Published in final edited form as:
Structure. 2018 September 04; 26(9): 1210–1225.e4. doi:10.1016/j.str.2018.06.003.

Zooming in on cadherin-23: Structural diversity and potential mechanisms of inherited deafness

Avinash Jaiganesh[a,b], **Pedro De-la-Torre**[a], **Aniket A. Patel**[a], **Domenic J. Termine**[a], **Florencia Velez-Cortes**[a], **Conghui Chen**[a], and **Marcos Sotomayor**[a,b,1]

[a]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210

[b]Biophysics Graduate Program, The Ohio State University, Columbus OH 43210

SUMMARY

Cadherin-23 (CDH23) is an essential component of hair-cell tip links, fine filaments that mediate inner ear mechanotransduction. The extracellular domain of CDH23 forms about three-fourths of the tip link with 27 extracellular cadherin (EC) repeats that are structurally similar but not identical to each other. Calcium (Ca^{2+}) coordination at the EC linker regions is key for tip-link elasticity and function. There are ~116 sites in CDH23 affected by deafness-causing mutations, many of which alter conserved Ca^{2+}-binding residues. Here we present crystal structures showing 18 CDH23 EC repeats, including the most and least conserved, a fragment carrying disease mutations, and EC repeats with non-canonical Ca^{2+}-binding motif sequences and unusual secondary structure. Complementary experiments show deafness mutations' effects on stability and affinity for Ca^{2+}. Additionally, a model of 9 contiguous CDH23 EC repeats reveals helicity and potential parallel dimerization faces. Overall, our studies provide detailed structural insight into CDH23 function in mechanotransduction.

---

[1] Corresponding Author and lead contact: Marcos Sotomayor, Department of Chemistry and Biochemistry, The Ohio State University, 484 W. 12th Avenue, Columbus OH 43210, USA, (614) 688-2070, sotomayor.8@osu.edu.

AUTHOR CONTRIBUTIONS

AJ performed cloning, protein expression, purification, crystallization, and X-ray structure solution for *hs* CDH23 EC6-8, *mm* CDH23 EC7-8, *mm* CDH23 EC19-21 (WT), *mm* CDH23 EC19-21 (S2064P), *mm* CDH23 EC24-25 and for disease mutants *mm* CDH23 EC19-21 (R2006W) and *mm* CDH23 EC19-21 (S2064P; D2125N). AJ did cloning, protein expression and purification for all mutants tested with DSF and Ca^{2+} titration experiments. PD performed cloning, protein expression, purification, and crystallization of *dr* Cdh23 EC1-3 and *mm* CDH23 EC17-18. PD and MS solved structures of *dr* Cdh23 EC1-3 and *mm* CDH23 EC17-18. FV performed cloning for *hs* CDH23 EC13-14, FV and CC expressed and purified this protein fragment, and FV, CC, and PD set up crystallization. PD and MS solved the X-ray structure for *hs* CDH23 EC13-14. AAP performed cloning, protein expression, purification, and crystallization of *hs* CDH23 EC21-23 and *mm* CDH23 EC22-24. AAP and AJ solved these X-ray structures. DJT performed cloning, protein expression, purification, and crystallization of *mm* CDH23 EC12-13 (S1316D), and AJ solved the X-ray crystal structure. AJ and MS wrote and edited the manuscript with feedback from PD and all co-authors. MS trained all co-authors, supervised, and assisted in cloning, crystal fishing/cryo-cooling, structure refinement, deposition, and data analysis presented in this work.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INTRODUCTION

The inner ear, containing the cochlea and the vestibular organs, is responsible for sound perception and for balance sensing in vertebrates (Manley and Clack, 2004). Inner-ear function is mediated by sensory receptors known as hair cells featuring bundles of hair-like projections called stereocilia (Fettiplace and Kim, 2014; Gillespie and Müller, 2009; Hudspeth, 2014). Electron microscopy (EM) images of the hair-cell apical surface show that stereocilia are arranged in rows from shortest to tallest, which defines their excitatory axis (Figure 1A) (Furness and Hackney, 1985; Hudspeth and Corey, 1977; Kachar et al., 2000). Forces from sound or head movements that deflect stereocilia along this axis are conveyed to transduction channels through a protein filament called the tip link. This filament is essential for hair-cell function across species and hair-cell types, directly gating transduction channels to mediate conversion of mechanical forces into electrical signals in a process termed mechanotransduction (Assad et al., 1991; Basu et al., 2016; Fettiplace and Kim, 2014; Gillespie and Muller, 2009; Pickles et al., 1984).

The tip link is thought to be made of cadherin-23 (CDH23) and protocadherin-15 (PCDH15) in rodent, chicken, human, frog, and fish hair cells (Ahmed et al., 2006; Basu et al., 2016; Indzhykulian et al., 2013; Jaiganesh et al., 2017; Kazmierczak et al., 2007; Lelli et al., 2010; Siemens et al., 2004; Söllner et al., 2004; Sotomayor et al., 2012). These extraordinarily large proteins belong to the cadherin superfamily of Ca^{2+}-dependent adhesion molecules (Gul et al., 2017). The tip-link architecture *in vivo* has not been fully established, but immunolabeling and EM images in rodents show that transient immature tip links are made of PCDH15 only (Indzhykulian et al., 2013; Rzadzinska and Steel, 2009), while in mature tip links PCDH15 forms the lower end and CDH23 forms the upper part of this 150-185 nm long filament (Figure 1B). These data also indicate that tip links have two parallel PCDH15 molecules (*cis* homodimer) interacting tip-to-tip with a *cis* homodimer of CDH23. Such a heterotetrameric complex is presumed to be robust in response to large forces from loud sound.

CDH23 is anchored to the cell membrane at its C-terminus through a transmembrane helix and corresponding cytosolic domain. This domain couples CDH23 to the cytoskeleton via various binding partners including Myo1c, Myo7a, Harmonin, and SANS (Ahmed et al., 2003a; Pan et al., 2009; Petit, 2001; Siemens et al., 2004; Wu et al., 2012; Yan et al., 2010), which are involved in transport and localization of the transduction apparatus, adaptation, and tip-link regeneration (Boëda et al., 2002; Holt et al., 2002; Michalski et al., 2009). CDH23's longest isoform has 27 extracellular cadherin (EC) repeats and one non-canonical 28th domain, constituting about three-fourths of the tip-link filament. Each EC repeat is ~110 residues long and the complete ectodomain spans ~3,000 residues. Linker regions between EC repeats are highly conserved in sequence and bind Ca^{2+} ions essential for tip-link function (Sakaguchi et al., 2009). CDH23 isoforms with varying ectodomain lengths are known and their localization and function have been partially explored (Bahloul et al., 2010; Boëda et al., 2002; Lagziel et al., 2005; Pepermans and Petit, 2015; Takahashi et al., 2016; Xu et al., 2008).

Initial structural studies of the N-terminal end of mouse CDH23 EC1 and EC2 (CDH23 EC1-2) showed a typical cadherin fold for each repeat and a canonical linker region with three bound Ca^{2+} ions (sites 1, 2, and 3; Figure 1C-E) (Elledge et al., 2010; Sotomayor et al., 2010). The three ion-binding sites showed varying affinity (K_d) for Ca^{2+} with values ranging from ~5 μM to 80 μM (Sotomayor et al., 2010). This is particularly relevant in the cochlea, where the Ca^{2+} concentration of the endolymph that surrounds tip links is tightly regulated around 20 to 40 μM in mammals (Salt et al., 1989). In addition, molecular dynamics (MD) simulations of the CDH23 EC1-2 fragment saturated with Ca^{2+} predicted that these repeats are stiff, suggesting that the entire CDH23 and PCDH15 tip link is not elastic (Sotomayor et al., 2010). However, EC repeats and linker regions vary in sequence throughout CDH23 and PCDH15 (Araya-Secchi et al., 2016; Powers et al., 2017), and structures and affinity for Ca^{2+} have not been determined for all EC repeats and linker regions. Some elastic properties may emerge only at low Ca^{2+} concentrations or as we probe the entire molecule or large fragments of it. The degree of CDH23 elasticity is important, as it determines how the tip link conveys force to the hair-cell transduction channel (Basu et al., 2016).

CDH23 was first identified as essential to hearing due to its involvement in hereditary deafness. Defective, null alleles of *CDH23* or *PCDH15* genes result in Usher syndrome characterized by profound deafness and progressive blindness (Ahmed et al., 2003a; Petit, 2001). There are also over 100 missense mutations affecting the CDH23 and PCDH15 extracellular domains that have been shown to cause either Usher syndrome or non-syndromic deafness in humans (reviewed in Jaiganesh et al., 2017). In addition, several mouse models have been established to study hearing impairment due to missense mutations in the CDH23 and PCDH15 ectodomains. The *salsa* (Schwander et al., 2009), *jera* (Manji et al., 2011), and *erlong* (Han et al., 2012; Noben-Trauth et al., 2003) mice carry missense mutations and in-frame deletions in *Cdh23*, while *noddy* mice (Geng et al., 2013) carry a missense mutation in *Pcdh15*. Variations in phenotype among these mouse models suggest that single missense mutations may disrupt tip-link function through different molecular mechanisms (Figure S1). However, structures and biochemical studies for most of the CDH23 and PCDH15 EC repeats have not been conducted and the effects of a majority of the deafness-causing mutations remain unknown.

To investigate the structural determinants of CDH23 function and the effects of potentially disruptive missense mutations, we solved the structures of 18 out of 27 EC repeats in CDH23. We found structural variations throughout the CDH23 ectodomain that are conserved in sequence among different species, implying functional significance and highlighting the need for structural variation among tip-link EC repeats. These structures allowed us to visualize the location of ~65% of the deafness mutations on CDH23 and to predict their effects on protein structure and function based on the features of the EC repeats that they may disrupt. We also tested our structure-based predictions by measuring the effects of some of these disease-related mutations on the structure and stability of a CDH23 fragment.

RESULTS

Protein sequences of the CDH23 ectodomain from up to 20 different species were compared through multiple sequence alignments revealing an overall identity of ~51% (Figure 1F, Table S1, and Data S1). These alignments also show significant variations in sequence conservation among EC repeats, with EC1-4, EC7, and EC28 being most conserved (> 62%) and EC19-20 being the least (~35% each). Variations among EC repeats from the same species were determined by comparing protein sequences for each of the EC1 to EC27 repeats of *Homo sapiens* (*hs*) CDH23 to each other (Figures 1G and S1). In contrast to the species comparison, the average sequence identity among all repeats was ~28% with maximum identity for EC5 and EC23 (43%) and minimum for EC3 and EC27 (~15%). Although EC repeats are relatively conserved throughout vertebrates, each EC repeat within a species is considerably different from the next. Visualizing the structural consequences of these sequence variations could reveal key features of CDH23 that drive tip-link mediated mechanotransduction. We therefore solved structures for protein fragments (not all contiguous) that cover two thirds of the CDH23 ectodomain, including the most and least conserved EC repeats across species.

Multiple CDH23 protein fragments of various sizes and from diverse species were used for crystallization trials (Chruszcz et al., 2008). Successful crystallization and structure determination was possible for *Danio rerio* (*dr*) Cdh23 EC1-3, *hs* CDH23 EC6-8, *Mus musculus* (*mm*) CDH23 EC7-8, *mm* CDH23 EC12-13 (S1316D), *hs* CDH23 EC13-14, *mm* CDH23 EC17-18, *mm* CDH23 EC19-21, *hs* CDH23 EC21-23, *mm* CDH23 EC22-24, and *mm* CDH23 EC24-25 (Figures 1H, 2-6, S2 and S3 and Table 1). Structures for these fragments show 18 of the 27 EC repeats of CDH23 including 16 previously unresolved EC repeats, and 13 of its 26 linker regions. Additionally, we solved the structures of two disease variants of *mm* CDH23 EC19-21 carrying mutations p.R2006W and p.D2125N that affect the EC19-20 linker region (residue numbering in the text and structures corresponds to processed of the longest isoform protein from their corresponding species unless otherwise stated, see Methods).

All CDH23 structures show the typical seven β-strand cadherin Greek-key motif topology (Figure 1C; β-strands labeled as A-G). The linker region between adjacent repeats usually contains conserved, acidic residues coordinating up to three Ca^{2+} ions with octahedral geometry (sites 1, 2, and 3; Figure 1D). The canonical cadherin Ca^{2+}-binding motif is represented as N-XEX-DXD-D(R/Y)E-XDX-DXNDN-C (Figure 1E). Canonical linker regions were observed in 8 out of the 13 solved. Interestingly, several EC repeats and linker regions display significant variations when compared to other cadherin structures. These variations include loops of different lengths, additional secondary structure elements, and non-canonical linker regions as described below.

VARIABILITY OF LOOPS AND SECONDARY STRUCTURE IN CDH23 EC REPEATS

Six loops connect the β-strands in each cadherin repeat (AB, BC, CD, DE, EF, and FG; Figure 1C, Data S1). Loops involved in Ca^{2+} binding include AB (XEX motif) and EF (DRE motif) from the preceding EC repeat as well as the BC loop (DXD motif and backbone coordination at site 3) and the top of β-strand F (XDX motif) from the subsequent EC repeat. Loops CD, DE, and FG rarely interact with the linker region. Interestingly, loops BC, CD, EF, and FG vary in length from one EC repeat to the next. In contrast, the AB and DE loops are the most invariant (Table S2). Here we analyze unusual structural features in the 16 previously unresolved CDH23 EC repeats presented in this work.

Structural features with deviations from those seen in canonical EC repeats include a β-strand B break in EC19, BC loops of EC7 and EC17, CD loops of EC13 and EC25, and FG loops of EC3, EC8, EC21, and EC23, which are all 2-7 amino acids off their median lengths for all CDH23 EC repeats (Figures 2A, 2D, 3A, 3D, 4A, 5A, 6A, and 6F; Table S2). In EC17, β-strands B and E are shifted towards the EC17-18 linker region where loop residues are usually observed (Figures 4A and 4B).

We also observe non-canonical loops with large length changes resulting in atypical secondary structure elements. The most unusual loop in the structures is the EF loop in *hs* CDH23 EC8, which contains 19 residues with 7 of these (824 to 830) folding into a 2-turn α-helix that is not seen in other cadherin structures (Figures 2D and S2A; Table S2). This helix points towards the EC8-9 linker region and its sequence is highly conserved among species (Data S1). The structure of *mm* CDH23 EC7-8 also shows an identical EC8 EF-loop helix to its human counterpart.

