Regulation of the Glial-specific JC Virus Early Promoter by the Transcription Factor Sp1*

John W. Henson†
From the Molecular Neuro-oncology Laboratory, Massachusetts General Hospital, Charlestown, Massachusetts 02129

The regulation of glial-specific JC virus early gene expression was addressed by functional dissection of a previously uncharacterized form of the JC virus promoter (MH-1). The MH-1 promoter directed 31-fold higher reporter gene expression in U87MG glioma cells than in HeLa cells in a transient transfection assay. Transfection of promoter constructs containing proximal or proximal plus upstream regions revealed that reporter gene expression was activated by both proximal and tandem repeat regions in glioma cells. The proximal region contains a guanine-rich sequence, the GA box, which was found to regulate the promoter, and was recognized specifically by the transcription factor Sp1. The GA box is also present in the promoters of three glial-specific cellular genes. Together with paired AP-1 and NF-1 sites in the tandem repeats, the GA box is part of a motif that is conserved between several glial-specific promoters, and is thus a potential determinant of glial-specific gene expression. These results delineate the promoter regions required for activation of the MH-1 JC virus promoter, suggest a new determinant of glial specificity, and establish a model for the investigation of combinatorial activation of a glial-specific viral promoter.

JC virus is a neurotropic papovavirus that was isolated from patients with a fatal brain disease, progressive multifocal leukoencephalopathy (PML), in which demyelination results from selective destruction of oligodendrocytes. The glial tropism of JC virus arises from cell-specific regulation of viral gene transcription. In transient transfection assays, the JC virus early promoter directs higher levels of reporter gene expression in glial-derived cell lines than in non-glial cell lines (Kenny et al., 1984), and a JC virus promoter/large T antigen transgene causes specific abnormalities of brain myelination in mice (Feigenbaum et al., 1992). Attempts to identify the molecular determinants of this specificity have resulted in the identification of several transcription factors that activate the JC virus early promoter, but their relevance to promoter specificity has either not been examined or is not clear (Amemiya et al., 1992; Ahmed et al., 1990; Tada and Khalili, 1992; Tamura et al., 1988; Wegner et al., 1993; Kumar et al., 1993). Furthermore, it is not known whether a single promoter region dominantly activates transcription in glial cells or whether multiple regions interact in a combinatorial fashion. Such a functional dissection has been difficult with the original, Mad-1 isolate, because the entire promoter region is duplicated in tandem.

Sequence analysis of numerous JC virus strains has shown that the Mad-1 type promoter is not commonly isolated from the brains of patients with PML (Henson et al., 1992; Martin et al., 1985). Unlike the Mad-1 strain, the promoter commonly found in PML brains contains a single TATA box in association with a guanine-rich, Sp1-like site, and two upstream tandem repeats. Promoters vary in the exact junction between the two repeats, while the junction between the repeats and the proximal promoter region is perfectly maintained (Henson et al., 1992). Fig. 1A shows that the structure of the MH-1 promoter, an example of the commonly isolated form, is very similar to that of the papovavirus SV40. Both MH-1 and SV40 promoters contain guanine-rich sequences (stippled) immediately upstream of a single TATA box (black), with the tandem repeats (open boxes) lying further upstream. Therefore, the structure of the MH-1 promoter lends itself to functional dissection. It can be divided between the proximal region containing the TATA box and guanine-rich region (“GA box”), and the upstream, tandemly repeated sequences.

Examination of the MH-1 JC virus promoter in this way suggests several new hypotheses. The GA box has sequence homology to the binding sites for the transcription factor Sp1 in the SV40 promoter. More interestingly, the promoters of two oligodendrocyte-specific cellular genes, myelin basic protein and proteolipid protein, contain GA boxes in their proximal regions, as shown in Fig. 1B (Gencic and Hudson, 1990; Devine-Beach et al., 1990). These homologies suggest that the GA box plays a role in the strength and specificity of the MH-1 promoter. The tandem repeats in the JC virus promoter, like those of the SV40 promoter, are also likely to have a major role in promoter strength, and possibly specificity. Here, the GA box is shown to regulate the JC virus promoter through interactions with the transcription factor Sp1. Furthermore, the tandem repeats are shown to cooperate with the GA box to activate gene expression in glioma cells. These observations suggest a potential, combinatorial determinant for the activation of glial-specific gene expression.

