JOURNAL OF VIROLOGY, Dec. 1992, p. 6885–6892
0022-538X/92/126885-08$02.00/0

Vol. 66, No. 12

A Novel Sequence-Specific DNA-Binding Protein, LCP-1, Interacts with Single-Stranded DNA and Differentially Regulates Early Gene Expression of the Human Neurotropic JC Virus

HIROOMI TADA AND KAMEL KHALILI*

Molecular Neurovirology Section, Jefferson Institute of Molecular Medicine, and Department of Biochemistry and Molecular Biology, Thomas Jefferson University, Philadelphia, Pennsylvania 19107-5541

Received 24 June 1992/Accepted 24 August 1992

We have identified a novel brain-derived single-stranded-DNA-binding protein that interacts with a region of the human neurotropic JC virus enhancer designated the lytic control element (LCE). This nuclear factor, LCP-1 (for lytic control element-binding protein 1), specifically recognizes the LCE, as determined by gel retardation assays. Alkylation interference showed that specific nucleotides within the LCE were contacted by LCP-1. Subsequent experiments revealed that point mutations within the LCE differentially affected LCP-1 binding. UV cross-linking and competition analysis suggested that the LCP-1 DNA-protein complexes were 50 to 52 and 100 to 120 kDa in size. Promoter mutations that affected LCP-1 binding reduced early mRNA transcription during the early phase of the lytic cycle. However, upon DNA replication in the presence of JC virus T antigen, when early mRNA initiation shifts to new locations indicative of the late phase, the LCP-1 mutations had no effect. We suggest that the JC virus early transcription unit is differentially regulated by LCP-1 prior to but not after DNA replication, suggesting a novel mechanism by which DNA structure regulates eukaryotic gene expression.

JC virus (JCV) is an excellent model system for examining the regulation of tissue-specific gene expression in the central nervous system. JCV is an opportunistic polyomavirus responsible for the fatal demyelinating disease progressive multifocal leukoencephalopathy (26, 27). Unlike simian virus 40 (SV40) and other polyomaviruses, JCV has an unusually narrow tissue tropism. In immunocompromised hosts, JCV lytically infects oligodendroglial cells, the myelin-producing cells in the central nervous system, while in cell culture, JCV replicates efficiently only in primary human fetal glial cells (26). The narrow tissue tropism of JCV can be attributed, at least in part, to the transcriptionally restricted expression of the early genes to glial cells, as examined by transient-transfection assays (8, 15, 37), cell fusion experiments (3), and experiments with transgenic mice (32, 39).

It has been shown previously that the viral enhancer contains a repeated pentanucleotide sequence within the OP_1 region, 5′-AAAAAAAAGGGAAGGGATGGCTG-3′, which downregulates transcription initiation from the viral late promoter as well as a heterologous promoter in glial cells (36). Others have found that this A+G-rich motif is required for T-antigen (T-Ag)-mediated replication of JCV DNA (23) and the related human polyomavirus BK virus (6). Multiple nuclear factors derived from glial cells have been found to interact with this sequence, in particular a 56-kDa protein (30, 36), although the functions of these proteins have not been determined. More recent results with a heterologous promoter have indicated that the pentanucleotide repeat may function as an orientation-dependent activator/repressor (35a), suggesting that this region may represent a novel multifunctional element capable of interactions with several nuclear proteins involved in the regulation of both viral transcription and DNA replication. Therefore, we have designated the pentanucleotide repeat sequence the lytic control element (LCE) for JCV.

In the present study, we have begun to identify, characterize, and purify the nuclear proteins derived from brain tissue which interact with the LCE. We have found several complexes which appear to interact with this region. In particular, we have identified a novel sequence-specific complex, lytic control element-binding protein 1 (LCP-1), which recognizes only single-stranded DNA. We have identified sets of point mutations that specifically affect the binding of LCP-1 to its target sequences within the LCE. These point mutations, when placed back into the full-length promoter, affected transcription from the early promoter prior to DNA replication (early-early mRNAs [E_E mRNAs]) but not after DNA replication (late-early mRNAs [L_E mRNAs]). We discuss how LCP-1 may be involved in the regulation of the early promoter during the course of the JCV lytic cycle in glial cells as well as the potential role of similar transcription factors in the regulation of cellular growth and gene expression.

MATERIALS AND METHODS

DNA-protein interactions. Oligonucleotides were synthesized commercially by Oligos Etc., Guilford, Conn., and gel purified by denaturing gel electrophoresis and UV shadowing prior to use. The sequences of the oligonucleotides were:

OP_{1L} 5′-GATCCAAAAAAAAGGGAAGGGATGGCTG-3′
3′- GTTTTTTTTCCCTTCCCTACCGACCTAG-5′ OP_{1E}
OP_{2L} 5′-GATCCAAAAAAAACAGATCTAATGGCTG-3′
3′- GTTTTTTTTGTCTAGATTACCGACCTAG-5′ OP_{2E}
OP_{3L} 5′-GATCCAAAAAAAATGGAACGGAG-3′
OP_{4L} 5′-GATCCAAAAAAAAGTGAAGCGAG-3′
OP_{5L} 5′-GATCCAAAAAAAAGGTAAGGCAG-3′

* Corresponding author.

Fresh monkey brain was obtained from Perkasave (Quakertown, Pa.) and kept frozen at −70°C until use. Nuclear extracts were prepared as described previously (1). The protein concentration was measured with a kit from Bio-Rad and determined to be 2 mg/ml.

Gel retardation experiments were performed essentially as described before (36). Briefly, double- and single-stranded DNAs were end labeled with ^{32}P and gel purified. Approximately 30,000 cpm of probe was incubated with 4 μg of nuclear extract on ice for 15 min prior to electrophoresis on native 9% polyacrylamide–0.5× TBE gels. In competition experiments, extracts were incubated with unlabeled DNAs on ice for 15 min before the probe was added.

For methylation and carbethoxylation interference assays, end-labeled oligonucleotides were alkylated with dimethyl sulfide (DMS) or diethylpyrocarbonate (DEPC) (Sigma) as described previously (35). Binding reaction mixes were scaled up fivefold, using 300,000 cpm of modified probe. Gel retardation was carried out as above, and the wet gel was exposed overnight. Free and bound oligonucleotides were isolated from the gel and cleaved for 30 min at 95°C in 1 M piperidine (Sigma). Following several rounds of lyophilization, the products were electrophoresed on a denaturing 20% polyacrylamide gel.

