

# A Mechanically Weak Extracellular Membrane-Adjacent Domain Induces Dimerization of Protocadherin-15

Pedro De-la-Torre,<sup>1</sup> Deepanshu Choudhary,<sup>1</sup> Raul Araya-Secchi,<sup>1,2</sup> Yoshie Narui,<sup>1</sup> and Marcos Sotomayor<sup>1,\*</sup>

<sup>1</sup>Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio and <sup>2</sup>Structural Biophysics, Section for Neutron and X-ray Science, Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

**ABSTRACT** The cadherin superfamily of proteins is defined by the presence of extracellular cadherin (EC) “repeats” that engage in protein-protein interactions to mediate cell-cell adhesion, cell signaling, and mechanotransduction. The extracellular domains of nonclassical cadherins often have a large number of EC repeats along with other subdomains of various folds. Protocadherin-15 (PCDH15), a protein component of the inner-ear tip link filament essential for mechanotransduction, has 11 EC repeats and a membrane adjacent domain (MAD12) of atypical fold. Here we report the crystal structure of a pig PCDH15 fragment including EC10, EC11, and MAD12 in a parallel dimeric arrangement. MAD12 has a unique molecular architecture and folds as a ferredoxin-like domain similar to that found in the nucleoporin protein Nup54. Analytical ultracentrifugation experiments along with size-exclusion chromatography coupled to multiangle laser light scattering and small-angle x-ray scattering corroborate the crystallographic dimer and show that MAD12 induces parallel dimerization of PCDH15 near its membrane insertion point. In addition, steered molecular dynamics simulations suggest that MAD12 is mechanically weak and may unfold before tip-link rupture. Sequence analyses and structural modeling predict the existence of similar domains in cadherin-23, protocadherin-24, and the “giant” FAT and CELSR cadherins, indicating that some of them may also exhibit MAD-induced parallel dimerization.

## INTRODUCTION

The cadherin superfamily of proteins can be broadly divided into three main subfamilies that include the classical cadherins (type I, II, and desmosomal), the clustered protocadherins ( $\alpha$ ,  $\beta$ , and  $\gamma$ ), and the nonclassical, nonclustered cadherins ( $\delta 1$  and  $\delta 2$  protocadherins along with various other atypical cadherins) (1–4). All of them feature two or more tandem extracellular cadherin (EC) “repeats” of similar fold (Greek key motifs with seven  $\beta$  strands) but distinct residue sequence (5–8). These EC repeats are often involved in protein-protein interactions essential for the diverse functions played by cadherins in cell-cell adhesion (2,9,10), signaling (11–15), and mechanotransduction (16–18).

The vertebrate classical cadherins were discovered as glycoproteins involved in calcium ( $\text{Ca}^{2+}$ )-dependent cellular adhesion (19,20), and the mechanism by which they facilitate adhesion is well understood (2). Their processed extra-

cellular domains (21), with five EC repeats, protrude from the surface of adjacent cells to engage in mostly homophilic (same type) binding across the cells (*trans*) (2). The individual *trans* bonds linking adjacent cells are weak, with the strength of cell-cell adhesion stemming from multiple *trans* bonds supported by *cis* interactions between extracellular domains protruding from the same cell (22,23). Biophysical and structural studies of classical cadherins suggest that their *trans* adhesive interactions are mediated by EC1 to EC1 contacts, whereas *cis* interactions of classical type I cadherins involve EC1 of one molecule interacting with EC2 of the neighboring molecule (23–31). In contrast, clustered protocadherins, with six EC repeats, use an antiparallel binding interface involving repeats EC1 to EC4 for adhesion (32–36), with promiscuous *cis* interactions mediated by the EC repeats closest to the membrane (EC5 and EC6) (37). The  $\delta 2$  protocadherins seem to use a similar mechanism for *trans* interactions with EC1–4 antiparallel interfaces (38), yet it is unclear if the  $\delta$  protocadherins engage in clustered-protocadherin-like *cis* interactions.

Less is known about binding mechanisms and oligomerization states of nonclassical, nonclustered, atypical

Submitted September 17, 2018, and accepted for publication November 5, 2018.

\*Correspondence: [sotomayor.8@osu.edu](mailto:sotomayor.8@osu.edu)

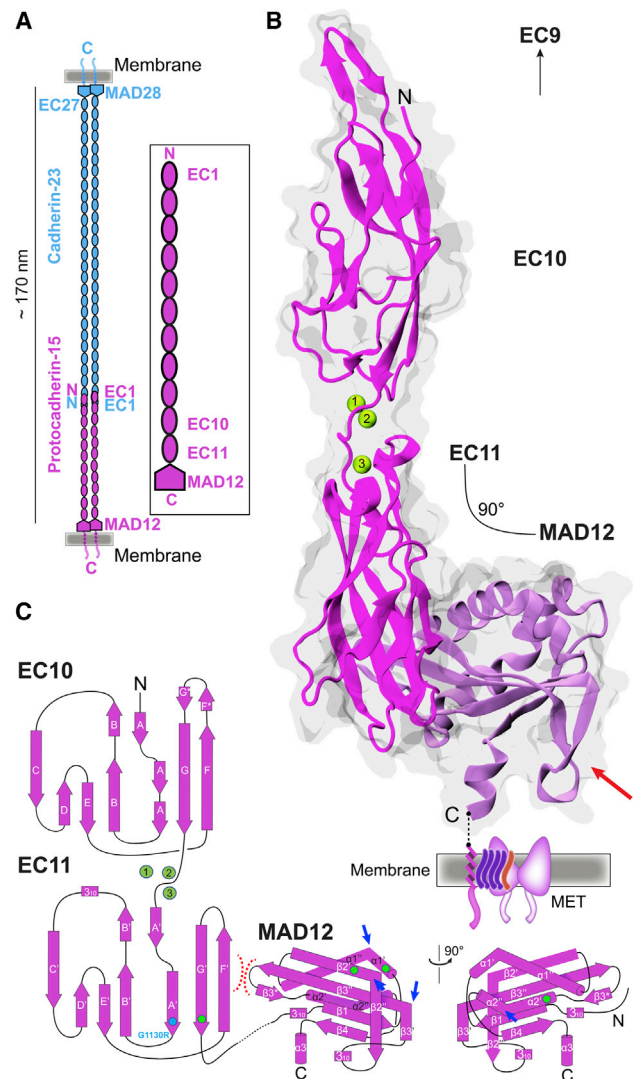
Editor: James Cole.

<https://doi.org/10.1016/j.bpj.2018.11.010>

© 2018 Biophysical Society.

cadherins, especially those that have enormous extracellular domains, such as protocadherin-15 (PCDH15) and cadherin-23 (CDH23), protocadherin-24 (PCDH24), the FAT cadherins, and CELSRs. The PCDH15 and CDH23 proteins, involved in hereditary deafness and blindness (18,39–41), use their EC1-2 tips to engage in an antiparallel, heterophilic *trans* “handshake” interaction essential for inner-ear function (42–44). Electron microscopy of negatively stained extracellular domains also suggests that their full-length extracellular domains form parallel *cis* homodimers (42). Although most PCDH15 and CDH23 fragments with up to four EC repeats have failed to conclusively show bona fide homophilic *cis* interactions (45–47), two recent structures show that PCDH15 can dimerize at EC2-3 (48) and at its C-terminal domain near the membrane (49). Less is known about the specific contacts leading to parallel dimerization of CDH23. Similarly, PCDH24 forms *trans* heterophilic bonds with the mucin-like cadherin CDHR5, and also forms *trans* homophilic bonds with itself (50). The PCDH24-CDHR5 interactions are essential for intermicrovillar link formation and important for brush-border function (51), but little is known about the EC repeats involved or about PCDH24’s oligomerization states. Further, the giant cadherins FAT4 and DCHS1 have been shown to self bend and form heterophilic complexes involved in cell signaling, with repeats EC1 to EC4 being sufficient for the interaction (52), yet details are still missing. The extracellular domains of FAT1-3 (13) and CELSR1-3 (53) have been the least studied biochemically, and their binding modes and oligomerization states remain to be fully elucidated.

Interestingly, long nonclassical, nonclustered, atypical cadherins feature extracellular domains with modules that do not fold as canonical cadherin repeats. PCDH15 and CDH23 have membrane adjacent domains (MADs) that are ~100-residues long and that had no predicted fold until recently (Fig. 1 A). For PCDH15, this domain has also been labeled as extracellular linker (EL) (49), protocadherin-15 interacting-channel associated domain (PICA) (48), or sea urchin sperm protein, enterokinase, agrin (SEA) like domain (54–56). PCDH24 has a similar putative MAD, whereas FAT and CELSR cadherins all have “unknown” domains between their cadherin repeats and additional extracellular domains with predicted EGF, EGF-like or LamG folds (details below). Here we present an x-ray crystal structure of the pig (*Sus scrofa* [ss]) PCDH15 including repeats EC10, EC11, and its MAD12 fragment (PCDH15 EC10-EC11-MAD12, hereafter referred to as PCDH15 EC10-MAD12) (Figs. 1 B and S1). This structure reveals a ferredoxin-like fold (49) for MAD12 that induces parallel *cis* homodimerization that is predicted to be mechanically weak using steered molecular dynamics (SMD) simulations and that might be present in CDH23, PCDH24, and the FAT and CELSR cadherins, as suggested by sequence analyses and structural modeling.



**FIGURE 1** PCDH15 architecture and structure of the EC10-MAD12 fragment. (A) Schematic of the cadherin tip link and the PCDH15 extracellular domain. Inset shows the location of the fragment studied here. (B) Ribbon diagram of ss PCDH15 EC10-MAD12 (magenta and purple) with transparent molecular surface. Ca<sup>2+</sup> ions are shown as green spheres. Red arrow indicates  $\beta 2''-3_{10}-\beta 3'$  hook. Dashed line indicates unresolved region. Membrane and mechanotransduction channel are indicated (not to scale; TMHS transmembrane helices are purple, TMIE transmembrane helix is orange, TMC1 dimer is light purple). (C) Topology diagram of PCDH15 EC10-MAD12. A typical cadherin fold is observed for EC10 and EC11. A ferredoxin-like fold is observed for MAD12. Intramolecular EC11-MAD12 contact is highlighted by red dashed lines. Blue arrows indicate kinks. Cyan and green circles highlight sites involved in hereditary deafness (causal and correlated mutations, respectively).

## MATERIALS AND METHODS

### Cloning, expression, and purification of bacterially expressed PCDH15 EC10-MAD12

Pig and mouse (*Mus musculus* [mm]) PCDH15 including EC10, EC11, and MAD12 (ss and mm PCDH15 EC10-MAD12) (Tables S1 and S2) were subcloned into NdeI and XhoI sites of the pET21a vector. These constructs were used for expression in Rosetta (DE3) competent cells (Novagen,

Madison, WI) cultured in terrific broth at 37°C and induced at  $OD_{600} = 0.6$  with 1 mM of isopropyl  $\beta$ -D-1-thiogalactopyranoside at 30°C for ~16 h. Cells were lysed by 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 incubated with Ni-Sepharose (GE Healthcare, Chicago, IL) for 1 h and eluted with denaturing buffer supplemented with 500 mM imidazole. Both pig and mouse PCDH15 EC10-MAD12 proteins were refolded at 4°C by fast or drop-wise dilution of 20 mL of eluted denatured protein (1–2 mg/mL) into 480 mL of refolding buffer (20 mM TrisHCl [pH 8.0], 150 mM KCl, 5 mM  $CaCl_2$ , 400 mM L-arginine, 2 mM dithiothreitol or 1 mM tris(2-carboxyethyl)phosphine (TCEP), and 10% glycerol). For fast dilution refolding, the denatured protein was dispensed into the refolding buffer over a 1-min period with vigorous stirring. For drop-wise refolding, the denatured protein was dispensed over a 20-min period with gentle stirring. In both cases, the protein solution was kept under constant and gentle stirring after mixing. All protein solutions were concentrated to 10 mL using Amicon Ultra-15 filters at 3000 revolutions per minute (rpm) with mixing every 20 min. 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, 5 mM  $CaCl_2$ , and 1 mM TCEP. Purity of the recombinant protein was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, after which the protein-containing fractions were pooled and used for further experiments. Pure proteins were concentrated by centrifuge ultrafiltration (Vivaspin-6 10 kDa) to 3–5 mg/mL for crystallization trials and biochemical assays. Protein concentrations were determined using the NanoDrop (Thermo Fisher Scientific, Waltham, MA).

### Cloning, expression, and purification of mammalian-expressed *mm\** PCDH15 EC10-MAD12

Use of mouse mammalian expressed protein is indicated by *mm\**. Suspension Expi293F cells (Thermo Fisher Scientific) were cultured in Expi293 medium at 37°C, 8%  $CO_2$ , and incubated on an orbital shaker at 125 rpm. The coding sequence for *mm* PCDH15 EC10-MAD12 was inserted into a pHis-N1 expression vector (modified version of the pEGFP-N1 vector from Clontech [Mountain View, CA] where the EGFP has been substituted for a hexahistidine tag; James Jontes, The Ohio State University) using XhoI and KpnI restriction sites. The native signal sequence was included before EC10. For a typical 30 mL expression, cells were prepared at a density of  $2.9 \times 10^6$  cells/mL and exchanged into 25.5 mL of fresh media. The DNA complexes were prepared by diluting 30  $\mu$ g of DNA up to 1.5 mL with Opti-MEM and diluting 81  $\mu$ L of ExpiFectamine 293 up to 1.5 mL. After 5 min, the two solutions were combined, gently mixed, and incubated for an additional 20 min at room temperature. The complexes were then added to the cells and incubated for 20 h. The next day, 150  $\mu$ L of ExpiFectamine 293 Transfection Enhancer 1 and 1.5 mL of ExpiFectamine 293 Transfection Enhancer 2 were added. Cells were grown for 4 days, and the conditioned media (CM) was collected. The CM was dialyzed overnight against 20 mM TrisHCl [pH 7.5], 150 mM KCl, 50 mM NaCl, and 10 mM  $CaCl_2$  to remove EDTA. The CM was concentrated using Amicon Ultra-15 10 or 30 kD concentrators and incubated with Ni-Sepharose beads for 1 h. The beads were washed three times with 20 mM TrisHCl [pH 8.0], 300 mM NaCl, 10 mM  $CaCl_2$ , and 20 mM imidazole, and the target protein was eluted with the same buffer containing 500 mM imidazole. The protein was further purified on a Superdex-200 column (GE Healthcare) in 20 mM TrisHCl [pH 7.5], 150 mM KCl, 50 mM NaCl, and 2 mM  $CaCl_2$ . The protein was concentrated and the final concentration was determined by measuring absorbance using the NanoDrop (Thermo Fisher Scientific).

