

Cloning and Characterization of a Human *c-myc* Promoter-Binding Protein

RATNA RAY AND DONALD M. MILLER*

Departments of Internal Medicine and Biochemistry, Comprehensive Cancer Center, University of Alabama at Birmingham, and the Birmingham VA Medical Center, Birmingham, Alabama 35294

Received 30 October 1990/Accepted 11 January 1991

A human cDNA clone encoding a *c-myc* promoter-binding protein was detected by screening a HeLa cell λ phage expression cDNA library. The library was screened by using an *XhoI-NaeI* human *c-myc* P₂ promoter fragment as a probe. The recombinant phage encoded a fusion protein, *myc*-binding protein 1 (MBP-1), which had an apparent molecular size of 40 kDa. A corresponding protein with a molecular size of 35 kDa was present in a HeLa cell extract. Sequence analysis of the cloned gene reveals an open reading frame of 1,038 bp with a 3' untranslated region of 378 bp. The predicted protein sequence contains a proline-rich region in the amino terminus but does not demonstrate a known DNA-binding domain. DNase I footprint analysis demonstrates that MBP-1 binds to the sequence just 5' of the TATA box sequence of the human *c-myc* P₂ promoter. MBP-1 cDNA hybridizes to a 1.4-kb mRNA from HeLa and HL-60 cells, indicating that the cDNA insert (1,416 bp) is a full-length clone. Coexpression of the MBP-1 protein represses transcription from the human *c-myc* promoter, suggesting that MBP-1 may act as a negative regulatory factor for the human *c-myc* gene.

Nuclear proto-oncogene products play an important role in the regulation of cell growth and differentiation. They also mediate signal transduction in the nucleus and participate in the regulatory processes controlling transcription and DNA replication (1, 6, 18, 63). The *c-myc* proto-oncogene appears to play a major role in the control of cell proliferation (9). The deregulation of *c-myc* expression, for example by amplification, chromosomal translocation, retroviral insertion, or point mutation, causes immortalization or transformation of normal cells (6, 12, 55). Regulation of *c-myc* gene expression has been shown to involve several mechanisms, including changes in transcription initiation and elongation, RNA stability, and RNA turnover and translation (5, 12, 36, 55). However, the cellular factors involved in the regulatory control of this gene have not been characterized.

Transcriptional regulation of several genes has been shown to involve the specific interaction of regulatory proteins with DNA sequences that are located in the 5'-flanking region (7). The human *c-myc* proto-oncogene contains two TATA box sequences separated by about 165 bp located near the 5' end of the first exon (3). A negative regulatory sequence has been localized to a 57-bp segment approximately 300 bp upstream of the human *c-myc* P₁ promoter (19). The transcription of *c-myc* from P₁ and P₂ is regulated by a composite of positive and negative elements located both upstream and downstream of these promoters (10, 19, 37, 59, 69). Recently, Asselin et al. (2) have identified the sequences upstream of the mouse P₁ and P₂ *c-myc* promoters that form specific complexes with the nuclear factors which are required for normal P₁ and P₂ transcription. Lipp et al. (37) suggested that the 142-bp fragment upstream of the 5' end of the *c-myc* gene acts as a positive regulatory region. Postel et al. (53) also showed that a 27-bp polypurine-polypyrimidine sequence present upstream of the human *c-myc* P₁ promoter interacts with a transcription initiation factor.

Sequence-specific interactions between DNA-binding pro-

teins and *cis*-acting elements have also been shown to play an important role in controlling gene expression (13, 14, 41). A combination of different DNA elements can generate various types of transcriptional regulation. Multiple factors can bind to the same DNA sequence or region and produce diversity in transcriptional control (1, 23, 26, 27, 28, 29, 34, 45, 46, 49, 57, 62). An example of this multifactorial regulation is the negative element of the *c-myc* gene (-350 to -296 bp relative to the transcription start site of P₁). Fos/Jun and octamer-binding proteins have also been shown to interact with this region (20, 67). The same DNA sequences may function as either positive or negative regulatory elements, depending on the DNA-binding factor(s) present in different cell types.

To characterize protein-DNA interactions involved in the control of *c-myc* expression, we have identified and cloned a *c-myc* promoter-binding protein (MPB-1) that forms a DNA-protein complex upstream of the transcriptional start site of the P₂ promoter. Structural and functional analyses suggest that this cDNA encodes a novel *trans*-acting factor that may be a negative regulator of *c-myc* expression.

MATERIALS AND METHODS

cDNA expression library. A HeLa cell cDNA library constructed in the λ zapII expression vector was obtained from Stratagene. The library contained approximately 3×10^{10} plaques and had an average insert size of 1 kb. *Escherichia coli* XL1-blue (Stratagene) was used as a host strain.

Plasmid. Plasmid pGEM-myc1 was constructed by subcloning the *SmaI-NaeI* fragment from the first-exon region of the human *c-myc* gene (55) into the *SmaI* site of the pGEM-3Z vector. The plasmid contained a 142-bp *XhoI-NaeI* fragment of the human *c-myc* first exon which has been suggested to function as a positive regulatory sequence of the human *c-myc* gene (37).

Probe. pGEM-myc1 was digested with restriction enzymes *XhoI* and *EcoRI*. The digested product was end labeled with [α -³²P]dATP and [α -³²P]dCTP, using Klenow enzyme, dGTP, and dTTP to fill in the ends of the restriction

* Corresponding author.

fragments. The 150-bp labeled fragment was separated by 8% native polyacrylamide gel electrophoresis, eluted from the excised gel, and purified by using ultrafree-MC filters (Millipore Corp.). This probe was used to screen the HeLa cell cDNA expression library.

Screening. The HeLa cell cDNA library was plated on LB agar and overlaid with nitrocellulose filters saturated with 10 mM isopropyl- β -D-thiogalactopyranoside (IPTG). A total of 2×10^6 plaques were screened with the radiolabeled 150-bp 5'-flanking *c-myc* probe as described by Singh et al. (60) and Vinson et al. (68).

Binding assay. MBP-1 cDNA was transferred to the pBluescript SK(+/-) (Stratagene) vector by in vivo excision. Extract for assaying DNA-binding activity of the fusion protein was prepared as follows. Cells were grown at 37°C to an optical density of 0.2 at 600 nm. IPTG was added to a final concentration of 10 mM, and the cells were induced for 1 h. The cells were pelleted and resuspended in lysis buffer (50 mM Tris [pH 8.0], 1 mM EDTA, 10% sucrose, 1 mM phenylmethylsulfonyl fluoride, 1% aprotinin). Lysozyme was added to a final concentration of 1 mg/ml, and the suspension was incubated for 10 min on ice. Subsequently, Nonidet P-40 was added to a final concentration of 0.1%, and incubation was continued for another 10 min. After a 1-h spin at $35,000 \times g$, the supernatant was quick-frozen and stored at -70°C. The binding reactions and gel electrophoresis were performed as previously described (54).

