) it contains overlapping information for at least two proteins, and in part for three, making use of different reading frames. Any deletion in this region, therefore, must affect at least two proteins. The "noncoding sequence" (65.4 to 73.8 map units) contains the viral origin of replication and three putative viral early RNA promoters (39). With such an intricate genetic organization, attempts to correlate cellular responses to specific polyoma viral gene products is complicated. Nonetheless, some recently isolated polyoma virus mutants have provided insight into viral growth properties and have suggested a crucial role for the viral middle T-antigen in cellular transformation (15, 16, 23, 26).

In this report we have determined the DNA sequence of four viable early deletion mutants that were isolated and partly characterized previously (16) and compared them with the wild-type DNA sequence (39). Two of the mutants (*dl*-6 and *dl*-28) have deletions from a region near the viral origin of replication, whereas the other two (*dl*-8 and *dl*-23) have deletions that are derived from a region which encodes parts of the large and middle T-antigens. We refer to the latter group of mutants as *mlt* mutants (i.e., mutants affected in the middle and large T-antigens). In addition to the *dl*-8 and *dl*-23 mutants, separate mutants of this type (mutant 45 [2] and mutants 1013-1015 [26]) have recently been isolated in other laboratories. The *dl*-8 and *dl*-23 mutants have altered transformation properties relative to the wild-type virus. Moreover, mutant *dl*-8 appears to replicate very poorly. The DNA sequences presented here allow some deductions to be made about the relationship between the coding properties of these mutants and the altered phenotypes observed.

MATERIALS AND METHODS

Viruses and viral DNA. Polyoma virus deletion mutants *dl*-6, *dl*-8, *dl*-23, and *dl*-28 were isolated and partly characterized previously (16). DNA from these mutants was extracted from virus infected 3T6 mouse cells by the Hirt procedure (20). Covalently closed circular DNA was initially purified by cesium chloride equilibrium density centrifugation followed by further purification on neutral sucrose velocity gradients by standard procedures (13).

Restriction enzymes. All restriction endonucleases were isolated and purified by standard procedures (31) in our laboratory and assayed using polyoma DNA as substrate (12-14).

Purification of bacteriophage T_4-induced DNA polymerase. The T_4-induced DNA polymerase was purified according to a modification of a method of R. Kamen (unpublished data). T_4amN82-infected *Escherichia coli* cells (80 g; New England Enzyme Center, Tufts University Medical School) were suspended in

† Present address: King's College, London WC2, England.

20 mM Tris-hydrochloride (pH 8.0)-5 mM 2-mercaptoethanol (ME)-10 mM $MgCl_2$-1 mM EDTA in 4 M NaCl. The cells were disrupted by sonication, and the suspension was centrifuged for 45 min at 50,000 rpm at 4°C. The supernatant was then supplemented with 0.5 volume of 30% polyethylene glycol (final concentration, 10%). After 30 min at 0°C with continuous stirring, the solution was centrifuged for 20 min at 15,000 rpm, and the supernatant was dialyzed against 4 liters of 20 mM Tris-hydrochloride (pH 8.0)-50 mM NaCl-1 mM EDTA containing 2 mM ME. After dialysis, the extract was loaded onto a DNA-cellulose column (1.5 by 15 cm), washed with dialysis buffer until the absorbancy at 280 nm was 0.05, and proteins which were absorbed were eluted with a linear step gradient from 0.15 M NaCl (100 ml) to 0.3 M NaCl (100 ml) in 20 mM Tris-hydrochloride (pH 8.0)-1 mM EDTA-2 mM ME. Fractions assayed to have DNA polymerase activity (see below) were pooled and applied to a hydroxylapatite column (1.5 by 15 cm). The column was equilibrated with 10 mM phosphate buffer (pH 6.8) containing 10 mM ME. Protein containing DNA polymerase activity was eluted from the column by a gradient from 10 to 300 mM phosphate buffer (pH 6.8)-10 mM ME in 10% glycerol. After dialysis, the material was loaded onto a DE-52 column (1.5 by 4 cm), and DNA polymerase was eluted from the column with a gradient from 0.05 M NaCl (50 ml) to 0.5 M NaCl (50 ml) in 25 mM phosphate buffer (pH 6.8)-10 mM ME-10% glycerol. The DNA polymerase activity eluted in a sharp peak and was applied to a Sephadex G-150 column (Pharmacia AB, Uppsala, Sweden) and eluted with 1 M NaCl in 50 mM Tris-hydrochloride (pH 7.5)-10 mM ME-10% glycerol. Fractions were assayed, and those with enzymatic activity were combined and dialyzed against 0.1 M phosphate buffer (pH 6.8) containing 10 mM ME and 50% glycerol. The final purified protein had an absorbancy at 280 nm of 2.5 (0.36 units of absorbancy at 280 nm per ml) and corresponded to 412 units, assayed in a manner similar to that described by Goulian et al. (11). It was free of detectable endonuclease activity and was used as described below.

Preparation of restriction fragments. The DNA from each mutant was cleaved with the *Hpa*II restriction enzyme in 6 mM Tris-hydrochloride (pH 7.5)-6 mM $MgCl_2$-1 mM dithiothreitol at 37°C for a sufficient time and with a sufficient amount of enzyme for complete digestion. After incubation, the DNA in 1 M NaCl was precipitated with 2 volumes of ethanol at −20°C overnight.

3′-Labelling of restriction fragments. The *Hpa*II restriction endonuclease fragments (14) were incubated in a total volume of 40 μl in 30 mM Tris-hydrochloride (pH 8.0)-30 mM $MgCl_2$-30 mM NaCl-5 mM dithiothreitol in the presence of [α-^{32}P]GTP (2,000 to 3,000 Ci/mmol) (Radiochemical Centre, Amersham, England), and three nonlabeled deoxyribonucleotide triphosphates, the latter being present in at least a twofold excess over the number of 3′ ends. The amount (in nanomoles) of radioactive dGTP used was at least equal to the number of nanomoles of 3′ ends present in the restricted fragments. T_4 DNA polymerase was added in the proportion 1 μl of enzyme per 7 to 10 μg of DNA in a total volume of 40 μl, and the solution was incubated for 30 min at 37°C. The reaction was stopped by the addition of 2 volumes of phenol. Carrier tRNA (50 μg) was added, and the aqueous layer was extracted three times with phenol. Traces of phenol were subsequently removed by three ether extractions. The labeled fragments were twice precipitated with ethanol as described above.

