THE EFFECTS OF AGGGAAGGGA PENTANUCLEOTIDE REPEAT ON JC VIRUS GENE EXPRESSION AND THE ISOLATION AND CHARACTERIZATION OF A cDNA ENCODING AN AGGGAAGGGA BINDING PROTEIN

MINGFENG LIU

THE EFFECTS OF AGGGAAGGGA PENTANUCLEOTIDE REPEAT ON JC VIRUS GENE EXPRESSION AND THE ISOLATION AND CHARACTERIZATION OF A cDNA ENCODING AN AGGGAAGGGA BINDING PROTEIN

BY

© MINGFENG LIU

A thesis submitted to the School of Graduate Studies in partial fulfilment of the requirements for the degree of Master of Science

Division of Basic Medical Science
Faculty of Medicine
Memorial University of Newfoundland

April, 1995

St. John's Newfoundland

ISBN 0-612-13920-4

ABSTRACT

Human JC virus (JCV) is the etiologic agent of progressive multifocal leukoencephalopathy (PML), a fatal demyelinating disorder of the central nervous system. Although JCV is a widely-distributed polyomavirus and is oncogenic in animals, its strict dependence for viral growth and expression of JCV on human fetal glial cells has hindered detailed studies of the regulation of JCV early and late gene expression. Previous reports have shown that gene expression of JCV is regulated by viral T antigen and cellular transactivating factors, such as NF1. To date, several cellular factors regulating JCV early and late gene expression have been identified and isolated, and their binding site sequences or motifs in JCV DNA have been characterized.

The AGGGAAGGGA pentanucleotide (penta) repeat motif of the JCV control region has been suggested to positively (stimulate)and negatively (inhibit) regulate JCV early and late gene expression. A similar sequence that negatively regulates the expression of c-myc was also identified in the upstream region of the c-myc gene. Several proteins ranging from 25 to 80 kDa have been identified which specifically interact with this penta repeat motif. Among them, a 56 kDa

protein has been suggested to act as a repressor for the JCV late gene expression and a 70 to 80 kDa protein has been shown to be required for the early gene expression. Thus, there are several proteins interacting with this penta repeat motif and differentially regulating the JCV early and late gene expression. Therefore, it would be very helpful to isolate and characterize such penta repeat motif binding proteins.

To study the regulation by the penta repeat motif for JCV expression, I cloned JCV promoter-enhancer fragments containing mutated penta repeat motif into a CAT reporter plasmid and transfected the plasmid into P19 cells differentiated by retinoic acid into glial cells and U87MG human glioblastoma cells. CAT assays showed that the pentanucleotide repeat motif within the JCV promoter-enhancer region upregulated JCV late expression and downregulated JCV early expression. Mobility shift assays revealed two protein complexes specifically bound to the wild type penta repeat motif in both glial and nonglial cell extracts. Two proteins of approximately 60 and 45 kDa were detected by UV crosslinking. The penta repeat motif was then used to isolate a penta repeat motif binding factor by screening a cDNA expression library from retinoic acid-differentiated P19 glial cells. One positive cDNA clone (RP1) was found to specifically bind to the penta repeat motif. To characterize RP1, Southwestern blots with a ß-galactosidase (β-gal) fusion

protein indicated that this RP1 cDNA encoded the protein that specifically binds to the penta sequence. Consistent with my mutation results, preliminary *in vivo* results showed that the protein encoded by this cDNA could negatively regulate JCV early gene expression. The sequence of RP1 cDNA was found to contain sequences that were predicted to code for a protein containing seven putative zinc finger motifs in the carboxyl terminus two-thirds of the protein. Based on my data and the results of others, I proposed a model for the interaction of RP1 with JCV penta repeat motif as a complex in glial and nonglial cells, suggesting that this interaction contributes to the glial cell specificity of JCV. This model would involve RP1-penta acting in concert with other sites and factors in the JCV promoter-enhancer to prevent JCV gene expression in nonglial cells.

ACKNOWLEDGEMENTS

First, I wish to sincerely thank my supervisors Drs. Mary M. Pater and Alan Pater, for their unfailing support and guidance during the entire thesis program. My appreciation is also extended to Drs. Gary Paterno and John Robinson for their helpful suggestions and encouragement as my supervisory committee members. I appreciate the financial support provided by the Office of Graduate Studies, Faculty of Medicine, Memorial University of Newfoundland.

I am grateful to Dr. Kotlo U. Kumar for his help at the beginning of my program. I appreciate my colleagues, Dr. Qi Sun, Mr. Xiaolong Yang and Mr. D.V. Reddy, for their friendship and valuable discussion.

I wish to thank Masuma Rahimtula and Ge Jin for their technical assistance. Thanks to all my colleagues and friends for their moral support and encouragement.

Last, but not least, I wish to express my sincere thanks to my parents for their blessing and affection. I especially thank my wife, Jingyu Diao, for her confidence, support and understanding that made this thesis possible.

TABLE OF CONTENTS

Page

ABSTRACT I

ACKNOWLEDGEMENTS iv

TABLE OF CONTENTS. v

LIST OF FIGURES. x

LIST OF ABBREVIATIONS. xii

CHAPTER 1 INTRODUCTION. 1

1.1 Mechanism of transcription 1

1.1.1 Promoters and enhancers 2

1.1.2 RNA polymerase II 3

1.1.3 General transcription factors and transcription initiation complex. 4

1.1.4 Sequence specific transcription factors 7

1.1.5 Coactivators and other accessary proteins. 11

1.2 JCV human papovavirus. 13

1.2.1 General features of JCV 13

1.2.2 Genome of JCV 16

1.2.2.1 Early and late coding region of JCV. 16

1.2.2.2 Regulatory region for JCV gene expression. 17

1.2.3 JCV proteins 23

1.2.3.1 Early proteins. 23
1.2.3.2 Late proteins. 25
1.3 Transcription factors that regulate JCV expression. 27
1.4 Cells used in this study. 35
1.5 Objective of this study. 36
CHAPTER 2 MATERIALS AND METHODS 38
2.1 Materials and reagents 38
2.2 Cell culture and transfection 39
2.2.1 Cell culture of P19 and U87 cells and the differentiation of P19 cells by RA or DMSO 39
2.2.2 The method of transfection. 40
2.3 Site directed mutagenesis and plasmid construction. 41
2.4 Chloramphenicol acetyltransferase (CAT) assays 42
2.4.1 Harvesting cells. 42
2.4.2 Chloramphenicol acetyltransferase assays 43
2.5 Preparation of nuclear and whole cell extracts 44
2.5.1 Preparation of nuclear extract. . . . 44
2.5.2 Preparation of whole cell extract . . 46
2.6 DNA binding analysis 46

2.6.1 Preparation of the DNA probe. 46
2.6.2 Mobility shift assays. 47
2.7 UV crosslinking assays 48
2.8 Screening cDNA expression library. 49
2.8.1 Preparation of DNA probe 49
2.8.2 Screening cDNA library. 50
2.9 Characterization of RP1 cDNA clone. 51
2.9.1 Preparation of RP1-β-galactosidase fusion protein. 51
2.9.2 Southwestern-blot assays 52
2.9.3 Total RNA preparation and Northern-blot assays. 54
2.9.3.1 Total RNA preparation. . . 54
2.9.3.2 Agarose gel electrophoresis of RNA. 55
2.9.3.3 Transfer of RNA to nitrocellulose membrane. 55
2.9.3.4 Prehybridization and hybridization. 56
2.9.4 cDNA sequence analysis. 57
CHAPTER 3 RESULTS 58
3.1 Effect of pentanucleotide (penta) repeat sequence on JCV promoter-enhancer activity. . . . 58
3.2 DNA-protein interaction in mobility shift assays 65

3.3 Analysis for interaction of penta repeat motif binding protein with T antigen. 66
3.4 Identification of the penta sequence binding factor by UV crosslinking 76
3.5 Isolation and characterization for RP1 cDNA, penta element binding factor 76
3.5.1 Screening of cDNA expression library. 77
3.5.2 Cell specificity and Northern-blot assays. 78
3.5.3 DNA binding analysis 81
3.5.4 DNA sequence analysis of RP1 cDNA to characterize the protein. 88
3.5.5 In vivo assays of repression by RP1 cDNA of JCV early gene expression 89
CHAPTER 4 DISCUSSION AND FUTURE DIRECTIONS. 96
4.1 Effect of penta repeat motifs on JCV promoter-enhancer in vivo activity 96
4.2 In vitro binding activity of penta repeat motif. 100
4.3 RP1 protein repression of JCV early expression and comparison with other penta repeat motif-binding factors 104
4.4 Significance of RP1 putative zinc finger motifs 108
4.5 Mechanism of regulation of JCV gene expression

by RP1 and penta repeat motif 110
EFERENCES. 117

LIST OF FIGURES

Page

Figure 1. Genome of JCV MadI strain 19

Figure 2. Diagram of wild type and mutated JCV sequences in expression plasmids 60

Figure 3. Effect of mutations on *in vivo* activity of JCV pentanucleotide repeat motif in glial cells. 62

Figure 4. *In vitro* binding of P19 RA glial cell factors with pentanucleotide repeat motif of JCV . 68

Figure 5. *In vitro* binding of U87 glial cell and DMSO-differentiated P19 cell factors with pentanucleotide repeat motif of JCV. . . . 70

Figure 6. *In vitro* binding of RA P19 glial cell and UD P19 cell proteins with penta repeat motif of JCV 72

Figure 7. *In vitro* binding of P19 RA cell factors with penta-nucleotide repeat motif of JCV in presence of T antigen 74

Figure 8. UV crosslinking to identify and determine size of the proteins binding to penta repeat motif of JCV 79

Figure 9. Specific binding of RP1 recombinant clone to penta repeat motif 82

Figure 10. Northern blot analysis of RP1 expression in

glial and nonglial cells. 84
Figure 11. DNA binding activity of RP1 86
Figure 12. RP1 cDNA sequence and predicted RP1 protein sequence. 91
Figure 13. Effect of RP1 on JCV gene expression *in vivo*. 94
Figure 14. Proposed mechanism of penta repeat motif-RP1 interaction for regulation of JCV early gene expression. 112

LIST OF ABBREVIATIONS

α-MEM	=	Alpha modification of Eagle's medium
AIDS	=	Acquired immunodeficiency syndrome
AP1	=	Activator protein 1
ATP	=	Adenosine triphosphate
BBS	=	BES(N,N-bis[2-hydroxyethyl]-2-amino-ethane-sulfonic acid) buffered saline
BKV	=	BK virus
bp	=	Base pairs
BRL	=	Bethesda Research Laboratories
CAT	=	Chloramphenicol acetyl transferase
cDNA	=	Complementary deoxyribonucleic acid
CMV	=	Cytomegalovirus
cpm	=	Counts per minute
CTD	=	Carboxyl terminal domain
CTF	=	CAAT transcription factor
DAB	=	DNA-TFIIA-TFIIB complex
DMEM	=	Dulbecco's modified Eagle medium
DMSO	=	Dimethylsulfoxide
DNA	=	Deoxyribonucleic acid
DNase I	=	Deoxyribonuclease I
DTT	=	Dithiothreitol
EC	=	Embryonal carcinoma
EDTA	=	Ethylenediaminetetraacetic acid
EGTA	=	Ethylene glycol-bis-(β-aminoethyl ether) N,N'-

tetraacetic acid

FCS = Fetal calf serum

β-gal = β-galactosidase

GTF = General transcription factor

GRE = Glucocorticoid response element

HEK = Human embryonic kidney cells

Hepes = H-(2-hydroxyethyl)-1-piperzinuthanesulfonic acid

HIV = Human immunodeficiency virus

HJC = Hamster fetal glial cells

HLH = α-Helix-loop-α-helix motif

HTH = α-Helix-turn-α-helix motif

Ig = Immunoglobulin

IgH = Immunoglobulin Heavy chain gene

IPTG = Isopropyl-1-thio-β-D-galactopyranoside

JCV = JC virus

JCV_E = JC virus early promoter-enhancer

JCV_L = JC virus late promoter-enhancer

kDa = Kilodaltons

MOPS = 3-(N-morpholino) propanesulfonic acid

mRNA = Messenger ribonucleic acid

NF1 = Nuclear factor 1

NP-40 = Nonidet P-40

nt = Nucleotide

ORF = Open reading frame

PAGE = Polyacrylamide gel electrophoresis

PBS	=	Phosphate buffered saline
PCR	=	Polymerase chain reaction
Penta	=	Pentanucleotide AGGGA sequence
PHFG	=	Primary human fetal glial cells
PML	=	Progressive multifocal leukoencephalopathy
PMSF	=	Phenylmethylsulfonyl fluoride
Pol	=	Polymerase
pRb	=	Retinoblastoma gene protein
RA	=	Retinoic acid
Rb	=	Retinoblastoma gene
RNA	=	Ribonucleic acid
RP1	=	Repressor protein 1
rpm	=	Revolutions per minute
RSV	=	Rous sarcoma virus
SD	=	Standard deviation
SDS	=	Sodium dodecyl sulfate
Sp1	=	Specificity protein 1
SSC	=	Sodium chlorid and sodium citrate buffer
SV40	=	Simian virus 40
t antigen	=	Small tumour antigen
T antigen	=	Large tumour antigen
TAF	=	TBP associated factor
TBP	=	TATA box binding protein
TBS	=	Tris buffered saline
TFII	=	Transcription factor for Pol II

TLC	=	Thin layer chromatography
UD	=	Undifferentiated
USA	=	Upstream factor stimulator activity
UTR	=	Untranslated region
UV	=	Ultraviolet
YB-1	=	Y-Box binding protein 1
YY1	=	Ying-yang 1 transcription factor

CHAPTER 1

INTRODUCTION

JC virus (JCV), the etiologic agent of progressive multifocal leukoencephalopathy (PML), is one of the human viruses that can cause tumours in other animals. Recently, PML has been frequently identified as a complication of AIDS. However, the studies of JCV have been hindered by the difficulty of growing the virus in culture and in animals, due to its strict restriction to glial cells of brain origin. Studies by the research group of Khalili have shown that the strict glial specificity of JCV is regulated at the transcriptional level. Subsequent studies have focused mainly on the transcriptional regulation of JCV that conferred the cell specificity of JCV.

1.1 Mechanism of transcription

Knowledge of protein-encoding gene transcription in eukaryotes by RNA polymerase II (RNAPII) has been intensively increased in the last decade, especially in the area of transcription initiation. Transcription initiation is a multi-step process involving several basal transcription factors. These factors assemble to form a multi-protein complex on the promoter-enhancer region, the central part of

transcriptional regulation. Because its nature is complex, the intricate protein-protein interactions of this machinery provide a fine-tuning system such that gene expression can be elaborately regulated. Through this machinery, sequence specific transcription factors can select genes to turn on or turn off at the start site of transcription. In the following sections, I will briefly summarize this transcription system.

1.1.1 Promoters and enhancers

Promoters have been defined as DNA sequences about 15 to 35 bp up from the initiation site. These sequences function as transcription signals to which basal transcription factors and RNAPII can bind and initiate mRNA synthesis (Zawel and Reinberg, 1993). Two kinds of promoters have been identified. One has the TATA box with or without the initiator (Inr) and the other has only the Inr. One or both elements are present in all protein coding genes and form minimal promoters. Either of them can form transcription initiation complexes (basic transcription machinery) with RNAPII and the general transcription factors (GTFs) to start basal transcription. In contrast, enhancers are elements with different sequences at variable distances from the initiation site and they are recognized by sequence specific DNA binding proteins. It is through interactions with these enhancer elements that transcription activators or repressors regulate transcription

activity. The combination of promoters and enhancers provides the specific characteristics of gene expression control.

1.1.2 RNA polymerase II

RNAPII transcribes the class II genes of the eukaryotic system and is a multi-subunit complex. In HeLa cells, RNAPII contains 10 subunits with molecular weights ranging from 240 to 10 kDa. The largest subunit of eukaryotic RNAPII contains a unique carboxyl terminal domain (CTD) that is not present in prokaryotic RNAP and eukaryotic RNAPI and RNAPIII. This CTD contains multiple consensus sequence repeats that are rich in serine and threonine residues and can be highly phosphorylated. The CTD of RNAPII can bind to DNA by intercalating into DNA nonspecifically to strengthen the binding of RNAPII and the stability of the initiation complex. CTD also has been suggested to mediate transcription activation of the upstream regulators based on Gal4-deletion mutants complementation experiments. The redundant heptapeptide repeat motifs present in CTD may act by interacting with the activation domains of some transcription factors. In this way, RNAPII could be accurately positioned in the promoter region and start transcription from a defined site by interacting with other GTFs.

1.1.3 General transcription factors and transcription

initiation complex

For the transcription of eukaryotic protein coding genes, about 20 proteins need to assemble at the promoter region to start the transcription initiation in addition to RNAPII. These proteins are GTFs, including TFIID, IIB, IIA, IIE, IIF, IIH and their associated factors.

TFIID is the only GTF that can specifically bind to TATA containing promoters of the class II genes. Recent studies have shown that TFIID is composed of a TATA-binding protein (TBP) and TBP-associated factors (TAFs) (Gill and Tjian, 1992). TBP is a eukaryotic protein that is highly conserved, especially in the carboxyl-terminal half of this protein that binds to the minor groove of the TATA element (Klug et al., 1993). This binding is the first step in the formation of the initiation complex. The amino-terminal domain of TBP, on the other hand, is highly divergent and forms the outer surface of the protein that is accessible for interaction with other transcription factors. In in vitro experiments, TBP can be separated from TAFs and supports basal transcription, but not activated transcription. However, the addition of purified TAFs can fully restore the activation response, suggesting that TAFs may act as coactivators or adaptors.

The human TFIIB is a 34.8 kDa protein. This protein can directly bind to the TBP-DNA complex or DNA-TFIIA-TFIIB complex (DAB complex). TFIIB is also required to recruit

RNAPII and TFIIF and to form the DAB-RNAPII-TFIIF (DAB-RNAPIIF) complex. TFIIB has two structural domains that correspond to its function: An amino-terminal putative zinc-finger domain with RNAPII recruiting activity and a carboxyl-terminal domain that can bind to the TBP-DNA complex (Ha et al., 1993; and Buratowski et al., 1994). In one case, TFIIB and RNA polymerase II alone can direct basal transcription on a supercoiled template DNA in the presence of Yin-Yang 1 transcription factor (Usheva and Shenk, 1994). Recent studies have suggested that TFIIB may play an important role in the interaction between the initiation complex and upstream activators.

TFIIF contains two subunits with molecular weights of 30 and 74 kDa, also known as RAP30 and RAP74. This factor can interact with RNAPII in the absence of other factors or DNA, suggesting that TFIIF is critical in recruitment of RNAPII to the assembling initiation complex. The recruitment is carried out by its small subunit via an interaction with TFIIB (Flores et al., 1991; and Ha et al., 1993). TFIIF also can affect transcription elongation that requires both subunits of TFIIF. Studies have showed that the activities of TFIID, TFIIB and TFIIF are enough to recruit RNAPII to form a stable initiation complex and to start transcription. This stable initiation complex was called minimal initiation complex or preinitiation complex. In the case of transcription of the immunoglobulin

heavy chain (IgH) gene *in vitro*, Parvin and Sharp (1993) found that TFIIF was not required for the IgH basal transcription, providing that the DNA template was negatively supercoiled.

TFIIE is a tetramer of two subunits encoded by two genes. It enters the initiation complex after the association of RNAPII/TFIIF with the DAB complex. The entrance of TFIIE may be mediated, in part, by directly interacting with RNAPII, TBP, TFIIF and TFIIH. The entrance of TFIIE is necessary for the subsequent incorporation of TFIIH into the complex (Flores *et al.*, 1991).

TFIIH is the last basal transcription factor incorporated into the preinitiation complex to form the complete initiation complex. It is a multi-subunit protein and possesses helicase, ATPase and CTD kinase activities. Therefore, the binding of TFIIH to the initiation complex will probably phosphorylate the carboxyl-terminal repetitive sequence of RNAPII. However, TFIIH is not required in all reconstituted basal transcription systems. Goodrich and Tjian (1994) used an abortive initiation assay to separate early and late initiation and showed that TFIIB, TFIIF and RNAPII can initiate mRNA synthesis. TFIIE, TFIIH and ATP were required to produce longer transcripts. This result suggested that TFIIE and TFIIH may play a role in promoter clearance, converting an initiation complex to an elongation complex. Interestingly, TFIIH also has been shown to function in

nucleotide excision repair of DNA and therefore, couples transcription and replication (Drapkin et al., 1994).

1.1.4 Sequence specific transcription factors

In addition to basal transcription factors, there are other transcription factors regulating transcription. These additional transcription factors can be classified as activators or repressors of transcription. By interacting with DNA in a sequence-specific manner, those factorsregulate gene expression, and thus cell differentiation andcell growth (Pabo and Sauer, 1992). To exert the above functions, a transcription factor usually containsaspecific DNA binding domain, a multimerization domain (either a homo- or hetero-domain) and a transcription activation domain.

To regulate transcription, the first step is the binding of an activator or repressor to DNA in a specific way through its DNA binding domain. Until now, six families of DNA-binding proteins have been identified according to the structure of their DNA binding domain. The first DNA binding motif identified was α-helix-turn-α-helix (HTH). In addition to the multimerization and activation domains, the proteins of this family usually contain an α-helix, a turn and a second α-helix in the carboxyl-terminal, such as seen in E. coli CAP protein and Cro protein. The homeodomain is another DNA binding motif existing in a large family of eukaryotic

regulatory proteins. This domain is usually a stretch of 60 amino acids that also contains an HTH motif. However, the homeodomain can bind to DNA independently. Two examples are Drosophila Antp and Oct-2 transcription factors (Scott et al., 1989). The zinc finger is a major structural motif for protein-DNA interaction first identified in transcription factor III A (TFIIIA) that participates in the transcription initiation of class III genes. Later, it was found that the zinc finger is widely distributed in transcription factors of eukaryotic organisms. Zinc finger proteins, such as Gal4, TFIIIA, YY1, usually contain sequence of Cys-X_{2-4}-Cys-X_{12}-His-X_{3-5}-His in 30 residue repeats in tandem. Later studies have extended this definition to any sequence motifs with a set of cysteine or histidine residues clustered in a short polypeptide chain. In this polypeptide chain, two cysteines and two histidines or cysteines coordinate a central zinc ion to form a zinc finger and bind the major groove of DNA (Pabo and Sauer, 1993; Rhodes and Klug, 1993).

Leucine zipper proteins belong to another family of transcription factors with the characteristic of forming dimers, either homodimers or heterodimers. The DNA binding domain is composed of two distinct subdomains in a short peptide of 60-80 residues: a leucine zipper region for dimerization and a basic region for DNA binding. The typical example of this family is AP1 proteins formed by a Jun and a

Fos protein. Similar to leucine zipper proteins, the proteins of the helix-loop-helix (HLH) family also have a DNA binding basic region and a dimer formation region composed of an α-helix, a loop and a second α-helix. An example is MyoD (Benezra et al., 1990; Pabo, 1992). Both types of proteins are playing important roles during differentiation and development. Furthermore, both types can function through the formation of heterodimers or homodimers.