DNase I footprinting analysis. A 68-bp (-47 to +21 bp relative to P₂ start site 1) DNA fragment was chosen for DNase I footprint analysis on the basis of preliminary results of a DNase protection assay using an *XhoI-NaeI* fragment. This 68-bp DNA fragment was amplified from pSV₂neomyc (55) plasmid DNA by using two synthetic oligonucleotide primers (5'-AGGGATCCGCGCTGAGTAT-3' and 5'-GCTGGAATTCCTACTACAGCG-3') by polymerase chain reaction. The sequence of the amplified DNA was verified by the Maxam-Gilbert sequencing method (39).

DNase I footprinting analysis of the human *c-myc* P₂ fragment with MBP-1 was performed as previously described (54). In brief, the radiolabeled fragment was incubated with crude bacterial extract for 20 min at room temperature. One unit of DNase I was added, and the reaction was terminated after 1 min with DNase stop buffer (100 mM Tris-HCl [pH 7.5], 100 mM NaCl, 20 mM EDTA). The sample was extracted with phenol-chloroform and precipitated with ethanol, and equal counts of all samples were analyzed on a 10% urea-polyacrylamide sequencing gel.

In vitro translation and hybrid arrest. In vitro transcription of the MBP-1 cDNA, cloned in the pBluescript vector, was performed by using T3 RNA polymerase (Stratagene) as instructed by the manufacturer (Promega Biotec). Rabbit reticulocyte lysate (Promega Biotec) was used for in vitro translation, and the reaction was carried out in the presence of [³⁵S]methionine as instructed by the manufacturer. Hybrid arrest translation was performed by using *HindIII*-digested plasmid DNA (MBP-1) as described by Paterson et al. (50).

Southwestern (DNA-protein) and Western immunoblot analyses. Bacterial extracts were electrophoresed on a 7.5 or 10% sodium dodecyl sulfate-polyacrylamide gel and transferred to a nitrocellulose filter. The filter was blocked with 5% (wt/vol) nonfat dry milk in 10 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES) buffer (pH 7.9) and incubated at room temperature for 1 h. The filter was incubated in binding buffer (10 mM HEPES [pH 7.9], 50 mM NaCl, 10 mM MgCl₂, 0.1 mM EDTA, 1 mM dithiothreitol,

0.25% nonfat dry milk) containing 10⁵ cpm of ³²P-labeled probe per ml for another hour (44). The filter was washed with binding buffer containing 0.3 M NaCl for 1 h and exposed to X-ray film for autoradiography.

Western blot analysis was performed by using an anti- β -galactosidase monoclonal antibody (Promega Biotec) or a polyclonal antibody (Cappel, Organon Teknika Co.) according to the standard procedure.

Sequencing. MBP-1 cDNA was subcloned into the pBluescript SK(+/-) vector. Sequencing of both strands was performed by the dideoxy-chain elongation method (56), using a Sequenase kit (United States Biochemicals), initiating SK and KS primers for chain elongation, and specific synthetic oligonucleotides as the continuing primers. The ambiguous sequences were confirmed by the chemical cleavage method (39).

CAT cotransfection analysis. The full-length MBP-1 was cloned into the pSV₂neo vector under the control of the simian virus 40 (SV40) early promoter in sense (pSV₂MBP-1) or antisense (pSV₂anti-MBP-1) orientation for use as an expression vector. The chloramphenicol acetyltransferase (CAT) reporter plasmid contained the human *c-myc* promoter (52) or chicken β -actin promoter (29). H-*myc* CAT and S-*myc* CAT (-2290 and -100, respectively, relative to the human *c-myc* transcription start site) reporter plasmid (5 μ g) and effector plasmid (5 μ g) were cotransfected by the lipofection reagent as described by the manufacturer (GIBCO BRL). A plasmid expressing the β -galactosidase gene from the Rous sarcoma virus long terminal repeat (RSV- β gal) (4 μ g) was included as an internal control for transfection efficiency. CV1 cells were grown in 10% newborn calf serum and plated at a density of 10⁶ cells per 100-mm dish and transfected with 14 μ g of total plasmid DNA. The level of CAT activity was determined 45 h after transfection, using equivalent levels of protein extract (16). CAT activity was measured by determining the percentage of acetylated ¹⁴C-labeled chloramphenicol. Both the acetylated and nonacetylated forms of ¹⁴C-labeled chloramphenicol were removed from the thin-layer chromatography plate, and the amount of radioactivity was analyzed to determine CAT activity.

Northern (RNA) blot analysis. Cytoplasmic RNA was isolated from HL-60 and HeLa cells for Northern blot analysis, using a 1.4-kb fragment of MBP-1 cDNA as a probe as described previously (55).

Nucleotide sequence accession number. The nucleotide sequence data reported here have been submitted to the GenBank data base and assigned accession number M55914.

RESULTS

Detection of a clone encoding a sequence-specific DNA-binding protein. The λ zapII HeLa cell cDNA expression library was screened with a double-stranded DNA fragment containing the *c-myc* P₂ promoter region. *c-myc* promoter-binding proteins expressed by the recombinant phages were detected by specific binding with the α -³²P-labeled 142-bp *XhoI-NaeI* fragment. Six recombinant phages whose products bound to this fragment were obtained from the primary screening. Four distinct positive clones were detected after the secondary and tertiary screenings, all of which bound specifically to the 142-bp *c-myc* probe. The positive clones were called MBP-1, MBP-2, MBP-3, and MBP-4. The filters were rescreened with specific (*c-myc*) and nonspecific probes to confirm the authenticity of recombinant phages. The screening strategy for the characterization of MBP-1

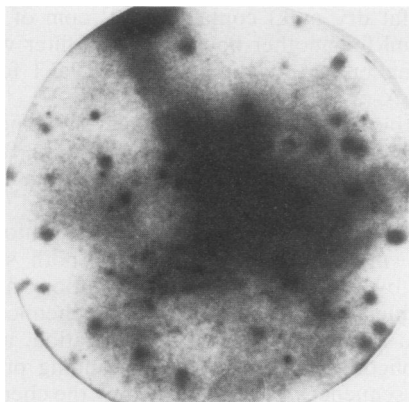


FIG. 1. Isolation of a λ recombinant clone binding specifically to the human *c-myc* P₂ promoter. A filter generated from tertiary stock of MBP-1 was probed with the 142-bp human *c-myc* P₂ promoter (5×10^5 cpm/ml).

demonstrated that this clone encoded a specific DNA-binding protein (Fig. 1). Phage DNA was transferred to the pBluescript SK(+/-) vector by in vivo excision. Digestion of the recombinant bacterial DNA with *Eco*RI showed that one of the positive clones (MBP-1) contained a 1.4-kb gene insert. Subsequent studies were performed with use of this clone.

