FEBS Letters 538 (2003) 8–13 FEBS 27026

Minireview

Structure–function relationships of the extracellular domain of the autosomal dominant polycystic kidney disease-associated protein, polycystin-1

Benjamin S. Weston[1], Ashraf N. Malhas, Robert G. Price*

Biochemistry Section, Department of Life Sciences, King's College London, London SE1 9NN, UK

Received 20 January 2003; revised 27 January 2003; accepted 28 January 2003

First published online 12 February 2003

Edited by Veli-Pekka Lehto

Abstract Polycystin-1 (PC-1) is a member of a novel family of proteins that have a multidomain structure. Although the C-terminal intracellular segments have been extensively studied, mainly with respect to their putative involvement in cell signalling, the potential function of the extracellular domains has received less attention. Mutations in PC-1 result in autosomal dominant polycystic kidney disease (ADPKD) which is characterised by perturbation of transport resulting in fluid accumulation, cell proliferation and modification of the extracellular matrix. The possibility that the interaction of a component of the extracellular matrix or some external factor with PC-1 may be important in the initiation or progression of ADPKD cannot currently be ruled out. The purpose of this review is to assess current evidence for the function of the PC-1 extracellular domains, and their potential implications for ADPKD.

Key words: Polycystic kidney disease; Polycystin-1; Extracellular matrix protein; Cell adhesion

1. Introduction

The localisation of the *PKD1* gene to chromosome 16p13.3 [1] and the subsequent characterisation of its protein product, polycystin-1 (PC-1), heralded an intense period of research. While many aspects of PC-1 biology, including its relationship with the related gene product polycystin-2, have been well studied by different groups (comprehensively reviewed in [2,3]), the extracellular PC-1 N-terminal received less attention. The *PKD1* gene product is an integral membrane protein (PC-1) that possesses multiple secondary structural motifs [4]. The putative extracellular domain sequences of both the human and mouse *PKD1* gene product sequences are similar and both exhibit homology with the product of the *Fugu rupribes PKD1* gene [5]. The evolutionary conservation of these domains indicates that they may play an important role in the function of PC-1.

There is now evidence to suggest that three of the PC-1 extracellular protein motifs are functional [6–9]. This review aims to focus on the possible functions of the PC-1 extracellular domains, and to outline their potential contribution to the function of this multifaceted molecule.

2. Domain organisation and potential function of PC-1

The *PKD1* gene product, PC-1, is a large 4302 amino acid protein predicted to possess an extracellular region of approximately 2579 residues, multiple transmembrane domains, and a highly charged 225 residue cytoplasmic carboxy-terminus [4]. The recent expression of full-length recombinant PC-1 has confirmed that the protein possesses a predicted mass of approximately 460 kDa, and a glycosylated mass of 520 kDa [10]. Full-length PC-1 has also been immunoaffinity-purified from the 293 embryonic kidney epithelial cell line [11]. The *PKD1* coding region begins with a hydrophobic 23 amino acid sequence that displays the characteristics of a signal peptide. This, coupled with the fact that PC-1 contains several well-defined peptide motifs found in extracellular locations in other cells, led to the suggestion that almost two thirds of the protein lie outside the cell [4,12], and that the N-terminal portion of PC-1 might be capable of binding ligands in the extracellular compartment [13]. The extracellular motifs are detailed in Fig. 1, and are discussed in the following sections.

3. Leucine-rich repeats

Two complete B-type leucine-rich repeats (LRRs) are encoded by *PKD1* exons 2 and 3 [1,4,12]. LRRs are characterised by the occurrence of a 24 amino acid consensus sequence, containing leucine or other aliphatic amino acids at key positions. The other amino acids distributed among the hydrophobic residues tend to be hydrophilic and it is thought that LRRs fold in an amphipathic structure [14]. B-type repeats, like those in PC-1, contain an asparagine at position 10. LRRs are usually flanked by cysteine-rich domains that are classified as being either amino- or carboxy-flanking [15]. The PC-1 flanking regions are encoded by *PKD1* exons 1 and 4 [4,12].

