Polycystin-1

DATABANKS

MEROPS name: polycystin-1
MEROPS classification: clan PB, subclan PB(T), family T6, peptidase T06.001
Tertiary structure: Available
Species distribution: phylum Chordata
Reference sequence from: Homo sapiens (UniProt: P98161)

MEROPS name: KIAA1879 protein
MEROPS classification: clan PB, subclan PB(T), family T6, peptidase T06.003
Species distribution: phylum Chordata
Reference sequence from: Homo sapiens

MEROPS name: polycystic kidney disease 1-like 3
MEROPS classification: clan PB, subclan PB(T), family T6, peptidase T06.004
Species distribution: superfamily Hominoidea
Reference sequence from: Homo sapiens
Name and History
Polycystin-1 has an essential function in renal tubular diameter control [1] and disruption of its function by mutations of its gene PKD1 causes autosomal dominant polycystic kidney disease (ADPKD) in humans [2]. ADPKD is characterized by the development of numerous fluid-filled tubular cysts in the kidneys, which destroy the renal parenchyma and eventually lead to renal failure. Polycystin-1 functions as a cell surface signaling receptor at cell-cell/cell-matrix junctions and as a mechano-sensor in renal primary cilia that activates signaling pathways involved in renal tubular differentiation [3].

Structural Chemistry
Polycystin-1 is a 4302 amino acid 11 transmembrane (TM) glycoprotein with a large N-terminal extracellular region of 3072 residues and a short cytoplasmic C-terminal tail of ~200 residues [4] (Figure 823.1A). The extracellular region contains a novel combination of motifs known to mediate protein-protein interaction and the ~1000 residue receptor for egg jelly (REJ) module of unknown function. Situated between the REJ module and the membrane is the G protein-coupled receptor (GPCR) proteolytic site, GPS, domain [5]. The GPS domain is ~50 residues long and contains the highly conserved tripeptide His-Leu-(Thr/Ser) and several other residues including either 2 or 4 cysteine residues [5] (Figure 823.1).

Activity and Specificity
Polycystin-1 Cleavage at GPS Domain
Polycystin-1 is constitutively cleaved at His-Leu↓Thr3049 within the GPS domain probably in the endoplasmic reticulum [6,7]. The cleavage results in the N-terminal fragment (NTF) and the C-terminal fragment (CTF) (Figure 823.1A), which remain tightly associated noncovalently. These presumably form a heterodimeric molecule, a property shared by other GPS-containing proteins [6,8]. Polycystin-1 is reportedly also cleaved at unknown positions within the third intracellular loop [9] and at the cytoplasmic C-terminus [10,11], generating distinct functional fragments. The mechanism of these cleavages is, however, unknown. This article focuses on polycystin-1 cleavage at the GPS domain. The position of the three residues of His-Leu↓Thr in respect to the cleavage site is designated respectively as −2, −1, and +1 in the text.

cis-Autoproteolytic Mechanism of Polycystin-1 Cleavage at GPS Domain
Proteolytic cleavage at the juxtamembrane position occurs to many membrane proteins [12] and serves a variety of functions ranging from channel activation [13] to receptor inactivation [14] to creation of a high-affinity binding pocket for ligands [15]. The cleavage is generally mediated by various proteases along the secretory pathway [16] or at the plasma membrane [17].

However, several types of self-catalyzed protein modifications that do not require the intervention of exogenous proteases have been identified [18–20]. The best-characterized mechanism among them is cis-autoproteolysis, a self-catalyzed protein rearrangement at the His-Xaa↓(Thr/Ser/Cys) involving an ester intermediate via N→O acyl rearrangement [18,21]. This process is essential for the biological function of a diverse group of proteins that include hedgehog, nucleoporin and N-terminal nucleophile hydrolases [22], and is reviewed in Chapter 818.

The sequence of the GPS domain of polycystin-1 does not match the consensus cleavage site of any known intracellular processing proteases, nor do various inhibitors of intracellular processing proteases affect the cleavage. The cleavage tripeptide sequence His-Leu↓Thr within the GPS domain matches the cleavage site sequence His-Xaa↓(Thr/Ser/Cys) of the cis-autoproteolytic proteins instead, although no overall sequence similarity outside of it could be recognized between them (Figure 823.2A).