The β-sheet or β-ribbon is another DNA binding motif that exists in some prokaryotic proteins, such as MetJ, Arc and Mnt repressors. Through the formation of two anti-parallel β-sheets, these proteins can symmetrically match the double-helical DNA and stably bind to a specific site (Kim, 1992).

There are other DNA binding motifs in other DNA binding proteins, like the POU domain and CCHC motif. However, their structures are not as clear as the above domains and will not be discussed here.

In contrast to the DNA binding domain, the structure of the activation domain is poorly understood. Several types of activation domains have been identified and are defined as acidic, glutamine rich and proline rich domains. Different classes of activation domain do not interact with the same target. For example, glutamine-rich and acidic rich activators are functionally distinct. By mutagenesis experiments, Gill et al. (1994) found that the most important

amino acid residue responsible for activation is not the glutamine or the acidic amino acid in those motifs. In some proteins, hydrophobic residues are important for activation, suggesting that those transcription factors activate transcription through hydrophobic forces. These kinds of weak forces could allow activators to contact multiple targets and form more stable complexes. Their activation domains showed no secondary structure when isolated. Thus, when activators interact specifically with other factors, they may undergo a conformation change to form a specific three-dimensional structure ("induced fit"), providing flexibility in the protein-protein interaction.

To regulate transcription, a sequence-specific transcription factor first binds to DNA and then regulates transcription through protein-protein interactions. Some auxiliary factors, also called accessory proteins, are required in the function of the sequence specific transcription factors. These auxiliary factors might act positively or negatively in regulating basal transcription, suggesting that these factors act in a multiple protein-protein interaction manner. Such interactions would affect the function of the sequence specific factor that binds the DNA.

The most obvious targets of these sequence-specific transcription factors are basal transcription factors. In

1990, Stringer et al. first identified such an interaction between the TBP and the acidic activation domain of VP16. Since then, many other factors have been found to bind to TBP. Subsequent studies found that other basal transcription factors, such as TFIIA, TFIIB, TFIIE, TFIIF and TFIIH, also could be the targets of activators. For example, Lin and Green (1991) found that VP16 can bind to TFIIB.

1.1.5 Coactivators and other factors

In 1990, Pugh and Tjian found that only partially purified TFIID, but not TBP, could support activated transcription, suggesting that the interaction between activators and basal transcription factors is not sufficient. Other factors present in the partially purified TFIID are also required for the activation. The finding of that TFIID is composed of TBP and TAFs indicates that the TAFs may act as coactivators linking the activation domain of the sequence-specific activator with the GTFs of the transcription initiation complex. These TAFs may also function to stabilize the binding of TBP to the TATA box (Lewin, 1990). Because there are at least eight human TAFs (250, 150, 110, 80, 60, 40, 30α, 30ß), different TAFs may interact with different transcription activators or repressors. Such interactions have been shown by Hoey et al. (1993) and Gill et al. (1992). They found that the glutamine-rich domain of TAF II110 and

activation domain of SP1 could interact directly. This interaction directly corresponds to the transcription activity. Moreover, the activation domain of VP16 was also found to interact with TAFII40. Furthermore, the disruption of this interaction abolished activated transcription but not basal transcription (Goodrich et al., 1993). These studies strongly demonstrated that TAFs are functioning as coactivators in the regulation of transcription. By using a temperature sensitive TAF mutant and activators SP1 and Gal4-VP16 fusion protein, Wang and Tjian (1994) demonstrated that while the basal transcription initiation complex can be stably reconstituted and functioning without TAFs, TBP-TAFs complexes are essential for activated transcription. They suggested that these factors, GTFs, TAFs and activators, may function at different times during the process of initiation and transcription.

Other protein-protein interactions also play an important role in regulation of transcription, such as the interactions between different sequence-specific transcription factors. One example is the interaction between the transcription regulator YY1 and c-myc. YY1 has been demonstrated to be an activator for transcription, a transcriptional repressor or a transcriptional initiator, depending on the context in different promoters. Using the yeast two-hybrid system, Shrivastava et al. (1993) found that YY1 could interact

directly with c-myc. The association of c-myc with YY1 could inhibit both activation and repression functions of YY1. It has been known that YY1 regulates many transcription factors, including c-myc, and c-myc could negatively autoregulate c-myc gene transcription. Therefore, c-myc may regulate its own transcription by associating with YY1. In another example, transcription factor E1A could relieve the repression of YY1 through direct interaction, mcdulating YY1's activation (Shi et al., 1991).

1.2 JCV human papovavirus

1.2.1 General features of JCV

JCV belongs to the polyomaviruses subfamily of the papovaviridae family. The papovaviridae family is composed of two subfamilies, the papillomaviruses and the polyomaviruses. I will discuss only the polyomaviruses, to which JCV belongs. The name of the polyomaviruses is derived from early studies showing that these viruses can produce tumours at multiple sites (Steward et al., 1957). Later, studies on this subfamily and the isolation of different polyomaviruses were well documented (Melnick et al., 1974). Polyomaviruses include mouse polyoma virus, simian virus 40 (SV40) and human BK and JC viruses (BKV and JCV). They do not cause tumours in their original hosts. However, they can cause tumours in experimental animals and induce transformation in cell

culture. Among these viruses, SV40 is the most extensively investigated virus. In 1960, it was discovered in polio vaccines made from rhesus monkey kidney cell culture. After infection, this virus can cause tumours in newborn hamsters. BKV and JCV are human viruses which were isolated in 1971 and were found to be related to SV40. Both viruses are widely distributed in the population and infect humans in childhood.

Polyomaviruses have a double stranded, covalently closed circular DNA genome of about 5,000 bp and an icosahedral capsid. Due to the small size of their genomes, they rely absolutely on cellular machinery for their replication and transcription. Thus, polyomaviruses provide a powerful system to study the cellular processes of DNA replication and RNA transcription.

JCV is one of the widely distributed polyomaviruses prevalent in the human population (Padgett et al., 1973). This virus is highly oncogenic and is one of the human viruses causing solid tumours in animals (Padgett et al., 1977; London et al., 1978). JCV was first isolated from the brain tissue of a patient with progressive multifocal leukoencephalopathy (PML), a rare fatal demyelinating brain disease usually linked with immunodeficient individuals (Padgett et al., 1976). PML is characterized by the presence of a large area of demyelination in the brain induced by JCV infection, mostly in an area of oligodendrocytes. Due to its consistent

association with PML, JCV is believed to be the etiologic agent of PML disease (Dorries, 1984). Recently, PML emerged as a frequent complication in AIDS patients. The infection of brain tissues by JCV and the ability of tat protein of HIV to activate JCV expression may explain this high incidence (Tada, et al., 1990). Recently, JC virus also has been detected in many human brain tissues from patients without PML, in human urine and kidney tissues (Arthur et al., 1989; Chesters et al., 1983; White et al., 1992; and Kitamura et al., 1990).

JCV is a strictly human virus. The host range restriction of the polyomaviruses is exaggerated in JC virus. Although it has been shown to be oncogenic in various animals, no other animals have been shown to be a JCV reservoir (Padgett et al., 1977). Primary human fetal glial (PHFG) cells, which were used for the initial isolation of JC virus, remain the major cell type used to grow this virus. However, the difficulties of obtaining and cultivating these cells and their mixed cell population of astrocytes and spongioblasts, the precursor of oligodendrocytes, prevented their general use. Thus, the unavailability of a convenient cell system has blocked the studies of JCV. Although JCV has been found to productively infect the myelin producing Schwann cells of the peripheral nervous system, the yields are very low and the viruses produced are defective (Assouline et al., 1991). Good yields of JC virus have also been obtained in human embryonic

kidney (HEK) cells. However, the distinct viral plaques contain only defective, complementing species of viral DNA (Yoshiike et al., 1982).

Considerable efforts have been directed towards finding a more suitable cell system to overcome the disadvantages of PHFG and other brain cells. Most approaches have followed the method used to derive COS cells by transforming monkey kidney cells with an origin defective mutant of SV40 (Gluzman, 1981). To derive a permissive cell line for JCV propagation, PHFG cells were transformed by origin defective mutants of SV40 or JCV to establish SVG, cPOS and POJ cell lines (Lynch et al., 1991; Major et al., 1985; and Mandl et al., 1987). These three cell lines are permissive for nondefective JCV and support DNA replication of JCV T antigen mutants and JCV origin containing plasmids. POJ cells also had increases in the duration of lytic cycle and time of the peak virus production. Although several such drawbacks still existed, including recombination between integrated genomes causing complications and generating altered virus, these three lines all have greatly facilitated the understanding of JCV and offered an important alternative for the studies of JCV.

1.2.2 Genome of JCV

1.2.2.1 Early and late coding region of JCV

The genome of polyomaviruses contain DNA of approximately

5,000 bp. This DNA is a double stranded, covalently linked circular DNA molecule. Because the viral infection cycle is divided into early and late stages corresponding to the pattern of gene expression, the viral genome is also divided into the early and late regions.

In the early stage of infection, polyomaviruses produce proteins called tumour antigens. These proteins encoded in the early region are different in size and are named large T, middle T and small t antigens. Because their coding regions are overlapping, these proteins are most probably produced by alternate splicing. However, only large T and small t antigens have been identified in JCV.

In the late stage of infection, polyomaviruses produce capsid proteins from the late region of the JCV genome named VP1, VP2 and VP3. Because the coding sequences of VP2 and VP3 are overlapping, VP3 has been considered as a subset of VP2. SV40 and the related viruses, BKV and JCV, encode another protein in the leader region of late mRNA. Since the function of this protein is still unknown, it has been named the agnoprotein (Eckhart, 1990). JCV is one of the polyomaviruses family members and its genome (Fig. 1) is very similar to other members such as SV40 and BKV (Frisque et al., 1984).

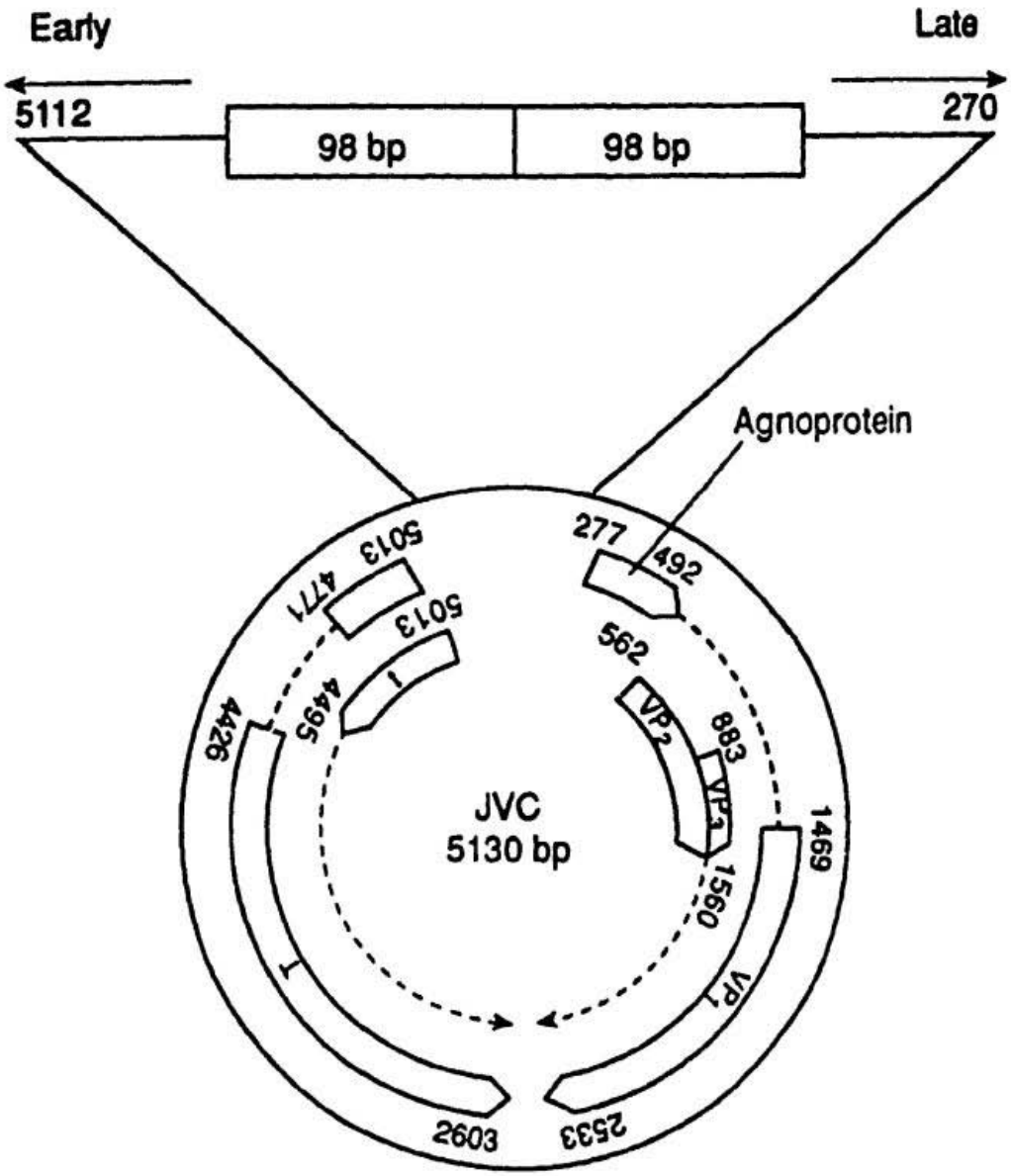
1.2.2.2 Regulatory region for JCV gene expression

Between the early and late coding regions, there is an

untranscribed 393 bp region spanning nucleotide positions (nt) 5112-276. This region contains the origin of replication (core ori), replication control elements and two classes of transcriptional elements: promoters and enhancers. Transcription proceeds bidirectionally from sites near control elements (Fig. 1). In contrast with papillomaviruses, which transcribes from a single strand in one direction, polyomaviruses transcribe from both strands of DNA in different orientations: early mRNA from one strand and late mRNA from the other strand (Frisque et al., 1984).

Fig. 1. Genome of JCV MadI strain

The top open boxes are the two 98 bp repeats of the JCV control region. Open boxes on the right side are open reading frames (ORFs) encoding the late proteins: agnoptotein, VP1, VP2, and VP3. On the left side, the open boxes show ORFs encoding early proteins: large tumour and small tumour antigen (T and t, respectively). The dashed lined represent the remaining early and late transcripts. The numbers indicate the nucleotide position. The drawing is modified from Frisque et al., 1984.


Early
Late
5112
270
98 bp
98 bp
Agnoprotein
277
492
562
VP2
883
VP3
1560
1469
VP1
2533
2603
T
4426
4495
T
5013
4771
5013
JVC
5130 bp

Comparing the sequences of JCV, BKV and SV40, it was found that the replication origin (including T antigen binding site) was the most conserved (82%) and the transcriptional element sequences were the least conserved (52%). In SV40, the noncoding region contains six copies of the GC rich binding site for the transcription factor, SP1, and two 72 bp tandem repeats. Unlike BKV and SV40, the JCV TATA box is duplicated in the two 98 bp tandem repeats. These two 98 bp repeats also contain multiple cellular transcription factor binding sites. Moreover, the SP1 site of BKV and SV40 is not present in the JCV Mad1 strain regulatory region. Furthermore, unlike for BKV and SV40, the propagation of JCV is strictly restricted to glial cells. As the promoter-enhancer region is the least conserved region among these three viruses, it has been suggested that the regulatory region of JCV is responsible for its cell and tissue specificity (Kenney et al., 1984).

The initially isolated JCV containing duplicated TATA sequences has been called the Mad1 prototype strain which was first isolated in Madison, WI. (Martin et al., 1985). Later studies had shown that many JCV variant strains exist (Grinnell et al., 1983). Different variants were recovered from the brain (GS/B) and kidney (GS/K) tissue of a single PML patient by Dörries (Dörries, 1984). Results showed that although the viral DNA isolated from one organ was the same in

size, the viral DNA from the brain and kidney tissues differed by about 120 nucleotides. Detailed examination of viral DNA sequence revealed only one minor difference in their coding region. All the other variations were in the regulatory regions (Loeber and Dörries, 1988). The most significant difference was that the GS/K genome contains only one tandem repeat containing the TATA box and the GS/B genome contains two tandem repeats. In another study, only one regulatory region was found in the JCV DNA from urine of normal people. This single regulatory region of JCV was identical to that of GS/K (Yogo et al., 1990). Thus, it was suggested that the JCV of GS/K strain represents the archetype strain of JCV and the prototype strain from the brain, represented by Mad1, might have evolved through regulatory region recombination and duplication of the GS/K strain. This additional promoter-enhancer element could explain the higher propagation and expression of JCV prototype strain Mad1 in brain cells compared with the low propagation of the archetype strain of JCV in kidney cells. However, it is also possible that GS/K strains were derived from GS/B by the loss of one copy of tandem repeat (Yoshiike et al., 1982). This discrepancy awaits further investigation to determine whether these changes are due to the adaption of JCV to propagate in brain cells.

1.2.3 JCV proteins

1.2.3.1 Early proteins

Polyomavirus early gene expression produces three T antigens through splicing from a single early mRNA molecule. After being synthesized in the cytoplasm, these antigens are distributed to different cellular positions. Among them, large T antigen mainly functions in the nucleus to participate in viral DNA replication and gene expression. Amino acid sequence analysis has found that SV40 T antigen contains a nuclear location signal. When attached to the other proteins, this signal can direct those proteins to the nucleus.

Large T antigen of the polyomaviruses is a multi-function protein. First, it can activate cellular gene expression by binding to the promoter-enhancer region of cellular genes and directly interacting with multiple factors in the transcription complex (Lane et al., 1985; and Gruda et al., 1993). These genes usually encode proteins regulating cellular DNA replication and the cell cycle. Thus, their expression could facilitate the replication of viruses and viral gene expression. Second, T antigen can specifically bind to viral DNA in the vicinity of replication origin and initiate replication. A single amino acid change can disrupt this function (Manos et al., 1985). Third, extensive studies of large T antigens of polyoma viruses and SV40 have shown that large T antigens have ATPase activity. Fourth, SV40

large T antigen has helicase activity required for viral DNA replication (Stahl et al., 1986). Fifth, studies also showed that the amino-terminal region of T antigen can activate cellular DNA and ribosomal RNA synthesis. This function is separable from its function for viral DNA replication (Clark et al., 1983). In addition, the large T antigen of SV40 and polyoma virus can immortalize primary cells in tissue culture. This is probably through T antigen's interaction with the tumour suppressor gene protein, p53, and retinoblastoma susceptibility gene (Rb) proteins (pRb), which are regulators of cell growth (Colbey and Shenk, 1982; DeCaprio et al., 1988; and Shohat et al., 1987).

Studies based on immunofluorescence, immunoprecipitated tryptic peptides and sequence analysis have demonstrated that the structure of large T antigen of JCV is closely related to the T antigen protein of SV40. Thus, the T antigen of JCV has been predicted to have many properties similar to those identified in SV40 T antigen (Frisque et al., 1980; Frisque et al., 1984; and Simmers et al., 1978). It has been confirmed that JCV T antigen has transforming and transregulating activities (Nakshatri et al., 1990). JCV T antigen also stimulates JCV DNA replication. However, all the activities demonstrated by JCV T antigen are not as effective as those for the SV40 T antigen. Comparing JCV T antigen with SV40 T antigen, Dyson et al. (1990) and Bollag et al. (1989)

showed that JCV T antigen is less efficient. The binding of JCV T antigen to DNA, p53 and pRb is not very stable and is lower than that of SV40 T antigen. These differences may lead to JCV's low transforming activity and its strict cell specificity comparing with those of other polyomaviruses.

Small t antigens of polyomaviruses are produced by alternative splicing of early transcripts. Due to the absence of good antisera to t antigen, small t antigen of JCV was identified only recently by immunoprecipitation. Therefore, the small t antigen of JCV has not been studied extensively. It was found that the small t antigens of JCV and SV40 are highly homologous in the region overlapping the large T antigen coding region, but only have 53% homology in the carboxyl-terminal region. By using a JCV/SV40 hybrid genome, Haggerty et al. (1989) found that the small t antigen of JCV is not as effective as the SV40 counterpart in enhancing transformation, most probably due to the poor expression of large T antigen of JCV. For SV40, its small t antigen can bind specifically to several cellular proteins and play a role in viral DNA accumulation.

1.2.3.2 Late proteins

The late regions of polyomaviruses encode the viral proteins, VP1, VP2 and VP3 and agnoprotein. VP1, VP2 and VP3 can enter the nucleus through their nuclear localization

signals. VP1 is encoded in the 3'-end of the late region in a mRNA of 16S. It is the major virus-encoded capsid protein with both structural and biological functions in the virion particle. After translation, polyoma and SV40 VP1 undergoes posttranslational modification by the host cell machinery, resulting in several VP1 species with distinct isoelectric points (pI). This has been suggested as a common feature of the papovaviruses (Bolen et al., 1981). VP1A, one of the VP1 species, is tightly associated with virion chromatin core. Comparing with the results of SV40 (Brady et al., 1981), Bolen et al., (1981) suggested that the function of VP1A is to maintain the viral DNA in a proper conformation to increase the potential accessibility to RNA polymerase. Other species such as VP1D, VP1E and VP1F can be phosphorylated and act for viral attachment during cell infections (Bolen et al., 1980). Although there is a strong crossreaction between the VP1 proteins of JCV, BKV and SV40, modifications similar to those in SV40 protein have not been identified in JCV VP1 protein (Shah et al., 1977).

VP2 and VP3 are minor capsid proteins that are encoded in the 5'-end of the late region and are translated from the same 18S to 19S mRNA. Thus, VP2 and VP3 have common C-terminal amino acid sequences. The functions of VP2/VP3 are not very clear. Studies have shown that the SV40 proteins do not associate with viral DNA (Brady et al., 1980). One possible

function of these two proteins is that they are the components of the 12 capsomeres lying at the vertices of the icosahedral shell. The functions of VP2 and VP3 have not been studied in JCV. Based on the sequence similarity to that of SV40, however, a similar function and modification has been predicted.

JCV, BKV and SV40 also encode another short polypeptide in the leader sequence of late region. This polypeptide contains 61 amino acids in SV40 and 71 amino acids in JCV and has been named the agnoprotein for its unknown function. The first two-third of the agnoprotein is highly homologous in all three viruses and the carboxyl third peptide is completely different. Studies of deletion mutants in this coding region of SV40 have suggested that the carboxyl third peptide was not essential for the propagation of JCV. Further studies based on genetics and biochemistry suggested that this peptide may function in the late lytic cycle (Jay et al., 1981; and Ng et al., 1985). Moreover, its capability of binding DNA suggested that this agnoprotein may affect the transcription of late mRNA (Alwine et al., 1982).

