NOTES

Mutation Affecting Late Gene Expression in Polyoma Virus Maps in the Late Region

ELAINE KINNEY THOMAS† AND J. DONALD HARE*

Department of Microbiology, University of Rochester Medical Center, Rochester, New York 14642

Received 15 December 1981/Accepted 21 December 1981

A mutation in polyoma virus strain 3049 which results in the overproduction of capsid proteins has been mapped to the late region of the genome between the *Hind*III site at 45.0 map units and the *Bam*HI site at 58.6 map units. This region contains the coding sequence for VP3 and a portion of VP2, but does not include the late promoters or the coding sequence for the late leaders. The possible role of VP2 or VP3 in the regulation of genetic expression in polyoma virus is discussed.

Several approaches are required to study the complex mechanisms by which eucaryotic organisms regulate the expression of genetic information. One such approach has been to study gene regulation in a simple genome, such as that of the smaller DNA viruses which depend entirely on the host for the enzymes that carry out transcription and post-transcriptional processing of mRNA. We have investigated the mechanism by which a strain of polyoma virus (3049) overproduces capsid protein giving rise to a visually recognizable phenotype, *cyc*, in which these proteins are overproduced and accumulate in the cytoplasm of infected cells (16, 27). Cells infected with this virus also display increased quantities of mRNA late in infection (23, 24) at a time when the replication of viral DNA and the quantity of viral transcription complexes extracted in Sarkosyl (4) are similar to those in cells infected with the wild-type virus (2).

To provide further insight into the mechanism of the altered expression of late genes, we have mapped the *cyc* mutation by constructing recombinant viral genomes containing specific restriction fragments derived from 3049 and a wild-type strain. The mutation has been mapped to a segment of the late region between the *Hind*III site at 45.0 and the *Bam*HI site at 58.6 map units which codes for VP2 and VP3. The function of these two virion proteins is as yet unclear, but they have been implicated in the regulation of late simian virus 40 transcription through an attenuation mechanism (20).

Virus stocks were prepared by infecting baby mouse kidney (BMK) cells with plaque-purified virus strains lpS, 3049, and lpD which have been

described previously (14, 15, 27). Strain A₂ (12, 26) was received from M. Fried. Viral DNA was prepared from BMK cell cultures infected for 48 to 72 h at a multiplicity of 1 PFU per cell. The DNA was extracted by the method of Hirt (17) and purified by equilibrium centrifugation in cesium chloride density gradients containing ethidium bromide (200 µg/ml) (22). Restriction endonucleases were from New England Biolabs or were the gift of G. A. Wilson and C. Duncan. Size markers were the replicative forms of φX174 (Bethesda Research Laboratories) treated with *Hpa*II and *Hae*III (25) and plasmid pMB9 (3) and polyoma strain A₂ (12) treated with *Hae*III.

To map the *cyc* mutation in strain 3049, a series of recombinant viruses composed of specific DNA fragments from the mutant and a second virus, which was wild type for this character, were constructed. To verify the actual DNA content of each recombinant, a virus strain was required which contained restriction fragments differing significantly from those of 3049. The restriction maps of two viruses, which were wild type for the *cyc* phenotype, i.e., lpS and lpD, were compared with 3049. The restriction sites in 3049 and lpS for the following enzymes were found to be identical: *Mbo*II, *Hind*III, *Sma*I, *Hinc*II, *Hha*I, *Xho*I, *Sal*I, *Pst*I, *Hpa*I, *Hae*III, *Kpn*I, *Alu*I, and *Sac*I. An insertion (0.9%) was found in the lpS genome mapping between 66.3 and 70.8 map units. The *Hpa*II patterns of 3049 and lpS, consisting of nine fragments, were similar to those found in strains p16 (8) and Csp (5). Both lacked the *Hpa*II fragment 1 of A₂ (13) owing to the presence of an extra *Hpa*II site at 39.6 map units, generating two smaller fragments.

† Present address: Genetic Systems Corporation, Seattle, WA 98121.

