MOLECULAR AND CELLULAR BIOLOGY, Apr. 1997, p. 2116–2126 Vol. 17, No. 4
0270-7306/97/$04.00+0
Copyright © 1997, American Society for Microbiology

# Human and Mouse MOK2 Proteins Are Associated with Nuclear Ribonucleoprotein Components and Bind Specifically to RNA and DNA through Their Zinc Finger Domains

VALÉRIE ARRANZ, FRANCIS HARPER, YVETTE FLORENTIN, EDMOND PUVION, MICHEL KRESS, AND MICHÈLE ERNOULT-LANGE*

*GMIFC CNRS-UPR9044, Institut de recherche sur le cancer, 94801 Villejuif Cedex, France*

Received 15 October 1996/Returned for modification 26 November 1996/Accepted 23 January 1997

**The human and murine *MOK2* ortholog genes that are preferentially expressed in brain and testis tissues encode two different Krüppel-like zinc finger proteins. In this paper, we show that the MOK2 proteins are mainly associated with nuclear ribonucleoprotein components, including the nucleoli and extranucleolar structures, and exhibit specific RNA homopolymer binding activities. Moreover, we have identified an identical 18-bp specific DNA binding sequence for both MOK2 proteins using a pool of random sequence oligonucleotides. The DNA binding domain is localized in the seven adjacent zinc finger motifs, which show 94% identity between human and murine proteins. Taken together, these results establish that the MOK2 proteins are able to recognize both DNA and RNA through their zinc fingers. This dual affinity and the subnuclear localization suggest that MOK2 may play roles in transcription, as well as in the posttranscriptional regulation processes of specific genes.**

The $C_2H_2$ zinc finger motif is unique in that it functions as both a specific RNA-binding module and a specific DNA-binding module. The prototype zinc finger protein is transcription factor IIIA (TFIIIA) from *Xenopus laevis*. TFIIIA binds to the internal control region of the 5S rRNA gene (14), but it is also capable of forming an alternative highly specific complex with the 5S rRNA itself (27, 46). A large zinc finger multigene family present in the genomes of different vertebrate organisms and made up of probably several hundred different genes has been characterized on the basis of structural similarity with the finger motif (13). A few mammalian $C_2H_2$ proteins have been characterized in detail, and all are transcriptional activators and/or repressors (5, 8, 10, 22, 28, 33, 54, 59, 65). These proteins participate in the regulation of cellular functions, such as embryonic development, cell proliferation, and differentiation. Several $C_2H_2$ proteins in *X. laevis* have been characterized as specific RNA-binding proteins. The p43 protein, which is structurally closely related to TFIIIA, is also a constituent of the cytoplasmic 5S rRNA storage particle in immature *Xenopus* oocytes (30). This protein shows no affinity for the internal control region of the 5S rRNA gene. The others (XFG 5-1, XFO 6, XFO 9-3, and Xfin) exhibit specific RNA homopolymer binding activities in vitro (1, 34, 35). Many clones encoding $C_2H_2$ proteins have not been fully characterized with respect to their binding sites and functional activities. At present, there exists no means of predicting a zinc finger protein's preference for binding to RNA or DNA.

The human and murine *MOK2* ortholog genes that are preferentially expressed in brain and testis tissues encode two different Krüppel-TFIIIA-related zinc finger proteins. The human MOK2 protein (hsMOK2) shows substantial differences from the murine MOK2 protein. The mouse MOK2 protein consists of seven tandem zinc finger motifs, with only five additional amino acids at the COOH-terminal end (16). The seven fingers motifs are highly similar to one another but are distinct from those of other zinc finger proteins. The structural feature of the murine MOK2 protein is present at the end of hsMOK2. Furthermore, the human protein contains three additional zinc finger motifs in tandem with the others and a nonfinger acidic domain of 173 amino acids at the $NH_2$-terminal end (15). We have shown that human *MOK2* RNA maturation results in three mRNAs with different 5′-untranslated exons. One of these three mRNAs encodes a smaller protein. This protein comprises 10 zinc finger motifs and a smaller $NH_2$-acidic domain made up of 76 amino acids. The human gene has been localized to band q13.2-q13.3 of chromosome 19, and the murine gene has been localized on chromosome 6 (2, 15).

As an approach to determine the biological functions of the human and mouse MOK2 proteins, we have determined their subcellular localizations by electron microscopy. These proteins are shown to be associated with nuclear ribonucleoprotein (RNP) components. In addition, we have demonstrated that the murine and human MOK2 proteins share specific RNA- and DNA-binding activities.

## MATERIALS AND METHODS

**Cells, viruses, and transient transfections.** L cells were grown in Dulbecco's modified Eagle's minimal essential medium (DMEM) supplemented with 10% fetal calf serum (FCS). HeLa cells were grown in DMEM with 10% newborn calf serum. *Autographa californica* nuclear polyhedrosis virus (AcNPV) and its derivative AcMOK2 were propagated in *Spodoptera frugiperda* cells (Sf9) at 28°C with TC100 medium (Gibco-Bethesda Research Laboratories) supplemented with 10% fetal calf serum (FCS). AcNPV viral DNA was isolated as described by Summers and Smith (60).

For the transient transfection assay, the cells were plated at $10^6$ cells per 100-mm petri dish 24 h prior to the addition of the recombinant plasmids by the calcium phosphate coprecipitation technique (17). Four hours before transfection, the medium was changed to DMEM containing 2.5% newborn calf serum and 2.5% FCS. The cells were incubated for 16 h at 37°C with a total of 30 μg of plasmid DNA per petri dish. They were then washed with phosphate-buffered saline (PBS) and incubated for 48 h in DMEM containing 2.5% newborn calf serum and 2.5% FCS.

**Recombinant vector constructions.** The recombinant baculovirus AcMOK2 contains the whole coding sequence of the murine *Mok2* gene (16). The 603-bp *Dde*I fragment, blunt ended with the Klenow fragment, was subcloned into the *Bam*HI/Klenow fragment-treated pAC701 vector (38). The transfer of the re-

* Corresponding author. GMIFC CNRS-UPR9044, Institut de recherche sur le cancer, 7 rue Guy Moquet, 94801 Villejuif Cedex, France.

sulting plasmid to the AcNPV viral genome and plaque purification of recombinant virus were performed according to the procedure described by Summers and Smith (60). To construct the β-galactosidase (β-Gal)–MOK2 and maltose-binding protein (MBP)–MOK2 fusion proteins, the blunt-ended *Dde*I fragment was inserted into the *Hin*dIII/Klenow fragment-treated pur292 β-Gal vector (56) and *Eco*RI/Klenow fragment-treated pMAL-P2 vector (New England Biolabs), respectively. The recombinant fusion proteins β-Gal-hsMOK2 and MBP-hsMOK2 were obtained by insertion of the blunt-ended *Aat*II-*Ssp*I fragment (1,152 bp) into *Bam*HI/Klenow fragment-treated pur292 β-Gal vector and *Xmn*I-digested pMAL-P2 vector, respectively. The *Aat*II-*Ssp*I fragment isolated from human *MOK2* cDNA contains the complete zinc finger domain and the 81 amino acids of the $NH_2$-nonfinger domain (15). The prokaryote expression vector pAX-XFG5-1 contains the complete *Xenopus* XFG 5-1 cDNA encoding the β-Gal–XFG 5-1 fusion protein (35).

The eukaryotic expression vector cytomegalovirus (CMV)-MOK2 contains the whole coding sequence and a part of the 3′ noncoding region of the murine *Mok2* gene. The *Sfa*NI-*Aat*II fragment (879 bp) was isolated from *Mok2* cDNA, blunt ended with the Klenow fragment, and ligated to *Not*I/Klenow fragment-treated pCMV vector. CMV-MOK2R contained the *Sfa*NI-*Aat*II fragment in reverse orientation. The eukaryotic expression vector pCMV-hsMOK2 contained the coding sequence that begins 31 bp before the third ATG and the 3′ noncoding region of the human *MOK2* gene. The *Eco*RI-*Sal*I fragment (2,600 bp) was isolated from the genomic GhsMOKcl1 clone, blunt ended with the Klenow fragment, and ligated to *Not*I/Klenow fragment-treated pCMV vector.

**Affinity purification of anti-MOK2 antibodies.** The β-Gal–MOK2 fusion protein was resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), purified, and injected into a New Zealand White rabbit. Affinity-purified MOK2 antibodies were obtained by elution of immunoglobulins bound to the mouse MOK2 protein expressed by the recombinant baculovirus Ac-MOK2. In brief, the MOK2 protein was resolved by SDS-PAGE and transfered to an Immobilon P filter (Millipore). The strip of Immobilon P filter that carries the MOK2 protein was incubated with the polyclonal antibody for 16 h at 4°C. After an extensive washing step, immunoglobulins bound to the protein were recovered by brief treatment with 0.1 M glycine (pH 2.8) followed by rapid neutralization with 0.1 volume of 1 M Tris-HCl (pH 8.0). The antibody was stored at 4°C after addition of 5 mg of bovine serum albumin per ml (57).

**Electron microscopy.** Cells were fixed in situ at 4°C in 4% formaldehyde (Merck, Darmstadt, Germany) in 0.1 M Sörensen phosphate buffer (pH 7.3). During the first h of fixation, the cells were scraped from the culture dish and centrifuged. The pellets were dehydrated by using increasing concentrations of methanol and were embedded in Lowicryl $K_4M$ (Chemische Werke Lowi, Waldkraibing, Germany). Polymerization was carried out for 5 days at −30°C under long-wavelength UV light. Ultrathin sections were collected on Formvar-carbon-coated gold grids (200 mesh). For the in situ immunodetection of structures containing proteins, grids bearing sections were incubated first in 5% bovine serum albumin and then for 30 min at room temperature on a 10-μl drop of affinity-purified MOK2 antibodies diluted 1/10 in PBS. After 30 min of washing in PBS, the grids were floated for 30 min on a 10-μl drop of a 1/50 dilution of goat anti-rabbit immunoglobulin G conjugated to 10-nm gold particles (Biocell Research Laboratories, Cardiff, United Kingdom). After a final PBS wash, the grids were rapidly rinsed with distilled water, air dried, and treated with the staining procedure described by Bernhard to reveal RNP structures (4).

For specific DNA staining, ultrathin sections of Lowicryl-embedded cells were mounted on naked gold grids and submitted to the Feulgen-like technique (9), which includes a 25-min treatment of sections with 5 N HCl at room temperature. After extensive washing, the grids were floated on osmium-amine-$SO_2$ for about 1 h at 37°C, rinsed with distilled water, and air dried. This staining was followed by immunolabeling as described above.