1.3 Transcription factors that regulate JCV expression

To identify transcription factors regulating the cell specificity of JCV, Khalili and coworkers initiated the search for such factors in 1988 (Khalili et al, 1988). They used

oligonucleotides spanning the JCV 98 bp repeat region, in regions they called A, B and C, to perform mobility shift assays and UV crosslinking assays with human fetal brain cell nuclear extracts and HeLa cell nuclear extracts. They demonstrated that four proteins ranging from 45 kDa to 230 kDa specifically bind to JCV control regions A, B and C. Of these regions, regions A and C were bound by proteins of the same size in both glial and nonglial HeLa cell nuclear extracts. However, region B, the central region of the JCV 98 bp repeats, showed tissue dependent diversity in binding proteins of HeLa cells and glial cells. The sizes of these proteins were 45 kDa in glial cells and 85 kDa in HeLa cells. The Khalili group went further to partially purify a 45 kDa protein from calf brain tissue (Ahmed et al., 1990). This brain specific protein bound specifically to the JCV nt 132-160 (region B) and stimulated JCV early expression in a nonglial HeLa cell extract. This result suggested that this protein may be responsible in part for the early promoter specificity of JCV.

Amemiya et al. (1989) and Tamura et al. (1988) performed similar studies using mobility shift assays and DNase protection assays with glial and nonglial cell extracts. In addition to the sites bound by similar proteins in both glial and nonglial cells, they identified three sites in the regulatory region that showed differences in binding to

factors from glial and nonglial cells. These sequences are located following the A-T rich region in region A and B in the 98 bp repeats, and in region C outside 98 bp repeats. Because all these sites showed sequence homology with the NF-1 binding motif of adenovirus 2, and because purified NF-1 can protect site B of JCV, they suggested that the protein interacting with all three sites is NF-1 or an NF-1-like protein. However, these three sites showed different affinities for interaction with this protein. In agreement with Khalili's work, they also observed that an oligonucleotide containing the consensus CTF binding motif was incapable of competing for the binding of those three sites. Therefore, they suggested that the JCV-binding NF1 may be a specific member of the family of CTF/NF1 related factors.

The experiments done by Nakshatri et al. (1990) in our laboratory extended the above results. They tested the activity of the regulatory sequences of JCV in P19 embryonal carcinoma (EC) cells containing the CAT reporter plasmid. Their experiments showed that the JCV enhancer only activated CAT expression in retinoic acid-differentiated P19 cells, a mixture of glial cells and neurons, but not in DMSO-differentiated cardiac and skeletal muscle cells or undifferentiated P19 cells. These results confirmed JCV's neural restricted cell specificity. DNase footprinting experiments were performed to further correlate JCV's promoter

and enhancer activity with the interaction of DNA-transcription factors. Consistent with Amemiya et al. (1989), three sites containing NF1 consensus sequences named I, II, III were protected (Nakshatri et al., 1990). The TATA box was also partially protected. These protections can only be observed when using RA-differentiated P19 cell nuclear extracts, but not with undifferentiated and DMSO-differentiated cell nuclear extracts. Moreover, oligonucleotide competition experiments showed that all three sites were bound with different affinities by the same protein. Furthermore, the competition experiments suggested that adjacent regions were bound cooperatively by NF1 or a NF1 like protein.

The studies conducted by Kumar et al. (1993) in our laboratory defined the function of these three NF1 cis-elements. These motifs within the JCV regulatory region re examined by measuring the CAT activity following site-directed mutagenesis. Results showed that while mutation of DNase I-protected region I (NF1 I) did not affect CAT activity, the mutation of NF1 II significantly lowered CAT expression in glial cells. If both region II and region III were mutated, the CAT activity in glial cells was further reduced. These results, confirmed by in vitro transcription assays, suggested that regions II and III containing 6 bp NF1 perfect inverted palindromic sequences were important for JCV

expression in glial cells. In contrast, the NF1 site I outside 98 bp repeats was not essential. Therefore, duplication of NF1 II/III in brain-type JCV was correlated directly with enhanced JCV expression, specifically in brain-type cells.

In addition to being activated by NF1 for early and late expression, JCV is also regulated by other positive and negative cellular and viral proteins. Wagner et al. (1993) presented evidence that cotransfection of glial transcription factor, Tst-1, can activate early and late gene expression of JCV. They found that Tst-1, a POU-domain family, activated gene expression by interacting with the JCV regulatory region. Due to the restriction of Tst-1 to the myelin-producing Schwann cells of the peripheral nervous system and oligodendrocytes of the central neurons system, the expression of this protein was strongly correlated with the specificity of JCV expression to glial cells and JCV infection to oligodendrocytes.

Recent discoveries that PML was repeatedly found as a complication in AIDS patients have promoted the thinking that HIV may have contributed to the pathogenesis of this disease. Tada et al. (1990) examined the possible role of an HIV regulatory protein in regulating JCV gene expression. They found that cotransfection of the tat gene, encoding an HIV-1 trans-regulating protein, could activate JCV late activity in

glial cells. In contrast, JCV T antigen can stimulate the JCV late promoter in glial and nonglial cells. This suggested that the tat protein and T antigen might activate JCV expression by different mechanisms. These data provided an interesting mechanism to explain the high incidence of PML in AIDS patients. Chowdhury et al. (1993) went further to identify the cis-element that may be recognized by the tat protein. By using deletion mutants, a region upstream of the JCV late RNA initiation site that was necessary for tat activation in glial cells was identified. Further experiments with DNA fragment spanning this region inserted in a heterologous promoter showed a GA/GC rich sequence, GGAGGCGGAGGC, functions by interacting with tat in glial cells (Chowdhury et al., 1993).

JCV T antigen is another major transregulating protein of JCV expression. The JCV promoter-enhancer fragment inserted into a CAT expression vector in the late orientation was expressed only in glial cells (Lashgari et al., 1989). However, in the presence of T antigen, CAT activity could be significantly increased in both glial and nonglial cells. Moreover, S1 nuclease analysis and nuclear run-on transcription experiments showed that this stimulation was due to an increase in the level of transcription. They suggested that T antigen might function to change the expression of cellular factors and to remove negative transregulators from

the late control region (Lashgari et al., 1989). Our laboratory also demonstrated similar results, showing that JCV T antigen suppressed JCV early expression, while JCV late expression was activated in RA-differentiated P19 cells (Nakshatri et al., 1990).

JCV T antigen also contributes to the restricted replication of JCV DNA (Chuke et al., 1986). However, JCV T antigen does not support the replication of SV40 DNA, suggesting that there is an element within the JCV DNA that discriminates this JCV T antigen-mediated JCV DNA replication. Through a comparative study of the replication activity with JCV/SV40 chimeric origin, Lynch and Frisque identified an AGGGA sequence in the JCV promoter-enhancer that stimulates the JCV, but not SV40, T antigen-mediated replication (Lynch and Frisque, 1990). Because this sequence was identified later to interact with cellular DNA binding proteins, it is possible that AGGGA binding proteins might contribute to JCV T antigen-mediated replication through a direct interaction with T antigen and cellular factors. The AGGGA sequence is not present in the SV40 origin but is repeated seven times in the JCV Mad1 strain regulatory region. In variant JCV genomes, although one AGGGA repeat is disrupted by a 23-nt insertion, the insertion itself contains a copy of AGGGA repeat (Martin et al., 1985). These variants replicate with similar efficiency to that of Mad1 in COS and POJ cells.

Interestingly, a copy of the AGGGA sequence present in the BKV origin does not support BKV replication mediated by JCV T antigen in COS and POJ cells (Deyerle et al., 1989; and Lynch and Frisque, 1990). These results suggested that AGGGA sequence and the AGGGA binding protein might be involved in the cell specificity of JCV (Lynch and Frisque, 1990).

As JCV propagates only in glial cells in tissue culture, this highly restricted tissue tropism should be contributed to, at least in part, by the negative regulation of the JCV early and late gene expression. To identify such negatively acting cis-elements, Tada et al. (1991) created a series of deletion mutants in the regulatory region of JCV and assayed the expression of CAT in the late orientation. CAT assays revealed that a silencer was located within the poly(A)/pentanucleotide (penta) repeat region. By using a synthesized oligonucleotide, spanning the region from nt 22-42 cloned into the pA10CAT plasmid, they tested the CAT activity and found that AGGGAAGGGA, a penta repeat motif, might be the negative acting cis-element. They went further by performing mobility shift assays and UV crosslinking assays and identified a protein of 56-60 kDa interacting specifically with the oligonucleotide containing the penta repeat motif. Similar sequences are present upstream of the rat embryonic myosin heavy chain gene (5'-AGGGAAGGAGAAAAAAAAA-3') and the c-myc promoter (5'-AGGGAAAGGGA-3'). These sequences have been

suggested to function negatively in regulating the associated genes. Thus, they speculated that the penta repeat motif might function to downregulate JCV late gene expression (Tada et al., 1991).

Another group (Sharma et al., 1991) also detected penta repeat motif binding proteins. They first identified a 53 kDa protein in C6 rat glioma cells and HeLa cells and speculated that this protein functioned negatively in regulating expression of the JCV early genes, because the expression of JCV early gene promoter-enhancer is more strictly brain specific than is the expression of the late gene promoter-enhancer. However, a later study by the same group found another penta repeat motif binding protein of 70-80 kDa (Kumar et al., 1994). Therefore, how the penta repeat motif and its binding proteins regulate JCV expression remains unclear.

1.4 Cells to be used in this project

Previously in our lab, Kumar et al. (1993) used P19 embryonal carcinoma cells to delineate in detail the function of the NF1 consensus sequences in the JCV regulating region immediately adjacent to the penta repeat motif. P19 cells have shown great advantages in studying the regulation of JCV gene expression due to the strict glial cell specificity of JCV. Upon treatment by retinoic acid (RA), P19 cells will differentiate to a mixture of neurons and glial cells.

However, P19 cells also can be differentiated into muscle cells following the treatment with DMSO. This provides an advantage for studing the effects of the penta repeat motif on the regulation of JCV gene expression in glial and nonglial cells, in which the cell specificity of the JCV and the effects of cis-elements on the regulation of JCV gene expression can be studied directly within the same cell system. However, as JCV expresses more efficiently in human glial cells, a human glioblastoma cell line, U87MG, is also employed in this project.

1.5 Objectives of this study

The first objective is to examine the regulatory role of penta sequences in the expression of JCV in P19 RA differentiated cells and U87MG human glioblastoma cells through site-directed mutagenesis and in vivo expression assays.

The second objective is to identify and characterize the proteins that binds to the penta repeat motif by mobility shift assays and UV cross-linking assays using a synthesized oligonucleotide containing the penta repeat motif as a probe for proteins in glial and nonglial cell extracts.

The third objective is to isolate and characterize the cDNA clone encoding the penta repeat motif binding factor by screening a cDNA expression library from P19 RA cells and

testing the regulatory role of cDNA clone in glial cells.

Finishing of all the objectives will help us to understand both positive and negative roles of pentanucleotide repeat motif in JCV expression in glial and nonglial cells, to understand how the interaction of this cis-element and cellular protein regulating JCV promoter-enhancer activity.

CHAPTER 2

MATERIAL AND METHODS

2.1 Materials and reagents

T4 DNA ligase, calf intestinal phosphatase and restriction endonucleases were supplied by Bethesda Research Laboratories (Gibco-BRL). Reverse transcriptase was purchased from Life Sciences. The ^{14}C chloramphenicol used is a product of ICN. α-P^{32} dCTP was supplied by Amersham. Acetyl coenzyme A was from Sigma. Thin layer chromatography (TLC) plates and X-ray films were supplied by Kodak. Nitrocellulose filters were purchased from the Mandel Scientific Company.

Synthetic oligonucleotides were supplied by General Synthesis and Diagnostics (GSD Oligos). The mutagenesis kit used is the product of Bio-Rad. Poly dI.dC was from Pharmacia. The nick translation kit used is a product of Gibco-BRL. The nick-column Sephadex G-50 column was supplied by Pharmacia. The JCV early promoter-enhancer CAT reporter plasmid, pJCECAT, was constructed by H. Nakshatri (Nakshatri et al., 1990). The pSV0CAT plasmid was a kind gift of Dr. B. Howard. The pBLCAT plasmid containing JC late promoter-enhancer, was a gift of Dr. B. Luckow. The JC virus promoter-enhancer region was cloned into pUC19 (pUC19JC) by K.U. Kumar (Kumar et al., 1993). The Sequenase Version 2.0 sequencing

kit was supplied by the United States Biochemical Corporation. pRc/CMV expression vector was purchased from the Invitrogen Corporation. The pBluescript KS cloning vector is a product of Stratagene. The protein assay kit was obtained from Bio-Rad.

Alpha-Modified Eagle's Medium (α-MEM) and Dulbecco's Modified Eagle's Medium (DMEM) were supplied by Flow Laboratories. Trypsin-EDTA and all-trans retinoic acid are products of Sigma. Dimethyl sulfoxide (DMSO) was purchased from Baker. Fetal calf serum is from Bockneck Laboratories. P19 embryonal carcinoma cell line was a kind gift of Dr. H. Hamada. The U87MG glioblastoma cell line was from American Type Culture Collection (ATCC No. HTB14).

2.2 Cell culture and transfection

2.2.1 The cell culture of the P19 and U87 cells and the differentiation of P19 cells by RA or DMSO

P19 cells were maintained in α-MEM medium supplemented with 10% heat-inactivated fetal calf serum (FCS) at 37°C. To induce the differentiation of P19 cells, confluent cells were trypsinized and resuspended in fresh medium containing 300 μM all-trans retinoic acid (RA) or 10% (v/v)of DMSO. The cells were plated in bacterial plates to grow in suspension for 48 hours. After discarding the old medium, the cells were resuspended in fresh medium containing RA or DMSO and were

cultured for another 48 hours. After RA or DMSO treatment, the cells aggregated and the medium was cloudy. The cells were then washed with phosphate buffered saline (PBS), trypsinized to resuspend in fresh medium without RA or DMSO and plated in tissue culture plates at a dilution of 1:4. These differentiated cells were incubated for eight hours and were then ready for transfection (Rudnicki and McBurney, 1987).

U87MG cells were maintained in DMEM medium supplemented with 10% (v/v) heat-inactivated FCS. Twenty-four hours before transfection, the cells were harvested, trypsinized and replated in 100 mm tissue culture plates with a cell concentration of 5×10^5 cells per plate. The cells were incubated overnight.

2.2.2 The method of transfection

For transfection, the modified calcium phosphate-mediated transfection method was used (Chen and Okayama, 1987). Briefly, 10-15 μg plasmid DNA were mixed with 0.5 ml of 0.25 M $CaCl_2$, 0.5 ml of 2× N,N-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid (BES) buffered saline (BBS), and incubated at room temperature for 15 minutes. The DNA mixture was added to cells in a dropwise manner and plates were mixed by gentle swirling. Plates were incubated for 18 hours at 37 °C in the presence of 3% CO_2. The cells were then washed twice

with PBS. After adding fresh medium, the cells were incubated at 37°C in the presence of 5% CO_2 for another 48 hours before harvesting.

2.3 Site directed mutagenesis and plasmid construction

The Muta-Gene R Phagemid In Vitro Mutagenesis kit from Bio-Rad was used to generate the desired mutations in the penta repeat motifs. Briefly, the JCV PvuII-HindIII fragment from nucleotide 270 to 5112 (Fig. 2) was used as template for mutagenesis. This fragment was from pUC19JC. The JCV fragment in pUC19JC was digested with SalI and BamHI and then isolated and inserted into the same sites of pTZ19U by incubating at 16°C overnight in the presence of T4 DNA ligase. The product, pTZ19UJC, was introduced into E.coli CJ 236 bacteria to form a uracil containing plasmid. The single strand of pTZ19UJC was rescued by helper phage infection.

The mutation primers were synthetic oligonucleotides with the following sequences:

wild type NF1 and mutant penta, 5'-GGCCAGCCATTAGATCTGTTTTTTTT-3';

mutant NF1 and mutant penta, 5'-GGCCAGTACTTAGATCTGTTTTTTTT-3'.

These primers were used to anneal with the template single stranded DNA template and generate mutations in regions II and III, as described in the manufacturer's instruction. The mutations were screened by digestion with BglII that can

specifically recognize the AGATCT sequence created in the mutants. These mutants were further confirmed by sequencing.

The mutated JCV promoter-enhancer fragments were cleaved from pTZ by HincII and SmaI, blunt-ended and cloned into the blunt-ended HindIII site of the pSV0CAT reporter plasmid. The plasmids were assayed to identify the JCV early promoter (pSVJCECAT) or JCV late promoter (pSVJCLCAT) from which the chloramphenicol acetyltransferase (CAT) gene is expressed, depending on the orientation (Nakshatri et al., 1990).

For the construction of the repressor cDNA expression plasmid, the cDNA was first isolated from the cDNA expression library made from P19 RA differentiated cells, by screening with a probe containing penta repeat motif of JCV as described in section 2.8. The positive recombinant phage, RP1, was digested with SalI and NotI to liberate the cDNA insert. This cDNA was then cloned into the same sites in the pBluescript KS vector. The RP1 cDNA fragment was further cloned into the NotI and HindIII sites of pRc/CMV vector, supplied by Invitrogen for the protein expression. The final product was sequenced to confirm insertion and orientation and was named pCMV-RP1.

2.4 Chloramphenicol acetyltransferase (CAT) assays

2.4.1 Harvesting the cells

Cells were harvested 48 hours after transfection according to the method of Gorman (Gorman et al., 1982).

After washing the cells with PBS, 1 ml of Tris-EDTA-NaCl solution (0.04 M Tris-HCl pH 7.4, 1 mM EDTA, 0.15 M NaCl) was added to the plates. Cells were then scraped into eppendorf tubes with rubber policemen. The cells were pelleted and resuspended in 100 μl 0.25 M Tris-HCl, pH 7.8. After vigorous vortexing, the cells were subjected to three cycles of freezing in liquid nitrogen and thawing in a 37°C water bath for 5 min each with vortexing after each cycle. The cell debris was pelleted by centrifugation for 5 min at 4°C in an eppendorf centrifuge and the supernatant was taken for CAT assays or frozen for future assay.

2.4.2 Chloramphenicol acetyltransferase assays

For CAT assays, the 30 μl final reaction mixture included 10 μl of cell extract, 1 μl ^{14}C chloramphenicol, 4 μl of 4 mM acetyl coenzyme A, 1 μl of distilled water and 14 μl of 1 M Tris-HCl, pH 7.8. The reactions were incubated for 1 hour at 37°C and stopped by adding 0.5 ml of ethyl acetate to extract chloramphenicol and its products. Tubes were vigorously votexed for 30 seconds. After 1 minute of centrifugation, the supernatant was transferred to fresh tubes and dried in a speed vacuum apparatus. The dried samples were dissolved in 15 μl of ethyl acetate and applied to thin layer chromatography (TLC) sheets. Acetylated and non-acetylated forms of chloramphenicol were separated by TLC with chloroform

: methanol (95:5). The dried TLC sheets were exposed to X-ray film at room temperature for 24 to 48 hours. X-ray films were developed and the results of acetylation were quantified by liquid scintillation counting of matching spots of acetylated and non-acetylated forms in the X-ray films to the TLC sheets. The results were expressed as the percentage of acetylated chloramphenicol to the total acetylated chloramphenicol plus non-acetylated chloramphenicol. The β-galactosidase (ß-gal) assay was performed as described (Nakshatri et al., 1990). This ß-gal expression plasmid was originally from pCH110 (Pharmacia) and modified by replacing the SV40 promoter-enhancer with the Rous sarcoma virus (RSV) promoter-enhancer sequences. This plasmid was cotransfected with the test plasmids and the β-gal activity served as a control for transfection efficiency. Thus, the corrected CAT activity was the CAT/ß-gal activities ratio.

2.5. Nuclear and whole cell extract preparation

2.5.1 Nuclear extract preparation

The nuclear extracts were prepared according to the method described by Hennighausen and Lubon (1987). Briefly, 40 plates of either undifferentiated, DMSO differentiated, RA differentiated P19 cells or U87MG cells were collected for nuclear extract preparation. The cells were first washed with PBS, and then harvested by scraping with rubber policemen.

Harvested cells were pelleted by centrifugation at 1,800 rpm for 10 min at 4°C. Pelleted cells were resuspended and washed again in 10 volumes of PBS. The pelleted cells were then resuspended in five pellet volumes of buffer A containing 0.3 M sucrose, 10 mM Hepes-KOH at pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.1 mM EGTA, 0.5 mM dithiothreitol (DTT), 0.5 mM phenylmethylsulfonyl fluoride (PMSF), and 2 μg/ml each of the protease inhibitors, antipain, pepstatin A, and aprotinin. Cells were lysed in an all-glass Dounce homogenizer by 10-12 strokes with a B pestle. The lysed cells were further given 1-2 strokes in the presence of 0.3-0.4% Nonidet P-40 (NP-40). The homogenate was examined by microscopy for completion of lysis and centrifuged at 2,700 rpm for 10 min at 4°C. The nuclear pellet was washed twice in buffer A containing 0.3 M sucrose without NP-40.

Nuclei were resuspended in a Dounce homogenizer by 10 strokes with B pestle in 2.5 nuclear pellet volumes of buffer B containing 400 mM NaCl, 10 mM Hepes-KOH, pH 7.9, 1.5 mM $MgCl_2$, 0.1 mM EGTA, 0.5 mM DTT, 5% glycerol and 0.5 mM PMSF. The suspension was stirred for 30 min at 4°C and centrifuged at 39,000 rpm for 60 min in rotor type 75 Ti. Supernatant was collected and dialysed for 4 hours at 4°C against 50 volumes of Buffer C containing 20 mM Hepes-KOH, pH 7.9, 75 mM NaCl, 0.1 mM EDTA, 0.5 mM DTT, 20% glycerol and 0.5 mM PMSF. The extract was further centrifuged for 20 min at 20,000 rpm and

4°C in the same rotor to remove any debris. The supernatant was then aliquoted, frozen in liquid nitrogen and stored at -70°C. The protein concentration of the nuclear extract was measured by the method of Bradford (1976).

2.5.2 Whole cell extract preparation

Whole cell extracts were prepared by the method of Tasset et al. (1990). For preparation, 40 plates of cells were washed with PBS and harvested by rubber policemen. Cells were recovered by centrifugation at 1,800 rpm and 4°C for 10 minutes. After washing again with PBS, the cells were resuspended in two pellet volumes of extraction buffer containing 10 mM Tris-HCl, pH 7.6, 400 mM KCl, 1.5 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 10 µg/ml of soya bean trypsin inhibitor, 1 mM PMSF, 2 µg/ml each of the protease inhibitors, antipain, pepstatin A, leupeptin and aprotinin, and 20% glycerol. The cells were then subjected to three cycles of freezing for 5 minutes in liquid nitrogen and thawing for 30 minutes in ice. After centrifugation for one hour at 39,000 rpm at 4°C, the supernatant was aliquoted, quick frozen in liquid nitrogen and stored at -70°C. The protein concentration was measured as described for nuclear extract preparation.

2.6 DNA binding analysis

2.6.1 Preparation of the DNA probe

The complementary oligonucleotides used in the *in vitro* binding assays are as following:
Penta wild type oligo, plus strand: 5'-GATCCAGGGAAGGGAG-3'; and minus strand: 3'-GTCCCTTCCCTCCTAG-5';
Penta mutant oligo, plus strand: 5'-AAAAAAAAGtcTaGaa-3' and minus strand: 3'-TTTTTTTTCagAtCttCTAG-5'.

