BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **252,** 433–439 (1998)
ARTICLE NO. RC989656

Analysis of a Human cDNA Containing a Tissue-Specific Alternatively Spliced LIM Domain

T. Putilina,*,[1] C. Jaworski,* S. Gentleman,* B. McDonald,† M. Kadiri,* and P. Wong*,†
**Laboratory of Retinal Cell and Molecular Biology, National Eye Institute, National Institutes of Health, Bethesda, Maryland; and †Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada, T6C 2E9*

Received September 21, 1998

A unique clone, isolated from a human pancreatic cDNA library, was sequenced and characterized. Northern blot analysis showed that the gene is active in a number of fetal and adult tissues, and immunoblots showed expression in nuclear and cytosolic cell fractions. The gene corresponding to the clone was localized to chromosome 13 by human/rodent somatic cell hybrid panels. The largest open reading frame contains a LIM domain, and the deduced peptide from the open reading frame appears to have the characteristics of a LIM-only protein, designated *LMO7*. RT-PCR and genomic sequence analyses indicate that expression of this gene product is subject to tissue-specific modulation by elimination of the LIM domain by alternative splicing in neural tissues.

The LIM domain is a cysteine-rich structural motif first described in the gene products of *Caenorhabditis elegans* lin-11 (1), rat ISL-1 (2), and *C. elegans* mec-3 (3), from which the domain derived its name. There are three general classes of LIM proteins (5). The first class described consists of proteins with two tandem repeats of the LIM domain followed by a homeodomain (e.g., lin-11); the other two classes have one to three LIM domains with the remainder of the various proteins being quite dissimilar (4-6). These classes include examples with protein kinase domains, actin-binding domains and in one case an apparent linker sequence between two unrelated LIM domains (6). In general, the LIM motif, which is a double zinc-finger domain (6,7), is the only common feature among the various LIM proteins. Members of all classes have been implicated in a variety of functions (e.g., transcription, differentiation, cytoskeletal interactions and signal transduction), but all act through protein-protein interactions (4-11).

In this communication, we characterize a novel human cDNA clone containing a single LIM domain, that undergoes alternative splicing in neural tissues. Genomic sequence analysis shows that the LIM domain is encoded in two exons; elimination of these exons by alternative splicing removes the zinc fingers of the LIM domain and additionally changes the downstream reading frame. With the exception of the LIM domain, which falls into a subclass of C-terminal LIM domains (5,6), this gene has no similarity to other known genes.

MATERIALS AND METHODS

cDNA clone isolation and DNA sequencing. Clones were isolated from a human pancreatic cDNA phage library in SWAJ-2 cloning vector (Clontech Laboratories, Inc., Palo Alto, CA) by standard protocols (12). Clones isolated were subcloned into PCR tm II plasmids with a TA cloning kit (Invitrogen Corp., San Diego, CA) prior to sequencing.

DNA sequencing was performed with a 373A automated sequencer using the PRISM ready reaction dye deoxy terminator cycle sequencing kit (Applied Biosystems, Foster City, CA). Some sequences were also determined by Lofstrand Laboratories (Rockville, MD).

Sequence analysis. DNA sequences were analyzed with programs in the Genetics Computer Group Sequence Analysis Software Package, Ver.8 (GCG, Inc., Madison, WI), the Lasergene DNAstar package (Madison, WI) and the BLAST network service at NCBI, NLM, Bethesda, MD.

Northern and Southern blot analysis. Southern blots of genomic DNA from multiple species and northern multiple tissue blots of adult and fetal human RNA were obtained from Clontech Laboratories, Inc. (Palo Alto, CA). In all cases, hybridization was carried out at 65 °C in Hybrisol II and washed at the same temperature through a series of washes to a final stringency of 0.1x SSC with 0.1% SDS for 15 min. Probes were made by random priming of clone 429186 (Research Genetics, Huntsville, AL) which is from the Soares fetal liver library, accession number AA005111. The sequence of this clone is 100% identical to bases 116 through 710 of LMO7.

Chromosomal localization. Chromosomal localization was performed by BIOS Laboratories (New Haven, CT) using the NIMGS human/rodent somatic hybrid mapping panels I and II and human/

[1] Present address: Institute for Environmental Medicine, University of Pennsylvania, 1 John Morgan Bldg., 3620 Hamilton Walk, Philadelphia PA 19104-6068.