Enzymatic digestions were carried out on ultrathin sections of Lowicryl-embedded cells mounted on Formvar-coated gold grids. Double digestions were performed by floating the grids on a drop of protease (Bacterial protease type VI; Sigma Chemical, St. Louis, Mo.) at 0.5 mg/ml in distilled water for 10 min at 37°C, followed by either DNase I (Worthington Biochemical, Freehold, N.J.) or RNase A (BDH Biochemical, Poole, United Kingdom) treatment at 1 mg/ml in 10 mM Tris-HCl (pH 7.3)–5 mM $MgCl_2$ for 1 h at 37°C. The RNase A reaction was performed without $MgCl_2$.

**RNA binding analysis.** For the production of β-Gal fusion proteins, transformed cells were grown in 50 ml of Luria-Bertani (LB) medium at 37°C to an optical density at 600 nm of 0.6. Fusion protein synthesis was induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 4 h. The cells were then harvested by centrifugation and disrupted by sonication in 2.5 ml of RNA extraction buffer (50 mM HEPES [pH 7], 150 mM NaCl, 0.05 mM $ZnCl_2$, 2 mM $MgCl_2$, 10% glycerol, 100 mM β-mercaptoethanol). After centrifugation at 10,000 rpm (Sorvall HB6 rotor) for 1 h, the cleared supernatant was used as a crude protein preparation for RNA binding analysis and was stored at −70°C. The protein concentrations were determined by the Coomassie blue protein assay (Pierce), and the fusion proteins were analyzed by SDS-PAGE.

The RNA binding analysis was performed as described by Nietfield et al. (42). In brief, β-Gal fusion proteins in crude protein preparations were further purified by immunoprecipitation with a mouse monoclonal antiserum raised against β-Gal (Promega) (5 μl per assay). Protein A-Sepharose beads (7.5 mg per assay) with immunoadsorbed proteins were incubated for 30 min at room temperature with $2 \times 10^5$ cpm (about 20 ng) of radiolabeled RNA homopolymer in a total volume of 600-μl binding buffer (30 mM Tris-HCl [pH 8], 1 mM EDTA, 15 mM β-mercaptoethanol, 100 mM NaCl). After four washings with binding buffer, the bound radioactive material was quantified in a scintillation counter.

RNA homopolymers (Sigma) were 5′ end labeled with [$\gamma$-$^{32}$P]ATP after dephosphorylation and were separated from unincorporated label as described previously (35).

**DNA binding site selection.** For the selection of DNA consensus binding sequences, we used the random oligonucleotides selection as described previously (48). For this, a 75-bp oligonucleotide with a central 35-bp random sequence and 20-bp flanking sequences was used. The nucleotide sequence of the flanking regions containing *Bam*HI or *Eco*RI sites was 5′-CAGGTCAGTTCA GCGGATCC$(N)_{35}$AGTTCCACTGAATTCGCCTC-3′. Twenty-base-pair oligonucleotides, which were identical or complementary to each end, were used as a forward primer and as a reverse primer. The complementary strand of the 75-bp oligonucleotide was synthesized by DNA polymerase (Klenow fragment; Boehringer) with the reverse primer in the presence of 10 μCi of [$\alpha$-$^{32}$P]dCTP and was used as a probe for binding selection.

The crude β-Gal or MBP fusion proteins were prepared as described for RNA binding analysis, but the bacteria were disrupted by sonication in the following DNA extraction buffer: 20 mM HEPES (pH 7.9), 150 mM KCl, 0.05 mM $ZnCl_2$, 0.1% Nonidet P-40, 0.5 mM dithiothreitol, 20% glycerol, 0.5 mM phenylmethylsulfonyl fluoride. The binding reaction (25-μl mixture) was performed in DNA extraction buffer containing 100 μg of crude β-Gal fusion protein preparations, 0.2 μg of poly(dI-dC), 2 μl of mouse monoclonal antiserum raised against the β-Gal domain, 1 mg of acetylated bovine serum albumin per ml, and 1 μl of oligonucleotide probe (0.2 ng). After seven rounds of binding and amplification by PCR, *Bam*HI/*Eco*RI-digested oligonucleotides were ligated into *Eco*RI/*Bam*HI-digested pBluescript $KS^+$ (Stratagene), cloned into *Escherichia coli* DH5α cells, and sequenced by using the Pharmacia T7 sequencing kit.

**Identification of potential target genes by the whole-genome PCR technique.** Murine and human genomic DNA sequences which bind to MOK2 proteins were isolated as previously described by Kinzler and Vogelstein (32), with slight modifications. DNA binding reactions were done with the DNA extraction buffer described above with 17 μl of crude β-Gal fusion protein preparations, 2 μl of poly(dI-dC) as competitor, and 500 ng of catch-linked DNA. After incubation for 30 min at room temperature, 200 μl of DNA extraction buffer, 5 μl of mouse monoclonal antiserum raised against the β-Gal domain, and 7.5 mg of protein A-Sepharose previously hydrated and equilibrated with DNA extraction buffer were added. The incubations were carried out overnight at 4°C with end-over-end mixing. After four cycles of selection, the amplified DNA cleaved with *Eco*RI was ligated into *Eco*RI-digested pBluescript $KS^+$ vector. The 5′-end-labeled oligonucleotide 5′-RCCTTRTCAGGGCCTTTR-3′ (R = A or G) corresponding to the MOK2-binding site was used as a probe for screening the libraries.

**Gel mobility shift assays.** End-labeled oligonucleotides (0.2 ng) were incubated for 20 min at room temperature in 12 μl of DNA extraction buffer containing 2 μg of poly(dI-dC) and 2.5 μg of crude MBP fusion protein preparations or 2 μl of rabbit reticulocyte lysate translation reaction. Complexes were analyzed by electrophoresis on a nondenaturing 4 or 6% premigrated polyacrylamide gel (acrylamide/bisacrylamide ratio of 19:1) in 0.5× TB buffer (1× TB buffer is 89 mM Tris-HCl [pH 8.3] and 89 mM borate) at 200 V for 4 h at 4°C. The gel was then dried and exposed to film overnight.

**In vitro translation.** The *Pst*I DNA fragment containing the full-length *Mok2* cDNA (16) was cloned into plasmid pSP64 (Promega) cut by *Pst*I to give the pSP64-MOK2 vector. *Mok2* transcript was prepared by in vitro transcription of *Eco*RI-cleaved pSP64-MOK2 with SP6 RNA polymerase (Promega). Transcripts were translated by using nuclease-treated rabbit reticulocyte lysates (Dupont-NEN) according to the conditions recommended by the manufacturer. Reactions were performed without labeled amino acids.

**DNase I footprinting.** The 58-bp oligonucleotide containing the binding site sequence GCCTTGTCAGGGCCTTTA and the flanking regions described for DNA binding selection was cloned into *Eco*RI/*Bam*HI-digested pBluescript $KS^+$. To obtain a single end-labeled fragment, the plasmid was digested by *Not*I or *Sal*I, dephosphorylated, 5′ end labeled with [$\gamma$-$^{32}$P]ATP by using polynucleotide kinase (Boehringer), and cut by *Sal*I or *Not*I, respectively. The 151-bp *Not*I-*Sal*I fragments labeled at either end were gel purified on 6% polyacrylamide nondenaturing gels. The DNase footprinting procedure was essentially the one described by Ohlsson and Edlund (43). Binding reactions were carried out with 50 μg of MBP fusion proteins in buffer containing 25 mM HEPES (pH 7.5), 150 mM KCl, 1 mM dithiothreitol, 0.01% Nonidet P-40, 10 μM $ZnSO_4$, and 20% glycerol for 15 min at room temperature. Poly(dI-dC) was used as the nonspecific competitor. The samples were incubated with 2.5 U of DNase I (Promega) for 1 min.

## RESULTS

**Localization of murine and human MOK2 proteins by electron microscopy.** The specific affinity-purified MOK2 antibody was used to determine the subcellular localizations of MOK2 proteins in a variety of transformed mouse cells which were

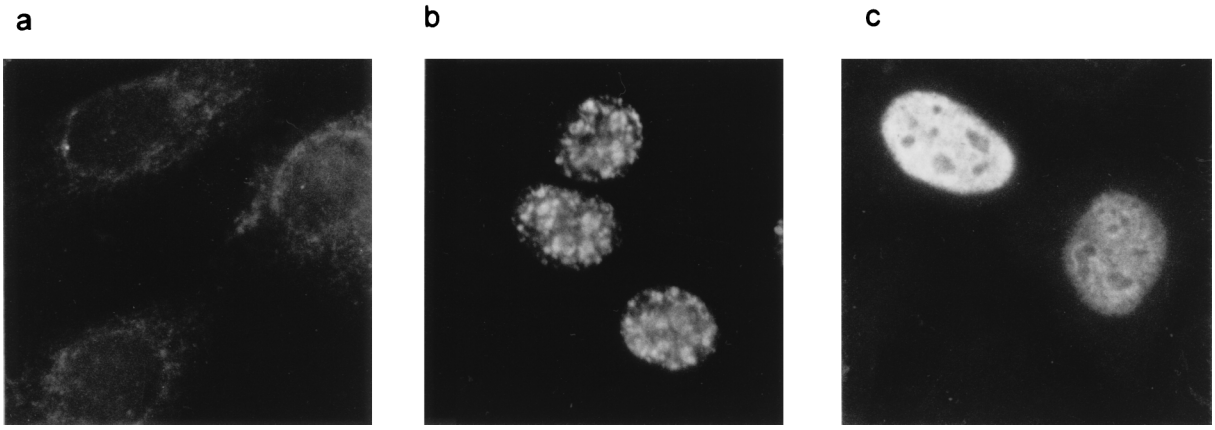


FIG. 1. Immunofluorescence on HeLa cells transfected with the reverse plasmid CMV-MOK2R (a), the coding plasmid CMV-MOK2 (b), and the coding plasmid CMV-hsMOK2 (c).

shown to express *Mok2* mRNA by indirect immunofluorescence. These attemps were unsuccessful, probably due to the very low levels of *Mok2* expression in these cells (16). To circumvent this problem, transient transfection into HeLa cells was performed with the eukaryotic expression vectors CMV-MOK2 or CMV-hsMOK2. Nuclear staining in HeLa cells transfected with the vectors expressing the murine and human proteins but not in HeLa cells transfected with the reverse plasmid CMV-MOK2R was detected (Fig. 1). Granular nuclear staining was detected with the murine MOK2 protein, whereas the staining was homogeneous with the human protein. These different immunofluorescence patterns suggested that these two proteins behaved differently when they were overexpressed in cells. In order to visualize the endogenous MOK2 proteins and to improve the subnuclear localization, we used in situ immunodetection and electron microscopy.

