JOURNAL OF VIROLOGY, Mar. 1981, p. 871-875
0022-538X/81/030871-05$02.00/0

Vol. 37, No. 3

Nucleotide Sequence Changes in Polyoma *ts-a* Mutants: Correlation with Protein Structure

PRESCOTT L. DEININGER, PATRICIA LAPORTE, AND THEODORE FRIEDMANN*
Department of Pediatrics, University of California, San Diego, La Jolla, California 92093

The mutations in three polyoma *ts-a* mutants have been determined. Two mutants, *ts*-25 and *ts*-52, have different single-base changes at the same position (2883) in the early region corresponding to a conserved glycine residue very near the C-terminus of the polyoma large T antigen. Mutant *ts*-48 has a single-base change at position 2341, as well as a second change at position 1228, in the region of large T antigen shared with medium T antigen.

A thorough understanding of the structure-function relationship of a biologically active protein requires detailed information of the enzymatic and other biological functions, knowledge of its primary structure through amino acid sequencing, a description of secondary and tertiary structural features through X-ray crystallography and other physical methods, the identification of functional domains and residues through characterization of mutants, and evolutionary studies of functionally and genetically related proteins to identify conserved regions of the proteins. Such analyses require ready sources of purified proteins and well-developed genetic systems in addition to assays for protein function. Within the past several years, the availability of purified simian virus 40 (SV40) large T antigen (13, 26) and the description of its protein kinase and ATPase activities (11, 27), together with the likely availability of other papovavirus tumor antigens, have begun to make possible both an examination of the structure-function relationships of the papovavirus tumor (T) antigens and an understanding of the mechanisms of cell transformation mediated by these proteins. Of major importance has been the development of methods for the rapid determination of the nucleotide sequence of large DNA molecules (17, 21), by which the nucleotide sequences of the entire genomes of SV40 (5, 19), polyoma virus (2, 8, 24), and BK virus (22) have been determined. These studies provided valuable information about the nature and structure of all of the virus-coded proteins, including the T antigens, long before purified proteins became available. Using nucleotide sequence information alone, we have identified functionally important regions of these molecules from an identification of conserved regions in the viral genomes (7). Other workers have examined the effects of mutations at a variety of positions in the genomes on the function of viral gene products.

We wish to present here the nucleotide sequence of polyoma group *A* mutants that affect the structure and expression of the polyoma large T antigens and a correlation of the altered sequences with characteristics of the protein structure.

MATERIALS AND METHODS

Wild-type strain 3 polyoma is the strain whose complete nucleotide sequence has been described and analyzed in detail in this laboratory (2, 8). The temperature-sensitive (*ts*) mutants *ts*-25, *ts*-48, and *ts*-52 were obtained from Walter Eckhart, The Salk Institute, La Jolla, Calif., and have been characterized extensively (3). All three mutants have the polyoma *ts-a* phenotype; i.e., they fail to transform cells at a nonpermissive temperature, and they are also conditionally defective in the induction of large T antigen and in the induction of viral DNA synthesis. All viral DNAs were prepared after infection of 3T6 cells by previously described methods (9).

Preparation of cloned fragments. DNA purified from the mutants was cleaved with *Mbo*I and cloned by ligation into the *Bam*HI site of plasmid pBR322; selection of ampicillin-resistant colonies and detection of appropriate polyoma inserts were done by the hybridization method of Grunstein and Hogness (12), using nick-translated (16) wild-type polyoma *Hae*III fragment 7 as a probe. Similarly, the *Pst*I fragments were cloned into the *Pst*I site of pBR322, and inserts were detected as above by selecting tetracycline-resistant colonies and by colony hybridization with wild-type polyoma *Hpa*II fragment 6.

Full-genome clones were prepared by cleaving mutant DNA with *Bam*HI and ligating into the *Bam*HI site of pBR322. The resulting colonies were screened as described above. Plasmid DNA was prepared in SF8 *Escherichia coli* (4) and purified in cesium chloride-ethidium bromide gradients (1, 20). *Hae*III and *Hpa*II fragments were prepared, labeled, and sequenced as described previously (9). All cloned materials were prepared and used in accordance with National Institutes of Health and University of California, San Diego, Biosafety Committee guidelines.