*Corresponding author.
E-mail address: robert.price@kcl.ac.uk (R.G. Price).

[1] Present address: Cell and Molecular Biology Section, Faculty of Medicine, Imperial College London, London SW7 2AZ, UK.

Abbreviations: ADPKD, autosomal dominant polycystic kidney disease; LRRs, leucine-rich repeats; ECM, extracellular matrix; LDL, low-density lipoprotein; REJ, receptor for egg jelly; GPS, G protein-coupled receptor proteolytic site; PC-1, polycystin-1

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)00130-3

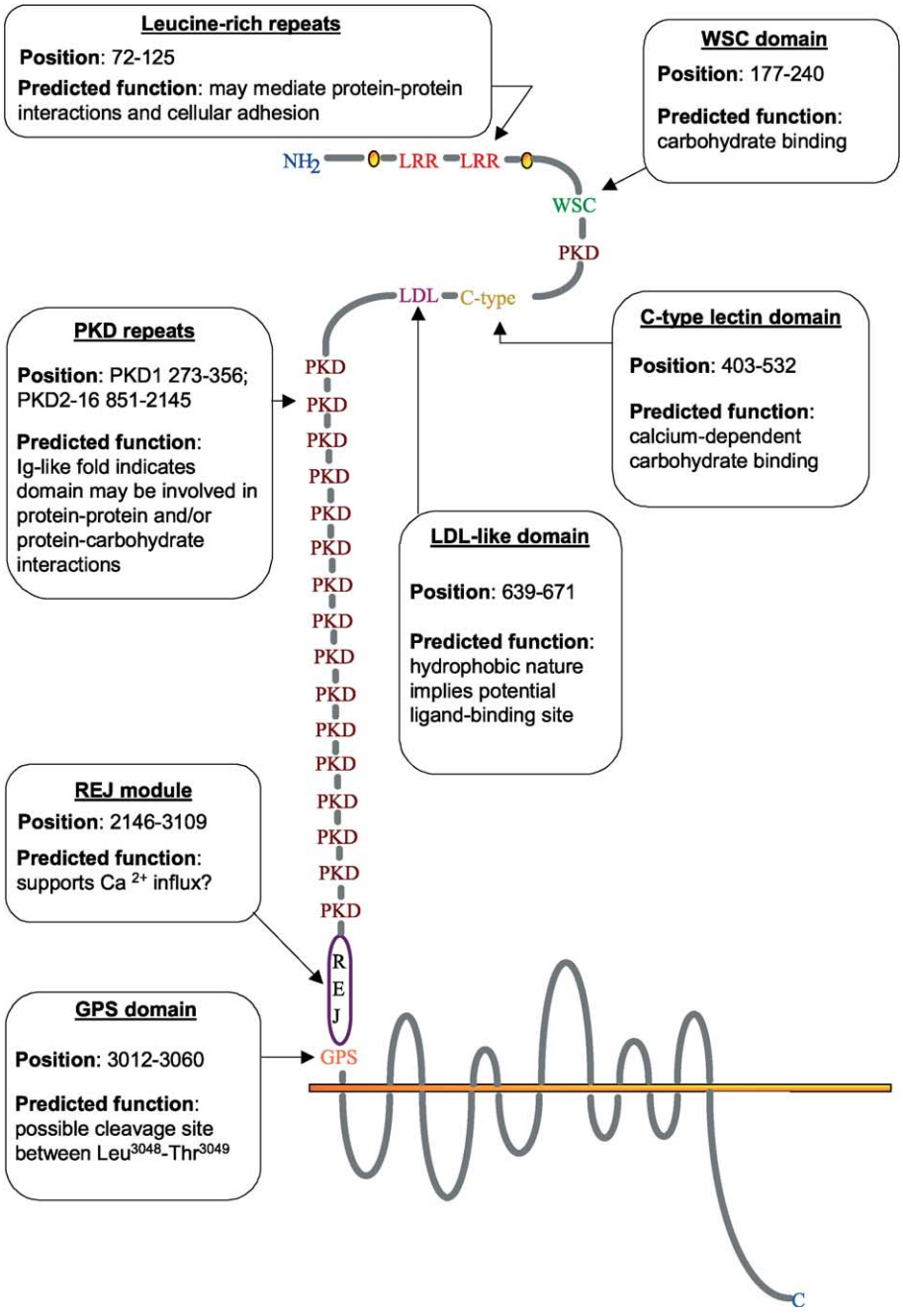

Fig. 1. Predicted structure of the *PKD1* gene product, PC-1, showing the positions and predicted functions of the extracellular protein motifs and domains. All domain positions were obtained using SwissProt using the accession number Q15141. Not to scale.

LRRs are found in a functionally and evolutionarily diverse set of proteins. The presence of such a motif is thought to provide a structural framework for the formation of protein–protein interactions, and it has been postulated that most proteins that contain LRRs are involved in protein–protein recognition processes [15]. Recently, Malhas et al. [9] used a PC-1 GST-LRR protein and demonstrated that the PC-1 LRRs bound the extracellular matrix (ECM) proteins collagen I, laminin and fibronectin. The fusion protein also interacted strongly with unidentified components of cyst fluid from autosomal dominant polycystic kidney disease (ADPKD) patients.

Over half of all proteins containing LRRs are involved in signal transduction pathways [15]. The PC-1 LRR fusion protein was shown to inhibit the proliferation of human brain astrocytoma and 293 cells, even under conditions that would otherwise be expected to stimulate proliferation [9]. There are two possible explanations for the PC-1 involvement in development and cystogenesis. The first assumes that PC-1 is part of a pathway that induces cell proliferation in response to binding ECM proteins. The addition of LRR fusion proteins to cells in culture competes with the LRRs in PC-1 resulting in the observed decrease in cell proliferation. The second assumes that the LRRs of PC-1 bind to a cell surface receptor

