COMMUNICATIONS BIOLOGY

ARTICLE

https://doi.org/10.1038/s42003-019-0586-0 OPEN

Identification of an adhesive interface for the non-clustered δ1 protocadherin-1 involved in respiratory diseases

Debadrita Modak[1] & Marcos Sotomayor[1]

Cadherins form a large family of calcium-dependent adhesive proteins involved in morphogenesis, cell differentiation, and neuronal connectivity. Non-clustered δ1 protocadherins form a cadherin subgroup of proteins with seven extracellular cadherin (EC) repeats and cytoplasmic domains distinct from those of classical cadherins. Non-clustered δ1 protocadherins mediate homophilic adhesion and have been implicated in various diseases including asthma, autism, and cancer. Here we present X-ray crystal structures of human Protocadherin-1 (PCDH1), a δ1-protocadherin member essential for New World Hantavirus infection that is typically expressed in the brain, airway epithelium, skin keratinocytes, and lungs. The structures suggest a binding mode that involves antiparallel overlap of repeats EC1 to EC4. Mutagenesis combined with binding assays and biochemical experiments validated this mode of adhesion. Overall, these results reveal the molecular mechanism underlying adhesiveness of PCDH1 and δ1-protocadherins, also shedding light on PCDH1's role in maintaining airway epithelial integrity, the loss of which causes respiratory diseases.

[1] Department of Chemistry and Biochemistry, The Ohio State University, 484 W 12th Avenue, Columbus, OH 43210, USA. Correspondence and requests for materials should be addressed to M.S. (email: sotomayor.8@osu.edu)

Cells are the basic units of life and selectively organize themselves to form tissues with the help of various adhesion proteins that engage in homophilic (same type) and heterophilic (different type) contacts. Cell–cell adhesion is of prime importance for the formation and maintenance of multicellular structures. Some of the main families of adhesion proteins include integrins, selectins, cadherins and members of the immunoglobulin superfamily[1–4].

Cadherins are calcium-dependent adhesion glycoproteins involved in a variety of biological processes such as cell differentiation and tissue morphogenesis[5], cell signaling[6–8], mechanotransduction[9–11], and brain morphogenesis and wiring[12–16]. These proteins have a unique structure including tandem extracellular cadherin (EC) repeats[5,17,18], followed by a transmembrane and a cytoplasmic domain. The highly similar EC repeats are around 100 amino acids long and feature highly conserved calcium-binding motifs that bind three calcium ions at the linker regions between the repeats.

Non-clustered protocadherins are non-classical cadherins expressed in various tissues all over the body and have a wide range of functions. They are typically placed in two groups: δ and ε protocadherins[19]. The δ-protocadherins have seven (δ1) or six (δ2) EC repeats, a single transmembrane helix, and a cytoplasmic domain containing highly conserved motifs of unclear function (CM1, CM2, and CM3 in δ1; CM1 and CM2 in δ2)[19,20]. These play a major role in neuronal tissue connectivity and brain wiring and are involved in various neurological disorders such as epilepsy, autism, and schizophrenia[20–22].

Protocadherin-1 (PCDH1) is a non-clustered δ1 protocadherin involved in respiratory diseases. It was first identified by Sano and co-workers[23] as PC42 in rat brain and retina tissue. Later on, it was found to be mainly expressed in the airway epithelium, bronchial epithelium, lungs, skin keratinocytes and also in kidney, neural, and glial cells[24–28]. PCDH1 has seven EC repeats (labeled EC1 to EC7 from N to C-terminus), a single pass transmembrane domain and a cytoplasmic domain with the three conserved motifs (CM1–CM3) typical of the δ1 subgroup. Immunofluorescence microscopy experiments showed that L cells transfected with PCDH1 DNA aggregated in a calcium-dependent manner and expressed the protein along the boundary of the cells at cellular contact sites[23]. These results suggested that PCDH1 mediated aggregation through calcium-dependent homophilic interactions.

PCDH1 has been associated with asthma, a chronic inflammatory disorder of the airways characterized by wheezing, coughing, breathlessness, chest-tightness, bronchial hyperresponsiveness, and obstruction to airflow. It is a widespread health problem affecting people of all ages, especially children, and is caused by interaction of genetic and environmental factors[29,30]. The epithelial layer of the airway mucosa and bronchia acts as a barrier and prevents the passage of allergens and pathogens. Dysfunction of this epithelial layer barrier is thought to play a role in asthma[31]. PCDH1 may bind bronchial epithelial cells and epithelial cells of the airway mucosa together by mediating homophilic adhesion and helping in forming the epithelial barrier. Loss of adhesion by PCDH1 or underexpression of the protein may lead to loss of epithelial integrity and pathogenesis[31]. However, how PCDH1 and other members of the δ1 protocadherin family can mediate cellular adhesion is unclear.

PCDH1 has also been found to be essential for New World hantaviruses infection[32]. These viruses are transmitted from rodents to humans and cause a fatal respiratory disease called hantavirus pulmonary syndrome[33,34]. There are no known vaccines or specific treatments for this disease in humans[35]. Results from in vitro and in vivo experiments indicate that direct recognition of PCDH1 EC1 by Gn/Gc glycoproteins on the virus envelope is a necessary step for infection[32]. However, the specific molecular mechanism by which PCDH1 EC1 interacts with the viral glycoproteins, and whether this interaction disrupts PCDH1 homophilic adhesion, remains to be elucidated.

The molecular mechanisms underlying adhesion by classical cadherins are well established. These proteins use their extracellular domains to form an intercellular (*trans*) bond that is mostly homophilic (between identical cadherins)[36,37]. The *trans*-homophilic bond is formed by contacts between the most membrane-distal EC repeats (EC1–EC1) coming from adjacent cells[36,38]. The EC1–EC1 contacts are mediated by the exchange of an N-terminal β-strand containing one or two tryptophan residues (Trp2 or Tpr2 and Trp4) that are docked into the hydrophobic pocket of the binding partner, thus forming a "strand-swapped" dimer.

Non-classical cadherins lack the tryptophans involved in classical adhesion and use different binding mechanisms. Cadherin-23 (CDH23) and protocadherin-15 (PCDH15) form a heterophilic bond essential for inner ear mechanotransduction[9,11]. The heterophilic CDH23–PCDH15 bond involves an antiparallel "extended handshake" arrangement that uses both EC1 and EC2 repeats[39]. In contrast, the clustered protocadherins (α, β, γ) and the non-clustered δ2 member PCDH19, all featuring six EC repeats, mediate *trans*-homophilic interactions using their first four extracellular repeats (EC1–4) arranged in an antiparallel bond, where EC1 interacts with EC4, EC2 with EC3, EC3 with EC2, and EC4 with EC1[40–45]. Missense mutations in PCDH19 that impair the formation of this bond are known to cause a form of epilepsy[42].

The calcium-dependent adhesion mechanisms for various other non-classical cadherins, including the non-clustered δ1 protocadherins, are still poorly understood. Here we use bead aggregation assays to show that repeats EC1 to EC4 form the minimal adhesive unit for human (*Homo sapiens* [*hs*]) PCDH1, despite its longer extracellular domain with seven EC repeats. In addition, we present crystal structures of *hs* PCDH1 EC1-4 and show that an antiparallel overlapped dimer mediates *hs* PCDH1 adhesion. The discovery of this mode of adhesion, likely used by all other members of the δ1 protocadherin family, along with our structures, might help in understanding the molecular mechanisms underlying PCDH1-related respiratory diseases such as asthma and hantavirus pulmonary syndrome.

Results

Binding assays determine EC repeats required for adhesion. Previous studies have shown that the full-length PCDH1 protein mediates adhesion in cell-based assays[23]. To determine which EC repeats are essential for PCDH1 homophilic adhesion we created a library of constructs from the human protein (hereafter referred to as PCDH1) by deleting EC repeats from the C- and N-terminus. Fc-tagged protein constructs were produced in HEK293T cells and binding assays were done with Protein G magnetic beads (see Methods). Results from bead aggregation assays of PCDH1 EC1-7Fc showed that PCDH1's full-length extracellular domain mediates calcium-dependent homophilic adhesion (Fig. 1a). Aggregation was abolished in the absence of calcium (2 mM EDTA) (Supplementary Fig. 1a). Next, to identify the minimum unit of homophilic adhesion mediated by the extracellular domain of PCDH1, bead aggregation assays were performed with PCDH1 EC1-6Fc, EC1-5Fc, EC1-4Fc, EC1-3Fc, EC1-2Fc, and EC2-7Fc (Fig. 1b–i). PCDH1 EC1-2Fc, and EC1-3Fc showed no aggregation while PCDH1 EC1-4Fc, EC1-5Fc, EC1-6Fc, and EC1-7Fc showed calcium-dependent aggregation (Fig. 1, Supplementary Fig. 1a–h). While bead aggregation assays are only semi-quantitative, bead aggregates mediated by PCDH1

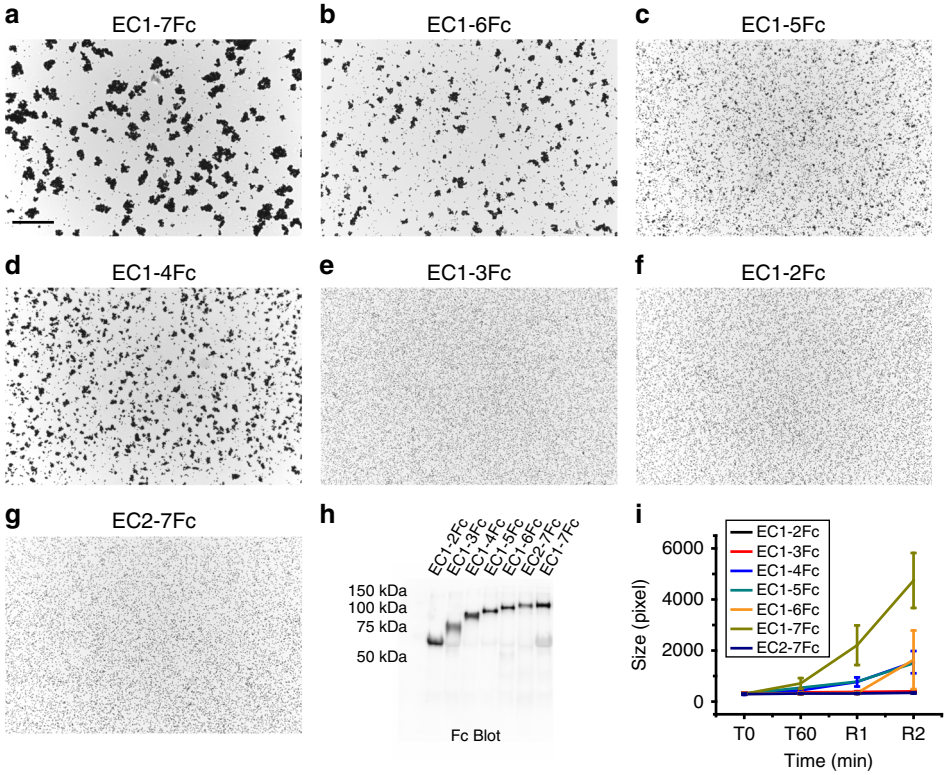

Fig. 1 Minimal adhesive unit of PCDH1 is EC1–4. **a**–**g** Protein G beads coated with full length (**a**) and truncated versions (**b**–**g**) of the PCDH1 extracellular domain imaged after incubation for 1 h followed by rocking for 2 min in the presence of calcium. Scale bar: 500 μm. **h** Western blot shows efficient expression of full length and truncated PCDH1 EC repeats. **i** Mean aggregate size for full length and truncated fragments of PCDH1 at T0 ($t = 0$ min), after 1 h of incubation, T60 ($t = 60$ min) followed by rocking for 1 min (R1) and 2 min (R2). Error bars are standard error of the mean ($n = 4$ independent experiments for all aggregation assays and constructs except for PCDH1 EC1-7Fc with $n = 3$ independent experiments)

EC1-7Fc were larger than those observed for truncated extracellular domains (Fig. 1i), a trend also seen in experiments with PCDH19 (ref. [42]). There was no aggregation when EC1 was truncated from the N-terminus (PCDH1 EC2-7Fc) (Fig. 1g). These experiments suggest that PCDH1 EC1-4 is the minimum unit of homophilic adhesion and that EC1 is required for adhesion.

