- 11. SMART database: http://coot.emblheidelberg.de/SMART/
- Sugita S, Ichchenko K, Khvotchev M, Südhof TC: α-Latrotoxin receptor CIRL/latrophilin 1 (CL1) defines an unusual family of ubiquitous G-proteinlinked receptors. J Biol Chem 1998, 273:32715-32724.
- Krasnoperov V, Bittner MA, Holz RW, Chepurny O, Petrenko AG: Structural requirements for α-latrotoxin binding and α-latrotoxin-stimulated secretion. J Biol Chem 1999, 274:3590-3596.
- Krasnoperov V, Bittner MA, Beavis R, Kuang Y, Salnikow KV, Chepurny OG, Little AR, Plotnikov AN, Wu D, Holz RW, *et al.*: or-Latrotoxin stimulates exocytosis by the interaction with a neuronal Gprotein-coupled receptor. *Neuron* 1997, 18:925-937.
- Ichtchenko K, Bittner MA, Krasnoperov V, Little AR, Chepurny O, Holz RW, Petrenko AG: A novel ubiquitously expressed αlatrotoxin receptor is a member of the CIRL family of G-protein-coupled receptors. J Biol Chem 1999, 274:5491-5498.
- Gray JX, Haino M, Roth MJ, Maguire JE, Jensen PN, Yarme A, Stetler-Stevenson MA, Siebenlist U, Kelly K: CD97 is a processed, seven-transmembrane, heterodimeric receptor associated with inflammation. *J Immunol* 1996, 157:5438-5447.
- Watnick TJ, Piontek KB, Cordal TM, Weber H, Gandolph MA, Qian F, Lens XM, Neumann HPH, Germino GG: An unusual pattern of mutation in the duplicated portion of *PKD1* is revealed by use of a novel strategy for mutation detection. *Hum Mol Genet* 1997, 6:1473-1481.
- Sandford R, Sgotto B, Aparicio S, Brenner S, Vaudin M, Wilson RK, Chissoe S, Pepin K, Bateman A, Chothia C, *et al.*: Comparative analysis of the polycystic kidney disease 1 (PKD1) gene reveals an integral membrane glycoprotein with multiple evolutionary conserved domains. *Hum Mol Genet* 1997, 6:1483-1489.
- Moy GW, Mendoza LM, Schultz JR, Swanson WJ, Glabe CG, Vacquier VD: The sea urchin sperm receptor for egg jelly is a modular protein with extensive homology to the human polycystic kidney disease protein, PKD1. J Cell Biol 1996, 133:809-817.
- Trimmer JS, Schackmann RW, Vacquier VD: Monoclonal antibodies increase intracellular Ca<sup>2+</sup> in sea urchin spermatozoa. Proc Natl Acad Sci USA 1986, 83:9055-9059.
- 21 Lu W, Peissel B, Babakhanlou H, Pavlova A, Geng, L, Fan X, Larson C, Brent G, Zhou J: Perinatal lethality with kidney and pancreas defects in mice with a targetted Pkd1 mutation. Nat Genet 1997, 17:179-181.
- Hughes J, Ward CJ, Aspinwall R, Butler R, Harris PC: Identification of a human homologue of the sea urchin receptor for egg jelly: a polycystic kidney disease-like protein. Hum Mol Genet 1999, 8:543-549.
- Mochizuki T, Wu G, Hayashi T, Xenophontos SL, Veldhuisen B, Saris JJ, Reynolds DM, Cai Y, Gavow PA, Pierides A, *et al.*: PKD2, a gene for polycystic kidney disease that encodes an integral membrane protein. *Science* 1996, 272:1339-1342.

- Tsiokas L, Kim E, Arnould T, Sukhatme VP, Walz G: Homo- and heterodimeric interactions between the gene products of *PKD1* and *PKD2*. *Proc Natl Acad Sci* USA 1997, 94:6965-6970.
- Brown N, Lai J: Consensus. http://www.bork.emblheidelberg.de/Alignment/consensus.html
- Rost B, Sander C: Prediction of protein structure at better than 70% accuracy. *J Mol Biol* 1993, 232:584-599.