on adjacent cells to reduce the rate of proliferation. Both explanations suggest that the PC-1 LRRs may be involved in cell–matrix interactions and might provide means of controlling the cell's proliferative state.

4. WSC domain

A putative WSC domain, named after the cell wall integrity and stress component proteins 1–4 found in *Saccharomyces cerevisiae*, is found in PC-1 between the LRRs and the first PKD domain (see below) [16]. While the function of WSC domains is not known, the *S. cerevisiae* WSC proteins have been shown to act as upstream regulators of the stress-activated protein kinase C1–mitogen-activated protein kinase cascade, and are required for the heat shock response and for maintenance of cell wall integrity [17]. The similarity of the PC-1 WSC domain with those from a hypothetical sulphotransferase protein, and a β-1,3-exoglucanase from *Trichoderma harzianum* [18], has led to the suggestion that it may be involved in carbohydrate binding [16]. Indeed the proximity of the WSC domain to the C-type lectin domain may suggest that this region has a role involving interaction with carbohydrate, possibly acting as a regulator of stress-activated pathways if the WSC domain behaves in a similar manner to that seen in *S. cerevisiae* WSC proteins.

5. C-type lectin domain

Analysis of the deduced amino acid sequence from exons 6 and 7 of the *PKD1* gene shows a high level of homology with a C-type lectin domain [4,12]. These motifs are compactly folded protein modules of between 115 and 130 amino acids, and are characterised by the occurrence of a common sequence motif of 14 invariant and 18 highly conserved residues [19]. C-type lectin carbohydrate recognition domains bind carbohydrate in a calcium-dependent manner [19]. The C-type lectin family binds a wide range of carbohydrate ligands, with different subtypes displaying affinity for different classes of carbohydrate, and C-type lectins mediate many different biological processes in vivo, including cell signalling and exocytosis [20].

