DNA AND CELL BIOLOGY
Volume 20, Number 3, 2001
Mary Ann Liebert, Inc.
Pp. 159–173

Alternative Promoter Usage and Splicing of *ZNF74* Multifinger Gene Produce Protein Isoforms with a Different Repressor Activity and Nuclear Partitioning

FRANCINE CÔTÉ,[1]* FRANÇOIS-MICHEL BOISVERT,[1,3]* BENOÎT GRONDIN,[1] MARTINE BAZINET,[1] CYNTHIA G. GOODYER,[2] DAVID P. BAZETT-JONES,[3] and MURIEL AUBRY[1]

ABSTRACT

We have previously shown that *ZNF74*, a candidate gene for DiGeorge syndrome, encodes a developmentally expressed zinc finger gene of the Kruppel-associated box (KRAB) multifinger subfamily. Using RACE, RT-PCR, and primer extension on human fetal brain and heart mRNAs, we here demonstrate the existence of six mRNA variants resulting from alternative promoter usage and splicing. These transcripts encode four protein isoforms differing at their N terminus by the composition of their KRAB motif. One isoform, ZNF74-I, which corresponds to the originally cloned cDNA, was found to be encoded by two additional mRNA variants. This isoform, which contains a KRAB motif lacking the N terminus of the KRAB A box, was devoid of transcriptional activity. In contrast, ZNF74-II, a newly identified form of the protein that is encoded by a single transcript and contains an intact KRAB domain with full A and B boxes, showed strong repressor activity. Deconvolution immunofluorescence microscopy using transfected human neuroblastoma cells and non-immortalized HS68 fibroblasts revealed a distinct subcellular distribution for ZNF74-I and ZNF74-II. In contrast to ZNF74-I, which largely colocalizes with SC-35 in nuclear speckles enriched in splicing factors, the transcriptionally active ZNF74-II had a more diffuse nuclear distribution that is more characteristic of transcriptional regulators. Taken with the previously described RNA-binding activity of ZNF74-I and direct interaction with a hyperphosphorylated form of the RNA polymerase II participating in pre-mRNA processing, our results suggest that the two ZNF74 isoforms exert different or complementary roles in RNA maturation and in transcriptional regulation.

INTRODUCTION

THE TRANSCRIPTIONAL AND POST-TRANSCRIPTIONAL regulation of gene expression is mediated by various types of nucleic acid binding factors (Ladomery, 1997). Among these factors, zinc finger proteins of the Cys_2-His_2 TFIIIA/Kruppel type are classically defined as belonging to the largest class of DNA-binding transcription factors (Klug and Schwabe, 1995). Furthermore, some of them have been suggested to function at both the transcriptional and post-transcriptional levels (Theunissen *et al.*, 1992; Larsson *et al.*, 1995; Bardeesy and Pelletier, 1998; Arranz *et al.*, 1997). It is estimated that between 300 and 700 human genes encode proteins containing zinc finger motifs of the Cys_2-His_2 type; a large proportion of them, called multifinger genes, encode more than five zinc finger motifs organized in tandem (Klug and Schwabe, 1995; Pieler and Bellefroid, 1994). The Cys_2-His_2 zinc finger proteins are often divided into subclasses according to the presence and identity of evolutionarily conserved domains located at the N-terminus of the encoded protein, the most common being the **Kr**uppel-**a**ssociated **b**ox (KRAB) box that is found exclusively in multifinger proteins. The KRAB/Cys_2His_2 proteins constitute the single largest class of potential regulators of gene expression (Mark *et al.*, 1999). Indeed, on the basis of hybridization studies, it is estimated that KRAB-containing genes represent one third of the Cys_2His_2 TFIIIA/Kruppel type of zinc finger genes

[1]Département de Biochimie, Université de Montréal, Montréal, Québec, Canada.
[2]Department of Pediatrics, McGill University, Montréal, Québec, Canada.
[3]Department of Cell Biology and Anatomy, Faculty of Medicine, University of Calgary, Calgary, Alberta, Canada.
*These authors contributed equally to this work.

(Mark *et al.*, 1999; Bellefroid *et al.*, 1991). The KRAB box is composed of two distinct subdomains, generally encoded by separate exons, named A box and B box. The A box can confer distance-independent transcriptional repression when fused to an heterologous DNA-binding domain (DBD) such as the yeast Gal4 DBD (references in Mark *et al.* [1999]; Ryan *et al.* [1999]; Vissing *et al.* [1995]). It has been suggested that such repression is mediated via the interaction of the KRAB box with the ubiquitous corepressor KAP-1 (also named TIF1-beta or KRIP1) (references in Ryan *et al.* [1999]).

We previously isolated a gene belonging to the large KRAB multifinger family, called *ZNF74* (Aubry *et al.*, 1993; Grondin *et al.*, 1996). Because of its consistent deletion in DiGeorge syndrome-affected patients (Aubry *et al.*, 1993; Demczuk *et al.*, 1995) and its embryonic expression in neural crest-derived tissues (Ravassard *et al.*, 1999), *ZNF74* is considered a candidate gene for involvement in this developmental disorder, which affects 1 in 3000 to 5000 newborns (Burn *et al.*, 1993).

If the presence of a KRAB box in the ZNF74 protein is suggestive of a role in transcriptional regulation, several lines of evidence indicate that this multifinger protein may also function as a regulator of mRNA maturation. First, it behaves as an RNA-binding protein associated with the nuclear matrix (Grondin *et al.*, 1996). Second, it interacts with a hyperphosphorylated form of the largest subunit of the RNA polymerase II (pol IIo), a form of the RNA polymerase functionally associated with mRNA processing (Grondin *et al.*, 1997). Finally, it colocalizes in the nuclear matrix with the pol IIo in subnuclear domains enriched in splicing factors (Grondin *et al.*, 1997).

The previously isolated ZNF74 full-length 3.9-kb cDNA encodes a multifinger protein with an N-terminal truncated KRAB A box called here ZNF74-I (Grondin *et al.*, 1996). In this study, we report that alternative promoter usage and splicing produce either isoforms with a truncated KRAB domain lacking repressor activity, such as ZNF74-I, or protein isoforms such as ZNF74-II with repressor activity. Interestingly, these two isoforms exhibit different partitioning in the cell nucleus. Only the isoform without repressive activity associates predominantly with domains enriched in splicing factors. Our results suggest that the isoforms of ZNF74 may have either different or complementary roles in transcriptional regulation and RNA processing.

MATERIALS AND METHODS

Oligonucleotides

The oligonucleotides corresponding to the *ZNF74* gene are described in Table 1. The other oligonucleotides used are the following:

5′-RACE anchor primer: 5′CUACUACUACUAGGCCACG-CGTCGACTAGTACGGGIIGGGIIGGGIIG-3′
3′-RACE adaptor primer: 5′GGCCACGCGTCGACTAGTAC-TTTTTTTTTTTTTTTT-3′
Universal adaptor primer: 5′GGCCACGCGTCGACTA GTAC-3′

RNA isolation

RNA was prepared from brain and heart tissues from 13.5-week human fetuses. Total RNA was extracted from the tissue samples using the TRIZOL reagent (GIBCO/BRL) according to the manufacturer's method.

5′- and 3′-end rapid amplification of cDNA ends and RT-PCR analyses

The 5′ and 3′ ends of *ZNF74* variants were obtained essentially according to the protocols from the 5′ and 3′ RACE sys-

Table 1. Oligonucleotides Used in this Study

Name	*Sequence*	*Position*[a]	*Orientation*
A	5′-TTTGTTGGGAGTCCGGTCTGTCC-3′	Exon Ia; nt 4–25	Sense
B	5′-ATGGAGATCCCTGCCCC-3′	Exon Ia; nt 442–458	Sense
C	5′-TATATCGGGTTGGGGTC-3′	Exon Ic; 1199–1215	Sense
D	5′-GACCCCAACCCGATATA-3′	Exon Ic; nt 1199–1215	Antisense
E	5′-ccctctagaGAATCGGTGAGTTTCAAG-3′	Exon IIa; nt 1236–1253 (and a cloning tail)	Sense
F	5′-CATCACATCCCGGTACAAGGCC-3′	Exon IIa; nt 1310–1331	Antisense
G	5′-CTGGACAGGGCCCTCT-3′	Exon IIb; nt 1444–1458	Antisense
H	5′-ACCGCCTTCAGCTCCCATT-3′	Exon III; nt 1459–1477	Antisense
I	5′-CTGCTGTTGAGAGGGCACCG-3′	Exon III; nt 1474–1493	Antisense
J	5′-CTTGCCACCTGTTTTCCTTGATG-3′	Exon III; nt 3273–3295	Sense
K	5′-AACTGTTGTATCCTCGGACGGG-3′	Exon III; nt 3657–3678	Sense
L	5′-GTCCACAGCCACATCCT-3′	Exon IIa; nt 1252–1268	Antisense
P1	5′-CCACACCAGGCTCAGC-3′	Exon Ia; nt 207–222	Antisense
P1′	5′-CTAAGGTTCCGCGGCCAGGAGCCAT-3′	Exon Ia; nt 98–122	Antisense
P2	5′-CTCTGGCTCCTTTGGAACCAT-3′	End of intron I	Sense
P2′	5′-CTGAAATGACATATAGGTTCTCAC-3′	End of intron I	Antisense
P3	5′-CCCACTTTCTCTCCACGAGCAG-3′	End of intron II	Sense
P3′	5′-GCTCGTGGAGAGAAGTGGG-3′	End of intron II	Antisense

[a]The position of the oligonucleotides in *ZNF74* exons is indicated relative to the nucleotide position in *ZNF74-2* cDNA (Accession No. X92715).

tem kits (GIBCO/BRL). In brief, first-strand cDNA was synthesized from 100 to 500 ng of total RNA using 125 ng of primer F or I (for 5′-RACE) or the 3′-RACE adapter primer (for 3′-RACE) and SuperScript™ II RNase H$^-$ reverse transcriptase. After degradation of the RNA template using RNase H, samples were passed through a Microcon-100 filter (Amicon). The recovered eluate containing the cDNA was either used directly for 3′-end PCR amplification or first submitted to a tailing reaction with dCTP and terminal deoxynucleotidyltransferase (TdT) prior to PCR amplification of the 5′-end fragments.

For amplification of 5′-end fragments, PCRs were carried out for 30 cycles in the presence of *Taq* DNA polymerase (Perkin-Elmer) using the template cDNA (generated from the equivalent of 20 to 100 ng of total RNA), as well as primer F or I and the 5′-RACE anchor primer. In some cases, 1% of the first-round PCR products was used to carry out a second round of PCR using the 5′-RACE anchor primer and a 3′ nested primer. For amplification of 3′-end fragments, a first PCR amplification was done with primer J and the universal adapter primer; a second amplification was done using primer K as a nested primer instead of primer J. To confirm the specificity of the reaction, PCR products were resolved on 1.5% (w/v) agarose gels, transferred to a Hybond-N+ membrane (Amersham), and hybridized with an appropriate oligonucleotide probe labeled with (γ-^{32}P]-ATP in the presence of T4 polynucleotide kinase. The 5′- and 3′-end fragments obtained by RACE were first subcloned and then sequenced using the dideoxynucleotide chain-termination method (T7 sequencing kit; USB). For RT-PCR analysis, the first-strand cDNA synthesis was carried out as described above using a $(dT)_{18}$ primer in the presence of SuperScript™ II RNase H$^-$ RT. The PCR was carried out directly after the reverse transcription using as cDNA template the equivalent generated from 20 ng of total RNA. Negative controls included reactions without RNA, primer, or RT and were negative on both agarose gel and Southern hybridization.