Another non-canonical feature is the 27-residue-long β-strand-C-CD-loop-β-strand-D fragment in EC21 (Figures 5A, 5B, 7C, and S2B), which is 12 residues longer than the median length for all CDH23 repeats for this region. As a result, β-strand C is 5 residues longer than the median value, and the remaining residues in this region fold back to make a longer β-strand D as well. This peculiar structure allows p.D2196 from the CD-loop of EC21 to interact with p.R2225 at the N**R**E motif in the EC21 EF-loop stabilizing the subsequent linker region. Another salt bridge between p.D2195 and p.R2199 on the same loop keeps the elongated β-strands C and D rigid (Figure 5F). The interactions of these loops with their corresponding linker regions may influence CDH23's response to force.

INTER-REPEAT LINKER REGIONS IN CDH23

Cadherin linker regions determine the shape and elasticity of cadherin ectodomains and typically bind three Ca^{2+} ions at sites 1, 2, and 3 (Figures 1D and 1E). We observe canonical cadherin linker regions at *dr* Cdh23 EC1-2 (Figure 1D), *hs* CDH23 EC6-7 (Figure S4A), *hs* and *mm* CDH23 EC7-8 (Figure 2F), *hs* CDH23 EC13-14 (Figure 3G), *mm* CDH23 EC17-18 (Figure 4C), EC19-20 (Figure 5D), EC22-23, and EC23-24 (Figures 6C and 6D). The typical, conserved DRE Ca^{2+}-binding sequence motif is present in linker regions of EC1-2, EC7-8, EC13-14, EC17-18, EC19-20, and EC23-24, while the less common DYE variation is found in EC6-7 and EC22-23. All these linker regions bind three Ca^{2+} ions except for EC7-8, where Ca^{2+} at site 1 is replaced by a sodium ion likely due to crystallization conditions (1.2 M NaCl; Table S3). However, not all linker regions observed have this canonical architecture.

Atypical linker regions are seen in EC2-3 (Figure 2C), EC12-13 (Figure 3C), EC20-21, EC21-22 (Figures 5E and 5F), and EC24-25 (Figure 6H). Canonical Ca^{2+}-binding motif sequences (XEX and DYE) are present in EC2, but the EC2-3 linker has an unusual DIQDM sequence (usually DXNDN) and EC3 has a non-canonical X**E**X motif (usually X**D**X) on its FG loop. The larger glutamine p.Q205 in DIQDM still coordinates the Ca^{2+} ion at site 3 but prevents the canonical BC loop backbone coordination from EC3 resulting in only five protein coordinating groups for this Ca^{2+} ion, with a water molecule helping to fulfill the coordination sphere (Figure 2C). To accommodate the longer glutamate p.E293 from the unusual X**E**X motif, the FG loop and β-strand G of EC3 are bent by 90°, facilitating Ca^{2+} coordination at site 3 and an unusual salt bridge between p.E293 and p.R244 from the BC loop (Figures 1C, 1D, and 2C). Although this unusual architecture does not prevent Ca^{2+} binding, it may alter Ca^{2+} affinity at the EC2-3 linker region.

The wild-type (WT) sequence of *mm* CDH23 EC12-13 hints that this fragment has non-canonical features. First, the canonical DXNDN linker is NELDE. Similarly, the DXD motif at the top of EC13 is present as **S**XD with a conserved serine. This residue (p.S1316) was mutated to an aspartate to resemble a canonical DXD motif, hence stabilizing the linker region and facilitating structure determination. As expected, the linker region in the *mm* CDH23 EC12-13 (p.S1316D) structure shows significant differences when compared to a canonical cadherin linker region, despite featuring an engineered DXD motif. The leucine p.L1284 in the middle of the NE**L**DE linker in place of the canonical asparagine reduces one of the canonical Ca^{2+}-coordinating groups, makes a hydrophobic barrier, and favors an incomplete coordination sphere for the Ca^{2+} ion at site 3 (Figures 3C and S2C). The incomplete coordination is consistent with electron density at site 3, which was assigned to a sodium ion given the crystallization conditions used (0.5 M NaCl; Table S3) and various features of the electron density map. In WT CDH23 EC12-13, this weakened coordination sphere would be further affected by the conserved serine, p.S1316 in the **S**XD motif at the top of EC13. The partial Ca^{2+}-binding site at the top of WT *hs* CDH23 EC13-14 (Figures 3D and 3E) shows that the ion bound here is displaced from the usual binding site and away from residues p.S1316 and p.D1318, which supports atypical coordination at this site.

The linker region between EC20-21 is largely canonical and binds three Ca^{2+} ions, despite the lack of the canonical backbone coordination to Ca^{2+} at site 3 (similar to EC2-3; Figures 5C and 5E). The next linker region (EC21-22) contains an NRE motif (canonically DRE) at its EF loop. This loop is stabilized by a salt bridge between p.D2196 (EC21 CD loop) and p.R2225 (EC21 N**R**E motif) and is possible due to the long CD loop of EC21 (Figure 5F).

The EC24-25 linker region is the most unconventional and only binds two Ca^{2+} ions at sites 2 and 3. Normally, the Ca^{2+} ion at site 1 is coordinated by the XEX and DRE motifs. In EC24, the DRE motif is absent and replaced by QSYE with the loop pointing away from the Ca^{2+} ions (Figures 6H and S2D). Therefore, two out of five coordinating groups responsible for binding Ca^{2+} at site 1 are missing. A similar and unusual, low-Ca^{2+}-affinity and flexible linker region with two bound Ca^{2+} ions has also been observed between EC3 and EC4 of *hs* PCDH15 (Powers et al., 2017), yet the EC24-25 linker region is the first partial Ca^{2+}-free linker region observed in CDH23. This suggests that some CDH23 linker regions might have lower affinity for Ca^{2+} and could be more flexible.

SITES OF DEAFNESS-CAUSING MUTATIONS ON THE CDH23 ECTODOMAIN

There are ~116 sites of missense mutations on the CDH23 extracellular domain and the location of 76 of these can be mapped onto the structures presented here (Table S4). The median number of mutations per repeat for all our structures was 4 ± 2 sites. Repeats with the most number of mutation sites are EC14, EC18, EC21, and EC23 (6-7 sites) and the least mutated repeat is EC8 with just one site. The 76 mutation sites observed in our structures can be classified into three broad categories according to their position on the EC repeat (Figure S5). The first category includes residues pointing inward towards the hydrophobic core (18 sites). In the second category, residues face outwards and are away from Ca^{2+}-binding sites (33 sites). The last category has residues that affect Ca^{2+} binding (25 sites). Mutations in each group may alter protein structure and function through different mechanisms, as discussed below.

The first group of mutations to inward-facing residues comprises 18 locations and includes 6 proline residues present at the top of β-strand A of their respective EC repeat (p.P536S, p.P646S, p.P1183R, p.P1826A, p.P1934S, and p.P2377S). As part of the conserved X**P**XF/L motif immediately after the DXNDN linker, mutations to these prolines may prevent proper folding of β-strand A, compromising EC repeat stability. The next subgroup of mutations includes 3 sites where small hydrophobic residues are changed to larger hydrophobic ones, namely, p.V116I (human numbering; p.I114 in zebrafish), p.V1885I, and p.V2612F. These mutations potentially cause steric hindrance at the protein core. Conversely, large hydrophobic residues are also mutated to smaller side chains that likely reduce packing efficiency at the hydrophobic core, destabilizing EC repeats (p.Y1223H, p.F1865S, and p.F2554S). The last six mutations in this group include: p.A235T, p.G1994S, p.G2306V, G2432V, and p.L2450P that may impair packing, and mutation p.V2337E (*jera*) on EC22 that is located just prior to the DYE Ca^{2+}-binding motif and would likely disrupt the structure of this loop and hence the linker region between EC22 and EC23.

The second group of mutations includes outward facing residues and constitutes the largest group with 33 sites mapped on our structures. These can be further classified into charge swap or loss of charge mutations (11 sites), mutations changing large polar residues (3 sites), mutations changing small polar residues (9 sites), mutations changing hydrophobic residues (8 sites), and mutations at the CDH23-PCDH15 interface (2 sites). In the first subgroup, p.E2415K located on EC23 is the only mutation that results in a charge swap. The next 8 out of the remaining 10 sites in this sub-group involve arginines and result in the loss of a positive charge and change in length of the side chain (p.R559Q, p.R1394W, p.R1484Q, p.R1723Q, p.R2148C, p.R2359Q, p.R2466H, p.R2585H, and p.R2585C). The p.R1394W, p.R2466H, and p.R2585H mutations result in bulky residues pointing to the surface and can lead to unfavorable solvent interactions or loss of potential salt bridges involved in *cis* interactions. The last two sites in this subgroup involve loss of negatively charged, surface exposed residues p.D1400N in EC14 and p.D2196N in EC21. The latter may result in destabilization of a nearby Ca^{2+}-binding motif by disrupting the p.D2196-p.R2225 salt bridge (N**R**E motif; Figure 5F). Charge swap or loss of charge mutations, other than a special case in the EC21 CD loop, may affect hydrogen bonding and salt bridges that are involved in the CDH23 *cis* homodimer interface.

The large polar residues of the next subgroup of outward facing residues mutated in deafness (p.H732Y, p.Q1473H, and p.Q2204P) are found on EC7 (Figure 2D), EC14 (Figure 3D), and EC21 (Figures 5A and 7C), respectively. Residue p.H732 is located on a 3_{10} helix on the FG loop of EC7 and its mutation may affect loop stability, possible *cis* interactions, or may allosterically affect the nearby EC6-7 linker region. On EC14, p.Q1473, p.D1487, and p.S1437 form a network of hydrogen bonds (Figure S4B). The p.Q1473H mutation may affect this network, destabilizing the EC repeat. Finally, p.Q2204 is located on the CD loop in EC21 and the ε-nitrogen hydrogen bonds with the backbone of p.P2202 (Figure S2B). Disruption of this bond by the addition of another proline may compromise the short β-strand made by this long loop. Mutations to large outward facing polar residues are mostly predicted to disrupt hydrogen bond networks on the surface of EC repeats that likely help stabilize the Greek key motif.

The subsequent subgroup of 9 outward facing, small polar residues mutated in deafness involves two glycines, five serines, and two threonines. Two of these, p.G728R and p.G1426D, are likely to cause steric hindrance to potential *cis*-dimerization interfaces in EC7 and EC14 due to significant changes in sidechain volume and charge. However, these are also proximal to the EC6-7 and EC13-14 linker regions (p.G728 is on the FG loop upstream of the XDX motif and p.G1426 is on the BC loop near the site-3 backbone coordination) and may impact Ca^{2+} binding.

The second to last subgroup of mutations to outward facing residues involves hydrophobic side chains at 8 sites: p.A156T on EC2 (p.S154 in zebrafish), p.P217L on EC3 (p.P215 in zebrafish), p.V780I on EC8, p.V1784M on EC17, p.A2107V on EC20, p.L2200W on EC21, and p.I2646V and p.M2682V on EC25. The p.L2200W mutation is on the long CD loop of EC21 and may alter its structure. The other outward facing hydrophobic residues might be implicated in protein-protein interactions, and mutations to these residues could affect formation of the putative CDH23 *cis* homodimer, weakening the tip-link response against large forces from loud sounds. The last subgroup includes mutations to outward facing residues at the CDH23 and PCDH15 interface: p.Y16H (p.Y14 in zebrafish) and p.G139R (p.G137 in zebrafish; Figures 2A, S4C), likely disrupting tip-link formation.

The last group of mutations includes residues that either directly coordinate Ca^{2+} ions or are in between two Ca^{2+}-binding residues at Ca^{2+}-binding motif sequences. There are 25 examples of these mutations categorized in four subgroups. The first subgroup (10 sites) includes Ca^{2+}-coordinating aspartates and glutamates changed to similarly sized neutral asparagines and glutamines, respectively. This may result in reduced Ca^{2+}-binding affinity. Another populous subgroup includes conserved aspartates or glutamates mutated to small side chains like glycine or valine (6 sites). A well characterized example is the p.D101G mutation to the EC1-2 linker showing a ~6 fold reduction in Ca^{2+} affinity and a loss of binding cooperativity (Sotomayor et al., 2010). The third subgroup contains mutations that swap charge at acidic residues (3 sites) or change them to large hydrophobic residues (2 sites). Swapping negatively charged residues to lysine or arginine would cause electrostatic repulsion and severe disruption of Ca^{2+}-binding sites. Similarly, substitutions to large hydrophobic residues in a highly hydrophilic environment would cause significant instability at the linker region. Finally, mutations to Ca^{2+}-binding motif residues that do not directly bind Ca^{2+} usually occur at the arginine residue of the D**R**E motif. There are 4 different examples of such sites where R is mutated to W, P, or Q with unknown effects. These mutations (and any that alter residues chelating Ca^{2+}) may compromise the rigidity and rupture force of the tip link.

EFFECTS OF DISEASE MUTATIONS ON CDH23 STABILITY AND LINKER REGION STRUCTURE

To test some of the mechanisms predicted for CDH23 mutations, we focused on the well-behaved EC19-21 fragment and measured its thermal stability and Ca^{2+} sensitivity in WT and mutant forms. We also solved structures of two disease mutants of this construct that affect residues in Ca^{2+}-binding motifs (p.R2006W and p.D2125N).

Mutations to the EC19-21 fragments tested include an inward facing residue (p.G1994S), outward facing residues (p.A2107V in EC20; p.L2200W and p.Q2204P on the EC21 CD-loop) and residues in Ca^{2+}-binding motifs (p.R2006W, p.D2125N at the EC19-20 linker region and p.D2179N at the EC20-21 linker region; Figures 5A and 7A, 7B, and 7C). All constructs tested did not display obvious problems in protein expression, purification, and refolding. We performed two differential scanning fluorimetry (DSF) -based assays to assess in a semi-quantitative way thermal stability (all mutants) and stability at different free-Ca^{2+} concentrations (Ca^{2+}-binding motif mutants).

Disease mutations can reduce thermal stability at saturating Ca^{2+} concentrations

Thermal stability of WT and mutant *mm* CDH23 EC19-21 constructs was determined via DSF at saturating Ca^{2+} concentrations (5 mM CaCl$_2$). The melting temperature T_m for WT was 61.8 ± 1.0 °C (Figure 7D, Table 2). The same experiment was carried out for mutants p.R2006W, p.D2125N, and p.D2179N altering linker regions, and mutants p.G1994S (on EC19), p.A2107V (on EC20), p.L2200W, and p.Q2204P (on EC21) that are not directly involved in or adjacent to Ca^{2+}-binding sites.