### Crystallization and micro-seeding

Initial needle-like crystals of *ss* PCDH15 EC10-MAD12 were grown by vapor diffusion at 4°C by mixing equal volumes (0.6  $\mu$ L) of protein (purified

without TCEP) and reservoir solutions (0.1 M HEPES [pH 7.5], 25% w/v PEG 2000 MME) and by equilibrating against 75  $\mu$ L of reservoir solution in a sitting drop setup. A seed stock was prepared by mixing of crystal-containing solution (2  $\mu$ L) with reservoir solution (50  $\mu$ L), followed by vortexing with microseed beads (MD2-14; Molecular Dimensions, Newmarket, UK). The resulting solution was diluted to prepare a seed stock and used for random micro-seeding (57). Crystallization droplets were composed of equal volumes (1  $\mu$ L) of protein solution (repurified without TCEP; 4.7 mg/mL) and reservoir solution along with 0.5  $\mu$ L of the seed stock. The droplets were equilibrated against 75  $\mu$ L of reservoir solution in a sitting-drop setup at 4°C. Final hits were obtained in 0.1 M HEPES [pH 7.5], and 20% v/v Jeffamine M-600.

### Data collection and structure determination

Crystals were cryoprotected in reservoir solution plus 25% v/v PEG400. All crystals were cryo-cooled in liquid  $N_2$ . X-ray diffraction data were collected as indicated in Table 1 and processed with HKL2000 (58). The final structure was determined by molecular replacement using PHASER (59) and Buccaneer (60). Initial search models were based on the structure of *Homo sapiens* (*hs*) PCDH15 EC10 (Protein Data Bank [PDB]: 4XHZ) (45) and a separate homology model for the *mm* PCDH15 EC11 repeat

**TABLE 1 Statistics for *ss* PCDH15 EC10-MAD12 Structure**

Data Collection	<i>Ss</i> PCDH15 EC10-MAD12
Space group	P2 <sub>1</sub>
Unit cell parameters	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	92.28, 49.25, 108.25
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 92.2, 90.0
Molecules per asymmetric unit	2
Beam source	APS-24-ID-C
Date of data collection	07-NOV-2017
Wavelength (Å)	0.9792
Resolution limit (Å)	2.10
Unique reflections	57,228
Completeness (%)	97.60 (85.30)
Redundancy	5.30 (2.70)
<i>I</i> / $\sigma$ ( <i>I</i> )	15.21 (2.45)
<i>R</i> <sub>merge</sub>	0.11 (0.44)
<i>R</i> <sub>meas</sub>	0.12 (0.53)
<i>R</i> <sub>pim</sub>	0.05 (0.30)
<i>CC</i> <sub>1/2</sub>	0.93 (0.75)
<i>CC</i> <sup>*</sup>	0.98 (0.93)
Refinement	
Resolution range (Å)	108.18–2.10 (2.15–2.10)
<i>R</i> <sub>work</sub> (%)	18.90 (23.80)
<i>R</i> <sub>free</sub> (%)	21.70 (26.90)
Residues (atoms)	657 (5175)
Water molecules	260
RMSs	
Bond lengths (Å)	0.011
Bond angles (°)	1.55
<i>B</i> -factor average	
Protein	45.51
Ligand/ion	49.85
Water	40.93
Ramachandran plot region (PROCHECK)	
Most favored (%)	90.7
Additionally allowed (%)	9.3
Generously allowed (%)	0.0
Disallowed (%)	0.0
PDB code	6BXZ

obtained using SWISS-MODEL (61) and the PDB: 5DZX (32) as a template. MAD12 was built de novo and refined by using COOT (62). REFMAC5 (63) was used for restrained translation-libration-screw-rotation refinement. Data collection and refinement statistics are provided in Table 1. Coordinates for *ss* PCDH15 EC10-MAD12 have been deposited in the PDB: 6BXZ.

## Analytical ultracentrifugation

Sedimentation velocity analytical ultracentrifugation (AUC) experiments were performed in a ProteomeLab XL-I AUC (Beckman Coulter, Brea, CA) following standard procedures (64–66). Briefly, size-exclusion chromatography (SEC) purified protein samples were loaded into AUC cell assemblies with Epon centerpieces and 12 mm path length. To achieve chemical and thermal equilibrium, the An-50 TI rotor with loaded samples was allowed to equilibrate for ~2 h at 20°C in the centrifuge. The rotor was spun at 50,000 rpm and data were collected using absorption optics. Data analysis was performed with the software SEDFIT (<http://sedfitsedphat.nibib.nih.gov>), using a continuous sedimentation coefficient distribution model  $c(S)$ . Standard values for buffer viscosity (0.01002 poise), density (1 g/mL), and partial specific volume (0.73 mL/g) were used, and confidence level was set to 0.68 as routinely done. The obtained  $c(S)$  distribution was loaded in GUSI (67). AUC experiments were done with at least two biological replicates, which were also accompanied by at least one duplicate to ensure data were obtained accurately.

## SEC-multiangle laser light scattering

SEC-multiangle laser light scattering (MALLS) data were collected by using an ÄKTAmicro system connected in series with a Wyatt miniDAWN TREOS system. The protein samples were separated on a Superdex-S200 3.2/30 column (GE Healthcare), and both the absorbance at 280 nm and the light scattering were monitored. The scattering information was subsequently converted into molecular weight using a rodlike model. SEC-MALLS experiments were done with at least two biological replicates and accompanied by at least one duplicate.

## SEC–small-angle x-ray scattering

Small-angle x-ray scattering (SAXS) data were collected at SIBYLS beamline 12.3.1 in the Advanced Light Source (ALS) (Berkeley, CA) as described (68,69) (Table S3). Purified samples of *ss* PCDH15 EC10-MAD12 (6.3 mg/mL) and *mm* PCDH15 EC10-MAD12 (5.3 mg/mL) were analyzed. Data were collected using an Agilent 1260 series HPLC with a Shodex KW-803 analytical column at a flow rate of 0.5 mL/min using 20 mM TrisHCl [pH 8.0], 150 mM KCl, 5 mM CaCl<sub>2</sub>, and 1 mM TCEP as a mobile phase at 20°C. Each 50  $\mu$ L sample was run through SEC and 3 s x-ray exposures were collected continuously during a ~40 min elution. Samples were examined with incident light at  $\lambda = 1.03$  Å and at a sample to detector distance of 1.5 m resulting in scattering vectors,  $q$ , ranging from 0.01 to 0.6 Å<sup>-1</sup>, where the scattering vector is defined as  $q = 4\pi \sin\theta/\lambda$  and  $2\theta$  is the measured scattering angle. SAXS frames recorded before the protein elution peak were used to subtract all other frames. Buffer subtraction and data reduction were performed at the beamline with SCATTER (70). Buffer matched controls were used for buffer subtraction.

Further data analysis of the merged SAXS data was carried out with PRIMUS (71) and the ATSAS program suite (72). Estimates of the radius of gyration ( $R_g$ ) from the Guinier region were measured with PRIMUS. Maximum dimension ( $D_{\max}$ ) of particles was estimated from an indirect Fourier transform of the SAXS profiles using GNOM (73). Values of  $D_{\max}$  between 120 and 140 Å provided the best solutions for both samples. The oligomeric state of the samples was assessed by estimating their molec-

ular weight using the method implemented in the SAXSMoW2 server (74). Three-dimensional spatial representations of *ss* and *mm* PCDH15 EC10-MAD12, classified as flat using DATCLASS and with an AMBIMETER shape ambiguity score of 2.3 for both samples (possibly ambiguous) (75), were obtained by ab initio modeling with DAMMIF (76) considering  $q < 0.2$ , P1 symmetry, and prolate anisotropy. For each sample, 30 models were generated and averaged with DAMAVER (77) with a mean normalized spatial discrepancy (NSD) of  $1.07 \pm 0.16$  (one model rejected) and  $0.98 \pm 0.13$  (two models rejected) for *ss* and *mm* PCDH15 EC10-MAD12, respectively. A final bead model refined against the SAXS data was produced with DAMMIN/DAMMSTART (78) with an estimated resolution of  $44 \pm 3$  and  $37 \pm 3$  Å (79). Model scattering intensities were computed from *ss* PCDH15 EC10-MAD12 (6BXZ) and fitted to the experimental SAXS data using FoXS (80). Search of conformational changes of the dimeric structure in solution was performed by normal modes analysis with SREFLEX (81).

## Molecular dynamics simulations

The structure of *ss* PCDH15 EC10-MAD12 (PDB: 6BXZ) was used to build monomeric and dimeric systems. Missing residues in one of the dimer subunits were included in the structure using coordinates obtained from the other superposed subunit. Final models included residues p.G1009–p.I1342. The psfgen, solvate, and autoionize visual molecular dynamics plugins were used to build all systems (Table S4) (82). Hydrogen atoms were automatically added to protein and crystallographic water molecules. Residues D, E, K, and R were charged. Histidine residues were neutral, and their protonation states were chosen to favor the formation of evident hydrogen bonds. Additional water molecules and randomly placed ions were used to solvate the systems (150 mM KCl) in boxes large enough to accommodate stretched states and to prevent interactions between periodic images.

Molecular dynamics simulations were performed using NAMD 2.12 (83), the CHARMM36 force field for proteins with the CMAP correction and the TIP3P model for water (84). A cutoff of 12 Å (with a switching function starting at 10 Å) was used for van der Waals interactions along with periodic boundary conditions. The Particle Mesh Ewald method was used to compute long-range electrostatic forces (grid point density of  $>1$  Å<sup>-3</sup>). A uniform 2 fs integration time step was used together with SHAKE. Langevin dynamics was utilized to enforce constant temperature ( $T = 300$  K;  $\gamma = 0.1$  ps<sup>-1</sup>). Constant pressure simulations ( $NpT$ ) at 1 atm were done using the hybrid Nosé-Hoover Langevin piston method with a 200 fs decay period and a 100 fs damping time constant.

Constant-velocity stretching simulations used the SMD method and the NAMD Tcl forces interface. Constant-velocity SMD simulations (85–88) were performed by attaching C $\alpha$  atoms of N- and C-terminal residues to independent virtual springs. For some simulations (Table S4), these springs were connected to virtual slabs connected to a third spring. All springs had stiffness  $k_s = 1$  kcal mol<sup>-1</sup> Å<sup>-2</sup>. The free ends of the stretching springs were moved away from the protein in opposite directions at a constant velocity. Applied forces were computed using the extension of the virtual springs. Maximum force peaks and their averages were computed from 50 ps running averages used to eliminate local fluctuations.

## Sequence analyses, structural modeling, and analysis tools

To compare protein sequences among various species, 18 sequences from 17 species were obtained for their longest PCDH15 isoforms from the National Center for Biotechnology Information protein database (Table S1) and were split into the corresponding EC10, EC11, and MAD12 fragments before alignment. These fragments were aligned as in (47). The ClustalW algorithm (89) on Geneious (90) was used to obtain percent sequence

identity for each fragment. Alignment files from Geneious were colored according to % of sequence identity in JalView with 45% conservation threshold (91). Similarly, sequences among MADs from different cadherin subfamilies (Table S2) were split for sequence alignments (colored according to % sequence identity in JalView with 1% conservation threshold), which were used as input into the Sequence Identity and Similarity (SIAS) server (92) to obtain sequence identity matrices. 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 *hs* PCDH15 isoform 1 precursor, National Center for Biotechnology Information Reference Sequence: NP\_001136235.1 (signal peptide with 26 residues). Details of interatomic interactions were analyzed using Maestro (93) and visual molecular dynamics (82). Molecular stereo images were prepared with PyMOL (94). Plots and curve fits were prepared with XMGrace or QtiPlot. Protein-protein contacts were analyzed with the Protein, Interfaces, Surfaces, and Assemblies (PISA) server (95). Coordinates for structural models of *hs* CDH23 MAD28, *hs* PCDH24 MAD10, *hs* FAT1 MAD35, *hs* FAT2 MAD34, *hs* CELSR1 MAD10, *hs* CELSR2 MAD10, *dm* FAT MAD35, as well as dimeric *hs* CDH23 EC26-MAD28 and *hs* PCDH24 EC8-MAD10 are available upon request.

### Protein-protein docking and MM-GBSA calculations

Homology models of *hs* CDH23 EC26-27 and *hs* PCDH24 EC8-9 were obtained using SWISS-MODEL (61) and published structures (PDB: 5TFK and 5SZR) (33,47) as templates. RaptorX MAD models were connected to models of the preceding EC repeats to build L-shaped structures using COOT (62) and *ss* PCDH15 EC10-MAD12 as a guide. Models were minimized to reduce steric clashes with MacroModel and the OPLS3e force field using standard parameters (implicit solvent, charges from forcefield, PRCG method with 2500 iterations in gradient conversion mode, convergence threshold of 0.05) without constraints (96). Resulting monomers were built as dimers based on *ss* PCDH15 EC10-MAD12 and further minimized (97). Monomeric protomers were then used for protein-protein docking with BioLuminate using the “Dimer mode” with 70,000 rotations for the ligand. Final docking poses were selected after automatized clustering and structural superposition to *ss* PCDH15 EC10-MAD12. Final models were further minimized and used as inputs for molecular mechanics-generalized Born surface area (MM-GBSA) calculations using Prime (96), the VSGB solvation model and the OPLS3e force field with standard parameters, the “side chain only” minimization mode, and no constraints.

