Journal of NeuroVirology (1996) 2, 87–100

Expression of multiple classes of the Nuclear Factor-1 family in the developing human brain: differential expression of two classes of NF-1 genes

Charlotte Sumner[1], Toshiya Shinohara[1], Linda Durham[1], Renee Traub[1], Eugene O Major[1] and Kei Amemiya[1,2]

[1]*Laboratory of Molecular Medicine and Neuroscience, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892;* [2]*IGEN, Inc., Rockville, Maryland 20852, USA*

Nuclear factor-1 (NF-1) is a multifunctional protein that participates in both transcription and replication. NF-1 proteins exist as a family of proteins that share some common structural and functional features but also demonstrate organ and cell type specific expression. Based upon these characteristics, the family of NF-1 proteins is divided into four classes, A, B, C and D. Several NF-1 binding sites have been identified in the regulatory sequences of the human polyomavirus, JCV, which multiplies most efficiently in glial cells derived from human fetal brain. Nuclear proteins from these cultures bind specifically to these NF-1 sites. It is not known, however, which member(s) of the NF-1 family is expressed in cells susceptible to JCV infection. We have examined glial cells as well as HeLa cells, which are not permissive to JCV, for NF-1 expression. By RT-PCR analysis, all four classes of NF-1 are expressed in human fetal glial cells and HeLa cells. However, by Northern analysis the expression of class D gene is much higher in the glial cells than HeLa cells. Expression of the class C gene, first identified in HeLa cells as NF-1/CTF1, is barely detectable in glial cells but highly expressed in HeLa cells. The screening of cDNA libraries from two early human brain tissues resulted in the identification of a number of clones which appear to be related and belong to a single class of the NF-1 family, class D. Nucleotide sequence of one clone, designated NF-1/AT1, confirms this. The NF-1/AT1 protein was overexpressed in *E coli* and found to bind specifically to an NF-1 probe by gel shift analysis. Southern analysis of human fetal glial cells indicates that the NF-1/AT1 gene, class D, is derived from a different gene than NF-1/CTF1. These results suggest the possibility that genes or viruses, like JCV, which use NF-1 for their expression in human brain derived cells may preferentially use the NF-1 class D protein.

Keywords: human brain; nuclear factor-1; DNA binding; cloning; transcription

Introduction

Our studies have focused on the nuclear factor-1 (NF-1) activator, which is a multifunctional cellular protein that participates in cellular and viral gene transcription and viral DNA replication. NF-1 was first isolated from HeLa cells and shown to be a host cell factor required for the initiation of adenovirus-2 DNA replication (Nagata *et al*, 1982, 1983). It was shown to mediate this effect by binding to a specific site in the adenovirus origin of DNA replication (Nagata *et al*, 1983; Rawlins *et al*, 1984). Analysis of this site and other NF-1 binding sites revealed that the consensus recognition sequence for NF-1 is TGGC/A(N)$_5$GCCAA (DeVries *et al*, 1985; Leegwater *et al*, 1985). NF-1 binding sites have been identified within the promoters of several viral (Hennighausen *et al*. 1986; Shaul *et al*, 1986; Jeang *et al*, 1987; Tamura *et al*, 1988a; Amemiya *et al*, 1989; Gloss *et al*, 1989; Nilsson *et al*, 1989; Sundsfjord *et al*, 1990; Chong *et al*, 1991) and cellular genes (Hennighausen *et al*, 1985; Oikarinen *et al*, 1987; Nowock *et al*, 1985; Jones *et al*, 1987; Lichsteiner *et al*, 1987; Rupp and Sippel, 1987; Gil *et al*, 1988; Rossi *et al*, 1988; Aoyama *et al*, 1990; Courtois *et al*, 1990). In addition, NF-1 has been shown to activate transcription from the promoter of the human α globin (Jones *et al*, 1987), human hepatitis B virus S gene (Shaul *et al*, 1986), and murine α2(I) collagen gene (Rossi *et al*, 1988).

Structural and functional analysis of the NF-1 protein has revealed that the protein could be divided into two functional domains (Mermod *et*

Correspondence: EO Major
Received 19 September 1995; revised 31 January 1996; accepted 5 February 1996

al, 1989). The N-terminal domain is required for site specific recognition of DNA, protein dimerization, and adenovirus DNA replication. The proline rich C-terminal domain is required for transcriptional activation. Attempts to purify NF-1 binding activity from HeLa cells has resulted in a number of polypeptides ranging in size from 52 to 66 kDa (Rosenfeld and Kelly, 1986). The results from further studies also suggest the existence of related but distinct CTF/NF-1 proteins (Celada *et al*, 1988; Chodosh *et al*, 1988; Dorn *et al*, 1987; Raymondjean *et al*, 1988). In addition, the existence of a family of NF-1 proteins has been supported by the isolation of several NF-1 cDNA clones from different species and tissues. These cDNA clones have a highly conserved nucleotide sequence in their 5′ region, and a divergent nucleotide sequence in their 3′ region (Gil *et al*, 1988; Inoue *et al*, 1990; Meisterenst *et al*, 1988; Paonessa *et al*, 1988; Rupp *et al*, 1990; Santoro *et al*, 1988). It is suggested that the diversity of these NF-1 cDNA clones may be generated by alternative splicing events (Santoro *et al*, 1988; Inoue *et al*, 1990) or by the expression of multiple genes coding for different NF-1 species (Gil *et al*, 1988; Rupp *et al*, 1990).

Our interest in NF-1 derives from our studies of JC virus (JCV), a human polyomavirus, that is the etiological agent of the demyelinating disease, progressive multifocal leukoencephalopathy (see reviews Major *et al*, 1992; Tornatore *et al*, 1994). The propagation of JCV in the laboratory has been limited principally to human fetal glial (HFG) cells. There are four NF-1 binding sites in the promoter/enhancer region of JCV (Tamura *et al*, 1988a; Amemiya *et al*, 1989, 1992) which are important for JCV glial cell specific expression (Tamura *et al*, 1988a). The importance of these sites has been demonstrated in several studies. In one study, mutations in two of these NF-1 sites resulted in decreased activity of the JCV regulatory region in p19 embryonic carcinoma cells (Kumar *et al*, 1993). In another study, mutations in the NF-1 site in a JCV promoter/enhancer construction containing only one NF-1 binding site resulted in decreased JCV DNA replication (Sock *et al*, 1991). In a third study, Caruso *et al* (1990) described a naturally occurring polyoma virus mutant which gained the ability to grow in neuroblastoma cells by acquiring a NF-1 binding site in its enhancer. Other studies have suggested that proteins from the nervous system which bind to these NF-1 sites may be different from the proteins which bind from other tissues. Examination of protein complexes which bind to the NF-1 sites in the JCV enhancer by electrophoretic mobility shift assays indicates that there are differences in the sizes of the protein(s) binding to the NF-1 site from human glial cells as compared to other cell types (Tamura *et al*, 1988a; Amemiya *et al*, 1989). Similarly, the analysis of protein complexes from brain and non-brain extracts binding to the NF-1 site in the myelin basic protein enhancer, another neural specific gene, indicates that there may be different proteins involved (Miura *et al*, 1989; Tamura *et al*, 1988b; Aoyama *et al*, 1990). Amemiya *et al* (1992) have noted that there are a number of genes expressed in the central (CNS) and peripheral nervous systems (PNS) which contain NF-1 binding sites within their promoter-enhancer regions. These studies suggest that there may be a unique species of NF-1 in the CNS and PNS which is responsible for cell specific gene expression.

In this report, we show, using class specific NF-1 PCR primers, that the expression of all four classes of the NF-1 family can be detected in both JCV permissive human fetal glial cells as well as non-permissive HeLa cells. However, when quantitated by Northern analysis, the expression of the class D gene is up regulated in primary HFG and down regulated in HeLa cells. We also describe the isolation of NF-1 cDNA clones from two early human brain cDNA libraries. The majority of the NF-1 cDNA clones are closely homologous to each other. Sequence analysis of the early human brain cDNA clones reveals that they belong to NF-1 class D, which is highly expressed in glial cells. Because of the differences between the 3′-end of the NF-1/CTF1 and NF-1/AT1 cDNAs, we suggest that they may interact differently with the transcription initiating complex. Our results suggest that the NF-1 activator used by genes or viruses expressed in early human developing brain cells belongs primarily to NF-1 class D.

Results

Classification of NF-1 family members

Rupp *et al* (1990) and Kruse *et al* (1991) have previously suggested that the NF-1-like nuclear proteins, which bind to a specific nucleotide sequence, belong to one of four groups or classes. Table 1 shows a listing of the members of each class of NF-1 cDNA clones and includes cDNA clones from recent studies (Inoue *et al*, 1990; Jung *et al*, 1993) and from this study. Not all of the cDNA clones of a single type have been included. In cases where multiple spliced species of a cDNA clone type have been isolated, only the least spliced species is listed (Santoro *et al*, 1988). In other cases, such as in this study, where it is likely the cDNA clones are related, the longest clone containing the putative coding region for the NF-1 protein is listed. Within each class of the NF-1 protein family, there are homologies of 85–93% at the nucleotide level and 95–97% at the amino acid level. The homologies between the classes are 44% and 50% on the nucleotide and amino acid levels, respectively. Within the first 180 amino acids of the members of the NF-1 family, however, the homology is 93%, reflecting the high conservation in the DNA binding domain. Figure 1 shows the genetic map of the NF-

Table 1 Classification of nuclear factor 1 family members

Class	*Member (species/tissue)*	*Reference*
A	CNF1-A1 (chicken/myeloid)	Rupp *et al.* (1990) *Nuclei Acids Res.* **18**, 2607–2616
	NF-1-L (rat/liver)	Paonessa *et al.* (1988). *EMBO J.* **7**, 3115–3123
	NF-1-L2 (rat/liver)	Jung *et al.* (1993) *Korean Biochem. J.* **26**, 312–316
	NF-1-B1 (mouse/cerebellum)	Inoue *et al.* (1990) J Biol Chem. **265**, 19065–19070
B	CNF1-B1 (chicken/liver)	Rupp *et al.* (1990) *Nucleic Acids Res.* **18**, 2607–2616
	CNF1-B2 (chicken/liver)	Rupp *et al.* (1990) *Nucleic Acids Res.* **18**, 2607–2616
	NF1Red1 (hamster/liver)	Gil *et al.* (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8963–8967
C	CNF1-C1 (chicken/myeloid)	Rupp *et al.* (1990) *Nucleic Acids Res.* **18**, 2607–2616
	pNF-1/CTF1 (pig/liver)	Meisterernst *et al.* (1988) *FEBS Lett.* **236**, 27–32 Meisterernst *et al.* (1989) *Biochemistry* **28**, 8191–8200
	NF-1/CTF1 (human/Hela)	Santoro *et al.* (1988) *Nature* **334**, 218–224
D	CNF1/x (chicken/promacrophage)	Kruse *et al.* (1991) *Nucleic Acids Res.* **19**, 6641
	NF-1/x (hamster/liver)	Gil *et al.* (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8963–8967
	NF-1AT1 (human/brain)	This study

1/CTF1 gene demonstrating the functional regions of all NF-1 family members: the amino terminal DNA binding domain and the carboxy terminal transcription activator region. For comparison, cDNA clones isolated in this study from human brain libraries are shown in Figure 1.