Mutations to Ca^{2+}-binding motif residues showed a greatly reduced T_m with p.R2006W melting at 51.9 ± 0.6 °C, p.D2125N melting at 49.4 ± 1.2 °C, and p.D2179N, which affects coordination of Ca^{2+} at both sites 2 and 3, melting at 53.0 ± 0.7 °C (Figure 7D, Table 2). In contrast, mutations to residues not involved in Ca^{2+} binding showed less drastic changes in thermal stability. Mutants p.G1994S (57.7 ± 0.1 °C), p.L2200W (58.3 ± 0.3 °C), and p.Q2204P (58.9 ± 0.5 °C) showed minor T_m deviations from WT. Mutation p.A2107V, melting at 61.7 ± 0.1 °C, had the least deviation from WT (Figure 7E, Table 2). These results show that mutations to Ca^{2+}-binding motif residues alter EC repeat stability even at saturating Ca^{2+} concentrations.

Ca^{2+}-dependent thermal stability varies for different Ca^{2+}-binding mutations

The WT *mm* CDH23 EC19-21 fragment and mutants p.R2006W, p.D2125N, and p.D2179N were Ca^{2+}-depleted to test their Ca^{2+}-dependent stability through DSF. The change in melting temperature ΔT_m from the Ca^{2+}-free control (T_m at 25 mM EGTA; Table 2) was monitored as a function of free Ca^{2+} concentration (from 20 μM to 2.25 mM Ca^{2+}). All protein fragments were stabilized by addition of Ca^{2+}. The WT *mm* CDH23 EC19-21 fragment showed a ΔT_m of ~11 °C at 2.25 mM free Ca^{2+} (Figure 7F), and mutants p.D2125N and R2006W showed similar changes in ΔT_m, with increasing amounts of Ca^{2+}. In contrast, the p.D2179N mutant was less sensitive to changes in Ca^{2+} concentration (ΔT_m ~6.5 °C at 2.25 mM free Ca^{2+}; Figure 7F). These results indicate that mutations to Ca^{2+}-binding motif residues modify Ca^{2+} binding differentially.

Structures of *mm* CDH23 EC19-21 R2006W and D2125N mutants

Two structures were obtained for *mm* CDH23 EC19-21 carrying deafness mutations at the EC19-20 linker region (Table 1). The first structure has an p.R2006W mutation that changes a conserved arginine in the D**R**E motif of EC19 (Figure 7G). The other has an p.D2125N mutation at the aspartate in the X**D**X^{top} motif of EC20 (Figure 7H), along with a second, structurally inconsequential mutation (p.S2064P) that facilitated structure determination (RMSD Cα between CDH23 EC19-21 WT and S2064P structures is 0.24 Å). Surprisingly, all three Ca^{2+} ions seen in the structure of the WT are still present in both mutant structures (Figure S3). However, the hydrogen bonds formed between p.R2006 with both the backbone of p.E2073 and the p.S2075 sidechain are lost (Figure 7G). For the p.D2125N mutation, the WT sidechain hydrogen bonds with the backbone nitrogen of p.L2131. When mutated to p.D2125N, the sidechain flips by 180° and the δ-nitrogen hydrogen bonds with the backbone of E2080. The subtle structural perturbations observed in both structures suggest that mutations to Ca^{2+}-binding motif residues primarily reduce the site's affinity for Ca^{2+} and may reduce the number of bound Ca^{2+} ions at physiological Ca^{2+} concentrations, hence destabilizing EC repeats. This would reduce CDH23’s ability to withstand force from large stimuli.

EMERGENT PROPERTIES FROM LONG CDH23 MODELS

The structures presented here allowed us to build a continuous model from EC19 to EC25 (all linker regions from structures), to which we added *mm* CDH23 EC17-18 (EC18-19 linker region modeled; Figure S6). This structural model with 9 continuous EC repeats was used to map disease mutation sites as well as predicted and known glycosylation sites on its surface (Vester-Christensen et al., 2013; Zielinska et al., 2010). Interestingly, they mostly fall on separate faces of the CDH23 EC17-25 model (Figure 8D). This suggests that one of the CDH23 faces is glycosylated and hence unavailable for *cis* interactions, while the other is available for homodimerization and requires specific residues for CDH23 function. This is further supported by sequence conservation and hydrophobicity analyses (Figure S7). None of the short CDH23 fragments we crystallized behaved as multimers in solution, nor were any *cis* interactions evident from crystal contacts. Perhaps *cis* dimerization can only be observed biochemically with longer CDH23 constructs where glycosylation and multiple contact points along the surface promote more robust interactions involving the CDH23 face affected by deafness mutations to surface accessible residues.

EM images of intact hair cells suggest that tip links adopt a helical structure when connecting two stereocilia with helical turns being repeated every 20 to 25 nm (Kachar et al., 2000; Tsuprun et al., 2004). The CDH23 EC17-25 model spans ~42 nm and appears to have two turns of a shallow helix (repeating every 4.5 EC repeats). The end-to-end distance for the first helical turn of CDH23 EC17-25 is ~27 nm. When the CDH23 EC17-25 model is aligned to the *x*-axis and the projections of the backbone coordinates in the *xy* and *xz* planes are plotted (Figure S7C), we observe an out of phase sinusoidal character for at least one of the two turns in the CDH23 EC17-25 structure (0-20 nm). While inter-repeat orientations from crystallographic structures and models (Figures S6D and S6E) should be cautiously interpreted, our data support the notion that the tip-link cadherins possess inherent helicity.

DISCUSSION

The structures of CDH23 presented here shed light on the heterogeneous architecture of its EC repeats and linker regions, and how these may influence CDH23 function. Two intriguing aspects of these structures are the lack of bona fide multimerization interfaces, and the presence of non-canonical linker regions with altered Ca^{2+}-binding sites. It is possible that multimerization requires analysis of longer constructs or posttranslational modifications. All our structures were obtained from short protein fragments expressed in bacteria and refolded, and as such, lack glycosylation. However, similar strategies have successfully identified homophilic interfaces mediating adhesion for clustered protocadherins and non-clustered delta-protocadherins (Cooper et al., 2016; Nicoludis et al., 2015, 2016). The molecular determinants of CDH23 multimerization and the role of glycosylation in CDH23 and tip-link function remain to be determined.

The presence of non-canonical linker regions with altered Ca^{2+}-binding sites in CDH23 suggests points of flexibility along its extracellular domain, most notably at EC2-3, EC12-13, EC20-21, and EC24-25. Molecular dynamics simulations and accurate measurements of Ca^{2+} dissociation constants for these linker regions should clarify their role in tip-link function *in vivo*, when exposed to the low Ca^{2+} concentrations of cochlear and vestibular endolymph (Salt et al., 1989).

Deafness-causing mutations on the CDH23 extracellular domain have been identified at ~116 distinct sites. A total of 76 sites can now be directly visualized in our structures, which allowed us to classify the corresponding mutations into three groups. These groups include mutations to residues pointing towards the hydrophobic core of the EC repeat, mutations of surface residues, and mutations to residues involved in Ca^{2+} binding (Figure S5A). Based on these classifications, thermal stability analyses, and Ca^{2+} sensitivity studies, we suggest five molecular mechanisms (Figures S5B and S8) through which inherited missense mutations may affect CDH23 function resulting in varying phenotypes.

The first mechanism involves destabilization of the hydrophobic core (e.g., p.F1865S in EC18 (Schultz et al., 2005), p.V2612F on EC25 (Brownstein et al., 2011)). The second mechanism includes mutations that affect *trans* interactions with PCDH15 potentially causing severe deafness (p.Y14H (Abdi et al., 2016) on CDH23 and p.R113G (Ahmed et al., 2008) on PCDH15) and additional vestibular dysfunction (p.G137R (Lenarduzzi et al., 2015) in CDH23 and p.I108N (Geng et al., 2013) in PCDH15) depending on whether *trans* interaction is impaired or completely abolished. The third mechanism corresponds to interactions altering residues on unique secondary structure elements that may affect

CDH23's force response. The fourth mechanism involves disruption of possible cis-dimer interfaces by mutations to outward facing residues (e.g., p.V780I on EC8 (Miyagawa et al., 2012)). In the last mechanism, destabilization of EC repeats is caused by mutations to residues at or near Ca^{2+}-binding sites. These mutations may cause progressive hearing loss (e.g., p.D101G (Astuto et al., 2002), p.E714V (Schwander et al., 2009)) or more severe deafness (p.V2337E (Manji et al., 2011)) depending on how they alter affinity for Ca^{2+}, linker region structure, and number of Ca^{2+}-binding sites perturbed. Decreased affinity for Ca^{2+} may result in a more flexible linker region that is unable to maintain rigidity to favor engagement with PCDH15 or to form proper *cis* dimerization (Sotomayor et al., 2010). Linker regions may also unfold more readily in this scenario, leading to flexibility and soft elasticity possibly preventing tip-link function in response to repetitive stimuli. Interestingly, none of the mutations that alter Ca^{2+}-binding sites seem to cause vestibular defects in humans and mice, perhaps because endolymph Ca^{2+} levels are higher in this system, or because mechanical stimuli are less demanding. Results from our thermal melting assays, which may correlate with mechanical strength of EC repeats, suggest that some mutations at Ca^{2+}-binding motif sequences can alter stability even at saturating Ca^{2+} concentrations, thus favoring the hypothesis that differences in mechanical stimuli underlie phenotype in these cases.

Expression of CDH23 and its isoforms has been detected in multiple organs, tissues, and cells, most notably in the eye photoreceptors (Ahmed et al., 2003b; Lagziel et al., 2009; Sahly et al., 2012; Schultz et al., 2011). Usher syndrome, in which patients have hearing loss with progressive blindness (retinitis pigmentosa), has been typically associated to null mutations of CDH23. However, in humans, multiple missense mutations have been definitively associated with Usher syndrome involving retinitis pigmentosa, 19 of which can be mapped on our structures (Table S4). It is unclear why some missense mutations only cause a defective hearing phenotype while others seem to affect both hearing and eyesight. Further biochemical and biophysical analyses will be needed to determine whether missense mutations altering the mechanics and stiffness of CDH23 at low Ca^{2+} concentrations, but not its trafficking, stability, and ability to bind partners, are responsible for limited non-syndromic phenotypes.

Initial simulation studies suggested that the tip-link cadherins CDH23 and PCDH15 are stiff (Sotomayor et al., 2010). Recent structures of the mouse and human PCDH15 EC9-10 Ca^{2+}-free linker region revealed a bent conformation (Araya-Secchi et al., 2016). This non-canonical linker region is predicted to remain bent under resting tension, providing soft elasticity and limited extensibility (4 nm) during unbending caused by physiological forces. While our CDH23 structures reveal some non-canonical linker regions, none of them exhibit bending as seen for PCDH15 EC9-10. Yet, at the low Ca^{2+} concentrations of the cochlear endolymph bending might be pronounced leading to flexibility at atypical CDH23 linkers, some of which still need to be structurally characterized (Figure S1). The helical structure observed in 9 consecutive EC repeats that could easily extend from 40 to 45 nm (assuming that resting tension does not keep CDH23 straight) also suggests that some elasticity and extensibility can be provided by CDH23.

The structures and analyses presented here provide a high-resolution view of CDH23 and have implications beyond inner-ear function, deafness, vision and blindness. Several other cadherins have unusually long extracellular domains. Protocadherin-24 (9 EC repeats) forms intermicrovillar links essential for gut function (Crawley et al., 2014) and its EC repeats may share some of the structural variations observed for CDH23. Similarly, the FAT cadherins (32-34 EC repeats), the flamingo cadherins (9 EC repeats), and the dachsous cadherins (22-27 EC repeats) are all involved in essential physiological processes that rely on their long extracellular domains (Tsukasaki et al., 2014), which feature several structural variations, some similar to the ones described here for CDH23.

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Marcos Sotomayor (sotomayor.8@osu.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Escherichia coli (*E. coli*) DH5α and XL10 – Gold cells were used for molecular cloning and mutagenesis. *E. coli* Codon Plus and Rosetta cells were used for expression of CDH23 protein fragments.

METHOD DETAILS

Multiple sequence alignment and residue numbering

To compare CDH23 extracellular domain sequences among various species, sequences were obtained for their longest CDH23 isoforms from the NCBI protein database (Table S1) and were split into 27 EC repeats before alignment. These fragments were then aligned using the ClustalW algorithm (Larkin et al., 2007) on Geneious (Kearse et al., 2012) to obtain percent sequence identity for each EC repeat (Data S1). Alignment files from Geneious were colored according to sequence conservation in JalView with 45% conservation threshold (Waterhouse et al., 2009). To compare sequences among different EC repeats of CDH23 for the same species, the *hs* CDH23 sequence was split into individual EC repeats and aligned using Clustal-Omega (Sievers et al., 2011), and the resulting alignment (Figure S1) was input into the Sequence Identity and Similarity (SIAS) server (Garcia-Boronat et al., 2008) to obtain a sequence identity matrix.

Residue numbering throughout the text and in the structures corresponds to the processed protein, which does not include the signal peptide. Numbering is based on entries *mm* CDH23 isoform 1 precursor, NCBI Reference Sequence: NP_075859.2 (signal peptide with 23 residues), *hs* CDH23 isoform 1 precursor, NCBI Reference Sequence: NP_071407.4 (signal peptide with 23 residues), and *dr* Cdh23 precursor, NCBI Reference Sequence: NP_999974.1 (signal peptide with 31 residues).

Cloning, protein expression and purification

All constructs coding for CDH23 fragments used for structure determination were subcloned into a pET21a+ vector containing a C-terminal hexahistidine tag using a 5'-end NdeI or NheI site and a 3'-end XhoI or HindIII site, as described in deposited structures. Constructs coding for *mm* CDH23 EC19-21 fragments (WT and mutants) for melting experiments used NdeI and XhoI pET21a+ sites and included residues p.N1932 to p.N2266 (all same length). Mutations were generated using the QuikChange Lightning kit. All constructs were sequence verified. All proteins were overexpressed in *E. coli* BL21-RIPL or Rosetta cells (Stratagene) cultured in TB (Fisher BioReagents). Overexpression was induced at OD_{600} ~ 0.6 with 0.2 - 1 mM of IPTG at 37 °C and then cells were grown at 30 °C for ~16 hr. Protein was harvested by first lysing cells via sonication in denaturing buffer (20 mM TrisHCl [pH 7.5], 6 M guanidine hydrochloride, 10 mM $CaCl_2$, and 20 mM imidazole). The cleared lysates were loaded onto Ni-Sepharose beads (GE Healthcare) and eluted with denaturing buffer containing 500 mM imidazole.