MATERIALS AND METHODS

Plasmid Construction—Inserts described below were ligated into the HindIII/Smal site of pA2PLUC, upstream of the firefly luciferase gene (gift from Aviva Symes, Molecular Neurobiology Laboratory, Massachusetts General Hospital; Maxwell et al. (1989)). pA2PLUC contained a triplet of SV40 polyadenylation signals immediately
upstream of the cloning sites, such that spuriously initiated transcripts would not be translated, resulting in low background. Plasmids MH-1LUC (called plasmid C in Fig. 2) and Mad-1LUC were constructed by ligating a HindIII/Msal restriction fragment of MH-1 or Mad-1 JC virus into pA3PLUC. The MH-1 JC virus promoter fragment used to construct MH-1LUC (plasmid C) was derived from plasmid p2 (Henson et al., 1991); the Mad-1 JC virus promoter fragment used to construct Mad-1LUC was derived from plasmid p1.

The remaining inserts (A, B, and D) were produced using plasmid p2 as a template for polymerase chain reaction (PCR) amplification. Plasmid A contained the TATA box and the single GA box of the wild-type promoter, and was constructed using a pair of primers (JC7 and JC21, see Table I) that amplified this segment of the wild-type promoter. Plasmid B was produced using JC7 and a 60-base-long upstream primer (JC27) that contained four tandem GA boxes (underlined in Table I), the 3’ end of which was complimentary to the template. PCR generated a product that contained all four GA boxes as well as a noncomplimentary Smal restriction site. Each GA box was 10–11 bases apart, as measured from the central adenine, thereby separating each by one integral turn of the double helix, and placing the four binding sites on the same face of the helix. Plasmid D was the wild-type JC virus promoter except that 4 point mutations have been introduced into the GA box (Ho et al., 1989). In brief, two primer pairs (JC7 and JC19m, JC6 and JC20m) were used to generate PCR products that overlapped across the GA box. Two 30-base-long primers overlying the GA box region each contained four mismatched bases. Both PCR products were purified in low melt agarose gels, and were added to a third PCR reaction containing primers JC6 and JC7. In this manner, using high template concentrations and low numbers of PCR cycles to avoid polymerase errors, full-length JC virus promoter was synthesized containing the four point mutations. Plasmid SV/JC-LUC was prepared by amplification of the tandem repeats of SV40, blunt-ending the PCR product with T4 DNA polymerase (New England BioLabs) and ligating the product into the Smal site upstream of the GA box in plasmid A. Plasmids were purified using column chromatography (Qiagen). All constructs were verified by restriction analysis and by dyeoxy DNA sequencing (U. S. Biochemical Corp.). Plasmids were analyzed on agarose gels to assure purity and supercoiling.

Transfections—HeLa or U87MG glioma cells were plated at approximately 75% confluence in 100-mm plastic dishes. 12 h later the medium was changed and incubated for 1 h. 15 μg of each plasmid DNA were mixed with 5 μg of SVCAT DNA per dish, and transfected according to the calcium phosphate method (Kingston, 1991). 12 h later 15% glycerol solution was added for 30 s, the cells were washed with serum-free medium and transferred to serum-containing medium. Forty-eight hours later the cells were washed with ice-cold PBS and harvested by scraping in 300 μl of luciferase lysis buffer which is 1% Triton X-100 and 1 mM dithiothreitol in 1 x luciferase buffer (4 mM EGTA, 25 mM glycyglycine, 15 mM MgSO4) at 4°C. Lysates were spun to pellet cell debris. 100 μl of lysate were added to 370 μl of complete buffer solution (300 μl of 1 x luciferase buffer, 62 μl of 0.1 M potassium phosphate, pH 7.8, 4 μl of 0.1 M dithiothreitol, and 4 μl of 0.2 M ATP). To determine luciferase activity, samples were placed in a luminometer (LKB), and 200 μl of luciferin substrate (200 μl luciferin, Sigma, in 1 x luciferase buffer) were added. The lumino- meter reports arbitrary light units (LU), which is a measure of the area under the curve of emitted light. The luciferase assay is sensitive over a 100,000-fold range (from 60 to 1,000,000 LU). pA3PLUC (promoterless) plasmid was transfected as a measure of background activity. Expression is linear over the 48-h incubation period (data not shown). Within each assay, LU values were normalized using CAT activity. CAT activity was measured in a standard two-phase partition assay, using butyryl-CoA (Pharmacia LKB Biotechnology Inc.) and [3H]chloramphenicol (DuPont) as substrate (Kingston, 1991). CAT activities of the samples were at least 5-fold above background and were within the linear range of the assay. Each experiment was performed in duplicate to assess the spread of data (Ci et al., 1983). Protein concentrations were determined by the Bradford assay (Bio-Rad). Mobility shift assays were performed using double-stranded oligonucleotides with sequences as shown in Table I. Oligonucleotides were annealed and end-labeled with [32P]ATP and poly nucleotide kinase (New England BioLabs), and purified on a 15% nondenaturing polyacrylamide gel. Assays were performed in a 20-μl