Characterization of the DNA-binding protein. Gel mobility shift analysis using a crude bacterial extract from MBP-1-transformed *E. coli* demonstrated the formation of a specific DNA-protein complex between MBP-1 and the *Xho*I-*Nae*I *c-myc* P₂ promoter fragment. This resulted in gel retardation in comparison with the free DNA probe which served as a negative control (Fig. 2A) under nondenaturing, low-ionic-strength polyacrylamide gel electrophoresis (54, 58). The recombinant protein from the bacterial extract bound to the 142-bp *c-myc* P₂ DNA fragment and formed protein-DNA complexes (Fig. 2A, lane 4). On the other hand, the extract from vector-transformed bacteria (BS), used as a negative control, did not show the formation of a DNA-protein complex (lane 1). Binding with crude HeLa nuclear extract and this fragment also formed a protein-DNA complex (lanes 2 and 5). This result suggested the specificity of binding of the recombinant bacterial fusion protein to the P₂ fragment of the human *c-myc* gene.

A competition experiment with the *c-myc* P₂ fragment and partially purified recombinant MBP-1 in a gel mobility shift assay showed that the DNA-protein complex is not formed in the presence of excess double-stranded 142-bp (*Xho*I-*Nae*I) unlabeled P₂ fragment (Fig. 2B, lane 2), whereas unrelated sequences did not block formation of the DNA-protein complex (lane 3). This result indicates that the recombinant protein expressed by the MBP-1 clone binds specifically to sequences contained in the *c-myc* P₂ fragment.

A DNase I protection assay was performed to determine the sequence specificity of MBP-1 binding to the human *c-myc* P₂ fragment (Fig. 3). MBP-1 protein showed binding to the *c-myc* fragment (Fig. 3A, lane 2; Fig. 3B, lanes 2 and 3), whereas an extract from vector-transfected cells (BS protein) did not react with that fragment (Fig. 3A, lane 3; Fig. 3B, lane 4). The fusion protein protects 30 bp just 5' of the *c-myc* P₂ transcription start site (-38 to -8 bp relative to the P₂ transcription start site). To determine the effect of mutation at the binding site of MBP-1, wild-type and mutant

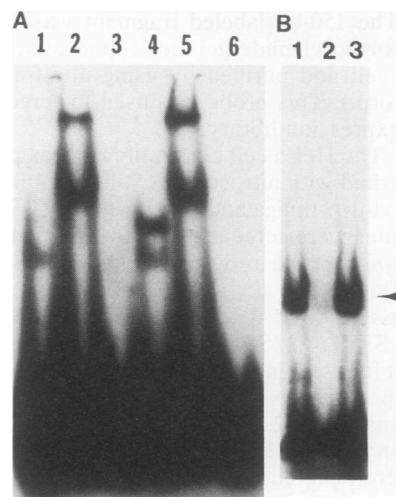


FIG. 2. Gel mobility shift analysis of the human *c-myc* P₂ promoter with the fusion protein encoded by MBP-1. (A) Extracts prepared from induced cultures of MBP-1 and BS (vector-transfected control) were incubated with a 142-bp *Xho*I-*Nae*I human *c-myc* P₂ fragment (20,000 cpm) and 5 μ g of poly(dI-dC). The reactions were resolved by 5% nondenaturing polyacrylamide gel electrophoresis. Assay mixtures contained 20 μ g of BS protein (lane 1), 15 μ g of crude nuclear extract (lanes 2 and 5), 20 μ g of MBP-1 protein (lane 4), or no protein (lanes 3 and 6). (B) Gel retardation MBP-1 binding assay with a 142-bp *Xho*I-*Nae*I human *c-myc* P₂ fragment. The gel mobility shift assay was performed as described for panel A. Lane 1 contained 20 μ g of MBP-1 protein with 5 μ g of poly(dI-dC). MBP-1 binding was completed in the presence of a 200-fold excess of the 142-bp *Xho*I-*Nae*I fragment (lane 2) or a 89-bp *Nco*I-*Nsi*I restriction fragment carrying the SV40 early promoter of plasmid PL1 (54) (lane 3). Competition was not observed with an excess of the 89-bp SV40 promoter fragment.

double-stranded oligonucleotides were synthesized. A single guanine residue within the binding site was replaced with pyrimidine in the mutant oligomer. A gel mobility shift assay showed binding of MBP-1 to the wild-type oligomer, whereas the mutant oligomer failed to form a protein-DNA complex (Fig. 4).

Analysis of in vitro translation products of the recombinant MBP-1 clone after transcription with T3 RNA polymerase suggested the presence of an open reading frame. The MBP-1 clone contains an open reading frame that yields a 37-kDa protein product (Fig. 5, lane 1), correlating with the 1.4-kb size of the cDNA. The specificity of this translated product was verified by hybrid arrest. The 1.4-kb DNA fragment from the MBP-1 clone completely blocked translation (lane 2). After melting of the RNA-DNA complex, the translation inhibition was reversed, with reappearance of the translated product (lane 3). This experiment provides further evidence that the 37-kDa translated protein product originates from the MBP-1 clone.

Southwestern blot analysis was performed to determine the specificity of DNA binding as well as the molecular weight of the recombinant fusion protein. In this experiment, the extents of binding of the *c-myc* P₂ promoter fragment to filter-bound proteins from HL-60, HeLa, CV1, and MBP-1-transformed *E. coli* cells were compared. The recombinant protein showed an apparent molecular size of 40 kDa (Fig. 6A). The same DNA probe recognized a 35-kDa protein band in the HL-60 and HeLa cell lysates (Fig. 6B, lanes 1 and 2) but not from an unrelated monkey kidney cell line