The PC-1 C-type lectin domain has been expressed as a GST fusion protein which bound carbohydrate in a Ca^{2+}-dependent manner in vitro, and interacted with different types of carbohydrate matrices [7]. Of interest to the current discussion is the observation that the PC-1 C-type lectin fusion protein bound to different ECM proteins, especially collagens I and IV. These interactions were Ca^{2+}-dependent and raise the interesting possibility that the PC-1 C-type lectin domain interacts with these ligands in vivo. However, as with all functional studies using fusion proteins, care must be taken to ensure that the recombinant protein assumes the correct fold, and that in vitro observations using isolated domains can be confirmed using approaches where the domain is present in the context of the whole protein.

6. LDL region

A low-density lipoprotein-A domain (LDL-A) is encoded by exon 10 of the *PKD1* gene [12]. LDL-A domains are cysteine-rich sequences of about 40 amino acid residues. They are present in extracellular portions of many proteins and are thought to be ligand-binding regions in LDL receptor-related molecules, predominantly because of their hydrophobic nature [21]. To date, no studies have focussed on the potential binding partners of the PC-1 LDL-A domain, perhaps because there is some confusion as to whether it occurs in PC-1 at all. The initial study by Hughes et al. [4] and the subsequent comparative study by Sandford et al. [5] failed to identify such a motif.

7. PKD repeats

A repeating module of approximately 80 amino acids, thought to represent a distinct protein family, is encoded by the *PKD1* gene. Both the human and *F. rupribes PKD1* genes are predicted to contain 16 copies of this 'PKD domain', one in exon 5, and 15 arranged in a tandem array in exons 11–15 [4,5]. Nuclear magnetic resonance structural analysis of the first PKD domain from human PC-1 revealed that it forms a characteristic Ig-like fold [22]. Similar patterns of amino acid residues in domains 2–16 also led to the prediction that these possess the same secondary structure as the first PKD domain [23]. Initial analysis of the PC-1 primary sequence showed that at least two of the PKD domains possess significant evolutionary similarity with the I-set of Ig domains, found in the cell adhesion molecules and cell adhesion receptors [4,24]. However, based on their structural data it was proposed that the PKD domains are completely distinct from this family [23]. What is clear is the importance of the PKD domains for PC-1 function, and perhaps domain 10 in particular [23]. The demonstration that the PKD domains interacted with each other in a calcium-dependent manner in vitro provided direct evidence that they may be important mediators of cell–cell interactions [8]. This study also showed that antibodies to these domains disrupt cell–cell interactions of Madin–Darby canine kidney (MDCK) cells, leading to the suggestion that loss of PKD domain interactions might be an important step in cyst formation in ADPKD. Additionally, when used as competitive inhibitors, peptides derived from the PKD repeats prevented normal ureteric bud branching in cultured kidney remnants [6]. These observations, coupled with the developmental expression patterns of PC-1 [25], suggest that PC-1, and in particular the activity of the PKD repeats, is important for normal development of the kidney.

8. REJ domain

The sea urchin receptor for egg jelly (suREJ) is a 210 kDa sperm membrane glycoprotein, and functions as a regulator of ion transport in the acrosome reaction. A large part of the REJ shares extensive sequence identity with PC-1 in a region downstream of PKD domains 2–16 [26]. This region, which has been termed the REJ module, is encoded by *PKD1* exons 15–27 [5]. More recently a human homologue of suREJ, termed PKDREJ, has been identified, which shares sequence similarity with PC-1 over 2000 amino acids, in the region corresponding to the REJ module [27]. The discovery of the REJ module supports the idea that, like suREJ, PC-1 functions to support Ca^{2+} influx [16,26,28].