Four classes of NF-1 genes are expressed in primary human fetal brain cells

In order to determine the expression of the different classes of the NF-1 family in primary HFG cells, cytoplasmic RNA was isolated from 15–16 week old primary HFG cells and poly(A)$^+$ mRNA was selected. PCR primers from a representative member of each class were selected from their respective 3′ regions. These primers were chosen to be homologous to the most conserved sequences within a class as well as the most divergent sequences between classes. Table 2 lists the representative member of each class of the NF-1 family chosen with the respective PCR primer sets used in this study. In addition, it lists the oligonucleotides used to verify the PCR product by Southern hybridization analysis. After the RT-PCR reaction with the class specific primers, the reaction products were analyzed by polyacrylamide gel electrophoresis. Figure 2A, panel A, shows the specific RT-PCR products that were obtained with mRNA extracted from primary HFG cells. The anticipated size of the RT-PCR product was obtained for each of the four classes A-D (463, 325. 421 and 301 bp, respectively) of the NF-1 family. Two control assays were performed to check for possible contamination of the poly(A)$^+$ mRNA with genomic DNA. A control primer set for human G3PDH mRNA was used and gave a product of expected size of 983 bp for the mature spliced mRNA (data not shown). In addition, PCR analysis of HFG cell genomic DNA with the class specific primer sets did not result in similar size products seen in the RT-PCR studies (data not shown). Southern hybridization analysis with class specific probes showed that the RT-PCR products were specific for their respective NF-1 class. Each blot containing all four classes of RT-PCR products was hybridized with each class specific probe (Figure 2A, panels b–e). In addition to analyzing primary HFG cells, HeLa cells were examined for the expression of the different classes of the NF-1. RT-PCR was performed with poly(A)$^+$ mRNA from HeLa cells with the four representative NF-1 primer sets as above. We obtained four specific RT-PCR products of the expected size with each of the class specific primer sets (Figure 2B, panel a). The specificity of each RT-PCR product was again verified by Southern hybridization analysis with the specific Southern blot probe (Figure 2B, panels b–e). In some cases, more than one RT-PCR product was obtained with the mRNA from either cell type (Figure 2A, panel c; Figure 2B, panels b and c). These results most likely indicate the presence of alternative spliced species and not the presence of another class, as they were identified by class specific Southern probes. The results of these studies suggest that at least four classes of the NF-1 family are expressed in these two diverse cell types. In addition, there appears to be some evidence that alternative spliced species of these classes are present.

Isolation of early human brain NF-1 cDNA clones

In order to characterize the type of NF-1 proteins present in early developing human brain cells, we took advantage of the homology in the DNA binding domain of NF-1 cDNAs (Gil *et al*, 1988; Rupp *et al*, 1990). A [^{32}P]-labeled 57 base oligonucleotide homologous to the DNA binding domain of NF-1/CTF1 (nuc 410–466 Santoro *et al*, 1988), a class C member, was used to screen two early human brain

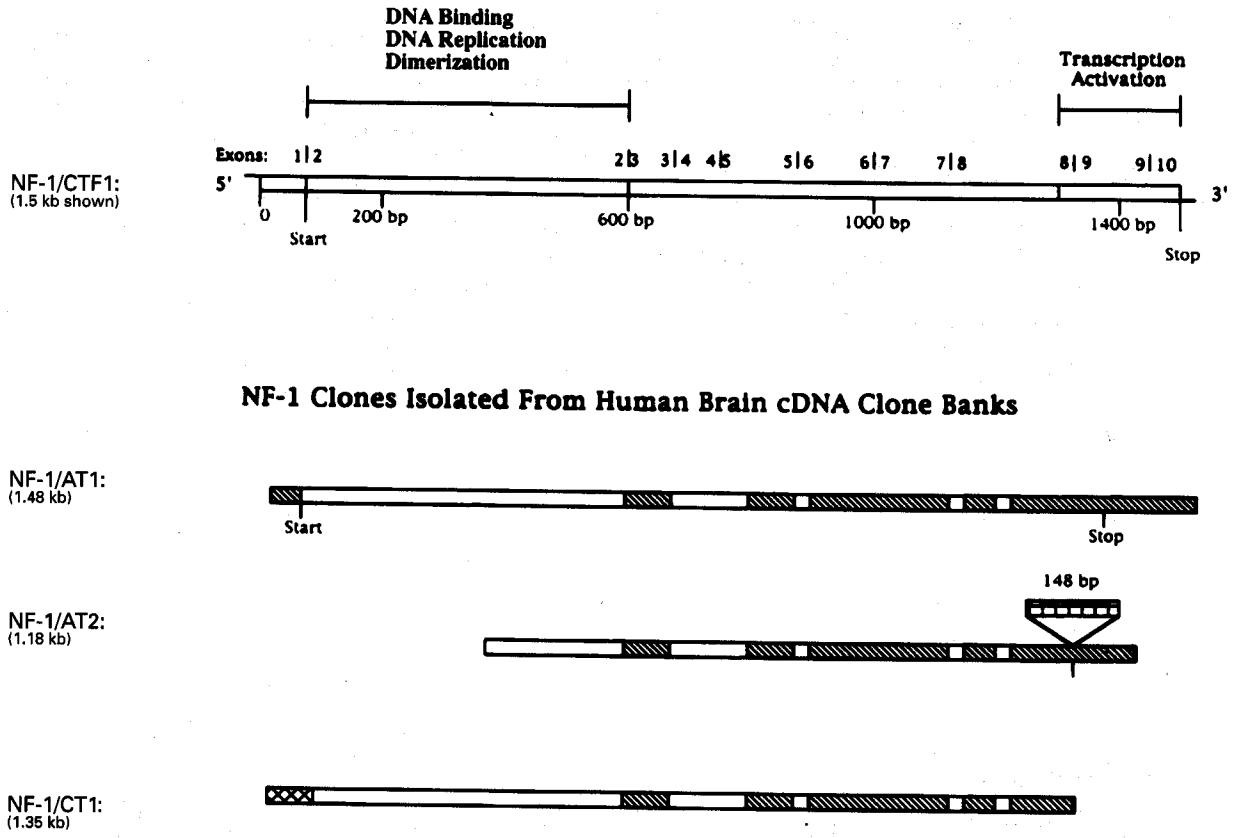

Figure 1 Graphic description of NF-1 functional domains and brain derived cDNA clones. The top portion shows structural and functional features of NF-1CTF1 which belongs to class C of the NF-1 family. The lines above the open box denote the regions involved in specific functional activities (Mermod *et al*, 1989). The numbers just above the open box denote the exon junctions for members of this class (Meisterernst *et al*, 1989). Only 1.5 kb of NF-1/CTF1 is shown. 'Start' and 'Stop' in figure denote beginning and end of the putative open reading frame. The bottom portion of the figure shows three cDNA clones isolated from early human brain cDNA libraries with their respective lengths. They are completely homologous with each other except for the very 5'-end of NF-1/AT1 and NF-1/CT1 and the 148 bp nonhomologous insertion in the 3'-end of NF-1/AT2. The open clear regions represent regions of homology with NF-1/CTF1 (88–93%), and the hatched regions denote regions of nonhomology.

Table 2 Oligonucleotides used to identify classes of nuclear factor-1

Class	*Representative member*	*RT-PCR Primer set*	*Southern blot probe*
A	NF-1-L[a]	5′ primer: nt 586–610 3′ primer: nt 1049–1025	nt 850–900
B	NF-1Red1[b]	5′ primer: nt 732–754 3′ primer: nt 1057–1036	nt 814–865
C	NF-1/CTF1[c]	5′ primer: nt 586–605 3′ primer: nt 1007–986	nt 900–950
D	NF-1AT1[d]	5′ primer: nt 599–618 3′ primer: nt 900-878	nt 700–725

[a]Paonessa *et al.* (1988). [b]Gil *et al.* (1988). [c]Santoro *et al.* (1988). [d]This study

cDNA libraries. Approximately 5×10^5 cDNAs were screened and two similar sized (1.35 kb) cDNA clones were isolated from the first cDNA library. One of these cDNA clones (Figure 1, NF-1/CT1) was further subcloned and the nucleotide sequence determined. A second early human brain cDNA library was screened (5×10^5 PFU) with the same [^{32}P]-labeled probe and two cDNA clones were isolated and their sequence determined (Figure 1, NF-1/AT1 and NF-1/AT2). The nucleotide sequences of these later two cDNAs were similar to each other, except for a 148 bp sequence present in the smaller of the two cDNA clones. In addition, the nucleotide sequence of both of these cDNA clones was similar to NF-1/CT1, except for a 100 bp sequence at the 5′ end. Rescreening the neonatal

brain cDNA library with the full length NF-1/AT1 cDNA clone resulted in the isolation of five additional clones ranging in size from 0.9 to 1.4 kb. Partial sequencing of these cDNA clones revealed that they were similar to NF-1/AT1. The human fetal brain cDNA library (approximately 5×10^5 PFU) was also screened with the 1.8 kb insert of pCTF1. One positive cDNA clone was isolated (0.9 kb). However, only a small region (150 bp) of this clone was homologous with NF-1/CTF1. The region of homology was located near the 3′ end of NF-1/CTF1, and it appears to coincide with the sequence belonging to the putative exon 7 (Figure 1, top line).