Nucleotide sequencing. Details of the sequence

determination of wild-type polyoma have been reported previously (2, 8, 9). The nucleotide sequences of the mutant DNAs were determined by the method of Maxam and Gilbert (17) and included the G-specific methylene blue reaction for G (6) and the potassium permanganate reaction for T (C. Houck, personal communication). Details of fragment preparation, labeling, and purification, gel electrophoresis and autoradiography, information storage, and sequence comparisons have been published previously from this laboratory (7-9).

RESULTS

Marker rescue experiments from other laboratories have previously mapped the mutations of several group mutants of polyoma (18). *ts-a* was localized to a fragment between the *Hin*dIII site at nucleotide position 1671 and an *Hha*I site at position 2331, whereas mutants *ts*-25, *ts*-48, and *ts*-52 have all been localized to the *Hha*I fragment spanning the region from positions 2331 to 2913 by the numbering convention used by Friedmann et al. (9). The sequence changes in the mutants examined in this study are summarized in Table 1. Each of the mutants showed only a single mutation within the marker rescue fragment, indicating that the mutations given in Table 1 are responsible for the *ts* phenotype. Mutant *ts*-48 is, in addition, known to be missing the *Hpa*II cleavage site at the junction of *Hpa*II fragments 7 and 8. We have sequenced across this site from the *Mbo*I site at position 1207 and have found a transition from the wild-type C to a T at position 1228 (W. Eckhart, S. Delbrück, P. Deininger, T. Friedmann, and T. Hunter, Virology, in press).

Figure 1 shows the amino acid distribution of the polyoma large T antigen and the positions of the mutations determined in this study. The positions at which all three papovavirus nucleotide sequences are identical in the large T antigen are indicated at the top of Fig. 1.

TABLE 1. *Nucleotide and amino acid sequence changes in polyoma ts-a mutants*

Mutant	Nucleotide no.	Nucleotide change	Amino acid change and position
ts-25	2883	G → T	Gly → Cys (large T antigen residue 778)
ts-48	2341	G → T	Ser → Ileu (large T antigen residue 597)
	1228	C → T	Thre → Ileu (large T antigen residue 597)
ts-52	2883	G → A	Gly → Ser (large T antigen residue 778)

DISCUSSION

The *ts-a* mutants characterized in this study were generated by nitrous acid mutagenesis of the polyoma genome, making it possible, and even likely, that some of these mutants contain multiple alterations in large T antigen and other viral proteins, as in the case of *ts*-48, in which we have found one alteration affecting only large T antigen, one affecting both the large T antigen and the medium T antigen (see below), and another altering the VP2,3 terminator (Eckhart et al., in press). The use of the marker rescue data and analysis of wild-type revertants makes it possible to determine the region containing the alterations responsible for the phenotypic changes brought about by an altered large T antigen (18). In the cases of all three mutants, there is only a single-base-pair change within the mapped region, indicating that, for these mutants, it is a single amino acid change that affects the *ts* phenotype and that other possible alterations outside the marker rescue region have little or no effect on T-antigen function.

The mutation in *ts*-48 that confers the *ts-a* phenotype is a transversion from a G to a T at position 2341, resulting in a substitution of an isoleucine residue for the normally occurring serine residue in large T antigen (Table 1). This mutation is in one of the most highly conserved regions of the papovavirus large T antigens, although the wild-type polyoma serine residue itself has not been conserved in the two other papovaviruses (Fig. 1). The serine residue is located amid a number of hydrophobic residues, mostly leucine and alanine, in a general region that also contains a number of the positively charged residues lysine and arginine (Fig. 1). The more short-range environment, however, in the immediate vicinity of the *ts*-48 mutation is devoid of charged residues over a stretch of 12 amino acids. Thus, the change of a serine to an isoleucine may enrich this already hydrophobic region.

