J Hum Genet (1999) 44:256–260

SHORT COMMUNICATION

Nobuhide Ueki · Naohiko Seki · Kazuhiro Yano
Yasuhiko Masuho · Toshiyuki Saito
Masa-aki Muramatsu

Isolation, tissue expression, and chromosomal assignment of a human LIM protein gene, showing homology to rat Enigma homologue (ENH)

Received: February 25, 1999 / Accepted: April 3, 1999

Abstract Rat ENH (Enigma homolog) is a LIM domain protein that associates with protein kinase C in an isoform-specific manner. We have identified a human cDNA which shares a significant sequence homology with rat *ENH*. The isolated cDNA clone, designated human *ENH* (*hENH*), was 3287 bp in length and encoded a predicted protein of 596 amino acids which had 88% overall identity to rat ENH protein. Northern blot analysis revealed that 1.9 kb of the *hENH* messenger RNA was predominantly expressed in heart and skeletal muscle, while 5.6 kb of the *hENH* messenger RNA was ubiquitously expressed in various human tissues. The chromosomal location of the gene was determined on chromosome 4q22 region, between markers WI-2900 and WI-3273, by polymerase chain reaction (PCR)-based analyses using both a human/rodent monochromosomal hybrid cell panel and a radiation hybrid mapping panel.

Key words LIM domain · Chromosome mapping · 4q22 · Enigma · Enigma homolog (ENH)

Introduction

The LIM domain is a Cys-rich domain composed of 50–60 amino acid residues with the consensus sequence (Cys-X_2-Cys-$X_{17–19}$-His-X_2-Cys)-X_2-(Cys-X_2-Cys-$X_{17–19}$-Cys-X_2-His/Asp/Cys) (Sanchez-Garcia and Rabbitts 1994; Dawid et al. 1995, 1998). LIM domain proteins have been implicated in developmental regulation, cellular differentiation, and actin-based cytoskeletal interaction (Sadler et al. 1992; Crawford et al. 1994; Sanchez-Garcia and Rabbitts 1994). Spectroscopic studies of LIM domains derived from a number of different proteins have revealed that the LIM domain specifically coordinates two zinc ions (Michelsen et al. 1993, 1994; Archer et al. 1994; Kosa et al. 1994). Although many zinc finger motifs are involved in DNA or RNA binding (Klug and Schwabe 1995), the LIM domain has been proposed to participate in protein-protein interactions (Sanchez-Garcia and Rabbitts 1994; Dawid et al. 1995, 1998).

LIM domain proteins are classified according to sequence relationships among LIM domains and on the overall structure of the proteins (Dawid et al. 1998). Group 1 proteins include LIM-homeodomain proteins which are developmentally regulated transcription factors. *C. elegans* Lin-11, rat Isl-1, and *C. elegans* Mec-3 were the first identified LIM proteins from which the term LIM (Lin-11, Isl-1 and Mec-3) is derived (Way and Chalfie 1988; Freyd et al. 1990; Karlsson et al. 1990). Group 2 proteins include cystein-rich intestinal protein (CRIP; accession number, M33146) (Liebhaber et al. 1990) and cystein-rich protein (CRP; accession number, M76375) (Wang et al. 1992), which contain a short additional conserved motif after the LIM domain. Group 3 proteins include Ril (accession number, X76454) (Kiess et al. 1995), paxillin (accession number, U14588) (Salgia et al. 1995), zyxin (accession number, X94991) (Macalma et al. 1996), Enigma (accession number, L35240) (Wu and Gill 1994), and ENH (accession number, U48247) (Kuroda et al. 1996). Although this group is heterogeneous, the LIM domains in this group are more closely related to each other than to those in groups 1 and 2.

We previously isolated a partial cDNA clone encoding a potential LIM domain during the course of screening for nuclear proteins (Ueki et al. 1998). We describe here the full sequence, expression profile, and chromosomal assignment of a human ENH cDNA.