9. GPS motif

PC-1 is also predicted to contain a G protein-coupled re-

ceptor proteolytic site (GPS) [16]. This domain is immediately downstream of the REJ module and also occurs in suREJ. Also known as latrophilin/CL-1-like domains, GPS domains are known to be sites of proteolytic cleavage in other proteins [29]. Despite only possessing one of the two putative disulphide bridges found in the CL-1 domains from other proteins Ponting et al. [16] predicted that PC-1 is endogenously cleaved at the Leu^{3048}-Thr^{3049} peptide bond to produce an amino-terminal chain of 3048 amino acids and a carboxy-terminal chain of 1255 amino acids. The demonstration that suREJ is cleaved at its GPS domain [30] added weight to the prediction. A recent study has now shown that PC-1 is indeed cleaved, most likely at the GPS domain, a process that requires the adjacent REJ module to be present [31]. Cleavage occurs rapidly after synthesis, although most of the N-terminal cleaved fragment remains at the cell surface. Known ADPKD mutations in the REJ module prevent cleavage and signal transduction through PC-1, leading to the suggestion that cleavage is vital for PC-1 to exhibit full biological activity, perhaps through the creation of a novel ligand-binding pocket [31].

What the consequences of cleavage at the GPS domain would mean for the N-terminal domains in PC-1 remains unclear at present. The fact that the PKD repeats are involved in mediating both cell–cell contact and renal development [6,8] and that the PC-1 LRRs inhibit cell proliferation in vitro [9] indicate that the extracellular portion of PC-1 is important for function. In addition, it has been recently demonstrated that mouse PC-1 and polycystin-2 located in the single non-motile primary cilium on kidney epithelial cells are important for calcium influx in response to physiological fluid flow [32]. The use of a blocking antibody raised against the first PKD domain of PC-1 abolished this activity, leading to the suggestion that the N-terminal portion of the PC-1 protein may act as a mechano-fluid stress sensor [31], further evidence that this region is functionally important, at least under certain circumstances.

Ectodomain shedding has been observed in a number of membrane-anchored proteins such as syndecan-1 and -4 [33]. The shed ectodomains of syndecans can still bind to their ligands and therefore compete for the same ligands as their membrane-bound proteins [33,34]. Whether protein–protein and/or protein–carbohydrate interactions mediated by the N-terminal portion of PC-1 are important post-cleavage remains to be determined.

