L27, a novel heterodimerization domain in receptor targeting proteins Lin-2 and Lin-7

Membrane-associated guanylate kinases (MAGUKs) are emerging as pivotal for the organization of cell-surface proteins and their interaction with the cytoskeleton[1]. They are involved in cell junction organization and tumour suppression. Recent work has indicated that the *Caenorhabditis elegans* MAGUK protein Lin-2 is crucial for the proper targeting of the worm growth factor receptor Let-23 to the basolateral surface of epithelial cells[2]. Lin-7 and Lin-10 are also required for the basolateral targeting of this receptor in worm and form a complex with Lin-2 (Ref. 3). Mammalian Lin-7 (or Veli)[4] has also been found to associate with several other smaller Lin-2-related MAGUK proteins including Dlg2, Dlg3, Pals1 and Pals2 (Ref. 5).

Lin-2- and Lin-7-related proteins contain a PDZ domain (i.e. domain present in PSD-95, dlg and ZO-1/2), which is not involved in complex formation with the other Lin proteins[3,4]. PSI-BLAST searches[6] with the N-terminal region preceding the PDZ domain[7] of Dlg2 retrieved all members of the Lin-2 family of MAGUKS. DOTPLOT[8] two-dimensional visual comparison analysis of these proteins indicates an internal duplication, which is confirmed by MACAW alignment analysis[9] (*P* value 10^{-50}). Using the second duplicate alone in Dlg2, additional BLASTP searches against the wormpep18 database show weak similarity (*E* value 0.1) to the Lin-7 protein. In Lin-7, the matching region is also located at the N-terminus and is followed by a PDZ domain. The

```
MAGUK
Cask_b          mm   405    AVQRAKEVLEEISCYPE[1]NDAKELKRILTQ    PHFMALLQTHDVVAHEVYSDEALR    O70589
P55T/PALS2_b    mm    56    NLELVNEILEDITPLIS[2]ENVAELVGILKE    PHFQSLLEAHDIVASKCYDSPPSS    AAD45009
Camguk_b        dm   405    AVGRCRDVLEQLSSTSG[7]YAKEELMRLLAA    PHMQALLHSHDVVARDVYGEEALR    Q24210
Dlg3_b          hs    68    AVALAEDVMEELQAASV[1]SDERELLQLLST    PHLRAVLMVHDTVAQKNFDPVLPP    Q13368
Dlg2_b          hs    91    NLELVQEILRDLAQLAE[2]STAAELAHILQE    PHFQSLLETHDSVASKTYETPPPS    Q14168
LIN2_b          ce   421    TSTLRKETLNQIDGLLG[2]PEALELRQLLNS    PHLASCVQALDVVVCEIRDPKNEA    P54936
PALS1_b hs           186    VQDLVQEVQTVLKPVHQ   KEGQELTALLNA    PHIQALLLAHDKVAEQEMQLEPIT    AF199008
HSZZ27178       hs    15    AAALADDLAEELQNKPL[1]SEIRELLKLLSK    PNVKALLSVHDTXAQKNYDPVLPP    AA322046
Cask_a          mm   346    AVSQVLDSLEEIHALTD[3]KDLDFLHSVFQD    QHLHTLLDLYDKINTKSSPQIRNP    O70589
P55T/PALS2_a    mm     1    .MQQVLENLTELPSSTG[3]IDLIFLKGIMEN    PIVKSLAKAHERLEDSKLEAVSDN    AAD45009
Camguk_a        dm   346    AVQRILDCLDDIYSLQD[2]VDADVLRDMLRD    NRLHQFLQLFDRIAATVVTSNGRA    Q24210
Dlg3_a          hs    10    LHETLALLTSQLRPDSN[2]EEMGFLRDDFSE    KSLSYLMKIHEKLRYYERQSPTPV    Q13368
Dlg2_a          hs    11    AMQQVLDNLGSLPSATG[3]LDLIFLRGIMES[24]KYMLKYFGAHERLEETKLEAVRDN    Q14168
LIN2_a          ce   371    KVLGSLDAINSLLDPNS[2]PGSTTFQKIHDD    GSVRNLLRLYDKIKALPCEPVVTE    P54936
PALS1a          hs   123    DVEDLFSSLKHIQHTLV[5]EDISLLLQLVQN    RDFQNAFKIHNAVTVHMSKASPPF    AF199008

LIN-7
LIN-7           ce   120    DVQRILELMEHVQKTGE[3]AKLASLQQVLQS    EFFGAVREVYETVYESIDADTTPE    CAA22459
LIN-7-BA        rn    15    DVARAIELLEKLQESGE[3]HKLQSLKKVLQS    EFCTAIREVYQYMHETITVNGCPE    Q9Z251
hypLIN7         sm     1    ............RCPE[3]SKLAALQRILQS    DFCDMIREVYEHIYTTVDINGSEE    O17458

HPT
rdea            dd    26    EKEFTFELLDSYISSVE   EHLPELLNSFEA [1]DLKGAVLHSHDIKGSSSYIGCEAV    O77083
EVGS_ECOLI      ec  1098    DLQLMQEILMTFQHETH   KDLPAAFQALEA [1]DNRTFHQCIHRIHGAANILNLQKL    P30855
BARA_ECOLI      ec   822    KTDLARDMLQMLLDFLP   EVRNKVEEQLVG [1]NPEGLVDLIHKLHGSCGYSGVPRM    P26607
TORS_ECOLI      ec   811    GTEKIHEWLVLFTQHAL   PLLDEIDIARAS [1]DSEKIKRAAHQLKSSCSSLGMHIA    P39453
CHEA_ECOLI      ec     8    FYQTFFDEADELLADME   QHLLVLQPEAPD [2]QLNAIFRAAHSIKGGAGTFGFSVL    P07363
YPD1            sc    24    DSDFSKGLIIQFIDQAQ   TTFAQMQRQLDG [2]NLTELDNLGHFLKGSSAALGLQRI    Q07688
ATHP3           at    38    NPDFVSQVVTLFFQDSD   RILNDLSLSLDQ [3]DFKKVDPHVHQLKGSSSSIGAQRV    Q9ZNV9
Consensus (80%)               .hthh.-.hpph.t..t....ph..L.t.hpp......ht.hhthaphhttt...s....
Sec.struc.pred.(2lin )        ..hHHHHHHHHHH.......HHHHHHHHHHH.....HHHHHHHHHHHHHHhhhh.......
Sec.struc.pred.(7lin )        .HHHHHHHHHHHHH......HHHHHHHHHHHH....HHHHHHHHHHHh............
Sec.struc.(1a0b)              ..HHHHHHHHHHHHHH...HHHHHHHHHHHH.....HHHHHHHHHHHHHHHHHH.....
```