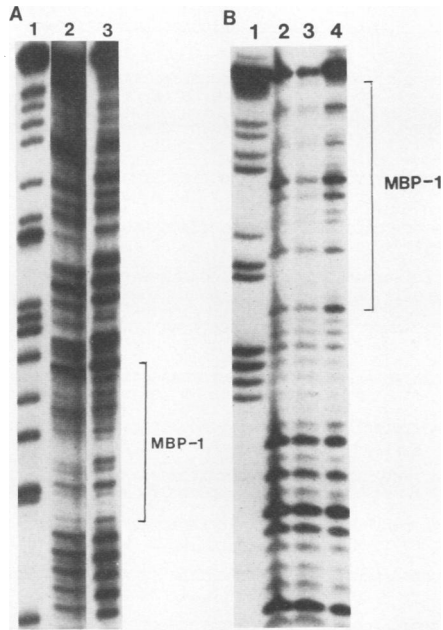


FIG. 3. DNase I protection assay using the human *c-myc* P₂ fragment (coding and noncoding) to identify the MBP-1-binding site. (A) A DNase I protection assay was performed by using a 142-bp *XhoI*-*NaeI* fragment 3' end labeled at the *XhoI* site. Lane 1 contained a Maxam-Gilbert (39) guanine-specific reaction. Labeled DNA was incubated with MBP-1 protein (lane 1, 30 μ g of protein) or BS protein (lane 2) and then digested with DNase I. (B) The DNase I protection assay was performed by using a 68-bp amplified fragment 3' end labeled at the *EcoRI* site. Labeled DNA was incubated with BS protein (lane 4) or MBP-1 protein (lanes 2 and 3, containing 15 and 30 μ g of protein, respectively) and then digested with DNase I. Lane 1 contained a Maxam-Gilbert (39) guanine-specific reaction.

(CV1) (lane 3). This difference may be due to the fusion sequence present in the recombinant bacterial protein. The high-molecular-weight bands probably represent nonspecific binding with probe.

Although the cloned DNA-binding protein is fused with the β -galactosidase gene, both polyclonal and monoclonal anti- β -galactosidase antibodies failed to recognize the fusion protein in Western blot analysis. The pBluescript SK(+/-) vector contains 36 amino acid residues from the N-terminal end of the β -galactosidase gene, and possibly the antibodies do not recognize any epitope(s), if present, within these 36 amino acid residues under the experimental conditions used.

Structural analysis of MBP-1. The MBP-1 cDNA clone was sequenced by the dideoxynucleotide chain elongation method. The nucleotide and predicted amino acid sequences are shown in Fig. 7. The MBP-1 gene contains an open reading frame of 1,038 bp and a 3' untranslated region of 378 bp. The nature of the fusion protein requires transcription of the entire open reading frame. The protein sequence derived from the entire open reading frame contains 336 amino acid residues and has a calculated molecular size of 37 kDa. This is in excellent agreement with the size characteristics of the *in vitro* translation product in a rabbit reticulocyte lysate system (Fig. 5) as well as size data obtained by Southwestern blot analysis (Fig. 6B).

The following features were noted upon analysis of the MBP-1 sequences. The predicted amino acid sequence con-

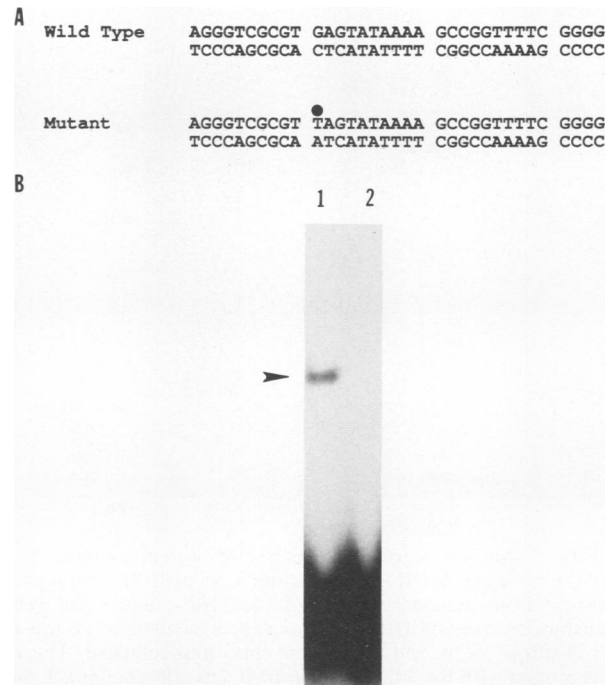


FIG. 4. Inhibition of protein binding by mutation of the MBP-1-binding site. (A) Sequences of the wild-type and mutant synthetic oligomers used for the mobility shift assay. The guanine-to-thymidine substitution is indicated (●). (B) Gel mobility shift analysis using wild-type (lane 1) and mutant (lane 2) oligomers. The protein-DNA complex is marked by an arrowhead.

tains 14% basic residues, is enriched with leucine and serine, and has a potential glycosylation site at the amino terminus. There is no zinc finger motif present in the MBP-1 sequence. The amino-terminal region of this protein has proline-rich residues. Computer analysis of the MBP-1 clone using the

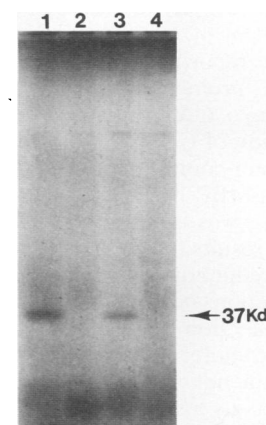


FIG. 5. *In vitro* translation and hybrid arrest using the MBP-1 clone. The MBP-1 clone was linearized with *HindIII*; RNA was transcribed by using T3 RNA polymerase and translated by using rabbit reticulocyte lysate. The protein product was analyzed by SDS-polyacrylamide gel electrophoresis. The translated product is shown in lane 1. Lane 2 shows the result of translation after hybridization of the transcript with the 1.4-kb MBP-1 fragment. Lane 3 shows the result obtained after melting of the transcript-MBP-1 hybrid. Lane 4 is the tRNA control.

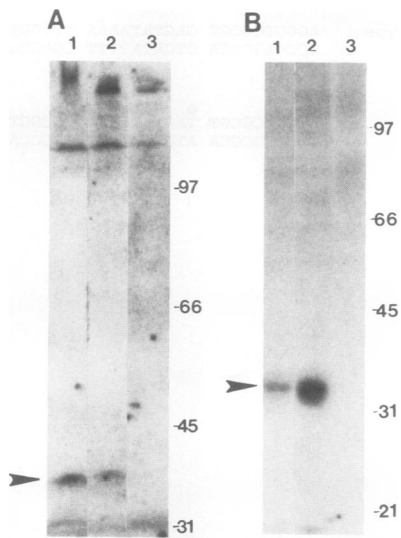


FIG. 6. Southwestern blot analysis of protein encoded by the MBP-1 clone and by HL-60, HeLa, and CV1 cells. (A) Total protein extracts from induced cultures of the MBP-1 clone and vector-transformed bacteria (BS) were analyzed by SDS-polyacrylamide gel electrophoresis and transferred onto nitrocellulose. The filter was probed with the 142-bp *c-myc* fragment (10^6 cpm/ml). Lanes 1 and 2 contained MBP-1 recombinant protein induced with 10 and 1 mM IPTG, respectively. Lane 3 contained vector-transformed (BS) protein as a negative control. (B) HL-60, HeLa, and CV1 cells were grown as 70% confluent cultures and analyzed similarly. The positions of molecular weight markers (in kilodaltons) are shown on the right.