In UV cross-linking experiments, binding reactions were carried out as above with 1 μg of total protein from partially purified fractions. Complexes were cross-linked for 30 min at room temperature with a hand-held long-wave UV light. The resulting complexes were either analyzed directly by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) or loaded onto native gels and then analyzed by SDS-PAGE as described in the text. The dried gel shown in Fig. 5 was exposed to a phosphor screen and the complexes were detected with a PhosphorImager (Molecular Dynamics), whereas the gels shown in Fig. 6 were visualized by standard autoradiography.

Plasmids and site-directed mutagenesis. All enzymes and reagents used in cloning and sequencing were purchased from either New England Biolabs, Boehringer Mannheim Biochemicals, or U.S. Biochemicals and used according to the supplier's recommendations. The plasmid $pBEL_2$-JC was constructed by replacing the SV40 control region in the vector $pBEL_2$ (41) with the control region of the Mad I strain of JCV. The *Nco*I fragment from plasmid pBJC (1) containing the early and late promoter regions was blunted with T4 DNA polymerase and gel purified. The SV40 sequences were removed from the vector with *Hin*dIII, and the ends were blunted with T4 DNA polymerase and ligated to the JCV fragment.

The same *Nco*I fragment was cloned into the *Sma*I site of M13 mp19 and mutagenized as described previously (19). The sequences of the mutagenic oligonucleotides (mutated bases are underlined) were: 3A, 5′-AAAAAAAATGGAA CGGATGGCTGCCAGCC-3′; 4A, 5′-AAAAAAAAGTGAAG CGATGGCTGCCAGCC-3′; and 5A, 5′-AAAAAAAAGGT AAGGCATGGCTGCCAGCC-3′. The resulting mutants were sequenced to identify clones containing the expected mutations only in the "A" 98-bp repeat (see Fig. 1). The *Hin*dIII-*Pvu*II fragment containing the origin and the 98-bp repeats from each mutant replaced the same fragment in the parent plasmid to yield $pBEL_2$-3A, -4A, and -5A. This entire insert was sequenced again to ensure that no additional mutations were present in the promoters of the $pBEL_2$-derived vectors.

Cell culture, transfections, and S1 nuclease protection. U87-MG cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (GIBCO) and antibiotics. Cells were plated at a density of 2×10^6 cells per 10-cm dish and given fresh medium 4 h prior to transfection by the calcium phosphate coprecipitation method (10). Twenty-five micrograms of each $pBEL_2$-derived plasmid was cotransfected with 5 μg of either pUC19 or pBJC-T, a plasmid encoding JCV T-Ag under the control of the herpes simplex virus ICP4 promoter (38).

Total RNA was isolated 40 h after transfection for S1 nuclease protection assays. Cells were washed twice with phosphate-buffered saline and lysed in a buffer containing 4 M guanidinium isothiocyanate, 10 mM EDTA, and 140 mM 2-mercaptoethanol. The lysates from two 10-cm plates were pooled and pelleted through a cushion of 5.7 M cesium chloride–10 mM EDTA in a TL-100 rotor (Beckman) for 4.25 h at 70,000 rpm. Input plasmid DNA was subsequently removed by digestion with RNase-free DNase I (Boehringer Mannheim).

The S1 nuclease protection probe was derived essentially as described previously (38). The fragment spanning the globin cDNA sequence from +108 through the JCV early region to nucleotide 111 was cloned into the *Sma*I site in M13 mp19. The universal primer was used to direct synthesis of a uniformly labeled single-stranded probe 622 nucleotides in length. Total RNA (30 μg) was resuspended with the probe (20,000 cpm) in 20 μl of 80% formamide–40 mM PIPES [piperazine-*N*,*N*′-bis(ethanesulfonic acid), pH 6.5]–400 mM NaCl–1 mM EDTA, denatured for 15 min at 70°C, and hybridized overnight at 37°C. S1 digestion was carried out at 37°C for 1 h by adding 180 μl of 30 mM sodium acetate (pH 4.6)–250 mM NaCl–1 mM $ZnSO_4$–30 μg of denatured pUC19 DNA per ml–120 U of S1 nuclease (Boehringer Mannheim). The reaction mixes were subsequently phenol-chloroform extracted, ethanol precipitated, and electrophoresed on a denaturing 6% polyacrylamide gel.

RESULTS

Identification of a novel DNA-binding protein, LCP-1. Our attention has focused on the LCE within the JCV enhancer, a region which may be involved in the regulation of both viral transcription and DNA replication. The region containing the LCE in the A 98-bp repeat proximal to the origin (Fig. 1) displays an unusual DNA structure that is highly sensitive to S1 nuclease and bromoacetaldehyde (2). Therefore, we asked whether nuclear proteins could recognize different structural forms of the LCE. Nuclear extracts from monkey brain were prepared and tested for specific DNA-binding activity with single- and double-stranded oligonucleotides containing these sequences (Fig. 2A). Both the duplex oligonucleotide OP_1ds and the early-coding-strand oligonucleotide OP_{1E} formed two complexes, A and B, of similar mobilities in a gel retardation assay (Fig. 2A, lanes 1 and 3). The late-coding-strand oligonucleotide OP_{1L}, which contains the AGGGAAGGGA sequence, formed a minor complex similar in mobility to complex A. Interestingly, the major DNA-protein complexes formed by OP_{1L}, a doublet labeled LCP-1α and LCP-1β (Fig. 2A, lane 2), were unique to this probe.

Competition experiments were performed with single- and double-stranded oligonucleotides containing either the wild-type sequence (OP_1) or a mutant variant (OP_2) in order to test both the structural and DNA sequence requirements of the complexes (Fig. 2B). In this assay, unlabeled competitor DNAs were incubated with nuclear extract prior to addition of the OP_{1L} probe. Unlabeled OP_{1L} (Fig. 2B, lanes 2 and 3)