FIG. 1. Map of restriction sites of polyoma strain 3049 compared with A₂. The restriction sites in the A₂ strain for *Hpa*II, *Hae*III, and those enzymes indicated outside the circle (except for *Bgl*II) are from Griffin (12). The *Dpn*II sites were determined for the 3049 and lpD strains by the exonuclease III-mapping technique (28). The *Hae*III fragments of strain 3049 were mapped by secondary cleavage of *Hind*III and *Bam*HI-*Eco*RI fragments purified from 1.2% agarose gels by electroelution (G. A. Wilson, personal communication) and modified to remove agarose impurities by passage over DEAE-cellulose (29). The placement of many smaller fragments (13 through 21-2) is not precise as indicated by the question mark. The 3049 fragments 10-1 and 11-1 are in the 43.6% *Hind*III fragment. Fragment 7-1 contains the *Hind*III and *Hha*I sites, which in turn determine the coordinates of fragment 7-1.

The *Dpn*II fragments of 3049 and lpD were mapped by a procedure described by Thomas et al. (28) with exonuclease III. Viral DNA (13.5 μ g) was linearized with either *Eco*RI or *Bam*HI at 37°C in TMM buffer (6 mM MgCl₂-6 mM β -mercaptoethanol-6 mM Tris-hydrochloride [pH 7.5]). Exonuclease III (9 U; New England Biolabs) was added at time zero, and 4 U was added at 3- to 5-min intervals throughout the incubation period at 33°C. Samples (1.5 μ g) taken at time intervals from 0 to 23 min were dispensed directly into tubes preincubated at 80°C and held for 15 min. The samples were then cooled, treated with *Dpn*II, and separated on agarose gels. There was a sequential loss of fragments (data not shown) as a function of the time of exonuclease III digestion, which in turn defined the order of the fragments from the ends of the linearized molecule. Figure 1 shows the location of the *Dpn*II sites in 3049 and lpD viruses which were identical.

Strain lpD displayed a *Hae*III restriction pattern significantly different from that of 3049 (Fig. 2). This pattern was, however, identical to that of the A₂ strain. Because *Hind*III, *Eco*RI, and *Hpa*II sites in lpD DNA were also identical to those in A₂, we concluded that these viruses are either identical or at least closely related. Other important sites were also mapped and found to be similar in all virus strains. For instance, in

FIG. 2. *Hae*III restriction fragments of 3049 and lpD strains of polyoma virus. Viral DNA (5 μ g) was digested with *Hae*III in 50 μ l of TMM buffer (6 mM MgCl₂-6 mM β -mercaptoethanol-6 mM Tris-hydrochloride [pH 7.5]) and separated on an 8% acrylamide slab gel (40 cm by 20 cm by 3 mm) in modified TPE buffer (40 mM Tris-1 mM EDTA, adjusted to pH 7.7 with H₃PO₄) from 1,400 to 1,800 V-h. The lpD fragments are identified by numbers assigned to A₂ fragments by Griffin (12). The 3049 fragments having mobilities identical to those of the lpD strain were assigned the same arabic numeral. Unique 3049 fragments were assigned the arabic numeral of the nearest lpD fragment of larger size followed by a hyphen and a second numeral starting with 1. The largest 3049 fragment (I) was larger than the largest lpD fragment (1). The sizes of the *Hae*III fragments of 3049 were as follows: I, 13.8%; 1, 13.6%; 1-1, 13.1%; 3, 10.2%; 5, 9.8%; 7, 6.0%; 7-1, 5.3%; 8, 3.9%; 10-1, 3.1%; 11, 2.7%; 11-1, 2.4%; 12, 2.3%; 13, 1.9%; 15, 1.5%; 16, 1.4%; 17, 1.3%; 18, 1.0%; 18-1, 1.0%; 19, 0.9%; 20, 0.8%; 21, 0.7%; 21-1, 0.6%; and 21-2, 0.4%.

lpD, lpS, and 3049, the *Bam*HI site was at 58.6 map units (58.5 map units in A₂ [12]) and the *Bgl*II site was at 72.6 map units (72.2 map units in A₂ [26]).