Primer extension analyses

For primer extension, total RNA from human fetal brain (10 to 30 μg) was coprecipitated with an antisense oligonucleotide primer (P1, P1′, P2, or P3) end-labeled with (γ-^{32}P)-ATP by T4 polynucleotide kinase (10^6 cpm) (about 2×10^5 cpm/ng of primer). The coprecipitate was then resuspended in 30 μl of annealing buffer (80% formamide, 400 mM NaCl, 40 mM PIPES, 1 mM EDTA, and 1 μl of RNase inhibitor (RNAguard; Pharmacia), incubated at 85°C for 10 min, and hybridized overnight at 30°C. As a negative control, the labeled primer was also hybridized with 5 μg of yeast tRNA. After precipitation of the hybridized products and resuspension in 20 μl of 50 mM Tris HCl, pH 8.3; 75 mM KCl, 3 mM $MgCl_2$, and 0.2 mM DTT, primer elongation was performed at 44°C for 2 h using SuperScript II RNase H$^-$ RT 10 units/μl (GIBCO/BRL). The extension products were extracted with phenol–chloroform, precipitated in the presence of 10 μg of tRNA, and analyzed by electrophoresis on a 5% acrylamide/urea sequencing gel next to a sequencing ladder corresponding to a sequencing reaction mixture prepared with the extension primer or another primer.

In vitro *transcription/translation*

For *in vitro* transcription, *ZNF74* constructs were obtained by subcloning sequences derived from *ZNF74-1* cDNA (Accession No. X71623) and/or 5′-end RACE-PCR fragments in the *Hin*dIII or *Sma*I site of pSP64 poly(A) transcription vector (Promega). To obtain the construct with the ATG_0 from exon Ia (Accession No. AF072558) as the first inframe start codon and encoding the ZNF74-II protein isoform, the region of *ZNF74-1* cDNA located 5′ to the *Xho*I site in exon 1c was replaced by subcloning into the *Xho*I site a 5′-end RACE-PCR product corresponding to the sequence of exon Ia starting at nt 541 (ATG_0 at position 568–570) (Accession No. AF072558) joined to exon 1c (Accession No. AF072560). The other constructs were previously described (Grondin *et al.,* 1996). They are derived from *ZNF74-1* cDNA and correspond to nt 1–2144, which encode the ZNF74-I protein isoform (initiation start at the ATG_1 from exon IIa, located at position 163–165), nt 222–2144 (first inframe ATG is ATG_2 from exon IIb at position 259–261), and nt 367–2144 (no inframe ATG used as initiation start). The ZNF74 protein isoforms were synthesized in the presence of translation-grade [^{35}S]-methionine from pSP64 poly(A) constructs using a TnT™ coupled transcription/translation rabbit reticulocyte lysate system (Promega) as indicated by the manufacturer. The labeled proteins were analyzed by SDS gel electrophoresis, transferred to a nitrocellulose membrane, and detected by autoradiography.

Immunofluorescence and digital deconvolution microscopy

The SK-N-SH human neuroblastoma cells or nonimmortalized HS68 human dermal fibroblasts were cultured directly on glass coverslips in JMEM or DMEM (respectively) supplemented with 10% fetal bovine serum (FBS). At 95% confluence, cells were transfected with 2 μg of the eukaryotic expression vector pCGN encoding either the ZNF74-I (Grondin *et al.*, 1996) or ZNF74-II isoforms and 5 μg of Lipofectamine 2000 (GIBCO/BRL) as suggested by the manufacturer. The pCGN plasmid encodes an N-terminal HA epitope under the control of the cytomegalovirus promotor. At 24 h after transfection, cells were fixed with 1% paraformaldehyde in PBS, pH 7.5, at room temperature for 5 min. Subsequently, cells were permeabilized for 5 min in PBS containing 0.5% Triton X-100. Cells were then labeled for 30 min with a rabbit polyclonal antibody recognizing the HA epitope (Y-11; Santa Cruz Biotechnology) along with a mouse monoclonal antibody specific for the Arg-Ser spliceosome assembly protein SC-35 (hybridoma number 1023768 from ATCC; Fu and Maniatis, 1990). Cells were then incubated with goat anti-rabbit (Cy3, Chemicon International) and goat anti-mouse (Alexa 488; Cedarlane Labs) as secondary antibodies. After rinsing, the samples were mounted in paraphenylenediamine 1 mg/ml in PBS/90% glycerol containing 4′,6-diamidino-2-phenylindole (DAPI) at 3 μg/ml as a DNA-specific dye.

Digital deconvolution immunofluorescence microscopy was performed using a 14-bit cooled CCD camera (Princeton Instruments) mounted on a Leica DMRE immunofluorescence microscope. VayTek Microtome digital deconvolution software

was used to remove out-of-focus contributions, and image stacks were projected into one image plane using Scion (NIH) Image software. False coloring and superimposition was done with Adobe Photoshop 5.0. Measurements of signal intensities were made on the raw images collected with the camera using ErgoVista 4.4 image analysis software (Atlantis Scientific).

Transcriptional activity analysis by a luciferase gene reporter assay

To express the DNA-binding domain of the Gal4 transcription factor, pRSV-Gal4 DBD (1-147), a Rous sarcoma virus enhancer-driven eukaryotic expression vector encoding the Gal4 DBD (aa 1–147) (kindly provided by Dr. Robert Rehfus) was used. To generate Gal4 DBD fusion proteins, PCR fragments amplified with oligonucleotide primers containing *Xba*I cloning sites were subcloned inframe at the 3′ end of the Gal4 DBD in a unique *Xba*I site of the pRSV-Gal4 DBD (1-147) vector. In each case, the 3′ primer also contained an inframe stop codon. The PCR fragments subcloned in pRSV-Gal4 DBD (1-147) corresponded to nt 1236–1458 (aa 41–114; encoded by the *F2* cDNA variant) and nt 1329–1458 (aa 1–43; encoded by the *F1* cDNA variant) of *ZNF74-2* cDNA (Accession No. X92715). These PCR products encode, respectively, the complete KRAB box of ZNF74 (KRAB AB^{ZNF74}), corresponding to exon IIa (KRAB A box) and exon IIb (KRAB B box), and the truncated KRAB box of ZNF74 (KRAB $AB^{trunc\ ZNF74}$) starting at a methionine located at position 1329–1331 in *ZNF74-2* cDNA. The KRAB box from the KRAB multifinger Kid1 (KRAB AB^{Kid1}) was used as a control for repression (nt 312–543; aa 1–77; Accession No. M96548) (Witzgall *et al.*, 1993). All subcloned PCR-derived sequences were confirmed by sequencing of both strands. A luciferase reporter carrying five Gal4-binding sites upstream of a thymidine kinase (TK) promoter with basal activity (−109 bp promotor) (p5xGal4-TK-luc) and a control luciferase reporter plasmid lacking the five Gal4-binding sites (pTK-luc) were used (Durocher *et al.*, 1997).

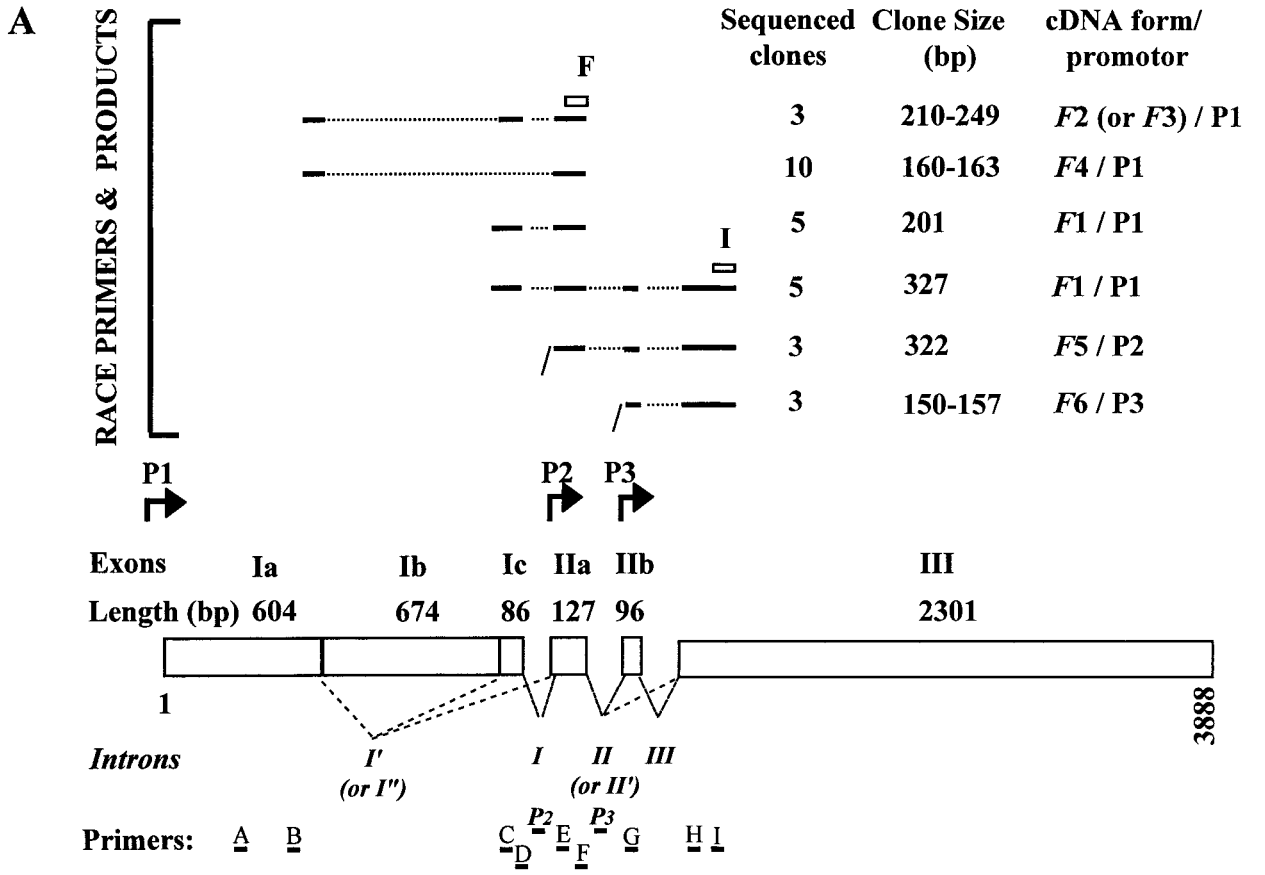
For the luciferase gene reporter assay, transiently transfected L cells or 293 T cells grown in 35-mm tissue culture dishes in DMEM supplemented with 10% FBS were used. Cells plated 24 h prior to transfection at 2.5×10^5 cells/dish were transfected using the calcium phosphate method (Grondin *et al.*, 1997) with 2 μg of p5xGal4-TK-luc or pTK-luc control luciferase reporter and 1 μg of pRSV-Gal4 DBD or of the tested effector construct. The relative transfection efficiencies were determined by cotransfection with β-galactosidase expression plasmid (0.25 μg of pRSV-Lac Z vector). In experiments where the amount of effector was varied, the pGEM vector was used to keep the total amount of transfected DNA constant (usually 7 μg). At 36 h post-transfection, luciferase and β-galactosidase activities were determined from cell extracts (Pott *et al.*, 1995). The luciferase activity was assayed with a Berthold LB953 luminometer. The expression of Gal4 fusion proteins of expected molecular mass was assessed by Western blot analysis using an anti-Gal4 antibody (Santa Cruz). All cotransfections were performed in duplicate for each experiment, and the experiments were repeated at least four times using at least two independent plasmid preparations.