A disulfide bond independent dimer mediates adhesion. In parallel to bead aggregation assays carried out with protein from HEK293T cells, we used PCDH1 EC1-4 and PCDH1 EC1-5 fragments refolded from *Escherichia coli* inclusion bodies for binding assays that required large sample amounts. Refolding resulted in two monodisperse peaks corresponding to a monomer and a dimer mediated by an intermolecular cysteine-disulfide bond (Supplementary Fig. 2a). To test if free cysteines were necessary for PCDH1 dimerization, cysteine residues (C375 and C548) that were presumed to not be involved in intramolecular disulfide bonds were mutated to serine, thus abolishing potential PCDH1 intermolecular disulfide bonds. Bead aggregation assays performed with PCDH1 EC1-7Fc C375S C548S showed that PCDH1 mediated bead aggregation independent of these cysteines (Fig. 2a, b, Supplementary Fig. 2b). Therefore, to avoid confounding effects all binding assays of PCDH1 EC1-4 and EC1-5 obtained from bacteria were done with proteins carrying the C375S (EC1-4, EC1-5) and C548S (EC1-5) mutations.

Analytical ultracentrifugation (AUC) sedimentation velocity experiments with a range of concentrations of PCDH1 EC1-4 C375S showed both monomeric and dimeric states in solution. The ratio of monomer to dimer was dependent on the concentration (Fig. 2c). Concentration-dependent monomeric and dimeric states were also observed in sedimentation velocity experiments of PCDH1 EC1-5 C375S C548S (Fig. 2d). Samples were collected after AUC experiments and were run on SDS-PAGE, which indicated lack of detectable degradation of the protein fragments (Supplementary Fig. 2c, d). Dimerization was observed for both PCDH1 EC1-4 C375S and PCDH1 EC1-5 C375S C548S fragments with the dimer dissociation constant (K_D) estimated to be less than 5 μM. These results show that a dimer not formed by disulfide bonds mediates bead aggregation and exist in solution.

X-ray crystal structures reveal the architecture of PCDH1. Human PCDH1 protein fragments produced by bacterial and mammalian cells were used for crystallization and structure determination attempts. Three crystal structures were determined through molecular replacement and fully refined. The first two, PCDH1 EC3-4bc at 3.05 Å and PCDH1 EC1-4bc at 2.85 Å, were obtained using bacterially produced protein. The third one, PCDH1 EC1-4mc at 3.15 Å, was obtained using protein produced by mammalian cells (Fig. 3, Table 1). The PCDH1 EC3-4bc structure (T216 to Q438) had one molecule in the asymmetric unit, had poor density for three loops in EC4 that were not modeled, and it was used as an initial model for phasing the PCDH1 EC1-4bc structure by molecular replacement. Both PCDH1 EC1-4 structures had one molecule in the asymmetric unit with good-quality electron density maps that allowed for the unambiguous positioning of side chains for most of the residues (Supplementary Fig. 3a, b). These included V3 to N443 in PCDH1 EC1-4bc and V3 to V439 in PCDH1 EC1-4mc, with both missing eight residues in their EC1 repeat, and with PCDH1 EC1-4bc missing four residues in its EC4 repeat. The root mean square

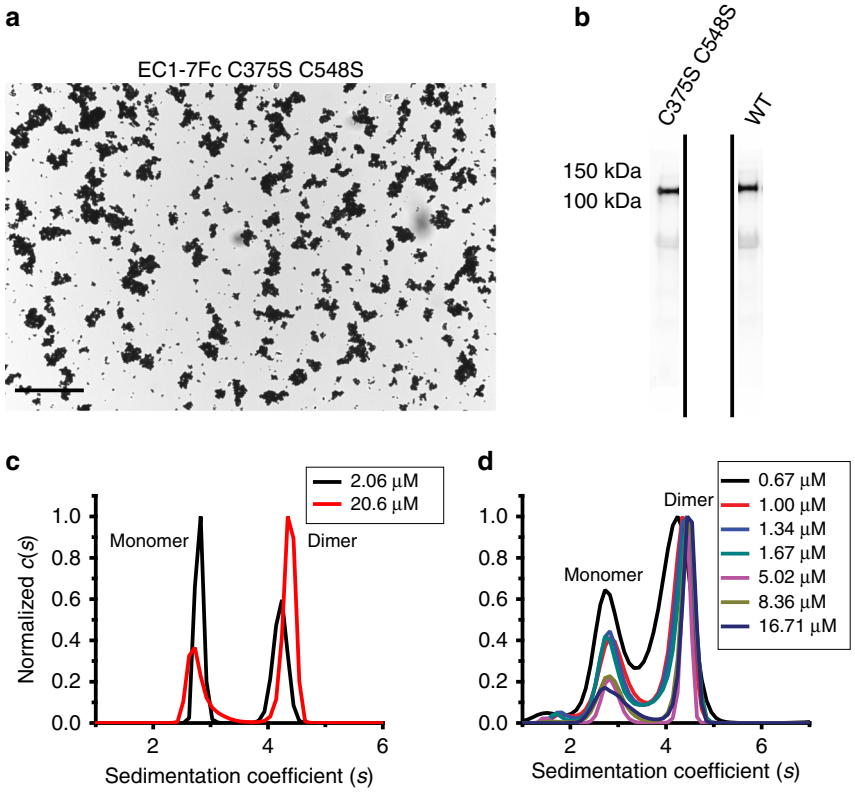

Fig. 2 PCDH1-mediated bead aggregation is independent of disulfide bond formation and PCDH1 EC1-4 C375S and PCDH1 EC1-5 C375S C548S form dimers in solution. **a** Protein G beads coated with the PCDH1 EC1-7Fc C375S C548S mutant ($n = 3$ independent experiments) imaged after incubation for 1 h followed by rocking for 2 min in the presence of calcium. Scale bar: 500 μm. Bead aggregation is evident. **b** Western blot shows efficient expression of wild-type PCDH1 EC1-7Fc and mutant PCDH1 EC1-7Fc C375S C548S. Full blot is shown in Supplementary Fig. 2b. **c** AUC data from bacterially produced PCDH1 EC1-4 C375S ($n = 2$ independent experiments) show both dimeric and monomeric peaks at 2.06 and 20.6 μM with normalized c(s) plotted against sedimentation coefficient (s). **d** AUC data from bacterially produced PCDH1 EC1-5 C375S C548S ($n = 2$ independent experiments) show both dimeric and monomeric peaks for a range of concentrations between 0.67 and 16.71 μM

deviation (RMSD) for the Cα atoms among the chains of the two structures was ~1 Å, suggesting that glycosylation was inconsequential for the models. The RMSD for the shared Cα atoms among the chains of PCDH1 EC1-4bc and PCDH1 EC3-4bc was also low (~1.5 Å). Since protomers were highly similar, we will describe features as seen in the PCDH1 EC1-4bc structure, unless otherwise explicitly stated.

PCDH1 EC1-4 has four tandemly arranged EC repeats with similar architecture (seven β-strands labeled A to G arranged in a Greek-key motif; Fig. 3a, b). Although PCDH1 EC1-4 is somewhat straight as compared to the classical cadherins, which are curved, the EC repeats have the same overall classical architecture. There are three calcium ions in each complete linker region (EC1–EC2, EC2–EC3, and EC3–EC4) and one calcium ion at the end of EC4. Calcium ions at complete linker regions are coordinated by highly conserved canonical calcium-binding motifs XEX^{BASE}, DXD, DRE/DYE, XDX^{TOP} and DXNDN (Supplementary Fig. 4). An additional calcium ion near the EC2–EC3 linker is unique, it is at a crystal contact, and it might be present due to purification and crystallization conditions with high calcium (Fig. 3b, d). The PCDH1 EC1-4mc structure does not have this additional calcium ion. Analyses of the structures and sequences of the linker regions of PCDH1 and various classical (*mm* CDH1 and *mm* CDH2), non-clustered (*dr* PCDH19) and clustered protocadherins (*mm* PCDHα7, *mm* PCDHβ6, and *mm* PCDHγB7) show that linkers' length and structure are similar across EC1–2, EC2–3, and EC3–4 for all proteins except classical cadherins, which have a somewhat longer and bent EC3–4 linker (Supplementary Table 1). This may explain the somewhat straighter conformation observed for protocadherins.

Another notable feature observed in the PCDH1 EC1-4 structure is the presence of a 10-residue long (C66-C75) disulfide loop in EC1, between β-strands E and F (Fig. 3c). Shorter disulfide loops (seven residues) have been previously seen in the clustered and δ2 protocadherin structures at this location. The presence of an extended insertion with 11 residues (I335-A345) within β-strand A in EC4, which we denote as "δ-insertion", is also unique and not seen in other clustered protocadherin structures (Fig. 3e, Supplementary Fig. 5). Similarly, a six-residue insertion in the DE loop of EC4 (G392-K397) is found only in PCDH1 sequences (Supplementary Fig. 5, Supplementary Table 2). Both insertions in EC4 are disordered in the PCDH1 EC3-4bc structure, suggesting that they become ordered in the crystallographic arrangement of the PCDH1 EC1-4 structures. These unique features might be relevant for function, as further discussed below.

There are three cysteine residues found in the EC1–4 sequence, two of which are involved in an intramolecular disulfide bond (C66–C75) as discussed above. The third one, C375, is partially buried in both structures (from bacterial and mammalian expressed proteins). This suggests that potential disulfide-mediated dimers observed in refolded PCDH1 EC1-4 might be rare or artefactual.