Because the *Hae*III sites in 3049 were significantly different from those in lpD, the positions of several *Hae*III fragments of 3049 were determined by using standard mapping techniques, i.e., combined digestions and fragment isolation followed by subcleavage with *Hae*III, and by the exonuclease III technique described here. The 3049 DNA contained a *Hae*III fragment I which was larger than the lpD fragment 1. Cleavage with *Bgl*II reduced the size of fragment I but *Dpn*II did not (data not shown). These data and size constraints place fragment I in the proximal early region and suggest that this results from the fusion of A₂ fragments 2 and 18', both of which are missing from the 3049 *Hae*III pattern. The *Hae*III fragments around the *Bam*HI site were also mapped by treating *Bam*HI-linearized DNA with exonuclease III at 26°C for up to 6 min, followed by *Hae*III digestion in the presence of 3 mM NaCl. The size of the *Hae*III fragments is summarized in Fig. 2. Several fragments unique to 3049 are identified which provide markers distributed around the genome which can be used to confirm the structure of recombinants made up of lpD and 3049 DNA.

Three sets of reciprocal lpD-3049 recombinant viruses were constructed, and single-plaque progeny were produced. *Hae*III restriction anal-

ysis of the recombinants established the identity of the DNA contributed by lpD and 3049 viruses. Table 1 lists the number of plaques identified on the basis of *Hae*III patterns and the phenotype of each group. Recombinant viruses Rec 1a and Rec 1b are subsets of the Rec 2a and Rec 2b viruses and thus are not critical to the mapping of the *cyc* mutation. They do, however, provide evidence for the location of the unique *Hae*III fragments 18-1, 21-1, and 21-2 in the *Bam*HI-*Bgl*II segment of 3049. In general, the entire set of recombinants provided mapping data that was totally consistent with the assignment of restriction sites in both 3049 and lpD DNA.

The phenotype of each pair of recombinants was then determined several times in a double-blind mode by using the indirect fluorescent-antibody technique. The reciprocal recombinants within each group consistently displayed the opposite phenotype (Fig. 3), whereas separate plaque isolates of each identical recombinant displayed the same phenotype. The genetic composition of the recombinants and each phenotype is shown in Fig. 4. The 3049 DNA sequence between the *Hind*III site at 45.0 map units and the *Bam*HI site at 58.6 map units and between the *Eco*RI site at 0 map unit and the *Hind*III site at 1.4 map units correlated with the *cyc* phenotype. It was felt that the region between 0 and 1.4 map units was unlikely to be involved in the *cyc* phenotype. To evaluate this region of the 3049 genome, the DNA between

TABLE 1. Infectious ligation products and number of plaques identified from in vitro ligation experiments using lpD and 3049 DNA

Ligation expt no.	Endonucleases used	Sizes of fragments (%) ^a	Infectious products identified ^b	No. of plaques identified	Source of fragments in infectious products						Phenotype
					14.0%	86.0%	41.4%	58.6%	43.6%	56.4%	
1	<i>Bam</i> HI, <i>Bgl</i> II	14.0, 86.0	lpD	4	lpD	lpD					wt ^c
			3049	4	3049	3049					<i>cyc</i> ^d
			Rec 1a ^e	1	3049	lpD					wt
			Rec 1b	1	lpD	3049					<i>cyc</i>
			Mixed	2							
2	<i>Eco</i> RI, <i>Bam</i> HI	41.4, 58.6	lpD	6			lpD	lpD			wt
			3049	8			3049	3049			<i>cyc</i>
			Rec 2a	5			3049	lpD			wt
			Rec 2b	2			lpD	3049			<i>cyc</i>
			Mixed	1							
3	<i>Hind</i> III	43.6, 56.4	lpD	9					lpD	lpD	wt
			3049	15					3049	3049	<i>cyc</i>
			Rec 3a	3					3049	lpD	wt
			Rec 3b	2					lpD	3049	<i>cyc</i>
			Mixed	2							

^a Expressed as percentage of full-length DNA.