Protein refolding was performed by different methods. Overnight dialysis (Spectrum Spectra/Por 7 2000 MWCO) of protein (< 1 mg/ml) against 20 mM Tris HCl [pH 8.0], 150 mM KCl, 400 mM arginine, and 5 mM $CaCl_2$ was done for *hs* CDH23 EC6-8 (+2 mM DTT), *mm* CDH23 EC7-8 (+2 mM DTT), *hs* CDH23 EC13-14, *mm* CDH23 EC19-21 WT and G1994S, R2006W, A2107V, D2125N, D2179N mutants (+2 mM DTT), *mm* CDH23 EC19-21 L2200W and Q2204P, *hs* CDH23 EC21-23, *mm* CDH23 EC22-24, and *mm* CDH23 EC24-25. Fragment *mm* CDH23 EC12-13 (S1316D) with 5 mM DTT was dialyzed in the same way against 20 mM Tris HCl [pH 8.0], 150 mM NaCl, 400 mM arginine, and 2 mM $CaCl_2$. Fragment *dr* Cdh23 EC1-3 was refolded by drop-wise dilution of 20 ml of eluted denatured protein (0.9 mg/ml) into 480 ml of refolding buffer (20 mM TrisHCl [pH 7.5], 150 mM KCl, 50 mM NaCl, 5 mM $CaCl_2$, 400 mM arginine, and 10% glycerol) over 20 minutes. Similarly, *mm* CDH23 EC17-18 (1.4 mg/ml with 40 mM DTT) was diluted into refolding buffer containing 100 mM HEPES [pH 7.5], 10 mM $CaCl_2$, 10% glycerol, and 5 mM TEAC. Refolded protein was kept at 4 °C overnight with gentle stirring. All protein solutions were concentrated to <10 ml.

Refolded protein was further purified on a Superdex-200 column (GE Healthcare) in holding buffer containing 20 mM TrisHCl [pH 8.0], 150 mM KCl, and 5 mM $CaCl_2$ (*hs* CDH23 EC6-8, *mm* CDH23 EC7-8, *hs* CDH23 EC13-14, *mm* CDH23 EC19-21 [WT and G1994S, R2006W, A2107V, D2125N, D2179N mutants], *mm* CDH23 EC19-21 L2200W and Q2204P +2 mM DTT, *hs* CDH23 EC21-23, *mm* CDH23 EC22-24, and *mm* CDH23 EC24-25); 20 mM Tris HCl [pH 8.0], 150 mM NaCl, 400 mM arginine, and 2 mM $CaCl_2$ (mm CDH23 EC12-13 (S1316D)); 20 mM TrisHCl [pH 7.5], 150 mM KCl, 50 mM NaCl, and 2 mM $CaCl_2$ (*dr* Cdh23 EC1-3); or 20 mM TrisHCl [pH 8.0], 150 mM KCl, 5 mM $CaCl_2$, and 2 mM DTT (*mm* CDH23 EC17-18). Pure proteins were concentrated by centrifuge ultrafiltration to 3-12 mg/ml for crystallization.

X-ray crystallography experimental details

All crystals were grown by vapor diffusion at 4 °C using the sitting-drop method. To perform initial crystal screening, protein was mixed with various ratios of precipitant in 96 well format trays. Successful crystallization conditions were refined by finely varying pH and precipitant concentrations (Table S3). In preparation for cryogenic storage, crystals were treated with 25% - 30% glycerol, MPD, or PEG-400 and then cryo-cooled in liquid nitrogen. X-ray diffraction data sets were processed with HKL2000 or HKL3000 (Minor et al., 2006; Otwinowski and Minor, 1997). All structures were determined by molecular replacement using a separate homology model for each repeat as an initial search model using PHASER (McCoy et al., 2007). Starting models for phasing used structures 2WHV, 2WD0, 2A62, 2EE0 (Patel et al., 2006; Sotomayor et al., 2010), and CDH23 structures presented in this study (phasing of 5WT4 used 2WHV; 5TFM used 5TFL; 5TFL used 5I8D; 5VH2 used 2WHV and 5TFK; 5WJ8 used 5VH2; 5WJM used 2WD0; 5TFK used 5I8D; 5VVM used 5I8D and 5UZ8; 5UZ8 used 5I8D; 5VT8 used 5UZ8 and 2EE0; 5UN2 and 5ULU used 5I8D, 5I8D used 2A62). Model building was done with COOT (Emsley et al., 2010) and restrained TLS refinement was performed with REFMAC5 (Murshudov et al., 2011; Winn et al., 2001). Data collection and refinement statistics are provided in Table 1.

DSF to measure protein stability

Protein stability measurements at saturating Ca^{2+} concentrations (5 mM) were done with stock protein concentrations of 1 mg/ml. For DSF experiments, protein was diluted in holding buffer to ~0.3 mg/ml (~8 μM) in all cases to a volume of 19 μl and loaded into a 96-well, 0.2 ml thin-wall PCR plate (BioRad). Sypro Orange dye (Molecular Probes) was diluted to 100x with holding buffer and 1 μl was added into each reaction for a final concentration of 5x and a total reaction volume of 20 μl. The plate was sealed with optical-quality sealing tape (BioRad). Thermal denaturing was performed using a CFX96 real time-PCR instrument (Bio-Rad) where temperature was increased in a step-wise manner from 10 to 95 °C in 0.2 °C/cycle increments (ramp rate ~0.6 °C/min) and with an equilibration time of 5 s at each temperature.

Ca^{2+}-dependent protein stability through DSF

Mm CDH23 EC19-21 protein fragments were first purified via gel filtration in holding buffer containing 5 mM $CaCl_2$ and concentrated to 4 mg/ml. Two standard buffers, buffer A (20 mM Tris [pH 8.0], 150 mM KCl, 5 mM EGTA [pH 7.0]) and buffer B (20 mM Tris [pH 8.0], 150 mM KCl, 5 mM $CaCl_2$) were prepared by precisely weighing out each component. Buffers A and B were mixed in different ratios to obtain various free Ca^{2+} concentrations in solution (from 20 μM to 2.25 mM free Ca^{2+}). Ratios of buffer A to buffer B were obtained using the Maxchelator program (Bers et al., 2010). Free Ca^{2+} concentration was verified using fura-2 (Molecular Probes) fluorescence. Protein holding buffer was treated with 40 mg/ml chelex resin (BioRad) for two days, two times. Protein (2 μl) and standard buffers (18 μl were mixed together in a thin wall PCR plate. Sypro orange at 100x (1 μl in Ca^{2+}-free holding buffer) was added to the reaction mix bringing final concentration to 5x. Thermal denaturing was performed using a CFX96 real time-PCR instrument (Bio-Rad) with the same melting protocol described above. Melting temperatures were monitored as a function of Ca^{2+} concentration from 20 μM to 2.25 mM free Ca^{2+} in solution. All PCR plates, tubes, and tips involved in Ca^{2+} titrations were pre rinsed in Ca^{2+}-free holding buffer twice and thoroughly dried prior to experiments.

DSF data analysis

Averaged fluorescence data output from DSF experiments were normalized over the temperature range tested (10 °C to 95 °0). The average for several runs for each construct of normalized fluorescence was plotted as a function of temperature along with standard deviation in signal. Temperature closest to half maximal fluorescence was determined to be the T_m for a given sample. For measurements involving Ca^{2+} titration, T_m was determined for each construct at each Ca^{2+} concentration and for each replicate. The shift in average melting temperature (compared to the Ca^{2+}-free control) was plotted as a function of Ca^{2+} concentration with standard deviations.

QUANTIFICATION AND STATISTICAL ANALYSIS

Measurements of melting temperatures were done with at least two biological replicates, both for saturating calcium concentrations and calcium titration experiments. Each biological replicate for DSF experiments was also accompanied by at least one duplicate to ensure data was obtained accurately.

DATA AND SOFTWARE AVAILABILITY

Coordinates for all CDH23 structures solved have been deposited in the Protein Data Bank with entry codes 5W4T (*dr* Cdh23 EC1-3), 5TFM (*hs* CDH23 EC6-8), 5TFL (*mm* CDH23 EC7-8), 5VH2 (*mm* CDH23 EC12-13-S1316D), 5WJ8 (*hs* CDH23 EC13-14), 5WJM (*mm* CDH23 EC17-18), 5TFK (*mm* CDH23 EC19-21), 5VVM (*hs* CDH23 EC21-23), 5UZ8 (*mm* CDH23 EC22-24), 5VT8 (*mm* CDH23 EC24-25), 5UN2 (*mm* CDH23 EC19-21-R2006W), 5ULU (*mm* CDH23 EC19-21-S2064P-D2125N), and 5I8D (*mm* CDH23 EC19-21-S2064P).

Data S1. CDH23 sequence alignment. Related to Figure 1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank Deepanshu Choudhary for initial training on DSF, Deryanur Kilic for initial trials with some CDH23 fragments, and other members of the Sotomayor research group for their feedback, training, and discussions. This work was supported by the Ohio State University and the National Institutes of Health - National Institute on Deafness and Other Communication Disorders (NIH/NIDCD grants K99/R00 DC012534 and R01 DC015271), and the Alfred P. Sloan foundation (FR-2015-65794). Use of the APS NE-CAT beamlines was supported by NIH (P41 GM103403 & S10 RR029205) and the Department of Energy (DE-AC02-06CH11357) through grants GUP 40277 and 49774. DSF experiments were performed at the Plant-Microbe Genomics facility at The Ohio State University. Molecular images were prepared with VMD and PyMOL (Humphrey et al., 1996; Schrodinger, 2015). Conserved protein surfaces were calculated using ConSurf (Ashkenazy et al., 2016) PD is a Pelotonia fellow

REFERENCES

Abdi S, Bahloul A, Behlouli A, Hardelin J-P, Makrelouf M, Boudjelida K, Louha M, Cheknene A, Belouni R, Rous Y, et al. (2016). Diversity of the Genes Implicated in Algerian Patients Affected by Usher Syndrome. PLoS One 11, e0161893. [PubMed: 27583663]

Ahmed Z, Riazuddin S, Riazuddin S, and Wilcox E (2003a). The molecular genetics of Usher syndrome. Clin. Genet 63, 431–444. [PubMed: 12786748]

Ahmed ZM, Riazuddin S, Ahmad J, Bernstein SL, Guo Y, Sabar MF, Sieving P, Riazuddin S, Griffith AJ, Friedman TB, et al. (2003b). PCDH15 is expressed in the neurosensory epithelium of the eye and ear and mutant alleles are responsible for both USH1F and DFNB23. Hum. Mol. Genet 12, 3215–3223. [PubMed: 14570705]

Ahmed ZM, Goodyear R, Riazuddin S, Lagziel A, Legan PK, Behra M, Burgess SM, Lilley KS, Wilcox ER, Riazuddin S, et al. (2006). The tip-link antigen, a protein associated with the transduction complex of sensory hair cells, is protocadherin-15. J. Neurosci 26, 7022–7034. [PubMed: 16807332]

Ahmed ZM, Riazuddin S, Aye S, Ali RA, Venselaar H, Anwar S, Belyantseva PP, Qasim M, Riazuddin S, and Friedman TB (2008). Gene structure and mutant alleles of PCDH15: nonsyndromic deafness DFNB23 and type 1 Usher syndrome. Hum. Genet 124, 215–223. [PubMed: 18719945]

Araya-Secchi R, Neel BL, and Sotomayor M (2016). An elastic element in the protocadherin-15 tip link of the inner ear. Nat. Commun 7, 13458 [PubMed: 27857071]

Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, and Ben-Tal N (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res

Assad J a, Shepherd GM, and Corey DP (1991). Tip-link integrity and mechanical transduction in vertebrate hair cells. Neuron 7, 985–994. [PubMed: 1764247]

Astuto LM, Bork JM, Weston MD, Askew JW, Fields RR, Orten DJ, Ohliger SJ, Riazuddin S, Morell RJ, Khan S, et al. (2002). CDH23 Mutation and Phenotype Heterogeneity: A Profile of 107 Diverse Families with Usher Syndrome and Nonsyndromic Deafness. Am. J. Hum. Genet 71, 262–275. [PubMed: 12075507]

Bahloul A, Michel V, Hardelin J-P, Nouaille S, Hoos S, Houdusse A, England P, and Petit C (2010). Cadherin-23, myosin VIIa and harmonin, encoded by Usher syndrome type I genes, form a ternary complex and interact with membrane phospholipids. Hum. Mol. Genet 19, 3557–3565. [PubMed: 20639393]

Basu A, Lagier S, Vologodskaia M, Fabella BA, and Hudspeth AJ (2016). Direct mechanical stimulation of tip links in hair cells through DNA tethers. Elife 5.

Bers DM, Patton CW, and Nuccitelli R (2010). A practical guide to the preparation of Ca(2+) buffers. Methods Cell Biol 99, 1–26. [PubMed: 21035681]

Boëda B, El-Amraoui A, Bahloul A, Goodyear R, Daviet L, Blanchard S, Perfettini I, Fath KR, Shorte S, Reiners J, et al. (2002). Myosin VIIa, harmonin and cadherin 23, three Usher I gene products that cooperate to shape the sensory hair cell bundle. EMBO J. 21, 6689–6699. [PubMed: 12485990]

Brownstein Z, Friedman LM, Shahin H, Oron-Karni V, Kol N, Abu Rayyan A, Parzefall T, Lev D, Shalev S, Frydman M, et al. (2011). Targeted genomic capture and massively parallel sequencing to identify genes for hereditary hearing loss in Middle Eastern families. Genome Biol. 12, R89. [PubMed: 21917145]

Chruszcz M, Wlodawer A, and Minor W (2008). Determination of protein structures--a series of fortunate events. Biophys. J 95, 1–9. [PubMed: 18441029]

Cooper SR, Jontes JD, and Sotomayor M (2016). Structural determinants of adhesion by protocadherin-19 and implications for its role in epilepsy. Elife 5.