![Diagram](image_url)

**FIG. 1.** A, comparison of the structures of SV40 early promoter and the Mad-1 and MH-1 JC virus early promoters. The early genes (large and small T antigen) are to the right. The stippled boxes represent the GA box in the MH-1 promoter and the six Spl binding sites in the SV40 promoter. The open boxes represent tandemly repeated sequences. Note the structural homology between the MH-1 and SV40 promoters. B, proximal promoter sequences of four glial-specific genes contain GA boxes. The stippled boxes represent the GA  box in the MH-1 promoter and the six Spl binding sites in the SV40 promoter. B, proximal promoter sequences of four glial-specific genes contain GA boxes.

**Table I**

| Oligonucleotide | Consensus Sequence |
|-----------------|-------------------|
| JC6             | 5’-AGAAGCTTCACTAGTGAAGACTCT-3’ |
| JC7             | 5’-CCATAGGCTTCATGAGGATCCC-3’ |
| JC19m           | 5’-AGGCTGAGTACTACATCTTCCTTCCC-3’ |
| JC20m           | 5’-AGGCTGAGTACTACATCTTCCTTCCC-3’ |
| JC21            | 5’-AGAAGCTTCACTAGTGAAGACTCT-3’ |
| JC27            | 5’-AGAAGCTTCACTAGTGAAGACTCT-3’ |
| Mobility shift  |                   |
| Sp1 consensus   | 5’-AGAAGCTTCACTAGTGAAGACTCT-3’ |
| JC              | 5’-AGAAGCTTCACTAGTGAAGACTCT-3’ |
| JC21            | 5’-AGAAGCTTCACTAGTGAAGACTCT-3’ |
| JC27            | 5’-AGAAGCTTCACTAGTGAAGACTCT-3’ |

* Base changes are in lower case letters.
+ 5’ refers to the plasmid site shown in brackets.
+ Sequences complimentary to the GA boxes are underlined.

**Figure: Diagram showing the comparison of the structures of SV40 early promoter and the Mad-1 and MH-1 JC virus early promoters.** The early genes (large and small T antigen) are to the right. The stippled boxes represent the GA box in the MH-1 promoter and the six Spl binding sites in the SV40 promoter. The open boxes represent tandemly repeated sequences. Note the structural homology between the MH-1 and SV40 promoters. B, proximal promoter sequences of four glial-specific genes contain GA boxes. The stippled boxes represent the GA box in the MH-1 promoter and the six Spl binding sites in the SV40 promoter.
total volume containing 25 mM Tris, pH 8.0, 6.25 mM MgCl₂, 0.5 mM EDTA, 50 mM KCl, 0.5 mM dithiothreitol, 10% glycerol, 2% polyvinyl alcohol, 1 μg of poly(dI-dC), and 50,000 cpm of labeled double-stranded oligonucleotide (approximately 1 pmol). 1 μl of nuclear extract (3 μg of U87MG protein and 1 μg of HeLa protein) was added, and the mixture was incubated for 20 min at 4 °C. Reactions were terminated by adding gel-loading buffer (0.4% Nonidet P-40, 2% glycerol, 0.05% TBE (Tris/boric acid/EDTA), and 0.04% bromphenol blue) and products were separated on a 4% polyacrylamide gel (30:1 acrylamide/bis ratio, 0.5 × TBE). Gels were dried and exposed to X-OMAT film (Kodak). For the supershift studies, 1 μg of an affinity-purified rabbit polyclonal antibody against Sp1 (PEP2, Santa Cruz) at a concentration of 10 μg/ml. Secondary antibody (Vector) was detected with ECL (Amersham Corp.) on autoradiography film (Kodak).