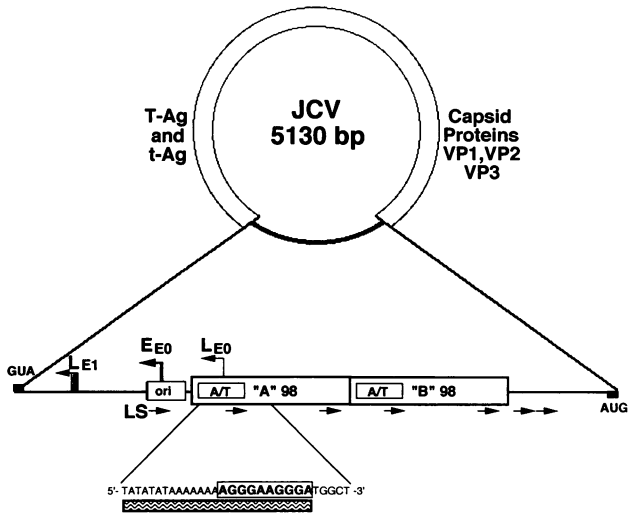

FIG. 1. JCV control region consists of an origin of DNA replication (ori) and a 98-bp direct repeat. The E_E and L_E mRNA initiation sites are shown above the control region, whereas the late mRNA initiation sites are shown below, as determined previously (14, 17). The box within each 98-bp repeat denotes the A+T-rich region. The sequences surrounding the LCE (box) are shown for the proximal A 98-bp repeat and are found on the late coding strand. The shaded box below the sequence denotes the region of hypersensitivity determined previously (2) that is found only in the A 98-bp repeat.

competed efficiently for the complex, whereas OP_1ds (Fig. 2B, lanes 4 and 5) was unable to sequester LCP-1. Comparison of OP_{1L} (Fig. 2B, lanes 2 and 3) with OP_{2L} (Fig. 2B, lanes 6 and 7) indicated that the AGGGAAGGGA sequence was required for the formation of the LCP-1 complex.

The complex comigrating with band A (shown in Fig. 2B) revealed wider sequence specificity than LCP-1. The intensity of this complex was diminished only when OP_{2L} was used at a higher concentration as a competitor. Complex A but not the LCP-1 complexes was observed when the OP_1ds probe was used (Fig. 2C, lane 1). The competition experiments indicated that complex A had a different sequence specificity than LCP-1. Both single- and double-stranded OP_1 oligonucleotides competed effectively for this complex (Fig. 2C, lanes 2 to 6). Moreover, OP_2ds (lanes 6 and 7) and OP_{1E} (not shown) also competed for this activity. Complex B was common to OP_{1E} and OP_1ds; however, it was not reproducibly observed in our studies. Therefore, it appears that the LCP-1 complexes were formed only on single-stranded DNA containing the AGGGAAGGGA sequence and may represent novel sequence-specific, single-stranded-DNA-binding complexes present in brain nuclear extracts.

Specific nucleotide contacts made by the LCP-1 complexes. Methylation and carbethoxylation interference assays were used to determine which purine nucleotides were in close contact with the LCP-1 proteins. DMS methylates DNA at N-7 of guanosyl and N-3 of adenosyl residues (G > A). DEPC carbethoxylates purines at N-7 (A > G). In these experiments, OP_{1L} was ^{32}P-end-labeled and chemically

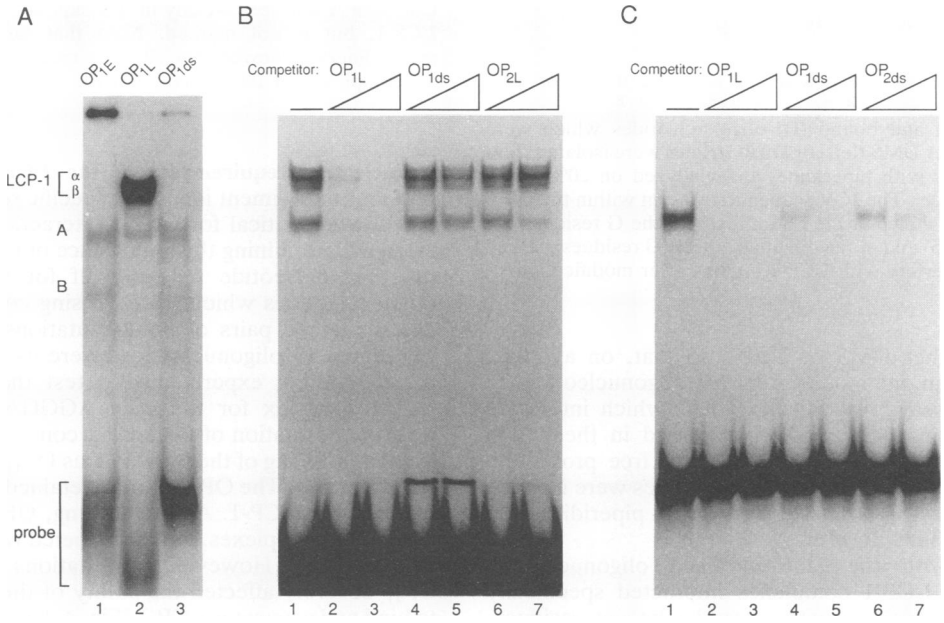

FIG. 2. (A) 5'-End-labeled single-stranded and duplex oligonucleotides were analyzed on native gels for their ability to form complexes with a protein present in monkey brain nuclear extract. OP_{1E} represents the early coding strand (lane 1), OP_{1L} represents the late coding strand (lane 2), and OP_1ds represents duplex probe (lane 3). The doublet labeled LCP-1α and LCP-1β was unique to OP_{1L}, and its sequence specificity is analyzed in panel B. Complex A was common to all three oligonucleotides, and its specificity is analyzed in panels B and C. Complex B was common to OP_{1E} and OP_1ds but was not reproducibly observed. The sequences of the oligonucleotides are shown in Materials and Methods. (B) Complexes formed by the OP_{1L} single-stranded probe (lane 1). The upper doublet is LCP-1α/β; the lower doublet is complex A. The specificity of these complexes was tested by preincubation of the extract with 10 or 100 ng of OP_{1L} (lanes 2 and 3, respectively), OP_1ds (lanes 4 and 5, respectively), and the mutant OP_{2L} (lanes 6 and 7, respectively). The new band that appeared in lanes 4 and 5 at the bottom of the gel is free duplex oligonucleotides. (C) Complex A formed by the OP_1ds oligonucleotide (lane 1). The specificity of complex A was tested by preincubation of the extract with 10 or 100 ng of single-stranded OP_{1L} (lanes 2 and 3, respectively), duplex OP_1ds (lanes 4 and 5, respectively), and the mutant duplex OP_2ds (lanes 6 and 7, respectively). This set of samples was run on the same gel as the samples in panel B, but the figure was separated for simplicity. The migration of complex A was identical to that in panel B.