0006-291X/98 $25.00

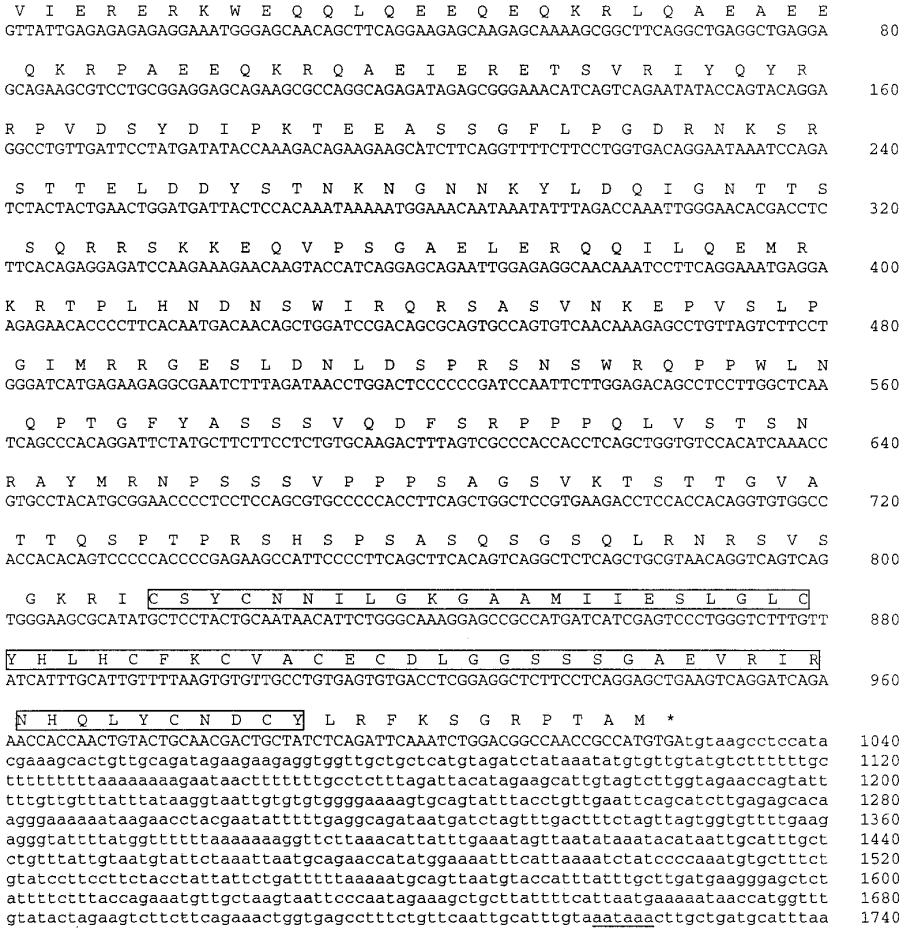

FIG. 1. Sequence of LMO7 cDNA. The nucleotide sequence of the partial cDNA for LMO7 is shown with coding sequence in uppercase and the 3'-untranslated region in lowercase. The deduced peptide sequence is in single letter code, with amino acids centered above each codon. A poly-adenylation signal is underlined. The boxed region contains sequence that fits the consensus for a zinc finger LIM domain.

rodent panels from the Coriell Institute. Probes derived from cloned inserts were prehybridized with excess amounts of human Cot-1 DNA (Life Technologies Inc., Gaithersburg, MD) prior to use on blots to suppress background hybridization by repetitive DNA sequences present within the probe sequence.

PCR analysis. For PCR of genomic templates, 200 ng of human genomic DNA (Clontech) was amplified with the Expand Long Template PCR system (Boehringer Mannheim), Indianapolis, IN) as suggested by the manufacturer, using primer concentrations of 0.3 μM. Samples were denatured at 94 °C for 10 seconds, followed by 30 seconds at 60 °C, then a 6 minute extension period at 68 °C, for a total of 10 cycles. An additional 20 cycles were carried out in which each successive cycle contained an additional 5 second increment for the extension step.