EMBL/GenBank data base suggested an overall 70% homology with human α -enolase at the protein level (17). However, enzymatic assay of enolase activity in the crude extract of recombinant (MBP-1) and vector-transformed (negative control) bacteria failed to show any difference in activities.

Transcriptional repression by MBP-1. To study the functional role of MBP-1, DNA-mediated gene transfer experiments were carried out. MBP-1 was cloned into the pSV₂neo vector under control of the SV40 early promoter and cotransfected with a reporter CAT gene under the control of the human *c-myc* promoter (52). Cotransfection of pSV₂ MBP-1 with H-*myc* CAT or S-*myc* CAT resulted in a significant reduction of CAT activity (Fig. 8, lanes 2 and 4) relative to the vector control (lanes 1 and 3). When antisense MBP-1 (pSV₂anti-MBP-1) (5 μ g) was cotransfected with H-*myc* CAT, there was no inhibition of CAT activity (lane 5). Averaging the results from four independent experiments yielded a sixfold reduction of *c-myc* promoter relative activity by cotransfection of the MBP-1-containing plasmid. However, cotransfection of pSV₂MBP-1 or pSV₂neo with a chicken β -actin promoter plasmid (p β Act) did not alter the CAT activity (data not shown). This result suggests that MBP-1 in vivo acts as a sequence-specific repressor.

Hybridization analysis with the cDNA segment of MBP-1. Expression of the MBP-1 gene was studied by Northern blot analysis (Fig. 9). A single 1.4-kb transcript was obtained with cytoplasmic RNA from HL-60 and HeLa cells. The cDNA insert of the recombinant clone appears to be a full-length sequence, since the open reading frame and the transcript size are nearly identical. The filter was reprobbed with tubulin and found to contain similar amounts of RNA in both lanes (data not shown).

```

1  AGGAATTCGGAAATTCGGAAATTCGGATGGATGGAACAGAAAATAAATCTAAGTTTGGTG
    *
    M E Q K I N L S L V
61  CGAAGCCATTCTGGGGGTGTCCTTGGCCGTGCAAAGCTGGTGGCCGTGAGAGGGGG
    R T P F W G C P L P S A K L V P L R R G
121 TCCCCTGTACCGCCACATCGGCTACTTGGCTGGCAACTCGAAGTCATCTGCCAGTCCC
    S P V P P H R V L G W Q L R S H P A S P
181 GCGTTCGAAGTGCATCATCAATGGCGGTCTCATGTGCGCAACAGCTGGCCATGCAGA
    G V Q V S S S M A V L M L A T S W P C R
241 GTCTGTCTCCAGTCGGTGCAGCAAACTCAGGGAAGCCATGCCGCATGGAGCAGAGGT
    V C P P S R C S K L R E A M P H W S R G
301 TTACCACAACCTGAAGAATGTCATCAAGGAGAAATATGGGAAAGATGCCACCAATGTGGG
    L P Q E E C H Q G E I W E R C H C T G G
361 GATTGCGCGGGTTTCTCCCAACATCTGGAGAATAAAGAAGCCCTGGAGCTGCTGAAG
    D L R G F A P N I L E N K E G L E L L K
421 ACTGCTATTGGAAGCCTGGCTCACTGTAAGGTGGTTCATGGCATGGAGTAGCGGCC
    T A I G K P G L H C K G G H G M D V A A
481 TCCGATCTCTCAGGTCAGGGAAGTACCTGGACTTCAAGTCTCCCGATGACCCCGCC
    S E F F R S G N Y D L D F K S P D D P S
541 AGGTACATCTCGCTGACCAGCTGGCTGACCTGTACAAGTCCCTCATCAAGACTACCCA
    R Y I S P D Q L A D L Y K S F I K D Y P
601 GTGGTCTATCGAAGATCCCTTGGACAGGATGACTGGGAGSTCAGAAGTTCACAGC
    V V S I E D P F D Q D D W F G A S E V H S
661 CAGTCAGGAATCCAGGTAGTGGGGGATGACTCAGTACAGTACCAACCCAAAGAGGATCGC
    Q C R N P G S G G M T H S D Q P K E D R
721 CAAGCGGTGAACGAGAAGTCCGCAACTGCCTCCTGCTCAAAGTCAACAGATTGGCTCC
    Q G V N E K S C N C L L L K V N Q I G S
781 GTGACCGAGTCTCTTTCAGGCGTCAAGTGGCCAGGCAATGGTTGGGGCGTCATGGTG
    V T E S L Q A C K L A Q A N G W G V M V
841 TCTCATGTTCCGGGGAGACTGAAGATACCTTTCATCGTACCTGGTTGTGGGGCTGTGC
    S H R S G E T E D T F I A D L V V G L G
901 ACTGGGGCAGATCAAGACTGGTGGCCCTTGGCCATCAGCGCTTGGCCAAAGTACAACAG
    T G A D Q D W C P L P I T R L A K Y N Q
961 CTCTCAGAATTGAAGAGGAGTGGGCAGCAAGGCTAAGTTTGGCCGAGCAAGTTCAGA
    L L R I E E E L G S K A K F A G R N F R
1021 AACCCCTTGGCCAAAGTAAAGCTGTGGGCAGGCAAGCCTTCGGTCACTGTTGGTACAGAC
    N P L A K *
1081 CCCTCCCTGGTGTGCTCAGCTCAGGAGCTCGAGGCCCGCCGACCAACTGCAGGGTCCC
1141 TGCTAGTTAGCGCCACCCTGGAGTTCGTACCGCTTCCCTAGAACTCTACAGAAGCC
1201 AAGCTCCCTGGAAGCCCTGTTGGCAGCTTAGCTTTCAGTGTGTAATGGCCCAAGTC
1261 ATTGTTTTCGCTTACTTCCACCAAGTGTCTAGAGTATGTGAGCCCTGTGTCATC
1321 TCCGGGTGGCCACAGGCTAGATCCCGGTGGTTTGTGCTCAAAATAAAGCCCTCAGT
1381 GACCATGAAAAAAGGAATTCGGAAATTCGG

```

FIG. 7. Nucleotide and predicted amino acid sequences of MBP-1. The first methionine of MBP-1 is marked with an asterisk. The possible glycosylation site containing the Asn-X-Ser sequence is underlined. The stop codon is marked with a double underline, and a poly (A) stretch is marked by a wavy underline.

