Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 8707–8710, November 1990
Biochemistry

Two genes encode factors with NF-κB- and H2TF1-like DNA-binding properties

(DNA-binding proteins/transcription factors/cDNA/major histocompatibility complex class I genes)

ANIL K. RUSTGI, LAURA J. VAN 'T VEER, AND RENE BERNARDS

Laboratory of Molecular Genetics, The Cancer Center of the Massachusetts General Hospital and Harvard Medical School, 149 13th Street, Charlestown, MA 02129

Communicated by Kurt J. Isselbacher, August 10, 1990

ABSTRACT **NF-κB and H2TF1 are transcription factors that bind to related DNA sequence motifs. NF-κB is a constitutive factor in mature B lymphocytes and is inducible in immature B cells and nonlymphoid cells. H2TF1 is a constitutive ubiquitous transcription factor that activates expression of major histocompatibility complex class I genes. We report here the isolation of two zinc finger genes. One, named MBP-2 (major histocompatibility complex-binding protein 2), encodes a factor that has DNA-binding properties similar, if not identical, to the H2TF1 transcription factor. The second, named KBP-1 (kappa-enhancer binding protein 1), encodes a factor that is similar to, but distinct from, NF-κB in its DNA-binding properties and methylation interference pattern, suggesting that at least two proteins can bind to the immunoglobulin κ gene enhancer.**

NF-κB is a constitutive transcription factor in mature B lymphocytes, where it binds to the enhancer of the immunoglobulin κ gene (1). In nonlymphoid cells, and in immature B cells, NF-κB is found in an inactive form in the cytoplasm bound to an inhibitor, IκB (2, 3). A variety of agents can cause the release of active NF-κB from the inactive complex, indicating that NF-κB is an inducible transcription factor in nonlymphoid cells (4). Apart from the immunoglobulin κ gene, a number of viral and cellular regulatory regions have been shown to contain NF-κB-binding sites. They include the major histocompatibility complex (MHC) class I heavy chain genes (5), the β interferon gene (6, 7), the interleukin 2 receptor gene (8), the long terminal repeat of human immunodeficiency virus (HIV-1) (8), simian virus 40 (5), and cytomegalovirus (9), suggesting that NF-κB or NF-κB-like factors can play a role in the regulation of these genes.

H2TF1 is a transcription factor that is related to NF-κB in that both factors bind to related DNA sequence motifs (5, 10). H2TF1 differs from NF-κB, however, in that H2TF1 is a constitutive transcription factor that functions to activate expression of MHC class I genes (5, 10). Apart from H2TF1, at least one additional transcription factor, KBF1, has been reported to bind the enhancers of MHC class I heavy chain genes and β_2-microglobulin with equal affinity (11, 12). In this respect, KBF1 appears to differ from H2TF1, which has been shown to bind the enhancer of MHC class I heavy chain genes with greater affinity than the β_2-microglobulin gene enhancer (13). This would suggest that H2TF1 and KBF1 are distinct factors, although it cannot be excluded at present that both factors are products of a single gene.

We have shown previously that in neuroblastoma N-*myc* overexpression suppresses the binding of an H2TF1-like transcription factor to the MHC class I gene enhancer, thereby causing a significant reduction in the expression of MHC class I genes (14, 15). To study the interaction between the N-*myc* protein and this H2TF1-like transcription factor in more detail, we have undertaken the cloning of genes whose products can bind to the MHC class I gene enhancer. We report here the isolation of two related genes. One encodes a factor that is in DNA-binding properties similar, if not identical, to the H2TF1 transcription factor. The second encodes a factor that is in DNA-binding properties similar to, but distinct from, NF-κB.*

MATERIALS AND METHODS

Isolation of Related Genes and DNA Sequence Analysis. cDNA libraries of adenovirus-transformed human retinal cells and the human B-cell lymphoma line BJAB were screened with a 900-base-pair (bp) cDNA fragment of MBP-1 as a probe (16). Hybridization was in 30% formamide, 5× SSCPE (1× SSCPE = 0.15 M NaCl/15 mM sodium citrate/10 mM NaH_2PO_4/1 mM EDTA), 1× Denhardt's solution (0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone), 0.5 mg of denatured salmon sperm DNA per ml, and 0.1% SDS for 48 hr at 42°C. Filters were washed in 2× SSC (0.3 M NaCl/30 mM sodium citrate) at 50°C. The areas of highest homology to MBP-1 were identified by Southern hybridization and subjected to DNA sequence analysis.