As a control, wild-type polyoma virus DNA (A2 strain) was also cleaved with *Hpa*II and terminally labeled as above.

Separation and extraction of radioactivity labeled fragments. The 3′-labeled fragments were separated by electrophoresis in 5% acrylamide-(bisacrylamide) (purchased from BDH Chemicals Ltd., Poole, England; gel electrophoresis grade) slab gels (20 by 40 cm) in Tris-acetate (pH 7.5) buffer (35). Electrophoresis was carried out at 30 mA until a bromphenol blue dye marker had migrated about two-thirds down the gel. Fragments were located by autoradiography, and those with altered mobilities relative to wild-type DNA fragments were excised and eluted by the procedure of Maxam and Gilbert (28).

Mutants *dl*-6 and *dl*-28 contained *Hpa*II-5 fragments which migrated faster than the wild-type *Hpa*II-5 fragment and overlapped the *Hpa*II-6 fragment. After excision and elution, the DNA was further cleaved with the *Hph*I restriction enzyme. This enzyme leaves the *Hpa*II-6 fragment intact and cleaves the *Hpa*II-5 fragment once (12). The *Hpa*II-5-*Hph*I subfragments were separated from *Hpa*II-6 by electrophoresis on a 10% acrylamide slab gel essentially as described above.

The DNA from mutant *dl*-8 has a *Hpa*II-4 restriction fragment which migrates faster than its counterpart from wild-type polyoma virus DNA. This fragment was isolated and further cleaved with the *Hae*III restriction enzyme (12). Subfragments were separated by electrophoresis and isolated as described above.

The *Hpa*II restriction enzyme digest from mutant *dl*-23 yields only seven fragments compared with eight found in wild-type DNA. The deletion in this mutant results in a fusion of the *Hpa*II-7 and *Hpa*II-8 fragments. The fragment unique to *dl*-23 was cleaved with the *Mbo*II restriction enzyme (16), and subfragments were again separated and isolated as described above.

DNA sequencing. All of the *Hpa*II fragments were cleaved with another restriction enzyme as described above to produce fragments that were labeled at one end only and which were separated on 8 or 10% acrylamide gels (see above). The appropriate fragments were excised and eluted, and DNAs were precipitated with ethanol. The Maxam-Gilbert DNA sequencing method (28) was used for determining the sequence of mutant specific restriction fragments. Electrophoresis was carried out on 20% denaturating acrylamide gels (20 cm by 40 cm by 1 mm) (32). Samples were loaded in several slots at different times to increase the length of DNA sequence which could be read from one gel. Autoradiography was carried out on Kodirex films (Kodak, Ltd.) at −20°C. Detailed descriptions of procedures used for sequencing polyoma virus DNA have been published elsewhere (10, 37).

RESULTS

The sequences of the polyoma virus deletion mutants *dl*-6, *dl*-8, *dl*-23, and *dl*-28 are reported

in this paper. All four of these mutants are viable; that is, they grow in the absence of helper virus. Their deletions are located in the early region of polyoma virus DNA (Fig. 1) and vary in size from 24 to 102 base pairs. The biological properties of the mutants are summarized in Table 1.

Deletion mutants from the coding region of the genome. Two viable deletion mutants of polyoma virus, *dl*-8 and *dl*-23, previously isolated in this laboratory (16) have transformation properties that differ markedly from those of the wild-type virus. The DNA sequence analysis of the *dl*-8 mutant is shown in Fig. 2 and 3. The deletion in this mutant was found to lie between 89.5 and 91.2 map units (14), in a region which appears to code for the body of the middle T-antigen and part of the large T-antigen of polyoma virus, with two different coding frames (10, 38). The data show that 90 deoxynucleotides are deleted in the mutant relative to the wild-type species. This should result in a decrease in molecular weight corresponding to 30 amino acids of both of these T-antigens. The size estimates of the T-antigens of *dl*-8 are consistent with the DNA sequencing data (15, 23). Extensive mapping by other restriction enzymes and sequence analysis of the region around the viral origin of replication (data not given) have shown no differences in other regions of *dl*-8 DNA as compared with wild-type A3 strain viral DNA (38). The *dl*-23 mutant has a deletion which lies between 92.1 and 94.3 map units. The sequence of this region of the viral genome is shown in Fig. 2 and 3. The deletion is 102 nucleotides long. It causes a decrease in molecular weight of middle and large T-antigens consistent with the loss of 34 amino acids (15, 23). Also, for this mutant, no changes in DNA sequence in other regions were observed by restriction enzyme analysis.

The deletions in the *dl*-8 and *dl*-23 mutants do not result in shifts of reading frames. The postulated composition of deleted amino acids in these mutants is shown in Table 2. From a consideration of the changes in amino acid composition of the mutant proteins, the deletions in *dl*-8 and *dl*-23 might be predicted to change considerably the properties of the middle and large T-antigens (see below).

Mutants with deletions in the noncoding region of the viral genome. Two mutants of polyoma virus, *dl*-6 and *dl*-28, were shown by sequence analysis to have deletions of 24 nucleotides relative to wild-type polyoma virus A2 strain DNA (Fig. 1) and to be identical. Their deletions are located entirely within the *Hpa*II-5 restriction endonuclease fragment, between 71.5 and 72.3 map units, near the viral origin of replication.

The DNA sequence between the origin of replication and the ATG initiation codon for the T-antigen is partly self-complementary and

TABLE 1. *Properties of viable early mutants of polyoma virus*

Mutant	Size of deletion in base pairs (% of genome)	Plaque size	DNA synthesis	Transformation
dl-6[a]	24 (0.45)	Large[b]	Normal	As wild-type virus
dl-28[a]	24 (0.45)	Large	Normal	As wild-type virus
dl-8[a]	90 (1.70)	Small	Diminished	More proficient than wild-type virus[c]
dl-23[a]	102 (1.93)	Large[d]	Normal	Deficient

[a] Data taken from Griffin and Maddock (16).
[b] Originally isolated as minute plaque.
[c] Data taken from Griffin et al. (15).
[d] Originally isolated as small plaque.