RESULTS

Identification of ZNF74 *mRNA variants generated by alternative promotor usage and splicing*

Previously, *ZNF74* was found to be expressed in human fetal tissues (brain, heart, lung, and kidney; 20–26 weeks) as a doublet of poly(A)$^+$ transcripts detected by Northern blot analysis and migrating at approximately 4.4 kb and 3.6 kb (Aubry *et al.*, 1993). To determine the identity of these two messages and identify potential new functional mRNA variants of *ZNF74*, we used human fetal (13.5 weeks) brain mRNA for 5′-end RACE and RT-PCR analysis, as summarized in Figure 1 and Table 2. RT-PCR products of the same length were identified

FIG. 1. ZNF74 and its various transcripts generated by alternative promotor usage and splicing as identified by 5′-end RACE PCR using fetal brain mRNA. The genomic organization of *ZNF74* was deduced from analysis of the previously isolated cosmids and *ZNF74-2* cDNA (Accession No. X92715) (Aubry *et al.*, 1993, Grondin *et al.*, 1996), as well as from analysis of 5′-end RACE-PCR products (this figure) and RT-PCR products (Table 2) amplified from human fetal brain mRNA. (**A**) The exon–intron structure of *ZNF74* (Accession Nos. AF072557–AF072568). The exons (open boxes) are named I (a,b,c), II (a, b), and III, and their lengths are indicated. The longest cDNA (3888 nt) includes at its 5′ end 130 nt that were not reported in the previously published *ZNF74-2* cDNA. Broken lines and broken dotted lines indicate intron regions from *ZNF74* that are removed by splicing and alternative splicing, respectively. Alternative promotor usage and splicing was suggested by the identification of new human fetal brain transcripts following the sequencing of 5′-end RACE-PCR products. The number of RACE products sequenced, their deduced exon structure, and their size, as well as the primers (either F or I) used to initiate reverse transcription, are indicated. Sequences of the RACE products suggested the existence of three promotors, P1 to P3, and several alternatively spliced transcripts. Sequences from intron I and intron II were found at the 5′ end of transcripts presumably initiated at the P2 and P3 promoters, respectively. The PCR and hybridization primers used (Table 2) to confirm *ZNF74* alternatively spliced exons are named and positioned along the *ZNF74* exon–intron map. (**B**) The various forms of *ZNF74* transcripts expressed in human fetal brain and heart are represented. Initiation of transcription at promotors P1 (5′ to exon I), P2 (5′ to exon IIa), or P3 (5′ to exon IIb) and removal of exon Ib, Ic, and/or IIb by virtue of alternative splicing generate six different possible transcripts (*F1* to *F6* cDNA forms). Untranslated (open boxes) and translated (gray boxes) exons, as well as spliced introns (dotted lines), are shown. The position of two Alu repetitive sequences identified in the 5′ and 3′ UTRs of the KRAB A and B boxes and of the multifinger domain are indicated. The six transcripts encode four protein isoforms (ZNF74-I to ZNF74-IV) that differ mainly in the presence or absence of the KRAB A and B boxes, the KRAB A box being either complete, when the first inframe ATG is the ATG_0, or truncated, when translation is initiated at ATG_1. ▼ = first inframe ATG; ▽ = downstream inframe ATG; trunc = truncated box.


A
RACE PRIMERS & PRODUCTS
F
I
Sequenced clones
Clone Size (bp)
cDNA form/ promotor
3 210-249 F2 (or F3) / P1
10 160-163 F4 / P1
5 201 F1 / P1
5 327 F1 / P1
3 322 F5 / P2
3 150-157 F6 / P3
P1
P2
P3
Exons Ia Ib Ic IIa IIb III
Length (bp) 604 674 86 127 96 2301
1
3888
Introns
I' (or I'')
I
II (or II')
III
Primers: A B C D P2 E F P3 G H I

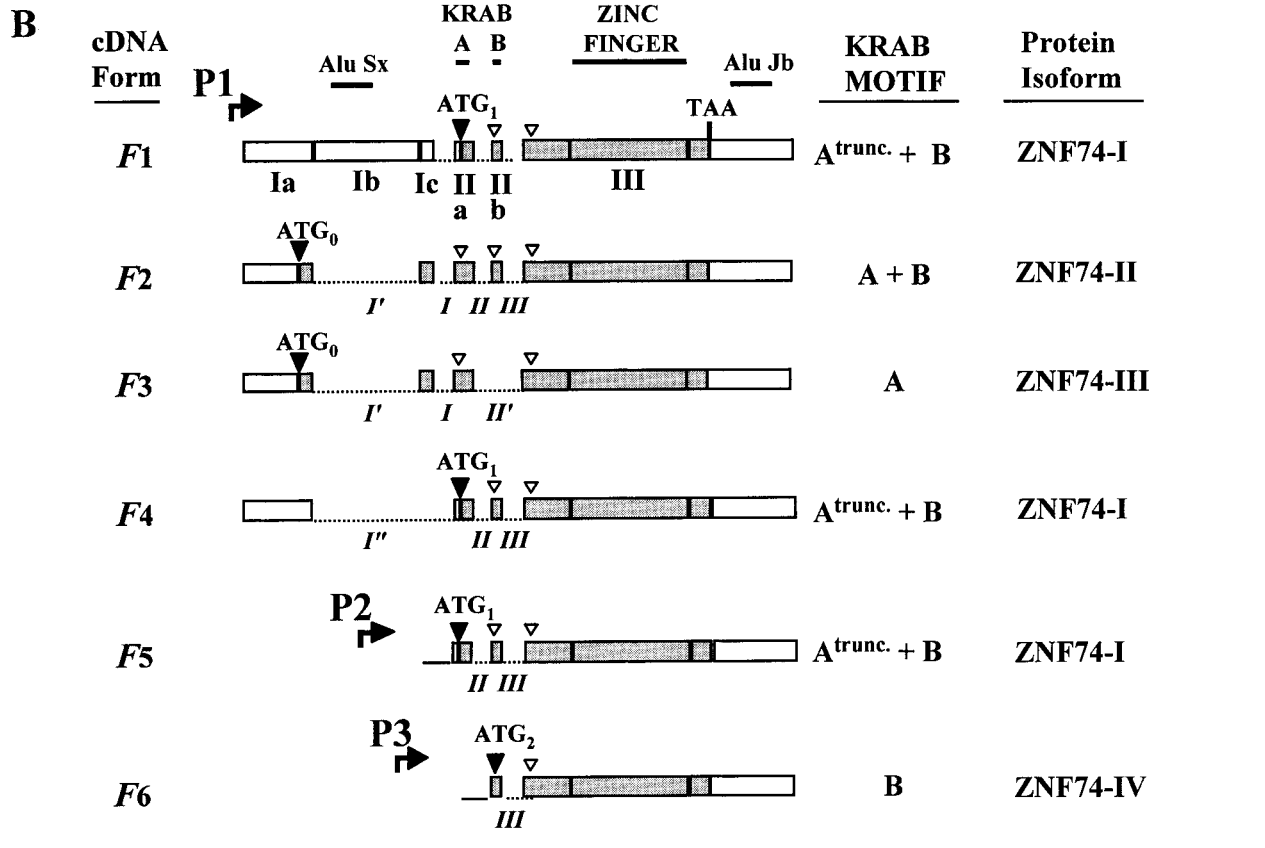

B
cDNA Form
P1
Alu Sx
KRAB A B
ZINC FINGER
Alu Jb
KRAB MOTIF
Protein Isoform
ATG1
TAA
F1
Ia Ib Ic II a II b III
A trunc. + B
ZNF74-I
ATG0
F2
I' I II III
A + B
ZNF74-II
F3
I' I II'
A
ZNF74-III
F4
I'' II III
A trunc. + B
ZNF74-I
P2
F5
II III
A trunc. + B
ZNF74-I
P3
ATG2
F6
III
B
ZNF74-IV

TABLE 2. CHARACTERIZATION OF ZNF74 cDNA FORMS AMPLIFIED BY PCR USING SELECT OLIGOMERS FOR PCR PRODUCT HYBRIDIZATION

PCR primers	*Hybridization oligomer*[a]	*cDNA form/PCR product expected size (bp)*	
		Detected	*Undetected*
B and **G**	**E**	***F4*/256**	*F1*/1016, *F2*/342[b]
B and **F**	**E**	***F2* (*or F3*)/215** ***F4*/129**	*F1*/889
A and **D**	**B**	***F2* (*or F3*)/537**	*F1*/1211[b]
C and **I**	**G**	***F2*/294**	None
	E	***F3*/198**	
E and **H**	**G**	***F2, F4 or F5*/241**	*F3*/145[b]
C and **F**	**E**	***F2 or F3*/132**	None
P2 and **I**	**G**	***F5*/303**	None
	E	***F5*/303**	None
P3 and **H**	**G**	***F6*/139**	None

[a]PCR primers and PCR product hybridization oligomers used to identify the various *ZNF74* cDNA forms expressed in human fetal brain are described in Table 1 and are positioned along *ZNF74* gene as shown in Figure 1A.

[b]Some cDNA fragments expected to be amplified remained undetected; this can be attributed to a bias toward amplification of shorter PCR products or to underrepresentation of some cDNA forms (such as F3).

by oligonucleotide hybridizations using either human fetal brain or heart template mRNA. Both sequence analysis of the obtained 5′-end RACE fragments (Fig. 1A) and oligonucleotide hybridization of the RT-PCR products (Table 2) suggested that transcription of *ZNF74* was initiated at three different promotors (P1 to P3) and that the primary transcript resulting from initiation at the most 5′ promotor (P1) was subjected to alternative splicing (Fig. 1). The P1 appeared to be the most prevalent promotor, as a majority of the 5′-end RACE clones randomly selected and sequenced were initiated at this promotor (Fig. 1A). Interestingly, the P1 promotor was positioned in the proximity of a GC-rich region that includes a *Not*I restriction site (Fig. 2) and resembled the CpG-rich islands that often colocalize with the 5′ ends of genes (Cross *et al.*, 1999).

On the basis of these 5′-end RACE and RT-PCR studies, we can deduce the existence of six cDNA forms, named *F1* to *F6*, as detailed in Figure 1B. Initiation at the P1 promotor produces four cDNA forms (*F1* to *F4*). The spliced *F1* form corresponded to the longest *ZNF74-2* cDNA previously isolated from a human fetal brain cDNA library (Accession No. X92715) (Grondin *et al.*, 1996). The *F2*, *F3*, and *F4* variants were subjected to alternative splicing of exons present at the 5′ end of the *F1* form. These variants were obtained by alternative splicing removal of, respectively, exon 1b, which contained an *Alu Sx* repetitive sequence (*F2* form); both exon 1b and exon IIb, the latter encoding the KRAB B box (*F3* form); or both exon 1b and 1c (*F4* form) (Fig. 1B). Initiation at the P2 and P3 promotors produced the *F5* and *F6* cDNA forms, respectively.

Comparison of the *ZNF74* genomic sequence (Accession Nos. AF072557 through AF072568) with that of the longest *ZNF74-2* cDNA and its spliced variants allowed us to position the exon–intron boundaries and to confirm the presence of the gt/ag consensus sequence at each 5′- and 3′-splice junction (Table 3 and Fig. 2).