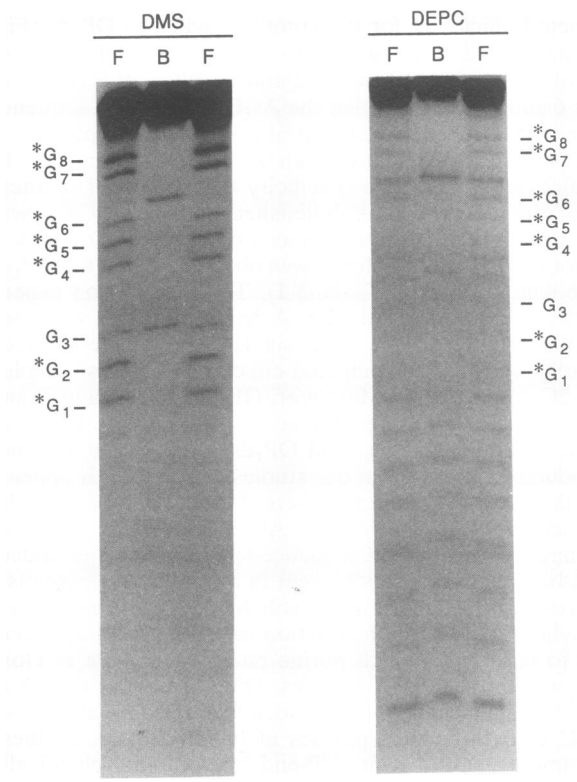

FIG. 3. Free (F) and bound (B) oligonucleotides which were premodified by either DMS (left) or DEPC (right) were isolated from native gels, cleaved with piperidine, and analyzed on 20% acrylamide sequencing gels. The JCV sequences present within the OP_{1L} oligonucleotide are shown at the bottom, where the G residues are numbered from the 5′ end of the oligonucleotide. G residues marked with an asterisk interfere with LCP-1 binding after modification.

modified with either DMS or DEPC so that, on average, there was less than one modification per oligonucleotide. In gel retardation assays, modified residues which interfered with LCP-1 binding were underrepresented in the bound population of oligonucleotides. Therefore, free probe and DNA bound in the LCP-1α and -1β complexes were isolated after native gel electrophoresis, cleaved with piperidine, and compared on sequencing gels.

The results with the DMS-modified oligonucleotide showed that the LCP-1 complexes interacted specifically with the LCE as well as with several adjacent residues. Comparison of the bound and free DNAs indicated that G-1, G-2, G-4, G-5, and G-6 of the pentanucleotide repeat as well as the 3′ residues G-7 and G-8 were in close contact with LCP-1 (Fig. 3). Similarly, the DEPC-modified oligonucleotide identified the same guanosyl nucleotides as being in contact with LCP-1. It does not appear that G-3 and the 3′-most G residue interfered with LCP-1 binding when modified at N-7, and they may not be in close contact with the complex. In addition, none of the adenosyl residues appear to make critical contacts with LCP-1 at either the N-3 or N-7 position.

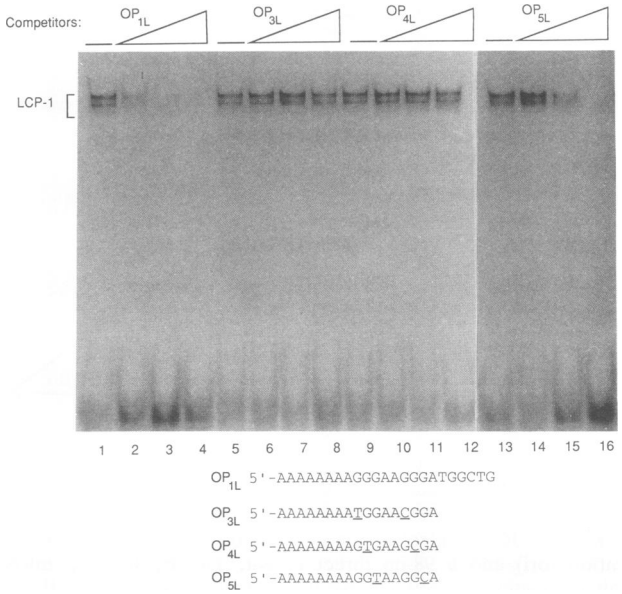

FIG. 4. Extracts were incubated with increasing concentrations of either single-stranded homologous or mutant competitor DNAs prior to addition of the OP_{1L} probe. The first lane in every set (lanes 1, 5, 9, and 13) contains probe alone; other lanes contained 1, 10, or 100 ng, respectively, of the homologous OP_{1L} (lanes 2 to 4) and the mutant OP_{3L} (lanes 6 to 8), OP_{4L} (lanes 10 to 12), and OP_{5L} (lanes 14 to 16) DNAs. The sequences of the competitor DNAs are shown at the bottom. This figure is a composite of two gels which were electrophoresed in parallel and represents equivalent exposures. Complex A was seen weakly in this exposure, migrating faster than LCP-1, but is not marked. Note that longer exposure of the autoradiogram showed that all three mutants effectively competed for complex A at 100 ng (not shown).

Nucleotide requirements of the LCP-1 complexes. The previous experiment identified specific residues within OP_{1L} which were critical for LCP-1 interactions. We were interested in determining the importance of the G residues within the pentanucleotide repeat itself for binding. Therefore, oligonucleotides which were missing contacts G-7 and G-8 and contained pairs of point mutations were synthesized. These mutant oligonucleotides were used as competitors in gel retardation experiments to test the specificity of the LCP-1 complex for the core AGGGAAGGGA sequence (Fig. 4). Formation of the LCP-1 complexes were abolished by 10 and 100 ng of the homologous OP_{1L} competitor (Fig. 4, lanes 3 and 4). The OP_{5L} mutant retained a reduced ability to compete for LCP-1. At 1 and 10 ng, OP_{5L} did not diminish the LCP-1 complexes, but it competed well at 100 ng (Fig. 4, lanes 14 to 16). However, the mutations present in OP_{3L} and OP_{4L} severely affected the ability of these oligonucleotides to bind and sequester LCP-1 (Fig. 4, lanes 6 to 8 and 10 to 12, respectively). Even at 100 ng, which represented a greater than 50-fold molar excess, no competition by these two mutants was observed.