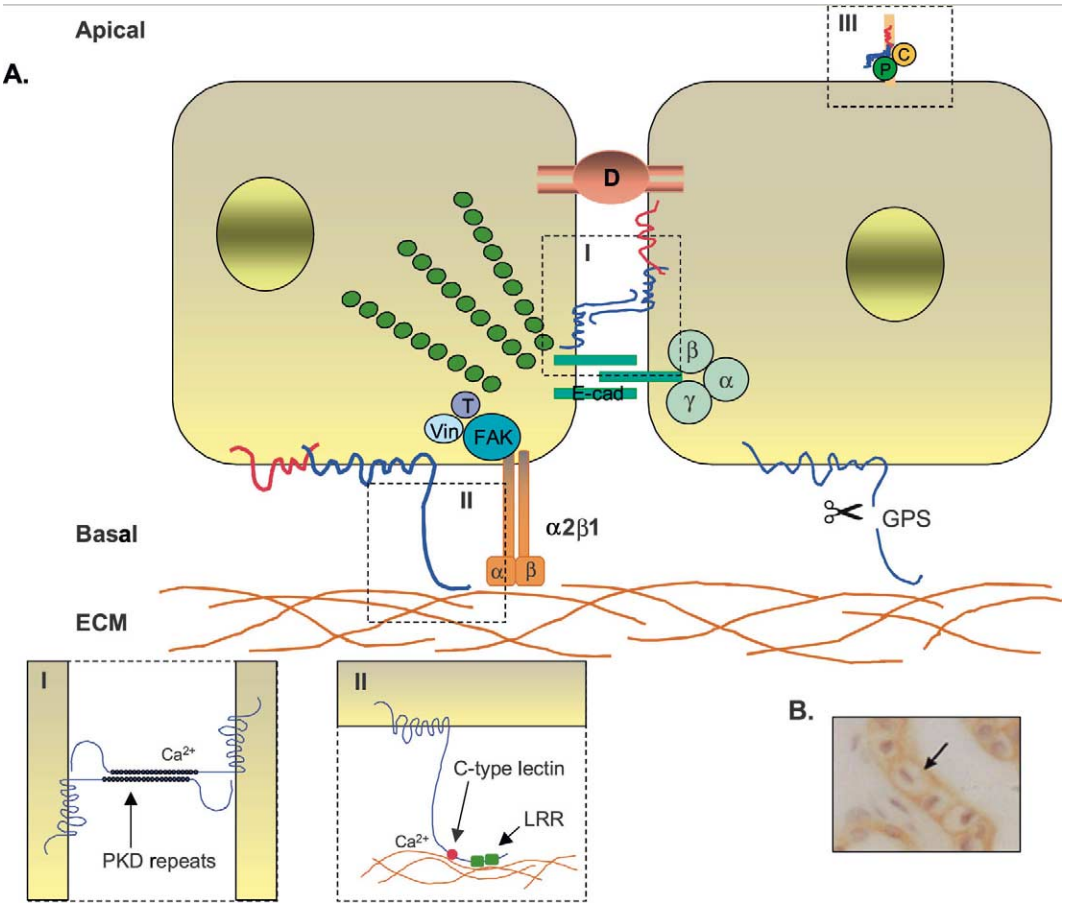

Fig. 2. Diagram showing the subcellular locations of polycystin-1 (in blue). A: Experimental evidence has been obtained that suggests that PC-1 exists in a complex with polycystin-2 (in red) and is localised (I) at points of cell–cell contact; (II) at points of cell–matrix contact; and (III) in primary cilia. Not to scale. B: Immunohistochemical staining for PC-1 showing the membrane distribution of the protein in normal renal tubular epithelial cells. The antibodies and conditions used are described in Weston et al. [25]. Abbreviations used in the figure: FAK, focal adhesion kinase; Vin, vinculin; T, talin; D, desmosome; α, β, γ, α, β and γ catenin; E-cad, E-cadherin; C, cystin; P, polaris; GPS, G protein-coupled receptor proteolytic site.

10. Polycystin subcellular localisation: at least two sides to every story

The predicted existence of multiple protein motifs towards the N-terminus of PC-1 [4] offered an insight into potential aspects of protein function, based on the assumptions that (a) the domains were active, and (b) they would display similar characteristics to identical domains that form part of other, unrelated proteins. However, the variety of protein motifs seen in the molecule has meant that there has always been an ambiguity about the location of PC-1 in the cell. There has never been any doubt about it being an integral membrane protein but whether it exists at the basal, lateral or even in the apical membrane has always proved to be less clear cut. The three do not have to be mutually exclusive and indeed, experimental evidence obtained specifically relating to the behaviour of the N-terminal motifs and the identification of possible binding/co-localising partners indicates that PC-1 may exist in at least two distinct membrane locations. If the PC-1 C-type lectin [7] and/or LRR domains [9] associate with ubiquitous ECM molecules in vivo, and PC-1 co-localises with $\alpha 2\beta 1$ integrin in focal clusters [35,36] then clearly PC-1 must exist in the basal membrane under some circumstances. Similarly, PC-1 must also exist in the lateral membrane if it co-associates with E-cadherin and the catenins [37], and desmosomes [38]. The fact that inhibition of the PKD repeats disrupts cell–cell interactions would also support this view [8]. Furthermore, more recent evidence also places both human and mouse PC-1 and lov-1, the *Caenorhabditis elegans* PC-1 homologue, in the primary cilia [32,39,40]. Our current knowledge of where PC-1 is localised within a cell is summarised in Fig. 2.

There is in vitro evidence to suggest that PC-1 may switch from predominantly basally associated to predominantly laterally associated membrane forms, when subconfluent cell monolayers reach confluence [36], introducing the idea that the location of PC-1 may perhaps be developmentally controlled. In addition further evidence for temporal regulation of location has been obtained by Bukanov et al. [41], who by monitoring hepatocyte growth factor-induced tubulogenesis from MDCK cysts found that PC-1 is only associated in the basolateral membrane at points of cell–cell contact at the time of cell polarisation and lumen formation, and not before. Using the same model this location was also seen to change during cystogenesis when a dramatic downregulation in PC-1 mRNA levels was accompanied by a shift in PC-1 to a solely cytoplasmic distribution [41]. Exactly what the potential role of PC-1 in these different subcellular locations is requires further investigation.