DNA-Binding Mobility Shift Assay. Partial cDNA clones spanning the zinc finger domain of MBP-1, MBP-2, and KBP-1 were inserted in the β-galactosidase gene of the inducible expression vector pBSK (Stratagene) and transformed in XL1-blue bacteria. The cDNA fragments (900 bp for MBP-1 and MBP-2 and 700 bp for KBP-1) were inserted into the pBSK vector so that after induction the bacteria would synthesize a β-galactosidase–transcription factor fusion protein. Protein extracts from isopropyl β-D-thiogalactoside-induced cultures were prepared as described (16) and incubated with the different end-labeled oligonucleotides (5000 cpm). Complexes were separated on a 4% polyacrylamide/90 mM Tris/64.6 mM boric acid/2.5 mM EDTA, pH 8.3 (TBE) gel as described (15). The following oligonucleotides, subcloned in the polylinker of pUC13, were used: MHC wild type: 5'-GGGGATTCCCC-3', MHC mutant: 5'-GCGGATTCCCG-3'; NF-κB wild type: 5'-GGGGACTTTCCG-3', NF-κB mutant: 5'-ATTCACTTTCCG-3', NF-κB (G → C) mutant: 5'-GGGGACTTTCCC-3', β_2-microglobulin: 5'-AGGGACTTTCCC-3'. Oligonucleotides were labeled by digesting plasmids harboring the oligonucleotides cloned in the pUC13 polylinker with *Eco*RI and *Hin*dIII. After labeling with [^{32}P]dCTP and Klenow enzyme, the ≈70-bp fragments harboring the pUC13 polylinker sequences

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: MHC, major histocompatibility complex.
*The partial sequences of the cDNAs encoding the amino acid sequences reported in this paper have been deposited in the GenBank data base (accession nos. M33919 and M33920).

```
        1                                                                                                                  70
MBP-2 --GKYICEECGIRCKKPSMLKKHIRTHTDVRPYVCKLCNFAFKTKGNLTKHMKSKAHMKKCLELGVSMTSVD
        |||||||||||||||||||||||||||||||| |  ||| |||||||||||||||| |||
MBP-1 --GKYICEECGIRCKKPSMLKKHIRTHTDVRPYHCTYCNFSFKTKGNLTKHMKSKAHSKKCVDLGVSVGLID
        ||| ||||||||||||||||||||||||||| |  | | ||||||||||||||||||||
KBP-1 --GKYVCEECGIRCKKPSMLKKHIRTHTDVRPYVCKHCHFAFKTKGNLTKHMKSKAHSKKCQETGVLEELEA
```

ZINC FINGER DOMAIN

```
        61                                                                                                                 140
MBP-2   DTETEEAENLEDLHKAAEKHSMSSISTDHQFSDAEESDGEDGDDNDDDDEDEDDFDDQGDLTPKTRSRST--

MBP-1   EQDTEESDEKQRFSYERSGY............DLEESDGPDEDDNENEDDDEDSQAESVLSATPSVTASP--

KBP-1   EEGTSDDLFQDSEGREGSEAVEEHQFS.....DLEDSDSDSDLDEDEDEDEEESQDELSRPPSEAPPPGP--
```

ACIDIC DOMAIN

FIG. 1. Comparison of predicted partial amino acid sequences of two transcription factor genes. The amino acid sequence of MBP-1 has recently been determined by others (13, 18, 19) and is shown for comparison. The zinc finger domain and acidic domain of all three peptide sequences are indicated. Conserved cysteine and histidine residues in the zinc finger domain are underlined. Homology outside the zinc finger domain is lower and is not indicated in the comparison. Gaps have been inserted in the amino acid sequences of MBP-1 and KBP-1 to align the acidic domains of all three proteins.

with the oligonucleotides inserted were purified on a 6% polyacrylamide gel.