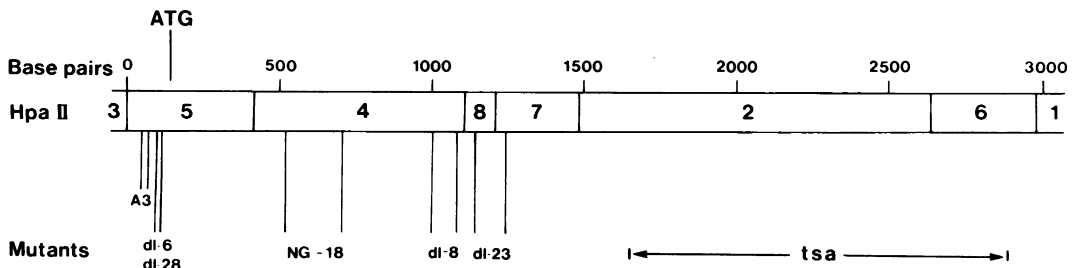

FIG. 1. *Physical map of the early region of polyoma virus A2 strain DNA showing the location of some mutants and a strain variant, A3. The wild-type HpaII physical map is taken from Griffin et al. (14). The base pair numbers and "initiation codon" (ATG) for the early proteins are taken from DNA sequence data (39), and the location of the A3 strain deletion, the deletion in the host range transformation mutant, NG-18, and the temperature sensitive (tsa) mutant lesions are taken from Soeda et al. (38), Soeda and Griffin (40), and Miller and Fried (29), respectively.*

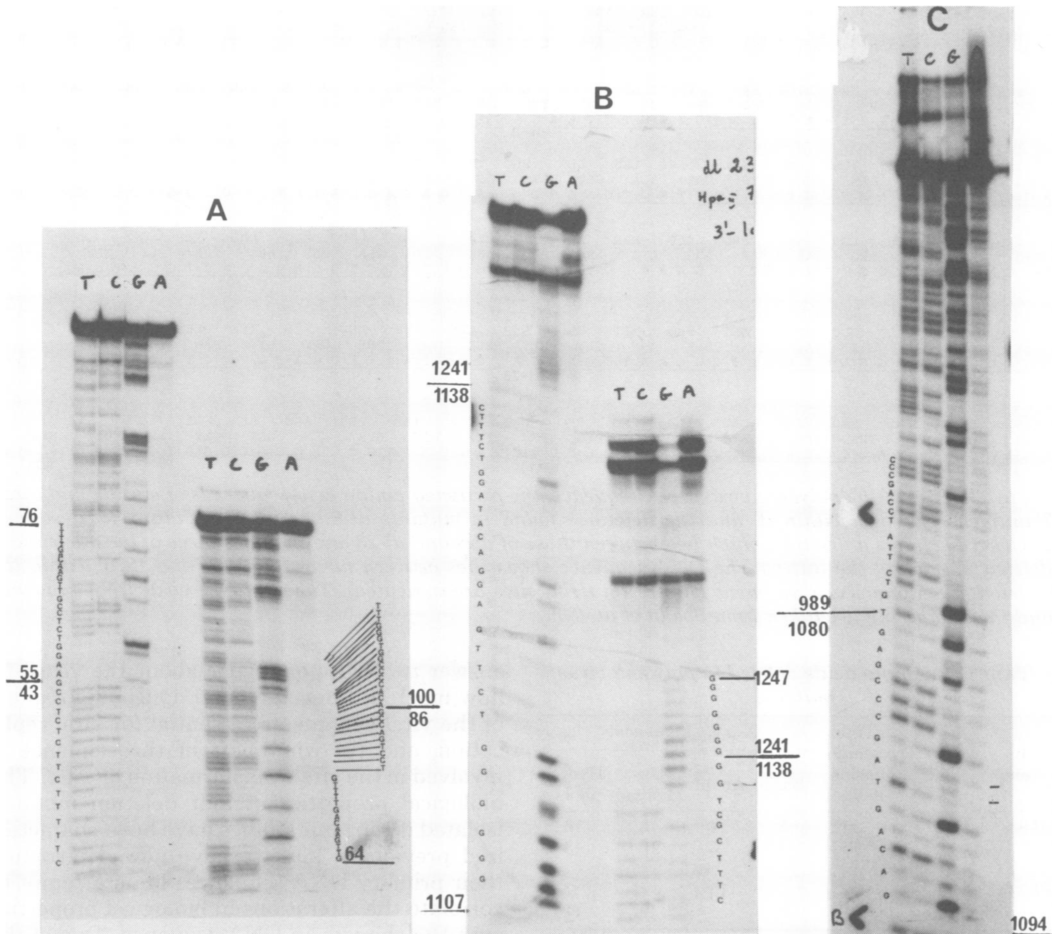

FIG. 2. ***Acrylamide gels from which the sequence of polyoma virus mutants were in part derived: A, dl-28 (or dl-6); B, dl-23; and C, dl-8. DNAs from all mutants were digested with the HpaII restriction enzyme. After 3' labeling, the fragments which migrated with mobilities different from those of wild-type (A3 strain) DNA were excised and cleaved with a second restriction enzyme as appropriate. Thus, fragment HpaII-5 of dl-28 (or dl-6) was further cleaved with HphI, HpaII-4 of dl-8 was further cleaved with HaeIII, and HpaII-7 of dl-23 was further cleaved with MboII. Details of sequencing and DNA analysis are given in the text. The numbers given in A correspond to nucleotides complementary to those shown in Fig. 5. The numbers given in B and C correspond to nucleotide numbers in Fig. 3, the gel sequence in B being, however, derived from the complementary DNA strand. Numbers placed above and below a single line indicate the site of the deletion (see Fig. 3).***

could adopt two hairpin structures. It is noteworthy that the regions around the origins of replication in simian virus 40 and BK viral DNAs permit the formation of two hairpin loops. One such structure in polyoma virus can be formed from the part of the DNA which has considerable (about 80%) homology with either simian virus 40 or BK virus DNA sequences and probably contains their origins of replication (38). The other, however (see Fig. 5), is formed in a region that is unique to polyoma virus DNA. According to present sequence data, the DNAs from the mutants *dl*-6 and *dl*-28 have deleted 13 nucleotides from the latter structure. However, the sequence analysis yields a slightly ambiguous result because of nucleotide duplications at both the 5' and 3' sides of the deletions. For this reason, although they appear to be identical mutants, *dl*-6 and *dl*-23 could have arisen by deletions of different sequences. It is noteworthy that of the two mutants, *dl*-6 was originally a minute-plaque isolate, whereas *dl*-28 was a large-plaque isolate. Since the sequence of both mutants proved to be essentially identical (see below) the plaque morphologies of the viral stocks from which the sequenced DNAs had