^b Identified by *Hae*III-cleavage pattern.

^c wt, Wild-type phenotype; capsid proteins accumulate only in nucleus.

^d *cyc*, Accumulation of capsid protein in cytoplasm and nucleus.

^e Rec, Recombinants.

FIG. 3. Indirect fluorescent-antibody staining of Rec 3a- and Rec 3b-infected cells. Recombinant viruses were prepared by mixing equal quantities (2.5 μ g) of DNA from 3049 and lpD strains, after which the DNA was treated with the enzymes reported in Table 1. The endonucleases were inactivated at 65°C for 5 min, and the products were ligated with T4 ligase (Bethesda Research Laboratories) in 20 mM Tris-hydrochloride (pH 7.6)–10 mM $MgCl_2$ –10 mM dithiothreitol–0.2 mM ATP (ATP was freshly prepared) for 14 h at 15°C. Mouse embryo cells were transfected with ligated DNA (11) and overlaid with agar medium for plaque formation. Individual plaques were picked after 10 days and amplified in BMK cells. Viral DNA for restriction enzyme analysis was prepared by the Hirt procedure (17), which was modified to allow the contents of one large (100 mm) plastic dish to be processed in microcentrifuge tubes. BMK cells infected with representative plaques as shown in Table 1 were stained by the indirect fluorescent-antibody technique (16) by using rabbit anti-whole virus or anti-disaggregated capsid serum and fluorescein-labeled goat anti-rabbit gamma globulin serum (Cappel Laboratories). (A) Rec 3b infection displaying the *cyc* phenotype. (B) Rec 3a infection displaying the wild-type phenotype.

the *EcoRI* and *HindIII* sites was sequenced by the dideoxynucleoside triphosphate technique of Maat and Smith (21). The sequence of the 3049 DNA matched exactly that reported for the A₂

strain (26), a virus which displays the wild-type phenotype. Thus, it is concluded that the *cyc* mutation maps in the *HindIII*–*BamHI* segment of 3049 DNA.

FIG. 4. Composite diagram of genetic regions common to recombinant viruses demonstrating *cyc* or wild-type phenotype. (a) The wild-type phenotype is expressed when recombinant viruses contain lpD DNA from map unit 0 to 1.4 in the early region and from map unit 45.0 to 58.6 in the late region. (b) The *cyc* phenotype is expressed when recombinant viruses contain 3049 DNA from map unit 0 to 1.4 in the early region and from map unit 45.0 to 58.6 in the late region. The early region was ruled out by demonstrating that the sequence of the DNA between 0 and 1.4 map units was identical to that in a wild-type virus, A₂. (—) lpD DNA; (---) 3049 DNA.

This portion of the late region contains the coding sequence for most of VP3 and that part of VP2 shared with VP3 and for the amino-terminal portion of VP1 and the 3' end of the VP1 mRNA intron. These details of genetic structure have been reviewed recently by Soeda et al. (26). It does not contain the late promoters, the 5' ends of the late mRNAs (26), or the reiterated leader sequences (18).

Several other polyoma mutations occur in this portion of the late region which also display altered expression of the viral genome. The first, *ts3*, maps in *Hpa*II fragment 3 (7) and the second, *ts59*, in *Hpa*II fragment 1 (10). The *ts3* is a D-type mutant, the effect of which can be bypassed by infection with purified DNA (6), suggesting that there exists a repressor which in *ts3* binds irreversibly to the origin of DNA replication at the nonpermissive temperature, thereby interfering with the initiation of DNA synthesis (20). The *Hpa*II fragment 3 contains the coding sequences for the amino-terminal ends of both VP2 and VP3 (26). The *ts59* mutation leads to reduced late RNA and protein synthesis at the nonpermissive temperature (10), an effect which is the reverse of that seen in the 3049 mutation. The *Hpa*II fragment 1 contains the coding sequences for the carboxy-terminal ends of VP2 and VP3 and the bulk of the VP1 coding sequences (26).