The PCDH1 EC1-4mc structure had regions in which electron density could only be explained by the presence of glycosylation. Residues N89, N248, and N346 are predicted to be N-linked glycosylation sites and sugars can be observed at N248 in the

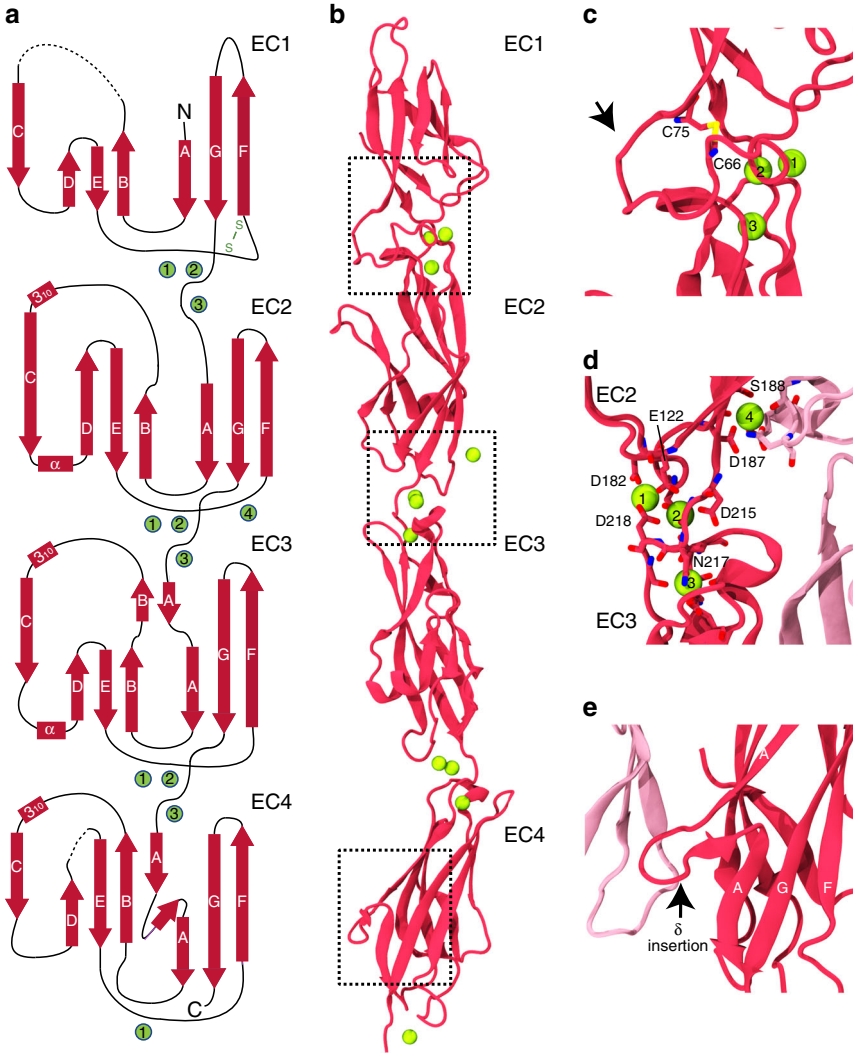

Fig. 3 X-ray crystal structure of PCDH1 EC1-4. **a** Topology diagram of PCDH1 EC1-4bc. A typical cadherin fold is observed for each EC repeat with seven β-strands labeled A to G. Calcium ions are shown as green circles. Dashed lines are missing loops. **b** Ribbon diagram of PCDH1 EC1-4bc (6BX7). Black dashed boxes indicate unique features of PCDH1 EC1-4 highlighted in panels **c**–**e**. **c** Details of a disulfide loop in EC1 (arrow). **d** Details of the calcium-binding site in the EC2–EC3 linker region. **e** Details of an extended δ insertion within β-strand A in EC4 not seen in clustered protocadherins (arrow)

PCDH1 EC1-4mc structure. C-linked glycosylation has been predicted for W186 but is not observed in our structure, while O-linked glycosylation has not been predicted but it is observed at S430 (Supplementary Fig. 3b). We could not detect structural differences caused by glycosylation, and thus the role of these glycosylation sites in PCDH1 function remains unclear.

Antiparallel interfaces in the crystal structures of PCDH1. Physiologically relevant adhesive interfaces have been observed in crystal structures of classical cadherins, of the PCDH15 and CDH23 heterophilic complex, and of clustered and non-clustered protocadherins[18,39,40,42,43,46,47]. Crystal contacts in the PCDH1 EC3-4bc structure are not expected to reveal a physiologically relevant interface as their areas are small (Supplementary Fig. 6a–d) and repeat EC1, required for *trans* binding, is absent. Nevertheless, some of these contacts could represent low-affinity transient intermediates for both *trans* and *cis* (from the same cell) interactions. Importantly, the crystal structures of PCDH1 EC1-4 obtained from both bacterial and mammalian cells, which have the same crystal packing, reveal one small and likely non-physiological or transient interface involving EC2 contacts (Supplementary Fig. 6e), as well as two large plausible *trans* interfaces of homophilic interaction that may mediate adhesion (Fig. 4).

The first large homophilic adhesion interface (PCDH1-I1) involves two PCDH1 EC1-4 molecules having complete antiparallel overlap. In this arrangement, EC1 of one molecule interacts with EC4 of the second molecule (EC1:EC4), EC2 with EC3 (EC2:EC3), EC3 with EC2 (EC3:EC2), and EC4 with EC1 (EC4:EC1) (Fig. 4a, b). The second possible interface of adhesion (PCDH1-I2) involved the opposite side of PCDH1 EC1-4, having antiparallel overlap of EC2 of one molecule with EC4 of the other (EC2:EC4), EC3 with EC3 (EC3:EC3), EC4 with EC2 (EC4:EC2), and potential EC1:EC5 contacts (with EC5 not present in the structure; Fig. 4c, d). These two possible interfaces were analyzed and tested to determine which one mediates adhesion in vitro and potentially in vivo.

Analyses of the interaction interfaces with the Protein, Interfaces, Surfaces and Assemblies (PISA)[48] server revealed large interface areas of 1539 Å^2 for PCDH1-I1 (385 Å^2 per repeat; 486 Å^2 for EC1:EC4 and EC4:EC1, 231 Å^2 for EC2:EC3 and EC3:EC2; Tables 2) and 1270 Å^2 for PCDH1-I2 (423 Å^2 per repeat; 461 Å^2 for EC2:EC4 and EC4:EC2, 190 Å^2 for EC3:EC3) respectively. Both interface areas are greater than 856 Å^2, an

Table 1 Data collection and refinement statistics[a]

	Hs PCDH1 EC1-4bc	*Hs* PCDH1 EC1-4mc	*Hs* PCDH1 EC3-4bc
Data collection			
Space group	$P6_122$	$P6_122$	$P4_22_12$
Cell dimensions			
a, *b*, *c* (Å)	146.98, 146.98, 147.69	147.23, 147.23, 149.37	125.57, 125.57, 45.29
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 90
Molecules per asymmetric unit	1	1	1
Resolution (Å)	2.85	3.15	3.05
$I/\sigma(I)$	18.3 (1.923)	58 (2.5)	12 (1.67)
R_{merge}	0.156 (1.627)	0.078 (1.762)	0.166 (1.027)
R_{meas}	0.164 (1.710)	0.079 (1.792)	0.175 (1.110)
R_{pim}	0.050 (0.521)	0.013 (0.321)	0.054 (0.401)
$CC_{1/2}$	0.998 (0.716)	1.004 (0.794)	0.938 (0.551)
CC^*	0.999 (0.913)	1.001 (0.941)	0.981 (0.843)
Completeness (%)	100.0 (100.0)	100.0 (99.9)	99.3 (90.6)
Redundancy	10.7 (10.6)	36.4 (28.2)	9.9 (6.5)
Refinement			
Resolution (Å)	48.26–2.85 (2.93–2.85)	48.53–3.15 (3.22–3.15)	44.43–3.05 (3.13–3.05)
No. of reflections	22,567	17,231	7363
R_{work}/R_{free}	22.1/24.3	21.5/26.3	21.9/26.6
No. of atoms			
Protein	3288	3284	1533
Ligand/ion	26	34	4
Water	25	0	6
B-factors			
Protein	78.75	152.67	64.93
Ligand/ion	87.18	127.24	60.20
Water	47.34	0	37.05
R.m.s. deviations			
Bond lengths (Å)	0.010	0.0095	0.0093
Bond angles (°)	1.484	1.503	1.536
Ramachandran plot (%)			
Most favored	87.8	83.2	86.0
Additionally allowed	11.7	16.8	14.0
Generously allowed	0.5	0.0	0.0
Disallowed	0.0	0.0	0.0
PDB ID code	6BX7	6MGA	6PIM

[a]One crystal was used for each structure. Values in parentheses are for highest-resolution shell

empirical cut-off to distinguish biological interfaces from crystal contacts. Therefore both are unlikely to be crystal packing artifacts[49]. Interestingly, the interface area per EC repeat is larger for PCDH1-I2.

Glycosylation sites, if present at the interface, may interfere with binding and thus reveal non-physiological interfaces as in the case of VE-cadherin[50]. Analyses of the N- and C-linked predicted glycosylation sites showed that residues N89 and N346 are not present in either of the interfaces. In contrast, residues N248 and W186 are near the PCDH1-I2 interface. The observed O-linked glycosylated residue S430 is also near the PCDH1-I2 interface. However, sugars observed at N248 and S430 did not interfere with this interface in the crystal.

Similarly, intermolecular disulfide bonds could interfere with binding thus favoring one interface over the other. Cysteine C375 in EC4 is located near the PCDH1-I2 interface (EC4:EC2), suggesting that an intermolecular disulfide bond at this location would block this interface and favor PCDH1-I1. Yet involvement of C375 in disulfide bond formation is not required for adhesion and might be irrelevant in vivo.

Next, we designed mutations predicted to break each interface as determined from an analysis of the structures. A residue was considered to be at the interface if its buried surface area was at least 20% of their accessible surface area according to PISA. Residues K398 in the PCDH1-I1 interface and V115 in the PCDH1-I2 interface met this criterion and were selected for mutagenesis. For the PCDH1-I1 interface, the K398 residue was found to form a salt bridge with E80 of the adjacent molecule. Therefore, we introduced the K398E mutation to prevent the interaction and the formation of the PCDH1-I1 interface. For the PCDH1-I2 interface, V115 was found to form hydrophobic interactions with V420, and therefore we introduced the V115R mutation to disrupt this interface (Fig. 5a, d).