N. Ueki · K. Yano · Y. Masuho · M. Muramatsu (✉)
Biological Technology Laboratory, Helix Research Institute, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel. +81-438-52-3966; Fax +81-438-52-3952
e-mail: mmasaaki@hri.co.jp

N. Seki · T. Saito
Genome Research Group, National Institute of Radiological Sciences, Anagawa, Chiba, Japan

M. Muramatsu
Department of Biological Cybernetics, Medical Research Institute, Tokyo Medical Dental University, Tokyo, Japan

Materials and methods

Isolation and sequencing of the human *ENH* gene

A partial cDNA clone encoding a potential LIM domain was initially isolated using the nuclear transportation trap method (Ueki et al. 1998). Specific primers were designed according to the cDNA sequence to obtain the full-length cDNA from a human fetal brain library, using GeneTrapper (GIBCO BRL, Gaithersburg, MD, USA). The nucleotide sequences of both strands were determined by the primer walking method. Homology searches in the public databases were performed routinely with the FASTA program of the UWGCG package.

Northern blot analysis

Northern blot filters containing adult human poly (A)+ RNAs (2 μg/lane) were purchased from Clontech (Palo Alto, CA, USA), and hybridization and washing were performed according to the manufacturer's instructions. The gene-specific polymerase chain reaction (PCR) primers were designed within the coding region (5′-TCT CCT GCT GTG TCC AAA GTC-3′, 5′-ACG CCA GTC TTC AGT ATC CTC-3′, with the PCR product size 486 bp). The PCR amplified fragment was labelled with [α-^{32}P] dCTP and used as a hybridization probe.

Chromosome mapping

Chromosomal assignment of the human *ENH* gene was done by PCR analysis of a human/rodent somatic cell hybrid panel (Mapping Panel #2; Coriell Cell Repositories, Camden, NJ, USA) and a radiation hybrid panel (Genebridge 4; Research Genetics, Huntsville, AL, USA). The gene-specific PCR primers were designed within the coding region of the gene (5′-GAC AAA CCC AGC CAA GTG ACC-3′, 5′-GCG CAC ATC GGA GTT CGT TTC-3′, with the PCR product size 82 bp). PCR was carried out in the same manner as that described in previous reports (Seki et al. 1997; Saito et al. 1998).

Results and discussion

The isolated cDNA clone (HFB230) was 3287 bp in length and the nucleotide sequence will appear in DDBJ, EMBL, and GenBank nucleotide sequence databases under the accession number AF061258. The HFB230 clone contained an open reading frame for a predicted protein of 596 amino acid residues with a calculated molecular weight of approximately 63 kiloDalton. The first ATG was preceded by an in-frame stop codon, indicating that the entire coding region is covered. Homology search for relevant amino acid sequences in the protein database (NIBI Protein Database and Swiss Protein Database) revealed that HFB230 protein was most homologous to rat ENH (accession number, U48247) (Kuroda et al. 1996), having 88% identity at the amino-acid level. Thus, we renamed this clone human ENH (Enigma homolog). The alignment of the deduced amino acid sequences of human ENH, rat ENH, Enigma, and a hypothetical protein KIAA613 is shown in Fig. 1. Two Pro/Ser-rich regions (residues 106–216 and 308–394) and three LIM domains (residues 415–585) were observed in both human and rat ENH proteins. Residues 415–585, including the LIM domain of hENH, also shared a high similarity to the corresponding domains of Enigma (Wu and Gill 1994, accession number, L35240) and of the hypothetical protein KIAA613 (Ishikawa et al. 1998, accession number, AB014513), which were 57% and 66% identical at the amino-acid level, respectively.

The rat ENH was initially identified during the screening for protein kinase C (PKC)-interacting protein by a yeast two-hybrid assay (Kuroda et al. 1996). Deletion analysis demonstrated that any single LIM domain of rat ENH can associate with PKCβI, γ, and ε, but not with PKCα, δ, and ζ (Kuroda et al. 1996). Interestingly, the LIM domains of Enigma also associated with the PKCα, β, and ζ but not with PKCγ, δ, and ε (Kuroda et al. 1996). Protein kinase C (PKC) defines a family of serine/threonine kinases which play critical roles in the regulation of cellular differentiation and proliferation in many cell types (Nishizuka 1992). Members of the PKC family are typically activated by the second messenger diacylglycerol and participate in cellular responses to various agonists, such as hormones, neurotransmitters, and growth factors (Bell and Burns 1991; Nishizuka 1992; Hug and Sarre 1993; Dekker and Parker 1994). The association of different LIM domains with PKC subtypes is expected to shed light on the understanding of specific functions of PKC.

