Isolation of a Polyomavirus with an Insertion of Foreign DNA in the Early Gene Promoter Region

KAREN L. CLARK, MARY M. BENDIG,[†] AND WILLIAM R. FOLK^{‡*}

Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109

Received 7 June 1984/Accepted 2 September 1984

We have isolated a polyomavirus with 134 base pairs of foreign DNA between the origin of replication and the early promoter. The insertion reduces the infectivity of the virus by interfering with events required for the initiation of infection. mRNA transcripts from the early region exhibit a marked heterogeneity of 5' termini.

Upstream of the initiation codon for the polyomavirus early proteins, sequences occur that are important for the initiation of transcription and for its regulation by large T-antigen (2, 14, 19, 22, 25, 34). We have previously reported the construction of polyomavirus DNAs containing deletions near this region (1, 2). Viral DNA (strain A3) was linearized by digestion with *Bgl*II, treated with S1 nuclease, and then transfected into mouse cells. Of several hundred viable plaque-forming mutants analyzed, one was found to have gained, rather than lost, DNA sequences. Analysis of this mutant by DNA sequencing revealed that a deletion of 3 nucleotides (nt) and an insertion of 134 nt had occurred at the *Bgl*II site (Fig. 1).

The DNA insertion in mutant 44 has no obvious homology with any region of the polyoma genome. Using a single-stranded cloned DNA containing the insert, we probed mouse genomic DNA blots at normal and low stringencies (T_m , -42°C, [21]). We were able to detect several weakly hybridizing bands of ca. 1.3 kilobases in *Bam*HI and *Eco*RI digests. DNA reassociation analyses with mouse genomic DNA as driver indicated that the insert is not complementary to highly repeated mouse sequences. Although neither result conclusively demonstrates that the insert is derived from the mouse genome, it is the most likely source. Ding et al. (9) have described another deletion-insertion mutant of polyomavirus; they also suggest that the source of its insert is the mouse genome. These viruses probably acquired cellular sequences by recombination during lytic infection (39). In the case of mutant 44, recombination between the viral genome and the mouse genome may have been promoted by linearization of the viral DNA.

Mutant 44 appears to grow more slowly than wild-type virus in mouse cells. Plaques produced by mutant 44 are somewhat smaller than wild-type plaques. A growth curve indicates that the yield of infectious mutant 44 virus is ca. 1/10 that of wild-type polyomavirus (Fig. 2a). However, the numbers of virus particles (measured by hemagglutination [20]) produced by cells infected with mutant or wild-type virus are approximately the same, when adjusted for the input multiplicity (Table 1). Similarly, equivalent amounts of intracellular viral DNA are found in cells infected by either of the two viruses (Table 1). In plaque assays, mutant 44 viruses are 1/15 to 1/20 as infectious as wild-type polyoma-

viruses (Table 1). Restriction analysis of the viral DNAs used for these infections indicates a low proportion of defective viral DNAs with genomic rearrangements. Thus, it appears that cells infected by mutant 44 support as much viral DNA replication and produce as many viruses as wild-type virus-infected cells, but the mutant 44 virus is only 1/15 to 1/20 as infectious as wild-type virus (Table 1). This could be due to one of several reasons: (i) mutant 44 DNA is intrinsically less infectious than wild type DNA, (ii) mutant 44 viruses are less stable or contain less infectious DNA than wild-type viruses, or (iii) mutant 44 viruses are less efficient at initiating a productive infection. The results of the following experiments support the third alternative.

The specific infectivity of DNA isolated from virus was measured by plaque assay with DEAE-dextran transfection (12). Equivalent infectivities were observed with mutant 44 and wild-type DNAs (ca. 10^5 PFU/ μ g). Similarly, DNAs excised from pBR322-polyoma clones of mutant 44 and wild type were equally infectious (3×10^4 PFU/ μ g). Thus, the reduced infectivity of mutant 44 is not observed when the normal processes of virion adsorption and penetration are bypassed.