Mutated nts were typed in small letters and wild type nts were typed in capital letters. These oligonucleotides were annealed to form double-stranded DNA and were end-labelled by filling the BamHI ends with α-[^{32}P] dCTP by reverse transcriptase (Sambrook *et al.*, 1989). Unincorporated dCTP was separated from end-labelled DNA by passing the labelling reaction mix through the Sephadex G-50 Nick columns, or by precipitating with two volumes of 100% ethanol and washing with 80% ethanol.

2.6.2 Mobility shift assays

Mobility shift assays were performed as previously described (Chodosh *et al.*, 1988). The binding reactions were performed in a 15 μl final volume, using either nuclear extracts or whole cell extracts. The proteins were incubated with binding buffers containing 5.0 μg poly(dI-dC)·(dI-dC), 3.0 μg denatured salmon sperm DNA, 12% glycerol, 12 mM Hepes-NaOH (pH 7.9), 60 mM KCl, 4 mM Tris-HCl pH 7.9, 1 mM EDTA, 0.6 mM dithiothreitol, 300 μg/ml bovine serum albumin and

approximately 20,000 cpm (0.1-0.3 ng) end-labelled DNA probes. One assay without extract was used as negative control. All assays were incubated at 30°C for 30 min. The protein-DNA complexes were resolved by electrophoresis in 4% low ionic strength native nondenaturing polyacrylamide gels. The electrophoresis was run at 4°C with circulating 22 mM Tris-borate, pH 7.9, 0.5 mM EDTA buffer at 5 volts/cm. Following electrophoresis, gels were transferred to a Whatman 3 MM filter paper, dried in a gel dryer and exposed to X-ray film.

2.7 UV crosslinking assays

Complementary strands of penta wild type oligonucleotides were annealed to produce double-stranded DNA fragments. The annealed DNA was labelled by nick translation with α-[^{32}P] dCTP. UV cross-linking was performed according to the method of Chodosh et al. (1986). In the assays, a total of 50 μl of reaction mix was buffered by the same reaction buffer used in mobility shift assay, and contained 10^5 cpm of nick translated penta probe, 10 μg of poly (dI.dC)·(dI.dC), and 60 μg of whole cell extract. After incubation at 30°C for 30 minutes, the reactions were transferred to NUNC vials and covered with plastic wrap. The vials were UV-irradiated under a Fotodyne UV lamp at the wave-length of 310 nm for 30 minutes. The DNA was not treated with DNase I as described in the method, since the size of the probe was less than 50 bp. The reactions were

stopped by adding equal volume of 2× SDS sample buffer containing 100 mM Tris-HCl, pH 6.8, 200 mM DTT, 4% SDS, 20% glycerol and 0.2% bromophenol blue. After boiling for 5 minutes, the samples were subjected to electrophoresis in 10% SDS polyacrylamide gels at 8 volts/cm. The gels were dried and subjected to autoradiography to visualize the cross-linked proteins.

2.8 Screening cDNA expression library

2.8.1 Preparation of DNA probe

The following complementary penta oligonucleotides were used for cDNA screening:

plus strand, 5'-GATCCAGGGAAGGGAG-3' and minus strand, 3'-GTCCCTTCCCTCCTAG-5'

The two oligonucleotides were first phosphorylated by bacteriophage T4 polynucleotide kinase. Phosphorylated oligonucleotides were mixed and annealed by incubating the mixture at the following temperatures and times:

85°C for 2 min,

65°C for 15 min,

37°C for 15 min,

room temperature for 15 min,

0°C for 15 min.

To concatemerize the probe after annealing, 4 μl of T4 DNA ligase and 1 μl of 50 mM ATP were added to the reaction

mixture and the mixture was incubated at 16°C for 12 hours. The ligation was stopped by adding 0.5 M EDTA to a final concentration of 5 mM and 20% SDS to a final concentration of 0.5%. The reaction was extracted with phenol/chloroform (1/1, volume/volume) and precipitated with three volumes of ethanol at 0°C. Then 500 ng of annealed, ligated concatemer penta DNA fragments were used for labelling by nick translation. The size of the radiolabelled DNA probe was analyzed by nondenaturing polyacrylamide gel electrophoresis.

2.8.2 Screening cDNA library

The screening of the cDNA expression library was performed by the Southwestern-blot method of Vinson et al. (1988). Freshly prepared E. coli Y1090R⁻ bacteria were infected with recombinant phage of the P19 RA λgt 22A library at approximately 5×10^4 plaque forming units per 150 mm plate. The infection was performed by incubating at 37°C for 30 min. Then, 7.5 ml of 0.7% top agarose was mixed with bacteria and the mixture was overlaid onto 1.5% T-Tyn agar plates. After incubating at 42°C for 4 hours, Isopropyl-1-thio-ß-D-galactopyranoside (IPTG)-impregnated nitrocellulose filters were overlaid onto the plates and the plates were incubated for additional 6 hours at 37°C. The filters were then removed from the plates, air-dried and immersed in binding buffer containing 6 M guanidine-HCl and 25 mM HEPES, pH 7.9, 3 mM

$MgCl_2$, 4 mM KCl and 1 mM DTT for 5 minutes at 4°C with gentle shaking. The above buffer was diluted with equal volumes of binding buffer in series dilution and the filters were immersed in diluted buffer for 5 minutes. The dilution and filter incubation processes were repeated for four more times to finish the denaturation and renaturation of the proteins. The filters were washed twice in the binding buffer without guanidine and they were blocked by incubating them in binding buffer containing 5% nonfat dried milk for 30 minutes at 4°C. After rinsing them in binding buffer containing 0.25% nonfat dried milk, the filters were probed with 10^6 cpm/ml of nick translated concatemer DNA probe in binding buffer containing 0.25% nonfat dried milk and 10 μg/ml sonicated and denatured salmon sperm DNA. Hybridization with this probe was performed at 4°C for 3 hours with gentle agitation. The filters were then washed three times for 15 min each time with binding buffer containing 0.25% nonfat dried milk. Finally, the filters were wrapped in the plastic wrap and exposed to X-ray film at -70°C overnight with an intensifying screen. Positive plaques were identified and rescreened until they were homgeneously positive.

2.9 Characterization of RP1 cDNA clone

2.9.1 Preparation of RP1-β-gal fusion protein lysate

The fusion protein was prepared as described by Huang et

al. (1989). Briefly, freshly prepared Escherichia coli strain Y1090⁻ were infected with λgt22 RP1 recombinant clone at 5× 10^5 plaques per 150× 15 mm plate. After incubating for 30 minutes at 37°C, the infected bacteria were plated on soft agarose containing 10 mM $MgSO_4$ and 50 μg/ml ampicillin. The plates were further incubated at 42°C for 4 hours or until the plaques were clearly visible, before induction. The fusion protein was induced by adding 6 ml of induction solution containing 10 mM IPTG and 10 mM $MgSO_4$ and 0.5× LB broth to each plate. The plates were further incubated at 37°C for 3 hours and the supernatant was harvested. The induction was repeated twice with 1 hour incubation intervals between the inductions. The supernatants were combined and precipitated with 47.6 g/100 ml ammonium sulfate at 4°C. The fusion protein was recovered by centrifugation at 20,000 rpm for 30 min and the pellet was resuspended in small volumes of buffer containing 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$ and 1 mM dithiothreitol. After dialysing against the same buffer, the fusion protein was aliquoted, flash frozen in liquid nitrogen and stored at -70°C until further use. The protein concentration lysate was determined by the method of Bradford (1976).

2.9.2 Southwestern-blot assay

50 μg of fusion protein lysate was mixed with an equal

volume of 2× SDS sample buffer and boiled for five minutes. Then, the samples were applied to a 10% denaturing SDS-polyacrylamide gel. The gel was run with the same conditions used by UV cross-linking. After electrophoresis, the proteins were transferred to a nitrocellulose membrane by semi-dry electrophoretic transfer with Tris-glycine buffer (containing 48 mM Tris, 39 mM glycine, pH 7.4, 0.037% SDS and 20% methanol) at 0.8 mA/cm^2 for 30 minutes using a Bio-Trans unit (Gelman Inc.). Southwestern-blotting was then performed by the method of Vinson et al. (1988). The membrane was first denatured by immersing in DNA binding buffer (25 mM HEPES-NaOH, pH 8.0, 5 mM $MgCl_2$, 0.5 mM DTT, 25 mM NaCl) containing 6 M guanidine-HCl. After five minutes of incubation with gentle shaking, the membrane was transferred to binding buffer containing diluted guanidine HCl (3M), made by adding an equal volume of binding buffer, and then incubated at 4°C for five minutes to initiate renaturation. This serial dilution and incubation was repeated three more times and the membrane was then rinsed twice in binding buffer without guanidine HCl. Then, the membrane was blocked by incubation in 5% nonfat dry milk in binding buffer at 4°C for two hours. After rinsing once in binding buffer containing 0.25% nonfat dry milk, the membrane was probed with 10^6 cpm/ml of nick translated penta concatemer DNA probe prepared for screening the cDNA library. Incubation was for two hours at 4°C in binding buffer

containing 0.25% nonfat dry milk and 10 μg/ml of denatured salmon sperm DNA. This was followed by three washes with binding buffer containing 0.25% nonfat dry milk for 15 minutes each at 4°C. Finally, the membrane was air-dried and subjected to autoradiography by expose to X-ray film.

2.9.3 Total RNA preparation and Northern-blot assays

2.9.3.1 Total RNA preparation

Total RNA was prepared according to the method of Glisin et al. (1974). Briefly, media were removed and cells were washed with PBS. Cells from 8 plates (100 mm) were harvested by adding a total of 3.5 ml guanidinium solution onto the monolayers. The lysed cells were recovered by scraping with a rubber policeman. Lysates were combined and passed through a 20-G needle three times to reduce the viscosity. Each preparation of 3.5 ml cell lysates was overlaid onto 1.5 ml of 5.7 M CsCl in a 13×51 mm ultracentrifugation tube to create a step gradient, followed by centrifugation for 20 hours at 35,000 rpm at 18°C in an SW-55 rotor. RNA pellets were recovered, suspended in 3 M sodium acetate at pH 5.2, and precipitated with 2 volumes of 100% ethanol and chilled at -70°C for 30 minutes. After centrifugation for 15 minutes at 4°C, RNA pellets were dissolved in water and stored at -70°C for future use.

2.9.3.2 Electrophoresis of RNA in agarose gels

To separate the above RNA in an agarose gel, 10 to 30 μg RNA in 4.5 μl was mixed with 2 μl of 5× formaldehyde gel running buffer (0.1 M 3-(N-Morpholino)propanesulphonlic acid (MOPS), pH 7.0, 40 mM sodium acetate and 5 mM EDTA, pH 8.0), 3.5 μl formaldehyde and 10 μl formamide. After being incubated at 65°C for 15 minutes, the samples were chilled on ice and 2 μl of Diethyl pyrocarbonate (DEPC)-treated formaldehyde gel-loading buffer was added (50% glycerol, 1 mM EDTA, pH 8.0, 0.25% bromophenol blue and 0.25% xylene cyanol FF). The RNA samples were then loaded onto the 1% agarose gel containing 2.2 M formaldehyde and 1× formaldehyde gel running buffer and the gel was run by electrophoresis at 5 V/cm in 1× formaldehyde gel running buffer. The gel has been prerun for 5 minutes at the same condition. At the end of the run, the gel was stained with ethidium bromide solution containing 10 mM sodium phosphate, pH 7.0 and 0.5 μg/ml of ethidum bromide for 10 minutes at room temperature. The gel was then examined and photographed with a UV light transilluminator.

2.9.3.3 Transfer of RNA to nitrocellulose membrane

The stained gel was washed twice for 20 minutes each with DEPC-treated water. The gel was then placed onto a sheet of nitrocellulose membrane in a VT-20 vacuum system (Tyler Research Instruments) and vacuum was applied to the system at

about 15 to 20 inches of water. A buffer containing 50 mM NaOH, 0.1 M Tris-HCl, pH 7.5 was added to just cover the surface of the gel. After 5 minutes, the solution on the top of the gel was gently removed. This step was repeated once with another buffer containing 0.1 M Tris-HCl pH 7.5. For transfer, 10× SSC was added to the gel at a level of twice the height of the gel and RNA was transferred to membrane by vacuum for two hours. After transfer, the nitrocellulose membranes were air dried at room temperature and the RNA was fixed to the membrane by baking at 80°C for 30 minutes. The gel was occasionally restained and examined for even and efficient transfer.

2.9.3.4 Prehybridization and hybridization

The membrane bearing RNA was hybridized with probe, as described by Sambrook *et al.*, 1989. Briefly, membrane was prehybridized with hybridization buffer containing 6× SSC, 2× Denhardt's reagent and 0.1% SDS at 60°C for 2 hours. The buffer was removed and a fresh hydridization buffer was added to the membrane. DNA probe (2×10^{8} cpm/μg) was denatured by incubation at 100°C for 5 minutes and immediately added to the hybridization reaction and incubated for 24 hours at 60°C with gentle agitation. The membrane was then washed twice with 1× SSC, 0.1% SDS at room temperature for 15 minutes each, and

once with 0.25× SSC, 0.1% SDS at 55-60°C for 15-30 minutes. Washed membranes were exposed to X-ray film at -70°C for 24-48 hours (Maniatis, 1989).

*Preparation of solutions:

50× Denhardt's reagent: 5 g of Ficoll, 5 g of polyvinyl-pyrrolidone, 5 g of bovine serum albumin and H_2O to 500 ml.

20× SSC: 175.3 g of NaCl and 88.2 g of sodium citrate in H_2O to 1 litre, adjusted pH to 7.0 with 10 N of NaOH.

2.9.4 cDNA sequence analysis

RP1 cDNA was cloned into pBluscript plasmid and sequenced using Sequenase Version 2.0 (United States Biochemical Corporation) sequencing kit, according to the manufacturer's instructions.

The sequence was compared with known sequences using Genbank (EMBL).

CHAPTER 3

RESULTS

3.1 Effect of penta repeat motif on JCV promoter-enhancer activity

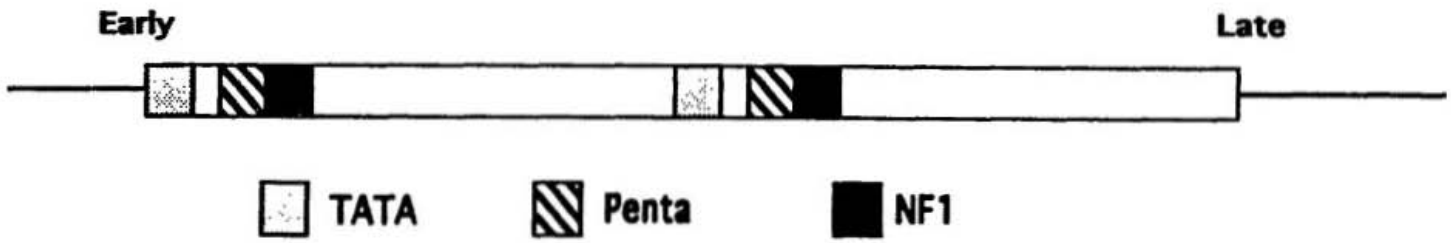
Previous results have shown that JCV early and late gene expression is under positive and negative regulation by a number of factors binding to different sequences. The AGGGA pentanucleotide repeat sequences (penta) and factor(s) binding to penta are the focus of my studies. In previous studies using a synthetic oligonucleotide containing the AGGGA repeat sequence cloned into the pA10CAT reporter plasmid, Tada et al., (1991) showed that the penta repeat motif functioned to downregulate CAT activity when inserted in the late orientation. Their further studies with mobility shift assays and UV crosslinking assays identified a protein of 56-60 kDa in glial cell extracts that specifically bound to the penta elements. The authors suggested that this protein may act as a silencer and interact with the penta repeat motif to repress JCV late (JCV_L) promoter activity (Tada et al., 1991). Since JCV T antigen stimulates JCV late promoter and does not bind to a specific motif, the activation role of T antigen in JCV late expression might involve its acting as an adaptor to displace the repressor. More recently, another group examined

the function of the same AGGGAAGGGA sequence in JCV early gene expression in C6 glial cells (Kumar et al., 1994). They found that this sequence was required for JCV early transcription. However, other than the 56-60 kDa protein, they identified a protein of 70-80 kDa binding to this sequence in the presence of the neighbouring cis DNA elements. All these results suggested that there are one or more proteins binding to AGGGAAGGGA elements and regulating JCV early and/or late gene expression in glial cells.

To examine whether that the penta motif cis-element was acting with a repressor to downregulate JCV gene expression in glial cells, I cloned nt 5112-270 of the JCV promoter-enhancer fragment into the pTZ vector. Site-directed mutagenesis was performed to introduce mutations into the two penta repeat motif in the promoter-enhancer region of JCV. The fragments containing mutations were cloned into the promoterless-enhancerless CAT reporter plasmid, pSV0CAT, in both early and late orientations (Fig. 2). The plasmids with the mutant JCV promoter-enhancer were then transfected into RA-differentiated P19 glial cells and U87MG (U87) glioblastoma cells to test the role of penta repeat motif using CAT expression assays. The mutations are diagrammed in Fig. 2. Mutations of both NF1 sites were similarly derived, as described in Fig. 2.

Fig. 2. Diagram of wild type and mutated JCV sequences in expression plasmids.

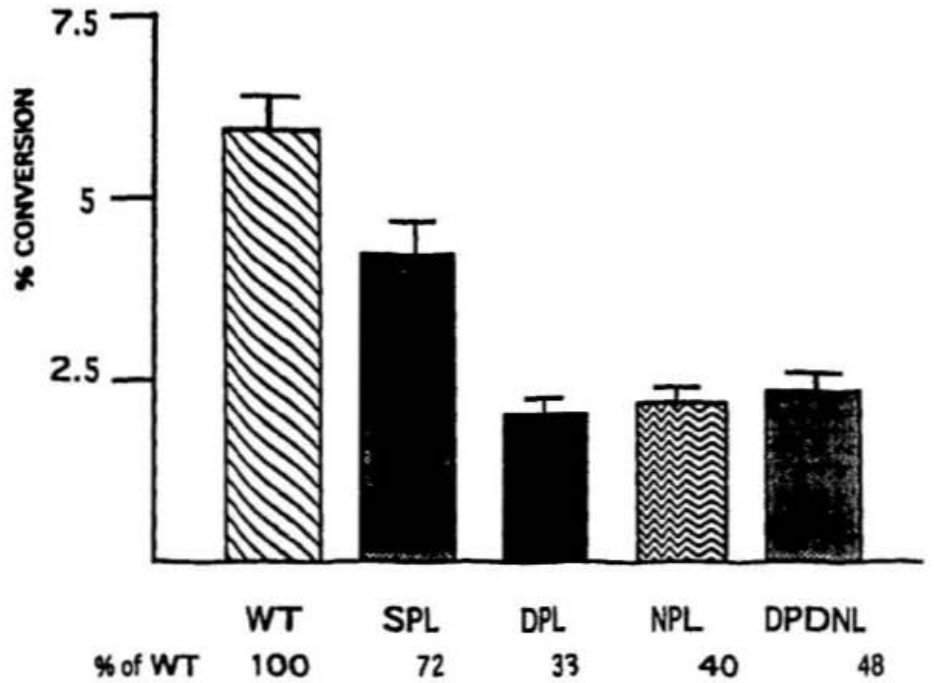
The 98 bp repeats of JCV are shown as two wide open boxes containing the small boxes for indicated motifs. Early and late directions of gene expression are indicated. Labels are: WT, wild type sequences of penta repeat motif and NF1 motif; S mutant, mutated penta motif close to NF1 II (Kumar et al., 1993) site; D mutant, mutated penta motifs in both repeats; D plus NF1, mutated penta and NF1 motifs in both repeats. Capital letters, dashes and blanks represent WT and lower case letters represent mutant nucleotides. WT penta motifs are bolded and WT NF1 motifs are underlined.

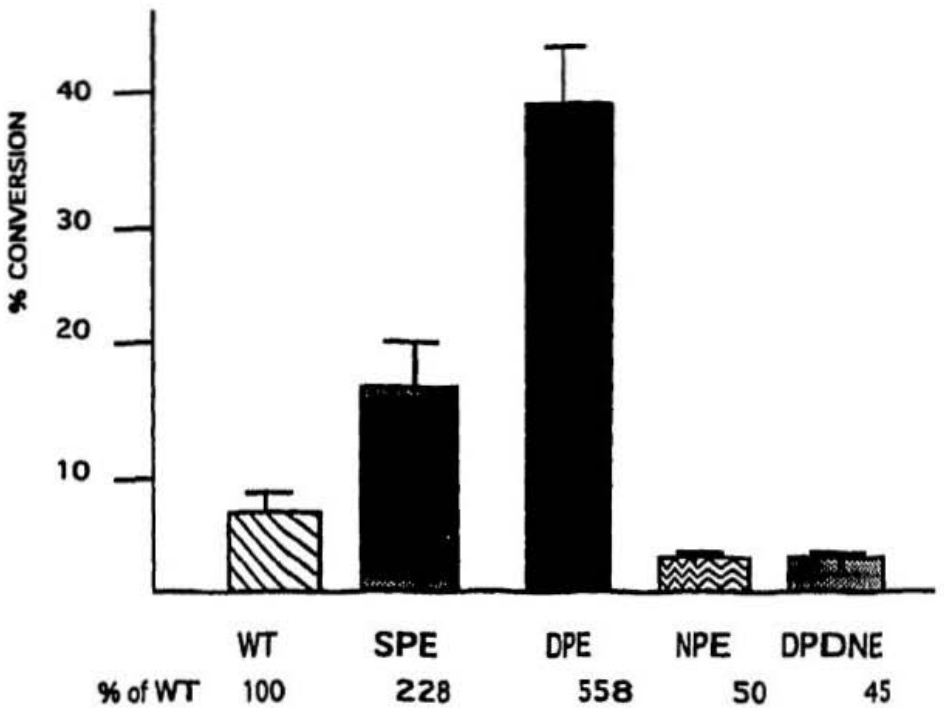


	penta NF1 III	penta NF1 II
WT	AGGGAAGGGATGGCTGCCAGCCA	AGGGAAGGGATGGCTGCCAGCCA
S mutant		-ca-tcta-
D mutant	-ca-tcta-	-ca-tcta-
D plus NF1	-ca-tcta-gta------ a-c	-ca-tcta-gta------ a-c

Fig. 3. Effect of mutations on in vivo activity of JCV penta repeat and NF1 motifs in glial cells.

Top panel: effect of penta repeat motif mutations on JCV late promoter activity. Lower panel: effect of penta repeat motif mutations on JCV early promoter activity. Labels: WT, wild type; L, late orientation; E, early orientation; S, single mutated penta motif close to NF1 II; D, double mutated penta motifs; N, NF1 sites mutation; DN, both penta motifs and NF1 sites mutation. The percent conversion of ^{14}C chloramphenicol to products for CAT activity is given.The results are the average of three experiments plus or minus standard deviation (SD).


