Molecular cloning and characterization of a novel human kinase gene, *PDIK1L*

LINGCHEN GUO, CHAONENG JI, SHAOHUA GU, KANG YING, HAIPENG CHENG, XIAOGHUA NI, JIANPING LIU, YI XIE and YUMIN MAO*

State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Sciences, Fudan University, Shanghai 200433, People's Republic of China

Abstract

We isolated a 4301-bp cDNA from a human foetal brain cDNA library by high-throughput cDNA sequencing. It encodes a protein of 341 amino acids, which shows 69% identity with the human kinase CLIK1 (AAL99353), which was suggested to be the CLP-36 interacting kinase. Bioinformatics analysis suggests that the putative kinase may interact with PDZ and LIM domain proteins. Therefore the protein and its cDNA were named 'PDLIM1 interacting kinase 1 like' (PDIK1L; nomenclature approved by the HUGO Gene Nomenclature Committee). Ensembl Genome Browser located *PDIK1L* to human chromosome 1p35.3. It spans about 13.7 kb and consists of four exons and three introns. Multiple-tissue cDNA panel PCR revealed that the gene is expressed widely in human tissues: liver, kidney, pancreas, spleen, thymus and prostate. The protein appears to be localized to the nucleus.

[Guo L., Ji C., Gu S., Ying K., Cheng H., Ni X., Liu J., Xie Y. and Mao Y. 2003 Molecular cloning and characterization of a novel human kinase gene, *PDIK1L*. *J. Genet*. **82**, 27–32]

Introduction

The PDZ domain is an 80–120-amino-acid domain that was first identified in the postsynaptic protein post-synaptic density 95 (PSD95) (Cho *et al*. 1992). Sequence analysis has shown that PDZ domains are common protein motifs that occur in a variety of proteins that interact with the cytoskeleton (Ponting and Phillips 1995). It was named for the three proteins in which it was first recognized, PSD95, discs large and zol tight junction protein (Fanning and Anderson 1996).

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-$X_{16–20}$-Cys-X_2-His/Asp/Cys) (where X represents any amino acid) and is found in various proteins (Sanchez-Garcia and Rabbitts 1994; Dawid *et al*. 1995). Carboxyl-terminal LIM domain proteins that cantain a PDZ domain at the amino terminal were recently identified (Xia *et al*. 1997). The members of this subclass of carboxyl-terminal LIM domain proteins include actin-associated LIM protein (ALP) (Xia *et al*. 1997), reversion-induced LIM protein (RIL) (Bashirova *et al*. 1998) and a carboxyl-terminal LIM domain protein of 36 kDa (CLP-36) (Kotaka *et al*. 1999). Studies have indicated that LIM domains are capable of protein–protein interactions and are found in a variety of proteins with different cellular functions (Sanchez-Garcia and Rabbitts 1994; Taira *et al*. 1995; Kuroda *et al*. 1996; Dawid *et al*. 1998; Guy *et al*. 1999).

The *CLP-36* gene was cloned from rat hepatocytes, and its expression was shown to be downregulated when chemical hypoxia was induced (Wang *et al*. 1995). hCLIM1 was found to be the human homologue of CLP-36, and studies led to the speculation that hCLIM1 may be involved in a signal transduction pathway or the cytoskeletal network through interaction with other proteins, forming a multiple-protein complex (Kotaka *et al*. 1999). An

L. Guo and C. Ji contributed equally to this work.
*For correspondence. E-mail: ymmao@fudan.edu.cn.

Keywords. *PDIK1L*; protein kinase; human; expression.

immunofluorescence study demonstrated that hCLIM1 colocalizes with alpha-actinin at the Z-discs in human myocardium. It suggests that hCLIM1 is a novel cytoskeletal protein and may act as an adapter that brings other proteins to the cytoskeleton (Kotaka *et al*. 2000). CLIK1, the CLP-36 interacting kinase, was identified to be a novel kinase that associates with CLP-36 PDZ–LIM. The association with CLP-36 led to relocalization of the otherwise nuclear CLIK1 kinase to actin stress fibres, where it disrupted the periodic staining pattern of CLP-36 (Vallenius and Makela 2002). The PDZ domain of CLP-36 mediates its stable association with nonmuscle forms of alpha-actinin (Vallenius *et al*. 2000). Here we report the cloning of the gene for a novel human kinase that shows 69% homology with the human CLIK1 protein, as well as tissue distribution

Figure 1. cDNA nucleotide sequence and deduced amino-acid sequence of *PDIK1L* (GenBank accession number AF411102). Its predicted amino acid sequence is shown below in single-letter code; The polyA tailing signal (aataaa) is underlined.

of its expression and subcellular localization of the protein.