## RESULTS AND DISCUSSION

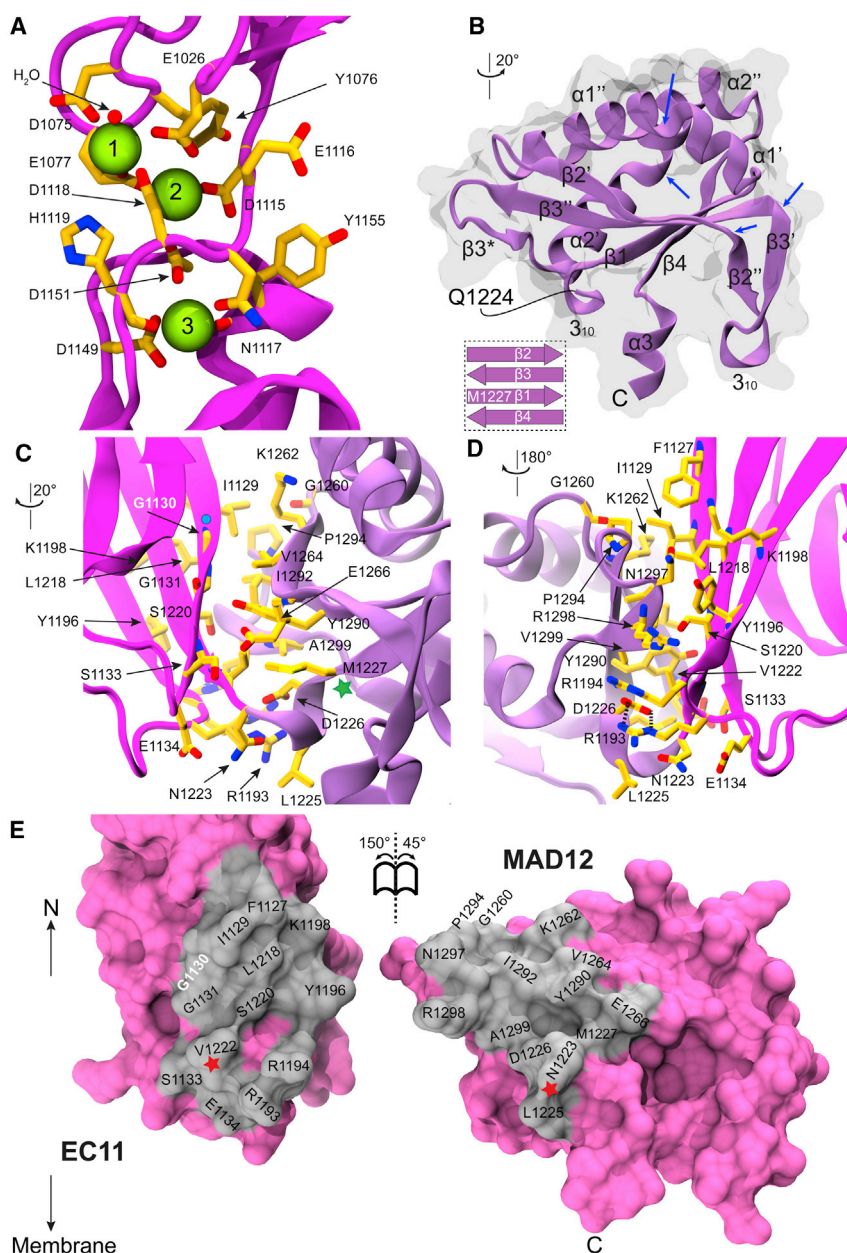
Multiple sequence alignments of the PCDH15 EC10-MAD12 from 17 different species reveal that repeat EC10 is less conserved compared to repeat EC11, which is moderately conserved, while MAD12 is most conserved, with sequence identities of 37.6, 47.7, and 58.3%, respectively (Fig. S2). Secondary structure predictions and domain identification algorithms are consistent in identifying PCDH15 EC10 and EC11 as EC repeats while failing to identify MAD12 as a domain of known fold. As expected, Ca<sup>2+</sup>-binding motifs for the linker region EC10-11, typically characterized by the sequence NTerm-XEX-DXD-D(R/Y)E-XDX-DXNDN-CTerm, are highly conserved, with a minor variation at the linker (DENXH). In contrast, Ca<sup>2+</sup>-binding motifs are absent in the EC11-MAD12 linker, despite the high sequence conservation of the last domain across mammals, birds, reptiles, and fish (Fig. S2). To determine the architecture and fold of

this fragment and to biochemically characterize it, we purified *ss* and *mm* PCDH15 EC10-MAD12 from bacterial inclusion bodies, as well as *mm* PCDH15 EC10-MAD12 produced in mammalian cells (labeled as *mm*\*). These protein fragments included residues p.H1007–p.E1353 (residue numbering corresponds to processed protein without its signal sequence). All protein fragments were well behaved in SEC experiments (Fig. S3) and were used for crystallization trials, structure determination, and biochemical and biophysical analyses, as described below.

### Crystal structure of pig PCDH15 EC10-MAD12 reveals a ferredoxin-like fold for MAD12

Diffraction crystals were obtained for the *ss* PCDH15 EC10-MAD12 protein fragment, for which an x-ray crystal structure was solved and refined to 2.1 Å resolution (Table 1). Two protomers were observed in the asymmetric unit, comprising residues p.G1009–p.I1342 and p.I1011–p.T1340, with missing residues not observed in the electron density map. These protomers adopted an L-shaped architecture (Fig. 1B) and interacted with each other in a parallel *cis* arrangement, as also recently reported in (49) (Fig. S4) and as discussed below. The long end of the L-shaped bent structure has two EC repeats (EC10 and EC11) with typical Greek key folds (seven β strands labeled A–G for EC10; A'–G' for EC11; Fig. 1C) and a canonical linker region with three bound Ca<sup>2+</sup> ions at sites 1, 2, and 3 (Figs. 1B and 2A; Fig. S1A). The *ss* PCDH15 EC10 repeat shares structural features with its orthologs from human and mouse (45), with some differences in its FG connector, which folds into a stable α helix in the presence of the preceding EC9 repeat to be part of a bent, Ca<sup>2+</sup>-free EC9-10 linker region. Instead, in one chain of our structure, this loop is extended forming a β hairpin (F\*G\*) involved in crystal contacts (Fig. 1, B and C; Fig. S4, A and B), whereas electron density is absent for the other chain. The *ss* PCDH15 EC11 repeat has a canonical fold with an atypical and highly conserved p.N1223-XL-p.D1226 linker that forms a 3<sub>10</sub> helix leading to MAD12 (Figs. 2B and S2), a domain that is positioned next to the EC11 face formed by β-strands A', F', and G'. MAD12 tucks against EC11 to form the short arm of the L-shaped protomer.

The overall architecture of MAD12 is similar to that observed for the ferredoxin-like fold in the Nup54 αβ domain (PDB: 5C2U, core RMSD 1.64 Å) (98). MAD12 has four β strands and two kinked α helices in a βαβαβαβ fold with kinks dividing α1 into α1' and α1", as well as α2 into α2' and α2" (Figs. 1C and 2B). Both MAD12 and Nup54 have a short β strand preceding α2' (labeled β3\*) that forms a β-hairpin with the end of β3. There are also key differences between these folds, the most significant being related to kinks in MAD12 that divide β2 into β2' and β2", and β3 into β3' and β3". The unique β2"–β3' hairpin, bridged by a 3<sub>10</sub> helix, bends to form a “hook” reaching in the direction of the membrane, thereby establishing



**FIGURE 2** Structural details of EC10-MAD12 fragment. (A) Detail of EC10-11 linker region.  $\text{Ca}^{2+}$  ions are shown in green and  $\text{Ca}^{2+}$ -coordinating side chains in yellow sticks. (B) MAD12 in ribbon representation with secondary structure elements labeled. Blue arrows indicate kinks. Inset shows  $\beta$  sheet arrangement. (C and D) A detailed view of the intramolecular interactions at the EC11-MAD12 interface. Residues are shown in yellow sticks and labeled; some backbone atoms are omitted for clarity. Cyan circle indicates site involved in inherited deafness (p.G1130R). Green star indicates the first residue of MAD12. (E) Interaction surface exposed (gray). Red star indicates connection point. Deafness mutation site p.G1130R is highlighted in white. Rotations in (C) and (E) are indicated with respect to view of *ss* PCDH15 EC10-MAD12 in Fig. 1 B. Rotation in (D) is with respect to view in (C).

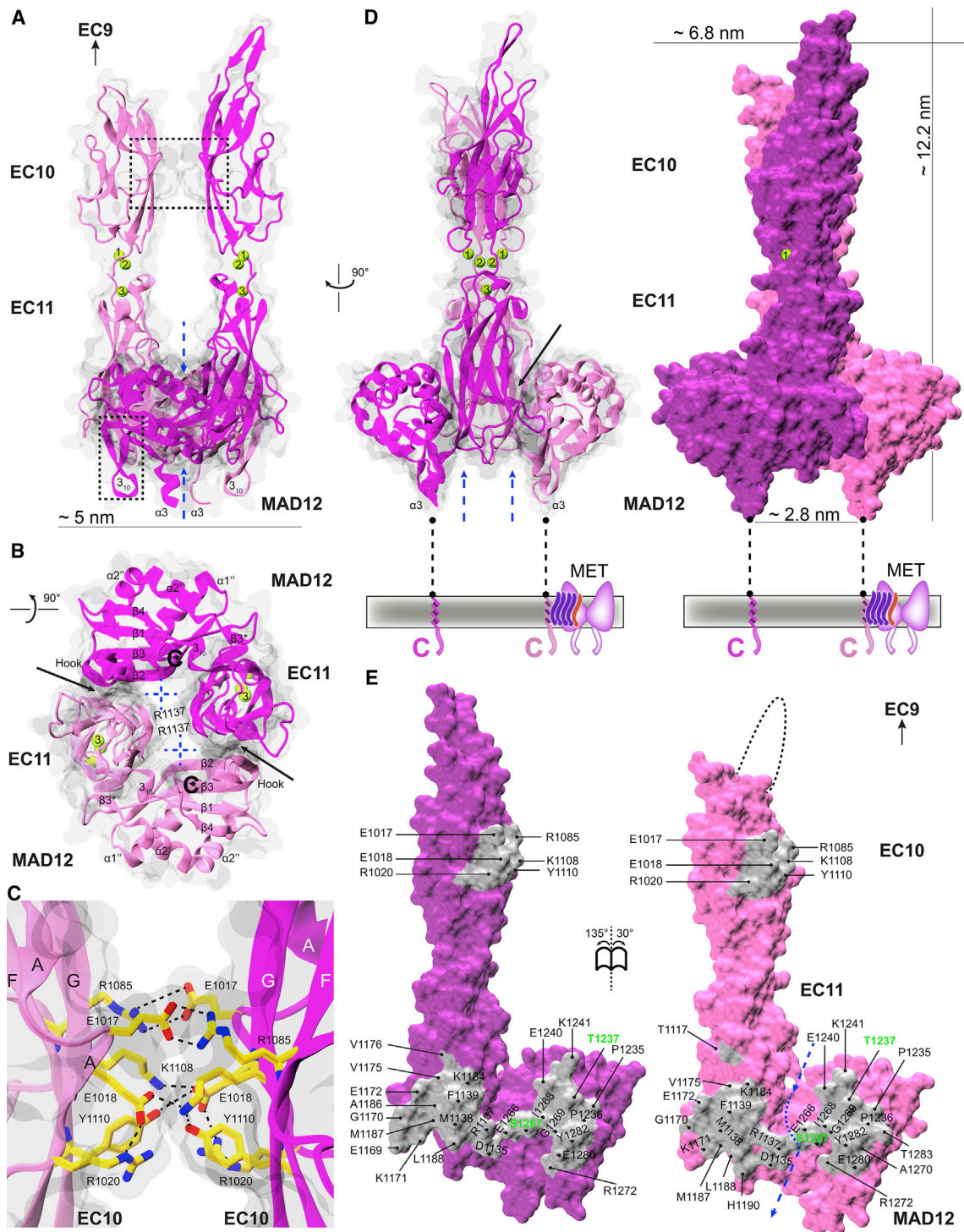
contacts that support the  $\alpha 3$  linker connecting to PCDH15's transmembrane helix (Figs. 1 B and 2 B; Fig. S1, B and D).

The L-shaped structure of the PCDH15 EC10-MAD12 protomer harbors various intramolecular contacts between EC11 and MAD12 involving both hydrophobic and hydrophilic surface residues located at the EC11-MAD12 linker, at the beginning of  $\beta 2'$ , at the end of  $\beta 3''$ , and at the  $\beta 3''$ - $\beta 3^*$ - $\alpha 2'$  connection (Fig. 2, C-E). This interaction surface is  $\sim 640 \text{ \AA}^2$  (for one protomer) and the hydrophobic nature of some residues (F1127, I1129, L1218, V1222, V1264, Y1290, I1292, P1294, A1299) that stabilize the interface between EC11 and MAD12 is strictly conserved across species (Figs. 2 C and S2), with some variations in sidechain size, which suggests that this intramolecular contact is somewhat robust to residue

substitutions as long as their hydrophobicity is maintained. Interestingly, a missense mutation (p.G1130R) that causes inherited deafness (99) alters a site at this interface (Figs. 2, C-E and S1 C). The sidechain of the mutated residue would not point toward the interface but would likely destabilize the hydrophobic core of EC11. This suggests that disruption of the structural integrity of EC11 and the intramolecular EC11-MAD12 contacts might be deleterious for PCDH15 function.