RT-PCR employed a solid-phase immobilized first strand cDNA produced as described elsewhere, and kindly provided by Dr. Ignacio Rodriguez (13). Equal amounts of template from various tissues were amplified using AmpliTaq (Perkin-Elmer Cetus Applied Biosystems, Foster City, Ca.) according to the manufacturer's instructions. Primers were used at final concentrations of 0.8 uM, and the PCR profile consisted of 10 seconds at 94 degrees, 30 seconds at 60 °C, and 1 minute at 72 °C for a total of 30 cycles. All primer sequences are available upon request.

Subcellular fractionation. Fresh Rhesus monkey lung tissue was homogenized in a 10X volume of 10 mM HEPES buffer (pH 7.2) containing 5 mM $MgCl_2$, 4% (w/v) sucrose and one Complete Inhibitor tablet/50 ml (Boehringer-Mannheim, Rockford, IL). The homogenate was centrifuged at 300xg for 10 minutes. The pellet (P1) was resuspended in buffer containing 0.5 M NaCl and incubated on ice for 30 minutes. The supernatant was centrifuged at 27,000xg for 20 minutes, and the supernatant fraction was reserved. The pellet (P2) was incubated in the buffer with the addition of 2% Triton-X100 on ice for 30 minutes and centrifuged for 20 minutes at 27,000xg. The detergent-soluble supernatant (P2d) was reserved, and the pellet fraction was extracted with buffer containing 0.5 M NaCl as previously and recentrifuged at 27,000xg for 20 minutes. The salt-soluble supernatant (P2s) was reserved and the final pellet discarded.

SDS polyacrylamide electrophoresis and immunoblotting. 12% polyacrylamide gels were prepared using a modification of buffer System 12 (14). The catholyte was 0.113 M bisTris/0.044 M TES/0.15 SDS (pH 7.25), the anolyte was 0.063 M bisTris/HCl (pH 5.9) and the gel buffer was 0.123 M bisTris/HCl (pH 6.61). Samples were solubilized in gel buffer containing 2% SDS and 25 mM β-mercaptoethanol at room temperature. 10 mM β-mercaptoethanol was added to the catholyte at the beginning of the electrophretic run. After electrophoretic separation, proteins were electroblotted onto Immobilon-P

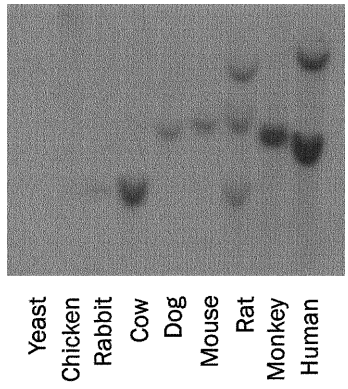

FIG. 2. Southern blot of multiple species genomic DNA probed with Research Genetics clone 429186. One week exposure. Hybridization conditions are described under Methods.

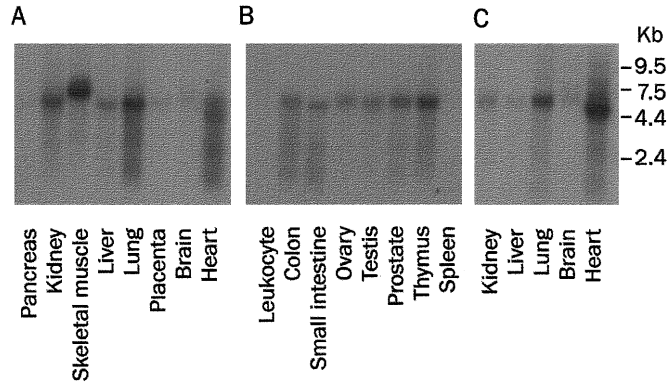

FIG. 3. Northern blots. (A) Human adult multiple tissue panel, MTN-1. (B) Human adult multiple tissue panel, MTN-2. (C) Human fetal multiple tissue panel. Probe and hybridization conditions as described in Figure 2.

membranes (Millipore Corp., Bedford, MA) and probed with anti-peptide antiserum. An antibody to a MAP peptide, consisting of residues 201-220 of the deduced peptide, was raised in rabbits (Anaspec Inc., San Jose, CA). The blots were developed with the Tropix Western Light Kit (Tropix PE Applied Biosystems, Bedford, MA), using CDP-Star as the substrate, according to the manufacturer's instructions.