7.5
5
2.5
% CONVERSION
WT
SPL
DPL
NPL
DPDNL
% of WT
100
72
33
40
48


40
30
20
10
% CONVERSION
WT
SPE
DPE
NPE
DPDNE
% of WT
100
228
558
50
45

Surprisingly, my results revealed that the penta repeat motifs were required for JCV_L expression (Fig. 3). When one penta site adjacent to the NF1 II site, named penta II, was mutated, late activity was reduced significantly in both RA P19 cells (data not shown) and U87 cells. When both penta II plus penta III adjacent to NF1 III were mutated, JCV late activity was further reduced in both cell lines (Fig. 3). In contrast, JCV early gene expression was increased in RA P19 cells (data not shown) and U87 cells when penta II was mutated, and the double mutation of penta II and penta III further increased the JCV early activity (Fig. 3). These results suggested that both penta II and penta III sites are important in regulating JCV early and late activities in glial cells, negatively regulating JCV early expression and positively regulating JCV late expression. The effect of the mutations was more significant in U87 cells than in P19 cells. This could be explained by the fact that U87 is composed of only human glioblastoma cells and P19 RA cells are a mixture of mouse neuronal and glial cells. To examine the possibility of neighbouring penta-NF1-binding proteins interactions, I mutated all four sites (Fig. 2). The results showed that JCV late and early promoter activities were approximately the same compared with the results with NF1 II and III mutation alone (Fig.3). The results for NF1 were consistent with previous results (Kumar et al., 1993). The absence of any significant

effect of the NF1 mutation on the mutant penta repeat motif and vice-versa, suggested that both functions required the other and the mutation of NF1 will abolish most of the JCV promoter-enhancer activity.

3.2 DNA-protein interaction as shown by mobility shift assays

Since previous reports and my mutagenesis analysis suggested that there are one or more proteins interacting with the penta repeat motif sequence and regulating JCV early and late gene expression, I analyzed whether such protein(s) were present in the RA P19 and U87 cells. Therefore, mobility shift assays were employed to evaluate DNA-protein binding. Using the synthetic oligonucleotide spanning the penta repeat motif as a probe, two DNA-protein complexes were observed with the nuclear extract of RA P19 cells (Fig. 4). To examine the specificity of this interaction, competition experiments were performed. Results showed that, while the unlabelled wild type oligonucleotide used as probe was able to compete for this binding efficiently (Fig.4, lane B+w), the mutant penta did not affect this DNA-protein interaction (Fig.4, lane B+m). This result and my CAT assay results suggested that there are one or more proteins in P19 RA cells interacting with the penta motif and regulating the JCV gene expression. Mobility shift assays were also performed using U87 cell nuclear extracts. However, only the lower mobility complex close to

the top was repeatedly observed with RA P19 and U87 extracts (Fig. 5). Although two complexes were detected in RA P19 cells and one complex was detected in U87 cells, whether the protein(s) interacting with the penta motif is a multimer or monomer is not clear at this time.

Because this protein(s) was involved in negative regulation of JCV early gene activity as suggested by my results, it may play a role in the cell specific expression of JCV. Thus, to test the binding of this protein(s) in glial and nonglial cells, mobility shift assays were performed using nuclear extracts from DMSO differentiated P19 cells composed of mainly muscle cells and from undifferentiated P19 cells. Again and at the same positions, two DNA-protein complexes were detected with the DMSO differentiated and undifferentiated P19 cell extracts, suggesting that this penta motif binding protein(s) is not cell specific (Fig. 5 and Fig. 6) respectively. The U87 results were also similar to those with P19 extracts, but the binding appeared to be less, especially for the lower, faster mobility complex (Fig. 5).

3.3 Penta repeat motif binding protein does not interact with T antigen

Previous studies have shown that the NF1 site and TATA box are separated by a penta repeat motif, a repressor is inhibiting late expression through interaction with this penta

motif, and T antigen can activate JCV late expression. These findings lead to a hypothesis that T antigen may be acting as an adaptor, functioning to remove the repressor protein, clearing the steric hindrance to NF1, and facilitating the interaction between NF1 and general transcription factors bound to the TATA box. To test this hypothesis, P19 RA cells were transfected with a JCV T antigen expression plasmid. Whole cell extracts from these cells was used for mobility shift assays with the penta repeat motif probe. The results showed that the expression of T antigen did not affect the interaction of penta repeat motif and penta-binding protein (Fig. 7). Incubation of the binding reaction with a specific T antigen antibody did not show a supershift that would have indicated a T antigen-penta protein interaction. Thus, the data suggested that T antigen was not activating JCV late expression by removing this penta motif binding protein through the direct interaction with T antigen. This result also suggests that T antigen negatively regulates JCV early promoter-enhancer activity by different mechanisms.

Fig. 4. *In vitro* binding of P19 RA glial cell factors with penta repeat motif of JCV.

Mobility shift assays are shown. Double-stranded oligonucleotides containing penta repeat was end-labelled and incubated with P19 RA differentiated cell nuclear extract. Labels are: F, free probe without nuclear extract; B, binding with P19 retinoic acid-differentiated cell nuclear extracts; w, homologous competitor; m, heterologous competitor. In each reaction, 5 μg protein was used. Arrows indicate the two specific DNA-protein complexes. Arrowhead indicates free probe.

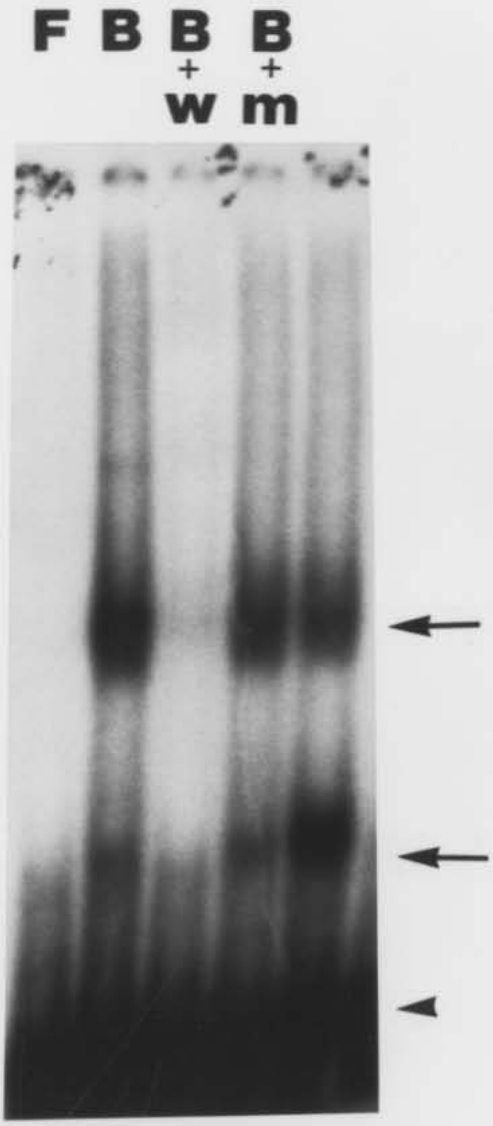

F
B
B + w
B + m

Fig. 5. <u>In vitro</u> binding of U87 glial cell factors and DMSO differentiated P19 cell factors with penta repeat motif of JCV.

Mobility shift assays are shown and 5 μg protein in extracts was used. Left panel was with U87 cell nuclear extract. Right panel was with P19 DMSO differentiated cell nuclear extract. Labels are: F, free probe without nuclear extract; B, binding with nuclear extracts of indicated cells; w, homologous competitor; m, heterologous competitor. Arrows indicate the specific DNA-protein complexes. Arrowhead indicates free probe.

U87
F
B
-
+
+
w
m
P19DM
F
B
-
+
+
w
m

Fig. 6. *In vitro* binding of RA P19 glial cell proteins and UD P19 cell proteins with penta repeat motif of JCV.

Mobility shift assays of the binding of penta repeat motif with nuclear extracts of undifferentiated P19 cells. The two lanes labelled (m) for RA indicate duplicate assays in RA P19 cells with the same mutant competitor. Labels are: F, free probe without nuclear extract; B, binding with nuclear extracts of indicated cells; w, homologous competitor; m, heterologous competitor. Arrows indicate the specific DNA-protein complexes. Arrowhead indicates free probe.

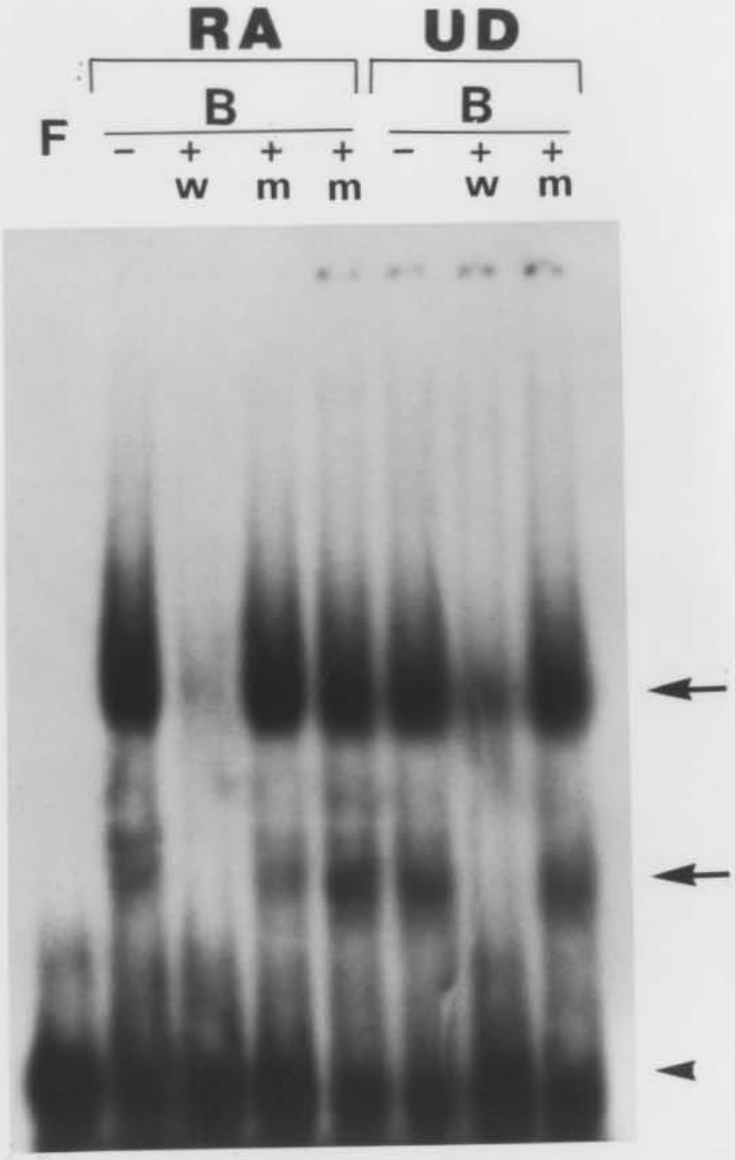
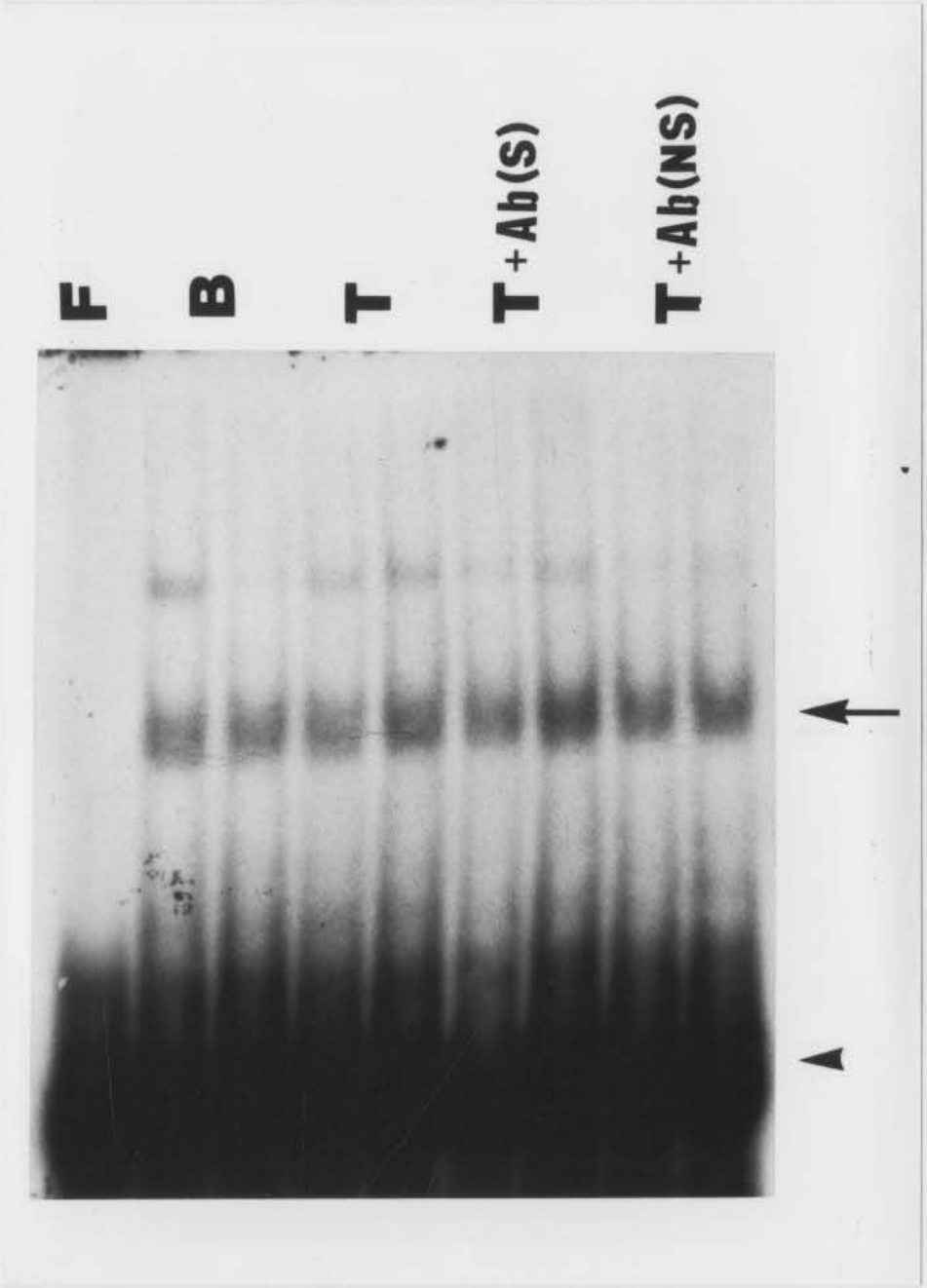

RA
UD
B
B
F
–
+
w
+
m
+
m
–
+
w
+
m

Fig. 7. In vitro binding of P19 RA cell factors with penta repeat motif of JCV in T antigen-expression plasmid-transfected cells

Mobility shift assays are shown using 5 μg whole extracts. Probe is the same as in Fig. 4. Labels are: F, free probe without extract; B, binding of the probe with P19 RA cell whole cell extract; T, whole cell extract of RA P19 cells transfected with 20 ug/plate of the JCV T antigen expression plasmid, pCMVJCT; Ab(S),specific T antigen antibody; Ab(NS), nonspecific antibody.


F
B
T
T+Ab(S)
T+Ab(NS)

3.4 Identification of the penta repeat motif binding factor by UV crosslinking

UV cross-linking assays were used to identify and determine the size of DNA binding proteins that interact with the penta repeat motif of JCV. The same probe used for mobility shift assays was labelled by nick translation and UV crosslinked with proteins in the P19 RA cell extract. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS/PAGE) and autoradiographic results showed that two proteins with apparent molecular weights of 65 kDa and 43 kDa interacted with the penta motif (Fig. 8). The DNA-protein interaction was specific, as unlabelled probe efficiently competed the binding of both complexes, while the unlabelled oligonucleotide containing the mutant penta motif did not affect the binding. Combined with the above mobility shift assay results and the previous data, the results provided here indicated that the cis-element, AGGGAAGGGA, in JCV regulates JCV early and late gene expression by interacting with cellular factors. It is possible that these two proteins bind separately to the penta motif to form the two complexes observed in the mobility shift assays, or they may form dimers, either homodimers or heterodimers, when binding to penta cis-elements.

3.5 Isolation and characterization the RP1 cDNA of the penta repeat motif binding factor

3.5.1 Screening RA P19 cell expression library for cDNA encoding the penta repeat motif binding factor

In vivo assays based on site-directed mutagenesis and in vitro binding in UV cross-linking experiment indicated that there are two cellular proteins regulating JCV early and late gene expression through the penta motif. These proteins are present in both glial and nonglial cells and are able to interact specifically and directly with the penta repeat motif. The isolation of a cDNA clone(s) encoding these factors would help us to understand the regulation of JCV expression.

A probe containing a multimer of the penta motif was used to screen approximately 2×10^{6} plaques from a P19 RA λgt22 expression library, following the Southwestern-blot method of Singh et al. (1988) as and Vinson et al. (1988) described in section 2.8. After primary screening and following purification of secondary and tertiary rounds of screening, two plaques were found to be positive. To test the specificity of these two positive clones, replica filters from plated recombinant phage were hybridized with nick translated oligonucleotide probes of the sequences of the wild type penta, mutant penta, wild type NF1 II/III motifs and glucocorticoid response element (GRE). Results showed that one clone encoded a protein that bound to all four sequences and was not further analyzed. This clone could represent one

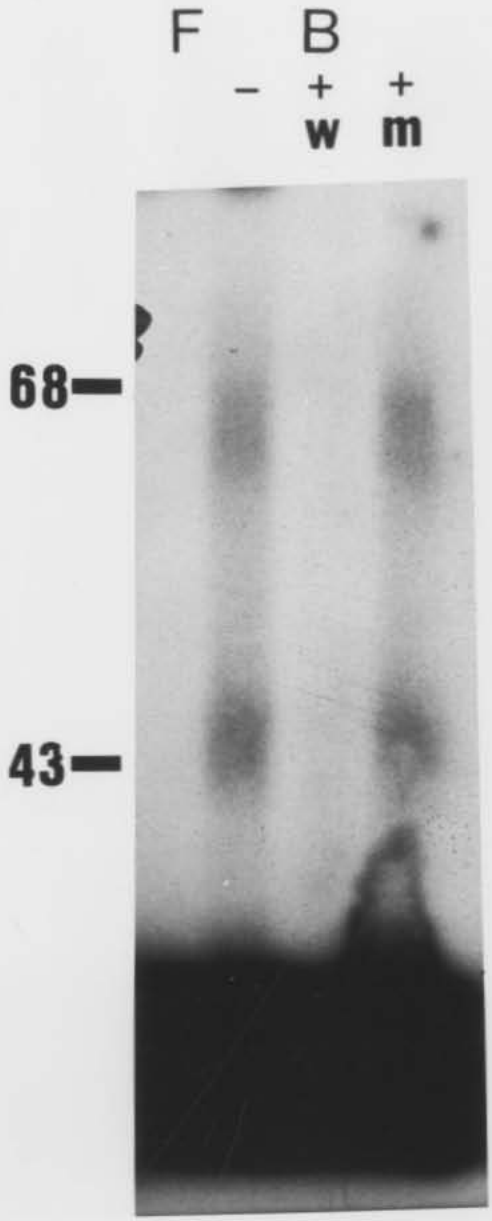
of the proteins binding to the penta motif. The second recombinant phage, RP1, however, encoded a specific DNA binding protein that only bound to wild type penta, but not to the mutant penta, NF1 or GRE oligonucleotides (Fig. 9).

3.5.2 Cell specificity and Northern-blot assays

The RP1 cDNA was isolated from a library of P19 RA cells containing two or more specifically and directly penta repeat binding proteins that are present in nonglial cells also. Therefore, I determined the size of this cDNA and evaluated if RP1 cDNA encodes a factor specifically expressed in glial cells. For this purpose, Northern-blot analysis was performed. Total RNA was isolated from undifferentiated (UD), DMSO differentiated (DM) and RA differentiated (RA) P19 cells (Fig. 10). The cDNA from the RP1 recombinant phage was excised by SalI and NotI. The DNA fragment was then nick translated to be used as a probe in Northern-blot assays. An mRNA species of about 2.2 kb was detected in all the cell types tested (Fig. 10). This is consistent with my <u>in vitro</u> binding assays which indicated that the penta motif binding protein(s) was expressed in RA P19 glial cells and U87 glioblastoma cells, and UD and DM P19 nonglial cells (Fig. 5; Fig. 6).

Fig. 8. UV crosslinking to identify and determine the size of the proteins binding to penta repeat motif of JCV.

The probe used for UV crosslinking was labelled by nick translation. Labels are: F, free probe without the P19 RA differentiated cell nuclear extract; B, binding of the penta repeat probe with the P19 RA cell nuclear extract; w, homogenous competitor; m, heterogenous competitor. Numbers on the left side indicate the protein size marker (Gibico-BRL, high range) analyzed in a lane in the same gel that is not shown.


F
B
–
+
w
+
m
68
43

3.5.3 DNA binding analysis

To reexamine whether the RP1 recombinant clone expressed a protein that specifically binds to the penta repeat motif, a β-galactosidase-RP1 fusion protein was prepared as described in section 2.9. The proteins from uninfected bacteria and infected bacteria that were not induced to produce the fusion protein were used as negative controls. The protein samples were resolved in SDS-PAGE, transferred to nitrocellulose membranes and analyzed by Southwestern-blot assays as described by Vinson et al. (1988). Using the nick translated penta repeat motif as a probe, a protein of approximately 145 kDa was detected from protein samples of the RP1 phage clone induced by IPTG (Fig.11). This suggested that the RP1 cDNA encodes a protein of about 30 kDa. The same protein was not detected in the samples from bacteria containing RP1 phage clone but not induced by IPTG, from bacteria containing a nonspecific phage clone or from uninfected bacteria (Fig. 11). This result confirmed that RP1 encodes a DNA-binding protein specifically recognizing the JCV penta repeat motif.

Fig. 9. Specific binding of RP1 recombinant clone to penta motif

Southwestern-blot of the purified phages containing the RP1 clone is shown. The positive recombinant phage clone was plated and transferred, as described in section 2.8.2. The nitrocellulose membrane bearing the fusion protein was treated with 6 M guanidine HCl solution. After renaturing by gradually diluting the guanidine HCl, the membrane quarter pieces were probed with end-labelled oligonucleotides containing penta motif (A), mutant penta motif (B), NF1 motif (C) and GRE motif (D).