We then examined the tissue distribution of *hENH* by northern blot analysis. Hybridization of human RNAs from multiple tissue blots (MTNblot, I, II; Clontech) with a *hENH* cDNA probe detected two mRNA species that migrated as 1.9-kb and 5.6-kb bands. The 1.9-kb band of the *hENH* messenger RNA was predominantly expressed in heart and skeletal muscle, while the 5.6-kb band of the *hENH* messenger RNA was ubiquitously expressed in various human tissues (Fig. 2). These observations are closely related to findings for rat *ENH* (Kuroda et al. 1996), and suggest that *hENH* is involved in a basic cellular function. We do not know at present the nature of the shorter band seen in heart and muscle. A search in database of Expressed Sequence Tags (dbEST), using the *hENH* gene as a query, identified several ESTs derived from heart and muscle cDNA libraries (accession numbers, W67269, W67963, AA086432, AA086443, AA644237, AA194513). Alignment of these ESTs and with that of the *hENH* gene suggested that these EST clones may be splicing variants lacking the coding region and the 3′ untranslated region. Thus, we speculate that the shorter band may represent the alternative spliced form of the *hENH* gene.

We determined the chromosomal location of the human *ENH* gene using a somatic cell hybrid panel. First, a specific amplified product was detected only from the hybrid con-