Addresses: \*National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda Maryland 20894, USA. \*MEMOREC Stoffel GmbH, Stöckheimer Weg 1, D-50829 Köln, Germany. <sup>‡</sup>European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69012 Heidelberg, Germany. <sup>§</sup>Max-Delbrück-Center for Molecular Medicine, D-13125 Berlin-Buch, Germany. E-mail: ponting@ncbi.nlm.nih.gov

# The PLAT domain: a new piece in the PKD1 puzzle Alex Bateman\* and Richard Sandford<sup>†</sup>

Autosomal dominant polycystic kidney disease (ADPKD) has a prevalence of 1 in 800 of the world's population and accounts for 10% of individuals who require renal replacement therapy, either dialysis or transplantation. Renal cyst formation occurs as part of a 'twohit' process in which inactivation of both alleles of ADPKD genes leads to abnormalities of cell proliferation, apoptosis and differentiation [1]. Of ADPKD cases, 85% are due to mutations in the PKD1 gene, which encodes a 4,302 amino acid protein, polycystin-1 (PKD1), of unknown function. Comparison of the PKD1 sequence with homologous sequences from mouse and Fugu predicts polycystin-1 to have a large extracellular region of 3,000 amino acid residues, a region containing 11 putative transmembrane segments and a short intracellular tail [2]. A well-defined extracellular domain structure is apparent; the presence of amino-terminal leucinerich repeats, a C-type-lectin domain and multiple PKD repeats suggests a role in cell-cell or cell-matrix interactions (Figure 1) [3]. So far, no extracellular ligands of polycystin-1 have been identified. Of the intracellular regions of PKD1, functional properties have been defined only for the short 198 amino acid carboxy-terminal region, which contains a predicted coiled-coil domain. These include a direct interaction with the carboxyl terminus of the protein encoded by PKD2, polycystin-2 [4], activation of

## Figure 1



Schematic diagram of the domain organisation of PLAT domain proteins. Information and alignments for each of the domains can be found in the Pfam database [14,15] using the following identifiers or accession numbers: leucine-rich repeats, PF00560; C-type lectin, PF00059; LDL-A, PF00057; PKD domains, PF00801; PLAT domain, PF01477; lipoxygenase, PF00305; Zn<sup>2+</sup>-dependent phospholipase C, PF00882; lipase, PF00151. Abbreviation: aa, amino acids. transcription factor AP-1 [5] and activation of heterotrimeric G proteins [6]. The last of these occurs via a motif present in one of polycystin-1's most highly conserved regions. Polycystin-1 may therefore act as a cell-surface receptor or form part of a large membrane-associated complex that is capable of signaling by several different pathways to control cell proliferation and differentiation. To further aid the understanding of this enigmatic protein, we have surveyed all the intracellular regions for potential domains that may suggest novel functions and identify further avenues for experiment.

## Figure 2

The current model of polycystin-1 topology suggests that there are four intracellular regions that are large enough to contain a discrete protein domain. These regions are between transmembrane (TM) helices TM1 and TM2 (residues 3,096-3,280), TM3 and TM4 (residues 3,344-3,558), TM5 and TM6 (residues 3,603-3,668) and the carboxy-terminal region after TM11 (residues 4,105–4,302). The high sequence conservation seen in these regions between the human and Fugu polycystin-1 suggest that they are functionally important. We have used each of these regions as a query for the sequence comparison program PSI-BLAST, using an

expectation-value (E-value) threshold of 0.001 [7]. For each of the four intracellular regions, PSI-BLAST returned the known PKD1 orthologues from human, mouse and Fugu. Only one region returned significant matches with PSI-BLAST. The first intracellular region between TM1 and TM2, which represents the most strongly conserved sequence region of PKD1 between human and Fugu [2], was found to match to 67 other sequences in SWISS-PROT version 37 and TrEMBL version 9 [8]. These sequences include mammalian lipoxygenases, triacylglycerol lipase and lipoprotein lipase. The common feature found in



A multiple sequence alignment of PLAT domains. The colouring scheme is that used in ClustalX. The secondary structures for the known structures are shown below the sequences; E denotes  $\beta$ -strand.

all these alignments was topological and sequence similarity to a  $\beta$ -sandwich domain. We call this new protein domain the PLAT domain (after polycystin-1, lipoxygenase and alpha toxin). The copy of the PLAT domain found in polycystin-1 therefore identifies an important new region of the protein.