RESULTS

Characterization of the LMO7 cDNA

The LMO7 cDNA sequence (Fig. 1) shows no apparent similarity with other known gene sequences; however, it matches more than forty overlapping expressed sequence tags (ESTs) from human cDNA libraries with 98-100% identity. The sequence from the original LMO7 cDNA clone was extended in both directions using overlapping ESTs from pancreatic cDNA libraries; this additional sequence is represented in Genbank entry U90654. The longest open reading frame for the sequence extends from base 3-1023; a polyadenylation signal is present at bases 1738-1743. However, the 5′ end is undetermined at this time.

A multiple species southern blot revealed high stringency cross-hybridization of the labeled clone with human and monkey, with lesser hybridization with other mammalian species (Fig. 2). The gene was localized to chromosome 13 on the Coriell Institute somatic mapping panel 2 and this assignment was confirmed with the NIGMS human/rodent panels 1 and 2. This gene has been designated LMO7 by the human nomenclature committee.

Expression of LMO7

Multiple tissue northern blots of human RNA were probed with Research Genetics Clone 429186 (Fig. 3) and showed a 6 Kb message widely expressed in adult and fetal tissues. Strong signals were seen in both adult and fetal lung and heart, with the heart showing also a 4.8 Kb message, which was the stronger signal in fetal heart. A 7 Kb message was seen only in skeletal muscle. Identical results were obtained in blots probed with the Cot-1-suppressed probe from the original clone (data not shown).

An antibody was made to a peptide from the predicted sequence of LMO7 and used to probe an immunoblot of Rhesus monkey lung subcellular fractions. Analysis showed immunoreactive bands with apparent molecular sizes of 31 kDa and 16.5 kDa in the nuclear-enriched fraction, and a band of 70 kDa in the cytosolic fraction (Fig. 4). The 16.5 kDa band of the nuclear fraction may be a proteolytic fragment. A faint 70 kDA band was also seen in the detergent-soluble membrane fraction.

LIM Domain

Examination of the predicted polypeptide sequence revealed the presence of a conserved cysteine-rich LIM domain (bases 814-990) (Fig. 1). The human LMO7 LIM domain is shown in alignment with the LIM do-

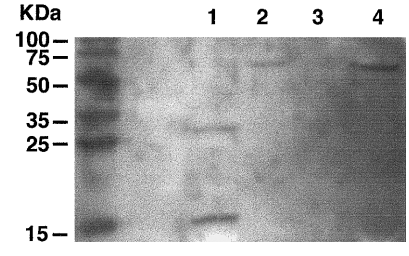

FIG. 4. Immunoblot of monkey lung subcellular fractions. Fractions were separated on a 12% SDS polyacrylamide gel as described. Lane 1, 300 xg pellet; lane 2, detergent extract of 27,000 xg pellet; lane 3, salt extract of 27,000 xg pellet; lane 4, 27,000 xg supernatant. Markers are Novagen Perfect Protein markers.

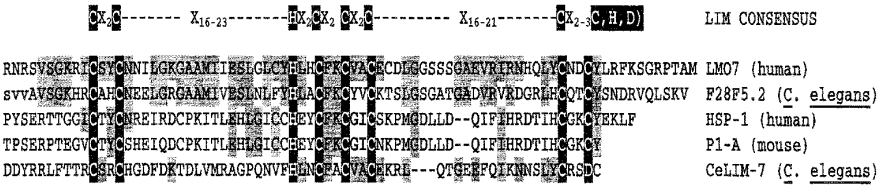

FIG. 5. LIM consensus sequences. The consensus sequence that defines LIM domains is shown above the deduced peptide sequence of LMO7. Key residues (cysteines and a histidine), indicated by black boxes, are all present in LMO7. Below, the LIM domain of human LMO7 is shown in alignment with the most similar of other LIM domains from proteins found in GenBank.

main consensus (upper portion of Fig. 5). The eleven residues conserved in all LIM domains are present in LMO7, and the length of distances between conserved residues are also in accordance with the consensus. The lower portion of Fig. 5 compares the LMO7 LIM domain with the most similar LIM domains found in Genbank. The best match, with 57% identity and 79% similarity, is from a putative protein from *C. elegans* (gi|470355). LIM domains of HSP1 (gi|2088537) and P1-A (gi|1401076), orthologous proteins from human and mouse, are 57 and 60% similar to LMO7, respectively. The LIM domain of CeLIM-7 (gi|1658400), another *C. elegans* protein, is 52% similar. All these proteins also have the LIM domain located in the C-terminal region. Thus, human LMO7 appears to be a novel gene, but the detailed design of its LIM domain is similiar to those of other proteins that are widely distributed in nature.