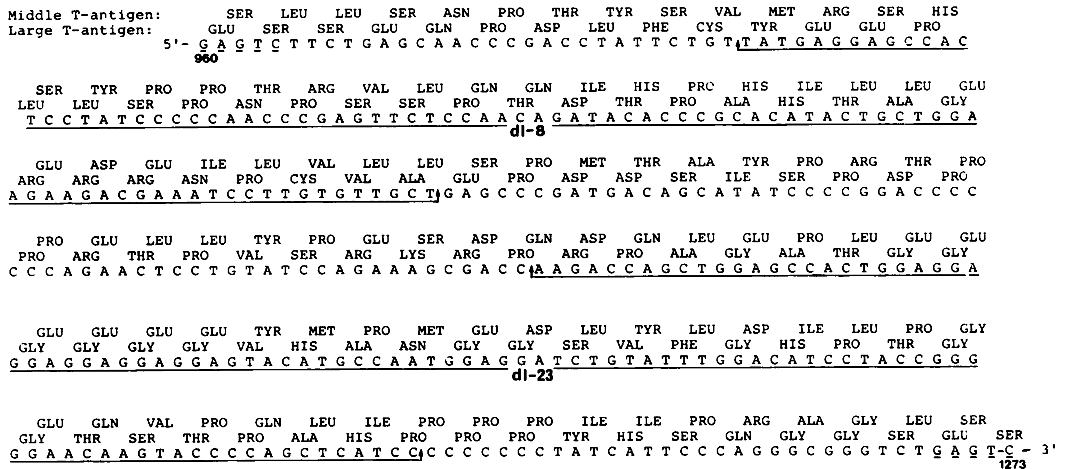

FIG. 3. *Part of the polyoma virus DNA sequence, the predicted amino acid sequences of middle and large T-antigens encoded within it, and the deletions found in mutants dl-8 and dl-23. (Mutant 45, described elsewhere [2], has a deletion which lies between those of dl-8 and dl-23 and includes three of the nucleotides deleted in each of the latter.) The DNA sequence shown lies between nucleotides 960 and 1273, within the restriction endonuclease fragment Hinf-8 (16) (Hinf sites are indicated). The wild-type viral DNA sequence and the numbering are taken from Soeda et al. (39).*

TABLE 2. *Amino acids deleted in polyoma virus mutants*

Mutant	T-antigen	Total amino acids deleted	No. of amino acids[a]			
			Basic	Acidic	Uncharged	Hydrophobic
dl-8	Middle	30	5	4	5	16
	Large	30	4	3	10	13
dl-23	Middle	34		13	4	17
	Large	34	4		19	11

[a] Amino acids are divided into four groups as follows: basic, arg, his, and lys; acidic, asp and glu; uncharged, asn, cys, gln, gly, ser, thr, and tyr; hydrophobic (nonpolar), ala, ile, leu, met, pro, trp, and val. There was no change of amino acid composition in the small T-antigen of either mutant (23).

been isolated were determined (Table 1). Stocks of both mutants gave large plaques. This is not the first instance in which plaque morphologies of mutants have appeared to change during propagation in culture. A similar observation was made in the case of the *dl*-23 mutant. Since mutant virus stocks employed here represent two passages of the original virus isolate in culture, it is possible that the original plaques may have contained more than one mutant species and that selection of the more viable one has occurred during viral growth. Alternatively, the mutants may have undergone some alteration during passage in their permissive host cell.

DISCUSSION

Deletion mutants of the early region of polyoma virus have been isolated in an attempt to answer two basic questions about the virus: (i) how much of the coding and noncoding regions of the viral genome are essential for viral replication, and (ii) what part of the sequence is involved in the viral transformation of cells? The biological properties of four deletion mutants isolated during our studies have been characterized previously, and we have now determined their primary DNA sequences in an attempt to correlate the alterations in biological properties with the changes in DNA sequence. Two of the mutants (*dl*-6 and *dl*-28) have growth and transformation properties similar to those of the wild-type virus. The other two mutants (*dl*-8 and *dl*-23) have transformation properties which differ markedly from those of the wild-type virus, and one of them (*dl*-8) has altered growth properties.

Mutants *dl*-8 and *dl*-23. The *dl*-8 and *dl*-23 *mlt* mutants are of special interest since they have altered transformation phenotypes relative to the wild-type virus (16). Moreover, the transformation phenotype of *dl*-23 differs considerably from that of *dl*-8, even though, as shown here, their sequence deletions lie in the same part of the genome, only sixty nucleotides apart (Fig. 3 and Table 1). Since analysis of the early proteins induced by wild-type and mutant viruses showed no apparent quantitative differences among the various T-antigens (23), the change in biological properties observed in the mutants would appear to be a direct result of protein sequence alterations.

A functional large T-antigen is essential for DNA replication and virus growth (8). Studies

on the large T-antigen of simian virus 40 have shown that it binds specifically to viral DNA in the region around the viral origin of replication (42), and it seems likely that polyoma virus large T-antigen can interact with viral DNA in an analogous fashion (38). The properties of the *dl*-8 mutant suggest that viral DNA replication is in some way impaired. On the other hand, this mutant appears to express properties associated with transformation and tumor formation to a greater extent than does the wild-type virus. From these and other considerations, we have postulated that impairing replication might enhance transformation (15). That is, in rat cells (which are semipermissive for expression of polyoma virus) replication and transformation might appear as antagonistic events. For example, an analysis of the wild-type DNA sequence missing in *dl*-8 shows that the deletion in this mutant (corresponding to 30 amino acids) removes a part of a very proline-rich region of large T-antigen (Fig. 3). Overall, polyoma virus large T-antigen is not especially rich in proline (37), but the sequence present in wild-type virus and deleted in *dl*-8 would code for a polypeptide that contains 20% of this amino acid. Since proline residues often play a special role in folding of both globular and fibrous proteins (9), a sequence alteration in a proline-rich part of a protein might be expected to modify tertiary structure and biological function more than might be expected from a similar deletion of other amino acids.