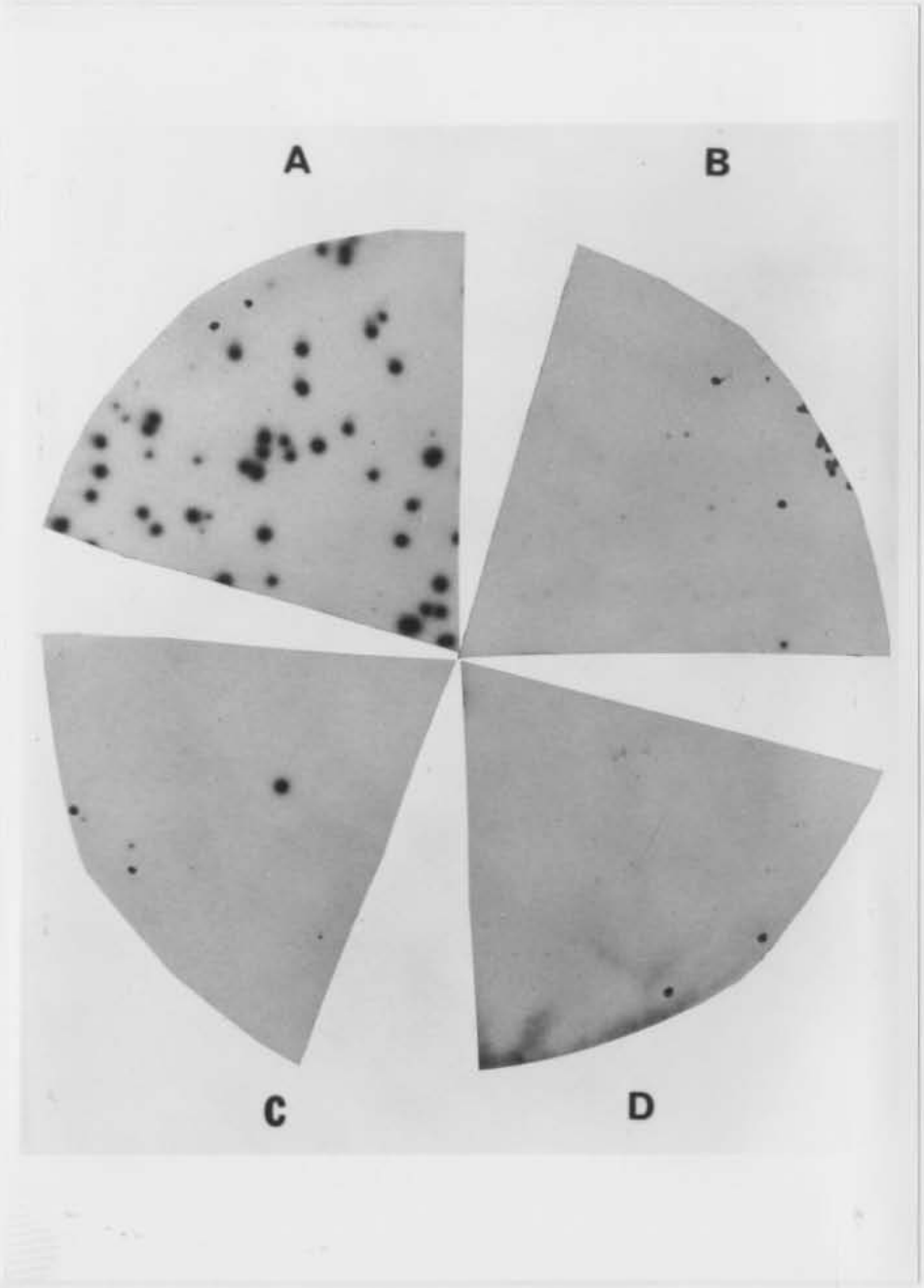

A
B
C
D

Fig. 10. Northern blot analysis of RP1 expression in glial and nonglial cells.

Total RNA from undifferentiated P19 cells (UD), RA differentiated P19 glial cells (RA), and DMSO differentiated P19 cells (DM) was prepared as described in section 2.9. Excised RP1 cDNA was nick translated and used as probe. The numbers on the left are RNA size markers in kb for the RNA ladder markers lane that is not shown.

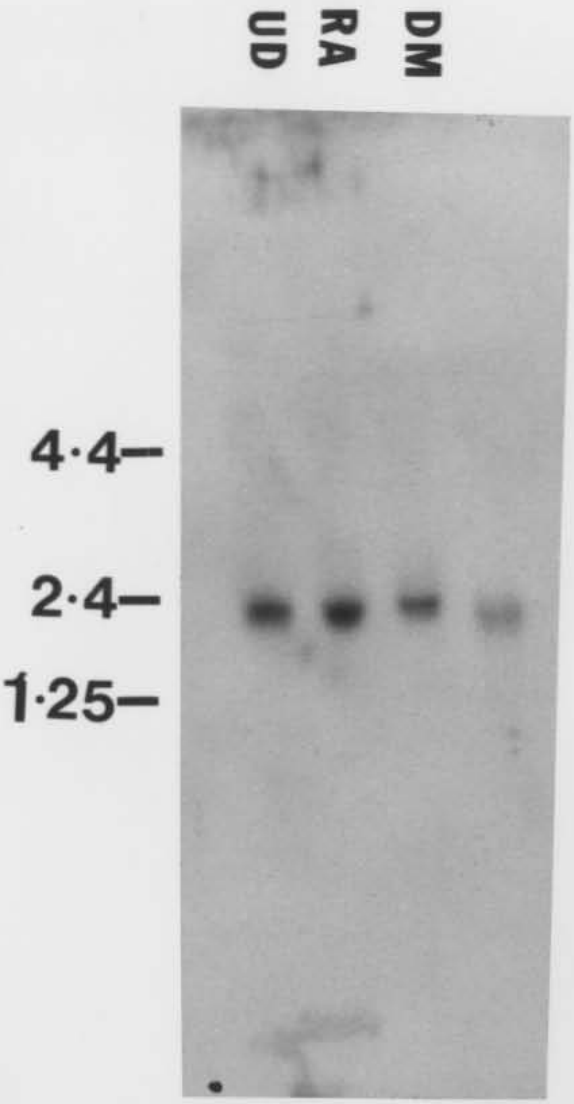
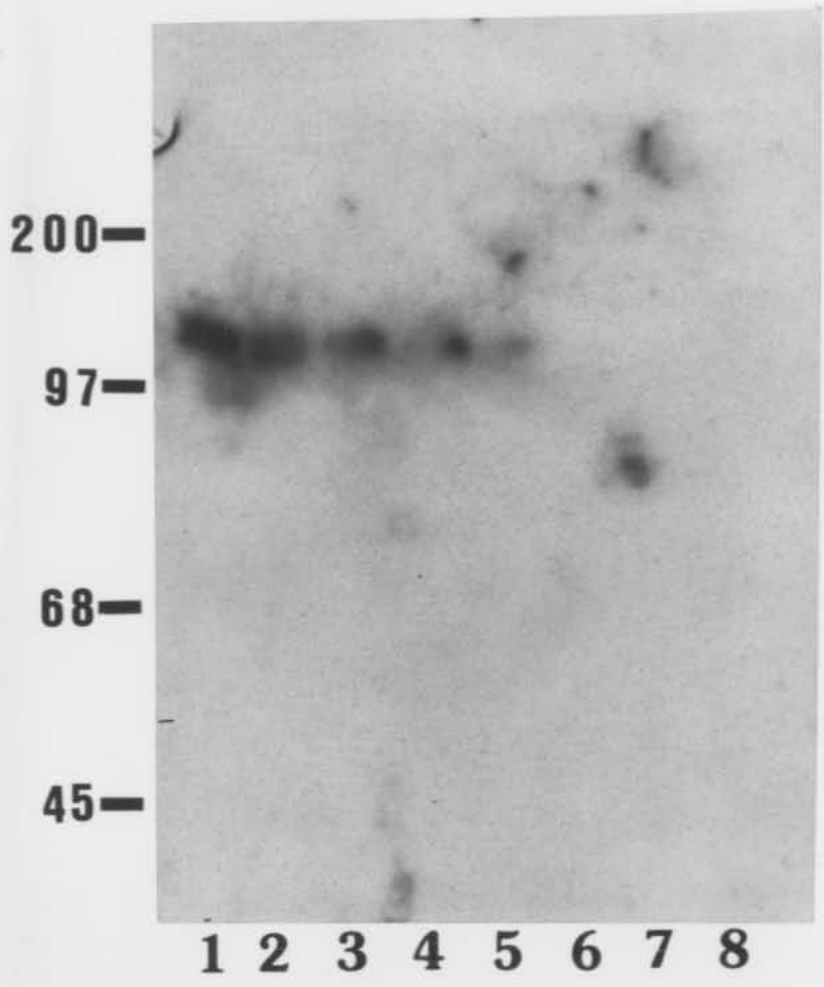

UD
RA
DM
4·4
2·4
1·25

Fig. 11. DNA binding activity of RP1.

Southwestern-blot analysis of an RP1 fusion protein is shown. An RP1-β-gal fusion protein was prepared by infecting E.coli Y1090 bacteria with RP1 recombinant phage clone and inducing with IPTG. The fusion protein was resolved by 10% SDS-PAGE. After transferring to a nitrocellulose membrane, the protein was denatured and renatured and probed by nick-translation labelled penta repeat motif oligonucleotide. Numbers on the left side are the protein size markers in kDa for a lane that is not shown. Lanes: 1-5, IPTG-induced bacteria containing RP1 clone using 100, 80, 60, 40 and 20 μg protein, respectively; 6, RP1 clone without induction; 7, uninfected induced with IPTG; 8, a nonspecific clone isolated from the same library with and induced with IPTG.


200
97
68
45
1 2 3 4 5 6 7 8

3.5.4 DNA sequence analysis of RP1 cDNA to characterize RP1 and the RP1 protein

After cloning into pBluescript, the 5' and 3' ends of the RP1 cDNA were partially sequenced. These sequences were searched in a databank (EMBL/DDBJ/GenBank) for homologous sequences. The search showed that this cDNA is 100% identical to a previously isolated cDNA named HSNABP, accession number L12693. HSNABP had been randomly isolated from a mouse liver cDNA library using a zinc finger homologous probe. Thus, its function was unknown. To further evaluate whether my cDNA is the same cDNA as that in the databank, several restriction enzyme recognizing sites in HSNABP were used to compare with sites in RP1. The same set of restriction sites were found in RP1 and therefore were consistent with the two cDNAs being similar (data not shown).

RP1 cDNA has 54 more nucleotides and 25 more nucleotides than HSNABP in the 5' end and 3' end respectively, the later including a poly A sequence. It also has an AATAAA polyadenylation signal at nt 1512-1517, a sequence usually appearing 10 to 20 nucleotides before the polyadenylate tail of most eukaryotic mRNAs (Proudfoot et al., 1974). The cDNA nucleotide sequence is shown on Fig. 12. From which, an ORF coding for a putative protein of 222 amino acid protein was deduced (Fig. 12). The RP1 does not have an ATG start codon at the 5' end. Because it is a fusion protein, the coding

region must be in frame with β-gal. From the 5' end of RP1, the sequence encoded a protein of 29 kDa in size that terminated at a TAA stop codon at nucleotide position 669. The 2.2 kb mRNA identified in Northern blot assay (Fig. 10) suggesting that about 500bp sequence at 5'end of my cDNA was missing. In combination with the Northern blot result, my cDNA sequence data suggests that RP1 cDNA most probably encodes the 43 kDa protein identified in my UV crosslinking experiment. From the amino acid sequence, I identified seven putative zinc finger motifs distributed throughout the sequences (Fig.12, underlined amino acids), suggesting that RP1 belongs to the zinc finger super family of transcription factors. Each RP1 zinc finger contains three cysteine residues and one histidine residue arranged in the following order: cysXXcysglyXXglyhisXXXXcys. All these putative zinc finger motifs have glycines immediately after the second cysteine and before the histidine residue; six of them have tyrosine in the second position. In addition to these putative zinc finger motifs, there are many paired basic and hydrophobic amino acids. However, the importance of these amino acids in RP1 and the potential sequence motif formed by these amino acids is awaiting further computer analysis.

3.5.5 *In vivo* assays of repression by RP1 protein of JCV early gene expression

After the RP1 cDNA was isolated, it was important to examine whether or not this cDNA encodes a protein that is functional *in vivo*. Therefore, RP1 cDNA was inserted into a pCMV expression vector in the sense direction. The inserted cDNA was designed to produce RP1 protein from the ATG start codon in the vector after transfection. This pCMV-RP1 expression plasmid was cotransfected into P19 RA cells and U87 cells with pJCLCAT or pJCECAT plasmids. Compared with pJCLCAT alone, pCMV-RP1 had no significant influence on JCV late expression in P19 RA cells and U87 cells (Fig. 13). This would have been due to the truncation of the cloned RP1. In P19 RA cells, CAT activity of JCV early CAT alone was slightly lower than normal, but not significantly. However, compared with pJCECAT alone, cotransfection of pCMV-RP1 and JCV early reporter plasmids in U87 cells showed significantly lower CAT activity (Fig. 13). This is not surprising, as site mutagenesis resulted in a more significant change in U87 cells than in P19 RA cells (Fig. 3). Also, the results are consistent with site directed mutagenesis and binding data, which indicated that the penta motif most probably is functioning to repress JCV early expression through the interaction with cellular factors (Fig. 4). Taken together, my data suggested that I isolated a cDNA encoding a cellular protein possibly involved in the regulation of JCV early gene expression and JCV late gene expression.

Fig. 12. RP1 cDNA nucleotide sequence and the predicted amino acid sequence of RP1 protein.

The total cDNA contains 1647 nt. The 222 amino acid of the predicted protein begin from the first codon. The seven putative zinc fingers are underlined. The polyadenylate signal AATAAA is in bolded italic type.

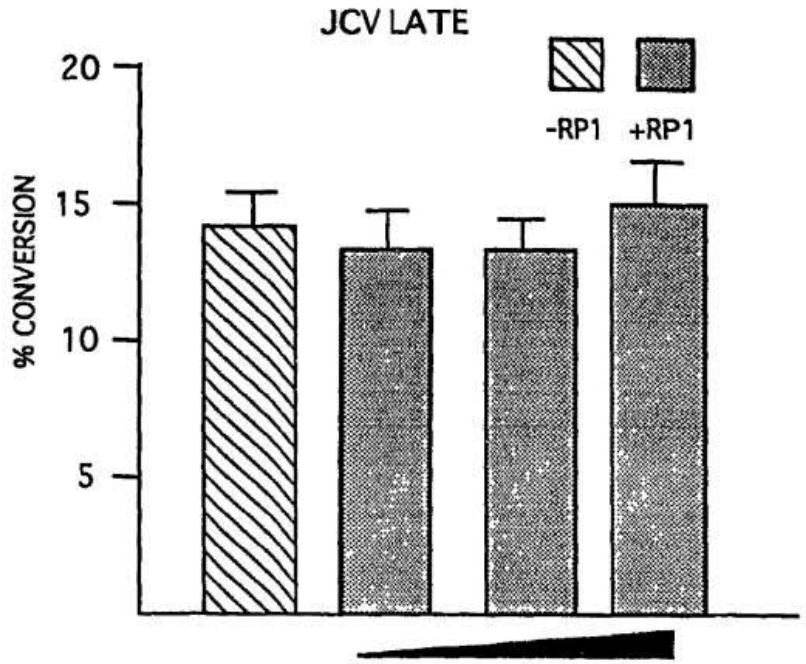
```
5'-   CGCGTCCGCTTGGATGGGGCCGTGTGCAGACCCGCGTGTGGCGCA
N -   ArgValArgLeuAspGlyAlaValCysArgProAlaCysGlyAla

 46   GGCAAGGACCCTGAAAATAAACAGCCGCTGCTTTGCGAGTCGCCT
 16   GlyLysAspProGluAsnLysGluArgLeuLeuCysGlnSerPro

 91   TCTTGGTTCTTCGTCCGAGTCTCCTCCGCTGTGGGCAGCTCAGAC
 31   SerTrpPhePheArgArgValSerSerValValGlySerSerAsp

136   GCCGAAGCTCTAACTGCAGCTATGAGCAGCAACGAATGCTTCAAG
 41   AlaGluAlaLeuThrAlaAlaMetSerSerAsnGluCysPheLys

181   TGTGGACGATCTGGCCACTGGGCCAGGGAGTGCCCTACTGGTGGA
 56   CysGlyArgSerGlyHisTrpAlaArgGluCysProThrGlyGly

226   GGTCGGGGTCGTGGAATGAGAAGCCGCGGCAGAGGGTTCCAGTTT
 61   GlyArgGlyArgGlyMetArgSerArgGlyArgGlyPheGlnPhe

271   GTTTCCTCGTCTCTCCCTGACATCTGCTACCGCTGTGGTGAGTCT
 76   ValSerSerSerLeuProAspIleCysTyrArgCysGlyGluSer

316   GGTCATCTTGCCAAGGATTGTGATCTGCAGGAGGATGCCTGCTAT
 91   GlyHisLeuAlaLysAspCysAspLeuGlnGluAspAlaCysTyr

361   AACTGCGGTAGAGGTGGCCACATTGCCAAGGACTGCAAGGAGCCC
121   AsnCysGlyArgGlyGlyHisIleAlaLysAspCysLysGluPro

406   AAGAGAGAGCGAGAGCAATGCTGCTACAATTGTGGCAAGCCAGGC
136   LysLysGluArgGluGlnCysCysTyrAsnCysGlyLysProGly

451   CATCTCGCTCGTGACTGTGACCACGCGGATGAGCAGAAGTGCTAT
151   HisLeuAlaArgAspCysAspHisAlaAspGluGlnLysCysTyr

496   TCTTGTGGTGAATTTGGACATATTCAAAAAGACTGCACCAAGGTG
166   SerCysGlyGlyPheGlyHisIleGlnLysAspCysThrLysVal

541   AAGTGCTATAGGTGTGGTGAAACTGGTCATGTAGCCATCAATTGC
181   LysCysTyrArgCysGlyGluThrGlyHisValAlaIleAsnCys

586   AGCAAGACAAGTGAAGTCAACTGTTACCGCTGTGGCGAGTCAGGG
196   SerLysThrSerGluValAsnCysTyrArgCysGlyGluSerGly

631   CATCTTGCACGGGAATGCACAATTGAGGCTACAGCCTAATTATTT
211   HisLeuAlaArgGluCysThrIleGluAlaThrAla

676   TCCTTTGTCG CCCCTCCTTT TTCTGATTGA TGGTTGTATT
716   ATTTTCTCTG AATCCTCTTC ACTGGCCAAA GGTTGGCAGA
756   TAGAGGCTGT TCCCAGGCCA GTGAGCTTTA CTTGCAGTGT
```

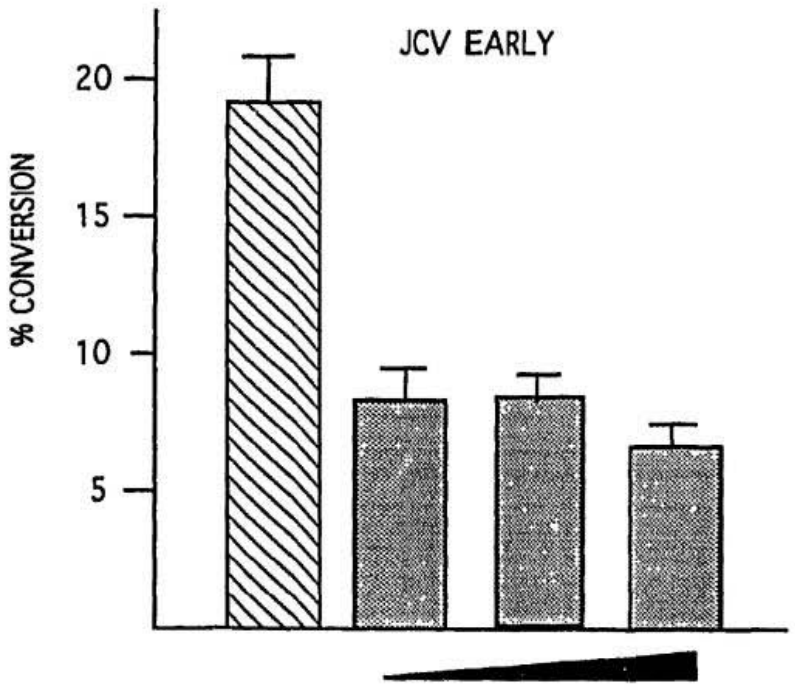
```
 796   AAAAGGAGGA AAGGGGTGGA AAAAACCGAA TTTCTGCATT
 836   TAACTACAAA AAAAGTTTAT GTTTAGTTTG GTAGAGGTGT
 876   TATGTATAAT GCTTTGTTAA AGAACCCCCT TTCCGTGCCA
 916   CTGGTGAATA GGGATTAATG AATGGGAAGA GTTCAGTCAG
 956   ACCAGTAAGC CCTTCTGGGT TTGAGTGTGT TCCCATGTAG
 996   GAGGTAAAAC CAATTCTGGA AGCATCTAAG CTTCCATAAA
1036   TAACTTTAAT TCTTAGCATA ATGACGGCCT TGGATTGTCT
1076   GACCTCAGTA GCTATTAAAT AACATCGAGT AACATCTGCA
1116   TCAGGCCCTC AGAATATACA GTTGAGTTGG GAGTAAACTG
1156   AAAAGACAAA TGTGTTGAAG GCTATGCCAG GGAATCTGGC
1196   TCAAAGCCTA ACACAGAAGC AGCTTCATCC CAGTGACGAT
1236   GCTGGACGTA CAGATGGTGA TGGCAAAGGT GTAGAACACA
1276   TTTTTTCAAA GACTAAATCT AAAACCCAGA GTAAACATCC
1316   GATGCTCAGA GTTAGCATAA TTTGGAGCTA TTCAGGAATT
1356   GCAGAGAAAT GCATTTTCAC AGAAATCAAG ATGTTATTTT
1396   TGTATACTAT ATCACTTAGA CAACTGTGTT TGCTTTGCTG
1436   TAATCAGTTT TTAAAAGTCA GATGGAAAAA GCAACTGAAG
1476   TCCTAGAAAA TAGAAAATGT AATTTTAAAC TATTCCAATA
1516   AAGCTGGAGG AGGAAGGGGA GTTTTGACTA AAGTTCCTTT
1556   TGTTTGTTTT AAATTTTCAT CAATGTATAT AGAACAAAAT
1596   ACCATATTAA AGAGGGGAAT GTGGAGGACT GAAAAAAAAA
1636   AAAAAAAAAA AA-3'
```

Fig. 13. Effect of RP1 on JCV gene expression *in vivo*

RP1-pCMV expression plasmid was cotransfected with either JCV early or JCV late CAT reporter plasmid into U87 cells. CAT activity was analyzed and plotted. Assay results with cotransfected JCV late and JCV early CAT expression plasmids are shown with increasing levels of RP1 plasmid. 10 μg of early or late reporter plasmid was used. The inclining solid symbols at the bottom indicate 2.5 μg, 5 μg, and 10 μg of RP1 used. Hatched lanes and -RP1 indicate no cotransfection with RP1 expression plasmid, shaded lanes and +RP1 indicate cotransfection with RP1 expression plasmid. The results are the average of three experiments plus or minus standard deviation (SD).