Materials and methods

***Molecular cloning and sequencing of PDIK1L cDNA*:** Human foetal brain polyA+ RNA was purchased from Clontech (Cat. No. 6525-1). Double-stranded cDNA was prepared with the SMART PCR cDNA Synthesis kit (Clontech, Cat. No. K1052-1) according to the manufacturer's recommendation. The cDNAs were digested with *Sfi*I, followed by size fractionation through a Sepharose CL-2B column. Then the cDNAs were directionally cloned into pBluescript II SK(+) vector. Both 5′ and 3′ ESTs were generated by either dye primer or dye terminator chemistries on ABI 377 sequencer using M13 consensus primers. Primer walking was performed as necessary.

***Bioinformatics analysis*:** DNA and deduced amino acid sequence comparisons were performed using BLAST-N and BLAST-X (http://www.ncbi.nlm.nih.gov/blast). Mapview analysis was performed using Ensembl Genome Browser (http://www.ensembl.org/Homo_sapiens/blastview). Comparison of amino acid sequences between different proteins was performed with GeneDoc. The predicted amino acid sequence was scanned against the profile entries in PROSITE to find the occurrence of known profiles (http://www.expasy.ch/pfscan). A protein structure analysis was performed using the PredictProtein server (http://cubic.bioc.columbia.edu/predictprotein). Some other sequence analysis was performed using GeneRunner software.

***Multiple-tissue cDNA panel PCR*:** Adult multiple-tissue cDNA (MTC) panel I, II were purchased from Clontech (Cat. No. K1420-1, K1421-1). The MTC-based RT-PCR was operated according to the manufacturer's recommendation. The sequences for human *PDIK1L*-specific primers were 5′-GTGCTTGGAGCCACAGTTTCCTAAGCG-3′ and 5′-CCTATTATGGGCAAGGCACTGTGTACG-AAG-3′. Thirtynine cycles of amplification (30 s at 94°C, 45 s at 60°C, and 45 s at 72°C) were performed using Advantage 2 PCR Enzyme Systems (Clontech, Cat. No. PT3281-1). *G3PDH* cDNA was used as the positive control.

***Subcellular localization of PDIK1L*:** The putative ORF of *PDIK1L* was cloned into pEGFP-C1 expression vector (Clontech, Cat. No. 6084-1) to allow expression of human *PDIK1L* as a fusion protein with green fluorescent protein (GFP). The fusion plasmid generated was used to transfect COS 7 cells as described previously (Chen and Okayama 1997). GFP served as control.

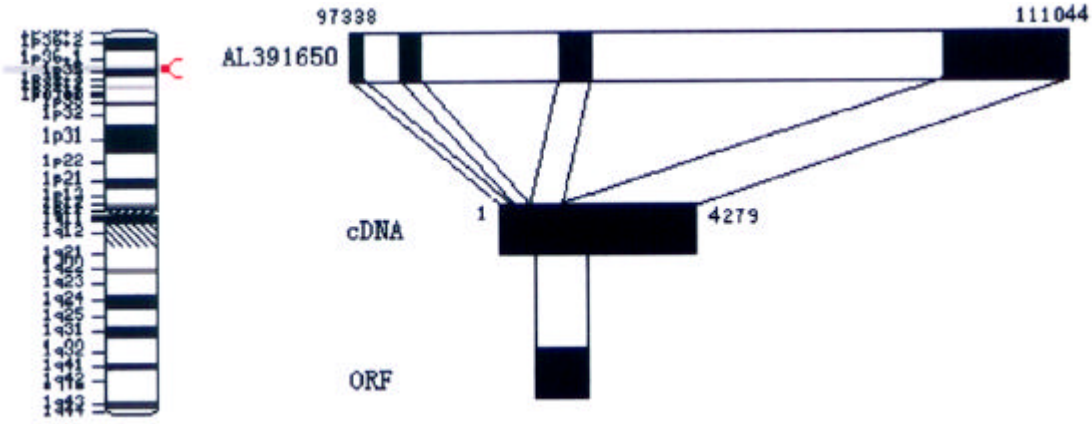

Figure 2. Chromosome alignment and gene organization of *PDIK1L*. *PDIK1L* corresponds to a finished genome sequence (AL391650), consists of four exons and three introns, and is located at human chromosome 1p35.3.

Table 1. Nucleotide sequence at exon–intron junctions of human PDIKIL. Intron and exon nucleotide sequences are shown in lower-case and upper-case letters respectively.