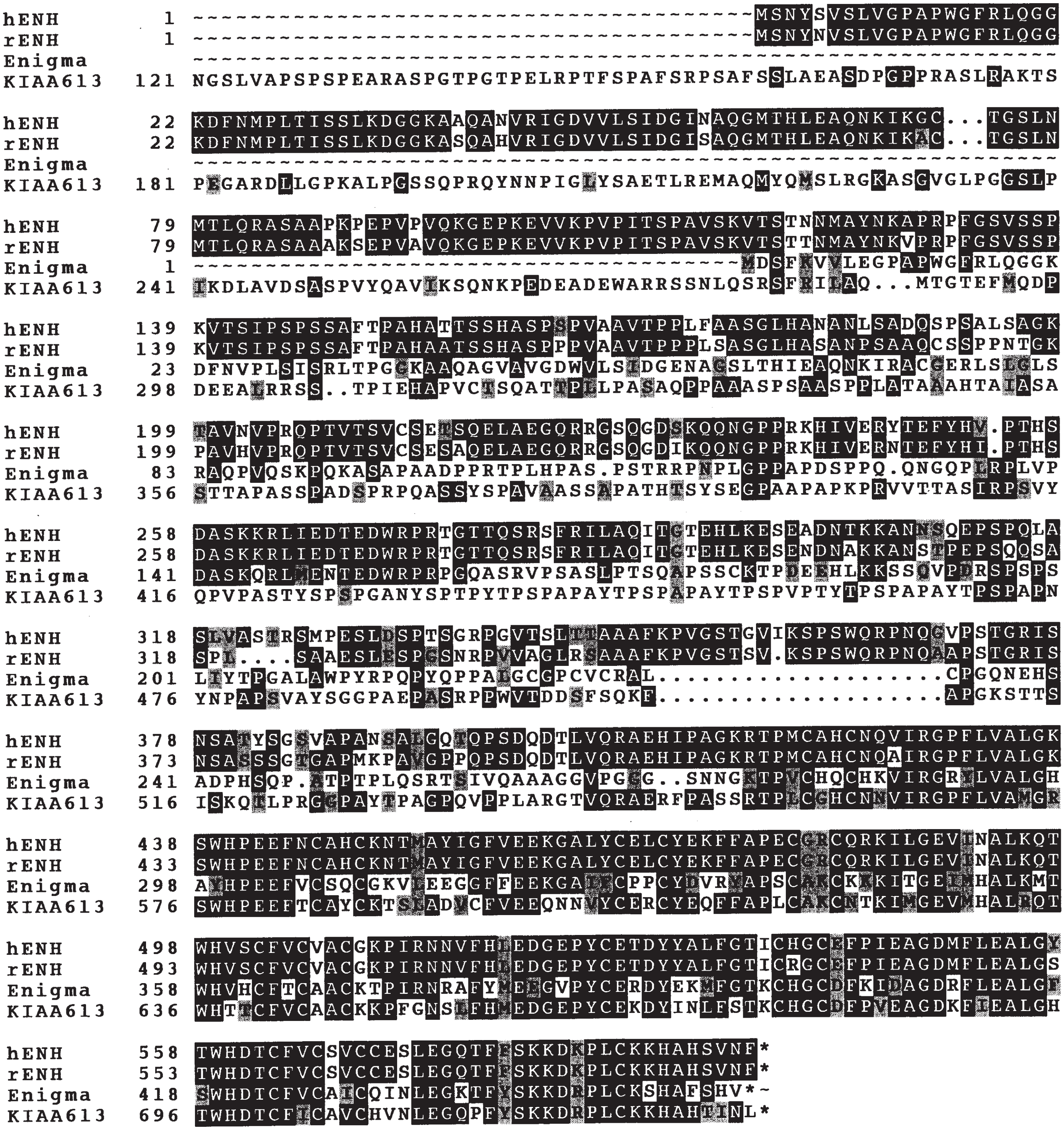

Fig. 1. Alignment of human ENH (hENH accession number AF061258), rat ENH (rENH accession number U48247), Enigma (accession number L35240), and KIAA613 (accession number AB014513). Identities are indicated by *black background*, and similar residues are *shadowed*. *Asterisks* denote termination codons

taining human chromosome 4 (data not shown). Then, we performed further mapping analysis, using a PCR-based radiation hybrid panel with the same primers as those used in the assay for the human/rodent somatic cell hybrid panel. Statistical analysis of the radiation hybrid data was performed using the RHMAPPER software package (http://carbon.wi.mit.edu:8000/cgi-bin/contig/rhmapper.pl). The data vector for the *hENH* gene was 0100000101 0000000110 0100010001 0000000100 1001000000 1010001000 0010011000 0011000001 0110001001 011 and the consequent report indicated that the gene was mapped between markers WI-2900 and WI-3273, both of which have been cytogenetically mapped to 4q22 (Fig. 3). The position of the gene was 2.84 cR proximal from WI-2900 (lod > 3.0).

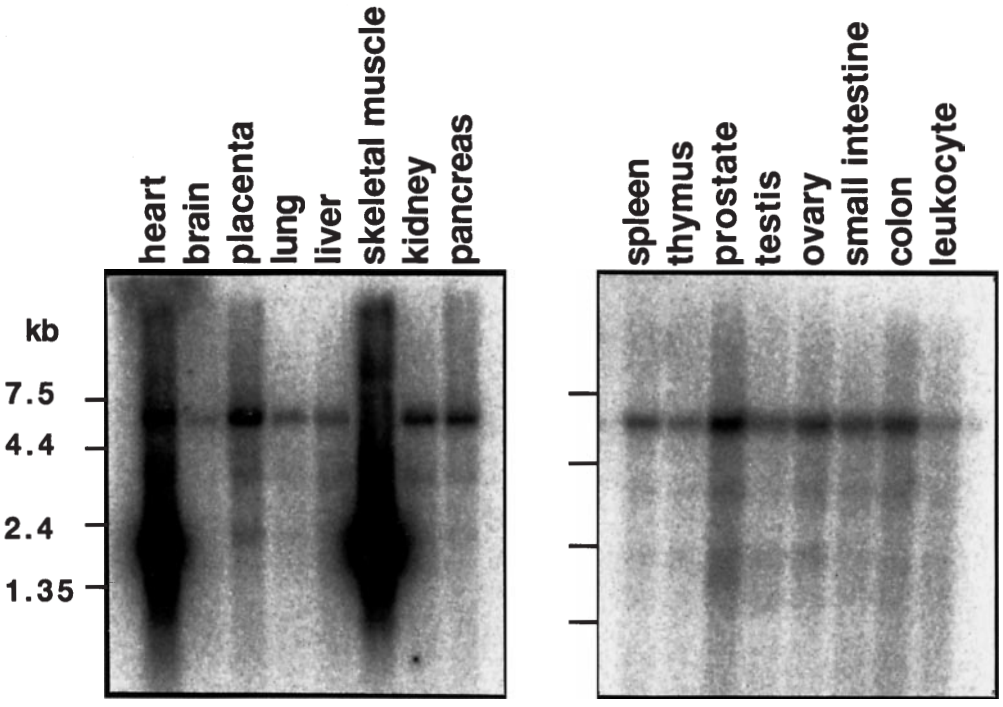

Fig. 2. Northern blot analysis of human *ENH*. Northern blot filters containing various adult human poly (A)+ RNAs (2 μg/lane) were hybridized to a *hENH* cDNA probe. Size markers (*left*) are in kilobases (*kb*)