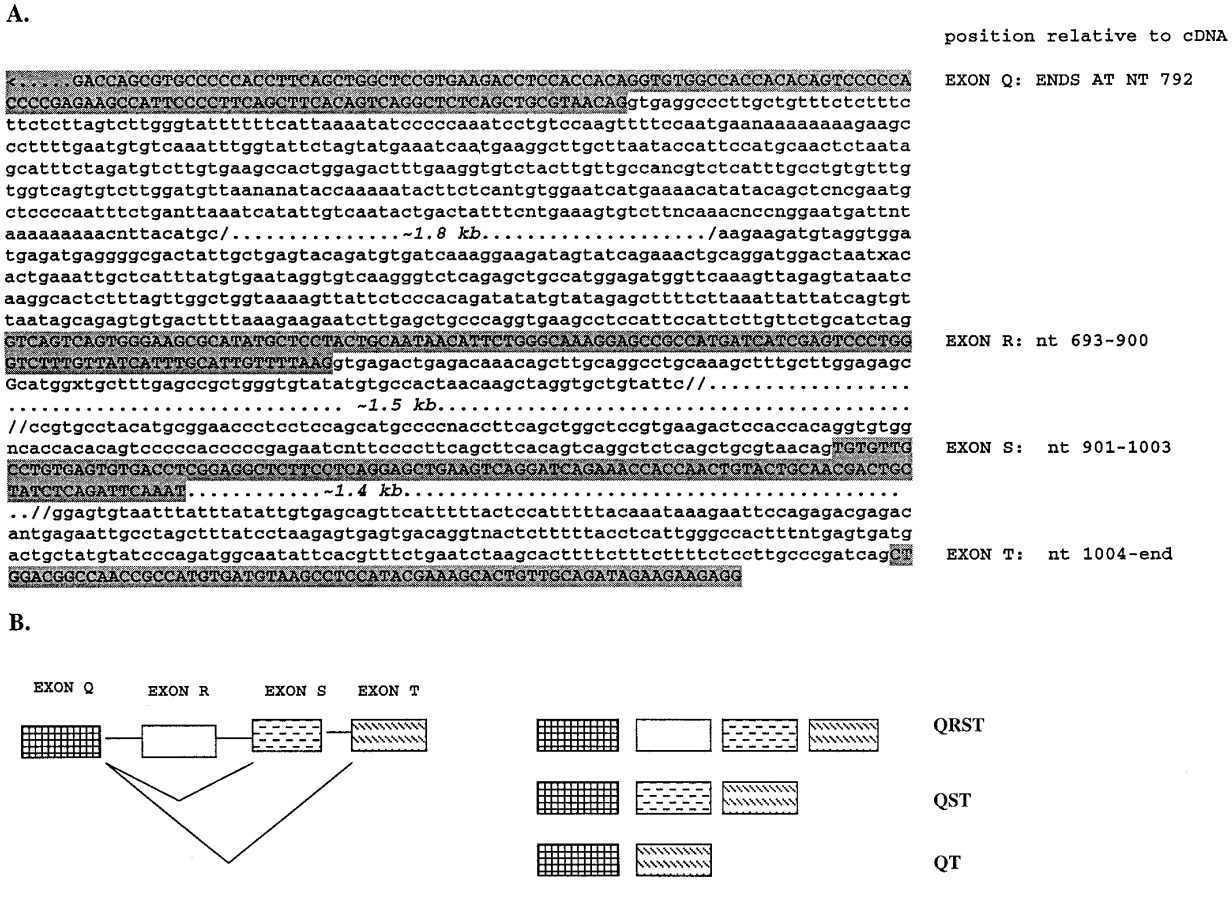

FIG. 6. (A) Genomic DNA sequence of the LIM-containing region. Coding sequences are highlighted and shown in uppercase; intron sequences are in lowercase. Accession numbers for the exons with intron sequence are: Q, AF092554; R, AF092555; S, AF092556; T, AF092557. The nucleotide positions covered by each exon, relative to the cDNA sequence, are indicated on the right. The 5′ boundary of exon Q is not determined. (B) Schematic view of splicing. The LMO7 transcripts produced by alternative splicing are depicted in schematic form, designated by the exons included in each form.