Bead aggregation assays of PCDH1 EC1-7Fc K398E revealed no aggregation of beads in the presence of calcium (Fig. 5b, h, i). Consistent with these results, sedimentation velocity AUC experiments with protein fragments produced in bacteria and with mutated cysteine residues to avoid confounding effects showed monomeric states in solution for the K398E mutant (PCDH1 EC1-4 C375S K398E and PCDH1 EC1-5 C375S C548S K398E; Fig. 5c, g and Supplementary Fig. 6f, g). In contrast, binding assays of PCDH1 EC1-7Fc V115R revealed aggregation of beads in the presence of calcium and AUC experiments showed both monomeric and dimeric states in solution for this mutant (PCDH1 EC1-5 C375S C548S V115R) (Fig. 5e–i). Melting experiments of the cysteine mutated PCDH1 EC1-5 C375S C548S, EC1-5 C375S C548S K398E and EC1-5 C375S C548S V115R showed that their stability is comparable (Supplementary Fig. 7a). Similar results were obtained for PCDH1 EC1-2, EC1-2 V115R, EC3-4 C375S, and EC3-4 C375S K398E (Supplementary Fig. 7b). Altogether, these data indicate that the PCDH1-I1 interface mediates *trans* homodimerization in solution (independently of possible *cis* dimer conformations induced by C-terminal EC repeats or the Fc tag) and that it is responsible for bead aggregation in vitro.

Homophilic interaction model for PCDH1 and δ1 protocadherins. Mutagenesis coupled with binding assays and AUC experiments of the wild type and mutant protein fragments strongly suggest that PCDH1-I1 is the interface that mediates homophilic interaction. A schematic of this mode of interaction is shown in Fig. 5j. This *trans* antiparallel dimer interface suggests a mechanism of adhesion involving EC repeats 1 to 4. This mechanism of interaction is different from the tryptophan exchange mode between EC1 repeats of classical cadherins[36] and the "handshake" mechanism of interaction between PCDH15 and CDH23 involving EC1 and EC2[39]. It is similar to the "forearm handshake" mechanism of interaction of the δ2 protocadherins and of the clustered protocadherins involving EC1 to EC4[42,43,45]. Although the overall mode of homophilic adhesion might be the same for δ1, δ2, and clustered protocadherins, there are important differences. The EC1:EC4 contacts are greater in the δ1 protocadherin PCDH1, covering almost 64% of the total interface area (486 Å^2 for EC1:EC4 and 231 Å^2 for EC2:EC3). In contrast, both PCDH19 (δ2 protocadherin) and clustered protocadherins EC2:EC3 interface areas are much larger than their EC1:EC4 contacts[42,43] (Table 2).

An analysis of sequence conservation across species of interfacial residues revealed that residues at EC1:EC4 contacts are more conserved than the ones in EC2:EC3 sites (Fig. 6a, b, Supplementary Figs. 8 and 9). About 43% of the interfacial area involves hydrophobic residues, ~23% are charged, and ~34% are hydrophilic residues (Supplementary Fig. 10). The EC2:EC3 contacts are more amphiphilic with ~42% hydrophobic, ~17% charged, and ~42% hydrophilic contacts as compared to EC1:EC4 contacts which are ~43% hydrophobic, ~26% charged, and ~30% hydrophilic. All salt-bridge interactions are present in the EC1:EC4 interface. Thus, the size and charged nature of the PCDH1 EC1:EC4 interface suggest that it is important for both strength and specificity of PCDH1 homophilic adhesion.

An interesting structural aspect of the PCDH1 EC1:EC4 interface is the involvement of the δ insertion within β-strand A of EC4. This 11-residue insertion, which seems to become ordered upon complex formation, contributes four residues that directly contact EC1 of the opposite protomer (Fig. 6c). Similar

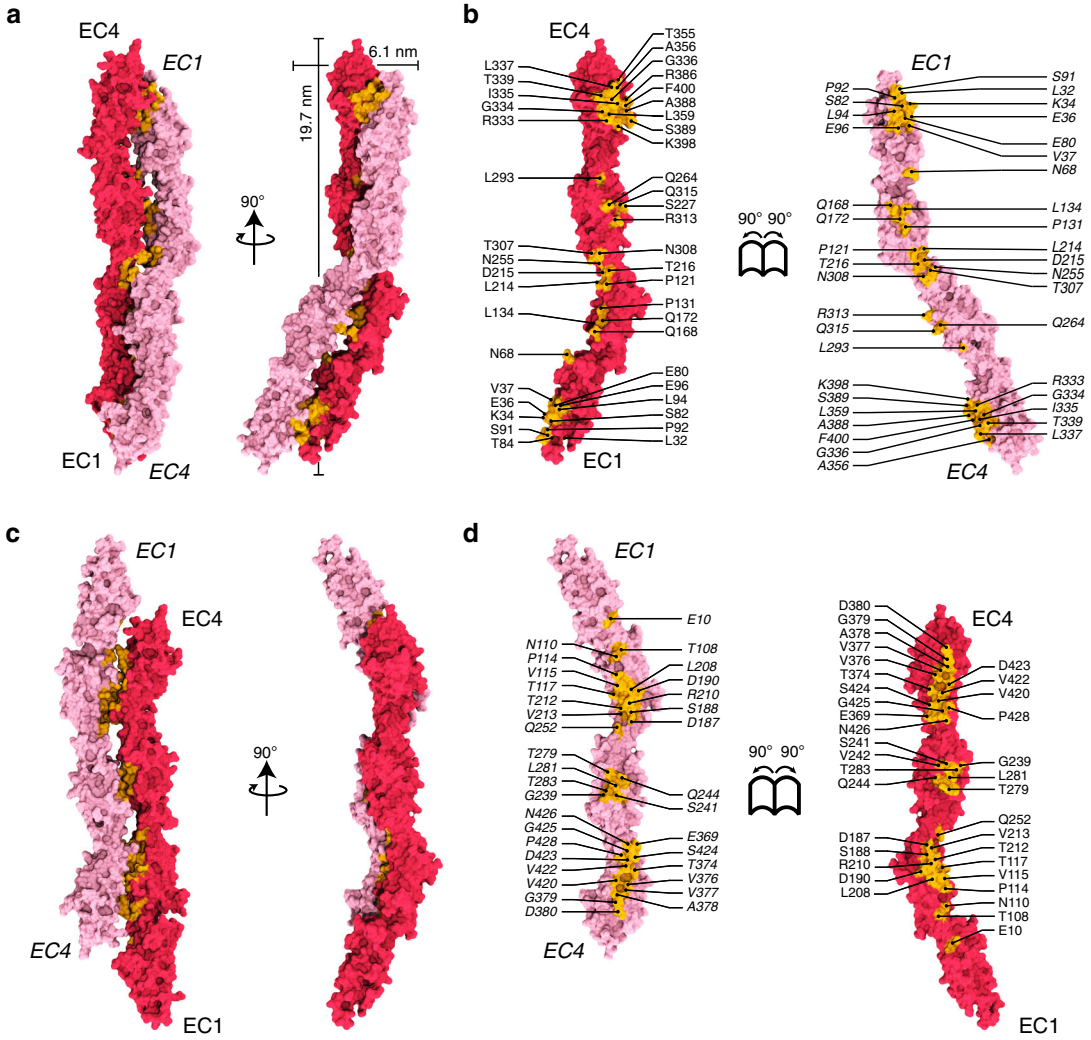

Fig. 4 Crystallographic antiparallel interfaces of PCDH1. **a** Molecular surface representation of two PCDH1 EC1-4 molecules in the crystal with the interaction interface formed by fully overlapped and antiparallel EC1-4 repeats (PCDH1-I1). Two perpendicular views are shown. **b** Interaction surface exposed with interfacing residues listed and shown in yellow. **c** Molecular surface representation of two PCDH1 EC1-4 molecules in an alternate crystallographic arrangement in which the interaction interface would be formed by fully overlapped and antiparallel EC1–5 repeats (PCDH1-I2). Two perpendicular views are shown. **d** Interaction surface exposed with interfacing residues listed and shown in yellow

Table 2 Interface areas per EC repeat for PCDH1 (δ1 protocadherin), PCDH19 (δ2 protocadherin), PCDHα7, PCDHβ6, and PCDHγB7 (clustered protocadherins)

Protocadherin	EC1:EC4 (Å^2)	EC2:EC3 (Å^2)	EC3:EC2 (Å^2)	EC4:EC1 (Å^2)
Hs PCDH1 (6BX7)	486 (34%)	231 (16%)	231 (16%)	486 (34%)
Dr PCDH19 (5IU9)	379 (24%)	446 (28%)	452 (28%)	326 (20%)
Mm PCDHα7 (5DZV)	501 (32%)	446 (29%)	381 (24%)	231 (15%)
Mm PCDHβ6 (5DZX)	531 (24%)	660 (30%)	491 (22%)	545 (24%)
Mm PCDHγB7 (5SZP)	179 (11%)	636 (37%)	474 (28%)	414 (24%)

The percentage of total interface areas that they cover is in parentheses

insertions are found in other mouse and human δ1 protocadherins and the δ2 protocadherin-8, but not in clustered protocadherins. The δ insertion, perhaps in coordination with the 10-residue long disulfide bond in EC1, may help slide and stop protomers during complex assembly, as has been proposed for the BC loop of EC2 in classical cadherins[51]. Alternatively, the δ insertion may modulate the affinity of the interaction by providing additional contacts specific to each type of δ protocadherin.

Given the sequence similarity among δ1 protocadherins (~50%) (Supplementary Figs. 11–14), the mode of interaction is likely to be the same for all other δ1 family members—PCDH7, PCDH9, and PCDH11X/Y. In support of this model, a very recent study showed biochemically that PCDH9 also uses its EC1–4 repeats to mediate homophilic adhesion[52]. A general mode of interaction in which all δ1 protocadherins, with seven EC repeats, use the same EC1–4 interface used by the δ2 and clustered protocadherins (6 EC repeats) is appealing, but rises the