3′ Splice acceptor	Exon	Size (bp)	5′ Splice donor	Intron	Size (bp)
(5′-end) GGCAGCGGAG	1	79	CCCTGAGgtaaatc	1	544
CctattcagGAAGGTGCTC	2	191	GATTTGAAGGgggtgggtg	2	1641
atttttccaAAACCTCGAC	3	302	TATTTACAGgtatgtgttg	3	7243
cttcttgtagCTTGTAGAAA	4	3707			

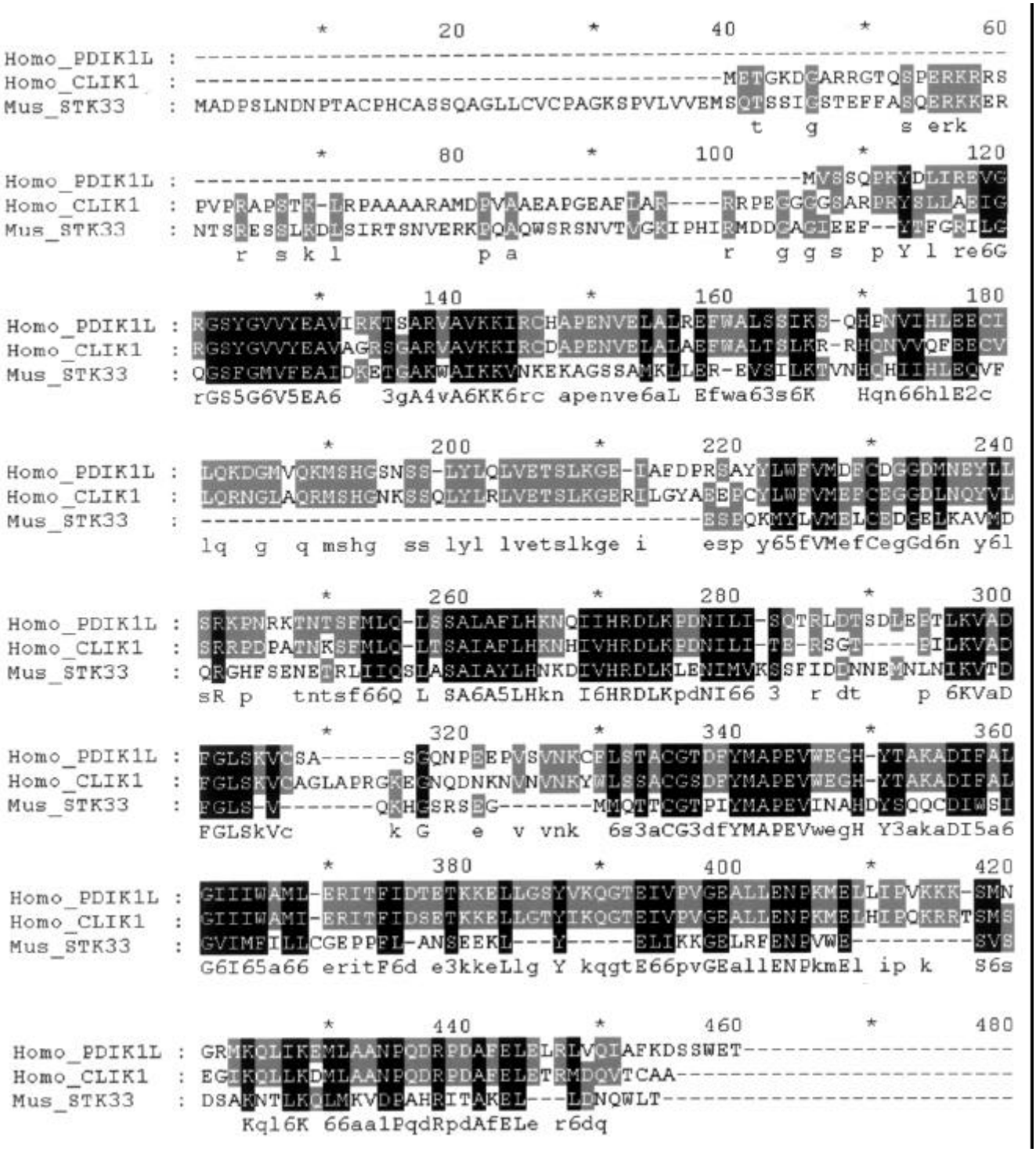

Figure 3. Protein sequence alignment of human PDIK1L (GenBank AAN03661), human CLIK1 (GenBank AAL99353) and mouse STK33 (GenBank CAC39171). The numbers indicate positions of amino acids. Identical and similar peptide sequences are shaded in dark and gray. Two motifs are discernible: protein kinase ATP-binding region (residues 120–144) and serine/threonine protein kinase active site (residues 270–282). Most of the whole sequence (residues 114–454) shows conservation of sequence.

