Structural, Biochemical and *In Silico* Studies of Atypical Cadherins Involved in Inner Ear Hair Cell Organization and Mechanotransduction

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Elakkiya Tamilselvan

Biophysics Graduate Program

The Ohio State University

2023

Dissertation Committee

Marcos Sotomayor, Advisor

James Jontes

Steffen Lindert

Kotaro Nakanishi

Abstract

Cadherins are a family of large transmembrane glycoproteins instrumental in facilitating organ formation during morphogenesis in vertebrates and invertebrates. At the cellular level, they are involved in adhesion, signaling, recognition, mechanotransduction, and motility. In the modern classification of the cadherin superfamily, classical cadherins with five extracellular cadherin (EC) repeats as well as clustered and non-clustered δ-protocadherins with six or seven EC repeats have been well-studied and their homophilic/heterophilic interactions with molecules on the same (*cis*) cell or opposite (*trans*) cells have been characterized. Complexity arises when the number of EC repeats increases with diverse Ca^{2+} coordination at linker regions between two consecutive EC repeats. In larger cadherins, such as cadherin-23 (CDH23), protocadherin-15 (PCDH15) and cadherin epithelial growth factor (EGF) Laminin-G (LAG or LamG) seven pass G-type receptor-1 (CELSR1), the structural flexibility afforded by different Ca^{2+} coordination plays determinant roles in their adhesion capacity during inner-ear mechanotransduction and planar cell polarity (PCP).

CDH23 and PCDH15, each with 27 and 11 EC repeats, connect two adjacent hair-like protrusions known as stereocilia together atop of a hair cell, the primary mechanosensory cell in the inner ear. Through heterophilic interactions between their first

two N-terminal EC repeats, CDH23 and PCDH15 form a filament known as the tip link. In response to sound, stereocilia undergo displacement and the tip link experiences tension, which opens the ion-conducting mechanotransduction channels on the tip-link's lower end to send signals to the brain. The heterophilic *trans* tetrameric complex formed by CDH23 and PCDH15, and the *cis* interactions along the length of PCDH15 have been well-studied in the past but full-length ectodomain structures and high-resolution structural models of complete CDH23 and PCHD15 ectodomain have not been resolved. Here, we present the structure of a PCDH15 segment and protocols for expression and purification of full-length ectodomains that allowed us to obtain several 2-D classes and a cryogenic-electron microscope (cryo-EM) envelope of PCDH15 along with cryo-EM images of CDH23. An *in silico* constructed model of PCDH15 fitted well to the *ab inito* envelope and 2D classes. We similarly use structures and AlphaFold2 (AF2) predictions to build a CDH23 ectodomain model that is representative of some of the conformations observed using cryo-EM. Unlike the PCDH15 dimeric ectodomain, the conformations of CDH23 are numerous and lack stable *cis* interactions. While this posed a difficulty, our high-resolution structural models of CDH23 and PCDH15 provide the opportunity to study these proteins in atomic-level detail for the first time in response to force as expected in the inner ear.

Another cadherin that is also found in the inner ear but functions in a different capacity is CELSR1. Forming a transmembrane protein complex with two other core PCP proteins in the apical region of cells, CELSR1 is involved in the organization of hair cells in the cochlea to maximize efficiency of signal transduction. CELSR1, an adhesion G-

protein coupled receptor (GPCR) family member, is one of the most highly conserved cadherins since multicellularity arose in evolution. With nine EC repeats in the ectodomain, CELSR1 is involved in homophilic interactions that help propagate PCP, a property in which cells undergo coordinated alignment along the tissue plane. We have used a myriad of biophysical and biochemical techniques along with crystal structures of EC1-4 and EC4-7 to show that CELSR1 is weakly adhesive and that there is a flexible linker region in its ectodomain. We also propose the existence of a *cis* mediated interaction at the C-terminal end of the CELSR1 cadherin ectodomain similar to what is observed for PCDH15. In addition, we have also built two hybrid models for CELSR1 that resemble *in vitro* states of its paralog, CELSR2, observed using high-speed atomic force microscopy. Our results support a signaling function for CELSR1.

Overall, our high-resolution structural models of CDH23, PCDH15, and CELSR1 based on cryo-EM images and constructed using structures of smaller fragments solved by X-ray crystallography and AF2 predictions provide a framework for studying other elongated cadherins in the superfamily.

Dedication

This work is dedicated to my mother for her love and affection, indomitable spirit, and unwavering support

Acknowledgments

A lot of people have supported me throughout the years, but none come close to my advisor Dr. Marcos Sotomayor, whom I would like to thank from the bottom of my heart. Like a typical advisor in academia, he has always been extremely busy. But even during those times, he has never not found the time to provide feedback or answer questions that I have had over the years. Forever equipped with a smile, his enthusiasm and dedication towards advancing research in the field of hearing is unparalleled. I also would like to express my sincere gratitude to my committee members, Dr. James Jontes, Dr. Steffen Lindert and Dr. Kotaro Nakanishi for their valuable feedback, support, and guidance. Their individual contributions to science have undoubtedly been an encouragement for me to persevere in this journey.

A wonderful lab environment with plenty of friendly faces past and present have made my time in Sotomayor lab memorable. A few honorable mentions include Dr. Deepanshu Choudhary, Dr. Pedro De-la-Torre, Dr. Yoshie Narui, Dr. Mich Gray, Dr. Brandon Neel, and Dr. Sharon Cooper without whom I would not be able to stand where I am today. They have been my second mentors after Dr. Sotomayor. I would also like to thank Dr. Sanket Walujkar for teaching me the basics in VMD and helping me set up

simulations in NAMD using the world-class supercomputers that we were fortunate enough to access throughout my whole time in the lab.

I would also like to thank my collaborators inside and outside the lab, Carter Wheat, Jasanvir Sandhu, Harsha Mandayam Bharathi, Harper Smith, Dr. Felix Rico (Aix-Marseille University), Dr. Tian-Min Fu, Dr. Wang Zheng (Harvard) and Dr. Jeffrey Holt (Harvard) for being patient and resourceful in our projects together. Working with them has opened my mind to their scientific mind prowess and instilled within myself a competitive spirit to succeed.

I would not have had the motivation if not for a few wonderful people, Jessica Trevail, Daisy Alvarado, Wei Hsiang Weng, Shounak Mukherjee, Dr. Collin Nisler, Joe Sudar, and Jeffrey Lotthammer who have made my time enjoyable and remarkable in Columbus. Thank you for tolerating my quirks and providing me with the emotional support throughout the years.

Last but not least, I would like to thank my wonderful family to whom I owe everything in life. A special shout out goes to my brother (Vijay Aadhithan), my aunt (Dr. Devaki Rajasekar), and my cousins (Vikram and Sri Aarani).

Vita

2011-2015 …….	B.Tech., Biotechnology, Anna University, India
2016-Present …….	Graduate Research Assistant, Department of Chemistry and Biochemistry, The Ohio State University

Publications

1. Tamilselvan E, Sotomayor M. CELSR1, a core planar cell polarity protein, features a weakly adhesive and flexible cadherin ectodomain. *Structure*. 2023 (Accepted for publication)

2. Zheng W, Rawson S, Shen Z, et al. TMEM63 proteins function as monomeric high-threshold mechanosensitive ion channels. *Neuron*. 2023:1-16. doi:10.1016/j.neuron.2023.07.006

3. Gray ME, Johnson ZR, Modak D, Tamilselvan E, Tyska MJ, Sotomayor M. Heterophilic and homophilic cadherin interactions in intestinal intermicrovillar links are species dependent. *PLOS Biol*. 2021;19(12):e3001463. doi:10.1371/JOURNAL.PBIO.3001463

4. Hudson JD, Tamilselvan E, Sotomayor M, Cooper SR. A complete Protocadherin-19 ectodomain model for evaluating epilepsy-causing mutations and potential protein interaction sites. *Structure*. 2021;29(10):1128-1143.e4. doi:10.1016/J.STR.2021.07.006

5. Choudhary D, Narui Y, Neel BL, et al. Structural determinants of protocadherin-15 mechanics and function in hearing and balance perception. *Proc Natl Acad Sci U S A*. 2020;117(40):4837-24848. doi:10.1073/pnas.1920444117

6. Sasidharan S, Achuthan T, Tamilselvan E, Raju S, Rengaraju B. Synthesis of copolymer from lactic acid-polyethylene terephthalate (LA-PET) copolymerization. *African J Biotechnol* . 2015;14(11):953-960. doi:10.5897/AJB2014.14301

Fields of Study

Major Field: Biophysics

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Evolution and function of cadherins

Cadherins form a superfamily of transmembrane glycoproteins with tandem extracellular cadherin (EC) repeats that primarily promote Ca^{2+}-dependent cell-cell adhesion. Since the evolutionary jump of animals to multicellularity 600 million years ago, the diversification of cadherins has become evident as the animal kingdom expanded to accommodate higher order functions such as morphogenesis, cell signaling, cell recognition, and mechanotransduction[1]. For instance, the *Trichoplax adhaerens* (*Ta*), the simplest animal, has only 8 cadherin genes whereas the modern-day *Homo sapiens* (*Hs*) has 113[2]. In ancestral metazoans, the cadherin family of proteins was comprised of only the classical (CDH), Flamingo (CELSR), Dachsous (Dchs), FAT and FAT-like subfamilies but in present-day chordates, the family also includes protocadherins (PCDH) that evolved from FAT and cadherin related (CDHR) proteins[3] (Fig. 1.1). Among superfamily members, classical cadherins have undergone extensive changes from their ancestral versions through a drastic reduction in the number of EC repeats from up to 32 to only 5 or 7 but with conserved cytoplasmic domains that interact with armadillo-type catenins across epithelial cell borders[1,3–5]. Together, classical cadherins play an important role during development and maintenance of adult tissue architecture. PCDHs (Clustered - α, β, γ and non-clustered - $\delta 1$, $\delta 2$), on the other hand, underwent the largest expansion in chordates, have either 6 or

7 EC repeats, and are involved in synaptic development, neural connectivity and cell proliferation[3,4,6]. FAT and Dchs are the largest of the superfamily with 34 and 27 EC repeats and are involved in planar cell polarity (PCP) during morphogenesis along with the flamingo proteins (CELSR)[7,8]. CDHR proteins play unique roles in promoting organ specific functions. While several of these cadherin family members are ubiquitously present in diverse cell types and execute their cell-specific function through cell-cell adhesion, my dissertation will be mainly focused on CDH23 (CDHR23) and PCDH15 (CDHR15), two proteins belonging to the CDHR family and CELSR1, belonging to the most conserved flamingo family in mammalian inner ear development and function.

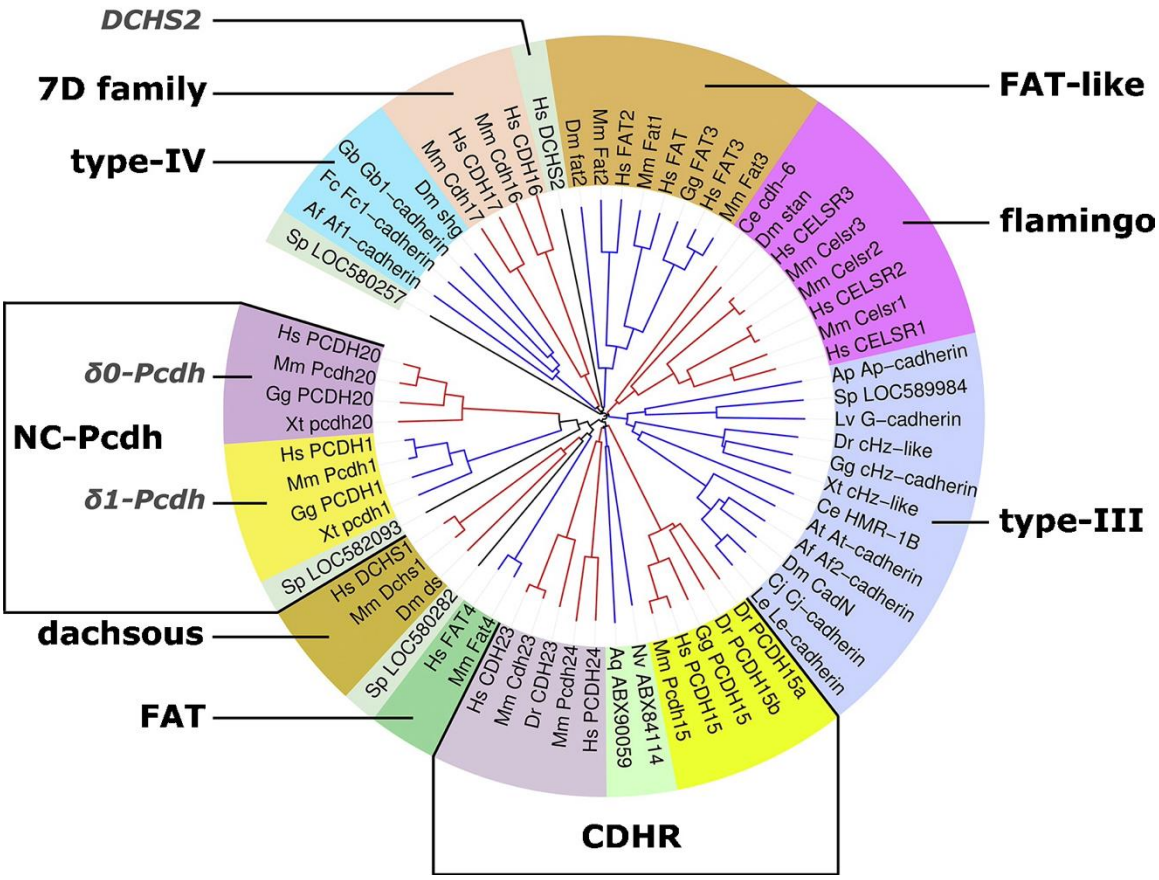


Figure 1.1. The cadherin superfamily

Phylogenetic tree of the cadherin superfamily based on the sequence homology of the first seven N-terminal EC repeats. Atypical cadherins, CDH23, PCDH15 and CELSR1 are classified under the CDHR and the flamingo subfamilies of cadherins, respectively. Adapted from Hulpiau *et al*. 2008[1].

Mammalian inner ear anatomy

Anatomically, the ear is made of a pinna or the external auditory canal connecting to the tympanic membrane or ear drum in the outer ear followed by the ossicles, malleus, incus, and stapes in the middle ear and the fluid-filled inner ear[9,10]. The inner ear within the temporal bone is made up of a bony labyrinth surrounded by petrosal bone and a membranous labyrinth (Fig. 1.2). The bony labyrinth has three distinct semicircular canals, the vestibule, and the cochlea. Within the bony labyrinth lies the membranous labyrinth with an inferior division, made of cochlear duct and saccule of the membranous vestibule and a posterior division, made of utricle of the vestibule and the three semi-circular canals[9,10]. In mammals, the inner ear mediates hearing (cochlea) and balance (vestibular system). The organ volume ranges from 1 mm^3 to over a 1,000 mm^3 and can detect frequencies ranging from below 20 Hz to above 20 kHz in the infrasonic and ultrasonic ranges[9].

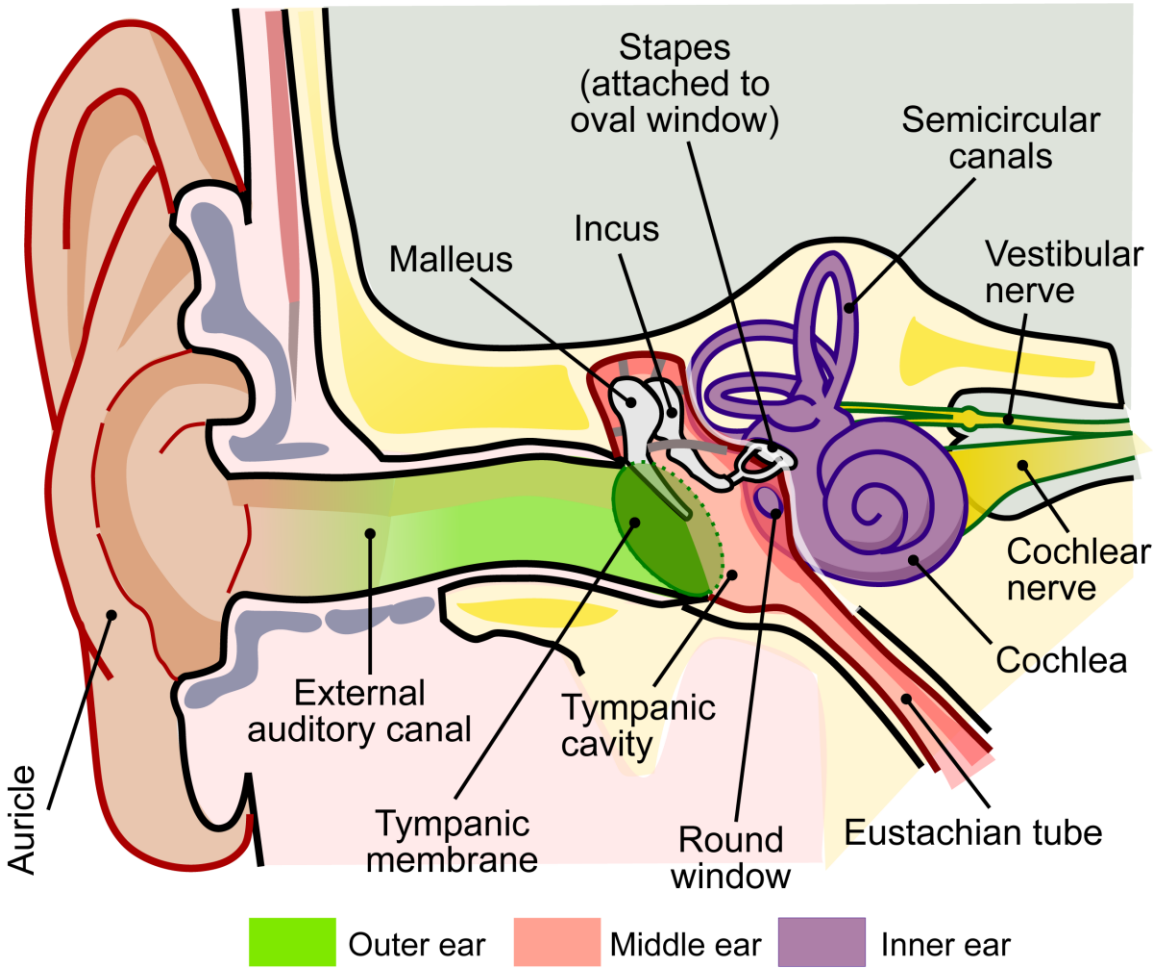


Figure 1.2. Ear anatomy

The human ear consists of the external auditory canal, the tympanic membrane followed by the middle ear made up of three ossicles, malleus, incus, and stapes, and the inner ear. In the inner ear, the spiral shaped cochlea houses the organ of Corti, the primary auditory sensing organ and the vestibular apparatus including the utricle, the saccule, and the semi-circular canals[9,10]. Sound vibrations that are transmitted from the external ear through vibrations of the tympanic membrane, the bone ossicles and the oval window, cause displacement of the fluid filled duct inside the cochlea. The sensory hair cells embedded within the organ of Corti detect the vibrations and generate a signal to be transmitted to the brain for sound perception through the cochlear nerve (Adapted from Wikipedia commons).

Organ of Corti and hair cells

The cochlea, the spiral shaped organ responsible for hearing in the inner ear, has three compartments, scala vestibuli and scala tympani separating the bony labyrinth from the cochlear duct, and the scala media[11]. The scala media is filled with endolymph while the scala vestibuli and scala tympani are filled with perilymph (Fig. 1.3. *A*). The stria vascularis, a vascularized epithelium along the cochlear walls, replenishes the endolymph. The perilymph is rich in sodium whereas the endolymph is rich in potassium. A high K^+ concentration results in a positive endocochlear potential (~ +80 mV in endolymph; ~ 0 mV in perilymph[12]) essential for proper ion conduction in response to a mechanical stimuli. Two membranes, basilar and Reissner's, separate the scala vestibuli and scala tympani from the scala media[11]. The organ of Corti, the primary sensory organ facilitating hearing, is placed atop the basilar membrane running along the spiral-shaped cochlea from the base to the apex (~ 34 mm long in humans). The basilar membrane has varying physical properties along the length of the cochlea with the membrane being relatively narrow, light, and taut at the base while at the apex it is broader, heavier, and more flaccid[13]. Throughout the length of the organ of Corti, two types of cells, mechanosensory and supporting cells, are present. The mechanosensory hair cells are arranged in three outer rows and a single inner row (Fig. 1.3. *B*). The stiffness and size of the hair cells along the length of the cochlea enable them to respond to distinct sound frequencies creating a tonotopic gradient with hair cells at the base responding to higher frequencies and hair cells at the apex responding to lower frequencies[14].

There are about ~ 3,500 inner (IHC) and ~ 12,000 outer hair cells (OHC) on the basilar membrane[15]. Anatomically, the IHC are bulbous, and each connect to ten to twenty different afferent nerve fibres while the OHC are cylindrical and connect to efferent neurons to adapt to functions of mechanosensation and signal amplification respectively[14].

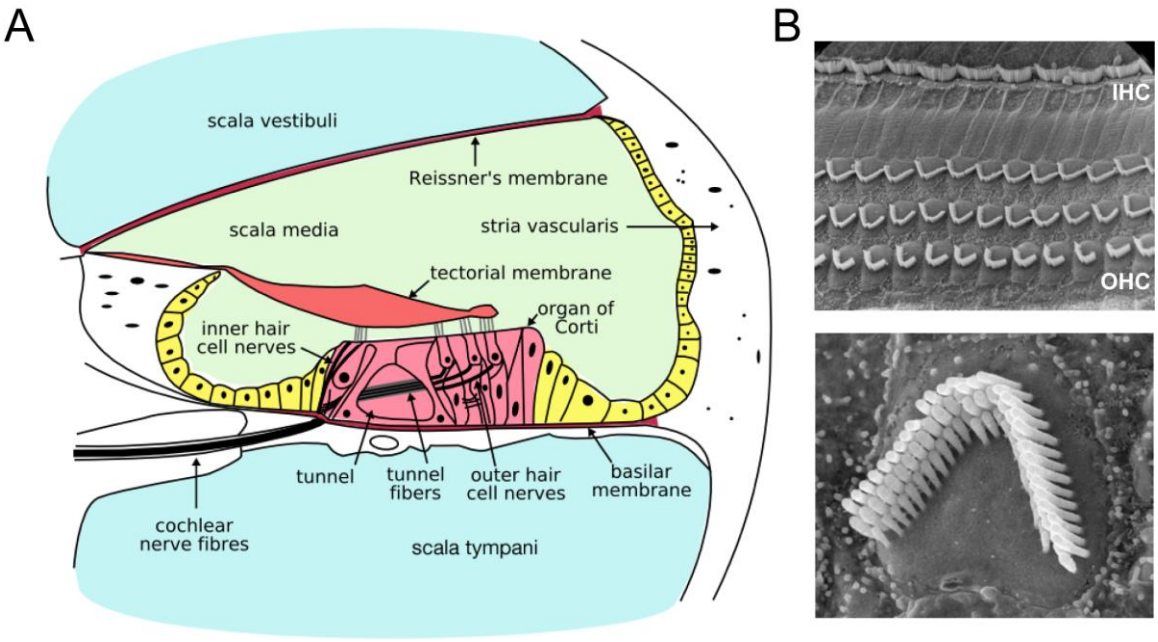


Figure 1.3. The organ of Corti houses sensory hair cells necessary for sound perception

(A) The cochlea in the middle ear contains the auditory organ, the organ of Corti, surrounded by two fluid filled ducts, scala vestibuli and scala tympani. The Reissner's membrane and the basilar membrane separate the organ of Corti filled with endolymph from scala vestibuli and scala tympani respectively (Adapted from Wikipedia commons). *(B) Top*: Inside the organ of Corti, sensory hair cells lie on top of the basilar membrane differentiated into three rows of outer hair cells (OHC) and a single row of inner hair cells (IHC). *Bottom*: On each sensory hair cell, actin filled stereocilia are present which exhibit a staircase like pattern arranged in a W-shape across OHCs and a flat shape in IHCs. The tallest stereocilia in the OHC are embedded into the tectorial membrane whereas the stereocilia in IHCs are free (Credit: Hair cell of inner ear, Dr David Furness).

Atop each hair cell lie key mechanosensory hair-like protrusions, the stereocilia, which are highly ordered actin-filled protrusions anchored to the cuticular plate on the apical surface (Fig. 1.3 *B*). The stereocilia are tapered at the bottom in the ankle regions to improve flexibility and to enable shearing during sound transduction[16]. IHCs have flat stereocilia bundles while the OHCs have W-shaped bundles with the tallest stereocilia row attached to the tectorial membrane. There are around 100 stereocilia in each cochlear hair cell arranged in a graded height with a staircase like pattern in three rows[17]. Stereocilia number and height vary inversely along the length of the cochlea from the base to apex. At the base, the stereocilia are shorter and larger in number while at the apex, the number decreases but with an increase in height[17]. Stereocilia are connected by a proteinaceous filament known as the tip link made up of atypical cadherins, cadherin-23 (CDH23) and protocadherin-15 (PCDH15)[18–23] (Fig. 1.4 *A*). The PCDH15 ectodomain extends from the top of a shorter stereocilium to form heterophilic interactions with CDH23 from the adjacent taller stereocilium[19] (Fig. 1.4 *B*). In the shorter stereocilium, PCDH15 connects to transmembrane channel-like (TMC1/2), lipoma HMGIC fusion partner-like 5 (LHFPL5) and transmembrane inner ear (TMIE) proteins to form the mechanotransduction (MET) complex necessary for hair cell mechanotransduction[24].

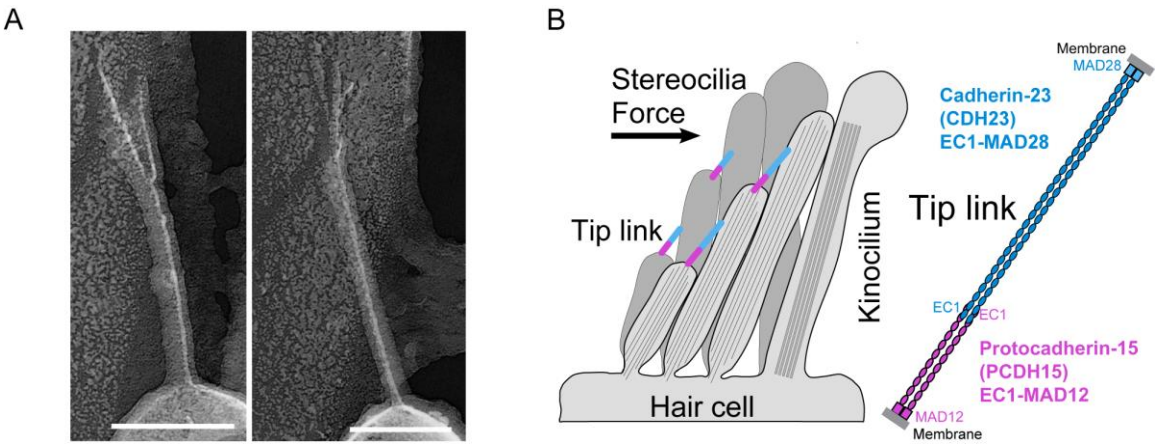


Figure 1.4. Tip links made of CDH23 and PCDH15 connect adjacent stereocilia together along the excitatory axis

(A) On top of the hair cell, a short stereocilium is connected to the adjacent taller stereocilium by a thin filamentous element known as the tip link[25]. Scale bar: 100 nm (Adapted from Kachar *et al*. 2000[25]) *(B)* The tip link made of atypical cadherins, CDH23 and PCDH15, experiences tension upon displacement of stereocilia in response to a sound signal to open mechanosensitive ion channels next to PCDH15 in the lower stereocilium.

Process of hearing

During sound perception, sound waves travel through the external auditory canal in the outer ear and vibrate the tympanic membrane. In the middle ear, the vibrations are transmitted to the malleus, incus, and stapes, which amplify sound signals due to a decrease in surface area resulting in a pressure increase that impinges on the oval window that acts as a piston to induce pressure changes in the fluid-filled cochlea. The pressure changes inside the cochlea oscillate the basilar membrane in an up and down motion[26–28]. The differential movement of the hair cells between the basilar and the tectorial membrane bend the stereociliary bundle on their apical layer[26–28]. The resulting displacement leads to stretching of tip links made of CDH23 and PCDH15 that ends up opening the MET complex downstream of PCDH15 allowing for hair cell depolarization due to an influx of potassium ions driven by voltage[15]. Hair-cell depolarization leads to the release of glutamate at the base of the hair cells resulting in the generation of an action potential in the nerve fibers that transmit signals to the brain (Fig. 1.5). The round window below the oval window acts as a pressure relief valve and accommodates for the inner ear expansion during sound transmission through the cochlea[26–28].

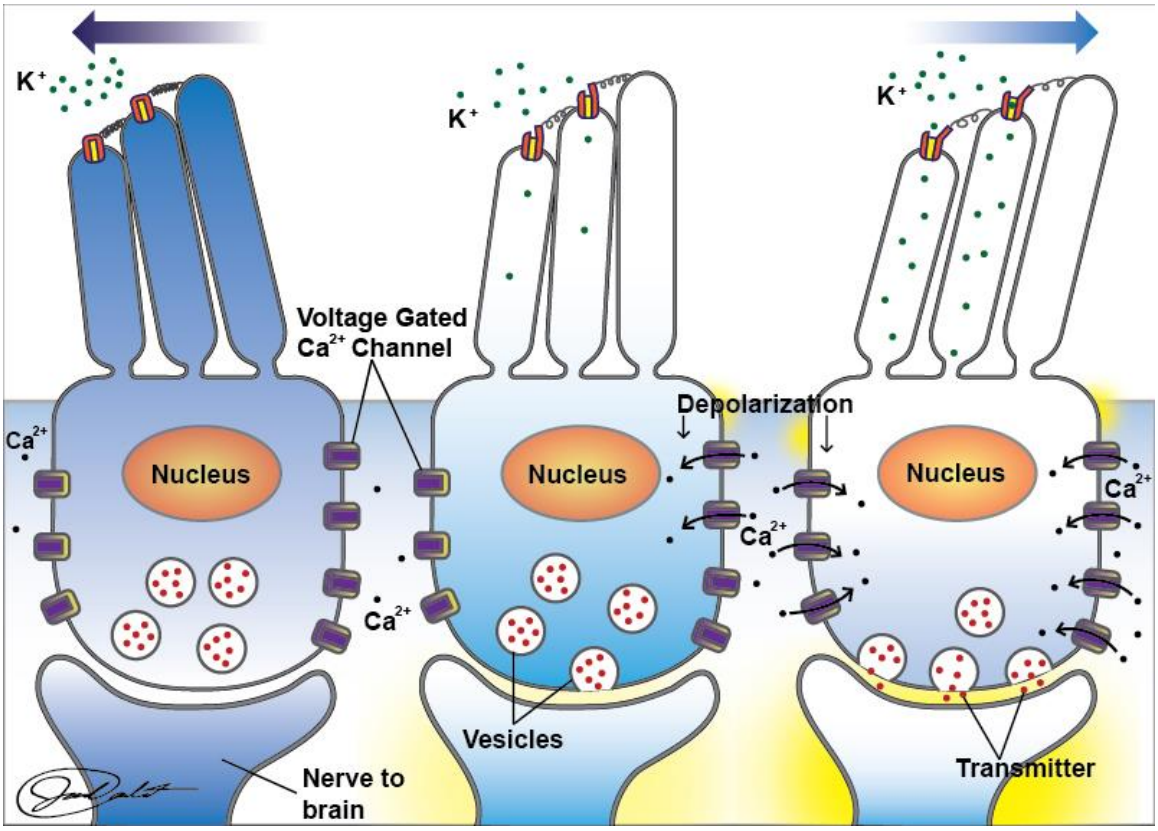


Figure 1.5. Schematic of mechanotransduction in hair cells

Under no external force, mechanosensitive channels on the shorter stereocilia are mostly closed (*middle*). In the presence of a sound signal, the stereocilia on the hair cells undergo displacement in the positive direction tugging on the tip links resulting in the opening of ion channels next to PCDH15 (*right*). Potassium ions are conducted leading to the depolarization of hair cells which in turn releases neurotransmitters thereby generating an action potential in the auditory nerve that connects to the brain. Bundle movement in the opposite direction completely closes channels (*left*). (Credit: Produced by BYU-Idaho student Jared Cardinet F14).

How CDH23 and PCDH15 play a role in the process of hearing

CDH23 and PCDH15, the components of tip links, have 38 extracellular cadherin (EC) repeats in total[20]. Each EC repeat has similar topology but varies in sequence. The EC repeats have 7 β-strands labeled A to G with a linker region connecting two consecutive repeats (Fig. 1.6 *A*). At the linker region, three Ca^{2+} ions are canonically coordinated by conserved motifs of negatively charged residues, XEX, DXE, DXNDN, DXD and XDX[29–31] (Fig. 1.6 *B*). In some linker regions along the length of the tip link, Ca^{2+} coordination is affected by substitution of amino acids in the canonical motifs rendering flexibility to those regions[32,33]. Functionally, Ca^{2+} protects cadherin repeats against degradation by proteases and offers rigidity to the structure[34,35]. Thus, in the inner ear, the functions of hearing and balance can be readily disrupted by the chelation of Ca^{2+} ions using EGTA or similar chelators.

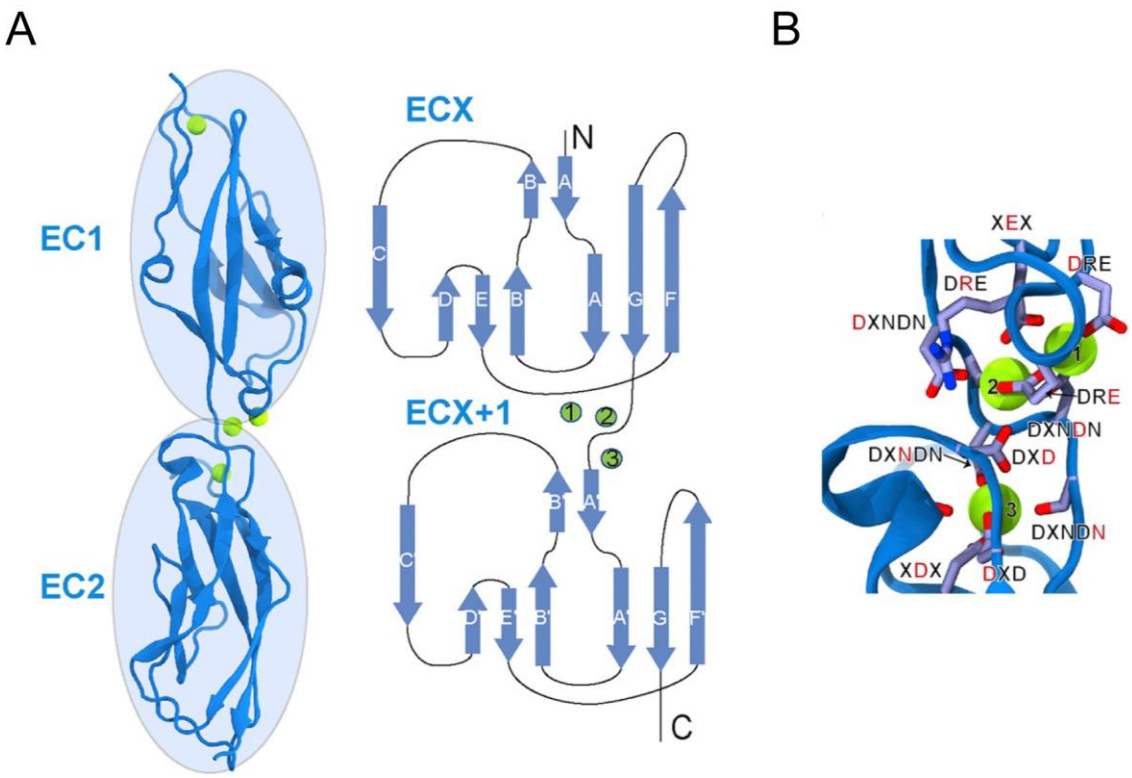


Figure 1.6. Typical EC structure and topology

(A) Each EC repeat has 7 β-strands labeled A to G and coordinate three Ca^{2+} ions in the linker region between two consecutive EC repeats[29]. *(B)* The canonical Ca^{2+} coordinating motifs, XEX, DXE, DXNDN, DXD and XDX are highlighted.

Structurally, the tip link is a 150 to 180 nm long filament made of a parallel *cis* dimer of PCDH15 on the lower side accounting for 1/3rd of the length of the tip link and two CDH23 molecules on the upper side accounting for 2/3rd of the length[18,19,25] (Fig. 1.7 *A*). The two most N-terminal repeats of CDH23 and PCDH15 interact in a heterophilic manner to form the tetrameric "handshake" connecting the two filaments together[36,37] (Fig. 1.7 *B*). Electron microscope images of PCDH15 show the structure to possess 2.5 helical turns[25] with two dimerization points along its length[37–40]. The *cis* dimer in PCDH15 is primarily promoted by a Ca^{2+}-dependent EC2-3 X-dimer conformation[37,38] and a Ca^{2+}-independent dimer formed by a membrane adjacent domain (MAD) in the C-terminus of the ectodomain[37,38,40,41] (Fig. 1.7 *C*). While the EC2-3 X-dimer facilitates interactions with CDH23 EC1-2 to form a heterotetramer, the dimerization induced by MAD adopting the ferredoxin like fold in the C-terminus of PCDH15 prevents the ends from being splayed apart[37].

Although there is less structural information available on the whole CDH23 ectodomain, structures of fragments of CDH23 (EC1, EC2, EC3, EC6, EC7, EC8, EC12, EC13, EC14, EC17, EC18, EC19, EC20, EC21, EC22, EC23, EC24 and EC25) have been published[29]. In solution, unlike the PCDH15 EC1-MAD12 ectodomain, CDH23 EC1-27 fragments overwhelmingly favor the existence of a monomer rather than a *cis* dimer[42,43] but lack of MAD28 in those in-solution experiments puts an inconclusive light on the *cis* dimer capabilities of the CDH23 ectodomain from their C-terminal end.

In the absence of mechanical displacement, a resting tension of ~ 10 to 40 pN acts on the tip-link filaments[44]. However, in the presence of a sound signal, according to the frequency of the sound and position along the tonotopic axis of the length of the cochlea, displacement of the stereocilia bundle adds a proportional tension on the tip-link filaments. The intensity of the sound signal and the resulting added tension along the tip link determine the open channel probability of the mechanosensitive ion channels downstream of PCDH15. Here, CDH23 and PCDH15, might form the gating spring component or be in series with it[15,45].

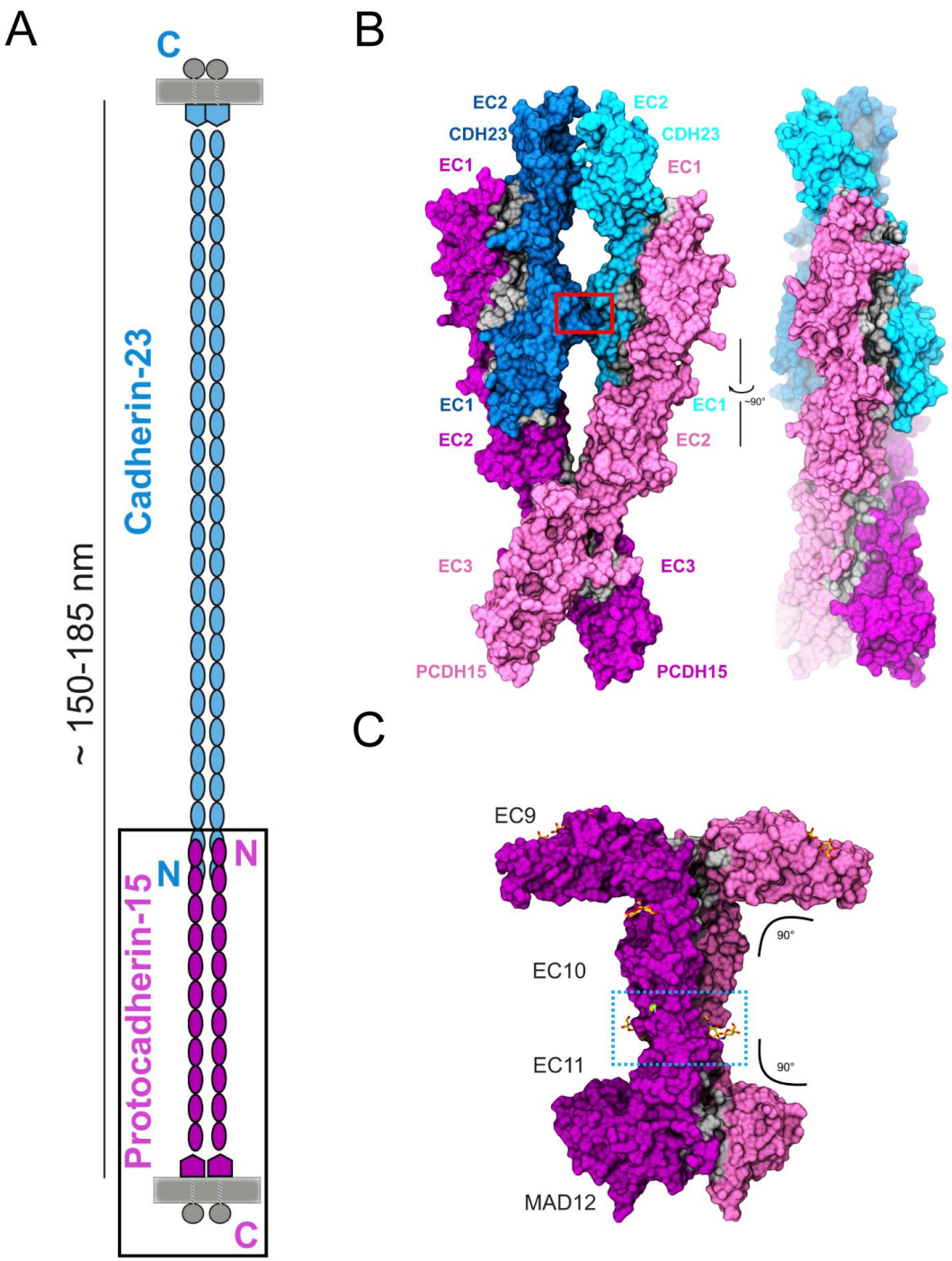


Figure 1.7. Handshake interactions within the tip link

(A) CDH23, consisting of 27 EC repeats and a C-terminal MAD and PCDH15, consisting of 11 EC repeats and a C-terminal MAD are around ~ 120 nm and ~ 50 nm in length (Adapted from Choudhary *et al*. 2020[37]). *(B)* The most N-terminal EC repeats of CDH23 and PCDH15 interact to form a heterophilic tetramer. The X-dimer formed by EC2-3 of

PCDH15 primarily facilitates this interaction. (*C*) *Cis* interactions formed between C-terminal EC10-MAD12 fragments in PCDH15 (Adapted from Choudhary *et al*. 2020[37]).

Inner ear morphogenesis and organization of hair cells

In addition to CDH23 and PCDH15, another atypical cadherin, CELSR1, is critical for the inner ear. CELSR1 is involved in the development of the inner ear morphology that gives rise to the intricate and uniform orientation of the hair cells in the cochlea and the vestibular organs. During morphogenesis, the inner ear originates from the otic placode, a thick epithelial layer on the surface of the non-neural ectoderm adjacent to the caudal part of the hindbrain (Fig. 1.8 *A*). Through a series of complex morphogenetic events that involve invagination driven by apical constriction, a closed otic vesicle leads to the formation of an otocyst that further develops into a dorsal vestibular and ventral auditory components seen in the mature inner ear structure[46] (Fig. 1.8 *B*). At the end of the otocyst stage, three primary cell types are established: neural, sensory and nonsensory cells. While neural and nonsensory cells develop into nerves and structures such as the endolymphatic duct and the semi-circular canals to support the function of the inner ear, sensory cells develop into hair cells, the primary mechanosensory cells enveloped by several non-sensory supporting cells[47]. The sensory cell types include OHCs and IHCs. The non-sensory supporting cell (SC) types include border cells (BC), inter phalangeal cells (IPhC), outer (Opcs) and inner pillar cells (Ipcs) and Deiter cells (Dc1,2). Along the length of the organ of Corti, hair cells surrounded by supporting cells are uniformly aligned along the medial-lateral (neural-abneural) axis, a phenomenon clearly visualized by the asymmetric orientation of the stereociliary tuft to the abneural edge[48] (Fig. 1.8 *C*). This arrangement allows for the detection of sound-induced vibrations in a uniform manner along the lateral axis of the cochlea. An alignment of sensory cells is also present in the vestibular organs,

where directional stimuli uniquely allows hair cells in specific regions of the semicircular canals, utricle, and saccule to be activated in response to a specific stimulus such as head rotation, gravity, and linear acceleration[49]. The tissue wide orientation of hair cells with respect to supporting cells manifested through the uniform alignment of the stereocilia is a clear case of planar cell polarity (PCP)[50–52].

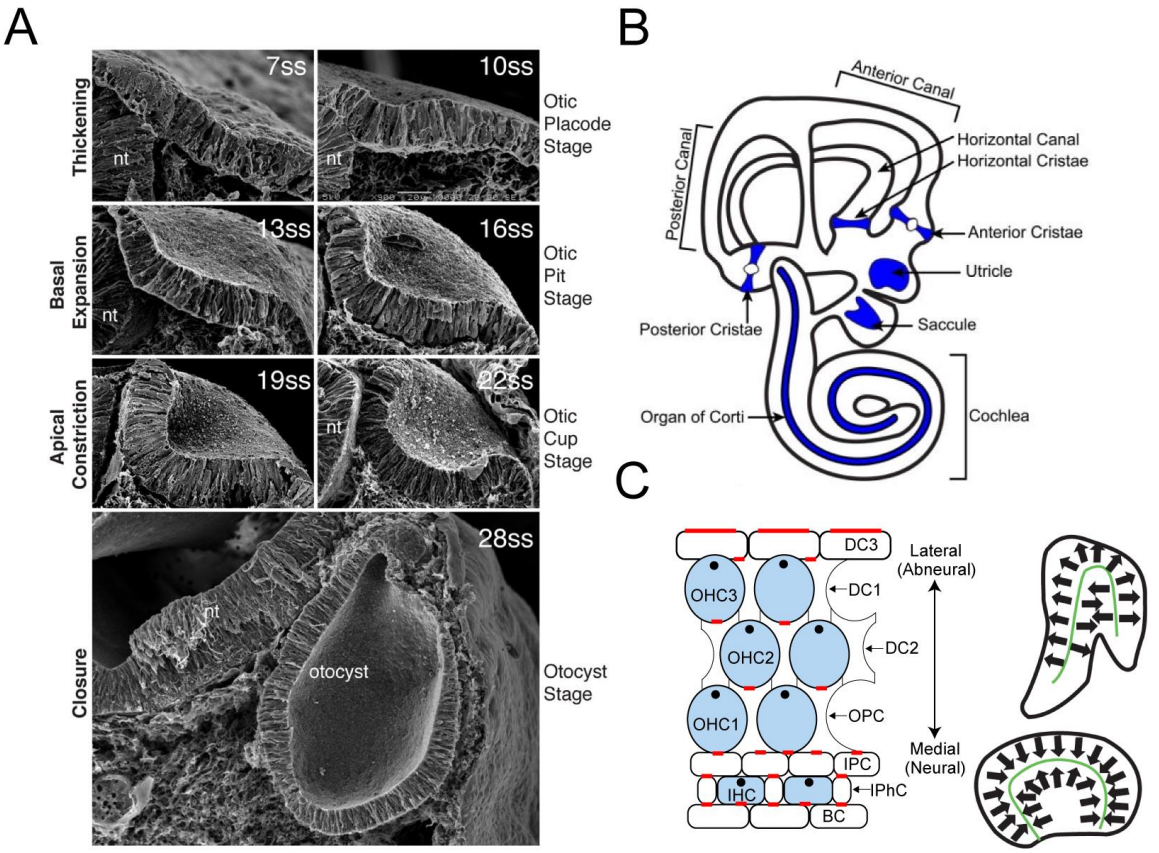


Figure 1.8. Inner-ear development and hair-cell alignment inside the mature inner ear

(A) Scanning electron micrographs of the chick otic placode at different stages of development. The cells on the ectoderm undergo a series of events (thickening, basal expansion, apical constriction, and invagination followed by closure) to form the otocyst, the precursor organ of the cochlea (Adapted from Sai *et al*. 2015[46]) *(B)* The mature inner ear is divided into the auditory sensitive cochlea perceiving sound and the vestibular organs (semicircular canals, utricule, and saccule) perceiving head movements, gravity, and linear acceleration. Hair cells lie along the blue shaded regions throughout both auditory and vestibular systems (Adapted from Duncan *et al*. 2017[49]). *(C)* Schematic of hair cell orientation inside the organ of Corti, saccule, and utricle. In the organ of Corti, the hair cells are oriented resulting in the localization of kinocilia (black circles) to the cell apex in the abneural or lateral edge. Non-sensory supporting cells that guide orientation are highlighted. In the saccule and utricle, the hair cells orient in opposite directions along the line of polarity reversal (Adapted from Wu *et al*. 2012[47], Rida *et al*. 2009[48]).

Planar cell polarity and its role in hair cell organization

In the mammalian inner ear, among the sensory hair cells, the initial orientation of the stereociliary bundle is directed by a unique cilium known as the kinocilium, made up of microtubules. This specialized primary cilium, which later disappears in mature auditory hair cells, defines PCP and serves as the guidepost for hair-cell bundle morphogenesis[47]. Initially, the kinocilium, which is present in the center of the apical surface, migrates towards the lateral side of the apical surface, influencing the surrounding microvilli to trace a similar path (Fig. 1.9 *A*). Microvilli, which are numerous and of the same height initially, adopt the gradual, gradient staircase like pattern to give rise to a either W-shaped or a flat-shaped stereociliary bundle seen in mature OHC and IHC, respectively. A similar mechanism is observed for vestibular hair cells (Fig. 1.9 *B*). The localization of the kinocilium that occurs in all types of hair cells is uniform and influenced by interactions with adjacent supporting cells facilitated by core PCP proteins[47]. In the event of a PCP disruption and in the absence of severe neural tube defects, the orientation of hair cells and hence the position of the resulting kinocilium and the orientation of the stereociliary bundle are affected (Fig. 1.9 *C*). Interestingly, while this does not affect the asymmetric localization of the stereocilia on the apical surface of hair cells or the kinocilium localization to the cell vertex, their lateral localization is affected resulting in misoriented bundles[47,48,51,52]. While this does not abolish hearing, the efficiency is reduced and abnormal head shaking behavior is observed[50].

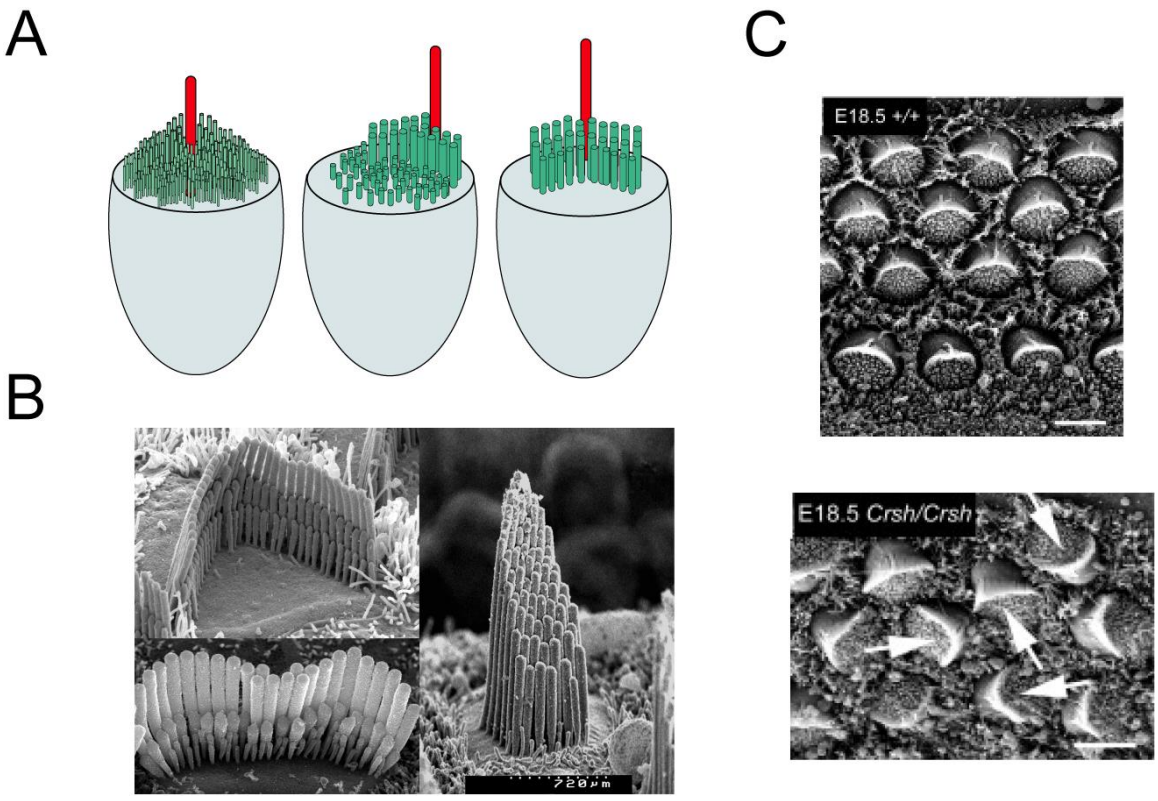


Figure 1.9. PCP in inner ear hair cell tissues

(A) Schematic of the localization of a kinocilium in the developing hair cell. A kinocilium gradually begins to migrate towards the lateral side of the cell guiding the developing stereocilia to orient in a similar fashion (Adapted from Wu *et al*. 2012[47]). *(B)* Electron micrographs showing bundles of different hair cell types, OHC (W-shaped bundle), IHC (flat bundle) and vestibular hair cell (Credit: Hair cell of inner ear, Dr David Furness). *(C)* Overall orientation of cochlear hair cells indicated by the position of the stereociliary tuft in OHC and IHC. Defects in PCP result in misorientation of stereociliary bundles (Adapted from Curtin *et al*. 2003[50]).

Components of PCP that contribute to hair cell alignment

PCP is mediated by a highly conserved pathway manifested throughout evolution in multicellular organisms that was first discovered in drosophila during studies on the orientation of hair wing epithelia. There are six core PCP proteins that regulate orientation by asymmetric localization. Of the six, CELSR1, Frizzled (Fzd3/6) and Dishevelled (Dvl1/2/3) localize to the lateral (abneural) side while CELSR1, Vang like (Vangl1/2) and Prickle (Pk1/2) localize to the medial (neural) side in a cochlear hair cell[48,51] (Fig. 1.10 A, *B*). The asymmetric localization of the complexes inside the cell is initiated by yet unknown factors to be discovered. But the leading candidates are thought to be members of the well-conserved Wingless (Wg)/Wnt family of morphogens that similarly exist in gradients and bind to Fzd receptors[48]. Inside the cell, the asymmetric complex localization is enhanced by antagonistic inhibition of the cytoplasmic proteins, Dvl and Pk to their preferred binding partners, Fzd and Vangl on the lateral and medial sides respectively[53,54]. Once the asymmetry is established inside the cell, the tissue wide asymmetry is communicated by CELSR1. CELSR1 is an atypical cadherin with nine EC repeats[50,55,56]. Unlike CDH23 and PCDH15, CELSR1 is thought to mediate homophilic adhesion between cells[56–58], but selectively recruiting opposing complexes to the adjacent cell border. This mechanism of differential recruitment by CELSR1 complexes arises from the preference of CELSR1 molecules to CELSR1: Fzd3/6 complexes in the opposing cells[59,60]. CELSR1 complexes may form supramolecular aggregates across cell borders that are inherently stable and display a higher degree of immobility compared to classical E-cadherins in the adherens junction[61]. Several mutations in the EC repeats of CELSR1 have been identified that

disrupt this junctional stability. The mutations, which exist close to a MAD in CELSR1, have been discovered to affect this function possibly by interfering with the *cis* interactions that exists between CELSR1 complexes on the same side of the cell[50,58]. But there is a lack of structural and biochemical information that could decipher *cis* and *trans* interactions at the molecular level. To completely understand the role of core PCP components that lead to hair-cell orientations inside the inner ear, it is necessary to understand the intricate behavior of CELSR1 that leads to the propagation of tissue-wide polarity across cells.