TiBS

Figure 1

Multiple alignment of L27 domains of the MAGUK Lin-2-related proteins, Lin-7-like proteins and selected members of histidine phosphotransfer (HPt)-domain-containing proteins. First column: protein names (repeated domains in the same protein are labeled a or b); second column: species names (at: *Arabidopsis thaliana*; ce: *Caenorhabditis elegans*; dd: *Dictyostelium discoideum*; dm: *Drosophila melanogaster*; ec: *Escherichia coli*; hs: *Homo sapiens*; mm: *Mus musculus*; rn: *Rattus norvegicus*; sc: *Saccharomyces cerevisiae*; sm: *Schistosoma mansonii*); third column: first amino acid of the domain; far-right column: database accession numbers. Conserved negatively charged residues are shown in red; conserved hydrophobic residues are shown in blue; other conserved residues are shown in bold. The consensus sequence (conserved in 80% of the sequences) shown below the alignment; h, p, a, t, s and – indicate hydrophobic, polar, aromatic, turn-like, small and negatively charged residues, respectively. Conserved residues, hydrophobicity patterns and secondary structure elements resemble the core of HPt-signalling domains[10] and are shown below the MAGUK/LIN alignment. The predicted secondary structure taken from the consensus of the Lin-2-related proteins (2lin), the consensus of Lin-7-like proteins (7lin) and the known secondary structure of the HPt-domain-containing proteins (1a0b) are shown (H, helix known or predicted with expected average accuracy >82%; h, helix predicted with expected average accuracy <82%)[11]. Three blocks of the multiple alignment are seperated by variable regions (represented by numbers in brackets). MACAW alignment analysis[9] shows significant similarity between both copies of lin-2 and lin-7 for three distinct blocks marked in red. The first block has a *P* value of 7.7×10^{-13} over 12 amino acids, the second block has a *P* value of 4.9×10^{-8} over six amino acids and the third block has a *P* value of 10^{-50} over 13 amino acids.