Crawley SW, Shifrin DA, Grega-Larson NE, McConnell RE, Benesh AE, Mao S, Zheng Y, Zheng QY, Nam KT, Millis BA, et al. (2014). Intestinal brush border assembly driven by protocadherin-based intermicrovillar adhesion. Cell 157, 433–446. [PubMed: 24725409]

Elledge HM, Kazmierczak P, Clark P, Joseph JS, Kolatkar A, Kuhn P, and Müller U (2010). Structure of the N terminus of cadherin 23 reveals a new adhesion mechanism for a subset of cadherin superfamily members. Proc. Natl. Acad. Sci. U. S. A 107, 10708–10712. [PubMed: 20498078]

Emsley P, Lohkamp B, Scott WG, and Cowtan K (2010). Features and development of Coot. Acta Crystallogr. D. Biol. Crystallogr 66, 486–501. [PubMed: 20383002]

Fettiplace R, and Kim KX (2014). The physiology of mechanoelectrical transduction channels in hearing. Physiol. Rev 94, 951–986. [PubMed: 24987009]

Furness DN, and Hackney CM (1985). Cross-links between stereocilia in the guinea pig cochlea. Hear. Res 18, 177–188. [PubMed: 4044419]

Garcia-Boronat M, Diez-Rivero CM, Reinherz EL, and Reche PA (2008). PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. Nucleic Acids Res. 36, W35–W41. [PubMed: 18442995]

Geng R, Sotomayor M, Kinder KJ, Gopal SR, Gerka-Stuyt J, Chen DH-C, Hardisty-Hughes RE, Ball G, Parker A, Gaudet R, et al. (2013). Noddy, a mouse harboring a missense mutation in protocadherin-15, reveals the impact of disrupting a critical interaction site between tip-link cadherins in inner ear hair cells. J. Neurosci 33, 4395–4404. [PubMed: 23467356]

Gillespie PG, and Müller U (2009). Mechanotransduction by hair cells: models, molecules, and mechanisms. Cell 139, 33–44. [PubMed: 19804752]

Gul IS, Hulpiau P, Saeys Y, and van Roy F (2017). Evolution and diversity of cadherins and catenins. Exp. Cell Res 358, 3–9. [PubMed: 28268172]

Han F, Yu H, Tian C, Chen HE, Benedict-Alderfer C, Zheng Y, Wang Q, Han X, and Zheng QY (2012). A new mouse mutant of the Cdh23 gene with early-onset hearing loss facilitates evaluation of otoprotection drugs. Pharmacogenomics J 12, 30–44. [PubMed: 20644563]

Holt JR, Gillespie SKH, Provance DW, Shah K, Shokat KM, Corey DP, Mercer J. a, and Gillespie PG (2002). A chemical-genetic strategy implicates myosin-1c in adaptation by hair cells. Cell 108, 371–381. [PubMed: 11853671]

Hudspeth AJ (2014). Integrating the active process of hair cells with cochlear function. Nat. Rev. Neurosci 15, 600–614. [PubMed: 25096182]

Hudspeth AJ, and Corey DP (1977). Sensitivity, polarity, and conductance change in the response of vertebrate hair cells to controlled mechanical stimuli. Proc. Natl. Acad. Sci. U. S. A 74, 2407–2411. [PubMed: 329282]

Humphrey W, Dalke A, and Schulten K (1996). VMD: visual molecular dynamics. J Mol Graph 14, 27-28-38.

Indzhykulian AA, Stepanyan R, Nelina A, Spinelli KJ, Ahmed ZM, Belyantseva IA, Friedman TB, Barr-Gillespie PG, and Frolenkov GI (2013). Molecular Remodeling of Tip Links Underlies Mechanosensory Regeneration in Auditory Hair Cells. PLoS Biol 11, e1001583. [PubMed: 23776407]

Jaiganesh A, Narui Y, Araya-Secchi R, and Sotomayor M (2017). Beyond Cell-Cell Adhesion: Sensational Cadherins for Hearing and Balance. Cold Spring Harb. Perspect. Biol a029280.

Kachar B, Parakkal M, Kurc M, Zhao Y. -d., and Gillespie PG (2000). High-resolution structure of hair-cell tip links. Proc. Natl. Acad. Sci 97, 13336–13341. [PubMed: 11087873]

Kazmierczak P, Sakaguchi H, Tokita J, Wilson-Kubalek EM, Milligan RA, Müller U, and Kachar B (2007). Cadherin 23 and protocadherin 15 interact to form tip-link filaments in sensory hair cells. Nature 449, 87–91. [PubMed: 17805295]

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. [PubMed: 22543367]

Lagziel A, Ahmed ZM, Schultz JM, Morell RJ, Belyantseva IA, and Friedman TB (2005). Spatiotemporal pattern and isoforms of cadherin 23 in wild type and waltzer mice during inner ear hair cell development. Dev. Biol 280, 295–306. [PubMed: 15882574]

Lagziel A, Overlack N, Bernstein SL, Morell RJ, Wolfrum U, and Friedman TB (2009). Expression of cadherin 23 isoforms is not conserved: implications for a mouse model of Usher syndrome type 1D. Mol. Vis 15, 1843–1857. [PubMed: 19756182]

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948. [PubMed: 17846036]

Lelli A, Kazmierczak P, Kawashima Y, Müller U, and Holt JR (2010). Development and regeneration of sensory transduction in auditory hair cells requires functional interaction between cadherin-23 and protocadherin-15. J. Neurosci 30, 11259–11269. [PubMed: 20739546]

Lenarduzzi S, Vozzi D, Morgan A, Rubinato E, D'Eustacchio A, Osland TM, Rossi C, Graziano C, Castorina P, Ambrosetti U, et al. (2015). Usher syndrome: an effective sequencing approach to establish a genetic and clinical diagnosis. Hear. Res 320, 18–23. [PubMed: 25575603]

Manji SSM, Miller KA, Williams LH, Andreasen L, Siboe M, Rose E, Bahlo M, Kuiper M, and Dahl H-HM (2011). An ENU-induced mutation of Cdh23 causes congenital hearing loss, but no vestibular dysfunction, in mice. Am. J. Pathol 179, 903–914. [PubMed: 21689626]

Manley GA, and Clack JA (2004). An outline of the evolution of vertebrate hearing organs. In Evolution of the Vertebrate Auditory System, pp. 1–26.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, and Read RJ (2007). Phaser crystallographic software. J. Appl. Crystallogr 40, 658–674. [PubMed: 19461840]

Michalski N, Michel V, Caberlotto E, Lefèvre GM, van Aken AFJ, Tinevez J-Y, Bizard E, Houbron C, Weil D, Hardelin J-P, et al. (2009). Harmonin-b, an actin-binding scaffold protein, is involved in the adaptation of mechanoelectrical transduction by sensory hair cells. Pflugers Arch 459, 115–130. [PubMed: 19756723]

Minor W, Cymborowski M, Otwinowski Z, and Chruszcz M (2006). 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. [PubMed: 16855301]

Miyagawa M, Nishio S, and Usami S (2012). Prevalence and clinical features of hearing loss patients with CDH23 mutations: a large cohort study. PLoS One 7, e40366. [PubMed: 22899989]

Murshudov GN, Skubák P, Lebedev A. a., Pannu NS, Steiner R. a., Nicholls R. a., Winn MD, Long F, and Vagin A. a. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. Sect. D Biol. Crystallogr 67, 355–367. [PubMed: 21460454]

Nicoludis JM, Lau S-Y, Schärfe CPI, Marks DS, Weihofen WA, and Gaudet R (2015). Structure and Sequence Analyses of Clustered Protocadherins Reveal Antiparallel Interactions that Mediate Homophilic Specificity. Structure 23, 2087–2098. [PubMed: 26481813]

Nicoludis JM, Vogt BE, Green AG, Schärfe CP, Marks DS, and Gaudet R (2016). Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1-4. Elife 5.

Noben-Trauth K, Zheng QY, and Johnson KR (2003). Association of cadherin 23 with polygenic inheritance and genetic modification of sensorineural hearing loss. Nat. Genet 35, 21–23.

Otwinowski Z, and Minor W (1997). Processing of X-ray diffraction data. Methods Enzym 276, 307–326.

Pan L, Yan J, Wu L, and Zhang M (2009). Assembling stable hair cell tip link complex via multidentate interactions between harmonin and cadherin 23. Proc. Natl. Acad. Sci. U. S. A 106, 5575–5580. [PubMed: 19297620]

Patel SD, Ciatto C, Chen CP, Bahna F, Rajebhosale M, Arkus N, Schieren I, Jessell TM, Honig B, Price SR, et al. (2006). Type II Cadherin Ectodomain Structures: Implications for Classical Cadherin Specificity. Cell 124, 1255–1268. [PubMed: 16564015]

Pepermans E, and Petit C (2015). The tip-link molecular complex of the auditory mechano-electrical transduction machinery. Hear. Res 330, 10–17. [PubMed: 26049141]

Petit C (2001). Usher syndrome: from genetics to pathogenesis. Annu. Rev. Genomics Hum. Genet 2, 271–297. [PubMed: 11701652]

Pickles JO, Comis SD, and Osborne MP (1984). Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction. Hear. Res 15, 103–112. [PubMed: 6436216]

Powers RE, Gaudet R, and Sotomayor M (2017). A Partial Calcium-Free Linker Confers Flexibility to Inner-Ear Protocadherin-15. Structure 25, 482–495. [PubMed: 28238533]

Rzadzinska AK, and Steel KP (2009). Presence of interstereocilial links in waltzer mutants suggests Cdh23 is not essential for tip Link formation. Neuroscience 158, 365–368. [PubMed: 18996172]

Sahly I, Dufour E, Schietroma C, Michel V, Bahloul A, Perfettini I, Pepermans E, Estivalet A, Carette D, Aghaie A, et al. (2012). Localization of Usher 1 proteins to the photoreceptor calyceal processes, which are absent from mice. J. Cell Biol 199, 381–399. [PubMed: 23045546]

Sakaguchi H, Tokita J, Müller U, and Kachar B (2009). Tip links in hair cells: molecular composition and role in hearing loss. Curr. Opin. Otolaryngol. Head Neck Surg 17, 388–393. [PubMed: 19633555]

Salt AN, Inamura N, Thalmann R, and Vora A (1989). Calcium gradients in inner ear endolymph. Am. J. Otolaryngol 10, 371–375. [PubMed: 2596623]

Schrödinger L (2015). The PyMOL Molecular Graphics System, Version-1.8

Schultz JM, Yang Y, Caride AJ, Filoteo AG, Penheiter AR, Lagziel A, Morell RJ, Mohiddin SA, Fananapazir L, Madeo AC, et al. (2005). Modification of human hearing loss by plasma-membrane calcium pump PMCA2. N. Engl. J. Med 352, 1557–1564. [PubMed: 15829536]

Schultz JM, Bhatti R, Madeo AC, Turriff A, Muskett JA, Zalewski CK, King KA, Ahmed ZM, Riazuddin S, Ahmad N, et al. (2011). Allelic hierarchy of CDH23 mutations causing non-syndromic deafness DFNB12 or Usher syndrome USH1D in compound heterozygotes. J. Med. Genet 48, 767–775. [PubMed: 21940737]

Schwander M, Xiong W, Tokita J, Lelli A, Elledge HM, Kazmierczak P, Sczaniecka A, Kolatkar A, Wiltshire T, Kuhn P, et al. (2009). A mouse model for nonsyndromic deafness (DFNB12) links hearing loss to defects in tip links of mechanosensory hair cells. Proc. Natl. Acad. Sci. U. S. A 106, 5252–5257. [PubMed: 19270079]

Siemens J, Lillo C, Dumont R. a, Reynolds A, Williams DS, Gillespie PG, and Müller U (2004). Cadherin 23 is a component of the tip link in hair-cell stereocilia. Nature 428, 950–955. [PubMed: 15057245]

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol 7, 539. [PubMed: 21988835]

Söllner C, Rauch G-J, Siemens J, Geisler R, Schuster SC, the Tübingen 2000 Screen Consortium, Müller U, and Nicolson T (2004). Mutations in cadherin 23 affect tip links in zebrafish sensory hair cells. Nature 428, 955–959. [PubMed: 15057246]

Sotomayor M, Weihofen WA, Gaudet R, and Corey DP (2010). Structural Determinants of Cadherin-23 Function in Hearing and Deafness. Neuron 66, 85–100. [PubMed: 20399731]

Sotomayor M, Weihofen WA, Gaudet R, and Corey DP (2012). Structure of a force-conveying cadherin bond essential for inner-ear mechanotransduction. Nature 492, 128–132. [PubMed: 23135401]

Takahashi S, Mui VJ, Rosenberg SK, Homma K, Cheatham MA, and Zheng J (2016). Cadherin 23-C Regulates Microtubule Networks by Modifying CAMSAP3’s Function. Sci. Rep 6, 28706. [PubMed: 27349180]

Tsukasaki Y, Miyazaki N, Matsumoto A, Nagae S, Yonemura S, Tanoue T, Iwasaki K, and Takeichi M (2014). Giant cadherins Fat and Dachsous self-bend to organize properly spaced intercellular junctions. Proc. Natl. Acad. Sci 111, 16011–16016. [PubMed: 25355906]

Tsuprun V, Goodyear RJ, and Richardson GP (2004). The Structure of Tip Links and Kinocilial Links in Avian Sensory Hair Bundles. Biophys. J 87, 4106–4112. [PubMed: 15377520]

Vester-Christensen MB, Halim A, Joshi HJ, Steentoft C, Bennett EP, Levery SB, Vakhrushev SY, and Clausen H (2013). Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. Proc. Natl. Acad. Sci. U. S. A 110, 21018–21023. [PubMed: 24101494]

Waterhouse AM, Procter JB, Martin DMA, Clamp M, and Barton GJ (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191. [PubMed: 19151095]

Winn MD, Isupov MN, and Murshudov GN (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr. Sect. D Biol. Crystallogr 57, 122–133. [PubMed: 11134934]

Wu L, Pan L, Zhang C, and Zhang M (2012). Large protein assemblies formed by multivalent interactions between cadherin23 and harmonin suggest a stable anchorage structure at the tip link of stereocilia. J. Biol. Chem 287, 33460–33471. [PubMed: 22879593]

Xu Z, Peng AW, Oshima K, and Heller S (2008). MAGI-1, A Candidate Stereociliary Scaffolding Protein, Associates with the Tip-Link Component Cadherin 23. J. Neurosci 28, 11269–11276. [PubMed: 18971469]

Yan J, Pan L, Chen X, Wu L, and Zhang M (2010). The structure of the harmonin/sans complex reveals an unexpected interaction mode of the two Usher syndrome proteins. Proc. Natl. Acad. Sci 107, 4040–4045. [PubMed: 20142502]

Zielinska DF, Gnad F, Wiśniewski JR, and Mann M (2010). Precision Mapping of an In Vivo N-Glycoproteome Reveals Rigid Topological and Sequence Constraints. Cell 141, 897–907. [PubMed: 20510933]

HIGHLIGHTS

- Structures of 18 CDH23 extracellular cadherin (EC) “repeats” are presented.
- Non-canonical linker regions are at EC2-3, EC12-13, EC20-21, EC21-22, and EC24-25.
- The CDH23 EC17-25 fragment shows inherent helicity and a possible dimer interface.
- CDH23 missense mutations may affect structure and function by 5 distinct mechanisms.