Methylation Interference Analysis. Methylation interference analysis was performed essentially as described by Baldwin (17). In short, ^{32}P-labeled oligonucleotides, prepared as described above, representing the H2TF1-binding motif or the NF-κB motif, were partially methylated on guanine residues using dimethyl sulfate. These oligonucleotides were then used in a preparative electrophoresis mobility shift assay with bacterial MBP-2 fusion protein (MHC class I gene enhancer oligonucleotide) or bacterial KBP-1 fusion protein (immunoglobulin κ enhancer oligonucleotide). Free probe and DNA–protein complexes were isolated. Purified DNA was cleaved with piperidine and electrophoresed on a 10% polyacrylamide/urea gel.

RESULTS

Isolation of MBP-1-Like Genes. Two years ago, Singh *et al.* (16) isolated a gene coding for a zinc finger protein that bound to the MHC class I gene enhancer and to the related NF-κB motif. Recent data indicate that this factor, named MBP-1, PRDII-BF1, or HIV-EP1, is probably distinct from NF-κB and H2TF1 (13, 18, 19). To isolate additional genes whose protein products could potentially bind to the NF-κB and H2TF1 recognition motifs, we explored the possibility that MBP-1 is a member of a gene family, whose members bind to closely related DNA sequences. To test this hypothesis, we screened several human cDNA libraries at low hybridization stringency with a MBP-1 cDNA fragment as a probe. As a result of this, we isolated two cDNAs that hybridized to MBP-1 at low stringency but not at high stringency. The first cDNA, named MBP-2, was 3.7 kilobases (kb) and was isolated from a library made from human adenovirus-transformed retinal cells. The second cDNA, named KBP-1, was 1.7 kb and was isolated from a human B-cell lymphoma cDNA library.

DNA Sequence Analysis. To determine the relationship of the two isolated genes to MBP-1, we first determined by Southern blot analysis which regions of the cDNAs shared the highest homology to MBP-1 (not shown). These fragments of the cDNAs where then subjected to DNA sequence analysis. The results of this indicated that both cDNAs contain an open reading frame that encodes a protein having a zinc finger structure (Fig. 1). Comparison of the zinc finger region of the two isolated genes to that of the MBP-1 transcription factor indicated that the three proteins share >90% amino acid homology in this domain (Fig. 1). Such zinc finger-like structures are presumed to be involved in DNA binding (20). Furthermore, all three genes were found to contain a cluster of highly acidic amino acids downstream from the presumed DNA-binding domain, suggesting that all three proteins are involved in transcription activation (21).