Recently, Llopis and Stark (20) have studied

several deletion mutants of simian virus 40 which map in the VP2 and VP3 proteins. These mutants display an altered transcriptional activity compared with the wild-type virus when transcriptionally active DNA complexes are isolated under isotonic conditions. They proposed (20) that VP2 or VP3, or both, are regulatory proteins whose function is to modulate late transcription through an attenuation mechanism.

At this point, we have no conclusive evidence to distinguish between a transcriptional or post-transcriptional model for the expression of the *cyc* mutation. Most of our data is compatible with either model. For instance, increased late protein (27) and mRNA (23, 24) levels, and the coordinate step-up in the synthesis of both VP1 and VP3 polypeptides (B. J. Pomerantz and J. D. Hare, unpublished data) are compatible with either mechanism. Blackburn and Hare (2) have shown recently that the number of viral transcription complexes extracted by Sarkosyl is identical in 3049 and wild-type virus infection. Since Sarkosyl extraction strips most proteins except the RNA polymerase from viral chromatin (9), the findings fail to rule out the possibility of an altered transcription-attenuation system in 3049.

Preliminary studies (Pomerantz and Hare, unpublished data) have found no significant difference between 3049 and lpS or lpD in the amount of either total nuclear viral RNA or in large, nuclear viral RNA transcripts. These findings, coupled with the fact that the late promoters do not map in the same region as *cyc*, comprise the only evidence to date which would be incompatible with a transcriptional initiation mechanism. A mechanism which operates through an altered attenuation of transcription is compatible with the data, however.

We are currently examining native viral transcription complexes to detect possible differences in the activation of transcription in 3049 and wild-type chromatin to evaluate the attenuation model. The availability of two other late mutants which map in the same region promise to provide invaluable insight into the mechanism by which polyoma virus regulates late expression, possibly through the function of two minor late gene products, VP2 and VP3. It is of interest (19) that these proteins are the products of the intron which is removed from the mRNA for the major late protein, VP1, and may be an example of an intron-coded protein which regulates the expression of the protein coded for by the mRNA from which it has been removed.

We thank G. Wilson and F. E. Young (University of Rochester) and R. Rothstein and R. Wu (Cornell University) for their generous gift of several restriction enzymes. We also express gratitude to M. Fried for communicating the *Mbo*I

restriction sites before they were published and to S. Zain for the use of her DNA-sequencing facility and expertise. We also thank G. G. Carmichael for suggestions concerning the preparation of this manuscript.

This work was initially supported by Public Health Service grant CA-07239 from the National Institutes of Health and, more recently, by the M. A. Hare Basic Research Fund.