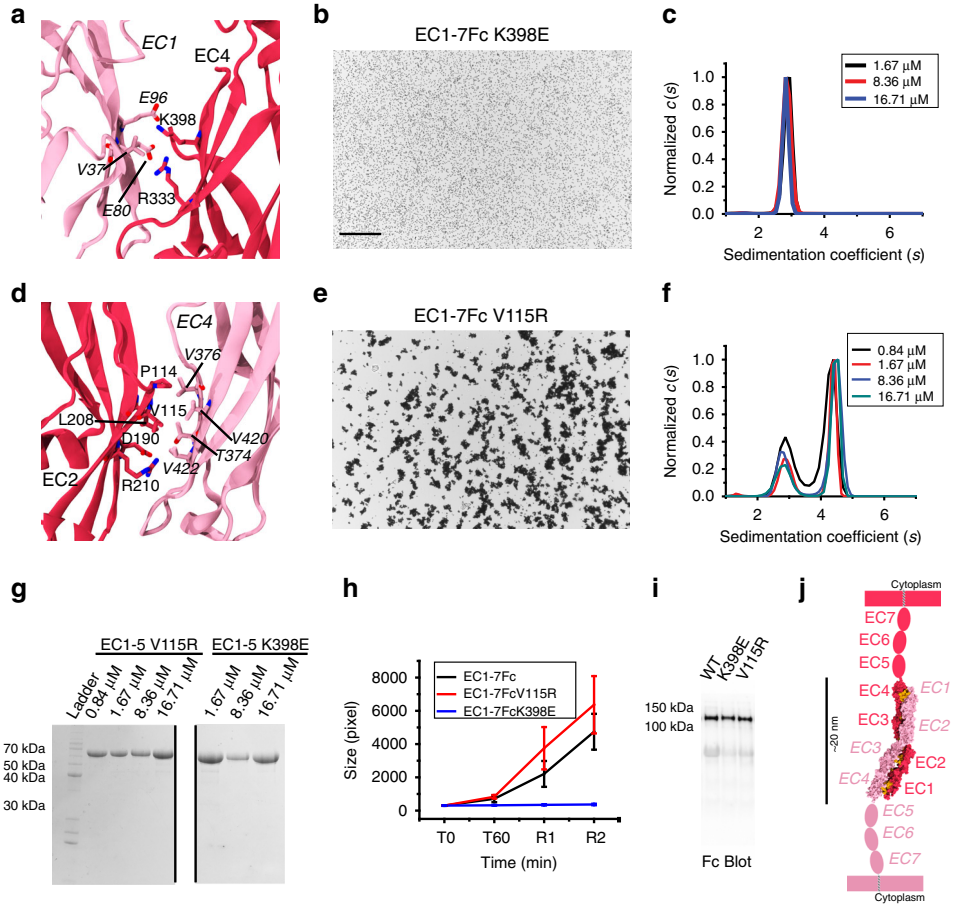

Fig. 5 Testing of crystallographic PCDH1 antiparallel interfaces. **a** Details of the PCDH1-I1 interface (EC1:EC4). Residue K398 interacting with E80 is highlighted. **b** Protein G beads coated with PCDH1 EC1-7Fc K398E mutant protein and imaged after incubation for 1 h followed by rocking for 2 min in the presence of calcium. Scale bar: 500 µm. No aggregation is observed. **c** AUC of bacterially produced PCDH1 EC1-5 C375S C548S K398E ($n = 2$ independent experiments) shows only a monomeric peak for a range of concentrations between 1.67 and 16.71 µM. **d** Details of the PCDH1-I2 interface (EC2:EC4). Residue V115 interacting with V420 is highlighted. **e** Protein G beads coated with PCDH1 EC1-7Fc V115R mutant protein and imaged after incubation for 1 h followed by rocking for 2 min in the presence of calcium. Shown as in **b**. Aggregation is observed. **f** AUC of bacterially produced PCDH1 EC1-5 C375S C548S V115R ($n = 2$ independent experiments) shows both dimeric and monomeric peaks for a range of concentrations between 0.84 and 16.71 µM. **g** SDS-PAGE shows samples for PCDH1 EC1-5 C375S C548S V115R and PCDH1 EC1-5 C375S C548S K398E after AUC. Different amounts of samples were loaded while running the gel, which is shown split for visualization purposes only. **h** Mean aggregate size for wild-type PCDH1 EC1-7Fc and mutants PCDH1 EC1-7Fc K398E, PCDH1 EC1-7Fc V115R at T0 ($t = 0$ min), after 1 h of incubation, T60 ($t = 60$ min) followed by rocking for 1 min (R1) and 2 min (R2). Error bars are standard error of the mean ($n = 4$ independent experiments for mutants, $n = 3$ independent experiments for wild type). **i** Western blot shows efficient expression of wild-type PCDH1 EC1-7Fc and mutants PCDH1 EC1-7Fc K398E and PCDH1 EC1-7Fc V115R. **j** Schematic of *trans* antiparallel binding of PCDH1 involving overlap of repeats EC1 to EC4

question of how the longer δ1 protocadherin complexes span the intercellular space when compared to the shorter δ2 and clustered protocadherins complexes. Interestingly, PCDH1 EC1-4 is more curved than the clustered protocadherins and PCDH19 EC1-4 fragments (Supplementary Fig. 15, Supplementary Movies 1 and 2). Comparison of the PCDH1-I1 complex with those formed by δ2 and clustered protocadherins also shows that the PCDH1 arrangement of protomers is more twisted, yet substantial tilting would still be required to span the same intercellular space that the other shorter members of the family can cover using the same EC1–4 binding mechanism.

Discussion

Non-classical cadherins are part of multiple and diverse structures in multicellular organisms, including adhesive bonds important for morphogenesis, sensory perception, and neuronal connectivity[11–13,36]. Molecular features of these cadherins are key determinants of their function. For instance, long non-classical cadherins such as protocadherin-24 (9 ECs), PCDH15 (11 ECs), and CDH23 (27 ECs), are part of long protein filaments essential for enterocyte brush border morphogenesis and inner ear mechanotransduction[11,53]. In contrast, the shorter clustered protocadherins, with six EC repeats, have been implicated in neuronal self-recognition and avoidance, and somewhat paradoxically have been shown to form weak and strictly homophilic *trans* adhesive bonds[13].

In the context of brain wiring, δ protocadherins are unique. First, unlike the clustered protocadherins, the cytoplasmic domains of δ1 and δ2 protocadherins have distinct sequence motifs that suggest a more diverse and versatile role in cellular signaling[19]. Second, a unique δ insertion in EC4 hints at a specialized role for this repeat in δ protocadherins. Last, the length of δ1 protocadherins' extracellular domains, with seven EC repeats, suggest a distinct set of functional roles yet to be determined. Given these differences and their differential expression in the

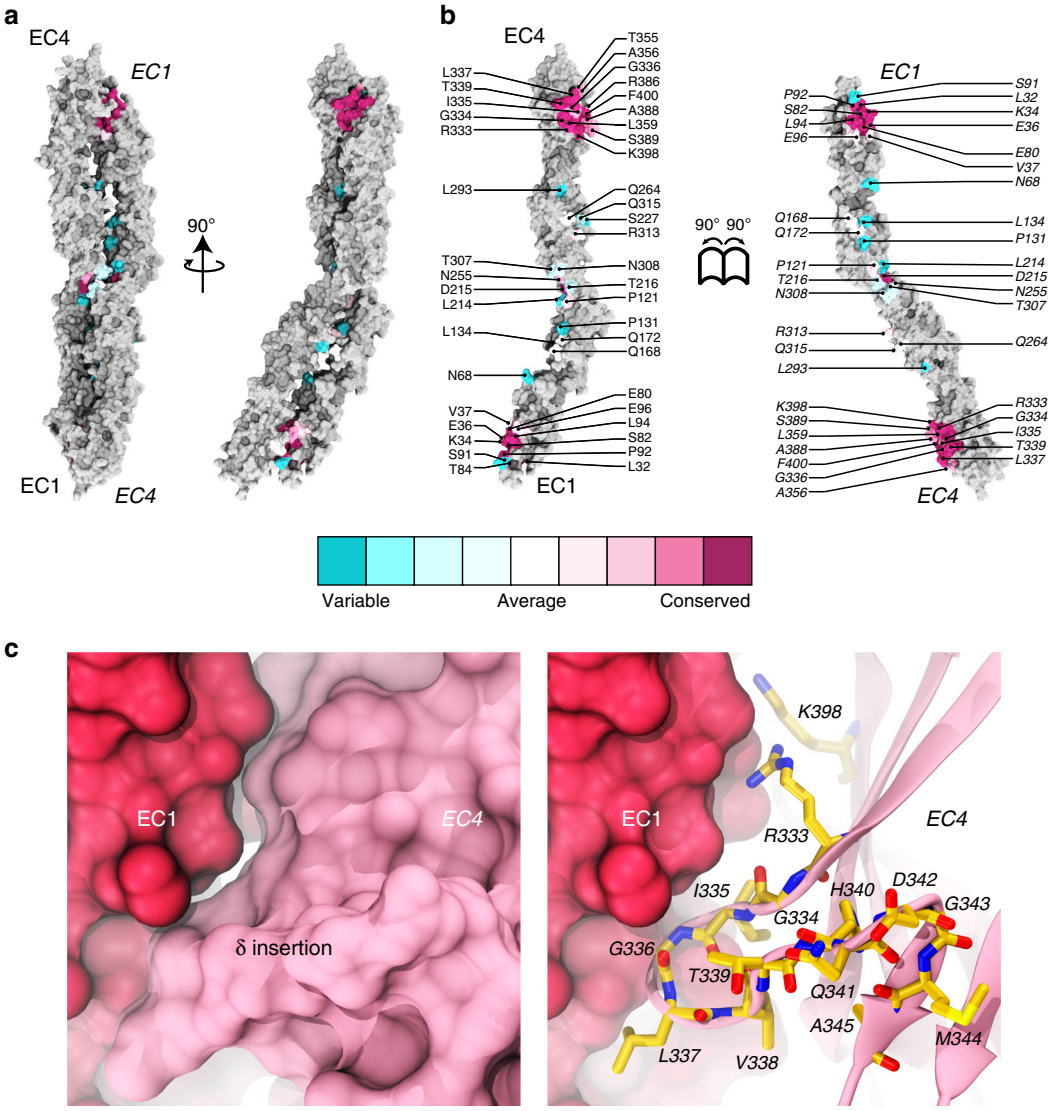

Fig. 6 Conservation of residues at the crystallographic PCDH1-I1 interface. **a** Molecular surface representation of PCDH1-I1 antiparallel dimer. Two perpendicular views are shown. **b** Interaction surface exposed with interfacing residues listed and colored according to sequence conservation among 87 species . The accession numbers for the sequences of 87 species are given in Supplementary Table 3. **c** Details of the δ insertion in EC4 shown in molecular surface representation (left) and in ribbon with relevant residues as sticks (right)

brain[20,26,27,54,55], an important question about δ protocadherins' function pertains to whether they are involved in neuronal adhesion and circuitry specificity or rather mediate neuronal self-recognition and avoidance as proposed for the clustered protocadherins[13,56].

Functional studies of PCDH19 indicate that this δ2 protocadherin is required for the formation of neuronal columns in the fish optic tectum, thus implying an adhesive role[57]. Structural studies, validated with analyses of epilepsy-causing mutations, indicate that PCDH19 uses an extensive but weak antiparallel EC1–4 adhesive interface ideally suited to encode specificity[42]. Clustered protocadherins use the same antiparallel EC1–4 interface[40,43–45], and sequence and evolutionary analyses suggest that at least some subgroups of clustered protocadherins use their EC1:EC4 contacts to modulate the bond affinity and their EC2:EC3 interactions to determine specificity[45].

Interestingly, our structures of PCDH1 and associated binding assays indicate that this δ1 protocadherin uses an antiparallel EC1–4 interface that is similar to that used by clustered and δ2 protocadherins, yet interactions at the EC2:EC3 contacts are minimal, and salt-bridges possibly encoding for specificity are at the much larger EC1:EC4 interface (Supplementary Movies 1 and 2). Binding assays of PCDH9, another δ1 protocadherin, suggest that this protein also uses an antiparallel EC1–4 interface, with the EC1:EC4 contact playing a more important role than EC2:EC3 interactions in bead aggregation[52]. Taken together, these results indicate that δ1 protocadherins use a unique and enlarged EC1:EC4 contact both to strengthen the adhesive bond and to determine its specificity. How the bond assembles, perhaps through induced-fit changes in protomer curvature (Supplementary Fig. 15c), remains unexplored, and whether δ1 protocadherins use this bond to trigger a repulsive signal mediating neuronal self-recognition, to form true adhesive and specific intercellular contacts, or both depending on molecular and cellular context, is still unclear.