Recent data indicate that polycystin-1 cleavage at GPS occurs indeed through a cis-autoproteolytic mechanism analogous to that of glycosylasparaginase, an Ntn hydrolase [7] (Figure 823.2B). Consistent with this mechanism, a nucleophile hydroxyl (Thr or Ser) or thiol (Cys) group at the +1 position is necessary for cleavage. Moreover, by using a slow-cleavage mutant with a Cys at +1 position, the cleavage rate can be increased ~20-fold by the strong nucleophile hydroxylamine. This finding indicates the formation of a slowly hydrolyzed thiolester intermediate in the course of its cleavage reaction, the hallmark of cis-autoproteolysis. In a model proposed for polycystin-1 cleavage (Figure 823.2B), the folding of newly synthesized polypeptide generates an energetically unfavorable tight strain at His-Leu-Thr in the GPS domain to confer Thr(+1) its cis-autoproteolytic potential. This potential is further augmented by deprotonation of its hydroxyl group via the penultimate His(−2) (I). The thereby activated Thr(+1) launches a nucleophilic attack on the preceding α-carbonyl of Leu(−1), resulting in the formation of a transitional tetrahedral intermediate (II). The intermediate collapses by protonation of the amino group on Thr(+1) to form a more reactive ester intermediate via a reversible N→O acyl shift (III). The subsequent attack on the ester by a second nucleophile, such as a water molecule, leads to the irreversible cleavage of the scissile bond between Leu (−1) and Thr(+1) (IV). A similar conclusion was made for the adhesion GPCR EMR2 protein [23]. Collectively, cis-autoproteolysis is probably a general property of all proteins cleaved at GPS.
FIGURE 823.1  Diversity of G protein-coupled receptor proteolytic site (GPS) domain. (A) Schematic diagram of the structure of polycystin-1. LRR: leucine-rich repeat; PKD: PKD repeats; CLD: C-type lectin domain; LDLA: low density lipoprotein class A module; REJ: receptor for egg jelly module; FnIII: fibronectin III domain; GPS: G protein-coupled receptor proteolytic site domain; TM: transmembrane domain; LH2: lipoxygenase homology 2 (beta barrel) domain, also called PLAT domain. The cleavage site at GPS is marked by an arrow. The two cleavage products, NTF and CTF, are depicted. The two arrowheads indicate additional cleavage within the third intracellular loop (between 5th and 6th TM) and at the carboxy cytoplasmic tail; (B) Multiple sequence alignment of GPS sequences of polycystin-1 family proteins using ClustalW. The sequence of the GPS domains of humans (polycystin-1, PKDREJ, PKD1-L1, -L2 and -L3), sea urchin (spREJ1-9) and C. elegans (ceLov1) proteins is shown. The cleavage site between Leu and Thr at the conserved tripeptide His-Leu-Thr (red box) is marked by an arrow. The position of cleavage and the position of the three residues in respect to it (−2, −1, and +1) is indicated at the top of alignment. Proteins containing an inactive GPS domain are shaded in red; (C) Multiple sequence alignment of GPS sequences of adhesion GPCR proteins using ClustalW. Only the GPS domain of those human adhesion GPCR groups discussed in the text are shown. Proteins from other species are indicated by a prefix: sp, Strongylocentrotus purpuratus (sea urchin); ce, Caenorhabditis elegans; tn, Tetraodon nigroviridis (pufferfish); dm, Drosophila melanogaster; nv, Nematostella vectensis (sea anemone). Proteins containing an inactive GPS domain are shaded in red; and proteins with possibly inactive GPS domains are shaded in green.
Biological Aspects

**Biological Function of Polycystin-1 Cleavage at GPS Domain**

PKD1-associated mutations in the GPS domain in the neighboring REJ module, as well as synthetic mutations at the cleavage site, disrupt the cleavage and concomitantly result in loss of the functional properties of polycystin-1 to activate the JAK2-STAT pathway and induce in vitro tubulogenesis of MDCK cells in three-dimensional culture [6]. The essential function of the cleavage in vivo has been directly demonstrated by using the Pkd1 knock-in (Pkd1/V/V) mouse (Figure 823.3A), in which polycystin-1 is no longer cleaved at GPS due to the mutation of the critical nucleophile Thr(+1) to Val by homologous recombination [24] (Figure 823.3B).