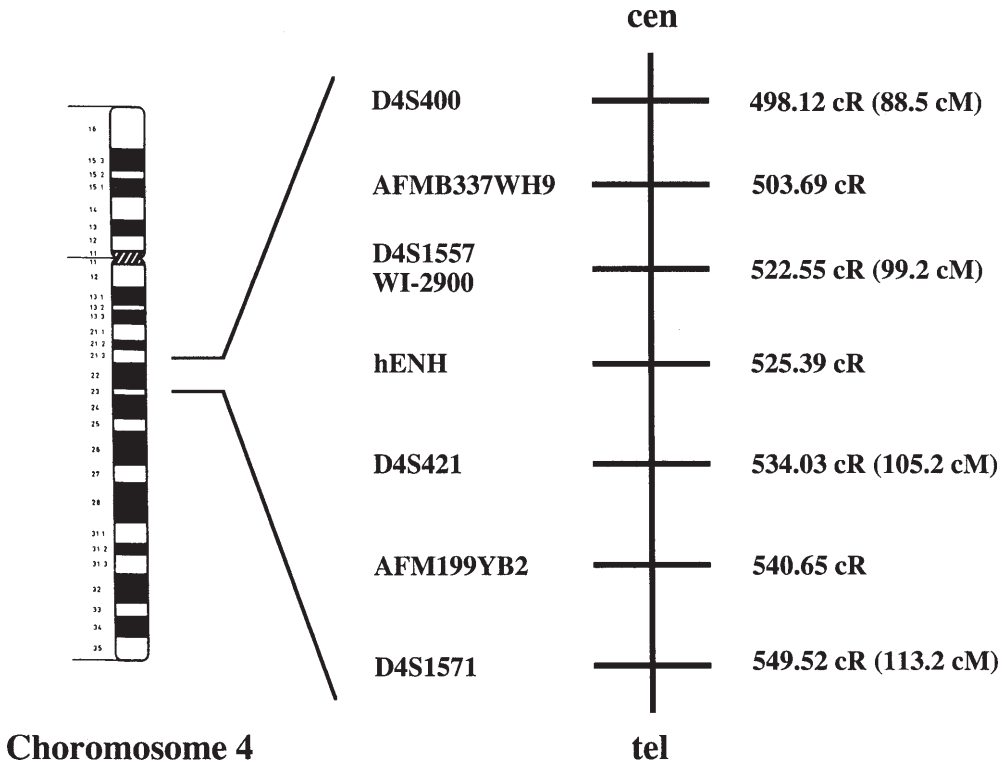

Fig. 3. Chromosomal placement of human *ENH* gene at a relative distance to framework markers on the Welcome To the Whitehead Institute for Biomedical Research/MIT Center for Genome Research (WICGR) radiation hybrid map of the human genome. The approximate corresponding cytogenetic location of the gene on the 4q22 region is indicated. Distances are in centirays (*cR*) and centimorgans (*cM*) from the top of the chromosome 4 linkage group

With regard to diseases, very few genetic disorders have been reported to be associated with 4q22, including deletion 4q21/4q22 syndrome (Nowaczyk et al. 1997). Determination of the chromosomal position and expression profile of the *hENH* gene may contribute toward ongoing positional candidate approaches for potential disease genes linked to this genomic locus.

References

Archer VE, Breton J, Sanchez-Garcia I, Osada H, Forster A, Thomson AJ, Rabbitts TH (1994) Cysteine-rich LIM domains of LIM-homeodomain and LIM-only proteins contain zinc but not iron. Proc Natl Acad Sci USA 91:316–320

Bell RM, Burns DJ (1991) Lipid activation of protein kinase C. J Biol Chem 266:4661–4664

Crawford AW, Pino JD, Beckerle MC (1994) Biochemical and molecular characterization of the chicken cysteine-rich protein, a developmentally regulated LIM-domain protein that is associated with the actin cytoskeleton. J Cell Biol 124:117–112

Dawid IB, Toyama R, Taira M (1995) LIM domain proteins. C R Acad Sci III 318:295–306

Dawid IB, Breen JJ, Toyama R (1998) LIM domains: multiple roles as adapters and functional modifiers in protein interactions. Trends Genet 14:156–162

Dekker LV, Parker PJ (1994) Protein kinase C — a question of specificity. Trends Biochem Sci 19:73–77

Freyd G, Kim SK, Horvitz HR (1990) Novel cysteine-rich motif and homeodomain in the product of the *Caenorhabditis elegans* cell lineage gene lin-11. Nature 344:876–879

Hug H, Sarre TF (1993) Protein kinase C isoenzymes: divergence in signal transduction? Biochem J 291:329–343