### Pig PCDH15 EC10-MAD12 forms parallel *cis* dimers *in crystallo*

The two protomers observed in the asymmetric unit form a parallel *cis* dimer with three contact points (Fig. 3). The



**FIGURE 3** Dimerization interface in *ss* PCDH15 EC10-MAD12. (*A* and *B*) Front and bottom views of *ss* PCDH15 EC10-MAD12 dimer in ribbon (magenta and mauve) with a transparent molecular surface. Black-dashed boxes indicate EC10-EC10 interface (*A*, top) and MAD12's hook (*A*, bottom). Dashed blue arrows indicate opening crossed by p.R1137. Solid black arrows indicate binding interface between protomers. (*C*) Detail of EC10-EC10 interface (black box in *A*). Protein backbone is shown in ribbons and relevant residues are shown as yellow sticks. (*D*) Side view of dimer in ribbon (left) and opaque surface (right) representations. (*E*) Interaction surfaces exposed and shown in gray with interfacing residues labeled. Missing loop in one protomer is indicated by dashed-black line. Deafness associated sites are labeled in green.

first one involves interactions between the two copies of repeat EC10, which face each other bringing together residues on the surface of their  $\beta$ -strands A, G, and F (Fig. 3, A, C, and E). The surface area of this point of contact is small ( $\sim 185 \text{ \AA}^2$ ) with interactions mediated by charged and polar residues forming hydrogen bonds and salt-bridges (p.E1017, p.E1018, p.R1020, p.R1085, p.K1108, and p.Y1110). There is strict conservation of charge for all five charged sites in the pig construct, whereas p.Y1110 is conserved only in mammals and replaced by hydrophobic residues in birds, reptiles, and some fish (Fig. S2). Under MAD12-induced dimerization, details of this contact may minimally vary across species, and although not required for dimerization (45,48,49), the conserved EC10-EC10 contacts may help stabilize a subset of possible conformations for the PCDH15 EC9-MAD12 subdomain (49). Interestingly, the dynamics of this EC10-EC10 *cis* interface might be allosterically controlled by  $\text{Ca}^{2+}$ -binding to the EC10-11 linker positioned right below it, thereby providing some  $\text{Ca}^{2+}$ -dependent rigidity at the lower end of the tip link.

The other two contact points, also recently reported in (49), are equivalent and involve the tail end of one of the L-shaped molecules interacting with the elbow of the other, together forming a closed ring arrangement with some additional contacts stemming from residues pointing toward the center of the ring on the membrane-proximal end of the dimer (Figs. 3, A–E and S5). Total surface area for these interfaces is  $\sim 975 \text{ \AA}^2$ , with hydrophobic and hydrophilic residues from EC11's A'B' loop (p.R1137, p.M1138, and p.F1139), C'D' loop (p.E1172),  $\beta$ -strand D' (p.V1175), the end of  $\beta$ -strand E' (p.K1184 and p.A1186) and the beginning of the E'F' loop (p.L1188), as well as residues from MAD12's  $\alpha 1'$  helix (p.P1236, p.T1237, and p.E1240) and its  $\beta$ -strands  $\beta 2'$  (p.E1266) and  $\beta 3'$  (p.E1280 and p.Y1282). Residue p.T1237, which forms a hydrogen bond with p.E1172 at this interface, is associated to inherited deafness. The mutation p.T1237P is correlated with a deafness phenotype, although causality has not been established (100). Interestingly, the ring-like arrangement (Fig. 3 B) features an opening exposed to solvent and crossed by p.R1137 residues from the two protomers. These sidechains interact with either the backbone carbonyl group of p.D1135 or with the sidechain of p.E1266 from the neighboring EC11. These interactions bring AB loops from each EC11 protomer together to stabilize the dimeric interface. Charge is again strictly conserved throughout all the species analyzed (Fig. S2), while hydrophobicity is less conserved for two of the sites (p.1138 and p.1188). The total surface area involved in the potential *cis* dimerization interface, including all three points of contact ( $\sim 1160 \text{ \AA}^2$ ), and the sequence conservation of residues involved in it suggests that parallel dimerization might be robust and relevant for PCDH15 function.

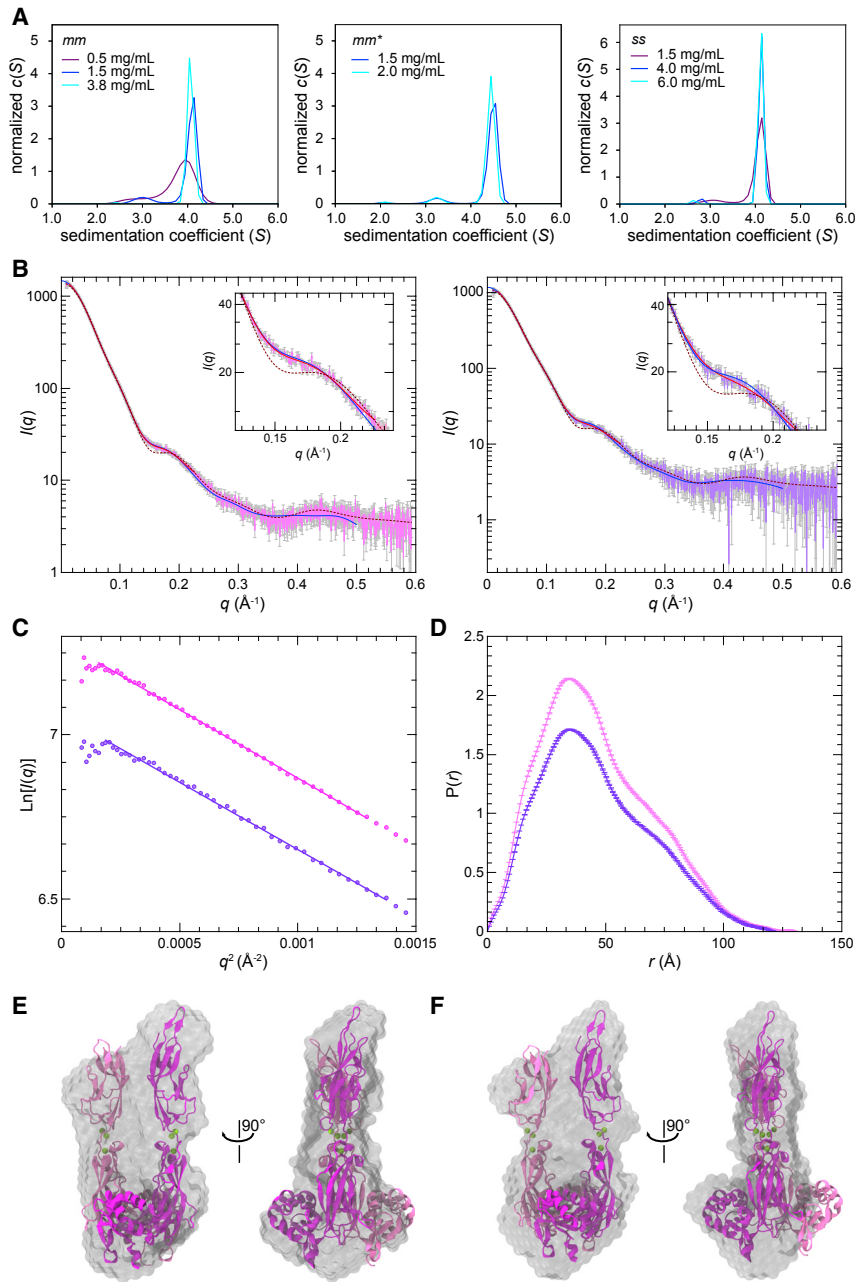
## Pig and mouse PCDH15 EC10-MAD12 form dimers in solution

To determine the oligomeric state of PCDH15 EC10-MAD12 in solution and to test the propensity of this fragment to dimerize, as seen in the x-ray crystal structure, we used pig and mouse protein fragments for AUC experiments, as well as SEC coupled to MALLS and SAXS. Sedimentation velocity AUC data for the refolded *ss* PCDH15 EC10-MAD12 (produced in bacteria; theoretical monomeric mass of 39.96 kDa) show a major peak with a sedimentation coefficient ( $S$ ) of  $\sim 4$  and a minor peak at  $S \sim 3$  (Fig. 4 A). From these AUC measurements, the molar mass of the major peak is  $\sim 60$  kDa, yet uncertainty in buffer density, viscosity, specific molar volume, and friction coefficients precluded a more accurate estimate. Nevertheless, we assigned these peaks to dimer and monomer forms, respectively. Refolded *mm* PCDH15 EC10-MAD12 (produced in bacteria; theoretical monomeric mass of 39.97 kDa) also showed similar peaks (Fig. 4 A). Consistent with AUC, although less conclusive, SEC-MALLS experiments suggest that the *ss* and *mm* PCDH15 EC10-MAD12 proteins are dimeric in solution ( $M_w = 62.13 \text{ kDa} \pm 0.058\%$  and  $61.13 \text{ kDa} \pm 0.043\%$ , respectively) (Fig. S6, A and B). Molecular weight estimates from SEC-SAXS experiments for both the pig and the mouse fragments (76.20 and 74.92 kDa; 4.7–6.3% discrepancy with sequence-derived molecular weight) are in better agreement with the expected mass for dimeric PCDH15 EC10-MAD12 in solution.

Similarly, glycosylated *mm\** PCDH15 EC10-MAD12 produced in a mammalian expression system behaves as a dimer in solution. AUC experiments show that *mm\** PCDH15 EC10-MAD12 is predominantly a dimer with a calculated mass of  $89.4 \pm 0.3 \text{ kDa}$ . The  $c(S)$  distribution (Fig. 4 A) also shows a small monomer peak corresponding to a mass of  $54.6 \pm 2.1 \text{ kDa}$ . The expected mass is  $\sim 55 \text{ kDa}$ , as determined by SDS-gel analysis, with the additional mass due to posttranslational modifications on the protein (Fig. S3). Estimates from SEC-MALLS ( $130.3 \text{ kDa} \pm 0.299\%$ ) (Fig. S6 A) are consistent with the AUC data and with the expected mass of a dimeric *mm\** PCDH15 EC10-MAD12 in solution. Together, this data strongly support a dimeric arrangement at the membrane proximal region of PCDH15, as also recently reported for similar protein fragments (48,49).

Data from SAXS experiments can also be used to obtain information about the size and shape of a protein in solution. The radius of gyration ( $R_g$ ) for both *ss* and *mm* PCDH15 EC10-MAD12 was estimated using two approaches. First, the Guinier analysis was used to obtain  $R_g$  values of  $35.37 \pm 0.27$  and  $35.53 \pm 0.53 \text{ \AA}$ , respectively (Fig. 4, B and C; Table S3). In the second approach, the indirect Fourier transform of the SAXS profile yielded  $R_g$  values of  $36.30 \pm 0.09 \text{ \AA}$  for the pig and  $36.34 \pm 0.08 \text{ \AA}$  for





**FIGURE 4** PCDH15 EC10-MAD12 forms a parallel dimer in solution. (A) AUC experiments for mouse protein refolded from bacteria (*mm*, left), mouse protein purified from mammalian cells (*mm\**, middle), and pig protein refolded from bacteria (*ss*, right). Peaks at  $S > 4$  represent dimeric states. (B) X-ray scattering intensity as a function of the scattering vector  $q$  (SAXS profile) for pig (magenta, left) and mouse (purple, right) PCDH15 EC10-MAD12. Predicted scattering intensities from the structure (6BXZ) obtained with FoXS are shown in maroon-dashed lines ( $\chi^2 = 5.09/3.42$ ), and theoretical scattering curves obtained from ab initio modeling (DAMMIF) and from flexible refinement with SREFLEX are shown in red ( $\chi^2 = 1.44/1.14$ ) and blue ( $\chi^2 = 1.64/1.23$ ), respectively. (C) Guinier plot of the low  $q$  region of the SAXS data for *ss* and *mm* (magenta and purple circles, respectively) PCDH15 EC10-MAD12. Magenta and purple solid lines show linear fits from which the gradient of the slope ( $-R_g^2/3$ ) was used to estimate  $R_g$ . (D) Real-space pair distribution function  $P(r)$  from SAXS data for *ss* and *mm* PCDH15 EC10-MAD12 (magenta and purple). (E and F) Superposition of the *ss* PCDH15 EC10-MAD12 structure (6BXZ) with refined low-resolution bead models obtained for pig ( $44 \pm 3 \text{ \AA}$ ) and mouse ( $37 \pm 3 \text{ \AA}$ ) PCDH15 EC10-MAD12 data.

the mouse protein, with maximum dimensions ( $D_{\max}$ ) of 131 and 122  $\text{\AA}$ , respectively (Fig. 4 D; Table S3). These estimates are in excellent agreement with each other and with the  $R_g$  obtained using the *ss* PCDH15 EC10-MAD12 structure (34  $\text{\AA}$ ), suggesting that the overall shape of the dimer observed in the crystallographic structure is maintained in solution. However, comparison of the SAXS data to x-ray intensities modeled from the crystal structure revealed some discrepancies reflected in large  $\chi^2$  values obtained for the fitting (5.09 and 3.42 for *ss* and *mm* PCDH15 EC10-MAD12, respectively) and by a clear dip in the  $q$ -region  $0.1\text{--}0.25 \text{ \AA}^{-1}$  observed in the modeled intensities

but absent in the experimental data (Fig. 4 B). These discrepancies may indicate loss of symmetry in solution and may arise from the relaxation and reaccommodation of the dimer interface and “wobbling” of the protein. Analysis of the Kratky plot (Fig. S6 C) indicates that while both proteins are folded in solution, they exhibit significant flexibility, likely arising from interrepeat motion or the presence of a C-terminal tail including residues p.I1342–p.E1353 (not seen in the structure) and a His-tag used for protein purification.

The presence of a flexible or conformationally heterogeneous dimer in the samples was further highlighted by

ab-initio modeling of the SAXS data. 30 models produced for shape reconstruction considering P1 symmetry and prolate anisotropy (see [Materials and Methods](#)) were averaged to generate final models that were refined against the SAXS data for each sample using DAMMIF/DAMMIN/DAMSTART (76,78). Predicted x-ray intensities of the resulting bead models are in good agreement with the experimental data ( $\chi^2 = 1.45$  for pig and  $\chi^2 = 1.15$  for mouse) and with the crystallographic structure (Fig. 4, E and F). However, the bead models superposed to the crystallographic structure highlight the lack of symmetry of the dimer in solution and the likely flexibility of the EC10-EC10 interface.

To determine if the crystallographic structure could be reaccommodated in a way that fits the experimental data, we carried out normal-mode analysis on it followed by fitting to the SAXS data with SREFLEX (81) for data from both pig and mouse samples. X-ray intensities calculated from these models fit the experimental data better than the crystal structure ( $\chi^2 = 1.64$  for pig and  $\chi^2 = 1.23$  for mouse) and lack the dip in the  $q$ -region  $0.1\text{--}0.25 \text{ \AA}^{-1}$ . The adjusted models have an overall root-mean-square-deviation (RMSD) for  $C\alpha$  atoms of  $\sim 7 \text{ \AA}$  when compared with the crystallographic structure. These models show changes in the orientation and rotation of the ECs with loss of symmetry and with one EC10 repeat tilted and leaning toward the adjacent, straighter EC10 (Fig. S6 D). Overall, SAXS data suggest a flexible and perhaps asymmetric arrangement of EC10 repeats while strongly supporting a dimeric PCDH15 EC10-MAD12 in solution.