A similar consideration of the coding sequence lost in the gene for middle T-antigen of *dl*-8, although it points to no unusual features in terms of single amino acids, shows that many of the amino acids in the missing fragment are essentially hydrophobic in nature (Fig. 3 and Table 2).

Analysis of the wild type viral DNA sequence missing in *dl*-23 is shown in Fig. 2 and 3. Since *dl*-23 grows at least as well as the wild-type virus, it would appear that the mutation has little functional effect upon the large T-antigen. On the other hand, the *dl*-23 mutant does not express all of the phenotypic properties normally associated with transformation in rat fibroblast cells. This might be due to the particular loss of amino acids from the middle T-antigen, since it would now appear that this protein plays an important role in the transformation of cells by polyoma virus (15, 23). In the *dl*-23 mutant, 13 acidic amino acids out of a total of 34 may be predicted to have been lost from the wild-type protein and a very acidic polypeptide, hexaglutamic acid [$(Glu)_6$], found only once in polyoma virus, is removed (38, 39). Since middle T-antigen is a membrane-associated protein (22) and changes of the charge characteristic of proteins known to bind to membranes can be shown to cause dramatic changes in membrane properties (3, 5), it is tempting to suggest that the unusual phenotype of the *dl*-23 mutant may be explained by its altered middle T-antigen. However, in this connection it should be noted that in the mutant middle T-antigen-related protein, membrane association is maintained (15). Several studies (6, 33, 36) have shown that a kinase activity might be associated with middle T-antigen and that in vivo this protein is phosphorylated on a tyrosine residue (6). Sequence analysis shows that the *dl*-23 mutant middle T-antigen loses two tyrosine residues, one of which is contiguous with the $(Glu)_6$ residue, and the protein has been found to be poorly phosphorylated (36). However, we note that an *mlt* mutant (*dl*-1015) recently isolated by Magnusson and Berg (26) was found to have properties similar to those of the *dl*-23 mutant, although it retains several of the DNA sequences which code for the acidic residues, including those which code for the polypeptide $(Glu)_6$. Therefore, it is not clear what correlation, if any, exists between the membrane association, kinase activity, and transformation.

The mutants *dl*-8 and *dl*-23 were isolated by transfection with heteroduplex DNA of polyoma virus strains A2 and A3 after treatment of the heteroduplex molecule with the single-strand specific nuclease S1. Since sequence differences between these strains are only known to exist in the region around the viral origin of replication (38), the region covered by mutants *dl*-8 and *dl*-23 and mutant 45 (described elsewhere [2]) may contain sites that are intrinsically S1 sensitive. The sequences from this portion of the genome do contain extensive regions of intrastrand complementarity that might allow the DNA to assume a very unusual secondary structure (Fig. 4). It is not suggested that in the DNA such a structure would stably exist under physiological conditions (although it might form part of a stable secondary structure in the viral mRNA's), but in the presence of the anomalously high Zn^{2+} concentration and the low pH employed for nuclease S1 treatment such a structure, or a portion of it, might conceivably exist as an intermittent alternative conformation that would cause the DNA to become susceptible to the enzyme. It is noteworthy that in recent studies aimed at probing the DNA structure of polyoma virus using electron microscopy, Wu et al. (45) have shown that the viral single-stranded DNA has considerable secondary structure in the region which maps between 92.9 ± 0.8 and 95.0 ± 0.7 map units, or roughly the region covered in Fig. 4. Moreover, as more DNA sequence data on different organisms become available, the