The three-dimensional structure of the PLAT domain is known for human pancreatic lipase [9], rabbit 15-lipoxygenase [10] and alpha toxin from Clostridium perfringens [11]. The domain is a  $\beta$ -sandwich composed of two sheets of four strands each. The sequence relationship of the alpha toxin to polycystin-1 can be demonstrated by using the sequence of the PLAT domain from the known structure as a query for PSI-BLAST. The search essentially converges to the same family, including the polycystin-1 PLAT domain. Soybean lipoxygenase L-1 [12] contains a domain structurally related to the PLAT domains. It is more distant in sequence to the rest of the family; PSI-BLAST is able to find relationships to the rest of the family for only a few sequences. Although structural similarities were noticed between these structures, it was not suggested that they share a common ancestor [11].

The most highly conserved regions in the alignment of known PLAT domains (Figure 2) coincide with the  $\beta$ -strands. Most of the highly conserved residues are buried residues. An exception to this is a surface lysine or arginine that occurs on the surface of the fifth  $\beta$ -strand of all the eukaryotic PLAT domains. In pancreatic lipase, the lysine in this position forms a salt bridge with the procolipase protein. The conservation of a charged surface residue may indicate the location of a conserved ligand-binding site within the PLAT domain.

The importance of PLAT domains is underlined by mutations that lead to human disease. Mutations in lipoprotein lipase lead to chylomicronaemia and mutations in triacylglycerol lipase lead to hepatic lipase deficiency [13]. In pancreatic lipase the PLAT domain is involved in binding to the procolipase protein. This interaction is required to bring the enzymatic active site of the lipase into close contact with its lipid substrate. In 15-lipoxygenase, a protein composed of an aminoterminal  $\beta$ -sandwich (PLAT) and a carboxy-terminal catalytic domain, the PLAT domain may function to localise the enzyme near its membrane or lipoprotein sequestered substrates, by analogy to the lipase-procolipase protein-protein interaction. It is also possible that the PLAT domain of 5-lipoxygenase, another member of the mammalian lipoxygenase family, mediates an interaction with the 5-lipoxygenase activating protein (FLAP), an integral membrane protein. For alpha toxin and plant lipoxygenases, it has been suggested that the PLAT domain interacts directly with the membrane in a Ca<sup>2+</sup> dependent manner. Although a Ca<sup>2+</sup>-binding region has been predicted for alpha toxin from crystallographic data and similarity to eukaryotic calcium-binding C2 domains [11], the conserved residues that form this region are not present in the PLAT domains identified in polycystin-1 or pancreatic lipase. PLAT domains may therefore be involved in protein-protein and protein-lipid interactions. The presence of the PLAT domain in the first cytoplasmic loop of polycystin-1 suggests that this region is important in mediating interactions with other membrane protein(s) involved in polycystin-1 function.

#### References

- Qian F, Watnick T, Onuchic L, Germino G: The molecular basis of focal cyst formation in human autosomal dominant polycystic kidney disease type I. *Cell* 1996, 87:979-987.
- Sandford R, Sgotto B, Aparicio S, Brenner S, Vaudin M, Wilson RK, Chissoe S, Bateman A, Chothia C, Hughes J, et al.: Comparative analysis of the polycystic kidney disease 1 (PKD1) gene reveals an integral membrane glycoprotein with multiple evolutionary conserved domains. *Proc Natl Acad Sci USA* 1997, 6:1483-1489.