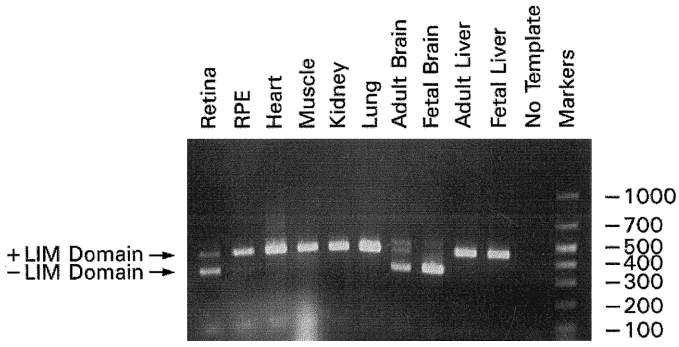

FIG. 7. RT-PCR of human tissue mRNAs. LMO7 expression was examined in various tissues by RT-PCR, using primers that flank the LIM domain. The upper band encompasses the complete LIM domain. Lower bands represent the 2 alternatively spliced forms that eliminate the LIM domain.

Splice Variants of LMO7

Comparison of the LMO7 cDNA with ESTs in GenBank suggests that the LMO7 transcript undergoes alternative splicing within the region encoding the LIM domain. Four ESTs from an infant human brain (accession numbers: N89937, T92191, D31521, and T94037) showed 100% identity with LMO7 from nucleotide 559 through base 1094 but were lacking the sequence from bases 792 to 900. Potential splice sites at bases 790-793 (-AGGT-) and 899-902 (-AGTG-) border the region missing in these ESTs, supporting the idea that the ESTs represent alternatively spliced forms of the gene.

We used two approaches to investigate the possibility of alternative splicing of LMO7. First, we examined the region of the LMO7 gene encoding the LIM domain. PCR products were generated from genomic DNA using oligonucleotide primers flanking and within the LIM domain region. These products were sequenced directly. The genomic sequence, shown in Fig. 6A, reveals the presence of 4 exons in this region, designated Q, R, S, and T (from 5′ to 3′). The intron/exon junctions for exon R precisely match the splice sites suggested by the EST data, and the QST form is identical to the sequences reported for infant brain ESTs. A schematic of the possible alternative splice forms in this region of the LMO7 gene is shown in Figure 6B.

In a second approach, LMO7 expression was examined in a number of tissues by RT-PCR, using primers flanking the LIM domain in order to detect the presence of alternatively spliced forms. As shown in Fig 7, LMO7 expression is widely distributed. However, the pattern of splicing differs markedly in a tissue-specific manner. In addition to the form encoding the entire LIM domain (QRST), 2 variants were found by RT-PCR and verified by DNA sequencing. In one, exon R is absent (QRT), and in the other, exons R and S are both eliminated (QT). In neural tissue (retina and brain) the