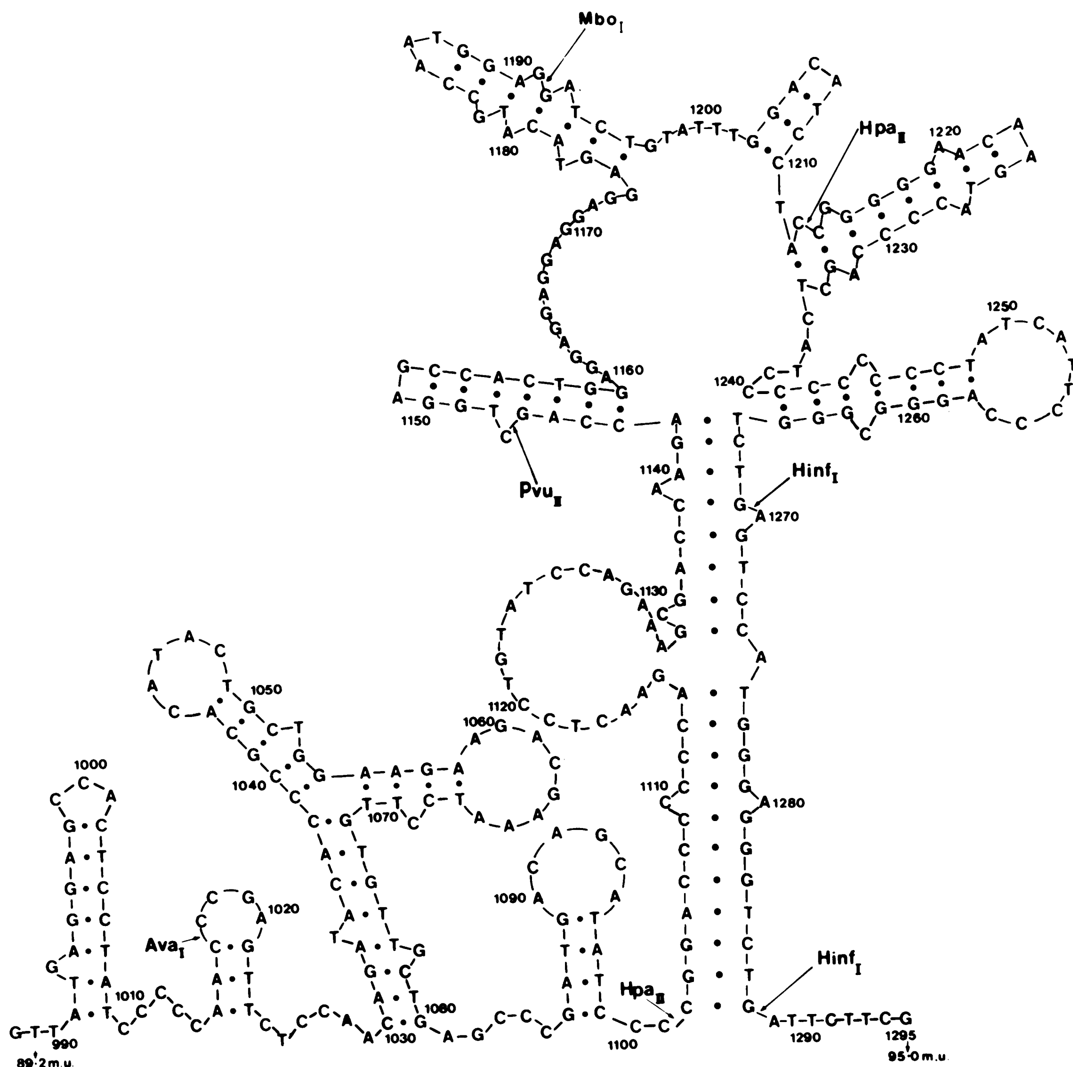

FIG. 4. *Secondary structure suggested by the DNA sequence of the part of polyoma virus which lies between nucleotides 990 and 1290 (89.2 to 95.0 map units) (38). Some restriction endonuclease sites which lie within this region are indicated. The dl-8 mutant lacks the sequence which includes nucleotides 990 through 1079, inclusive, and the dl-23 mutant lacks the sequence which includes nucleotides 1139 through 1240. Studies on single-stranded polyoma virus DNA by electron microscopy (45) revealed a hairpin "fold back" structure between 92.9 ± 0.8 to 95.0 ± 0.7 map units, which includes most of the largest loop shown in this figure.*

potential for such intra-strand complementary structures to exist and possibly even to play functional roles becomes more apparent. For example, there are obvious similarities between the secondary structure proposed for the region around the origin of replication in *Salmonella typhimurium* (46) and that shown in Fig. 4.

Mutants *dl*-6 and *dl*-28. Mutants *dl*-6 and *dl*-28, as originally isolated, differed in their plaque morphologies, but according to the present DNA sequence data (Fig. 2) they have either the same or a very similar deletion in their DNAs. The mutants are both derivatives of the A3 (large-plaque) strain of polyoma virus that lacks 11 base pairs relative to the more commonly studied large-plaque A2 strain (38). This deletion lies very near the viral origin of replication. The *dl*-6 and *dl*-28 mutants also lack a further 13 base pairs, relative to the A2 DNA, which are located about midway between the origin of replication and the ATG triplet that specifies the initiation codon for the viral early mRNA's. The region deleted has a unique (as opposed to repeated) sequence and includes the

two *Hae*II restriction endonuclease sites and the single *Bgl*I site found in polyoma virus DNA. These mutants replicate and transform in a manner practically indistinguishable from that of the wild-type virus (Table 1).

The exact sequence removed from the A3 strain DNA to generate mutants *dl*-6 and *dl*-28 could not be deduced by DNA sequence analysis, since, as shown in Fig. 5, a small C-G dinucleotide repeat occurs before and after the deletions. A similar situation has been observed with some of the other polyoma virus deletion mutants as well as with mutants with deletions within the noncoding part of the late region of simian virus 40. For example, an apparently similar sequence is deleted in viable simian virus 40 mutants 893 and 894 and in strains Rh911 and 1802, but ambiguities which result from short sequence repeats on both the 3′ and 5′ sides of the deletions make it impossible to assign these deletions to specific nucleotides in the wild-type DNA (43).

Several other viable mutants with deletions in the noncoding part of the early region of polyoma virus have been characterized recently (1, 26, 44). The deletions all seem to cluster around the *Hae*II and *Bgl*I sites, so it must be assumed that this region is either nonessential for the virus or serves a function not yet identified. Viable mutants with deletions near the origin of replication in simian virus 40 usually remove one copy of a tandemly repeated sequence (41). This is not the case in polyoma virus, in which the sequences removed in the mutants do not appear in other parts of the DNA. So far, however, no polyoma virus viable mutants have been isolated that contain a deletion within the striking hairpin or four-stranded structures that could occur close to the origin of replication (38). Some of the mutants isolated by Bendig et al. (1) grow poorly and have deletions which presumably approach the boundaries of the noncoding and nonessential region of polyoma virus. As might be predicted, however, it would appear that there are strictly limited portions of the DNA that can be deleted and at the same time allow the viral phenotype to be apparently essentially retained. Since in several other parts of the polyoma virus genome, two, and sometimes even three, reading frames are used for coding (10, 38–39), it would be surprising if extensive viral sequences could be removed without a concomitant change in some function. Moreover, it is possible that in these mutants a virally induced cellular function may have been altered, and we (and others) would not have detected this in our studies.

A final interesting observation arises from a consideration of the DNA sequences of mutants described in this manuscript as well as other