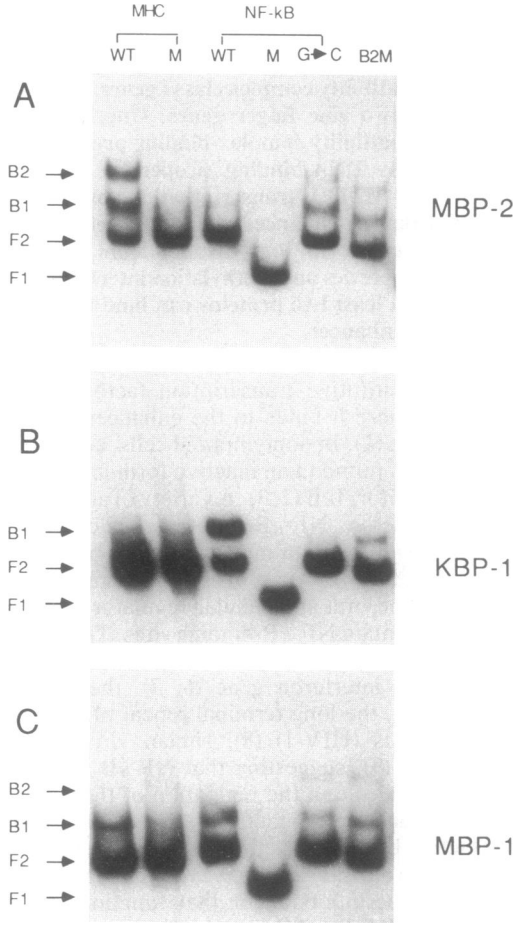

FIG. 2. Electrophoresis mobility shift assay on bacterial fusion proteins. Partial cDNA clones spanning the zinc finger domain of MBP-1, MBP-2, and KBP-1 were inserted in the β-galactosidase expression vector pBSK. Protein extracts from induced cultures (*A*, 0.2 μg; *B*, 10 μg; and *C*, 10 μg) were incubated with the different end-labeled oligonucleotides (5000 cpm) and complexes were separated on a 4% polyacrylamide/TBE gel. Arrows F1 and F2 indicate the positions of the free oligonucleotides; arrows B1 and B2 indicate retarded complexes. The mutant NF-κB oligonucleotide was shorter, due to different polylinker sequences flanking the oligonucleotide; as a result the free probe migrates faster on the acrylamide gel (indicated by arrow F1) than the other oligonucleotides (indicated by arrow F2). The following oligonucleotides, subcloned in the polylinker of pUC plasmids, were used: MHC wild type (WT): 5′-GGGGATTCCCC-3′, MHC mutant (M): 5′-GCGGATTCCCG-3′, NF-κB wild type (WT): 5′-GGGGACTTTCCG-3′, NF-κB mutant (M): 5′-ATTCACTT-TCCG-3′, NF-κB (G → C) mutant: 5′-GGGGACTTTCCC-3′, β_2-microglobulin (B2M): 5′-AGGGACTTTCCC-3′. No complexes were detected when extracts from bacteria harboring the β-galactosidase expression vector were used (data not shown).

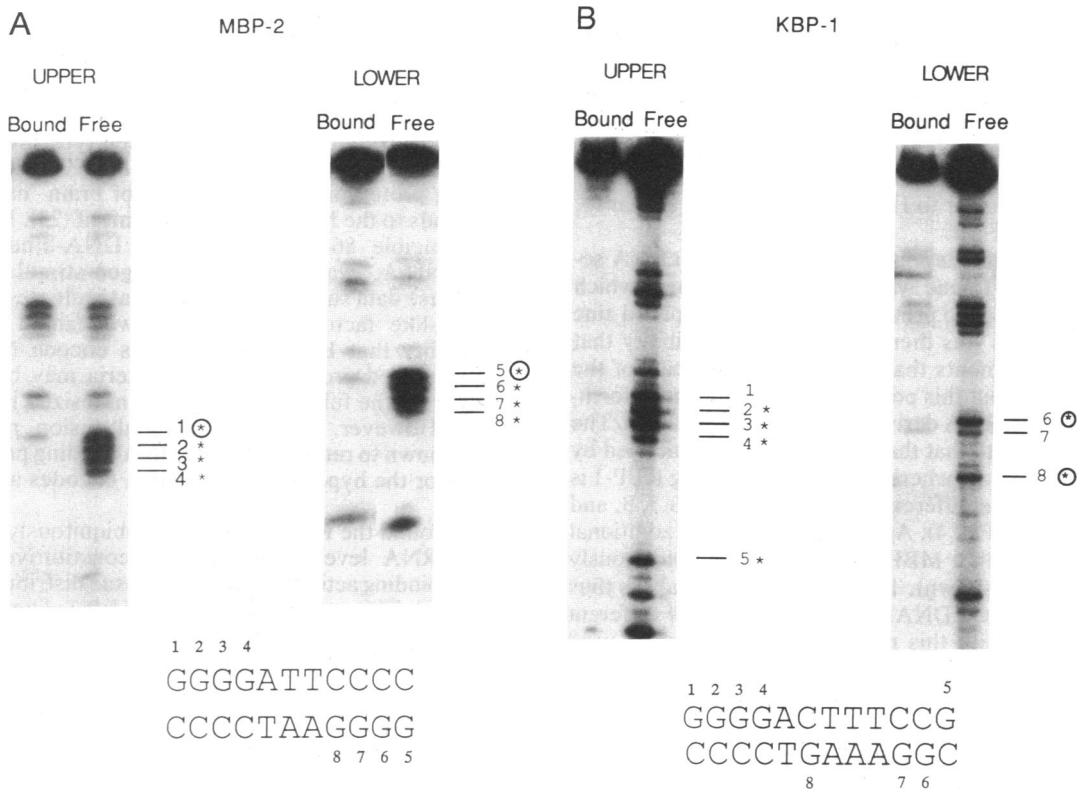

FIG. 3. Methylation interference analysis. (*A*) Methylation interference analysis of MBP-2 protein on the MHC class I gene enhancer. The panel labeled "Upper" represents the methylation interference pattern on the upper strand of the MHC class I gene enhancer; the panel labeled "Lower" represents the lower strand. The methylation interference pattern shown is that of the B2 retarded complex. The methylation interference pattern for the B1 complex was identical to that of B2 (not shown). (*B*) Methylation interference analysis of KBP-1 fusion protein on the immunoglobulin κ enhancer motif. The eight guanine residues in each of the enhancers are numbered 1–8 and their relative positions in the acrylamide gel are indicated. Asterisks indicate guanines whose methylation interferes with binding, circled asterisks indicate guanines whose methylation results in partial interference with binding.