RESULTS

MH-1 JC Virus Early Promoter Directs Glial-specific Expression—Previous experiments demonstrating the glial specificity of the JC virus promoter in transient transfections assays have employed the Mad-1 type promoter. To test whether the MH-1 JC virus promoter can direct glial-specific gene expression, the promoter was ligated into a reporter plasmid upstream of the coding region of the firefly luciferase (Schleicher & Schuell), the membrane was allowed to dry for 30 min and was blocked in blotto. The filter was incubated with affinity-purified rabbit polyclonal antibodies against Sp1 (PEP2, Santa Cruz) at a concentration of 10 μg/ml. Secondary antibody (Vector) was incubated with affinity-purified rabbit polyclonal antibodies against Sp1 (PEP2, Santa Cruz Biotechnology) was incubated with nuclear extract under the above conditions for 10 min, labeled double-stranded oligonucleotides were added, and the mixture was incubated for an additional 10 min, terminated, and separated on a gel. For Western analysis, 50 μg of nuclear extract were boiled in 2% SDS loading buffer and separated by 8% SDS-polyacrylamide gel electrophoresis. Following transfer of proteins onto nitrocellulose membrane it was allowed to dry to 30 min and was blocked in blotto. The membrane was incubated with affinity-purified rabbit polyclonal antibodies against Sp1 (PEP2, Santa Cruz) at a concentration of 10 μg/ml. Secondary antibody (Vector) was detected with ECL (Amersham Corp.) on autoradiography film (Kodak).

GA Box Activates Transcription from the MH-1 Promoter—The GA box has sequence and positional homology to the Spl gene and transfected into U87MG glioma cells and HeLa cells. The U87MG glioma cell line has been used in similar previous experiments with the Mad-1 promoter, and the HeLa cell line chosen as a non-glial cell control (Tada and Khalili, 1992). Table II shows that the MH-1 promoter stimulated 31-fold higher levels of luciferase expression in U87MG glioma cells than in HeLa cells. The Mad-1 promoter was also tested in this assay and it stimulated 40-fold more luciferase expression in U87MG glioma cells than in HeLa cells. These results demonstrate that the MH-1 promoter can direct glial-specific gene expression in a transient transfection assay. The full-length MH-1 promoter (plasmid C) directed 81-fold stronger activity than plasmid A. Four point mutations in the GA box (plasmid D) reduced expression 3–4-fold compared to the full-length, wild-type promoter. These results demonstrate that the proximal promoter is regulated by the GA box. In addition, it is clear that the GA box acts in concert with elements in the tandem repeat region, suggesting a combinatorial mechanism of transcriptional activation of the JC virus early promoter.

Dulators of Glial Specificity—The proximal region of the proteolipid protein, and myelin basic protein and JC virus promoters contain GA boxes (Fig. 1B), suggesting that the GA box acts in concert with other factors to activate transcription in glial cells.

Table II

| MH-1-11UC | CAT activity | N | LU × N | Ratio |
|-----------|--------------|---|--------|-------|
| U87MG     | 57,094       | 182,005 | 0.56   | 31,973 | 31   |
| HeLa      | 1,036        | 102,568 | 1.0    | 1,036  | 1    |
| Mad-1-11UC| U87MG        | 125,199 | 189,250| 0.22   | 27,543| 40   |
| HeLa      | 686          | 41,394  | 1.0    | 686    | 1    |

FIG. 2. Deletion and mutational analysis of the MH-1 JC virus promoter demonstrates that two promoter regions activated expression in U87MG glioma cells. Relative levels of luciferase expression directed by various promoter constructs are reported, with expression directed by the proximal promoter equal to 1 (plasmid A, 906 LU). Values were derived from a single transfection assay, performed in duplicate, and light units were normalized to CAT activity. Plasmid A contained only the promoter sequence downstream of the tandem repeats. Plasmid B was the same as A except that it contained four tandem GA boxes. Plasmid C was full-length, wild-type promoter. Plasmid D contained four point mutations within the GA box but was otherwise identical to plasmid C.
Glial-specific Promoter Regulation