A comparison of the human brain cDNA clones including the nucleotide sequence of one clone, NF-1/AT1 (Figure 3), with the previous isolated human NF-1/CTF1 cDNA clone (Santoro *et al*, 1988; Mermod *et al*, 1989) shows that the brain NF-1 cDNA clones can be divided into two regions (Figure 1). One region was located at the 5′ end of the cDNA clones. This region has high nucleotide sequence homology (>90%) with NF-1/CTF1 and was demonstrated by Mermod *et al* (1989) to contain the domain involved in DNA binding, DNA replication, and protein dimerization. Interestingly, this region appears to be located in a single exon. This is based on the exon structure proposed

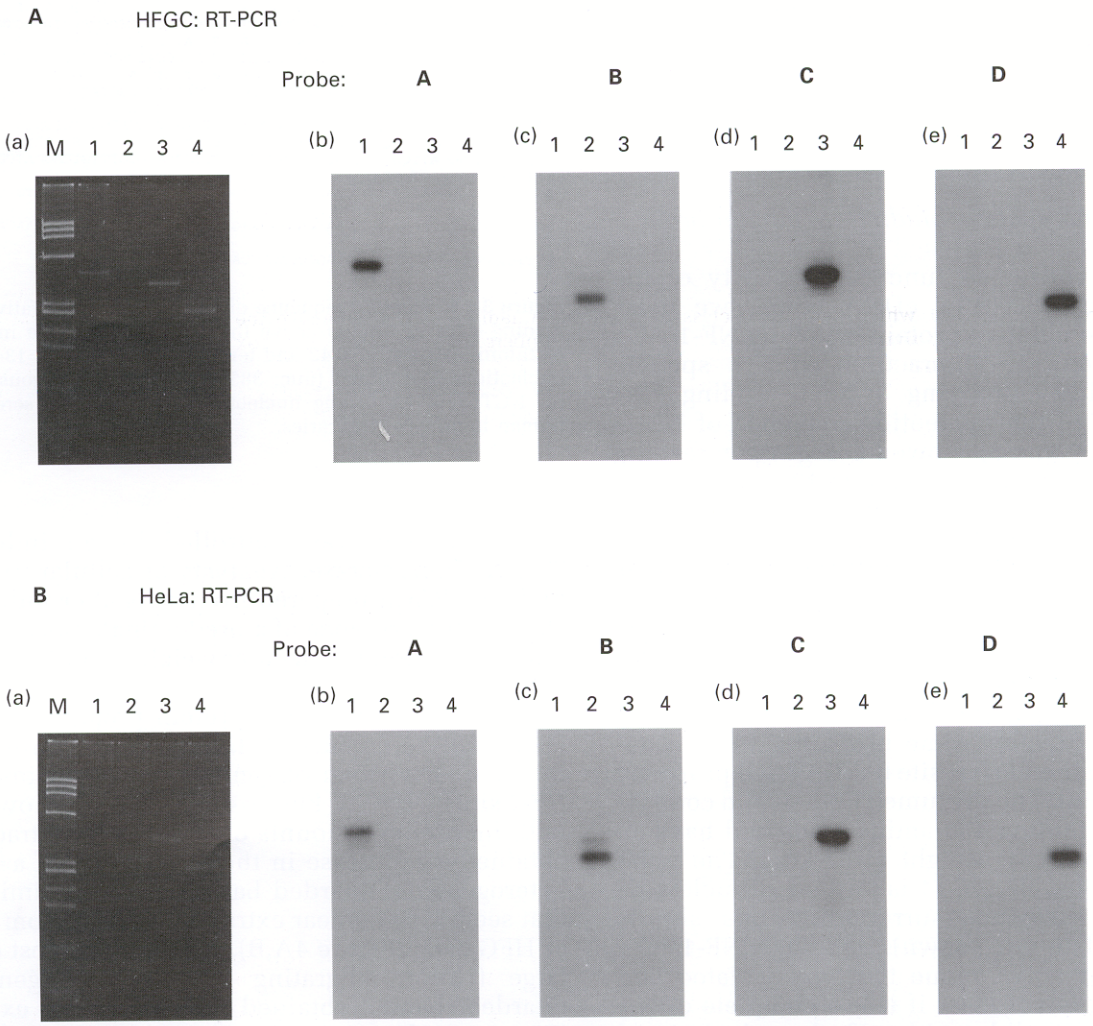

Figure 2 Expression of four classes of the NF-1 activator family in primary HFG and HeLa cells. Poly(A)$^+$ mRNA was isolated from primary HFG (A) or HeLa cells (B) and reversed transcribed (RT). PCR analysis was performed using the RT product with specific NF-1 class primer sets as listed in Table 2. Panel a of each cell type shows the PCR product produced with each class specific primer set (lanes 1–4, class A, B, C and D, respectively). Panels b–c show Southern analysis of RT-PCR products probed with class specific (see Table 2) Southern probes (lanes 1–4, A, B, C and D probes, respectively). A positive control RT-PCR reaction for human G3PDH cDNA was performed with each analysis (not shown). M represents size markers.

by Meisterernst *et al* (1989). In the 3′ end of the cDNA clones, the DNA homology diverges between the brain cDNA clones and the cDNA clone isolated from HeLa cells. However the sequence is very similar to the previously isolated class D cDNA clones (see Table 1). These brain derived cDNA clones can, therefore, be classified as class D NF-1. It is this 3′ region which was reported to contain the transcriptional activation domain of NF-1/CTF1 (Mermod *et al*, 1989) and thus may confer differential functions between NF-1 classes. As with other NF-1 cDNA clones previously isolated (Santoro *et al*, 1988; Meisterernst *et al*, 1989; Inoue *et al*, 1990; Rupp *et al*, 1990), there appears to be evidence for alternative splicing events to explain the difference between the brain cDNA clones we isolated. There is a nonhomologous 148 bp sequence present in the 3′ region of NF-1/AT2 which is not present in NF-1/AT1. This may represent an extra exon. There is also a region of nonhomology at the 5′ end of NF-1/AT1 and NF-1/CT1. This region just precedes the conserved region containing the putative DNA binding domain and this border has been proposed to be a splice junction between exons 1 and 2 by Meisterernst *et al* (1989) for class C NF-1 activators.

Expression and specific binding by human brain NF-1/AT1

In order to verify the functional activity of the human brain NF-1/AT1 cDNA clone, we overexpressed the putative protein coded by NF-1/AT1 and examined the interaction with a specific oligonucleotide containing a NF-1 binding site. Examination of the nucleotide sequence of NF-1/AT1 (Figures 1 and 3) showed a large open reading frame (ORF) starting at nucleotide 42 and ending at nucleotide 1340. This start site is similar to the putative translational start site for NF-1/CTF1 (Santoro *et al*, 1988). Furthermore, this site is located at the beginning of the sequence coding for highly conserved DNA binding domain and at the junction of the proposed exon 1 and 2 splice site for NF-1/CTF1 (Meisterernst *et al*, 1989). In addition, the putative translational start site for NF-1/AT1 (TCAATGG) contains some homology to the consensus translational start site (ACCATGG) proposed by Kozak (1986). The presumed ORF would code for a protein containing 433 amino acids and have an estimated molecular weight of 52 kdal. The putative start site for NF-1/CT1 is located at nucleotide 87 and the sequence surrounding this site is CGCATGC (data not shown). Although NF-1/AT1 is the longest cDNA clone that was isolated and sequenced, it is not clear if it is a complete cDNA clone because we did not identify the poly A signal (AATAAAA) in the 3′ region of NF-1/AT1.