DNA-Binding Properties. To determine the DNA-binding specificity of the factors encoded by the two isolated genes, we inserted the presumed DNA-binding domains of both genes in bacterial expression vectors. Extracts from induced bacteria were then used in a DNA-binding mobility shift assay employing oligonucleotides that represent NF-κB and H2TF1 recognition motifs as probes.

As shown in Fig. 2*A*, one of the genes encodes a protein that binds with high affinity to the MHC class I gene enhancer, with somewhat lower affinity to the β_2-microglobulin enhancer, and with even lower affinity to the NF-κB recognition motif. These findings prompted us to name this gene MBP-2 (MHC-binding protein 2). Our data thus indicate that at least two distinct genes, MBP-1 and MBP-2, encode proteins that bind to the MHC class I and β_2-microglobulin enhancers. MBP-2 differs from MBP-1, however, in that MBP-1 binds the NF-κB site with high affinity, whereas MBP-2 binds this site with much lower affinity (Fig. 2 *A* and *C*). In this respect, MBP-2 resembles the H2TF1 transcription factor, which has also been shown to bind the MHC class I gene enhancer with at least 10-fold greater affinity than the NF-κB motif (5). Furthermore, it should be noted that substitution of a single base in the NF-κB motif changes this site into a high-affinity binding site for MBP-2 (Fig. 2*A*, lane 5). This base substitution has previously been shown to change the NF-κB motif into a high-affinity site for H2TF1 (22), again indicating that MBP-2 encodes a factor whose DNA-binding properties are very similar or identical to those reported for H2TF1.

The protein encoded by the second gene binds with high affinity to the NF-κB motif and with low affinity to the β_2-microglobulin enhancer. No binding was observed to the MHC class I gene enhancer or to the NF-κB (G $\rightarrow$ C) mutant (Fig. 2*B*, lanes 1 and 5). This pattern of binding is distinct from that observed for NF-κB, which has been shown to bind the immunoglobulin κ (G $\rightarrow$ C) mutant enhancer and the MHC class I gene enhancer with equally high affinity (5, 22). Because of this, we named this gene KBP-1 (kappa-enhancer binding protein 1). The binding properties of MBP-1 are shown for comparison and are distinct from KBP-1 and MBP-2 (Fig. 2*C*). We conclude from the data shown above that MBP-1, MBP-2, and KBP-1 are members of a gene family whose products bind to related enhancer motifs. It seems likely that amino acid substitutions in the region surrounding the second pair of cysteines in the zinc finger domain cause the differences in DNA-binding specificity of these factors (Fig. 1).

Methylation Interference Analysis. To further evaluate the relationship between the two isolated transcription factor genes and the NF-κB and H2TF1 factors, we performed methylation interference analysis. Fig. 3*A* shows that binding of MBP-2 to the MHC class I gene enhancer is perturbed by methylation of guanines 2, 3, 4, 6, 7, and 8, whereas methylation of guanines 1 and 5 caused a partial interference of transcription factor binding. This pattern of methylation interference is virtually indistinguishable from that of H2TF1 and KBF1 (10, 11).

Fig. 3*B* shows that guanines 2, 3, 4, and 5 of the NF-κB recognition motif interfere strongly with the binding of KBP-1 to DNA, whereas guanines 6 and 8 are interfering weakly. We also found that some of the guanine residues flanking the NF-κB motif were protected by the KBP-1 fusion protein. This pattern of interference was highly reproducible and is probably a result of interactions of the bacterial β-galactosidase protein, to which

the KBP-1 DNA-binding domain is fused, with flanking nucleotides. It has previously been shown that guanines 2, 3, 4, 6, 7, and 8 interfere with NF-κB binding to its cognate site (1). This pattern of interference is different from that of KBP-1, indicating that KBP-1 is distinct from NF-κB. The methylation interference data thus appear to corroborate the conclusions drawn from the DNA-binding assays in that they suggest that MBP-2 is identical or very similar to H2TF1, and KBP-1 is similar to, but distinct from, NF-κB.