The versatility of PCDH1 in mediating adhesion and signaling might be enhanced by the existence of various isoforms that modify its extracellular and cytoplasmic domains[58]. The possible

existence of a PCDH1 isoform with a shorter extracellular domain starting at EC4 (NM_001278615.1/NP_001265544.1) indicates that a pseudo-heterophilic bond could be formed and be mediated by the EC1:EC4 interface alone. Surface expression levels of this isoform could modulate the strength of the adhesion between two cells expressing PCDH1.

The discovery of an antiparallel EC1–4 interface for PCDH1 and δ1 protocadherins with seven ECs poses an interesting set of questions regarding the arrangement of this adhesive complex in vivo. The intercellular space can vary wildly among different types of cellular junctions. For instance, occludin, claudins, and JAM proteins mediate very close contacts in tight junctions, while the intercellular space occupied by classical cadherins (five ECs interacting through EC1s in a curved and tilted arrangement) is ~25 nm[47], and by clustered and δ2 protocadherins (six ECs with an EC1–4 overlap) is <40 nm[42,44,45]. A straight δ1 protocadherin complex that uses the EC1–4 antiparallel interface and that has seven ECs per protomer (<50 nm) may not fit. Tilting of the complex, as observed for PCDH15 (ref. [59]), might solve this conundrum. It is also possible that the longer δ1 protocadherins are used to explore and form initial transient cell–cell contacts while other shorter cadherins and protocadherins are recruited to enable more permanent adhesion. On the contrary, a longer and straight δ1 complex might prevent adhesion mediated by shorter and tilted classical complexes.

The similarities in the overall mode of adhesion for δ1, δ2, and clustered protocadherins suggest that they may have a common ancestry[60]. It is thus tempting to speculate that δ protocadherins may engage in *cis* interactions as the clustered protocadherins do using their EC5 and EC6 repeats[41,61]. Whether EC7 in δ1 protocadherins is also involved in this type of contacts is unknown. Interestingly, PCDH19, a δ2 protocadherin, interacts with the classical N-cadherin[42,62,63]. Perhaps PCDH1 interacts with E-cadherin in a similar fashion, as they co-localize in airway epithelial cells. Since the protocadherins are known to mediate weak interactions, *cis* dimerization with classical members of the family may help in forming more robust adhesion complexes.

PCDH1 is also expressed in several non-neuronal tissues, including the cochlea, skin keratinocytes, lungs, and airway epithelia, all of which are subjected to mechanical stimuli[25,28,31,64–66]. The twisted adhesive bond of PCDH1 may provide some elasticity to these tissues, and its length might allow for additional stretching of cell–cell contacts without complete rupture of the adhesive contact. In addition, PCDH1 is known to help in forming the airway epithelial barrier by mediating adhesion between epithelial cells, the dysfunction of which leads to pathogenesis[31]. PCDH1 expression increases during differentiation of airway and bronchial epithelial cells, and it has been localized to cell–cell contact sites along with E-cadherin[31]. Our structural model of PCDH1 homophilic adhesion partially reveals how the airway epithelial barrier is formed, and suggests mutation sites that can be used to disrupt PCDH1 adhesion, which should help to elucidate the roles played by PCDH1's extracellular and cytoplasmic domains in vivo.

The structure of PCDH1 EC1–4 and the homophilic binding mechanism presented here also have important implications in relation to New World Hantavirus infectivity. PCDH1 EC1 has been recently shown to be critical for cellular attachment and entry of these viruses[32], yet molecular details are still unclear. It is possible that the hantavirus glycoproteins interact with PCDH1 EC1 and disrupt homophilic binding by preventing contacts with the partnering PCDH1 EC4. In this scenario, PCDH1 EC4 could be used and engineered to block interactions with the virus. On the other hand, native PCDH1 EC1 interactions with hantavirus glycoproteins may involve residues not engaged in homophilic binding, in which case PCDH1 EC1 itself could be used to develop hantavirus blockers. In either case, the molecular details provided by our PCDH1 EC1-4 structure may guide drug discovery strategies aimed at reducing or blocking hantavirus infections.

Methods

Cloning and mutagenesis. Sequences enconding for human PCDH1 repeats EC1–2, EC3–4, EC1–4, and EC1–5 were subcloned using the *Nde*I and *Xho*I sites of the pET21a vector for bacterial expression (without signal peptide), and using the *Xho*I and *Bam*HI sites of the CMV:N1-Fc and CMV:N1-His vector[42] for mammalian expression. Constructs encoding for truncated versions of PCDH1 (PCDH1EC1-7, PCDH1EC1-6, PCDH1EC1-5, PCDH1EC1-4, PCDH1EC1-3, PCDH1EC1-2, PCDH1EC2-7) were created by PCR subcloning of the native signal peptide and appropriate EC domains into CMV:N1-Fc. Truncation boundaries for EC repeats are as in Supplementary Fig. 4. Mutations were created in both the bacterial and the mammalian expression constructs by site-directed mutagenesis using the Quikchange Lightning Site-directed Mutagenesis Kit (Agilent). All constructs were sequence verified.

Bacterial expression, purification, and refolding. PCDH1 fragments were expressed in BL21 CodonPlus (DE3)-RIPL cells (Stratagene) that were cultured in LB (EC1–2, EC3–4, EC1–4, and EC1–5 WT and mutants), induced at OD_{600} ~ 0.6 with 200 μM IPTG, and grown at 30 °C for ~18 h. Cells were lysed by sonication in denaturing buffer (20 mM Tris-HCl [pH 7.5], 6 M guanidine hydrochloride, 10 mM $CaCl_2$, and 20 mM imidazole). The cleared lysates were loaded onto Ni-Sepharose (GE Healthcare) and eluted with denaturing buffer supplemented with 500 mM imidazole. PCDH1 fragments for crystallography were refolded by overnight dialysis against 20 mM TrisHCl [pH 8], 150 mM KCl, 50 mM NaCl, 400 mM arginine, 2 mM $CaCl_2$, and 2 mM DTT using MWCO 2000 membranes at 4 °C. Refolded proteins were further purified by size exclusion chromatography on a Superdex200 column (GE Healthcare) in 20 mM Tris-HCl [pH 8.0], 150 mM KCl, 50 mM NaCl, 2 mM $CaCl_2$, and 2 mM DTT. Mutated PCDH1 fragments used for AUC experiments and differential scanning calorimetry were refolded and purified using the same buffers without DTT.

Mammalian expression and purification. The PCDH1 EC1-4His fragment was expressed in Expi293™ cells (ThermoFischer Scientific A14527). Transfection was done with 30 μg of plasmid and ExpiFectamine™ 293. Media was collected on the fourth day after transfection by pelleting cells and collecting the supernatant to which 10 mM $CaCl_2$ was added. The mix was dialyzed overnight against 20 mM TrisHCl [pH 8.0], 150 mM KCl, 50 mM NaCl, and 10 mM $CaCl_2$ using MWCO 2000 membranes. The PCDH1 EC1-4His fragment was then purified by nickel affinity chromatography under native conditions with buffer containing 20 mM Tris-HCl [pH 8.0], 300 mM NaCl, 10 mM $CaCl_2$, and 20 mM imidazole. The protein was eluted with the same buffer containing 500 mM imidazole. It was further purified by size exclusion chromatography on a Superdex200 column (GE Healthcare) in 20 mM Tris-HCl [pH 8.0], 150 mM KCl, 50 mM NaCl, and 2 mM $CaCl_2$.

Analytical ultracentrifugation. Sedimentation velocity experiments were performed in a ProteomeLab™XL-I analytical ultracentrifuge (Beckman Coulter). Samples were prepared by concentrating purified proteins from size exclusion chromatography using Sartorious Vivaspin 20 concentrators (MWCO 10 kDa). Protein concentrations were measured using a Nanodrop 2000c spectrophotometer (Thermo Scientific). Buffer containing 20 mM Tris-HCl [pH 8.0], 150 mM KCl, 50 mM NaCl, and 2 mM $CaCl_2$ was used as blank. All samples were run at 50,000 r.p.m. at 20 °C and detection was by UV at 230 nm for samples with concentration <8 μM and at 280 nm for others. Solvent density was assumed to be 1 g cc^{-1} and the partial specific volume of solute ($\bar{v}$) was set to 0.73 cc g^{-1}. Data were analyzed using SEDFIT[67,68]. Experiments were repeated twice ($n = 2$) with distinct samples.

Crystallization, data collection, and structure determination. Crystals of PCDH1 EC1-4 fragments from bacterial and mammalian expression systems were grown by vapor diffusion at 4 °C by mixing equal volumes of protein (~6 mg ml^{-1}) and reservoir solution (0.1 M NaHEPES [pH 7.5], 0.2 M $CaCl_2$, 15% PEG 400, 15% Glycerol for PCDH1 EC1-4bc; 0.05 M TrisHCl [pH 8.5], 0.1 M $MgCl_2$, 30% PEG550 MME for PCDH1 EC1-4mc). Crystals of PCDH1 EC3-4 fragments from bacterial expression systems were also grown by vapor diffusion at 4 °C by mixing equal volumes of protein (~8.5 mg ml^{-1}) and reservoir solution of 0.1 M MES [pH 6.5], 20% PEG 4000, and 0.6 M NaCl (cryo-protected with 25% glycerol). Crystals were cryo-cooled in liquid N_2. X-ray diffraction data was collected at the Advanced Photon Source beam line 24-ID-E (Argonne National Laboratory) and processed with HKL2000 (ref. [69]). X-ray wavelength was set to 0.97921 Å and diffraction experiments were carried out at 100 K. The PCDH1 EC3-4bc structure was determined by molecular replacement using a homology model based on PCDHα7 (5DZV[43]) as an initial search model in PHASER[70] (Table 1). Model building was done with COOT[71] and restrained TLS refinement was performed with REFMAC5 (ref. [72]). The PCDH1 EC1-4bc structure was determined through molecular

replacement using the PCDH1 EC3-4bc structure. Likewise, the PCDH1 EC1-4mc structure was determined through molecular replacement using the PCDH1 EC1-4bc structure as the initial search model in PHASER. Data collection and refinement statistics are provided in Table 1.

Inspection of 2Fo–Fc electron density maps during initial stages of refinement of the PCDH1 EC1-4bc and PCDH1 EC1-4mc model revealed unassigned densities at the linker regions. Calcium was present at concentrations of at least 2 mM in the protein solution buffers, so we assigned calcium ions in the PCDH1 EC1-4bc model. However, the crystallization buffer for PCDH1 EC1-4mc had 0.1 M $MgCl_2$. We tried to assign both calcium as well as magnesium ions in the linker regions of the PCDH1 EC1-4mc model, but only the assignment of calcium ions was compatible with the final 2Fo−Fc map and was supported by an analysis of distances to coordinating atoms and B factor values of the ion and surrounding residue atoms. All molecular images were generated with VMD[73].