Mapping of the 5′ and 3 ends of ZNF74 *mRNA*

In order to estimate the size of the six mRNA variants, we searched for the 5′ transcription initiation sites of *ZNF74* pro-

FIG. 2. Sequence of the 5′-end region of *ZNF74* gene subjected to alternative promotor usage and splicing and deduced amino acid sequence of the N-terminal portion of the encoded protein isoforms. Exons are in uppercase (except for exon Ib, represented here as an alternatively spliced region also referred as intron *I′*) (Accession Nos. AF072558–AF07260, AF072563, AF072565, and AF072568). Introns I′, I″, I, II, and III are in lowercase (Accession Nos. AF072558, AF072561, AF072562, AF072564, AF072566, and AF072567). Intron–exon junctions are shown, and the 5′-gt and 3′-ag ends of the introns are in boldface type. The sizes of introns and exons are within brackets. Encoded amino acids (single-letter code) are shown below the coding regions. Potentially inframe methionines that are in a good context for initiation of translation have been called M_0, M_1, M_2, and M_3 (boldface type and underlined). For the transcript initiated at the P1 promotor, the first inframe methionine is either M_o or M_1, depending on the splicing pattern. The M_1 and M_2 are the first inframe methionines when P2 and P3 promotors are used, respectively. For transcripts initiated at P2, an inframe stop codon (TAA) and an ATG codon in an unfavorable context for translation initiation are found just upstream of M_1 and are in lowercase and underlined. The KRAB A (exon IIa) and (exon IIb) boxes are shaded. The promotors P1, P2, and P3 are indicated by arrows. Noticeably, P1 is a CpG-rich promotor, and P3 exhibits potential TATA (at −22) and CAAT (at −62) boxes. ▼ = 5′ end of the primer extension longest products; ◆ and ◇ = 5′ end of the P1′ primer extension products; ● = 5′ end of *ZNF74-2* cDNA (Accession No. X92715).

```
acccatgcagcggcactacactgatccgcacacggagatacacagcgacatccgcccgcaagagc
atggtgggtgccggagctcccggctgtgggcccaggaactacatttcctagaaagatgtgctagat
  P1>                                  [▼ ExonIa (604bp)                ♦
actcgttcgcgtcgggacgcaagcaccggcccGAACTCGCTCTAGGAAATGGAGTCTGAGCCTGC
GCGCGGCAAACGCTCCCGGGAGGTGTAGTTTGCGCCTATTCGCTGAGGCCGCGTTTCCCGCAGCG
NotI                               •     ◊
GCCGCCTGCTGCTCTTTGTGGCAGTCGCAGTCCTTTTGTGGGAGTCCGGTCTGTCCACTTGCCGG
TCCCTCAGACCGTCGGCGGTCTCTGTCCGCTTCGGGACCTGTCCGCTGGTCGCTCCGCGTCCGAT
GGCTCCTGGCCGCGGAACCTTAGGCCTGGCCCTGGTCTCCGAGCGCGGGTTCGCCGGGAGGAGCG
TGTGGCGGGGGTGTGCCGGGGCGTGAGTGCGCCGAGCATGGGGCTGAGCCTGGTGTGGGGAGTGG
GTATCTGCGGAGCCGGCCTGAACCCCACCTCAGCCGGGCGCGGGGAGGGGGCTCCGTGCGTGTGA
TCGTGCAGCTGTGAGCGCGTGGCCGCCCCGCGGGGCTCCGCTGCAGGCCCCTCAGCCCCAGGAGC
AGTACTCGCTCTTCAGGGCCTGCCCTGGATCCTGGAGGCTACACAGCTGCCCACTCCTCCTGGGG
                                                    Exon Ia][Exon Ib(674bp)
AGGCTGCCGTGGAGGCCATGGAGATCCCTGCCCCGGAGCCCGAGAAGACAGgtacagcttcactc
                  M0 E I P A P E P E K T A

                                                                  Exon Ib]
ttgtagtcagta.Intron I'= Exon Ib(674bp)..ccttgactcatttctccttcgccag
[Exon Ic (86bp)
CTCTTTCCTCTCAGGATCCTGCTCTTTCCCTGAAAGAGAATCTCGAGGATATATCGGGTTGGGGT
  L S S Q D P A L S L K E N L E D I S G W G
                Exon Ic][
CTTCCCGAAGCCAGGTCCAAGgtgagt... .Intron I (4.9kb)....tcccaccgcaagt
L P E A R S K
tgccagggccctgtggtgtctattcagacctgctttcccaatctagaaggtggtgtgggaatgca
tgccaattcccagccccctccctgcccctggtccccgaattcagtcatggtcagtgtctacag
  P2>                                  ▼
cacaaccttctggcatgggagaaggtcacctgaggttaaccctttacttgtagggtccttaatca
atctctggctcctttggaaccatggacacagaatcagtcctgggtgagaacctatatgtcatttc
 ][Exon IIa (127bp)
agGAATCGGTGAGTTTCAAGGATGTGGCTGTGGACTTCACCCAGGAGGAGTGGGGTCAACTAGAC
  E S V S F K D V A V D F T Q E E W G Q L D
                                                          Exon IIa]
TCCCCTCAGAGGGCCTTGTACCGGGATGTGATGTTGGAGAACTACCAGAACCTTCTTGCCCTAG
S P Q R A L Y R D V M1 L E N Y Q N L L A L G
[Intron II (516bp)
gtaaaatccccccagccccaggctgggcgtcggctgtcagccttcttctacatggtagatattaa
ggagaatcccttcagtgccggccctggtgccagatatatcatggggtgggaagcaaggcctgtgc
cttggggcactcatggccaggggctcggcacgccagaagtccccaagggacagcgtgggagctgg
tagcatccttgcttggtgtttcccctggtcctgggataggggcagggagagaagagaagaaggtg
ggaccctggagcccagctggctctggaacagtcctcaagcaatgaagtaaggacgtcagcacacg
  P3>                                  ▼
tggggtcttcagagtatacactgggatttgggttccagtctgagaccctcactgcttttggccca
gaggacctcctgttcccagagagggtgtccacttcacaggcattgtccttgttaggggcggcct
gtaaggtcgactgggcctggagagctgcagcatgcccagcccactttctctccacgagcag]
[Exon IIb (96bp)
GACCTCCACTGCACAAGCCAGATGTGATCTCTCATCTGGAACGAGGCGAGGAGCCATGGAGCATG
  P P L H K P D V I S H L E R G E E P W S M2
                                Exon IIb][
CAGAGGGAAGTCCCCAGAGGGCCCTGTCCAGgtgagca...Intron III (4.6kb)...tat
Q R E V P R G P C P E
      ][Exon III (2301bp)
cttacagAATGGGAGCTGAAGGCGGTGCCCTCTCAACAGCAGGGCATTTGCAAAGAAGAACCGGC
         W E L K A V P S Q Q Q G I C K E E P A
                                                          Exon III ]
CCAGGAGCCCATCATGGAGCGGCCCCTCGGCGGGGCGCAGGCGTGGGGGCGCCAGGCAGGT....
 Q E P I M3 E R P L G G A Q A W G R Q A G
```

TABLE 3. SEQUENCE OF EXON/INTRON JUNCTIONS FOR ALTERNATIVELY SPLICED mRNA VARIANTS OF *ZNF74*

5′ Splice donor	*Intron name/size*	*3′ Splice acceptor*	*RNA variant/isoform/ amino acid interrupted by the intron*
1235 Exon Ic ↓ AGGTCCAAG**gt**gagt . . R S K_{40}	*I*/4.9 kb	1236 ↓ Exon IIa . . . tttc**ag**GAATCGGTG E_{41} S V	*F1*/ZNF74-I/noncoding *F2*/ZNF74-II/none *F3*/ZNF74-III/none
475 Exon Ia ↓ AGAAGACAG**gt**acag . . K T A_{12}	*I′*/0.67 kb	1150 ↓ Exon Ic . . . cgcc**ag**CTCTTTCCT L_{13} S S	*F2*/ZNF74-II/Ala-12 *F3*/ZNF74-III/Ala-12
475 Exon Ia ↓ AGAAGACAG**gt**acag . .	*I″*/5.7 kb	1236 ↓ Exon IIa . . . tttc**ag**GAATCGGTG	*F4*/ZNF74-1/noncoding
1362 Exon IIa ↓ TTGCCCTAG**gt**aaaa . . A L $G_{12 \text{ or } 83}$	*II*/0.52 kb	1363 ↓ Exon IIb . . . gagc**ag**GACCTCCAC $P_{13 \text{ or } 84}$	*F1*, *F4*, *F5*/ZNF74-1/Gly-12 *F2*/ZNF74-11/Gly-83
1362 Exon IIa ↓ TTGCCCTAG**gt**aaaa . . A L E_{83}	*II′*/5.2 kb	1459 ↓ Exon III . . . ttac**ag**AATGGGAGC W_{84} E L	*F3*/ZNF74-III/Glu-83
1458 Exon IIb ↓ CCTGTCCAG**gt**gagc C P $E_{42, 44 \text{ or } 115}$	*III*/4.6 kb	1459 ↓ Exon III . . . ttac**ag**AATGGGGAG $W_{13, 45 \text{ or } 116}$	*F1*, *F4*, *F5*/ZNF74-I/Glu-44 *F2*/ZNF74-II/Glu-115 *F6*/ZNF74-IV/Glu-12

Note: Exon sequences are in uppercase letters, and intron (or spliced exon) sequences are in lowercase letters. The last or first exon nucleotide is marked by an arrow, and the position compared with *ZNF74-2* cDNA (Accession No. X92715) is indicated. The size of introns is indicated. The gt/ag splice sites in each intron are in boldface. The amino acids at exon boundaries are shown below the nucleotide sequences and numbered from the first inframe initiating methionine in the various *ZNF74* variants (as deduced from Figs. 1B and 2).

motors and the 3′ poly(A) addition site. Initial attempts to map the major P1 promotor by 5′-end RACE PCR using several primers from exon Ia led to several prematurely ended products (not shown), probably because of the high GC content of exon Ia (Fig. 2). When primer extension analysis was performed with two different primers on total RNA from human fetal brain, two major extended products ending in close proximity in exon Ia were observed. The first 223-bp product, extending from the *P1′* primer (Fig. 3), ended at position 29 (see ◆ in exon Ia, Fig. 2), and the 351-bp product, extending from the *P1* primer (not shown), ended at position 1 of *ZNF74* cDNA (see ▼ in exon Ia, Fig. 2). These positions map, respectively, 101 and 129 nt upstream of the 5′ end (see ● in exon Ia, Fig. 2) from the previously isolated *ZNF74-2* cDNA clone (Grondin *et al.*, 1996). Using the *P1′* primer (Fig. 3), a second, shorter (116-bp) extended product ending very close to the 5′ end of *ZNF74-2* cDNA (position 7 of *ZNF74-2* cDNA) was also detected (see ◇ in exon Ia, Fig. 2). For P2 and P3, the most prominent products extended by RT ended at G residues lying 98 bp upstream of exon IIa and 153 bp upstream of exon IIb, respectively (Fig. 3 and positions identified by ▼ in Fig. 2).

To localize precisely the poly(A) site addition, and thus the 3′ end of ZNF74 mRNA, we performed 3′-end RACE PCR. The polyadenylation site was found 17 bp downstream from the AATAAAA poly(A) signal at position 3739–3745 in *ZNF74-2* cDNA, as revealed by sequencing of the 3′-end RACE-PCR products obtained. From the most 5′-end position of *ZNF74* mRNAs defined by primer extension and on the 3′-end position determined by RACE, the size of the *F1* (3.89 kb), *F2* (3.21 kb), *F3* (3.12 kb), *F4* (3.13 kb), *F5* (2.63 kb), and *F6* (2.55 kb) variants was estimated. These values do not take into account the length of the poly(A) tail (0.1–0.4kb) added to each of these variants. Thus, the *F1* cDNA form initiated at the P1 promotor most likely corresponds to the *ZNF74* poly(A)$^+$ RNA band previously detected by Northern blot around 4.2 to 4.4 kb using a unique exon III-derived probe (Aubry *et al.*, 1993). The *F2*, *F3*, and *F4* cDNA forms, also initiated at the P1 promotor, should correspond to the second RNA band detected with *ZNF74* probe around 3.4 to 3.6 kb (Aubry *et al.*, 1993). No mRNA species was detected by Northern blot around 2.5 to 3.0 kb, indicating that the *F5* and *F6* forms, initiated at the P2 and P3 promotors, respectively, are less abundant in the human fetal tissues tested (20–26 weeks of fetal age) than the forms initiated at the P1 promotor.