The *ts*-52 mutation responsible for the altered phenotype is a G-to-A transition at position 2883, very near the portion of the sequence encoding the C-terminus of the T antigen at position 2919. This region in polyoma virus shows relatively poor conservation with the other papovaviruses, but the alteration itself does represent the change of a conserved glycine (Fig. 1) to serine. Mutant *ts*-25 contains a G → T transversion at the same position as the *ts*-52 mutation, resulting in a change from glycine to cysteine in the mutant protein. These alterations

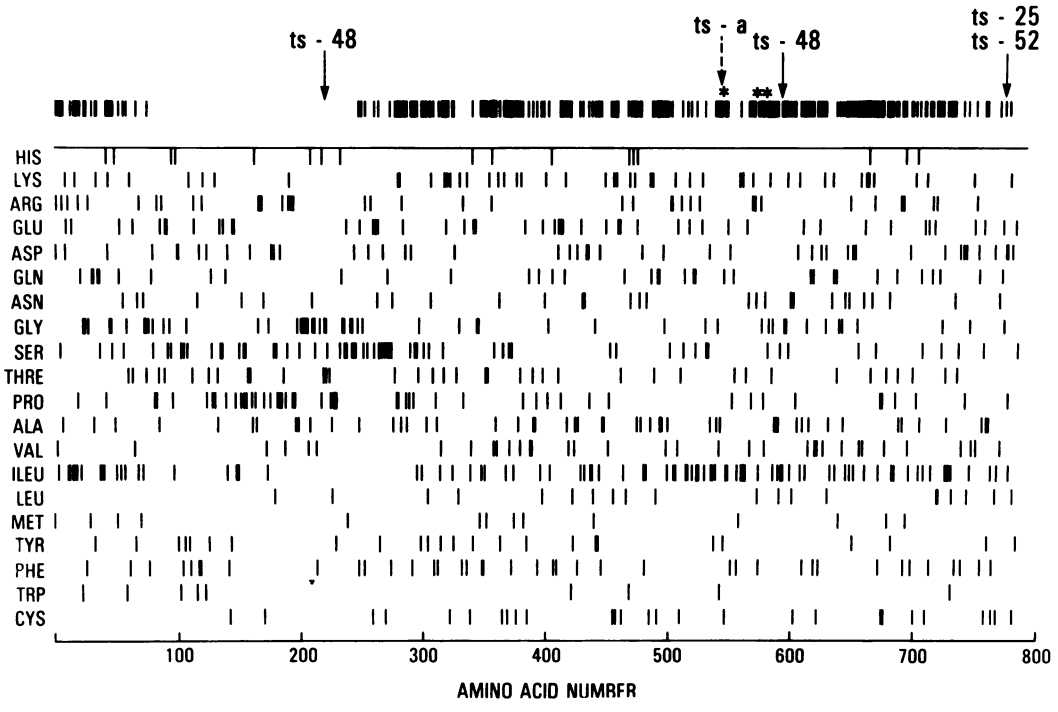

FIG. 1. *Amino acid distribution of the polyoma large T antigen. The arrows indicate the position of the mutations in ts-25, ts-48, and ts-52. The positions that are conserved in all three papovavirus large T antigens are indicated at the top. The large gap represents the unique portion of the polyoma large T antigen. The asterisks indicate the positions of the sequence changes in four SV40 tsA mutants (23), and the dashed arrow indicates the ts-a polyoma mutant characterized by Thomas et al. (25). The ts-25 sequenced by Thomas et al. is mutated at the same site as the ts-25 sequence in our study.*

occur in the relatively highly negatively charged C-terminus of the large T antigen and place a serine residue within a negatively charged cluster. The C-terminus of the wild-type large T antigen contains eight negatively charged residues and only one positively charged residue.

We have detected in *ts*-48 a second early region point mutation outside the marker rescue region, a C-to-T transition at position 1228 which causes a change from threonine to isoleucine in large T antigen at amino acid position 219 (Fig. 1), and a change from proline to serine in medium T antigen. This mutation lies within a domain of the polyoma large T antigen that is unusual in several respects. Most notable are the very marked clustering of glycine and serine residues, the large number of proline and threonine residues, and the deficiency of charged residues. This region contains a portion of the coding region for the polyoma medium T antigen, and since there is no equivalent region in the BK virus or SV40 genomes, it can be considered an insert, relative to those genomes. The marker rescue results indicate that neither the change in medium T antigen nor the change in the large T antigen caused by the mutation at position 1228 is responsible for the *ts-a* phenotype. However, it is possible that the medium T antigen, not required for lytic infection of permissive cells, or even large T antigen functions are altered in a manner not detected by the lytic infection assay used for marker rescue.