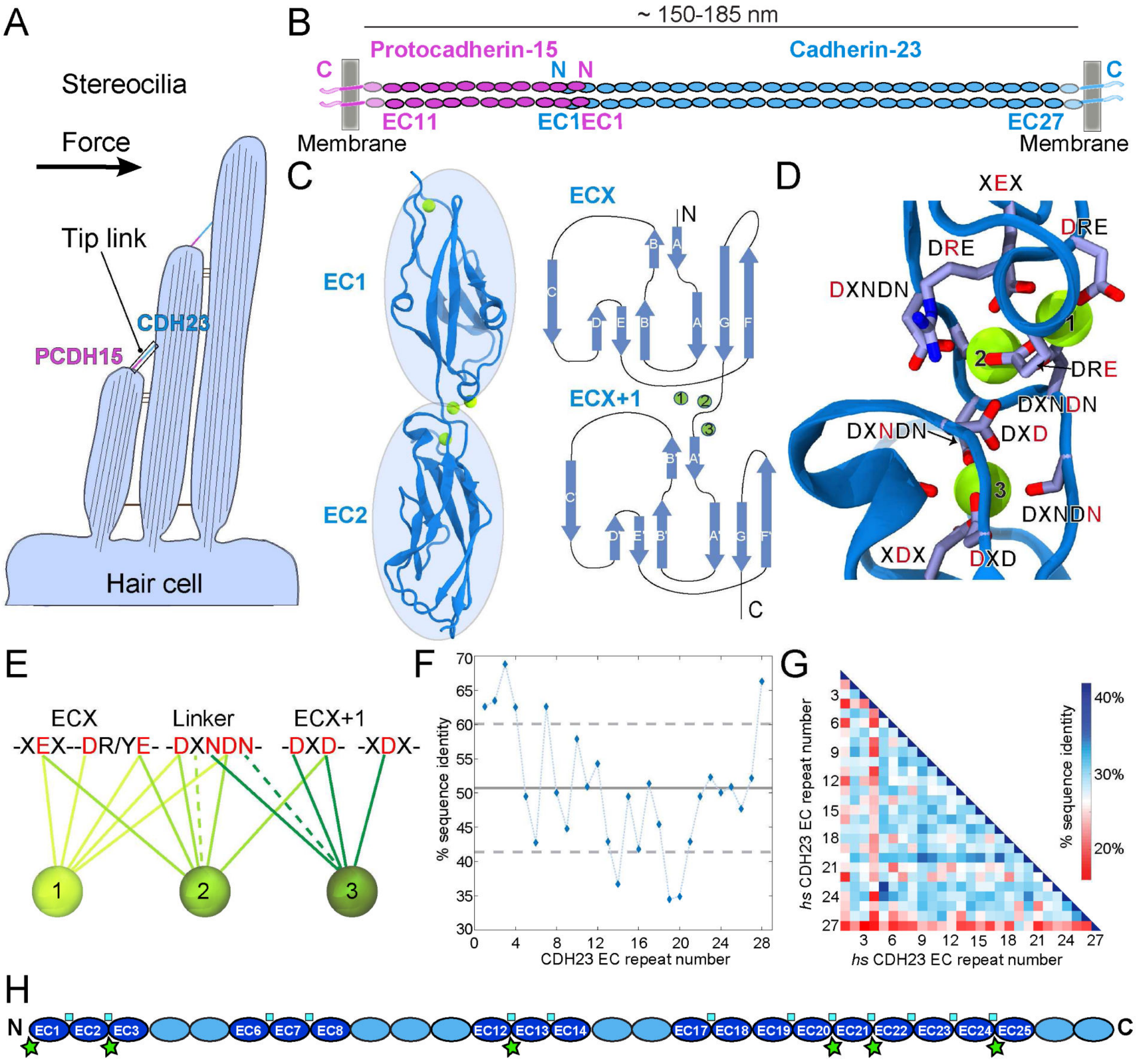

Figure 1. Hair cells, tip links, and CDH23 architecture.

(A) Schematic representation of a cochlear hair-cell stereocilia bundle highlighting the tip link location. (B) The tip link, formed by the tip-to-tip (*trans)* interaction of PCDH15 and CDH23. Both proteins are thought to form parallel (*cis)* dimers. (C) Ribbon diagram of *Danio rerio* (*dr*) Cdh23 EC1-2 showing Ca^{2+}-binding sites and the 7 β-strand cadherin fold (left). Topology diagram shows generic cadherin secondary structure (right). (D) Canonical Ca^{2+}-binding sites at the EC1-2 linker region of *dr* Cdh23 with conserved Ca^{2+}-chelating residues in motifs highlighted. (E) Schematic representation of a Ca^{2+}-binding motif between cadherin EC repeats. Solid lines show side chain coordination and dotted lines show backbone coordination. (F) Plot of percent sequence identity of each CDH23 EC repeat among up to 20 species against EC repeat number (Data S1). The average sequence identity is 51 ± 9%. (G) Heat map of a sequence identity matrix among EC-repeats for *Homo sapiens* (*hs*) CDH23. (H) Schematic representation of all CDH23 EC repeats. Dark blue ovals indicate structures presented in this work (cyan squares highlight structurally resolved linker regions); green stars indicate unusual Ca^{2+}-binding sites. See also Figure S1, Table S1, and Data S1.

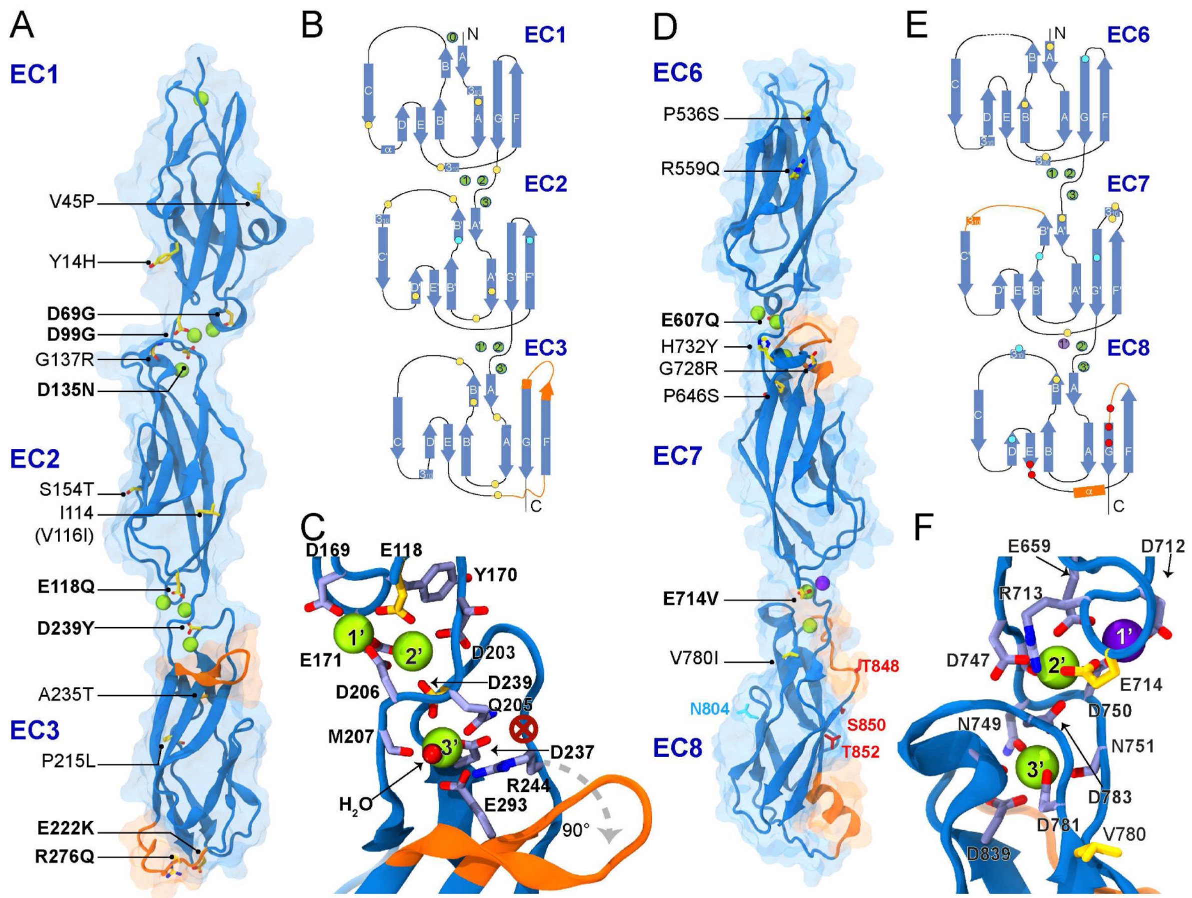

Figure 2. Structures of *Danio rerio* (*dr*) Cdh23 EC1-3 and *Homo sapiens* (*hs*) CDH23 EC6-8.

(A) Ribbon diagram of *dr* Cdh23 EC1-3 in blue with unusual features highlighted in orange and with a transparent molecular surface representation. Disease-related mutations are in yellow sticks and listed (bolded residues are at conserved Ca^{2+}-binding motifs). Ca^{2+} ions are in green. (B) Topology diagram of *dr* Cdh23 EC1-3. Yellow circles mark locations of disease-related mutations, cyan circles are N-glycosylation sites, and unusual features on the structure are in orange. (C) The non-canonical linker region between EC2 and EC3. Residues directly binding Ca^{2+} are in grey-blue sticks. Residues implicated in deafness are in yellow sticks. Red cross indicates missing backbone coordination from BC loop. (D) Ribbon diagram of *hs* CDH23 EC6-8 as in A. A sodium ion at the EC7-8 linker region is in purple. (E) Topology diagram of *hs* CDH23 EC6-8 as in B. Experimentally verified and predicted O-glycosylation sites are in red (Vester-Christensen et al., 2013). Dashed lines denote missing residues. (F) Detail of the EC7-8 linker region with the location of the *salsa* mutation (p.E714V) (Schwander et al., 2009). See also Figures S2A and S4A.

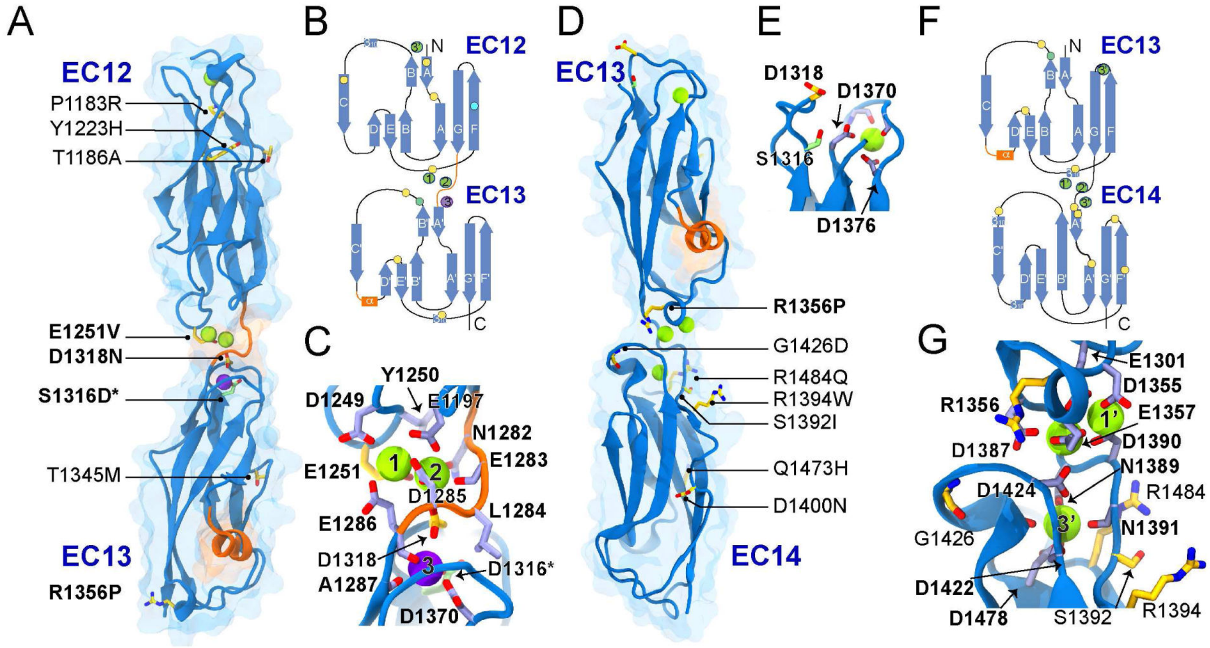

Figure 3. Structures of *Mus musculus* (*mm*) CDH23 EC12-13 S1316D and *Homo sapiens* (*hs*) CDH23 EC13-14.

(A) Structure of *mm* CDH23 EC12-13 as in Figure 2A. The engineered p.S1316D mutation is in lime green. (B) Topology diagram of *mm* CDH23 EC12-13 (S1316D) as in Figure 2B. Location of an engineered mutation is marked by a lime green circle. (C) Non-canonical *mm* CDH23 EC12-13 linker region. Engineered mutation p.S1316D is in lime green and highlighted with an asterisk. (D) Structure of *hs* CDH23 EC13-14. (E) Partial Ca^{2+}-binding site at the top of EC13. Residue p.S1316 is highlighted in lime green. (F) Topology diagram of *hs* CDH23 EC13-14. (G) Detail of the EC13-14 linker region as in C. See also Figures S2C and S4B.

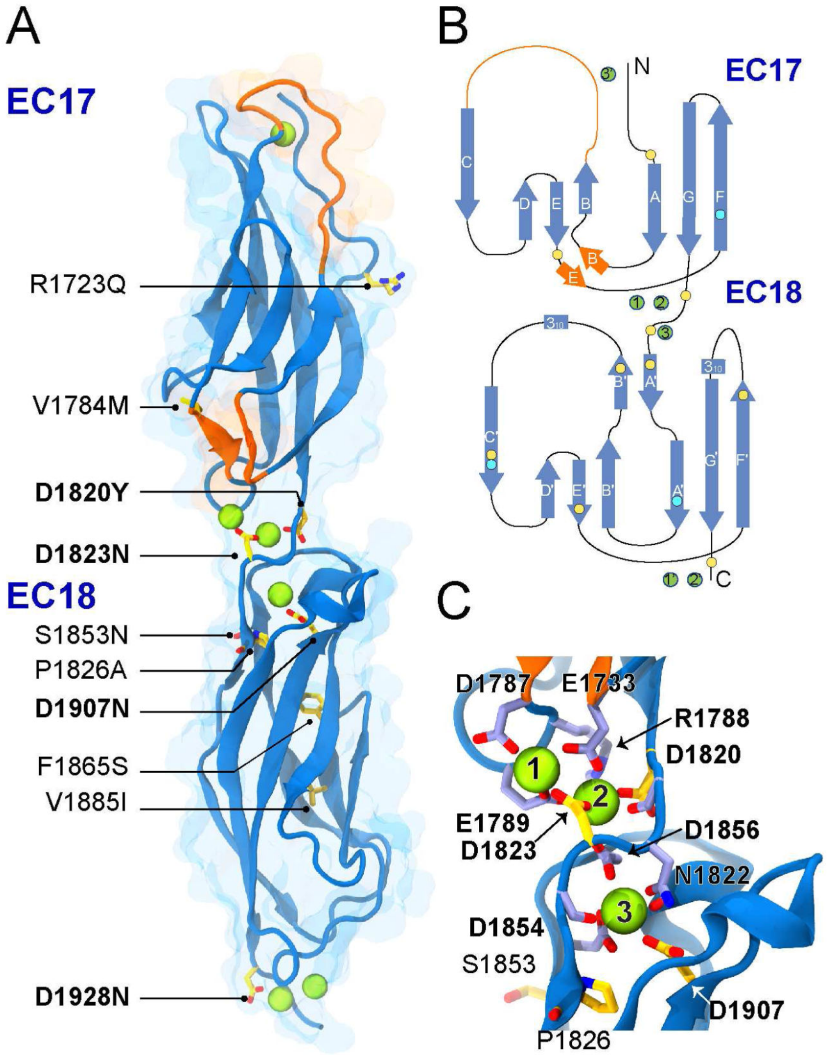

Figure 4. Structure of *Mus musculus* (*mm*) CDH23 EC17-18.