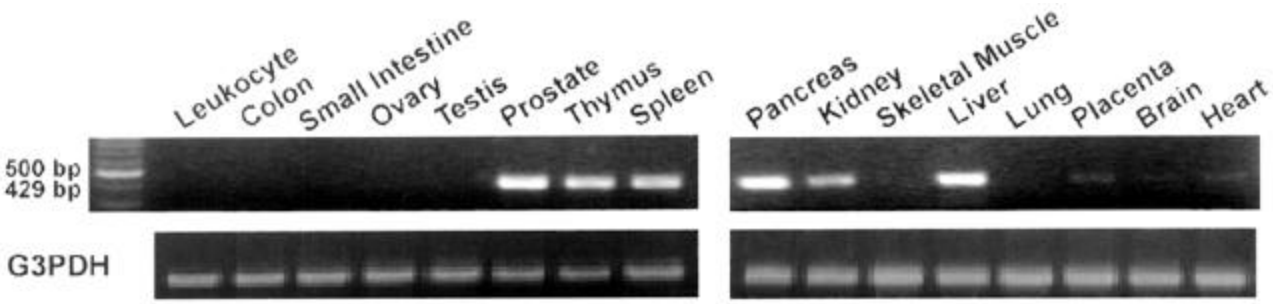

Figure 4. Expression pattern of *PDIK1L* examined by adult multiple-tissue cDNA (MTC) panel PCR. The numbers at left are lengths of DNA marker (above) and the PCR product of *PDIK1L* cDNA (below). Expression was detected in liver, kidney, pancreas, spleen, thymus, prostate, placenta, heart and brain.

Results and discussion

Molecular cloning and identification of human PDIK1L cDNA

A 4301-bp cDNA which contains a 1023-bp ORF (nucleotides 288–1313) that encodes a putative protein with 341 amino acids was obtained by high-throughput cDNA sequencing (figure 1). The predicted protein has molecular mass of 38.5 kDa and isoelectronic point of 6.7.

To find out the potential function of the gene, a BLAST analysis of the peptide sequence was performed in the GenBank CDS translation database; the human serine/threonine kinase CLIK1 (AAL99353), which showed 69% homology with the putative protein, was found. The result also showed that the putative protein contains the catalytic domain of serine/threonine protein kinases and is homologous to the CLIK1 kinase which interacts with PDZ–LIM protein. Therefore we named the protein, and its cDNA, 'human PDLIM1 interacting kinase 1 like' (PDIK1L; nomenclature approved by the HUGO Gene Nomenclature Committee). The cDNA sequence has been deposited in GenBank with accession number AF411102.

Bioinformatics analysis of human PDIK1L

Mapviewer analysis of *PDIK1L* cDNA aligned the gene on human chromosome 1p35.3 (figure 2). *PDIK1L* corresponds to genome sequence AL391650, which suggests that *PDIK1L* consists of four exons and three introns and spans 13.7 kb (table 1).

Human PDIK1L protein shows 69% and 27% identity with human CLIK1 (GenBank accession number AAL99353) and mouse STK33 (GenBank accession number CAC39171), respectively. Protein motif analysis revealed that human PDIK1L protein of 341 amino acids contains a motif of protein kinase domain throughout the whole peptide, located from residue 8 to 334, a serine/threonine protein kinase active site signature located at residues 160–172, and a protein kinase ATP-binding region signature located at residues 14–38. The serine/threonine protein kinase consensus motif of PDIK1L and its flanking residues show 100% and 84% similarity to the motif of human CLIK1 and mouse STK33, respectively (figure 3).

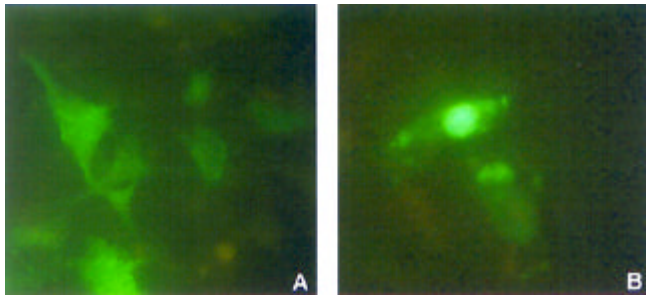

Figure 5. Subcellular distribution of PDIK1L. COS 7 cells were transfected with pEGFP (A) and pEGFP-C1-PDIK1L (B). Product from the positive control pEGFP is distributed throughout the cell as expected. GFP–PDIK1L fusion product is localized to the nucleus.