### Steered molecular dynamics simulations predict the mechanics of *ss* PCDH15 EC10-MAD12

The topology and structure of MAD12 has important implications for the transmission of force to the inner-ear transduction channel. Unlike the  $\beta$ -sandwich Greek-key fold of EC repeats, in which N- and C-terminal points are at opposite ends of the folding unit, the N-terminus of MAD12's  $\beta\alpha\beta\beta\alpha\beta$  fold is spatially nearby its C-terminal end (within  $\sim 5 \text{ \AA}$ ). While stretching and mechanical unfolding of an EC repeat requires large forces to shear-out  $\beta$ -strand A or G (101–103), unfolding of MAD12, when pulling from the N- and C-termini, would require unzipping of  $\beta 4$  from  $\beta 1$  and subsequent unraveling of  $\alpha 2$ . Experiments and simulation have suggested that unzipping of  $\beta$  strands is easier than shearing (104–108), indicating that unfolding of MAD12 may require less force than the force needed to unfold EC repeats.

To explore the elastic response of *ss* PCDH15 EC10-MAD12, its unfolding pathway, and its mechanical strength, we carried out SMD simulations in which force was applied through independent springs attached to the ends of the protein fragment (86–88,109,110). The free ends of the springs were in turn moved at constant speed in opposite directions, mimicking in vivo force application. Simulation systems,

encompassing up to  $\sim 279,000$  atoms (Table S4), included explicit solvent and ions along with either a monomer or a dimer of *ss* PCDH15 EC10-MAD12. Stretching of the monomer at  $0.1 \text{ nm/ns}$  revealed quick unraveling of  $\alpha 3$ , followed by opening of the EC11-MAD12 interface (separation of p.R1144 from p.E1244) and unrolling of MAD12 accompanied by a  $90^\circ$  swing of MAD12's hook (Figs. 5 A and S7, A and B; Simulation S1d; Table S4; Video S1). A final unfolding event involved both unzipping of  $\beta 4$  from  $\beta 1$  with rupture of backbone hydrogen bonds, and splitting of two salt bridges formed during the SMD trajectory (p.D1286–p.R1334 and p.R1271–p.E1332) (Fig. 5, A and B; Fig. S7, A–G). Rupture of this “electrostatic lock,” which linked  $\beta 4$  with  $\beta 2$  (hook) and with  $\beta 3$ , correlated with the largest force peak ( $477.9 \text{ pN}$ ) monitored throughout the simulation (Fig. S7 C). Further unfolding of MAD12's  $\alpha 2$  ensued at drastically smaller forces without noticeable unfolding of EC10 or EC11. A similar scenario was observed in the slowest stretching simulation at  $0.02 \text{ nm/ns}$  (Table S4, S1e), with the magnitude of force peaks decreasing at slower stretching speeds as expected (111). Monomer force peaks at all stretching speeds were associated to MAD12 unfolding and were smaller in magnitude than unfolding forces reported for other CDH23 and PCDH15 EC repeats stretched under similar simulation conditions (Fig. 5 G) (45,46,103). In addition, MAD12's unfolding force peaks were smaller than predicted unbinding force peaks for the PCDH15-CDH23 monomeric complex (43,112), suggesting that a single tip-link handshake bond formed by EC1-2 repeats of CDH23 and PCDH15 can withstand similar or larger forces than MAD12.

Stretching of the *ss* PCDH15 EC10-MAD12 dimer was carried out using two constant-velocity SMD protocols and revealed a more complex response. First, we applied forces in opposite directions through independent springs attached to each of the four ends of the protein fragments (Fig. 5 C). At the slowest stretching speed ( $0.1 \text{ nm/ns}$ ) (Table S4, S2d), one of the subunits'  $\beta 4$  detached from  $\beta 1$  as MAD12 was separating out from EC11 and unrolling, which correlated with a broad and rather undefined force peak (data not shown). In the other subunit MAD12 separated from EC11 and unrolled after unzipping  $\beta 4$  from  $\beta 1$  and rupture of the electrostatic lock described above, which correlated with a clear force peak (Fig. 5 D). In a second approach, pairs of N- or C-termini were connected to virtual slabs through independent springs (Fig. 5 E). The slabs were in turn connected through springs to SMD atoms moving at constant speeds in opposite directions. In the simulations that used this approach at the slowest stretching speeds ( $0.1$  and  $0.02 \text{ nm/ns}$ ) (Table S4, S2g and S2h), MAD12 from one subunit separated from EC11 and unrolled before  $\beta 4$  detached from  $\beta 1$ , while MAD12 from the other subunit remained partially attached to EC11, with unzipping of  $\beta 4$  from  $\beta 1$  happening first (Figs. 5 E and S8; Video S2). In all cases, additional intermolecular interactions formed

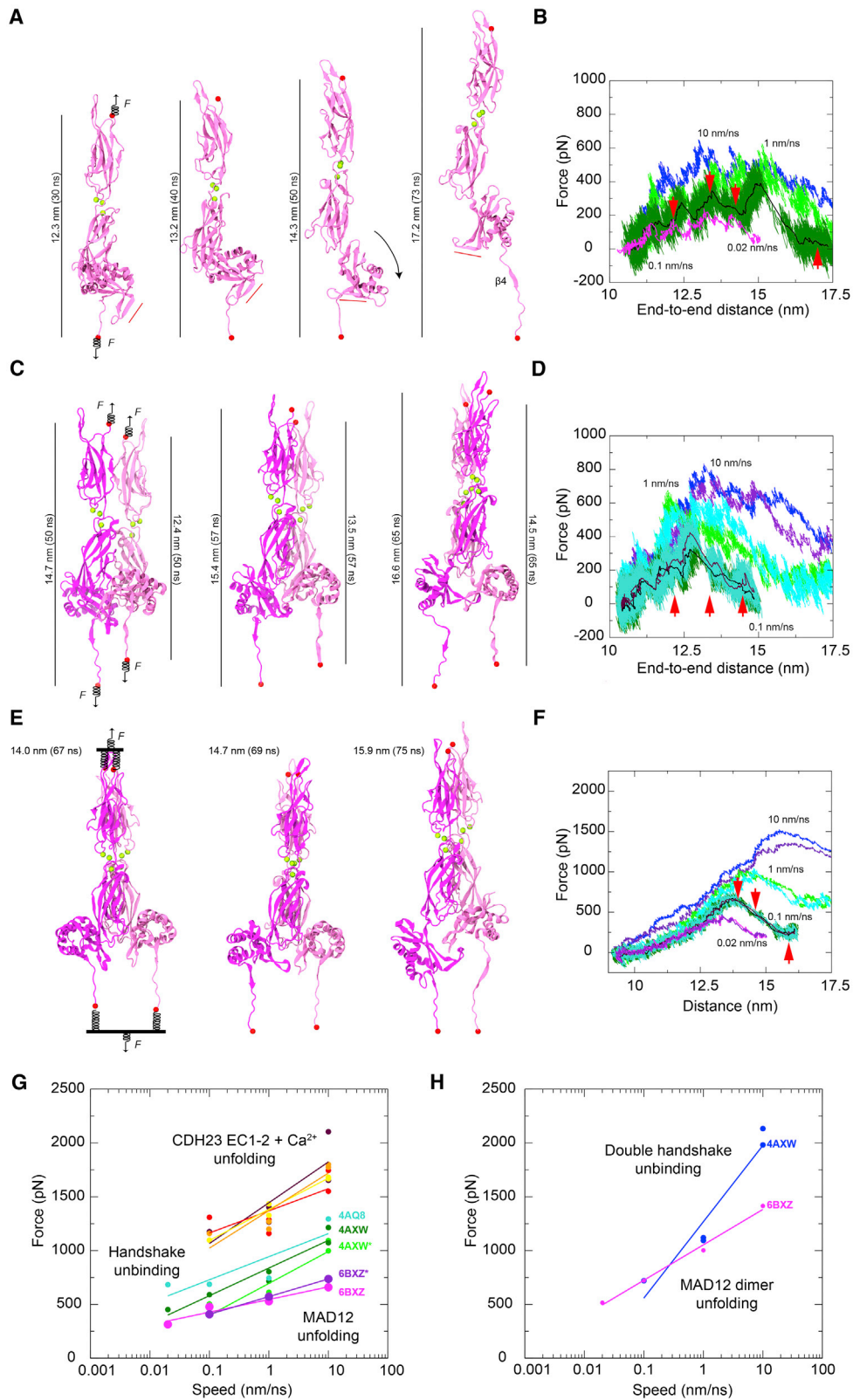


FIGURE 5 Constant-velocity SMD simulations of *ss* PCDH15 EC10-MAD12. (A) Snapshots of monomeric *ss* PCDH15 EC10-MAD12 stretching from Simulation S1d (0.1 nm/ns; Table S4). Protein is shown in mauve ribbon representation,  $\text{Ca}^{2+}$  ions as green spheres, and stretched terminal  $\text{C}\alpha$  atoms as red spheres. Springs indicate position and direction of applied forces. Red bar illustrates direction of PCDH15's MAD12 Hook. (B) Force applied to C terminus

(legend continued on next page)

throughout trajectories and MAD12's hook swung as monitored during stretching of the monomer of *ss* PCDH15 EC10-MAD12. Average unfolding force peaks of MAD12 were smaller than unfolding force peaks of PCDH15 EC repeats and unbinding force peaks of the PCDH15 EC1-2 + CDH23 EC1-2 bond (43,45,46,103,112), again suggesting that MAD12 is mechanically weak, with the tip-link bond being of similar or larger strength at the stretching speeds tested in this study (Fig. 5, G and H).

## Secondary structure analyses and modeling predict MAD folds in other cadherins

The topology and structure of MAD12 has important implications for other cadherins as well. The high structural similarity observed between MAD12 and Nup54 (98), despite their very poor sequence identity (Figs. S9 and S10), led us to search for other  $\beta\alpha\beta\beta\alpha\beta$  domains that might be hidden within the cadherin superfamily (Fig. 6). We looked for extracellular segments that lacked predicted domain architecture, referred to as unknown regions according to SMART (113). In parallel, we predicted the secondary structure of sequences from all cadherins in the human genome using RaptorX (114), which had correctly predicted the secondary structure of MAD12 and identified Nup54 (PDB: 5C2U) (98) as a good template for “homology” modeling. We then focused on sequences predicted to have the  $\beta\alpha\beta\beta\alpha\beta$  fold, and found the pattern in CDH23, PCDH24 (also known as CDHR2), all FAT cadherins, and all CELSRs (Figs. S9 and S11). These predicted  $\beta\alpha\beta\beta\alpha\beta$ -fold containing segments were classified as unknown regions by SMART.

The potential presence of MADs in CDH23 and PCDH24 is intriguing (MAD28 and MAD10, respectively) (Fig. 6). In both cases, these domains would be located close to their membrane insertion points, linking cadherin EC repeats to their associated transmembrane helices. Both proteins had been identified as members of the Cr-2 subfamily based on sequence similarity of their extracellular tips (1,4,103,115), and the presence of a ferredoxin-fold domain in both these proteins further confirms their structural rela-

tionship. Perhaps even more intriguing is the existence of potential MAD folds in human FAT and CELSR cadherins. Analysis of the *Drosophila melanogaster* (*dm*) FAT also shows the potential existence of a similar  $\beta\alpha\beta\beta\alpha\beta$  fold (MAD35) (Fig. S11). In all these cases, the predicted “unknown region” is between the last, C-terminal EC and subsequent EGF or LAMG domains (MAD35 for FAT1; MAD34 for FAT2; MAD33 for FAT3; MAD35 for FAT4; and MAD10 for CELSR1-3) (Fig. 6 A).

To gain further insights into the structural similarities among cadherin MADs, we used RaptorX to create structural models for each of these domains in the human “MAD bearing” cadherins (PCDH15, CDH23, PCDH24, FAT1-4, and CELSR1-3) and the fly sequences (*dm* FAT) mentioned above (11 models) (Figs. 6, B–E and S11, A–G). Visual inspection confirmed the  $\beta\alpha\beta\beta\alpha\beta$  fold for eight of these models (*hs* PCDH15, *hs* CDH23, *hs* PCDH24, *hs* FAT1, *hs* FAT2, *hs* CELSR1, *hs* CELSR2, and *dm* FAT). In addition, we classified models according to the hydrophobicity of their core. Hydrophobic cores without polar or charged residues pointing into it were observed for PCDH15 MAD12, CDH23 MAD28, FAT1 MAD35, and CELSR1-2 MAD10, which also had their charged residues distributed on their surfaces and loops, as expected. Models for PCDH24 MAD10, FAT2 MAD34, and *dm* MAD35 were satisfactory, with one or two polar residues lying at the interface between their hydrophobic cores and loops, which could be easily accommodated by rotamer rearrangement. A structural alignment of these models to *ss* PCDH15 MAD12 using STAMP (116,117) revealed high structural conservation at the core secondary structure elements, and poor structural conservation of loops for the eight models mentioned above (Fig. 6).