Human HeLa cells (Fig. 2A and B) or murine L cells (data not shown) were always weakly labeled. However, the gold particles were clearly localized in the nucleus. The combination of the immunolabeling and EDTA staining reveals the presence of MOK2 antigenic sites and RNP components, respectively. These MOK2 antigenic sites were found over the perichromatin fibrils distributed within the interchromatin space. The latter extended between clumps of condensed chromatin, which appeared bleached by this staining procedure. Gold particles were also found over some but not all interchromatin granule clusters and their associated zones (Fig. 2B). No significant labeling of condensed chromatin was observed. In addition, a few gold particles were constantly present over the nucleoli without any specificity toward fibrillar or granular components (Fig. 2A and B).

The labeling was increased by transfection of HeLa cells (Fig. 3A and B) or L cells (data not shown) with either CMV-hsMOK2 or CMV-MOK2 vector. The different behaviors of overexpressed human and murine MOK2 proteins observed by immunofluorescence were equally visible by electron microscopy, and this was independent of cell type. Transfections with CMV-hsMOK2 carried out in HeLa cells (Fig. 3A) dramatically increased the labeling without changing the general distribution of labeled nuclear RNP components. This situation concerned about 20% of the cells which actually displayed very strong accumulations of gold particles over the nucleoli and extranucleolar RNP components. This included the perichromatin fibrils, the clusters of interchromatin granules, and their associated zones. In contrast to the above observations, cells overexpressing the murine MOK2 protein exhibited strong alterations of nuclear structure. These alterations concerned 20 to 40% of transfected cells that showed a high degree of immunolabeling. This consisted of condensation of nucleoplasmic components into aggregates of variable sizes and forms composed of densely stained coiled ribbons. The formation of these heavily immunolabeled aggregates occurred concomitantly with the almost complete disappearance of perichromatin fibrils (Fig. 3B). The clusters of interchromatin granules were still present without significant morphological changes. In addition, the nucleolar components were partially segregated while being heavily labeled.

In order to verify the RNP nature of these transfection-induced masses, two series of experiments were performed. Immunolabeling of cells and application of the Feulgen-like method (9), which specifically reveals DNA, clearly confirmed that the aggregates do not contain DNA and that the chromatin was not labeled (Fig. 4). Finally, the RNP nature of the aggregates was further confirmed by their almost complete disappearence following digestions of the thin sections with protease and RNase, whereas the aggregates remained highly contrasted following protease and DNase digestions (data not shown).

**RNA homopolymer binding analysis.** On account of the subcellular association of murine and human MOK2 proteins with nuclear RNP components, we sought to investigate whether these proteins possess an RNA-binding activity. To address directly this question, we performed an in vitro binding assay using bacterially expressed fusion proteins and radiolabeled RNA homopolymers as described by Nietfeld et al. (42). The MOK2 proteins encoded by the murine and human cDNAs were expressed as a β-Gal fusion protein in *E. coli*. Fusion proteins were purified from the soluble fraction of crude extracts by immunoprecipitation with an antibody directed against the β-Gal domain. These immunoprecipitates were then incubated with different radiolabeled RNA homopolymers, and the bound fractions were quantified after extensive washing. The β-Gal protein and the β-Gal-XFG 5-1 fusion protein were used as negative and positive controls, respectively (35). The binding obtained with the β-Gal protein gave the background level which was high only with poly(G) in our experiments. The *Xenopus* XFG 5-1 fusion protein was found to bind to poly(U) and to poly(G) to the degree previously reported by Köster et al. (35) (Table 1). The murine and human MOK2 proteins were found to bind significantly to poly(U) and poly(G) homopolymers (Table 1). Furthermore, the mouse protein (β-Gal–MOK2) binds more efficiently to

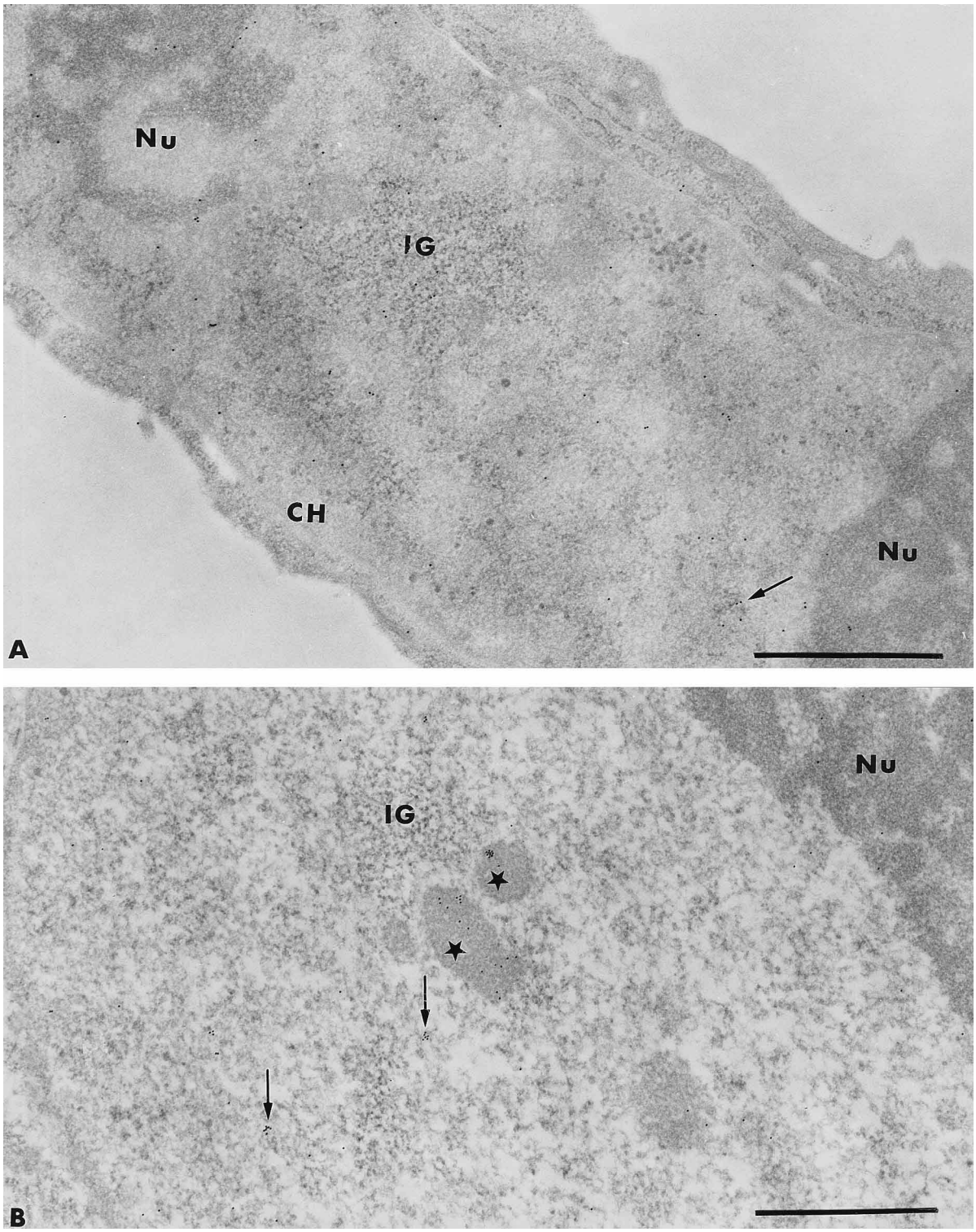


FIG. 2. Localization of the human hsMOK2 protein in HeLa cells by in situ immunodetection. (A) Gold particles observed over perichromatin fibrils (arrows). Interchromatin granules (IG) and the nucleoli (Nu) were lightly labeled. Chromatin clumps (CH) were bleached following EDTA staining. (B) Interchromatin granule-associated zones (indicated by stars). Scale bars, 1 μm.

poly(U) and poly(G) than the human protein (β-Gal–hsMOK2). These two fusion proteins did not bind to poly(A) or poly(C). Similar levels of poly(U) and poly(G) binding were obtained when we varied the salt concentration between 50 and 200 mM NaCl (data not shown). Increasing the salt concentration in the binding reaction mixtures to higher than 300 mM NaCl reduced binding to background levels.