The Pkd1/V/V mice exhibit a hypomorphic phenotype that differs from that of Pkd1 knock-out (Pkd1/-/-) mice in important ways (Figure 823.3C). In contrast to Pkd1/-/- mice, which die in utero with severe polycystic cystic kidney with dilation in all nephron segments and with variable multi-organ defects [25,26], the Pkd1/V/V mice are viable with macroscopically normal appearing kidneys at birth. However, they develop rapid and progressive tubular cystic dilation at the postnatal period and die of kidney failure 4–5 weeks after birth (Figure 823.3D). Remarkably, the cystic dilation affects only the distal nephron segments, with the proximal nephron segments remaining intact (Figure 823.3E). GPS cleavage of polycystin-1 is therefore not required for the embryonic development and proximal nephron segments, but is essential for proper structure and function of the distal nephron of the kidney.

The significance of the GPS cleavage can be further understood in light of the cleavage pattern of wild-type polycystin-1 in vivo [24]. The cleavage of polycystin-1 is incomplete in the tissues, including the kidney, with a small amount of uncleaved full-length polycystin-1 (polycystin-1FL) coexisting along with the more abundant cleaved products (Figure 823.3B). The non-cleavable polycystin-1V expressed in the Pkd1/V/V mice is at comparable levels as the polycystin-1FL counterpart in the wild-type mice [24]. This correspondence has led to the notion that polycystin-1V is functionally comparable, if not equivalent, to polycystin-1FL, providing the critical activity of the latter to the Pkd1/V/V mice. This supposition is supported by the nature of the mutation designed to exclusively inhibit the cis-autoproteolytic cleavage of polycystin-1. The Thr to Val mutation at +1 position changes the side chain with a nucleophile –OH group to one with a similarly sized hydrophobic –CH3 group. While effectively preventing cis-autoproteolysis, this smallest possible difference would not be expected to significantly alter the conformation surrounding the cleavage site or the overall conformation in polycystin-1V. Structural analysis of glycosylasparaginase has shown minimal effects of the mutations at the +1 position to the conformation [21]. These considerations imply differential functions of polycystin-1FL and cleaved polycystin-1, with polycystin-1FL playing an important role for the embryonic development and proximal nephron segments, and the cleaved products being essential for the distal nephron segments.
Cellular Function of GPS Cleavage

GPS cleavage may be required for intracellular trafficking and/or cell surface expression of polycystin-1. A mutation of Leu(-1) to His at His-Leu-Thr abolishes the cleavage and results in the loss of the cell surface and ciliary localization of polycystin-1 [27]. Similar findings have been described for latrophilin/CIRL [28] and GPR56 [29,30]. It should be cautioned, however, that the mutations analyzed in the studies might also influence other properties such as folding and affect cleavage secondarily. The relationship between GPS cleavage and cell surface and ciliary targeting is likely more complex, as GPS cleavage is not required for cell surface and ciliary targeting for some other polycystin-1 family members. PKDREJ is naturally uncleaved (see below), but is still distributed to the plasma membrane of the sperm [31]. Similarly, Pkd1-L1 is not cleaved at the GPS domain but colocalizes with polycystin-2 to the nodal cilia, where they function as a cilia-specific, stress-responsive channel at the node [32]. It is likely that GPS cleavage can affect the protein functions at multiple levels.

Related Peptidases

The GPS domain was first demonstrated to be the internal cleavage site for the neuronal GPCR, latrophilin/CIRL, whereby the actual cleavage site is determined to be at the tripeptide His-Leu[Thr [8]. The GPS domain was later recognized at the same position in adhesion GPCRs, which are characterized by having an unusually large and complex ectodomain [33–36]. This group of proteins is identified from invertebrates to mammals and has diverse functions ranging from clearance of apoptotic cells (BAI1) [37] to establishing planar cell polarity (Flamingo/Celsr) [38,39].