Bead aggregation assays. Bead aggregation assays were modified from those described previously[42,63,74,75]. The PCDH1Fc fusion constructs were transfected into HEK293T cells (ATCC® CRL-1573™) using calcium-phosphate transfection[76–78]. Briefly, solution A (10 µg of plasmid DNA and 250 mM $CaCl_2$) was added drop-wise to solution B (2× HBS) while mildly vortexing, and the final transfection solution was added drop-wise to one 100-mm dish of cultured HEK293T cells (two plates per assay). The next day, cells were rinsed twice with 1× PBS and serum-free media. Cells were allowed to grow in the serum-free media for 2 days before collecting the media containing the secreted Fc fusions. The media was concentrated (Amicon 10 kDa concentrators) and incubated with 1.5 µl of protein G Dynabeads (Invitrogen) while rotating at 4 °C for 2 h. The beads were washed in binding buffer (50 mM TrisHCl [pH 7.5], 100 mM NaCl, 10 mM KCl, and 0.2% BSA) and split into two tubes with either 2 mM EDTA or 2 mM $CaCl_2$. Beads were allowed to aggregate in a glass depression slide in a humidified chamber for 60 min without motion, followed by 1 min and then 2 min of rocking (8 oscillations per min). Images were collected upon adding EDTA or $CaCl_2$, after 60 min incubation, and after each rocking interval using a microscope (Nikon Eclipse Ti) with a 10× objective. Bead aggregates were quantified using ImageJ software as described previously[63,74]. Briefly, the images were thresholded, the area of the detected aggregate particles was measured in units of pixels, and the average size was calculated. Assays were repeated four times (except for PCDH1 EC1-7Fc and PCDH1 EC1-7Fc C375S C548S with $n = 3$) from distinct protein samples and their mean aggregate size (±SEM) at each time point was plotted.

Western blots were performed using part of the beads containing the Fc fusion proteins after the bead aggregation assay to confirm expression and secretion of the protein. The protein was eluted with sample loading dye, boiled for 5 min, and loaded onto SDS-PAGE gels (BioRad) for electrophoresis. Proteins were transferred to PVDF membrane (GE Healthcare) and blocked with 5% nonfat milk in TBS with 0.1% Tween-20 before incubating overnight at 4 °C with goat anti-human IgG (1:200 Jackson ImmunoResearch Laboratories, Catalog-109-025-003, Lot-117103). After several washes with 1× TBS, the blot was incubated with mouse anti-goat HRP-conjugated secondary (1:5000; Santa Cruz Biotechnology, Catalog-sc-2354, Lot-A2017) for 1 h at room temperature, washed with 1× TBS and developed with ECL Select Western Blot Detection (GE Healthcare) for chemiluminescent detection (Omega Lum G).

Differential scanning fluorimetry. The bacterially expressed PCDH1 EC1-5 C375S C548S, EC1-5 C375S C548S K398E, EC1-5 C375S C548S V115R, EC1-2, EC1-2 V115R, EC3-4 C375S, and EC3-4 C375S K398E fragments were purified as described above and used for differential scanning fluorimetry (DSF)[79–81]. Experiments were done in triplicates and repeated three independent times for PCDH1 EC1-5 C375S C548S, EC1-5 C375S C548S K398E, and EC1-5 C375S C548S V115R using distinct protein samples at 0.2 mg ml^{-1} and two independent times for EC1-2, EC1-2 V115R, EC3-4 C375S, and EC3-4 C375S K398E using distinct protein samples at 0.5 mg ml^{-1}. All protein samples were mixed in buffer (20 mM TrisHCl [pH 8.0], 150 mM KCl, 50 mM NaCl, and 2 mM $CaCl_2$) with SYPRO Orange dye (20×, Invitrogen). Fluorescent measurements were performed in a CFXConnect RT-PCR instrument (BioRad) using hard-shell 96-well thin wall PCR plates (BioRad) while samples were heated from 10° to 95° in 0.2° steps. Melting temperatures were defined as those recorded when the normalized fluorescence reached 0.5.

Sequence analysis. For analysis of PCDH1 residue conservation across species, 87 sequences were obtained from the NCBI protein database and processed manually to include the extracellular domain from EC repeats 1 to 7. The accession numbers for sequences of each of the 87 species are listed in Supplementary Table 3. These PCDH1 sequences were aligned using Clustal X2 (ref. [82]) and the alignment file was put into ConSurf[83] to calculate relative conservation of each residue and to categorize the degree of conservation into nine color bins. Sequences for δ1 protocadherins *hs* PCDH1, *mm* PCDH1, *hs* PCDH7, *mm* PCDH7, *hs* PCDH9, *mm* PCDH9, *hs* PCDH11X, and *mm* PCDH11X were obtained from the NCBI protein database, processed manually to include only the extracellular domain through the end of EC7, and aligned using the Clustal Omega server[84]. Identity and similarity were calculated using the Sequence Identity And Similarity SIAS server[85]. The accession numbers of the sequences are listed in Supplementary Table 4.

Prediction of glycosylation and glycation sites. Potential PCDH1 glycosylation sites were predicted using the following servers: NetNGlyc 1.0 (N-glycosylation, GlcNAc-β-Asn), NetOGlyc 4.0 (O-glycosylation, GalNAC-α-Ser/Thr)[86,87], and NetCGlyc 1.0 (C-glycosylation, Man-α-Trp)[88].

Orientation of EC repeats and azimuthal angles calculation. Relative orientation of EC repeats was determined using *mm* CDH23 (2WHV) as a reference. Repeat EC1 of *mm* CDH23 was aligned to the *z*-axis in VMD. Then, the N-terminal EC repeat for a pair of ECs of interest was aligned to *mm* CDH23 EC1 using COOT. The principal axis of the C-terminal EC repeat was calculated using the Orient plugin in VMD. The projection of the principal axis on the *xy*-plane was plotted to identify the azimuthal angle φ. Repeat CDH23 EC2 was used to define $\varphi = 0°$.

Statistics and data reproducibility. Details of all experiments and protocols are described in the Methods section for reproducibility. Bead aggregation assays were done at least three times with independent samples. AUC experiments were repeated twice with independent samples. DSF experiments were done at least twice with independent samples.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data Availability

Coordinates for *hs* PCDH1 EC3-4bc, *hs* PCDH1 EC1-4bc, and *hs* PCDH1 EC1-4mc have been deposited in the Protein Data Bank with entry codes 6PIM, 6BX7, and 6MGA, respectively. Other raw data and data sets generated and/or analyzed during the current study are available from the corresponding author on reasonable request. Figures 1, 2, and 5 have associated source data files available in Supplementary Data 1 and 2. Supplementary Fig. 7 has an associated source file in Supplementary Data 3.

Received: 10 May 2019 Accepted: 21 August 2019
Published online: 30 September 2019

References

1. Steinberg, M. S. Adhesion in development: an historical overview. *Dev. Biol.* **180**, 377–388 (1996).
2. Shapiro, L., Love, J. & Colman, D. R. Adhesion molecules in the nervous system: structural insights into function and diversity. *Annu. Rev. Neurosci.* **30**, 451–474 (2007).
3. Hynes, R. O. Integrins: bidirectional, allosteric signaling machines. *Cell* **110**, 673–687 (2002).
4. Tucker, R. P. & Chiquet-Ehrismann, R. Teneurins: a conserved family of transmembrane proteins involved in intercellular signaling during development. *Dev. Biol.* **290**, 237–245 (2006).
5. Takeichi, M. Morphogenetic roles of classic cadherins. *Curr. Opin. Cell Biol.* **7**, 619–627 (1995).
6. Wheelock, M. J. & Johnson, K. R. Cadherin-mediated cellular signaling. *Curr. Opin. Cell Biol.* **15**, 509–514 (2003).
7. Suzuki, S. T. Structural and functional diversity of cadherin superfamily: are new members of cadherin superfamily involved in signal transduction pathway? *J. Cell. Biochem.* **61**, 531–542 (1996).
8. Sotomayor, M., Gaudet, R. & Corey, D. P. Sorting out a promiscuous superfamily: towards cadherin connectomics. *Trends Cell Biol.* **24**, 524–536 (2014).
9. Gillespie, P. G. & Müller, U. Mechanotransduction by hair cells: models, molecules, and mechanisms. *Cell* **139**, 33–44 (2009).
10. Pruitt, B. L., Dunn, A. R., Weis, W. I. & Nelson, W. J. Mechano-transduction: from molecules to tissues. *PLoS Biol.* **12**, e1001996 (2014).
11. Jaiganesh, A., Narui, Y., Araya-Secchi, R. & Sotomayor, M. Beyond cell-cell adhesion: sensational cadherins for hearing and balance. *Cold Spring Harb. Perspect. Biol.* a029280, https://doi.org/10.1101/cshperspect.a029280 (2017)
12. Hirano, S. & Takeichi, M. Cadherins in brain morphogenesis and wiring. *Physiol. Rev.* **92**, 597–634 (2012).
13. Mountoufaris, G., Canzio, D., Nwakeze, C. L., Chen, W. V. & Maniatis, T. Writing, reading, and translating the clustered protocadherin cell surface recognition code for neural circuit assembly. *Annu. Rev. Cell Dev. Biol.* **34**, 471–493 (2018).
14. Lefebvre, J. L., Kostadinov, D., Chen, W. V., Maniatis, T. & Sanes, J. R. Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature* **488**, 517–521 (2012).