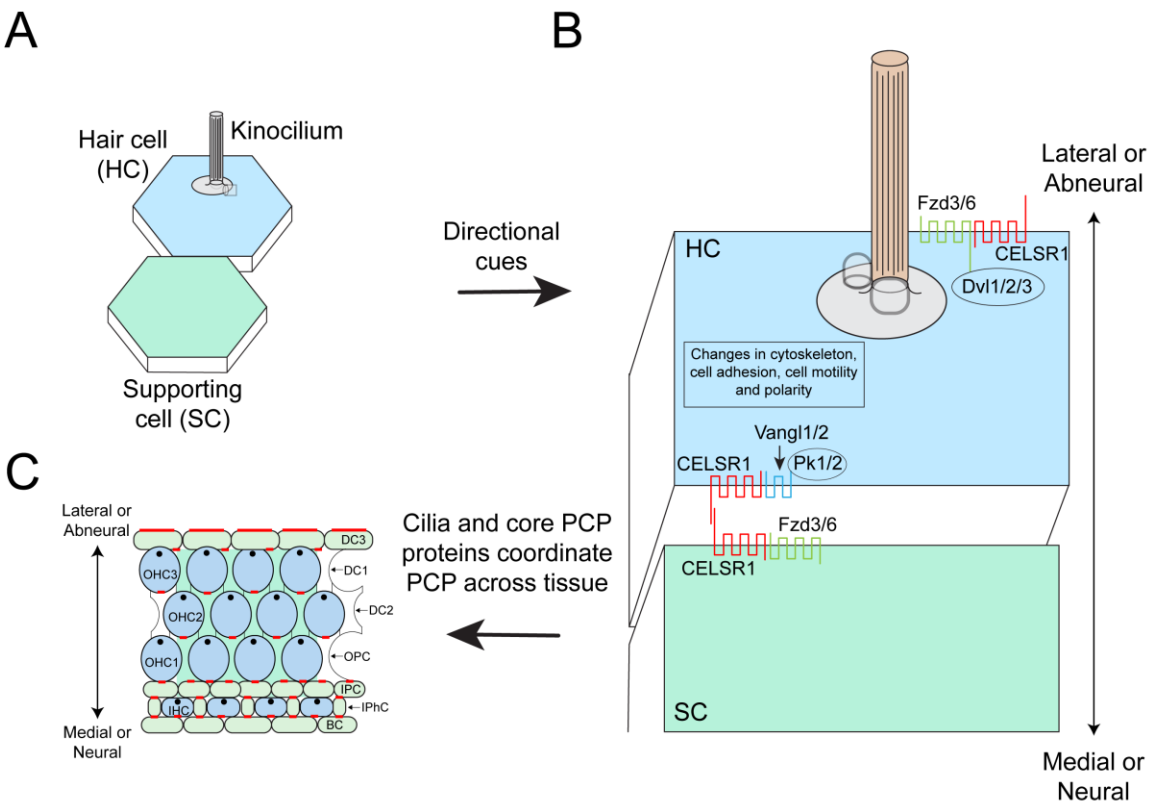


Figure 1.10. Directional cues in the inner ear and PCP

(A) Schematic of a hair cell with a kinocilium and a supporting cell (Adpated from Rida *et al*. 2009[48]). *(B)* In the presence of morphogenic gradients, the CELSR1:Fzd3/6 complex localizes to the lateral or the abneural edge of the hair cells and supporting cells recruiting CELSR1:Vangl1/2 in the opposing cell and vice versa. Here, the kinocilium serves as the guidepost by moving toward the lateral sides in each hair cell (Adpated from Rida *et al*. 2009[48]). *(C)* Overall orientation of hair cells in the sensory epithelium. Red bars indicate the localization of the core PCP complex, CELSR1:Fzd3/6 in the supporting and hair cells (Adpated from Rida *et al*. 2009[48]).

Dissertation outline

The structural and biochemical characterization of atypical cadherins CDH23, PCDH15, and CELSR1 has been the primary focus of my work. In the inner ear, each of these unique cadherins perform diverse functions: CDH23 and PCDH15 are directly involved in mechanotransduction while CELSR1 is involved in PCP driving organogenesis. Although these cadherins have long ectodomains, variations in sequence, number of EC repeats, and additional subunits make them unique and suited to their specific functions.

In chapter 2, the structure of an isoform (CD2-1) of PCDH15 covering repeats EC3-5, cryogenic electron microscopy (cryo-EM) images of PCDH15 and CDH23 ectodomains are presented[37]. Biochemical analysis describing the *cis* interactions of PCDH15 are also described. The structural and architectural information derived from this study were successfully used to build *in silico* full-length structural models of the tip link. Molecular dynamics simulations and steered molecular dynamics simulations of the tip link complex are providing coveted predictions of the stiffness and elastic response of the components that can be used to study the effect of disease mutations involved in hearing and balance.

In chapter 3, structures of *hs* CELSR1 fragments are presented along with extensive biochemical and computational analysis to describe their unique adhesive interactions. Two hybrid full-length ectodomain models based on our structures and AlphaFold2[62] predictions have enabled us to reproduce in-solution architectures observed in the paralog

CELSR2[63]. Based on the results of our study, we propose a mechanistic model for CELSR1 with weak adhesive interactions primarily promoting a signaling function rather than cell-cell adhesion function observed in other cadherins such as PCDH1, PCDH19, and PCDH24.

Published work on a project focused on a mechanosensitive protein (TMEM63B) also present in the inner ear, including my contributions, is in appendix A.

Chapter 2. Structural models of CDH23 and PCDH15 involved in hearing and deafness[1]

Abstract

The inner ear of mammals, housing mechanosensory hair cells with stereociliary tufts on their surface, are adapted to function in hearing and balance. Central to sensory perception is the tip link protein complex, made of cadherin-23 (CDH23) and protocadherin-15 (PCDH15), two atypical, large cadherins with 11 and 27 extracellular cadherin (EC) repeats. In auditory hair cells, the displacement of stereociliary bundles caused by sound exerts a tension on tip links leading to transduction channel opening necessary for generating an action potential. Here, we present cryogenic electron microscopy (cryo-EM) and high-speed atomic force microscopy (HS-AFM) images of ectodomains of CDH23 and PCDH15. Based on high-resolution structures of smaller fragments of CDH23 and PCDH15, we have built a high-resolution model of the full-length tetrameric ectodomain. These results reveal the structural determinants of tip link-mediated

[1] This chapter contains data obtained in collaboration with Deepanshu Choudhary and Yoshie Narui with certain sections on PCDH15 published in PNAS, doi:10.1073/pnas.1920444117 (2020). Elakkiya Tamilselvan and Deepanshu Choudhary obtained the *hs* PCDH15 EC3-5 ex12a+ fragment. Yoshie Narui prepared cryo-EM grids and obtained images of *mm* PCDH15 EC1-MAD12 and *mm* CDH23 EC1-MAD28 variant and performed data analysis. Jasanvir Sandhu, Carter Wheat, Elakkiya Tamilselvan, Travis Harrison-Rawn, and Marcos Sotomayor built the *in silico* model of *mm* CDH23 ectodomain. Jasanvir Sandhu, Joe Sudar, Pedro De-la-Torre, Marissa Boyer, and Travis Harrison-Rawn resolved structures of CDH23 fragments. Marcos Sotomayor cloned the *mm* CDH23 EC1-MAD28 variant. Daisy Alvarado helped with midi-preps of the *mm* CDH23 EC1-MAD28 variant. Qurat Ashraf provided the AlphaFold2 predictions of MAD28 fragment. Felix Rico and his team at Aix-Marseille obtained high-speed atomic force microscopy images of the *mm* CDH23 EC1-MAD28 variant.

inner-ear sensory perception and elucidate the structural and adhesive properties of CDH23 and PCDH15 relevant in disease.

Introduction

In the vertebrate inner ear, hair cells host the primary mechanotransduction apparatus that converts sound stimuli to electrical signals detected by the brain. On top of each hair cell lie actin-filled stereocilia that are arranged in order of descending heights with taller stereocilia connected to the shorter ones by a thin proteinaceous filament known as the tip link[25,64] (Fig 2.1 *A*). The tip link is made of two atypical cadherins of extraordinary length, cadherin-23 (CDH23) and protocadherin-15 (PCDH15), that form parts of its upper and lower ends[19,21–23,25]. CDH23 has 27 extracellular cadherin (EC) repeats and PCDH15 has 11 EC repeats, both interacting with each other at the N-terminal tips through EC1-2 repeats[36,37] (Fig 2.1). Both proteins possess ferredoxin-like domains at their C-terminal ends that are predicted to be structurally weaker and more flexible than the rigid EC repeats[37,40]. In PCDH15, this membrane adjacent domain (MAD) induces *cis* dimerization[40,41]. The role of MAD in CDH23 is unclear. Previous scanning-electron micrographs (SEM) and freeze-fracture transmission electron micrographs (TEM) of hair cell stereocilia, however, reveal the presence of *cis* dimers in both CDH23 and PCDH15[19,61]. Hence, at the level of the tip link, CDH23 and PCDH15, form a dimer of heterodimers, reinforced by an X-dimer interaction between EC2-3 of PCDH15[37,38]. Although recent in-solution studies of CDH23 revealed a monomeric fragment for EC1-27[42,43], the absence of MAD28 may inhibit potential *cis* dimer formation. PCDH15, in the

lower end of the tip link, connects to membrane proteins forming the MET complex that conduct potassium ions in response to mechanical stimuli. The MET complex is made of TMC1/2, TMIE, and LHFPL5[24,65,66].

Under resting conditions, the tip link experiences a tension of 10 pN to 40 pN[44]. Displacement of hair bundles occur in response to mechanical stimuli and lead to a tug exerting more tension on tip links. Force pulls on the heterotetrameric CDH23 and PCDH15 leading to TMC1/2 opening. Thus, the tip link forms an essential component of the transduction apparatus that couples displacement to channel opening[67]. Biophysical measurements suggests that there is a soft element, called the gating spring, that opens the channel[45,68]. The tip link might be this soft element or might be in series with it. Previous studies suggested that the tip link by itself may be too stiff for being the sole component[69]. With an estimated stiffness of around 0.5 to 4 mN/m for the gating spring[44], CDH23-like EC repeats were shown to be too stiff[70]. However, there are components of PCDH15 that are predicted to be flexible such as its EC3-4, EC5-6, and EC8-9 linkers[37,71]. These linker regions hold only two, one, or zero Ca^{2+} ions rather than three as observed in canonical linkers. This property, in turn, would make the regions flexible under force. In fact, steered molecular dynamics (SMD) simulations carried out with CDH23 at different Ca^{2+} occupancies in the linker regions of *Mus musculus* (*mm*) CDH23 EC1-2 and CDH23 EC1-2 D101G mutant predicted softer elements[69]. SMD simulations carried out with the *mm* PCDH15 EC1-MAD12 + CDH23 EC1-3 heterotetrameter at 0.02 nm/ns predict a stiffness of $\sim$ 24 mN/m[37]. In these simulations, MAD unfolding was observed without unbinding

of the tetramer, which emphasizes its mechanical weakness compared to EC repeats[37,40]. But the models lacked the CDH23 EC4-MAD28 region. Lack of dimerization points and flexibility imposed by MAD and certain regions of CDH23 ectodomain complicates the estimation even though CDH23 linker regions are mostly canonical.

In single-molecule force spectroscopy measurements[72], tip link fragments were observed to unbind at lower forces at 50 μM Ca^{2+} compared to 2 mM Ca^{2+}. The concentration of Ca^{2+} in the endolymph that surrounds tip links varies along the length of the cochlea between ~ 20 μM (base) to ~ 40 μM (apex)[73,74]. Dissociation constants of Ca^{2+} in the canonical linker region of CDH23 EC1-2 were found to vary from ~ 2 μM at site 3 up to ~ 80 μM at site 1[69]. At any point over the length of the cochlea, we can assume that only the Ca^{2+} at site 3 would be reliably occupied with the other sites having a lower likelihood of Ca^{2+} occupancy. Understanding tip-link mechanics under such conditions would require the availability of structural models of complete ectodomains of both CDH23 and PCDH15. In a quest to obtain low-resolution envelopes of ectodomains of CDH23 and PCDH15, we have attempted to use cryogenic electron microscopy (cryo-EM) and high-speed atomic force microscopy (HS-AFM) to understand the structural and adhesive properties of the tip link. A structural model of the complete ectodomain of the tetramer based on high-resolution structures of CDH23 was also built to study the mechanical properties of the tip link filament *in silico*. Here, I describe my contributions to obtain these models in collaboration with Deepanshu Choudhary (structure of *Homo sapiens* (*hs*) PCDH15 EC3-5 ex12a+), Yoshie Narui (cryo-EM imaging of *mm* PCDH15

and *mm* CDH23 ectodomains and 2D classification), Felix Rico (HS-AFM imaging of *mm* CDH23 ectodomain), Jasanvir Sandhu, Carter Wheat and Travis Harrison-Rawn (*in silico* construction of *mm* CDH23 ectodomain model). Structures of *mm* CDH23 fragments solved by Jasanvir Sandu, Joseph C. Sudar, Pedro De-la-Torre, Marissa Boyer, and Travis Harrison-Rawn were used to build the *mm* CDH23 ectodomain. AlphaFold2 (AF2) prediction of MAD28 was done by Qurat Ashraf.

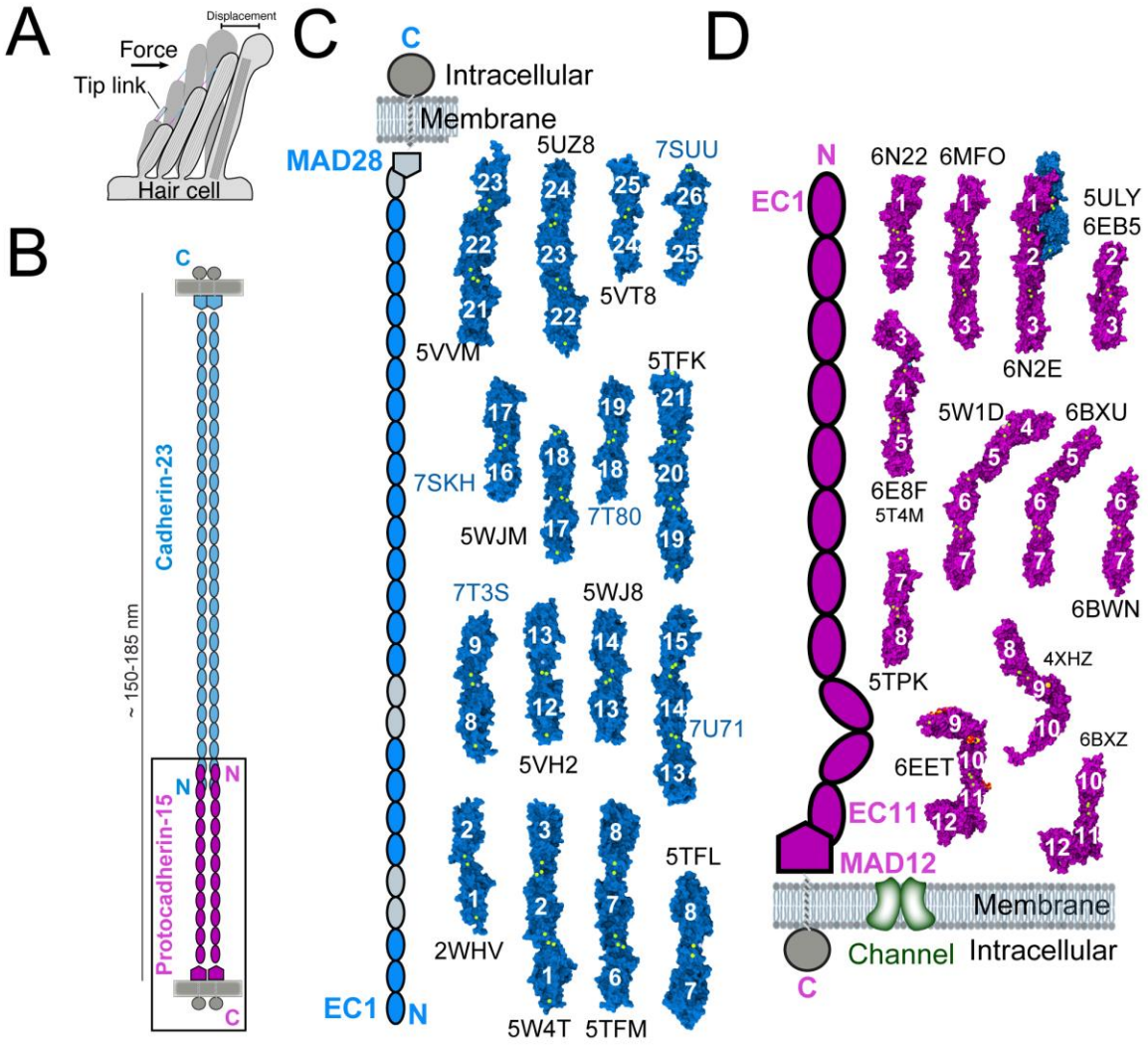


Figure 2.1. CDH23 and PCDH15 form the tip link

(*A*) A hair cell in the inner ear with stereocilia arranged in a staircase like pattern. Tip link connects two adjacent stereocilia and experiences tension when the stereocilia undergoes displacement in response to a mechanical stimulus. (*B*) Tip link made of *cis* dimers of CDH23 and PCDH15 each containing 27 and 11 EC repeats with a MAD at their C-terminal ends. The *cis* dimers form an heterotetrametric complex. (*C*) EC fragments of CDH23 for which X-ray crystal structures are available. Grayed out EC repeats in schematics denote those with unavailable structures. PDBs of unpublished structures, EC8-9 (Jasanvir Sandhu, Pedro De-la-Torre), EC13-15 (Pedro De-la-Torre, Jasanvir Sandhu, Florencia Velez-Cortes), EC16-17 (Joseph C. Sudar, Jasanvir Sandhu), EC18-19 (Travis Harrison-Rawn), and EC25-26 (Marissa Boyer, Jasanvir Sandhu) are indicated in blue. (*D*) Crystal structures of PCDH15 ectodomain fragments[37]. Structures are shown with their PDB accession codes. Ca^{2+} ions are shown as green spheres.

Results

Structure of *hs* PCDH15 EC3-5 ex12a+ found in PCDH15 CD2 isoforms

A previous structure of *Homo sapiens* (*hs*) PCDH15 EC3-5 ex12a− (PDB ID: 5T4M) revealed two bound Ca^{2+} ions at the EC3-4 linker region and a canonical Ca^{2+}-binding site at the EC4-5 linker region[75]. Monomers in the asymmetric unit displayed different orientations of EC3 with respect to EC4. However, this structure lacked exon 12a, which encodes for a seven-residue insertion p.V(414+1) PPSGVP (414+7) near the EC3-4 linker region. Given that structural variations due to isoform-specific (in-frame) insertions or deletions might be relevant for function, we sought to obtain a structure with exon 12a, present in some of the CD2 isoforms thought to be essential for inner-ear mechanosensation[21,76,77]. We solved the *hs* PCDH15 EC3-5 ex12a+ structure (PDB ID: 6E8F) (Fig 2.2 *A*), which has three molecules in the asymmetric unit having distinct EC3-4 interrepeat conformations (Fig 2.2 *C*) and with two of the linkers with only one bound Ca^{2+} ion at site 3[37]. The insertion enlarges the EC4 BC loop, which projects away from the repeat without altering its folding. This insertion may affect Ca^{2+}-binding affinity at the nearby EC3-4 linker. Overall, our *hs* PCDH15 EC3-5 ex12a+ structure is consistent with enhanced flexibility at the EC3-4 linker[37].

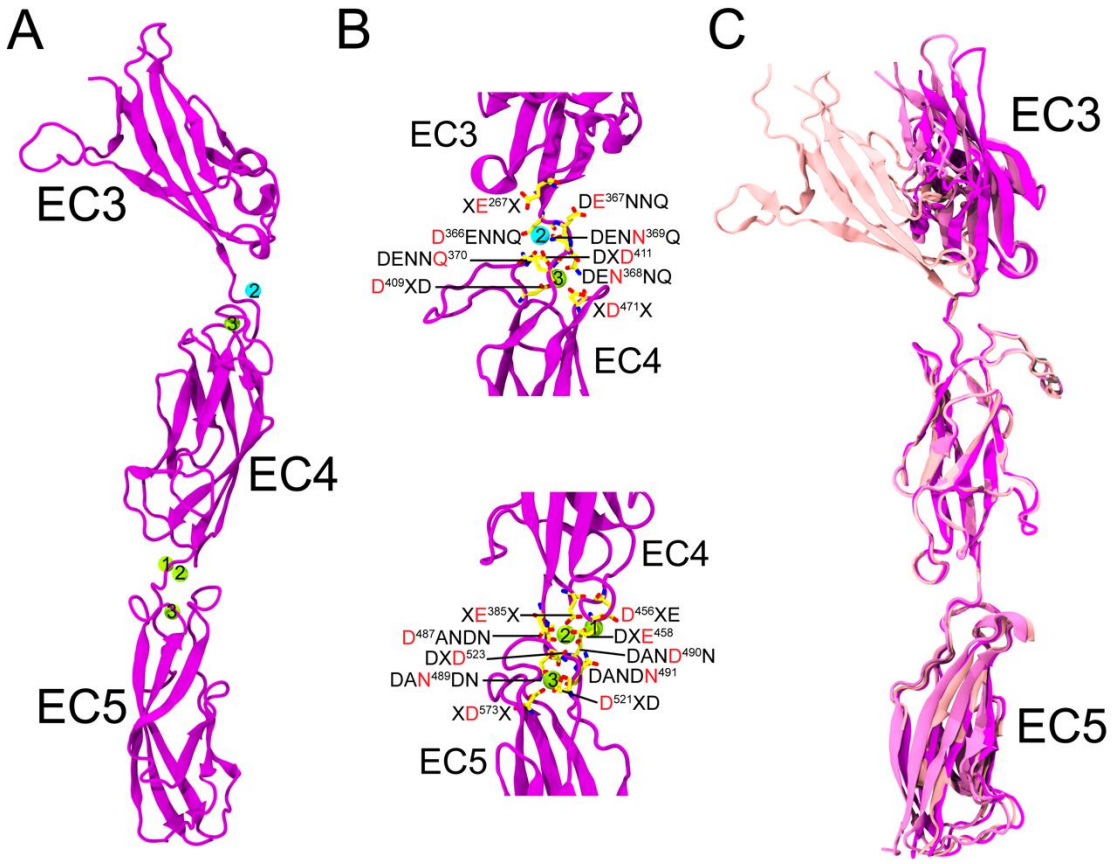


Figure 2.2. Structure of *hs* PCDH15 EC3-5 ex12a+

(*A*) Structure of *hs* PCDH15 EC3-5 ex12a+ revealed a non-canonical linker in the EC3-4 region with a Na^+ at site 2 and no ion coordinated at site 1[37]. (*B*) The linker regions in the *hs* PCDH15 EC3-5 ex12a+ show residues coordinating Ca^{2+} and Na^+ at EC3-4 (*top*) and EC4-5 (*bottom*). (*C*) Three chains in the crystal structure of *hs* PCDH15 EC3-5 ex12a+ aligned at the EC4 repeat show orientational variabilities of the EC3 repeat. Protein is shown in cartoon representation, Ca^{2+} and Na^+ ions are shown as green and cyan spheres. Residues involved in Ca^{2+} coordination are labeled and shown in stick representation.

Orientational analysis of EC4 with respect to EC3 in *hs* PCDH15 EC3-5 ex12a- and ex12a+ structures

In the asymmetric units of both *hs* PCDH15 EC3-5 ex12a- and ex12a+ crystal structures, two and three chains were observed with wide orientational conformations of EC4 with respect to EC3[37,71,75]. To quantitatively analyze the difference in orientation of EC4 with respect to EC3 in the different chains, we calculated the angle (θ) between EC3, which was aligned to EC1 of CDH23 (PDB ID: 2WHV) parallel to the *z*-axis, and EC4. The orientation of CDH23 EC2 with respect to EC1 was used as a reference (φ = 0°) (Fig 2.3). Although EC2 of CDH23 had an orientation angle of θ < 30°, the angle between EC3 and EC4 in all chains of both isoforms of *hs* PCDH15 EC3-5 had values between θ = 30° and θ = 90° (Fig 2.3 *A*). The most bent conformation (chain C) with only one Ca^{2+} at site 3 in the EC3-4 linker region of PCDH15 EC3-5 ex12a+ and the relatively less bent conformation with Ca^{2+} at site 2 in the EC3-4 linker region of PCDH15 EC3-5 ex12a- show a striking difference on the effect of Ca^{2+} in maintaining structural rigidity (Fig 2.3 *B*). Interestingly, in the full-length ectodomain model of *hs* PCDH15 ex12a+, the orientation angle of EC4 with respect to EC3 lies between the extreme conformations of the ex12a- and the ex12a+ structures of the *hs* PCDH15 EC3-5 fragment (Fig 2.3 *A*).

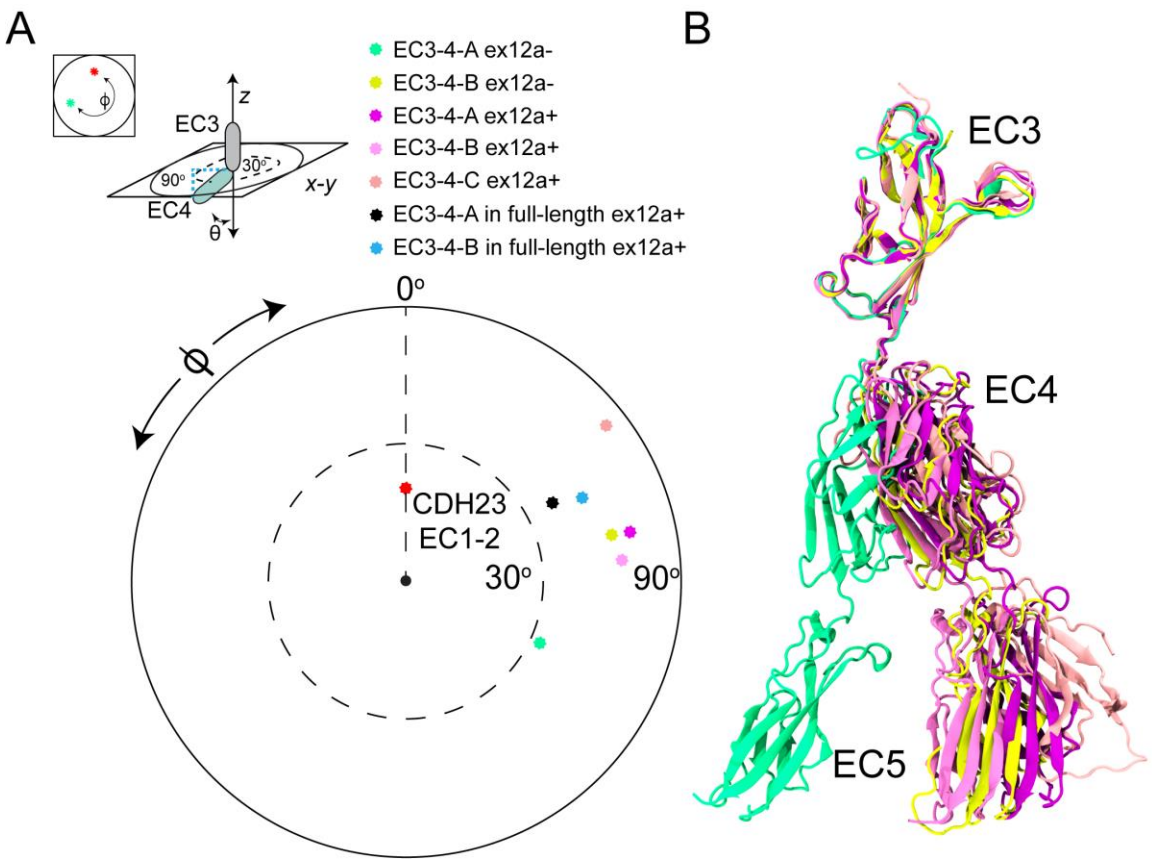


Figure 2.3. *hs* PCDH15 EC3-5 ex12a- and ex12a+ show varying degrees of bending at the EC3-4 linker

(*A*) Orientation of EC4 with respect to EC3 aligned along *z* to the CDH23 EC1 repeat (PDB ID: 2WHV). In all chains, EC4 lies at an angle > 30° with respect to EC3. (*B*) EC3 in all chains of *hs* PCDH15 EC3-5 ex12a- (PDB ID: 5T4M)[33] and ex12a+ (PDB ID: 6E8F)[37,75] aligned to each other to show orientational variabilities of EC4-5 with respect to EC3. The EC3-4 linker in ex12a- coordinates 2 Ca^{2+} ions at site 2 and 3 in both chains.

PCDH15 ectodomain exists as a Ca^{2+}-independent dimer in solution

An intriguing feature of the PCDH15 ectodomain is the presence of a ferredoxin-like domain succeeding the EC11 repeat known as MAD12 at the C-terminal end[39,40]. Unlike the EC repeats made of 7 β-strands, the secondary structure of MAD12 has four β-strands with two kinked α helices that form a βαββαβ fold. In solution, the shorter pig PCDH15 EC10-MAD12 fragment existed as an asymmetric dimer and SMD simulations of the PCDH15 ectodomain revealed that MAD12 is mechanically weak and exhibited unrolling and unfolding in response to force[37,40].

In addition to MAD induced dimerization, an X-dimer interaction at PCDH15 EC2-3 forms the second dimerization contact along the length of the ectodomain[37,38]. Here, the EC2-3 linker region plays a prominent role in stabilizing the interaction. While EC2-3 X-dimer interactions are expected to be Ca^{2+}-dependent, the MAD induced dimerization at the PCDH15 C-terminal end is expected to be Ca^{2+}-independent. Hence, characterizing the *cis* dimer promoted through both MAD12 and the X-dimer in the EC2-3 interface in the presence and absence of Ca^{2+} might reveal clues as to the overall architecture of the PCDH15 ectodomain under physiological concentrations of Ca^{2+} in the endolymph.

To produce the full-length ectodomain of *mm* PCDH15 EC1-MAD12 ex12a-, a mammalian expression system including Expi293 cells was used. The protein was purified successfully in size exclusion chromatography (SEC) where it eluted in a single peak (Fig. 2.4 *A*). In the presence of 5 mM Ca^{2+}, multi-light angle scattering (MALS) coupled to SEC

revealed a single peak with molecular weight calculated at ~ 377 kDa for the peak, indicating the presence of a dimer (Fig. 2.4 *B*). On incubation with 5 mM EDTA, the protein eluted with a shift to the right, but the molecular weight remained similar at ~ 356 kDa (Fig. 2.4 *B*)[37]. MALS data can be used to estimate the molecular weight of the molecule. The PCDH15 ectodomain in the presence of 5 mM Ca^{2+} would be a complete *cis* dimer bolstered by interactions facilitated by the EC2-3 X-dimer and the EC10-MAD12 fragment. However, in the presence of 5 mM EDTA, the EC2-3 mediated *cis* interactions would have been abolished, but MAD-induced dimerization would remain (Fig. 2.4 *C)*. The discrepancy in the molecular weight in MALS between Ca^{2+} and EDTA samples would have resulted from the disruption of the linear architecture of the EC repeats assuming a different shape from the rod configuration presumably observed for PCDH15 in the presence of Ca^{2+}. *Cis* interactions still maintained by MAD12 in the presence of EDTA show dimerization capacity that is independent of Ca^{2+} unlike the EC2-3 X dimer, although Ca^{2+}-dependent dimerization of EC2-3 has not been tested.

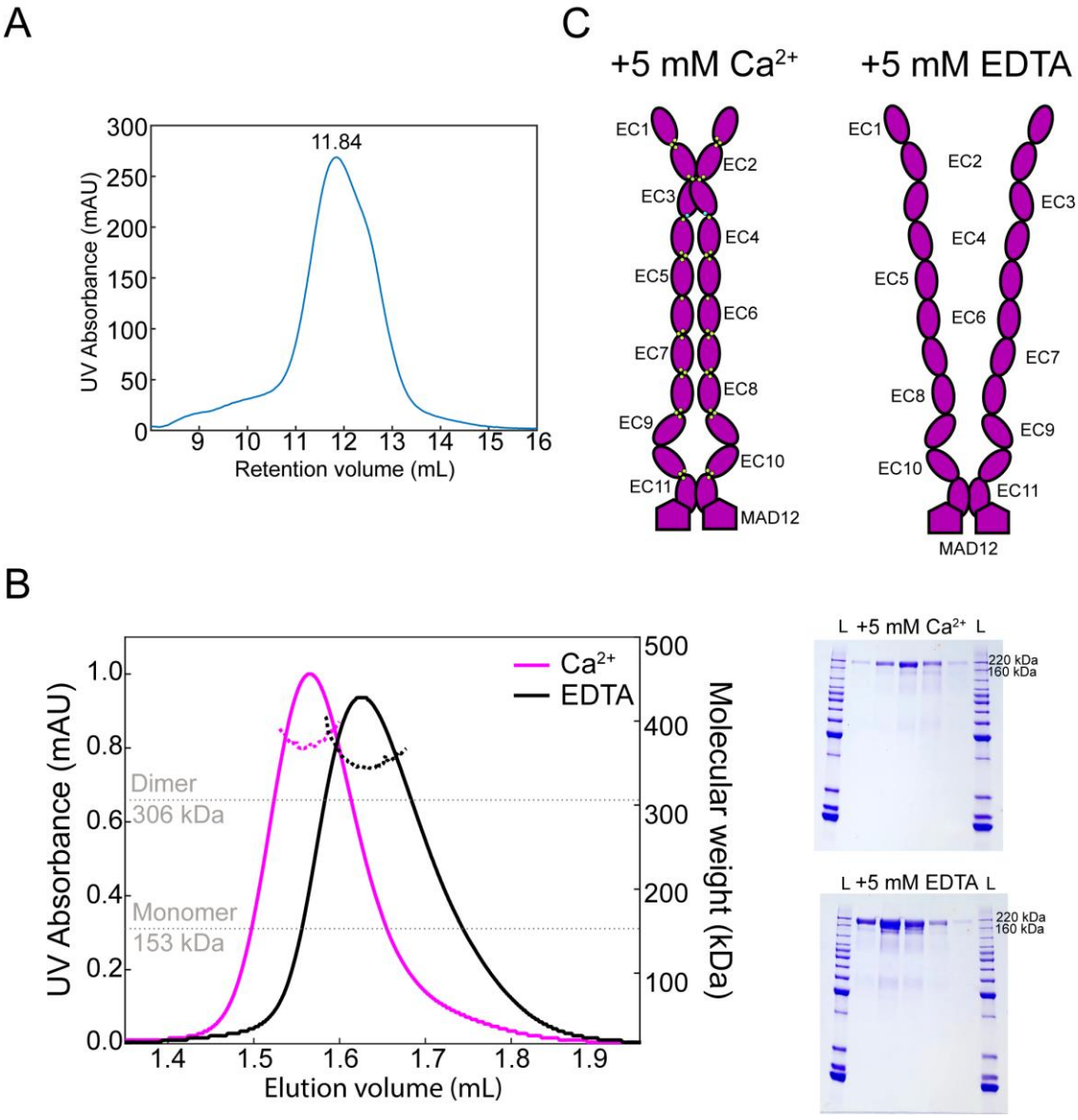


Figure 2.4. In-solution oligomerization state of *mm* PCDH15 EC1-MAD12 ex12a- in the presence and absence of Ca^{2+}

(*A*) SEC profile of *mm* PCDH15 EC1-MAD12 ex12a- produced from mammalian cells shows a single homogeneous peak. (*B*) SEC-MALS of *mm* PCDH15 EC1-MAD12 ex12a- show dimer peaks in the presence of 5 mM Ca^{2+} and 5 mM EDTA. SDS PAGE gels show the presence of protein. (*C*) Schematic of ectodomain of *mm* PCDH15 EC1-MAD12 ex12a- in the presence of 5 mM Ca^{2+} and 5 mM EDTA

Cryo-EM of PCDH15 ectodomain reveals *cis* dimer conformations

To visualize the architecture of the entire PCDH15 EC1-MAD12 ex12a-ectodomain, we used cryo-EM. PCDH15 ectodomains were produced in mammalian cells and were distributed evenly throughout grids for cryo-EM. All filaments of PCDH15 were observed in a dimeric configuration. A 2D classification of the filaments revealed several conformations but most of the observed 2D classes had two prominent dimerization points along the length of the ectodomain, likely at EC2-3 and MAD12 (Fig. 2.5). The 90º kink between EC9 and EC10 was not seen and instead a bulge was observed around the region in some of the 2D classes that showed a slightly bent conformation. A cryo-EM map (~ 20 Å resolution) obtained from the 2D classification was used to fit in the ectodomain model constructed previously[37] (Fig. 2.5). Even though the resolution of the ectodomain was low owing to the low number of particles in the twelve 2D classes obtained from cryo-EM images, some of the conformations observed resemble the high-resolution structural model constructed previously using smaller fragments.

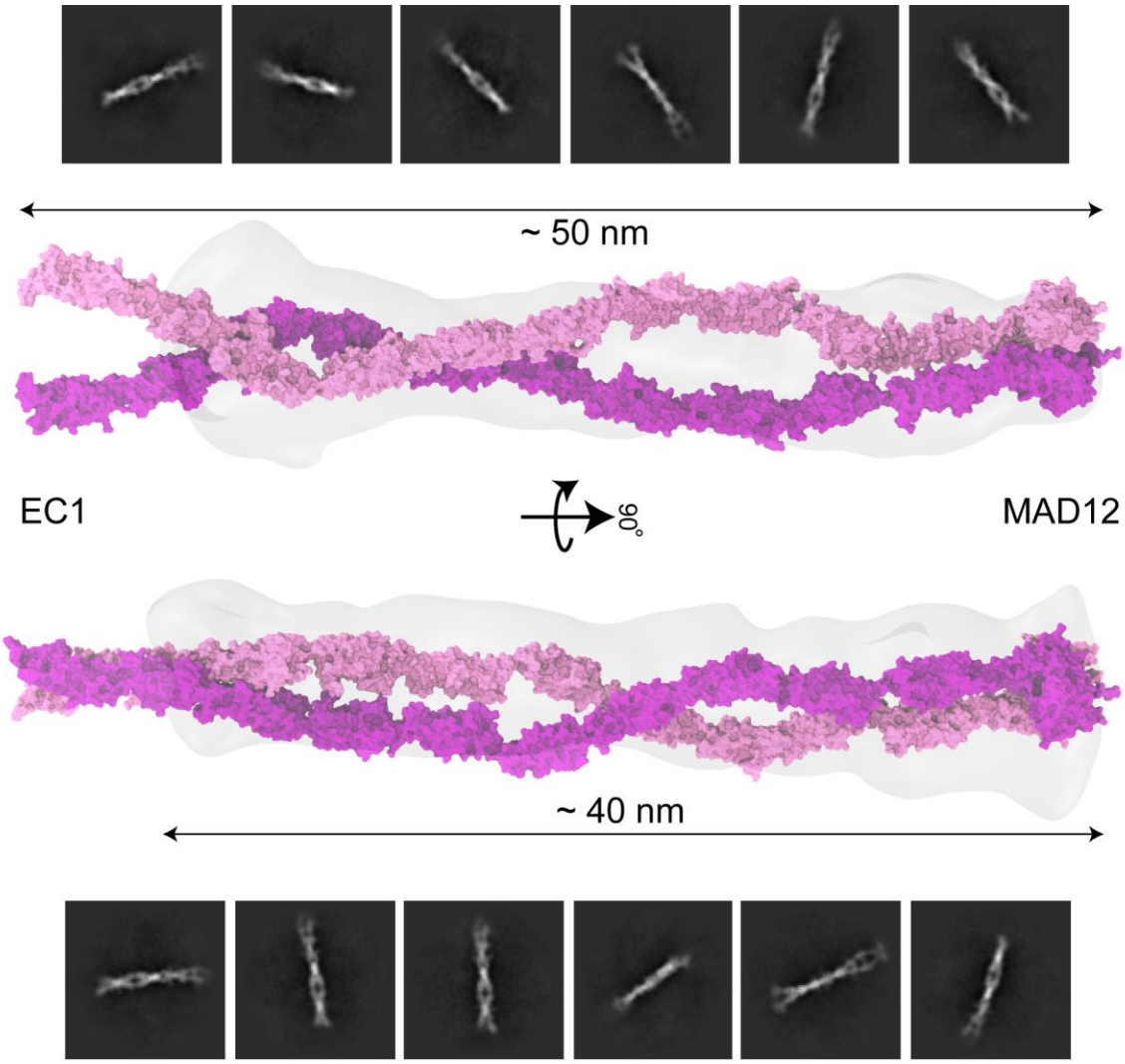


Figure 2.5. Cryo-EM images of the *mm* PCDH15 EC1-MAD12 ex12a- ectodomain

2D templates (*upper* and *lower* panels) of *mm* PCDH15 EC1-MAD12 ex12a- used for *ab initio* reconstruction of the 3D envelope (*Middle* panel). The *mm* PCDH15 EC1-MAD12 ex12a+ ectodomain model was fit into the 3D envelope. Due to unbound EC1-2, less contrast was seen for the region in all 2D classes and hence a missing density was observed in the 3D envelope. Here, the *mm* PCDH15 EC1-MAD12 ex12a- protein was produced by E.T, cryo-EM grids, data acquisition and image analysis were done by Yoshie Narui.

Cryo-EM and HS-AFM images of CDH23 ectodomain reveal multiple conformations

Similar to PCDH15, a mammalian expression system including Expi293F cells was used to express the complete ectodomain of a CDH23 variant (*mm* CDH23 EC1-MAD28, p.Q1 to p.D3044, 356NL357-, V976A, and S2064P). Unlike PCDH15, which was purified in 5 mM $CaCl_2$, CDH23 was purified using Ni-NTA beads in the absence of Ca^{2+} and incubated with 5 mM EDTA before SEC, after which the protein was purified in the presence of 5 mM $CaCl_2$. This gave us a better yield of presumably properly folded protein (Fig. 2.6 *A*). The protein was subsequently used for cryo-EM which gave us multiple monomer and dimer conformations (Fig. 2.6 *B, C*). Due to difficulties encountered in picking particles, the 2D classification was not possible. Similarly, HS-AFM imaging of CDH23 revealed several monomeric and a few dimeric conformations (Fig. 2.6 *D)*. Most of the dimeric conformations appeared to have dimerized towards either the N-terminal or the C-terminal end with the other end splayed open similar to some TEM images obtained from directly imaging stereocilia[25]. Overall, unlike for the PCDH15 ectodomain, we could not discern images from cryo-EM data that clearly showed specific dimerization points along the length of the CDH23 ectodomain.

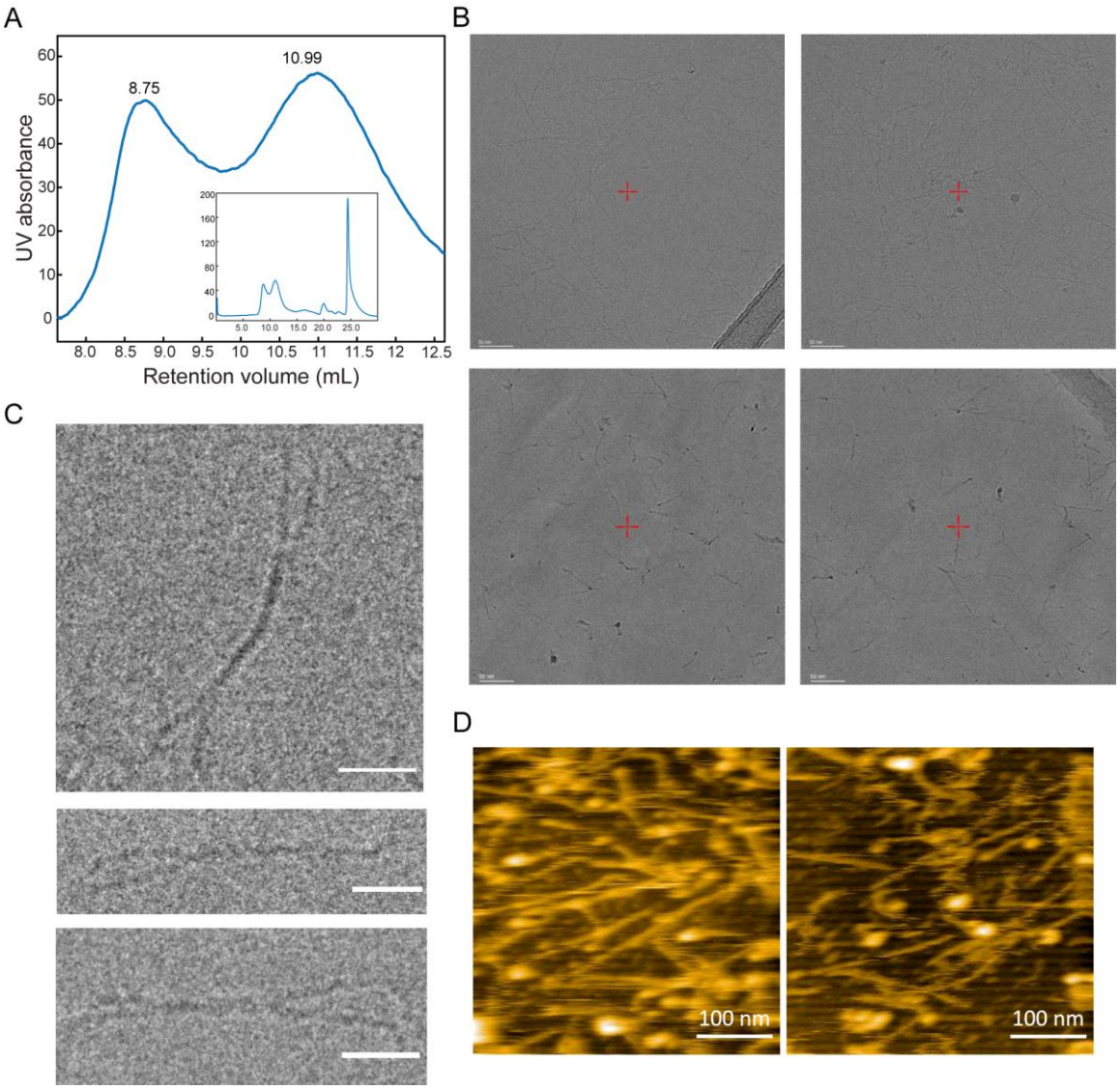


Figure 2.6. Purification, cryo-EM images, and HS-AFM imaging of a *mm* CDH23 EC1-MAD28 ectodomain variant

(*A*) SEC profile of *mm* CDH23 EC1-MAD28 variant produced from mammalian cells shows a protein peak eluted at 10.99 mL (zoom of the inset). (*B*) Cryo-EM images at 92,000x magnification. Both monomeric and dimeric filaments are observed. Scale bar: 50 nm. (*C*) Zoomed in images of isolated CDH23 fragments. Scale bar: 25 nm. Cryo-EM data obtained by Yoshie Narui (*D*) HS-AFM images show mostly single, monomeric filaments. Data obtained by Dr. Felix Rico and his team. Images shown with his permission.

Constructing a full-length tip-link model

To construct the full-length ectodomain of CDH23, we followed the protocol devised to construct the *in silico* model of the PCDH15 ectodomain[37]. High-resolution X-ray crystallography structures of smaller CDH23 fragments EC1-2 (PDB: 2WHV), EC3 (PDB: 5W4T), EC6 (PDB: 5TFM), EC7-8 (PDB: 5TFL), EC9 (PDB: 7T3S), EC12-13 (PDB: 5VH2), EC13-15 (PDB: 5WJ8), EC16-21 (PDB: 7SKH, 5WJM, 7T80, 5TFK), EC21-22 linker (PDB: 5VVM) and EC22-26 (PDB: 5UZ8[29], 5VT8[29], 7SUU) were stitched together through alignment of common regions. AF2[62] was used to generate models for EC4-5, EC10-11 and EC27-MAD28, which were used for fragments for which structures were not available (Fig. 2.7 *A*). For the construction of the tetramer, the EC1-3 fragment of a full-length CDH23 ectodomain that was equilibrated for 30 ns was aligned to one of the CDH23 EC1-3 fragments in the unstretched PCDH15 EC1-MAD12 ex12a+ + CDH23 EC1-3 model obtained previously[37]. To obtain the tetrameric model, we followed the same procedure but aligned the full-length CDH23 ectodomain to the other CDH23 EC1-3 fragment of the unstretched PCDH15 EC1-MAD12 ex12a+ + CDH23 EC1-3 model (Fig. 2.7 *C*). The *in silico* constructed CDH23 ectodomain monomeric and tetrameric complexes resemble some of the conformations observed in cryo-EM images, SEM and TEM[25,65].

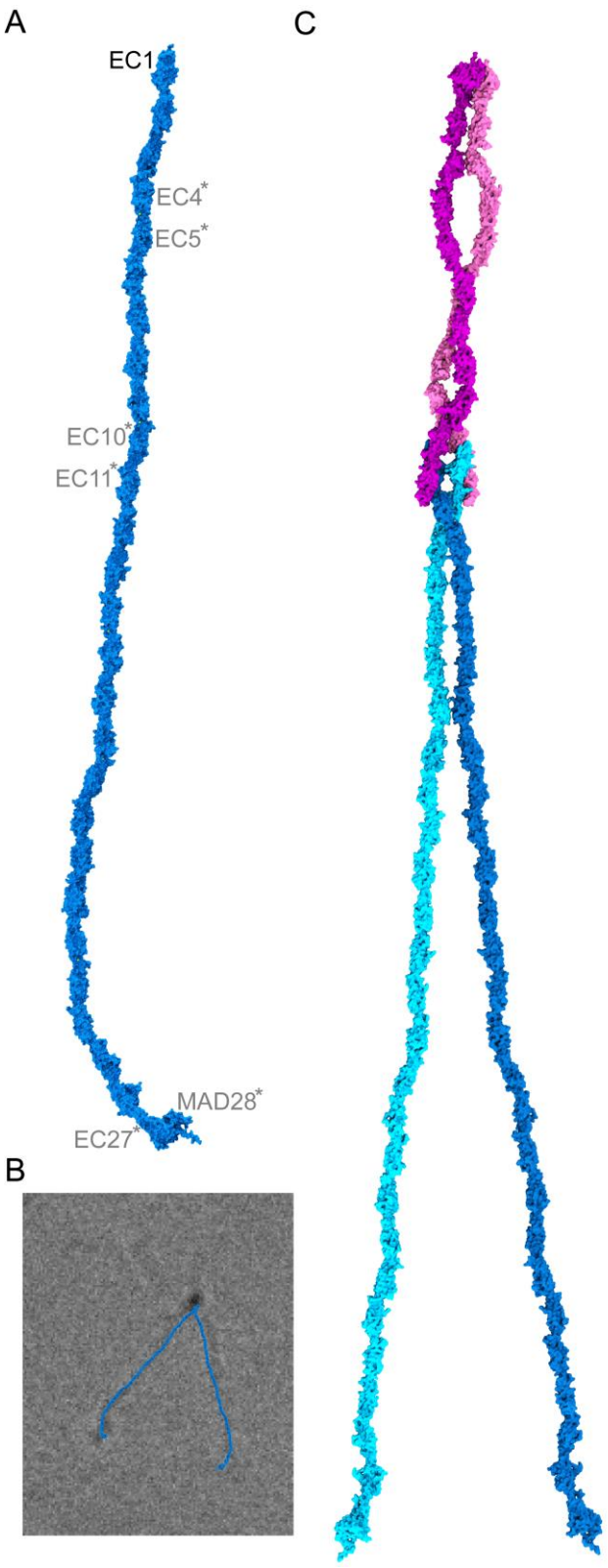

A
EC1
EC4*
EC5*
EC10*
EC11*
MAD28*
EC27*
B
C

Figure 2.7. CDH23 EC1-MAD28 ectodomain and PCDH15-CDH23 models

(*A*) *mm* CDH23 EC1-MAD28 monomer model constructed from X-ray crystal structures and AF2 models. This model was constructed in collaboration with Marcos Sotomayor, Jasanvir Sandhu, Carter Wheat and Travis Harrison-Rawn. Published[29] and unpublished structures, EC8-9 (Jasanvir Sandhu, Pedro De-la-Torre), EC13-15 (Pedro De-la-Torre, Jasanvir Sandhu, Florencia Velez-Cortes), EC16-17 (Joseph C. Sudar, Jasanvir Sandhu), EC18-19 (Travis Harrison-Rawn), and EC25-26 (Marissa Boyer, Jasanvir Sandhu) were used along with AF2[78] predictions of EC4, EC5, EC10, EC11 and MAD28 (Qurat Ashraf). AF2 predicted models are labeled in gray. (*B*) The CDH23 monomer models fit into the zoomed in micrograph. (*C*) Tetramer model of *mm* CDH23-PCDH15 built using the monomer models of CDH23 EC1-MAD28 and the previously constructed *mm* PCDH15 EC1-MAD12 ex12a+ ectodomain.

Conclusions

Our cryo-EM and HS-AFM images of CDH23 and PCDH15 reveal their dynamic and multiple conformations, especially for CDH23 which is ~ 120 nm in length. The diversity in conformational differences between CDH23 protomers and PCDH15 *cis* dimers reveal how they might contribute to elastic responses *in vivo*. Unlike PCDH15[38], *mm* CDH23 EC1-MAD28 variant seems to lack stable *cis* dimerization points along its length in agreement with in-solution studies of CDH23 EC1-27 fragment[42,43]. The Ca^{2+} independent dimerization observed for PCDH15 through MAD12[40,41], not observed for the protomers in CDH23 through MAD28, reveal important functional implications for the tip link during transduction. It is tempting to hypothesize that compared to PCDH15, the more variable nature of CDH23 might play an important role in modulating the stiffness of the tip link.

An important property of the tip link is its architecture and morphology under low Ca^{2+} concentrations. *In vivo* conditions for tip links vary significantly across organs and species. For example, Ca^{2+} concentration near tip links will greatly depend on the organ in which hair cells are located. The vestibular endolymph Ca^{2+} concentration is 90 to 150 μM[79], while the bulk cochlear endolymph Ca^{2+} concentration varies from 20 to 40 μM[73,74]. Intriguingly, Ca^{2+} concentrations near hair bundles in the cochlear subtectorial space might be significantly higher than previously thought (> 300 μM), both because of the action of stereocilia Ca^{2+} pumps functioning in a restricted space and because of the buffering effect of the tectorial membrane[80,81]. Similarly, resting tension and physiological forces will vary

within and across organs and in the cochlea due to tonotopy, whereas glycosylation and differential expression of various tip-link PCDH15 isoforms[82] may also be diverse. All these properties indicate the versatility of the tip link over the tonotopic length of the cochlea. Our structural models will be well-suited to study these scenarios by providing atomic-level details under the application of force in SMD simulations.

The complexity of CDH23 and PCDH15 ectodomains revealed by our structures and the various sets of environments in which it functions during mechanotransduction indicate that the tip link might be a versatile and multimodal protein that can be tuned to display distinct elastic responses[37]. It is important to understand the elastic responses in different Ca^{2+} concentrations of the entire tip link for which our high-resolution tetramer model can be altered *in silico* to perform SMD simulations. Determining the stiffness of the tip link under such conditions coupled to the mechanotransduction complex would offer insights into the open channel probability that might be otherwise difficult to study *in vivo*.

Methods

Expression and purification of bacterially expressed *hs* PCDH15 EC3-5 ex12a+ fragment

hs DNA sequence encoding for the protein fragment *hs* PCDH15 EC3-5 ex12a+ was subcloned into *NdeI* and *XhoI* sites of the pET21a vector. This bacterial construct had a starting methionine residue and lacked the native signal peptide. Residue numbering in the text and structure corresponds to processed protein without signal peptide (starting methionine residue is labeled M0). The DNA construct was sequence verified. Protein fragment was expressed in *Escherichia coli* BL21 CodonPlus (DE3)-RIPL (Agilent) cells, which were cultured in LB media, induced at OD_{600} ~0.6 with 200 μM IPTG and grown at 30° C for ~16 h. Cells were lysed by sonication in denaturing buffer (20 mM Tris HCl, pH 7.5, 6 M guanidine hydrochloride, 10 mM $CaCl_2$, and 20 mM imidazole). The cleared lysates were loaded onto Ni-Sepharose (GE Healthcare), eluted with denaturing buffer supplemented with 500 mM imidazole and refolded in 20 mM Tris HCl, pH 8.0, 10 mM $CaCl_2$, 400 mM L-Arg, and 1 mM GSSG by overnight dialysis using MWCO 2000 membranes (Spectra/Por). Prior to starting refolding reactions, elution solution for the PCDH15 fragment was diluted to ~ 0.5 mg/mL using the denaturing buffer, and then reduced by adding 2 mM DTT to the diluted sample. Refolded protein was concentrated using Amicon 15 centrifugal concentrators (10 kDa molecular weight cutoff) and further double-purified on Superdex S200 column (GE Healthcare) in 20 mM Tris HCl, pH 8.0, 150 mM KCl, 50 mM NaCl, and 5 mM $CaCl_2$ buffer. The protein sample was concentrated by ultrafiltration to ~ 5 mg/ml for crystallization[37].