The high similarity between serine/threonine protein kinase active domains of human PDIK1L and CLIK1 suggests that they may have similar functions.

Expression analysis

MTC panel PCR revealed that *PDIK1L* is expressed in several human tissues: liver, kidney, pancreas, spleen, thymus and prostate; very weak signals were detected in placenta, heart and brain (figure 4). The size of PCR products was 429 bp, which is the expected length from the specific primer pair used.

Subcellular localization of PDIK1L

In transfected COS 7 cells, GFP protein was distributed in both nucleus and cytoplasm as expected. The GFP–PDIK1L fusion product was detected only in the nucleus of COS 7 cells. (figure 5). This suggests that, like CLIK1 (Vallenius *et al*. 2002), PDIK1L may also have a nuclear

localization and function. Further studies can throw light on the function of this protein.

Acknowledgements

This research is a part of Project 30070383 supported by the National Natural Science Foundation of China.

References

Bashirova A. A., Markelov M. L., Shlykova T. V., Levshenkova E. V., Alibaeva R. A. and Frolova E. I. 1998 The human *RIL* gene: mapping to human chromosome 5q31.1, genomic organization and alternative transcripts. *Gene* **210**, 239–245.

Chen C. and Okayama H. 1987 High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**, 2745–2752.

Cho K. O., Hunt C. A. and Kennedy M. B. 1992 The rat brain postsynaptic density fraction contains a homolog of the Drosophila discs-large tumour suppressor protein. *Neuron* **9**, 929–942.

Dawid I. B., Toyama R. and Taira M. 1995 LIM domain proteins. *C. R. Acad. Sci. III* **318**, 295–306.

Dawid I. B., Breen J. J. and Toyama R. 1998 LIM domains: multiple roles as adapters and functional modifiers in protein interactions. *Trends Genet.* **14**, 156–162.

Fanning A. S. and Anderson J. M. 1996 Protein-protein interactions: PDZ domain networks. *Curr. Biol.* **6**, 1385–1388.

Guy P. M., Kenny D. A. and Gill G. N. 1999 The PDZ domain of the LIM protein enigma binds to beta-tropomyosin. *Mol. Biol. Cell* **10**, 1973–1984.

Kotaka M., Ngai S. M., Garcia-Barcelo M., Tsui S. K., Fung K. P., Lee C. Y. and Waye M. M. 1999 Characterization of the human 36-kDa carboxyl terminal LIM domain protein (hCLIM1). *J. Cell. Biochem.* **72**, 279–285.

Kotaka M., Kostin S., Ngai S., Chan K., Lau Y., Lee S. M., Li H., Ng E. K., Schaper J., Tsui S. K., Fung K., Lee C. and Waye M. M. 2000 Interaction of hCLIM1, an enigma family protein, with alpha-actinin 2. *J. Cell. Biochem.* **78**, 558–565.

Kuroda S., Tokunaga C., Kiyohara Y., Higuchi O., Konishi H., Mizuno K., Gill G. N. and Kikkawa U. 1996 Protein-protein interaction of zinc finger LIM domains with protein kinase C. *J. Biol. Chem.* **271**, 31029–31032.

Ponting C. P. and Phillips C. 1995 DHR domains in syntrophins, neuronal NO synthases and other intracellular proteins. *Trends Biochem. Sci.* **20**, 102–103.

Sanchez-Garcia I. and Rabbitts T. H. 1994 The LIM domain: a new structural motif found in zinc-finger-like proteins. *Trends Genet.* **10**, 315–320.

Taira M., Evrard J. L., Steinmetz A. and Dawid I. B. 1995 Classification of LIM proteins. *Trends Genet.* **11**, 431–432.

Vallenius T. and Makela T. P. 2002 Clik1: a novel kinase targeted to actin stress fibers by the CLP-36 PDZ-LIM protein. *J. Cell Sci.* **115**, 2067–2073.

Vallenius T., Luukko K. and Makela T. P. 2000 CLP-36 PDZ-LIM protein associates with nonmuscle alpha-actinin-1 and alpha-actinin-4. *J. Biol. Chem.* **275**, 11100–11105.

Wang H., Harrison-Shostak D. C., Lemasters J. J. and Herman B. 1995 Cloning of a rat cDNA encoding a novel LIM domain protein with high homology to rat RIL. *Gene* **165**, 267–271.

Xia H., Winokur S. T., Kuo W. L., Altherr M. R. and Bredt D. S. 1997 Actinin-associated LIM protein: identification of a domain interaction between PDZ and spectrin-like repeat motifs. *J. Cell Biol.* **139**, 507–515.

Received 13 January 2003; in revised form 19 May 2003