The specificity of these interactions was tested in the pres-

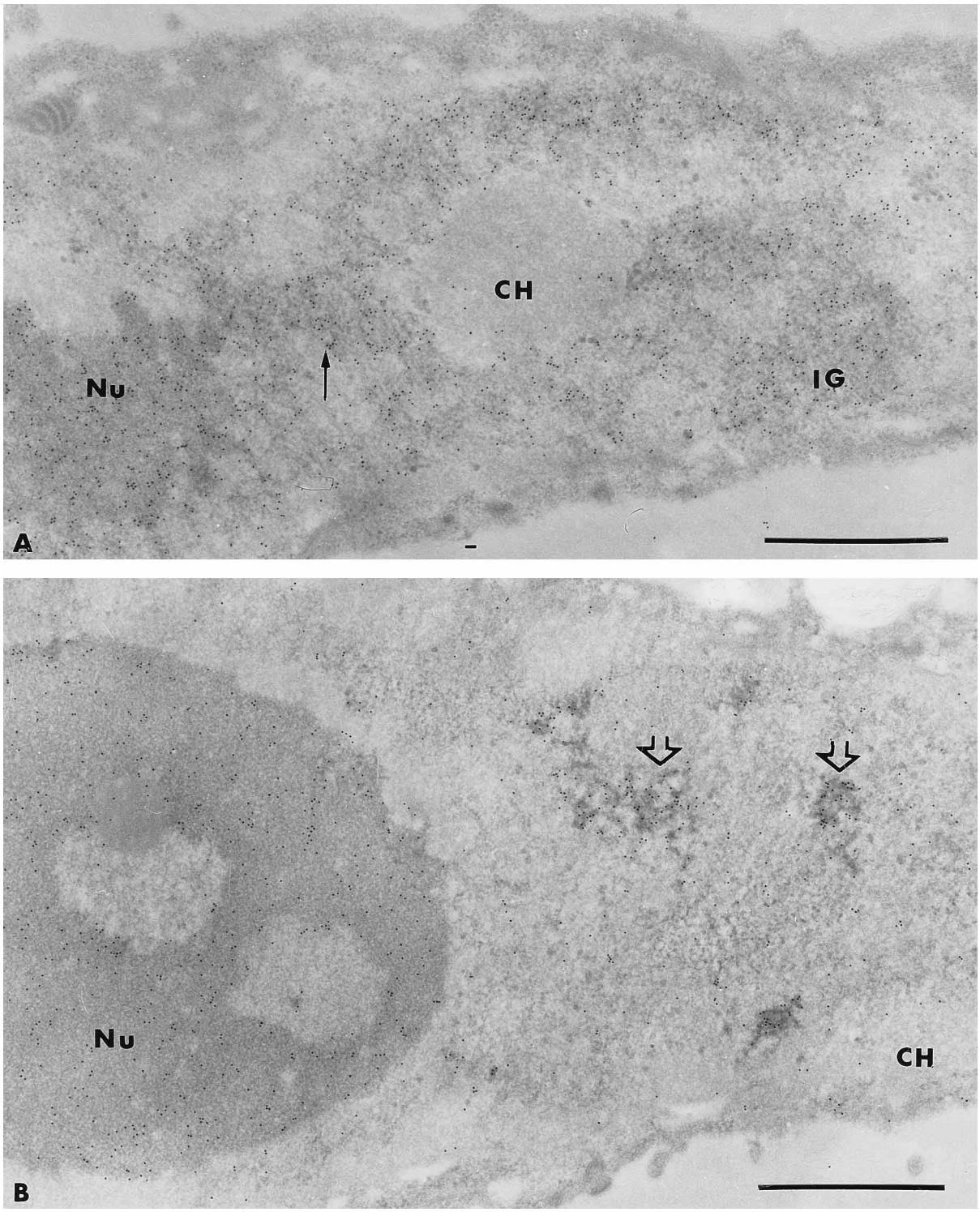


FIG. 3. Immunodetection of MOK proteins in HeLa cells transfected with expression vectors CMV-hsMOK2 (A) and CMV-MOK2 (B). In panel A, intense labeling in perichromatin fibrils (arrow), interchromatin granule clusters (IG), and nucleoli (Nu) was observed. In panel B, the nucleoplasmic labeling accumulated over dense aggregates (thick open arrows), while the rest of the nucleoplasm was only lightly labeled. The nucleolus (Nu), which displayed partial segregation of its components, was heavily labeled. In both panels, no labeling of the bleached chromatin (CH) was observed. Scale bars, 1 μm.

ence of increasing amounts of unlabeled specific and unspecific competitors. Some homopolymers were not used for competition because these homopolymers did not allow us to distinguish between the competition effect and RNA duplex formation. Poly(G) binding by murine and human MOK2 proteins was specifically competed out by an excess of unlabeled poly(G) but not by the addition of poly(A), tRNA, or double-stranded DNA fragments (Fig. 5A and B, left panel). Poly(U) binding by

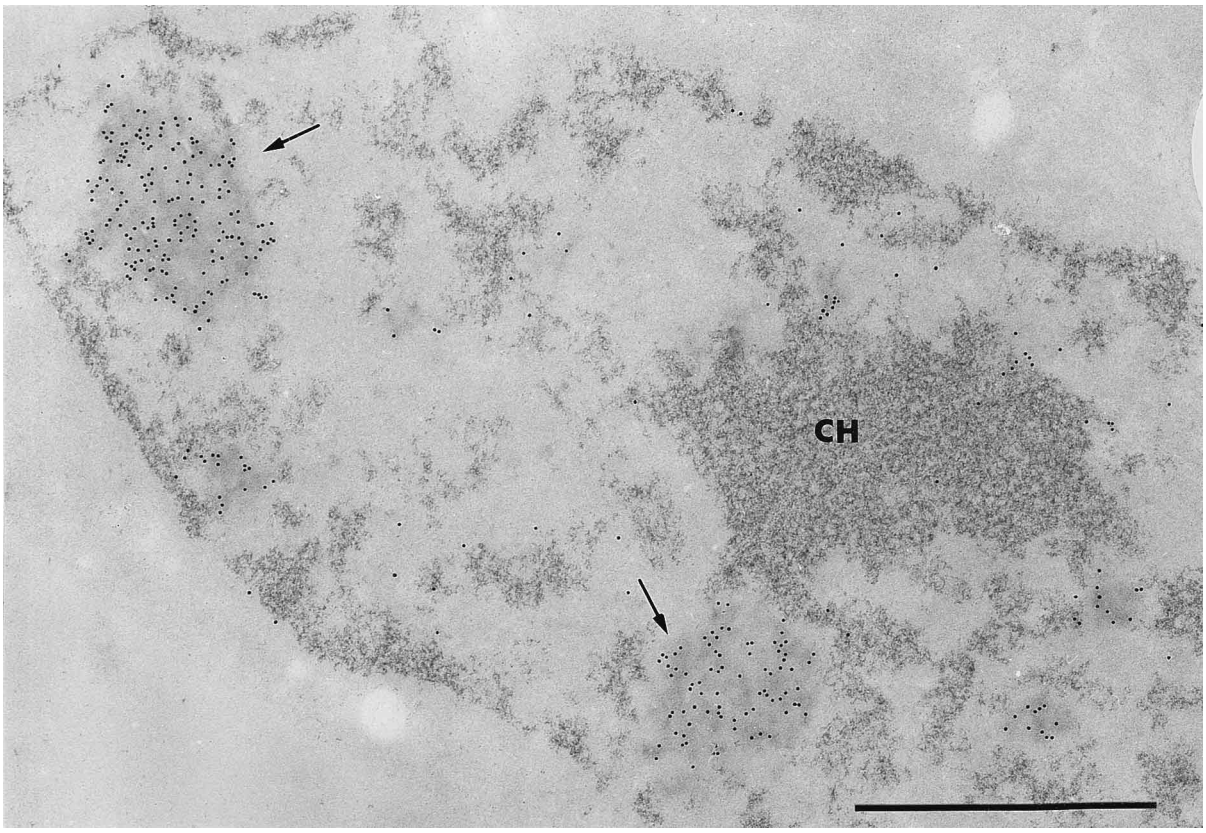


FIG. 4. Specific DNA staining and immunodetection of the MOK2 protein in HeLa cells transfected with CMV-MOK2 vector. The condensed and more dispersed chromatins (CH) were specifically stained. The labeling accumulated over unstained aggregates, which consequently do not contain DNA (arrows). Scale bar, 1 μm.

murine and human proteins was specifically competed out by an excess of unlabeled poly(U), whereas the addition of poly(C) or double-stranded DNA fragments did not influence the binding to poly(U) (Fig. 5A and B, right panel). Poly(U) binding of murine MOK2 protein was reduced by half in the presence of an excess of tRNA (Fig. 5A, right panel). This reduction was not found with the human hsMOK2 protein (Fig. 5B, right panel).

**Identification of a consensus DNA binding site for the murine and human MOK2 proteins.** Electron microscopy and RNA homopolymer binding showed clearly that the murine and human MOK2 proteins are RNA binding proteins mainly associated with nuclear RNP components including the nucleoli and extranucleolar structures. However, the morphological methods used were not sufficient to exclude the possibility of DNA binding of MOK2 proteins to a few dispersed chromatin fibers. Therefore, we set up experiments to determine if murine and human MOK2 proteins could also be DNA binding proteins.

In order to identify a consensus DNA binding site for the two proteins, the β-Gal fusion proteins were chosen as a source of MOK2 and hsMOK2, and an anti-β-Gal antibody was used to select specific DNA-protein complexes. This eliminated the problem of using the anti-MOK2 antibody, which would have been expected to prevent DNA-binding activity. Crude extract proteins were used for the binding reactions with a labeled oligonucleotide that contained a core of 35 random nucleotides (nt) and defined flanking sequences as described in Materials and Methods. Specific DNA-protein complexes were purified by immunoprecipitation, and the recovered DNA was amplified by PCR. The β-Gal protein was used in a parallel selection procedure as a negative control.

Following seven rounds of binding and amplification, the PCR products were used in mobility shift assays with β-Gal fusion proteins. Oligonucleotides contained in the shifted bands were extracted from the dried gel and amplified by PCR (data not shown). The selected populations of amplified DNA were cloned and sequenced. The sequences of 24 and 33 unique clones selected by murine and human MOK2 proteins respectively, are shown in Fig. 6. A consensus binding sequence of 18 nt was readily detectable in the sequences selected by the two proteins, i.e., 5′-(A/G)CCTT(A/G)TCAGG GCCTTTA-3′ for MOK2 and 5′-(A/G)CCTT(A/G)TCAGGG CCTTT(A/G)-3′ for hsMOK2. The frequencies of the individual nucleotides at each of the 18 positions of the derived consensus sequences are shown. The majority of specific nucleotide bases occur with a frequency of more than 70%, and for

TABLE 1. RNA homopolymer binding activities of murine and human MOK2 fusion proteins and relative controls[a]

| Protein | RNA binding (%) | | | |
|---|---|---|---|---|
| | Poly(A) | Poly(U) | Poly(C) | Poly(G) |
| β-Gal | 0 | 0.5 | 0 | 10 |
| β-Gal–MOK2 | 0 | 15 | 3 | 56 |
| β-Gal–hsMOK2 | 0 | 6 | 1 | 34 |
| β-Gal–XFG 5-1 | 0 | 20 | 2 | 56 |

[a] The indicated values are percentages of nucleic acid bound by each protein in the in vitro binding assay as described in Materials and Methods. The percentages shown are the averages of five independent experiments. Experiments were performed with at least two independent protein preparations.

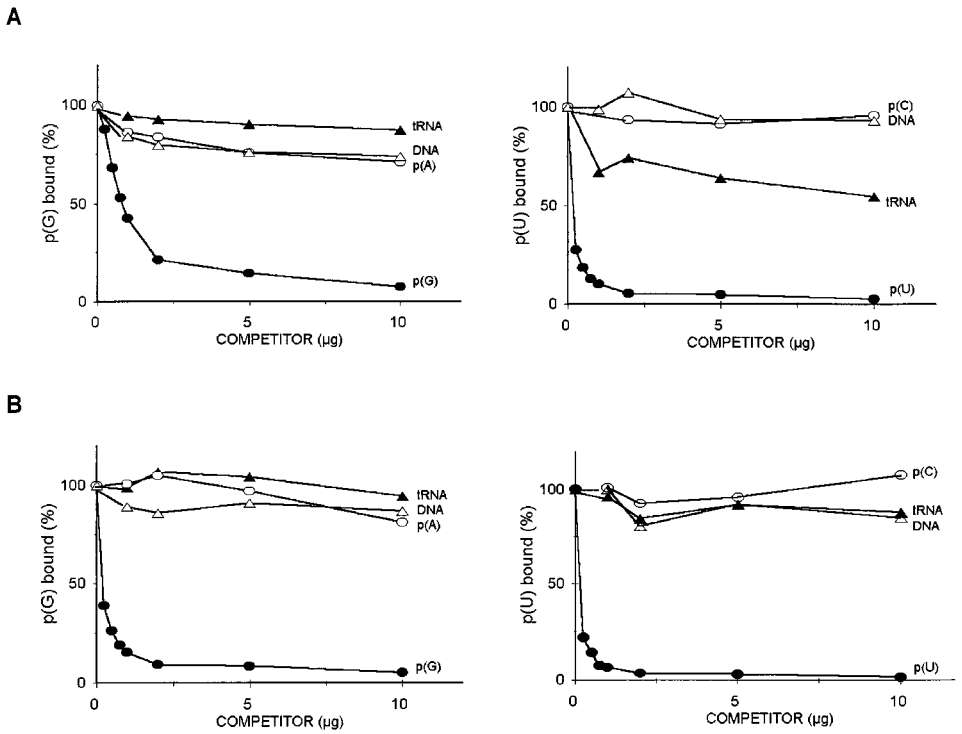


FIG. 5. Competition experiments to test the poly(G) and poly(U) binding specificities of fusion proteins β-Gal–MOK2 (A) and β-Gal–hsMOK2 (B). The immunoprecipitated fusion proteins were incubated with 5′-end-labeled poly(G) [p(G)] or poly(U) [p(U)] in the presence of increasing amounts of the indicated competitors (RNAs or DNA). Four independent experiments with two independent protein preparations were performed for each point.

certain bases this was even higher (90 to 100%). The zinc finger domains of MOK2 and hsMOK2 appear to select DNA consensus binding sites of the same length, with a slight difference at position 18. This involved an A rather than A/G in the case of the mouse protein. The deletion or insertion of base pairs found in some sequences is probably due to PCR mispriming. The consensus sequences do not show any preference for purine (39%) or pyrimidine (61%) bases.