In Vivo Effect of RP-1





CHAPTER 4

DISCUSSION AND FUTURE DIRECTIONS

4.1. Effect of penta repeat motifs on JCV promoter-enhancer activity

A number of reports showed that JCV early and late gene expression is subject to negative as well as positive regulation. The negative regulation of the JCV early promoter-enhancer was first identified by Beggs et al. (1988). Using somatic cell hybridization techniques, they found that T antigen expression declined after JCV-transformed hamster glial cells were fused with mouse fibroblasts. This decrease reached the maximal level 24-36 hours after fusion. The decrease in the level of T antigen depended on the ratio of glial to fibroblast nuclei in the multinucleated heterokaryons. Similar results were observed when such JCV-transformed cells were fused with mouse Ltk cells, suggesting the negative regulation of JCV by mouse cellular factors in nonglial cells. Moreover, reexpression of T antigen occurred in some of the subclones of the Ltk hybrids, concomitant with the loss of mouse chromosomes and, presumably, the loss of the mouse cell factors involved in negative regulation. Bollag et al. (1989) identified the cis-elements interacting with the predicted negative regulators of the JCV regulatory region.

It was concluded that this negative regulation was at the transcription level. Also, Kumar et al. (1991) identified a protein of 53 kDa in both glial and nonglial cells interacting with the penta repeat motif in the JCV control region and suggested that this protein may play a role in negative regulation of JCV early gene expression in nonglial cells. Further, JCV T antigen has been shown both to suppress JCV early expression in cotransfection assays (Nakshatri et al., 1990) and to stimulate JCV late gene expression, suggesting a role in the late stages of infection when early gene expression is repressed and late gene expression is stimulated. However, no specific cis element recognized by T antigen was identified that was involved in the negative regulation of JCV early gene expression.

Tada et al. (1991) provided other evidence that the JCV late promoter is under negative control. They used pJCLCAT plasmids containing different deletions in the JCV promoter-enhancer region. CAT assays using hamster fetal glial cells (HJC) showed that a cis-acting element present within the poly(A)-penta repeat motif region (nt 12-35) of the 98 bp repeat, which is distal to the late region, may act as a silencer motif. The same mechanism probably applied to the proximal repeat. The remaining sequences were involved in the positive regulation of JCV late gene expression. To define this silencer sequence, they used synthetic complementary

oligonucleotides spanning nt 22-42 which contained either a wild type AGGGAAGGGA (OP1) or mutant (OP2) penta repeat motif. After insertion into the enhancerless pA10CAT$_2$ plasmid in the late orientation, CAT activity was tested in HJC cells. Results showed that OP1 decreased while OP2 increased CAT activity. Tada et al. (1991) concluded that the penta repeat motif immediately adjacent to the positive-acting poly(A) tract was involved in JCV late gene negative regulation.

Mobility shift assays with OP1 using nuclear extracts from HJC glial cells showed three distinct DNA-protein complexes. Two were formed by proteins interacting with penta repeat motif. UV crosslinking identified three proteins of 56-60, 28 and 25 kDa that were bound to the penta repeat motif. The 56-60 kDa protein appeared to be in contact with the three G residues of the penta repeat motif. Taken together, Tada et al. (1991) suggested that this penta repeat motif may act as a negative regulator of JCV late gene expression in the JCV early lytic cycle. However, it remained unclear how the late promoter was activated after DNA replication was initiated. A more recent report by this group added more complication to this factor by showing that the penta-binding proteins are involved in T antigen-mediated JCV DNA replication (Chang et al., 1994).

To define the function of this penta repeat motif further, I introduced a set of nucleotide substitutions that

changed the AGGGAAGGGA penta repeat motif within the promoter-enhancer of JCV to ACAGATCTAA. Wild type and mutant whole promoter-enhancer fragments were inserted into the promoterless-enhancerless pSV0CAT in early and late orientations. The plasmids were then transfected into P19 RA cells and U87 cells to analyze the effect of the mutations on CAT activities. In contrast to previous results (Tada et al., 1991), my results showed that this cis-acting sequence negatively regulated JCV early expression and upregulated JCV late gene expression (Fig. 3). This result is not unexpected due to the following reasons. First, as JCV early gene expression is restricted mostly to glial cells and virus infected oligodendrocytes in PML patients, an element might be acting in cis to prevent expression of the JCV early genes in nonglial cells and during the late stages of infection. Second, similar sequences upstream of the c-myc and myosin heavy chain gene have been implicated in regulating these genes negatively. The penta repeat motif in JCV might also have the same function, as suggested previously (Kumar et al., 1991). Third, in my experiments I used the whole promoter-enhancer of JCV Mad1 strain that contains two 98 bp repeats. Therefore, the function of penta was tested in its natural context. However, in a previous study, a synthetic oligonucleotide was used and inserted into a heterologous promoter and the CAT activity was tested in the hamster fetal

glial cells (Tada *et al.*, 1991). They could not explain why the insertion of the mutant penta repeat motif in pA10CAT could activate the CAT activity. Fourth, differences in the cell types used might also contribute to the different results. In my data, the mutants were tested in P19 RA cells and U87 cells. The same result was observed in both cell types, except that the change was much more significant in U87 cells than in P19 cells. The mixed cell type of mouse P19 RA cell compared with the single glial cell type of human U87MG and the more efficient expression of JCV is in human glial cells might be responsible for the difference. It is possible that the difference between these results might be caused by different proteins recruited into the transcription machinery in different cell systems. These proteins would regulate JCV gene expression differently during their interaction with the penta repeat motifs. Thus, to clarify these issues, it was necessary to identify and isolate proteins that interacted with the penta repeat motif.

4.2 *In vitro* binding activity to penta repeat motif

My mobility shift assay results identified two DNA-protein complexes in P19 RA cells and only one in U87 glial cells. The results of P19 RA cells agreed with that of Tada *et al.* (1991) showing two DNA complexes (A and C) with the penta repeat motif. My mobility shift assays using extracts

from P19 DM cells and P19 UD cells also showed the same pair of complexes (Fig. 5; Fig. 6), suggesting that the proteins interacting with penta repeat motif were present in glial and nonglial cells. My results support the suggestion of Kumar et al. (1991) that the penta repeat motif acting as a negative regulator prevents JCV expression in nonglial cells. In a UV crosslinking experiment, Tada et al. (1991) identified three proteins of 56-60, 28 and 25 kDa that were bound specifically to penta repeat motifs. My experiments using P19 RA cell extracts identified two proteins of 60 and 43 kDa (Fig. 8). Thus, my data and the data of others clearly demonstrated that there are multiple proteins interacting with the penta repeat motif. These proteins may play different roles in the JCV early and late gene expression. It also would be interesting to identify the proteins binding to penta motif in P19 DM cells and P19 UD cells to determine if they have the same size in UV cross-linking assays as those from P19 RA cells. This would further clarify the negative role of the penta repeat motif and the penta binding proteins in JCV early gene expression.

At the time I finished my experiments, Kumar et al. (1994) published their identification of another protein interacting with the penta repeat motif. The protein had a different size compared with their previous results and the previous data of others (Kumar et al., 1992 and Tada et al.,

1991). The new protein of Kumar et al. (1994) had a size of 70-80 kDa and was bound to the penta motif. However, the binding was observed only in the presence of the penta and the neighbouring sequences. Thus, to date, at least five proteins of 70-80, 53 or 56-60, 43, 28 and 25 kDa have been identified that interact with penta repeats. Among these, the protein of 53 or 56-60 kDa was detected both in glial and nonglial cells. Most probably, the five or more proteins bind to the penta repeat motif with different affinity. For example, the 43 and 53-56-60 kDa proteins may bind penta repeat motif with higher affinity than the 70-80 kDa protein, as the 70-80 kDa protein binds to the penta motif only in the presence of neighbouring sequences. However, the 43 and 53-56-60 kDa proteins could bind to this motif in the absence of the additional sequences, suggesting the negative role of these proteins in regulating JCV early gene expression in nonglial cells. Their higher affinity in binding to the penta motif might serve to displace the 70-80 kDa protein.

In an earlier study, Kumar (1994) showed that JCV T antigen facilitated the binding of NF1 to the NF1 II/III sites in vitro. He found that the presence of T antigen increased the binding of NF1 to the NF1 II/III sites in mobility shift assays using P19 RA cell extracts expressing T antigen. This increase was blocked by incubating the reaction with a specific T antigen antibody, but not a nonspecific antibody.

He suggested that an increase in binding of NF1 to the NF1 II/III sites might be involved in the mechanism by which T antigen activates JCV late expression. The interaction of the TEF-1 and the basal transcriptional machinery has been shown to be essential for activation of the late promoter transcription in the SV40 system, and SV40 T antigen has been suggested to function as a coactivator through its interaction with TEF-1 and TBP (Coulombe et al., 1992; Gruda et al., 1993). Because penta repeat motifs are located between the NF1 II/III sites and the TATA box in JCV, the binding of a protein to the penta repeat motif might block NF1-GTFs interaction and repress JCV late promoter activity. Based on these studies, Kumar (1994) proposed a model for the mechanism of the JCV late promoter activation by T antigen (Kumar, 1994). In this model, either T antigen would remove the 56 kDa protein from the penta repeat motif or it would prevent the binding of the 56 kDa protein. Thus, T antigen would increase the binding of NF1 to NF1 II/III, and facilitate the interaction of NF1 with the GTFs assembled on the TATA box, resulting in the activation of the JCV late promoter by NF1.

According to the model of Kumar (1994), T antigen inactivates the transcriptional function of the 56 kDa protein by sequestering it or blocking its binding to penta repeat motif. To directly test this model, I performed mobility shift assays using extracts from P19 RA cells that had been

transfected with T antigen using the conditions described by Kumar (Kumar, 1994). I used an oligonucleotide containing the wild type penta repeat motif as the probe. In my experiment, the specific binding observed in the absence of T antigen was also observed in the presence of T antigen. Pre-incubation of the reaction mixture with an antibody specific to T antigen or a nonspecific antibody had no effect on binding (Fig. 7). This result clearly indicated that T antigen was not directly interacting with the penta repeat motif-binding factors. In the original study, the binding of NF1 to NF1 II/III showed no supershift in the presence of T antigen and T antigen specific antibody. A recent study also demonstrated that, for polyoma virus, T antigen was unable to activate the late promoter in the absence of DNA replication. T antigen could activate late expression of replication-incompetent genomes only in the presence of actively replicating polyoma virus genomes (Liu et al., 1993). Thus, the activation of JCV late genes by T antigen most probably involves replication-related processes, but not through T antigens' interacting with NF1 or penta repeat motif-binding proteins. However, the possibility that T antigen modifies NF1 function through interactions that were too weak to allow my detection of them cannot completely be excluded at this time.

4.3 Isolation of a cDNA clone encoding a protein that

represses JCV early expression and comparison with other penta repeat motif-binding factors.

Data from our lab and the data of others indicated that there are at least five proteins in both nonglial and glial cells binding to the penta repeat motif in the control region of JCV (Tada et al., 1991; Kumar et al., 1992 and Kumar et al., 1994). To further understand the regulation of JCV expression, the isolation of the cDNA encoding such proteins emerged as a necessity. The isolation of cDNA clones using protein-DNA interactions has been well documented. This method was initially established by Singh et al. (1988) when they isolated a cDNA clone encoding the immunoglobulin enhancer binding protein, and was modified later by Vinson et al. (1988), Clerc et al. (1988) and Staudt et al. (1988). A modified version of this technique uses a concatenated probe containing multiple sequences of the binding site. Another important step in this technique is to denature and renature the fusion proteins produced in the bacteria with a high concentration of guanidine HCl. This process allows the fusion protein to fold into a conformation suitable for binding to the DNA ligand. Following this method, I screened a cDNA expression library from mRNA of P19 RA cells and isolated two cDNA clones using a concatenated probe containing multiple penta repeat motifs. The first was not studied further, since it also interacted with nonpenta sequences in

plaque hybridization assays. However, the possibility that it represents the other protein binding to penta in UV cross-linking assays (Fig. 8) cannot be excluded. The second cDNA clone specifically bound to penta repeat motif (Fig. 9). Protein expressed from this cDNA, named RP1, bound to the penta repeat motifs specifically, as demonstrated by its binding only to the wild type penta repeat motif, but not to the mutant penta repeat motif and not to other nonspecific DNA probes. The DNA binding activity was also demonstrated in the Southwestern blot assays. In this assay, only proteins produced from IPTG-induced RP1 recombinant phage-containing bacteria showed the ability to bind to the penta repeat motif probe (Fig. 11).

My early studies shoed that the penta repeat motif-binding protein(s) existed in both glial and nonglial cells. Thus, it was also important to analyze the expression and presence of RP1 in these cells. Northern blots were performed using the whole cDNA as a probe. RP1 expression was detected in P19 RA glial cells, P19 UD cells and P19 DM muscle cells (Fig. 10). This data again supported my previous mobility shift assay results showing penta-protein complexes with extracts from all three cell types (Fig. 5 and 6). Moreover, in cotransfection assays, RP1 significantly repressed JCV early promoter activity in U87 cells (Fig. 13). Taken together, these data suggested that the cDNA I isolated

encoded a functional cellular protein involved in regulation of JCV early gene expression. Furthermore, the cotransfection results were also consistent with my previous site-directed mutagenesis data showing that penta repeats motif act in cis to repress JCV early expression (Fig. 3). The effects of mutation on JCV gene expression were more significant in U87 cells than in P19 RA cells.

At the time I was preparing this thesis, the same laboratory that found that the penta repeat motif acted to repress JCV late expression also isolated a cDNA encoding a DNA binding factor named YB-1 (Kerr et al., 1994). This factor was isolated from a HeLa cell cDNA library using a probe for region B of the JCV control region spanning the NF1, penta motifs, and poly(A) stretch. The binding of YB-1 could be specifically competed by the penta repeat motif. In contrast with their previous result showing that proteins binding to penta acted as a silencer, YB-1 positively regulated JCV late gene expression. This finding supports my result that penta motif is required for JCV late gene expression. Moreover, in agreement with my results that demonstrated the ubiquitous presence of penta repeat motif binding protein (Fig. 5; Fig. 6 and Fig. 10), YB-1 message was also detected in various cells and tissues, including brain cells and cells of other tissues, both glial and nonglial cells. According to my results, penta repeat motif showed

dual functions by activating JCV late gene expression and repressing JCV early gene expression (Fig. 3). Thus, a factor acting in trans through interaction with penta repeat motif would most probably possese the same dual functions. Interestingly, a more recently study by the group of Khalili showed that YB-1 could bind to the opposite strand of the binding site of Purα, a strong activator of early gene transcription binding to single strand DNA, and reduce the binding of Purα. Further, Purα and T antigen could increase the binding of YB-1 to DNA. Thus, YB-1 could act as a weak activator to repress the JCV early gene expression through a cooperative action with Purα and T antigen (Chen et al., 1995).

4.4 The significance of RP1 zinc finger motifs

Analysis of the RP1 cDNA sequence revealed that its 5' end was not present. The sequence that was cloned encodes a predicted protein of 222 amino acids from the fusion with β-gal to the stop codon at nt 673. Northern blots showed that the mRNA of this cDNA is 2.2 kb, that would potentially code for approximately 370 amino acids, excluding the 5' untranslated region (UTR) sequences. Therefore, RP1 probably encodes the 43 kDa protein identified by UV crosslinking. RP1 contains seven putative zinc finger motifs with an amino acid sequence of CysXXCysXXXXHisXXXXXCys, suggesting that RP1

belongs to the superfamily of zinc finger proteins. The zinc finger is the most common DNA binding motif. It was initially identified in the transcription factor IIIA (TFIIIA) from *Xenopus laevis* (Miller et al., 1985). Later studies found that homologous zinc fingers existed in many proteins regulating eukaryotic gene transcription. For instance, the zinc finger motif can be found in many signal transduction proteins, proto-oncogenes and general transcription factors (Klug et al., 1987; Berg, 1990; Kaptein, 1991 and Klug, 1993). Zinc finger proteins thus comprise a superfamily of transcription factors. The central characteristic of this family was initially described as proteins that usually have tandem repeats of about 30 amino acids containing the Cys-X_{2-4}-Cys-X_{12}-His-X_{3-5}-His/Cys sequence (Miller et al., 1985). The cysteine and histidine pairs within the motif are folded together to bind a zinc ion like a finger (Pabo, 1992). However, later studies have defined the zinc finger as any sequence that has a set of cysteine and histidine residues clustered in a short polypeptide region (Pabo, 1992). A zinc finger protein can contain from as few as 2 to as many as 37 tandem fingers. The different combinations of zinc fingers would allow zinc finger proteins to specifically recognize a DNA sequence. Thus, the identification of the seven putative zinc fingers in my isolated RP1 was consistent with its DNA-

binding ability.

4.5 The mechanism of regulation of JCV gene expression by RP1 and penta repeat motif

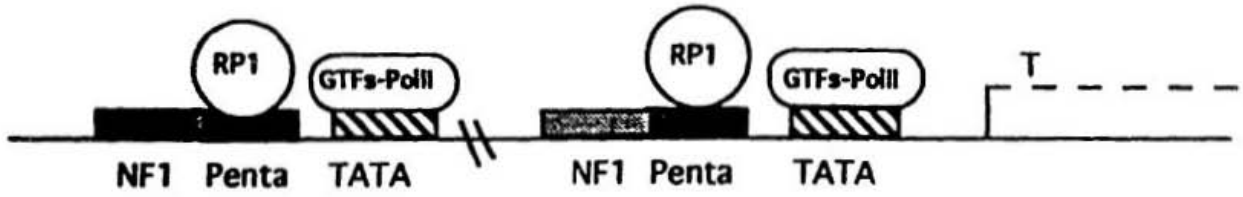
JCV is the most strictly cell-specific virus that propagates only in brain cells of glial origin. As demonstrated by Khalili et al. (1988), this restriction is due to the regulation of JCV gene transcription. The replication of JCV is strongly dependent on the function of T antigen. Because T antigen is expressed from the early orientation of JCV and is required during early infection, the expression of JCV early genes should be more active in newly infected cells than late genes (Kumar, 1994). After the expression of T antigen, most probably in glial cells, T antigen initiates replication that activates JCV late gene expression. Previous data have shown that NF1 is specifically present only in glial cells (Kumar et al., 1993). Thus, NF1 may act as the major factor in the onset of JCV early gene expression. My results clearly indicated that the penta repeat motif functions in cis to repress JCV early expression. Moreover, the penta motif binding protein, RP1, is present in both glial and nonglial cells. Furthermore, the major early gene transcription initiation sites are located about 20 to 30 bp downstream from the TATA boxes that are close to the penta motif and NF1 site. All these results suggest that the penta motif may act in cis

to repress JCV early expression in nonglial cells and possibly in late infections of glial-type oligodendrocytes.

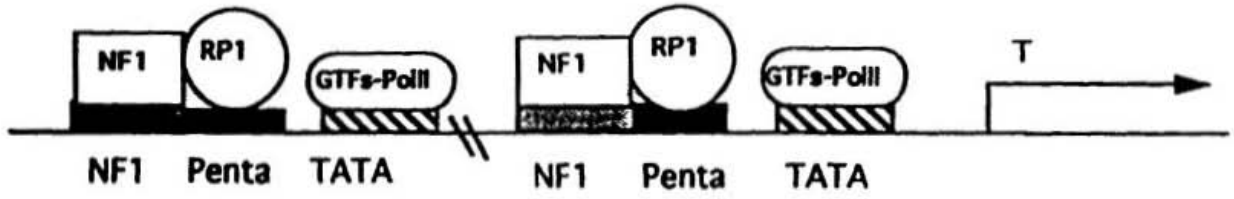
Fig. 14. Proposed mechanism for regulation of JCV early gene expression by penta motif-RP1 interaction.

The duplicated TATA box, penta motif, mutated penta motif (mpenta) and NF1 elements are as indicated and symbolized by the respective rectanges. Panels are: A, in nonglial cells, the binding of RP1 to the penta motif and the absence of NF1 repress JCV early gene transcription; B, In glial cells, the glial-specific NF1 interacts with NF1 sites and only moderately activates transcription, due to the interference of RP1; C, In glial cells, the mutation of the penta motifs abolishes the binding of RP1 and NF1 fully activates JCV early transcription. The dashed line indicates no transcription of T antigen (T). The thin arrow indicates moderately activated transcription. The thick arrow indicates fully activated transcription.

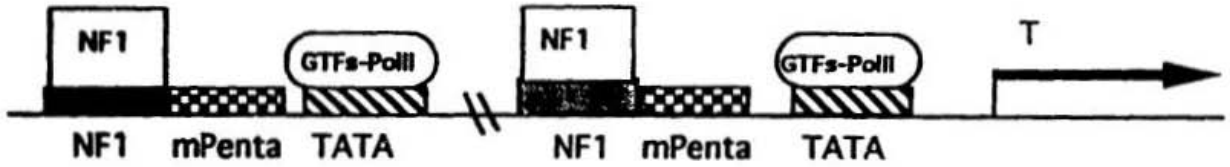
A. Nonglial cells



B. Glial cells



C. Glial cells



Based on the observations described above, I propose the following model to explain the function of the penta repeat motif and the RP1 protein in the regulation of JCV early gene expression (Fig. 14). In this model, first, the penta repeat motif binds RP1 which exists ubiquitously (Fig. 14 A). In the absence of NF1 in nonglial cells, RP1 binds to the penta repeat motif and represses JCV early expression. Second, the binding of the glial cell-specific NF1 to the NF1 sites II/III immediately adjacent to the penta repeat motif partially, but not completely, relieves the repression by RP1 (Fig. 14 B), as suggested by the increase in the expression of the JCV early promoter-enhancer in glial cells after mutation of the penta repeat motif (Fig. 14 C). Thus, the function of NF1 is to allow JCV to express JCV T antigen at a moderate level. Subsequently, the induction of T antigen in glial cells stimulates the JCV replication and late gene expression leading to early gene repression. This could be examined by my system by cotransfection T antigen expression plasmid with pJCCAT reporter plasmid containing wild type and mutant penta repeat motifs.

This model explains the poor propagation of JCV in nonglial cells and why JCV expresses moderately even in glial cells. However, as the mechanism by which NF1 activates JCV early expression is unknown, it is impossible to clarify how NF1 might relieve the repression of RP1. One possibility is

that NF1 functions by interacting directly or via T antigen with the downstream general transcription initiation complex, causing the GTFs to change the conformation of the initiation complex. Another possibility is that NF1 directly interacts with RP1, as suggested by the closely proximate of their binding sites. Further investigation of NF1-RP1 protein-protein interactions in vivo and in vitro should clarify this issue. For example, both NF1 and RP1 cDNAs isolated in our lab could be expressed in HeLa cells and their ability to affect JCV expression can be tested by CAT assays in glial and nonglial cells. Also, the binding of NF1 and RP1 to NF1 and penta repeat motifs can be tested by mobility shift and supershift assays using extracts of various cell types. In this experiment, we would examine how these two proteins affect each other. To directly test the possible interaction between NF1 and RP1, the yeast two hybrid system can also be employed (Fields and Song, 1989).