A BINDING SITE

PROTEIN  JC  AJC

- H G G G H G

COMPETITOR  a b c --

B BINDING SITE

PROTEIN  JC

- H G H G

ANTIBODY  Sp1  Sp1  Sp1

1 2 3 4 5 6 7 8

FIG. 3. A, mobility shift assay shows that similar proteins in HeLa and U87MG nuclear extracts bind to the GA box. The labeled oligonucleotide binding site in lanes 1-6 contained the sequence of strongly activate expression. An alternative explanation could be a repressor acting at the basal promoter in nonglial cells, or an antiterminator acting in glial cells. In the former situation, transcription factors in both cell types could activate the promoter, but transcription would be greatly reduced in nonglial cells by the action of a repressor on the basal promoter. To test this possibility, a chimeric promoter was constructed which placed the two tandem repeats from the SV40 promoter upstream of the proximal JC virus promoter (pSV/JC-LUC). If a repressor was operating at the basal promoter region, the SV/JC-LUC promoter should still be more active in U87MG cells than in HeLa cells. However, the chimeric promoter activated luciferase expression strongly in both cell types (data not shown). This result suggests that glial specificity is not due to the action of a repressor in nonglial cells.

Sp1 Binds to the GA Box—The GA box has sequence homology to the consensus binding site for the zinc finger transcription factor, Sp1. Recombinant HeLa Sp1 binds to the GA box (Henson et al., 1992), but the identity of binding proteins in nuclear extracts has not been established. To characterize these proteins, nuclear extracts from U87MG and HeLa cells were prepared, incubated with [32P]-labeled oligonucleotides containing the GA box sequence, and analyzed by mobility shift assay. Lanes 1-6 of Fig. 3A employed a labeled oligonucleotide containing the JC virus GA box sequence (see Table I). Lanes 2 and 3 show that two distinct bands bind as a doublet to the GA box from both HeLa and U87MG nuclear extracts. The doublet was competed out by a 20-fold excess of cold Sp1 consensus oligonucleotide (lane 4), but not by an oligonucleotide containing an AP-1 site (lane 5) nor by an oligonucleotide of irrelevant sequence (lane 6). Point mutations in the GA box abolished binding of the doublet (lanes 7 and 8).

Fig. 3A suggested that Sp1 is one of the GA box binding proteins in the nuclear extracts. To determine if Sp1 is in fact producing the retarded bands, antibodies to Sp1 were added to the mobility shift assay (Fig. 3B). The upper band of the doublet in HeLa (lane 4) and U87MG (lane 6) extracts was further retarded ("supershifted") by anti-Sp1 antibodies. These results show that Sp1 from U87MG and HeLa cell nuclei binds to the GA box in an identical fashion. The lower band of the doublet was not supershifted and probably represents binding of a recently identified Sp1 family member, Sp3 (Kingsley and Winoto, 1992). Western analysis of the U87MG and HeLa nuclear extracts showed two protein bands of expected M, representing differentially phosphorylated species of Sp1 (data not shown; Jackson et al., 1990). Thus, postranslational modification of Sp1 is similar in the two cell lines. Since the GA box regulates the MH-1 promoter, it is a functional Sp1 binding site.

DISCUSSION

The MH-1 JC virus promoter is a new model for studies of glial-specific gene expression. It is composed of a well defined, 300-bp-long DNA segment and thus is more amenable to the GA box (JC), and the GA box with four point mutations in lanes 7 and 8 (AJC). Lanes 2 and 7 contained HeLa nuclear extract (H). Lanes 3-6 and 8 contained U87MG nuclear extract (G). Lane 1 contained no protein. 20-fold molar excess cold Sp1 consensus oligonucleotide was added as competitor to lane 4 (a), an AP-1 binding site oligo in lane 5 (b), and an oligo of irrelevant sequence to lane 6 (c). B, Sp1 binds to the GA box. Antibodies to Sp1 supershifted the upper band of the doublet in HeLa (lane 4) and U87MG (lane 6) nuclear extracts. Lane 1 contained anti-Sp1 antibody without nuclear extract; irrelevant polyclonal antibody was added to lane 2 with U87MG nuclear extract. The non-supershifted band probably represents Sp3, a recently discovered family member of Sp1.
than in glial cells, a repressor could be operating in the basal promoter region (Levine and Manley, 1989). However, a chi-squared test indicates that the hypothesis that the combination of a GA box with paired regions presented in this paper is consistent with this hypothesis.