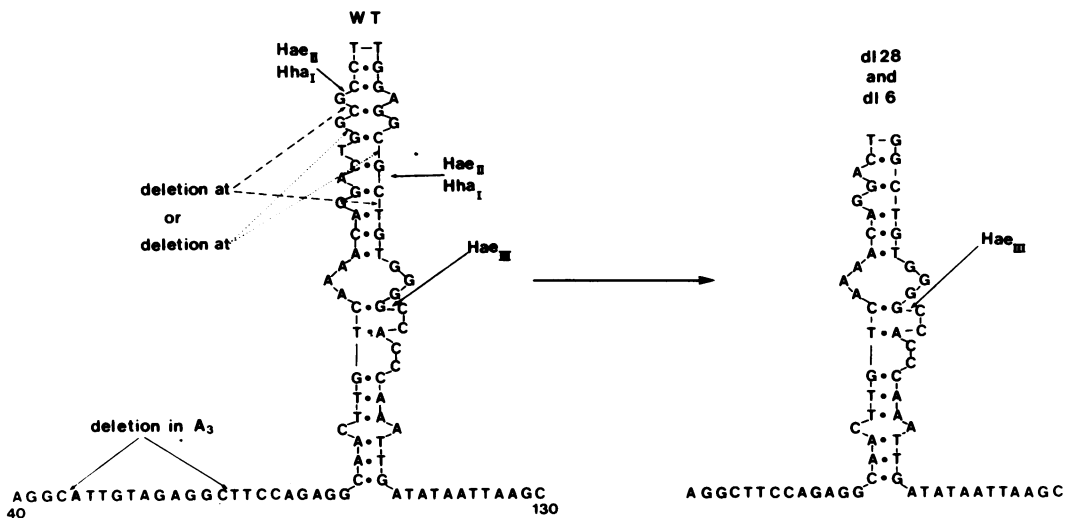

FIG. 5. *Secondary structure possible within the early noncoding region of polyoma virus wild-type DNA and altered in some deletion mutants. Virus mutants dl-6 and dl-28 have deletions of 24 base pairs relative to the wild-type A2 strain and appear to be derived from the A3 strain. The deletions in these mutants remove the single BglI recognition site (between nucleotides 87 and 97) found in polyoma virus DNA, both the HaeII sites (as shown), one of the GGAGG pentanucleotides present five times, and one of the AGAGGCC heptanucleotides present four times in the noncoding region of wild-type viral DNA (38); they retain a CTTCC possible ribosomal binding site. Because of a small sequence duplication, the deletions cannot be precisely defined and may be located as indicated either by the dotted or dashed lines. Either deletion gives a mutant with the same structure. The dashed line shows the more symmetrical of the two possible deletions. Numbers are taken from the DNA sequence (39).*

TABLE 3. *Ambiguities in DNA sequence in deletion mutants of polyoma virus*[a]

Virus	Nucleotide sequence	Reference
A3	G A↓G↓G↓C (A T T.... ↓G↓G↓C) T T C C A	(38)
NG-18	T A G T↓A (A A T.... G C↓A) T G C C T	(40)
A8	A G A C↓A (C C T.... C A↓A) A T G T G	(19)
B2	C T↓G↓C↓A (G A A.... ↓G↓C↓A) C A G C G	(19)
SD15	A A G↓G↓C (T A C.... A↓G↓C) A A C A T	(4)
dl-6 *dl*-28	C T G↓G↓C (G C C.... C↓G↓C) T G T G G	
dl-8	T C T G↓T (T A T.... G C↓T) G A G C C	
dl-23	C G A↓C↓C (A A G.... T↓C↓C) C C C C C	

[a] Deletions can be sequences given within brackets or alternatively those indicated by arrows. A similar ambiguity has been observed in the *dl*-1015 mutant (26), though apparently not in mutant 45 (2).

polyoma virus mutants. In many mutants (Table 3), there is an ambiguity in the assignment of sequences around the deletion sites, i.e., the deletions that have occurred are found at sites which have repeating sequences. In some cases these are simple dinucleotide repeats, but in other cases they are longer. Similar repetitions are apparent in DNA sequence studies on simian virus 40 deletion mutants (17, 34, 43). In this connection, it is interesting to note that comparable results have apparently been observed in genes coding for various human β-globins (27). Further, in *E. coli*, spontaneous deletions account for a large proportion of spontaneous mutations, and they occur at regions of short sequence repeats (7).

ACKNOWLEDGMENTS

We thank Anne Leigh-Brown for assistance in preparation of DNA and for plaque assays, and R. Kamen for the T4 bacteriophage DNA polymerase protocol.