The role of the penta repeat motif in JCV late gene expression regulation remains unclear. My studies have shown that the penta repeat motif acts differently in regulating early and late genes (Fig. 3). At least two proteins were identified that interacted with penta elements, a 43 kDa protein and a 60 kDa protein (Fig. 8). However, my isolated RP1 appeared to encode the 43 kDa species and only affected early gene expression of JCV. This prompts me to suggest that

these proteins are functioning differently for the expression of JCV late and early genes. I speculate that the early and late gene transcription machinery may recruit different proteins. For instance, RP1 may regulate early expression while the 60 kDa protein regulates late expression. The recently-isolated YB-1 activates JCV late gene expression through an element overlapping one penta repeat motif (Kerr et al., 1994). Although it is not clear if YB-1 activates or represses early gene expression of JCV, the difference of YB-1 and RP1 supports the suggestion that different proteins interact with the JCV penta repeat motif and each plays a role in JCV early or late gene expression. It is also possible that the binding affinity of these proteins is different, and that the binding of one protein will affect the binding of the other protein. Thus, further studies of the binding affinity of RP1 and YB-1 to the penta repeat motif and the protein-protein interactions between YB-1 and RP1 should help to clarify this issue. As a possible explanation for the different sequence structures of the two penta binding proteins, it should be noted that YB-1 is isolated from human cells and RP1 is isolated from mouse cells. Thus, the activation by YB-1 and the repression by RP1 in late expression of JCV late viral proteins may explain why JCV virus propagates efficiently only in human cells.

REFERENCES

Ahmed, S., J. Rappaport, H. Tada, D. Kerr, and K. Khalili. 1990. A nuclear protein derived form brain cells stimulates transcription of the human neurotropic virus promoter, JCVE, in vitro. J. Biol. Chem. *265*, 13899-13905.

Alwine, J.C. 1982. Evidence of simian virus 40 late transcriptional control: Mixed infectious of wild-type simian virus 40 and a late leader deletion mutant exhibit trans effects on late viral RNA synthesis. J. Virol. *42*, 798-803.

Amemiya, K., R. Traub, L. Durhan, and E.O. Major. 1989. Interaction of a nuclear factor-1-like protein with the regulatory region of the human polyomavirus JC virus. J. Biol. Chem. *264*, 7025-7032.

Arthur, R.R., S. Dagostin, and K. Shah. 1989. Detection of BK virus in urine and brain tissue by the polymerase chain reaction. J. Clini. Micro. *27*, 1174-1179.

Assouline, J.G., and E.O. Major. 1991. Human fetal Schwann cells support JC virus multiplication. J. Virol. *65*, 1002-1006.

Beggs, A.H., R.J. Frisque, and G.A. Scangos. 1988. Extinction of JC virus tumour antigen expression in glial cell-fibroblast hybrids. Proc. Natl. Acad. Sci. USA *85*, 7632-7636.

Benezra. R., R.L., Davis, D., Lockshon, D.L., Turner, and H., Weintraub. 1990. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell *61*, 49-59.

Berg, J.M. 1990. Zinc finger domains: hypotheses and current knowledge. Ann. Rev. Biophs. Chem. *19*, 405-421.

Bollag, B., W.-F. Chuke, and R.J. Frisque. 1989. Hybrid genomes of the polyomaviruses JC virus, BK virus, and simian virus 40: Identification of sequences important for efficient transformation. J. Virol. *63*, 863-872.

Bolen, J.B., D.G. Anders, J. Trempy, and R.A. Consigli. 1981. Differences in the subpopulation of the structural proteins

of polyoma virions and capsid: Biological functions of the V species. J. Virol. *37*, 80-91.

Bouvaget, P.F., E.E. Strehler, G.E. White, M.A. Strelher-Page, B. Nadalginard, and V. Mahcavi. 1987. Multiple positive and negative 5'-regulatory elements control the cell-type-specific expression of the embryonic skeletal myosin heavy-chain gene. Mol. Cell Biol. *7*, 4377-4389.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry *72*, 248-254.

Brady, J.N., C. Lavialle, and N.P. Salzman. 1980. Efficient transcription of a compact nucleoprotein complex isolated from purified simian virus 40 virions. J. Virol. *35*,371-381.

Brady, J.N., C. Lavialle, M. Radonovich, and N.P. Salzman. 1981. Simian virus 40 maturation: chromatin modifications increase the accessibility of viral DNA to nuclease and RNA polymerase. J. Virol. *39*, 603-611.

Buratowski, S. 1994. The basics of basal transcription by RNA polymerase II. Cell. *77*, 1-3.

Chang, C.F., H. Tada, and K. Khalili. 1994. The role of a pentanucleotide repeat sequence, AGGGAAGGGA, in the regulation of JC virus DNA replication. Gene. *148*, 309-314.

Chen, N.N., C.F. Chang, G.L. Gallia, D.A. Kerr, E.M. Johnson, C.P. Krachmarov, S.M. Barr, R.J. Frisque, B. Bollag, and K. Khalili. 1995. Cooperative action of cellular proteins YB-1 and Purα with the tumor antigen of the human JC polyomavirus determines their interaction with the viral lytic control element. Proc. Natl. Acad. Sci. USA. *92*, 1087-1091.

Chen, C., and H. Okayama. 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. *7*, 2745-2752.

Chesters, P.M., J. Heritage, and D.J. McCance. 1983. Persistence of DNA sequences of BK virus and JC virus in normal human tissues and in diseased tissues. J. Infect. Diseases *147*, 676-684.

Chodosh, L.A., R.W. Charthew, and P.A. Sharp. 1986. A single

polypeptide possesses the binding and transcription activities of the adenovirus major latetranscription factor. Mol. Cell Biol. *6*, 4723-4733.

Chodosh, L.A., R.W. Charthew, and P.A. Sharp. 1988. Human CCAAT-binding proteins have heterologous subunits. Cell *33*, 11-24.

Chowdhury, M., J.P. Taylor, H. Tada, J. Rappaport, F.Wong-Staal, S. Amini, and K. Khalili. 1990. Regulation of the human neurotropic virus promoter by JCV-T antigen and HIV-1 tat protein. Oncogene *15*, 1737-1742.

Chowdhury, M., M. Kundu, and K. Khalili. 1993. GA/GC-rich sequence confers tat responsiveness to human neurotropic virus promoter, JCV_L, in cells derived from central nervous system. Oncogene *8*, 887-892

Chuke, W.F., D.L. Walker, L.B. Peitzman, and R.J. Frisque. 1986. Construction and characterization of hybrid polymomavirus genomes. J. Virol. *60*, 960-971.

Clark, R., K. Peden, J.M. Pipas, D. Nathans, and R. Tjian. 1983. Biochemical activities of T antigen proteins encoded by simian virus 40 A gene deletion mutants. Mol. Cell Biol. *3*, 220-228.

Clerc, R.G., L.M.Corcoran, J.H. LeBowitz, D. Baltimore, and P.A. Sharp. 1988. The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. Genes Dev. *2*, 1570-1581.

Colby, W.W., and T. Shenk. 1982. Fragments of the SV40 transforming gene facilitate transformation of rat embryo cells. Proc. Natl. Acad. Sci. USA *79*, 5189-5193.

Coulombe, J., L. Berger, D.B. Smith, R.K. Hehl, and A. G. Wildeman. 1992. Activation of simian virus 40 transcription in vitro by T antigen. J. Virol. *66*, 4591-4596.

Daniel, A.M., and R. Frisque. 1993. Transcription initiation sites of prototype and variant JC virus early and late messenger RNAs. Virology. *194*, 97-109.

DeCaprio, J.A., J.W. Ludlow, and J. Figge. 1988. SV40 large tumour antigen forms a specific complex with the product of the retinoblastoma susceptibility gene. Cell *54*, 275-283.

Deyerle, K.L., F.G. Sajjadi, and S. Subramani. 1989. Analysis of origin of DNA replication of human papovavirus BK. J. Virol. *63*, 356-365.

Dorries, K. 1984. Progressive multifocal leucoencephalopathy: analysis of JC virus DNA from brain and kidney tissue. Virus Res. *1*, 25-38.

Dyson, N., R. Bernards, S.H. Friend, L.R. Gooding. 1990. Large T antigens of many polyomaviruses are able to form complexes with the retinoblastoma protein. J. Virol. *64*, 1353-1356.

Eckhart, W. 1990. Polyomavirinae and their replication. Virology, pp. 1593-1607, Raven Press. Ltd. Second edition.

Feigenbaum, L. K. Khalili, E. Major, and G. Khoury. 1987. Regulation of the host range of human papovavirus JCV. Proc. Natl. Acad. Sci. USA. *84*, 3695-3698.

Fields, S., and O.K. Song. 1989. A novel genetic system to detect protein-protein interactions. Nature *340*, 245-246.

Frisque, R.L., 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.

Frisque, R.J., G.L. Bream, and M.T. Cannella. 1984. Human polyomavirus JC virus genome. J. Virol. *51*, 458-469.

Gill, G., and R. Tjian. 1992. Eukaryotic coactivators associated with the TATA box binding protein. Opinion in Genetics and Development *2*, 236-242.

Gill, G., E., Pascal, Z., Tseng, and R., Tjian. 1994. A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the $dTAF_{II}$ 110 component of the Drosophila TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. USA. *91*, 192-196.

Glisin, V., R. Crkvenjakov, and C. Byus. 1974. Ribonucleic acid isolated by cesium chloride centrifugation. Biochemistry *13*, 2633-2637.

Gluzman, Y. 1981. SV40 transformed cells support the replication of early SV40 mutants. Cell *23*, 175-182.

Goodrich, J.A., T. Hoey, C. Thut, A. Admon, and R. Tjian. 1993. Drosophila TAFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. Cell *75*, 519-530.

Goodrich, J.A., and R. Tjian. 1994. Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. Cell *77*, 145-156.

Gorman, C.M., M.F. Moffat, and B.H. Howard. 1992. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. *2*, 1044-1051.

Grinnell, B.W., B.L. Padgett, and D.L. Walker. 1983. Comparison of infectious JC virus DNAs cloned from human brain. J. Virol. 45, 299-308.

Gruda, M.C., J.M. Zabolotny, J.H. Xiao, I. Davidson, and J.C. Alwine. 1993. Transcriptional activation by simian virus 40 large T antigen: Interactions with multiple components of the transcription complex. Mol. Cell. Biol. *13*, 961-969.

Ha, I., S. Roberts, E. Maldonado, X. Sun, L.U. Kin, M. Green, and D. Reinberg. 1993. Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. Genes Dev. *7*, 1021-1032.

Haggerty, S., D.L. Walker, and R.J. Frisque. 1989. JC virus-simian virus 40 genomes containing heterologous regulatory signals and chimeric early regions: identification of regions restricting transformation by JC virus. J. Virol. *63*, 2180-2190.

Hennighausen, L., and H. Luban. 1987. Interaction of protein with DNA in vitro. Methods in Enzymology *152*, 721-735.

Hoey, T., R.O.J. Weinzierl, G. Gill, J.L. Chen, B.D. Dynlacht, and R. Tjian. 1993. Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. Cell *72*, 247-260.

Huang, S.H., J.M. Tomich, H. Wu, A. Jong, and J. Holcenberg. 1989. Human deoxycytidine kinase, Sequence of cDNA and analysis of expression in cell lines with and without enzyme activity. J. Biol. Chem. *264*, 14762-14768.

Jay, G., S. Nomura, C.W. Anderson, and G. Khoury. 1981. Identification of the SV40 agnogene product: A DNA binding protein. Nature *291*, 346-349.

Kakkix, E., K.J. Riggs, W. Gillespie, and K. Calamer. 1989. A transcriptional repressor of c-myc. Nature *339*, 718-721.

Kenny, 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.

Kenny, S., V. Natarajan, and N.P. Salzman. 1986a. Mapping 5'termini of JC virus late RNA. J. Virol. *58*, 216-219.

Kenny, S., V. Natarajan, G. Selzer, and N.P. Salzman. 1986b. Mapping 5' termini of JC virus early RNAs. J. Virol. *58*, 651-654.

Kerr, D., C.F. Chang, N.cChen, G. Gallia, G. Raj, B. Schwartz, and K. Khalili. 1994. Transcription of a human neurotropic virus promoter in glial cells: Effect of YB-1 on expression of the JC virus late gene. J. Virol. *68*, 7637-7643.

Khalili, D., J. Brady, and G. Khoury. 1987. Evidence for a shift in 5'-termini of early viral RNA during the lytic cycle of JC virus. Virol. *158*, 469-472.

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.

Kim, S.H., 1992. β-Ribbon: A new DNA recognition motif. Science *255*, 1217-1218.

Kitamura, T., Y. Aso, N. Kuniyoshi, K. Hara, and Y. Yogo. 1990. High incidence of urinary JC virus excretion in nonimmunosuppressed older patients. J. Infect. Diseases *161*, 1128-1133.

Klug, A., and D. Rhodes. 1987. Zinc fingers: A novel protein motif for nucleic acid recognition. Trends in Biochemical Sciences *12*, 464-469.

Klug, A. 1993. Opening the gateway. Nature *365*, 486-487.

Kumar, G., D.K. Srivastava, A.K. Sharma, W. Tefera, and D. Patel. 1992. Evidence that AGGGA domain, in its natural

context is involved in the positive transcriptional regulation of JC virus early promoter. J. Cell. Biochem. (Suppl.) *16* (E), 165.

Kumar, K.U., A. Pater, and M.M. Pater. 1993. Human JC virus perfect palindromic nuclear factor 1-binding sequences important for glial cell-specific expression in differentiating embryonal carcinoma cells. J. Virol. *67*, 572-576.

Kumar, K.U. 1994. Transcriptional regulation of human neurotropic papovavirus, JCV. Ph.D thesis, Memorial University of Newfoundland.

Kumar, G., D.K. Srivastava, and W. Tefera. 1994. A 70- to 80-kDa glial cell protein interacts with the AGGGAAGGGA domain of the JC virus early promoter only in the presence of the neighbouring cis DNA elements. Virol. *203*, 116-124.

Lane, D.P., V. Simanis, R. Bartsch, J. Yewdell, J. Gannon, and S. Mole. 1985. Cellular targets for SV40 large T-antigen. Proc. R. Soc. Lond. B *226*, 25-42.

Lashgari, M.S., H. Tada, S. Amini, and K. Khalili. 1989. Regulation of JCV_L promoter function: Transactivation of JCV_L promoter by JCV and SV40 early proteins. Virol. *170*, 292-295.

Lewin, B. 1990. Commitment and activation at Pol II promoters: A axil of protein-protein interactions. Cell *61*, 1161-1164.

Lin, Y.S., and M.R. Green. 1991. Mechanism of action of an acidic transcriptional activator in vitro. Cell *64*, 971-981.

Liu, Z., and G.G., Carmichael. 1993. Polyoma virus early-late switch: Regulation of late RNA accumulation by DNA replication. Proc. Natl. Acad. Sci. USA *90*, 8494-8498.

Loeber, G., and K. Dörries. 1988. DNA rearrangements in organ-specific variants of polyomavirus JC strain GS. J. Virol. *62*, 1730-1735.

London, W., S. Houff, D. Madden, D. Fuccillo, M. Gravel, W. Wallan, A. Palmer, J. Sever, B. Padgett, D. Walker, G. ZuRheim, and T. Ohashi. 1978. Brain tumours in owl monkeys inoculated with a human polyomavirus (JC virus). Science *201*,

1246-1249.

Lynch, K.J., and R.J. Frisque. 1991. Factors contributing to the restricted DNA replicating activity of JC virus. Virol. *180*, 306-317.

Major, E.O., A.E. Miller, P. Mourrain, R.G. Traub, E. deWidt, and J. Sever. 1985. Establishment of a line of human fetal glial cells that supports JC virus multiplication. Proc. Natl. Acad. Sci. USA. *82*, 1257-1261.

Mandl, C., D.L. Walker, and R.J. Frisque. 1987. Derivation and characterization of POJ cells, transformed human fetal glial cells that retain their permissivity for JC virus. J. Virol. *61*, 755-763.

Manos, M.M., and Y. Gluzman. 1985. Genetic and biochemical analysis of transformation-competent, replication defective simian virus 40 large T antigen mutants. J. Virol. *53*, 120-127.

Martin, J.D., D.M. King, J.M. Slauch, and R.J. Frisque. 1985. Differences in regulatory sequences of naturally occurring JC virus variants. J. Virol. *53*, 306-311.

Miller, J., A.D. Mclachlan and A. Klug. 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from xenopus oocytes. EMBO. J. *4*, 1609-1614.

Nakshatri, H., M.M. Pater, and A. Pater. 1990. Activity andenhancer binding factors for JC virus regulatory elementsin differentiating embryonal carcinoma cells. Virol. *177*, 784-789.

Ng, S.C., J.E. Mertz, W.S. Sanden, and M. Bina. 1985. Simian virus 40 maturation in cells harbouring mutants deleted in the agnogene. J. Biol. Chem. *260*, 1127-1132.

Pabo, C.O., and R.T. Sauer. 1992. Transcription factors: Structural families and principles of DNA recognition. Ann. Rev. Biochem. *61*, 1053-1095.

Padgett, B.L., and D.L. Walker. 1973. Prevalence of antibodies in human sera against JC virus, an isolate from a case of progressive multifocal leukoence-phalopathy. J. Infect. Dis. *127*, 467-470.

Padgett, B.L., D.L. Walker, G.M. ZuRhein, and J.N. Varakis. 1977. Differential neurooncogenicity of strains of JC virus, a human polyoma virus, in newborn Syrian hamsters. Cancer Res. *37*, 718-720.

Parvin, J.D., and A. Sharp. 1993. DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. Cell *73*, 533-540.

Proudffoot, N.J., and G.G. Brownlee. 1974. Sequence at the 3'end of globin mRNA shows homology with immunoglobulin light chain mRNA. Nature *252*, 359-362.

Pugh, B.F., and R. Tjian. 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell *61*, 1187-1197.

Rudnicki, M.A., and M.W. McBurney. 1987. Cell culture methods and induction of differentiation of embryonal carcinoma cells. In "Teratocarcinoma and embryonic stem cells: A practical aproach" PP. 19-49, (Ed, e.J. Robertson). IRL press, Oxford.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning: A laboratory manual. CSH press, Cold Spring Harbor, New York.

Schwabe, J.W.R., and D. Rhodes. 1991. Beyond zinc fingers: Steroid hormone receptors have a novel structural motif for DNA recognition. Trends in Biochemical Sciences *16*, 291-296.

Scott, M.P., J.W., Tamkun, and G.W., Hartzell. 1989. The structure and function of homeodomain. Biochim. Biophys. Acta *989*, 25-48.

Shah, K.V., R.W. Daniel, and T.J. Kelly, Jr. 1977. Immunological relatedness of papovaviruses of the simian virus 40-polyoma subgroup. Infect. Immun. *18*, 558-560.

Shah, K.V., H.L. Ozer, H.N. Ghazey, and T.J. Kelly, Jr. 1977. Common structural antigen of papovaviruses of simian virus 40-polyoma subgroup. J. Virol. *21*, 179-186.

Sharma, A.K., and G. Kumar. 1991. A 53 kDa protein binds to the negative regulatory region of JC virus early promoter. FEBS Lett. *28*, 272-274.

Shi, Y., E. Seto, L.S. Chang, and T. Shenk. 1991. Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein. Cell *67*, 377-388.

Shohat, O., Greenberg, D. Reisman, M. Oren, and V. Rotter. 1987. Inhibition of cell growth mediated by plasmids encoding p53 antisense. Oncogene *1*, 277-283.

Shrivastava, A., S. Saleque, G.V. Kalpana, S. Artandi, S.P. Goff, and K. Calame. 1993. Inhibition transcriptionalregulator Yin-Yang-1 by association with c-Myc. Science *262*, 1889-1892.

Simmons, D.T., and M.A. Martin. 1978. Common methionine-tryptic peptides near the amino-terminal end of primate papovavirus tumour antigens. Proc. Natl. Acad. Sci. USA *75*, 1131-1135.

Singh, H., J.H. Lebowitz, A.S. Baldwin Jr, and P.A. Sharp. 1988. Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA. Cell *52*, 415-423.

Stahl, H., P. Droge, and R. Knippers. 1986. DNA helicase activity of SV40 large tumour antigen. EMBO. J. *5*, 1939-1944.

Staudt, L.M., R.G. Clerc, J. Singh, J.H. Lebowitz, P.A. Sharp, and D. Baltimore. 1988. Cloning of a lymphoid-specific cDNA encoding a protein binding the regulatory concatemer DNA motif. Science *241*, 577-580.

Stewart, S.E., B.E. Eddy, A.M. Gochenour, N.G. Borgese, andG.E. Grubbs. 1957. The induction of neoplasms with a substance released from mouse tumours by tissue culture. Virol. *3*, 380-400.

Stringer, K.F., C.J. Ingles, and J. Greenblatt. 1990. Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. Nature *345*, 783-786.

Tada, H., J. Rappaport, M. Lashgari, S. Amini, F. Wong-Ataal, 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.

Tada, H., M.S. 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. Virol. *180*, 327-338.

Tamura, T., T. Inoue, K. Nagata, and K. Mikoshiba. 1988. Enhancer of human polyoma JC virus contains nuclear factor 1 binding sequences; analysis using mouse brain nuclear extract. Biochem. Biophys. Res. Commun. *157*, 419-425.

Tasset, D., L. Tora, C. Fromental, E. Scheer, and P. Chambon. 1990. Distinct classes of transcription activating domains function by different mechanisms. Cell *62*, 1177-1187.

Tjian, R. and T. Maniatis. 1994. Transcriptional activation: A complex puzzle with few easy pieces. Cell *77*, 5-8.

Usheva, A., and T. Shenk. 1994. TATA-binding protein independent initiation: YY1, TFIIB, and RNA polymerase II direct basal transcription on supercoiled template DNA. Cell *76*, 1115-1121.

Vacante, D.A., R. Traub, and E.O. Major. 1989. Extension of JC virus host range to monkey cells by insertion of a simian virus 40 enhancer into the JC virus regulatory region. Virol. *170*, 353-361.

Vinson, C.R., K.L. LaMarco, P.F. Johnson, W.H. Landschulz, and S.L. McKnight. 1988. In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Dev. *2*, 801-806.

Wang, E.H., and R. Tjian. 1994. Promoter-selective transcriptional defect in cell cycle mutant ts 13 rescued by hTAFII250. Science *263*, 811-814.

Wagner, M., D.W. Drolet, and M.G. Rosenfeld. 1993. Regulation of JC virus by POU-domain transcription factor Tst-1: Implications for progressive multifocal leukoencephalopathy. Proc. Natl. Acad. Sci. USA *90*, 4743-4747.

White III, F.A., M. Ishaq, G.L. Stoner, and R.J. Frisque. 1992. JC virus DNA is present in many human brain samples from patients without progressive multifocal leukoencephalopathy. J. Virol. *66*, 5726-5734.

Yogo, Y., T. Kitamura, C. Sugimoto, T. Ueki, Y. Aso, K. Hara, and F. Taguchi. 1990. Isolation of a possible archetypical JC virus DNA sequence from nonimmuno-compromised individuals. J. Virol. *64*, 3139-3143.

Yoshiike, K., T. Miyamura, H.W. Chan, and K.K. Takemoto. 1982. Two defective DNAs of human polyomavirus JC adapted to growth in human embryonic kidney cells. J. Virol. *42*, 395-401.

Zawel, L. and D. Reinberg. 1993. Initiation of transcription by RNA polymerase II: A multi-step process. Prog. Nucl. Acids Res. *44*, 67-108.