The hypothesis is supported by the fact that cooperative interactions between the GA box and upstream promoter regions. Spl is a newly described participant in virus promoter regulation, and by binding to a site (GA box) which is conserved in several glial-specific promoters, appears to play a role in promoter strength and specificity.

Fig. 4 shows the locations of sites at which transcription factors are known to regulate or bind to the JC virus promoter. Examination of the proximal and upstream promoter regions as diagrammed in Fig. 4 reveals an interesting homology with several brain-specific cellular gene promoters. The promoters of JC virus, myelin basic protein and proteolipid protein each contain a GA box in their proximal regions, and adjacent AP-1 and NF-1 sites in their upstream promoter regions (see above; Amemiya et al., 1992). Paired AP-1 and NF-1 sites are present in the promoters of several other genes that are selectively expressed in the nervous system, including the human glial fibrillary acidic protein promoter, the mouse neurofilament L promoter, and the human proenkephalin promoter (Tada, 1992). Therefore, by acting in a combinatorial manner with other transcription factors in glial cells, Sp1 may contribute to promoter specificity.

Acknowledgments—I am indebted to Henry M. Furneaux and Steven A. Reeves for helpful discussions.

REFERENCES

Ahmed, S., Rappaport, J., Tada, H., Kerr, D., and Khalili, K. (1990) J. Biol. Chem. 265, 13899-13905

Amemiya, K., Traub, R., Durham, L., and Major, E. O. (1992) J. Biol. Chem. 267, 14204-14211

Devine-Beach, K., Lasheen, M. S., and Khalili, K. (1990) J. Biol. Chem. 265, 13830-13835

Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) Nucleic Acids Res. 11, 1475-1489

Feigenbaum, L., Hinrichs, S. H., and Jay, G. (1992) J. Virol. 66, 1176-1182

Ganci, S., and Hudson, L. D. (1990) J. Neurosci. 10, 117-124

Gidoni, D., Kadonaga, J. T., Barrera-Saldana, H., Takahashi, K., Chambon, P., and Tjian, R. (1986) Science 230, 511-517

Henson, J., Saffer, J., and Furneaux, H. (1992) Ann. Neurol. 32, 72-77

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pese, L. R. (1989) Gene 77, 51-59

Jackson, S. P., MacDonald, J. J., Lees-Miller, S., and Tjian, R. (1990) Cell 63, 155-165

Kenny, N., Natarajan, V., Strike, D., Khoury, G., and Salzman, N. P. (1984) Science 226, 1533-1534

Kingsley, M. D., and Winoto, A. (1992) Mol. Cell. Biol. 12, 4251-4261

Kingston, R. E. (1991) in Current Protocols in Molecular Biology (Ausbel, F. M., et al., eds) John Wiley & Sons, New York

Kumar, K. U., Pater, A., and Pater, M. (1995) J. Virol 79, 572-576

Levine, M., and Manley, J. L. (1989) Cell 60, 405-408

Martin, J. D., King, D. M., Slauch, J. M., and Priace, R. J. (1985) J. Virol. 53, 306-311

Maxwell, I. H., Harrison, G. S., Wood, W. M., and Maxwell, F. (1988) Bio-Techniques 7, 276-280

Saffer, J. D., Jackson, S. P., and Thurston, S. J. (1990) Genes & Dev. 4, 659-666

Saffer, J. D., Jackson, S. P., and Annarella, M. B. (1991) Mol. Cell. Biol. 11, 2189-2199

Su, W., Jackson, S., Tjian, R., and Echols, H. (1991) (Genes & Dev. 5, 820-826

Tada, H., and Khalili, K. (1992) J. Virol. 66, 6885-6892

Tamura, T., Inoue, T., Nagata, K., and Miki, H. (1988) Biochem. Biophys. Res. Commun. 157, 419-425

Wegner, M., Drolet, D. W., and Rosenfeld, M. G. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 4743-4747

3 J. W. Henson, unpublished observations.