Crystallization and structure determination

Crystals were grown by vapor diffusion at 4° C by mixing protein and reservoir solutions (0.6 μL + 0.6 μL 40% MPD, 0.2 M lithium chloride, 0.01 M ATP Disodium). Crystals were cryo-cooled in liquid N_2. X-ray diffraction data sets were collected as indicated in Table 2.1 and processed with HKL2000[83]. Structures were determined by molecular replacement using PHASER[84] and the *hs* PCDH15 EC3-5 ex12a- structure (PDB: 5T4M). Refinement of this structure used the amplitude-based "Twin Refinement" option in REFMAC5[85] after achieving an R_{free} value of ~27%. Model building was done with COOT[86] and restrained refinement was performed with REFMAC5[85]. Data collection and refinement statistics are provided in Table 2.1. The structure was further analyzed using Procheck[87], Whatcheck, and Checkmymetal[88] prior to deposition. In the *hs* PCDH15 EC3-5 ex12a+ structure, initial assignment of Ca^{2+} at site 2 of the EC3-4 linker region of chain A resulted in a positive value of the F_o-F_c density at the location of the ion. The protein solution buffer has 5 mM $CaCl_2$ and 50 mM NaCl while the crystallization buffer had 200 mM LiCl. Thus, we tried placing a Li^+ ion, which resulted in a large positive value of F_o-F_c density at the site of the ion. In contrast, Na^+ at the location was compatible with the $2F_o$-F_c and F_o-F_c electron density maps, its β factor was similar to that of surrounding atoms, and it is also the cation with the closest ionic radius to Ca^{2+}. Thus, a Na^+ ion was placed at this site in the final model for the structure. Protein chains were generally modeled using a cutoff of 1.5 rmsd for the contour level of the $2F_o$-F_c map in COOT[89]. In some loops, weak electron density was observed at a contour level of 1 rmsd, which allowed us to fit the residues in the density but resulted in higher β-factor values for those regions. In

our structure, we observed no density for the N- and C- terminal loops even at a contour level of 0.5 rmsd. Such regions of the protein were not built. Missing residues in our model include the BC loop of EC3 in all monomers and the EC5 β-strand A along with the connecting AB loop in chain C of the *hs* PCDH15 EC3-5 ex12a+ structure.

In silico analyses

Individual chains of the *hs* PCDH15 EC3-5 ex12a- and ex12a+ structures were aligned to each other using C_α atoms in the regions of EC4 (Fig. 2.2 *C*) and EC3 (Fig. 2.3 *B*) using VMD[90]. The VMD orient plugin was used to compute principal axes and to calculate the relative orientation of EC repeats with respect to each other within a pair by using *mm* CDH23 EC1-2 (2WHV)[69] as a reference. The EC3 repeats of *hs* PCDH15 EC3-5 ex12a- and ex12a+ structures were aligned to CDH23 EC1 in the *z* direction and the relative orientation of EC4 to EC3 was calculated using the projection of the longest principal axes on the *x-y* plane. The calculated azimuthal angles (ϕ) were plotted in reference to the orientation of CDH23 EC2 to EC1, which was defined as 0°. Molecular figures were made using VMD. Plots were prepared using MATLAB[91].

Expression and purification of mammalian expressed PCDH15 and CDH23 ectodomains

Mouse PCDH15 EC1-MAD12 ex12a- (p.Q1 to p.G1327) and a CDH23 EC1-MAD28 variant (p.Q1 to p.D3044, 356NL357-, V976A, and S2064P) were subcloned into a pHis-N1 vector (a modified version of the pEGFP-N1 vector from Clontech donated generously by Dr. James Jontes where the EGFP has been substituted for a hexahistidine

tag) using *XhoI* and *KpnI* restriction sites. The respective native signal sequences were included before the start of EC1 for PCDH15 and CDH23, and all constructs were sequence verified. All protein fragments were expressed by transient transfection of Expi293 cells using ExpiFectamine. After 4-5 days of expression, the conditioned media (CM) was collected and dialyzed overnight against 20 mM Tris HCl, pH 7.5, 150 mM KCl, 50 mM NaCl, and 10 mM $CaCl_2$ to remove EDTA for PCDH15. The CM for both CDH23 and PCDH15 were concentrated using Amicon 10 kDa or 30 kDa concentrators and incubated with Ni-Sepharose beads for 1 h. The beads were washed 3 times with 20 mM Tris HCl, pH 8.0, 200 mM NaCl (with 10 mM $CaCl_2$ for *mm* PCDH15 EC1-MAD12 ex12a-), and 20 mM imidazole, and the target protein was eluted with the same buffer containing 500 mM imidazole. Additionally, the *mm* CDH23 EC1-MAD28 variant was incubated with 5 mM EDTA for 15 min after elution. Both the *mm* PCDH15 EC1-MAD12 ex12a- and *mm* CDH23 EC1-MAD28 variant proteins were purified on a Superose 6 10/300 column in 20 mM Tris HCl, pH 8.0, 150 mM KCl, and 5 mM $CaCl_2$ and concentrated for SEC-MALS.

SEC-MALS

SEC-MALS experiments were done using an AKTAmicro system connected in series with a Wyatt miniDAWN TREOS system. Mammalian expressed *mm* PCDH15 EC1-MAD12 ex12a- (~1.3 mg/mL) was separated on a Superose 6 3.2/30 column in 20 mM Tris HCl, pH 8.0, 150 mM KCl, with either 5 mM $CaCl_2$ or 5 mM EDTA. Absorbance at 280 nm and light scattering were monitored. The scattering information was subsequently converted into molecular weight using a rod-like model and a theoretical

extinction coefficient of 0.678 $(mg/ml)^{-1}cm^{-1}$. The SEC-MALS curves were adjusted for connecting tubing length before plotting. Measurements listed in Table 2.2 were taken from distinct samples.

Cryo-EM of *mm* PCDH15 EC1-MAD12 ex12a- and *mm* CDH23 EC1-MAD28 variant

For cryo-EM studies, *mm* PCDH15 EC1-MAD12 ex12a- and the *mm* CDH23 EC1-MAD28 variant were suspended in a buffer containing 20 mM Tris HCl, pH 8.0, 150 mM KCl and 5 mM $CaCl_2$ before plunge freezing. A PELCO easiGlow (TedPella, Redding, CA, USA) was used to glow-discharge the Quantifoil grids for 60 s. A Vitrobot Mark IV (Thermo Scientific, Waltham, MA, USA) was used to blot the samples (3 µL) on to the grid for 4 s at 100% humidity followed by plunging into liquid ethane. The blotted grids were clipped and stored in liquid nitrogen until data collection. The grids were screened on a Glacios cryo-EM (Thermo Scientific, Waltham, MA, USA) equipped with a Falcon 3EC direct electron detector at a pixel size of 1.13 Å for homogeneous sample distribution. The optimal grids were subsequently used to collect data on a Titan Krios G3i (Thermo Fisher Scientific, Waltham, MA, USA) operated at 300 kV, equipped with a K3 direct electron detector, a Bioquantum energy filter, and a Cs image corrector. For PCDH15, a total of 586 movies were collected at a magnification of 81,000× corresponding to a pixel size of 0.4495 Å in super-resolution mode at a defocus range of –0.5 µm to –2.5 µm with a total electron dose of 60 $e^{-}/Å^2$. For CDH23, screening of the grids was performed corresponding to a pixel size of 0.4495 Å at a defocus range of –0.5 µm to –2.5 µm with a total electron dose of 60 $e^{-}/Å^2$. Data collection for both samples were performed using EPU software

(Thermo Fisher Scientific, Waltham, MA, USA). Data collection statistics are shown in Table 2.3.

Image processing and 3D reconstruction

All raw movies of *mm* PCDH15 EC1-MAD12 were motion corrected and the contrast transfer function (CTF) parameters were estimated using the Patch CTF estimation module in cryoSPARCv3.2[92]. Particle (blob) picking followed by an initial round of inspection and extraction from the micrographs was used to generate templates. The templates were subsequently used to refine and improve 2D classes, that were used to develop an *ab initio* reconstruction of the PCDH15 molecule. A homogeneous refinement was performed to sharpen the map obtained from the previous step. Data collection statistics, image processing, and the refinement summary of models are shown in Table 2.3.

HS-AFM of the *mm* CDH23 EC1-MAD28 variant

Experiments were carried out by Dr. Felix Rico and his team. Methods for these experiments will be described in an upcoming manuscript.

Table 2.1. Statistics for *hs* PCDH15 EC3-5 ex12a+ structure

Data collection	***hs* PCDH15 EC3-5 ex12a+**
Space group	$P222_1$
Unit cell parameters	
a, *b*, *c* (Å)	95.74, 95.91, 258.60
α, β, γ (°)	90.0, 90.0, 90.0
Molecules per asymmetric unit	3
Beam source	APS-24-ID-C
Date of data collection	29-MAR-2018
Wavelength (Å)	0.9792
Resolution limit (Å)	2.99
Unique reflections	48,387
Completeness (%)	99.9 (99.1)
Redundancy	10.2 (7.9)
$I / \sigma(I)$	9.3 (2.4)
R_{merge}	0.21 (0.81)
R_{meas}	0.22 (0.87)
R_{pim}	0.07 (0.30)
$CC_{1/2}$	0.94 (0.78)
CC^*	0.99 (0.94)
Refinement	
Resolution range (Å)	47.95 – 2.99 (3.07 – 2.99)
R_{work} (%)	17.4 (23.4)
R_{free} (%)	21.4 (27.9)
Residues (atoms)	1021 (7,966)
Water molecules	8
Rms deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.198
B-factor average	
Protein	89.12
Ligand/ion	85.04
Water	58.58
Ramachandran Plot Region (PROCHECK)	
Most favored (%)	88.1
Additionally allowed (%)	11.7
Generously allowed (%)	0.2
Disallowed (%)	0.0
PDB ID	**6E8F**

Table 2.2. Oligomerization state of PCDH15 fragments

System	**Predicted Mass Monomer / Dimer kDa**	**SEC-MALS kDa†**	**State**
*mm** PCDH15 EC1-MAD12 ex12a- (*n* = 2)	152.4 / 304.8	377.1 (23.2%)	Dimer
*mm** PCDH15 EC1-MAD12 ex12a-EDTA¶ (*n* = 2)	152.4 / 304.8	355.8 (16.3%)	Dimer

† Values in parenthesis indicate percentage deviation from predicted molecular mass.

* Mammalian expressed protein. The predicted mass represents that of the protein sequence without glycosylation. The calculated mass from SEC-MALS experiments is higher than the predicted mass due to the presence of sugar moieties on the protein.

¶ EDTA was present in the SEC column but not directly added to the sample.

Table 2.3. Data collection statistics for cryo-EM of *mm* PCDH15 EC1-MAD12 ex12a- and the *mm* CDH23 EC1-MAD28 variant

Map	**PCDH15**	**CDH23**
Data collection		
Microscope	FEI Titan Krios G3i	Thermo Glacios
Voltage (kV)	300	200
Detector	Gatan K3	Ceta 16M
Automation software	EPU	EPU
Energy filter slit width (eV)	20	20
sRecording mode	Super-resolution	Super-resolution
Magnification (nominal)	81,000	92,000
Movie micrograph pixel size (Å)	0.899	0.899
Total Dose rate (e^{-}/Å^2)	60	60
Defocus range (μm)	–0.5 to –2.5	–0.5 to –2.5
Spherical aberration (mm)	0.01	0.01
Movies	586	
Total extracted particles	12,280	
Total # of refined particles	6,550	

Chapter 3. CELSR1, a core planar cell polarity protein, features a weakly adhesive and flexible cadherin ectodomain[2]

Abstract

Planar cell polarity (PCP), essential to multicellular developmental processes, arises when cells polarize and align across tissues. Central to PCP is CELSR1, an atypical cadherin featuring a long ectodomain with nine extracellular cadherin (EC) repeats, a membrane adjacent domain (MAD10), and several characteristic adhesion GPCR domains. Cell-based aggregation assays have demonstrated CELSR1's homophilic adhesive nature, but mechanistic details are missing. Here, we investigate the possible adhesive properties and structures of CELSR1 EC repeats. Our bead aggregation assays do not support strong adhesion by EC repeats alone. Consistently, EC1-4 only dimerizes at high concentration in solution. Crystal structures of human CELSR1 EC1-4 and EC4-7 reveal typical folds and a non-canonical linker between EC5 and EC6. Simulations and experiments using EC4-7 indicate flexibility at EC5-6, and solution experiments show EC7-MAD10-mediated dimerization. Our results suggest weak homophilic adhesion by CELSR1 cadherin repeats and provide mechanistic insights into the structural determinants of CELSR1 function.

[2]This chapter is adapted from a manuscript accepted for publication in *Structure* (2023). Elakkiya Tamilselvan did cloning, protein expression and purification, and carried out all binding and bead aggregation assays, crystallization trials, simulations, and sequence alignments. Marcos Sotomayor trained E.T, supervised work, and assisted with data analysis. E.T. and M.S. solved crystal structures, prepared figures, and wrote the manuscript.

Introduction

The compartmentalization of organs and organ systems in multicellular organisms coincides with their ability to perform cohesive but diverse and specific functions needed for survival. At the heart of successful organogenesis lies the collective polarization and alignment of cells along an axis parallel to the tissue plane. This process, termed planar cell polarity (PCP), regulates oriented cell division, cell migration, the orientation of cytoskeletal elements, and positioning of cell extensions such as cilia and axons[93–95]. At the tissue level, PCP facilitates proper embryonic development during morphogenesis and organogenesis, defects of which can lead to improper neural tube closure and congenital heart disorders[96–99].

In response to global sensory cues, PCP is established by the asymmetric localization of two transmembrane complexes to opposite sides on the apical region of the cell: Cadherin epithelial growth factor (EGF) Laminin-G (LAG or LamG) seven pass G-type receptor (CELSR1) + Frizzled (Fzd3/6) and CELSR1 + Vang like (Vangl1/2)[60,100–103]. Initially, the complexes are randomly distributed throughout the apical region. However, Fzd3/6 and Vangl1/2 through interaction with their cytoplasmic partners, Dishevelled (Dsh) and Prickle (Pk), amplify asymmetry through antagonistic inhibition of oppositely oriented complexes[53,99,104]. This leads to the re-distribution of CELSR1 + Fzd3/6 and CELSR1 + Vangl1/2 resulting in the formation of a uniform asymmetrical lattice across cell borders.

CELSR1, localizing to both sides of the cell, is both an atypical cadherin belonging to the larger cadherin superfamily and an adhesion G-protein coupled receptor (adhesion GPCR) family member. Cell-based aggregation assays and *in vitro* experiments have shown that CELSR1 adheres cells together through intercellular homophilic interactions[56,57,105]. During events leading to PCP, CELSR1 controls the resulting planarity at the tissue level by complementary recruitment of either only CELSR1 + Fzd3/6 or only CELSR1 + Vangl1/2 complexes across the cell junction from neighboring cells[56,59,106–108].

Structurally, CELSR1 has nine extracellular cadherin (EC) repeats, a membrane adjacent domain (MAD10)[40], also referred to as SEA[109], eight EGF domains interspersed with two LamG domains, and an EGF-Lam domain followed by a hormone receptor binding (HormR) domain, a GPCR autoproteolysis-inducing (GAIN) domain, seven transmembrane domains, and a cytoplasmic domain[55,110,111]. Cadherin ectodomains are known to mediate cell-cell adhesion across intercellular junctions through their first few N-terminal EC repeats either by homophilic or heterophilic interactions[112–116]. The function of MADs is less clear, but MAD12 induces parallel dimerization in PCDH15[37,40]. In other adhesion GPCRs, the EGF domains are involved in adhesion, the HormR is a potential ligand binding domain, and the GAIN domain cleaves itself into two fragments to regulate its function[117,118]. In CELSR1, the auto-proteolytic action of the GAIN domain has been discovered to be deficient[119] but the minimal set of EC repeats that mediates homophilic

interactions to promote cell-cell adhesion, the function of MAD10 and EGF domains, and the binding partners of the HormR domain remain unknown[55].

In cadherins, each EC repeat has ~110 amino acids and adopts a common immunoglobulin-like fold with seven ß-strands that resemble a Greek-key motif[120]. The EC repeats are arranged in series with the linker region between two consecutive repeats canonically coordinating three Ca^{2+} ions. In non-canonical cases, the linker region coordinates less than three Ca^{2+} ions often leading to flexibility and a bending in the structure[33,121,122]. These non-canonical linker regions may regulate adhesion by facilitating the accommodation of large cadherin ectodomains in the small extracellular space between adjacent cells[123]. CELSR1, which localizes to the adherens junction during PCP events[124–127], has an extracellular region (> 50 nm) that is too long to be accommodated linearly within the junction (~15-30 nm between cells). A predicted non-canonical linker region between CELSR1 EC5 and EC6 could induce bending but how this protein forms homophilic complexes under such conditions to mediate cell-cell adhesion is unclear.

Here, we use bead aggregation assays to test the adhesiveness of the *Homo sapiens* (*hs*) CELSR1 cadherin ectodomain (EC1-EC9) with and without MAD10. We also present crystal structures of *hs* CELSR1 EC1-4 and EC4-7, and we use molecular dynamics (MD) simulations to study the flexibility of the *hs* CELSR1 EC5-6 non-canonical linker region *in silico*. In addition, we use small angle X-ray scattering (SAXS) to get a low-resolution envelope of EC4-7 to determine its in-solution architecture. Size exclusion

chromatography-multiangle light scattering (SEC-MALS) and analytical ultracentrifugation (AUC) suggest in-solution oligomerization states of fragments EC1-2, EC1-4, EC1-4 C407S, EC1-4 C407A, EC4-7, and EC7-MAD10. Collectively, our results indicate a weak adhesion mediated by CELSR1 EC repeats alone, that might require the presence and *cis* dimerization of domains downstream of EC6 to exhibit the robust adhesion observed in cell-based aggregation assays[56,117].

Results

Bead aggregation assays do not support strong adhesion by CELSR1 cadherin repeats

To determine the adhesive properties of the CELSR1 cadherin ectodomain and identify a potential minimal adhesive EC repeat unit, we used bead aggregation assays[128]. Like classical cadherins, CELSR1 has a prodomain at the N-terminal tip preceding the first EC repeat (Fig. 3.1 *A*). Classical cadherin prodomains are known to inhibit adhesion and are cleaved[129]. The prodomain region in *hs* CELSR1, spanning residues L(-214) to R(-3) (residue numbering corresponds to processed protein with negative values for signal peptide and prodomain), is followed by a furin protease cleavage site, -3RARR0 (RX[K/R]R). The cleavage and the subsequent absence of the prodomain in the secreted protein were tested *in vitro* by the expression of *hs* CELSR1 EC1-5Fc fragment in HEK293T cells (Fig. 3.8 *A*). Because prodomain cleavage was apparent, we created a library of constructs encoding only for the EC repeats without the prodomain and fused to the native signal peptide (SP) at the N-terminal end. These protein fragments (SP-EC1-9 to SP-EC1-2 and SP-EC2-9) fused with an Fc tag at the C-terminal end were produced in

HEK293T cells (see Methods) and used along with protein G magnetic beads for aggregation assays. The SP label for protein fragments used in bead aggregation assays will be omitted below as these are expected to be cleaved in the mature, secreted proteins.

Control experiments carried out with *hs* PCDH1 EC1-7Fc[113] showed clear Ca^{2+}-dependent bead aggregation after 1 h of incubation and 2 min of rocking (R2) (Fig. 3.1 *B*). However, beads coated with the *hs* CELSR1 EC1-9Fc fragment exhibited poor Ca^{2+}-dependent bead aggregation at R2 (Fig. 3.1 *D*). For the consecutive C-terminal EC repeat deletion series, the largest bead aggregate area was seen for the EC1-5Fc fragment (Fig. 3.1 *H*) in the presence of Ca^{2+} followed by EC1-4Fc, EC1-9Fc, and EC1-8Fc (Fig. 3.1 *I*, *D*, and *E*). Negligible bead aggregation was observed for EC1-7Fc, EC1-6Fc, EC1-3Fc, and EC1-2Fc fragments under the conditions tested (Fig. 3.1 *F*, 1*G*, 1*K*, and 1*L*). Deletion of EC1 in the N-terminal end abolished the poor Ca^{2+}-dependent bead aggregation (Fig. 3.1 *M*). In the absence of Ca^{2+}, bead aggregation was abolished for all fragments (Fig. 3.8 *B*-*L*). Unlike control *hs* PCDH1 EC1-7Fc, the size of the bead aggregates did not increase for CELSR1 fragments between 0 min and R2 (Fig. 3.1 *O* and *P*). Quantitatively, we observed no statistically significant differences ($p < 0.05$) in bead aggregate sizes when comparing results obtained using different CELSR1 fragments and their corresponding negative controls with EDTA (Fig. 3.1 *P*, 3.8 *M*-*N*).

We also tested the adhesiveness of selected *hs* CELSR1 fragments (EC1-MAD10, EC1-9, EC1-5, EC1-4 and EC7-MAD10) in the presence of 2 mM Mg^{2+} and 2 mM Ca^{2+}.

Quantitative analysis revealed results comparable to those obtained in experiments with 2 mM Ca^{2+} (Fig. 3.9 *B-I*). In addition, we tested the adhesiveness of selected *hs* CELSR1 fragments using a protocol similar to that used for testing CELSR2 adhesiveness[63]. We did not detect significant differences between results obtained using this protocol and our standard protocol (Fig. 3.10 *B-H*). Under the conditions tested, bead aggregation assays cannot detect Ca^{2+}-dependent adhesion mediated by CELSR1 EC repeats, which suggest that *trans* homophilic CELSR1 adhesion is weak when compared to PCDH1 and other cadherins, or that other receptors or co-factors are needed to observe the robust adhesion detected in cell-based aggregation assays.

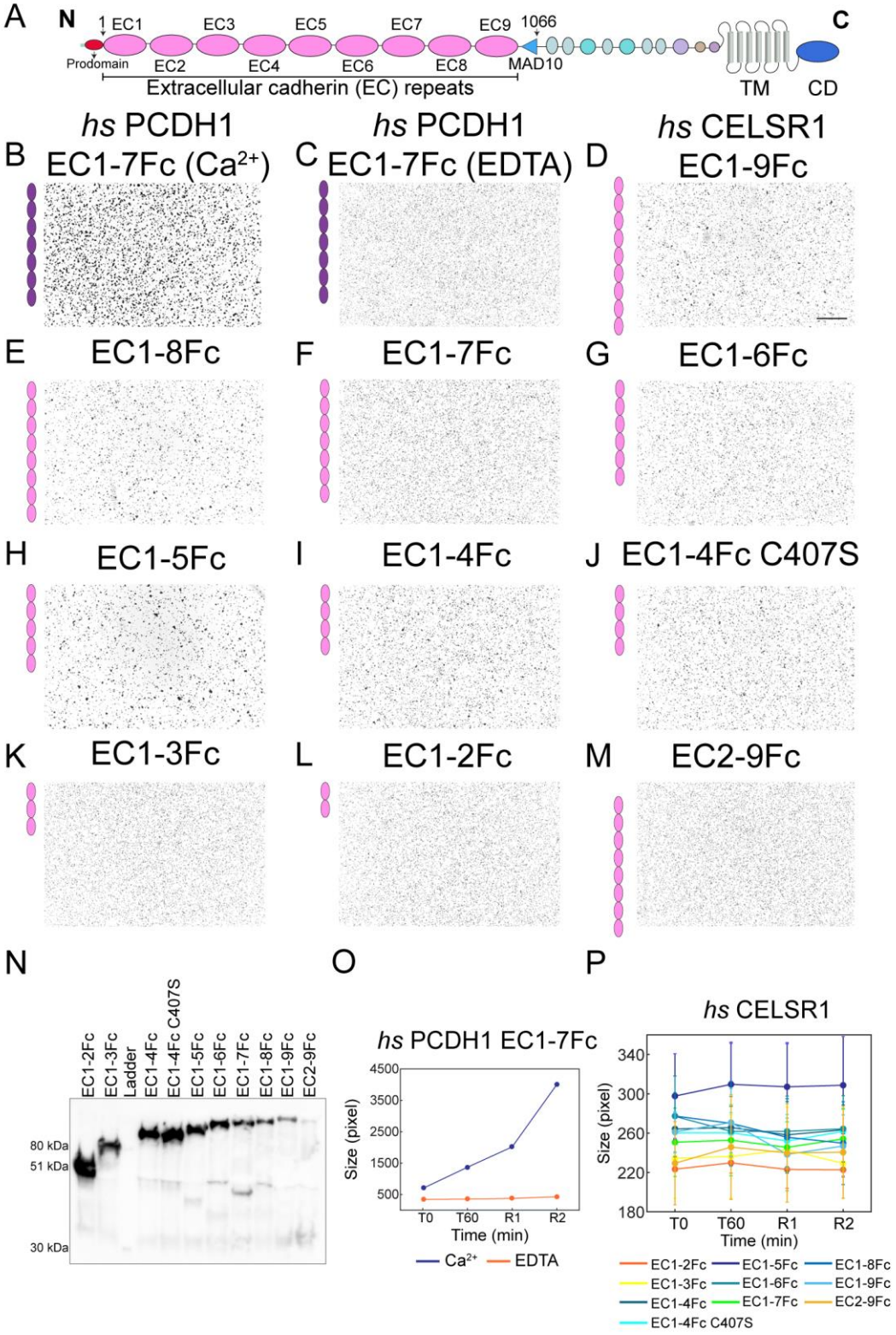

A
N
1 EC1
EC3
EC5
EC7
EC9
1066
C
Prodomain
EC2
EC4
EC6
EC8
MAD10
Extracellular cadherin (EC) repeats
TM
CD
B
hs PCDH1
EC1-7Fc (Ca2+)
C
hs PCDH1
EC1-7Fc (EDTA)
D
hs CELSR1
EC1-9Fc
E
EC1-8Fc
F
EC1-7Fc
G
EC1-6Fc
H
EC1-5Fc
I
EC1-4Fc
J
EC1-4Fc C407S
K
EC1-3Fc
L
EC1-2Fc
M
EC2-9Fc
N
EC1-2Fc
EC1-3Fc
Ladder
EC1-4Fc
EC1-4Fc C407S
EC1-5Fc
EC1-6Fc
EC1-7Fc
EC1-8Fc
EC1-9Fc
EC2-9Fc
80 kDa
51 kDa
30 kDa
O
hs PCDH1 EC1-7Fc
Size (pixel)
4500
3500
2500
1500
500
T0
T60
R1
R2
Time (min)
Ca2+
EDTA
P
hs CELSR1
Size (pixel)
340
300
260
220
180
T0
T60
R1
R2
Time (min)
EC1-2Fc
EC1-5Fc
EC1-8Fc
EC1-3Fc
EC1-6Fc
EC1-9Fc
EC1-4Fc
EC1-7Fc
EC2-9Fc
EC1-4Fc C407S

Figure 3.1. CELSR1 bead aggregation assay results

(*A*) Schematic representation of *hs* CELSR1 with its cadherin ectodomain, transmembrane (TM) and cytoplasmic domains (CD) highlighted. Signal peptide and prodomain are identified by a green line and red oval. Residues G1 to P1066 include the entirety of EC repeats and MAD10. Residue numbering corresponds to the processed protein. (*B*-*C*) Protein-G beads coated with *hs* PCDH1 EC1-7 fused with a C-terminal Fc tag were used as positive control for comparison with *hs* CELSR1 ECX fragments. Experiments were done in the presence of Ca^{2+} and EDTA. Panel (*C*) is also in Figure S1B. (*D*-*M*) Protein-G beads coated with truncated EC fragments of CELSR1 with a C-terminal Fc tag were imaged to detect aggregation in the presence of 2 mM Ca^{2+} after 1 h of incubation and rocking after 1 and 2 min. Images are representative of different sample areas and biological replicates. Scale bar: 500 µm. Results from control experiments in the absence of Ca^{2+} are shown in Fig. 3.8. (*N*) Western blot indicates the presence of the Fc-tagged CELSR1 proteins during the assay. (*O*) Size analysis of the aggregated beads in pixels for *hs* PCDH1 EC1-7Fc at time points 0 min (T0), 60 min (T60), after 1 min of rocking (R1) and after 2 min of rocking (R2). (*P*) Size analysis of the aggregated beads in pixels for *hs* CELSR1 fragments. Error bars indicate standard deviations from three independent repeats for each fragment, except for EC1-5Fc, which was repeated four times.

The crystal structure of *hs* CELSR1 EC1-4 reveals a canonical cadherin architecture

To gain further insights into the cadherin region of the CELSR1 ectodomain, we produced *hs* CELSR1 EC1-4 in bacteria for crystallization and structure determination. These protein fragments did not include the SP and prodomain. The crystal structure of *hs* CELSR1 EC1-4 was refined at 3.5 Å resolution with two molecules in the asymmetric unit, which included residues G4 to V445 in chain A and S3 to N446 in chain B. All EC repeats adopted the typical cadherin Greek-key fold with seven ß-strands labeled A to G (Fig. 3.2 *A* and *B*). Both chains had missing loop densities in EC4 between ß-strands C and D, which were not modeled. Otherwise, good quality electron density maps allowed us to unambiguously position residue side chains. Chains A and B had a root-mean-square-deviation (RMSD) of 2.46 Å after backbone alignment (Fig. 3.11 *A*). Linker regions in the EC1-4 structure between consecutive EC repeats were canonical, coordinating either Na^+ or Ca^{2+} at site 1 and Ca^{2+} at sites 2 and 3 (Fig. 3.2 *B*-*E*, 3.11 *B* and *C*). The Ca^{2+}-binding sites 1 and 2 are proximal to the pre-linker EC and site 3 is buried in the post-linker EC (Fig. 3.2 *C*-*E*). Ions at site 1 are typically exposed to and partially coordinated by solvent and may have a weaker binding affinity[69]. In our EC1-4 structure, the ion at site 1 was found to be Na^+ in linker regions between EC1 and EC2 in chain A and B (Fig. 3.2 *C* and 3.11 *B*) and between EC2 and EC3 in chain B (Fig. 3.11 *C*), likely due to the high salt (3 M NaCl) crystallization condition and the solvent exposed nature of the site. Both Ca^{2+} and Na^+ were coordinated by highly conserved acidic residue motifs DXE and XEX^{base} from the pre-linker ECs, DXNDN in the linker region, and DXD and XDX^{top} from the post-linker ECs (Fig. 3.2 *C*-*E*). Chain A had three smaller A ß-strands in its EC1 repeat, which

were absent and replaced by a single, longer A ß-strand in chain B. We also observed a large density that was left unmodelled at the end of the EC4 repeat in chain B, likely from the hexahistidine tag used for purification (Fig. 3.11 *D*).

An analysis of crystallographic contacts using the protein interfaces, surfaces, and assemblies (PISA) server[130] revealed five interfaces (Fig. 3.12 *A*, *C* and *F-H*). None of the crystal contacts had buried surface area (BSA) greater than 856 Å^2, an empirical threshold that distinguishes biologically relevant interactions from crystallographic artifacts that occur due to packing[131]. The largest interface with a BSA of 829.6 Å^2 was observed for an anti-parallel dimer between EC3 and EC4 of chains A and B (Fig. 3.12 *A* and *B*). The second largest interface was observed for an X-dimer between EC1 and EC2 for chains A and B with a BSA of 782.9 Å^2 (Fig. 3.12 *C* and *D*). Analytical ultracentrifugation-sedimentation velocity (AUC-SV) experiments, however, showed that *hs* CELSR1 EC1-2 does not dimerize even at high concentrations (83 µM; Fig. 3.12 *E*). Overall, the crystal structure of *hs* CELSR1 EC1-4 did not reveal potential adhesive interfaces and showed a canonical cadherin architecture for this protein fragment.

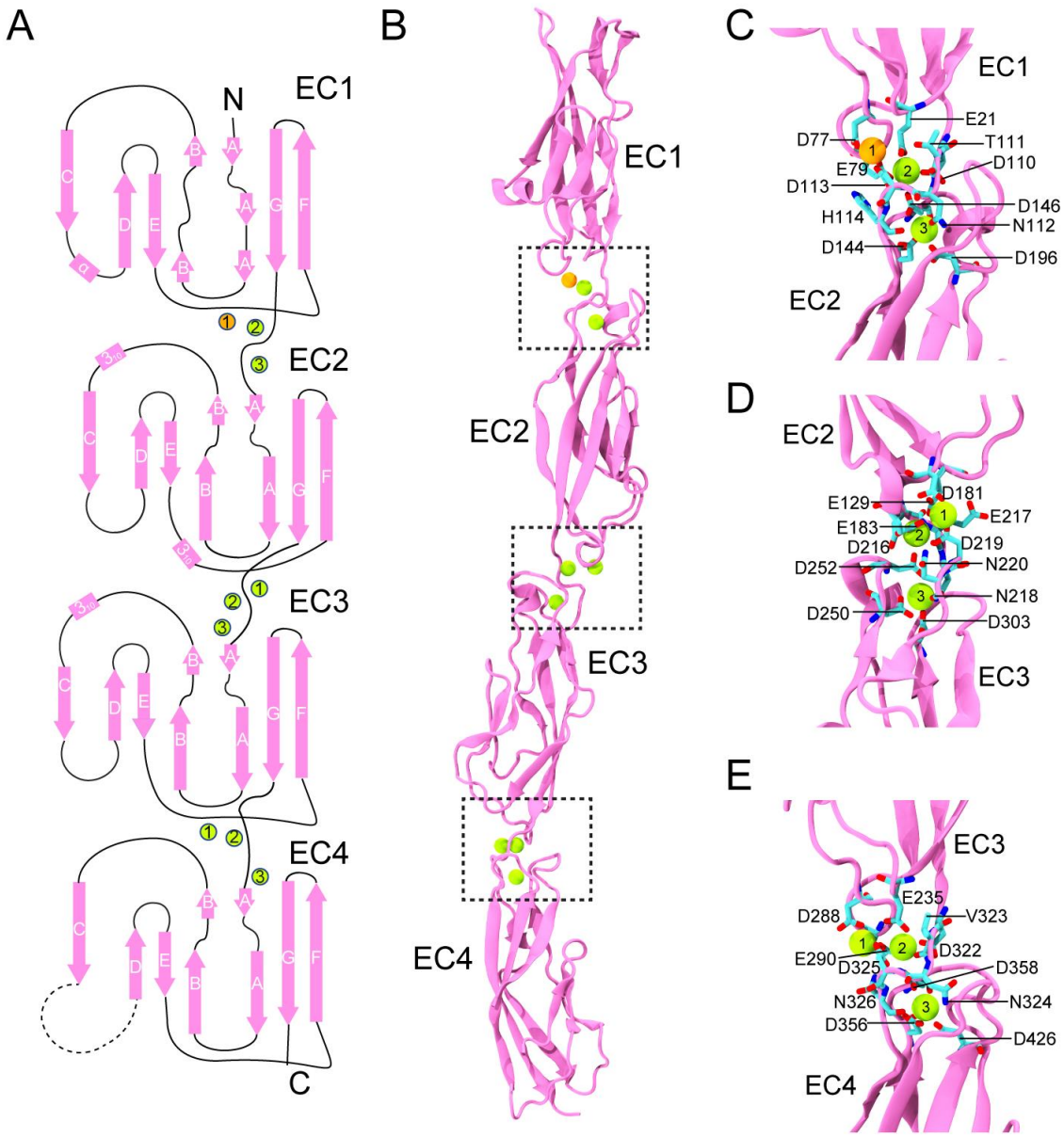


Figure 3.2. Crystal structure of hs CELSR1 EC1-4

(*A*) Topology diagram representing the secondary structure of each cadherin repeat. Missing loop (371-391) in EC4 is shown as a dashed line. Ca^{2+} are shown as green circles and Na^{+} as an orange circle. (*B*) The *hs* CELSR1 EC1-4 structure (chain A) in cartoon representation. Ca^{2+} and Na^{+} are shown as green and orange spheres, respectively. (*C-E*) Linker regions EC1-2, EC2-3, and EC3-4 are highlighted. Ca^{2+} coordinating side chains are shown in blue sticks and labelled.

Hs CELSR1 EC1-4 does not form high-affinity complexes in solution

To further investigate the oligomeric state of *hs* CELSR1 EC1-4 in solution, we used AUC-SV. At concentrations ranging from 1 to 22 μM, we observed a concentration-dependent increase in a small potentially dimeric fraction of the sample (Fig 3.3 *A*). In SEC-MALS, at concentrations ranging from 26 to 83 μM, we observed a shift in the molecular weight from a monomer to that of a dimer (Fig. 3.3 *B*). The native EC1-4 construct had a cysteine in EC4 repeat pointing outside. To test whether this residue was promoting intermolecular disulfide bond formation at high concentrations, we mutated the cysteine to a serine. The *hs* CELSR1 EC1-4 C407S fragment did not show a dimeric species in AUC-SV at 16 μM but a broader peak was observed (Fig. 3.3 *C*). In SEC-MALS, a concentration-dependent molecular weight increase was seen for the mutant fragment. At a concentration of 153 μM, the molecular weight of EC1-4 C407S corresponded to that of the dimer (Fig. 3.3 *D*). To preclude the possibility of C407S causing electrostatic repulsions between the probable dimer interface that might have caused no visible dimer peaks in AUC-SV, we mutated the cysteine to an alanine. Like *hs* CELSR1 EC1-4 C407S, *hs* CELSR1 EC1-4 C407A exhibited broader peaks at concentrations > 10 μM in AUC-SV (Fig. 3.3 *E*). These results indicate a rapid monomer-dimer equilibrium exchange of the *hs* CELSR1 EC1-4 fragment in solution, also evidenced by a single eluting species in SEC-MALS. We therefore suspect a weak dimer for *hs* CELSR1 EC1-4 fragments with a higher K_D value (> 50 μM) compared to other cadherins in the superfamily.

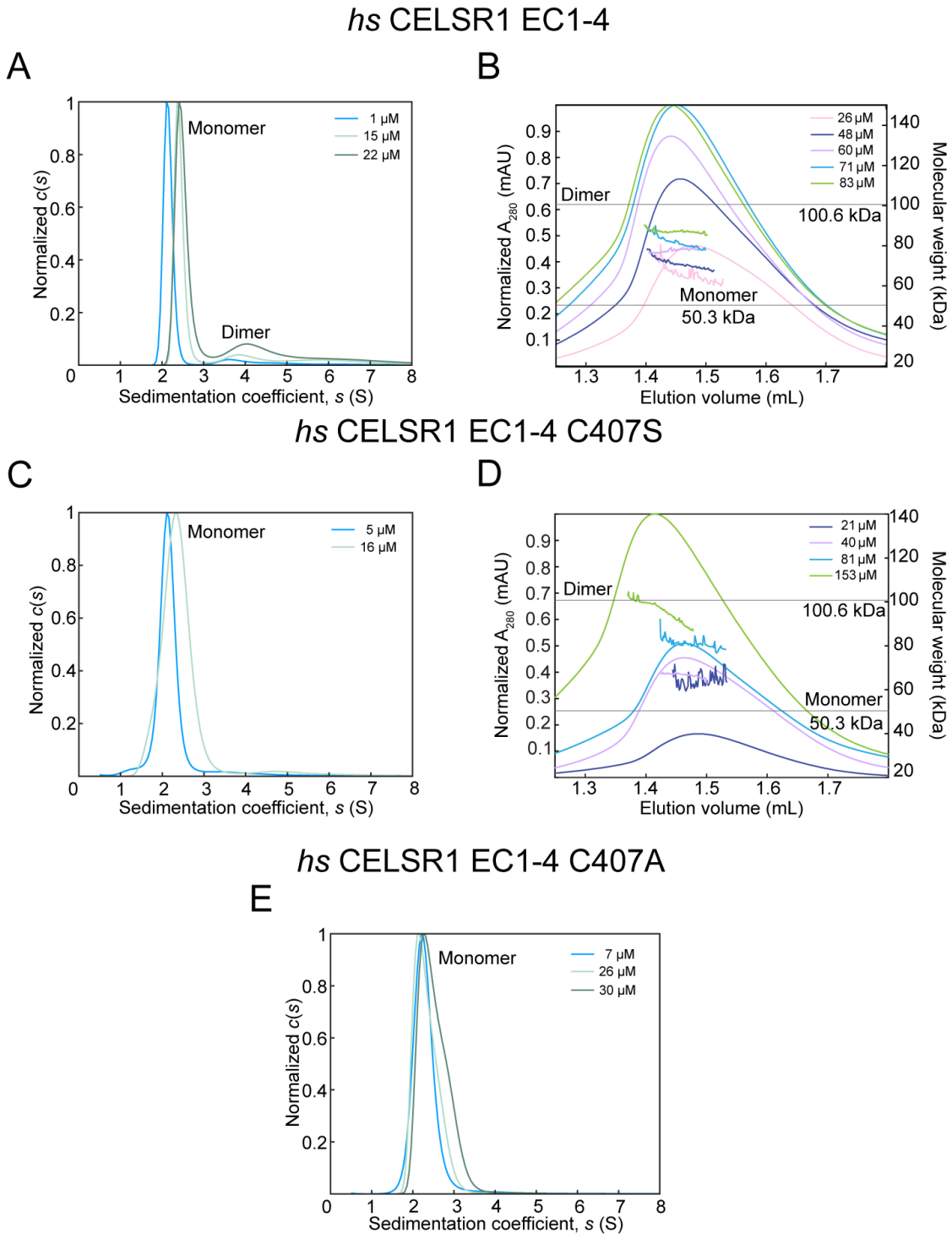


Figure 3.3. Oligomerization state of *hs* CELSR1 EC1-4 in solution

(*A*) AUC-SV results for *hs* CELSR1 EC1-4 at 1 μM, 15 μM, and 22 μM (10 °C) show two peaks, a monomer and a potential dimer based on SEDFIT analysis. (*B*) SEC-MALS of *hs* CELSR1 EC1-4 at concentrations ranging from 26 μM to 83 μM. (*C*) AUC-SV results for *hs* CELSR1 EC1-4 C407S at 5 μM and 16 μM (10 °C). (*D*) SEC-MALS of *hs* CELSR1 EC1-4 C407S at concentrations ranging from 21 μM to 153 μM. Estimated molecular weights increase with concentration. (*E*) AUC-SV results for *hs* CELSR1 EC1-4 C407A at 7 μM, 26 μM, and 30 μM (10 °C).

Hs CELSR1 EC4-7 has a non-canonical linker region between EC5 and EC6

CELSR1 is predicted to have a non-canonical linker region between EC5 and EC6, where the DXE motif is 516DYK518, and the DXNDN motif is 546DANTH550. The non-canonical substitutions are conserved across species and in orthologs, CELSR2 and CELSR3 (Fig. 3.21). To understand how the non-canonical linker region between EC5 and EC6 in CELSR1 affects the rigidity and architecture of the extracellular domain in CELSR1, bacterially expressed EC4-7 fragments were purified and crystallized for structure determination. The crystal structure of *hs* CELSR1 EC4-7 was refined at 2.3 Å resolution. The solved structure had two molecules in the asymmetric unit, that included residues V323 to N758 in chain A and D322 to N758 in chain B with all EC repeats adopting the seven ß-strand Greek-key motif (Fig. 3.4 *A* and *B*). Good quality electron density maps allowed us to unambiguously position most residue side chains except for missing loop densities in EC4 between ß-strands C and D in both chains. The two protein chains had an RMSD of 4.50 Å after backbone alignment (Fig. 3.13 *A*). The EC4-5 and EC6-7 linker regions in chains A and B were canonical with conserved Ca^{2+}-binding motifs DXE, XEX^{base}, DXNDN, DXD, and XDX^{top} coordinating three Ca^{2+} at sites 1, 2, and 3 (Fig. 3.4 *C* and *E*). As expected, the non-canonical EC5-6 linker region lacked a bound ion at site 1 and featured a Na^{+} at site 2 in both chains (Fig. 3.4 *D* and 3.13 *B*). The electron density, coordination distances, and B factor values of the ion and its surrounding residues supported the placement of Na^{+} instead of Ca^{2+} (Fig. 3.13 *C* and *D*).

Although we hypothesized that the non-canonical EC5-6 linker would be bent, chains A and B in the EC4-7 crystal structure were partially straight and formed an antiparallel dimer with a BSA of 2,036 $Å^2$ (Fig. 3.14 *A*). This is much larger than the empirical threshold of 856 $Å^2$ that distinguishes biologically relevant interactions from crystallographic artifacts[131]. Two salt-bridge interactions between R497 in EC5 chain A and E581 in EC6 chain B and vice versa stabilized this dimer interaction along with several electrostatic and buried hydrophobic interactions (Fig. 3.14 *B*). There were six additional crystallographic interfaces (Fig. 3.14 *C*, *E*-*G*) with one antiparallel dimer having an interface area of 898 $Å^2$ (Fig. 3.14 *C*), albeit at a slightly different configuration when compared to the top scoring dimer. A Mg^{2+} ion at this interface was found between EC6 in chain B and EC4 in chain A (Fig. 3.14 *D*) but bead aggregate sizes in the presence of Mg^{2+} were comparable to tests conducted with only 2 mM Ca^{2+} ruling out this divalent as possible co-factor (Fig. 3.1 *D*-*M*, 3.9 *B*-*C*). Regardless, this protein fragment was monomeric in solution in SEC-MALS experiments. At concentrations of 52 μM and 131 μM, the protein always eluted with a molecular weight closer to that of the theoretical weight of the monomer (49 kDa) (Fig. 3.14 *H*). Hence, we inferred that the interfaces seen in the crystal structure of EC4-7 were due to crystal packing and that the EC4-7 fragment is monomeric in solution.

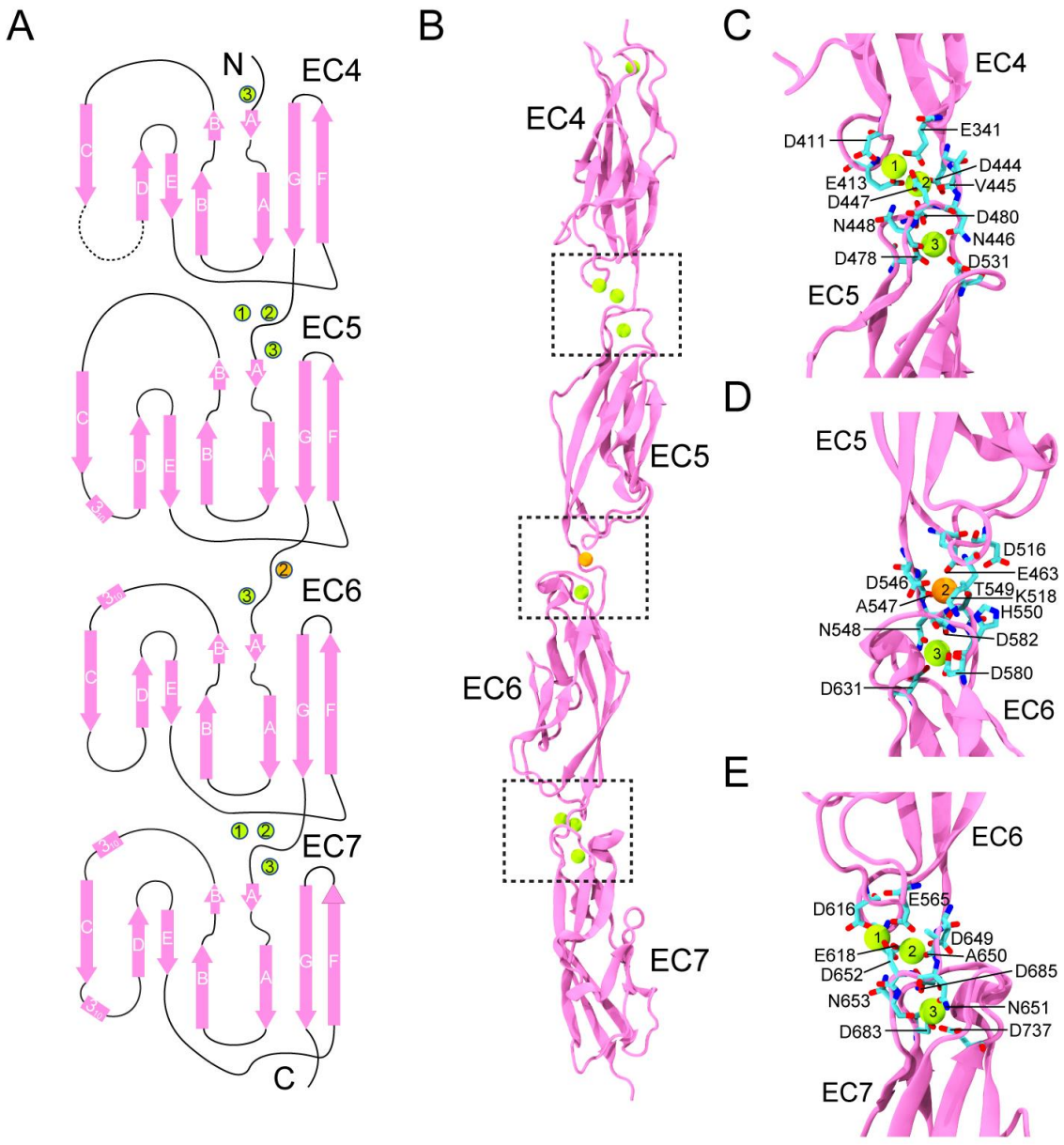


Figure 3.4. Crystal structure of *hs* CELSR1 EC4-7

(*A*) Topology diagram representing the secondary structure of each cadherin repeat shown as in Figure 3.2 *A*. (*B*) The *hs* CELSR1 EC4-7 structure (chain A) in cartoon representation. (*C-E*) Linker regions EC4-5, EC5-6, and EC6-7 are highlighted and shown as in Figure 3.2. The EC5-6 linker region is non-canonical and binds Ca^{2+} at site 3 and Na^{+} at site 2.

***In silico* analysis and in-solution envelope of CELSR1 EC4-7 reveal a flexible linker**

To analyze the conformations of the monomeric EC4-7 protein fragment, we used molecular dynamics (MD) simulations[132]. Chain A of the *hs* CELSR1 EC4-7 structure was used to generate a complete model including the missing EC4 loop. Two equilibrium MD simulations (sim1 and sim2; Table 3.1) lasting 200 ns each showed conformational variabilities at the EC5-6 linker region throughout the trajectories. Simulations sim1 and sim2 differed on the ion occupancy at site 2 of the EC5-6 linker region, which was occupied by a Na^+ in sim1 (experimental occupancy) and a Ca^{2+} in sim2 (manually placed instead of Na^+). At the end of 200 ns in sim1, the molecule adopted a bent, L-shaped conformation at an angle close to 90º around the EC5-6 linker region (Fig. 3.5 *A*). The Na^+ ion left site 2 after 155 ns (Fig. 3.5 *A* and *B*). In simulation sim2, the molecule was bent at an angle > 30º in some frames, but Ca^{2+} still occupied site 2 between EC5-6 (Fig. 3.15 *A*), which might have prevented more orientational dynamics.

Although both sim1 and sim2 involved large systems with box dimensions of 167 × 166 × 196 $Å^3$, *hs* CELSR1 EC4-7 in sim1 started to rotate and came close to within 12 Å of its neighbor in the periodic box at ~ 62 ns (Fig. 3.5 *B*). With a van der Waals cutoff of 12 Å, we wanted to exclude the possibility of EC4-7 adopting the bent conformation in sim1 due to periodic boundary contacts. Hence, a third smaller system (sim3) was built (228 × 99 × 100 $Å^3$) in which we constrained four C_α atoms from EC4 (residues 329, 355, 365, and 425; k_c = 1 kcal/mol/$Å^2$) to prevent overall rotations. Like in sim1, the *hs* CELSR1

EC4-7 fragment in sim3 showed a similar bent conformation towards the middle of the simulation and an empty 2nd site at the EC5-6 linker region (Fig. 3.16 *A*).

We also monitored the orientation of EC5, EC6, and EC7 relative to EC4, EC5, and EC6, throughout the simulation trajectories obtained for all three systems (Fig. 3.5 *C-E*, 3.15 *C-E* and 3.16 *C-E*). Some flexibility was observed for the EC4-5 and EC6-7 linker regions. The EC5-6 linker region was the most flexible in sim1 and sim3. The EC6 repeat was bent with respect to EC5 at angles > 30°, sometimes reaching 90° in sim1 (Fig. 3.5 *D*). A gradual decrease in the end-to-end distance from 200 Å to 100 Å between C_α atoms of residues V323 in EC4 and N758 in EC7 was seen in sim1 after Na^+ vacated the EC5-6 linker region (Fig. 3.5 *B*). In sim3, a similar trend in end-to-end distance decrease was observed from 200 Å to 150 Å for up to 50 ns when Na^+ left the EC5-6 linker region (Fig. 3.16 *B*). Unlike sim1 and sim3, the end-to-end distance in sim2 was close to 200 Å throughout the run due to negligible bending around the EC5-6 linker region because of Ca^{2+} coordination at site 2 (Fig. 3.15 *B*). These results suggest that the EC5-6 linker region can be flexible and bend.

Table 3.1. Summary of MD simulations

Label	**System (PDB ID)**	**Harmonic constraints (k = 1 kcal/mol/Å^2)**	**Time (ns)**	**System size (# atoms)**	**Box dimensions (Å^3)**
Sim1	*hs* CELSR1 EC4-7 + Na^+	None	200	525,915	167 × 166 × 196
Sim2	*hs* CELSR1 EC4-7 + Ca^{2+}	None	200	525,917	167 × 166 × 196
Sim3	*hs* CELSR1 EC4-7 + Na^+	C_α atoms of residues 329, 355, 365, and 425	110	216,407	228 × 99 × 100

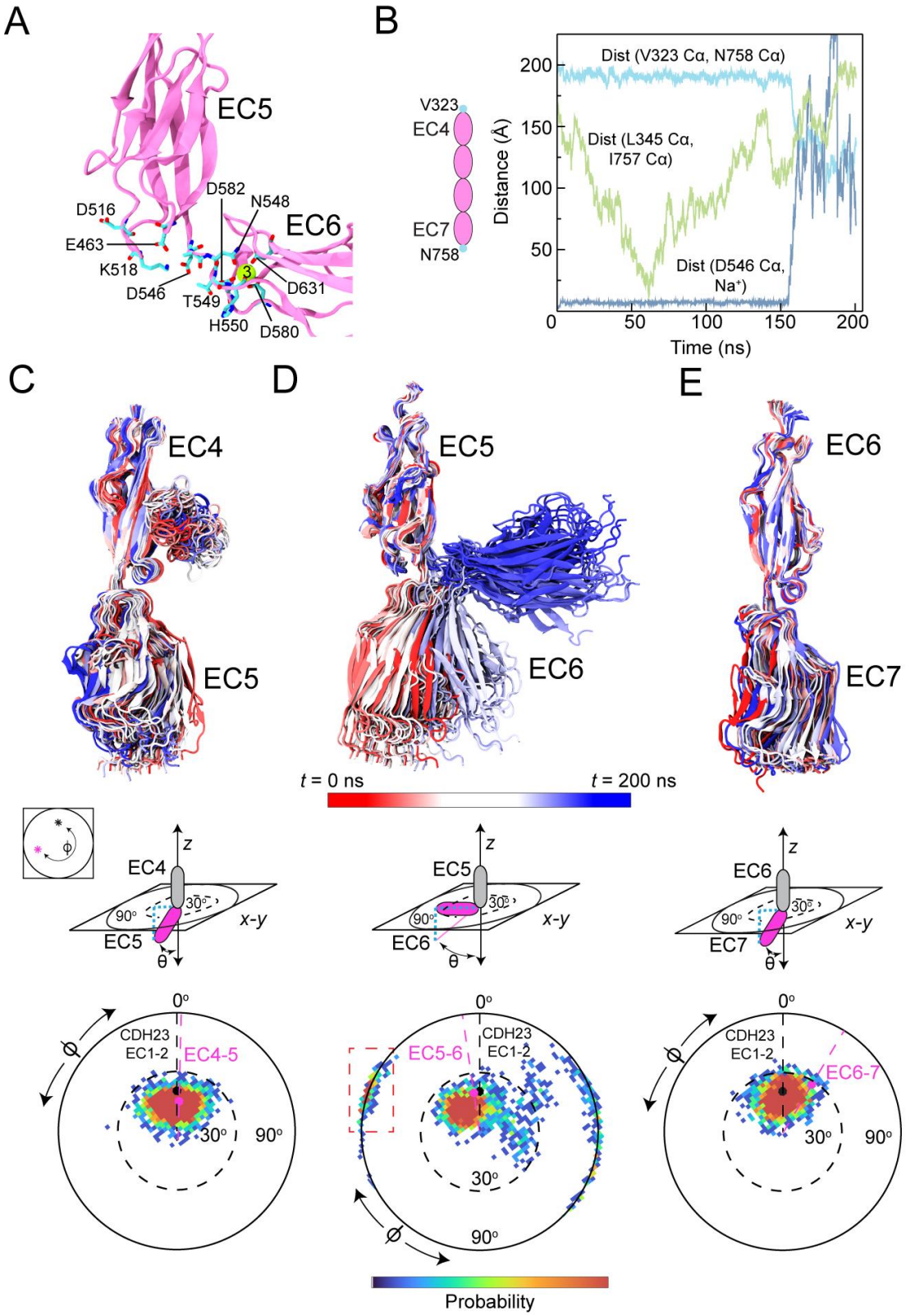

A
EC5
EC6
D582
N548
D516
E463
K518
D546
T549
H550
D580
D631
B
V323
EC4
EC7
N758
Dist (V323 Cα, N758 Cα)
Dist (L345 Cα, I757 Cα)
Dist (D546 Cα, Na+)
Distance (Å)
Time (ns)
C
EC4
EC5
D
EC5
EC6
E
EC6
EC7
t = 0 ns
t = 200 ns
z
x-y
θ
ϕ
90°
30°
CDH23 EC1-2
EC4-5
EC5-6
EC6-7
0°
30°
90°
Probability

Figure 3.5. *In silico* orientation of EC repeats with Na+ in the EC5-6 linker region

(*A*) The EC5-6 linker region after 200 ns of equilibration (sim1). (*B*) Distances between C_α atoms of residues V323 in EC4 and N758 in EC7 (cyan), between C_α of L345 in EC4 and I757 in EC7 across the periodic cell boundary (olive green); and between C_α of D546 and Na^+ in the EC5-6 linker region (blue). (*C-E*) Top panels show superposition of conformations every 5 ns for repeats *hs* CELSR1 EC4-5, EC5-6, and EC6-7 (sim1). Red represents the initial conformation at $t = 0$ ns and blue indicates the final conformation at $t = 200$ ns. Bottom panels show the projections of the longest principal axes of *mm* CDH23 EC2 and *hs* CELSR1 EC5, EC6, and EC7 on the *x-y* plane calculated every 100 ps and plotted after alignment of EC4, EC5, and EC6 to *mm* CDH23 EC1 oriented in the *z* direction (PDB ID: 2WHV). The initial projections of *mm* CDH23 EC2 (black), *hs* CELSR1 EC5, EC6, and EC7 (mauve) are indicated in circles. Color indicates probability.