These data appeared to be consistent with the results of the methylation and carbethoxylation interference experiments. All three mutant oligonucleotides were missing contact points G-7 and G-8. In addition, OP_{3L} and OP_{4L} had transversion mutations at two contact points in the core sequence (G-1 plus G-4 and G-2 plus G-5, respectively). Therefore, only three of seven contact points for LCP-1 remained intact in these mutants. On the other hand, OP_{5L}

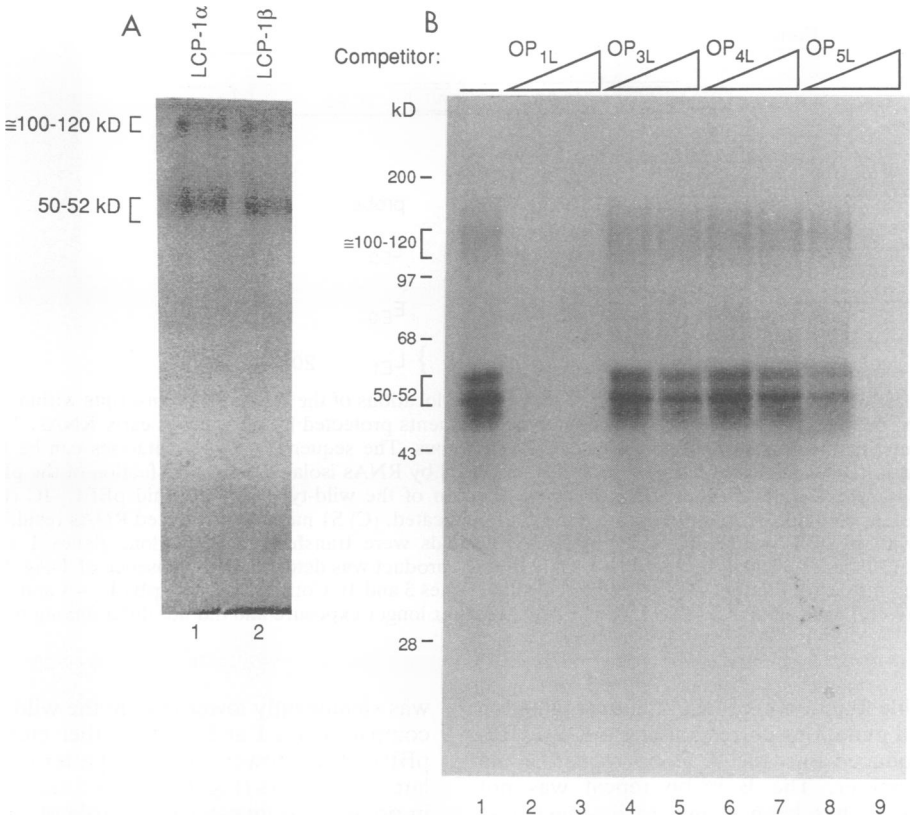

FIG. 5. (A) The LCP-1 complexes were UV cross-linked to the OP_{1L} probe prior to loading on a native gel. LCP-1α (lane 1) and LCP-1β (lane 2) were cut separately from the wet gel, eluted, and loaded onto SDS–10% PAGE gels. The dried gel was exposed to a phosphor screen, and the complexes were detected with a PhosphorImager (Molecular Dynamics). (B) Partially purified LCP-1 was covalently cross-linked to labeled OP_{1L} probe by UV light in the absence (lane 1) and presence (lanes 2 to 9) of 10 (lanes 2, 4, 6, and 8) and 100 (lanes 3, 5, 7, and 9) ng of the unlabeled homologous or mutant competitor DNAs used in Fig. 4. The resulting DNA-protein complexes were analyzed directly on SDS–10% PAGE gels and visualized by autoradiography. Two resolvable complexes of 50 to 52 kDa and a larger complex of 100 to 120 kDa are labeled. The smear below the labeled complexes may be degradation products.

contains transversion mutations at G-3 and G-6, but only one of the mutations affected a contact point within the LCE (G-6). We suggest that since OP_{5L} contained four of seven contact points, it retained a reduced ability to bind LCP-1, whereas the three contact points present in OP_{3L} and OP_{4L} were not sufficient to bind LCP-1.

Characterization of the protein components of LCP-1. LCP-1 was partially purified from monkey brain nuclear extracts by ion-exchange chromotography in order to study these proteins biochemically. The LCP-1 complexes were isolated from the other OP_1 region-binding proteins with a purification of approximately 17-fold (35a). These partially purified protein fractions were used to characterize the protein components of LCP-1 by UV cross-linking and competition.

Initially, UV cross-linking was used to determine whether there were differences between the protein components of LCP-1α and LCP-1β. Proteins were cross-linked to end-labeled OP_{1L} by exposure to long-wave UV light for 30 min prior to analysis by gel retardation. LCP-1α and LCP-1β were cut separately from the wet gel, eluted into protein sample buffer, and analyzed by SDS-PAGE. Our results showed two cross-linked species present in both LCP-1α and LCP-1β, a smaller complex of 50 to 52 kDa and a larger one of 100 to 120 kDa (Fig. 5A). Although it was difficult to resolve a clear size difference in this particular experiment, it appeared that the 50- to 52-kDa component of LCP-1β migrated slightly faster than that of LCP-1α. Moreover, a faint minor band above the 50- to 52-kDa band was detected in the LCP-1α complex. The integrity of this band has yet to be determined. Later experiments in which cross-linked proteins were loaded directly on protein gels resolved two distinct complexes in this size range.

The sequence specificity of these complexes was tested by performing competition experiments with the OP_3, OP_4, and OP_5 mutant oligonucleotides. These unlabeled competitors were incubated with the partially purified protein fraction prior to addition of OP_{1L} probe, cross-linked, and analyzed immediately by SDS-PAGE. Resolvable 50- and 52-kDa complexes as well as the 100- to 120-kDa band were detected (Fig. 5B, lane 1). The smeared bands smaller than 50 kDa appear to be degradation products. The homologous competitor abolished binding of all of the proteins in the sample at 10 and 100 ng (Fig. 5B, lanes 2 and 3). The OP_{5L} oligonucleotide did not diminish protein binding much at 10 ng but competed well at 100 ng (Fig. 5B, lanes 8 and 9). Neither OP_{3L} (Fig. 5B, lanes 4 and 5) nor OP_{4L} (Fig. 5B, lanes 6 and 7) competed at either concentration of DNA. These results mirrored the gel retardation results in Fig. 4 and confirmed that these DNA-protein complexes were components of LCP-1.