Ishikawa K, Nagase T, Suyama M, Miyajima N, Tanaka A, Kotani H, Nomura N, Ohara O (1998) Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro. DNA Res 5:169–176

Karlsson O, Thor S, Norberg T, Ohlsson H, Edlund T (1990) Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. Nature 344:879–882

Kiess M, Scharm B, Aguzzi A, Hajnal A, Klemenz R, Schwarte-Waldhoff I, Schafer R (1995) Expression of ril, a novel LIM domain gene, is down-regulated in Hras-transformed cells and restored in phenotypic revertants. Oncogene 10:61–68

Klug A, Schwabe JW (1995) Protein motifs 5. Zinc fingers. FASEB J 9:597–604

Kosa JL, Michelsen JW, Louis HA, Olsen JI, Davis DR, Beckerle MC, Winge DR (1994) Common metal ion coordination in LIM domain proteins. Biochemistry 33:468–477

Kuroda S, Tokunaga C, Kiyohara Y, Higuchi O, Konishi H, Mizuno K, Gill GN, Kikkawa U (1996) Protein protein interaction of zinc finger LIM domains with protein kinase C. J Biol Chem 271:31029–31032

Liebhaber SA, Emery JG, Urbanek M, Wang XK, Cooke NE (1990) characterization of a human cDNA encoding a widely expressed and highly conserved cysteine-rich protein with an unusual zinc-finger motif. Nucleic Acids Res 18:3871–3879

Macalma T, Otte J, Hensler ME, Bockholt SM, Louis HA, Kalff-Suske M, Grzeschik KH, von der Ahe D, Beckerle MC (1996) Molecular characterization of human zyxin. J Biol Chem 271:31470–31478

Michelsen JW, Schmeichel KL, Beckerle MC, Winge DR (1993) The LIM motif defines a specific zinc-binding protein domain. Proc Natl Acad Sci USA 90:4404–4408

Michelsen JW, Sewell AK, Louis HA, Olsen JI, Davis DR, Winge DR, Beckerle MC (1994) Mutational analysis of the metal sites in an LIM domain. J Biol Chem 269:11108–11113

Nishizuka Y (1992) Intracellular signaling by hydrolysis of phospholipids and activation of protein kinase C. Science 258:607–614

Nowaczyk MJ, Teshima IE, Siegel-Bartelt J, Clarke JT (1997) Deletion 4q21/4q22 syndrome: two patients with de novo 4q21.3q23 and 4q13.2q23 deletions. Am J Med Genet 69:400–405

Sadler I, Crawford AW, Michelsen JW, Beckerle MC (1992) Zyxin and cCRP: two interactive LIM domain proteins associated with the cytoskeleton. J Cell Biol 119:1573–1587

Saito T, Seki N, Yamauchi M, Tsuji S, Hayashi A, Kozuma S, Hori T (1998) Structure, chromosomal location, and expression profile of EXTR1 and EXTR2, new members of the multiple exostoses gene family. Biochem Biophys Res Commun 243:61–66

Salgia R, Li JL, Lo SH, Brunkhorst B, Kansas GS, Sobhany ES, Sun Y, Pisick E, Hallek M, Ernst T (1995) Molecular cloning of human paxillin, a focal adhesion protein phosphorylated by P210BCR/ABL. J Biol Chem 270:5039–5047

Sanchez-Garcia I, Rabbitts TH (1994) The LIM domain: a new structural motif found in zinc-finger-like proteins. Trends Genet 10:315–32

Seki N, Nimura Y, Ohira M, Saito T, Ichimiya S, Nomura N, Nakagawara A (1997) Identification and chromosome assignment of a human gene encoding a novel phosphatidylinositol-3 kinase. DNA Res 4:355–358

Ueki N, Oda T, Kondo M, Yano K, Noguchi T, Muramatsu M (1998) Selection system for genes encoding nuclear-targeted proteins. Nat Biotechnol 16:1338–1342

Wang X, Lee G, Liebhaber SA, Cooke NE (1992) Human cysteine-rich protein. A member of the LIM/double-finger family displaying coordinate serum induction with c-myc. J Biol Chem 267:9176–9184

Way JC, Chalfie M (1988) mec-3, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans*. Cell 54:5–16

Wu RY, Gill GN (1994) LIM domain recognition of a tyrosine-containing tight turn. J Biol Chem 269:25085–25090