Due to the flexible nature of *hs* CELSR1 EC4-7 *in silico*, we wanted to observe the in-solution architecture of the fragment, so we used SEC-SAXS (Fig 3.6 and Table 3.2). The same construct used for crystallization was used for experiments at a concentration of 170 μM. Analysis of the intensity data (Fig. 3.6 *A*) using the Guinier plot indicated an R_g value of 48.76 Å (Fig. 3.17 *A*). The Kratky plot indicated that the protein was folded in solution but with some flexibility that might have arisen due to inter-repeat motion or due to the C-terminal hexahistidine tag used for purification (Fig. 3.17 *B*). Fourier transform analysis of the data using a D_{max} value of 168.8 Å (Fig. 3.6 *B*) gave an R_g value of 50.81 Å, in good agreement with the Guinier analysis. An *ab initio* analysis of the data resulted in an envelope that was L-shaped (Fig. 3.6 *C*). Out of fifteen models generated by DAMMIF, fourteen with a mean normalized spatial discrepancy (NSD) value of 1.057 were accepted. X-ray scattering intensities calculated from the filtered bead model (Fig. 3.6 *C*) generated by DAMMIN/DAMMIF had a χ^2 value of 1.064, indicating a good fit with the experimental scattering intensities (Fig. 3.6 *A*). A concurrent MALS analysis and a molecular weight prediction from the intensity data using the SAXS MoW2 server gave molecular weights of 52.7 kDa and 50.8 kDa, confirming that the protein fragment was monomeric (theoretical molecular weight of 49 kDa). When we attempted to calculate the theoretical scattering intensities of the straight *hs* CELSR1 EC4-7 crystal structure (chain A), we obtained a χ^2 value of 23.64, indicating a bad fit[133,134]. Hence, the L-shaped conformation adopted by the molecule at ~173 ns in sim1 (Fig. 3.5 *D*) was fit into the filtered bead model and the envelope acquired from SAXS. The χ^2 value of 3.63 indicated

a better fit but also implied that further simulation sampling would be needed to capture the specific bent conformation observed in solution (Fig. 3.6 *C*).

Table 3.2. Data collection parameters

Data collection parameters	
Protein	***hs* CELSR1 EC4-7**
Instrument	SIBYLS beamline
Beam geometry	Point focus
Wavelength (Å)	1.127
q range $Å^{-1}$	0.008-0.46
Frame slicing (s)	2
Exposure times (min)	25
Concentration (μM)	170
Temperature (K)	298
Structural parameters	
R_g (Å) [from $P(r)$]	50.8
R_g (Å) [from Guinier]	48.8
D_{max} (Å)	168.8
Molecular mass determination using SAXS MoW2	
Molecular Mass (kDa)	50.8
Calculated monomer mass from sequence (kDa)	49
Discrepancy (%)	3.7
Oligomeric state	Monomer
Software employed	
Primary data reduction	Beamline software
Data processing	PRIMUS, GNOM
Ab initio analysis	DAMMIN
Validation and averaging	DAMAVER
Computation of model intensities	FoXS

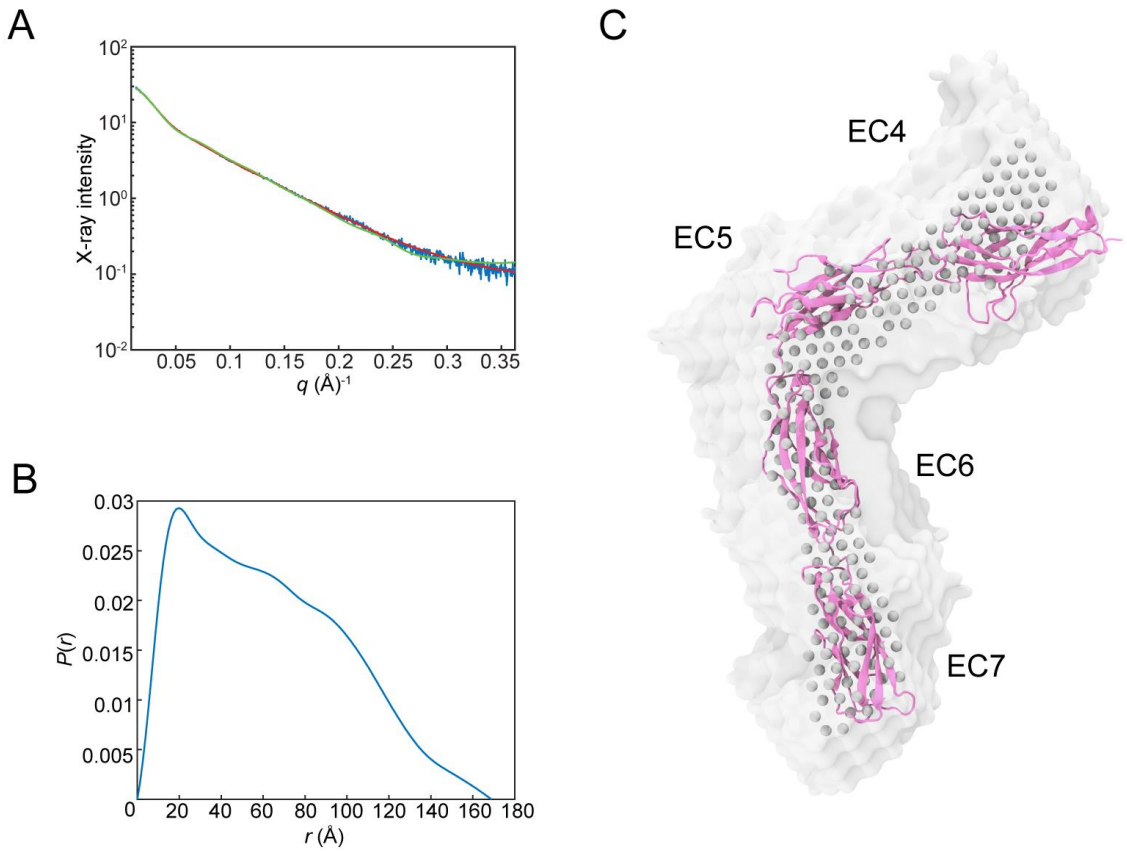


Figure 3.6. Oligomeric state and conformations of *hs* CELSR1 EC4-7 in solution

(*A*) SAXS intensity plot of *hs* CELSR1 EC4-7 at 170 μM. Red line is the fit of a bead model shown in (*C*) to scattering intensity data with a χ^2 of 1.064. Green line is the fit of the *in silico* model shown in (*C*) with a χ^2 of 3.63. (*B*) Fourier transform of the intensity data using a D_{max} of 168.8 Å. (*C*) Filtered bead model (gray) from 15 runs with an average envelope derived from the DAMMIN/DAMMIF analysis. An *in silico* model (sim1, ~ 173 ns) fitted to the filtered *ab initio* bead model is shown in mauve.

Parallel CELSR1 dimerization downstream of EC6 between EC7 and MAD10

Next, we investigated repeats EC7-9 with a membrane adjacent domain (MAD10) found immediately downstream of EC9. In PCDH15, MAD12 induces Ca^{2+}-independent parallel *cis* dimerization[37,40]. We hypothesized a similar role for MAD10 in CELSR1. To test this hypothesis, we bacterially expressed and purified *hs* CELSR1 EC7-MAD10. Although we obtained two peaks in SEC experiments, the presumable dimer peak was inseparable from the aggregate peak. The *hs* CELSR1 EC7-MAD10 fragment has two cysteine residues in EC9 (C891 and C956). We assumed that the poor behavior in SEC experiments was due to undesirable intermolecular disulfide bridge formation during folding and mutated the cysteines (C891S and C956A) to circumvent the issue. These cysteine residues are predicted to point away from each other and are unlikely to form a disulfide bond in the WT protein. The chromatograms obtained using the mutant and the WT protein fragments were similar, and upon re-purification of the mutant, we obtained more protein and two distinct peaks, fractions of which were separately collected and concentrated at 18 μM and 24 μM (Fig. 3.18 *A*).

During AUC-SV (Fig. 3.18 *B*), the SEC monomeric fraction had large monomeric and minor dimeric species (35%). In contrast, the SEC dimeric fraction only had a minor monomeric species (15%). Both samples had higher-order oligomeric species. These results show that, unlike the *hs* CELSR1 EC1-4 fragment, the dimer formation in *hs* CELSR1 EC7-MAD10 is robust and independent of cysteine residues. Bead aggregation assays testing for *trans* interactions showed no aggregation for beads coated with WT *hs*

CELSR1 EC7-MAD10 (Fig. 3.9 *G* and 3.10 *G*), implying that the dimerization of mutant EC7-MAD10 fragments observed in solution is due to a parallel *cis* dimer formation (Fig. 3.18 *A*-*B*). Whether this *cis* dimerization is physiologically relevant is unclear.

Constructing a hybrid structural model of the entire CELSR1 ectodomain

To determine whether we could use AlphaFold2 (AF2)[135] to build a hybrid model of CELSR1, we first computed backbone RMSD values of individual EC repeats as well as of successive EC (ECX, ECX+1) repeats using crystal structures and AF2 predictions[62,135,136]. AF2 did exceptionally well for EC1-7, as evidenced by low RMSD values (< 2 Å) when comparing predictions to experimentally determined structures (Fig. 3.19 and Table 3.4). Therefore, we used AF2 to predict the entire ectodomain. In the predicted model, we identified an intriguing interaction between EC9, MAD10, and EGF-like 1 domain, where EGF-like 1 was nested between EC9 and MAD10 (Fig. 3.18 *C*). We also observed a disordered region formed by residues P2049-G2099, an exposed cleavage-deficient GAIN site, and a transmembrane domain with the characteristic seven α helices of GPCRs.

We used our crystal structures of *hs* CELSR1 EC1-4 and EC4-7, as well as our bent EC4-7 model from sim1 along with the AF2 prediction of the ectodomain to build two hybrid monomeric structural models (M1 and M2) of the CELSR1 protein up to its transmembrane region (Fig. 3.7 *A* and *B*). In these models, orientation of domains downstream of EC7 and the interaction between EC9-MAD10 and EGF-like 1 are not

experimentally validated. However, the overall shape and length (~ 45 nm for M1) of the CELSR1 ectodomain closely resembles those recently obtained for CELSR2 using high-speed atomic force microscopy (HS-AFM) imaging[63], including the length for EC1-8 estimated here to be 38.8 nm and measured as 39.1 ± 0.5 nm. In addition, bent conformations similar to those of M2 were also observed[63] for CELSR2 suggesting that our CELSR1 models can represent in-solution states. Although we were not able to predict and build parallel CELSR1 dimeric models, simulated AFM views[137] of the monomeric conformations in M1 and M2 are highly consistent with the CELSR2 HS-AFM data[63] (Fig. 3.7 *A* and *B* insets).

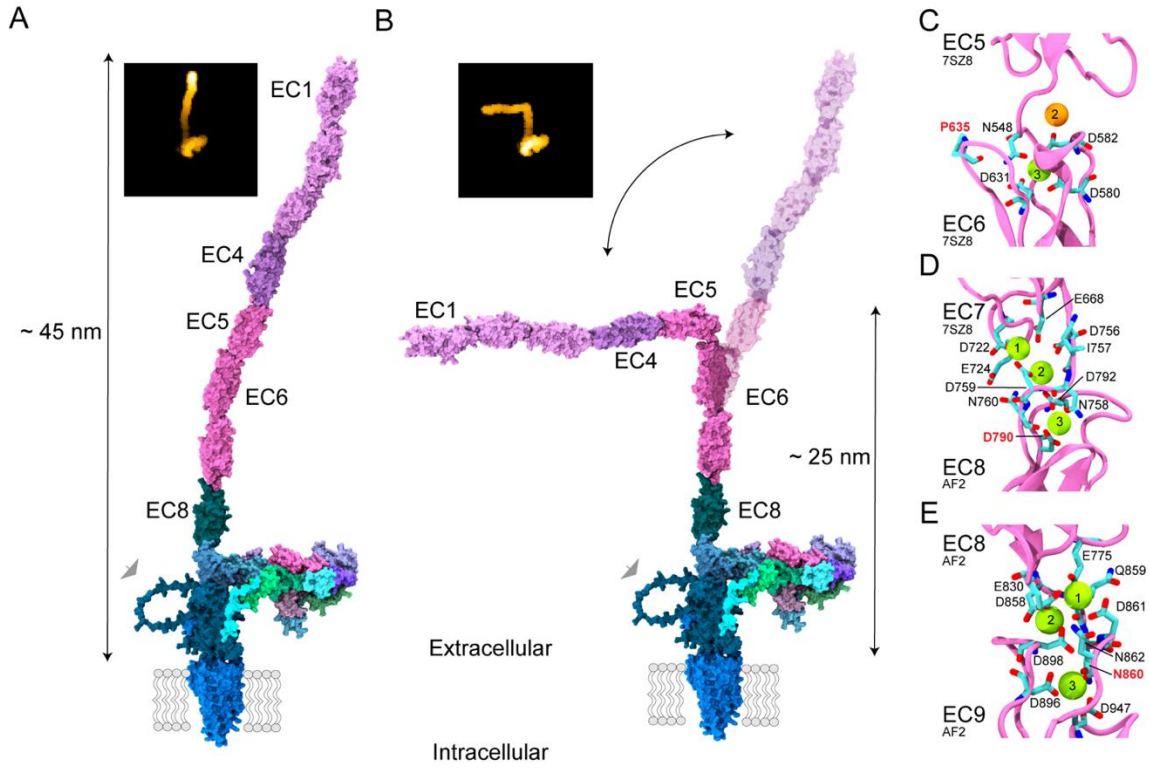


Figure 3.7. Hybrid structural models of *hs* CELSR1

(*A*) A hybrid model (M1) of *hs* CELSR1 with its ectodomain and transmembrane domain based on the crystal structures of *hs* CELSR1 EC1-4 and EC4-7, as well as AF2 predictions. Gray arrowhead indicates disordered region. (*B*) A hybrid model (M2) of *hs* CELSR1 as in (A) but using a bent conformation for EC4-7 from sim1 (~ 173 ns). Transparent molecular surface shows the straight conformation in (*A*). Insets in (*A*) and (*B*) show simulated AFM views[137] of our CELSR1 models M1 and M2 without the transmembrane domain (residues G4-I2235). These are similar to recently reported HS-AFM images of CELSR2[63]. (*C-E*) Location of CELSR1 residues involved in disease and altered morphogenesis mapped to our hybrid model, including: P635 with P635L (equivalent to P870L) implicated in congenital heart and neural tube defects in humans and zebrafish (*C*); D790 with D790G (equivalent to D1040G) causing defects in inner ear PCP and neural tube closure in the mouse (*D*); and N860 with N860K (equivalent to N1110K) causing a D790G-like phenotype in the mouse.

Discussion

Among core proteins facilitating PCP, CELSR1 adheres cells together through homophilic interactions of its extracellular domain, but molecular-level details of the interaction remain unclear[56,63]. Previous studies have shown the significance of the CELSR1 ectodomain in mediating adhesion through cell-based aggregation assays[56]. Our bead aggregation assays performed only with EC repeats or with EC repeats and MAD10 show that these did not cause significant Ca^{2+}-dependent bead aggregation (Fig. 3.1 *D*-*M*), in contrast to observations for other cadherins such as N-cadherin, PCDH19, PCDH24, CDHR5, and PCDH1 (Fig. 3.1 *B*)[113,114,138,139]. These results are consistent with similar bead aggregation assays done with the *hs* CELSR2 full-length ectodomain and its smaller cadherin fragments[63]. Interestingly, both sets of experiments, ours with CELSR1 and Nishiguchi's with CELSR2, suggest that EC1-4 could be the minimal adhesive unit and that EC5 might be important for robust, yet weak adhesion. We speculate that these results are due to inherently low binding strength, which might be appropriate for a signaling function.

The largest cadherin *trans* interfaces observed so far involve repeats EC1-4, although biophysical studies suggest a larger interface for fly N-cadherin[121] and for human CELSR2[63]. For CELSR1, we observed a single peak with concentration-dependent molecular weight increase for bacterially expressed EC1-4, EC1-4 C407S, and EC1-4 C407A fragments in solution suggesting a rapid monomer-dimer equilibrium exchange (Fig. 3.3). The crystal structure of the same fragment had an EC3-EC4 anti-parallel

interface (Fig. 3.12 *A* and *B*) but with two evolutionarily conserved N-linked glycosylation sites near it (N311 in EC3, and N399 in EC4, Fig. 3.12 *B*, Fig. 3.21)[140]. Glycosylation of these residues *in vivo* could regulate or interfere with the EC3-EC4 interface (potential EC1-6 antiparallel overlap) seen in the crystal structure. With the recent discovery of potential EC1-8 and EC1-5 *trans* dimer interfaces in human CELSR2[63], it remains to be discovered if *hs* CELSR1 mediates adhesion using similar *trans* dimer interfaces. An analysis of our CELSR1 ectodomain model indicates a length of ~ 45 nm. A non-overlapped dimer would be ~ 90 nm, but HS-AFM images of CELSR2 suggest that the anti-parallel dimer is ~ 67 nm. A ~ 23 nm overlap would be consistent with an EC1-5 anti-parallel dimer, which coincides with the fragments that have the largest bead aggregate size in our experiments.

In cadherins, Ca^{2+} binding to linker regions that connect two consecutive EC repeats regulates adhesive function by impacting structural rigidity[30,141,142]. In CELSR1, the EC5-6 linker region coordinates only one Ca^{2+} at site 3 and a Na^{+} at site 2 (Fig. 3.4 *D*). The sequence motifs of this non-canonical linker region are evolutionarily conserved across species and among paralogs (Fig. 3.21). This feature of CELSRs make the linker region more flexible leading to a bent conformation in solution (Fig. 3.6 *C*). Equilibrium MD simulations of *hs* CELSR1 4-7 also revealed the propensity of the initially linear crystal configuration of this fragment to rapidly adopt similar bent conformations after 150 ns (Fig. 3.5 *D*). In terms of adhesion, the flexible EC5-6 linker region might partially explain our bead aggregation results. Typically, protein fragments with larger number of

EC repeats than the minimal adhesive unit induce rapid aggregation[113,114,138]. However, for CELSR1, flexibility imposed by the EC5-6 linker region might have introduced multiple states (i.e., bent, partially bent, and straight) between adjacent molecules, which would have lowered the probability of the molecule existing in a form amenable for adhesion. These dynamic states might explain bead aggregation results when using the EC1-5Fc fragment where the sizes of bead aggregates were larger compared to those obtained using other fragments with the EC5-6 linker region included (Fig. 3.1 *D*-*H*). Interestingly, the mutation P635L identified in EC6 (Fig. 3.7 *C*) induces neural tube defects in a zebrafish model[58,98,143]. In our structure, P635 lies in a loop coordinating Ca^{2+} at site 3 in the EC5-6 linker region (Fig. 3.7 *C*). Mutation of proline to leucine might indirectly interfere with Ca^{2+} binding, flexibility, or with yet uncharacterized adhesion interfaces.

Another intriguing structural feature of CELSR1 is the ferredoxin-like domain labeled MAD10 downstream of EC9. A similar MAD domain induces *cis* dimerization in PCDH15[39,40]. AUC-SV experiments of *hs* CELSR1 EC7-MAD10 showed that the protein sedimented as a potential dimer at low concentrations compared to the *hs* CELSR1 EC1-4 fragment. A previous study established that the *cis*-dimerization-mediating EC7 repeat is important for the formation of stable, evenly spaced, punctate clusters[58]. Mutations in mouse CELSR1, Crsh (D790G) and Spin cycle (N860K), affecting Ca^{2+} coordination at site 3 in the EC7-8 and EC8-9 linker regions (Fig. 3.21) affected the *cis* dimer stability of CELSR1 in tissues[50,58]. The mutated forms of CELSR1 gave rise to disorganized, non-punctate regions that disrupted PCP in cells[58]. In structural terms, we hypothesize that the

disruption of Ca^{2+} coordination at site 3 between EC7 and EC8, which already lacks the XDX motif (Fig. 3.7 *D*) would make the region flexible like EC5-6 and possibly abolish this *cis* dimer interaction. Similarly, the disruption of Ca^{2+} coordination at site 3 between EC8 and EC9 would enhance flexibility and alter *cis* dimerization. The higher turnover of the Crsh mutant also underscores the importance of *cis* interactions in stabilizing the complex *in vivo*[58]. Here, we postulate that the CELSR1 complex stabilization through EC7-MAD10 and the rest of the ectodomain leading to lateral clustering *in vivo* would stabilize and improve *trans* dimer formation by reinforcing an optimal adhesion configuration. Such a prerequisite could also explain the robust aggregation of cells expressing CELSR1 compared to beads coated with only the EC1-MAD10Fc fragment in our experiments. Furthermore, while the CELSR1 sequence is heavily conserved across species, the EC5-MAD10 fragment shows a higher degree of conservation (> 75%) than EC1-4 (< 75%) (Fig. 3.20 *A*). Even among paralogs (CELSR1, 2, and 3), the EC7-MAD10 fragment has a higher percentage identity when compared to EC1-4 (Fig. 3.20 *B*). These results underscore the possible significance of EC7-MAD10 *cis* interactions.

Limitations of our study include the inability of bead aggregation assays to clearly characterize weak Ca^{2+}-dependent adhesion and the possibility that lack of eukaryotic posttranslational modifications confound the interpretation of results obtained with bacterially expressed protein for some of the structural studies. However, a direct comparison between bead aggregation enabled by PCDH1 and CELSR1 ectodomains produced in mammalian cells unequivocally highlights a stark contrast on aggregate size,

and our previous work on PCDH1 has demonstrated that bacterial and mammalian expression systems produce equivalent cadherin ectodomain proteins with proper disulfide bonds[113]. Based on our data, we propose that the relatively weak adhesion exhibited by the N-terminal cadherin repeats of CELSR1 might facilitate its dynamic redistribution during initial events that lead to PCP and might be necessary for a signaling function. During the establishment of complete asymmetry across cell-cell junctions, *cis* dimerization downstream of EC7 might stabilize the complex in the junctional lattice[58]. Recent data indicate that CELSR1, CELSR2, and CELSR3 express differentially in the mouse brain and suggest that CELSR1 and CELSR3 functions are not dependent on auto-proteolytic cleavage of the ectodomain[119]. In addition, cell-based assays suggest that CELSR1 and CELSR2 contribute differently to epidermal PCP[61]. It is possible that the biophysical properties of the CELSR ectodomains vary across paralogs to enable their different functions, or that all CELSR ectodomains mediate weak adhesion tailored for a signaling function.

Methods

Cloning and mutagenesis

DNA sequences encoding for *hs* CELSR1 (NG_030466.2) were subcloned into pET21a and CMV: N1-Fc[114] vectors for bacterial and mammalian expression. The protein sequence (NCBI Ref. NP_001365257.1) of EC1-2 and EC1-4 used for bacterial expression started from residue G1 (EC1) after the predicted prodomain cleavage site and ended at the conserved Ca^{2+}- binding motif DXND**N** of either EC2 or EC4. For the EC4-7 and EC7-MAD10 fragments, the sequences started at the **D**XNDN site of the preceding EC repeat (EC3 or EC6) and ended at the DXND**N** site of EC7 or residue F1067 of MAD10. The constructs were subcloned into the pET21a vector using *NdeI* and *XhoI* restriction sites with a hexahistidine tag at the C-terminus. The truncated constructs EC1-2, EC1-3, EC1-4, EC1-4C407S, EC1-5, EC1-6, EC1-7, EC1-8, EC1-9, EC1-MAD10, EC2-9 and EC7-MAD10 used for mammalian expression were subcloned with the native *hs* CELSR1 signal peptide but without the prodomain using *XhoI* and *BamHI* restriction sites and with an Fc tag at the C terminus. The mature protein sequences started from residue G1 at EC1, residue S115 for EC2 and residue A654 for EC7 ending at L989 for EC1-9 and EC2-9, at the DXND**N** motif for EC1-2 to EC1-8 and at F1067 for MAD10 with an additional GGRDPPV linker for all constructs before the Fc tag. The mutated constructs EC1-4 C407S, EC1-4 C407A and EC7-MAD10 C891S C956A were generated by site directed mutagenesis using the Quikchange lightning site-directed mutagenesis kit (Agilent). All constructs were sequence verified.

Bacterial expression and purification of *hs* CELSR1 fragments

DNA constructs were used for protein expression in BL21 codonPlus (DE-3)-RIPL cells. Transformed cells cultured in LB or TB media were induced at OD_{600} ~ 0.4-0.6 with 800 μM IPTG and grown for ~18 h at 30 °C. Cells were lysed by sonication in a denaturing buffer (20 mM Tris [pH 7.5], 6 M guanidine hydrochloride, 10 mM $CaCl_2$ and 20 mM imidazole). The cleared lysates were loaded onto Ni-Sepharose beads (GE Healthcare) and eluted with denaturing buffer supplemented with 500 mM imidazole. The *hs* CELSR1 EC1-4, EC4-7 and EC7-MAD10 fragments were refolded overnight at 4 °C by dropwise dilution of 30 mL of eluted denatured protein (0.5 mg/mL) into 500 mL of refolding buffer (20 mM Tris [pH 8.0], 200 mM NaCl, 5 mM $CaCl_2$, 400 mM L-arginine, 1 mM tris(2-carboxyethyl) phosphine [TCEP] and 10% glycerol). Refolded proteins were concentrated to 2 mL by centrifugal ultrafiltration (Amicon ultra-15) and purified by size-exclusion chromatography in two steps using a Superdex200 column (cytvia) to a homogenous single peak in 20 mM Tris [pH 8.0], 200 mM NaCl, 5 mM $CaCl_2$ and 1 mM TCEP buffer. The *hs* CELSR1 EC1-2, EC1-4 C407S, EC1-4C407A and EC7-MAD10 C891S C956A fragments were refolded overnight by dropwise dilution and purified in Superdex200 column but without 1 mM TCEP in the buffers. The purified proteins were concentrated to 0.1-5 mg/mL by ultrafiltration (Vivaspin 10 kDa) for crystallization and biochemical assays.

Crystallization, data collection and structure determination

Crystals were grown by sitting-drop vapor diffusion at 4 °C by mixing protein (0.6 μL of *hs* CELSR1 EC1-4 at 4 mg/mL and 1 μL of CELSR1 EC4-7 at 6 mg/mL) with 0.6

μL of reservoir solution (0.1 M Tris [pH 8.5], 3 M NaCl and 0.1 M MES [pH 7.3], 0.1 M $MgCl_2$, 30% PEG400). The *hs* CELSR1 EC1-4 crystals were cryoprotected by exchanging to a reservoir buffer with 25% glycerol. Crystals were cryo-cooled in liquid N_2, and X-ray diffraction data were collected as indicated in Table 3.3. Datasets were processed with HKL2000[144], and structures were obtained by molecular replacement using PHASER[145]. We used RaptorX to generate models for molecular replacement. The final structures were refined using COOT[89] and REFMAC[85] and verified with procheck[87], whatcheck[146] and checkmymetal[88]. Data collection and refinement statistics are provided in Table 3.3 and coordinates for *hs* CELSR1 EC1-4 and EC4-7 have been deposited in the protein data bank (ID: 8D40 for *hs* CELSR1 EC1-4 and ID: 7SZ8 for *hs* CELSR1 EC4-7).

Bead aggregation assays

Bead aggregation assays were modified from those described previously[114,147,148]. The CELSR1-Fc fused DNA constructs were delivered to HEK293 cells (ATCC® CRL-1573™) via calcium phosphate transfection[149]. Solutions of DNA at a concentration of 10 μg dissolved in 61 μL of 250 mM $CaCl_2$ were added dropwise to 2XHBS while under mild vortexing. The resulting DNA-calcium phosphate complex was added dropwise to a 100 mm dish with HEK293 cells cultured in DMEM supplemented with FBS, L-glutamine and PenStrep, that were 60-70% confluent. Two such plates were used for constructs EC1-2 to EC1-7 and 4 plates for EC1-8 and EC1-9 due to low expression. Cells were washed twice with 1XPBS buffer the next day to remove any excess DNA complex. Fresh serum-free DMEM media was added to the plates after washing and the cells were allowed to grow

for two more days until the media was collected and used for bead aggregation assays. The media with the secreted Fc-fused construct was concentrated to ~500 μL using Amicon-15 ultra concentrators and incubated with 1.5 μL of Protein G beads (Invitrogen) for 1.5 h at 4 °C under constant rotation. The Fc bound beads were washed and resuspended in binding buffer (20 mM Tris [pH 7.5], 200 mM NaCl and 0.2% BSA) and split into two equal volumes for addition of either 2 mM $CaCl_2$ (or 2 mM $CaCl_2$ + 2 mM $MgCl_2$) or 2 mM EDTA. This was followed by additional incubation with rotation (40 rpm) for 30 min at 4 °C. The beads were then allowed to aggregate at room temperature in depression slides with five images covering ~ 60% of the slide area taken at t = 0 min, 60 min, and after mild rocking (10 oscillations/min) for 1 and 2 min using a Nikon Eclipse Ti microscope with a 10X objective. Bead aggregation assays were also carried out using a protocol in which 1.5 μL of Protein G beads coated with 500 μL of protein fragments suspended in buffer (20 mM Tris [pH 7.5], 200 mM NaCl and 0.2% BSA) with either 2 mM $CaCl_2$ or 10 mM EDTA were transferred to 24-well suspension culture plates[63]. After a brief storage period of 10 min, the plates were rotated for 10 min at 180 rpm in a horizontal shaker followed by a resting period of 10 min and subsequent imaging with a 10X objective. Representative images are shown. All images were individually quantified through calculation of aggregated bead areas at different time points with ImageJ software as described previously[128,148]. Images were thresholded to exclude background with area of detected particles measured in pixels. Average aggregate size (±SEM) for each construct from three independent biological replicates were plotted using MATLAB, except for fragment EC1-5Fc, for which there were four biological replicates.

Western blotting

Western blots were performed to verify the presence of secreted protein before and after the bead aggregation assays. Equal volumes of protein constructs mixed with 4X SDS and β-mercaptoethanol were boiled for 5 min at 95 °C. The denatured proteins were loaded onto an SDS-PAGE gel followed by blotting to PVDF membrane for 1 h at 4 °C. The membrane was blocked with 5% non-fat milk in 1XTBS (20 mM Tris [pH 7.5], 150 mM NaCl + 0.1% TWEEN-20) for 1 h and incubated overnight with goat anti-human IgG (1:200; Jackson ImmunoResearch Laboratories, Catalog-109-025-003, Lot-117103, 135223) at 4 °C. After washing with 1XTBS thrice to remove excess primary antibodies, the membrane was incubated with mouse anti-goat HRP-conjugated secondary antibody (1:5,000; Santa Cruz Biotechnology, Catalog-sc-2354, Lot-A2017, J1718) for 1 h at room temperature. After additional washing with 1XTBS thrice, the membrane was developed with ECL select western blot detection kit (GE healthcare) for chemiluminescent detection (Omega Lum G).

Analytical ultracentrifugation

Sedimentation velocity AUC experiments were performed using a ProteomeLabXL-I analytical ultracentrifuge (Beckman Coulter). Double SEC-purified protein samples were concentrated to 420 μL using centrifugal ultrafiltration (Vivaspin 10 kDa). Protein concentrations were measured using a Nanodrop 2000c spectrophotometer with molecular weight and extinction coefficient calculated using Protparam[150]. The concentrated samples were loaded into AUC cell assemblies with an Epon centerpiece that

had a 12 mm path length. Buffer solutions without the protein (20 mM Tris [pH8.0], 200 mM NaCl, 5 mM $CaCl_2$, (+ 1 mM TCEP for *hs* CELSR1 EC1-4 WT)) were used as blanks. The An-50 TI rotor was pre-equilibrated to 10 °C for ~2 h to achieve thermal equilibrium. The rotor was spun at 50,000 rpm and the sedimentation was observed using absorption optics. For sample concentrations less than 8 μM, UV absorption at 230 nm was used instead of 280 nm. The sedimentation data were analyzed by SEDFIT using a continuous sedimentation coefficient distribution model $c(s)$[151,152]. Sednterp[152] was used to calculate solvent density (1.00675 g/mL) and buffer viscosity (1.3318 cp) at 10 °C. The partial specific volume of the protein was set to 0.73 mL/g. The resulting $c(s)$ distributions were plotted using MATLAB to obtain sedimentation plots.

SEC- MALS

SEC- MALS data were collected using an ÄKTAmicro systems coupled to a Wyatt miniDAWN TREOS system. The double SEC-purified protein samples at increasing concentrations were passed subsequently through a Superdex-S200 3.2/30 column (cytvia) at 0.05 mL/min. The absorbance at 280 nm and light scattering were monitored throughout the runs and the molecular weight was calculated by using the theoretical extinction coefficient of the monomer (0.811 for *hs* CELSR1 EC1-4 and EC1-4 C407S) and a rodlike model using the Wyatt Astra 6 software.

SEC- SAXS

Data were collected at the SIBYLS beamline 12.3.1 in the Advanced Light Source facility (Berkeley, CA) as described[153,154]. Double SEC-purified protein solutions at different concentrations measured using a Nanodrop 2000c spectrophotometer were passed through an Agilent 1260 series HPLC with a Shodex KW-803 analytical column at a 0.65 mL/min flow rate in the buffer (20 mM Tris [pH8.0], 200 mM NaCl, 5 mM $CaCl_2$, 1 mM TCEP). Half of the protein eluent was used for scattering by X-ray synchrotron radiation and the other half was passed through an agilent 1260 series multiple wavelength detector (MWD), Wyatt Dawn HELOS MALS, Wyatt DyanPo Titan quasi-electric light scattering detector (QELS) and Wyatt Optilab rEX refractometer. Sample eluent directed to X-rays were analyzed using SAXS at $\lambda = 1.127$ Å for the incident light at a detector distance of 2.1 m. Exposure time for each frame was 2 s, with data collected over the course of 25 min. Buffer matched controls were used for subtraction and the final merged data analysis was performed at the beamline using SCÅTTER[155].

An analysis of the merged data was carried out using PRIMUS[156] through the ATSAS[157] program suite which gave us estimates of radius of gyration (R_g) from the Guinier region. An indirect Fourier transform of the data was used to estimate the maximum dimension D_{max} of the particle using GNOM[158]. *Ab inito* modeling was carried out using DAMMIF[159] considering $q < 0.35$ to generate 15 models of which a filtered bead model and an average envelope was developed by DAMAVER[160] with a mean normalized spatial discrepancy (NSD) of > 1. The molecular weight of the protein was calculated

using the SAXSMoW2 server [161] and a Guinier fitting of the intensity data from 0.018 - 0.027 Å and a $q \times R_g$ of 1.302 with q_{max} calculated using the formula $I(0)/I(q_{max}) = 10^{2.25}$ (Table 3.2). Model scattering intensities were calculated using the FoXS server[134]. Data has been deposited in SASBDB with draft ID 5312.

MD simulations

The structure of *hs* CELSR1 EC4-7 (PDB ID: 7SZ8) - chain A was used for MD simulations. Missing residues 373-388 were built using COOT. The visual molecular dynamics (VMD)[90] plugins psfgen, solvate, and autoionize were used to build three systems labeled sim1, sim2 and sim3. Systems sim1 and sim2 were similar except for a crystallographic Na^+ at site 2 in the EC5-6 linker region, which was replaced with a Ca^{2+} (using COOT) in sim2. System sim3 was smaller and the C_α atoms of residues 329, 355, 365 and 425 were constrained (k_c =1 kcal/mol/Å^2). In all systems, hydrogen atoms were added to the structure with crystallographic water molecules; residues D, E, R and K were assumed to be charged; and the protonation states of histidine residues were chosen for optimal coordination with residues in the neighboring chemical environment. Additional water molecules and ions were added to solvate and ionize the system (150 mM NaCl). The box dimensions of each system are indicated in Table 3.1.

MD simulations were performed using NAMD 2.12[162] and the CHARMM36 force field with the CMAP correction and the TIP3P model for water. A van der Waals cutoff of 12 Å (with a switching function starting at 10 Å) was used with periodic boundary

conditions. The particle mesh Ewald method was used to calculate long range electrostatic interactions without cutoff and with a grid point density >1 $Å^3$. An integration timestep of 2 fs was used with SHAKE. Langevin dynamics was enforced to maintain the temperature of the system at 300 K with a damping coefficient of 0.1 ps^{-1}. The *NpT* ensemble at 1 atm was simulated using a hybrid Nosé-Hoover Langevin piston method with a decay period of 200 fs and 100 fs damping time constant. The systems were minimized for 1,000 steps with 200 ns of equilibration each except for sim3 for which equilibration was carried out for 110 ns.

In silico analyses

The RMSD values obtained when comparing chains in the *hs* CELSR1 EC1-4 and EC4-7 crystal structures were calculated using the VMD RMSD plugin after aligning the protein backbone. Only backbone atoms were included in the calculation. For distance plots, the distances between C_α atoms of residues V323-N758, L345-I757, and between the C_α atom of D546 and Na^+ were calculated every 100 ps. The VMD orient plugin[90] was used to compute principal axes and to calculate the relative orientation of EC repeats with respect to each other within a pair by using *mm* CDH23 EC1-2 (2WHV)[69] as a reference. Consecutive EC repeats of *hs* CELSR1 EC4-7 were aligned to CDH23 EC1 in the *z* direction and the relative orientations of EC5, EC6, and EC7 to EC4, EC5, and EC6 were calculated using the projection of the longest principal axes on the *x*-*y* plane. The calculated azimuthal angles (ϕ) were plotted in reference to the orientation of CDH23 EC2 to EC1

which was defined as 0°. Molecular figures were made using VMD[90] and stereo images using PyMOL[163]. Plots were prepared using MATLAB[91] and XMGrace[164].

Comparison of crystal structures to AF2 models

Crystal structures were compared to models predicted using AlphaFold 2.3.1[135]. The backbone of a given EC repeat from the structure and model were aligned to calculate RMSD values. During a single round of predictions, five models of *hs* CELSR1 EC1-7 were generated, and the top-scoring structure was used for RMSD evaluation. For building the CELSR1 ectodomain, we predicted fifty models without the prodomain and signal sequence, with the transmembrane domain, and without the cytoplasmic domain. To build M1 and M2, the top-scoring model from the ectodomain prediction was chosen and the EC1-7 repeats from the model were replaced with crystal structures of *hs* CELSR1 EC1-4 and EC4-7 for M1, and crystal structure of *hs* CELSR1 EC1-4 and EC4-7 from sim1 (~173 ns) for M2. Models were constructed from the TM domain to the N-terminal EC repeat. The RMSD value computed for the EC4 repeats from the structures was 0.9 Å after alignment. We used EC4 from the *hs* CELSR1 EC4-7 crystal structure for M1. In M1 and M2, Ca^{2+} ions were added to the EC7-8 and EC8-9 linker regions based on the distance and coordination geometry of canonical Ca^{2+}-binding residues in those linker regions upon visual inspection.

Simulated AFM images

Hybrid models M1 and M2 for the human CELSR1 ectodomain (residues G4-I2235) were used to generate simulated HS-AFM images with the BioAFMviewer[137]. Simulated scanning of the molecular structures was done with a scan step of 1.0 nm, a cone angle of 5°, and a tip radius of 1 nm.

Sequence alignments and analyses

Sequence alignments of CELSR1, CELSR2 and CELSR3 were carried out using the MUSCLE[165] algorithm on geneious (Geneious Prime 2022) using sequences obtained from NCBI (Table 3.5). Alignment files from geneious were imported to JalView[167] and colored according to % of sequence identity. N- and O-linked glycosylation sites were predicted using NetNGlyc 1.0 and NetOgly 4.0[140].

Quantification and Statistical Analyses

Statistical analyses for bead aggregation assays were performed on data from independent replicates of all fragments. Average bead aggregate size (± SEM) from Ca^{2+} and EDTA images at R2 (Fig. 3.1 *O* and 3.8 *L*) of each replicate were used in a *t*-test (two-sample assuming unequal variances) to compute significance ($p < 0.05$).

Table 3.3. X-ray diffraction data collection and refinement statistics

Data collection	***hs* CELSR1 EC1-4**	***hs* CELSR1 EC4-7**
Space group	*C* 1 2 1	*P* 1
Unit cell parameters		
a, *b*, *c* (Å)	305.65, 90.36, 94.95	42.38, 76.94, 92.74
α, β, γ (°)	90.00, 96.17, 90.00	89.83, 87.78, 76.24
Molecules per asymmetric unit	2	2
Beam source	APS 24-ID-E	APS 24-ID-C
Wavelength (Å)	0.97910	0.97911
Resolution limit (Å)	3.55	2.34
Unique reflections	31,176	47,823
Completeness (%)	98.7 (98.2)	85.3 (82.5)
Redundancy	5.7 (5.3)	3.3 (3.2)
Average I / $\sigma(I)$	10.1 (2.5)	15.4 (8.8)
R_{merge}	0.22 (1.3)	0.06 (0.1)
Refinement		
Resolution range (Å)	50.00-3.55 (3.61-3.55)	50.00-2.35 (2.39-2.35)
R_{work} (%)	21.6	21.3
R_{free} (%)	25.2	27.9
Protein residues (atoms)	849 (6,673)	837 (6,655)
Ligands/ions	18	20
Water molecules	8	180
RMS deviations		
Bond lengths (Å)	0.01	0.01
Bond angles (°)	1.93	1.75
B-factor average		
Protein	125.82	40.15
Ligand/ion	98.39	36.18
Water	94.86	28.38
Ramachandran Plot Region (PROCHECK)		
Most favored (%)	77.4	87.6
Additionally allowed (%)	22.1	12.0
Generously allowed (%)	0.5	0.3
Disallowed (%)	0	0.1
PDB ID code	**8D40**	**7SZ8**

Table 3.4. Backbone RMSD values for individual and adjacent pairs of EC repeats when comparing AF2 models and CELSR1 crystal structures

Repeat	Crystal Structure	RMSD (Å)
EC1	EC1-4	1.53
EC2	EC1-4	0.74
EC3	EC1-4	0.80
EC4	EC1-4	0.85
EC4	EC4-7	1.59
EC5	EC4-7	0.59
EC6	EC4-7	0.41
EC7	EC4-7	0.94
EC1-2	EC1-4	1.39
EC2-3	EC1-4	1.48
EC3-4	EC1-4	1.25
EC4-5	EC4-7	1.27
EC5-6	EC4-7	1.95
EC6-7	EC4-7	0.96

Table 3.5. Accession numbers of CELSR1 sequences from 19 species and human CELSR2 and CELSR3 used for sequence alignment

S/N	Common name (CELSR1)	Scientific name	Accession number NCBI
1	Human	*Homo sapiens*	NP_001365257.1
2	Mouse	*Mus musculus*	NP_034016.2
3	Bat	*Molossus molossus*	XP_036134132.1
4	Tiger	*Panthera tigris*	XP_042849484.1
5	Dog	*Canis lupus dingo*	XP_035552455.1
1	Oryx	*Oryx dammah*	XP_040089293.1
7	Asian arowana fish	*Scleropages formosus*	XP_029105791.1
8	Mexican fish	*Astyanax mexicanus*	KAG9272416.1
9	Milk fish	*Chanos chanos*	XP_030628227.1
10	Greater amberjack	*Seriola dumerili*	XP_022605139.1
11	Viviparous lizard	*Zootoca vivipara*	XP_034983593.1
12	Sand lizard	*Lacerta agilis*	XP_033019419.1
13	Eastern brown snake	*Pseudonaja textilis*	XP_026558732.1
14	Pit viper	*Protobothrops mucrosquamatus*	XP_015679241.1
15	Chicken	*Gallus gallus*	XP_040526552.1
16	Duck	*Aythya fuligula*	XP_032056934.1
17	Pheasant	*Phasianus colchicus*	XP_031459718.1
18	Sperm whale	*Physeter catodon*	XP_023982919.1
19	Drosophila	*Drosophila melanogaster*	NP_724962.3
	(CELSRX)		
20	Human CELSR2	*Homo sapiens*	NP_001399.1
21	Human CELSR3	*Homo sapiens*	NP_001398.2

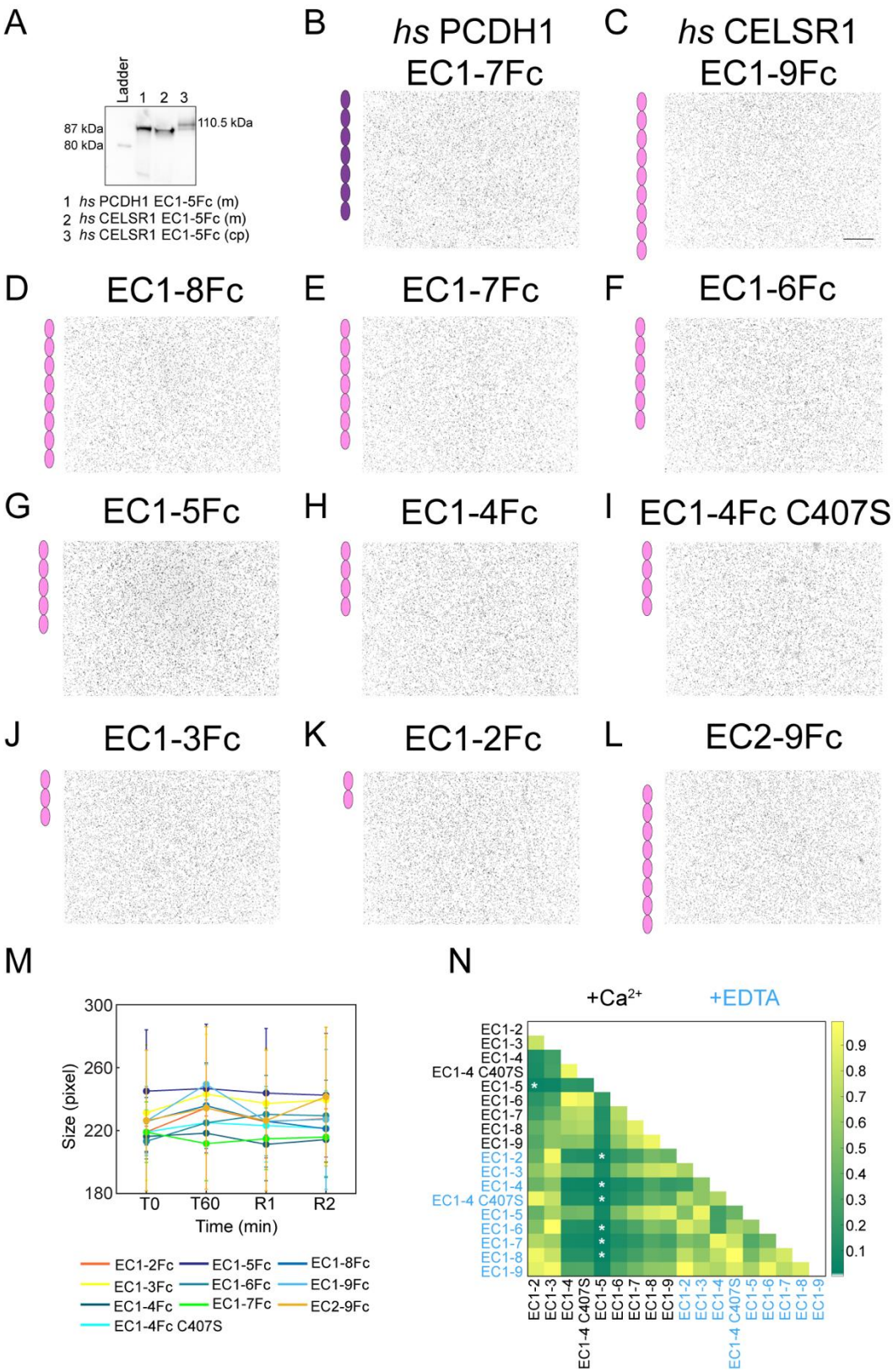

A
Ladder
1 2 3
87 kDa
80 kDa
110.5 kDa
1 hs PCDH1 EC1-5Fc (m)
2 hs CELSR1 EC1-5Fc (m)
3 hs CELSR1 EC1-5Fc (cp)
B
hs PCDH1 EC1-7Fc
C
hs CELSR1 EC1-9Fc
D
EC1-8Fc
E
EC1-7Fc
F
EC1-6Fc
G
EC1-5Fc
H
EC1-4Fc
I
EC1-4Fc C407S
J
EC1-3Fc
K
EC1-2Fc
L
EC2-9Fc
M
Size (pixel)
300
260
220
180
T0
T60
R1
R2
Time (min)
EC1-2Fc
EC1-5Fc
EC1-8Fc
EC1-3Fc
EC1-6Fc
EC1-9Fc
EC1-4Fc
EC1-7Fc
EC2-9Fc
EC1-4Fc C407S
N
+Ca2+
+EDTA
EC1-2
EC1-3
EC1-4
EC1-4 C407S
EC1-5
EC1-6
EC1-7
EC1-8
EC1-9
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1

Figure 3.8. CELSR1 bead aggregation assay results in the presence of EDTA

(*A*) Western blot shows cleavage of the processed *hs* CELSR1 EC1-5 protein secreted in media (m). Media secreted *hs* PCDH1 EC1-5Fc was used as a reference. The *hs* CELSR1 EC1-5Fc extracted from cell pellet (cp) after lysis shows a higher molecular weight band corresponding to the pre-processed protein with a non-cleaved prodomain. (*B*-*L*) Protein-G beads coated with *hs* PCDH1 EC1-7 (also shown in Fig. 3.1 *C*) and truncated EC fragments of *hs* CELSR1 with a C-terminal Fc tag were imaged to detect aggregation in the presence of 2 mM EDTA. Scale bar: 500 μm. (*M*) Size analysis of the aggregated beads in pixels for each fragment at time points 0 min (T0), 60 min (T60), after 1 min of rocking (R1), and after 2 min of rocking (R2). Error bars indicate standard deviations from independent repeats for each protein fragment as indicated in Fig. 3.1. (*N*) *t*-tests performed on comparisons of bead aggregate size values for R2 images obtained using different CELSR1 fragments showed no statistically significant differences ($p < 0.05$) for experiments carried out in the presence of Ca^{2+} and the corresponding negative controls with EDTA. Asterisks show statistically significant differences ($p < 0.05$) for some comparisons involving *hs* CELSR1 EC1-5 and various CELSR constructs, but not for its control with EDTA.

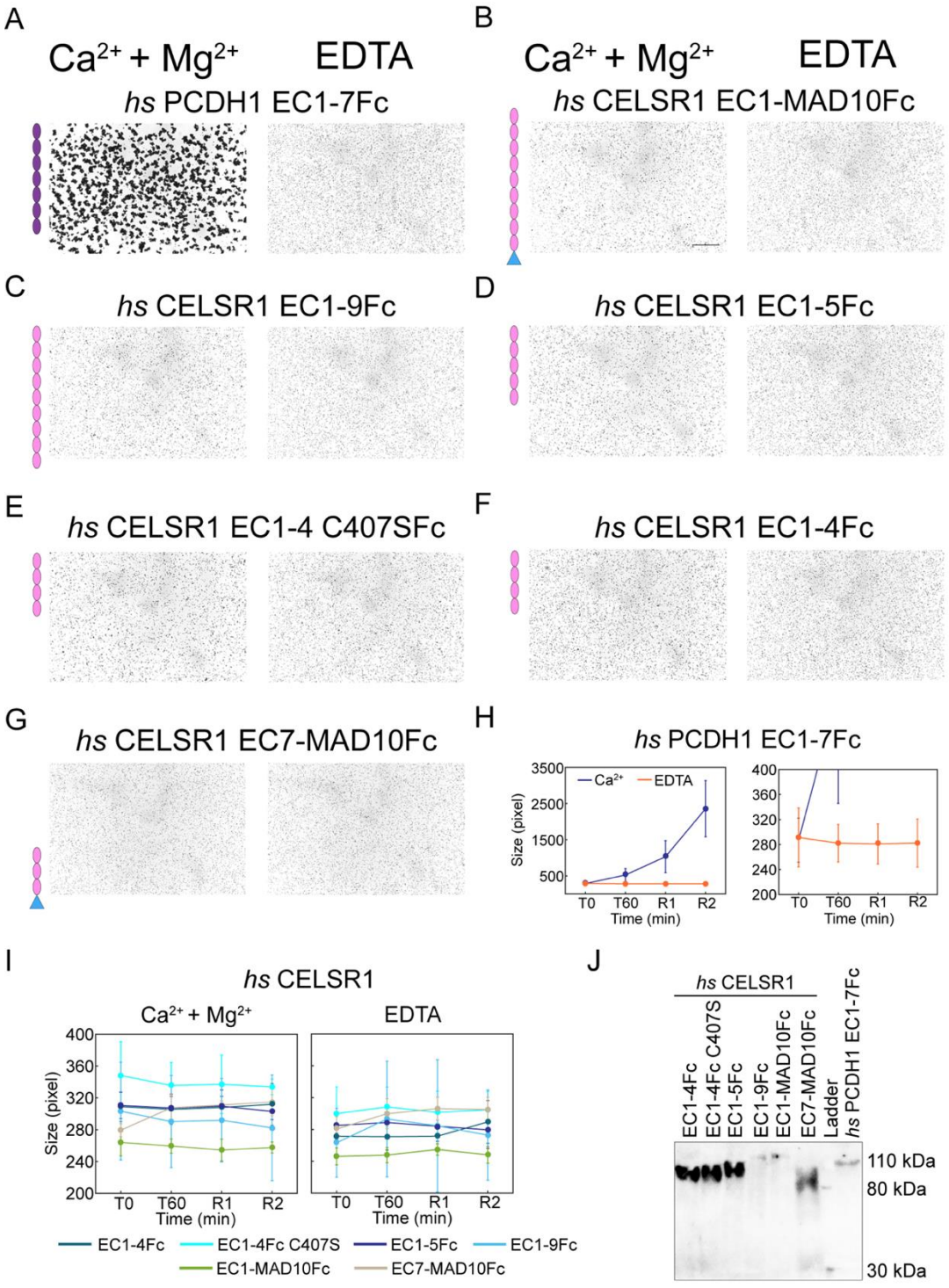


Figure 3.9. CELSR1 bead aggregation assay results in the presence of 2 mM Ca^{2+} + 2 mM Mg^{2+} or 2 mM EDTA

(*A*-*G*) Protein-G beads coated with *hs* PCDH1 EC1-7Fc and truncated EC fragments of *hs* CELSR1 with a C-terminal Fc tag were imaged to detect aggregation in the presence of (2 mM Ca^{2+} + 2 mM Mg^{2+}) or 2 mM EDTA. Scale bar: 500 µm. (*H*) Left: Size analysis of aggregated beads in pixels for *hs* PCDH1 EC1-7Fc in the presence of (2 mM Ca^{2+} + 2 mM Mg^{2+}) or 2 mM EDTA at time points 0 min (T0), 60 min (T60), after 1 min of rocking (R1), and after 2 min of rocking (R2). Error bars indicate standard deviations from three independent repeats as indicated in Fig. 3.1. Right: Zoomed in view of the EDTA control. (*I*) Size analysis of aggregated beads in pixels for truncated CELSR1 fragments. Error bars indicate standard deviations from three independent repeats for each protein fragment. (*J*) Western blot indicates the presence of the Fc-tagged PCDH1 and CELSR1 proteins during the assays.