In order to examine the protein coded by NF-1/AT1, the NF-1/AT1 ORF was subcloned into the expression vector pET15b, in which the expression

```
                                                  M  D  E  F  H  P  F
CTCAAAATCGGGCGTTGGGCTTTGCCGGGTGCTTCAGATCAATGGATGAGTTCCACCCGT   60
  I  E  A  L  L  P  H  V  R  A  F  S  Y  T  W  F  N  L  Q  A
TCATCGAGGCACTGCTGCCTCACGTCCGCGCTTTCTCCTACACCTGGTTCAACCTGCAGG  120
  R  K  R  K  Y  F  K  K  H  E  K  R  M  S  K  D  E  E  R  A
CGCGGAAGCGCAAGTACTTCAAGAAGCATGAAAAGCGGATGTCGAAGGACGAGGAGCGGG  180
  V  K  D  E  L  L  G  E  K  P  E  I  K  Q  K  W  A  S  R  L
CGGTGAAGGACGAGCTGCTGGGCGAGAAGCCCGAGATCAAGCAGAAGTGGGCATCCCGGC  240
  L  A  K  L  R  K  D  I  R  P  E  F  R  E  D  F  V  L  T  I
TGCTGGCCAAGCTGCGCAAGGACATCCGGCCCGAGTTCCGCGAGGACTTCGTGCTGACCA  300
  T  G  K  K  P  P  C  C  V  L  S  N  P  D  Q  K  G  K  I  R
TCACGGGCAAGAAGCCCCCCTGCTGCGTGCTCTCCAACCCCGACCAGAAGGGCAAGATCC  360
  R  I  D  C  L  R  Q  A  D  K  V  W  R  L  D  L  V  M  V  I
GGCGGATTGACTGCCTGCGCCAGGCTGACAAGGTGTGGCGGCTGGACCTGGTCATGGTGA  420
  L  F  K  G  I  P  L  E  S  T  D  G  E  R  L  Y  K  S  P  Q
TTTTGTTTAAGGGGATCCCCCTGGAAAGTACTGATGGGGAGCGGCTCTACAAGTCGCCTC  480
  C  S  N  P  G  L  C  V  Q  P  H  H  I  G  V  T  I  K  E  L
AGTGCTCGAACCCCGGCCTGTGCGTCCAGCCACATCACATTGGAGTCACAATCAAAGAAC  540
  D  L  Y  L  A  Y  F  V  H  T  P  E  S  G  Q  S  D  S  S  N
TGGATCTTTATCTGGCTTACTTTGTCCACACTCCGGAATCCGGACAATCAGATAGTTCAA  600
  Q  Q  G  D  A  D  I  K  P  L  P  N  G  H  L  S  F  Q  D  C
ACCAGCAAGGAGATGCGGACATCAAACCACTGCCCAACGGGCACTTAAGTTTCCAGGACT  660
  F  V  T  S  G  V  W  N  V  T  E  L  V  R  V  S  Q  T  P  V
GTTTTGTGACTTCCGGGGTCTGGAATGTGACGGAGCTGGTGAGAGTATCACAGACTCCTG  720
  A  T  A  S  G  P  N  F  S  L  A  D  L  E  S  P  S  Y  Y  N
TTGCAACAGCATCAGGGCCCAACTTCTCCCTGGCGGACCTGGAGAGTCCCAGCTACTACA  780
  I  N  Q  V  T  L  G  R  R  S  I  T  S  P  P  S  T  S  T  T
ACATCAACCAGGTGACCCTGGGGCGGCGGTCCATCACCTCCCCTCCTTCCACCAGCACCA  840
  K  R  P  K  S  I  D  D  S  E  M  E  S  P  V  D  D  V  F  Y
CCAAGCGCCCCAAGTCCATCGATGACAGTGAGATGGAGAGCCCTGTTGATGACGTGTTCT  900
  P  G  T  G  R  S  P  A  A  G  S  S  Q  S  S  G  W  P  N  D
ATCCCGGGACAGGCCGTTCCCCAGCAGCTGGCAGCAGCCAGTCCAGCGGGTGGCCCAACG  960
  V  D  A  G  P  A  S  L  K  K  S  G  K  L  D  F  C  S  A  L
ATGTGGATGCAGGCCCGGCTTCTCTAAAGAAGTCAGGAAAGCTGGACTTCTGCAGTGCCC 1020
  S  S  Q  G  S  S  P  R  M  A  F  T  H  H  P  L  P  V  L  A
TCTCCTCTCAGGGCAGCTCCCCGCGCATGGCTTTCACCCACCACCCGCTGCCTGTGCTTG 1080
  G  V  R  P  G  S  P  R  A  T  A  S  A  L  H  F  P  S  T  S
CTGGAGTCAGACCAGGGAGCCCCCGGGCCACAGCATCAGCCCTGCACTTCCCCTCCACGT 1140
  I  I  Q  Q  S  S  P  Y  F  T  H  P  T  I  R  Y  H  H  H  H
CCATCATCCAGCAGTCGAGCCCGTATTTCACGCACCCGACCATCCGCTACCACCACCACC 1200
  G  Q  D  S  L  K  E  F  V  Q  F  V  C  S  D  G  S  G  Q  A
ACGGGCAGGACTCACTGAAGGAGTTTGTGCAGTTTGTGTGCTCGGATGGCTCGGGCCAGG 1260
  T  G  Q  H  S  Q  R  Q  A  P  P  L  P  T  G  L  S  A  S  D
CCACCGGACAGCATTCGCAACGACAGGCGCCTCCTCTGCCAACCGGTTTGTCAGCATCGG 1320
  P  G  T  A  T  F
ACCCCGGGACGGCAACTTTCTGAACATCCCACAGCAGTCTCAGTCCTGGTTCCTCTGATA 1380

AGATCGACAAAAGAAACAACAAAATGAGAAGAAGAGGTTCCTCGAAAGGGGGGAGAAGAA 1440

TTTTGAGAATGGAAAAATCCCCCAGCCCAGCCCAGCCCCACCGAAA 1486
```

Figure 3 Nucleotide sequence of NF-1/AT1. The putative open reading frame is shown above the sequence with the initiating methionine at position 42 and termination condon at 1341. The nucleotides underlined (nuc. 383–440) are homologous to the NF-1/CTF1 57 base long nucleotide probe used to screen the human brain cDNA libraries.

of the protein was controlled by an inducible T7 RNA polymerase. Analysis of cellular extracts by SDS–PAGE prepared from cells after induction showed the presence of a predominate polypeptide with an apparent molecular weight of 55 kDa in the pellet fraction.

After activation of the protein preparation from cells containing pETNF-1/AT1, we examined its interaction with an oligonucleotide containing an NF-1 site by EMSA. Figure 4B lanes 5–9, shows that with increasing amounts of the protein extract, we obtained an increase in the appearance of a broad heterogeneous retarded band. This was similar to that seen with nuclear extracts prepared from HeLa or HFG cells (Figure 4A,B). In addition, most of the large (slower migrating bands), heterogeneous, retarded bands obtained with nuclear extracts (Figure 4A, lanes 3 and 6; Figure 4B, lane 3), and with the overexpressed protein preparation (Figure 4B, lane 10) were competed by a specific NF-1 oligonucleotide. At the same time, most of the retarded bands were not as affected by an oligonu-

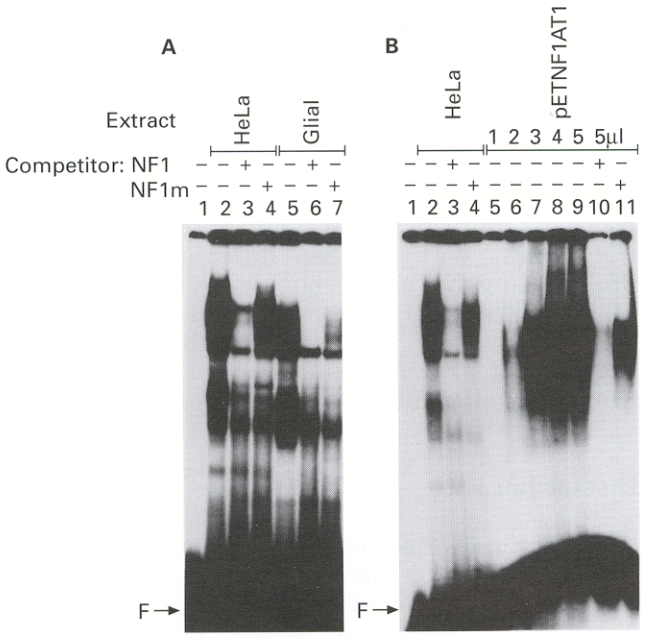

Figure 4 Specific binding by nuclear proteins from HeLa, primary HFG cells, and recombinant NF-1/AT1 analyzed by EMSA. A. Binding by 10 μg of extract from HeLa (lanes 2–4) or primary HFG (lanes 5–7) cells was examined with no competitor (lanes 2 and 5), 50 × homologous competitor (lanes 3 and 6), or 50 × mutant competitor (lanes 4 and 7). B. Binding by 10 μg of HeLa nuclear extract (lanes 2–4) or increasing amounts of recombinant NF-1/AT1 (0.1 μg, 1.0 μg, 5 μg, 10 μg, 20 μg in lanes 5–10 and 10 μg in lane 11) was examined. Either 50 × of homologous NF-1 competitor (lanes 3 and 10) or 50 × of mutant NF-1 competitor (lanes 4 and 11) was used. Lane 1 in both A and B represents labeled probe alone. F represents free probe. Anomalous migration of the free probe in B (lanes 6–11) is the result of residual guanidine-HCl in the sample.

cleotide competitor containing a mutated NF-1 bindings site (Figure 4A, lanes 4, 7; Figure 4B, lanes 4, 11). The heterogeneity in the retarded bands obtained with nuclear extracts is most likely the result of the presence of different classes of NF-1 proteins and proteolysis of some of these proteins. On the other hand, the heterogeneity seen in the retarded bands with the *E coli* protein extracts is most likely the result of proteolysis of the over-expressed protein. Similar results were previously seen with the overexpressed protein preparation from pETCTF1 transformed *E coli* (Amemiya *et al*, 1994). We were, therefore, able to demonstrate the expression of a specific protein with an apparent molecular weight of 65 kDa, which was close to the estimated size (52 kDa) based on the ORF of NF-1/AT1. We were also able to demonstrate specific binding to an NF-1 nucleotide sequence with an extract prepared from cells transformed with NF-1/AT1 cDNA. The specific binding results and the primary structure of NF-1/AT1 suggest that NF-1/AT1 is a member of the NF-1 family.