(A) Structure and disease mutations of *mm* CDH23 EC17-18 as in Figure 2A. (B) Topology diagram of *mm* CDH23 EC17-18 as in Figure 2B. (C) Detail of the EC17-18 linker region as in Figure 2C. See also Figures S8A and S8B.

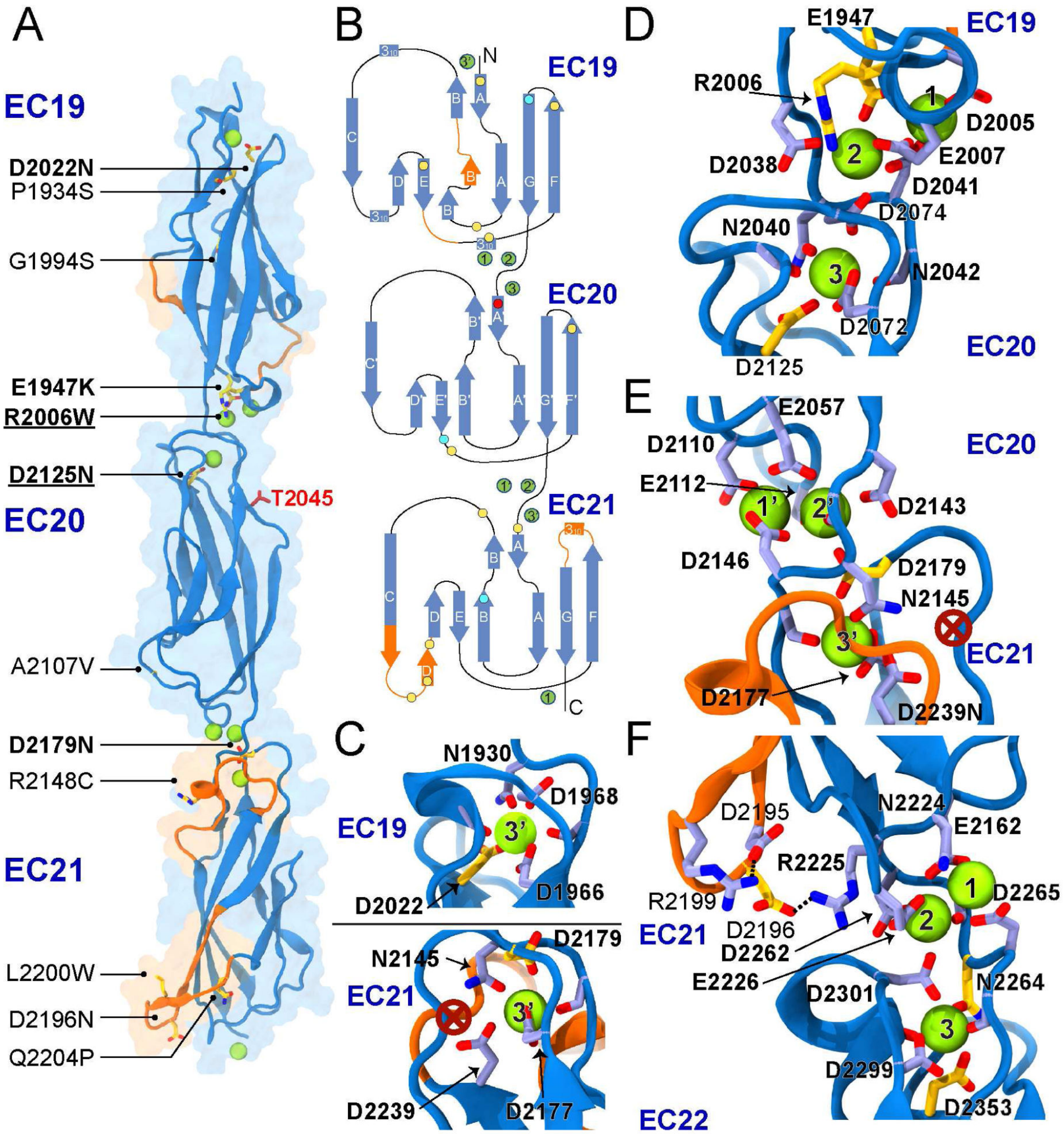

Figure 5. Structural features of *Mus musculus* (*mm*) CDH23 EC19-21 and an unusual EC21-22 linker region.

(A) Structure of *mm* CDH23 EC19-21 as in Figure 2A. Underlined residues indicate mutations for which structures are presented in this work (Figure 7). (B) Topology diagram of *mm* CDH23 EC19-21 as in Figure 2B. (C) Canonical Ca^{2+} binding at the top of EC19 (top) and non-canonical Ca^{2+} binding at the top of EC21 (bottom) where a backbone carbonyl from the BC loop group is missing in the coordination sphere (red 'X'). (D-F) Details of linker regions of *mm* CDH23 EC19-20 and EC20-21, and of *hs* CDH23 EC21-22. The EC21-22 linker region is unique due to a long CD-loop from EC21. Residues p.D2195-p.R2199 and p.D2196-p.R2225 form salt bridges. See also Figure S2B.

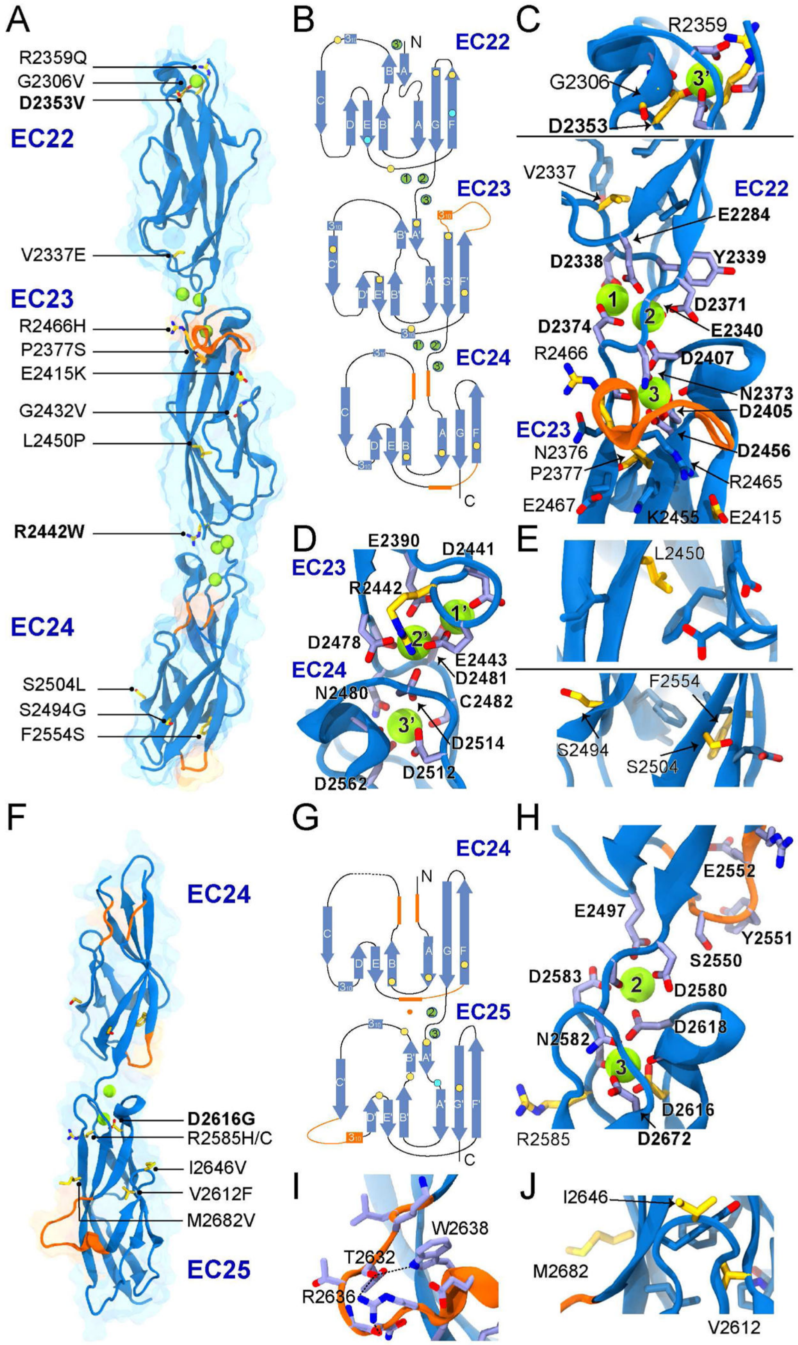

Figure 6. Structures of *Mus musculus* (*mm*) CDH23 EC22-24 and EC24-25.

(A) Structure of *mm* CDH23 EC22-24 as in Figure 2A. (B) Topology diagram of *mm* CDH23 EC22-24 as in Figure 2B. (C) Ca^{2+}-binding site at the top of EC22 (top). Details of the EC22-23 linker region (bottom), including the V2337E *jera* mutation (Manji et al., 2011), as in Figure 2C. (D) The EC23-24 linker region. (E) Location of selected inward and outward facing mutations in EC23 (top) and EC24 (bottom). (F) Ribbon diagram of *mm* CDH23 EC24-25 as in A. (G) Topology of *mm* CDH23 EC24-25 as in B. (H) Non-canonical EC24-25 linker region with missing DRE motif (replaced by QSYE). (I) Unusual CD loop on EC25. (J) Detail of disease mutation sites on EC25. See also Figures S2D and S8C, S8D, S8E, and S8F.

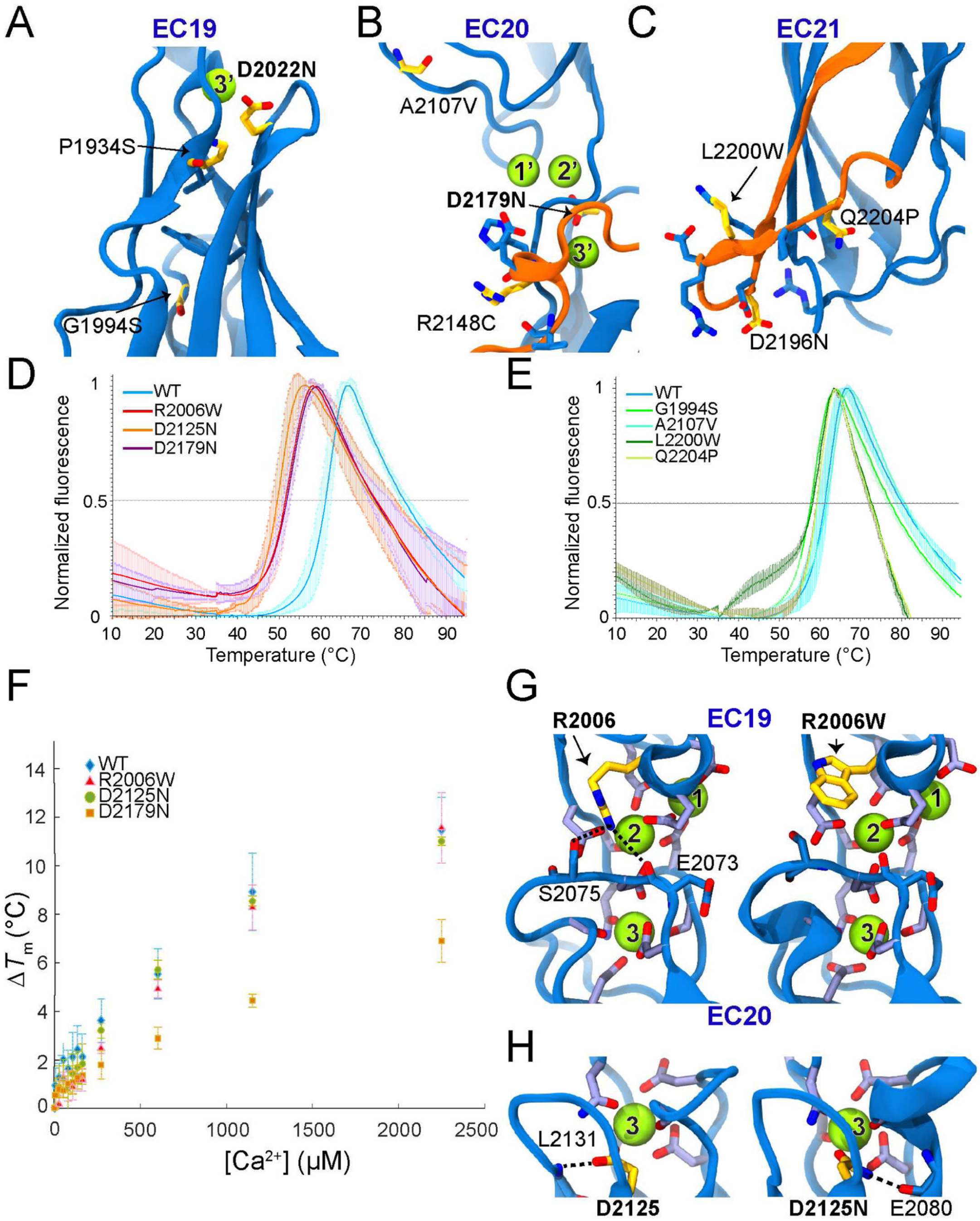

Figure 7. Stability and structures of mutant *Mus musculus* (*mm*) CDH23 EC19-21 fragments.
(A-C) Details showing disease mutation sites (yellow sticks) on EC19, EC20, and EC21 respectively. (D) Thermal melting of WT *mm* CDH23 EC19-21 and mutants R2006W, D2125N, and D2179N at conserved Ca^{2+}-binding motifs. Intersection with grey line defines melting temperature (T_m). Shadow of the melting trace denotes standard deviation. (E) Thermal melting of WT *mm* CDH23 EC19-21 (same as in D) and mutants away from Ca^{2+}-binding sites (Table 2). (F) Shift in melting temperature (ΔT_m) as a function of Ca^{2+} concentration for WT *mm* CDH23 EC19-21 and mutants R2006W, D2125N, and D2179N

(*n* = 3 for each). Error bars are standard deviation. (G) Detail of WT (left) and R2006W (right) EC19-20 linker regions. (H) Detail of WT (left) and D2125N (right) EC19-20 linker regions. See also Figure S3.