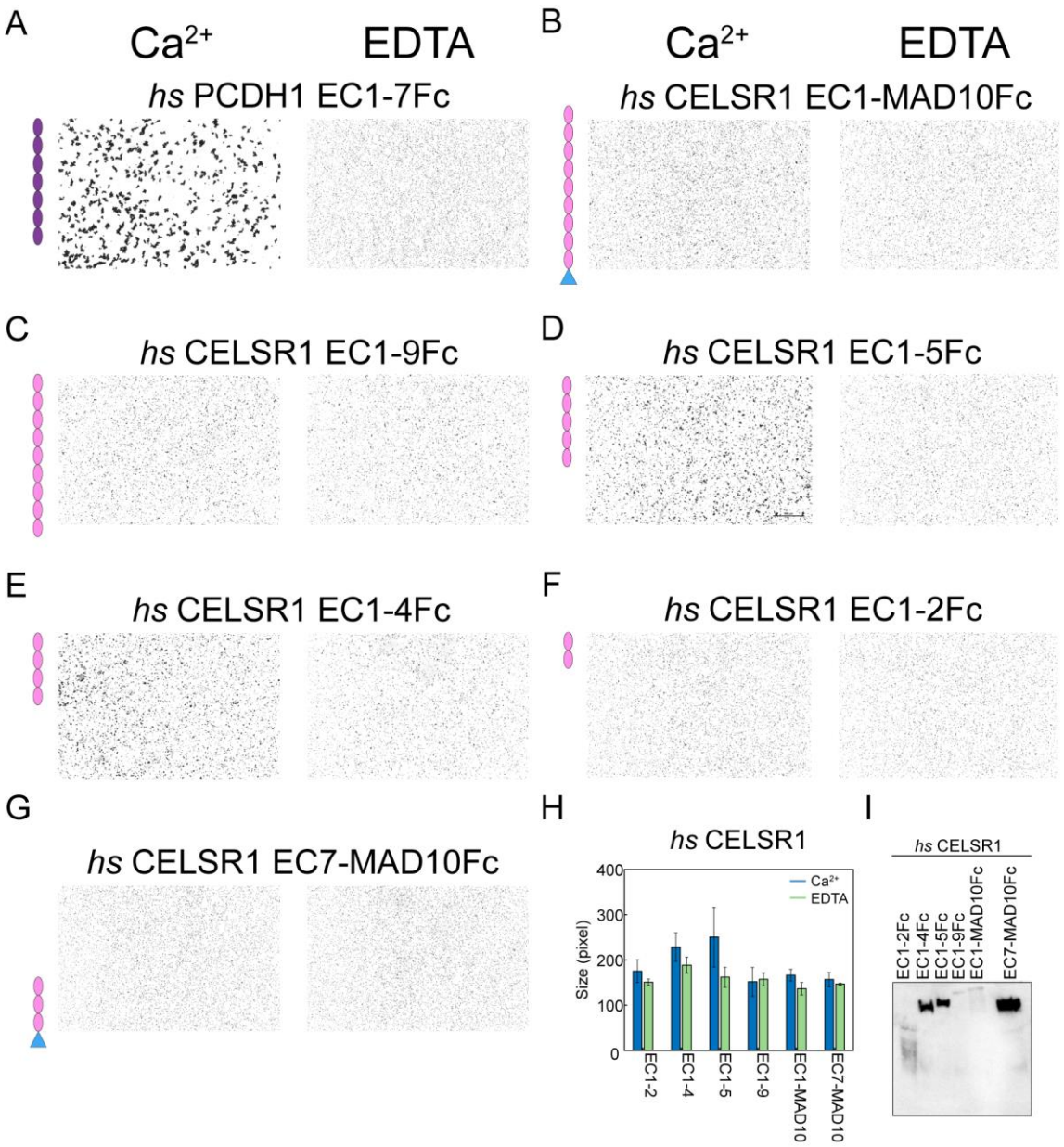


Figure 3.10. CELSR1 bead aggregation assay results in the presence of 2 mM Ca^{2+} or 10 mM EDTA at 180 rpm

(*A*-*G*) Protein-G beads coated with *hs* PCDH1 EC1-7Fc and truncated EC fragments of *hs* CELSR1 with a C-terminal Fc tag were imaged to detect aggregation in the presence of 2 mM Ca^{2+} or 10 mM EDTA. Scale bar: 500 µm. (*H*) Size analysis of the aggregated beads in pixels for each *hs* CELSR1 fragment. Error bars indicate standard deviations from three independent repeats for each protein fragment. (*I*) Western blot indicates the presence of the Fc-tagged CELSR1 proteins during the assays.

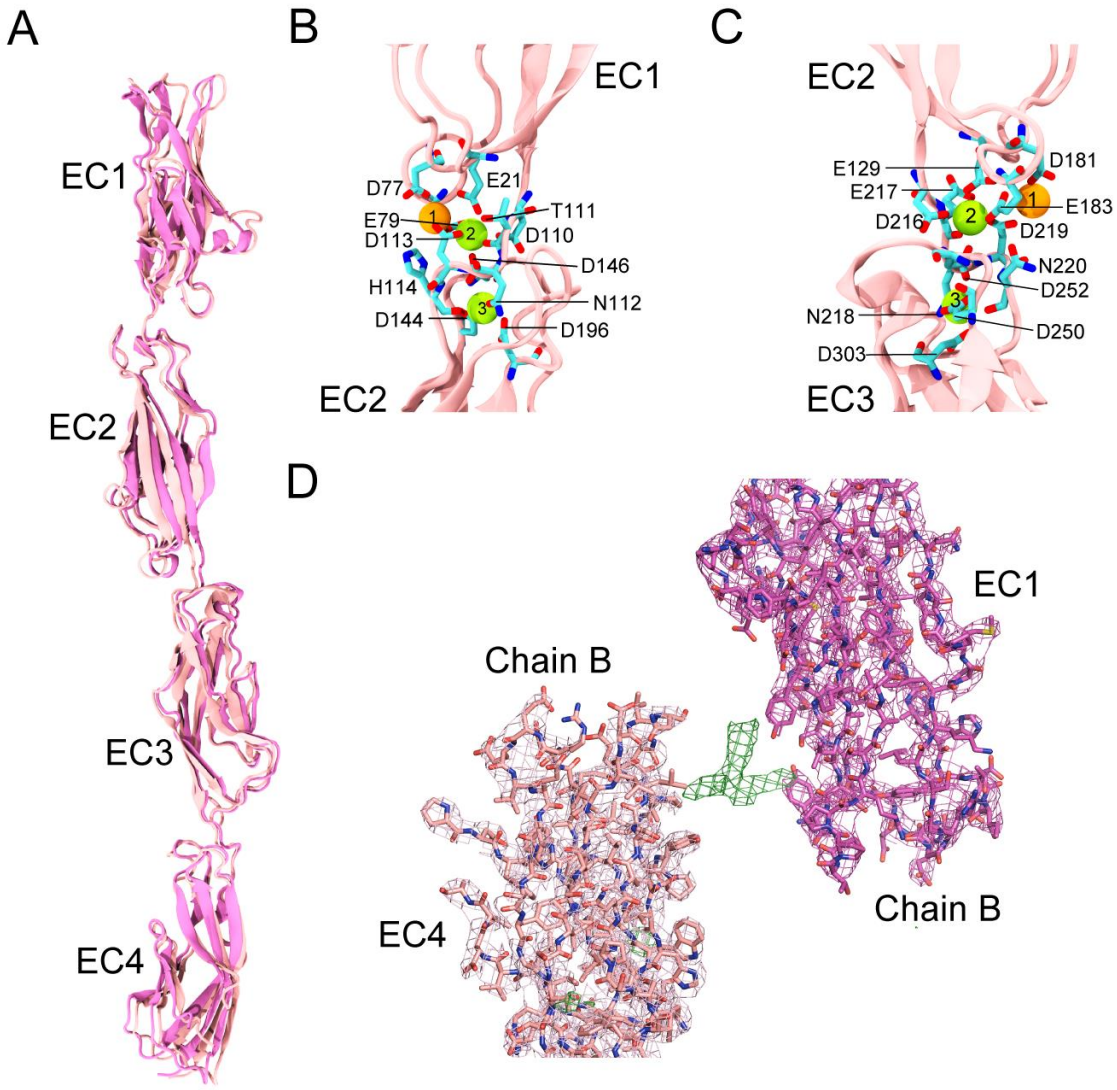


Figure 3.11. Structural features of *hs* CELSR1 EC1-4

(*A*) Structural alignment of *hs* CELSR1 EC1-4 using the two monomers in the asymmetric unit. (*B*, *C*) Highlight of the canonical Ca^{2+}-binding linker regions with Na^{+} at sites 1 in the EC1-2 and EC2-3 linker regions due to excess of Na^{+} in the crystallization condition. (*D*) A $2f_o$-f_c electron density map at $\sigma = 2.0$ contour level showing EC4 and EC1 of chain B in adjacent unit cells. A f_o-f_c electron density map at $\sigma = 5.0$ contour level showing unmodelled density between EC4 and EC1 of chain B in adjacent unit cells.

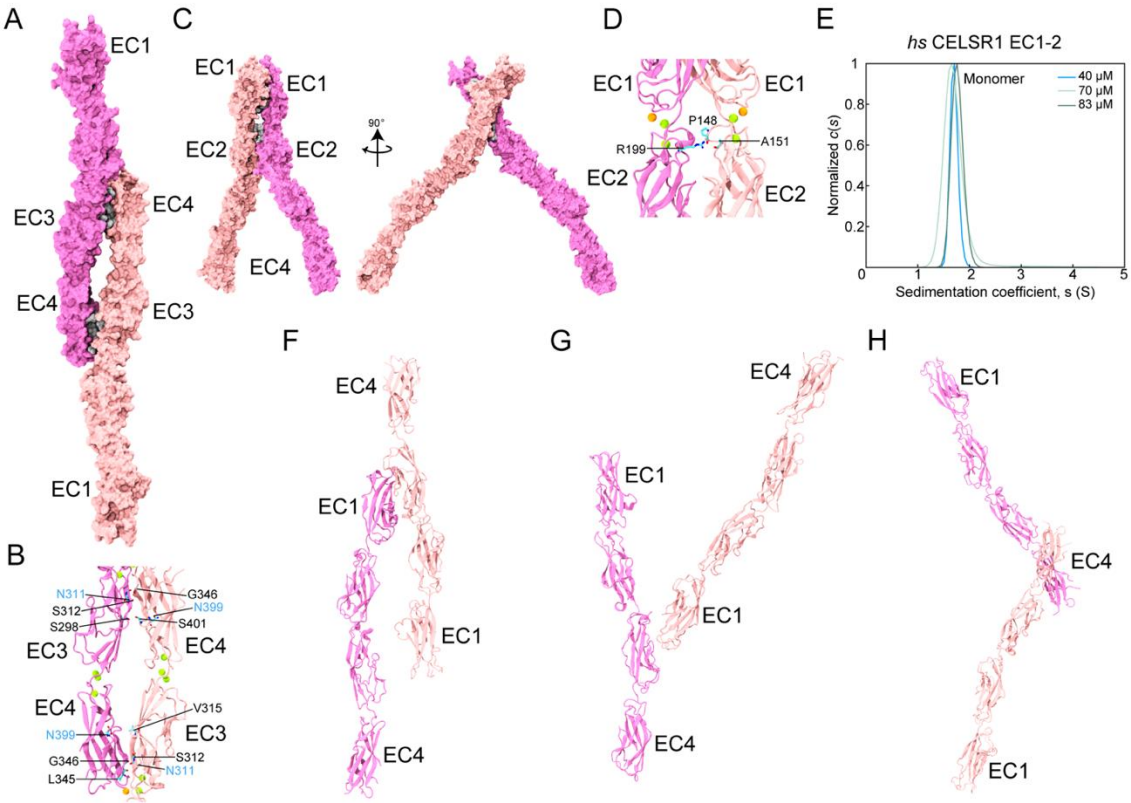


Figure 3.12. Crystallographic interfaces in the *hs* CELSR1 EC1-4 crystal structure

(*A*) Interface 1 with a BSA of 829.6 Å^2. (*B*) Residues found between EC3-4 and EC4-3 in the antiparallel interface shown in (*A*). Predicted glycosylation sites N311 and N399 are labeled in blue. Chain A is in mauve and chain B is in light pink. (*C*) Interface 2 with a BSA of 782.9 Å^2. (*D*) Residues found in the EC1-2 interface shown in (C). (*E*) AUC-SV results for *hs* CELSR1 EC1-2 at 40 μM, 70 μM, and 83 μM show a single peak based on SEDFIT analysis. (*F-H*) Other smaller crystallographic interfaces in the *hs* CELSR1 EC1-4 structure.

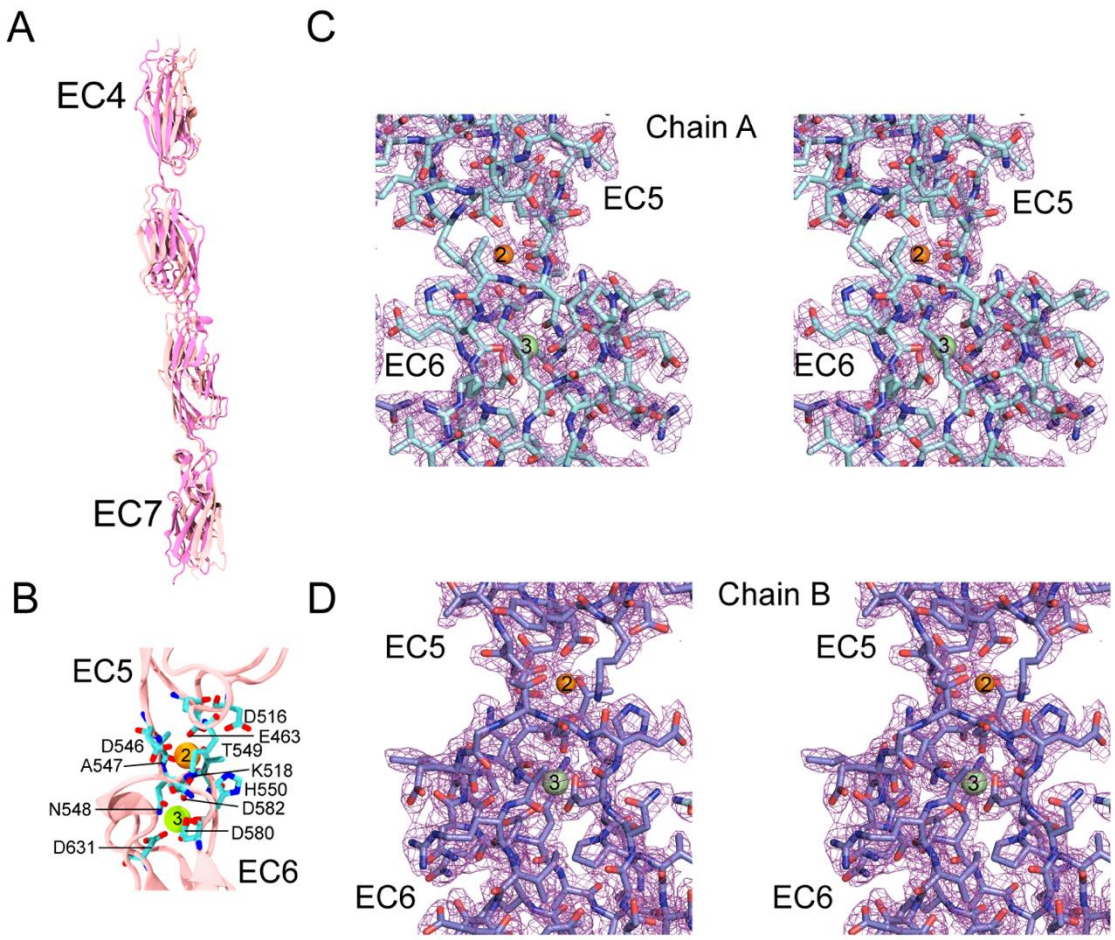


Figure 3.13. Structural features of the *hs* CELSR1 EC5-6 linker region

(*A*) Structural alignment of *hs* CELSR1 EC4-7 using the two monomers in the asymmetric unit. (*B*) The *hs* CELSR1 EC5-6 linker region in chain B binds a Ca^{2+} (green) at site 3 and a Na^{+} (orange) at site 2. (*C, D*) Stereo views of $2f_o$-f_c electron density maps at σ = 2.0 contour level showing the *hs* CELSR1 EC5-6 linker region in chains A and B.

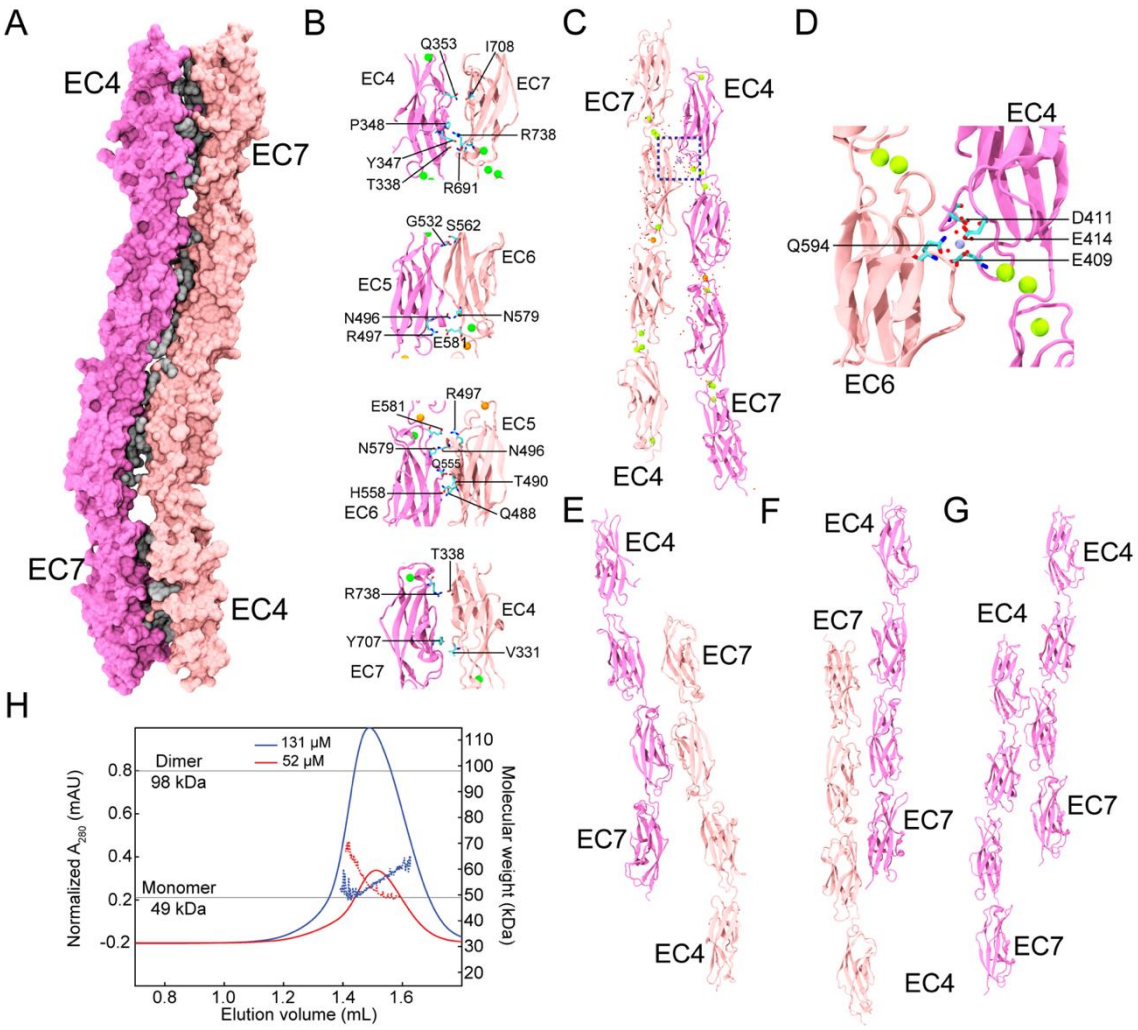


Figure 3.14. Crystallographic interfaces in the *hs* CELSR1 EC4-7 crystal structure

(*A*) The largest crystallographic interface with a BSA of 2,036 Å^2 in the crystal structure of *hs* CELSR1 EC4-7. (*B*) Residues found in the largest crystallographic interface shown in (*A*). (*C*) The second largest crystal interface with a BSA of 898 Å^2. A Mg^{2+} ion (purple) was found at the interface between EC6 of chain B and EC4 of chain A (box). (*D*) Mg^{2+} interface. Coordinating residues are in sticks and labeled. Water molecules are in red. (*E-G*) Other smaller crystallographic interfaces in the *hs* CELSR1 EC4-7 structure. (*H*) SEC-MALS of *hs* CELSR1 EC4-7 at 52 µM and 131 µM. Protein elutes as a monomer corresponding to the theoretical molecular weight of 49 kDa.

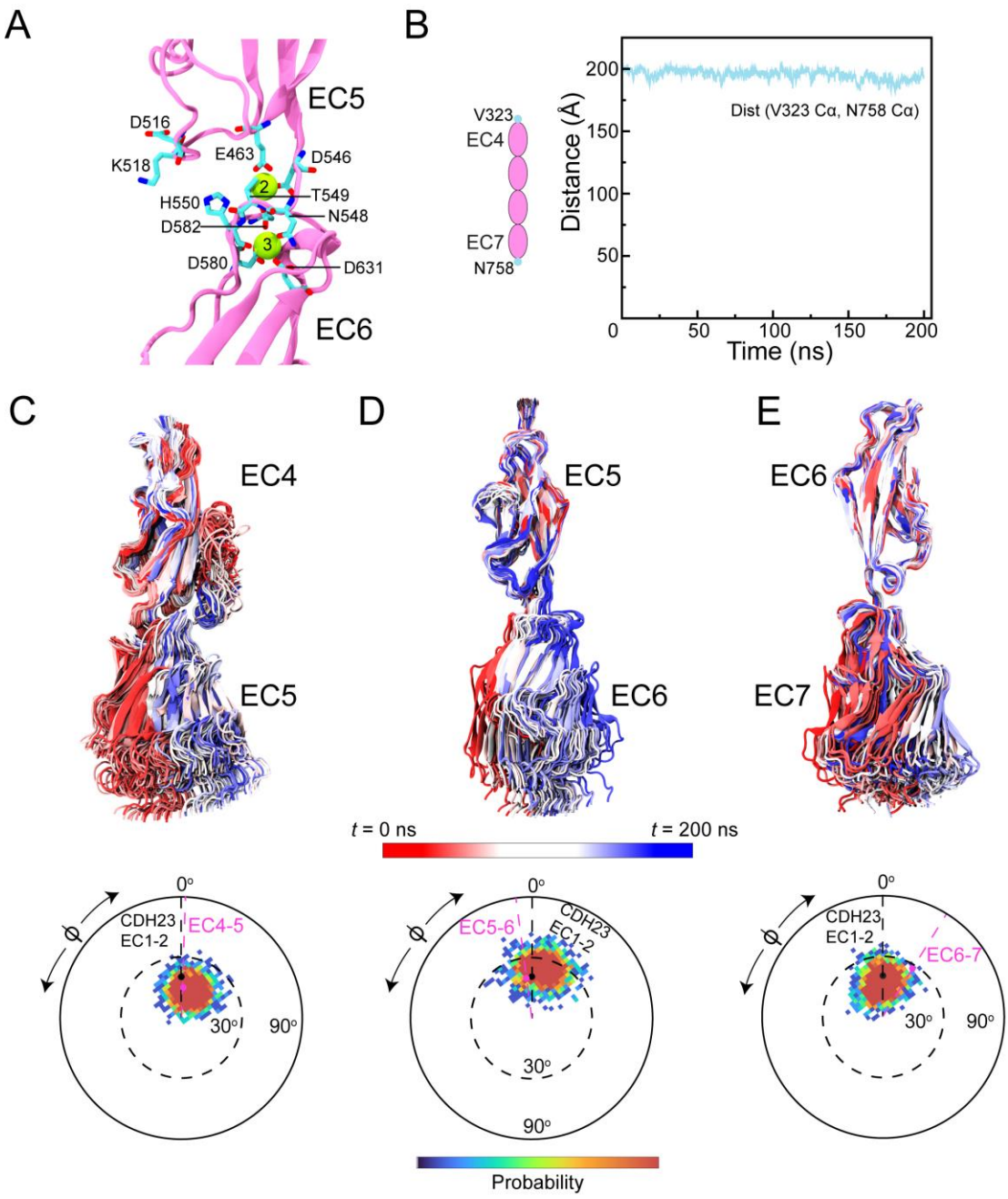


Figure 3.15. *In silico* orientation of EC repeats with Ca2+ at the EC5-6 linker region

(*A*) The EC5-6 linker region after 200 ns of equilibration (sim2). (*B*) Distance between C_α atoms of residues V323 in EC4 and N758 in EC7. (*C-E*) Top panels show the superposition of conformations every 5 ns for repeats *hs* CELSR1 EC4-5, EC5-6, and EC6-7 (sim2). Red represents the initial conformation at t = 0 ns and blue indicates the final conformation at t = 200 ns. Bottom panels show the projections of the longest principal axes of *mm* CDH23 EC2 and *hs* CELSR1 EC5, EC6, and EC7 on the *x-y* plane calculated every 100 ps and

plotted after alignment of EC4, EC5, and EC6 to *mm* CDH23 EC1 (PDB ID: 2WHV) oriented in the *z* direction. The initial projections of *mm* CDH23 EC2 (black), *hs* CELSR1 EC5, EC6, and EC7 (mauve) are indicated in circles. Color indicates probability.

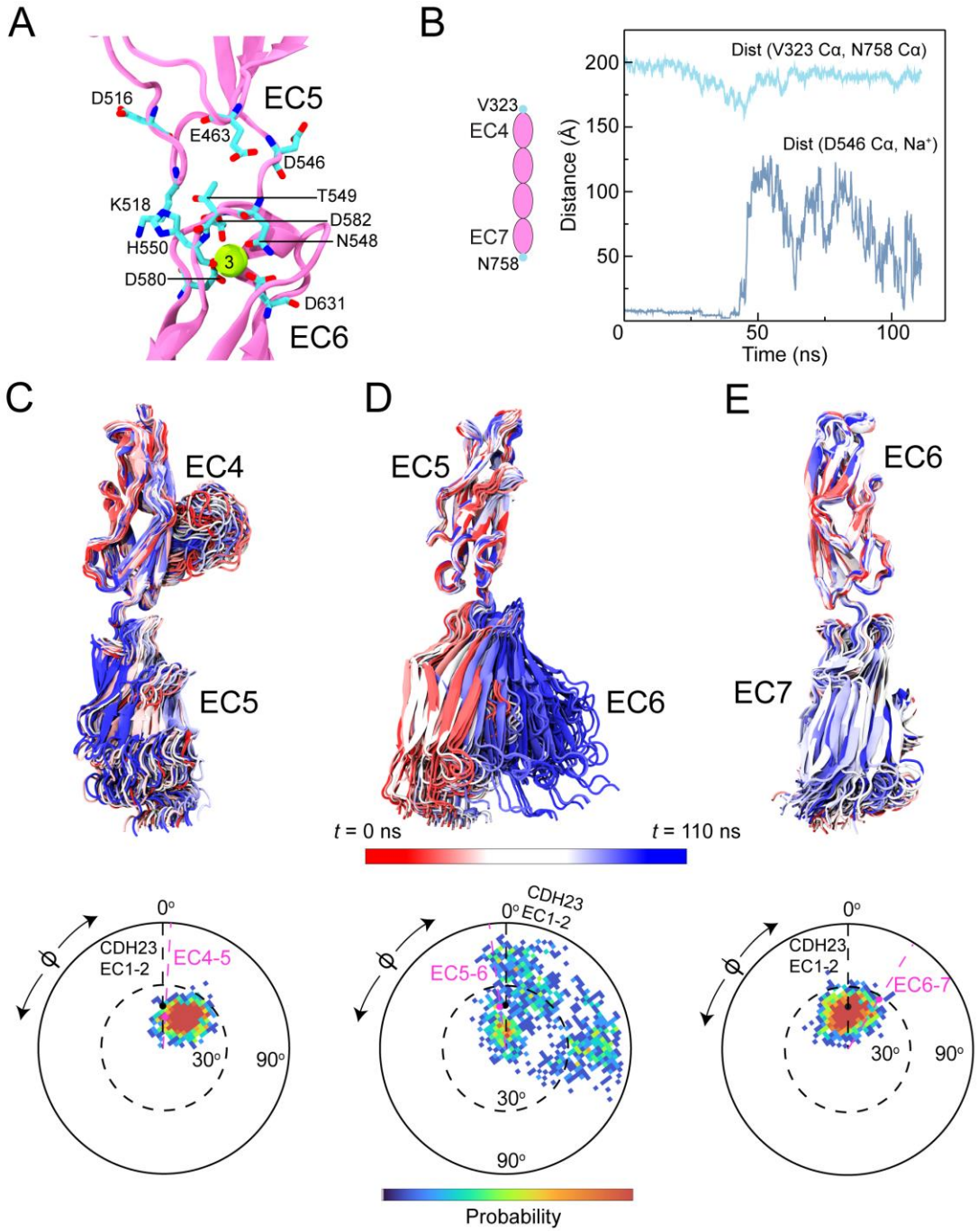


Figure 3.16. *In silico* orientation of EC repeats with Na+ at the EC5-6 linker region and EC4 constraints

(*A*) The EC5-6 linker region after 110 ns of equilibration (sim3). Na^+ is absent. (*B*) Distances between C_α atoms of residues V323 in EC4 and N758 in EC7 (cyan) and between C_α of D546 and Na^+ in the EC5-6 linker region (blue). (*C-E*) Top panels show the superposition of conformations every 5 ns for repeats *hs* CELSR1 EC4-5, EC5-6, and EC6-

7 (sim3). Red represents the initial conformation at t = 0 ns and blue indicates the final conformation at t = 110 ns. Bottom panels show the projections of the longest principal axes of *mm* CDH23 EC2 and *hs* CELSR1 EC5, EC6, and EC7 on the *x*-*y* plane calculated every 100 ps and plotted after alignment of EC4, EC5, and EC6 to *mm* CDH23 EC1 (PDB ID: 2WHV) oriented in the *z* direction. The initial projections of *mm* CDH23 EC2 (black), *hs* CELSR1 EC5, EC6, and EC7 (mauve) are indicated in circles. Color indicates probability.

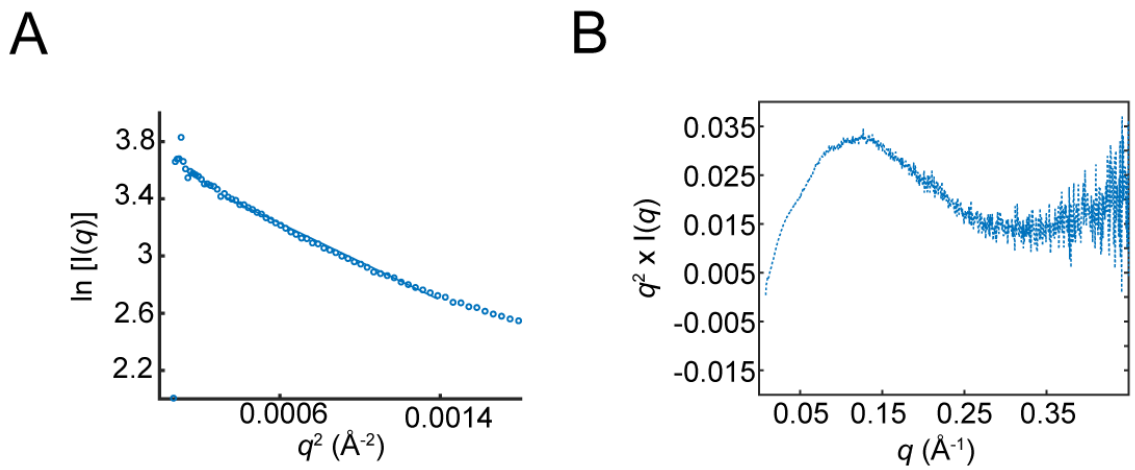


Figure 3.17. Guinier and Kratky analyses show that *hs* CELSR1 EC4-7 is folded in solution

(*A, B*) Guinier and Kratky plots from SAXS data on *hs* CELSR1 EC4-7. Kratky plot indicates a folded but flexible protein in solution.

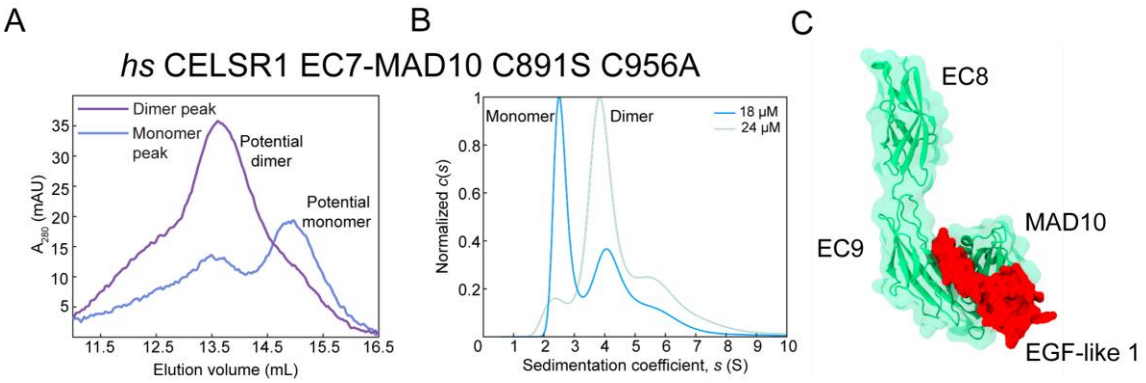


Figure 3.18. SEC and AUC of *hs* CELSR1 EC7-MAD10 suggests the presence of a monomer and a dimer in solution

(*A*) Resolved potential monomer and dimer peaks of *hs* CELSR1 EC7-MAD10 C891S C956A after repurification. (*B*) AUC-SV results for *hs* CELSR1 EC7-MAD10 C891S C956A at 18 μM and 24 μM (10 °C) show three peaks, a monomer, a potential dimer, and a higher-order oligomer peak based on SEDFIT analysis. (*C*) AF2-based full-length ectodomain model features an interaction between EC8-MAD10 (green) and the subsequent EGF-like 1 domain (red).

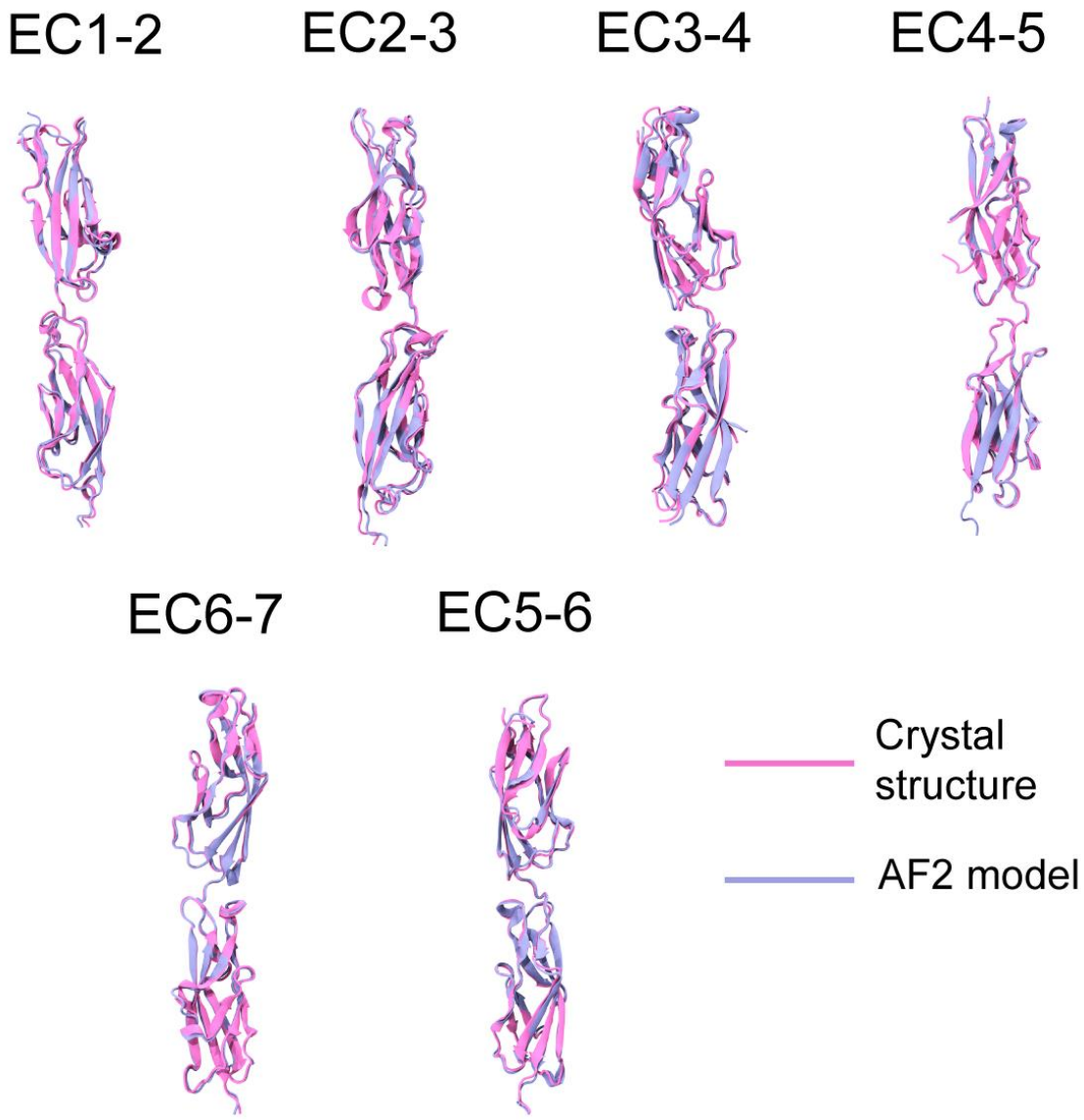


Figure 3.19. Alignment of consecutive EC repeats from crystal structures to AF2 models

Structural alignment of consecutive EC repeats to AF2 predicted models show good agreement for EC1-2, EC2-3, EC3-4, EC4-5, EC5-6, and EC6-7 (Table 3.4).

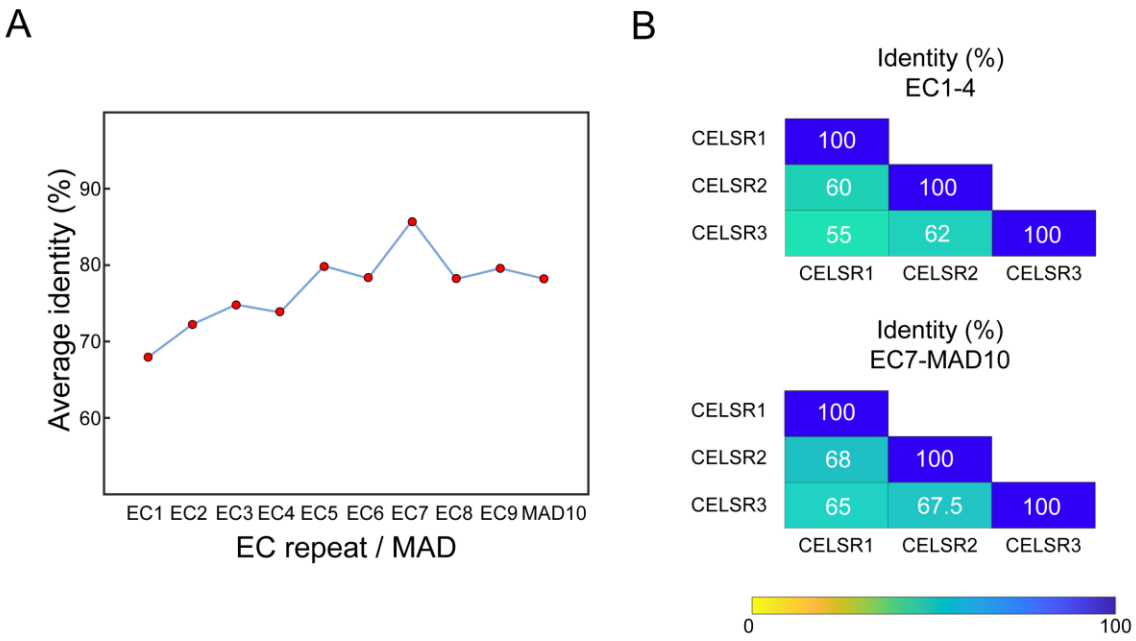
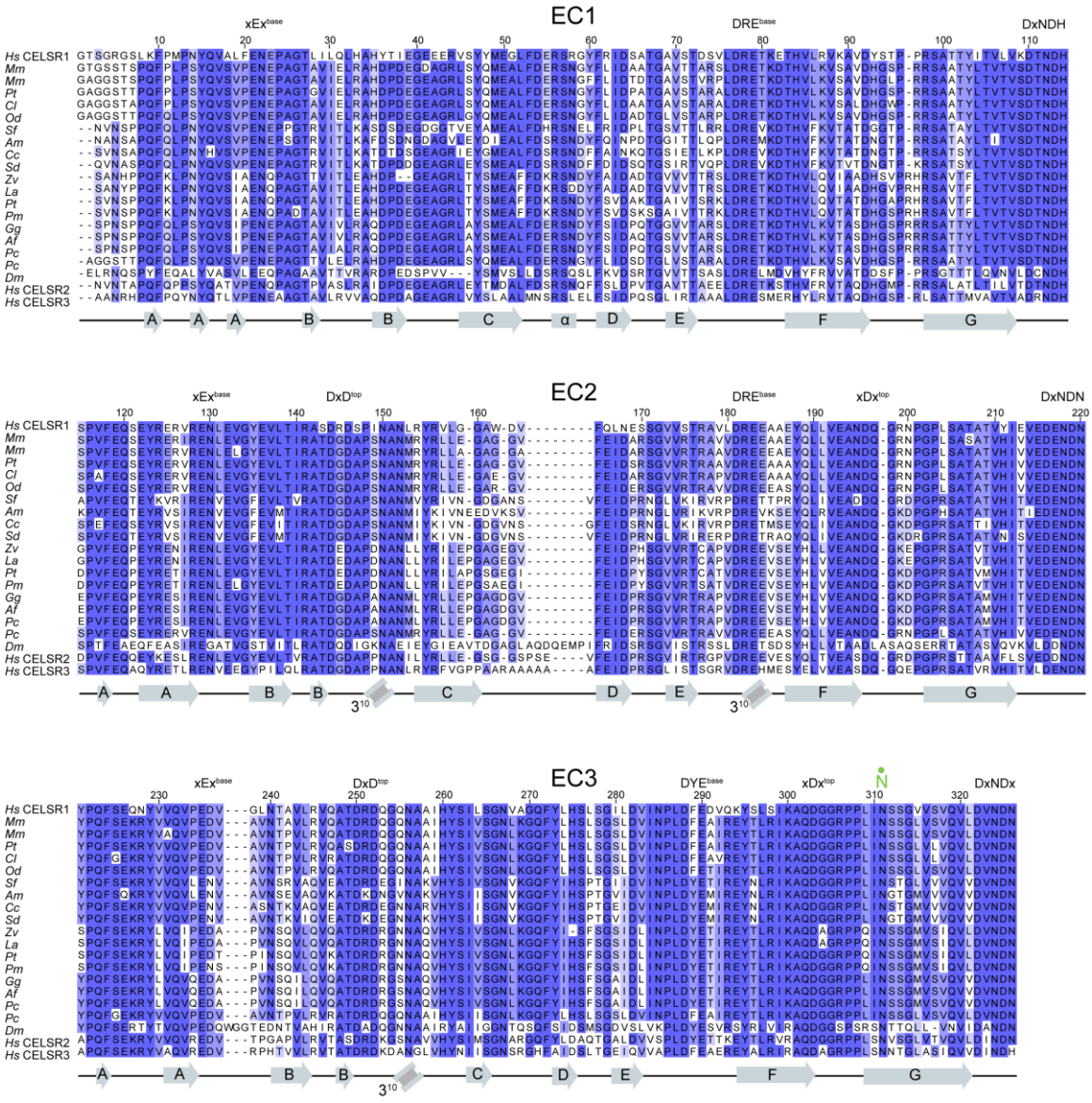
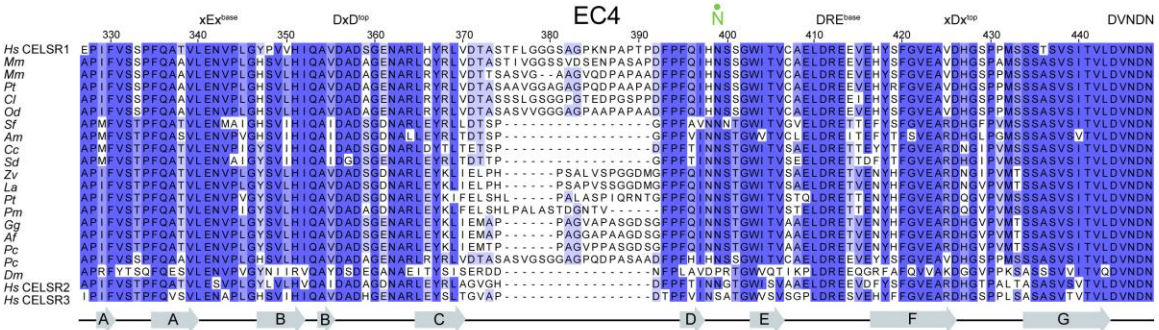
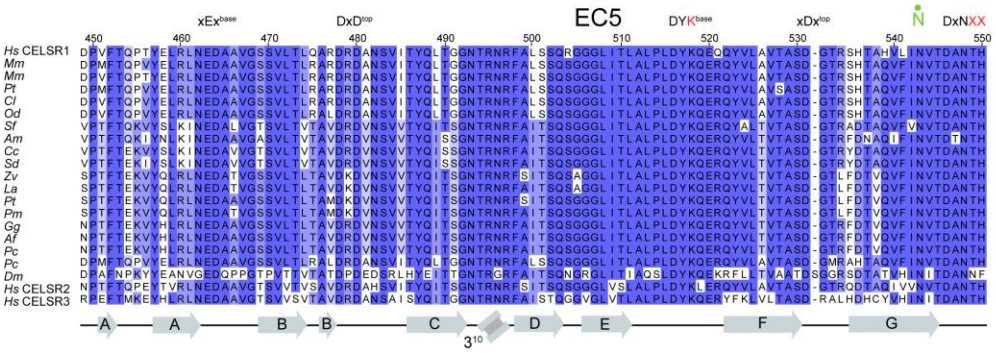


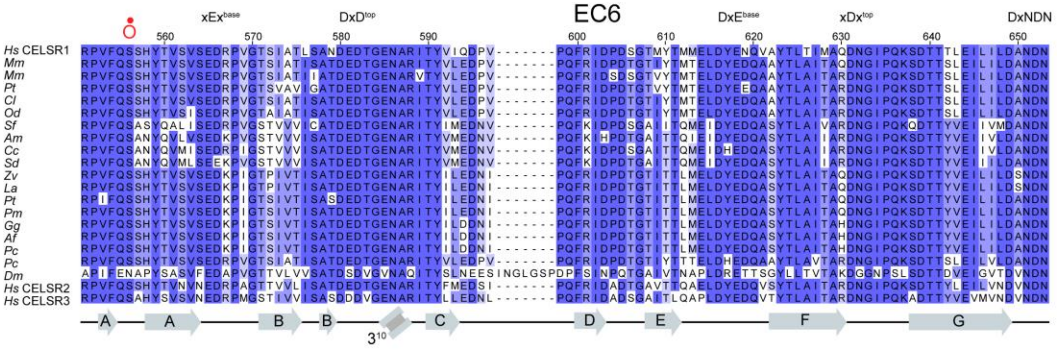
Figure 3.20. Sequence identity of EC repeats in CELSR1 across species and paralogs

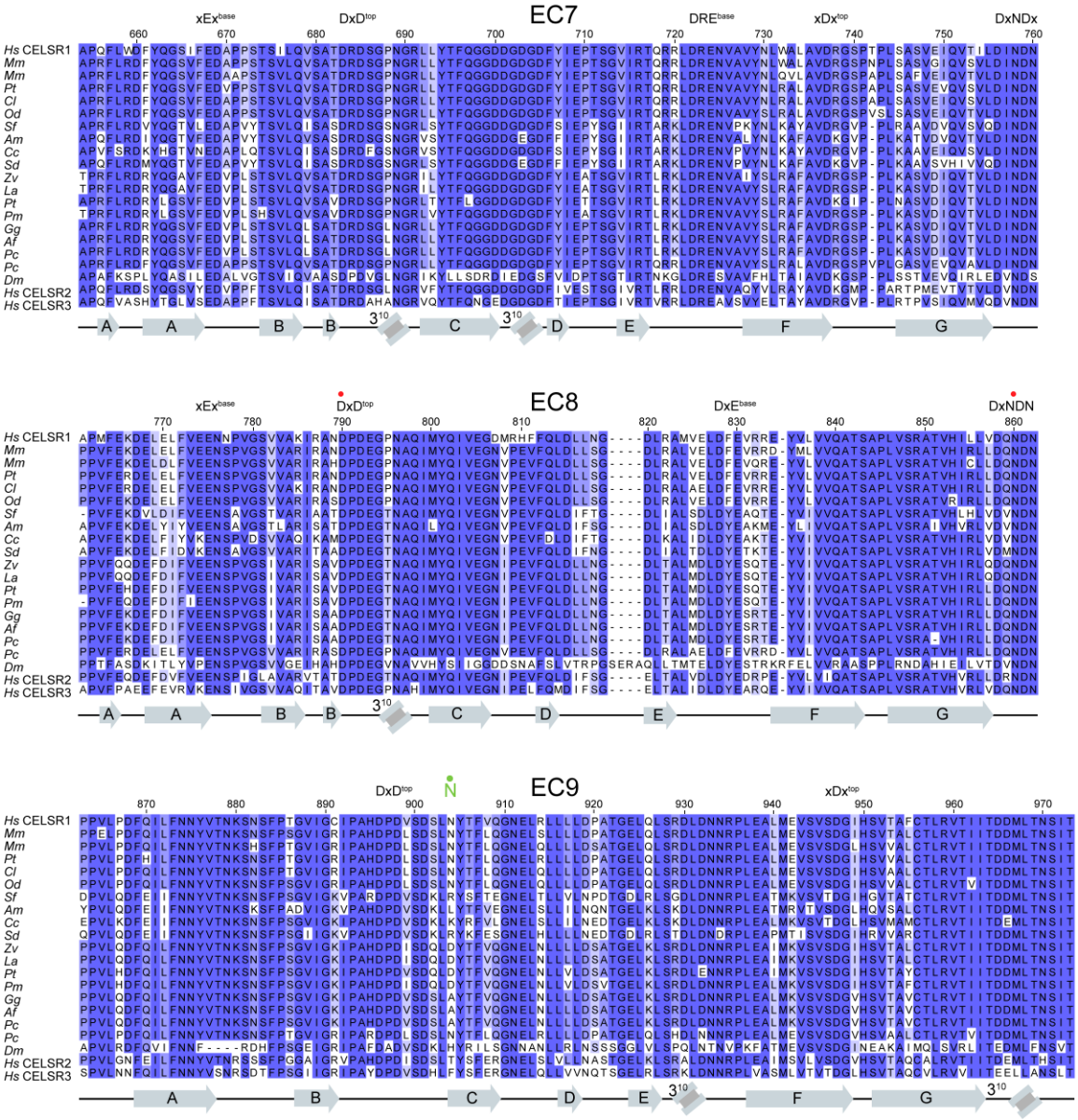
(*A*) Average protein sequence identity per EC repeat/MAD in CELSR1 across 19 species and its human paralogs, CELSR2 and CELSR3. (*B*) Protein sequence identity of human EC1-4 and EC7-MAD10 sequences across CELSR1, CELSR2 and CELSR3.


EC1
xEx^base
DRE^base
DxNDH
EC2
xEx^base
DxD^top
DRE^base
xDx^top
DxNDN
EC3
xEx^base
DxD^top
DYE^base
xDx^top
N
DxNDx


EC4
xEx^base
DxD^top
N
DRE^base
xDx^top
DVNDN
330
340
350
360
370
380
390
400
410
420
430
440
Hs CELSR1
Mm
Mm
Pt
Cl
Od
Sf
Am
Cc
Sd
Zv
La
Pt
Pm
Gg
Af
Pc
Pc
Dm
Hs CELSR2
Hs CELSR3
A
A
B
B
C
D
E
F
G


EC5
xEx^base
DxD^top
DYK^base
xDx^top
N
DxNXX
450
460
470
480
490
500
510
520
530
540
550
Hs CELSR1
Mm
Mm
Pt
Cl
Od
Sf
Am
Cc
Sd
Zv
La
Pt
Pm
Gg
Af
Pc
Pc
Dm
Hs CELSR2
Hs CELSR3
A
A
B
B
C
3^10
D
E
F
G


EC6
O
xEx^base
DxD^top
DxE^base
xDx^top
DxNDN
560
570
580
590
600
610
620
630
640
650
Hs CELSR1
Mm
Mm
Pt
Cl
Od
Sf
Am
Cc
Sd
Zv
La
Pt
Pm
Gg
Af
Pc
Pc
Dm
Hs CELSR2
Hs CELSR3
A
A
B
B
3^10
C
D
E
F
G


EC7
xEx^base
DxD^top
DRE^base
xDx^top
DxNDx
660
670
680
690
700
710
720
730
740
750
760
Hs CELSR1
Mm
Mm
Pt
Cl
Od
Sf
Am
Cc
Sd
Zv
La
Pt
Pm
Gg
Af
Pc
Pc
Dm
Hs CELSR2
Hs CELSR3
A
A
B
B
3^10
C
3^10
D
E
F
G
EC8
xEx^base
DxD^top
DxE^base
DxNDN
770
780
790
800
810
820
830
840
850
860
Hs CELSR1
Mm
Mm
Pt
Cl
Od
Sf
Am
Cc
Sd
Zv
La
Pt
Pm
Gg
Af
Pc
Pc
Dm
Hs CELSR2
Hs CELSR3
A
A
B
B
3^10
C
D
E
F
G
EC9
DxD^top
N
xDx^top
870
880
890
900
910
920
930
940
950
960
970
Hs CELSR1
Mm
Mm
Pt
Cl
Od
Sf
Am
Cc
Sd
Zv
La
Pt
Pm
Gg
Af
Pc
Pc
Dm
Hs CELSR2
Hs CELSR3
A
B
C
D
E
3^10
F
G
3^10

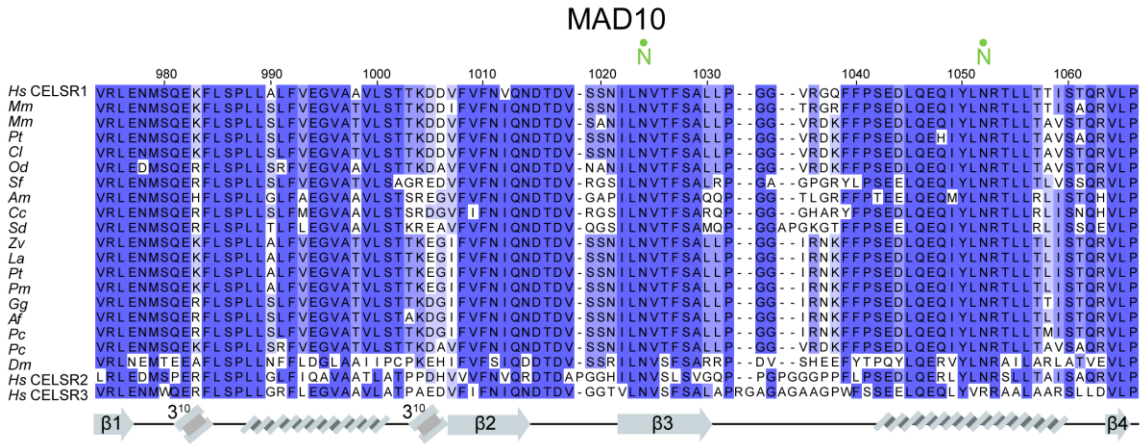


Figure 3.21. Sequence alignment of *hs* CELSR1 EC1-MAD10 across species

Sequence alignment of individual CELSR1 EC1-9 repeats and MAD10 across 19 different species. Paralogs *hs* CELSR2 and *hs* CELSR3 are also included (Table 3.5). The Crsh (D790G) and Spin cycle (N860K) mutations from mouse are highlighted with a red circle. N- and O-linked glycosylation sites are indicated with a circle and a N or O label.