The GPS domain is present in the entire polycystin-1 family, which has four other members in humans, all with 11 transmembrane and a large ectodomain (PKDREJ, PKD1L1, PKD1L2 and PKD1L3) [40], and as many as 10 members (spREJs) in sea urchin genome [41] (Figure 823.1B). They play a role in ion transport phenomena in different systems. Polycystin-1 modulates the channel function of polycystin-2 in renal primary cilia [42,43]. Its homolog in Caenorhabditis elegans (Lov-1) is involved in the regulation of the acrosome reaction [45]. PKDREJ interacts with polycystin-2 in the nodal cilia and is required for establishing left-right asymmetry [32,46]. PKD1L3 interacts with PKD2L1, forming a cation channel in taste buds for the sensation of sour [47,48].

GPS cleavage is also critical for the adhesion GPCR GPR56, as mutations in its GPS domain dramatically impair the cleavage and cause bilateral frontoparietal polymicrogyria (BFPP), a congenital brain malformation.
in humans [49]. Cleavage of CD97 receptor, a leukocyte-restricted adhesion GPCR, is also important for the homotypic cell aggregation through upregulation of the N-cadherin expression [50]. Collectively, GPS cleavage is probably essential for the function of all proteins with a functional GPS domain.

GPS cleavage may affect distribution of the proteins to membrane subdomains. The uncleaved GPR56 as well as its N-terminal α-subunit localize exclusively to the no-raft portion, while the C-terminal β-subunit is targeted to lipid raft regions [51].

Another intriguing and complex relationship between the adhesion-GPCR subunits is discussed for latrophilin/CIRL [52], with the two subunits behaving as separate membrane proteins. The dynamic NTF provides temporally/spatially restricted docking sites for recruitment of its cognate as well as complementary CTF fragments from distinct adhesion GPCRs and activates multiple signaling pathways [53].

**Factors Affecting cis-Autoproteolysis at the GPS Domain**

cis-Autoproteolysis is based on the ability of nucleophile threonine, serine, or cysteine to initiate a proximal and reversible N–O or N–S acyl shift [19] (Figure 823.2B). Proteins are not cleaved at every one of them; they acquire cis-autoproteolytic activity only when certain conditions are met that favor the N–O or N–S shift and drive the equilibrium toward the formation of the (thio)ester. Studies of glycosylasparaginase have identified some critical parameters: (1) a highly strained and energetically unfavorable tight turn at the scissile peptide bonds; and (2) precise arrangement of distant residues around the cleavage site to generate oxygen hole and facilitate proton transfer networks [19,21]. Factors that affect these cis-autoproteolytic parameters should impinge on the property and efficiency of GPS cleavage. In fact, cleavage property varies considerably among GPS-containing proteins analyzed so far.

SuREJ3 localizes to the acrosomal region of sea urchin sperm and is completely cleaved at the GPS domain [54]. PKDREJ is, however, not cleaved [31,55], indicating that GPS cleavage is not required for its functionality in the sperm. Some adhesion GPCRs, such as GPR56 [56], BA12 [57] and Flamingo [39], are incompletely cleaved in vivo, similar to polycystin-1. Celsr2 shows a developmental stage specific cleavage pattern: it is partially cleaved in rat testis at postnatal day 7, but appears completely cleaved afterwards [58]. While differential stability of the protein products may explain these findings, several factors might affect cis-autoproteolytic cleavage at the GPS domain itself.

1. Diversity of GPS Sequence

Mutagenesis studies of polycystin-1 and others have shown that a nucleophile residue (Thr, Ser or Cys) at the +1 position of His-Leu-Thr is essential for cleavage, in accordance with the cis-autoproteolytic mechanism [6,7]. A survey of all 36 GPS-containing proteins present in the human RefSeq collection reveals that five (14%) of them contain a GPS domain that does not have such a residue at the position and are predicted to be cis-autoproteolytically inactive (Figure 823.1C, shaded in red). Of the five human polycystin-1 family members, PKDREJ and PKD1-L1 in fact belong to this category (Figure 823.1B). The sea urchin member suREJ2 also contains an inactive GPS domain but the experimental data are inconclusive [59].