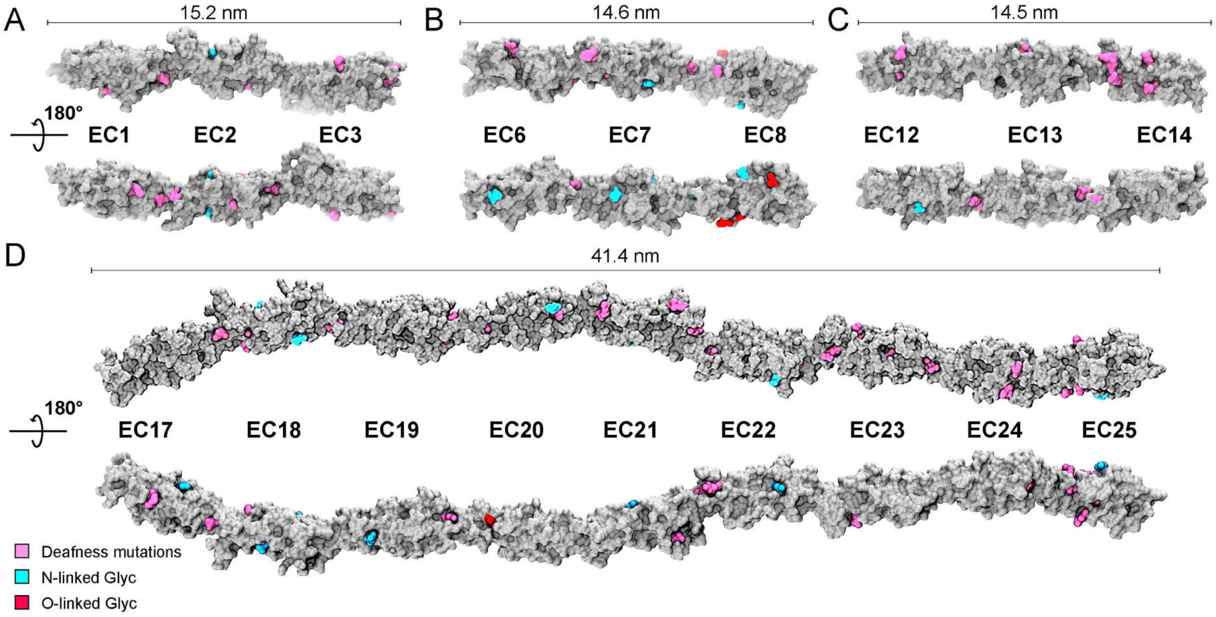

Figure 8. Structural features of long CDH23 fragments.

(A-C) Surface representation of *Danio rerio Cdh23* EC1-3, *Homo sapiens* CDH23 EC6-8, and *Mus musculus* (*mm*) CDH23 EC12-14. Disease mutation sites are highlighted in pink, N-linked glycosylation in cyan, and O-linked glycosylation in red. Top and bottom panels are 180° rotated from each other. (D) Surface representation of *mm* CDH23 EC17-25 obtained by fitting overlapping structures and shown as in A-C. Disease mutations and glycosylation sites tend to cluster on opposing surfaces. See also Figures S5, S6, and S7.

Table 1.

Statistics for CDH23 Structures

	***dr* Cdh23 EC1-3**	***hs* CDH23 EC6-8**	***mm* CDH23 EC7-8**	***mm* CDH23 EC12-13 (S1316D)**	***hs* CDH23 EC13-14**
Data Collection					
Space group	$P2_1$	$P3_212$	$P4_132$	P1	$P6_222$
Unit cell parameters					
a, b, c (Å)	178.1, 115.6, 67.9	93.4, 93.4, 122.4	185.8, 185.8, 185.8	43.7, 47.2, 113.0	74.7, 74.7, 213.5
α, β, γ (°)	90.0, 89.9, 90.0	90.0, 90.0, 120.0	90.0, 90.0, 90.0	89.9, 90.0, 115.8	90.0, 90.0, 120.0
Molecules per a.u.	4	1	2	4	1
Beam source	APS-24-ID-E	APS-24-ID-E	APS-24-ID-C	MicroMax-003	APS-24-ID-C
Wavelength (Å)	0.9792	0.9792	0.9792	1.542	0.9792
Resolution limit (Å)	2.65	2.92	3.56	2.84	1.86
Unique reflections	80,843	13,518	13,716	19,387	30,738
Completeness (%)	99.4 (99.2)	99.9 (99.9)	100.0 (100.0)	93.9 (89.3)	94.4 (83.3)
Redundancy	3.3 (3.1)	4.2 (4.1)	35.5 (23.7)	1.6 (1.6)	10.2 (7.1)
$I/\sigma(I)$	7.1 (2.0)	17.1 (2.0)	20.0 (2.5)	4.3 (2.1)	26.8 (2.0)
R_{merge}	0.16 (0.69)	0.08 (0.59)	0.36 (1.59)	0.17 (0.38)	0.08 (0.54)
R_{meas}	0.20 (0.82)	0.09 (0.68)	0.37 (1.62)	0.23 (0.53)	0.08 (0.58)
R_{pim}	0.11 (0.45)	0.04 (0.34)	0.06 (0.33)	0.17 (0.38)	0.02 (0.20)
$CC_{1/2}$	0.92 (0.70)	0.97 (0.88)	0.95 (0.75)	0.88 (0.72)	0.99 (0.91)
CC*	0.98 (0.91)	0.99 (0.97)	0.99 (0.93)	0.97 (0.92)	1.00 (0.98)
Refinement					
Resolution (Å)	178.05–2.65 (2.70–2.65)	80.89–2.92 (3.00–2.92)	131.39–3.56 (3.66–3.56)	33.99–2.84 (2.89–2.84)	71.17–1.86 (1.91–1.86)
R_{work} (%)	19.2 (31.8)	19.8 (33.3)	19.2 (31.2)	23.8 (27.4)	21.7 (30.7)
R_{free} (%)	22.7 (33.2)	24.6 (35.9)	22.1 (38.1)	27.2 (31.7)	24.8 (30.2)
Residues (atoms)	1,262 (9,891)	320 (2,479)	400 (3,084)	828 (6,376)	210 (1,632)
Water molecules	646	24	3	0	133
RMSDs					
Bond lengths (Å)	0.013	0.009	0.013	0.011	0.010
Bond angles (°)	1.464	1.428	1.365	1.475	1.500
B-factor average					
Protein	29.34	89.57	91.55	24.39	46.96
Ligand/ion	31.35	76.45	73.57	23.08	48.51
Water	22.39	58.55	42.46	-	48.77
Ramachandran plot					
Most favored (%)	89.6	88.8	89.0	89.8	90.9
Additionally allowed (%)	10.1	11.2	11.0	10.2	8.6
Generously allowed (%)	0.3	0.0	0.0	0.0	0.5
Disallowed (%)	0.0	0.0	0.0	0.0	0.0

	***dr* Cdh23 EC1-3**	***hs* CDH23 EC6-8**	***mm* CDH23 EC7-8**	***mm* CDH23 EC12-13 (S1316D)**	***hs* CDH23 EC13-14**
PDB ID	5W4T	5TFM	5TFL	5VH2	5WJ8

	***mm* CDH23 EC17-18**	***mm* CDH23 EC19-21**	***hs* CDH23 EC21-23**	***mm* CDH23 EC22-24**
Data Collection				
Space group	$P4_32_12$	C2	C2	C2
Unit cell parameters				
a, b, c (Å)	122.5, 122.5, 43.6	77.4, 66.9, 112.8	116.7, 166.8, 47.9	89.5, 57.2, 114.7
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 98.7, 90.0	90.0, 102.5, 90.0	90.0, 112.6, 90.0
Molecules per a.u.	1	1	2	1
Beam source	APS-24-ID-E	APS-24-ID-E	MicroMax-003	APS-24-ID-E
Wavelength (Å)	0.9792	0.9792	1.542	0.9792
Resolution limit (Å)	2.90	2.81	3.54	1.85
Unique reflections	7,803	14,163	10,856	45,949
Completeness (%)	100.0 (99.5)	97.6 (85.4)	99.4 (99.6)	99.5 (97.9)
Redundancy	6.0 (4.6)	3.6 (3.2)	2.9 (2.7)	3.3 (2.8)
$I/\sigma(I)$	7.2 (1.7)	17.3 (2.0)	5.95 (2.2)	12.8 (1.75)
R_{merge}	0.25 (0.83)	0.07 (0.50)	0.16 (0.43)	0.09 (0.66)
R_{meas}	0.27 (0.94)	0.08 (0.60)	0.20 (0.53)	0.11 (0.81)
R_{pim}	0.11 (0.42)	0.04 (0.32)	0.11 (0.31)	0.06 (0.46)
$CC_{1/2}$	0.91 (0.62)	0.99 (0.92)	0.92 (0.77)	0.94 (0.70)
CC*	0.97 (0.88)	1.00 (0.98)	0.98 (0.93)	0.98 (0.91)
Refinement				
Resolution (Å)	86.60–2.90 (2.98–2.90)	111.51–2.81 (2.87–2.81)	94.08–3.54 (3.63–3.54)	50.0–1.85 (1.90–1.85)
R_{work} (%)	20.0 (31.1)	20.1 (36.7)	30.3 (46.3)	17.1 (26.6)
R_{free} (%)	26.6 (45.4)	24.4 (37.3)	34.6 (52.2)	18.8 (29.2)
Residues (atoms)	216 (1,692)	335 (2,566)	612 (4,658)	317 (2,527)
Water molecules	23	10	-	541
RMSDs				
Bond lengths (Å)	0.012	0.010	0.010	0.011
Bond angles (°)	1.512	1.546	1.349	1.496
B-factor average				
Protein	28.58	82.99	68.58	22.86
Ligand/ion	33.28	80.79	50.51	23.71
Water	12.90	55.92	-	33.52
Ramachandran plot				
Most favored (%)	85.3	88.4	85.0	92.7
Additionally allowed (%)	14.7	11.3	14.6	6.9
Generously allowed (%)	0.0	0.3	0.4	0.4
Disallowed (%)	0.0	0.0	0.0	0.0
PDB ID	5WJM	5TFK	5VVM	5UZ8

	***mm* CDH23 EC24-25**	***mm* CDH23 EC19-21 (R2006W)**	***mm* CDH23 EC19-21 (S2064P-D2125N)**	***mm* CDH23 EC19-21 (S2064P)**
Data Collection				
Space group	$P2_1$	C2	C2	C2
Unit cell parameters				
a, b, c (Å)	27.6, 99.4, 157.4	189.5, 73.9, 49.4	84.2, 65.4, 114.7	77.7, 67.8, 112.8
α, β, γ (°)	90.0, 90.1,90.0	90.0, 96.6, 90.0	90.0, 98.6, 90.0	90.0, 99.1,90.0
Molecules per a.u.	4	1	1	1
Beam source	APS-24-ID-E	MicroMax-003	APS-24-ID-C	APS-24-ID-E
Wavelength (Å)	0.9792	1.542	0.9792	0.9792
Resolution limit (Å)	2.92	2.96	2.85	2.69
Unique reflections	18,772	14,294	14,561	16,091
Completeness (%)	94.3 (93.8)	91.5 (95.3)	95.7 (77.3)	98.4 (94.4)
Redundancy	1.9 (1.9)	3.8 (3.6)	3.1 (2.3)	2.5 (2.5)
$I/\sigma(I)$	6.4 (1.8)	7.9 (2.0)	15.7(3.4)	16.4(2.0)
R_{merge}	0.11 (0.42)	0.16 (0.55)	0.07 (0.18)	0.06 (0.45)
R_{meas}	0.15 (0.57)	0.19 (0.63)	0.08 (0.22)	0.07 (0.57)
R_{pim}	0.10 (0.38)	0.09 (0.29)	0.04 (0.13)	0.04 (0.34)
$CC_{1/2}$	0.93 (0.77)	0.97 (0.88)	0.99 (0.97)	0.98 (0.86)
CC*	0.98 (0.93)	0.99 (0.97)	1.00 (0.99)	0.99 (0.96)
Refinement				
Resolution (Å)	47.39–2.92 (2.98–2.92)	50.00–2.96 (3.04–2.96)	30.34–2.85 (2.92–2.85)	111.36–2.69 (2.76–2.69)
R_{work} (%)	20.8 (24.5)	19.6 (35.6)	19.5 (37.2)	22.1 (41.0)
R_{free} (%)	22.1 (39.8)	23.6 (34.8)	24.4 (35.9)	25.3 (45.6)
Residues (atoms)	821 (6,432)	335 (2,571)	320 (2,443)	335 (2,562)
Water molecules	-	59	7	9
RMSDs				
Bond lengths (Å)	0.010	0.010	0.011	0.011
Bond angles (°)	1.438	1.440	1.456	1.446
B-factor average				
Protein	49.14	49.84	106.03	76.19
Ligand/ion	44.47	42.46	87.82	71.47
Water	-	29.86	80.73	45.74
Ramachandran plot				
Most favored (%)	87.4	87.8	86.3	87.4
Additionally allowed (%)	12.5	12.2	13.0	11.6
Generously allowed (%)	0.1	0.0	0.7	1.0
Disallowed (%)	0.0	0.0	0.0	0.0
PDB ID	5VT8	5UN2	5ULU	5I8D

Table 2.

Melting temperature for *mm* CDH23 EC19-21 fragments

CONSTRUCT	T_m [*]	*n*
WT	61.8 ± 1.0 °C	8
p.G1994S	57.7 ± 0.1 °C	4
p.R2006W	51.9 ± 0.6 °C	8
p.A2107V	61.7 ± 0.1 °C	4
p.D2125N	49.4 ± 1.2 °C	4
p.D2179N	53.0 ± 0.7 °C	8
p.L2200W	58.3 ± 0.3 °C	3
p.Q2204P	58.9 ± 0.5 °C	3
WT (Ca^{2+}-free) [**]	45.1 ± 1.1 °C	3
p.R2006W (Ca^{2+}-free)	41.0 ± 2.2 °C	3
p.D2125N (Ca^{2+}-free)	40.6 ± 0.8°C	3
p.D2179N (Ca^{2+}-free)	45.5 ± 1.0°C	3

[*] WT ~ A2107V ~ Q2204P ~ L2200W ~ G1994S > D2179N ~ R2006W ~ D2125N

[**] When Ca^{2+} was added to the EGTA-treated, Ca^{2+}-free WT CDH23 EC19-21, T_m never reached back values of the fully saturated fragment.