Chapter 4. Conclusions

Cadherins are ubiquitously present in multicellular organisms engaging in a variety of cellular functions such as cell-cell adhesion, cell division, cell motility, mechanotransduction, and morphogenesis[4,6,7,20,168]. Among the cadherin family members, atypical cadherins with unusually large number of EC repeats in their ectodomain are often overlooked. The elongated architecture of their ectodomains and the repetitive nature of the EC repeats render these cadherins difficult to study using conventional structural methods such as X-ray crystallography, cryo-EM, or nuclear magnetic resonance. Hence, in our study, we used a combination of structural tools with a gamut of biophysical and biochemical techniques to recognize unique identifiers inherent to specific EC repeats along the length of long cadherin ectodomains. The versatility of *in vitro* and in-solution approaches combined with *in silico* studies have allowed us to describe cadherin-23 (CDH23), protocadherin-15 (PCDH15) and cadherin epithelial growth factor (EGF) Laminin-G (LAG or LamG) seven pass G-type receptor (CELSR1) ectodomains and predict some of their mechanical properties in detail.

CDH23 and PCDH15, two atypical cadherins with 27 and 11 EC repeats, are found in the inner ear and are responsible for mechanotransduction[19,20,25,64]. Mutations in either of the genes encoding for CDH23 or PCDH15 have been implicated in non-syndromic

deafness and Usher syndrome types 1D and 1F[169]. CELSR1, with nine EC repeats, also found in the inner ear, orchestrates the coordinated alignment of hair cells that results in oriented stereociliary bundles[50,52,55,170]. The coordinated alignment, termed planar cell polarity (PCP), is manifested during organ formation. Defects in PCP due to homozygous mutations in either of the core PCP proteins including CELSR1 have been implicated in severe neural tube defects that lead to embryo abortion[55,143]. Heterozygous mutations, on the other hand, cause local, organ specific defects. In the inner ear, they lead to misoriented hair bundles in the vestibular and cochlear sensory tissues reducing hearing efficiency and causing abnormal head shaking behavior[50].

The mechanism of *trans* heterophilic adhesion by CDH23-PCDH15 has been well-studied in the past[36]. However, the molecular structure of the entire tip link has been missing. On the other hand, for CELSR1, molecule-level details on their homophilic interactions are sparse with cell-based aggregation assays only establishing their homophilic nature[56].

In chapter 2, we present full-length cryogenic electron microscopy (cryo-EM) and high-speed atomic force microscopy (HS-AFM) images of CDH23 and PCDH15 along with a structural model of the entire tip link complex obtained *in silico*. Our *in silico* model serves as an approximate representation of some of the conformations observed *in vitro* as evidenced by the similar nature to the stained scanning-electron micrographs (SEM), freeze-fracture transmission electron micrograph (TEM) and cryo-EM images[39,61,72] .

We were, however, unable to acquire high-resolution cryo-EM data of CDH23. A bottle neck in obtaining high-resolution structures of full-length ectodomains of CDH23 and PCDH15 is their elongated, filamentous nature and size. In the cryo-EM images that we have obtained for PCDH15, we observed a uniform dispersion of particles on the grid which was not the case for CDH23. PCDH15 was also more stable owing to *cis* dimer formation. One way to overcome CDH23 instability would be to co-purify CDH23 and PCDH15 together. The stable *trans* dimer formation between EC1-2 of CDH23 and PCDH15 would reduce the entropic states of CDH23 in solution. But the C-terminal end of CDH23 would still be a problem which can be remedied by expression of the extracellular domain with its transmembrane domain and embedding the protein in a liposome. Trapping CDH23 molecules to a liposome would reduce the diffusion rate to only two dimensions along the membrane plane, thus locking any *cis* transient interactions between adjacent CDH23 molecules. The K_D of CDH23 *trans* dimer is higher than the K_D of the CDH23-PCDH15 *trans* dimer[43]. Optimistically, this would lead to higher populations of the tetramer than the CDH23 *trans* dimer complex between liposomes, which is a possibility. To further stabilize the conformations of CDH23, the molecules could be stretched. Tightly regulating the distance between CDH23 and PCDH15 can be exercised by embedding both CDH23 and PCDH15 transmembrane domains with their extracellular domains into liposomes. Through regulation of Ca^{2+} concentration, multiple conformational states between the complexes could be introduced which can be used to study their behavior in physiological conditions during transduction.

Using our full-length *in silico* model of the CDH23 and PCDH15 ectodomains can provide the best avenue to understand several properties of the tip link such as elasticity, Ca^{2+} dependence, and the effect of deafness mutations. Planned and ongoing molecular dynamics (MD) and steered molecular dynamics (SMD) simulations of the tip-link complex are starting to elucidate the elasticity of the tip link in response to force. Future studies focused on utilizing our model with alternative Ca^{2+} occupancies along the length of the tip link will help identify elastic regions that might be relevant at physiological Ca^{2+} concentrations. In-depth *in silico* studies have been performed individually on various components of the transduction apparatus but studies with the entirety of tip link and the transduction complex through coarse grained simulations could elucidate mechanical properties associated with channel opening due to tip-link tension.

In chapter 3, we addressed the adhesive function of CELSR1 restricted to the nine EC repeats and its membrane adjacent domain-10 (MAD10). We used bead aggregation assays to initially estimate and identify the N-terminal EC repeats needed for inducing adhesion. However, to our surprise, bead aggregation even in the presence of additional co-factors such as Mg^{2+} in addition to Ca^{2+} showed that the adhesion promoted by CELSR1 EC repeats is weak compared to other well-studied cadherins such as PCDH1[171], PCDH24[138], and PCDH19[114]. Recently published studies on CELSR2[63] have shown that weak adhesion might be prominent even among paralogs. Compared to cell-based aggregation assays with the complete Fmi (drosophila homolog of CELSR1) and *Mus musculus* (*mm*) CELSR1 ectodomain where robust aggregation is observed[56,58], it is

important to point out the presence of a co-factor that might be getting lost in bead aggregation assays during the washing steps.

Our structure of *hs* CELSR1 EC1-4 did not reveal a *trans* dimer but complementary in-solution studies show that this fragment might exist as a dimer only at high concentrations (> 100 μM). A lack of double peaks in analytical ultracentrifugation and multi-angle light scattering show that this fragment might exhibit a rapid monomer-dimer equilibrium exchange. On the other hand, *hs* CELSR1 EC1-5 exhibits two peaks with the lower elution volume assumed to be that of a potential dimer. Combined with the bead aggregate sizes monitored for EC1-4 and EC1-5 fragments, it is becoming clear that the EC5 repeat more readily stabilizes potential *trans* interaction promoted by EC1-4 fragment. To discern the individual capacity of these EC repeats, it would be interesting to perform cell-based aggregation assays with just the EC1-5 fragment in the ectodomain. A similar adhesion capacity comparable to bead aggregation assays would be helpful in the verification of the presence of a co-factor(s), if needed for the *trans* dimer formation.

MAD10 in CELSR1 is another interesting structural subdomain that is briefly discussed in my dissertation. Like PCDH15 MAD12[37,39,40], the CELSR1 MAD10 also seems to induce *cis* dimerization in solution but in a more robust manner compared to the EC1-4 fragment. This is noteworthy because Crsh and spin cycle mutations[55] that disrupt Ca^{2+} binding in the EC7-8 and EC8-9 linker regions show less propensity of the CELSR1 molecules to be established in a stable punctate in cell-cell junctions[50]. While we could not

clarify whether potential MAD10-induced *cis* dimerization is relevant for function *in vivo*, it can be postulated that MAD10 might stabilize the CELSR1 molecules in the punctate. MAD10-induced *cis* dimerization either could reinforce the weak *trans* interactions between EC1-4 or be involved in another *cis* interaction downstream of MAD10. In the bead aggregation assays that we performed, we could not unfortunately evaluate the function of MAD10 and its effect in promoting *trans* EC1-4 interactions. When EC1-MAD10 fragment was used, we were unable to determine if *cis* interactions were formed between adjacent EC7-MAD10 on the same bead surface along with the EC1-4 *trans* interactions formed between molecules on different beads. In the HS-AFM images of CELSR2[63], a lack of observation of both *cis* and *trans* dimers between adjacent molecules weakens our claim but similar to bead aggregation assays, these were performed with only the ectodomain and not in the context of a cell. Future assays specifically designed to evaluate both interactions in the context of the cell through fluorescence based *in vitro* techniques could be helpful but would suffer from low resolution. Using liposomes to express the full-length CELSR1 is tempting, but if a cellular co-factor is required for stable *cis* or *trans* interactions, we would still observe a weak binding between molecules similar to what was observed in bead aggregation assays. On the other hand, if the *cis* dimer formation is only limited by diffusion between adjacent CELSR1 molecules, then observation of both *cis* and *trans* dimers might be possible. Using cryo-electron tomography (cryo-ET) would be an excellent choice to observe these different conformational states in the context of the cell, especially in the junctional lattice that is proposed to form between cells during the later stages of PCP.

In conclusion, the work described in this dissertation offers data to study elongated atypical cadherins for which structures are not available and not easily obtained. Our combined use of biochemical, biophysical, and computational techniques shows that beyond cell-cell adhesion, variable mechanical properties of these cadherins are altered through substitution of the residues in the Ca^{2+} linker regions. A good agreement between computational, structural, and biochemical techniques in our studies underscores the importance and advocates for its usage in the field of cadherin biology and the study of their mechanics.

Bibliography

1. Hulpiau P, van Roy F. Molecular evolution of the cadherin superfamily. *Int J Biochem Cell Biol*. 2009;41(2):349-369. doi:10.1016/j.biocel.2008.09.027
2. Hulpiau P, Van Roy F. New insights into the evolution of metazoan cadherins. *Mol Biol Evol*. 2011;28(1):647-657. doi:10.1093/molbev/msq233
3. Gul IS, Hulpiau P, Saeys Y, van Roy F. Evolution and diversity of cadherins and catenins. *Exp Cell Res*. 2017;358(1):3-9. doi:10.1016/j.yexcr.2017.03.001
4. Halbleib JM, Nelson WJ. Cadherins in development: Cell adhesion, sorting, and tissue morphogenesis. *Genes Dev*. 2006;20(23):3199-3214. doi:10.1101/gad.1486806
5. Gallin WJ. Evolution of the "classical" cadherin family of cell adhesion molecules in vertebrates. *Mol Biol Evol*. 1998;15(9):1099-1107. doi:10.1093/oxfordjournals.molbev.a026017
6. Punovuori K, Malaguti M, Lowell S. Cadherins in early neural development. *Cell Mol Life Sci*. 2021;78(9):4435-4450. doi:10.1007/s00018-021-03815-9
7. Hale R, Strutt D. Conservation of Planar Polarity Pathway Function Across the Animal Kingdom. *Annu Rev Genet*. 2015;49:529-551. doi:10.1146/annurev-genet-112414-055224
8. Brooun M, Klimovich A, Bashkurov M, Pearson BJ, Steele RE, McNeill H. Ancestral roles of atypical cadherins in planar cell polarity. *Proc Natl Acad Sci U S A*. 2020;117(32):19310-19320. doi:10.1073/pnas.1917570117
9. Ekdale EG. Form and function of the mammalian inner ear. *J Anat*. 2016;228(2):324-337. doi:10.1111/joa.12308
10. Lawrence M. Structure and function of the ear and auditory nervous system. *Environ Health Perspect*. 1982;Vol. 44(April):9-13. doi:10.1289/ehp.82449
11. Patuzzi R, Robertson D. Tuning in the mammalian cochlea. *Physiol Rev*. 1988;68(4):1009-1082. doi:10.1152/physrev.1988.68.4.1009
12. Nin F, Hibino H, Doi K, Suzuki T, Hisa Y, Kurachi Y. The endocochlear potential depends on two K+ diffusion potentials and an electrical barrier in the stria vascularis of the inner ear. *Proc Natl Acad Sci U S A*. 2008;105(5):1751-1756. doi:10.1073/pnas.0711463105
13. Liu W, Atturo F, Aldaya R, et al. Macromolecular organization and fine structure of the human basilar membrane - RELEVANCE for cochlear implantation. *Cell Tissue Res*. 2015;360(2):245-262. doi:10.1007/s00441-014-2098-z
14. Lim DJ. Functional structure of the organ of Corti: a review. *Hear Res*. 1986;22(1-3):117-146. doi:10.1016/0378-5955(86)90089-4

15. Hudspeth AJ. Integrating the active process of hair cells with cochlear function. *Nat Rev Neurosci*. 2014;15(9):600-614. doi:10.1038/nrn3786
16. McGrath J, Roy P, Perrin BJ. Stereocilia morphogenesis and maintenance through regulation of actin stability. *Semin Cell Dev Biol*. 2017;65:88-95. doi:10.1016/j.semcdb.2016.08.017
17. Reichenbach T, Hudspeth AJ. The physics of hearing: Fluid mechanics and the active process of the inner ear. *Reports Prog Phys*. 2014;77(7). doi:10.1088/0034-4885/77/7/076601
18. Pickles JO. A model for the mechanics of the stereociliar bundle on acousticolateral hair cells. *Hear Res*. 1993;68(2):159-172. doi:10.1016/0378-5955(93)90120-P
19. Kazmierczak P, Sakaguchi H, Tokita J, et al. Cadherin 23 and protocadherin 15 interact to form tip-link filaments in sensory hair cells. *Nature*. 2007;449(7158):87-91. doi:10.1038/nature06091
20. Müller U. Cadherins and mechanotransduction by hair cells. *Curr Opin Cell Biol*. 2008;20(5):557-566. doi:10.1016/j.ceb.2008.06.004
21. Ahmed ZM, Goodyear R, Riazuddin S, et al. The tip-link antigen, a protein associated with the transduction complex of sensory hair cells, is protocadherin-15. *J Neurosci*. 2006;26(26):7022-7034. doi:10.1523/JNEUROSCI.1163-06.2006
22. Siemens J, Lillo C, Dumont RA, et al. Cadherin 23 Is a component of the tip link in hair-cell stereocilla. *Nature*. 2004;428(6986):950-955. doi:10.1038/nature02483
23. Söllner C, Nicolson T, Rauch GJ, et al. Mutations in cadherin 23 affect tip links in zebrafish sensory hair cells. *Nature*. 2004;428(6986):955-959. doi:10.1038/nature02484
24. Holt JR, Tobin M, Elferich J, et al. Putting the Pieces Together: the Hair Cell Transduction Complex. *JARO - J Assoc Res Otolaryngol*. 2021;22(6):601-608. doi:10.1007/s10162-021-00808-0
25. Kachar B, Parakkal M, Kurc M, Zhao YD, Gillespie PG. High-resolution structure of hair-cell tip links. *Proc Natl Acad Sci U S A*. 2000;97(24):13336-13341. doi:10.1073/pnas.97.24.13336
26. Stöver T, Diensthuber M. Molecular biology of hearing. *Laryngorhinootologie*. 2011;90(SUPPL. 1). doi:10.1055/s-0030-1270444
27. Brownell WE. How the ear works: Nature's solutions for listening. *Volta Rev*. 1999;99(5):9-28. /pmc/articles/PMC2888317/. Accessed September 12, 2023.
28. National Library of Medicine. Basics of Sound, the Ear, and Hearing. *Hear Loss Determ Eligibility Soc Secur Benefits*. 2005;(DC):1-19. https://www.ncbi.nlm.nih.gov/books/NBK207834/. Accessed September 12, 2023.
29. Jaiganesh A, De-la-Torre P, Patel AA, et al. Zooming in on Cadherin-23: Structural Diversity and Potential Mechanisms of Inherited Deafness. *Structure*. 2018;26(9):1210-1225.e4. doi:10.1016/j.str.2018.06.003
30. Nagar B, Overduin M, Ikura M, Rini JM. Structural basis of calcium-induced E-cadherin rigidification and dimerization. *Nature*. 1996;380(6572):360-364. doi:10.1038/380360a0

31. Shapiro L, Weis WI. Structure and biochemistry of cadherins and catenins. *Cold Spring Harb Perspect Biol*. 2009;1(3). doi:10.1101/cshperspect.a003053
32. Jin X, Walker MA, Felsövályi K, et al. Crystal structures of Drosophila N-cadherin ectodomain regions reveal a widely used class of Ca 2+-free interdomain linkers. *Proc Natl Acad Sci U S A*. 2012;109(3):E127-E134. doi:10.1073/PNAS.1117538108/SUPPL_FILE/PNAS.1117538108_SI.PDF
33. Araya-Secchi R, Neel BL, Sotomayor M. An elastic element in the protocadherin-15 tip link of the inner ear. *Nat Commun*. 2016;7:13458. doi:10.1038/ncomms13458
34. Takeichi M, Hatta K, Nose A, Nagafuchi A. Identification of a gene family of cadherin cell adhesion molecules. *Cell Differ Dev*. 1988;25(SUPPL.):91-94. doi:10.1016/0922-3371(88)90104-9
35. Pokutta S, Herrenknecht K, Kemler R, Engel J. Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding. *Eur J Biochem*. 1994;223(3):1019-1026. doi:10.1111/j.1432-1033.1994.tb19080.x
36. Sotomayor M, Weihofen WA, Weihofen WA, Corey DP. Structure of a force-conveying cadherin bond essential for inner-ear mechanotransduction. *Nature*. 2012. doi:10.1038/nature11590
37. Choudhary D, Narui Y, Neel BL, et al. Structural determinants of protocadherin-15 mechanics and function in hearing and balance perception. *Proc Natl Acad Sci U S A*. 2020;117(40):4837-24848. doi:10.1073/pnas.1920444117
38. Dionne G, Qiu X, Rapp M, et al. Mechanotransduction by PCDH15 Relies on a Novel cis-Dimeric Architecture. *Neuron*. 2018;99(3):480-492.e5. doi:10.1016/j.neuron.2018.07.006
39. Ge J, Elferich J, Goehring A, Zhao H, Schuck P, Gouaux E. Structure of mouse protocadherin 15 of the stereocilia tip link in complex with LHFPL5. *Elife*. 2018;7. doi:10.7554/eLife.38770
40. De-la-Torre P, Choudhary D, Araya-Secchi R, Narui Y, Sotomayor M. A Mechanically Weak Extracellular Membrane-Adjacent Domain Induces Dimerization of Protocadherin-15. *Biophys J*. 2018;115(12):2368-2385. doi:10.1016/j.bpj.2018.11.010
41. Ge J, Elferich J, Goehring A, Zhao H, Schuck P, Gouaux E. Structure of mouse protocadherin 15 of the stereocilia tip link in complex with LHFPL5. *Elife*. 2018;7:1-24. doi:10.7554/eLife.38770
42. Kaur V, Ghosh SK, Bhatia T, Rakshit S. Redefining the Structure of Tip Links in Hair Cells. *Biochemistry*. 2023;62(15):2244-2251. doi:10.1021/ACS.BIOCHEM.3C00161/SUPPL_FILE/BI3C00161_SI_001.PDF
43. Srinivas CS, Singaraju GS, Kaur V, et al. Transient interactions drive the lateral clustering of cadherin-23 on membrane. *Commun Biol 2023 61*. 2023;6(1):1-11. doi:10.1038/s42003-023-04677-6
44. Tobin M, Chaiyasitdhi A, Michel V, Michalski N, Martin P. Stiffness and tension gradients of the hair cell's tip-link complex in the mammalian cochlea. *Elife*. 2019;8. doi:10.7554/eLife.43473

45. Howard J, Hudspeth AJ. Compliance of the hair bundle associated with gating of mechanoelectrical transduction channels in the Bullfrog's saccular hair cell. *Neuron*. 1988;1(3):189-199. doi:10.1016/0896-6273(88)90139-0
46. Sai X, Ladher RK. Early steps in inner ear development: Induction and morphogenesis of the otic placode. *Front Pharmacol*. 2015;6(FEB):1-8. doi:10.3389/fphar.2015.00019
47. Wu DK, Kelley MW. Molecular mechanisms of inner ear development. *Cold Spring Harb Perspect Biol*. 2012;4(8). doi:10.1101/cshperspect.a008409
48. Rida PCG, Chen P. Line up and listen: Planar cell polarity regulation in the mammalian inner ear. *Semin Cell Dev Biol*. 2009;20(8):978-985. doi:10.1016/j.semcdb.2009.02.007
49. Duncan JS, Stoller ML, Francl AF, Tissir F, Devenport D, Deans MR. Celsr1 coordinates the planar polarity of vestibular hair cells during inner ear development. *Dev Biol*. 2017;423(2):126-137. doi:10.1016/j.ydbio.2017.01.020
50. Curtin JA, Quint E, Tsipouri V, et al. Mutation of Celsr1 disrupts planar polarity of inner ear hair cells and causes severe neural tube defects in the mouse. *Curr Biol*. 2003;13(13):1129-1133. doi:10.1016/S0960-9822(03)00374-9
51. Kelly M, Chen P. Shaping the mammalian auditory sensory organ by the planar cell polarity pathway. *Int J Dev Biol*. 2007;51(6-7):535-547. doi:10.1387/ijdb.072344mk
52. May-Simera H, Kelley MW. *Planar Cell Polarity in the Inner Ear*. Vol 101. 1st ed. Elsevier Inc.; 2012. doi:10.1016/B978-0-12-394592-1.00006-5
53. Devenport D. The cell biology of planar cell polarity. *J Cell Biol*. 2014;207(2):171-179. doi:10.1083/JCB.201408039
54. Simons M, Mlodzik M. Planar Cell Polarity Signaling: From Fly Development to Human Disease. *Annu Rev Genet*. 2008. doi:10.1146/annurev.genet.42.110807.091432
55. Wang XJ, Zhang DL, Xu ZG, et al. Understanding cadherin EGF LAG seven-pass G-type receptors. *J Neurochem*. 2015. doi:10.1111/jnc.12955
56. Usui T, Shima Y, Shimada Y, et al. Flamingo, a seven-pass transmembrane cadherin, regulates planar cell polarity under the control of Frizzled. *Cell*. 1999. doi:10.1016/S0092-8674(00)80046-X
57. Shimada Y, Usui T, Yanagawa S ichi, Takeichi M, Uemura T. Asymmetric colocalization of Flamingo, a seven-pass transmembrane cadherin, and Dishevelled in planar cell polarization. *Curr Biol*. 2001;11(11):859-863. doi:10.1016/S0960-9822(01)00233-0
58. Stahley SN, Basta LP, Sharan R, Devenport D. Celsr1 adhesive interactions mediate the asymmetric organization of planar polarity complexes. *Elife*. 2021;10:1-26. doi:10.7554/eLife.62097
59. Strutt H, Strutt D. Differential Stability of Flamingo Protein Complexes Underlies the Establishment of Planar Polarity. *Curr Biol*. 2008. doi:10.1016/j.cub.2008.08.063
60. Strutt DI. Asymmetric localization of frizzled and the establishment of cell polarity

in the Drosophila wing. *Mol Cell*. 2001;7(2):367-375. doi:10.1016/S1097-2765(01)00184-8
61. Basta LP, Sil P, Jones RA, Little KA, Hayward-Lara G, Devenport D. Celsr1 and Celsr2 exhibit distinct adhesive interactions and contributions to planar cell polarity. *Front Cell Dev Biol*. 2023;10:2398. doi:10.3389/fcell.2022.1064907
62. AlQuraishi M. Protein-structure prediction revolutionized. *Nat 2021 5967873*. 2021;596(7873):487-488. doi:10.1038/d41586-021-02265-4
63. Nishiguchi S, Kasai RS, Uchihashi T. Antiparallel dimer structure of CELSR cadherin in solution revealed by high-speed atomic force microscopy. *Proc Natl Acad Sci U S A*. 2023;120(118):e2302047120. doi:10.1073/pnas.2302047120
64. Pickles JO, Comis SD, Osborne MP. Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction. *Hear Res*. 1984;15(2):103-112. doi:10.1016/0378-5955(84)90041-8
65. Elferich J, Clark S, Ge J, Goehring A, Matsui A, Gouaux E. Molecular structures and conformations of protocadherin-15 and its complexes on stereocilia elucidated by cryo-electron tomograph. *Elife*. 2021;10:1-20. doi:10.7554/eLife.74512
66. Zheng W, Holt JR. The Mechanosensory Transduction Machinery in Inner Ear Hair Cells. *Annu Rev Biophys*. 2021;50:31-51. doi:10.1146/annurev-biophys-062420-081842
67. Assad JA, Shepherd GMG, Corey DP. Tip-link integrity and mechanical transduction in vertebrate hair cells. *Neuron*. 1991;7(6):985-994. doi:10.1016/0896-6273(91)90343-X
68. Hudspeth AJ, Gillespie PG. Pulling springs to tune transduction: Adaptation by hair cells. *Neuron*. 1994;12(1):1-9. doi:10.1016/0896-6273(94)90147-3
69. Sotomayor M, Weihofen WA, Gaudet R, Corey DP. Structural Determinants of Cadherin-23 Function in Hearing and Deafness. *Neuron*. 2010;66(1):85-100. doi:10.1016/j.neuron.2010.03.028
70. Sotomayor M, Corey DP, Schulten K. In Search of the Hair-Cell Gating Spring: Elastic Properties of Ankyrin and Cadherin Repeats. *Structure*. 2005;13(4):669-682. doi:10.1016/J.STR.2005.03.001
71. Araya-Secchi R, Neel BL, Sotomayor M. An elastic element in the protocadherin-15 tip link of the inner ear. *Nat Commun 2016 71*. 2016;7(1):1-14. doi:10.1038/ncomms13458
72. Mulhall EM, Ward A, Yang D, Koussa MA, Corey DP, Wong WP. Single-molecule force spectroscopy reveals the dynamic strength of the hair-cell tip-link connection. *Nat Commun*. 2021;12(1):1-15. doi:10.1038/s41467-021-21033-6
73. Salt AN, Inamura N, Thalmann R, Vora A. Calcium gradients in inner ear endolymph. *Am J Otolaryngol Neck Med Surg*. 1989;10(6):371-375. doi:10.1016/0196-0709(89)90030-6
74. Bosher SK, Warren RL. Very low calcium content of cochlear endolymph, an extracellular fluid. *Nature*. 1978;273(5661):377-378. doi:10.1038/273377a0
75. Powers RE, Gaudet R, Sotomayor M. A Partial Calcium-Free Linker Confers Flexibility to Inner-Ear Protocadherin-15. *Structure*. 2017;25(3):482-495.

doi:10.1016/j.str.2017.01.014
76. Webb SW, Grillet N, Andrade LR, et al. Regulation of PCDH15 function in mechanosensory hair cells by alternative splicing of the cytoplasmic domain. *Development*. 2011;138(8):1607-1617. doi:10.1242/DEV.060061
77. Pepermans E, Michel V, Goodyear R, et al. The CD2 isoform of protocadherin-15 is an essential component of the tip-link complex in mature auditory hair cells. *EMBO Mol Med*. 2014;6(7):984-992. doi:10.15252/EMMM.201403976
78. Tunyasuvunakool K, Adler J, Wu Z, et al. Highly accurate protein structure prediction for the human proteome. *Nat 2021 5967873*. 2021;596(7873):590-596. doi:10.1038/s41586-021-03828-1
79. Mammano F, Bortolozzi M, Ortolano S, Anselmi F. Ca2+ signaling in the inner ear. *Physiology*. 2007;22(2):131-144. doi:10.1152/physiol.00040.2006
80. Strimbu CE, Prasad S, Hakizimana P, Fridberger A. Control of hearing sensitivity by tectorial membrane calcium. *Proc Natl Acad Sci U S A*. 2019;116(12):5756-5764. doi:10.1073/pnas.1805223116
81. Yamoah EN, Lumpkin EA, Dumont RA, Smith PJS, Hudspeth AJ, Gillespie PG. Plasma membrane Ca2+-ATPase extrudes Ca2+ from hair cell stereocilia. *J Neurosci*. 1998;18(2):610-624. doi:10.1523/jneurosci.18-02-00610.1998
82. Narui Y, Sotomayor M. Tuning Inner-Ear Tip-Link Affinity Through Alternatively Spliced Variants of Protocadherin-15. *Biochemistry*. 2018;57(11):1702-1710. doi:10.1021/acs.biochem.7b01075
83. Otwinowski Z, Minor W. [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 1997;276:307-326. doi:10.1016/S0076-6879(97)76066-X
84. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007;40(Pt 4):658-674. doi:10.1107/S0021889807021206
85. Murshudov GN, Skubák P, Lebedev AA, et al. REFMAC5 for the refinement of macromolecular crystal structures. *urn:issn:0907-4449*. 2011;67(4):355-367. doi:10.1107/S0907444911001314
86. Emsley P, Cowtan K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr Sect D Biol Crystallogr*. 2004;60(12 I):2126-2132. doi:10.1107/S0907444904019158
87. Laskowski RA, MacArthur MW, Moss DS, Thornton JM, IUCr. PROCHECK: a program to check the stereochemical quality of protein structures. *urn:issn:0021-8898*. 1993;26(2):283-291. doi:10.1107/S0021889892009944
88. Zheng H, Cooper DR, Porebski PJ, Shabalin IG, Handing KB, Minor W. CheckMyMetal: a macromolecular metal-binding validation tool. *urn:issn:2059-7983*. 2017;73(3):223-233. doi:10.1107/S2059798317001061
89. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 2010;66(Pt 4):486-501. doi:10.1107/S0907444910007493
90. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol*

Graph. 1996;14(1):33-38. doi:10.1016/0263-7855(96)00018-5
91. MathWorks - Makers of MATLAB and Simulink - MATLAB & Simulink. https://www.mathworks.com/?s_tid=gn_logo. Accessed September 6, 2022.
92. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods*. 2017;14(3):290-296. doi:10.1038/nmeth.4169
93. Yang Y, Mlodzik M. Wnt-Frizzled/Planar Cell Polarity Signaling: Cellular Orientation by Facing the Wind (Wnt). *Annu Rev Cell Dev Biol*. 2015;31:623-646. doi:10.1146/annurev-cellbio-100814-125315
94. Baena-López LA, Baonza A, García-Bellido A. The orientation of cell divisions determines the shape of Drosophila organs. *Curr Biol*. 2005;15(18):1640-1644. doi:10.1016/j.cub.2005.07.062
95. Ségalen M, Johnston CA, Martin CA, et al. The Fz-Dsh Planar Cell Polarity Pathway Induces Oriented Cell Division via Mud/NuMA in Drosophila and Zebrafish. *Dev Cell*. 2010;19(5):740-752. doi:10.1016/j.devcel.2010.10.004
96. Wang M, Marco P de, Capra V, Kibar Z. Update on the Role of the Non-Canonical Wnt/Planar Cell Polarity Pathway in Neural Tube Defects. *Cells*. 2019;8(10):1-21. doi:10.3390/cells8101198
97. Phillips HM, Rhee HJ, Murdoch JN, et al. Disruption of planar cell polarity signaling results in congenital heart defects and cardiomyopathy attributable to early cardiomyocyte disorganization. *Circ Res*. 2007;101(2):137-145. doi:10.1161/CIRCRESAHA.106.142406
98. Qiao X, Liu Y, Li P, et al. Genetic analysis of rare coding mutations of CELSR1–3 in congenital heart and neural tube defects in Chinese people. *Clin Sci*. 2016;130(24):2329-2340. doi:10.1042/CS20160686
99. Butler MT, Wallingford JB. Planar cell polarity in development and disease. *Nat Rev Mol Cell Biol*. 2017;18(6):375-388. doi:10.1038/nrm.2017.11
100. Goodrich L V., Strutt D. Principles of planar polarity in animal development. *Development*. 2011;138(10):1877-1892. doi:10.1242/dev.054080
101. Strutt H, Strutt D. Asymmetric localisation of planar polarity proteins: Mechanisms and consequences. *Semin Cell Dev Biol*. 2009;20(8):957-963. doi:10.1016/j.semcdb.2009.03.006
102. Peng Y, Axelrod JD. *Asymmetric Protein Localization in Planar Cell Polarity: Mechanisms, Puzzles, and Challenges*. Vol 101. 1st ed. Elsevier Inc.; 2012. doi:10.1016/B978-0-12-394592-1.00002-8
103. Hakanen J, Ruiz-Reig N, Tissir F. Linking cell polarity to cortical development and malformations. *Front Cell Neurosci*. 2019;13(June):1-22. doi:10.3389/fncel.2019.00244
104. Harrison C, Shao H, Strutt H, Strutt D. Molecular mechanisms mediating asymmetric subcellular localisation of the core planar polarity pathway proteins. *Biochem Soc Trans*. 2020;48(4):1297-1308. doi:10.1042/BST20190404
105. Chen PL, Clandinin TR. The Cadherin Flamingo Mediates Level-Dependent Interactions that Guide Photoreceptor Target Choice in Drosophila. *Neuron*.

2008;58(1):26-33. doi:10.1016/j.neuron.2008.01.007
106. Tang X, Zhang L, Ma T, et al. Molecular mechanisms that regulate export of the planar cell-polarity protein Frizzled-6 out of the endoplasmic reticulum. *J Biol Chem*. 2020;295(27):8972-8987. doi:10.1074/jbc.RA120.012835
107. Tatin F, Taddei A, Weston A, et al. Planar cell polarity protein Celsr1 regulates endothelial adherens junctions and directed cell rearrangements during valve morphogenesis. *Dev Cell*. 2013;26(1):31-44. doi:10.1016/j.devcel.2013.05.015
108. Devenport D, Fuchs E. Planar polarization in embryonic epidermis orchestrates global asymmetric morphogenesis of hair follicles. *Nat Cell Biol*. 2008;10(11):1257-1268. doi:10.1038/ncb1784
109. Pei J, Grishin N V. Expansion of divergent SEA domains in cell surface proteins and nucleoporin 54. *Protein Sci*. 2017;26(3):617-630. doi:10.1002/pro.3096
110. Formstone CJ, Moxon C, Murdoch J, Little P, Mason I. Basal enrichment within neuroepithelia suggests novel function(s) for Celsr1 protein. *Mol Cell Neurosci*. 2010;44(3):210-222. doi:10.1016/j.mcn.2010.03.008
111. Sotomayor M, Gaudet R, Corey DP. Sorting out a promiscuous superfamily: Towards cadherin connectomics. *Trends Cell Biol*. 2014. doi:10.1016/j.tcb.2014.03.007
112. Harrison OJ, Brasch J, Katsamba PS, et al. Family-wide Structural and Biophysical Analysis of Binding Interactions among Non-clustered δ-Protocadherins. *Cell Rep*. 2020;30(8):2655-2671.e7. doi:10.1016/j.celrep.2020.02.003
113. Modak D, Sotomayor M. Identification of an adhesive interface for the non-clustered δ1 protocadherin-1 involved in respiratory diseases. *Commun Biol*. 2019;2(1):354. doi:10.1038/s42003-019-0586-0
114. Cooper SR, Jontes JD, Sotomayor M. Structural determinants of adhesion by protocadherin-19 and implications for its role in epilepsy. *Elife*. 2016;5(OCTOBER2016). doi:10.7554/eLife.18529
115. Medina E, Easa Y, Lester DK, Lau EK, Sprinzak D, Luca VC. Structure of the planar cell polarity cadherins Fat4 and Dachsous1. *Nat Commun 2023 141*. 2023;14(1):1-11. doi:10.1038/s41467-023-36435-x
116. Nicoludis JM, Vogt BE, Green AG, Schärfe CPI, Marks DS, Gaudet R. Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1-4. *Elife*. 2016;5(JULY). doi:10.7554/ELIFE.18449
117. Vizurraga A, Adhikari R, Yeung J, Yu M, Tall GG. Mechanisms of adhesion G protein–coupled receptor activation. *J Biol Chem*. 2020;295(41):14065-14083. doi:10.1074/JBC.REV120.007423
118. Araç D, Boucard AA, Bolliger MF, et al. A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoproteolysis. 2012;31(6):1364-1378. https://onlinelibrary.wiley.com/doi/full/10.1038/emboj.2012.26. Accessed March 17, 2023.
119. Bui DLH, Roach A, Li J, et al. The adhesion GPCRs CELSR1-3 and LPHN3

engage G proteins via distinct activation mechanisms. *bioRxiv*. April 2023. doi:10.1101/2023.04.02.535287
120. Overduin M, Harvey TS, Bagby S, et al. Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. *Science (80-)*. 1995;267(5196):386-389. doi:10.1126/science.7824937
121. Jin X, Walker MA, Felsövályi K, et al. Crystal structures of Drosophila N-cadherin ectodomain regions reveal a widely used class of Ca 2+-free interdomain linkers. *Proc Natl Acad Sci U S A*. 2012;109(3):E127-34. doi:10.1073/pnas.1117538108
122. Powers RE, Gaudet R, Sotomayor M. A Partial Calcium-Free Linker Confers Flexibility to Inner-Ear Protocadherin-15. *Structure*. 2017;25(3):482-495. doi:10.1016/j.str.2017.01.014
123. Tsukasaki Y, Miyazaki N, Matsumoto A, et al. Giant cadherins Fat and Dachsous self-bend to organize properly spaced intercellular junctions. *Proc Natl Acad Sci U S A*. 2014;111(45):16011-16016. doi:10.1073/pnas.1418990111
124. Fawcett DW, McNutt NS. The ultrastructure of the cat myocardium. I. Ventricular papillary muscle. *J Cell Biol*. 1969;42(1):1-45. doi:10.1083/jcb.42.1.1
125. Scott McNutt N, Weinstein RS. Membrane ultrastructure at mammalian intercellular junctions. *Prog Biophys Mol Biol*. 1973;26(C):45-101. doi:10.1016/0079-6107(73)90017-5
126. Miyaguchi K. Ultrastructure of the zonula adherens revealed by rapid-freeze deep-etching. *J Struct Biol*. 2000. doi:10.1006/jsbi.2000.4244
127. Nishimura T, Honda H, Takeichi M. Planar cell polarity links axes of spatial dynamics in neural-tube closure. *Cell*. 2012. doi:10.1016/j.cell.2012.04.021
128. Emond MR, Jontes JD. Bead aggregation assays for the characterization of putative cell adhesion molecules. *J Vis Exp*. 2014;(92):e51762. doi:10.3791/51762
129. Koch AW, Farooq A, Shan W, Zeng L, Colman DR, Zhou MM. Structure of the Neural (N-) Cadherin Prodomain Reveals a Cadherin Extracellular Domain-like Fold without Adhesive Characteristics. *Structure*. 2004;12(5):793-805. doi:10.1016/J.STR.2004.02.034
130. Krissinel E, Henrick K. Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol*. 2007;372(3):774-797. doi:10.1016/j.jmb.2007.05.022
131. Ponstingl H, Henrick K, Thornton JM. Discriminating between homodimeric and monomeric proteins in the crystalline state. *Proteins Struct Funct Genet*. 2000;41(1):47-57. doi:10.1002/1097-0134(20001001)41:1<47::AID-PROT80>3.0.CO;2-8
132. Karplus M, Petsko GA. Molecular dynamics simulations in biology. *Nature*. 1990;347(6294):631-639. doi:10.1038/347631a0
133. Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res*. 2016;44(W1):W424-W429. doi:10.1093/NAR/GKW389
134. Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. Accurate SAXS profile computation and its assessment by contrast variation experiments.

2013;105(4):962-974. http://www.cell.com/article/S0006349513008059/fulltext. Accessed August 8, 2022.
135. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2
136. Tunyasuvunakool K, Adler J, Wu Z, et al. Highly accurate protein structure prediction for the human proteome. *Nature*. 2021;596(7873):590-596. doi:10.1038/s41586-021-03828-1
137. Amyot R, Flechsig H. BioAFMviewer: An interactive interface for simulated AFM scanning of biomolecular structures and dynamics. *PLoS Comput Biol*. 2020;16(11):1-12. doi:10.1371/journal.pcbi.1008444
138. Gray ME, Johnson ZR, Modak D, Tamilselvan E, Tyska MJ, Sotomayor M. Heterophilic and homophilic cadherin interactions in intestinal intermicrovillar links are species dependent. *PLOS Biol*. 2021;19(12):e3001463. doi:10.1371/JOURNAL.PBIO.3001463
139. Biswas S, Emond MR, Jontes JD. Protocadherin-19 and N-cadherin interact to control cell movements during anterior neurulation. *J Cell Biol*. 2010;191(5):1029-1041. doi:10.1083/jcb.201007008
140. Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput*. 2002:310-322. doi:10.1142/9789812799623_0029
141. Cailliez F, Lavery R. Cadherin Mechanics and Complexation: The Importance of Calcium Binding. *Biophys J*. 2005;89(6):3895. doi:10.1529/BIOPHYSJ.105.067322
142. Kim SA, Tai CY, Mok LP, Mosser EA, Schuman EM. Calcium-dependent dynamics of cadherin interactions at cell-cell junctions. *Proc Natl Acad Sci U S A*. 2011;108(24):9857-9862. doi:10.1073/PNAS.1019003108/SUPPL_FILE/PNAS.201019003SI.PDF
143. Allache R, De Marco P, Merello E, Capra V, Kibar Z. Role of the planar cell polarity gene CELSR1 in neural tube defects and caudal agenesis. *Birth Defects Res Part A - Clin Mol Teratol*. 2012;94(3):176-181. doi:10.1002/bdra.23002
144. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 1997;276:307-326. doi:10.1016/S0076-6879(97)76066-X
145. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007;40(4):658-674. doi:10.1107/S0021889807021206
146. Hooft RWW, Vriend G, Sander C, Abola EE. Errors in protein structures [3]. *Nature*. 1996;381(6580):272. doi:10.1038/381272a0
147. Sivasankar S, Zhang Y, Nelson WJ, Chu S. Characterizing the initial encounter complex in cadherin adhesion. *Structure*. 2009;17(8):1075. doi:10.1016/J.STR.2009.06.012
148. Emond MR, Biswas S, Blevins CJ, Jontes JD. A complex of Protocadherin-19 and

N-cadherin mediates a novel mechanism of cell adhesion. *J Cell Biol*. 2011;195(7):1115-1121. doi:10.1083/JCB.201108115
149. Kwon M, Firestein BL. DNA transfection: Calcium phosphate method. *Methods Mol Biol*. 2013;1018:107-110. doi:10.1007/978-1-62703-444-9_10/COVER
150. Walker JM, Gasteiger E, Hoogland C, et al. Protein Identification and Analysis Tools on the ExPASy Server. *Proteomics Protoc Handb*. 2005:571-607. doi:10.1385/1-59259-890-0:571
151. Schuck P. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem*. 2003;320(1):104-124. doi:10.1016/S0003-2697(03)00289-6
152. Schuck P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys J*. 2000;78(3):1606-1619. doi:10.1016/S0006-3495(00)76713-0
153. Dyer KN, Hammel M, Rambo RP, et al. High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. *Methods Mol Biol*. 2014;1091:245-258. doi:10.1007/978-1-62703-691-7_18
154. Classen S, Hura GL, Holton JM, et al. Implementation and performance of SIBYLS: a dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the Advanced Light Source. *J Appl Crystallogr*. 2013;46(Pt 1):1-13. doi:10.1107/S0021889812048698
155. Tully MD, Tarbouriech N, Rambo RP, Hutin S. Analysis of sec-saxs data via efa deconvolution and scatter. *J Vis Exp*. 2021;2021(167):1-16. doi:10.3791/61578
156. Konarev P V., Volkov V V., Sokolova A V., Koch MHJ, Svergun DI. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *urn:issn:0021-8898*. 2003;36(5):1277-1282. doi:10.1107/S0021889803012779
157. Franke D, Petoukhov M V., Konarev P V., et al. ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr*. 2017;50(4):1212-1225. doi:10.1107/S1600576717007786
158. Svergun DI, IUCr. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *urn:issn:0021-8898*. 1992;25(4):495-503. doi:10.1107/S0021889892001663
159. Franke D, Svergun DI. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *urn:issn:0021-8898*. 2009;42(2):342-346. doi:10.1107/S0021889809000338
160. Volkov V V., Svergun DI. Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Crystallogr*. 2003;36(3 I):860-864. doi:10.1107/S0021889803000268
161. Piiadov V, Ares de Araújo E, Oliveira Neto M, Craievich AF, Polikarpov I. SAXSMoW 2.0: Online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale. *Protein Sci*. 2019;28(2):454-463. doi:10.1002/PRO.3528
162. Phillips JC, Braun R, Wang W, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005;26(16):1781-1802. doi:10.1002/JCC.20289

163. PyMOL | pymol.org. https://pymol.org/2/. Accessed September 6, 2022.
164. Grace Home. https://plasma-gate.weizmann.ac.il/Grace/. Accessed September 6, 2022.
165. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32(5):1792-1797. doi:10.1093/NAR/GKH340
166. Geneious. https://help.geneious.com/hc/en-us. Accessed September 6, 2022.
167. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics*. 2009;25(9):1189-1191. doi:10.1093/bioinformatics/btp033
168. King N. The unicellular ancestry of animal development. *Dev Cell*. 2004;7(3):313-325. doi:10.1016/j.devcel.2004.08.010
169. Mathur P, Yang J. Usher syndrome: Hearing loss, retinal degeneration and associated abnormalities. *Biochim Biophys Acta - Mol Basis Dis*. 2015;1852(3):406-420. doi:10.1016/j.bbadis.2014.11.020
170. Montcouquiol M, Kelley MW. Development and patterning of the cochlea: From convergent extension to planar polarity. *Cold Spring Harb Perspect Med*. 2020;10(1). doi:10.1101/cshperspect.a033266
171. Modak D, Sotomayor M. Identification of an adhesive interface for the non-clustered δ1 protocadherin-1 involved in respiratory diseases. *Commun Biol 2019 21*. 2019;2(1):1-13. doi:10.1038/s42003-019-0586-0
172. Kefauver JM, Ward AB, Patapoutian A. Discoveries in structure and physiology of mechanically activated ion channels. *Nature*. 2020;587(7835):567-576. doi:10.1038/s41586-020-2933-1
173. Cox CD, Bavi N, Martinac B. Biophysical Principles of Ion-Channel-Mediated Mechanosensory Transduction. *Cell Rep*. 2019;29(1):1-12. doi:10.1016/j.celrep.2019.08.075
174. Douguet D, Honoré E. Mammalian Mechanoelectrical Transduction: Structure and Function of Force-Gated Ion Channels. *Cell*. 2019;179(2):340-354. doi:10.1016/j.cell.2019.08.049
175. Jin P, Jan LY, Jan YN. Mechanosensitive Ion Channels: Structural Features Relevant to Mechanotransduction Mechanisms. *Annu Rev Neurosci*. 2020;43:207-229. doi:10.1146/annurev-neuro-070918-050509
176. Murthy SE, Dubin AE, Patapoutian A. Piezos thrive under pressure: Mechanically activated ion channels in health and disease. *Nat Rev Mol Cell Biol*. 2017;18(12):771-783. doi:10.1038/nrm.2017.92
177. Bootha IR, Blount P. The MscS and MscL families of mechanosensitive channels act as microbial emergency release valves. *J Bacteriol*. 2012;194(18):4802-4809. doi:10.1128/JB.00576-12
178. Procko C, Murthy S, Keenan WT, et al. Stretch-activated ion channels identified in the touch-sensitive structures of carnivorous droseraceae plants. *Elife*. 2021;10. doi:10.7554/eLife.64250
179. Walker RG, Willingham AT, Zuker CS. A Drosophila mechanosensory transduction channel. *Science (80-)*. 2000;287(5461):2229-2234.

doi:10.1126/science.287.5461.2229
180. Yan Z, Zhang W, He Y, et al. Drosophila NOMPC is a mechanotransduction channel subunit for gentle-touch sensation. *Nature*. 2013;493(7431):221-225. doi:10.1038/nature11685
181. Zhang W, Yan Z, Jan LY, Jan YN. Sound response mediated by the TRP channels NOMPC, NANCHUNG, and INACTIVE in chordotonal organs of Drosophila larvae. *Proc Natl Acad Sci U S A*. 2013;110(33):13612-13617. doi:10.1073/pnas.1312477110
182. Brohawn SG. How ion channels sense mechanical force: Insights from mechanosensitive K2P channels TRAAK, TREK1, and TREK2. *Ann N Y Acad Sci*. 2015;1352(1):20-32. doi:10.1111/nyas.12874
183. Natale AM, Deal PE, Minor DL. Structural Insights into the Mechanisms and Pharmacology of K2P Potassium Channels. *J Mol Biol*. 2021;433(17):166995. doi:10.1016/j.jmb.2021.166995
184. Plant LD. A role for K2P channels in the operation of somatosensory nociceptors. *Front Mol Neurosci*. 2012;5(MARCH):20861. doi:10.3389/fnmol.2012.00021
185. Bagriantsev SN, Gracheva EO, Gallagher PG. Piezo proteins: Regulators of mechanosensation and other cellular processes. *J Biol Chem*. 2014;289(46):31673-31681. doi:10.1074/jbc.R114.612697
186. Young M, Lewis AH, Grandl J. Physics of mechanotransduction by Piezo ion channels. *J Gen Physiol*. 2022;154(7). doi:10.1085/jgp.202113044
187. Pan B, Akyuz N, Liu XP, et al. TMC1 Forms the Pore of Mechanosensory Transduction Channels in Vertebrate Inner Ear Hair Cells. *Neuron*. 2018;99(4):736-753.e6. doi:10.1016/j.neuron.2018.07.033
188. Pan B, Géléoc GS, Asai Y, et al. TMC1 and TMC2 are components of the mechanotransduction channel in hair cells of the mammalian inner ear. *Neuron*. 2013;79(3):504-515. doi:10.1016/j.neuron.2013.06.019
189. Murthy SE, Dubin AE, Whitwam T, et al. OSCA/TMEM63 are an evolutionarily conserved family of mechanically activated ion channels. *Elife*. 2018;7. doi:10.7554/eLife.41844
190. Yuan F, Yang H, Xue Y, et al. OSCA1 mediates osmotic-stress-evoked Ca2+ increases vital for osmosensing in Arabidopsis. *Nature*. 2014;514(7522):367-371. doi:10.1038/nature13593
191. Hou C, Tian W, Kleist T, et al. DUF221 proteins are a family of osmosensitive calcium-permeable cation channels conserved across eukaryotes. *Cell Res*. 2014;24(5):632-635. doi:10.1038/cr.2014.14
192. Jojoa-Cruz S, Saotome K, Murthy SE, et al. Cryo-EM structure of the mechanically activated ion channel OSCA1.2. *Elife*. 2018;7. doi:10.7554/eLife.41845
193. Liu X, Wang J, Sun L. Structure of the hyperosmolality-gated calcium-permeable channel OSCA1.2. *Nat Commun*. 2018;9(1):1-9. doi:10.1038/s41467-018-07564-5
194. Maity K, Heumann JM, McGrath AP, et al. Cryo-EM structure of OSCA1.2 from Oryza sativa elucidates the mechanical basis of potential membrane

hyperosmolality gating. *Proc Natl Acad Sci U S A*. 2019;116(28):14309-14318. doi:10.1073/pnas.1900774116
195. Zhang M, Wang D, Kang Y, et al. Structure of the mechanosensitive OSCA channels. *Nat Struct Mol Biol*. 2018;25(9):850-858. doi:10.1038/s41594-018-0117-6
196. Medrano-Soto A, Moreno-Hagelsieb G, McLaughlin D, Ye ZS, Hendargo KJ, Saier MH. Bioinformatic characterization of the Anoctamin Superfamily of Ca2+-activated ion channels and lipid scramblases. *PLoS One*. 2018;13(3):e0192851. doi:10.1371/journal.pone.0192851
197. Dang S, Feng S, Tien J, et al. Cryo-EM structures of the TMEM16A calciumactivated chloride channel. *Nature*. 2017;552(7685):426-429. doi:10.1038/nature25024
198. Paulino C, Kalienkova V, Lam AKM, Neldner Y, Dutzler R. Activation mechanism of the calcium-activated chloride channel TMEM16A revealed by cryo-EM. *Nature*. 2017;552(7685):421-425. doi:10.1038/nature24652
199. Brunner JD, Lim NK, Schenck S, Duerst A, Dutzler R. X-ray structure of a calcium-activated TMEM16 lipid scramblase. *Nature*. 2014;516(7530):207-212. doi:10.1038/nature13984
200. Jeong H, Clark S, Goehring A, et al. Structures of the TMC-1 complex illuminate mechanosensory transduction. *Nature*. 2022;610(7933):796-803. doi:10.1038/s41586-022-05314-8
201. Fukumura S, Hiraide T, Yamamoto A, et al. A novel de novo TMEM63A variant in a patient with severe hypomyelination and global developmental delay. *Brain Dev*. 2022;44(2):178-183. doi:10.1016/j.braindev.2021.09.006
202. Tonduti D, Mura E, Masnada S, et al. Spinal cord involvement and paroxysmal events in “Infantile Onset Transient Hypomyelination” due to TMEM63A mutation. *J Hum Genet*. 2021;66(10):1035-1037. doi:10.1038/s10038-021-00921-1
203. Yan H, Helman G, Murthy SE, et al. Heterozygous Variants in the Mechanosensitive Ion Channel TMEM63A Result in Transient Hypomyelination during Infancy. *Am J Hum Genet*. 2019;105(5):996-1004. doi:10.1016/j.ajhg.2019.09.011
204. Yan H, Ji H, Kubisiak T, et al. Genetic analysis of 20 patients with hypomyelinating leukodystrophy by trio-based whole-exome sequencing. *J Hum Genet*. 2021;66(8):761-768. doi:10.1038/s10038-020-00896-5
205. Vetro A, Balestrini S, Pelorosso C, et al. Stretch-activated ion channel TMEM63B associates with developmental and epileptic encephalopathies and progressive neurodegeneration. *medRxiv*. 2022;22:2022.11.22.22282283. doi:10.1101/2022.11.22.22282283
206. Tábara LC, Al-Salmi F, Maroofian R, et al. TMEM63C mutations cause mitochondrial morphology defects and underlie hereditary spastic paraplegia. *Brain*. 2022;145(9):3095-3107. doi:10.1093/brain/awac123
207. Du H, Ye C, Wu D, et al. The Cation Channel TMEM63B Is an Osmosensor

Required for Hearing. *Cell Rep*. 2020;31(5):107596. doi:10.1016/j.celrep.2020.107596
208. Lan C, Stulz A, Barthes NPF, et al. Designed membrane protein heterodimers and control of their affinity by binding domain and membrane linker properties. *Nanoscale*. 2021;13(48):20692-20702. doi:10.1039/d1nr06574b
209. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nat 2021 5967873*. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2
210. Romero LO, Massey AE, Mata-Daboin AD, et al. Dietary fatty acids fine-tune Piezo1 mechanical response. *Nat Commun*. 2019;10(1):1-14. doi:10.1038/s41467-019-09055-7
211. Battle AR, Ridone P, Bavi N, Nakayama Y, Nikolaev YA, Martinac B. Lipid-protein interactions: Lessons learned from stress. *Biochim Biophys Acta - Biomembr*. 2015;1848(9):1744-1756. doi:10.1016/j.bbamem.2015.04.012
212. Smart OS, Neduvelil JG, Wang X, Wallace BA, Sansom MSP. HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph*. 1996;14(6):354-360. doi:10.1016/S0263-7855(97)00009-X
213. Zubcevic L, Lee SY. The role of π-helices in TRP channel gating. *Curr Opin Struct Biol*. 2019;58:314-323. doi:10.1016/j.sbi.2019.06.011
214. Peters CJ, Gilchrist JM, Tien J, et al. The Sixth Transmembrane Segment Is a Major Gating Component of the TMEM16A Calcium-Activated Chloride Channel. *Neuron*. 2018;97(5):1063-1077.e4. doi:10.1016/j.neuron.2018.01.048
215. Dhaka A, Viswanath V, Patapoutian A. TRP ion channels and temperature sensation. *Annu Rev Neurosci*. 2006;29:135-161. doi:10.1146/annurev.neuro.29.051605.112958
216. Kashio M, Tominaga M. TRP channels in thermosensation. *Curr Opin Neurobiol*. 2022;75:102591. doi:10.1016/j.conb.2022.102591
217. Delmas P, Parpaite T, Coste B. PIEZO channels and newcomers in the mammalian mechanosensitive ion channel family. *Neuron*. 2022;110(17):2713-2727. doi:10.1016/j.neuron.2022.07.001
218. Albuisson J, Murthy SE, Bandell M, et al. Dehydrated hereditary stomatocytosis linked to gain-of-function mutations in mechanically activated PIEZO1 ion channels. *Nat Commun*. 2013;4(1):1-9. doi:10.1038/ncomms2899
219. Fettiplace R. Hair cell transduction, tuning, and synaptic transmission in the mammalian cochlea. *Compr Physiol*. 2017;7(4):1197-1227. doi:10.1002/cphy.c160049
220. Jeng G, Aggarwal M, Yu WP, Chen TY. Independent activation of distinct pores in dimeric TMEM16A channels. *J Gen Physiol*. 2016;148(5):393-404. doi:10.1085/jgp.201611651
221. Lim NK, Lam AKM, Dutzler R. Independent activation of ion conduction pores in the double-barreled calcium-activated chloride channel TMEM16A. *J Gen Physiol*. 2016;148(5):375-392. doi:10.1085/jgp.201611650
222. Akyuz N, Karavitaki KD, Pan B, et al. Mechanical gating of the auditory

transduction channel TMC1 involves the fourth and sixth transmembrane helices. *Sci Adv*. 2022;8(28):1126. doi:10.1126/sciadv.abo1126
223. Feng S, Dang S, Han TW, et al. Cryo-EM Studies of TMEM16F Calcium-Activated Ion Channel Suggest Features Important for Lipid Scrambling. *Cell Rep*. 2019;28(2):567-579.e4. doi:10.1016/j.celrep.2019.06.023
224. Goehring A, Lee CH, Wang KH, et al. Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat Protoc*. 2014;9(11):2574-2585. doi:10.1038/nprot.2014.173
225. Tien J, Lee HY, Minor DL, Jan YN, Jan LY. Identification of a dimerization domain in the TMEM16A calcium-activated chloride channel (CaCC). *Proc Natl Acad Sci U S A*. 2013;110(16):6352-6357. doi:10.1073/pnas.1303672110
226. Zheng W, Hussein S, Yang JW, et al. A novel PKD2L1 C-terminal domain critical for trimerization and channel function. *Sci Rep*. 2015;5(1):1-12. doi:10.1038/srep09460
227. Zheng W, Shen C, Wang L, et al. pH regulates potassium conductance and drives a constitutive proton current in human TMEM175. *Sci Adv*. 2022;8(12):1568. doi:10.1126/sciadv.abm1568
228. Last NB, Kolmakova-Partensky L, Shane T, Miller C. Mechanistic signs of double-barreled structure in a fluoride ion channel. *Elife*. 2016;5(JULY). doi:10.7554/eLife.18767
229. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods*. 2017;14(4):331-332. doi:10.1038/nmeth.4193
230. Rohou A, Grigorieff N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol*. 2015;192(2):216-221. doi:10.1016/j.jsb.2015.08.008
231. Wagner T, Merino F, Stabrin M, et al. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun Biol*. 2019;2(1):1-13. doi:10.1038/s42003-019-0437-z
232. Scheres SHW. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol*. 2012;180(3):519-530. doi:10.1016/j.jsb.2012.09.006
233. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods*. 2017;14(3):290-296. doi:10.1038/nmeth.4169
234. Zivanov J, Nakane T, Scheres SHW. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ*. 2019;6(1):5-17. doi:10.1107/S205225251801463X
235. Morin A, Eisenbraun B, Key J, et al. Collaboration gets the most out of software. *Elife*. 2013;2013(2). doi:10.7554/eLife.01456
236. Pettersen EF, Goddard TD, Huang CC, et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci*. 2021;30(1):70-82. doi:10.1002/pro.3943

237. Adams PD, Afonine P V., Bunkóczi G, et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr*. 2010;66(2):213-221. doi:10.1107/S0907444909052925
238. Afonine P V., Poon BK, Read RJ, et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr Sect D Struct Biol*. 2018;74(6):531-544. doi:10.1107/S2059798318006551
239. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605-1612. doi:10.1002/jcc.20084
240. Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv*. March 2022:2021.10.04.463034. doi:10.1101/2021.10.04.463034
241. Phillips JC, Hardy DJ, Maia JDC, et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys*. 2020;153(4):44130. doi:10.1063/5.0014475
242. Buck M, Bouguet-Bonnet S, Pastor RW, MacKerell AD. Importance of the CMAP correction to the CHARMM22 protein force field: Dynamics of hen lysozyme. *Biophys J*. 2006;90(4):L36-L38. doi:10.1529/biophysj.105.078154
243. Gumbart J, Khalili-Araghi F, Sotomayor M, Roux B. Constant electric field simulations of the membrane potential illustrated with simple systems. *Biochim Biophys Acta - Biomembr*. 2012;1818(2):294-302. doi:10.1016/j.bbamem.2011.09.030
244. Zheng W, Yang X, Hu R, et al. Hydrophobic pore gates regulate ion permeation in polycystic kidney disease 2 and 2L1 channels. *Nat Commun*. 2018;9(1):1-14. doi:10.1038/s41467-018-04586-x
245. Nikolaev YA, Cox CD, Ridone P, et al. Mammalian TRP ion channels are insensitive to membrane stretch. *J Cell Sci*. 2019;132(23). doi:10.1242/jcs.238360

Appendix A. TMEM63 proteins function as monomeric high threshold mechanosensitive ion channels[3]

Abstract

OSCA/TMEM63s form mechanically activated (MA) ion channels in plants and animals, respectively. OSCAs and related TMEM16s and transmembrane channel-like (TMC) proteins form homodimers with two pores. Here, we uncover an unanticipated monomeric configuration of TMEM63 proteins. Structures of TMEM63A and TMEM63B (referred to as TMEM63s) revealed a single highly restricted pore. Functional analyses demonstrated that TMEM63s are bona fide mechanosensitive ion channels, characterized by small conductance and high thresholds. TMEM63s possess evolutionary variations in the intracellular linker IL2, which mediates dimerization in OSCAs. Replacement of OSCA1.2 IL2 with TMEM63A IL2 or mutations to key variable residues resulted in monomeric OSCA1.2 and MA currents with significantly higher thresholds. Structural analyses

[3] This chapter is adapted from a paper published in Neuron, https://doi.org/10.1016/j.neuron.2023.07.006. Wang Zheng, Shaun Rawson, Zhangfei Shen, Elakkiya Tamilselvan, Harper E. Smith, Julia Halford, Chen Shen, Swetha E. Murthy, Maximilian H. Ulbrich, Marcos Sotomayor, Tian-Min Fu, and Jeffrey R. Holt. W.Z. performed the cryo-EM, electrophysiological, and biochemical aspects of the work. S.R. performed the cryo-EM data processing and map generation. Z.S. did initial grid screening and preliminary data processing during sample optimization. W.Z., T.-M.F., and C.S. did the model building and refinement. W.Z. and T.-M.F. performed the data analysis and prepared the figures. E.T., H.E.S., and M.S. contributed to structure refinement and carried out simulations and data analysis. M.H.U. performed single-molecule imaging and data analysis. J.H. and S.E.M. maintained the knockin mouse model and prepared the brain tissue. W.Z., T.-M.F., and J.R.H. conceived the study, and J.R.H. provided supervision and funding throughout the project. W.Z., T.-M.F., and J.R.H. wrote the paper with input from all authors.

revealed substantial conformational differences in the mechano-sensing domain IL2 and gating helix TM6 between TMEM63s and OSCA1.2. Our studies reveal that mechanosensitivity in OSCA/TMEM63 channels is affected by oligomerization and suggest gating mechanisms that may be shared by OSCA/TMEM63, TMEM16, and TMC channels.