LITERATURE CITED

1. **Bendig, M. M., T. Thomas, and W. R. Folk.** 1980. Regulatory mutants of polyoma virus defective in DNA replication and the synthesis of early proteins. Cell **20**: 401–409.
2. **Bendig, M. M., T. Thomas, and W. R. Folk.** 1980. Viable deletion mutant in the medium and large T-antigen-coding sequences of the polyoma virus genome. J. Virol. **33**:1215–1220.
3. **Carasco, L.** 1977. The inhibition of cell functions after viral infection. FEBS Lett. **76**:11–15.
4. **Carmichael, G. G., and T. L. Benjamin.** 1980. Identification of DNA sequence changes leading to loss of transforming ability in polyoma virus. J. Biol. Chem. **255**:230–235.
5. **Durham, A.** 1978. A single protein key to cancer? New Sci. **77**:860–862.
6. **Eckhart, W., M. A. Hutchinson, and T. Hunter.** 1979. An activity phosphorylating tyrosine in polyoma T-antigen immunoprecipitates. Cell **18**:925–933.
7. **Farabaugh, P. J., and J. H. Miller.** 1978. Genetic studies of the *lac* repressor. VII. On the molecular nature of spontaneous hotspots in the *lacI* gene of *Escherichia coli*. J. Mol. Biol. **126**:847–863.
8. **Francke, B., and W. Eckhart.** 1973. Polyoma gene function required for viral DNA synthesis. Virology **55**: 127–135.
9. **Freedman, R.** 1979. Proline and folding proteins. Nature (London) **279**:756–757.
10. **Friedmann, T., A. Esty, P. La Porte, and P. Deininger.** 1979. The nucleotide sequence and genome organization of the polyoma early region: extensive nucleotide and amino acid homology with SV40. Cell **17**:715–724.
11. **Goulian, M., S. J. Lucas, and A. Kornberg.** 1968. Enzymatic synthesis of deoxyribonucleic acid. Purification and properties of deoxyribonucleic acid polymerase induced by infection with phage T4. J. Biol. Chem. **243**:627–638.
12. **Griffin, B. E.** 1977. Fine structure of polyoma virus DNA. J. Mol. Biol. **117**:447–471.
13. **Griffin, B. E., and M. Fried.** 1976. Structural mapping of the DNA of an oncogenic virus (polyoma viral DNA). Methods Cancer Res. **12**:49–86.
14. **Griffin, B. E., M. Fried, and A. Cowie.** 1974. Polyoma DNA: a physical map. Proc. Natl. Acad. Sci. U.S.A. **71**: 2077–2081.
15. **Griffin, B. E., Y. Ito, U. Novak, N. Spurr, S. Dilworth, N. Smolar, R. Pollack, K. Smith, and D. Rifkin.** 1980. Early mutants of polyoma virus (dl-8 and dl-23) with altered transformation properties: Is polyoma virus middle-T-antigen a transforming gene product? Cold Spring Harbor Symp. Quant. Biol. **44**:271–282.
16. **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.
17. **Gruss, P., C.-J. Lai, R. Dhar, and G. Khoury.** 1979. Splicing as a requirement for biogenesis of functional 16S mRNA of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **76**:4317–4321.
18. **Hassell, J. A., W. C. Topp, D. B. Rifkin, and P. E. Moreau.** 1980. Transformation of rat embryo fibroblasts by cloned polyoma virus DNA fragments containing only part of the early region. Proc. Natl. Acad. Sci. U.S.A. **77**:3978–3982.
19. **Hattori, J., G. G. Carmichael, and T. L. Benjamin.** 1979. DNA sequence alterations in hr-t deletion mutants of polyoma virus. Cell **16**:505–513.
20. **Hirt, B. J.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26**:365–369.
21. **Israel, M. A., D. T. Simmons, S. L. Hourihan, W. P. Rowe, and M. A. Martin.** 1979. Interrupting the early region of polyoma virus DNA enhances tumorigenicity. Proc. Natl. Acad. Sci. U.S.A. **76**:3713–3716.
22. **Ito, Y., J. R. Brocklehurst, and R. Dulbecco.** 1977. Virus-specific proteins in the plasma membrane of cells lytically infected or transformed by polyoma virus. Proc. Natl. Acad. Sci. U.S.A. **74**:4666–4670.
23. **Ito, Y., N. Spurr, and B. E. Griffin.** 1980. Middle T-antigen as primary inducer of full expression of the phenotype of transformation by polyoma virus. J. Virol. **35**:219–232.
24. **Kamen, R., D. M. Lindstrom, H. Shure, and K. W. Old.** 1975. Virus-specific RNA in cells productively infected or transformed by polyoma virus. Cold Spring Harbor Symp. Quant. Biol. **39**:187–198.
25. **Lania, L., D. Gardini-Attardi, M. Griffiths, B. Cooke, D. de Cicco, and M. Fried.** 1980. The polyoma virus 100K large T-antigen is not required for the maintenance of transformation. Virology **101**:217–232.
26. **Magnusson, G., and P. Berg.** 1979. Construction and analysis of viable deletion mutants of polyoma virus. J. Virol. **32**:523–529.
27. **Marotta, C. A., J. T. Wilson, B. G. Forget, and S. M. Weissman.** 1977. Human β-globin messenger RNA. III. Nucleotide sequences derived from complementary DNA. J. Biol. Chem. **252**:5040–5053.
28. **Maxam, A., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74**:560–564.

29. **Miller, L. K., and M. Fried.** 1976. Construction of a genetic map of the polyoma genome. J. Virol. **18**:824-832.
30. **Novak, U., S. M. Dilworth, and B. E. Griffin.** 1980. Coding capacity of a 35% fragment of the polyoma virus genome is sufficient to initiate and maintain cellular transformation. Proc. Natl. Acad. Sci. U.S.A. **77**:3278-3282.
31. **Roberts, R. J.** 1976. Restriction endonucleases. C.R.C. Crit. Rev. Biochem. **4**:123-164.
32. **Sanger, F., and A. R. Coulson.** 1978. The use of thin acrylamide gels for DNA sequencing. FEBS Lett. **87**: 107-110.
33. **Schaffhausen, B. E., and T. L. Benjamin.** 1979. Phosphorylation of polyoma T antigens. Cell **18**:935-946.
34. **Seif, J., G. Khoury, and R. Dhar.** 1980. A rapid enzymatic DNA sequencing technique: determination of sequence alterations in early simian virus 40 temperature sensitive and deletion mutants. Nucleic Acids Res. **8**:2225-2240.
35. **Sharp, P. A., B. Sugden, and J. Sambrook.** 1973. Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose ethidium bromide electrophoresis. Biochemistry **12**: 3055-3063.
36. **Smith, A. E., R. Smith, B. E. Griffin, and M. Fried.** 1979. Protein kinase activity associated with polyoma virus middle T-antigen *in vitro*. Cell **18**:915-924.
37. **Soeda, E., J. R. Arrand, and B. E. Griffin.** 1979. Polyoma virus. The early region and its T-antigens. Nucleic Acids. Res. **7**:839-857.
38. **Soeda, E., J. R. Arrand, N. Smolar, and B. E. Griffin.** 1979. Sequence from early region of polyoma virus DNA containing viral replication origin and encoding small, middle and (part of) large T-antigens. Cell **17**:357-370.
39. **Soeda, E., J. R. Arrand, N. Smolar, J. E. Walsh, and B. E. Griffin.** 1980. Coding potential and regulatory signals of the polyoma virus genome. Nature (London) **283**:445-453.
40. **Soeda, E., and B. E. Griffin.** 1978. Sequences from the genome of a nontransforming mutant of polyoma virus. Nature (London) **276**:294-298.
41. **Subramanian, K. N.** 1979. Segments of simian virus 40 DNA spanning most of the leader sequence of the major late viral messenger RNA are dispensable. Proc. Natl. Acad. Sci. U.S.A. **76**:2556-2560.
42. **Tjian, R.** 1978. The binding site on SV40 for a T-antigen related protein. Cell **13**:165-179.
43. **Van Heuverswyn, H., and W. Fiers.** 1979. Nucleotide sequence of the Hind-C fragment of simian virus 40 DNA. Eur. J. Biochem. **100**:51-60.
44. **Wells, R. D., M. A. Hutchinson, and W. Eckhart.** 1979. Isolation and characterization of polyoma virus genomes with deletions between the origin of viral DNA replication and the site of initiation of translation in the early region. J. Virol. **32**:517-522.
45. **Wu, M., H. Manor, and N. Davidson.** 1979. Secondary structures in polyoma DNA. J. Virol. **32**:334-338.
46. **Zyskind, J. W., and D. W. Smith.** 1980. Nucleotide sequence of the Salmonella typhimurium origin of DNA replication. Proc. Natl. Acad. Sci. U.S.A. **77**:2460-2464.