Differential expression of NF-1/AT1 and NF-1/CTF1 in human fetal brain cells

To further examine the expression of NF-1/AT1 in primary HFG and HeLa cells, Northern hybridization analysis was performed. Total RNA was extracted from both primary HFG and HeLa cells and Northern hybridization analysis performed with [^{32}P]-labeled restriction fragments from the 3′ variable region of the cDNA clones of NF-1/AT1 and NF-1/CTF1. Figure 5A shows a large difference in the amount of RNA from primary HFG and HeLa cells hybridizing to the NF-1/AT1 or class D labeled probe. At least two related species of NF-1/AT1 RNA can be seen in the primary HFG cell RNA, a minor 7.8 kb species and a major 6.6 kb species. In the HeLa cell RNA, on the other hand, only a 7.8 kb species can be seen. There was a greater than 20-fold difference in the amount of NF-1/AT1 specific RNA seen in the primary HFG cells when compared to that seen in HeLa cells. The amount of β-actin mRNA served as an internal control in these studies (see botton of Figure 5A, B). A much different result was obtained with the NF-1/CTF1 or NF-1 class C3′ labeled probe. Two similar RNA species can be seen in both cell types, a more abundant 7.5 kb species and a lesser 5.8 kb species. In this later case, however, there was approximately a 2-fold greater amount of NF-1/CTF1 specific RNA in HeLa cells when com-

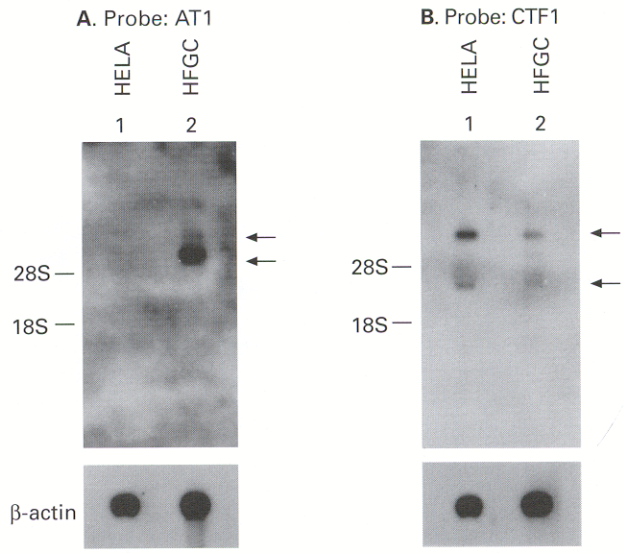

Figure 5 Northern analysis of NF-1/AT1 (class D) and NF-1/CTF1 (class C) mRNA expression in primary HFG and HeLa cells. Total RNA was extracted from HeLa and primary HFG (8 weeks) cells and separated (15 μg) on a 1% agarose-formaldehyde gel (lanes 1 and 2, respectively). Labeled 3′ probes from NF-1/AT1 (A) and NF-1/CTF1 (B) were prepared and hybridized to each RNA set. On the left of each panel shows the rRNA markers, and the arrows on the right show the mRNA species. Below each panel shows the β-actin mRNA for each sample.

pared to that in primary HFG cells (Figure 5B). More interestingly, when the expression of NF-1/AT1 was compared with that of NF-1/CTF1 in primary HFG cells, NF-1/AT1 was expressed approximately 5-fold higher than NF-1/CTF1 in these cells. In conclusion, the results of the Northern hybridization analysis suggest that the gene for NF-1/AT1 is much more highly expressed (20-fold) in primary HFG cells than that in HeLa cells. Although the gene for NF-1/CTF1 is expressed in both primary HFG and HeLa cells, it is expressed in double the concentration in HeLa cells than in primary HFG cells. The NF-1/AT1 gene is more highly expressed than the NF-1/CTF1 gene in primary HFG cells.

NF-1/AT1 and NF-1/CTF1 are derived from different genes

There are two explanations for the generation of the diverse classes of the NF-1 family of proteins. One model would predict that all classes of NF-1 proteins are derived from one gene and that the diversity of NF-1 species results solely from alternative splicing events. The second model predicts that the different NF-1 classes are expressed from different genes whose products may also undergo alternative splicing events. In order to determine if one or more genes codes for NF-1 proteins, primary HFG cellular DNA was digested with three different restriction enzymes and Southern blot hybridization analysis performed (Figure 6, panels A–C, *AvaI, BglI* or *SacI*, respectively). The Southern blots were probed with [^{32}P]-labeled fragments obtained from the 3′ region of the cDNA clone of NF-1/CTF1 (1023 bp *SalI/Eco*RI) or NF-1/AT1 (842 bp *Afl*III/*Eco*RI). The Southern hybridization results with the two different NF-1 cDNA probes show that each probe gave a unique pattern of hybridization (Figure 6). These results taken together with the Northern hybridization results, suggest that NF-1/AT1 and NF-1/CTF1 are derived from different genes.

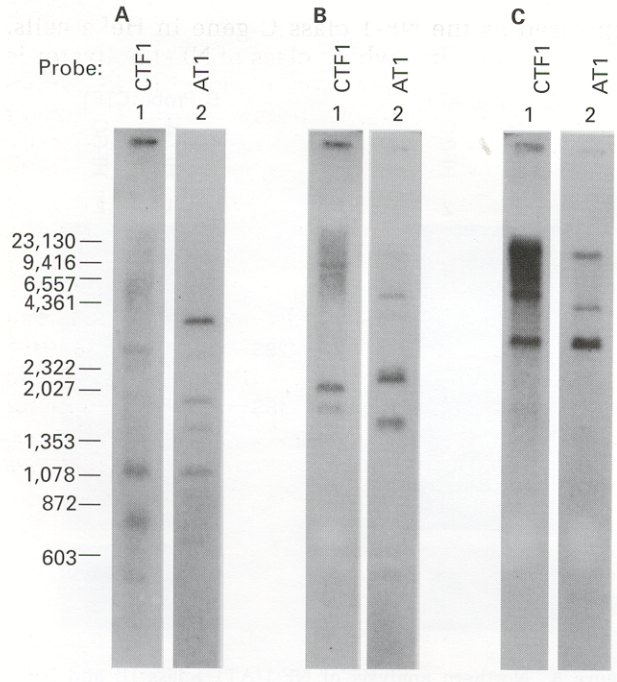

Figure 6 Southern analysis of NF-1/CTF1 (class C) and NF-1/AT1 (class D) genes. DNA was extracted from primary HFG cells and digested with either *AvaI* (A), *BglI* (B), or *SacI* (C). After separation and transferring the digested DNA to a filter, each filter set was probed with a 3′-labeled fragment from either NF-1/CTF1 (lane 1) or NF-1/AT1 (lane 2). The numbers on the left show the position of DNA markers.

Discussion

It was first suggested by Gil *et al* (1988) and later by Rupp *et al* (1990) that there are at least three or possibly four genes which are responsible for the diversity of the NF-1 family of proteins in mammalian or vertebrates species. Table 1 lists the NF-1-like cDNA clones which have been reported to date. Although they are derived from a wide variety of species and sources, one common property shared by all NF-1 members is their high homology (at least 93% at the amino acid level) in the N-terminal region. In contrast, the sequence in the 3′ region is heterogeneous in both primary structure and length. This diversity in the 3′ region is the basis for classification of the NF-1 family into four classes. As this is the region presumed to be responsible for transcription activation, the different classes may have significantly different functions in gene expression.

There are a number of examples of activators which could mediate the interaction of NF-1 with the formation of the transcription complex. In the promoter of the herpes simplex virus thymidine kinase (HSV-TK) gene, there is a SP1 binding site just upstream from the TATA box (Jones *et al*, 1985). Adjacent to this SP1 binding site is a NF-1 binding site followed by another SP1 binding site. The SP1 protein, which has been characterized as a guanine-rich activator, has been recently shown to interact with the Drosophila TBP-associated factor $dTAF_{II}110$ (Hoey *et al*, 1993). The interaction takes place through two domains within the guanine-rich region of SP1, and this interaction was important for mediating transcriptional activation (Gill *et al*, 1994). Another example is in the mouse mammary tumor virus long terminal repeat (MMTV-LTR), where an octamer protein (Oct1) can bind between the NF-1 and TFIID binding sites (Lee and Archer, 1994). In this case it appears that binding of NF-1 and Oct1 proteins is dependent on the accessibility of their respective binding sites. In a transient template binding of NF-1 and Oct1 can occur independent of hormone. In a stable inserted

template the binding by the two activators is hormone dependent. Access to the binding sites in these two different templates appears to be governed by the presence of nucleosome structures in the stable template. Binding of TBP on either template, on the other hand, is hormone dependent. Oct1 is an ubiquitously expressed POU protein, which can recognize a specific octamer motif and shares with the lymphoid specific Oct2 (Clec *et al*, 1988) the ability to bind to the same octamer motif (Tanaka and Herr, 1990). How Oct1 might interact with the transcription complex is not known. This type of interaction might be involved in the expression of many genes which have NF-1 binding sites or putative NF-1 binding sites much further upstream from the TBP binding site or have multiple activator binding sites between the NF-1 and the TBP binding sites. Some examples of this type of promoter/enhancer are the human histone H2B gene (LaBella *et al*, 1988), the human growth hormone gene (Courtois *et al*, 1990), the human β-globin gene (deBoer *et al*, 1988), the adenovirus 2 early promoter (Rosenfeld *et al*, 1987), the human papillomavirus type 16 promoter (Apt *et al*, 1993), and the cytomegalovirus major immediate early gene (Hennighausen and Fleckenstein, 1986). Like the MMTV promoter, the expression of these genes/viruses may be affected by the chromatin structure in the promoter/enhancer region and the possible presence of histones (Wolfe *et al* 1994). The combinatorial binding of multiple activators may only be effective in the context of a chromatin structure which may bring these activators close to the components involved in the formation of the transcription initiation complex.