We measured the amounts of viral DNA extracted from mutant 44 and wild-type viruses and determined them to be equivalent (Table 1). Additionally, DNAs in both types of viruses were equally resistant to DNase treatment (Table 1), and the viruses were equally resistant to heating at 60°C (Fig. 2b). This suggests that the mutant 44 viruses have the normal complement of DNA and are quite stable. Therefore, the dramatically reduced infectivity of mutant 44 virus is most likely attributable to a defect in a step during the process of infection. The 131-base-pair (bp) net increase in genome size of mutant 44 is the largest insertion in a viable mutant described to date and may approach the upper size limit of the polyoma genome which can be encapsidated. We speculate that the interaction between the larger genome of mutant 44 and its capsid proteins reduces its ability to be recognized as a true virion, perhaps at the level of uptake into the cell (4, 28) or during the organization of the intracellular viral genome into chromatin.

The inserted DNA in mutant 44 occurs in a region of the polyoma genome which is known to be important for early gene transcription. The principal early mRNA's of wild-type polyomavirus are initiated ca. 24 to 30 nt downstream of a TATA box sequence (19, 25). The 131-bp insertion occurs 25 nt upstream of this sequence, and as such, separates it from potential upstream regulatory signals such as the CAAT box (26) and the enhancer element (8) (Fig. 1).

* Corresponding author.

[†] Present address: Celltech Ltd., Slough SL1 4DX, United Kingdom.

[‡] Present address: Department of Microbiology, University of Texas at Austin, Austin, TX 78712.

FIG. 1. Early noncoding region of wild-type and mutant 44 polyomaviruses. The uppermost line is a schematic of the noncoding early region of the polyoma genome. Underlining arrows indicate candidates for large T-antigen binding sites (14, 25, 34). Arrows above the schematic indicate the primary early transcription start sites (19, 25). Early and late genes are transcribed in the directions indicated. The PyA3 sequence is the same as that determined by Deininger et al. (7) beginning at nucleotide 76. The DNA insert in mutant 44 was sequenced by the procedure of Maxam and Gilbert (29) with a pBR322-mutant 44 clone and by the dideoxy chain termination technique (37) with an M13 clone (31) containing the mutant *HpaII* -5 fragment. The boxed nucleotides in the PyA3 and the Py44 sequences are the CAAT and the TATA motifs (25). The three underlined bases are deleted in the mutant 44 genome. The insert is indicated by wavy underlining. *BglI* refers to the sequence recognized and cleaved by *BglI* (indicated by a vertical arrow).

FIG. 2. Comparison of mutant 44 and wild-type polyomaviruses. (A) Growth curves. Mouse cells were infected with mutant 44 or wild-type polyomavirus at similar multiplicities. At the indicated times, virus was harvested by repeated freeze-thawing and titered by plaque assay at 37°C. Symbols: ○, wild type; and ●, mutant 44. (B) Thermal inactivation. Virus diluted to about 10⁷ PFU/ml was heated at 60°C for the indicated times, and the remaining infectivity was determined by plaque assay. In this experiment, the mutant 44 plaques were somewhat smaller than the wild-type plaques. Symbols are as in 2A.

We mapped the 5' termini of early transcripts isolated 14 h postinfection by primer extension to determine whether the insert affects transcription initiation. The major termini determined by primer extension analysis were confirmed by S1 nuclease analysis of similar RNAs isolated at 24 h postinfection. The principal 5' termini of mutant 44 RNAs coincide with those of wild-type virus (Fig. 3, lanes 2 and 6). In addition to these major 5' termini around nt 165, mutant 44 has several minor upstream 5' termini within and beyond the insert. Wild-type polyoma also has some upstream 5' termini but not of the abundance or frequency that mutant 44 has. RNAs of both viruses appear to exhibit 5' termini in the region of the TATA box at about nt 140 (a band of ca. 240 bp in Fig. 3). Mutant 44 has additional unique 5' termini within its insert, particularly near each junction (Fig. 3, at ca. 400 and 270 bp). Two prominent alternate upstream 5' termini are common to both viruses. One is located around nt 30 in the polyoma genome (Fig. 3, at ca. 370 bp in lane 2 and ca. 500 bp in lane 6) and another is about 20 bp beyond this site, near a TATA-like sequence in the viral DNA. This region may have secondary structure which could produce these bands in the primer extension analysis, but Cowie et al. (5) have demonstrated that RNAs which initiate downstream from this TATA-like sequence are capped. Bands on the late side of the origin are observed in RNAs of both viruses (Fig. 3). Because late transcripts initiate in this area of the