- Hughes J, Ward CJ, Peral B, Aspinwall R, Clark K, San Millan JL, Gamble V, Harris PC: The polycystic kidney disease 1 (PKD1) gene encodes a novel protein with multiple cell recognition domains. Nat Genet 1995, 10:151-160.
- Qian F, Germino F, Cai Y, Zhang X, Somlo S, Germino G: PKD1 interacts with PKD2 through a probable coiled-coil domain. *Nat Genet* 1997, 16:179-183.
- Arnould T, Kim E, Tsiokas L, Jochimsen F, Gruning W, Chang J, Walz G: The polycystic kidney disease 1 gene product mediates protein kinase C alphadependent and c-Jun N-terminal kinasedependent activation of the transcription factor AP-1. *J Biol Chem* 1998, 273:6013-6018.
- Parnell S, Magenheimer B, Maser R, Rankin C, Smine A, Okamoto T, Calvet J: The polycystic kidney disease-1 protein, polycystin-1, binds and activates heterotrimeric G-proteins in vitro. *Biochem Biophys Res Commun* 1998, 251:625-631.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997, 25:3389-3402.
- Bairoch A, Apweiler R: The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1999. Nucleic Acids Res 1999, 27:49-54.
- van Tilbeurgh H, Egloff MP, Martinez C, Rugani N, Verger R, Cambillau C: Interfacial activation of the lipaseprocolipase complex by mixed micelles revealed by X-ray crystallography. *Nature* 1993, 362:814-820.
- Gillmor SA, Villasenor A, Fletterick R, Sigal E, Browner MF: The structure of mammalian 15-lipoxygenase reveals similarity to the lipases and the determinants of substrate specificity. Nat Struct Biol 1998, 4:1003-1009.
- Naylor CE, Eaton JT, Howells A, Justin N, Moss DS, Titball RW, Basak AK: Structure of the key toxin in gas gangrene. Nat Struct Biol 1998, 5:738-746.
- Minor W, Steczko J, Stec B, Otwinowski Z, Bolin JT, Walter R, Axelrod B: Crystal structure of soybean lipoxygenase L-1 at 1.4 Å resolution. *Biochem* 1996, 35:10687-10701.
- 13. Hegele R, Tu L, Connelly P: Human hepatic lipase mutations and polymorphisms. *Hum Mutat* 1992, 1:320-324.
- Bateman A, Birney E, Durbin R, Eddy SR, Finn RD, Sonnhammer ELL: Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. Nucleic Acids Res 1999, 27:260-262.
- 15. Pfam: Protein families database of alignments and HMMs. http://www.sanger.ac.uk/Software/Pfam/

Addresses: \*The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. †Cambridge Institute for Medical Research, Addenbrooke's Hospital, Cambridge CB2 2XY, UK. E-mail: agb@sanger.ac.uk

## Supplementary material

## A latrophilin/CL-1-like GPS domain in polycystin-1 C.P. Ponting, K. Hofmann and P. Bork Current Biology 26 August 1999, 9:R585–R588

## The WSC domain

A PSI-BLAST [S1,S2] search of the region of PKD1 lying between the leucine-rich repeat flanking regions and the first PKD domain of PKD1 (amino acid residues 181-272) using an expect value (E-value) inclusion threshold of E < 0.01 revealed significant similarity to two regions of Trichoderma harzianum  $\beta$ -1,3-exoglucanase [S3]  $(E = 1 \times 10^{-5} \text{ and } 6 \times 10^{-3})$  by iteration 1, and nine additional proteins (Figure S1) by convergence after five itera-Independent evidence for these tions. domain homologues and those discussed below was provided by generalised profile analysis [S4], in which the significance was better than p < 0.01 in all cases.

We term these homologues WSC domains, after *S. cere-visiae* WSC1 (cell-wall integrity and stress-response component 1; also called Slg1 and Hcs77), WSC2, WSC3 and WSC4 proteins, which each contain a single such domain [S5]. *S. cerevisiae* WSC1–3 proteins are localised to the plasma membrane and function upstream of the PKC1–MPK1 pathway [S5–S7]. The WSC domains in these proteins are predicted to be extracellular and are

amino-terminal to serine- and threonine-rich regions, a single predicted transmembrane sequence and divergent carboxy-terminal cytoplasmic sequences. The functions of these WSC domains are unknown but the WSC1–3 proteins are suggested to act as sensors of environmental stress [S5–S7], and WSC4 (also called Yhc8p) is implicated in protein translocation to the endoplasmic reticulum [S8].