The dimeric structure of *ss* PCDH15 EC10-MAD12 suggests that other MADs, and their preceding EC repeats, could form *cis* homodimers as well. However, part of the dimer interface involves the  $\beta 2''$ – $3_{10}$ – $\beta 3'$  hooks that are not well predicted by RaptorX (see “control” prediction in Fig. 6 C). Nevertheless, the  $\beta 2''$ – $3_{10}$ – $\beta 3'$  hook observed in *ss* PCDH15 MAD12 might be present in CDH23 MAD28. Although the hook is not observed in our model,

---

versus end-to-end distance for constant velocity stretching of monomeric *ss* PCDH15 EC10-MAD12 at 10 nm/ns (S1b, blue), 1 nm/ns (S1c, light green), 0.1 nm/ns (S1d, dark green; 1-ns running average shown in black) and 0.02 nm/ns (S1e, 1-ns running average shown in magenta). Red arrowheads indicate time-points for S1d illustrated in (A). (C) Snapshots of dimeric *ss* PCDH15 EC10-MAD12 stretching from simulation S2d as in (A), with subunits in magenta and mauve. (D) Force applied to N- and C-termini of one subunit (mauve in C) versus end-to-end distance for constant-velocity stretching of dimeric *ss* PCDH15 EC10-MAD12 at 10 nm/ns (S2b, blue and indigo), 1 nm/ns (S2c, light green and cyan), and 0.1 nm/ns (S2d, dark green and turquoise; 1-ns running average shown in black and maroon). Red arrowheads indicate time points illustrated in (C). (E) Snapshots of dimeric *ss* PCDH15 EC10-MAD12 stretching from simulation S2g as in (C). Stretching was carried out by attaching two slabs to springs that were in turn attached to the terminal ends of each protein. Slabs were moved in opposite directions through individual springs. (F) Force applied to each of the slabs versus slab separation for constant-velocity stretching of dimeric *ss* PCDH15 EC10-MAD12 at 10 nm/ns (S2e, blue and indigo), 1 nm/ns (S2f, light green and cyan), 0.1 nm/ns (S2g, dark green and turquoise; 1-ns running average shown in black and maroon), and 0.02 nm/ns (S2h, 1-ns running average shown in magenta and violet). (G) *In silico* force peak maxima versus stretching speed for CDH23 EC1-2 unfolding (red, maroon, yellow, and orange) (103), for CDH23 EC1-2 and PCDH15 EC1-2 handshake unbinding (cyan-4AQ8, dark green-4AXW, and light green-4AXW\* [after 1  $\mu$ s-long equilibration]) (43,112), and for MAD12 unfolding (magenta-6BXZ for simulations S1b–e; purple-6BXZ\* for averages obtained for simulations S2b–d). (H) *In silico* force peak maxima versus stretching speed for unbinding of pairs of CDH23 EC1-2 and PCDH15 EC1-2 handshakes (blue) (43) and for unfolding of MAD12 using slabs (magenta, simulations S2e–h).

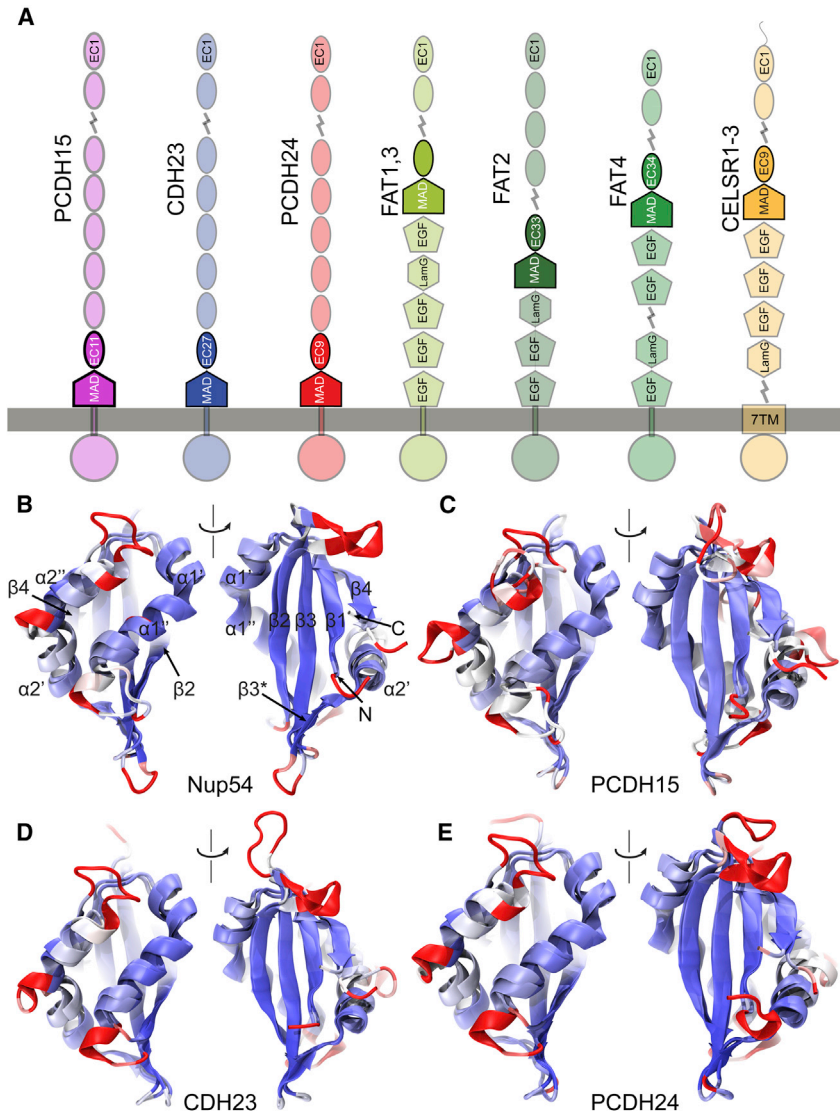


FIGURE 6 Presence of putative MADs in the cadherin superfamily. (A) Domain organization of human cadherins predicted to feature domains similar to PCDH15 MAD12. FAT cadherins are grouped according to post-MAD segment architecture. Preceding EC repeats are not labeled for FAT1 (EC34) and FAT3 (EC32). (B) Structural comparison between *xl* Nup54  $\alpha\beta$  domain (5C2U) (98) and *ss* PCDH15 MAD12 (PDB: 6BXZ; front and back views). Molecules are colored according to structural homology per residue ( $Q_{res}$ ) (117). Blue color implies high structural conservation, whereas red implies poor structural conservations. Beginning of  $\beta1$  and  $\alpha3$  were truncated out for better structural alignment. (C–E) Structural models of MADs from atypical cadherins (*hs* PCDH15 MAD12; *hs* CDH23 MAD28; *hs* PCDH24 MAD10) superposed on the crystal structure of *ss* PCDH15 MAD12 (6BXZ) are shown, as in (B).

CDH23 MAD28 has similar length and distribution of charged residues in this region (Fig. S9). Similarly, in the secondary structure prediction by RaptorX (Fig. S9), the well predicted short  $\beta3^*$  strand interacting with the preceding EC repeat exist in CDH23 and CELSR1 and 2, thus suggesting an architecture that might be similar to the L-shaped PCDH15 MAD12 favoring dimerization.

To determine whether parallel dimerization facilitated by MADs in other cadherins is possible *in silico*, we built models of *hs* CDH23 EC26-MAD28 and *hs* PCDH24 EC8-MAD10 and computed their interaction energies. Homology models for EC repeats were coupled to RaptorX models for MADs using the *ss* PCDH15 EC10-MAD12 structure as a guide. The resulting, models with L-shaped protomers were minimized as dimers and used for “blind” molecular docking to predict binding poses for CDH23 and PCDH24, which resulted in *cis* parallel dimers that are similar to *ss* PCDH15 EC10-MAD12 (Fig. S12). In

both cases, the L-shaped architecture of *hs* CDH23 EC26-MAD28 and *hs* PCDH24 EC8-MAD10 favors the formation of a closed ring-like arrangement involving several hydrophobic residues that could stabilize their dimeric interfaces. In addition, free energy binding calculations (via MM-GBSA) predict strong binding for these admittedly biased dimers (Table S5), suggesting that MAD-induced parallel dimerization is possible in CDH23 and PCDH24. Taken together, our modeling analyses suggest that MADs are present in all the aforementioned proteins, with some variations in fold at loops and at the  $\beta2''$ - $\beta3$  and  $\beta3''$ - $\beta3^*$  connections that might determine their oligomerization states and function.

## CONCLUSIONS

The structural, biochemical, and computational results presented here shed light on three aspects of PCDH15 function.

First, the unique architecture and fold of PCDH15's MAD12, along with SMD simulations of its elastic response, suggest that this domain is mechanically weak, which has implications on how PCDH15 and the tip link may respond to large mechanical stimuli (*e.g.*, loud sound). Second, MAD12 induces parallel dimerization at the membrane proximal end of PCDH15's extracellular domain, thus establishing constraints on models of mechanotransduction and how the tip link may interact with other parts of the transduction machinery. Last, simulations suggest mechanisms by which force-induced deformation of MAD12 can trigger conformational changes in associated membrane proteins, thus providing testable hypotheses on how PCDH15 may gate the inner-ear transduction channel should they be directly coupled to each other.

Determining the structure and mechanical properties of PCDH15 is a necessary step to understand its function in the context of inner-ear mechanotransduction (18,40,42,118–123). The  $\beta\alpha\beta\beta\alpha\beta$  ( $\alpha/\beta$ ) fold adopted by PCDH15's MAD12, revealed here and in (49), is distinct and unique when compared to those of all other EC repeats in PCDH15, which are all  $\beta$ -sandwich modules (all- $\beta$ ). Previous single-molecule force spectroscopy experiments and simulations have suggested the existence of a hierarchy of mechanical strengths that depends on both the secondary structure elements of a protein module and the geometry of stretching (104–108,124). All- $\beta$  proteins in which mechanical unfolding necessitates shearing of  $\beta$  strands are often strong, followed by all- $\beta$  proteins in which unzipping of  $\beta$  strands is mechanically easier. In contrast,  $\alpha/\beta$  proteins are mechanically weaker than all- $\beta$  proteins, and all- $\alpha$  modules seem to be the weakest of all. Consistent with this hierarchy, our SMD simulations of the *ss* PCDH15 EC10-MAD12 fragment show that MAD12 unrolls and separates from EC11 at low force, with ensuing unfolding at moderate forces facilitated by the unzipping of MAD12's  $\beta_4$  from  $\beta_1$  and rupture of an electrostatic lock formed by conserved charged residues at  $\beta_4$ ,  $\beta_2''$ , and  $\beta_3''$ . Peak unfolding forces for the  $\alpha/\beta$  MAD12 at all stretching speeds tested here were smaller than those reported for unfolding of various all- $\beta$  CDH23 and PCDH15 EC repeats with bound  $\text{Ca}^{2+}$  (45,46,103). At the low  $\text{Ca}^{2+}$  concentrations of the cochlea (125,126), MAD12 may still stochastically unfold before other weakened EC repeats without bound  $\text{Ca}^{2+}$  (at any of the sites 1, 2, or 3) when tip links are exposed to extreme mechanical stimuli. Further simulations mimicking these conditions may help elucidate the  $\text{Ca}^{2+}$ -dependent mechanics of PCDH15.

Intriguingly, our simulations predict that MAD12 unfolds at similar or smaller forces than those required to break a single PCDH15-CDH23 handshake bond in silico (43), suggesting that under some conditions, tip links may lengthen significantly due to unfolding without PCDH15-CDH23 unbinding. A stretch of  $\sim 120$  unfolded residues could easily account for  $>40$  nm lengthening, but we do not

know whether MAD12 would refold back into its ferredoxin-like fold in time for the next mechanical stimulus. While it is unclear whether unfolding could occur before unbinding at significantly slower stretching speeds (equivalent to low-frequency, normal to faint sound) than those used in our simulations (close to high-frequency, loud sound), it provides a plausible explanation to observations of long tip links that might have been artificially stretched during staining or freezing for electron-microscopy imaging. Unfolding also provides a more compelling explanation to electrophysiological experiments in which tip links are extended over 100 nm without compromising transduction (127). It is possible that both unfolding of MAD12 and unbinding of PCDH15 from CDH23 are used to provide distinct safety mechanisms that are perhaps adapted to different sets of extreme stimuli to prevent rupture of other transduction machinery components.

Parallel dimerization mediated by MAD12 presents additional constraints on how PCDH15 stretches and how it communicates force to its transmembrane protein partners. Predicted unfolding forces of MAD12 in the context of the dimer were not significantly different than those predicted for the monomer, and they were smaller than those required to break a double handshake stretched in silico, suggesting that unfolding can occur before unbinding. It remains to be determined whether a heterotetrameric handshake bond in which parallel contacts at PCDH15 EC2-3 (48) and between the CDH23 chains may strengthen the handshake even further, thus favoring unfolding over unbinding at all stretching speeds in heterotetrameric configurations. Intriguingly, deformation and unfolding of MAD12 during stretching simulations of the dimer were in some cases asymmetric. Whether unrolling and separation of MAD12 from EC11 occur before or after unzipping of MAD12's  $\beta_4$  from  $\beta_1$  in an asymmetric fashion might be controlled by dimerization, direction of the applied stretching force, and binding partners.

Asymmetry in parallel dimerization of PCDH15 observed in our SAXS data is consistent with raw data from negative staining and electron cryomicroscopy (48,49). Although averaged structural models may not fully reflect asymmetry and flexibility, SAXS data suggest that PCDH15 exhibits asymmetry and that in some cases each of the subunits in the dimer may adopt its own conformation. In the absence of force, the likely flexibility of the EC10-EC10 interface may indicate that the conformations at the lower end of PCDH15 are allosterically controlled and stabilized by MAD12 and by the EC10-EC11  $\text{Ca}^{2+}$ -binding linker and not necessarily by EC9-EC9 or EC10-EC10 contacts between both protomers (49). Molecular dynamics simulations also suggest asymmetric unrolling in the MAD12-induced dimer upon force application, indicating that in some cases force might be conveyed by each protomer independently and in slightly different ways to their respective binding partners.

Interestingly, PCDH15 seems to interact promiscuously with multiple transmembrane proteins, including TMCs, various TRP channels, and TMHS (128–131). The most compelling data on these interactions comes from the recent structural and biophysical characterization of the PCDH15 MAD12-TMHS complex (49). While transmembrane domains mediate most of the interaction between TMHS and PCDH15, there are contacts between the extracellular loops of TMHS and the  $\alpha 3$  helix and  $\beta 2''-3_{10}-\beta 3'$  hook of PCDH15's MAD12, as well as with the E'F' loop of PCDH15 EC11. These interactions suggest that TMHS may stabilize the base of MAD12, perhaps preventing unraveling of PCDH15's  $\alpha 3$  helix upon force application and facilitating membrane tensioning by tip links, which in turn could result in gating of nearby transduction channels (132–134). Alternatively, it is tempting to speculate that unrolling of MAD12 and swinging of its hook, as observed in our simulations upon force application, could induce conformational changes, either in TMHS or directly in TMC1, that result in “uncorking” of one of these proteins and the subsequent gating of the yet-to-be fully characterized inner-ear transduction channel (135–139). Further structural studies and SMD simulations of the entire complex, including the mobility constraints imposed from a more realistic membrane environment, will shed light on the molecular details of this process.