Another potential NF-1 activator interaction can be found in some genes or viruses which are expressed in the CNS or PNS. We had previously noted that a number of genes (myelin basic protein, glial fibrillary acidic protein, proteolipid protein, S100β, mouse neurofilament-L and human proenkephalin) expressed in the CNS or PNS contained adjacent or overlapping NF-1 and activator protein binding sites just upstream from the TBP binding site (Amemiya *et al*, 1992). Because of the commonality of these binding sites upstream from the mRNA initiation site in these genes expressed in the CNS/PNS, we have referred to these binding sites as the 'glial-neurobox'. The activator associated with the 'glial-neurobox' in addition to NF-1 appears to be related to c-Jun or activator protein-2 (AP2). Both of these proteins have a proline rich domain at their N-terminus, which has been demonstrated to be responsible for their transcriptional activity (Bohmann and Tjian, 1989; Williams and Tjian, 1991). Although c-Jun was noted to have a number of Ser-Pro-X-X and Thr-Pro-X-X sequences (Suzuki, 1989), neither of these two proteins appear to have a consensus heptapeptide repeat unit in their proline rich domain. Interestingly, Mitchell *et al* (1991) have found that AP2 is expressed in neural crest cell lineage's during mouse embryogenesis. Neural crest cells are precursors to cells involved in the development of the PNS, face, limbs, skin and nephritic tissues. We would suggest that because of the proximity of the activator protein binding site to the TBP binding site in these genes, there may be an interaction of these proteins with TBP or another component of the transcription initiation complex. Furthermore, we have some indication that there may also be an interaction between NF-1 and c-Jun, which we have seen in binding assays with the JCV promoter-enhancer region (Amemiya *et al*, 1992).

Based on our findings, we suggest that in early developing human brain cells, the class of NF-1 activator involved in the expression of genes which require this activator belongs to NF-1 class D. During the screening of two early human brain cDNA libraries, the primary type of NF-1 cDNA isolated belonged to class D. This suggests that the other classes of the NF-1 family, although present, are not as highly expressed as the NF-1 class D gene in the early human brain cells. Support for this conclusion comes from the results of the Northern analysis, which suggests that the NF-1 class D gene is highly expressed in human fetal glial cells in comparison to the NF-1 class C gene. On the other hand, the NF-1 class D gene is not as highly expressed as the NF-1 class C gene in HeLa cells. In order to confirm which class of NF-1 activator is involved in the expression of specific CNS/PNS genes in early developing human brain cells, class specific antibodies need to be used to help identify the NF-1 class involved. Antibodies are currently being made in rabbits to unique peptides for each class of NF-1 proteins. Use of the class specific antibodies in gel shift assays will help identify which class member(s) binds the JCV sequences. These antibodies potentially could be used in competitive binding assays to determine relative levels of binding affinities. Furthermore, in order to determine the importance of an NF-1 protein for expression of JCV in brain or other tissues, additional experiments involving functional assays need to be done. Such studies are in progress.

Methods

Cell culture

Cultures of primary HFG cells (gestational ages 8–16 weeks) were prepared as previously described (Tornatore *et al*, 1991). Both primary HFG and HeLa cells were maintained in Eagle's minimum essential medium containing 10% fetal bovine serum (V/V), L-glutamine (0.3 mg/ml), and antibiotics. Cells were grown at 37°C with 5% CO_2 in 75 cm^2 vented culture flasks. Flasks used for primary HFG cells were first treated with poly-D-lysine (0.1 mg/ml).

Detection of multiple classes of NF-1 by RT-PCR
Total cytoplasmic RNA extraction from primary HFG or HeLa cells was done as previously described by Tornatore *et al* (1991). Poly(A)$^+$ mRNA was selected with a Oligotex-dT mini kit (Qiagen). Synthesis of cDNA was carried out as recommended by the manufacturer (Perkin Elmer) in a total volume of 20 μl of a solution consisting of 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 5 mM $DMgCl_2$, 1 mM each deoxynucleotide (dNTP), 20 U of RNAse inhibitor, 50 U of Moloney murine leukemia virus reverse transcriptase (RT), 0.12 – 0.30 μg of template mRNA, and 1 μg of an antisense poly dT17 oligonucleotide. The reaction mixture was incubated at 42°C for 1 h and then incubated at 95°C for 5 min. PCR amplification of the cDNA was carried out in a final volume of 100 μl as recommended by the manufacturer (Perkin Elmer) with 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 2 mM $MgCl_2$, 0.2 mM each dNTP, 2.5 U of Ampli-*taq* DNA polymerase and 15 pmoles of each PCR primer. A positive control PCR reaction was included with each RT-PCR assay with PCR primers (Clontech) for the human G3PDH cDNA. An expected product of 983 bp would be produced from the appropriate cDNA template, however, in the presence of genomic DNA a PCR product of 3315 bp would be produced. Table 2 lists the RT-PCR primer sets and internal probes used to identify the expression of the different classes of NF-1. The PCR primers used for class A NF-1 members were derived from the NF-1-L cDNA clone (Paonessa *et al*, 1988). The upstream primer consisted of nucleotides 586 – 610 (5′-CCAAGTGATGCCTGACATTAAGGACC - 3′), and the downstream primer from nucleotides 1049 – 1025 (5′-CGGTGATGCTGTGTGAATGCTGTTC-3′). The PCR primers for class B NF-1 members were derived from the NF-1/Red1 cDNA clone (Gil *et al*, 1988). The upstream primer consisted of nucleotides 732 – 754 (5′-ACCCAGCCAGCCATACTATCATG-3′) and the downstream primer from nucleotides 1057 – 1036 (5′-TGTGTGCGACTCCAGGTATTCC-3′). The PCR primers for class C NF-1 members were derived from the NF-1/CTF1 cDNA clone (Santoro *et al*, 1988). The upstream primer was consisted of nucleotides 585 – 605 (5′-TACTTCGTCGTGAGCGAGATG-3′) and the downstream primers from nucleotides 1007 – 986 (5′-ACGAGATGCCTCCTTCCATGTC-3′). The PCR primers for class D NF-1 were derived from the NF-1/AT1 cDNA (this study). The upstream primer consisted of nucleotides 599 – 618 (5′-AAACCAGCAAGGAGATGCGG-3′) and the downstream primer from nucleotides 900 – 878 (5′-AGAAC ACGTCATCAACAGGGCTC-3′). The PCR program consisted of 95°C for 1 min, 55°C for 1 min and 72°C for 2 min, for a total of 30 cycles in a DNA thermal cycler (Perkin-Elmer Cetus). 20 μl of each PCR product was run on a 6% polyacylamide gel, stained with ethidium bromide, visualized with UV light, and photographed.

For southern transfer 10 μl of each sample was also run on a 3% 3:1 Nusieve agarose (FMC BioProducts), transferred to a Nytran Plus filter, and DNA cross-linked by UV irradiation. The filters were hybridized as described above. Each filter was probed with a specific [^{32}P]-labeled oligonucleotide (1×10^6 cpm/ml) listed in Table 2. The probe for NF-1 class A was from NF-1-L (Paonessa *et al*, 1988) and consisted of 5′-T GGACAG T CC T GG T GAAGAACCA T T T TA-CACAGGCGAAGGGCGTCCCCA-3′. The probe for NF-1 class B was from NF-1Red1 (Gil *et al*, 1988) and consisted of 5′-CCAAAACTATATCCATAGATGA-AAATATGGAGCCAAGTCCTACAGGAGACT - 3′. The probe for NF-1 class C was from NF-1/CTF1 (Santoro *et al*, 1988) and consisted of 5′-TGGAG-GAAGACG T G GA CA C GAG CCC T GGCGGCGA T-TACTACACTTCGCCCA-3′. The probe for NF-1 class D was from NF-1/AT1 (this study) and consisted of 5′-TGAGAGTATCACAGACTCCTGTTGCA-3′.

Northern hybridization analysis
Total RNA wax extracted from primary HFG (8 week) and HeLa cells by a modification of that of Chomcynski and Sacchi (1987). Cells grown in two 75 cm^2 flasks were each homogenized in a solution of 4M guanidine thiocyanate, 25 mM sodium citrate (pH 7.0), 0.5% sarcosyl and 0.1 M 2-mercaptoethanol. The mixture was extracted after adding 0.1 volume of 2 M sodium acetate (pH 4.0), an equal volume of acid phenol (saturated with water)/chloroform and shaken vigorously for 10 s before storing the mixture on ice for 15 min. The sample was then centrifuged at 12 000 r.p.m. for 20 min and upper aqueous phase collected and mixed with an equal volume of isopropanol. After storage for more than 1 h, the sample was centrifuged at 12 000 r.p.m. for 15 min at 4°C, and pellet dissolved in 20 μl water. The RNA was quantitated by absorbance at 260 nm.

For Northern analysis, 15 μg of each RNA was electrophoresed on a 1% agarose-formaldehyde gel and RNA transferred to a Nytran Plus filter. The filter was prehybridized and hybridized with a solution consisting of 3 × sodium chloride-sodium citrate (SSC) (1 × SSC is 0.15 M sodium chloride-0.015 M sodium citrate, pH 7.0), 5 × Denhardts, 100 μg/ml salmon sperm DNA and 50% formamide. The probes (1×10^6 cpm/ml) were the same [^{32}P]-labeled 3′-end fragments from NF-1/AT1 and NF-1/CTF1 used for the southern hybridization analysis. After hybridization, the filters were washed twice with 2 × SSC at room temperature for 30 min and once with 0.2 × SSC with 0.1% SDS at 55°C for 15 min. After autoradiography, the filters were stripped and rehybridized with a human β-actin probe (2 kb cDNA), Clontech) to serve as a RNA control.