Tandem WSC domains are also found in a hypothetical human protein (KIAA0523) that appears to be a member of a novel sulphotransferase subfamily, and in the *T. harzianum*  $\beta$ -1,3-exoglucanase [S3]. The latter occurrences indicate that the PKD1 WSC domain, and its homologues in general, might possess a carbohydrate-binding role. This would complement the predicted protein-, lipoprotein- and carbohydrate-binding functions conferred by the PKD, LDL $\alpha$  and C-type lectin domains of PKD1.

## The LH2 domain

The cytoplasmic region between the first and second transmembrane regions of PKD1 was found to contain a  $\beta$ -barrel domain, homologous to a non-catalytic domain of

## Figure S1

Multiple sequence alignments of WSC domains. Amino acid residues are coloured according to a 80% consensus (calculated using the program Consensus [S13]: + indicates positively charged residues (H, K and R, green); - indicates negatively charged residues (D and E, green); a indicates aromatic residues (F, H, W and Y, highlighted in yellow); b indicates big residues (E, F, I, K, L, M, Q, R, W and Y, grey or yellow); c indicates charged residues (D, E, H, K and R, green); h indicates hydrophobic residues (A, C, F, H, I, L, M, V, W and Y, highlighted in yellow); I indicates aliphatic residues (I, L and V, highlighted in yellow); o indicates alcohol residues (S and T, pink), p indicates polar residues (D, E, H, K, N, Q, R, S and T, dark blue); s indicates small residues (A, C, S, T, D, N, V, G and P, light blue); u indicates tiny residues (A, G and S, light blue). Residues that are predicted to form disulphide bridges are shown as white-onblack. Predicted [S14] secondary structures are indicated below the alignment (e/E, extended or  $\beta$ -strand structure; h/H, helix); lowercase letters represent predictions that have expected accuracies of > 72% and uppercase letters represent predictions that have expected accuracies of > 82%. Residue numbers and GenBank identifiers are shown