Our results have implications beyond PCDH15 and inner-ear mechanotransduction as we predict the presence of MADS in other large cadherins such as CDH23, PCDH24, FATs, and CELSRs. Both CDH23 and PCDH24 are expected to be tensioned under physiological conditions because CDH23 is part of the inner-ear tip link and PCDH24 forms brush-border intermicrovillar links (50). Possible MAD-induced dimerization may play a critical role in CDH23 and PCDH24 function. It is conceivable that “MAD-bearing” cadherins form MAD-mediated heterodimers, perhaps relevant for maturation of tip links (during which PCDH15 gets replaced by CDH23) (140), for adhesion specificity, or other physiologically relevant processes. Interestingly, PCDH15 and CDH23 have been recently implicated in brain circuit assembly (141), whereas FATs and CELSRs are known to be involved in regulation of cell polarity and migration. Perhaps all “MAD-bearing” cadherins play a mechanical role, either in sensory perception, epithelial morphogenesis and function, or in brain development and wiring (3,142).

## SUPPORTING MATERIAL

Twelve figures, five tables, and two videos are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(18\)31256-6](http://www.biophysj.org/biophysj/supplemental/S0006-3495(18)31256-6).

## AUTHOR CONTRIBUTIONS

P.D. did protein expression, purification, crystallization, and x-ray data collection of *ss* PCDH15 EC10-MAD12. P.D. and M.S. solved the x-ray

crystal structure. R.A.S. did cloning, protein expression, purification, and initial crystallization attempts of *ss* PCDH15 EC10-MAD12, which generated crystals for seed stocks. P.D. and D.C. prepared the seed stock for crystallization. P.D. and D.C. carried out protein expression and purification of *ss* and *mm* PCDH15 EC10-MAD12 for SEC-SAXS experiments. R.A.S. analyzed the SAXS data and generated SAXS models. P.D. and D.C. did protein expression and purification of *ss* and *mm* PCDH15 EC10-MAD12 for SEC-MALLS and AUC experiments. D.C. carried out SEC-MALLS and AUC experiments and did data analysis. Y.N. did cloning, expression, and purification of *mm\** PCDH15 EC10-MAD12 in mammalian cells, carried out AUC and SEC-MALLS with this sample, and did data analysis. P.D. and D.C. did structural and sequence analyses of MADs and Nup54. D.C. did bioinformatics analyses and generated models of MADs. P.D. did bioinformatics analyses, built homology models of *hs* CDH23 EC26-MAD28 and *hs* PCDH24 EC8-MAD10, and did molecular docking and MM-GBSA calculations. M.S. did molecular dynamics simulations. M.S. and P.D. wrote and edited the manuscript with help from all coauthors. M.S. trained and supervised all coauthors and assisted in crystal fishing and cryo-cooling, structure refinement, deposition, and data analysis presented in this work.

## ACKNOWLEDGMENTS

The authors thank Michal Hammel and Daniel Rosenberg for data collection and initial SEC-SAXS data analysis, Brandon L. Neel for assistance with sequence alignments, and 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 R01 DC015271) and by the National Science Foundation through XSEDE (XRAC MCB140226). Simulations were performed at the Pittsburgh Supercomputer Center, Bridges and Ohio Supercomputer Center, Owens (PAS1037) supercomputers. Use of the APS NE-CAT beamlines was supported by NIH (P30 GM1241653 and S10 RR029205) and the Department of Energy (DE-AC02-06CH11357) through grants GUP 49774 and 59251. SAXS data collection at the Advanced Light Source SIBYLS beamline is funded through: DOE BER Integrated Diffraction Analysis Technologies program and NIH grants NIGMS P30 GM124169 ALS-ENABLE and S10OD018483. P.D. is a Pelotonia Fellow and R.A.S. was a Pelotonia fellow.

## SUPPORTING CITATIONS

References (143–145) appear in [Supporting Material](#).

## REFERENCES

- Hulpiau, P., and F. van Roy. 2009. Molecular evolution of the cadherin superfamily. *Int. J. Biochem. Cell Biol.* 41:349–369.
- Brasch, J., O. J. Harrison, ..., L. Shapiro. 2012. Thinking outside the cell: how cadherins drive adhesion. *Trends Cell Biol.* 22: 299–310.
- Hirano, S., and M. Takeichi. 2012. Cadherins in brain morphogenesis and wiring. *Physiol. Rev.* 92:597–634.
- Sotomayor, M., R. Gaudet, and D. P. Corey. 2014. Sorting out a promiscuous superfamily: towards cadherin connectomics. *Trends Cell Biol.* 24:524–536.
- Ringwald, M., R. Schuh, ..., S. Mayer. 1987. The structure of cell adhesion molecule uvomorulin. Insights into the molecular mechanism of Ca<sup>2+</sup>-dependent cell adhesion. *EMBO J.* 6:3647–3653.
- Shapiro, L., A. M. Fannon, ..., W. A. Hendrickson. 1995. Structural basis of cell-cell adhesion by cadherins. *Nature.* 374:327–337.

7. Overduin, M., T. S. Harvey, ..., M. Ikura. 1995. Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. *Science*. 267:386–389.
8. Nagar, B., M. Overduin, ..., J. M. Rini. 1996. Structural basis of calcium-induced E-cadherin rigidification and dimerization. *Nature*. 380:360–364.
9. Takeichi, M. 1990. Cadherins: a molecular family important in selective cell-cell adhesion. *Annu. Rev. Biochem.* 59:237–252.
10. Gumbiner, B. M. 2005. Regulation of cadherin-mediated adhesion in morphogenesis. *Nat. Rev. Mol. Cell Biol.* 6:622–634.
11. Vlemminckx, K., and R. Kemler. 1999. Cadherins and tissue formation: integrating adhesion and signaling. *BioEssays*. 21:211–220.
12. Junghans, D., I. G. Haas, and R. Kemler. 2005. Mammalian cadherins and protocadherins: about cell death, synapses and processing. *Curr. Opin. Cell Biol.* 17:446–452.
13. Sadeqzadeh, E., C. E. de Bock, and R. F. Thorne. 2014. Sleeping giants: emerging roles for the fat cadherins in health and disease. *Med. Res. Rev.* 34:190–221.
14. Avilés, E. C., and L. V. Goodrich. 2017. Configuring a robust nervous system with Fat cadherins. *Semin. Cell Dev. Biol.* 69:91–101.
15. Rübsam, M., J. A. Broussard, ..., C. M. Niessen. 2018. Adherens junctions and desmosomes coordinate mechanics and signaling to orchestrate tissue morphogenesis and function: an evolutionary perspective. *Cold Spring Harb. Perspect. Biol.* 10:a029207.
16. Leckband, D. E., and J. de Rooij. 2014. Cadherin adhesion and mechanotransduction. *Annu. Rev. Cell Dev. Biol.* 30:291–315.
17. Pruitt, B. L., A. R. Dunn, ..., W. J. Nelson. 2014. Mechano-transduction: from molecules to tissues. *PLoS Biol.* 12:e1001996.
18. Jaiganesh, A., Y. Narui, ..., M. Sotomayor. 2018. Beyond cell-cell adhesion: sensational cadherins for hearing and balance. *Cold Spring Harb. Perspect. Biol.* 10:a029280.
19. Kemler, R., C. Babinet, ..., F. Jacob. 1977. Surface antigen in early differentiation. *Proc. Natl. Acad. Sci. USA*. 74:4449–4452.
20. Takeichi, M. 1977. Functional correlation between cell adhesive properties and some cell surface proteins. *J. Cell Biol.* 75:464–474.
21. Ozawa, M., and R. Kemler. 1990. Correct proteolytic cleavage is required for the cell adhesive function of uvomorulin. *J. Cell Biol.* 111:1645–1650.
22. Brasch, J., O. J. Harrison, ..., L. Shapiro. 2011. Structure and binding mechanism of vascular endothelial cadherin: a divergent classical cadherin. *J. Mol. Biol.* 408:57–73.
23. Harrison, O. J., X. Jin, ..., B. Honig. 2011. The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. *Structure*. 19:244–256.
24. Boggon, T. J., J. Murray, ..., L. Shapiro. 2002. C-cadherin ectodomain structure and implications for cell adhesion mechanisms. *Science*. 296:1308–1313.
25. Patel, S. D., C. Ciatto, ..., L. Shapiro. 2006. Type II cadherin ectodomain structures: implications for classical cadherin specificity. *Cell*. 124:1255–1268.
26. Harrison, O. J., F. Bahna, ..., L. Shapiro. 2010. Two-step adhesive binding by classical cadherins. *Nat. Struct. Mol. Biol.* 17:348–357.
27. Ciatto, C., F. Bahna, ..., L. Shapiro. 2010. T-cadherin structures reveal a novel adhesive binding mechanism. *Nat. Struct. Mol. Biol.* 17:339–347.
28. Wu, Y., J. Vendome, ..., B. Honig. 2011. Transforming binding affinities from three dimensions to two with application to cadherin clustering. *Nature*. 475:510–513.
29. Sivasankar, S., Y. Zhang, ..., S. Chu. 2009. Characterizing the initial encounter complex in cadherin adhesion. *Structure*. 17:1075–1081.
30. Zhang, Y., S. Sivasankar, ..., S. Chu. 2009. Resolving cadherin interactions and binding cooperativity at the single-molecule level. *Proc. Natl. Acad. Sci. USA*. 106:109–114.
31. Leckband, D., and S. Sivasankar. 2012. Cadherin recognition and adhesion. *Curr. Opin. Cell Biol.* 24:620–627.
32. Goodman, K. M., R. Rubinstein, ..., L. Shapiro. 2016. Structural basis of diverse homophilic recognition by clustered  $\alpha$ - and  $\beta$ -protocadherins. *Neuron*. 90:709–723.
33. Goodman, K. M., R. Rubinstein, ..., L. Shapiro. 2016.  $\gamma$ -protocadherin structural diversity and functional implications. *eLife*. 5:e20930.
34. Nicoludis, J. M., S. Y. Lau, ..., R. Gaudet. 2015. Structure and sequence analyses of clustered protocadherins reveal antiparallel interactions that mediate homophilic specificity. *Structure*. 23:2087–2098.
35. Nicoludis, J. M., B. E. Vogt, ..., R. Gaudet. 2016. Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1–4. *eLife*. 5:e18449.
36. Rubinstein, R., C. A. Thu, ..., B. Honig. 2015. Molecular logic of neuronal self-recognition through protocadherin domain interactions. *Cell*. 163:629–642.
37. Goodman, K. M., R. Rubinstein, ..., L. Shapiro. 2017. Protocadherin cis-dimer architecture and recognition unit diversity. *Proc. Natl. Acad. Sci. USA*. 114:E9829–E9837.
38. Cooper, S. R., J. D. Jontes, and M. Sotomayor. 2016. Structural determinants of adhesion by Protocadherin-19 and implications for its role in epilepsy. *eLife*. 5:e18529.
39. Quint, E., and K. P. Steel. 2003. Use of mouse genetics for studying inner ear development. *Curr. Top. Dev. Biol.* 57:45–83.
40. Peperman, E., and C. Petit. 2015. The tip-link molecular complex of the auditory mechano-electrical transduction machinery. *Hear Res.* 330:10–17.
41. Friedman, T. B., J. M. Schultz, ..., C. C. Brewer. 2011. Usher syndrome: hearing loss with vision loss. *Medical Genetics in the Clinical Practice of ORL*. Karger, pp. 56–65.
42. Kazmierczak, P., H. Sakaguchi, ..., B. Kachar. 2007. Cadherin 23 and protocadherin 15 interact to form tip-link filaments in sensory hair cells. *Nature*. 449:87–91.
43. Sotomayor, M., W. A. Weihofen, ..., D. P. Corey. 2012. Structure of a force-conveying cadherin bond essential for inner-ear mechanotransduction. *Nature*. 492:128–132.
44. Geng, R., M. Sotomayor, ..., K. N. Alagramam. 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.
45. Araya-Secchi, R., B. L. Neel, and M. Sotomayor. 2016. An elastic element in the protocadherin-15 tip link of the inner ear. *Nat. Commun.* 7:13458.
46. Powers, R. E., R. Gaudet, and M. Sotomayor. 2017. A partial calcium-free linker confers flexibility to inner-ear protocadherin-15. *Structure*. 25:482–495.
47. Jaiganesh, A., P. De-la-Torre, ..., M. Sotomayor. 2018. Zooming in on cadherin-23: structural diversity and potential mechanisms of inherited deafness. *Structure*. 26:1210–1225.e4.
48. Dionne, G., X. Qiu, ..., L. Shapiro. 2018. Mechanotransduction by PCDH15 relies on a novel cis-dimeric architecture. *Neuron*. 99:480–492.e5.
49. Ge, J., J. Elferich, ..., E. Gouaux. 2018. Structure of mouse protocadherin 15 of the stereocilia tip link in complex with LHFPL5. *eLife*. 7:e38770.
50. Crawley, S. W., D. A. Shifrin, Jr., ..., M. J. Tyska. 2014. Intestinal brush border assembly driven by protocadherin-based intermicrovillar adhesion. *Cell*. 157:433–446.
51. Crawley, S. W., M. S. Mooseker, and M. J. Tyska. 2014. Shaping the intestinal brush border. *J. Cell Biol.* 207:441–451.
52. Tsukasaki, Y., N. Miyazaki, ..., M. Takeichi. 2014. Giant cadherins Fat and Dachsous self-bend to organize properly spaced intercellular junctions. *Proc. Natl. Acad. Sci. USA*. 111:16011–16016.
53. Goffinet, A. M., and F. Tissir. 2017. Seven pass cadherins CELSR1-3. *Semin. Cell Dev. Biol.* 69:102–110.
54. Bork, P., and L. Patthy. 1995. The SEA module: a new extracellular domain associated with O-glycosylation. *Protein Sci.* 4:1421–1425.