TABLE 1. Comparison of mutant 44 and wild-type polyomavirus infections

Virus	Initial MOI (PFU/cell) ^a	Virus yield ^b		Amt of DNA in cells ^c	Amt of DNA in viruses ^d	
		Hemagglutination assay (HAU)	Plaque assay (PFU/ml)		Without DNase	With DNase
Mutant 44	0.26	54	6.7 × 10 ⁴	0.53	0.83	0.41
Wild-type	1.0	194	5.0 × 10 ⁶	3.54	3.40	1.96
Ratio (wild type/ mutant 44)	4.0	3.6	75.4	6.7	4.1	4.8

^a MOI, Multiplicity of infection. Duplicate plates of mouse cells were infected as previously described (12), and 72 h later DNA was isolated by the Hirt extraction procedure (20) or the cells and media were pooled and subjected to three cycles of freeze-thawing.
^b Virus was released from infected cells by repeated freeze-thawing and then titered by hemagglutination (6) and by plaque assay (12). HAU, Hemagglutination assay units.
^c *HpaII*-digested DNA from a Hirt extract was separated on a 1.5% agarose gel, blotted onto nitrocellulose (39), and hybridized with a nick-translated (36) pBR322-polyoma clone. A set of serial dilutions of polyoma DNA was dot-blotted (24) and included in the same hybridization. Viral DNA was quantitated by densitometry of the *HpaII* 4 fragment. The amounts are expressed in arbitrary units relating the amount of hybridizable DNA to the grain density of the autoradiogram.
^d Virus-associated DNA was isolated by phenol-chloroform extraction of virus (in some cases after treatment with DNase I [5 µg/ml] for 30 min at 37°C). The DNA was dot-blotted (24) and hybridized with a single-stranded probe (35) of polyoma sequences from the *BclI* site through the origin to the *EcoRI* site. Quantitation was as described for intracellular DNA.

genome, these bands do not definitively locate 5' termini, but they do indicate that transcripts of at least that length are present.

Late in infection, a shift in the initiation sites of early gene transcription has been observed with simian virus 40 (SV40) (15, 18) and has been suggested for polyomavirus (11). To assay for such a shift, we performed primer extension analysis on RNAs isolated 44 h postinfection. The abundance of the major 5' termini around nt 165 decreases relative to the upstream 5' termini late in wild-type polyoma virus infection (Fig. 3, lanes 2 and 3). A similar shift is seen with mutant 44 virus (Fig. 3, lanes 6 and 7). For RNAs of both viruses, those bands seen 14 h postinfection are also

present 44 h postinfection, although they have different relative intensities. Unique to mutant 44 RNA is the presence of multiple termini between the primary sites at nt 165 and the upstream sites around nt 25 (Fig. 3, lane 7). However, we do not know whether the new 5' termini in any of these viral RNAs represent genuine initiation sites or reflect cleavage of longer RNAs.

The pattern of heterogeneous termini is similar to that observed with the polyomavirus late promoter, which lacks a TATA box (23). Deletion of the early promoter TATA box, nearby sequences, or both also induces heterogeneous 5' termini and results in reduced transcription efficiency in vitro (22, 23, 25; W. Farmerie and W. Folk, unpublished data). The heterogeneous pattern of 5' termini in mutant 44 may indicate a similarly weak promoter.

The polyomavirus and SV40 early promoters differ in sequence and topology. Like many other RNA polymerase II promoters, the SV40 promoter has a GC-rich A domain (the 21-bp repeats) important for transcription efficiency 50 to 70 nt upstream of the cap sites (for a review, see reference 3). The spatial arrangement of this domain is important for promoter efficiency, as insertion of greater than 120 bp of *Escherichia coli* DNA between the A domain and the TATA box greatly reduces early gene expression (32). Furthermore, the SV40 TATA box is contained within sequences required for the initiation of DNA replication (16, 33, 38). In contrast, the polyomavirus TATA box is separated from the origin of DNA replication by approximately 150 bp, and the intermediary sequences have no distinguishable GC-rich A domain. As these two viruses are closely related, the differences between their early promoters are striking. Possibly, the polyomavirus promoter evolved from a more compact SV40-like form by the acquisition of foreign sequences, by a recombinational event similar to that which occurred with mutant 44. Alternatively, it may have formed by a process of duplication of the origin sequences (13, 17) and subsequent divergence.

We thank Bill Farmerie for advice on the transcription studies, Thesia Thomas for performing the early sequence analyses of mutant 44, Sarah Scanlon for help with some of the replication experiments, and Judy Birk for preparing mouse cells.