The region of the polyoma medium T antigen affected by the *ts*-48 mutation at position 1228 is within the unusual glutamic acid-, aspartic acid-, and proline-rich region of the molecule that characterizes the C-terminal half of medium T antigen. The mutation also lies within the deleted region of *dl*1013 and *dl*23 and near the deletions of *dl*1014 and *dl*1015, a series of mutants that generally reduce the transformation efficiency or the transformation phenotype in rat cells (10, 14). We do not know whether aspects of the defect in cell transformation of *ts*-48 may be related to the amino acid substitutions in medium T antigen.

Analysis of the conserved regions of the large T antigens of the three papovaviruses reveals further common and possibly important structural features. In the 79 amino acid residue

region of the large T antigen N-terminal to the proximal large T splice, none of the 13 amino acid differences between BK virus and SV40 occur at positions that are conserved in polyoma virus and SV40. The latter conserved residues therefore seem to be relatively important. Similarly, throughout the rest of the large T antigen molecules, most of the amino acid differences between SV40 and BK virus also occur at positions not conserved in polyoma virus and SV40, indicating that the conserved sites are important for protein function. Of the 22 amino acid differences between BK virus and SV40 at residues conserved in polyoma virus and SV40, 14 are functionally conservative, i.e., Lys → Arg, Asp → Glu, hydrophobic → hydrophobic. These findings suggest that most of the amino acid residues of the large T antigens play important and similar structure-function roles and are therefore conserved. There are two small regions that may be exceptions, one just N-terminal to the *ts*-48 mutation at nucleotide 1228, where several of the polyoma-SV40 conserved residues are altered in BK virus, and the other at the C-terminal tail in the vicinity of the *ts*-52 mutation.

Thomas and her colleagues have determined the nucleotide sequence alterations responsible for the *ts* phenotype in the prototypical group *A* mutant *ts-a* and in *ts*-25 (25). Their *ts*-25 shows a change identical to that described here, whereas *ts-a* contains a change at position 2193, corresponding to amino acid 547 in Fig. 1.

The mutations in four *tsA* mutants of SV40 have been identified by Seif et al. (23), and the positions of those mutations are indicated by asterisks in Fig. 1. All four fall in part of the sequence near that of polyoma mutants *ts*-48 and *ts-a*, as identified by Thomas et al. (25), suggesting particular importance for that portion of the molecule. It is of interest to note that, whereas the mutations in *ts*-48 and *ts*-25 cause changes to more hydrophobic amino acid residues, the mutation in *ts*-52 results in a change from glycine to serine, two residues with little or no difference in their hydrophobicity (15).

ACKNOWLEDGMENTS

We thank William Folk for allowing us to see his results before publication and Walter Eckhart, Tony Hunter, and Russell Doolittle for their helpful discussions. We also thank H. U. Bernard for the *ts*-52 whole-genome clone, Abby Esty for technical assistance, and Dorothy Miller for administrative help.

This work was supported by Public Health Service grant CA24288 from the National Cancer Institute, and P.L.D. was supported by Public Health Service training grant GM07199 from the National Institutes of Health.