Functional analysis of LCP-1 promoter mutations. In order

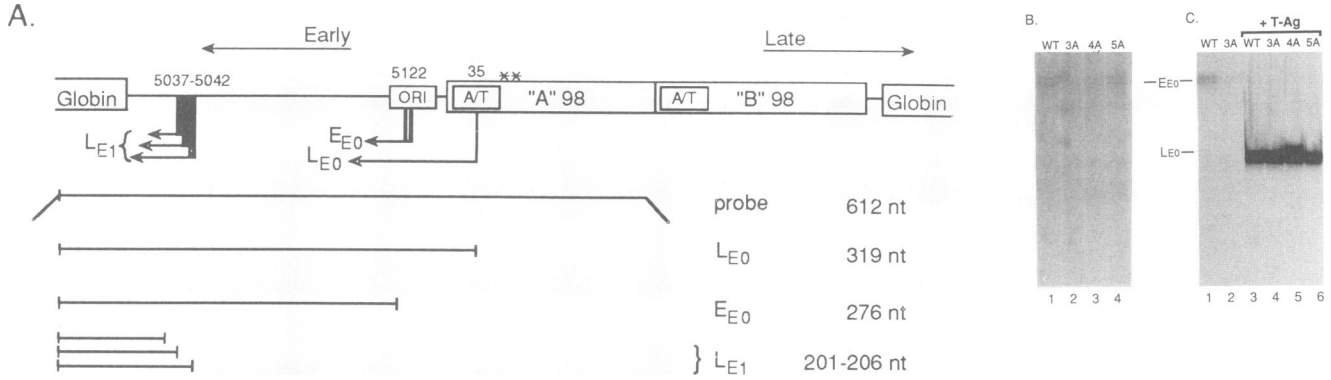

FIG. 6. (A) Structures of the $pBEL_2$-derived plasmids, showing the locations of the E_E and L_E transcripts within the JCV control region. Below are indicated the full-length S1 probe and the sizes of the fragments protected by the various early RNAs. The two asterisks above the A 98-bp repeat show the positions of the 3A, 4A, and 5A mutations. The sequences of the mutations can be found in Materials and Methods. nt, nucleotides. (B) S1 nuclease digestion products protected by RNAs isolated from transfection of the $pBEL_2$-derived plasmids into U87-MG glial cells. The E_{E0} mRNAs resulting from transfection of the wild-type (wt) plasmid $pBEL_2$-JC (lane 1) or the mutants $pBEL_2$-3A (lane 2), $pBEL_2$-4A (lane 3), and $pBEL_2$-5A (lane 4) are indicated. (C) S1 nuclease-protected RNAs resulting from cotransfection of $pBEL_2$ plasmids with pBJC-T. $pBEL_2$-JC and $pBEL_2$-3A plasmids were transfected either alone (lanes 1 and 2, respectively) or cotransfected with pBJC-T (lanes 3 and 4, respectively). Only the E_{E0} product was detected in the absence of T-Ag (lanes 1 and 2), whereas in the presence of T-Ag, the early initiation sites shifted as indicated (lanes 3 and 4). Cotransfection of $pBEL_2$-4A and $pBEL_2$-5A with pBJC-T (lanes 5 and 6, respectively) gave similar results. L_{E0} was detected after longer exposure and did not differ among the samples (not shown).

to identify the possible function(s) of LCP-1 in the regulation of JCV, site-directed mutations corresponding to OP_3, OP_4, and OP_5 were introduced into the A 98-bp repeat of the full-length JCV promoter. The B 98-bp repeat was not mutagenized because it has been shown to be much less sensitive to S1 nuclease and bromoacetaldehyde (2) and thus less likely to be unpaired and capable of binding LCP-1. The wild-type and mutant promoters were placed into plasmid $pBEL_2$ (13, 41), a vector which contains two divergent copies of the rabbit β-globin gene.

These plasmids were transfected either alone or with a plasmid encoding JCV T-Ag into U87-MG cells. This cell line was derived from a human glioblastoma which contained a protein complex indistinguishable from LCP-1 (35a). Total RNA was extracted at 40 h posttransfection and hybridized to a uniformly labeled S1 probe complementary to the globin cDNA and the JCV early region (Fig. 6A). This probe mapped the early RNAs according to their initiation sites, whereas the late RNAs protected a single fragment corresponding to the break point between the globin and late leader sequences.

Analysis of RNA derived from transfection of the wild-type construct $pBEL_2$-JC alone showed one protected species, marked E_{E0}, which corresponded to the major E_E mRNAs (Fig. 6B, lane 1). RNA isolated from cells transfected with the mutant plasmids $pBEL_2$-3A and $pBEL_2$-4A reduced the levels of the E_{E0} transcripts (Fig. 6B, lanes 2 and 3). Transfection with the $pBEL_2$-5A mutant gave approximately wild-type levels of the E_{E0} transcription products (Fig. 6B, lane 4). These results were consistent with those of the binding studies and suggested that LCP-1 regulates E_E transcription initiation of JCV.

Cotransfection of $pBEL_2$-JC with T-Ag, which initiates DNA replication, mimics the late phase of the lytic cycle by causing a shift in the early mRNA initiation sites from E_{E0} to two new locations, L_{E1} and L_{E0} (17). Interestingly, the early RNAs isolated from the cotransfection of $pBEL_2$-3A with T-Ag showed the same pattern and abundance of late-early transcripts as in the wild type (Fig. 6C, compare lanes 3 and 4), whereas in the absence of T-Ag, the E_{E0} product level was significantly lower than in the wild type (Fig. 6B and C, compare lanes 1 and 2). The other mutants, $pBEL_2$-4A and $pBEL_2$-5A, showed the same pattern and abundance of the late-early RNAs (Fig. 6C, lanes 5 and 6). Further studies are in progress to investigate the role of LCP-1 in transcription of the JCV late promoter before and after DNA replication (4).

DISCUSSION

We have presented evidence that a single-stranded-DNA-binding protein in brain nuclear extract specifically recognized the LCE. The binding studies showed that LCP-1 recognized a specific sequence within the enhancer, and point mutations that differentially affected LCP-1 binding to that sequence were determined. These LCE point mutations, when placed in the context of the full-length promoter, affected JCV early gene transcription initiating from the E_{E0} but not from either the L_{E0} (shown in Fig. 6C) or the L_{E1} (not shown). The correlation between the binding and functional assays suggested that LCP-1 regulates early gene transcription during the early phase of the lytic cycle, prior to DNA replication. However, during the late phase following replication, the mutations had no effect on L_E transcription, suggesting a novel mechanism by which changes in DNA structure may affect the interaction between a transcription factor and its cognate binding site to regulate transcription initiation.