R AKYIGCVLDDT-QSRALRGVSFDDYKK--MTIFRG-QDNGA-ERGYLVGGLEFGAECVGGHKIQATN SAVERGCFRRP--DNLSLALPVTAAMLN--MSVDKC-VDFGT-EKEYPLAALA-GTACHCGFPTTRFP R GTYIGCFSDG-HERTLKGAVFYDLRK--MTVSHG-QDAGA-ERSYVYAGLEAGAECVGGNRLPAVS TATYRCCFRLP--ENITHAPPSSLIQAN--VTVGTG-SGFGS-QKEFPLAILR-GFECGAYPTRFN GNNFRGCYTDS-VNARALIAESVPNGPSS-MTIEAG-QSVGK-GLGYTLAGLEYADECVGGNSLANGA GNNFLGCYSDN-VNGRTLANQVQVAGGASAMSIEAG-ETASE-SAGYTLAGVEYAGEGVGDTKFONGG AYEYVNCFSS--LFSDFSKADSYNW----QSSSHG-NSEGS-AKGASYFALVNHSEGVGGDTNPSGS QFTYKACYSADJRKLGUTYKGVFY----QSVSYG-QNECP---GQAVVALFNGTCGVGGSVAQLQ DFNYEGCYSADJICKLGUTYKGVFY----QSVSYG-QNCP--ESAVVALFNGTCGVGGSVAQLQ DFNYEGCYSSO--ITATTDGVRNQF----QSNSGC--GHQFALLNGCCGGSVASLFLT R ATQSVCSSQN---TATTDGVRNQF----QSNSGC--GHAVALINGCCGSDSEPSTQ D TYLGCYSSD-ALSGTKKNSYTY'---LDPAYGVNNGGSDNIAFVALFNGTCGVGGSTLTATE D LNYVACKES-SSGGAERSESVTF(7)FTRQC-NSVGF-TASHRVGCLGGRECCSTNSPNF GE EYVACLPDN-SSGTQAAVSFSAHEGLQPEAG-SAFG-STGQCLAALSEQGWCLCGAAQPSA KIAA0789 Hs KIAA0789 Hs 2 KIAA0523 Hs 1 2 1 2 KIAA0523 Hs β-1,3exoGluc β-1,3exoGluc Th Th WSC1/Sla1 Sc WSC2/YNL283c WSC3/YOL105c Sc Sc WSC4/YHL028w Sc Hp-WSC SPBC30B4.01c Hp Sp PKD1 PKD1 Hs Consensus/80% .hpa.uCb..s....sh...s.h.h...bo. .hsul.. .CaCu.. .C.ps.C... 2-structure (PHD) -EEEEEEe eeeee hHHH h EEEEe eEee DMECKG---ERGSVCG-GANRLSVYRLQ 127-219 (3882299) AQKCSA---EEFESCG-TPSYFIVYQTQ 230-324 (3882299) KIAA0789 VS----EAEC LHDREDEQLC Hs KIAA0789 Hs AQKGSA---EFFESG-TPSVFIVYQCQ 230-324 (3882299) NHECKG---EKGSVG-AVDRLSVYRVD 35-127 (3043570) GQDPEAQ--RLAEYGE-VYQTPVQDTRC 138-233 (3043570) NNNAGG---NAAETG-GYNRLDYSLA 816-914 (2924313) TMTOSG---YSENGG-GEDAYSVYQLD 21-110 (1711396) DVSCAG---WFYQNGG-GSSAMNVYINN 25-118 (1730647) GTKOSG--WFYQNGG-GSSAMNVYINN 39-132 (1077307) SGTCPG---YGYEDGGNADKDLFGYIXI 25-110 (731607) TTSCTG---YPMEKG-GSDSYSVYDE 15-101 (465861) TTPOFG--YGSLMGG-GSDSYSVYDE 15-101 (465861) TTPOFG---YGSLMGG-GLYWSVIC 1-89 (3417427) SAACTNP--QVMEVIG-ALPLPIQNVHS 198-296 (2627436) VG----LEEC LRDAMDSSVC TIAPDGNAGC KIAA0523 Hs 1 2 KIAA0523 β-1,3exoGluc Hs 1 Th B-1.3exoGluc Th 2 GPASDG**S**AQC EST---**S**SSC WSC1p/Slg1p WSC2/YNL283c Sc Sc SLTOVDSSK OVSAG -- WPYQNGG GSSANWYINN SGTC9G -- YGYEDCGNADKDLFGYIYI TTSCTG -- YGYEDCGNADKDLFGYIYI TTSCTG -- YGYEKGG GSDS%SYYDE SAACTNP - QVMEVIG GDLYMSYYITG SAACTNP -- QVMEVIG ALPLPIQNVHS WSC3/YOI 105c SLTKSTDSNC TS----VGDC WSC4/YHL028w Hp-WSC Sc dН P----DGS SPBC30B4.01c Sp VS----**S**SLC IS----ESQ 198-296 (2627436 PKD1 ------SFACLSLCSGPPAPPAPTCR-GPTLLQHVFPA .....p..Cs..C.u....p.CG.u.p.b.hh... PKD1 Hs 177-271 (1730587) Consensus/80% eeeeee 2-structure (PHD) Current Biology

following the alignment. Abbreviations: Fr, Fugu rubripes; Hp, Hansenula polymorpha; Hs, Homo sapiens; Sc, Saccharomyces cerevisiae; Sp, Schizosaccharomyces pombe; Th, Trichoderma harzianum;  $\beta$ -1,3exoGluc,  $\beta$ -1,3 exoglucanase.