Long before the identity of PC-1 was known extensive analysis had revealed that the three classical defects in ADPKD cyst-lining epithelial cells were altered proliferation, fluid accumulation and ECM remodelling [42]. These secondary events assume an even greater importance in view of our current understanding of the primary events associated with cystogenesis. Whether an altered expression pattern of PC-1 in ADPKD directly contributes to any of these remains to be elucidated. It has been proposed that PC-1 may act to influence cell function by controlling both the formation and activation of key signalling complexes [13]. With this in mind it will be necessary to identify which extracellular ligands can initiate polycystin-mediated signalling pathways, in order to better understand the potential consequences for altered signal transduction in ADPKD.

References

[1] European Polycystic Kidney Disease Consortium (1994) Cell 77, 881–894.
[2] Harris, P.C. (2002) Curr. Opin. Nephrol. Hypertens. 11, 309–314.
[3] Igarashi, P. and Somlo, S. (2002) J. Am. Soc. Nephrol. 13, 2384–2398.
[4] Hughes, J., Ward, C.J., Peral, B., Aspinwall, R., Clark, K., San-Millan, J-L., Gamble, V. and Harris, P.C. (1995) Nat. Genet. 10, 151–160.
[5] Sandford, R., Sgotto, B., Aparicio, S., Brenner, S., Vaudin, M., Wilson, R.K., Chissoe, S., Bateman, A., Chothia, C., Hughes, J. and Harris, P. (1997) Hum. Mol. Genet. 6, 1483–1489.
[6] van Adelsberg, J. (1999) Dev. Genet. 24, 299–308.
[7] Weston, B.S., Bagnéris, C., Price, R.G. and Stirling, J.L. (2001) Biochim. Biophys. Acta 1536, 161–176.
[8] Ibraghimov-Beskrovnaya, O., Bukanov, N.O., Donohue, L.C., Dackowski, W.R., Klinger, K.W. and Landes, G.M. (2000) Hum. Mol. Genet. 9, 1641–1649.
[9] Malhas, A.N., Abuknesha, R.A. and Price, R.G. (2002) J. Am. Soc. Nephrol. 13, 19–26.
[10] Boletta, A., Qian, F., Onuchic, L.F., Bragonzi, A., Cortese, M., Deen, P.M., Courtoy, P.J., Soria, M.R., Devuyst, O., Monaco, L. and Germino, G.G. (2001) Am. J. Kidney Dis. 38, 1421–1429.
[11] Malhas, A.N., Abuknesha, R.A. and Price, R.G. (2001) FEBS Lett. 505, 313–316.
[12] International Polycystic Kidney Disease Consortium (1995) Cell 81, 289–298.
[13] Wilson, P.D. (2001) J. Am. Soc. Nephrol. 12, 834–845.
[14] Buchanan, S.G. and Gay, N.J. (1996) Prog. Biophys. Mol. Biol. 65, 1–44.
[15] Kobe, B. and Kajava, A.V. (2001) Curr. Opin. Struct. Biol. 11, 725–732.
[16] Ponting, C.P., Hofmann, K. and Bork, P. (1999) Curr. Biol. 9, 585–588.
[17] Lodder, A.L., Lee, T.K. and Ballester, R. (1999) Genetics 152, 1487–1499.
[18] Cohen-Kupiec, R., Broglie, K.E., Friesem, D., Broglie, R.M. and Chet, I. (1999) Gene 226, 147–154.
[19] Drickamer, K. and Taylor, M.E. (1993) Annu. Rev. Cell Dev. Biol. 9, 237–264.
[20] Gabius, H-J. (1997) Eur. J. Biochem. 243, 543–576.
[21] Krieger, M. and Herz, J. (1994) Annu. Rev. Biochem. 63, 601–637.
[22] Bork, P., Holm, L. and Sander, C. (1994) J. Mol. Biol. 242, 309–320.
[23] Bycroft, M., Bateman, A., Clarke, J., Hamill, S.J., Sandford, R., Thomas, R.L. and Chothia, C. (1999) EMBO J. 18, 297–305.
[24] Harpaz, Y. and Chothia, C. (1994) J. Mol. Biol. 238, 528–539.
[25] Weston, B.S., Jeffery, S., Jeffrey, I., Sharaf, S.A., Carter, N., Saggar-Malik, A. and Price, R.G. (1997) Histochem. J. 29, 847–856.
[26] Moy, G.W., Mendoza, L.M., Schulz, J.R., Swanson, W.J., Glabe, C.G. and Vacquier, V.D. (1996) J. Cell Biol. 133, 809–817.
[27] Hughes, J., Ward, C.J., Aspinwall, R., Butler, R. and Harris, P.C. (1999) Hum. Mol. Genet. 8, 543–549.
[28] Ikeda, M. and Guggino, W.B. (2002) Curr. Opin. Nephrol. Hypertens. 11, 539–545.
[29] Krasnoperov, V.G., Bittner, M.A., Beavis, R., Kuang, Y., Salnikow, K.V., Chepurny, O.G., Little, A.R., Plotnikov, A.N., Wu, D., Holz, R.W. and Petrenko, A.G. (1997) Neuron 18, 925–937.
[30] Mengerink, K.J., Moy, G.W. and Vacquier, V.D. (2002) J. Biol. Chem. 277, 943–948.
[31] Qian, F., Boletta, A., Bhunia, A.K., Xu, H., Liu, L., Ahrabi, A.K., Watnick, T.J., Zhou, F. and Germino, G.G. (2002) Proc. Natl. Acad. Sci. USA 99, 16981–16986.
[32] Nauli, S.M., Alenghat, F.J., Luo, Y., Williams, E., Vassilev, P., Li, X., Elia, A.E., Lu, W., Brown, E.M., Quinn, S.J., Ingber, D.E. and Zhou, J. (2003) Nat. Genet. 33, 129–137.