To determine whether the derivated DNA binding consensus sequences obtained by using the β-Gal fusion proteins were representative of the DNA binding site of MOK2 proteins, we performed electromobility shift assays with in vitro-translated murine MOK2 protein and with the maltose fusion protein MBP (Fig. 7A and B). As illustrated in Fig. 7, the in vitro-translated MOK2 protein and the murine and human MBP fusion proteins specifically bound a double-stranded oligonucleotide containing the binding site sequence. No DNA-protein complex was found with rabbit reticulocyte lysats without MOK2 and MBP. The binding specificity of the MOK2 protein for the selected motif was tested in a competition experiment. An excess of unlabeled homologous double-stranded oligonucleotide effectively abolishes the specific shift seen with the probe (Fig. 7A). This specificity of binding was confirmed by DNase I footprinting assays (Fig. 8). The murine and human MOK2 proteins protected the entire sequence on both DNA strands corresponding to the determined DNA binding site.

We have tested whether the MOK2 proteins are also able to bind to single-stranded DNA or RNA containing the target sequence. For this purpose, we used the two complementary strands of the 18-nt synthetic oligonucleotide or RNA transcribed in vitro in both orientations. Electromobility shift assays with maltose murine or human MOK2 fusion proteins showed that no complexes were observed with labeled DNA or RNA probes (data not shown). These results show that the MOK2 proteins are unable to recognize the primary target sequence in single-stranded DNA or RNA.

**Isolation of murine and human genomic DNA sequences which bind to MOK2 proteins.** Identification of an 18-bp DNA binding site by the randomized oligonucleotide approach provides a tool to isolate potential DNA target genes for human and murine MOK2 proteins. For this purpose, we performed a previously described technique in which whole-genome PCR is used to amplify sequences selected by binding to a specific protein (32). After a total of four rounds of binding to human and murine β-Gal fusion proteins and PCR, the selected populations were cloned into plasmid vectors. We screened 500 recombinant clones for each protein with the 18-bp radiolabeled, double-stranded oligonucleotide 5′-RCCTTRTCAGG GCCTTTR-3′ (R = A or G). Approximately 15% of clones hybridized with the probe corresponding to the consensus MOK2 binding site characterized by the randomized oligonucleotide procedure. Six murine and ten human clones that were strongly positive were isolated and sequenced. Alignment of the nucleotide sequences of genomic clones is shown in Fig. 9. For each species, we show only the clones which presented at least 5% difference between themselves. All clones contained sequences similar to that of the 18-bp MOK2-binding site identified. Comparison of human and murine sequences shows that identities between the murine and human clones were found. The Mm1 clone shared 96.9% identity with the Hs1 clone, and the Mm2 clone shared 98.6% identity with Hs2. This result indicates that the genes corresponding to these fragments are conserved between humans and mice. The sequence alignment of genomic fragments shows that an adenosine is found at the second position of the 18-bp MOK2 binding site. At this position, a preference for a cytosine was found in the consensus sequence determined by the random oligonucleotide technique. The human and murine homolog clone 2 has an

**A**

```
 1                        g caactG TCAGGGCCTTTA ATCGGGCAGTGGGGTGTGGTggatcc
 2   ggatccCACGCTCGTAGAGACT TCCTTA TCAGAGCGTTTA Tagttgc
 3                 gcaactTGAGA ACCACG TCAGGTCCTTGA AGTTTTCATGCTggatcc
 4 gcaactGTACGAGTGACACTGCAC ACCTTG TCAGGTCTTTCg gatcc
 5          gcaactATTACGAGCAC AGCTTG TCAGGGCCTTTA CCTACTggatcc
 6  gcaactGCCATGGAAGGAGTTTG GACTTA TCAGAGCCTTTG ggatcc
 7   ggatccCGGGATAGGATGGATG ACCTTCTTCAGTGCCTTTa gttgc
 8       ggatcccGATAACCACGG CCCTTA TCAGCGCGCTCA GTCCTagttgc
 9    ggatcccGGTCCGCCAGGCGG GCCTTG TCCGAGCCTTTA CCagttgc
10                   ggatccG GACTTG TCAGGGCGTTGA GCGCCTAAGGGGTCTTagttgc
11                 gcaactTGG GCCACG TCCGAGCCTTGA CTTTTGGACCTTTAggatcc
12             ggatccTACGACG ACCTTA TCAGAGCCTTGT TTTGGCTGTTagttgc
13                 gcaactGGGA GCCTTA TCAGTGCACTAC TTGCAATCATACCggatcc
14                   ggatccG GCCTTA TCAGCGCCTTTT GCACGGGCGCCGTCCTagttgc
15              gcaactAAAAAC GCCAGCAACGCGGCCTTTT ACGGTTCCTggatcc
16          ggatccAATCGCGAAA GGCTTA TCAGGTCCTTTA TGCGCGTagttgc
17                 gcaactTGAGA ACCACG TCAGGTCCTTGA AGTTTTCATGCTggatcc
18              gcaactATACGG GACGTG TCAGAGCCTTGA TGTGCCGTCTAggatcc
19                   ggatcccG ACCTTG TCAGGGCCTTAG ATTAAGTGGCGCTTCagttgc
20         ggatcccCTCGGGCACAG GCCTTA TCAGTGCCTTGA ACCTGTagttgc
21                 ggatccGGTG ACCTTAGTGTGGACCTTTA CCCACCGCTGGCagttgc
22            ggatccGATCACAC ACCTTTTTCAGGACCTTTG GGATGGCCagttgc
23            gcaactACTACTTA ACCTTG TCAGCGCGTTTG AAAGCCCCCggatcc
24                       gga tccTTG TCAGTGCCTTGA CTTTATCGGGAGCTGGGTGTagttgc
```

**B**

```
 1  gcaactATAACGGTGGCAATACAC GCCTTA TCAGGGCCGTTg gatcc
 2                     ggatcc ACCTTA TCAGGGCGTTTG GCACCGGTACGCagttgc
 3          ggatccAACGGACTCGA ACCTTAGTTGGGACCTTTA CGTGTagttgc
 4  ggatcccTGCTCCCATGGTCACTT TACTTA TCAGGGCGTTTa gttgc
 5           ggatcccACGCGGTGG CACTTG TCAGGTCCTTTA GGTGGCCagttgc
 6 ggatcccGGAGGTACCTCCACGCGGG ACCTTG ACAGTGCGTTag ttgc
 7      ggatccGGCTACGTAGCACTC TCCTTG TAAGTGCCTTTG CTagttgc
 8                 ggatcccCAAG GCGTTG TCAGGACCTTTA TGGTCTTGGCTGTagttgc
 9          gcaactTATTAGCAAGA GCCTTG TCAGGGCCTTga tcc
10         ggatcccCAGGGCAAACGT TCCACG TCAGAGCCTTTA CCCGTagttgc
11                 ggatccACGC GCCTAG TCAGGGCCTTTA ATTCGGCGGTCGTagttgc
12              ggatccTGCCGT ACCTTG TCAGTGCCTTTG TTCACTGTTTTagttgc
13           ggatccGGTGCGAAGG ACCTTT TCAGGGCCTTGG TCTTTGTagttgc
14                       gga tccTTA TCAGCACCTTTG CACAGAGCGTCGCTGTACagttgc
15          ggatccAAAGTGCTATA TCCGGA GCCGAGCCTTTA Tagttgc
16  ggatcccCACCACCGGCTTTGGCAG GCCTTG TAAGAGCGCTTa gttgc
17             ggatccGACAGCCC ACCTTA TCAGGACATTTG CACGGATTTagttgc
18              ggatccGGAATTGG GCCTTG TCAGGTCGTTAG TCGTTATTTagttgc
19    ggatccGGCGTGCAGTAGCCGTG GCCTTG TCAGTTCCTTTa gttgc
20              ggatccGGGGTGAC ACCTTGATCTGGTCCTTTA CAGTAGTCTagttgc
21     ggatccGAAGACGCTGCCGATT ACCTTC TCAGGTCTTTTG Cagttgc
22      ggatcccCACCGGAGTGTGCAG GACTTG TCAGAGCATGCC GATagttgc
23 gcaactGGAGGATGCCGACATCACG ACCTTA TCAGCGCGTTga tcc
24                ggatccTGCGG CCCTTA TCAGTGCGTTAA ATTTCTTCGTCTagttgc
25                         ga tccTTC ACAGGGCGCCTG CGTGCTACGTCAGCATCTTTagttgc
26          ggatcccCGACCAGCAC TCCTTATTCAGTGCGTTTG TGCCCagttgc
27      ggatcccCACAGCAACAGATTG GCCTGC TCAGGTCCTTTA CTagttgc
28                  ggatccACA GCCTTTATCAGAGCCTTTA CCACGGTTGGCCCagttgc
29            gcaactAGAACCCTC ACGTTAGTCTGAGCCTTTG GTCCGgatcc
30                         gga tccTTG TCAGTGCGTT G TGCGGCCCAGGTTGTagttgc
31                     gcaactAG ACCTTG TCAGCACCTTGG GCCGGTCTACGGAATgatcc
32 gcaactGATCCACGGAGCAGTGCCG GCCTTG TCAGTGCGTTga tcc
33       ggatcccCCCAGCTGCATAGT GCCTTG TCAGGTCATTTA GTTagttgc
```

**DERIVED CONSENSUS SEQUENCE FOR MOUSE MOK2 PROTEIN**

```
A           A
G C C T T G T C A G G G C C T T T A
42          38
42 75 96 75 83 50 96 96 83 96 46 75 100 75 92 100 50 63
```

**DERIVED CONSENSUS SEQUENCE FOR HUMAN hsMOK2 PROTEIN**

```
A           A                       A
G C C T T G T C A G G G C C T T T G
33          33                      52
36 91 94 94 88 52 91 91 88 100 48 64 100 55 91 94 70 45
```

FIG. 6. Identification of a consensus DNA binding sequence for the murine and human MOK2 proteins. The sequences of 24 oligonucleotides selected by the mouse MOK2 protein (A) and the sequences of 33 oligonucleotides selected by the human hsMOK2 protein (B) were aligned for maximum matching. Lowercase, nucleotides derived from the flanking constant region; uppercase, nucleotides from the 35 random core; boxes, regions used to derive the DNA consensus binding site for the murine and human proteins. MOK2 and hsMOK2 consensus sequences were derived from all of the selected clones. The nucleotide frequencies in each position of the derived consensus sequences are indicated at the bottom. Bases with a frequency of greater than 90% are indicated by boldface type.

insertion of cytosine (Fig. 9). Moreover, an extensive region of homology was found on each side of the consensus sequence binding site. A longer consensus sequence, including a *Pst*I site on the left of the 18-bp MOK2 binding site, could then be derived. This *Pst*I site is conserved in the human and murine homolog clones 1 and 2.