DNA Southern hybridization analysis
DNA was extracted from 15 week old HFG cells as described by Tornatore *et al* (1991). After trypsini-

zation and washing the cells with phosphate buffered saline (PBS), the cells were treated with proteinase K (100 μg/ml) and 0.5% SDS and incubated overnight at 37°C. The mixture was then extracted twice with phenol/$CHCl_3$, and twice with $CHCl_3$ before the addition of 2–3 volumes of 100% ethanol and 0.1 volume 3 M sodium acetate, pH 6.0. The preparation was stored at −20°C overnight, and DNA was removed and suspended in 10 mM Tris-HCl, pH 8.0, with 1 mM EDTA. 10 μg of HFG cell DNA was digested with 8–10-fold excess of restriction enzyme according to the manufacturers specification at 37°C overnight. The digested DNA was analyzed on a 1% agarose gel, transferred to a Nytran Plus membrane (New England Nuclear/Dupont), and DNA crosslinked (Stratagene) onto the filter. DNA and RNA hybridization probes were prepared from pCTF1 and pBRNF-1AT1. A 1023 bp 3′ probe was obtained from pCTF1 after digestion of the plasmid with *SalI* and *Eco*RI, and probe gel purified. Similarly, a 842 bp 3′ probe was isolated from pBRNF-1/AT1 after digestion of the plasmid with *AflII* and *Eco*RI, and probe gel purified. Both 3′ probes were labeled with [α-^{32}P] ATP by nick translation (Loftstrand Labs Limited). Filters containing the digested DNA were pre-hybridized in a solution containing 5 × Denhardts (1 × Denhardts is 0.02% Ficol 400, 0.02% polyvinylpryrrolidine and 0.02% bovine serum albumin), 5 × SSPE (1 × SSPE is 0.15 M NaCl, 0.1 M NaH_2PO_4, 0.001 M EDTA-Na_2), 0.5% SDS, 20 μg/ml sheared and denatured salmon sperm DNA and 50% formamide, for 3 h at 42°C. After pre-hybridization of the filters, each filter was incubated with fresh pre-hybridization solution with 2×10^6 cpm/ml of the labeled 3′ probe from pCTF1 or pBRNF-1/AT1 at 42°C for 48 h. The filters were then washed twice with 6 × SSPE with 0.5% SDS for 15 min each at 22°C, washed once with 1 × SSPE with 0.5% SDS for 15 min at 22°C and washed twice for 30 min at 42°C with 1 × SSPE with 0.5% SDS. The filters were then set up for autoradiography.

Screening human brain cDNA libraries

Two early human brain cDNA libraries were screened for NF-1 cDNAs. One of these cDNA libraries (Clontech) was prepared from fetal brain (26 weeks) and constructed in λgt11.

Approximately 5×10^5 cDNA clones were examined (Sambrook *et al*, 1989) with a [^{32}P]-labeled 57 base oligonucleotide which was homologous to a region (nt 410–466) in the DNA binding domain of NF-1/CTF1 (Santoro *et al*, 1988). Two positive cDNA clones with 1.4 kb inserts were plaque purified and one of the inserts, after both of the inserts were verified by Southern hybridization, was subcloned into pBR322 (NF-1/CT1) and M13mp19 and DNA sequenced by the dideoxy method using a commercial kit (United States Biochemicals). Another early human brain cDNA library (American Type Culture Collection) prepared from the brain stem of a 1 day old patient was screened (5×10^5 cDNA clones) as described above, and 2 cDNA clones were isolated with 1.5 kb (NF-1/AT1) and 1.2 kb (NF-1/AT2) inserts, subcloned and sequenced. Subsequently, approximately 5×10^5 cDNA clones from the later library was screened with the [^{32}P]-labeled 1.5 kb insert of NF-1/AT1. At least five additional cDNA clones were isolated, subcloned into pBluescript II KS$^+$ (Stratagene), and DNA sequence partially determined.

Cloning and overexpression of NF-1/AT

The procedure for subcloning and overexpression of NF-1/AT1 was as described previously for NF-1/CTF1 (Amemiya *et al*, 1994). The 1.5 kb cDNA insert of NF-1/AT1 was subcloned into the *Xho*I site of the expression vector pET15b (Novagen, Inc.) after creating *Xho*I sites preceding nucleotide 42 and following nucleotide 1344 of NF-1/AT1 by the polymerase chain reaction (PCR) as described by the manufacturer (Perkin Elmer). After digestion of the PCR product with *Xho*I (New England Biolabs), the resulting DNA fragment was gel purified and subcloned into pET15b, and recombinant plasmid selected in *Escherichia coli* DH5α (GIBCO BRL). The recombinant plasmid was transformed into *E coli* BL21 (DE3) (Novagen, Inc.), which is a lysogen of λDE23 which carries the T7 RNA polymerase gene under *lac* UV5 control.

In order to overexpress the NF-1/AT1 protein, the cells were grown as previously described (Amemiya *et al*, 1994). The supernatant and pellet fractions were analyzed on a 10% sodium dodecyl sulfate-polyacrylamide gels (SDS–PAGE) with a 3% polyacrylamide stacking gel (Sambrook *et al*, 1989). After washing the pellet fraction, which contained the binding activity, with wash buffer [20 mM Tris-HCl, pH 7.8; 50 mM KCl; 10 mM MgCl; 1 mM EDTA; 5 mM dithiothreitol (DTT); and 0.5 mM phenylmethylsulfonyl fluoride (PMSF); and 1% Triton X-100], the pellet was resuspended in wash buffer plus 25% glycerol (V/V) and 6 M guanidine-HCl without Triton X-100 at a protein concentration of 50 μg/200 μl. After incubating the pellet fraction at room temperature for 1 h, it was dialyzed against wash buffer containing 4 M guanidine-HCl with 25% glycerol at 4°C overnight. The dialyzed fraction was stored in small aliquots at −70°C.

Electrophoretic mobility shift assays (EMSA)

In order to monitor the specific binding activity of the recombinant NF-1/AT1 protein, EMSA was performed as previously described with 6% polyacrylamide gels (Amemiya *et al*, 1989; Amemiya *et al*, 1994). Nuclear extracts from primary HFG and HeLa cells were prepared as previously described (Amemiya *et al*, 1992). EMSA with nuclear extracts

contained 10 μg of extract and 2 μg of poly (dI-dC)-poly (dI-dC) (Pharmacia). A [^{32}P]-labeled 34 bp oligonucleotide containing a NF-1 protein recognition site (5′-TGGCTGCCAGCCAA-3′) from the enhancer region of JC virus (Amemiya *et al*, 1992) was used as the probe and unlabeled oligonucleotide used as the wild type competitor. For the mutant oligonucleotide competitor, the NF-1 site contained the sequence 5′-TTACTGCCAGCTGA-3′ (Amemiya *et al*, 1992).

Note added in proof

Sequence data described in this paper have been submitted to the GenBank data bank under accession numbers U18759, U18760 and U18761, for NF-1/AT1, NF-1/AT2 and NF-1/CT1, respectively.

Acknowledgements

We would like to thank Naoko Tanese and Robert Tjian for pCTF1 and Carlo Tornatore for technical advice in the extraction of RNA from HFG cells. We would also like to thank Jim Nagle and Lev Goldfarb (NINDS) for help in sequencing of cDNA clones and Pauline Ballew for preparation and editing the manuscript. Charlotte Sumner was a Howard Hughes Medical Research Scholar at the NIH.

References

Amemiya K, Durham L, Major EO (1994). Overexpression and reactivation of binding activity of the recombinant CTF/NF-1 transcription factor. *Protein Express Purificat* **5:** 57–64.

Amemiya K, Traub R, Durham L, Major EO (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.

Amemiya K, Traub R, Durham L, Major EO (1992). Adjacent nuclear factor-1 and activator protein binding sites in the enhancer of the neurotropic JC virus. *J Biol Chem* **267:** 14204–14211.

Aoyama A, Tamura T, Mikoshiba K (1990). Regulation of brain-specific transcription of the mouse myelin basic protein gene: function of the NF-1-binding site in the distal promoter. *Biochem Biophys Res Commun* **167:** 648–653.

Apt D, Chong T, Liu Y, Bernard HU (1993). Nuclear factor I and epithelial cell-specific transcription of human papillomavirus. *J Viol* **67:** 4455–4463.

Bohmann D, Tjian Y (1989). Biochemical analysis of transcriptional activation by Jun: differential activity of c- and v-Jun. *Cell* **59:** 709–717.

Caruso M, Iacobini C, Passananti C, Felsani A, Amati P (1990). Protein recognition sites in polyomavirus enhancer: formation of a novel site for NF-1 factor in an enhancer mutant and characterization of a site in enhancer D domain. *EMBO J* **9:** 947–955.

Celada A, Shiga M, Imagawas M, Kop J, Maki RA (1988). Identification of a nuclear factor that binds to a conserved sequence of the I-$A_B$gene. *J Immunol* **140:** 3995–4002.

Chodosh LA, Baldwin AS, Carthew RW, Sharp PA (1988). Human CCAAT-binding proteins have heterologous subunits. *Cell* **53:** 11–24.

Chomczynski P, Sacchi N (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162:** 156–159.

Chong T, Apt D, Gloss B, Isa M, Bernard H (1991). The enhancer of human papillomavirus type 16: binding sites for the ubiquitous transcription factors oct 1, NFA, TEF-2, NF-1 and AP-1 participate in epithelial cell-specific expression. *J Virol* **65:** 5933–5943.

Clec R, Corcoran LM, LeBowitz JH, Baltimore D, Sharp PA (1988). The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. *Genes & Dev* **2:** 1570–1581.

Courtois SJ, Lafontaine DA, Lemaigre FP, Durviaux SM, Rousseau GG (1990). Nuclear factor-1 and activator protein-2 bind in a mutually exclusive way to overlapping promoter sequences and trans-activate the human growth hormone gene. *Nuc Acid Res* **18:** 57–64.

deBoer E, Antoniou M, Mignotti V, Wall L, Grosveld F (1988). The human β-globin promoter: nuclear protein factors and erythroid specific induction of transcription. *EMBO J* **7:** 4203–4212.

DeVries E, Van Driel W, Tromp M, Van Boom J, Van der Vliet PC (1985). Adenovirus DNA replication *in vitro*: site-directed mutagenesis of the nuclear factor I binding site of the Ad 2 origin. *Nucleic Acids Res* **13:** 4935–4952.

Dorn A, Bollekens J, Staub A, Benoist C, Mathis D (1987). A multiplicity of CCAAT box-binding proteins. *Cell* **50:** 863–872.

Gil G, Smith JR, Goldstein JL, Slaughter CA, Orth K, Brown MS, Osborne TF (1988). Multiple genes encode nuclear factor 1-like proteins that bind to the promoter for 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Proc Natl Acad Sci USA* **85:** 8963–8967.