DISCUSSION

We report the identification of a *c-myc* P₂ promoter-binding protein (MBP-1) cloned from a HeLa cell cDNA library in λ zapII expression vector. This clone was found to have a large open reading frame of 1,007 bp translating to a polypeptide of 37 kDa. The recombinant protein, MBP-1, binds specifically upstream to the human *c-myc* P₂ transcription start site (-38 to -8 relative to the P₂ transcription start site) and presumably plays a role in transcription regulation of the *c-myc* gene. However, the precise role of this protein in human *c-myc* gene regulation remains to be determined.

MBP-1 protein binds to sequences at the human *c-myc* P₂ promoter that include the TATA box rather than to regions further upstream. The binding site sequence is not present in the 5'-flanking region of closely related genes, such as *N-myc* or *L-myc*. Transcription factor TFIID has also been shown to bind to the TATA box sequences (21, 30, 51); however, sequence analysis confirmed that MBP-1 is quite distinct from TFIID. There are very few reports of proteins, other than RNA polymerase II (8), which bind to the

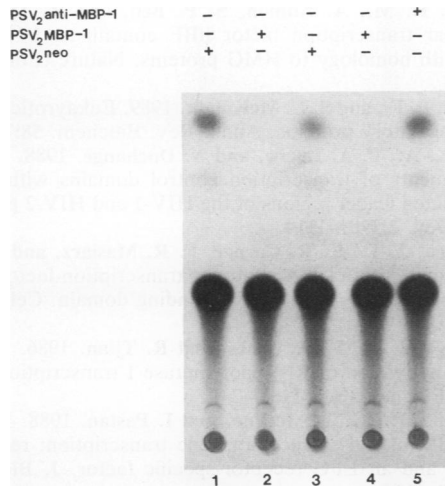


FIG. 8. Effect of MBP-1 on human *c-myc* expression in CV1 cells. A reporter construct was cotransfected with effector plasmids or the control vector pSV₂neo by the lipofection transfection procedure. Lanes: 1 and 3, effects of pSV₂neo on the H-*myc* CAT and S-*myc* CAT reporter plasmids, respectively; 2 and 4, effects of the pSV₂MBP-1 effector plasmid on the H-*myc* CAT and S-*myc* CAT reporter plasmids, respectively; 5, effect of the pSV₂anti-MBP-1 effector plasmid on the H-*myc* CAT reporter plasmid.

transcription start site. Although TFIIB binds in the vicinity of the initiation site, it does not appear to bind DNA directly (42), but rather interacts with other proteins already bound to the promoter region. On the other hand, a protein isolated from the HeLa cell nuclear extract, the dihydrofolate reductase initiator element termed housekeeping initiation protein 1, binds to the transcription initiation site and acts as a positive regulatory element (42). Similar DNA-binding proteins have been identified in other systems and have been shown to play a regulatory role in transcription. The herpes simplex virus ICP4 protein binds to its own transcription initiation site in a negative, autoregulatory manner (47). A cellular protein, designated leader binding protein 1, also spans the sequence from nucleotide -35 to +27 of the human immunodeficiency virus type 1 promoter. Like the

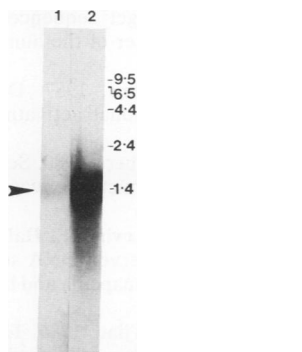


FIG. 9. Northern blot analysis of MBP-1 expression. Cytoplasmic RNA was prepared from HL-60 and HeLa cells and analyzed in a 1.2% formaldehyde-agarose gel. A Northern blot was probed with 1.4 kb of MBP-1 cDNA. Lanes 1 and 2 represent 20 μ g of cytoplasmic RNA from HL-60 and HeLa cells, respectively. The positions of RNA molecular weight markers are indicated on right in kilobases.

human *c-myc* promoters, the human immunodeficiency virus type 1 TATA element appears to be responsible for positioning transcription initiation (24).

The sequence of MBP-1 does not contain any known DNA-binding motifs such as a zinc finger structure (25, 43), leucine zipper (32), POU-specific regions (11, 15, 64-66), or homeoboxes (33, 40, 42). Although there is some similarity between MBP-1 and human α -enolase, the amino-terminal sequences are quite different and human α -enolase has not been shown to be a sequence-specific DNA-binding protein. It has been shown that the basic amino-terminal domains are a common site of homology among DNA-binding proteins. The MBP-1 DNA-binding site does not have any significant sequence homology to binding sites of the DNA-binding proteins AP1, AP3, ATF, USF, CTF, E2F, and EBP20, the octabox transcription factors SP1, 5'Mf1, 5'Mg1, 5'Mg2, 5'Mg3, ME1a1, and ME1a2, or nuclear transcription factor hUBF (2, 4, 22, 23, 25, 26, 31, 45, 46). Proline-rich regions, also known as alpha-helix breakers, have been shown in other mammalian transcription factors, including AP-2, Jun, Oct-2, and SRF (45). A proline-rich sequence is also present in the p53 protein (35) and has been predicted to be a transactivating domain. Interestingly, a proline-rich domain is present in the amino terminus of MBP-1 and may play a similar role. Further study is necessary to map the DNA-binding domain(s) of MBP-1 by in vitro mutagenesis.

Several groups have investigated the sequence requirement for initiation at the P₁ and P₂ start sites of the human *c-myc* gene (10, 19, 37, 48, 59). Sequences between -60 and -37 relative to the human *c-myc* P₁ start site are essential for its activity in microinjected *Xenopus laevis* oocytes (48). Sequences between -101 and -353 relative to P₁ have been reported to possess both positive (37) and negative (19) effector activity on the human *c-myc* P₁ and P₂ start sites. A 123-bp sequence upstream of the murine *c-myc* P₂ start site (-7 to +116 relative to P₁) was shown to be necessary for its activity and implicated the involvement of ME1a1 and ME1a2 proteins in P₂ regulation (2). Transcription of the human *c-myc* P₂ start site was also shown to require about 100 bp of the 5'-flanking sequence (37). Deletion of the ME1a1-binding site affected murine *c-myc* transcription by reducing activity of the P₂ promoter, suggesting that the P₂ promoter is inherently stronger than P₁ and may normally interfere with P₁ activity (2).