Introduction

Mechanosensitivity, the ability to detect and respond to mechanical stimuli, is fundamental to many physiological processes in all living organisms[172]. Mechanically activated (MA) ion channels convert external or internal mechanical cues into electrical signals, enabling cells to respond rapidly to their physical environment[173–176]. MA ion channels are widely expressed in all domains of life, including bacteria, plants, and animals, and play critical roles in diverse biological processes. The bacterial MscS and MscL proteins were the first identified MA ion channels and are implicated in osmotic stress responses[177]. In carnivorous plants, FLYC1 and FLYC2 are expressed in mechanosensory cells within touch-sensitive hairs triggered for capturing prey[178]. In invertebrates, such as *Drosophila melanogaster*, NompC, a member of the transient receptor potential (TRP) channel family, is involved in touch sensation and sound detection[179–181]. In mammals, the K2P potassium channel family may contribute to mechanical nociception[182–184], whereas the Piezo family of MA channels is implicated in a wide range of physiological processes, including gentle touch sensation, blood pressure regulation,

and vascular development[176,185,186]. Lastly, TMC1 from the transmembrane (TM) channel-like (TMC) family forms the pore of the mechanosensory transduction complex in inner ear hair cells and is critical for hearing[66,187,188].

Despite progress in identifying MA ion channels, the largest family of MA ion channels OSCA/TMEM63 was only recently identified[189,190]. OSCA channels exist in plants with 15 paralogs identified in *Arabidopsis thaliana*[189–191]. OSCA1.1 mutations in *A. thaliana* lead to impaired rapid osmotic stress-induced Ca^{2+} elevation, suggesting a role of OSCAs as osmosensors in plants[190,191]. Electrophysiological studies demonstrated that OSCAs are bona fide pore-forming mechanosensitive ion channels with preferred selectivity for cations[189]. Cryogenic electron microscopy (cryo-EM) analysis of OSCA1.1, OSCA1.2, and OSCA3.1[192–195] revealed that OSCA channels form homodimers with one ion permeation pathway in each monomer. The OSCA monomer can be divided into a TM domain (TMD), containing 11 TM segments labeled TM0 to TM10, and an intracellular domain (ICD). The ICD is mainly composed of a long intracellular linker between TM2 and TM3, known as IL2, and a C-terminal tail. A notable feature of IL2 is that it contains two long amphipathic α helices with an intervening hook domain inserted into the membrane, which was proposed to sense membrane tension[192]. Sequence analysis indicated that the OSCA/TMEM63 family is distantly related to the TMEM16/anoctamin and TMC families[196]. Structural analysis of TMEM16 proteins, such as mTMEM16A (calcium-activated chloride channel in mouse)[197,198],

*nh*TMEM16 (lipid scramblase in fungus *Nectria haematococca*)[199], and TMC-1 in *C. elegans*[200] revealed that they all share a dimeric architecture with TM domain arrangements similar to OSCA channels. TMEM63 proteins, including A, B, and C members, are animal orthologs of the plant OSCAs, with recently identified physiological functions and genetic pathologies. TMEM63A heterozygous variants have been identified in young human patients with hypo-myelinating leukodystrophies, characterized by myelin deficits, and global developmental delays with seizures[201–204]. Ten distinct TMEM63B heterozygous variants were recently reported in 16 unrelated human patients with a broad range of brain function abnormities, including severe early-onset developmental and epileptic encephalopathy, intellectual disability, and severe motor and cortical visual impairment[205]. Biallelic variants of TMEM63C were also identified in individuals with hereditary spastic paraplegias[206]. These clinical reports and genetic studies have demonstrated critical roles for TMEM63s in normal central nervous system development and neurodegenerative diseases. TMEM63B deficiency in mice was also reported to lead to deafness, and it was suggested to act as an osmosensor in outer hair cells of the inner ear[207]. When expressed in heterologous cells, TMEM63A and TMEM63B evoke small MA currents, whereas no current was detected for TMEM63C[189]. This raised fundamental questions as to whether TMEM63s are bona fide pore-forming MA ion channels. Furthermore, the structures and mechano-gating mechanisms of the entire TMEM63 family remain unexplored.

Here, we reveal an unexpected monomeric configuration of TMEM63s *in vitro* and *in vivo*, which contrasts with the dimeric assembly of OSCAs, TMEM16s, and TMC1. We also present cryo-EM structures of human TMEM63A and TMEM63B with nominal resolutions of 3.8 and 3.6 Å respectively. Structural and biochemical analysis reveal that the monomeric configuration of TMEM63s is governed by the IL2 domain. Electrophysiological recordings reveal that TMEM63A and TMEM63B are pore-forming MA ion channels with small conductance and high-threshold mechanosensitivity, compared with dimeric OSCA1.2. Moreover, by leveraging monomeric OSCA1.2 mutants, we find that the two protomers in an OSCA dimer are mechanically gated in an independent manner, and monomeric OSCA1.2 exhibits higher thresholds than dimeric OSCA1.2. Finally, structural analysis uncovers distinct conformations of mechano-sensing domain IL2 and gating helix TM6 between monomeric TMEM63s and dimeric OSCA1.2. Based on these data, we propose a mechanistic model that explains how protein oligomerization affects mechanosensitivity of the OSCA/TMEM63 channels.

Results

TMEM63s are monomeric proteins

We first examined the oligomeric configuration of overexpressed human TMEM63 proteins in HEK293T cell lysates using fluorescence-detection size exclusion chromatography (FSEC) (Fig. A. 1*a*). With four membrane proteins as controls, we generated a standard curve, in which the oligomeric molecular weight

(MW, kDa) showed a linear function of the retention time T (min) (MW = 242.6 3 T + 2,391, Fig. A. 1*b*). Compared with *A. thaliana* OSCA1.2, human TMEM63A and TMEM63B exhibited right-shifted peak profiles, with calculated MWs of 114 and 130 kDa, respectively (Fig. A. 1*b*), consistent with the size of glycosylated monomers (Fig. A. 9*b* and 10*b*). The monomeric nature of TMEM63s was further supported by cross-linking assays in cell lysates (Fig. A. 1*c* and 9*a*) and intact cells (Fig. A. 1*d*) and by blue native polyacrylamide gel electrophoresis (PAGE) (Fig. A. 1*c*). Human TMEM63C protein also displayed a monomeric configuration. Formation of heterodimers among TMEM63A, TMEM63B, and TMEM63C proteins was not observed (Fig. A. 8*b*). These data showed that overexpressed TMEM63s in HEK cells exist as monomers.

Next, we assessed oligomeric status of endogenous TMEM63s in mouse brain using denaturing and blue native PAGE. As controls, acid-gated ASIC1 channels and voltage- gated HCN1 channels were first confirmed to assemble as trimers and tetramers, respectively (Fig. A. 1*e* and 1*f*). In contrast, TMEM63A forms monomers with an apparent MW close to 100 kDa, consistent with overexpressed TMEM63A in HEK cells (Fig. A. 1*e* and 1*f*). Due to the lack of specific antibodies against TMEM63B, we evaluated the oligomeric status of TMEM63B by employing a Tmem63b$^{HA/HA}$ knockin mouse model, in which an HA tag was fused to the N terminus of TMEM63B[207]. Similar to TMEM63A, TMEM63B also exists as

monomers (Fig. A. 1*e* and 1*f*). As such, endogenous TMEM63A and TMEM63B in the brain are monomeric.

Finally, we assessed the oligomeric state of TMEM63s *in situ* in the cell surface membrane. Overexpressed TMEM63s or OSCA1.2 with a C-terminal GFP tag in Chinese hamster ovary (CHO) cells was subjected to single-molecule analysis using total internal reflection fluorescence (TIRF) microscopy. For TMEM63A-GFP, we observed diffusing fluorescent spots but also a considerable amount of background (Fig. A. 1*g*). In contrast, OSCA1.2-GFP and TMEM63B-GFP showed less background (Fig. A. 1*g*), allowing more reliable single-molecule analysis. We observed that many fluorescent spots of OSCA1.2-GFP exhibited two-step photobleaching, whereas TMEM63A-GFP or TMEM63B-GFP exclusively showed only a single bleaching step (Fig. A. 1*h*), indicating the monomeric nature of TMEM63s. We then assessed the oligomeric states of OSCA1.2-GFP and TMEM63B-GFP by quantifying distribution of their fluorescence intensities before and after photobleaching. For OSCA1.2-GFP, the intensity histograms before photobleaching showed a peak for a single GFP, but with a pronounced shoulder at about twice the intensity (Fig. A. 1*i*). After 7.5 s, the shoulder disappeared, supporting the notion that a majority of the GFP tags had bleached and one tag at most was left (spots with both tags bleached are invisible and therefore do not contribute to the histogram) (Fig. A. 1*i*). For TMEM63B-GFP, the shoulder at higher intensities was not visible (Fig. A. 1*i*). These observations suggested that OSCA1.2

forms dimers, but TMEM63B remains monomeric. We made similar observations for control proteins containing a monomeric TMD with either one or two C-terminal GFP tags[208] (Fig. A. 8*c*). Although the mean intensities of the spots from OSCA1.2-GFP or the control protein with two GFP tags decreased by factors of 1.48 ± 0.05 (n = 10 cells) and 1.53 ± 0.04 (n = 9), respectively, the mean intensities from TMEM63B-GFP and the control with one GFP both decreased by 1.27 ± 0.02 (n = 12, n = 9) (Fig. A. 1*j*). Taken together, these data support that TMEM63s are monomers in living cells, whereas OSCA1.2 forms a dimer.

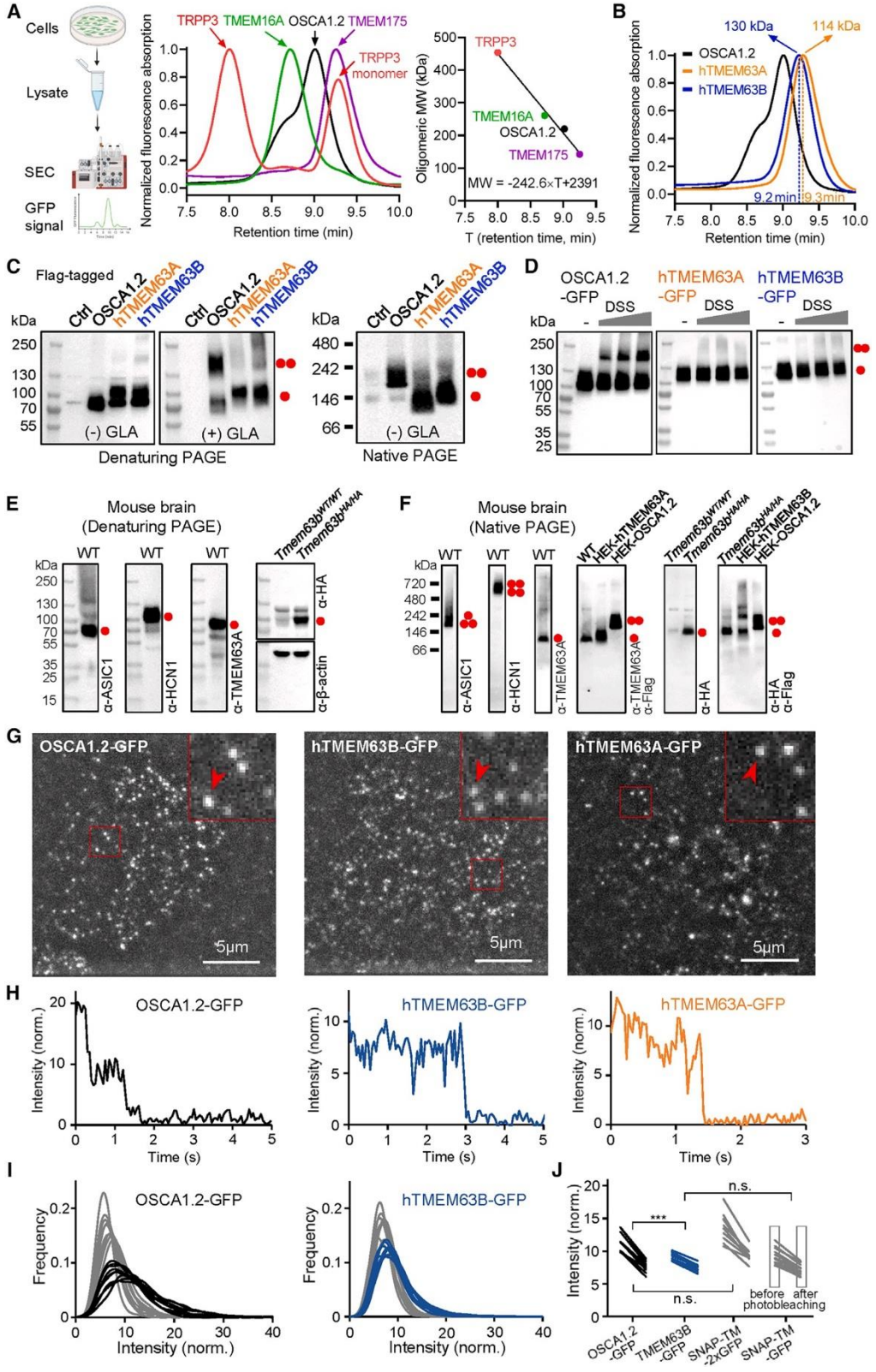

A
Cells
Lysate
SEC
GFP signal
TRPP3 TMEM16A OSCA1.2 TMEM175
TRPP3 monomer
Normalized fluorescence absorption
Retention time (min)
Oligomeric MW (kDa)
MW = -242.6×T+2391
T (retention time, min)
B
130 kDa
114 kDa
OSCA1.2
hTMEM63A
hTMEM63B
9.2 min
9.3 min
C
Flag-tagged
Ctrl OSCA1.2 hTMEM63A hTMEM63B
(-) GLA
(+) GLA
Denaturing PAGE
Native PAGE
D
OSCA1.2 -GFP
hTMEM63A -GFP
hTMEM63B -GFP
DSS
E
Mouse brain (Denaturing PAGE)
α-ASIC1
α-HCN1
α-TMEM63A
α-HA
α-β-actin
F
Mouse brain (Native PAGE)
HEK-hTMEM63A
HEK-OSCA1.2
HEK-hTMEM63B
α-Flag
G
OSCA1.2-GFP
hTMEM63B-GFP
hTMEM63A-GFP
5μm
H
Intensity (norm.)
Time (s)
I
Frequency
Intensity (norm.)
J
n.s.
***
before photobleaching
SNAP-TM -2xGFP
SNAP-TM -GFP

Figure A.1. TMEM63A and TMEM63B exist as monomers

(*a*) *Left*: schematic diagram of FSEC; *middle*, normalized FSEC traces of GFP-tagged membrane proteins; *right*: oligomeric MW (with GFP) versus retention time fit deduced from middle panel data. (*b*) Normalized FSEC traces of OSCA1.2, human TMEM63A, and TMEM63B. Calculated oligomeric MWs (without GFP) are shown. (*c*) Representative denaturing and blue native PAGE of HEK293T cell lysates without (control) or with indicated protein overexpression. The cell lysate was treated without (-) or with (+) 10 mM glutaraldehyde (GLA). Monomeric and dimeric protein bands are indicated. (*d*) Representative western blots of HEK293T cell lysates with overexpression of GFP-tagged proteins. Cells were treated without or with membrane permeable disuccinimidyl suberate (DSS) (in 0.0312, 0.0625, and 0.125 mM) before lysate preparation. (*e*) Representative denaturing PAGEs of whole brain tissue of $Tmem63b^{WT/WT}$ (WT) or $Tmem63b^{HA/HA}$ knockin mice. (*f*) Representative blue native PAGEs of membrane fraction of mouse brain tissues. HEK293T cell lysates with overexpression of FLAG-tagged proteins were also included. (*g*) Single-molecule images from the first frames of CHO cells expressing OSCA1.2-GFP, hTMEM63B-GFP, and hTMEM63A-GFP. (*h*) Intensity time courses of marked spots from (*g*). (*i*) Intensity distribution before (black, blue) and after (gray) photobleaching for OSCA1.2-GFP and hTMEM63B-GFP. (*j*) Mean intensity values before and after photobleaching for OSCA1.2-GFP, hTMEM63B-GFP, and the control proteins containing one or two GFP tags. ***$p < 0.001$. See also Fig. A. 8.

Structures of TMEM63A and TMEM63B

To understand the monomeric nature of TMEM63s, we determined cryo-EM structures of TMEM63A (3.8 Å) and TMEM63B (3.6 Å) in lipid nanodiscs and lauryl neopentyl glycol (LMNG) / cholesteryl hemisuccinate (CHS), respectively (Fig. A. 9 and 10; Table A. 1). Purified TMEM63A and TMEM63B proteins are monomeric (Fig. A. 8*d*) and are glycosylated, as confirmed by PNGase F treatment (Fig. A. 9*b* and 10*b*). The high-resolution maps of the TMDs of both proteins allowed us to build an atomistic model unambiguously (Fig. A. 9*j*, 10*j*, and 11). We were also able to identify extra densities of conserved N-linked glycosylation motifs in both TMEM63A (^{38}NST^{40} and ^{450}NVT^{452}) and TMEM63B (^{462}NVT^{464}) (Fig. A. 11). In contrast, only secondary structures could be distinguished from the relatively low-resolution density map of cytosolic IL2 (Fig. A. 9*j*, 10*j*, and 11), indicating its flexibility or multiple conformational states. By docking the AF2[209] model of the IL2 domain into cryo-EM densities, we were able to build complete models of TMEM63A and TMEM63B.

Different from the dimeric assembly of OSCAs[192–195], both TMEM63A and TMEM63B (hereafter referred as TMEM63s due to their similar structures) were monomers without any symmetry (Fig. A. 2*a*-*g*). The TMD of TMEM63s contains 11 TM helices with an overall arrangement resembling OSCAs, TMEM16s, and TMC1 (Fig. A. 2*c*, 2*e*, and 12*a*). The ICD of TMEM63s is mainly composed of the IL2 domain that contains a four-stranded anti-parallel β-sheet (IL2b1–IL2b4) and five α helices (IL2H1–IL2H5) (Fig. A. 2*c*). Notably, IL2H2 and IL2H3, two long amphipathic helices

shared by TMEM63s and OSCAs, are parallel to the membrane with an intervening loop, known as the ‘‘hook’’ domain, inserted into the membrane. These features may allow IL2 to sense membrane tension.

Additionally, non-protein densities within a cavity of TMDs of TMEM63s were observed and tentatively assigned as lipids (Fig. A. 2*a* and 2*b*). Given the critical roles of lipids in regulation/gating of mechanosensitive channels such as TRAAK, Piezo1, and MscS[210,211], it is possible that lipids may also affect functional aspects of TMEM63s.

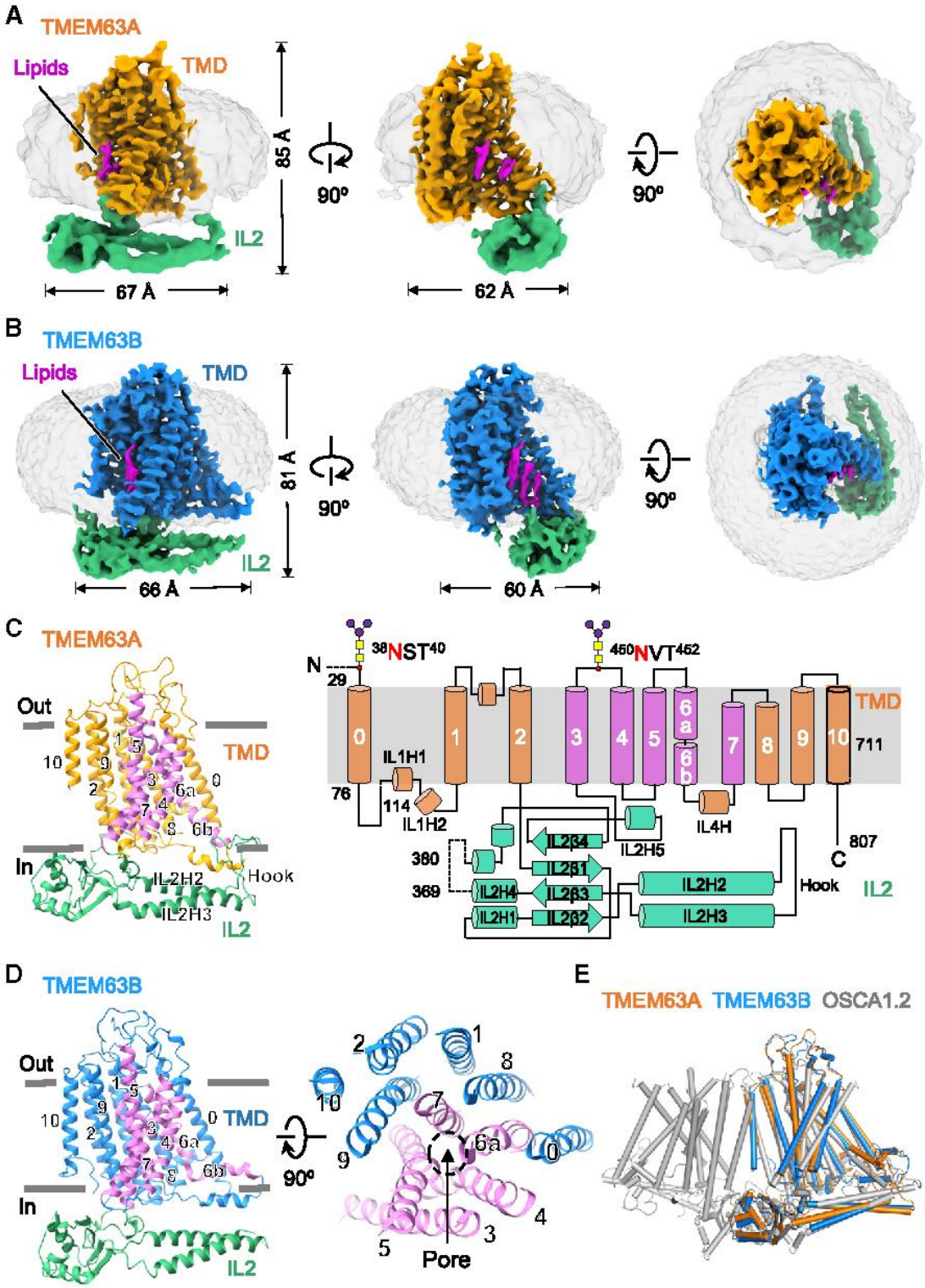

A
TMEM63A
Lipids
TMD
IL2
85 Å
67 Å
90°
62 Å
90°
B
TMEM63B
Lipids
TMD
IL2
81 Å
66 Å
90°
60 Å
90°
C
TMEM63A
Out
In
TMD
IL2
IL2H2
IL2H3
Hook
N
29
38NST40
450NVT452
76
114
IL1H1
IL1H2
IL2H5
IL4H
IL2β4
IL2β1
IL2β3
IL2β2
IL2H4
IL2H1
360
369
807
711
C
Hook
IL2
D
TMEM63B
Out
In
TMD
IL2
90°
Pore
E
TMEM63A TMEM63B OSCA1.2

Figure A.2. Cryo-EM structures of human TMEM63A and TMEM63B

(*a* and *b*) Side and top views of EM density maps of TMEM63A in nanodisc (*a*) and TMEM63B in detergent LMNG / CHS (*b*). Densities for the transmembrane domain (TMD) are colored in either orange (TMEM63A) or blue (TMEM63B). Densities of intracellular linker 2 (IL2) are colored green. Densities of nanodisc and detergent are gray. Lipid densities are highlighted in magenta. (*c*) *Left*: side view of TMEM63A in ribbon representation, with the pore-lining helices (TM3–TM7) colored in magenta. *Right*: topology diagram of TMEM63A with major structural components highlighted. Dashed lines denote regions excluded in model building due to missing density. Two N-linked glycosylation sites were predicted based on the protruding densities near N38 and N450 (Fig. A. 11). (*d*) *Left*: side view of TMEM63B with the pore-lining helices (TM3–TM7) colored in magenta. *Right*: top view of the TM segments of TMEM63B with the location of the pore indicated. (*e*) Structural alignments of monomeric human TMEM63A and TMEM63B with one subunit of dimeric OSCA1.2 (PDB: 6MGV). See also Fig. A. 9–11 and 15 and Table A. 1.

Ion permeation pores of TMEM63s

Like OSCAs, TMEM16s, and TMC1, we identified an ion permeation pore in TMEM63s at a similar position formed by TM3–TM7 (Fig. A. 3*a* and 12*a*) using the HOLE program[212]. Compared with OSCA1.2, the putative pore in TMEM63s exhibited two remark- able features. First, the TMEM63 pore contains a drastically elongated constriction region, known as the ‘‘neck,’’ which extends over 50 Å (Fig. A. 3.*a*, *b*). Comparison of pore domains between OSCA1.2 and TMEM63s revealed noticeable differences in the pore-lining TM4 and TM6 (Fig. A. 13*b*). Although an open pore conformation has not been resolved, it is highly likely that TMEM63s possess a more constricted open pore than OSCA1.2. Second, an N-terminal domain (NTD) in TMEM63s (S30-V45 in TMEM63A and P19-L36 in TMEM63B), absent in OSCAs, folds into a hairpin on the external side of the putative pore (Fig. A. 2*a*, 2*b*, 12*c*, and 12*d*). Collectively, these pore features may significantly limit ion flux, leading to low channel conductance for TMEM63s.

In line with the structural observations, electrophysiological recordings revealed small stretch-activated currents in heterologous HEK293T cells expressing TMEM63A (I_{max} = 33.5 ± 7.5 pA; n = 6) or TMEM63B (I_{max} = 17.4 ± 2.3 pA; n = 5), whereas OSCA1.2 were 10-fold larger (I_{max} = 310.2 ± 85.1 pA; n = 8) (Fig. A.3*c* and 3*d*). The TMEM63A currents were dramatically reduced by Gd^{3+} (76% reduction in I_{max}; Fig. A. 8*e*), which excluded the possibility of pressure-induced leak. In addition, robust single-channel currents were evident in OSCA1.2-transfected cells (with a conductance of 129

± 3 pS, see Fig. A. 5*l* and 5*m*) but we were not able to resolve single-channel currents of TMEM63s, which were probably too small to be detected. Of note, although TMEM63s had small current amplitudes, they lacked inactivation (Fig. A. 3*c* and A. 8*f*), which allows them to conduct similar numbers of charges/ions to OSCA1.2 and Piezo1 under sustained mechanical stimuli (Fig. A. 8*g*).

Pore-lining residues in TMEM63s are mostly hydrophobic within the neck region and hydrophilic at either end of the neck (Fig. A. 3*f*). In particular, the acidic residues at both extracellular and intracellular sides (D551 / E571 in TMEM63A and D538 / D564 / D584 in TMEM63B) of the neck may provide an attractive electrostatic environment for cations. Positively charged residues were previously found at an equivalent pore location of TMEM16A to account for a selectivity for anions[197,198]. As most pore-lining residues in the neck region are located in TM4 and TM6 (Fig. A. 3*f*), the pore opening of TMEM63s probably involves movement of TM4, TM6, or both. Like OSCA1.2, TMEM63s possess a pair of π helices in TM5 and TM6, respectively (Fig. A. 3*f*). The π helix is energetically unstable, and the transition between π helix and α helix in pore-lining helices has been implicated in gating of many ion channels, including TRPs and TMEM16A[198,213]. To experimentally investigate the channel pore, we replaced the putative pore-lining acidic residues in TM6 of TMEM63A (E571) and TMEM63B (D584) with positively charged lysines. These substitutions had minimal effect on protein surface expression (Fig. A. 8*h*) but dramatically reduced the maximal stretch-activated currents for both TMEM63A and TMEM63B (Fig. A. 3*g*), suggesting involvement of

this acidic residue in cation permeation in TMEM63s. A similar substitution in OSCA1.2 (E531K) also drastically reduced single-channel current (see Fig. A. 16*f*), consistent with a previous claim that E531 contributes to the ion permeation pathway of OSCA1.2[192].

To further characterize the ion permeation pore of TMEM63s, we carried out equilibrium (0 mV) and non-equilibrium (-500 mV) all-atom molecular dynamics (MD) simulations[132] on models based on our structure of TMEM63B (Fig. A. 12*e*), on an AF2 model of TMEM63B, and on the structure of OSCA1.2 (Table A. 2). Our structural model of TMEM63B was simulated with or without the NTD (Fig. A. 12*f*) and incorporated missing regions from AF2 prediction, including two conformations (C1 and C2) for the first intracellular linker (IL1) region (Fig. A. 12*g*). The AF2 model contains two loops connecting TM3/TM4 and TM5/TM6, respectively, that show prominently different conformations from our experimentally derived models (Fig. A. 12*h*). Equilibrium simulations (200 ns each) revealed a stable TMD for all systems simulated (root-mean-square deviation [RMSD] < 2 Å; Fig. A. 13), whereas IL2 domains were more dynamic. Analysis of water density throughout equilibrium simulations did not reveal persistent hydration of the putative pore in TMEM63B models, regardless of the presence of NTD, IL1 conformation, or membrane composition (Fig. A. 3*h* and A. 14). Nevertheless, we were able to identify potential permeation pathways that match those visualized with the HOLE program using the static structures (Fig. A. 3*i*). Transient hydration was more readily observed for the AF2 model of TMEM63B and for the OSCA1.2 channel (Fig. A. 3*i*). At -500 mV, we did not observe any ion permeation

events over 100 ns of simulation for each system. A pore with a conductance of ~120 pS, as expected for OSCA1.2, should conduct up to ~ 37 ions under the simulated conditions. These data supported the fact that our resolved structures of TMEM63s, as well as the published OSCA1.2 model[192], represent closed states and that the pore of TMEM63s is substantially constricted and of low conductance.

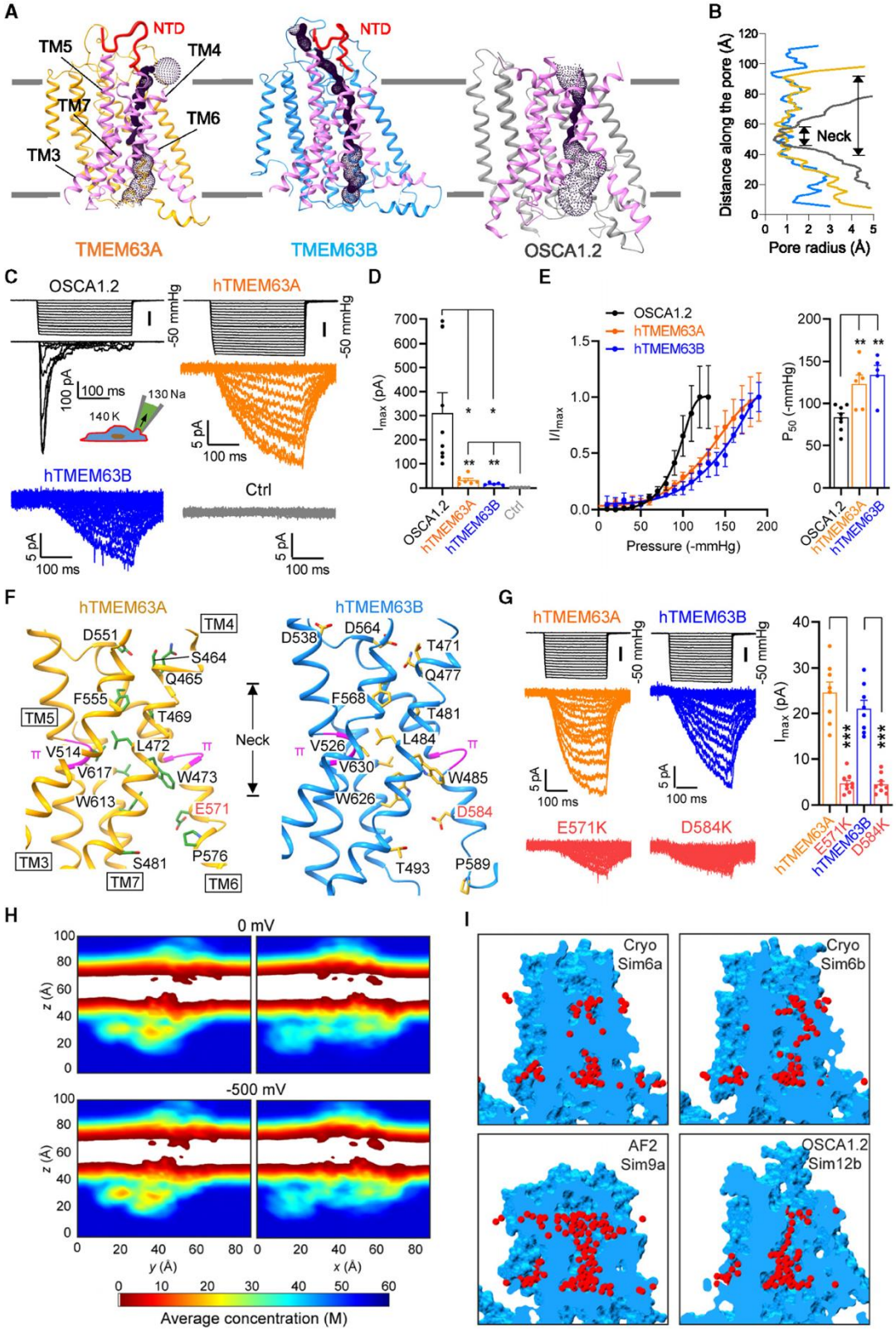

A
NTD
TM5
TM4
TM7
TM6
TM3
TMEM63A
TMEM63B
OSCA1.2
B
Distance along the pore (Å)
Neck
Pore radius (Å)
C
OSCA1.2
hTMEM63A
hTMEM63B
Ctrl
-50 mmHg
100 pA
100 ms
130 Na
140 K
5 pA
D
I_{max} (pA)
E
OSCA1.2
hTMEM63A
hTMEM63B
I/I_{max}
Pressure (-mmHg)
P_{50} (-mmHg)
F
hTMEM63A
hTMEM63B
D551
S464
Q465
F555
T469
V514
L472
V617
W473
W613
E571
S481
P576
D538
D564
T471
Q477
F568
T481
L484
V526
V630
W485
W626
D584
T493
P589
Neck
G
E571K
D584K
I_{max} (pA)
H
0 mV
-500 mV
z (Å)
y (Å)
x (Å)
Average concentration (M)
I
Cryo Sim6a
Cryo Sim6b
AF2 Sim9a
OSCA1.2 Sim12b

Figure A.3. Ion permeation pores in TMEM63s

(*a*) Predicted location of the pore in TMEM63A, TMEM63B, and OSCA1.2 (PDB: 6MGV). Pore-lining TMs are colored in magenta and N-terminal domains (NTDs) in TMEM63A and TMEM63B are accentuated by thickening line and colored in red. (*b*) van der Waals pore radii against distance along the pore, with neck regions indicated. (*c*) Representative traces of stretch-activated currents (-80 mV) from HEK293T cells without (control) or with indicated protein overexpression. A diagram is shown to indicate main cations (mM) used in the bath and pipette solutions. (*d*) Quantification of mean ± SEM maximal current amplitude (I_{max}) from individual patches. $^{*}p < 0.05$, $^{**}p < 0.01$. (*e*) *Left*: mean ± SEM and normalized pressure-response current curves fitted with a Boltzmann equation. OSCA1.2 (n = 8), hTMEM63A (n = 6), and hTMEM63B (n = 5). *Right*: mean SEM P_{50} values. **p < 0.01. (*f*) Pore-lining residues in TMEM63A (left) and TMEM63B (right) from HOLE analysis. (*g*) *Left*: representative traces of stretch-activated currents from HEK293T cells expressing hTMEM63A, hTMEM63B, or indicated mutants; *right*: quantification of I_{max}. $^{***}p < 0.001$. (*h*) Two-dimensional slices showing time- and space-averaged water concentrations for different all-atom MD simulations (Sim1a, top; Sim1b, bottom; see also Fig. A. 14). (*i*) Examples of transient water channels formed in simulations of our experimentally derived TMEM63B model (Sim6a and Sim6b), of an AF2 TMEM63B model (Sim9a), and of an experimentally derived model of OSCA1.2. See also Fig. A. 12–13 and Table A. 2.

TMEM63s form high-threshold mechanosensitive ion channels

Stretch-activated currents evoked by TMEM63 expression in HEK293T cells exhibited not only small amplitudes but also high thresholds. As shown in Fig. A. 3*e*, the pressure required for half-maximal activation (P_{50}) of human TMEM63A (-123.3 ± 11.2 mmHg; n = 6) or TMEM63B (-134.2 ± 11.4 mmHg; n = 5) was significantly larger than that for OSCA1.2 (-85.9 ± 5.2 mmHg; n = 8). To examine whether monomeric TMEM63s are bona fide pore-forming mechanosensitive ion channels with high thresholds, we performed electrophysiological recordings in proteoliposomes (Fig. A. 4*a*). Reconstitution of purified OSCA1.2 proteins into liposomes induced robust macroscopic stretch-activated currents (I_{max} = 155.4 ± 20.9 pA; n = 10) (Fig. A. 4*b* and 4*c*), confirming that OSCA1.2 is an intrinsically mechanosensitive ion channel[189]. Single-channel current recordings revealed a conductance of 133 pS (Fig. A. 8*i*), consistent with measurements from HEK293T cells. As a negative control, little current was detected in the empty liposomes (I_{max} = 2.6 ± 0.2 pA; n = 5) (Fig. A. 4*b* and 4*c*). Using the established protocol, we next detected small but significant stretch-activated currents for both TMEM63A (I_{max} = 18.7 ± 3.1 pA; n = 6) and TMEM63B (I_{max} = 16.3 ± 1.8 pA; n = 6) (Fig. A. 4*b* and 4*c*). No single-channel current was resolved, consistent with our assumption of small conductance for TMEM63s. Akin to recordings in HEK293T cells, the P_{50} of TMEM63A (-142.6 ± 9.1 mmHg; n = 6) or TMEM63B (-143.9 ± 7.3 mmHg; n = 6) was substantially higher than OSCA1.2 (-86.9 ± 6.4 mmHg; n = 10) (Fig. A. 4*d*), supporting the characteristic high-threshold mechanosensitivity for monomeric

TMEM63s. Furthermore, TMEM63s from other species, including mouse and fruit fly, also adopted a monomeric configuration (Fig. A. 8*j* and 8*k*) and exhibited high-threshold mechanosensitivities in HEK293T cells (Fig. A. 8*l* and 8*m*).

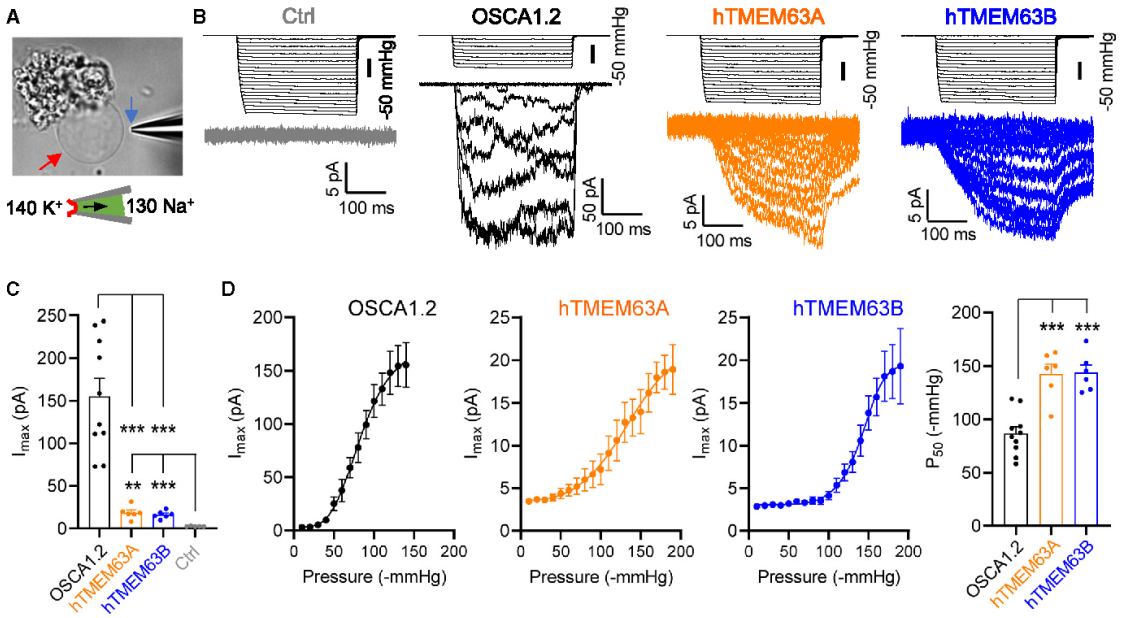


Figure A.4. TMEM63A and TMEM63B form high-threshold mechanosensitive ion channels

(*a*) Representative image showing a proteoliposome, indicated by a red arrow, grown from the lipid cloud. A recording pipette, indicated by a blue arrow, is also shown for excised patch recording. (*b*) Representative traces of stretch-activated currents from unilamellar liposomes reconstituted without (control) or with indicated purified proteins. (*c*) Quantification of mean ± SEM I_{max} for OSCA1.2 (n = 10), hTMEM63A (n = 6), hTMEM63B (n = 6), or control (n = 5). $^{**}p < 0.01$, $^{***}p < 0.001$. (*d*) Mean ± SEM pressure-response current curves fitted with a Boltzmann equation for OSCA1.2 (n = 10), hTMEM63A (n = 6), and hTMEM63B (n = 6). The right panel shows quantification of mean ± SEM P_{50} values. $^{***}p < 0.001$. See also Fig. A. 8.

Oligomeric configuration affects mechanosensitivity of OSCA/TMEM63s

We hypothesized that the differential mechanosensitivity between TMEM63s and OSCA1.2 is a result of their distinct oligomeric configuration. To examine this hypothesis, we first sought to unveil the structural basis underlying their distinct oligomerization. In OSCA1.2, the dimerization is solely mediated by the IL2 domain (Fig. A. 5*a* and 5*b*). Sequence alignments revealed that residues mediating interactions within the dimerization interface of OSCA1.2 are variable among TMEM63s (Fig. A. 5*c* and 15). Moreover, TMEM63s contain an additional inserted fragment of 20 amino acids (aa) in the middle of dimerization region (Fig. A. 5*c*). This 20-aa fragment was not resolved in our structures and probably forms a flexible loop. Conceivably, this 20-aa loop might clash with the dimerization interface.

To verify the importance of IL2 in protein oligomerization, we first generated a chimeric protein, OSCA1.2$_{63A\ IL2}$, in which the IL2 domain of OSCA1.2 (S194-P367) was replaced with that of TMEM63A (N224-Q414) (Fig. A. 5*d*). FSEC and cross-linking assays revealed a monomeric configuration of OSCA1.2$_{63A\ IL2}$ (Fig. A. 5*e* and 5*f*), suggesting that the IL2 domain is a key determinant of dimerization. Of note, dimeric TMEM63A was not achieved by replacement with OSCA1.2 IL2 or further with the OSCA1.2 C terminus (Fig. A. 16*a* and 16*b*); therefore, OSCA1.2 IL2 itself is not sufficient to induce dimerization. Furthermore, replacements of 5 dimerization interfacial residues in OSCA1.2 (W331G, V335G, Q338G, T339G, and R343A) lead to a monomeric mutant OSCA1.2$_{5Mu}$ (containing all five mutations)

(Fig. A. 5*g* and 16*c*). Additionally, engineering the 20-aa inserted loop of TMEM63s into OSCA1.2 led to a monomeric configuration (Fig. A. 5*h*). Collectively, these data support that genetic variation in the IL2 domain led to distinct oligomerization status between OSCAs and TMEM63s.

To examine whether the oligomeric configuration affects the mechanosensitivity of OSCA/TMEM63s, we functionally characterized the OSCA1.2 monomeric variants, OSCA1.2$_{5Mu}$ and OSCA1.2$_{63A\ IL2}$. In HEK293T cells, OSCA1.2$_{5Mu}$ expression evoked robust stretch-activated currents (I_{max} = 171.6 ± 25.5 pA; n = 12) with rapid activation and fast inactivation, akin to OSCA1.2 (Fig. A. 5*i* and 5*j*), demonstrating that monomeric OSCA1.2 preserves mechanosensitive channel function. Notably, compared with OSCA1.2, OSCA1.2$_{5Mu}$ showed a substantially higher P_{50} (-109.8 ± 4.1 mmHg; n = 12, Fig. A. 5*k*), suggesting that switching from a dimer to monomer leads to high-threshold mechanosensitivity. Similarly, monomeric OSCA1.2$_{63A\ IL2}$ expression also induced high-threshold (P_{50} = -133.1 ± 7.7 mmHg; n = 6) stretch-activated currents (I_{max} = 120.8 ± 33.2 pA; n = 6) (Fig. A. 5*i*–5*k*). Unlike OSCA1.2$_{5Mu}$, OSCA1.2$_{63A\ IL2}$ showed little inactivation (Fig. A. 5*i*), even with prolonged application of pressures (Fig. A. 8*f*). Of note, OSCA1.2$_{5Mu}$ and OSCA1.2$_{63A\ IL2}$ exhibited smaller I_{max} than OSCA1.2 (Fig. A. 5*j*), at least partially due to lower total and cell surface protein expression (Fig. A. 16*d*). Together, these data demonstrate the influence of oligomeric status on mechanosensitivity of OSCA/TMEM63 channels.

We next examined the influence of the oligomeric configuration on single-channel behavior. In HEK293T cells, OSCA1.2 induced two types of stretch-activated unitary currents, a dominant fully open current with a conductance of 129 ± 3 pS (n = 4) and a transient smaller current with approximately half amplitude (Fig. A. 5*l* and 5m). This transient current was suggested to be the single- pore current, and the fully open current was thought to arise from the simultaneous opening of two pores[189,192]. Thus, the two subunits in OSCA1.2 were assumed to function cooperatively. However, our recordings of monomeric OSCA1.2$_{5Mu}$ showed a similar fully open current with a conductance of 126 ± 5 pS (n = 4) and a small current of half amplitude (Fig. A. 5*l* and 5*m*), which suggests that the fully open current observed in OSCA1.2 is from a single pore and two subunits in OSCA1.2 function in a largely independent manner. The half-amplitude current likely represents a sub-conductance state, i.e., a partial open state, of a single pore. This presumption of independent gating of two OSCA1.2 subunits is further supported by single-channel recordings of a OSCA1.2 concatemer, containing two protomers of different conductance in tandem (E531K mutant and wild type [WT]) (Fig. A. 16*e*). This OSCA1.2 concatemer induced a mix of two levels of unitary current of WT (11 pA) and E531K mutant (3 pA), instead of one averaged level (7 pA) (Fig. A. 16*f*). Our recordings from monomeric OSCA1.2$_{63A\ IL2}$ only showed currents resembling the assumed sub-conductance currents, with a conductance of 57 ± 5 pS (n = 5, Fig. A. 5*l* and 5*m*). One possible explanation is that the IL2 domain from TMEM63A

only induces activation of the OSCA1.2 pore to a partially open state. Comparison of current traces between OSCA1.2 and OSCA1.2$_{5Mu}$ revealed that the mean open-dwell time of the pore in OSCA1.2 (t = 3.5 ± 0.3 ms; n = 4) is longer than that in OSCA1.2$_{5Mu}$ (t = 1.8 ± 0.2 ms; n = 4) (Fig. A. 5*l* and 16*g*). Therefore, the oligomeric status of OSCA/ TMEM63s barely affects the pore conductance but has an influence on the pore open-dwell time.