 PII: S0968-0004(00)01599-1

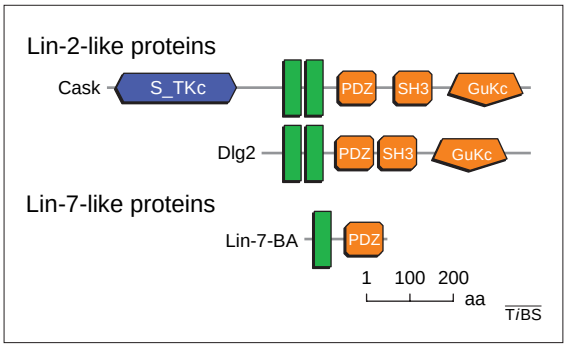

Figure 2

Domain architecture of proteins containing the L27 domain (green). Examples of proteins with distinct modular organizations are shown. The domain names are according to the Simple Modular Architecture Research Tool[12] (http://smart.embl-heidelberg.de). GuKc, guanylate kinase domain; S_TKc, serine/threonine protein kinase domain; SH3, src homology 3 domain; PDZ, domain present in PSD-95, dlg and ZO-1/2.

significance of the similarity is not only supported by the biological context, but also by analysis of the multiple alignment (for details, see legend of Fig. 1).

The single occurrence in Lin-7 and the duplication in Lin-2 (see Fig. 2) indicate the presence of an independent domain that we named L27, after the better characterized MAGUK subfamily (Lin-2) and its binding partners (Lin-7).The L27-domain is generally ~50 residues in length and its main features are conserved negatively charged residues and a conserved aromatic amino acid (Fig. 1). Interestingly, in MAGUKS the second copy of the domain contains a conserved histidine residue (except Lin-2 in *Caenorhabditis elegans*). It has been confirmed experimentally that this region N-terminal to the PDZ domain in Lin-2 and Pals1 can mediate Lin-7 binding[3–5]. These insights lend credence to the hypothesis that the binding process of Lin-2-related MAGUKs and Lin-7 are based on a domain-specific heterodimer interaction and the novel L27 domain reported here is the basis for the Lin-2–Lin-7 heterodimerization.

References

1 Anderson, J.M. (1996) Cell signalling: MAGUK magic. *Curr. Biol.* 6, 382–384

2 Hoskins, R. *et al.* (1996) The *C. elegans* vulval induction gene lin-2 encodes a member of the MAGUK family of cell junction proteins. *Development* 122, 97–111

3 Kaech, S.M. *et al.* (1998) The LIN-2/LIN-7/LIN-10 complex mediates basolateral membrane localization of the *C. elegans* EGF receptor LET-23 in vulval epithelial cells. *Cell* 94, 761–771

4 Butz, S. *et al.* (1998) A tripartite protein complex with the potential to couple synaptic vesicle exocytosis to cell adhesion in brain. *Cell* 94, 773–782

5 Kamberov, E. *et al.* (2000) Molecular cloning and characterization of Pals proteins associated with mLin-7. *J. Biol. Chem.* 275, 11425–11431

6 Altschul, S.F. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402

7 Ponting, C.P. *et al.* (1997) PDZ domains: targeting signalling molecules to sub-membranous sites. *Bioessays* 19, 469–479

8 Sonnhammer, E.L. and Durbin, R. (1995) A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* 29, GC1–10

9 Schuler, G.D. *et al.* (1991) A workbench for multiple alignment construction and analysis. *Proteins* 9, 180–190

10 Kato, M. *et al.* (1997) Insights into multistep phosphorelay from the crystal structure of the C-terminal HPt domain of ArcB. *Cell* 88, 717–723

11 Rost, B. *et al.* (1994) PHD – an automatic mail server for protein secondary structure prediction. *CABIOS* 10, 53–60

12 Schultz, J. *et al.* (2000) SMART: A Web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* 28, 231–234

TOBIAS DOERKS AND PEER BORK

EMBL, 69012 Heidelberg, Meyerhofstr. 1 and Max-Delbrueck-Centrum, Germany.
Email: Tobias.Doerks@EMBL-Heidelberg.de

EMMANUEL KAMBEROV, OLGA MAKAROVA, SEONOK MUECKE AND BEN MARGOLIS

HHMI, Depts of Internal Medicine and Biological Chemistry, University of Michigan Medical School, 1150 West Medical Center Dr., Ann Harbor, MI 48109-0650, USA.