The majority of eukaryotic transcription factors represent a class of proteins which recognize and bind specific DNA sequences to activate or repress transcription (11, 12, 21, 22, 34). The DNA target sequences for this growing class of proteins have been fairly well characterized, but much less is known about potential alternative structures of these *cis* elements that may influence DNA-protein interactions. From the existence of distinct DNase I- and S1-sensitive sites within the genome, it has been suggested that subclasses of sequence-specific proteins which recognize non-B-DNA structures may exist (20). For example, DNA bending has been shown to strengthen the interaction of the

Drosophila zinc finger protein *suppressor of hairy wing* [*su(HW)*] with the octamer motif within the gypsy element (33). The transcription factor MF-3 has recently been shown to recognize both double- and single-stranded DNAs containing both the MCAT and CArG motifs present in the promoters of some muscle-specific genes (28). The protein ssARS-T from *Saccharomyces cerevisiae* has been identified as a sequence-specific single-stranded-DNA-binding protein which regulates the initiation of DNA replication in yeast cells (29). Thus, it appears likely that new classes of sequence-specific DNA-binding proteins might recognize alternative DNA structures to regulate both transcription and replication.

The LCE has been previously shown to adopt a unique DNA structure, described as a non-B-DNA right-handed helix (2). Interestingly, the LCE in the A 98-bp repeat proximal to the origin and early-early mRNA initiation site is highly sensitive to S1 nuclease and bromoacetaldehyde in negatively supercoiled DNA. In contrast, the same sequence in the distal (B) 98-bp repeat shows little sensitivity to these reagents, suggesting that neighboring sequences, such as the origin, may influence the unusual structure of this region. This difference may have functional consequences, since the E_E mRNAs initiate predominantly from the A repeat TATATA box (17). Thus, while other factors may also be involved, the structure of the LCE may be an important determinant in the positioning of the E_E mRNA initiation complex.

The LCE lies between the TATATA box and the recognition site for a brain-specific activator protein (1, 16, 18). We hypothesize that the affinity of LCP-1 for single-stranded DNA contributes to the assembly of the E_E RNA initiation complex at the A repeat TATATA box. We suggest that the LCE in negatively supercoiled input DNA is single stranded, allowing LCP-1 to bind and subsequently interact with the transcription machinery. Following DNA replication, the LCP-1 binding site may no longer be accessible, possibly due to relaxation of the template by topoisomerase or the replication machinery. Failure to bind LCP-1 may then signal the transcription complex to assemble at new locations on both the early and late promoters. Therefore, the LCE may regulate the early-to-late shift by changing the structure of the region upon DNA replication, signaling transcription to initiate at other locations within the viral control region. Thus, the structure of the LCE may be important in regulating JCV gene expression throughout the viral lytic cycle.

What role does JCV T-Ag play in mediating the shift of the early mRNA initiation sites? We have already discussed a possible indirect role for T-Ag via stimulation of viral DNA replication, resulting in an altered chromatin structure that is no longer recognized by LCP-1. In the case of SV40, it has been suggested that there may be competition between the E_E and L_E start sites for transcription initiation, so that occupancy of the origin by T-Ag sterically hinders initiation at the E_E site (25, 34). We do not favor this hypothesis, since the presence of JCV T-Ag in the absence of DNA replication, as observed in the hamster glioma cell line HJC (8, 9), is not sufficient to induce the shift from the early initiation site in vivo and in vitro (17; unpublished data). On the other hand, it is possible that JCV T-Ag may play a more direct role, such as inactivating or sequestering LCP-1 so that it can no longer bind DNA, as has been shown for SV40 T-Ag and AP-2 (24).

Viruses represent interesting and convenient systems with which to study the relationship between DNA structure and mRNA transcription in eukaryotic cells. Our results suggest a novel mechanism of transcription regulation in which the interaction of a transcription factor with its binding site may be directly influenced by the structure of that region. Elucidation of the mechanisms by which the structure of *cis* elements regulates viral mRNA transcription will have important implications for the regulation of cellular mRNA synthesis, particularly in rapidly proliferating cells and in differentiating tissues.

At present, we have not identified any cellular genes which may be regulated by LCP-1, but we have found related DNA motifs in a number of cellular promoters. The promoters of myelin basic protein and proteolipid protein contain proximal elements that resemble the LCE sequence, differing by a single nucleotide (5, 7). The neurofilament heavy-chain gene promoter contains an A+G-rich element, Pal-1, that appears to be involved in regulation of neurofilament heavy-chain expression (21a). Interestingly, this sequence is present as a palindrome which can potentially alter DNA structure and, perhaps, gene expression.

ACKNOWLEDGMENTS

We thank Robert Lazzarini for communicating his results before publication, Richard Kuhn for providing the *dut ung* mutant *Escherichia coli* BW313, Alan Wildeman for providing the $pBEL_2$ plasmid, Darwin Prockop for the use of the PhosphorImager, Frances Feigl for art work, and members of the Molecular Neurovirology Section of the Jefferson Institute of Molecular Medicine for their helpful discussions and comments.

This work was supported by grants CA47996 and A128272, awarded by the National Institutes of Health, and grant RG2054A1, awarded by the National Multiple Sclerosis Society, to K. Khalili.

REFERENCES

1. **Ahmed, S., J. Rappaport, H. Tada, D. Kerr, and K. Khalili.** 1990. A nuclear protein derived from brain cells stimulates transcription of the human neurotropic virus promoter, JCV_E, *in vitro*. J. Biol. Chem. **265**:13899–13905.
2. **Amirhaeri, S., F. Wohlrab, E. O. Major, and R. D. Wells.** 1988. Unusual DNA structure in the regulatory region of the human papovavirus JC virus. J. Virol. **62**:922–931.
3. **Beggs, A. H., R. J. Frisque, and G. A. Scangos.** 1988. Extinction of JC virus tumor-antigen expression in glial cell-fibroblast hybrids. Proc. Natl. Acad. Sci. USA **85**:7632–7636.
4. **Chang, C.-F., et al.** Unpublished data.
5. **deFerra, F., H. Engh, L. Hudson, J. Kamholz, C. Puckett, S. Molineaux, and R. A. Lazzarini.** 1985. Alternative splicing for the four forms of myelin basic protein. Cell **43**:721–727.
6. **Del Vecchio, A. M., R. A. Steinman, and R. P. Ricciardi.** 1989. An element of the BK virus enhancer required for DNA replication. J. Virol. **63**:1514–1524.
7. **Diehl, H. J., M. Schaich, R. M. Budzinski, and W. Stoffel.** 1986. Individual exons encode the integral membrane domains of human myelin proteolipid protein. Proc. Natl. Acad. Sci. USA **83**:9801–9811.
8. **Feigenbaum, L., K. Khalili, E. O. Major, and G. Khoury.** 1987. Regulation of the host range of human papovavirus JCV. Proc. Natl. Acad. Sci. USA **84**:3695–3698.
9. **Frisque, R. J., D. B. Rifkin, and D. L. Walker.** 1980. Transformation of primary hamster brain cells with JC virus and its DNA. J. Virol. **35**:265–269.
10. **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.
11. **He, X., and M. G. Rosenfeld.** 1991. Mechanisms of complex transcriptional regulation: implications for brain development. Neuron **7**:183–196.
12. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **88**:799–839.
13. **Kelly, J. J., J. M. Munholland, and A. G. Wildeman.** 1989. Comeasurement of simian virus 40 early and late promoter