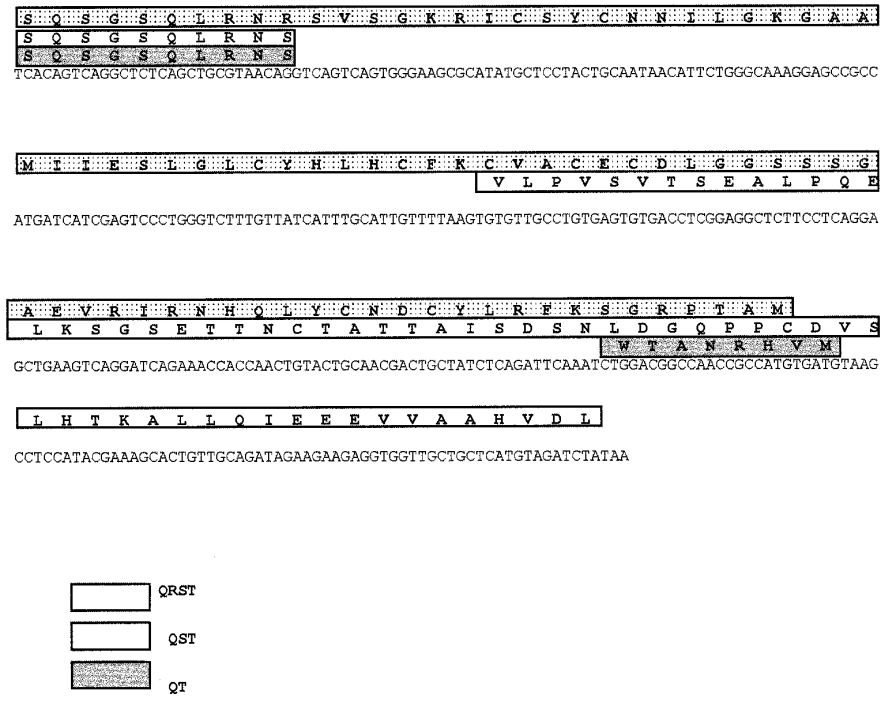

FIG. 8. Sequences of alternative transcripts. The DNA sequences of each LMO7 transcript are shown above. The deduced peptide sequences of the alternatively spliced products are listed below. The LIM domain sequence is encoded in exons R and S.

alternatively spliced variants that remove the LIM domain are predominant, while non-neural tissues appear only to express the form containing the entire LIM domain.

The number of nucleotides in exons R and S are not multiples of 3; therefore, the reading frame shifts if either of these exons is omitted (Fig. 8). The loss of exon R results in elimination of the LIM domain, giving a deduced protein about 1 kDa smaller than the canonical form, and removal of both exons R and S resulted in a deduced protein about 7 kDa smaller.

DISCUSSION

In this report we describe a clone coding for a novel LIM protein, LMO7. The LIM domain of the putative peptide deduced from this clone fits the LIM consensus sequence with absolute conservation of the seven cysteines and histidine defining the two thiolate coordination Zn(II) sites. In addition, aromatic residues before the histidine, after the third cysteine and after the seventh cysteine and a leucine in the fourth position past the fifth cysteine are conserved (4-8). Although the loops of these sites are somewhat atypical compared with the majority of LIM proteins (longer and of dissimilar lengths), they are nonetheless similar to LIM domains from other human, mouse and *C. elegans* proteins (see Fig. 5). The deduced peptide is also similar to zyxin and other cytosolic LIM domain proteins in that the LIM domain is located in the carboxyl terminus (5,6,9,10). In addition, the antibody recognition of an upstream peptide sequence (see Fig. 4) further supports the conclusion that a protein with the deduced sequence is indeed expressed.

LIM domains play roles in protein activity and targeting through protein-protein interactions (16-19). This is demonstrated by the in vitro expression of the Xlim1 gene of *Xenopus*. The homeodomain activity of this gene is enhanced by mutation or deletion of the LIM domains, which act as negative modulators through their ability to bind to associated proteins in the transcriptional complex (16). Likewise, LIM proteins of the cytoplasmic classes have often been shown to associate through the LIM domain with specific protein targets. For example, the second LIM domain of the muscle LIM protein (MLP) is specifically responsible for interaction with the actin cytoskeleton (18), and both MLP and zyxin exhibit LIM-LIM interactions with CRPs (cysteine-rich proteins) (17,18). These characteristics of LIM proteins suggest that LIM domains have specific roles in the assembly of multiprotein complexes. We have identified a LIM domain subgroup likely to have similar targeting, although the dissimilarity of these proteins outside the LIM domain suggests very different functions for the proteins overall.