LITERATURE CITED

1. Betts, R. F., T. G. Tachovsky, and J. D. Hare. 1972. Studies on the mechanism of cytoplasmic antigen accumulation following infection with a new variant of polyoma virus. *J. Gen. Virol.* **16**:29–38.
2. Blackburn, C., and J. D. Hare. 1981. DNA synthesis and template activity in a mutant of polyoma virus with altered expression of late viral proteins. *J. Virol.* **40**:142–149.
3. Bolivar, F., R. G. Rodriguez, M. C. Betlach, and H. W. Boyer. 1977. Construction and characterization of new cloning vehicles. I. Ampicillin-resistant derivatives of the plasmid pMB9. *Gene* **2**:75–93.
4. Condit, R. C., A. Cowie, R. Kamen, and F. Birg. 1977. Polyoma virus transcription *in vitro*. *J. Mol. Biol.* **116**:215–235.
5. Crawford, L. V., A. K. Robbins, P. M. Nicklin, and K. Osborn. 1974. Polyoma DNA replication: location of the origin in different virus strains. *Cold Spring Harbor Symp. Quant. Biol.* **39**:219–225.
6. Eckhart, W., and R. Dulbecco. 1974. Properties of the ts3 mutant of polyoma virus during lytic infection. *Virology* **60**:359–369.
7. Feunteun, J., L. Sompayrac, M. Fluck, and T. Benjamin. 1976. Localization of gene functions in polyoma virus DNA. *Proc. Natl. Acad. Sci. U.S.A.* **73**:4169–4173.
8. Fried, M., B. Griffin, E. Lund, and D. L. Roberson. 1974. Polyoma virus—a study of wild type, mutant and defective DNAs. *Cold Spring Harbor Symp. Quant. Biol.* **39**:45–52.
9. Gariglio, P., and S. Mousset. 1975. Isolation and partial characterization of a nuclear RNA polymerase-SV40 DNA complex. *FEBS Lett.* **56**:149–155.
10. Gibson, W., T. Hunter, B. Cogen, and W. Eckhart. 1977. Altered virion proteins of a temperature-sensitive mutant of polyoma virus, ts59. *Virology* **80**:21–41.
11. Graham, F. L., and A. J. Van Der Eb. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**:456–467.
12. Griffin, B. E. 1977. Fine structure of polyoma virus DNA. *J. Mol. Biol.* **117**:447–471.
13. Griffin, B. E., M. Fried, and A. Cowie. 1974. Polyoma DNA: a physical map. *Proc. Natl. Acad. Sci. U.S.A.* **71**:2077–2081.
14. Hare, J. D. 1964. Transplant immunity to polyoma virus induced tumors. I. Correlations with biological properties of virus strains. *Proc. Soc. Exp. Biol. Med.* **115**:805–810.
15. Hare, J. D. 1967. Transplant immunity to polyoma-induced tumor cells. IV. A polyoma strain defective in transplant antigen induction. *Virology* **31**:625–632.
16. Hare, J. D. 1970. A new type of variation among the polyoma viruses characterized by the cytoplasmic accumulation of capsid antigen. *Virology* **40**:978–988.
17. Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26**:365–369.
18. Legon, S., A. J. Flavell, A. Cowie, and R. Kamen. 1979. Amplification in the leader sequence of late polyoma mRNAs. *Cell* **16**:373–388.
19. Lewin, B. 1980. Alternatives to splicing: an intron-coded protein. *Cell* **22**:645–646.
20. Llopis, R., and G. R. Stark. 1981. Two deletions within genes for simian virus 40 structural proteins VP2 and VP3 lead to formation of abnormal transcriptional complexes. *J. Virol.* **38**:91–103.
21. Maat, J., and J. H. Smith. 1978. A method for sequencing restriction fragments with dideoxynucleoside triphosphates. *Nucleic Acids Res.* **5**:4537–4545.
22. Radloff, R., W. Bauer, and T. Vinograd. 1967. A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: the closed circular DNA in HeLa cells. *Proc. Natl. Acad. Sci. U.S.A.* **57**:1514–1521.
23. Rutherford, R. B., and J. D. Hare. 1974. Evidence for a regulatory function related to the expression of the polyoma genome following the onset of virus DNA replication. *Biochem. Biophys. Res. Commun.* **58**:839–846.
24. Rutherford, R. B., and J. D. Hare. 1975. The synthesis and turnover of virus-specific polyadenylated RNA in polyoma-infected cells. *Biochem. Biophys. Res. Commun.* **62**:789–797.
25. Sanger, F., G. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, J. C. Fiddes, C. A. Hutchinson III, P. M. Slocumbe, and M. Smith. 1977. Nucleotide sequence of bacteriophage ϕ X174 DNA. *Nature (London)* **265**:687–695.
26. 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.
27. Tachovsky, T. G., and J. D. Hare. 1975. Polyoma virus strain with enhanced synthesis of capsid protein. *J. Virol.* **16**:116–122.
28. Thomas, E. K., E. M. Hailparn, G. A. Wilson, and F. E. Young. 1976. Physical mapping of bacteriophage DNA by exonuclease III and endodeoxyribonuclease *Bam*HI, p. 605–610. *In* D. P. Nierlich, W. J. Rutter, and D. F. Fox (ed.), *Molecular mechanisms in the control of gene expression*. Academic Press, Inc., New York.
29. Zain, B. S., and R. J. Roberts. 1979. Sequences from the beginning of the fiber mRNA of adenovirus 2. *J. Mol. Biol.* **131**:341–352.