Since MBP-1 binds just upstream of the initiation start site of the human *c-myc* P₂ promoter, it may play an important regulatory function in the control of *c-myc* expression. Results from the cotransfection experiment suggest that MBP-1 may play a role as a negative regulatory factor in human *c-myc* expression. It is possible that MBP-1 competes with TFIID or other transcription factors for specific DNA binding and represses the transcription activity of human *c-myc* gene. We plan to purify the MBP-1 protein and determine its precise role in human *c-myc* gene expression.

ACKNOWLEDGMENTS

We thank Jeff Holt for the Myc-CAT plasmids, Bruce Paterson and Junita Eldridge for the actin-CAT plasmid, and Allen Naftilan for technical advice concerning the CAT assay. We also thank Shelia Thomas for technical assistance and Barbara Wilson for preparing the manuscript.

This work was supported by grants from NIH (RO1 CA42664 and RO1 CA42337) and the VA Medical Research Service. Donald M. Miller is a Clinical Investigator in the VA Career Development Program.

REFERENCES

1. Ariga, H., Y. Imamura, and S. M. M. Iguchi-Ariga. 1989. DNA replication origin and transcriptional enhancer in *c-myc* gene share the *c-myc* protein binding sequences. *EMBO J.* **8**:4273-4279.
2. Asselin, C., A. Nepveu, and K. B. Marcu. 1989. Molecular requirements for transcription initiation of the murine *c-myc* gene. *Oncogene* **4**:549-558.
3. Battey, J., C. Moulding, R. Tabu, W. Murphy, T. Stewart, M. Potter, G. Lenoir, and P. Leder. 1983. The human *c-myc* oncogene: structural consequences of translocation resulting from chromosome translocation in B-lymphoid tumors. *EMBO J.* **2**:2375-2383.
4. Benerza, 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.
5. Bentley, D. L., and M. Groudine. 1986. A block to elongation is largely responsible for decreased transcription of *c-myc* in differentiation HL-60 cells. *Nature (London)* **321**:702-706.
6. Bishop, J. M. 1987. The molecular genetics of cancer. *Science* **235**:305-311.
7. Breathnach, R., and P. Chambon. 1981. Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**:349-383.
8. Buratowski, S., S. Hahn, L. Guarente, and P. A. Sharp. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**:549-561.
9. Campisi, J., H. E. Gray, A. B. Pardee, M. Dean, and G. E. Sonenshein. 1984. Cell-cycle control of *c-myc* but not *c-ras* expression is lost following chemical transformation. *Cell* **37**:241-247.
10. Chung, J., E. Shin, R. R. Reed, and P. Leder. 1986. Transacting elements modulate expression of the human *c-myc* gene in Burkitt lymphoma cells. *Proc. Natl. Acad. Sci. USA* **83**:7918-7922.
11. 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.
12. Cole, M. 1986. The *myc* oncogene: its role in transformation and differentiation. *Annu. Rev. Genet.* **20**:361-384.
13. Dynan, W. S., and R. Tjian. 1983. The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* **35**:79-87.
14. Dynan, W. S., and R. Tjian. 1985. Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. *Nature (London)* **316**:774-778.
15. Finney, M., G. Ruvkun, and H. R. Horvitz. 1988. The *C. elegans* cell lineage and differentiation gene *unc-86* encodes a protein with a homeodomain and extended similarity to transcription factors. *Cell* **55**:757-769.
16. Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**:1044-1051.
17. Giallonga, A., F. Salvatore, R. Moore, C. M. Croce, and L. C. Showe. 1986. Molecular cloning and nucleotide sequence of a full-length cDNA for human α enolase. *Proc. Natl. Acad. Sci. USA* **83**:6741-6745.
18. Greenberg, M. E., L. A. Green, and E. B. Ziff. 1985. Nerve growth factor and epidermal growth factor induce rapid transient changes in protooncogene transcription in PC12 cells. *J. Biol. Chem.* **260**:14101-14111.
19. Hay, N., J. M. Bishop, and D. Levens. 1987. Regulatory elements that modulate expression of human *c-myc*. *Genes Dev.* **1**:659-671.
20. Hay, N., M. Takimoto, and J. M. Bishop. 1989. A *fos* protein is present in a complex that binds a negative regulator of *myc*. *Genes Dev.* **3**:293-303.
21. Hoffman, A., E. Sinn, T. Yamamoto, J. Wang, A. Roy, M. Horikoshi, and R. G. Rodere. 1990. Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFID). *Nature (London)* **346**:387-390.
22. Jantzen, H. M., A. Admon, S. P. Bell, and R. Tjian. 1990. Nucleolar transcription factor hBF contains a DNA-binding motif with homology to HMG proteins. *Nature (London)* **344**:830-836.
23. Johnson, P. F., and S. L. McKnight. 1989. Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem.* **58**:799-839.
24. Jones, K. A., P. A. Luciw, and N. Duchange. 1988. Structural arrangements of transcription control domains within the 5'-untranslated leader regions of the HIV-1 and HIV-2 promoters. *Genes Dev.* **2**:1101-1114.
25. Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. Tjian. 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* **51**:1079-1090.
26. Kadonaga, J. T., K. A. Jones, and R. Tjian. 1986. Promoter-specific activation of RNA polymerase I transcription by Sp1. *Trends Biochem. Sci.* **11**:20-23.
27. Kageyama, R., G. T. Merlino, and I. Pastan. 1988. Epidermal growth factor (EGF) receptor gene transcription: requirement for Sp1 and an EGF receptor-specific factor. *J. Biol. Chem.* **263**:6329-6336.
28. Kageyama, R., G. T. Merlino, and I. Pastan. 1988. A transcription factor active on the epidermal growth factor receptor gene. *Proc. Natl. Acad. Sci. USA* **85**:5016-5020.
29. Kageyama, R., and I. Pastan. 1989. Molecular cloning and characterization of a human DNA binding factor that represses transcription. *Cell* **59**:815-825.
30. Kao, C. C., P. M. Lieberman, M. C. Schmidt, Q. Zhou, R. Pie, and A. J. Berk. 1990. Cloning of a transcriptionally active human TATA binding factor. *Science* **248**:1646-1650.
31. Landschulz, W. H., P. F. Johnson, E. Y. Adashi, B. J. Graves, and S. L. McKnight. 1988. Isolation of a recombinant copy of the gene encoding C/EBP. *Genes Dev.* **2**:786-800.
32. Landschulz, W. H., P. F. Johnson, and S. L. McKnight. 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240**:1759-1764.
33. Laughon, A., and M. P. Scott. 1984. Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. *Nature (London)* **310**:25-31.
34. Lenardo, M. J., L. Staudt, P. Robbins, A. Kuang, R. C. Mulligan, and D. Baltimore. 1989. Repression of the IgH enhancer in teratocarcinoma cells associated with a novel octamer factor. *Science* **243**:544-546.
35. Levine, A. J. 1990. The p53 protein and its interactions with the oncogene products of the small DNA tumor viruses. *Virology* **177**:419-426.
36. Levine, R. A., J. E. McCormack, A. Buckler, and G. E. Sonenshein. 1986. Transcriptional and posttranscriptional control of *c-myc* gene expression in WEHI 231 cells. *Mol. Cell. Biol.* **6**:4112-4116.
37. Lipp, M., R. Schilling, S. Wiest, G. Laux, and G. W. Bornkamm. 1987. Target sequences for *cis*-acting regulation within the dual promoter of the human *c-myc* gene. *Mol. Cell. Biol.* **7**:1393-1400.
38. Ma, J., and M. Ptashne. 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* **48**:847-853.
39. Maxam, A., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499-560.
40. McGinnis, W., M. S. Levine, E. Hafen, A. Kuroiwa, and W. J. Gehring. 1984. A conserved DNA sequence in homeotic genes of the *Drosophila* antennapedia and bithorax complexes. *Nature (London)* **308**:428-433.
41. McKnight, S., and R. Tjian. 1986. Transcriptional selectivity of viral genes in mammalian cells. *Cell* **46**:795-805.
42. Means, A. L., and P. J. Farnham. 1990. Transcription initiation from the dihydrofolate reductase promoter is positioned by HIP1 binding at the initiation site. *Mol. Cell. Biol.* **10**:653-661.
43. Miller, J., A. D. McLachlan, and A. Klug. 1985. Repetitive zinc-binding domains in the protein transcription factor IIA from *Xenopus* oocytes. *EMBO J.* **4**:1609-1614.
44. Miskimins, W. K., M. P. Roberts, A. McClelland, and F. H.