LITERATURE CITED

1. **Clewell, D., and D. Helinski.** 1969. Supercoiled DNA-protein complex in *E. coli*. Proc. Natl. Acad. Sci. U.S.A. **62**:1159–1166.
2. **Deininger, P., A. Esty, P. LaPorte, and T. Friedmann.** 1979. Nucleotide sequence and genetic organization of the polyoma late region: features common to the polyoma early region and SV40. Cell **18**:771–779.
3. **Eckhart, W.** 1974. Properties of temperature sensitive mutants of polyoma virus. Cold Spring Harbor Symp. Quant. Biol. **39**:37–40.
4. **Enea, V., G. Vovis, and M. Zinder.** 1975. Genetic studies with heteroduplex DNA of bacteriophage F1. J. Mol. Biol. **96**:495–509.
5. **Fiers, W., R. Contreras, G. Haegeman, R. Rogiers, A. Van de Voorde, H. Van Heuverswyn, J. Van Herreweghe, G. Volckaert, and M. Ysebaert.** 1978. Complete nucleotide sequence of SV40 DNA. Nature (London) **273**:113–120.
6. **Friedmann, T., and D. Brown.** 1978. Base-specific reactions useful for DNA sequencing. Nucleic Acids Res. **5**:615–622.
7. **Friedmann, T., R. Doolittle, and G. Walter.** 1978. Amino acid sequence homology between polyoma and SV40 tumor antigens deduced from nucleotide sequences. Nature (London) **274**:291–293.
8. **Friedmann, T., A. Esty, P. LaPorte, and P. Deininger.** 1979. The nucleotide sequence and genome organization of the polyoma early region: extensive nucleotide and amino acid homology with SV40. Cell **17**:715–724.
9. **Friedmann, T., P. LaPorte, and A. Esty.** 1978. Nucleotide sequence studies of polyoma virus. J. Biol. Chem. **253**:6561–6567.
10. **Griffin, B., and C. Maddock.** 1979. New classes of viable deletion mutants in the early region of polyoma virus. J. Virol. **31**:645–656.
11. **Griffin, J., G. Spangler, and D. Livingston.** 1979. Protein kinase activity associated with simian virus 40 T antigen. Proc. Natl. Acad. Sci. U.S.A. **76**:2610–2614.
12. **Grunstein, M., and D. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. U.S.A. **72**:3961–3965.
13. **Kriegler, M., J. Griffin, and D. Livingston.** 1978. Phenotypic complementation of the SV40 TS A mutant defect in viral DNA synthesis following microinjection of SV40 antigen. Cell **14**:983–994.
14. **Magnusson, G., and P. Berg.** 1979. Construction and analysis of viable deletion mutants of polyoma virus. J. Virol. **32**:523–529.
15. **Manavalan, P., and P. Ponnuswamy.** 1978. Hydrophobic character of amino acid residues in globular proteins. Nature (London) **275**:673–674.
16. **Maniatis, T., A. Jeffrey, and D. Kleid.** 1975. Nucleotide sequence of the rightward operator of phage λ. Proc. Natl. Acad. Sci. U.S.A. **72**:1184–1188.
17. **Maxam, A., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74**:560–564.
18. **Miller, L., and M. Fried.** 1976. Construction of the genetic map of the polyoma genome. J. Virol. **18**:824–832.
19. **Reddy, V., B. Thimmappaya, R. Dhar, K. Subramanian, B. Zein, J. Pan, J. Celina, and S. Weissman.** 1978. The genome of SV40. Science **200**:494–502.
20. **Redloff, R., W. Bauer, and J. Vinograd.** 1967. A dye-buoyant density method for the detection and isolation of closed circular duplex DNA. Proc. Natl. Acad. Sci. U.S.A. **57**:1514–1521.
21. **Sanger, F., S. Nicklen, and A. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
22. **Seif, I., G. Khoury, and R. Dhar.** 1979. The genome of human papovavirus BKV. Cell **18**:963–977.
23. **Seif, I., G. Khoury, and R. Dhar.** 1980. A rapid enzy-

matic DNA sequencing technique. Nucleic Acids Res. **8**:2225–2240.

24. **Soeda, E., J. Arrand, N. Smolar, J. Walsh, and B. Griffin.** 1980. Coding potential and regulatory signals of the polyoma virus genome. Nature (London) **283**: 445–453.
25. **Thomas, T., P. Vollmer, and W. R. Folk.** 1980. Nucleotide sequence changes in polyoma virus *A* gene mutants. J. Virol. **37**:1094–1098.
26. **Tjian, R.** 1978. The binding site on SV40 DNA for a T antigen-related protein. Cell **13**:165–179.
27. **Tjian, R., and A. Robbins.** 1979. Enzymatic activities associated with purified simian virus 40 T antigen related protein. Proc. Natl. Acad. Sci. U.S.A. **76**:610–614.