Transcription Factor Expression. The complete cDNA sequence of the MBP-1 gene was recently determined, which indicated that its encoded protein has two widely spaced zinc finger motifs (18). It was therefore a formal possibility that the two cDNA fragments that we isolated were part of the same gene. To rule out this possibility, we performed Northern blot analysis on RNA derived from several cell lines. The results of this indicate that the MBP-2 protein is encoded by a 10-kb mRNA (Fig. 4), whereas the gene encoding KBP-1 is transcribed into three different mRNA species of 3.5, 8, and 12 kb, respectively (Fig. 4). Analysis of RNA from additional cell lines indicated that MBP-1 an KBP-1 are ubiquitously expressed (data not shown). The Northern blot analysis thus suggests that the two cDNAs are the products of different genes. Consistent with this notion, we have recently used panels of somatic cell hybrids to show that the MBP-2 and KBP-1 genes map to different human chromosomes (A.K.R. and R.B., unpublished data).

DISCUSSION

Several DNA-binding factors have been described that bind the MHC class I gene enhancer, including H2TF1, KBF1, and NF-κB (5, 10, 11). The methylation interference pattern of the MBP-2 transcription factor is similar to that of H2TF1 and KBF1 (10, 11). However, the DNA-binding specificity of MBP-2 resembles more closely that of H2TF1, which has been shown to have greater affinity for the MHC class I gene enhancer than for the β_2-microglobulin enhancer (13), whereas KBF1 has been shown to have approximately equal affinity for both enhancer motifs (11). The finding that MBP-1 and MBP-2 can bind the enhancers of the MHC class I heavy chain and β_2-microglobulin genes suggests that at least two distinct transcription factors can regulate the coordinate expression of these genes.

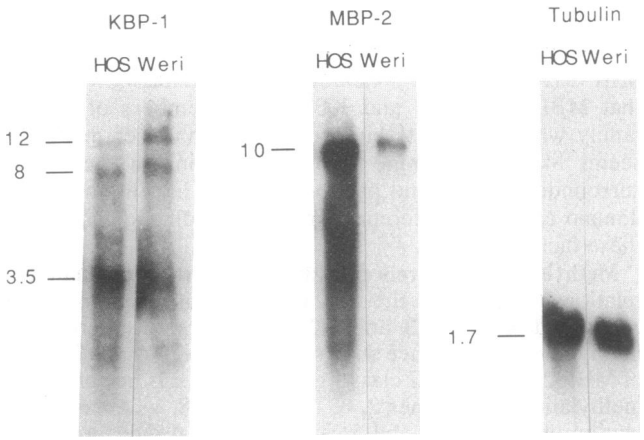

FIG. 4. Northern blot analysis of human tumor cell RNAs. Five micrograms of poly(A)-selected RNA derived from human osteosarcoma cell line HOS or retinoblastoma Weri-1 was electrophoresed through a 1% agarose/formaldehyde gel and transferred to nitrocellulose as described (14). Replicate filter strips were then hybridized to ^{32}P-labeled probes of either KBP-1 (*Left*) or MBP-2 (*Center*) or, as a control, with α-tubulin (*Right*). Filters were hybridized in 5× SSCPE/50% formamide and washed in 2× SSC at 55°C. Sizes are indicated in kb.

The KBP-1 transcription factor described here is distinct from NF-κB in its DNA-binding specificity and in its methylation interference pattern on the NF-κB site. The existence of NF-κB-like factors with limited cell type distribution has been described. Thus, Korner *et al.* (23) found a DNA-binding protein in the gray matter of brain, named BETA, that binds to the NF-κB recognition motif (23). Furthermore, an inducible 86-kDa NF-κB-like DNA-binding protein, HIVEN86A, was found in mitogen-stimulated T cells (8). These data support the notion that multiple genes encode NF-κB-like factors. At present, we cannot exclude the possibility that KBP-1 in fact does encode NF-κB, since the truncated protein made in bacteria may behave differently from the full-length protein synthesized in mammalian cells. However, as other bacterial fusion proteins have been shown to retain legitimate DNA-binding properties (24), we favor the hypothesis that KBP-1 encodes an NF-κB-like factor.