Several LMO7 ESTs from infant brain lack the LIM domain, suggesting that LMO7 undergoes alternative splicing. We have demonstrated that alternative splicing does occur in LMO7 transcripts and further shown that the alternative splicing is specific to neural tissue. Two exons make up the LIM domain in LMO7, and sequencing of RT-PCR products revealed the presence of two alternatively spliced forms. In both forms the LIM domain is entirely eliminated, although with different consequences to the carboxyl terminal of the resulting polypeptide (see Figs. 6 and 8). The elimination of the LIM domain could radically change the distribution and/or stability of the LMO7 protein in vivo. Thus, the alternative transcripts of LMO7 which determine the presence or absence of the LIM domain, and the differential expression of these transcripts, are intriguing in this respect.

The LIM domains of the human HSP1 and the mouse P1-A orthologues are also composed of two exons and have multiple transcripts, possibly resulting from alternative splicing (15). This characteristic stands in contrast to genes such as the LIM-homeobox genes Lhx1, Lhx3 and Lhx5, in which the domain is contained within one exon (20). Therefore, regulation of LIM-dependent associations by alternative splicing may be a general characteristic in LMO7-type LIM proteins.

ACKNOWLEDGMENTS

The authors thank Dr. Peter Fitzgerald (DCRT, NIH) for help and advice on GenBank data analysis and Drs. Paul Russell and Graeme Wistow (National Eye Institute, NIH) for many helpful discussions and suggestions. We are particularly grateful to Dr. Ignacio Rodriguez for supplying immobilized cDNA libraries and assistance in sequencing. Dr. Barbara Wiggert (National Eye Institute, NIH) has been most generous in support and encouragement in the writing of the manuscript.

REFERENCES

1. Freyd, G., Kim, S. K., and Horvitz, H. R. (1990) *Nature* **344,** 876–879.
2. Karlsson, O., Thor, S., Norberg, T., Ohlsson, H., and Edlind, T. (1990) *Nature* **344,** 879–882.
3. Way, J. C., and Chalfie, M. (1988) *Cell* **54,** 5–16.
4. Sanchez-Garcia, I., and Rabbitts, T. H. (1994) *Trends Genetics* **10,** 315–320.
5. Dawid, I. B., Toyama, R., and Taira, M. (1995) *C. R. Acad. Sci Paris* **318,** 295–306.
6. Dawid, I. B., Breen, J. J., Toyama, R. (1998) *Trends Genetics* **14,** 156–162.
7. Michelsen, J. W., Schmeichel, K. L., Beckerle, M. C., and Winge, D. R. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 4404–4408.
8. Archer, V. E. V., Breton, J., Sanchez-Garcia, I., Osada, H., Forster, A., Thomson, A. J., and Rabbitts, T. H. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 316–320.

9. Arber, S., Halder, G., and Caroni, P. (1994) *Cell* **79,** 221–231.
10. Sadler, I., Crawford, A. W., Michelsen, J. W., and Beckerle, M. C. (1992) *J. Cell Biol.* **119,** 1573–1587.
11. Nix, D. A., and Beckerle, M. C. (1997) *J. Cell Biol.* **138,** 1139–1147.
12. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY.
13. Schoen, T. J., Mazuruk, K., Chader, G. J., and Rodriguez, I. R. (1995) *Biochem. Biophys. Res. Commun.* **213,** 181–189.
14. Chrambach, A., and Jovin, T. M. (1983) *Electrophoresis* **4,** 190–204.
15. Taira, M., Otani, H., Saint-Jeannet, J.-P., and Dawid, I. B. (1994) *Nature* **372,** 677–679.
16. Schmeichel, K. L., and Beckerle, M. C. (1994) *Cell* **79,** 211–219.
17. Arber, S., and Caroni, P. (1996) *Genes Devel.* **10,** 289–300.
18. Macalma, T., Otte, J., Hensler, M. E., Bockholt, S. M., Louis, H. A., Kalff-Suske, M., Grzeschik, K.-H., von der Ahe, D., and Beckerle, M. C. (1996) *J. Biol. Chem.* **271,** 31470–31478.
19. Heiss, N. S., Gloeckner, G., Baechner, D. Kioschis, P., Klauck, S. M., Hinzmann, B., Rosenthal, A., Herman, G. E., and Poustka, A. (1997) *Genomics* **43,** 329–338.
20. Bertuzzi, S., Sheng, H. Z., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Taira, M., Dawid, I. B., and Westphal, H. (1996) *Genomics* **36,** 234–239.