15. Schreiner, D. & Weiner, J. A. Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion. *Proc. Natl. Acad. Sci. USA* **107**, 14893–14898 (2010).
16. Weiner, J. A. & Jontes, J. D. Protocadherins, not prototypical: a complex tale of their interactions, expression, and functions. *Front. Mol. Neurosci.* **6**, 4 (2013).
17. Shapiro, L. et al. Structural basis of cell-cell adhesion by cadherins. *Nature* **374**, 327–337 (1995).
18. Nagar, B., Overduin, M., Ikura, M. & Rini, J. M. Structural basis of calcium-induced E-cadherin rigidification and dimerization. *Nature* **380**, 360–364 (1996).
19. Redies, C., Vanhalst, K. & Roy, F. delta-Protocadherins: unique structures and functions. *Cell. Mol. Life Sci.* **62**, 2840–2852 (2005).
20. Kim, S.-Y., Yasuda, S., Tanaka, H., Yamagata, K. & Kim, H. Non-clustered protocadherin. *Cell Adh. Migr.* **5**, 97–105 (2011).
21. Redies, C., Hertel, N. & Hübner, C. A. Cadherins and neuropsychiatric disorders. *Brain Res.* **1470**, 130–144 (2012).
22. Hirabayashi, T. & Yagi, T. Protocadherins in neurological diseases. *Adv. Neurobiol.* **8**, 293–314 (2014).
23. Sano, K. et al. Protocadherins: a large family of cadherin-related molecules in central nervous system. *EMBO J.* **12**, 2249–2256 (1993).
24. Bononi, J., Cole, A., Tewson, P., Schumacher, A. & Bradley, R. Chicken protocadherin-1 functions to localize neural crest cells to the dorsal root ganglia during PNS formation. *Mech. Dev.* **125**, 1033–1047 (2008).
25. Koning, H. et al. Characterization of protocadherin-1 expression in primary bronchial epithelial cells: association with epithelial cell differentiation. *FASEB J.* **26**, 439–448 (2012).
26. Blevins, C. J., Emond, M. R., Biswas, S. & Jontes, J. D. Differential expression, alternative splicing, and adhesive properties of the zebrafish δ1-protocadherins. *Neuroscience* **199**, 523–534 (2011).
27. Redies, C., Heyder, J., Kohoutek, T., Staes, K. & Van Roy, F. Expression of protocadherin-1 (Pcdh1) during mouse development. *Dev. Dyn.* **237**, 2496–2505 (2008).
28. Mortensen, L. J., Kreiner-Møller, E., Hakonarson, H., Bønnelykke, K. & Bisgaard, H. The PCDH1 gene and asthma in early childhood. *Eur. Respir. J.* **43**, 792–800 (2014).
29. Blumenthal, J. B. & Blumenthal, M. N. Genetics of asthma. *Med. Clin. North Am.* **86**, 937–950 (2002).
30. Bateman, E. D. et al. Global strategy for asthma management and prevention: GINA executive summary. *Eur. Respir. J.* **31**, 143–178 (2008).
31. Kozu, Y. et al. Protocadherin-1 is a glucocorticoid-responsive critical regulator of airway epithelial barrier function. *BMC Pulm. Med.* **1**, 1–12 (2015).
32. Jangra, R. K. et al. Protocadherin-1 is essential for cell entry by New World hantaviruses. *Nature* **563**, 559–563 (2018).
33. Zaki, S. R. et al. Hantavirus pulmonary syndrome. *Pathog. Emerg. Infect. Dis. Am. J. Pathol.* **146**, 552–579 (1995).
34. Macneil, A., Nichol, S. T. & Spiropoulou, C. F. Hantavirus pulmonary syndrome. *Virus Res.* **162**, 138–147 (2011).
35. Garrido, J. L. et al. Two recombinant human monoclonal antibodies that protect against lethal Andes hantavirus infection in vivo. *Sci. Transl. Med.* **10**, eaat6420 (2018).
36. Brasch, J., Harrison, O. J., Honig, B. & Shapiro, L. Thinking outside the cell: how cadherins drive adhesion. *Trends Cell Biol.* **22**, 299–310 (2012).
37. Brasch, J. et al. Homophilic and heterophilic interactions of type ii cadherins identify specificity groups underlying cell-adhesive behavior. *Cell Rep.* **23**, 1840–1852 (2018).
38. Vendome, J. et al. Molecular design principles underlying β-strand swapping in the adhesive dimerization of cadherins. *Nat. Struct. Mol. Biol.* **18**, 693–700 (2011).
39. Sotomayor, M., Weihofen, W. A., Gaudet, R. & Corey, D. P. Structure of a force-conveying cadherin bond essential for inner-ear mechanotransduction. *Nature* **492**, 128–132 (2012).
40. Nicoludis, J. M. et al. Structure and sequence analyses of clustered protocadherins reveal antiparallel interactions that mediate homophilic specificity. *Structure* **23**, 2087–2098 (2015).
41. Rubinstein, R. et al. Molecular logic of neuronal self-recognition through protocadherin domain interactions. *Cell* **163**, 629–642 (2015).
42. Cooper, S. R., Jontes, J. D. & Sotomayor, M. Structural determinants of adhesion by Protocadherin-19 and implications for its role in epilepsy. *Elife* 5, https://doi.org/10.7554/eLife.18529 (2016).
43. Goodman, K. M. et al. Structural basis of diverse homophilic recognition by clustered α- and β-protocadherins. *Neuron*, https://doi.org/10.1016/j.neuron.2016.04.004 (2016)
44. Goodman, K. M. et al. γ-Protocadherin structural diversity and functional implications. *Elife* 5, https://doi.org/10.7554/eLife.20930 (2016).
45. Nicoludis, J. M. et al. Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1-4. *Elife* 5, https://doi.org/10.7554/eLife.18449 (2016).
46. Ciatto, C. et al. T-cadherin structures reveal a novel adhesive binding mechanism. *Nat. Struct. Mol. Biol.* **17**, 339–347 (2010).
47. Harrison, O. J. et al. The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. *Structure* **19**, 244–256 (2011).
48. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
49. Ponstingl, H., Henrick, K. & Thornton, J. M. Discriminating between homodimeric and monomeric proteins in the crystalline state. *Proteins* **41**, 47–57 (2000).
50. Brasch, J. et al. Structure and binding mechanism of vascular endothelial cadherin: a divergent classical cadherin. *J. Mol. Biol.* **408**, 57–73 (2011).
51. Dalle Vedove, A., Lucarelli, A. P., Nardone, V., Matino, A. & Parisini, E. The X-ray structure of human P-cadherin EC1-EC2 in a closed conformation provides insight into the type I cadherin dimerization pathway. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **71**, 371–380 (2015).
52. Peng, X. et al. Affinity capture of polyribosomes followed by RNAseq (ACAPseq), a discovery platform for protein-protein interactions. *Elife* **7**, https://doi.org/10.7554/eLife.40982 (2018).
53. Crawley, S. W., Mooseker, M. S. & Tyska, M. J. Shaping the intestinal brush border. *J. Cell Biol.* **207**, 441–451 (2014).
54. Kim, S.-Y., Chung, H. S., Sun, W. & Kim, H. Spatiotemporal expression pattern of non-clustered protocadherin family members in the developing rat brain. *Neuroscience* **147**, 996–1021 (2007).
55. Vanhalst, K., Kools, P., Staes, K., van Roy, F. & Redies, C. delta-Protocadherins: a gene family expressed differentially in the mouse brain. *Cell. Mol. Life Sci.* **62**, 1247–1259 (2005).
56. Jontes, J. D. The Cadherin superfamily in neural circuit assembly. *Cold Spring Harb. Perspect. Biol.* **10**, a029306 (2018).
57. Cooper, S. R. et al. Protocadherins control the modular assembly of neuronal columns in the zebrafish optic tectum. *J. Cell Biol.* **211**, 807–814 (2015).
58. Koning, H. et al. Mouse protocadherin-1 gene expression is regulated by cigarette smoke exposure in vivo. *PLoS One* **9**, e98197 (2014).
59. Ge, J. et al. Structure of mouse protocadherin 15 of the stereocilia tip link in complex with LHFPL5. *Elife* 7, https://doi.org/10.7554/eLife.38770 (2018).
60. Oda, H. & Takeichi, M. Evolution: structural and functional diversity of cadherin at the adherens junction. *J. Cell Biol.* **193**, 1137–1146 (2011).
61. Goodman, K. M. et al. Protocadherin *cis* -dimer architecture and recognition unit diversity. *Proc. Natl. Acad. Sci. USA* **114**, E9829–E9837 (2017).
62. Biswas, S., Emond, M. R. & Jontes, J. D. Protocadherin-19 and N-cadherin interact to control cell movements during anterior neurulation. *J. Cell Biol.* **191**, 1029–1041 (2010).
63. Emond, M. R., Biswas, S., Blevins, C. J. & Jontes, J. D. A complex of Protocadherin-19 and N-cadherin mediates a novel mechanism of cell adhesion. *J. Cell Biol.* **195**, 1115–1121 (2011).
64. Koppelman, G. H. et al. Identification of PCDH1 as a novel susceptibility gene for bronchial hyperresponsiveness. *Am. J. Respir. Crit. Care Med.* **180**, 929–935 (2009).
65. Faura Tellez, G. et al. Protocadherin-1 localization and cell-adhesion function in airway epithelial cells in asthma. *PLoS One* **11**, e0163967 (2016).
66. Lin, J. et al. Anatomical expression patterns of delta-protocadherins in developing chicken cochlea. *J. Anat.* **221**, 598–608 (2012).
67. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).
68. Schuck, P. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal. Biochem.* **320**, 104–124 (2003).
69. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution-from diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 859–866 (2006).
70. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
71. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
72. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 355–367 (2011).
73. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
74. Emond, M. R. & Jontes, J. D. Bead aggregation assays for the characterization of putative cell adhesion molecules. *J. Vis. Exp.* e51762. https://doi.org/10.3791/51762 (2014)
75. Sivasankar, S., Zhang, Y., Nelson, W. J. & Chu, S. Characterizing the initial encounter complex in cadherin adhesion. *Structure* **17**, 1075–1081 (2009).
76. Kwon, M. & Firestein, B. L. DNA transfection: calcium phosphate method. *Methods Mol. Biol.* **1018**, 107–110 (2013).

77. Barry, J. et al. Ankyrin-G directly binds to kinesin-1 to transport voltage-gated Na+ channels into axons. *Dev. Cell* **28**, 117–131 (2014).
78. Jiang, M. & Chen, G. High Ca2+-phosphate transfection efficiency in low-density neuronal cultures. *Nat. Protoc.* **1**, 695–700 (2006).
79. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221 (2007).
80. Lavinder, J. J., Hari, S. B., Sullivan, B. J. & Magliery, T. J. High-throughput thermal scanning: a general, rapid dye-binding thermal shift screen for protein engineering. *J. Am. Chem. Soc.* **131**, 3794–3795 (2009).
81. Choudhary, D., Kumar, A., Magliery, T. J. & Sotomayor, M. Using thermal scanning assays to test protein-protein interactions of inner-ear cadherins. *PLoS One* **12**, e0189546 (2017).
82. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
83. Ashkenazy, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* https://doi.org/10.1093/nar/gkw408 (2016).
84. Sievers, F. & Higgins, D. G. Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* **1079**, 105–116 (2014).
85. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
86. Hansen, J. E. et al. NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconj. J.* **15**, 115–130 (1998).
87. Steentoft, C. et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* **32**, 1478–1488 (2013).
88. Julenius, K. NetCGlyc 1.0: prediction of mammalian C-mannosylation sites. *Glycobiology* **17**, 868–876 (2007).

Acknowledgements

We thank members of the Jontes and Sotomayor laboratories; Dr. Sharon R. Copper and Dr. Deepanshu Choudhary for training, assistance, and discussions. This work was supported in part by the Ohio State University. Use of APS NE-CAT beamlines was supported by NIH (P30 GM124165 & S10 RR029205) and the Department of Energy (DE-AC02-06CH11357) through grants GUP 49774 and 59251. D.M. is a Pelotonia fellow and M.S. was an Alfred P. Sloan fellow (FR-2015-65794).

Author Contributions

D.M. did cloning, protein expression, and protein purification, and carried out all binding and bead aggregation assays, crystallization trials, differential scanning fluorimetry experiments, and sequence alignments. M.S. trained D.M. and assisted with crystal fishing and cryo-cooling. M.S. supervised work and assisted with data analysis. D.M. and M.S. solved crystal structures, prepared figures, and wrote the manuscript.

Additional information

Supplementary information accompanies this paper at https://doi.org/10.1038/s42003-019-0586-0.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.