Interesting GPS diversity exists within the Flamingo orthologs. The three mammalian orthologs, Celsr1–3 have distinct functions in regulating cell planar polarity and in nervous system development [38,60]. In contrast to Celsr2, Celsr1 and Celsr3 both contain an inactive GPS domain and are predicted to be uncleaved at the site (Figure 823.1C). GPS cleavage is probably not required for their functions. Two Flamingo orthologs are identified in pufferfish [36], with one of them (tnCelsr2-1) predicted to be uncleavable. CeFmi-1, the only Flamingo ortholog in *C. elegans*, also contains an inactive GPS domain. Consistent with this notion, the GPS domain is not specifically required in either pioneer or follower axon navigation [61]. In contrast, the only Flamingo in *Drosophila* [39] and its homolog in *Nematostella vectensis* [36] contain an active GPS domain. Analysis of the GPS cleavage of the Flamingo family may shed light on their biological functions.

The other three adhesion GPCRs with an inactive GPS domain are: EMR1, GPR115 and GPR124 (Figure 823.1C). GPR124 regulates CNS angiogenesis [62] and undergoes ectodomain shedding upon VEGF stimulation [63], suggesting an inducible sheddase likely being responsible to cleave the protein at a juxtamembrane position distinctive from the GPS. GPR111 and GPR125 have a GPS domain with an unconventional residue at the −2 position (Leu and Ser respectively). It is unclear whether these two GPS domains are inactive too.

2. Adjacent Sequence

Even an active GPS domain does not function as an autonomous cleavage signal. The GPS domain alone is not sufficient to support the cleavage of polycystin-1 and requires the adjacent REJ module for the reaction to occur [6]. The REJ module is not a single-folded domain but instead contain at least four FnIII domains [64]. It might be responsible for generating a proper conformation and/or establishing a proton transfer network required for cis-autoproteolysis at the GPS. This could explain why PKD1
missense mutations or in-frame deletions in the neighboring REJ module can disrupt cleavage at distance [6,65].

Adhesion GPCRs do not have an REJ module, but contain a structurally undefined region of various length immediately N-terminal to the GPS domain. This region may have a similar function in setting up cleavage competency of the GPS domain. A naturally occurring isoform of EMR2, which lacks 11 residues in the mucin-like stalk adjacent to the GPS domain by alternative splicing, is interestingly uncleaved [66]. A recent study identified 53 human splice variants among the adhesion GPCRs [67]. Several of the functional splice variants lack one or more of the functional domains near the GPS domain, while the others lack a GPS domain entirely. Alternative splicing may therefore provide a means to determine the cleavage property of the proteins, whereby controlling the relative abundance of the GPS-cleaved and uncleaved forms of the proteins in vivo.

3. Other Factors

Polycystin-1 has intrinsically limited cis-autoproteolytic capacity. The pulse-labeled recombinant polycystin-1 expressed in the mammalian cells begins cleavage after 15 min of chase and gradually decreases the amount to a level equal to that of increasing cleavage product NTF within 2 h, a ratio that remained unchanged for at least another 3 h in the cells [6]. This finding suggests that about half of the nascent polycystin-1 molecules complete cleavage within 2 h, whereas the other half remains uncleavable for a prolonged period of time before being degraded. At steady state, the uncleaved molecules might be mostly in this uncleavable state. This consideration has led to a model in which newly synthesized polycystin-1 can proceed through two competing pathways: the ‘cleavage’ pathway, which leads to irreversible cis-autoproteolytic cleavage, and the ‘non-cleavage’ pathway, which leads to trapping in the blocked state (Figure 823.2C) [7,68]. N-glycosylation at specific sites in the mucin-like stalk region and GPS domain of EMR2 has a profound effect on its cleavage [68]. Certain N-glycosylation patterns can promote protein folding to a more favorable conformation for the GPS cleavage and possibly commit the protein to the ‘cleavage’ pathway. N-glycosylation may therefore regulate the GPS cleavage leading to the production of either cleaved or uncleaved molecules. Other cellular factors that may affect cleavage efficiency include ligand binding or mechanical force in the case of polycystin-1.

Acknowledgments

This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) grants R01 DK062199 and P30 DK090868.

Further Reading

Reviews of polycystic kidney disease have been published by Harris & Torres [2] and Menezes & Germino [3]. Reviews of cis-autoproteolysis are published by Paulus [19] and Nathan Aronson in this volume (Chapter 818). Reviews of adhesion GPCRs have been published by Yona et al. [34] and Bjarnadóttir et al. [69].

References


Feng Qian
Department of Medicine, Division of Nephrology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. Email: fjshupkd@jhmi.edu