Protein isoforms

The six mRNA variants are predicted to encode four protein isoforms that differ mainly in the presence or absence of KRAB

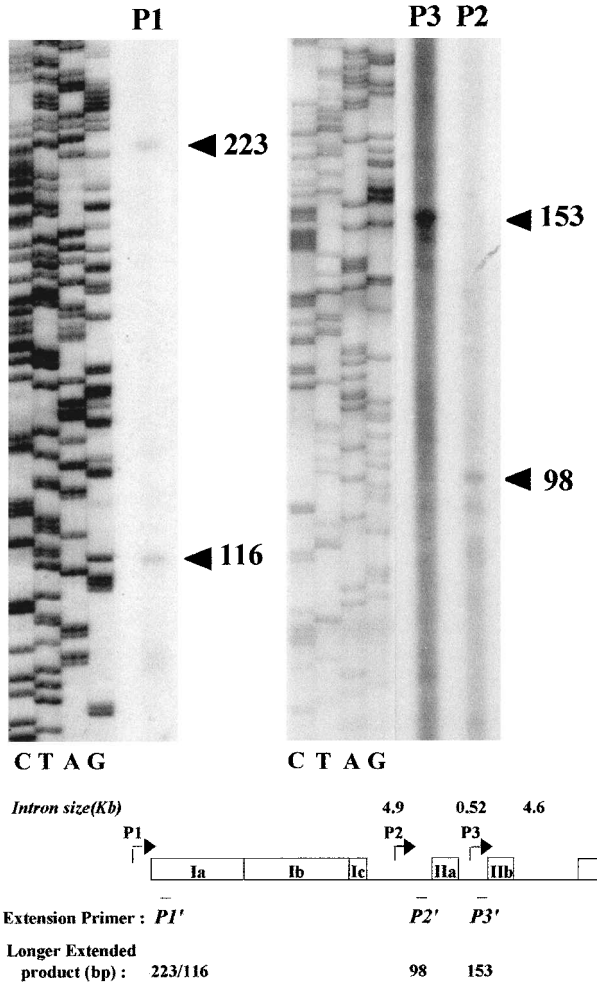


FIG. 3. Mapping of the 5′ end of ZNF74 mRNAs by primer extension analysis. Primer extension was performed using 10 μg of human fetal brain RNA as template. The primers used in the region of the P1, P2, and P3 promotors are represented relative to their position in *ZNF74*. The size of the extension product was deduced from the sequence ladders. Major (223- and 116-bp) products were detected for the P1 promotor. One major extended product of 98 and 153 bp was detected for the P2 and P3 promotors, respectively; shorter and less abundant primer extension products probably resulted from abortive reverse transcription. The ends of the P1, P2, and P3 extended products shown in this figure are positioned in Figure 2 (see triangle and diamond symbols). Lanes CTAG are sequence ladders that were extended from primer L for the P1 promotor and from primer E for the P2 and P3 promotors using the *ZNF74* gene. These sequences are part of the sequence presented in Figure 2.

A and B boxes, the potential KRAB A repressor motif being either intact or truncated (see Fig. 1B). For both the *F2* and *F3* cDNA forms, initiation of translation is predicted to occur at the ATG_o that encodes the first inframe methionine (M_o) found in exon Ia (see Figs. 1B and 2). The ATG_0 is in a context compatible with the Kozak consensus sequence (Kozak, 1991) and is efficiently used for initiation of translation, as seen by *in vitro* transcription/translation of *F2* cDNA in a reticulocyte lysate (Fig. 4, lane 1). The size of the resulting major translation product is in agreement with the molecular mass predicted for the ZNF74-II isoform (72 kDa; pI = 8.43) that encodes intact KRAB A and B boxes. A minor and slightly shorter period that most likely corresponds to alternative initiation at the ATG_2 was also seen. This conclusion was suggested by comparing its molecular mass with that of translation products generated from a truncated cDNA (Fig. 4, lane 3) previously shown to allow

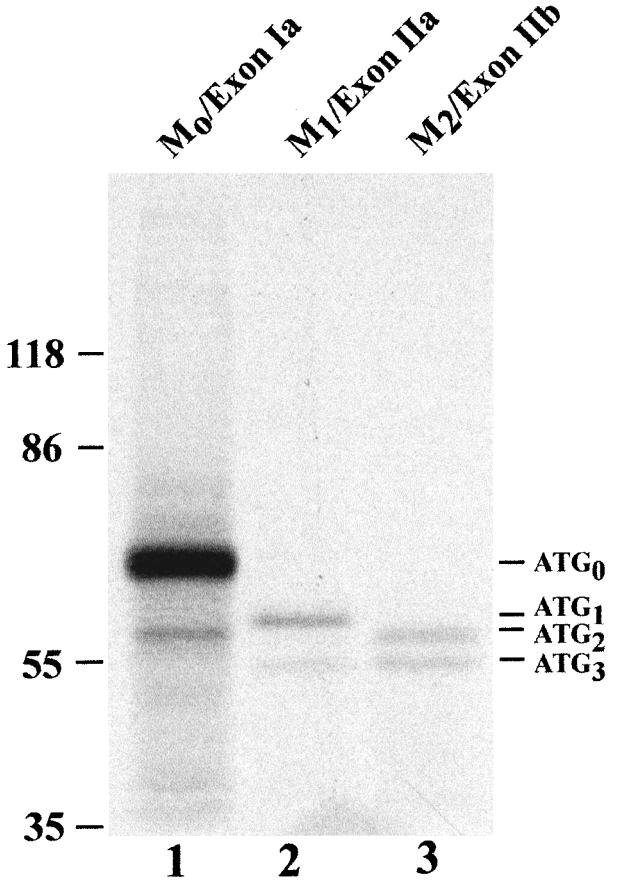


FIG. 4. *In vitro* translation of ZNF74-I and ZNF74-II protein isoforms that have a truncated and a complete KRAB box, respectively. The *ZNF74* constructs cloned under the control of an SP6 promotor were transcribed using SP6 polymerase and translated in the presence of [^{35}S]-methionine. The ZNF74-II protein isoform (643 amino acids), with a complete KRAB motif (A and B boxes), was synthesized from a construct including as a first inframe start codon, the ATG_o from exon 1a (lane 1). The ZNF74-I protein isoform (572 amino acids), with a truncated KRAB motif, was synthesized from a construct containing the ATG_1 from exon IIa as the first inframe translation initiation codon (lane 2). The ATG_2 and ATG_3, found in exon IIb and exon III, respectively, are in an appropriate context to be used alternatively (ATG_2 in lane 1 and ATG_3 in lane 2). Protein products initiated at ATG_2 and ATG_3 are also observed with a deletion construct where the ATG_2 from exon IIb is the first inframe ATG (lane 3) (as shown previously in Grondin *et al.* [1996]), no protein product is observed with a deletion construct starting just after the ATG_3). The calibrated molecular mass of prestained marker proteins (Bio-Rad) is indicated in kDa.

initiation at ATG_2 and ATG_3 (Grondin *et al.*, 1996). The ATG_2 from exon II b is the first inframe methionine (M_2) found in the *F6* variant and most likely serves for its translation initiation. In the *F1*, *F4*, and *F5* variants (Fig. 1B), the first inframe methionine (M_1) is in exon IIa (Figs. 1B and 2). As seen in Figure 4 (lane 2), translation can be initiated from ATG_1. The resulting protein isoform (64 kDa; pI = 8.84), called here ZNF74-I, corresponds to the protein with a truncated KRAB A box that we previously described (Grondin *et al.*, 1996, 1997).

Differential nuclear partitioning of ZNF74-I and ZNF74-II isoforms

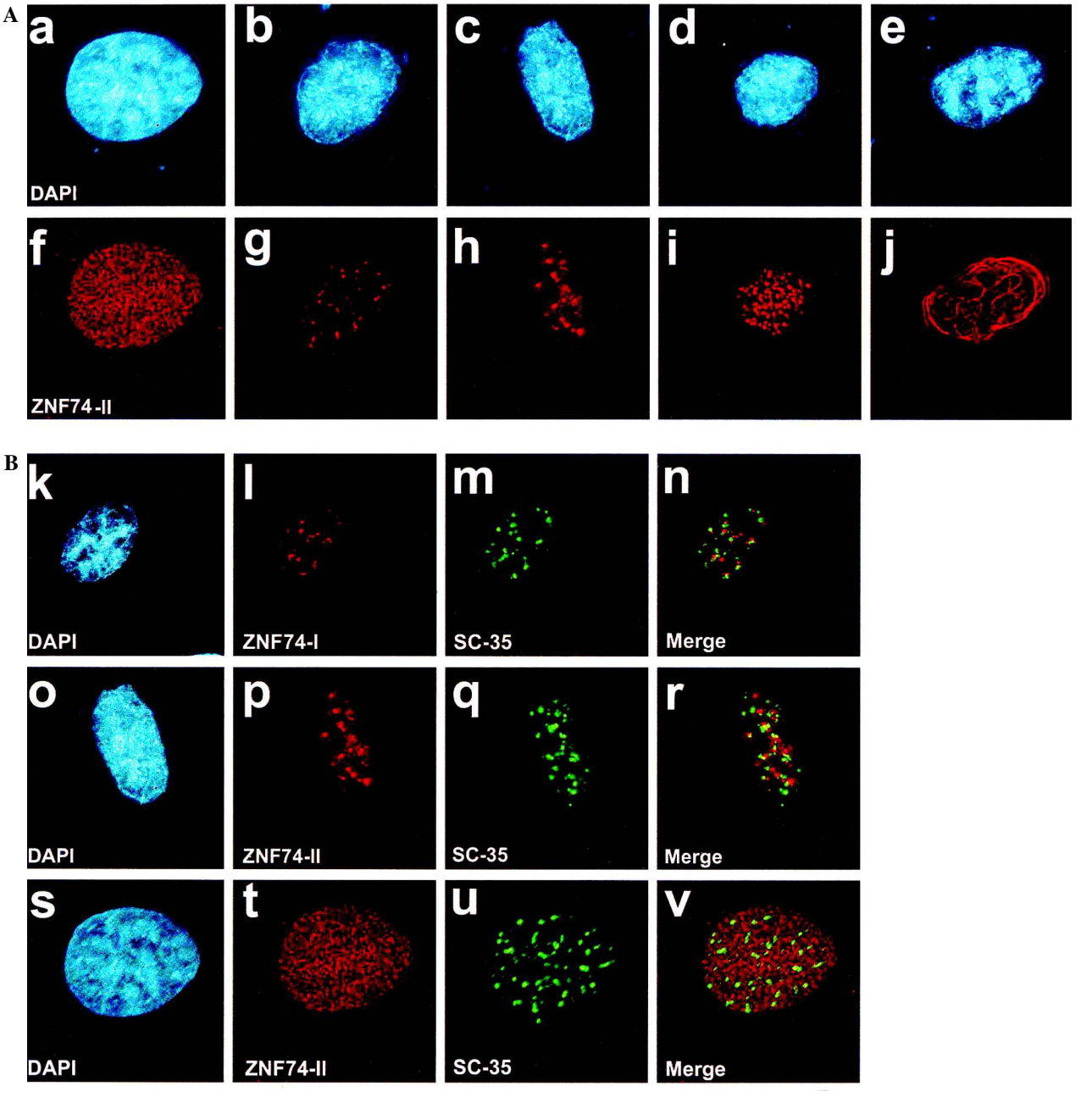
By performing conventional immunofluorescence microscopy on DNase I-treated cells as well as on DNase- and RNase-treated cells, we previously demonstrated that ZNF74-I (isoform with a truncated A box) is tightly bound to the nuclear matrix (Grondin *et al.*, 1996). Double immunofluorescence experiments using transfected monkey COS-7 cells also revealed that this protein isoform colocalizes in the nuclear matrix with a hyperphosphorylated form of the large subunit of the RNA polymerase II as well as with SC-35 splicing factor in irregularly shaped nuclear speckles enriched in splicing factors that are termed "interchromatin granule clusters" (ICGs) (Grondin *et al.*, 1997). In contrast to the results previously obtained with ZNF74-I, our initial studies, performed by conventional immunofluorescence microscopy, revealed no specific association of ZNF74-II (isoform with both intact KRAB A and B boxes) with subnuclear domains enriched in splicing factors (not shown). Here, we therefore used a quantitative digital deconvolution immunofluorescence microscopy approach to compare the subnuclear location of the newly identified ZNF74-II with that of the ZNF74-I isoform. We performed indirect double-labeling immunofluorescence to identify both HA-tagged ZNF74 and SC-35 antigen in transfected SK-N-SH human neuroblastoma cells. The ZNF74 subnuclear location was scored by classifying its typically observed nuclear patterns into five categories. As illustrated for ZNF74-II (Fig. 5A), these five nuclear distribution patterns were defined as uniformly distributed foci (panel f), micropunctuate (~200-nm dots) (panel g), SC-35 like (panel h), punctuate (0.75–1 μm dots) (panel i), and meshed (panel j). The nuclear distribution was classified as SC-35 like in cells in which ZNF74 was mainly found to colocalize or to be in very close association with SC-35 speckles (Fig. 5B, panels k–r). When the relative intensity of the signal was evaluated separately in cells presenting each of the five nuclear distribution profiles, the level of the ZNF74 signal was within the same range in cells presenting uniformly distributed foci as well as micropunctuate, SC-35-like, and punctuate patterns (Fig. 5C). It was, however, much higher (4–12 times) in cells of the meshed type, indicating strong overexpression of ZNF74 in these cells (Fig. 5C). Correlating with this observation, the DAPI staining of the latter cells consistently revealed an abnormal chromatin reorganization and condensation (Fig. 5A, panel e), suggestive of cell death.