## Figure S2



Multiple alignment of LH2 domains, coloured according to the consensus scheme of Figure S1. Predicted [S14] secondary structures are indicated below the alignment (e/E, extended or β-strand structure; h/H, helix); lowercase letters represent predictions that have expected accuracies of > 72% and uppercase letters represent predictions that have expected accuracies of > 82%. Residue numbers and GenBank identifiers are shown following the alignment. Abbreviations: Ce, Caenorhabditis elegans; Fr, Fugu rubripes; Gm, Glycine max (soybean); Hs, Homo sapiens; Mm, Mus musculus; Oc, Oryctolagus cuniculus (rabbit); Rn, Rattus norvegicus; LOX1, arachidonate 15-lipoxygenase; LIP2, pancreatic lipase-related protein 2.

lipoxygenases [S9]. A PSI-BLAST search [S1,S2] using PKD1 residues 3,096–3,280 as query revealed significant similarity ( $E < 10^{-6}$ ) by iteration 1 to the amino-terminal domain of mammalian arachidonate 5-lipoxygenases (Figure S2). This domain has an eight-stranded  $\beta$ -barrel fold [S9] and is highly similar in structure to the carboxy-terminal domains of mammalian lipases [S10,S11] and the amino-terminal domains of plant 15-lipoxygenases [S12] (Figure S2). We term these domains lipoxygenase homology 2 (LH2) domains. The current model for the function of LH2 domains is that they facilitate binding of lipase and lipoxygenase substrates to the enzymes' active sites [S9]. Thus the presence of this domain in the first cytoplasmic loop of PKD1 suggests a role in lipid-mediated modulation of PKD1 function.

## Supplementary references

- S1.Altschul SF, Madden TL, Shaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997, 25:3389-3402.
- S2. Altschul SF, Koonin EV: Iterated profile searches with PSI-BLAST – a tool for discovery in protein databases. *Trends Biochem Sci* 1998, 23:444-447.
- S3. Cohen-Kupiec R, Broglie KE, Friesem D, Broglie RM, Chet I: Molecular characterization of a novel β-1,3-exoglucanase related to mycoparasitism of *Trichoderma harzianum*. *Gene* 1999, 226:147-154.
- S4. Bucher P, Karplus K, Moeri N, Hofmann, K: A flexible search technique based on generalized profiles. Comput Chem 1996, 20:3-24.
- S5. Verna J, Lodder A, Lee K, Vagts A, Ballester R: A family of genes required for maintenance of cell wall integrity and for the stress response in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 1997, **94**:13804-13809.
- S6. Gray JV, Ogas JP, Kamada Y, Stone M, Levin DE, Herskowitz I: A role for the Pkc1 MAP kinase pathway of *Saccharomyces cerevisiae* in bud emergence and identification of a putative upstream regulator. *EMBO J* 1997, 16:4924-4937.
- S7. Jacoby JJ, Nilius SM, Heinisch JJ: A screen for upstream components of the yeast protein kinase C signal transduction pathway identifies the product of the *SLG1* gene. *Mol Gen Genet* 1998, 258:148-155.
- S8. Mamoun CB, Beckerich J-M, Gaillardin C, Kepes F: Disruption of YHC8, a member of the TSR1 gene family, reveals its direct involvement in yeast protein translocation. J Biol Chem 1999, 274:11296-11302.

- S9. Gillmor SA, Villasenor A, Fletterick R, Sigal E, Browner MF: The structure of mammalian 15-lipoxygenase reveals similarity to the lipases and the determinants of substrate specificity. *Nat Struct Biol* 1997, 4:1003-1009.
- S10. Roussel A, Yang Y, Ferrato F, Verger R, Cambillau C, Lowe M: Structure and activity of rat pancreatic lipase-related protein 2. *J Biol Chem* 1998, 273:32121-32128.
- S11. Winkler FW, D'Arcy A, Hunziker W: Structure of human pancreatic lipase. Nature 1990, 343:771-774.
- S12. Boyington JC, Gaffney BJ, Amzel LM: The three-dimensional structure of an arachidonic acid 15-lipoxygenase. *Science* 1993, 260:1482-1486.
- S13. Brown N, Lai J: Consensus. http://www.bork.emblheidelberg.de/Alignment/consensus.html
- S14. Rost B, Sander C: Prediction of protein structure at better than 70% accuracy. J Mol Biol 1993, 232:584-599.