55. Pei, J., and N. V. Grishin. 2017. Expansion of divergent SEA domains in cell surface proteins and nucleoporin 54. *Protein Sci.* 26:617–630.
56. Hayward, A. N., E. J. Aird, and W. R. Gordon. 2018. SNAPS: a toolkit for studying cell surface shedding of diverse transmembrane receptors. *bioRxiv*, Posted online October 5, 2018 <https://doi.org/10.1101/436592>.
57. D'Arcy, A., T. Bergfors, ..., M. Marsh. 2014. Microseed matrix screening for optimization in protein crystallization: what have we learned? *Acta Crystallogr. F Struct. Biol. Commun.* 70:1117–1126.
58. Otwinowski, Z., and W. Minor. 1997. Processing of X-ray diffraction data. *Methods Enzymol.* 276:307–326.
59. McCoy, A. J., R. W. Grosse-Kunstleve, ..., R. J. Read. 2007. Phaser crystallographic software. *J. Appl. Crystallogr.* 40:658–674.
60. Cowtan, K. 2006. The *Buccaneer* software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* 62:1002–1011.
61. Waterhouse, A., M. Bertoni, ..., T. Schwede. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46:W296–W303.
62. Emsley, P., B. Lohkamp, ..., K. Cowtan. 2010. Features and development of coot. *Acta Crystallogr. D Biol. Crystallogr.* 66:486–501.
63. Murshudov, G. N., P. Skubák, ..., A. A. Vagin. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* 67:355–367.
64. Dam, J., C. A. Velikovsky, ..., P. Schuck. 2005. Sedimentation velocity analysis of heterogeneous protein-protein interactions: lamm equation modeling and sedimentation coefficient distributions  $c(s)$ . *Biophys. J.* 89:619–634.
65. Schuck, P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* 78:1606–1619.
66. Zhao, H., C. A. Brautigam, ..., P. Schuck. 2013. Overview of current methods in sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. *Curr. Protoc. Protein Sci.* Chapter 20:Unit20.12.
67. Brautigam, C. A. 2015. Calculations and publication-quality illustrations for analytical ultracentrifugation data. *Methods Enzymol.* 562:109–133.
68. Dyer, K. N., M. Hammel, ..., G. L. Hura. 2014. High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. *Methods Mol. Biol.* 1091:245–258.
69. Classen, S., G. L. Hura, ..., J. A. Tainer. 2013. Implementation and performance of SIBYLS: a dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the advanced light source. *J. Appl. Cryst.* 46:1–13.
70. Rambo, R. P., and J. A. Tainer. 2011. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers.* 95:559–571.
71. Konarev, P. V., V. V. Volkov, ..., D. I. Svergun. 2003. PRIMUS: a windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* 36:1277–1282.
72. Franke, D., M. V. Petoukhov, ..., D. I. Svergun. 2017. ATLAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Cryst.* 50:1212–1225.
73. Svergun, D. I. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* 25:495–503.
74. Fischer, H., M. de Oliveira Neto, ..., A. F. Craievich. 2010. Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J. Appl. Crystallogr.* 43:101–109.
75. Petoukhov, M. V., and D. I. Svergun. 2015. Ambiguity assessment of small-angle scattering curves from monodisperse systems. *Acta Crystallogr. D Biol. Crystallogr.* 71:1051–1058.
76. Franke, D., and D. I. Svergun. 2009. DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering. *J. Appl. Cryst.* 42:342–346.
77. Volkov, V. V., and D. I. Svergun. 2003. Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* 36:860–864.
78. Svergun, D. I. 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* 76:2879–2886.
79. Tuukkanen, A. T., G. J. Kleywegt, and D. I. Svergun. 2016. Resolution of *ab initio* shapes determined from small-angle scattering. *IUCrJ.* 3:440–447.
80. Schneidman-Duhovny, D., M. Hammel, ..., A. Sali. 2013. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.* 105:962–974.
81. Panjkovich, A., and D. I. Svergun. 2016. Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *Phys. Chem. Chem. Phys.* 18:5707–5719.
82. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28..
83. Phillips, J. C., R. Braun, ..., K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.
84. Huang, J., and A. D. MacKerell, Jr. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* 34:2135–2145.
85. Isralewitz, B., J. Baudry, ..., K. Schulten. 2001. Steered molecular dynamics investigations of protein function. *J. Mol. Graph. Model.* 19:13–25.
86. Grubmüller, H. 2005. Force probe molecular dynamics simulations. *Methods Mol. Biol.* 305:493–515.
87. Gräter, F., J. Shen, ..., H. Grubmüller. 2005. Mechanically induced titin kinase activation studied by force-probe molecular dynamics simulations. *Biophys. J.* 88:790–804.
88. Sotomayor, M., and K. Schulten. 2007. Single-molecule experiments in vitro and in silico. *Science.* 316:1144–1148.
89. Larkin, M. A., G. Blackshields, ..., D. G. Higgins. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics.* 23:2947–2948.
90. Kearse, M., R. Moir, ..., A. Drummond. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 28:1647–1649.
91. Waterhouse, A. M., J. B. Procter, ..., G. J. Barton. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* 25:1189–1191.
92. Garcia-Boronat, M., C. M. Diez-Rivero, ..., P. A. Reche. 2008. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. *Nucleic Acids Res.* 36:W35–W41.
93. Schrodinger, L.L.C. Small-Molecule Discovery Suite 2018–2: New York, New York.
94. Schrödinger, L.L.C. 2015. The PyMOL Molecular Graphics System, Version-1.8.
95. Krissinel, E., and K. Henrick. 2007. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372:774–797.
96. Schrodinger, L.L.C. Maestro Release 2018–3: New York, New York.
97. Sastry, G. M., M. Adzhigirey, ..., W. Sherman. 2013. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* 27:221–234.
98. Chug, H., S. Trakhanov, ..., D. Görlich. 2015. Crystal structure of the metazoan Nup62•Nup58•Nup54 nucleoporin complex. *Science.* 350:106–110.
99. Miyagawa, M., T. Naito, ..., S. Usami. 2013. Targeted exon sequencing successfully discovers rare causative genes and clarifies the molecular epidemiology of Japanese deafness patients. *PLoS One.* 8:e71381.
100. Chen, S., C. Dong, ..., Y. Liu. 2016. Targeted next-generation sequencing successfully detects causative genes in Chinese patients with hereditary hearing loss. *Genet. Test. Mol. Biomarkers.* 20:660–665.

101. Sotomayor, M., D. P. Corey, and K. Schulten. 2005. In search of the hair-cell gating spring elastic properties of ankyrin and cadherin repeats. *Structure*. 13:669–682.
102. Oroz, J., A. Valbuena, ..., M. Carrión-Vázquez. 2011. Nanomechanics of the cadherin ectodomain: “canalization” by Ca<sup>2+</sup> binding results in a new mechanical element. *J. Biol. Chem.* 286:9405–9418.
103. Sotomayor, M., W. A. Weihofen, ..., D. P. Corey. 2010. Structural determinants of cadherin-23 function in hearing and deafness. *Neuron*. 66:85–100.
104. Carrion-Vazquez, M., A. F. Oberhauser, ..., J. M. Fernandez. 2000. Mechanical design of proteins studied by single-molecule force spectroscopy and protein engineering. *Prog. Biophys. Mol. Biol.* 74:63–91.
105. Brockwell, D. J., G. S. Beddard, ..., S. E. Radford. 2005. Mechanically unfolding the small, topologically simple protein L. *Biophys. J.* 89:506–519.
106. Li, M. S. 2007. Secondary structure, mechanical stability, and location of transition state of proteins. *Biophys. J.* 93:2644–2654.
107. Hughes, M. L., and L. Dougan. 2016. The physics of pulling polyproteins: a review of single molecule force spectroscopy using the AFM to study protein unfolding. *Rep. Prog. Phys.* 79:076601.
108. Carrion-Vazquez, M., H. Li, ..., J. M. Fernandez. 2003. The mechanical stability of ubiquitin is linkage dependent. *Nat. Struct. Biol.* 10:738–743.
109. Izrailev, S., S. Stepanians, and B. Isralewitz. 1998. Steered molecular dynamics. In *Computational Molecular Dynamics: Challenges, Methods, Ideas*. P. Deuffhard, J. Hermans, and R. D. Skeel, eds, pp. 39–65.
110. Karplus, M., and G. A. Petsko. 1990. Molecular dynamics simulations in biology. *Nature*. 347:631–639.
111. Evans, E., and K. Ritchie. 1997. Dynamic strength of molecular adhesion bonds. *Biophys. J.* 72:1541–1555.
112. Narui, Y., and M. Sotomayor. 2018. Tuning inner-ear tip-link affinity through alternatively spliced variants of protocadherin-15. *Biochemistry*. 57:1702–1710.
113. Letunic, I., and P. Bork. 2018. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* 46:D493–D496.
114. Källberg, M., H. Wang, ..., J. Xu. 2012. Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* 7:1511–1522.
115. Elledge, H. M., P. Kazmierczak, ..., U. Müller. 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. USA*. 107:10708–10712.
116. Russell, R. B., and G. J. Barton. 1992. Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins*. 14:309–323.
117. Roberts, E., J. Eargle, ..., Z. Luthey-Schulten. 2006. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics*. 7:382.
118. Alagramam, K. N., H. Yuan, ..., R. J. Smith. 2001. Mutations in the novel protocadherin PCDH15 cause Usher syndrome type 1F. *Hum. Mol. Genet.* 10:1709–1718.
119. Ahmed, Z. M., R. Goodyear, ..., T. B. Friedman. 2006. The tip-link antigen, a protein associated with the transduction complex of sensory hair cells, is protocadherin-15. *J. Neurosci.* 26:7022–7034.
120. Alagramam, K. N., R. J. Goodyear, ..., G. P. Richardson. 2011. Mutations in protocadherin 15 and cadherin 23 affect tip links and mechanotransduction in mammalian sensory hair cells. *PLoS One*. 6:e19183.
121. Kachar, B., M. Parakkal, ..., P. G. Gillespie. 2000. High-resolution structure of hair-cell tip links. *Proc. Natl. Acad. Sci. USA*. 97:13336–13341.
122. Corey, D. P., D. Ó. Maoiléidigh, and J. F. Ashmore. 2017. *Mechanical Transduction Processes in the Hair Cell*. Springer, Cham, Switzerland, pp. 75–111.
123. Basu, A., S. Lagier, ..., A. J. Hudspeth. 2016. Direct mechanical stimulation of tip links in hair cells through DNA tethers. *eLife*. 5:e16041.
124. Takahashi, H., F. Rico, ..., S. Scheuring. 2018.  $\alpha$ -Helix unwinding as force buffer in spectrins. *ACS Nano*. 12:2719–2727.
125. Bosher, S. K., and R. L. Warren. 1978. Very low calcium content of cochlear endolymph, an extracellular fluid. *Nature*. 273:377–378.
126. Salt, A. N., N. Inamura, ..., A. Vora. 1989. Calcium gradients in inner ear endolymph. *Am. J. Otolaryngol.* 10:371–375.
127. Shepherd, G. M., and D. P. Corey. 1994. The extent of adaptation in bullfrog saccular hair cells. *J. Neurosci.* 14:6217–6229.
128. Maeda, R., K. S. Kindt, ..., T. Nicolson. 2014. Tip-link protein protocadherin 15 interacts with transmembrane channel-like proteins TMC1 and TMC2. *Proc. Natl. Acad. Sci. USA*. 111:12907–12912.
129. Xiong, W., N. Grillet, ..., U. Müller. 2012. TMHS is an integral component of the mechanotransduction machinery of cochlear hair cells. *Cell*. 151:1283–1295.
130. Morgan, C. P., H. Zhao, ..., P. G. Barr-Gillespie. 2018. TRPV6, TRPM6 and TRPM7 do not contribute to hair-cell mechanotransduction. *Front. Cell. Neurosci.* 12:41.
131. Beurg, M., W. Xiong, ..., R. Fettiplace. 2015. Subunit determination of the conductance of hair-cell mechanotransducer channels. *Proc. Natl. Acad. Sci. USA*. 112:1589–1594.
132. Kung, C. 2005. A possible unifying principle for mechanosensation. *Nature*. 436:647–654.
133. Kung, C., B. Martinac, and S. Sukharev. 2010. Mechanosensitive channels in microbes. *Annu. Rev. Microbiol.* 64:313–329.
134. Katta, S., M. Krieg, and M. B. Goodman. 2015. Feeling force: physical and physiological principles enabling sensory mechanotransduction. *Annu. Rev. Cell Dev. Biol.* 31:347–371.
135. Corey, D. P., and J. R. Holt. 2016. Are TMCs the mechanotransduction channels of vertebrate hair cells? *J. Neurosci.* 36:10921–10926.
136. Wu, Z., and U. Müller. 2016. Molecular identity of the mechanotransduction channel in hair cells: not quiet there yet. *J. Neurosci.* 36:10927–10934.
137. Ballesteros, A., C. Fenollar-Ferrer, and K. J. Swartz. 2018. Structural relationship between the putative hair cell mechanotransduction channel TMC1 and TMEM16 proteins. *eLife*. 7:e38433.
138. Pan, B., N. Akuz, ..., J. R. Holt. 2018. TMC1 forms the pore of mechanosensory transduction channels in vertebrate inner ear hair cells. *Neuron*. 99:736–753.e6.
139. Barr-Gillespie, P. G. 2018. Honing in on TMC as the hair cell’s transduction channel. *Neuron*. 99:628–629.
140. Indzhukulian, A. A., R. Stepanyan, ..., G. I. Frolenkov. 2013. Molecular remodeling of tip links underlies mechanosensory regeneration in auditory hair cells. *PLoS Biol.* 11:e1001583.
141. Libé-Philippot, B., V. Michel, ..., C. Petit. 2017. Auditory cortex interneuron development requires cadherins operating hair-cell mechano-electrical transduction. *Proc. Natl. Acad. Sci. USA*. 114:7765–7774.
142. Barnes, J. M., L. Przybyla, and V. M. Weaver. 2017. Tissue mechanics regulate brain development, homeostasis and disease. *J. Cell Sci.* 130:71–82.
143. Astuto, L. M., J. M. Bork, ..., W. J. Kimberling. 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.
144. Bork, J. M., L. M. Peters, ..., R. J. Morell. 2001. Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23. *Am. J. Hum. Genet.* 68:26–37.
145. Oshima, A., T. Jaijo, ..., W. J. Kimberling. 2008. Mutation profile of the CDH23 gene in 56 probands with Usher syndrome type I. *Hum. Mutat.* 29:E37–E46.