- Ruddle. 1985. Use of a protein-blotting procedure and a specific DNA probe to identify nuclear proteins that recognize the promoter region of the transferrin receptor gene. *Proc. Natl. Acad. Sci. USA* **82**:6741-6744.
45. Mitchell, P. J., and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**:371-378.
 46. Muller, M. M., S. Ruppert, W. Schaffner, and P. Matthias. 1988. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature (London)* **336**:544-551.
 47. Muller, M. T. 1987. Binding of the herpes simplex virus immediate-early gene product ICP4 to its own transcription start site. *J. Virol.* **61**:858-865.
 48. Nishikura, K. 1986. Sequences involved in accurate and efficient transcription of human *c-myc* genes microinjected into frog oocytes. *Mol. Cell. Biol.* **6**:4093-4098.
 49. O'Neill, E. A., C. Fletcher, C. R. Burrow, N. Heintz, R. G. Roeder, and T. J. Kelly. 1988. Transcription factor OTF-1 is functionally identical to the DNA replication factor NFIII. *Science* **241**:1210-1213.
 50. Paterson, B. M., B. E. Roberts, and E. Kuff. 1977. Structural gene identification and mapping by DNA-mRNA hybrid arrested cell free translation. *Proc. Natl. Acad. Sci. USA* **74**:4370-4374.
 51. Peterson, M. G., N. Tanese, F. Pugh, and R. Tjian. 1990. Functional domains and upstream activation properties of cloned human TATA binding protein. *Science* **248**:1625-1630.
 52. Pietenpol, J. A., J. T. Holt, R. W. Stein, and H. L. Moses. 1990. Transforming growth factor β 1 suppression of *c-myc* gene transcription: role in inhibition of keratinocyte proliferation. *Proc. Natl. Acad. Sci. USA* **87**:3758-3762.
 53. Postel, E. H., S. E. Mango, and S. J. Flint. 1989. A nuclease-hypersensitive element of the human *c-myc* promoter interacts with a transcription initiation factor. *Mol. Cell. Biol.* **9**:5123-5133.
 54. Ray, R., R. C. Snyder, S. Thomas, C. A. Koller, and D. M. Miller. 1989. Mithramycin blocks protein binding and function of the SV40 early promoter. *J. Clin. Invest.* **83**:2003-2007.
 55. Ray, R., S. Thomas, and D. M. Miller. 1989. Mouse fibroblasts transformed with the human *c-myc* gene express a high level of mRNA but a low level of *c-myc* protein and are non-tumorigenic in nude mice. *Oncogene* **4**:593-600.
 56. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
 57. Scheidereit, C., J. A. Cromlish, T. Gerster, K. Kawakami, C. G. Balmaceda, R. A. Currie, and R. G. Roeder. 1988. A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homeobox protein. *Nature (London)* **336**:551-557.
 58. Sen, R., and D. Baltimore. 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* **46**:705-716.
 59. Siebenlist, U., L. Hennighausen, J. Battey, and P. Leder. 1984. Chromatin structure and protein binding in the putative regulatory region of the *c-myc* gene in Burkitt lymphoma. *Cell* **37**:381-391.
 60. Singh, H., J. H. LeBowitz, A. S. Baldwin, 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**:514-523.
 61. Singh, H., R. Sen, D. Baltimore, and P. A. Sharp. 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature (London)* **319**:154-158.
 62. Staudt, L. M., H. Singh, R. Sen, T. Wirth, P. A. Sharp, and D. Baltimore. 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. *Nature* **323**:640-643.
 63. Studzinski, G. P., Z. S. Brelvi, S. C. Feldman, and R. A. Watt. 1986. Participation of *c-myc* protein in DNA synthesis of human cells. *Science* **236**:467-470.
 64. Sturm, R. A., T. Baumruker, B. R. Franza, Jr., and W. Herr. 1987. A 100kd HeLa cell octamer binding protein (OBP100) interacts differently with two separate octamer-related sequences within the SV40 enhancer. *Genes Dev.* **1**:1147-1160.
 65. Sturm, R. A., G. Das, and W. Herr. 1988. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes Dev.* **2**:1582-1599.
 66. Sturm, R. A., and W. Herr. 1988. The POU domain is a bipartite DNA binding structure. *Nature (London)* **336**:640-643.
 67. Takimoto, M., J. P. Quinn, A. R. Farina, L. M. Staudt, and D. Levens. 1989. Fos/jun and octamer-binding protein interact with a common site in a negative element of the human *c-myc* gene. *J. Biol. Chem.* **264**:8992-8999.
 68. 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.
 69. Zajac-Kaye, M., E. P. Gelmann, and D. Levens. 1988. A point mutation in the *c-myc* locus of a Burkitt lymphoma abolishes binding of a nuclear protein. *Science* **2**:1776-1780.