A search of a nucleotide sequence data base (GenBank R97 [December 1996]) indicated that only one human clone had been previously reported. Clone Hs4 (103 bp) is 100% identical to (nt 6722 to 6827) intron 2 of the human interstitial retinol-binding protein gene (*IRBP*), and the 50-bp sequence of clone Hs4 shown in Fig. 9 is 85% identical to intron 2 of bovine *IRBP* (nt 7233 to 7283). Moreover, the 18-bp MOK2 binding site has been found in intron 7 of the human *PAX3* gene (nt 7061 to 7096). In this intron, the MOK2 binding site is present in reverse orientation compared to that for the *IRBP* gene. The MOK2-DNA binding site found in the *PAX3* gene contains, as do the nine genomic clones isolated, an adenosine at position 2. DNase I footprint analysis of MOK2 binding to a fragment of the *PAX3* gene containing the MOK2 binding site did not show significant differences from the footprint presented in Fig. 8, which was obtained with a fragment containing the 18-bp synthetic double-stranded oligonucleotide (data not shown).

## DISCUSSION

In this paper, we have shown that murine and human MOK2 zinc finger proteins are able to bind specifically to RNA as well as to DNA. First, our results clearly show that both proteins are mainly associated with nuclear RNP components, including the nucleoli and extranucleolar structures. It is well known that the nucleoli are sites of ribosome biogenesis. This biogenesis involves transcription of rRNA genes, association of nascent transcripts with specific proteins, and the multistep of matura-

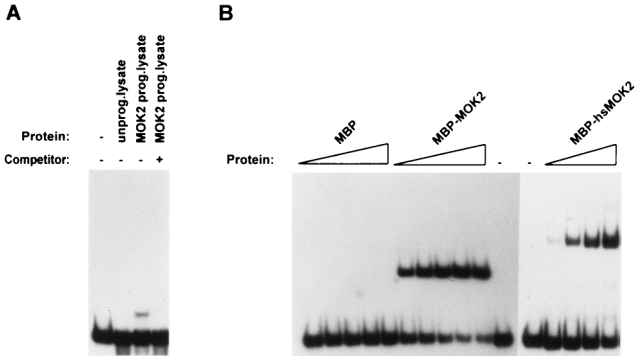


FIG. 7. DNA binding activities of the MOK2 proteins detected by the gel mobility shift assay. Protein extracts were incubated with 0.2 ng of labeled double-stranded oligonucleotide probe containing the consensus sequence GC CTTGTCAGGGCCTTTA. (A) The protein used was a rabbit reticulocyte lysate programmed with (prog) or without (unprog) murine *Mok2* transcript. The competitor used (concentration of 25 ng per reaction) was an unlabeled oligonucleotide corresponding to the probe (+). DNA-protein complexes were analyzed on a native 6% polyacrylamide gel. (B) MBP fusion proteins containing the murine or human MOK2 proteins were produced as described in Materials and Methods. The binding reactions were performed with 2.5, 5, 7.5, 10, or 12.5 μg of crude bacterial extracts. Controls included the oligonucleotide alone (−) or MBP protein. DNA-protein complexes were analyzed on a native 4% polyacrylamide gel.

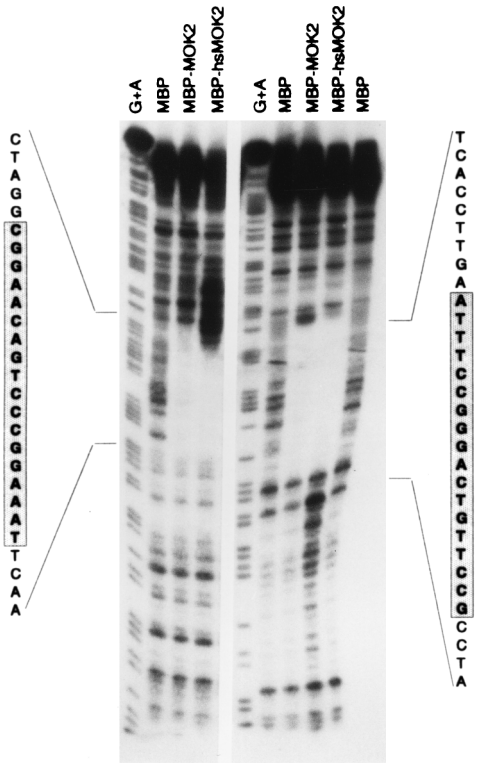


FIG. 8. DNase I footprint analysis of MOK2 binding to the consensus sequence. Single end-labeled probes were incubated with 50 μg of crude extract of MBP fusion proteins. The control was done with MBP. After digestion with DNase I, samples were separated on a urea-6% polyacrylamide gel. A/G ladders were prepared by the Maxam-Gilbert sequencing reaction.

tion of primary transcripts that lead to the formation ot the preribosomal subunits (24, 26). The main and permanent constituents of the nucleolar body are the dense fibrillar and granular RNP components (DFC and GC, respectively). In addition, the DFC encircles fibrillar regions of lower contrast, named fibrillar centers (FC), in a large variety of cells. Both FC and DFC are considered sites of transcription of the rRNA precursors, while the subsequent rRNA processing reactions occur in the GC (51). Therefore, the homogeneous distribution of labeling observed over DFC and GC in this study indicates that the MOK2 proteins may interfere with each of the successive steps of preribosome production. In the nucleoplasm of normal nontransfected and hsMOK2-transfected cells, the proteins were associated with perichromatin fibrils and clusters of interchromatin granules, two structures involved in transcription and processing of pre-mRNA (63, 64). This localization was indirectly confirmed in MOK2-transfected cells, in which the perichromatin fibrils apparently aggregate to form abnormal highly contrasted RNP masses. A number of previous studies demonstrated that perichromatin fibrils are formed at sites of transcription at the border of condensed chromatin, where they correspond to in situ nascent pre-mRNA transcripts associated with hnRNP core proteins (20) and spliceosomes (20, 50, 63). These fibrils subsequently migrate toward the interchromatin space while their RNA is being matured so that they also represent the main substrate of splicing (19, 49). On the other hand the clusters of interchromatin granules which also contain snRNP are not considered active sites of transcription and splicing. They are more likely to be involved in pre- and/or postsplicing events such as spliceosome assembly and sorting of spliced molecules (64). It is very interesting in this context that MOK2 and hsMOK2 were localized in the interchromatin granule-associated zones. These well-delineated fibrillar domains were previously shown to contain the $U_1$ snRNP but not the $U_2$ snRNP. They were also shown to possess proteins of nuclear bodies, especially p80 coilin (52) and PML (53). Unlike the clusters of interchromatin granules, no poly(A) RNA could be detected at that site. Without excluding some other unknown function, the interchromatin granule-associated zones are considered to be a nuclear domain indirectly involved in splicing events such as those involved in the maturation of the $U_1$ snRNP. Therefore, the presence of hsMOK2 and MOK2 in this structure might mean that the proteins are associated with the $U_1$ snRNP particle. This possibility would explain the presence of MOK2 proteins on perichromatin fibrils. This would not exclude their possible association with the pre-mRNA.

The subcellular localization of murine and human MOK2 proteins with nuclear RNP components suggested that these proteins might be able to bind RNA. Thus, we investigated the RNA binding activity in vitro of MOK2 proteins. Our results show that the MOK2 proteins specifically interact with poly(U) and poly(G) RNA. The RNA homopolymer binding properties have previously been studied for *Xenopus* zinc finger proteins (35). These studies had shown that poly(G) and poly(U) binding can reflect a specific RNA binding activity for zinc finger proteins. Recently, it has been reported that another mammalian zinc finger protein, called ZNF74, that is associated with the nuclear matrix, showed RNA homopolymer binding activities identical to those of MOK2 proteins (25). Several proteins of hnRNP have been characterized and distinguished by their RNA homopolymer binding properties (62). Swanson and Dreyfuss (62) have found that some proteins of hnRNP bound poly(G) and/or poly(U) at low salt concentration. The poly(U) binding abilities of other RNA binding proteins have also been described. These different observations show that the affinities of MOK2 and hsMOK2 for poly(U) and poly(G) homopolymers correspond to a specific RNA binding activity. It is evident that poly(G) and poly(U) may not be the natural targets of MOK2 proteins. Therefore, the true physiological target of MOK2 proteins remains to be determined. That the special structure of the murine MOK2 protein consists solely of the

```
Mm1 (129 bp)  .....GGTTTTCCTTGCTGCAGGAC-TTGTCAGGGCCTTTAACATGCTTATGTTCATTG.....
Mm2 (84 bp)   .....AGGCTTCTCC-CTGCA-GACCTTGTCAGAGCCTTTGGAACTCCAAAGTATCTGT.....
Mm3 (169 bp)  .....TGGTTTCTTCACAGCAGGAC-TTGTCAGGGCCTTTAATACACCAATGGACACGG.....
Mm4 (296 bp)  .....GGGTTGTGT---TGCAGGAC-TTGTCAGAGCCTTTGGTATGCTATCGAGCATGG.....
Mm5 (235 bp)  .....ATGGGCGGTGAGTGCAGGAC-TTCTCAGAGCCTTTAATACACAGAAGGGCATCT.....

Hs1 (87 bp)   .....GGTTTTCTTTGCTGCAGGAC-TTGTCAGGGCCTTTAACATGCTTATGTTCATTG.....
Hs2 (79 bp)   .....AGGCTTCTCC-CTGCA-GACCTTGTCAGGGCCTTTGGAACTCCAAAGTATCTGT.....
Hs3 (121 bp)  .....GGGTCTTTGCCC-GCA-GAC-TTGTCAGGGCCTTTAAGAAGTTAGTGTACAGGA.....
Hs4 (103 bp)         GTTTCTGA-CTGCAGGAC-TTGTCAGGGCCTTTAATATATTAACATGTATTA.....

Human  IRBP   .....AAGTTTCTGA-CTGCAGGAC-TTGTCAGGGCCTTTAATATATTAACATGTATTA.....
Bovine IRBP   .....AAGTTCCAAG-CTGCAGGAC-TTG-CAGAGCCTTTAATATAACATACAATATGT.....
Human  PAX3   .....TGTTTTCATACCT-CAAGAC-TTGTCAGGGCCTTTAGTAGGCTAATGTGCACTC.....

                                PstI
Consensus                      CTGCAGGAC-TTGTCAGGGCCTTTAA-A
                                                   A      GG
```