Gill G, Pascal E, Tseng ZH, Tjian R (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.

Gloss B, Yeo-Gloss M, Meisterernst M, Rogge L, Winnacker EL, Bernard HU (1989). Clusters of nuclear factor I binding sites identify enhancers of several papillomaviruses but alone are not sufficient for enhancer function. *Nuc Acids Res* **17:** 3517–3532.

Hennighausen L, Fleckenstein B (1986). Nuclear factor 1 interacts with five DNA elements in the promoter region of the human cytomegalovirus major immediate early gene. *EMBO J* **5:** 1367–1371.

Hennighausen L, Siebenlist U, Danner D, Leder P, Rawlins D, Rosenfeld P, Kelly T Jr. (1985). High-affinity binding site for a specific nuclear protein in the human IgM gene. *Nature* **314:** 289–292.

Hoey T, Weinzierl ROJ, Gill G, Chen JL, Dynlacht BD, Tjian R (1993). Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. *Cell* **72:** 247–260.

Inoue T, Tamura T, Furuichi T, Mikoshiba K (1990). Isolation of complementary DNAs encoding a cerebellum-enriched nuclear factor I family that activates transcription from the mouse myelin basic protein promoter. *J Biol Chem* **265:** 19065–19070.

Jeang K-T, Rawlins DR, Rosenfeld PJ, Shero JH, Kelly TJ, Hayward GS (1987). Multiple tandemly repeated binding sites for cellular nuclear factor 1 that surround the major immediate-early promoters of simian and human cytomegalovirus. *J Virol* **61:** 1559–1570.

Jones KA, Kadonaga TK, Rosenfeld PJ, Kelly TJ, Tjian R (1987). A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell* **4:** 79–89.

Jones KA, Yamamoto KR, Tjian R (1985). Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. *Cell* **42:** 559–572.

Jung H, Ahn KJ, Park JS (1993). Cloning and analysis of a rat cDNA encoding a nuclear factor-like protein (NF-1-L). *Korean Biochem J* **26:** 312–316.

Kozak M (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44:** 283–292.

Kruse U, Qian F, Sippel AE (1991). Identification of a fourth nuclear factor 1 gene in chicken by cDNA cloning: NF-1-X. *Nuc Acids Res* **19:** 6641.

Kumar KU, Pater A, Pater MM (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.

LaBella F, Sive HL, Roeder RG, Heintz N (1988). Cell-cycle regulation of a human histone H2b gene is mediated by the H2b subtype-specific consensus element. *Genes & Dev* **2:** 32–39.

Lee HY, Archer TK (1994). Nucleosome-mediated disruption of transcription factor-chromatin initiation complexes at the mouse mammary tumor virus long terminal repeat in vivo. *Mol Cell Biol* **14:** 32–41.

Leegwater PA, van Driel W, van der Vliet PC (1985). Recognition site of nuclear factor I, a sequence-specific DNA-binding protein from HeLa cells that stimulates adenovirus DNA replication. *EMBO J* **4:** 1515–1521.

Lichsteiner S, Wuarin J, Schibler U (1987). The interplay of DNA-binding proteins in the promoter of the mouse albumin gene. *Cell* **51:** 963–973.

Major EO, Amemiya K, Tornatore CS, Houff SA, Berger JR (1992). Pathogenesis and molecular biology of progressive multifocal leukoencephalopathy, the JC virus-induced demyelinating disease of the human brain. *Clin Microbiol Rev* **5:** 49–73.

Meisterernst M, Rogge L, Donath C, Gander I, Lottspeich F, Mertz R, Dobner T, Fockler R, Stelzer G, Winnacker EL (1988). Isolation and characterization of the porcine nuclear factor I (NF-1) gene. *Febs Letters* **236:** 27–32.

Meisterernst M, Rogge L, Foeckler R, Karaghiosoff M, Winnacker EL (1989). Structural and functional organization of a porcine gene coding for nuclear factor I. *Biochemistry* **28:** 8191–8200.

Mermod N, O'Neill EA, Kelly TJ, Tjian R (1989). The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. *Cell* **58:** 741–753.

Mitchell PJ, Timmons PM, Hebert JM, Rigby PWJ, Tjian R (1991). Transcription factor AP-2 is expressed in neural crest cell lineages during mouse embryogenesis. *Genes & Dev* **5:** 105–119.

Miura M, Tamura T, Aoyama A, Mikoshiba K (1989). The promoter elements of the mouse myelin basic protein gene function efficiently in NG108-15 neuronal/glial cells. *Gene* **75:** 31–38.

Nagata K, Guggenheimer RA, Enomoto T, Lichy JH, Hurwitz J (1982). Adenovirus DNA replication in vitro: identification of a host factor that stimulates synthesis of the preterminal protein-dCMP complex. *Proc Natl Acad Sci USA* **79:** 6438–6442.

Nagata K, Guggenheimer RA, Hurwitz J (1983). Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. *Proc Natl Acad Sci USA* **80:** 6177–6181.

Nilsson P, Hallberg B, Thornell A, Grundstroem T (1989). Mutant analysis of protein interactions with a nuclear factor I binding site in the SL3-3 virus enhancer. *Nuc Acids Res* **17:** 4061–4075.

Nowock J, Borgmeyer U, Peuschel AW, Rupp RAW, Sippel AE (1985). The TGGCA-binding protein binds to the MMTV-LTR, the adenovirus origin of replication, and the BK virus enhancer. *Nuc Acids Res* **13:** 2045–2061.

Oikarinen J, Hatamochi A, de Chombrugghe B (1987). Separate binding sites for nuclear factor 1 and a CCAAT DNA binding factor in the mouse α_2(I) collagen promoter. *Chem* **262:** 11064–11070.

Paonessa G, Gaunari F, Frank R, Cortese R (1988). Purification of a NF-1-like DNA binding protein from rat liver and cloning of the corresponding cDNA. *EMBO J* **7:** 3115–3123.

Rawlins DR, Rosenfeld PJ, Wides RJ, Challberg MD, Kelly TJ Jr. (1984). Structure and function of the adenovirus origin of replication. *Cell* **37:** 309–319.

Raymondjean M, Cereghini S, Yaniv M (1988). Several distinct 'CCAAT' box binding proteins coexist in eukaryotic cells. *Proc Natl Acad Sci USA* **85:** 757–761.

Rosenfeld PJ, Kelly TJ (1986). Purification of nuclear factor I by DNA recognition site affinity chromatography. *J Biol Chem* **261:** 1398–1408.

Rosenfeld PJ, O'Neill EO, Wides RJ, Kelly TJ (1987). Sequence-specific interactions between cellular DNA-binding proteins and the adenovirus origin of DNA replication. *Mol Cell Biol* **7:** 875–886.

Rossi P, Karsenty G, Roberts AB, Roche NS, Sporn MB, De Crombrugghe B (1988). A nuclear factor 1 binding site mediates the transcriptional activation of a type I collagen promoter by transforming growth factor-beta. *Cell* **52,** 405–414.

Rupp RA, Kruse U, Multhaup G, Gobel U, Beyreuther K, Sippel AE (1990). Chicken NF-1/TGGCA proteins are encoded by at least three indpendent genes: NF-1A, NF-1B and NF-1C with homologues in mammalian genomes. *Nuc Acids Res* **18:** 2607–2616.

Rupp RA, Sippel AE (1987). Chicken liver TGGCA protein purified by preparative shift electrophoresis (PMSE) shows a 36.8 to 29.8 kD microheterogeneity. *Nuc Acid Res* **15:** 9707–9726.

Sambrook J, Fritsch EF, Maniatis T (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Santoro C, Mermod N, Andrew PC, Tjian R (1988). A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. *Nature* **334:** 218–224.

Shaul Y, Ben-Levy R, De Medina T (1986). The high affinity binding site for nuclear factor I next to the hepatitis B virus S gene promoter. *EMBO J* **5:** 1967–1971.

Sock E, Wegner M, Grummt F (1991). DNA replication of human polyomavirus JC is stimulated by NF-1 *in vivo*. *Virology* **182:** 298–308.

Sundsfjord A, Johansen T, Flaegstad T, Moens U, Villand P, Subramani S, Traavik T (1990). At least two types of control regions can be found among naturally occurring BK virus strains. *J Virol* **64:** 3864–3871.

Suzuki M (1989). SPXX, a frequent sequence motif in gene regulatory protein. *J Mol Biol* **207:** 61–84.

Tamura T, Inoue T, Nagata K, Mikoshiba K (1988a). Enhancer of human polyoma JC virus contains nuclear factor-1-binding sequences: analysis using mouse brain nuclear extracts. *Biochem Biophys Res Commun* **157:** 419–425.

Tamura T, Miura M, Ikenaka K, Mikoshiba K (1988b). Analysis of transcription control elements of the mouse myelin basic protein gene in HeLa cell extracts: demonstration of a strong NF-1 binding motif in the upstream region. *Nuc Acids Res* **16:** 11441–11459.

Tanaka M, Herr W (1990). Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. *Cell* **60:** 375–386.

Tornatore C, Amemiya K, Atwood W, Conant K, Major EO, Berger J (1994). JC virus: current concepts and controversies in the molecular virology and pathogenesis of progressive multifocal leucoencephalopathy. *Rev Med Virol* **4:** 197–219.

Tornatore C, Nath A, Amemiya K, Major EO (1991). Persistent human immunodeficiency virus type 1 infection in human fetal glial cells reactivated by T cell factor(s) or by cytokines tumor necrosis factor alpha and interleukin-1 beta. *J Virol* **65:** 6094–6100.

Williams T, Tjian R (1991). Analysis of the DNA-binding and activation properties of the human transcription factor AP-2. *Genes & Dev* **5:** 670–682.

Wolffe AP (1994). Transcription: in tune with the histones. *Cell* **77:** 13–16.