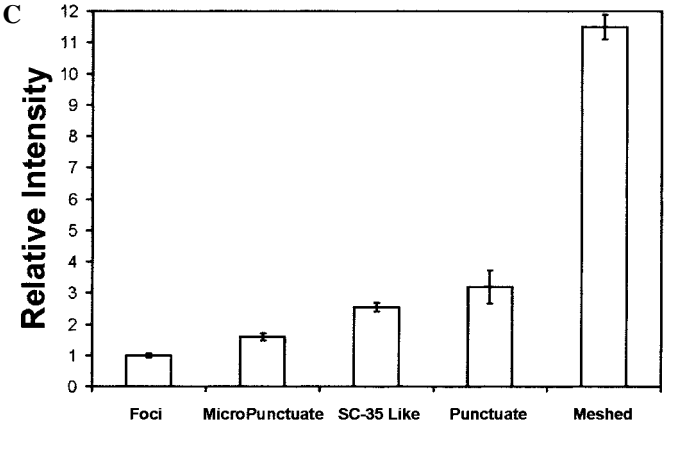
Each of the five nuclear distribution patterns observed with ZNF74-II (Fig. 5A) was also detected in cells expressing ZNF74-I isoforms (Fig. 5D). Clearly, however, the partitioning of ZNF74-II was quite different from that of ZNF74-I (Fig. 5D). In the majority (~60%) of cells, ZNF74-II was uniformly distributed in small foci with no specific association with SC-35 (Fig. 5B, panels s–v). The proportion of ZNF74-II-expressing cells that exhibited an SC-35-like pattern (10%) or any of the three other patterns ranged from 7% to 11%. In contrast, ZNF74-I-expressing cells exhibited two main distribution patterns in 70% of the cells. The protein was found in association with SC-35-enriched domains (~30% of the cells), as previously observed in a large proportion of COS-7 cells (Grondin *et al.*, 1997). In a smaller proportion of the cells than ZNF74-II, the ZNF74-I isoform was also found to be uniformly distributed in small foci (~40% of the cells). This differential partitioning of ZNF74-II and ZNF74-I was not unique to SK-N-SH human neuroblastoma cells, as it was also observed in nonimmortalized human HS68 fibroblasts (not shown).

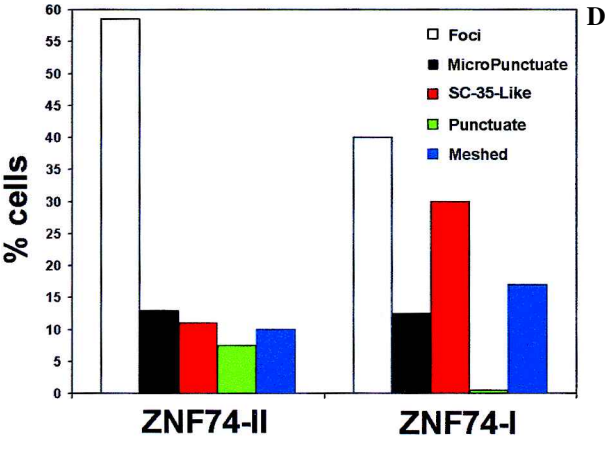
Repression by the KRAB box

Previous studies have shown that the KRAB box domain of a few characterized multifinger proteins is able to repress transcription when heterologously targeted to promotors (Witzgall *et al.*, 1994a; Margolin *et al.*, 1994; Vissing *et al.*, 1995; Pengue and Lania, 1996). To assess the potential repressor activity of the ZNF74 isoforms, the KRAB domain derived from both ZNF74-II, with a full-length KRAB box, and ZNF74-I, with a

FIG. 5. Differential nuclear partitioning of ZNF74-I and ZNF74-II isoforms as examined by digital deconvolution immunofluorescence microscopy. The SK-N-SH cells expressing HA epitope-tagged ZNF74-II or ZNF74-I were double-labeled with a rabbit anti-HA antibody (Cy3; green) and the mouse monoclonal antibody SC35 (Alexa 488; red). (**A**) Five different nuclear distribution profiles for ZNF74 protein are observed as illustrated for ZNF74-II isoform and are defined as foci (f), micropunctuate (g), SC-35 like (h), punctuate (i), and meshed (j). The corresponding staining of the cell nucleus with DAPI DNA-binding dye is shown (a–e). (**B**) HA-ZNF74 (1, p, t), SC-35 (m, q, u), and DAPI staining (k, o, s) are shown for cells where ZNF74-I (k–n) and ZNF74-II (o–r) show a SC-35-like distribution or where ZNF74-II is uniformly distributed in foci (s–v). In superimposed images (n, r, v), the colocalization of Cy3-labeled and Alexa 488-labeled structures appears yellow. (**C**) As an estimate of ZNF74 expression in the cells, the intensity of fluorescence was measured in a total of 400 cells (200 cells expressing ZNF74-I and 200 cells expressing ZNF74-II). The fluorescence intensity (±SD) was found to vary according to ZNF74 nuclear distribution profiles. Cells were grouped according to the five distribution profiles of ZNF74 illustrated in panel A (foci: N = 197; micropunctuate: N = 51; SC-35 like: N = 82; punctuate: N = 16; meshed: N = 54). For these five groups, cells expressing ZNF74-II were pooled with those expressing ZNF74-I, as their fluorescence intensity was not significantly different. (**D**) The percentage of cells exhibiting each of the five distribution profiles was determined for cells expressing ZNF74-II (N = 200) and ZNF74-I (N = 200).


A
a
b
c
d
e
DAPI
f
g
h
i
j
ZNF74-II
B
k
l
m
n
DAPI
ZNF74-I
SC-35
Merge
o
p
q
r
DAPI
ZNF74-II
SC-35
Merge
s
t
u
v
DAPI
ZNF74-II
SC-35
Merge


C
Relative Intensity
0
1
2
3
4
5
6
7
8
9
10
11
12
Foci
MicroPunctuate
SC-35 Like
Punctuate
Meshed


D
% cells
0
5
10
15
20
25
30
35
40
45
50
55
60
Foci
MicroPunctuate
SC-35-Like
Punctuate
Meshed
ZNF74-II
ZNF74-I

truncated A box, were fused in-phase with the DBD (aa 1–147) of the Gal4 transcription factor. Expression plasmids for these Gal4-DBD fusion proteins were transfected into L cells with a luciferase reporter carrying five Gal4-binding sites upstream of a TK promotor (p5xGal4-TK-luc). This reporter exhibited a basal level of transcriptional activity when transfected alone (not shown) or with the control plasmid encoding the Gal4 DBD alone (Fig. 6A). Transcriptional activity was significantly repressed by Gal4 DBD fused to the full-length KRAB domain of ZNF74-II (fourfold to fivefold) as well as to the KRAB domain of Kid1 (~16 fold) (Fig. 6A), a KRAB multifinger protein that was previously found to strongly repress transcription (Witzgall *et al.*, 1994b). However, the fusion protein with the truncated KRAB A box of ZNF74-I did not significantly repress transcription (Fig. 6A). As previously found for Kid1 and other Krab multifinger proteins (Margolin *et al.*, 1994; Kim *et al.*, 1996; Friedman *et al.*, 1996; Moosmann *et al.*, 1996), the repression obtained with Gal4 DBD fused to the full-length KRAB domain of ZNF74 was dependent on site-specific DNA binding, as demonstrated using a control reporter plasmid lacking Gal4-binding sites (pTK-luc) (Fig. 6B). Western analysis confirmed that all Gal4 fusion proteins exhibited the expected molecular mass and equivalent expression levels (not shown).

DISCUSSION

The present study revealed the existence of multiple transcripts for *ZNF74*, a member of the large and poorly characterized KRAB multifinger family, encoding a KRAB repressor domain and a tandem of 12 Cys_2His_2 zinc finger motifs with nucleic acid-binding activity. The protein isoforms encoded by these transcripts differ mainly in the presence of either a full-length and functional or a truncated and inactive KRAB repressor domain. Interestingly, while the ZNF74-I isoform, with a truncated KRAB domain, was previously found to associate with subnuclear domains enriched in splicing factors, the repressive isoform, ZNF74-II, was shown to partition differently and more diffusely in the cell nucleus. The differential localization of these ZNF74 isoforms with repressive and nonrepressive transcriptional activity is suggestive of distinct or complementary roles of the two isoforms in transcriptional and post-transcriptional regulation.