[33] Subramanian, S.V., Fitzgerald, M.L. and Bernfield, M. (1997) J. Biol. Chem. 272, 14713–14720.
[34] Fitzgerald, M.L., Wang, Z., Park, P.W., Murphy, G. and Bernfield, M. (2000) J. Cell Biol. 148, 811–824.
[35] Wilson, P.D., Geng, L., Li, X. and Burrow, C.R. (1999) Lab. Invest. 79, 1311–1323.
[36] Geng, L., Burrow, C.R., Li, H. and Wilson, P.D. (2000) Biochim. Biophys. Acta 1535, 21–35.
[37] Huan, Y. and van Adelsberg, J. (1999) J. Clin. Invest. 104, 1459–1468.
[38] Scheffers, M.S., van der Bent, P., Prins, F., Spruit, L., Breuning, M.H., Litvinov, S.V., de Heer, E. and Peters, D.J. (2000) Hum. Mol. Genet. 9, 2743–2750.
[39] Yoder, B.K., Hou, X. and Guay-Woodford, L.M. (2002) J. Am. Soc. Nephrol. 13, 2508–2516.
[40] Barr, M.M., DeModena, J., Braun, D., Nguyen, C.Q., Hall, D.H. and Sternberg, P.W. (2001) Curr. Biol. 11, 1341–1346.
[41] Bukanov, N., Husson, H., Dackowski, W.R., Lawrence, B.D., Clow, P.A., Roberts, B.L., Klinger, K.W. and Ibraghimov-Beskrovnaya, O. (2002) Hum. Mol. Genet. 11, 923–936.
[42] Grantham, J.J. (1992) Hosp. Pract. (Off. Edn.). 27, 51–59.