We found the KBP-1 gene to be ubiquitously expressed at the mRNA level. However, no constitutive NF-κB-like DNA-binding activity with a wide tissue distribution has been reported. This would suggest that KBP-1, like NF-κB, is an inducible transcription factor. The cloning of the NF-κB-like gene should allow us to study the regulation of these factors in more detail.

We thank Harinder Singh, Albert Baldwin, and Phillip Sharp for the gift of the MBP-1 cDNA and Ranjan Sen for oligonucleotides. This work was supported in part by the Searle Scholarship Foundation and the Edward J. Mallinckrodt Foundation. L.J.V. was supported by a fellowship from the Netherlands Organization for Scientific Research (NWO). A.K.R. was supported by fellowships from the Aid for Cancer Research Foundation, National Institutes of Health Training Grant T32DK07191, and the American Gastroenterological Association.

1. Sen, R. & Baltimore, D. (1986) *Cell* **46**, 705–716.
2. Bauerle, P. A. & Baltimore, D. (1988) *Cell* **53**, 211–217.
3. Bauerle, P. A. & Baltimore, D. (1988) *Science* **242**, 540–546.
4. Lenardo, M. J. & Baltimore, D. (1989) *Cell* **58**, 227–229.
5. Baldwin, A. S. & Sharp, P. A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 723–727.
6. Goodbourn, S. & Maniatis, T. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1447–1451.
7. Lenardo, M. J., Fan, C. M., Maniatis, T. & Baltimore, D. (1989) *Cell* **57**, 287–294.
8. Böhnlein, E., Lowenthal, J. W., Siekevitz, M., Ballard, D. W., Franza, B. R. & Greene, W. C. (1988) *Cell* **53**, 827–836.
9. Boshart, M., Werber, F., Posch-Hasler, G. J., Fleckenstein, B. & Schaffner, W. (1985) *Cell* **41**, 521–530.
10. Baldwin, A. S. & Sharp, P. A. (1987) *Mol. Cell. Biol.* **7**, 305–313.
11. Israel, A., Kimura, A., Kieran, M., Yano, O., Kanellopoulos, J., Le Bail, O. & Kourilsky, P. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2653–2657.
12. Yano, O., Kanellopoulos, J., Kieran, M., Le Bail, O., Israel, A. & Kourilsky, P. (1987) *EMBO J.* **6**, 3317–3324.
13. Baldwin, A. S., LeClair, K. P., Singh, H. & Sharp, P. A. (1990) *Mol. Cell. Biol.* **10**, 1406–1414.
14. Bernards, R., Dessain, S. K. & Weinberg, R. A. (1986) *Cell* **46**, 667–674.
15. Lenardo, M., Rustgi, A. K., Schievella, A. R. & Bernards, R. (1989) *EMBO J.* **8**, 3351–3355.
16. Singh, H., LeBowitz, J. H., Baldwin, A. S. & Sharp, P. A. (1988) *Cell* **52**, 415–423.
17. Baldwin, A. S. (1990) in *Current Protocols in Molecular Biology*, eds. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (Green, Brooklyn, NY), pp. 1231–1235.
18. Fan, C. M. & Maniatis, T. (1990) *Genes Dev.* **4**, 29–42.
19. Maekawa, T., Sakura, H., Sudo, T. & Ishii, S. J. (1989) *J. Biol. Chem.* **264**, 14591–14593.
20. Evans, R. M. & Hollenberg, S. M. (1988) *Cell* **52**, 1–3.
21. Mitchell, P. J. & Tjian, R. (1989) *Science* **245**, 371–378.
22. Mauxion, F. & Sen, R. (1989) *Mol. Cell. Biol.* **9**, 3548–3552.
23. Korner, M., Rattner, A., Mauxion, F., Sen, R. & Citri, Y. (1989) *Neuron* **3**, 563–572.
24. Singh, H., Clerc, R. G. & LeBowitz, J. H. (1989) *Bio/techniques* **7**, 252–261.