Support was provided by Public Health Service grant CA 13978, by grant MV 132 from the American Cancer Society, and by training grant NIH5-T32 GM 0744 to K.L.C.

LITERATURE CITED

1. Bendig, M. M., and W. R. Folk. 1979. Deletion mutants of polyoma virus defining a nonessential region between the origin of replication and the initiation codon for early proteins. *J. Virol.* 32:530-535.
2. 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.
3. Bensimhon, M., J. Gabarro-Arpa, R. Ehrlich, and C. Reiss. 1983. Physical characteristics in eukaryotic promoters. *Nucleic Acids Res.* 11:4521-4540.
4. Bolen, J. B., and R. A. Consigli. 1979. Differential adsorption of polyoma virions and capsids to mouse kidney cells and guinea pig erythrocytes. *J. Virol.* 32:679-683.
5. Cowie, A., P. Jat, and R. Kamen. 1982. Determination of sequences at the capped 5' ends of polyoma virus early region transcripts synthesized in vivo and in vitro demonstrates an unusual microheterogeneity. *J. Mol. Biol.* 159:225-255.
6. Crawford, L. V. 1962. The adsorption of polyoma virus. *Virology* 18:177-181.
7. Deininger, P. L., A. Esty, P. LaPorte, H. Hsu, and T. Friedmann. 1980. The nucleotide sequence and restriction enzyme sites of the polyoma genome. *Nucleic Acids Res.* 8:855-860.

FIG. 3. Primer extension mapping of early transcripts. Polyadenylic acid-containing RNA was isolated from RNA harvested as described (10) at 14 or 44 h postinfection. Polyadenylic acid-containing RNA from 14 h postinfection (1 μ g) or from 44 h postinfection (0.5 μ g) was hybridized with an excess of labeled primer, extending from the *AccI* site at nt 383 (the 5' end) to the *DdeI* site at nt 200, which was made single stranded by hybridization of its complement to the plus strand of an M13mp8 genome containing the corresponding polyoma sequence. The primer used in lanes 3, 5, 6, and 7 was 5' labeled with polynucleotide kinase (29), and the primer used in lane 2 was 3' labeled by filling in the end with T4 DNA polymerase. The hybridized primers were then extended by the procedure described by McKnight and Kingsbury (30). The nucleic acids were precipitated and fractionated on a 6% polyacrylamide-8 M urea gel, and the gel was autoradiographed with flashed film (27). Lane 1, *HpaII*-digested pBR322 (marker for lane 2); lane 2, wild type at 14 h postinfection; lane 3, wild type at 44 h postinfection; lane 4, *HpaII*-digested pBR322 (marker for lane 3); lane 5, RNA without polyadenylic acid; lane 6, mutant 44 at 14 h postinfection; lane 7, mutant 44 at 44 h postinfection; and lane 8, *HpaII*-digested pBR322 (marker for lanes 6 and 7). Schematics represent the genomes of wild-type and mutant 44 polyomaviruses. The hatched box indicates the DNA insertion.