FIG. 9. Alignment of nucleotide sequences of genomic clones isolated by the whole-genome PCR technique. Five different murine (Mm1 to Mm5) and four human (Hs1 to Hs4) sequences from the 18-bp MOK2 binding site were aligned. The sizes of different clones are indicated. The 18-bp MOK2 binding site is boxed. The *Pst*I site indicated in the consensus sequence is present in five clones. The sequences around the MOK2 DNA binding site found in intron 2 of human and bovine *IRBP* genes and intron 7 of the *PAX3* gene are shown. These sequences are localized between nt 6720 and 6771 in human *IRBP*, nt 7233 and 7283 in bovine *IRBP*, and nt 7112 and 7061 in *PAX3*. The accession numbers for these sequences are J05253, M20748, and U12259, respectively. The accession numbers for Hs3, Mm1, Mm2, Mm3, Mm4, and Mm5 are, respectively, Y11085, Y11086, Y11087, Y11088, Y11089, and Y11090.

zinc finger domain indicates that MOK2 can bind RNA through its zinc finger domain. Moreover, the nonfinger domain of the human hsMOK2 does not contain RNA binding domains such as the RNP domain, the double-stranded RNA binding domain, or the K homology domain (40). This is different for Wilms' tumor protein (WT1), which has been found to be associated with spliceosomes and coiled bodies (7, 36). In all WT1 isoforms, an evolutionarily conserved N-terminal RNA recognition motif similar to that in the constitutive splicing factor U1A has been identified (31).

The murine and human MOK2 proteins show different behaviors. The human hsMOK2 protein containing three additional zinc finger motifs and a nonfinger region is unable to aggregate the perichromatin fibrils when it is overexpressed in the cells. In addition, this protein shows a RNA binding activities lower than those of the murine MOK2 protein. These results suggest either (i) that MOK2 proteins might have differential affinities for the same physiological RNA targets or (ii) that MOK2 proteins might bind different physiological RNA targets.

In addition to this RNA binding activity, our results show that MOK2 proteins are equally DNA binding proteins like the transcription factor TFIIIA. The human and murine MOK2 proteins recognize the same 18-bp specific sequence in duplex DNA. This result indicates that the DNA binding domain is certainly localized in the seven adjacent zinc finger motifs, which show 94% identity between human and murine proteins (15). These seven zinc finger motifs are highly similar to one another (16). Structural analysis of three different zinc finger-DNA complexes by X-ray crystallography has demonstrated that each zinc finger interacts with 3 to 5 bp of DNA (18, 44, 45). Assuming this, probably only six of the seven common MOK2 zinc fingers make contact with the isolated 18-bp binding consensus sequence. On the basis of known recognition sequences, a series of rules have been proposed for binding (12, 29, 41, 55, 61). The potential DNA contact residues in the seven zinc finger domains of the murine MOK2 protein, reading in the carboxy- to amino-terminal direction, are INQ IKQ IDQ INQ INQ IKQ VSR (zinc fingers 7 to 1). In the human hsMOK2 protein, two changes were found in these residues. In fingers 7 and 1, the sequences SNQ and ISR, respectively are found. It should be noted that the same number of triplets beginning with an isoleucine amino acid is conserved in both proteins. Six of the seven triplets contain an Ile residue at the first position. To our knowledge, no rule has been described for triplet sequences that contain an Ile residue at the first position. However, as the MOK2 protein contained a number of identical triplet sequences, we would have expected to find a repeated DNA motif in the DNA recognition sequence. This was not the case. Indeed, no identical base pair triplets with the correct spacing (3 bp) were found in the consensus sequence obtained by binding-amplification. Thus, it is not possible to assign a specific zinc finger to a specific base pair triplet. DNA binding sites for only a small fraction of the sequenced zinc finger proteins are known. The MOK2 zinc finger domain does not belong to the previously described zinc finger motif family, since it contains a large number of Ile residues instead of the typically conserved Arg residue. Identification of a DNA binding site for MOK2 proteins will provide a valuable tool to help elucidate the rules governing DNA recognition by zinc finger domains. Further investigations of the MOK2 zinc finger domain with its DNA binding site will allow a better understanding of this important class of DNA binding motifs.

A computer search in the GenBank database with either the 18-bp MOK2-binding site or genomic fragments isolated shows that only two different known human genes contain the MOK2 binding site. Surprisingly, this sequence is found in the intron and not in the promoter region, as shown for the binding sites of several other zinc finger proteins. It is interesting to note that these two potential target genes for the MOK2 protein play a role in the brain, in which the *MOK2* gene is preferentially expressed. The first one codes for the human IRBP, which is expressed specifically in the retina and pineal gland. It is thought to be involved in the visual cycle in the vertebrate retina (6, 37, 47). The MOK2 DNA binding site is localized in intron 2 of the human and bovine *IRBP* genes, which are highly conserved (21). It has been suggested by the authors that introns 2 and 3 of the *IRBP* gene might contain important regulatory elements for *IRBP* gene expression. Interestingly, DesJardin et al. (11) have suggested that a negative regulatory element affecting mRNA elongation might be involved in controlling *IRBP* gene expression during fetal retinal development. The second one is the human *PAX3* gene, which is a transcription factor expressed during brain development (23, 58). In this gene, the MOK2-binding site is localized in the last intron (39). This intron is of particular interest, since it is disrupted by translocation in human alveolar rhabdomyosarcomas (3). Thus, the MOK2 protein might play an important role in transcriptional regulation of the *IRBP* and *PAX3* genes. The identification by the whole-genome PCR technique of six other genomic DNA sequences which bind to MOK2 proteins shows that the MOK2 proteins might be able to regulate at least six additional genes.

In conclusion, the human and murine MOK2 proteins are factors able to recognize both DNA and RNA through their zinc finger motifs. The only other factor with these properties is the transcriptional factor TFIIIA. These dual-function binding properties and subnuclear localization suggest that MOK2 plays roles in transcription as well as in the posttranscriptional regulation processes of specific genes.

## ACKNOWLEDGMENTS

We thank T. Pieler for the kind gift of the pAX-XFG 5-1 vector and P. Hughes for linguistic revision of the manuscript.

This research was supported by a grant from the Association de la Recherche sur le Cancer.

## REFERENCES

1. **Andreazzoli, M., S. Delucchini, M. Costa, and G. Barsacchi.** 1993. RNA binding properties and evolutionary conservation of the Xenopus multifinger protein Xfin. Nucleic Acids Res. **21:**4218–4225.
2. **Arranz, V., M. Kress, and M. Ernoult-Lange.** 1996. Localization of zinc finger *Mok2* gene to mouse chromosome 6, a new region of homology with human chromosome 19. Mamm. Genome **7:**77–78.
3. **Barr, F. G., N. Galili, J. Holick, J. A. Biegel, G. Rovera, and B. S. Emanuel.** 1993. Rearrangement of the PAX3 paired box gene in the paediatric solid tumour alveolar rhabdomyosarcoma. Nature Genet. **3:**113–117.
4. **Bernhard, W.** 1969. A new staining procedure for electron microscopical cytology. J. Ultrastruct. Res. **27:**250–265.
5. **Bushmeyer, S., K. Park, and M. L. Atchison.** 1995. Characterization of functional domains within the multifunctional transcription factor, YY1. J. Biol. Chem. **270:**30213–30220.
6. **Chader, G.** 1989. Interphotoreceptor retinoid-binding protein (IRBP): a model protein for molecular biological and clinically relevant. Invest. Ophthalmol. Vis. Sci. **30:**7–22.
7. **Charlieu, J. P., S. Larsson, K. Miyagawa, V. Vanheyningen, and N. D. Hastie.** 1995. Does the Wilms' tumour suppressor gene, WT1, play roles in both splicing and transcription? J. Cell Sci. **19**(Suppl.)**:**95–99.
8. **Chavrier, P., U. Janssen-Timmen, M. G. Mattei, M. Zerial, R. Bravo, and P. Charnay.** 1989. Structure, chromosome location, and expression of the mouse zinc finger gene *Krox-20*: multiple gene products and coregulation with the proto-oncogene c-*fos*. Mol. Cell. Biol. **9:**787–797.
9. **Coghati, R., and A. Gautier.** 1973. Mise en évidence de l'ADN et des polysaccharides à l'aide d'un nouveau réactif "de type Schiff." C. R. Acad. Sci. **276:**3041–3044.
10. **Crosby, S. D., J. J. Puetz, K. S. Simburger, T. J. Fahrner, and J. Milbrandt.** 1991. The early response gene NGFI-C encodes a zinc finger transcriptional activator and is a member of the GCGGGGGCG (GSG) element-binding