activity in HeLa and 293 cells in the presence of T antigen. J. Virol. **63**:383–391.
14. **Kenney, S., V. Natarajan, and N. P. Salzman.** 1986. Mapping 5′ termini of JC virus late RNA. J. Virol. **58**:216–219.
15. **Kenney, S., V. Natarajan, V. Strika, G. Khoury, and N. P. Salzman.** 1984. JC virus enhancer-promoter active in human brain cells. Science **226**:1337–1339.
16. **Kerr, D., and K. Khalili.** 1991. A recombinant cDNA derived from human brain encodes a DNA binding protein that stimulates transcription of the human neurotropic virus JCV. J. Biol. Chem. **266**:15876–15881.
17. **Khalili, K., L. Feigenbaum, and G. Khoury.** 1987. Evidence for a shift in 5′-termini of early viral RNA during the lytic cycle of JC virus. Virology **158**:469–472.
18. **Khalili, K., J. Rappaport, and G. Khoury.** 1988. Nuclear factors in human brain cells bind specifically to the JCV regulatory region. EMBO J. **7**:1205–1210.
19. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**:488–492.
20. **Larsen, A., and H. Weintraub.** 1983. An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. Cell **29**:609–622.
21. **Latchman, D. S.** 1990. Eukaryotic transcription factors. Biochem. J. **270**:281–289.
21a. **Lazzarini, R.** Personal communication.
22. **Levine, M., and J. L. Manley.** 1989. Transcriptional repression of eukaryotic promoters. Cell **59**:405–408.
23. **Lynch, K. J., and R. J. Frisque.** 1990. Identification of critical elements within the JC virus DNA replication origin. J. Virol. **64**:5812–5822.
24. **Mitchell, P. J., C. Wang, and R. Tjian.** 1987. Positive and negative regulation of transcription *in vitro*: enhancer-binding protein AP-2 is inhibited by SV40 T-Ag. Cell **50**:847–861.
25. **Myers, R. M., D. C. Rio, A. K. Robbins, and R. Tjian.** 1981. SV40 gene expression is modulated by the cooperative binding of T antigen to DNA. Cell **25**:4445–4449.
26. **Padgett, B. L., and D. L. Walker.** 1976. New human papovaviruses. Prog. Med. Virol. **22**:1–35.
27. **Padgett, B. L., D. L. Walker, G. M. ZuRhein, R. J. Eckroade, and B. H. Dessel.** 1971. Cultivation of a papovavirus-like virus from human brain with progressive multifocal leukoencephalopathy. Lancet **i**:1257–1260.
28. **Santoro, I. M., T. M. Yi, and K. Walsh.** 1991. Identification of single-stranded-DNA-binding proteins that interact with muscle gene elements. Mol. Cell. Biol. **11**:1944–1953.
29. **Schmidt, A. M. A., S. U. Herterich, and G. Krauss.** 1991. A single-stranded DNA binding protein from *S. cerevisiae* specifically recognizes the T-rich strand of the core sequence of ARS elements and discriminates against mutant sequences. EMBO J. **10**:981–985.
30. **Sharma, A. K., and G. Kumar.** 1991. A 53 kDa protein binds to the negative regulatory region of JC virus early promoter. EEBS **281**:272–274.
31. **Small, J. A., G. Khoury, G. Jay, P. M. Howley, and G. A. Scangos.** 1986. Early regions of JC virus and BK virus induce distinct and tissue-specific tumors in transgenic mice. Proc. Natl. Acad. Sci. USA **83**:8288–8292.
32. **Small, J. A., G. A. Scangos, L. Cork, G. Jay, and G. Khoury.** 1986. The early region of human papovavirus JC induces dysmyelination in transgenic mice. Cell **46**:13–18.
33. **Spana, C., and V. G. Corces.** 1990. DNA bending is a determinant of binding specificity for a *Drosophila* zinc finger protein. Genes Dev. **4**:1505–1515.
34. **Struhl, K.** 1991. Mechanisms for diversity in gene expression patterns. Neuron **7**:177–181.
35. **Sturm, R., T. Baumruker, B. R. Franza, and W. Herr.** 1987. A 100-kD HeLa cell octamer binding protein (OBP100) interacts differently with two separate octamer-related sequences within the SV40 enhancer. Genes Dev. **1**:1147–1160.
35a. **Tada, H.** 1992. Ph.D. thesis. Thomas Jefferson University, Philadelphia, Pa.
36. **Tada, H., M. Lashgari, and K. Khalili.** 1991. Regulation of JCV_L promoter function: evidence that a pentanucleotide "silencer" repeat sequence AGGGAAGGGA down-regulates transcription of the JC virus late promoter. Virology **180**:327–338.
37. **Tada, H., M. Lashgari, J. Rappaport, and K. Khalili.** 1989. Cell type-specific expression of JC virus early promoter is determined by positive and negative regulation. J. Virol. **63**:463–466.
38. **Tada, H., J. Rappaport, M. Lashgari, S. Amini, F. Wong-Staal, and K. Khalili.** 1990. *trans*-Activation of the JC virus late promoter by the tat protein of type 1 human immunodeficiency virus in glial cells. Proc. Natl. Acad. Sci. USA **87**:3479–3483.
39. **Trapp, B. D., J. A. Small, M. Pulley, G. Khoury, and G. A. Scangos.** 1988. Dysmyelination in transgenic mice containing JC virus early region. Ann. Neurol. **23**:38–48.
40. **Wasylyk, B., C. Wasylyk, H. Matthes, M. Wintzerith, and P. Chambon.** 1983. Transcription from the SV40 early-early and late-early overlapping promoters in the absence of DNA replication. EMBO J. **2**:1605–1611.
41. **Wildeman, A. G.** 1989. *trans*-Activation of both early and late simian virus 40 promoters by large tumor antigen does not require nuclear localization of the protein. Proc. Natl. Acad. Sci. USA **86**:2123–2127.