8. deVilliers, J., and W. Schaffner. 1981. A small segment of polyoma virus DNA enhances the expression of a cloned β -globin gene over a distance of 3400 bp. *Nucleic Acids Res.* **9**:6251–6264.
9. Ding, D., M. D. Jones, A. Leigh-Brown, and B. E. Griffin. 1982. Mutant din-21, a variant of polyoma virus containing a mouse DNA sequence in the viral genome. *EMBO J.* **1**:461–466.
10. Favaloro, J., R. Treisman, and R. Kamen. 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimensional nuclease S1 mapping. *Methods Enzymol.* **65**:718–749.
11. Fenton, R. G., and C. Basilico. 1982. Changes in the topography of early region transcription during polyoma virus lytic infection. *Proc. Natl. Acad. Sci. U.S.A.* **79**:7142–7146.
12. Folk, W. R. 1973. Induction of virus synthesis in polyoma-transformed BHK-21 cells. *J. Virol.* **11**:424–431.
13. Folk, W. R., and B. R. Fishel. 1975. Tandem repetition of the origin of DNA replication in defective polyoma virus DNAs. *Virology* **64**:447–453.
14. Gaudray, P., C. Tyndall, R. Kamen, and F. Cuzin. 1981. The high-affinity binding site on polyoma virus DNA for the viral large-T protein. *Nucleic Acids Res.* **9**:5697–5710.
15. Ghosh, P. K., and P. Lebowitz. 1981. Simian virus 40 early mRNA's contain multiple 5' termini upstream and downstream from a Hogness-Goldberg sequence; a shift in 5' termini during the lytic cycle is mediated by large T antigen. *J. Virol.* **40**:224–240.
16. Gluzman, Y., R. Frisque, and J. Sambrook. 1980. Origin-defective mutants of SV40. *Cold Spring Harbor Symp. Quant. Biol.* **44**:293–300.
17. Griffin, B. E., and M. Fried. 1975. Amplification of a specific region of the polyoma virus genome. *Nature (London)* **256**:175–179.
18. Hansen, U., D. G. Tenen, D. M. Livingston, and P. A. Sharp. 1981. T antigen repression of SV40 early transcription from two promoters. *Cell* **27**:603–612.
19. Heiser, W. C., and W. Eckhart. 1982. Polyoma virus early and late mRNA's in productively infected mouse 3T6 cells. *J. Virol.* **44**:175–188.
20. Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26**:365–369.
21. Howley, P., M. Israel, M.-F. Law, and M. Martin. 1979. Rapid method for detecting and mapping homology between heterologous DNAs. *J. Biol. Chem.* **254**:4876–4883.
22. Jat, P., U. Novak, A. Cowie, C. Tyndall, and R. Kamen. 1982. DNA sequences required for specific and efficient initiation of transcription at the polyoma virus early promoter. *Mol. Cell. Biol.* **2**:737–751.
23. Jat, P., J. Roberts, A. Cowie, and R. Kamen. 1982. Comparison of the polyoma virus early and late promoters by transcription in vitro. *Nucleic Acids Res.* **10**:871–887.
24. Kafatos, F. C., C. W. Jones, and A. Efstratiadis. 1979. Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. *Nucleic Acids Res.* **7**:1541–1552.
25. Kamen, R., P. Jat, R. Treisman, J. Favaloro, and W. Folk. 1982. 5' termini of polyoma virus early region transcripts synthesized in vivo by wild-type virus and viable deletion mutants. *J. Mol. Biol.* **159**:189–224.
26. Katinka, M., and M. Yaniv. 1982. Deletions of N-terminal sequences of polyoma virus T-antigens reduce but do not abolish transformation of rat fibroblasts. *Mol. Cell. Biol.* **2**:1238–1246.
27. Laskey, R. A., and A. D. Mills. 1975. Quantitative film detection of ^3H and ^{14}C in polyacrylamide gels by fluorography. *Eur. J. Biochem.* **56**:335–341.
28. Mackay, R. L., and R. A. Consigli. 1976. Early events in polyoma virus infection: attachment, penetration, and nuclear entry. *J. Virol.* **19**:620–636.
29. Maxam, M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499–559.
30. McKnight, S. L., and R. Kingsbury. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. *Science* **217**:316–324.
31. Messing, J., and J. Vieira. 1982. A new pair of M13 vectors for selecting either strand of double-digest restriction fragments. *Gene* **19**:269–276.
32. Moreau, P., R. Hen, B. Wasyluk, R. Everett, M. P. Gaub, and P. Chambon. 1981. The SV40 72 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic Acids Res.* **9**:6047–6067.
33. Myers, R. M., and R. Tjian. 1980. Construction and analysis of SV40 origins defective in tumor antigen binding and DNA replication. *Proc. Natl. Acad. Sci. U.S.A.* **77**:6491–6495.
34. Pomerantz, B. J., C. R. Mueller, and J. A. Hassell. 1983. Polyoma large T antigen binds independently to multiple, unique regions on the viral genome. *J. Virol.* **47**:600–610.
35. Ricca, G. A., J. M. Taylor, and J. E. Kalinyak. 1982. Simple rapid method for the synthesis of radioactively labeled cDNA hybridization probes utilizing bacteriophage M13mp7. *Proc. Natl. Acad. Sci. U.S.A.* **74**:724–728.
36. Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* **113**:237–251.
37. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**:5463–5467.
38. Shortle, D. R., and D. Nathans. 1979. Regulatory mutants of SV40 constructed with base substitutions of the origin of DNA replication. *J. Mol. Biol.* **131**:801–817.
39. Türlér, H. 1977. Interaction of polyoma and mouse DNAs. IV. Time course and extent of integration of polyoma DNA into mouse DNA during lytic infection. *J. Virol.* **23**:272–285.