protein family. Mol. Cell. Biol. **11:**3835–3841.
11. **Des,Jardin, L. E., A. M. Timmers, and W. W. Hauswirth.** 1993. Transcription of photoreceptor genes during fetal retinal development. Evidence for positive and negative regulation. J. Biol. Chem. **268:**6953–6960.
12. **Desjarlais, J. R., and J. M. Berg.** 1993. Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. Proc. Natl. Acad. Sci. USA **90:**2256–2260.
13. **El-Baradi, T., T. Bouwmeester, R. Giltay, and T. Pieler.** 1991. The maternal store of zinc finger protein encoding messenger rNAs in fully grown *Xenopus*-oocytes is not required for early embryogenesis. EMBO J. **10:**1407–1413.
14. **Engelke, D. R., S. Y. Ng, B. S. Shastry, and R. G. Roeder.** 1980. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. Cell **19:**717–728.
15. **Ernoult-Lange, M., V. Arranz, M. Leconiat, R. Berger, and M. Kress.** 1995. Human and mouse Kruppel-like (MOK2) orthologue genes encode two different zinc finger proteins. J. Mol. Evol. **41:**784–794.
16. **Ernoult-Lange, M., M. Kress, and D. Hamer.** 1990. A gene that encodes a protein consisting solely of zinc finger domains is preferentially expressed in transformed mouse cells. Mol. Cell. Biol. **10:**418–421.
17. **Ernoult-Lange, M., P. May, P. Moreau, and E. May.** 1984. Simian virus 40 late promoter region able to initiate simian virus 40 early gene transcription in the absence of the simian virus 40 origin sequence. J. Virol. **50:**163–173.
18. **Fairall, L., J. W. R. Schwabe, L. Chapman, J. T. Finch, and D. Rhodes.** 1993. The crystal structure of a 2 zinc-finger peptide reveals an extension to the rules for zinc-finger DNA recognition. Nature **366:**483–487.
19. **Fakan, S.** 1994. Perichromatin fibrils are *in situ* forms of nascent transcripts. Trends Cell Biol. **4:**86–90.
20. **Fakan, S., G. Leser, and T. E. Martin.** 1984. Ultrastructural distribution of nuclear ribonucleoproteins as visualized by immunocytochemistry on thin sections. J. Cell Biol. **98:**358–363.
21. **Fong, S. L., W. B. Fong, T. A. Morris, K. M. Kedzie, and C. D. Bridges.** 1990. Characterization and comparative structural features of the gene for human interstitial retinol-binding protein. J. Biol. Chem. **265:**3648–3653.
22. **Funabiki, T., B. L. Kreider, and J. N. Ihle.** 1994. The carboxyl domain of zinc fingers of the Evi-1 myeloid transforming gene binds a consensus sequence of GAAGATGAG. Oncogene **9:**1575–1581.
23. **Gérard, M., M. Abitbol, A. Delezoide, J. Dufier, J. Mallet, and M. Vekemans.** 1995. PAX-genes expression during human embryonic development, a prelimary report. C.R. Acad. Sci. **318:**57–66.
24. **Goessens, G.** 1984. Nucleolar structure. Int. Rev. Cytol. **87:**107–158.
25. **Grondin, B., M. Bazinet, and M. Aubry.** 1996. The KRAB zinc finger gene ZNF74 encodes an RNA-binding protein tightly associated with the nuclear matrix. J. Biol. Chem. **271:**15458–15467.
26. **Hadjiolov, A. A.** 1985. The nucleolus and ribosome biogenesis. Springer, Berlin, Germany.
27. **Honda, B. M., and R. G. Roeder.** 1980. Association of a 5S gene transcription factor with 5S RNA and altered levels of the factor during cell differentiation. Cell **22:**119–126.
28. **Hydederuyscher, R. P., E. Jennings, and T. Shenk.** 1995. DNA binding sites for the transcriptional activator/repressor YY1. Nucleic Acids Res. **23:**4457–4465.
29. **Jacobs, G. H.** 1992. Determination of the base recognition positions of zinc fingers from sequence analysis. EMBO J. **11:**4507–4517.
30. **Joho, K. E., M. K. Darby, E. T. Crawford, and D. D. Brown.** 1990. A finger protein structurally similar to TFIIIA that binds exclusively to 5S RNA in Xenopus. Cell **61:**293–300.
31. **Kennedy, D., T. Ramsdale, J. Mattick, and M. Little.** 1996. An RNA recognition motif in Wilms' tumour protein (WT1) revealed by structural modelling. Nature Genet. **12:**329–331.
32. **Kinzler, K. W., and B. Vogelstein.** 1989. Whole genome PCR: application to the identification of sequences bound by gene regulatory proteins. Nucleic Acids Res. **17:**3645–3653.
33. **Kinzler, K. W., and B. Vogelstein.** 1990. The *GLI* gene encodes a nuclear protein which binds specific sequences in the human genome. Mol. Cell. Biol. **10:**634–642.
34. **Klöcke, B., M. Köster, S. Hille, T. Bouwmeester, S. Bohm, T. Pieler, and W. Knöchel.** 1994. The far domain defines a new Xenopus laevis zinc finger protein subfamily with specific RNA homopolymer binding activity. Biochim. Biophys. Acta. **1217:**81–89.
35. **Köster, M., U. Kuhn, T. Bouwmeester, W. Nietfeld, T. El-Baradi, W. Knöchel, and T. Pieler.** 1991. Structure, expression and in vitro functional characterization of a novel RNA binding zinc finger protein from Xenopus. EMBO J. **10:**3087–3093.
36. **Larsson, S. H., J. P. Charlieu, K. Miyagawa, D. Engelkamp, M. Rassoulzadegan, A. Ross, F. Cuzin, V. Vanheyningen, and N. D. Hastie.** 1995. Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. Cell **81:**391–401.
37. **Liou, G., L. Geng, and W. Baehr.** 1991. Interphotoreceptor retinoid-binding protein: biochemistry and molecular biology. Prog. Clin. Biol. Res. **362:**115–137.
38. **Luckow, V. A., and M. D. Summers.** 1988. Trends in the development of baculovirus expression vectors. Bio/Technology **6:**47–55.
39. **Macina, R. A., F. G. Barr, N. Galili, and H. C. Riethman.** 1995. Genomic organization of the human PAX3 gene: DNA sequence analysis of the region disrupted in alveolar rhabdomyosarcoma. Genomics **26:**1–8.
40. **Nagai, K.** 1996. RNA-protein complexes. Curr. Opin. Struct. Biol. **6:**53–61.
41. **Nardelli, J., T. J. Gibson, C. Vesque, and P. Charnay.** 1991. Base sequence discrimination by zinc-finger DNA-binding domains. Nature **349:**175–178.
42. **Nietfeld, W., H. Mentzel, and T. Pieler.** 1990. The Xenopus laevis poly(A) binding protein is composed of multiple functionally independent RNA binding domains. EMBO J. **9:**3699–3705.
43. **Ohlsson, H., and T. Edlund.** 1986. Sequence-specific interactions of nuclear factors with the insulin gene enhancer. Cell **45:**35–44.
44. **Pavletich, N. P., and C. O. Pabo.** 1991. Zinc finger-DNA recognition—crystal structure of a Zif268-DNA complex at 2.1-A. Science **252:**809–817.
45. **Pavletich, N. P., and C. O. Pabo.** 1993. Crystal structure of a 5-finger GLI-DNA complex—new perspectives on zinc fingers. Science **261:**1701–1707.
46. **Pelham, H. R., and D. D. Brown.** 1980. A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. Proc. Natl. Acad. Sci. USA **77:**4170–4174.
47. **Pepperberg, D., T. Okajima, B. Wiggert, H. Ripps, R. Crouch, and G. Chader.** 1993. Interphotoreceptor retinoid-binding protein (IRBP). Molecular biology and physiological role in the visual cycle of rhodopsin. Mol. Neurobiol. **7:**61–85.
48. **Pollock, R., and R. Treisman.** 1990. A sensitive method for the determination of protein-DNA binding specificities. Nucleic Acids Res. **18:**6197–6204.
49. **Puvion, E., and G. Moyne.** 1978. Intranuclear migration of newly synthesized extranucleolar ribonucleoproteins. A high resolution quantitative autoradiographical and cytochemical study. Exp. Cell Res. **115:**79–88.
50. **Puvion, E., A. Viron, C. Assens, E. Leduc, and P. Jeanteur.** 1984. Immunocytochemical identification of nuclear structures containing snRNPs in isolated rat liver cells. J. Ultrastruct. Res. **87:**180–189.
51. **Puvion-Dutilleul, F., J. P. Bachellerie, and E. Puvion.** 1991. Nucleolar organization of HeLa cells as studied by in situ hybridization. Chromosoma **100:**395–409.
52. **Puvion-Dutilleul, F., S. Besse, E. K. Chan, E. M. Tan, and E. Puvion.** 1995. p80-coilin: a component of coiled bodies and interchromatin granule-associated zones. J. Cell Sci. **108:**1143–1153.
53. **Puvion-Dutilleul, F., M. K. Chelbi-Alix, M. Koken, F. Quignon, E. Puvion, and H. de The.** 1995. Adenovirus infection induces rearrangements in the intranuclear distribution of the nuclear body-associated PML protein. Exp. Cell Res. **218:**9–16.
54. **Reddy, J. C., and J. D. Licht.** 1996. The WT1 Wilms' tumor suppressor gene: How much do we really know? Biochim. Biophys. Acta. **1287:**1–28.
55. **Rosenfeld, R., and H. Margalit.** 1993. Zinc fingers—conserved properties that can distinguish between spurious and actual DNA-binding motifs. J. Biomol. Struct. Dyn. **11:**557–570.
56. **Ruther, U., and B. Muller-Hill.** 1983. Easy identification of cDNA clones. EMBO J. **2:**1791–1794.
57. **Sambrook, J., E. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
58. **Stoykova, A., and P. Gruss.** 1994. Roles of Pax-genes in developing and adult brain as suggested by expression patterns. J. Neurosci. **14:**1395–1412.
59. **Sukhatme, V. P., X. M. Cao, L. C. Chang, C. H. Tsai-Morris, D. Stamenkovich, P. C. Ferreira, D. R. Cohen, S. A. Edwards, T. B. Shows, and T. Curran.** 1988. A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization. Cell **53:**37–43.
60. **Summers, M. D., and G. E. Smith.** 1988. A manual of methods for baculovirus vectors and insect cell culture procedure. Texas Agricultural Experiment Station, College Station, Tex.
61. **Suzuki, M., M. Gerstein, and N. Yagi.** 1994. Stereochemical basis of DNA recognition by Zn fingers. Nucleic Acids Res. **22:**3397–3405.
62. **Swanson, M. S., and G. Dreyfuss.** 1988. Classification and purification of proteins of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities. Mol. Cell. Biol. **8:**2237–2241.
63. **Visa, N., F. Puvion-Dutilleul, J. P. Bachellerie, and E. Puvion.** 1993. Intranuclear distribution of U1 and U2 snRNAs visualized by high resolution in situ hybridization: revelation of a novel compartment containing U1 but not U2 snRNA in HeLa cells. Eur. J. Cell Biol. **60:**308–321.
64. **Visa, N., F. Puvion-Dutilleul, F. Harper, J. P. Bachellerie, and E. Puvion.** 1993. Intranuclear distribution of poly(A) RNA determined by electron microscope *in situ* hybridization. Exp. Cell Res. **208:**19–34.
65. **Werner, H., F. J. Rauscher, V. P. Sukhatme, I. A. Drummond, C. T. Roberts, and D. Leroith.** 1994. Transcriptional repression of the insulin-like growth factor I receptor (IGF-I-R) gene by the tumor suppressor Wt1 involves binding to sequences both upstream and downstream of the IGF-I-R gene transcription start site. J. Biol. Chem. **269:**12577–12582.