Protein isoforms resulting from alternative promotor usage and splicing of ZNF74

A systematic study of the various *ZNF74* cDNA forms obtained by cDNA library screening (Grondin *et al.*, 1996),

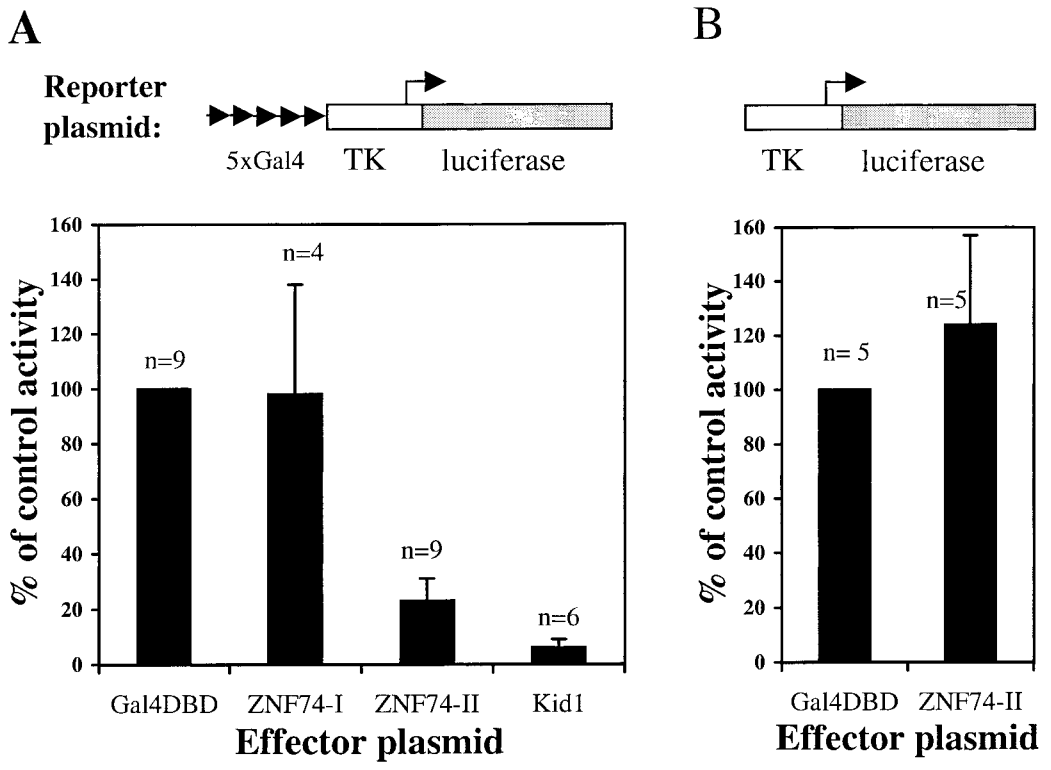


FIG. 6. Repressor activity of Gal4-ZNF74 fusion proteins. The L cells were transiently cotransfected with Gal4-DBD (aa 1–147) fusion constructs as effector plasmids and TK promoter-luciferase reporter plasmids. The effector plasmids include either Gal4 DBD alone (control) or in fusion with the truncated KRAB domain of ZNF74-I, the full-length KRAB domain of ZNF74-II, and the full-length KRAB domain of Kid1. The luciferase activity measured in cell extracts is normalized for transfection efficiency and is expressed relative to the luciferase activity obtained with the control Gal4-DBD (aa 1–147) effector plasmid, taken as 100%. The error bars indicate standard deviation from the mean. The number of independent experiments (N) is indicated. (**A**) Cotransfections of effector plasmids with p5xGal4-TK-luc reporter containing five copies of the Gal4 binding site placed in front of the basally active TK promotor. (**B**) Cotransfections of effector plasmids with pTK-luc control reporter plasmid.

RACE, and RT-PCR led to the identification of six transcripts. These transcripts, generated by usage of three possible promotors and alternative splicing, ultimately give rise to four protein isoforms ranging from 61 to 72 KDa. The previously identified *ZNF74-2* cDNA (Grondin *et al.*, 1996), which corresponds to the longest RNA species detected by Northern blot (Aubry *et al.*, 1993), encodes the ZNF74-I protein isoform characterized by the presence of a truncated KRAB box at its N terminus. Here, ZNF74-I was shown to be devoid of transcriptional repressive activity. The results presented here also indicate that this isoform can, in addition, be generated by alternative splicing of the longest RNA species initiated at the main P1 promotor or by alternative usage of another promotor (P2). We now also report that alternative splicing of the longest mRNA species, through skipping of an Alu sequence-containing exon, can generate ZNF74-II, a newly identified protein isoform containing a full-length KRAB box with transcriptional repressive activity. Furthermore, a transcript generated through the additional skipping of the KRAB B box exon was found to encode an isoform, ZNF74-III, with a full-length KRAB A box but no KRAB B box. The presence of a functional repressive KRAB domain and of zinc fingers in ZNF74 is suggestive of a physiologic role as a DNA-binding transcriptional regulator. Surprisingly, however, despite extensive efforts, no double-stranded specific consensus sequence has been identified for ZNF74 (Guévremont, 1999) or for any of the characterized KRAB multifinger proteins. Binding to heteroduplex DNA in a sequence-independent manner has, however, been described for the kid1 KRAB multifinger protein (Elser *et al.*, 1997). It thus remains to ascertain whether ZNF74, as well as other KRAB multifinger proteins, can function as *bona fide* DNA-binding transcriptional factors.

While the generation of isoforms without a KRAB B box by alternative splicing has been described for a few other KRAB multifinger proteins (references in Mark *et al.*, [1999]; Rosati *et al.*, [1991]), the generation of isoforms with a nonrepressive truncated KRAB box by alternative splicing as well as by alternative promotor usage is a precedent for members of the KRAB multifinger family. Even more significant is the fact that this isoform, in addition to being nonfunctional in terms of transcriptional repressive activity, has a different location from the repressive ZNF74-II, suggesting a different or complementary role. As suggested from our previous studies (Grondin *et al.*, 1996, 1997), and as discussed below, ZNF74 and, more specifically, ZNF74-I may participate in RNA metabolism.

Differential subnuclear localization of ZNF74 protein isoforms

Nuclear factors are often localized in discrete domains associated with specific cellular events such as DNA replication, transcription, or pre-mRNA splicing (references in Spector [1993]). In this study, we compared the subnuclear location of the newly identified ZNF74-II isoforms that encodes a full-length KRAB box with that of the previously identified ZNF74-I isoform by using deconvolution immunofluorescence microscopy. To express the ZNF74-I isoform with a truncated KRAB domain and determine its nuclear localization, we previously used COS-7 cells transformed with the SV40 virus lacking the viral origin of replication (Grondin *et al.*, 1997). Because the cloning vector (pCGN) for ZNF74 isoforms contains an SV40 origin of replication, and because episomal replication may affect the nuclear organization, revealing larger domains for replication, transcription, and splicing in a process similar to acute SV40 infection (references in Larsson *et al.* [1995]), we used human SK-N-SH neuroblastoma cells as well as nonimmortalized HS68 dermal fibroblasts that are not transformed with SV40 virus. As in COS-7 cells, the ZNF74-I truncated isoform was found in a large proportion of cells in association with the SC-35 antigen, a typical marker of speckles enriched in splicing factors (Spector, 1993). Although some studies suggest that these nuclear speckles may simply represent sites of storage for splicing factors (Zhang *et al.*, 1994; Jackson *et al.*, 1993; Wansink *et al.*, 1993; Cmarko *et al.*, 1999; Zeng *et al.*, 1997), others indicate that multiple functions related to RNA metabolism can take place within or just at the periphery of such domains (Xing *et al.*, 1995; Smith *et al.*, 1999; Spector, 1993; Huang and Spector, 1996; Jimenez-Garcia and Spector, 1993; Carter *et al.*, 1991). Thus, proteins such as ZNF74-I that are associated with SC-35 nuclear speckles, where components of the splicing machinery are concentrated, may participate in cotranscriptional or post-transcriptional RNA processing or RNA transport and stabilization (Larsson *et al.*, 1995; Smith *et al.*, 1999). The ZNF74-I isoform was found either to strictly colocalize with or to be just at the periphery of SC-35 nuclear speckles. This result is interesting, as it has been hypothesized from various observations that individual nuclear speckles are subcompartmentalized, with splicing of specific RNAs occurring at the periphery and other processes related to splicing complex assembly and RNA transport in the interior of the domain (references in Smith *et al.* [1999]). In contrast to the ZNF74-I isoform, ZNF74-II, the isoform with a full-length KRAB box, was preferentially expressed in uniformly diffuse foci distinct from nuclear speckles. In agreement with the proposed role of ZNF74-II in transcriptional regulation, this distribution is reminiscent of that observed for transcriptional regulators that exist in hundreds of small nuclear foci occurring both in association with and away from chromatin (Hendzel *et al.*, 1998). This difference in nuclear partitioning between ZNF74-I and ZNF74-II, although clearly evident from quantitative data, is not absolute. Such a marked but not absolute correlation between subnuclear localization and isoforms is reminiscent of the differential subnuclear localization of WT1 isoforms encoded by the Wilms' tumor zinc finger gene that depends on the presence or absence of a three-amino acid motif, KTS, located between the third and four zinc finger motifs (Larsson *et al.*, 1995; Ladomery *et al.*, 1999); only the +KTS isoforms preferentially colocalize with splicing factor-enriched nuclear speckles. In the case of WT1, this result represented the first indication that some isoforms of WT1 might have a post-transcriptional role. This idea was later corroborated by the finding of a stronger association of the +KTS isoforms with the U2AF65 splicing factor (Davies *et al.*, 1998) and RNA-binding activity of the WT1 zinc finger domain (Caricasole *et al.*, 1996). In contrast to the isoform of WT1, which loses its association with speckles after treatment with RNases, the ZNF74-I isoform, which has also been shown to have RNA-binding activity (Grondin *et al.*, 1996), remains associated with domains enriched in splicing factor after DNase and RNase treatment (Grondin *et al.*, 1997), suggesting that this associa-

tion is not nucleic acid dependent and could depend on protein–protein interaction.

A ubiquitously expressed corepressor, KAP-1, interacting with KRAB domains and mediating KRAB repression, has been shown to be localized in small punctuate heterochromatic and euchromatic domains throughout the nucleus (Ryan *et al.*, 1999; Friedman *et al.*, 1996). Future studies will address the possibility that ZNF74-II's predominant localization to uniformly diffuse domains unrelated to SC-35 nuclear speckles depends on its potential targeting to KAP-1 enriched domains. At present, it is not known if the presence of an intact KRAB box is sufficient for the predominant localization ZNF74-II to uniformly diffuse foci or if the conformation of part or all of the molecule is affected by the presence of the full-length KRAB motif, resulting in new molecular interactions and different subnuclear partitioning.

In recent years, it has been suggested that substantial crosstalk may take place between transcriptional and post-transcriptional processes (Misteli and Spector, 1998; Carmo-Fonseca *et al.*, 1999). In agreement with this notion, a few proteins with nucleic acid-binding activity have emerged as examples of potential multifunctional proteins with transcriptional and post-transcriptional roles (e.g., TFIIIA, WT1, hnRNP K, PTB) (reviewed in Ladomery [1997]). The data presented here suggest that, through its various isoforms, ZNF74 may also have a dual function in transcription and in RNA metabolism, a trait that may be shared by other members of the large KRAB multifinger family.

ACKNOWLEDGMENTS

We thank Dr. Marie Trudel and Karl Bergeron for critically reading this manuscript. This work was supported by grants from the Medical Research Council of Canada (MRC) (to MA and to DPB-J) and the Heart and Stroke Foundation of Canada (to MA). The Microscopy and Imaging Facility (University of Calgary) is supported by a maintenance grant from the MRC. FC was supported by an MRC fellowship, F-MB by an AHFMR studentship, BG by a Fonds de la Recherche en Santé du Québec (FRSQ) studentship, and MA by a scholarship from the FRSQ.

REFERENCES

ARRANZ, V., HARPER, F., FLORENTIN, Y., PUVION, E., KRESS, M., and ERNOULT-LANGE, M. (1997). Human and mouse MOK2 proteins are associated with nuclear ribonucleoprotein components and bind specifically to RNA and DNA through their zinc finger domains. Mol. Cell. Biol. **17,** 2116–2126.

AUBRY, M., DEMCZUK, S., DESMAZE, C., AIKEM, M., AURIAS, A., JULIEN, J.P., and ROULEAU, G.A. (1993). Isolation of a zinc finger gene consistently deleted in DiGeorge syndrome. Hum. Mol. Genet. **2,** 1583–1587.

BARDEESY, N., and PELLETIER, J. (1998). Overlapping RNA and DNA binding domains of the wt1 tumor suppressor gene. Nucleic Acids Res. **26,** 1784–1792.

BELLEFROID, E.J., PONCELET, D.A., LECOCQ, P.J., REVELANT, O., and MARTIAL, J.A. (1991). The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. Proc. Natl. Acad. Sci. USA **88,** 3608–3612.

BURN, J., TAKAO, A., WILSON, D., CROSS, I., MOMMA, K., WADEY, R., SCAMBLER, P., and GOODSHIP, J. (1993). Conotruncal anomaly face syndrome is associated with a deletion within chromosome 22q11. J. Med. Genet. **30,** 822–824.

CARICASOLE, A., DUARTE, A., LARSSON, S.H., HASTIE, N.D., LITTLE, M., HOLMES, G., TODOROV, I., and WARD, A. (1996). RNA binding by the Wilms' tumor suppressor zinc finger proteins. Proc. Natl. Acad. Sci. USA **93,** 7562–7566.

CARMO-FONSECA, M., CUSTODIO, N., and CALADO, A. (1999). Intranuclear trafficking of messenger RNA. Crit. Rev. Eukaryot. Gene Expression **9,** 213–219.

CARTER, K.C., TANEJA, K.L., and LAWRENCE, J.B. (1991). Discrete nuclear domains of poly(A) RNA and their relationship to the functional organization of the nucleus. J. Cell Biol. **115,** 1191–1202.

CMARKO, D., VERSCHURE, P.J., MARTIN, T.E., DAHMUS, M.E., KRAUSE, S., FU, X.D., VAN DRIEL, R., and FAKAN, S. (1999). Ultrastructural analysis of transcription and splicing in the cell nucleus after bromo-UTP microinjection. Mol. Cell. Biol. **10,** 211–223.

CROSS, S.H., CLARK, V.H., and BIRD, A.P. (1999). Isolation of CpG islands from large genomic clones. Nucleic Acids Res. **27,** 2099–2107.

DAVIES, R.C., CALVIO, C., BRATT, E., LARSSON, S.H., LAMOND, A.I., and HASTIE, N.D. (1998). WT1 interacts with the splicing factor U2AF65 in an isoform-dependent manner and can be incorporated into spliceosomes. Genes Dev. **12,** 3217–3225.

DEMCZUK, S., LEVY, A., AUBRY, M., CROQUETTE, M.F., PHILIP, N., PRIEUR, M., SAUER, U., BOUVAGNET, P., ROULEAU, G.A., THOMAS, G., and AURIAS, A. (1995). Excess of deletions of maternal origin in the DiGeorge/velo-cardio-facial syndromes: A study of 22 new patients and review of the literature. Hum. Genet. **96,** 9–13.

DUROCHER, D., CHARRON, F., WARREN, R., SCHWARTZ, R.J., and NEMER, M. (1997). The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. EMBO J. **16,** 5687–5696.

ELSER, B., KRIZ, W., BONVENTRE, J.V., ENGLERT, C., and WITZGALL, R. (1997). The Kruppel-associated box (KRAB)-zinc finger protein Kid-1 and the Wilms' tumor protein WT1, two transcriptional repressor proteins, bind to heteroduplex DNA. J. Biol. Chem. **272,** 27908–27912.

FRIEDMAN, J.R., FREDERICKS, W.J., JENSEN, D.E., SPEICHER, D.W., HUANG, X.P., NEILSON, E.G., and RAUSCHER, F.J. III. (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev. **10,** 2067–2978.

FU, X.D., and MANIATIS, T. (1990). Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. Nature **343,** 437–441.

GRONDIN, B., BAZINET, M., and AUBRY, M. (1996). The KRAB zinc finger gene ZNF74 encodes an RNA-binding protein tightly associated with the nuclear matrix. J. Biol. Chem. **271,** 15458–15467.

GRONDIN, B., CÔTÉ, F., BAZINET, M., VINCENT, M., and AUBRY, M. (1997). Direct interaction of the KRAB/Cys2-His2 zinc finger protein ZNF74 with a hyperphosphorylated form of the RNA polymerase II largest subunit. J. Biol. Chem. **272,** 27877–27885.

GUÉVREMONT, M. Interactions de la protéine à multidoigt de zinc ZNF74 avec les acides nucléiques (1999). (Dissertation, Université de Montréal).

HENDZEL, M.J., KRUHLAK, M.J., and BAZETT-JONES, D.P. (1998). Organization of highly acetylated chromatin around sites of heterogeneous nuclear RNA accumulation. Mol. Biol. Cell **9,** 2491–2507.

HUANG, S., and SPECTOR, D.L. (1996). Intron-dependent recruitment of pre-mRNA splicing factors to sites of transcription. J. Cell Biol. **133,** 719–732.

JACKSON, D.A., HASSAN, A.B., ERRINGTON, R.J., and COOK, P.R. (1993). Visualization of focal sites of transcription within human nuclei. EMBO J. **12,** 1059–1065.

JIMENEZ-GARCIA, L.F., and SPECTOR, D.L. (1993). In vivo evidence that transcription and splicing are coordinated by a recruiting mechanism. Cell **73,** 47–59.

KIM, S.-S., CHEN, Y.-M., O'LEARY, E., WITZGALL, R., VIDAL, M., and BONVENTRE, J.V. (1996). A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. Proc. Natl. Acad. Sci. USA **93,** 15299–15304.

KLUG, A., and SCHWABE, J.W. (1995). Protein motifs 5: Zinc fingers. FASEB J. **9,** 597–604.

KOZAK, M. (1991). An analysis of vertebrate mRNA sequences: Intimations of translational control. J. Cell Biol. **115,** 887–903.

LADOMERY, M. (1997). Multifunctional proteins suggest connections between transcriptional and post-transcriptional processes. BioEssays **19,** 903–909.

LADOMERY, M.R., SLIGHT, J., MCGEE, S., and HASTIE, N.D. (1999). Presence of WT1, the Wilms' tumor suppressor gene product, in nuclear poly(A)(+) ribonucleoprotein. J. Biol. Chem. **274,** 36520–36526.

LARSSON, S.H., CHARLIEU, J.-P., MIYAGAWA, K., ENGELKAMP, D., RASSOULZADEGAN, M., VAN HEYNINGEN, V., and HASTIE, N.D. (1995). Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. Cell **81,** 391–401.

MARGOLIN, J.F., FRIEDMAN, J.R., MEYER, W.K., VISSING, H., THIESEN, H.J., and RAUSCHER, F.J., 3rd. (1994). Kruppel-associated boxes are potent transcriptional repression domains. Proc. Natl. Acad. Sci. USA **91,** 4509–4513.

MARK, C., ABRINK, M., and HELLMAN, L. (1999). Comparative analysis of KRAB zinc finger proteins in rodents and man: Evidence for several evolutionarily distinct subfamilies of KRAB zinc finger genes. DNA Cell Biol. **18,** 381–396.

MISTELI, T., and SPECTOR, D.L. (1998). The cellular organization of gene expression. Curr. Opin. Cell Biol. **10,** 323–331.

MOOSMANN, P., GEORGIEV, O., LE DOUARIN, B., BOURQUIN, J.P., and SCHAFFNER, W. (1996). Transcriptional repression by RING finger protein TF1-beta that interacts with the KRAB repressor domain of KOX1. Nucleic Acids Res. **24,** 4859–4867.

PENGUE, G., and LANIA, L. (1996). Kruppel-associated box-mediated repression of RNA polymerase II promoters is influenced by the arrangement of basal promoters elements. Proc. Natl. Acad. Sci. USA **93,** 1015–1020.

PIELER, T., and BELLEFROID, E. (1994). Perspectives on zinc finger protein function and evolution: An update. Mol. Biol. Rep. **20,** 1–8.

POTT, U., THIESEN, H.J., COLELLO, R.J., and SCHWAB, M.E. (1995). A new Cys2/His2 zinc finger gene, rKr2, is expressed in differentiated rat oligodendrocytes and encodes a protein with a functional repressor domain. J. Neurochem. **65,** 1955–1966.

RAVASSARD, P., CÔTÉ, F., GRONDIN, B., BAZINET, M., MALLET, J., and AUBRY, M. (1999). ZNF74, a gene deleted in DeGeorge syndrome, is expressed in human neural crest-derived tissues and foregut endoderm epithelia. Genomics **62,** 82–85.

ROSATI, M., MARINO, M., FRANZE, A., TRAMONTANO, A., and GRIMALDI, G. (1991). Members of the zinc finger protein gene family sharing a conserved N-terminal module. Nucleic Acids Res. **19,** 5661–5667.

RYAN, R.F., SCHULTZ, D.C., AYYANATHAN, K., SINGH, P.B., FRIEDMAN, J.R., FREDERICKS, W.J., and RAUSCHER, F.J., III (1999). KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: A potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. Mol. Cell. Biol. **19,** 4366–4378.

SMITH, K.P., MOEN, P.T., WYDNER, K.L., COLEMAN, J.R., and LAWRENCE, J.B. (1999). Processing of endogenous pre-mRNAs in association with SC-35 domains is gene specific. J. Cell Biol. **144,** 617–629.

SPECTOR, D.L. (1993). Macromolecular domains within the cell nucleus. Annu. Rev. Cell Biol. **9,** 265–315.

THEUNISSEN, O., RUDT, F., GUDDAT, U., MENTZEL, H., and PIELER, T. (1992). RNA and DNA binding zinc fingers in Xenopus TFIIIA. Cell **71,** 679–690.

VISSING, H., MEYER, W.K., AAGAARD, L., TOMMERUP, N., and THIESEN, H.J. (1995). Repression of transcriptional activity by heterologous KRAB domains present in zinc finger proteins. FEBS Lett. **369,** 153–157.

WANSINK, D.G., SCHUL, W., VAN DER KRAAN, I., VAN STEENSEL, B., VAN DRIEL, R., and DE JONG, L. (1993). Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. J. Cell Biol. **122,** 283–293.

WITZGALL, R., O'LEARY, E., GESSNER, R., OUELLETTE, A.J., and BONVENTRE, J.V. (1993). Kid-1, a putative renal transcription factor: Regulation during ontogeny and in response to ischemia and toxic injury. Mol. Cell. Biol. **13,** 1933–1942.

WITZGALL, R., O'LEARY, E., LEAF, A., ONALDI, D., and BONVENTRE, J.V. (1994a). The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. Proc. Natl. Acad. Sci. USA **91,** 4514–4518.

WITZGALL, R., VOLK, R., YEUNG, R.S., and BONVENTRE, J.V. (1994b). Genomic structure and chromosomal location of the rat gene encoding the zinc finger transcription factor Kid-1. Genomics **20,** 203–209.

XING, Y., JOHNSON, C.V., MOEN, P.T.J., MCNEIL, J.A., and LAWRENCE, J.B. (1995). Non-random gene organization: Structural arrangements of specific pre-mRNA transcription and splicing with SC-35 domains. J. Cell Biol. **131,** 1635–1647.

ZENG, C., KIM, E., WARREN, S.L., and BERGET, S.M. (1997). Dynamic relocation of transcription and splicing factors dependent upon transcriptional activity. EMBO J. **16,** 1401–1412.

ZHANG, G., TANEJA, K.L., SINGER, R.H., and GREEN, M.R. (1994). Localization of pre-mRNA splicing in mammalian nuclei. Nature **372,** 809–812.

Address reprint requests to:
Dr. Muriel Aubry
Département de Biochimie
Université de Montréal
C.P. 6128, Succursale Centre-Ville
Montréal H3C 3J7, Canada

E-mail: Muriel.Aubry@UMontreal.ca

Received for publication November 28, 2000; accepted December 19, 2000.