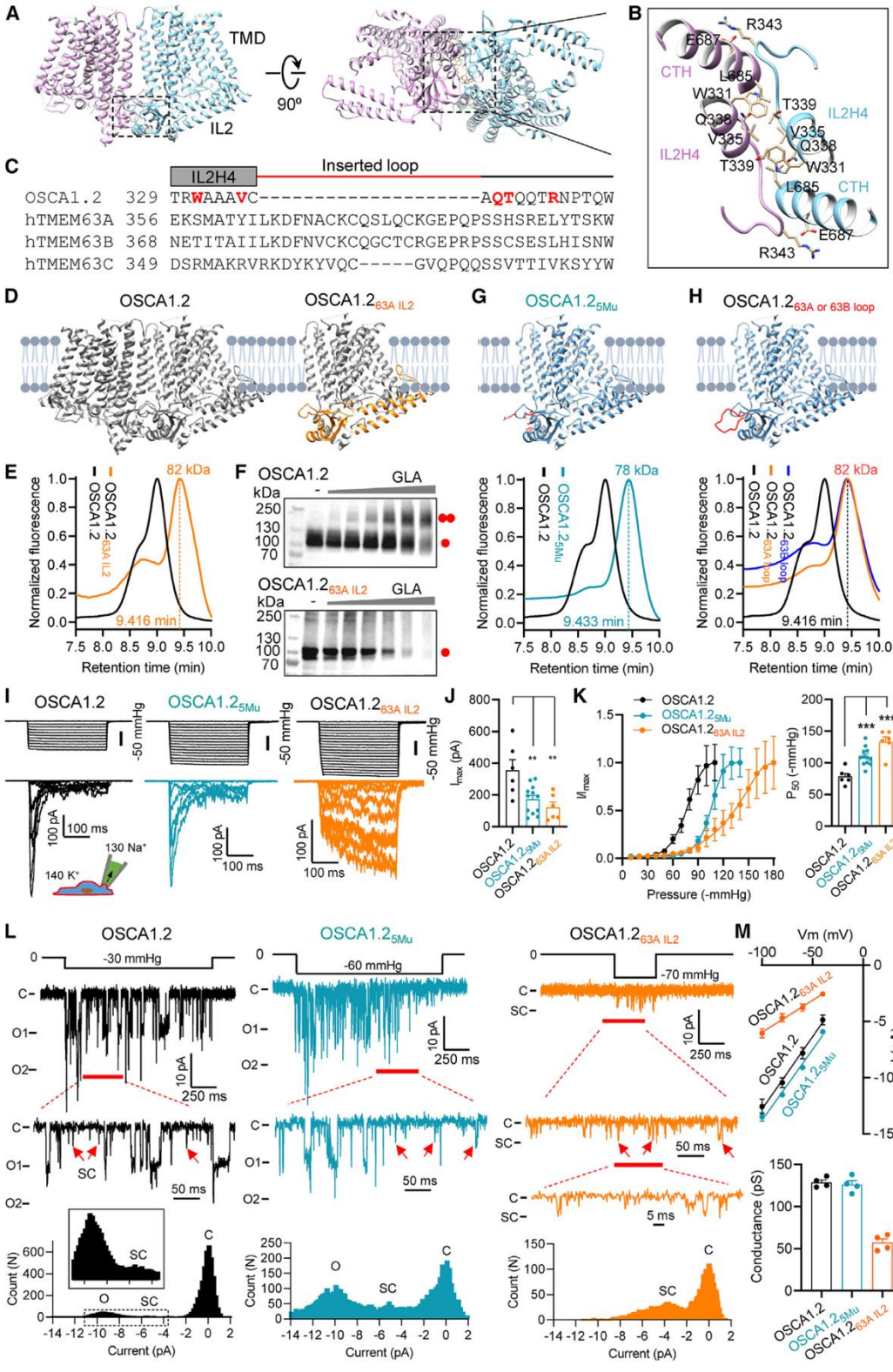

A
TMD
IL2
90º
B
R343
E687
CTH
L685
W331
T339
IL2H4
Q338
V335
C
IL2H4
Inserted loop
OSCA1.2 329 TRWAAAVC----------AQTQQTRNPTQW
hTMEM63A 356 EKSMATYILKDFNACKCQSLQCKGEPQPSSHSRELYTSKW
hTMEM63B 368 NETITAIILKDFNVCKCQGCTCRGEPRPSSCSESLHISNW
hTMEM63C 349 DSRMAKRVRKDYKYVQC-----GVQPQQSSVTTIVKSYYW
D
OSCA1.2
OSCA1.2 63A IL2
G
OSCA1.2 5Mu
H
OSCA1.2 63A or 63B loop
E
82 kDa
9.416 min
Normalized fluorescence
Retention time (min)
F
OSCA1.2
GLA
kDa
250
130
100
70
OSCA1.2 63A IL2
78 kDa
9.433 min
I
OSCA1.2
OSCA1.2 5Mu
OSCA1.2 63A IL2
-50 mmHg
100 pA
100 ms
130 Na+
140 K+
J
K
Pressure (-mmHg)
L
OSCA1.2
OSCA1.2 5Mu
OSCA1.2 63A IL2
-30 mmHg
-60 mmHg
-70 mmHg
C
O1
O2
SC
10 pA
250 ms
50 ms
5 ms
Count (N)
Current (pA)
M
Vm (mV)
I (pA)
Conductance (pS)

Figure A.5. The oligomeric configuration affects mechanosensitivity of OSCA/TMEM63s

(*a*) Side and bottom views of dimeric OSCA1.2 (PDB: 6MGV). The rectangle indicates the dimerization interface in IL2. (*b*) Enlarged dimerization interface in OSCA1.2, with interacting residues shown. CTH, C-terminal helix (L685-R698). (*c*) Sequence alignment of the region within IL2 that mediates the dimerization of OSCA1.2 (UniProt accession numbers of sequences used: OSCA1.2: Q5XEZ5; hTMEM63A: O94886; hTMEM63B: Q5T3F8; hTMEM63C: Q9P1W3). Residues that directly mediate OSCA1.2 dimerization are highlighted in red. The inserted loop in TMEM63s is indicated by the red line. (*d*) Illustrations of dimeric OSCA1.2 and monomeric chimera protein $OSCA1.2_{63A\ IL2}$. (*e*) Normalized FSEC traces of GFP-tagged OSCA1.2 and $OSCA1.2_{63A\ IL2}$, with MW of the $OSCA1.2_{63A\ IL2}$ (without GFP) indicated. (*f*) Representative western blots of GFP-tagged OSCA1.2 (*upper*) or $OSCA1.2_{63A\ IL2}$ (*lower*). Cell lysates were treated without or with GLA at various concentrations (in 0.3125, 0.625, 1.25, 2.5, 5, and 10 mM). (*g*) Illustration of monomeric $OSCA1.2_{5Mu}$ mutant containing 5 mutations, W331G, V335G, Q338G, T339G, R343A (*upper*), and normalized FSEC traces (*lower*). (*h*) Illustration of monomeric OSCA1.2 with inserted 20-aa loop from hTMEM63A ($OSCA1.2_{63A\ loop}$) or hTMEM63B ($OSCA1.2_{63B\ loop}$) (*upper*) and normalized FSEC traces (*lower*). (*i* and *j*) Representative traces of stretch-activated currents recorded from HEK293T cells (I) and quantification of mean ± SEM I_{max} (*j*). $^{**}p < 0.01$. (*k*) *Left*: mean ± SEM and normalized pressure-response current curves fitted with a Boltzmann equation for OSCA1.2 (n = 6), $OSCA1.2_{5Mu}$ (n = 12), $OSCA1.2_{63A\ IL2}$ (n = 6); *right*: quantification of mean ± SEM P_{50} from individual patches. $^{***}p < 0.001$. (*l*) Single-channel currents recorded at -80 mV. The closed (C) and fully open (O) states are indicated. Red arrows mark the assumed sub-conductance (SC) state. Current amplitude histogram for each trace is shown at the bottom. (*m*) *Upper*: averaged current-voltage relationship. OSCA1.2 (n = 4), $OSCA1.2_{5Mu}$ (n = 4), $OSCA1.2_{63A\ IL2}$ (n = 4); *lower*: mean ± SEM single-channel conductance from individual patches. See also Fig. A.16.

Oligomeric configuration affects conformation of IL2 and TM6

To mechanistically understand how the oligomeric configuration affects mechanosensitivity of OSCA / TMEM63, we compared structures of monomeric TMEM63s and dimeric OSCA1.2. Although the TMD exhibited large similarity between TMEM63s and OSCA1.2, we observed prominent conformational differences of the IL2 domain, which exhibited more intimate contact with the membrane in OSCA1.2 (Fig. A. 6*a*). Isolated IL2 domains were superimposed between TMEM63s and OSCA1.2; thus, we speculated that dimerization of OSCA1.2 induces a rigid-body movement of the IL2 domain toward the membrane (Fig. A. 6*a*) and stabilizes the IL2 domain in a conformation that allows membrane association of the hypothetic mechano-sensing elements, including amphipathic helices IL2H2 / IL2H3 and the hook domain (Fig. A. 6*a*). We deleted the hook domain in OSCA1.2 (Q272-K287), TMEM63A (L298-W317), and TMEM63B (M310-Q329 and found that the deletion had little effect on the mechanosensitivity of these channels (Fig. A. 16*h*), demonstrating that the hook domain is dispensable for mechano-sensation. We thus propose that the amphipathic helices IL2H2 / IL2H3 may serve as the mechanosensor, given the prevalence of the amphipathic helix in other mechanosensitive ion channels[172,200]. Compared with OSCA1.2, the IL2 domain in TMEM63s is largely mobile (Fig. A. 13) and mainly adopts a conformation in which IL2H2 / IL2H3 is farther from the internal membrane surface (Fig. A. 6*a*), which could confer high-threshold mechanosensitivity.

The pore opening (gating) of a mechanosensitive ion channel involves both sensation of mechanical force and transduction of the force to open the pore. In the Ca^{2+}-activated chloride channel TMEM16A, the pore-lining TM6 directly binds with Ca^{2+} and undergoes a dramatic rearrangement[197,198] (Fig. A. 6*b*). Thus, TM6 was proposed to act as a gating helix to bridge Ca^{2+} binding and pore opening in TMEM16A[214]. Interestingly, a similar conformational rearrangement in TM6 was observed in the structural alignment of TMEM63A, TMEM63B, and OSCA1.2 (Fig. A. 6*b*). The cytoplasmic half of TM6, named TM6b, in dimeric OSCA1.2 is oriented to establish stable physical contact with IL2H2 (Fig. A. 6*c*), whereas such a physical link was not observed in TMEM63s (Fig. A. 6*c*). This conformation of TM6b in OSCA1.2 places it in a favorable position to transduce force sensed by IL2H2 / IL2H3 to promote pore opening, whereas in TMEM63A and TMEM63B, pore opening may require greater force. Interestingly, transient physical contact between TM6b and IL2H2 in TMEM63B was observed in our MD simulation studies (Fig. A. 17), implicating the IL2-TM6 interaction in TMEM63 gating. We propose that these features of monomeric TMEM63s have evolved to fill a functional niche, namely, as high-threshold mechanosensitive ion channels (Fig. A. 6*d*).

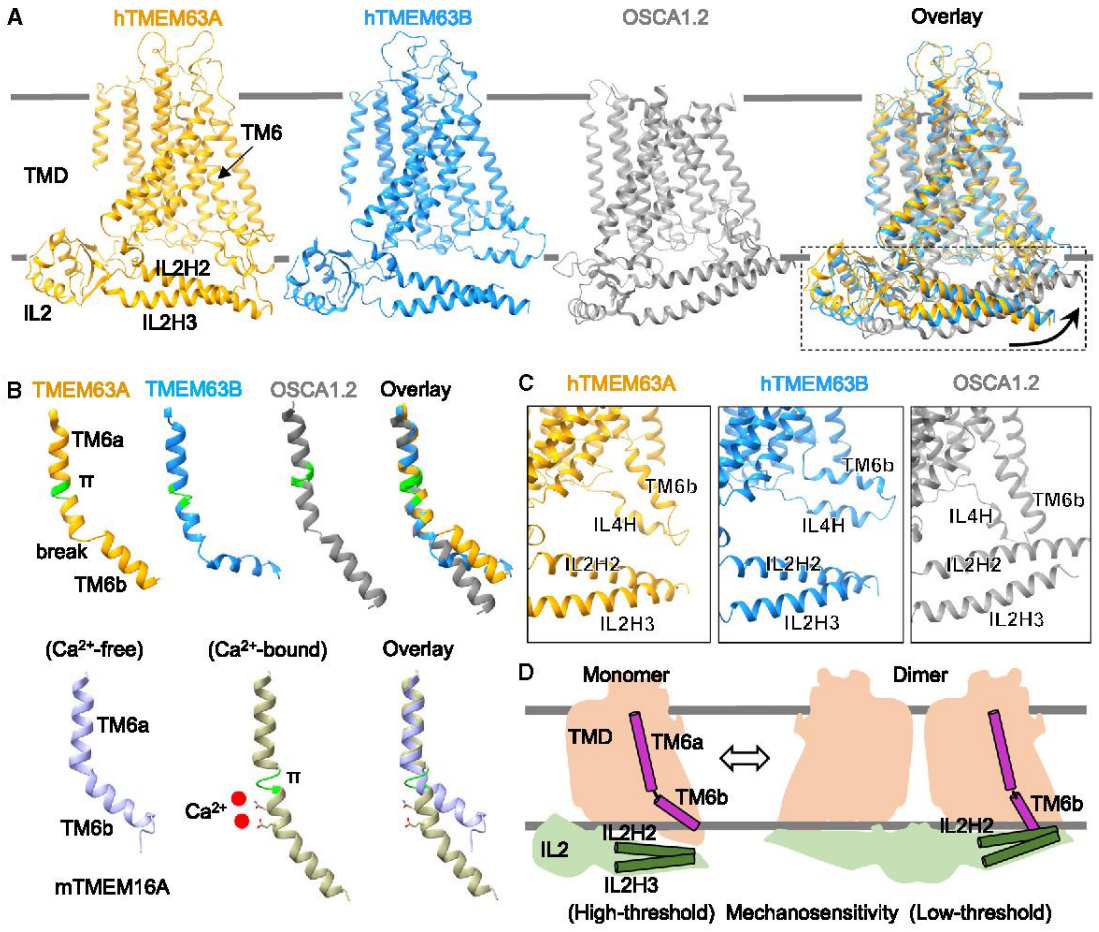


Figure A.6. The oligomeric configuration affects conformations of mechano-gating elements, including IL2 and TM6, in OSCA/TMEM63 family

(*a*) Structural comparisons among hTMEM63A, hTMEM63B, and OSCA1.2 (one subunit, PDB: 6MGV) based on the alignment of TMDs. Arrows indicate the movement of IL2 domain from the conformation in hTMEM63A or hTMEM63B to that in OSCA1.2. (*b*) *Upper*: ribbon representations of TM6 from hTMEM63A, hTMEM63B, and OSCA1.2 (PDB: 6MGV). A π helix in the TM6 is colored in green and a break point in the TM6 is indicated. *Lower*: TM6 from mouse TMEM16A in Ca^{2+}-free (PDB: 5OYG) and Ca^{2+}-bound (PDB: 5OYB) conformations. The π helix and Ca^{2+} ions in the Ca^{2+}-bound state are shown. (*c*) Structural comparisons showing the relative positions of IL2H2 / IL2H3 to TM6b / IL4H in hTMEM63A, hTMEM63B, and OSCA1.2. The hook domain is not shown for clarity. (*d*) Schematic diagram showing distinct conformations of IL2 and TM6 in monomeric and dimeric OSCA/TMEM63 proteins. See also Fig. A. S16 and 17.

DISCUSSION

Like thermosensation in mammals, which is mediated by various TRP channels detecting a wide range of temperatures[215,216], mechanosensation may also involve distinct molecules with sensitivities that span a broad physiological range of mechanical forces encountered in the biological world. To date, Piezo proteins (piezo1 and piezo2)[176,217] and TMC1[66] are the predominant mechanosensitive ion channels identified in mammals with low activation thresholds. We propose that TMEM63s evolved to function as high-threshold mechanosensitive ion channels, complementary to Piezo and TMC1 channels. Human patients with TMEM63B variants exhibit chronic hemolytic anemia[205], characterized by dysmorphic red blood cells (RBCs) due to dehydration. Interestingly, Piezo1 mutations are known to cause similar symptoms in humans with dehydrated hereditary stomatocytosis (DHS)[218]. RBCs are highly deformable cells that experience a broad range of mechanical stimuli when circulating through blood vessels of variable sizes. It is plausible that Piezo1 and TMEM63B may work complementarily in RBCs to sense mild to intense mechanical force, thereby ensuring effective mechanosensory responses of RBCs. TMEM63B has also been suggested to function as an osmosensor in cochlear outer hair cells, and TMEM63B deficiency in mice leads to outer hair cell death and deafness[207]. Outer hair cells act as mechanical actuators capable of amplifying sound signals in the ear through a process known as cochlear amplification[219]. The mechanical actuator protein, prestin, is densely packed in the outer hair cell basolateral membrane and its electromotility activity likely evokes significant

mechanical force within the basolateral membrane of outer hair cells. Therefore, the high-threshold activation of TMEM63B may be beneficial and could allow the channel to remain in a closed state during normal electromotility function of outer hair cells. Opening of TMEM63B would only be triggered when outer hair cells experience severe osmotic stress. Alternatively, it is plausible that prestin activity may generate sufficient force to activate TMEM63B directly. Because prestin is voltage dependent, it is also plausible that TMEM63B activation may provide feedback to modulate prestin activity, outer hair cell electromotility, and cochlear amplification.

Unlike well-established ion channels containing one central pore formed by multiple subunits, OSCA, TMEM16, and TMC channels are dimers with one pore in each subunit (Fig. A. 12*a*). Because each subunit of the dimeric channels contains a complete ion permeation pathway, several fundamental questions arise. Can the monomeric protein function as a channel? If yes, why is dimerization needed? Our data, recorded from TMEM63A and TMEM63B in HEK293T cells and liposomes (Fig. A. 3*c* and 4*b*), demonstrate that monomeric proteins form functional channels, which is further supported by recordings from monomeric OSCA1.2$_{5Mu}$ and OSCA1.2$_{63A\ IL2}$ (Fig. A. 5*i*). Comparison of single-channel currents between OSCA1.2 and monomeric OSCA1.2$_{5Mu}$ (Fig. A. 5*l*) and recordings from a OSCA1.2 concatemer (Fig. A. 16*e* and 16*f*) further suggest that the two subunits in the dimeric OSCA1.2 function in a largely independent manner. Independent gating of

TMEM16A subunits while in a dimeric complex was previously deduced from measurement of Ca^{2+}-induced activation in a fused protein with two subunits of different Ca^{2+} sensitivities in tandem[220,221]. Independent function may extend to other members in this superfamily, including TMCs. Structural comparisons between monomeric TMEM63s and dimeric OSCA1.2 revealed distinct conformations of the mechano-sensing domain IL2 and the gating helix TM6 (Fig. A. 6*a* and 6*b*). We suggest that dimerization may affect the conformation of key structural elements (Fig. A. 6*d*), thus leading to new functional characteristics. For example, the dimeric OSCA1.2 showed mechanosensitivity with a lower threshold, compared with monomeric $OSCA1.2_{5Mu}$, $OSCA1.2_{63A\ IL2}$, and TMEM63s (Fig. A. 3*e* and 5*k*). Therefore, we suggest that monomeric protein subunits in OSCA/TMEM63, TMEM16, and TMC families are functional, and the two subunits in the dimeric protein can operate independently; dimerization could lead to distinct biophysical properties, thereby diversifying functionalities in this superfamily of ion channel proteins.

TMEM63 structures displayed similar TMD architectures with OSCA1.2, TMEM16A, and TMC-1 (Fig. A. 12*a*). It is unclear whether they share a common gating mechanism. Structural studies[198] in mouse TMEM16A revealed that Ca^{2+} binding induced a dramatic rearrangement in TM6 (Fig. A. 6*b*), highlighting a critical role of TM6 in the gating of TMEM16A. Indeed, mutations in TM6 of TMEM16A resulted in significant alterations in Ca^{2+} sensitivity[214]. A recent

functional study[222] in mouse TMC1 also showed that mutations in TM6 can affect the mechanosensitivity of sensory hair cells in the inner ear, which implies an involvement of TM6 in the gating process of TMC1. Interestingly, our structural studies revealed a conformational difference in TM6 between monomeric TMEM63s and OSCA1.2, akin to the Ca^{2+}-induced conformational change in TM6 of TMEM16A (Fig. A. 6*b*). Given the mechanosensitivity difference between TMEM63s and OSCA1.2 (Fig. A. 3*e* and 4*d*), it is highly likely that TM6 functions as a gating helix in the OSCA/TMEM63 family. In support of this hypothesis, pathogenic variants in humans have been identified in TM6 of both TMEM63A and TMEM63B[201–205] (Fig. A. 7). In particular, variants of the glycine residue in the π helix of TM6 (G567 in TMEM63A and G580 in TMEM63B) were reported in high frequency. Based on this evidence, it is plausible that TM6 is a common gating helix among OSCA/TMEM63, TMEM16, and TMC families. How TM6 couples stimulus-induced local conformational changes to pore opening may be diverse in this ion channel superfamily and remains to be elucidated. The density map of TMEM63A in nanodisc suggested a membrane distortion near IL2H2 / IL2H3 and TM6b. Similar distortion has been observed in TMEM16F, a scramblase with ion channel function, and has been shown to be related to TMEM16F scramblase activity[223]. Whether TMEM63A possesses scramblase activity is yet to be determined. It will also be interesting to see whether this membrane distortion is involved in the mechanogated pore opening process of TMEM63.

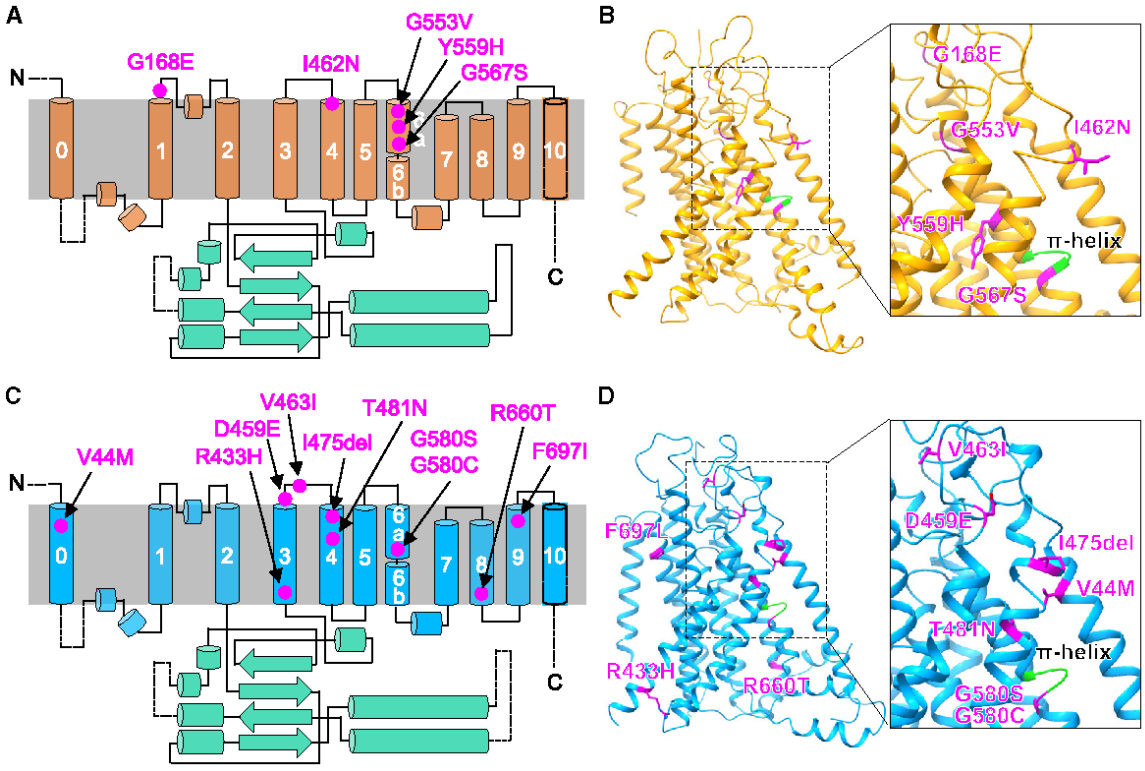


Figure A.7. Disease-associated variants in humans identified in TMEM63A and TMEM63B

(*a* and *c*) Location of five mutations of TMEM63A (*a*) and 10 mutations of TMEM63B (*c*) on a topology diagram. (*b* and *d*) Mutations mapped to structures of TMEM63A (*b*) and TMEM63B (*d*). See also Fig. A. 15.

Methods

Cultured cell lines

Suspended Expi293F cells were grown in Expi293 Expression Medium with 8% CO2 environment at a rotation speed of 100 rpm. When the cell density reached 1-2 x 10^6 cells/mL, constructs expressing TMEM63A or TMEM63B were transiently transfected with polyethylenimine Max 40k. For 1L of cell culture, 1 mg construct and 3 mg PEI Max 40k was used. Adherent HEK293T cells were cultured in Dulbecco's modified Eagle's medium plus 10% fetal bovine serum and 10 mg/mL ciprofloxacin hydrochloride, with 5% CO2 environment. Transient transfections were performed with lipofectamine 3000 based on the manufacturer's instructions.

Mice

All animal procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at Oregon Health & Science University (protocol IP00002957). The Tmem63b$^{HA/HA}$ line has been maintained on a C57BL/6 background and is genotyped using the following primers: Forward- TCA ACA GCA GCA ACC CGAAG, Reverse- CAC ATG AAG TCC AGA GCCAG. Brain tissue from Tmem63b$^{HA/HA}$ knockin or littermate Tmem63b$^{WT/WT}$ mice was collected at post-natal day 21 (P21), and snap-frozen in liquid nitrogen before use. Tissues were collected from animals of both sexes.

cDNA constructs and mutagenesis

The cDNA constructs of human TMEM63A (OHu08349C, GenBank: NM_014698.3), TMEM63B (OHu13152C, GenBank: NM_018426.3) and TMEM63C

(OHu27782C, GenBank: NM_020431.4) were purchased from Genscript (Piscataway, NJ). The TMEM63A, B or C coding sequence was then subcloned into pEG BacMam vector[224], a kind gift from Eric Gouaux (OHSU Vollum Institute, OR). An EGFP tag was placed at the C-terminus of TMEM63 proteins with a 3C protease cleavage site in the middle. A TwinStrep tag was added to the EGFP C-terminus for affinity purification. The OSCA1.2-PP-EGFP and mouse TMEM63A-GFP constructs was kindly provided by Ardem Patapoutian (Scripps Research Institute, CA). Flag-tagged OSCA1.2, human TMEM63A or TMEM63B constructs were generated by cloning the coding sequence to pEG BacMamvector with a C-terminal Flag tag. Drosophila TMEM63c DNA was purchased from Addgene (Cat# 136598) and cloned into pEG BacMam with a C-terminal GFP tag. pcDNA3.1 (+) construct expressing mouse TMEM63B cDNA (OMu53492C, GenBank: XM_006524138.3) with a GFP tag in the C-terminus was purchased from Genscript. The TMEM16A-GFP construct[225] was provided by Lily Yeh Jan (University of California at San Francisco, CA). Constructs expressing TRPP3-GFP and TMEM175-GFP were described previously[226,227]. Chimeric constructs were generated with NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs, Ipswich, MA). The OSCA1.2 concatemer was generated by cloning the coding sequence of OSCA1.2 E531K mutant into the upstream of OSCA1.2-PP-EGFP with a linker fragment in the middle that contains a glycophorin A helix, as described before[228]. Point mutations, insertions or deletions were introduced with QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, La Jolla, CA). All constructs were confirmed by sequencing.

Fluorescence-detection size-exclusion chromatography (FSEC)

Constructs expressing GFP-tagged proteins were transfected into HEK293T cells in 6-well plates with lipofectamine^{TM} 3000 (Thermo Fisher Scientific). Around 30 h following the transfection, cells were washed twice with ice-cold PBS buffer (pH 7.4) and lysed in Tris buffer, supplemented with 1% DDM (Anatrace) and Protease Inhibitor Cocktail (Thermo Fisher Scientific), for 30 min in cold room. The cell lysates were then cleared by centrifugation at maximum speed and 40 mL samples were applied to FSEC (Xbridge Protein BEH SEC Column, 450 Å, 3.5 mm, 7.8 mm X 150 mm, SKU176003598, Waters Corporation, Milford, MA) in Tris buffer supplemented with 0.05% DDM (Anatrace) and 1 mM Tris (2-carboxyethyl) phosphine (TCEP, P1020-25, Ubiquitin-Proteasome Biotechnologies, Dallas, TX).

Cross-linking assay

For the cross-linking of purified proteins, 7 mg of purified OSCA1.2-GFP, TMEM63A-GFP or TMEM63B-GFP protein was used in 30 mL reaction buffer (50 mM HEPES, 150 mM NaCl, pH 7.4) supplemented with 0.05% DDM (Anatrace). The cross-linker glutaraldehyde (Cat# G5882, Sigma-Aldrich, St. Louis, MO) was then added to the final concentration of 10 mM. The mixture was incubated at room temperature for 5 min. The reaction was then quenched by adding 3 mL of 1 M Tris (pH 7.4). The proteins were then subjected to SDS-PAGE and stained with Coomassie Brilliant Blue R-250 (Bio-Rad, Hercules, CA).

For the cross-linking of over-expressed proteins in HEK293T cell lysates, cells in a 6-well plate were first washed with PBS buffer (pH 8.0) twice. The cell lysates were then prepared with 500 mL reaction buffer supplemented with 1% DDM and Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific, Waltham, MA). Of note, DDM was used for all cell lysate preparations in our study unless otherwise specified. Cell lysates (30 mL) were incubated with glutaraldehyde (Sigma-Aldrich) at indicated concentrations for 5 min at room temperature. 3 mL 1 M Tris (pH 7.4) was then added to quench the reaction. The mixture was then subjected to western blot with antibodies against GFP (A-31852, Thermo Fisher Scientific), Flag (F7425, Sigma-Aldrich) and β-actin (sc-47778, Santa Cruz Biotechnology, Dallas, TX).

For cross-linking in intact cells, HEK293T cells expressing indicated proteins in 6-well plates were washed three times with PBS buffer (pH 8.0). At the last wash, the cross-linker disuccinimidyl suberate (DSS, A39267, Thermo Fisher Scientific) was added to a variety of concentrations, as indicated. The reaction mixture was incubated for 10 min at room temperature. 1 M Tris (pH 7.5) solution was then added to a final concentration of 50 mM and incubated for 10 min at room temperature to quench the reaction. The cells were then washed twice with Tris buffer (20 mM Tris, 150 mM NaCl, pH 7.4) and lysed in the same buffer supplemented with 1% DDM (Anatrace) and Protease Inhibitor Cocktail (Thermo Fisher Scientific). The cell lysates were then subjected to western blot with anti-GFP antibody (A-31852, Thermo Fisher Scientific).

Blue native PAGE

Blue native PAGE followed by immunoblotting was performed as described[226] with modifications of sample preparation. Cell lysates from HEK293T cells expressing OSCA1.2, TMEM63A or TMEM63B were prepared with NativePAGE Sample Prep Kit (Thermo Fisher Scientific) according to the manufacturer's manual. To prepare samples of mouse brain tissue for blue native PAGE, the whole brain of Tmem63b$^{WT/WT}$ or Tmem63b$^{HA/HA}$ knockin mice was first washed with ice-cold PBS solution (pH 8). The brain tissue was then homogenized on ice in 4 mL of NativePAGE 1X sample buffer contained in the NativePAGE Sample Prep Kit. The mixture was next centrifuged at a maximal speed to collect the pellet (cell membrane fraction). The pellet was washed with 2 mL of NativePAGE 1X sample buffer twice. The washed pellet was then solubilized in 2 mL of sample buffer supplemented with 1% DDM (Anatrace) and Protease Inhibitor Cocktail (Thermo Fisher Scientific) at cold room for 2 hrs, followed by centrifugation at a maximal speed. The supernatant was collected and diluted 5-fold. When preparing brain tissue samples for denaturing western blot, the whole brain was directly homogenized and solubilized in 2 mL of CelLytic MT lysis buffer (Sigma-Aldrich) with 1% DDM and Protease Inhibitor Cocktail. To run blue native gel, Coomassie blue G-250 was added to the supernatants from HEK293T cells or mouse brain tissue at 4:1 of detergent: G-250 ratio. Protein complexes were separated at 150 V for 100 min at room temperature using NuPAGE Novex 4-16% Bis-Tris protein gels (Thermo Fisher Scientific). For immunoblotting, separated proteins were transferred to polyvinylidene difluoride membrane using NuPAGE Novex transfer buffer (Thermo Fisher Scientific). The membrane was stained with ponceau red to visualize the protein ladder, which was then

marked with ballpoint pen. Next, the membrane was blocked with 3% skimmed mild in PBS buffer with 0.1% tween-20 for 30 min and then incubated with antibodies against Flag (F7425, Sigma-Aldrich), HCN1 (180092, Addgene), ASIC1 (184198, Addgene), TMEM63A (HPA068918, Sigma-Aldrich) or HA (ab9110, Abcam, Cambridge, UK).

Live-cell single-molecule imaging and tracking

Single-molecule imaging was done in transfected CHO cells using Total Internal Reflection (TIRF) Microscopy at an Olympus IX71 microscope equipped with a 100 x NA1.7 objective and a 488 nm laser. Movies of 200 frames were recorded at 20 Hz, a power density of 110 W/cm^2 (hTMEM63A-GFP was recorded at 33 Hz and a power density of 360 W/cm^2 for a better quality) and 100 nm pixel size with an EMCCD camera (Andor iXon DV-897 BV). Tracking was done with the particle tracking tool from the MOSAIC suite in ImageJ, and intensities were extracted from the identified spots after background subtraction. Intensity time courses were normalized to the different frame rates and power densities to yield comparable brightness values in Fig. A. 1*h*–1*j* and 8*c*.

Protein expression and purification

The human TMEM63A or TMEM63B in the pEG BacMam vector was expressed in Expi293F cells by transient transfection using PEI Max 40k (polysciences). 36-48 h after transfection, cells from 1 L culture were collected for protein purification. The cell pellet was washed with ice cold Tris buffer (50 mM Tris, 150 mM NaCl, pH 7.6) and was then resuspended and sonicated in 50 mL Tris buffer plus 1 mM tris (2-carboxyethyl)phosphine (TCEP) and protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO). The membrane

fraction was collected via ultracentrifugation for 30 min at 40,000 rpm. To purify proteins in detergent LMNG, the membrane fraction was first homogenized in 60 mL Tris buffer supplemented with protease inhibitor cocktail, 1 mM TCEP and 0.7% LMNG / 0.07% CHS (Anatrace, Maumee, OH). The homogenized mixture was then incubated at 4 °C for 2.5 h on a nutating mixer, followed by ultracentrifugation at 40,000 rpm for 1 h. The supernatant was then incubated with 1 mL Strep-Tactin XT 4Flow resin (IBA Lifesciences, Gottingen, Germany) at 4 °C for 3 h. The resin was washed with 10 column volumes of buffer W (IBA Lifesciences) plus 1 mM TCEP and 0.005% LMNG / 0.0005% CHS. The GFP-tagged human TMEM63A or TMEM63B was eluted with buffer BXT (IBA Lifesciences) plus 1 mM TCEP and 0.005% LMNG / 0.0005% CHS. The eluate was concentrated with a 100 kDa MWCO Amicon Ultra Centrifugal filter (MilliporeSigma, Burlington, MA) and the GFP tag was removed by treatment with 3C protease (Sigma-Aldrich, 1:100 in molar ratio) overnight at 4 °C. The TMEM63A or B protein was then further purified by size-exclusion chromatography (SEC) using a Superose 6 Increase column (GE Healthcare, Chicago, IL) equilibrated with Tris buffer supplemented with 1 mM TCEP and 0.005% LMNG / 0.0005% CHS. Fractions corresponding to the TMEM63A or TMEM63B were pooled and concentrated to around 2.5 mg/mL. The OSCA1.2 protein was expressed and purified as described previously[192] with the exception that Expi293F cells were used.

To obtain purified proteins in lipid nanodiscs, the TMEM63A or TMEM63B was first purified in detergent n-Dodecyl-b-D-maltoside (DDM, Anatrace) using the same purification procedures with LMNG replaced by DDM. In the membrane fraction solubilization, the 1% DDM/0.1% CHS was used, while in other steps, 0.05% DDM /

0.005% CHS was used. After elution and concentration, purified TMEM63A or TMEM63B with GFP tag was mixed with MSP2N2 scaffold protein and lipid (POPC: POPE: POPG in a 3:1:1 ratio, Anatrace) at a molar ratio of TMEM63 monomer: MSN2P2: lipid = 1: 4: 30. The mixture was incubated on ice for 1 h. To remove the DDM detergent, around 10-20 mg Bio-Beads SM2 (Bio-Rad, Hercules, CA) were added to the mixture, which was then incubated at 4 °C with constant rotation for 1 h. Another batch of Bio-Beads SM2 was added to the system for 1 h rotation at 4 °C. A third batch of Bio-Beads SM2 together with 3C protease was added and the mixture was incubated overnight at 4 °C. The Bio-Beads SM2 were then removed, and the sample was applied to SEC with a Superose 6 Increase column equilibrated with Tris buffer plus 1 mM TCEP. Fractions corresponding to the TMEM63A or TMEM63B nanodiscs were collected and reapplied to the SEC. The nanodisc fractions were then pooled and concentrated to around 2 mg/mL.

Cryo-EM sample preparation and data collection

A 3 μL drop of purified TMEM63A or TMEM63B protein in either detergent LMNG or nanodisc was applied to a glow-discharged Quantifoil R1.2/1.3 400-mesh gold grid (Electron Microscopy Sciences, Hatfield, PA). The grid was then blotted for 6 s in 100% humidity at 6 °C with a wait time of 3 s. The blotted grid was plunged into nitrogen-cooled liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). All grids were screened using a FEI Talos Arctica microscope at the Harvard Cryo-EM Center for Structural Biology. The data was then collected using a 300-kV Titan Krios microscope (FEI) equipped with a K3 direct electron detector with Gatan Quantum Image Filter (Gatan) at the Harvard Cryo-EM Center. For TMEM63A in nanodisc, images were

collected in a counting mode at a nominal magnification of 105,000x with a pixel size of 0.825 Å. 50 frames were collected per movie for a total accumulated dose of 50.7 electrons per Å using a defocus range of -1.0 to -2.0 μm. For TMEM63A in LMNG, images were collected with a 200 kV Talos Arctica at a nominal magnification of 36,000x with a pixel size of 1.1 Å. 47 frames were collected per movie for a total accumulated dose of 52 electrons per Å with a defocus range of -1.4 to -2.5 μm. For TMEM63B in either nanodisc or LMNG, images were collected in a counting mode at a nominal magnification of 105,000x with a pixel size of 0.825 Å. 50 frames were collected per movie for a total accumulated dose of 53 electrons per Å with a defocus range of -0.8 to -2.1 μm.

Cryo-EM imaging processing

The image processing for both TMEM63A and TMEM63B is shown in Fig. A. 9*g* and 10*g*, respectively. In short, dose fractionated images were motion corrected with MotionCor[229] (v1.2.6) followed by CTF estimation using CTFFIND4[230]. Particle picking was carried out using crYOLO[231] resulting in 4,552,788 and 2,433,802 particles for TMEM63A and TMEM63B respectively. Classification steps were carried out in RELION[232] (v3.0) unless otherwise noted. TMEM63A particles were subjected to multiple rounds of 3D classification until a clear secondary structure in the transmembrane region was observed, resulting in 198,144 particles being selected. CTF parameters were re-estimated for these particles using patch CTF estimation within cryoSPARC prior to non-uniform refinement[233] in cryosparc, resulting in a 3.8 Å reconstruction.

TMEM63B particles then underwent 2D classification resulting in 1,994,464 particles before a single round of full particle 3D classification giving 550,379 well aligning particles. A subsequent masked classification excluding the micelle was carried out giving a single class of 64,399 with high-resolution features. These particles were polished in RELION[234] and CTF parameters were re-estimated in cryoSPARC patch CTF refinement resulting in a reconstruction at 3.64 Å from cryoSPARC non-uniform refinement. Map quality and resolution was subsequently slightly improved using cryosparc local refinement excluding the micelle giving a final map at 3.62 Å. Structural biology applications other than cryosparc used in this project were compiled and configured by SBGrid[235].

Cryo-EM model building and refinement

The sharpened EM density maps obtained from cryoSPARC[233] were used for atomic model building. For both TMEM63A and TMEM63B, the AF2[209]-predicted model served as an initial template. The model was first fitted into the density map in UCSF ChimeraX[236]. Then, each individual residue was manually examined and adjusted to fit the map in Coot[86]. To build a model for the IL2 domain, the IL2 structure predicted by AF2 was docked into the density map with rigid body fitting and adjusted to fit the density map in secondary structural level using UCSF ChimeraX. Due to the missing densities, several fragments in TMEM63A (M1-D29, W76-C114, A369-E380, C711-A807) and TMEM63B (M1-N18, R71-F125, V307-V330, V381-E392, Y741-Q832) were excluded from final models. Generated models were subsequently refined in PHENIX using phenix.real_space_refine[237,238] and manually inspected in Coot for several iterations. The

model geometry was evaluated with the MolProbity Web Server. The ion permeation pore was predicted using the HOLE program[212]. Figures were generated with Pymol (Schrodinger), UCSF ChimeraX or Chimera[239].

Molecular dynamics simulations

A cryo-EM model that lacked IL1 (residues R71-F125) was used to build a complete human TMEM63B model for simulation studies. This model had a RMSD C_α of 0.8 Å with respect to the deposited structure (PDB: 8EHX) and retained a IL2 generated during refinement. A trimmed TMEM63B sequence (residues P19-N740) was submitted to AlphaFold2 v2.1.0[209], which made 50 predictions without relaxation in both monomer and multimer mode[240]. The top-ranking structures differed in the number of helical elements in IL1 (Fig. A. 12*g*): multimer mode predicted two helices (conformation C1; ipTM + pTM 0.18) while monomer mode predicted one (conformation C2; mean pLDDT 79.19). Residues V65-G70 and C126 were removed from the cryo-EM structure and then IL1 of each top-ranked prediction was positioned using a MatchMaker alignment of 20 residues flanking IL1 in ChimeraX[236]. TMEM63B systems based on the cryo-EM structure were built with and without residues P19-L36 (+/- NTD; see Fig. A. 12*f*). Additional systems using the complete AF2-monomer prediction described above were also built. All TMEM63B systems have a N-terminal protonated proline and C-terminal methylamidation, as well as disulfide bonds involving residues C320-C327 and C384-C389. Systems with the NTD (+ NTD) have an additional disulfide bond involving residues C23-C559. A monomeric OSCA1.2 was modeled using a 3.1 Å resolution structure in nanodisc (PDB: 6MGV). All non-terminal missing loops (residues P51-M70, V123-Y155,

I402-K419, and N491-I501) were replaced based on an AF2-multimer prediction of full-length dimeric OSCA1.2 (ipTM + pTM 0.83). Here, loops were placed using a MatchMaker alignment of 10 flanking residues.

After selecting neutral histidine protonation states to favor evident hydrogen bonding, our initial protein models were oriented and positioned in the membrane plane using the PPM 3.0 webserver (https://opm.phar.umich.edu/ppm_server3). Each protein was embedded in either pure POPC (VMD Membrane Builder[90]) or a more-realistic soybean membrane comprising seven types of phospholipids and two plant sterols (https://user.eng.umd.edu/ jbklauda/memb.html). In all cases, the initial patch was patterned in a regular array, trimmed to a final *x*- and *y*-extent of 130 Å, and lipids with head groups within 1.4 Å of protein were removed. The protein-lipid system was solvated with 18 Å box padding in the *z*-direction using VMD solvate. The systems were neutralized, and ions were added to achieve a final concentration of 150 mM KCl with VMD autoionize.

All MD simulations were performed using NAMD 2.14[241] and CHARMM36 force fields with CMAP correction and the TIP3P model for water[242]. A van der Waals cutoff of 12 Å (with a switching function starting at 10 Å) was used with periodic boundary conditions. The particle mesh Ewald method was used to calculate long range electrostatic interactions without cutoff and with a grid point density > 1 $Å^{-3}$. An integration timestep of 2 fs was used with the SHAKE algorithm. Electrostatic interactions were computed every other time step. Equilibrium simulations to obtain a well-packed and disordered

membrane bilayer were carried out in steps that included a minimization of 1,000 steps followed by 0.5 ns of free dynamics with everything fixed except for lipid tails, followed by a second minimization of 1,000 steps and 0.5 ns of free dynamics with only the protein constrained (k = 1 kcal/mol/Å^2). This step was followed by a 0.5 ns-long equilibration with the protein constrained except for IL1 and a final step with 0.5 ns of free dynamics with no constraints. All these steps used a Langevin damping coefficient of γ = 1 ps^{-1}. The rest of the equilibrium and voltage simulations, lasting up to 200 ns and 100 ns respectively, were performed using γ = 0.1 ps^{-1}. Langevin dynamics was used to maintain the temperature of the system at 310 K. The *NpT* ensemble at 1 atm was simulated using a hybrid Nose-Hoover Langevin piston method with a decay period of 200 fs and 50 fs damping time constant. Simulations with a normalized applied electric field to generate a -500 mV membrane potential were carried out following Gumbart et al[243].

To obtain water density plots, systems were centered according to the protein position at each frame and sequentially unwrapped and wrapped using VMD PBCTools. The number density of water oxygen atoms throughout each trajectory was computed and converted to concentration using VMD's volmap plugin with a resolution of 1 Å using frames saved every 50 ps. Density plots show averages computed using a box (90 x 90 x 100 Å^3) centered at the protein's center-of-mass. MATLAB surfaces were generated by a custom script with (Fig. A. 3*h*) or without (Fig. A. 14) interpolated shading. RMSD values were calculated every saved frame (5 ps) in VMD for a given selection after aligning that same selection to its initial conformation. Selections for TMEM63B include Cα atoms of transmembrane helices (residues 40-66, 148-178, 201-226, 427-458, 477-501, 506-545,

566-585, 620-638, 643-667, 674-702, and 707-727) and IL2 (residues 240-412). Same regions were defined for OSCA1.2 based on a structural alignment. Contacts between IL2 and TM6b / IL4H were quantified using buried surface area (BSA) computed every 5 ps. To compute the BSA between selections A and B, the VMD measure sasa command was used to calculate $BSA_{AB} = ½ (SASA_A + SASA_B - SASA_{AB})$. Selections for TMEM63B were residues 589-614 (TM6b/ILH4) and residues 278 to 354 (IL2). Molecular figures from simulations were rendered using VMD.

Surface protein biotinylation assay

The biotinylation assay for detecting proteins in cell surface membrane was performed as described previously[244]. In brief, HEK293T cells in 6-well plates were transfected with OSCA1.2 constructs using lipofectamine 3000 (Thermo Fisher Scientific). After 48 h, cells were washed twice with ice-cold PBS solution (pH 8), followed by incubation with sulfo-NHS-SS-Biotin (0.5 mg/mL in 500 mL PBS, Piece, Rockford, IL) at room temperature for 15 min. The excess nonreacted biotin was then quenched with NH_4Cl. Next, cells were washed once with PBS and lysed using CelLyticMlysis buffer (Sigma-Aldrich) supplemented with proteinase inhibitor (Thermo Fisher Scientific). Cell lysates were then incubated with 50 mL Streptavidin bead (Pierce) for 4 h at 4 °C. The captured surface proteins were washed with PBS buffer plus 0.05% DDM (Anatrace) and were then subjected to western blot and detected with GFP antibody (Thermo Fisher Scientific) and b-actin antibody (sc-47778 HRP, Santa Cruz Biotechnology).

Proteoliposome reconstitution

Liposome reconstitution of purified membrane proteins were performed with dehydration/rehydration (D/R) method as previously described[245]. Specifically, 20 mg azolectin (P5638, Sigma-Aldrich) was first dissolved in 1 mL chloroform in a glass test tube. The chloroform was then completely removed by nitrogen gas (around 15 min). We then added 2 mL D/R buffer (5 mM HEPES, 200 mM KCl, pH 7.2) to the glass tube to make a final concentration of 10 mg/mL for azolectin. To dissolve the azolectin, the mixture was vortexed for 10 min, followed by 30 min of bath sonication (around 6 cycles of 5 min sonication with 2 min off) until the mixture looked translucent. We took 200 μL of the resulting lipid solution to a 1.5 mL Eppendorf tube and added fresh purified protein (in DDM detergent) with a desired weight ratio (1: 500 for OSCA1.2 and 1:200 for TMEM63A or TMEM63B). The mixture was then rotated gently for 1 h at room temperature. Next, the DDM detergent was removed by incubating with BioBeads SM2 (Bio-Rad) for 3 h at room temperature. After removal of the BioBeads, the mixture was centrifuged at 40,000 rpm for 45 min at 4 °C (MLS 50 rotor, Optima MAX Ultracentrifuge, Beckman Coulter Brea, CA). The supernatant was discarded, and the pellet was resuspended in 80 μL D/R buffer. Four drops with each 20 μL of the resuspension were placed on a clean glass slide and were kept in a vacuum desiccator overnight for dehydration. Next day, 20 μL D/R buffer was added to each dehydrated spot and kept it at 4 °C (cold room) without disturbance for 24 h before carrying out electrophysiological recordings.

Electrophysiological recordings

Proteoliposome recordings were adapted based on previous description[245]. We added 5 μL of the rehydrated drop to the bath solution (in mM): 140 KCl, 10 HEPES, 1 $MgCl_2$, 10 glucose, pH 7.3 adjusted with KOH) in the recording chamber and waited for 5-10 min to allow the lipids to settle down. The unilamellar liposome membrane would then be visible under the microscope and ready to be recorded. The recording pipette was made from borosilicate glass with an outer diameter of 1.5 mm (Warner Instruments, Hamden, CT) using a micropipette puller (model P-2000, Sutter Instruments, Novato, CA) and polisher (ALA Scientific Instruments). The polished pipette was back-filled with internal solution containing (in mM): 130 NaCl, 5 KCl, 10 HEPES, 10 TEA-Cl, 1 $CaCl_2$, 1 $MgCl_2$, pH 7.3 (with NaOH). The pipette resistance was in the range of 3-5 MΩ when filled with the internal solution. The stretch-activated current was evoked in excised configuration with stepwise, 300 ms negative pressure pulses (Δ10 mmHg with 2 s between stimuli) using a high-speed pressure clamp system (HSPC-1, ALA Scientific Instruments, Farmingdale, NY). The membrane potential inside the patch was held at - 80 mV. Currents were recorded using an Axopatch 200B patch-clamp amplifier (Molecular Devices, Union City, CA), filtered at 10 kHz through a low-pass analog filter, and sampled at a rate of 20 kHz with a digidata 1322A digitizer (Molecular Devices). The pClamp 10 software (Axon Instruments, Union City, CA) was used for data acquisition and analysis. Recordings were not corrected for liquid junction potential. For single-channel recordings, macroscopic currents were first recorded to find out the pressure threshold for stretch-induced activation. The single-channel openings were then resolved with the pressure threshold using the same conditions described above.

Electrophysiological recordings on cultured cells were performed as previously described[185]. HEK293T cells in 6-well plates were transfected with indicated constructs using lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's instructions. Cells were seeded 24-48 h following transfection onto matrigel-coated coverslips (BD Bioscience, Billerica, MA) and used for recordings. 4-12 h later, the stretch-activated macroscopic or single-channel currents were recorded with cell-attached configurations. Experimental conditions were similar to what was described above for proteoliposome recordings. The stretch-activated peak current (I) against pressure (P) was fitted with a Boltzmann equation $I = I_{min} + (I_{max}-I_{min}) / (1 + \exp(-Z^*(P-P_{50}))$, where the I_{min} is the minimal stretch-activated current and I_{max} is the maximal stretch-activated current; P_{50} is the pressure required for half-maximal activation and Z represents the slope of the activation curve.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data were analyzed and plotted using Clampfit and GraphPad Prism 8.2.1 (GraphPad Software Inc., La Jolla, CA) and expressed as means ± SEM (standard error of the mean) as noted in the text and figure legends. Number (n) of patches recorded is indicated for each mean value in the text and figure legends. Statistical analyses were carried out using Student's t tests when comparing two groups or one-way or two-way ANOVA for three or more groups, with corrections for multiple comparisons. A probability value (p) of less than 0.05, 0.01, and 0.001 was considered statistically

significant and indicated by *, **, and ***, respectively. The investigators were not blinded to the experimental groups.

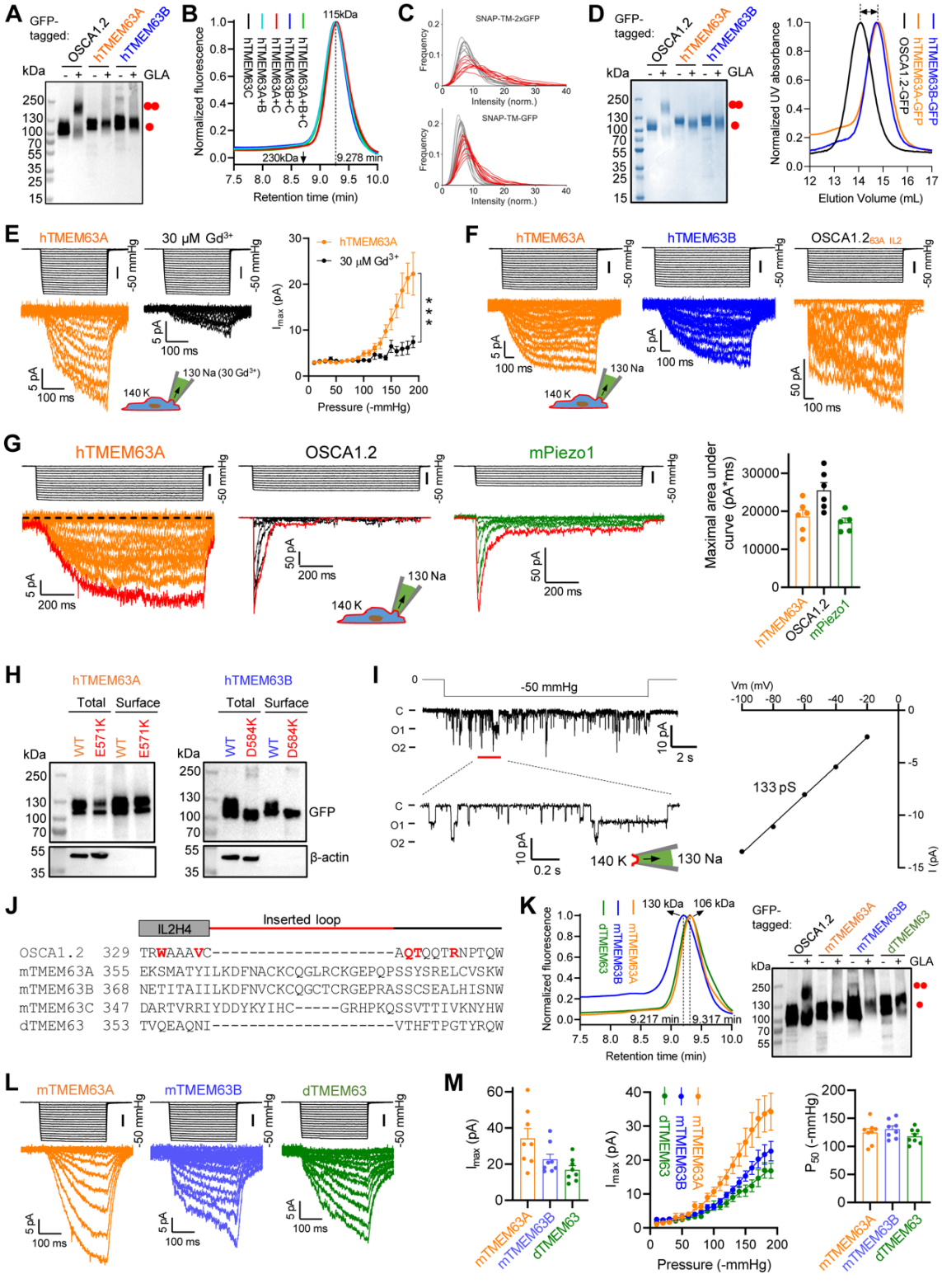

A
GFP-tagged:
OSCA1.2
hTMEM63A
hTMEM63B
kDa
GLA
250
130
100
70
55
35
25
15
B
115kDa
hTMEM63C
hTMEM63A+B
hTMEM63A+C
hTMEM63B+C
hTMEM63A+B+C
Normalized fluorescence
230kDa
9.278 min
Retention time (min)
C
SNAP-TM-2xGFP
SNAP-TM-GFP
Frequency
Intensity (norm.)
D
GFP-tagged:
OSCA1.2
hTMEM63A
hTMEM63B
kDa
GLA
Normalized UV absorbance
OSCA1.2-GFP
hTMEM63A-GFP
hTMEM63B-GFP
Elution Volume (mL)
E
hTMEM63A
30 μM Gd3+
-50 mmHg
5 pA
100 ms
130 Na (30 Gd3+)
140 K
Imax (pA)
Pressure (-mmHg)
F
hTMEM63A
hTMEM63B
OSCA1.2R83A IL2
130 Na
140 K
50 pA
G
hTMEM63A
OSCA1.2
mPiezo1
200 ms
130 Na
140 K
Maximal area under curve (pA*ms)
H
hTMEM63A
hTMEM63B
Total
Surface
WT
E571K
D584K
kDa
GFP
β-actin
I
-50 mmHg
C
O1
O2
10 pA
2 s
0.2 s
140 K
130 Na
Vm (mV)
133 pS
I (pA)
J
IL2H4
Inserted loop
OSCA1.2 329 TRWAAAVC----------AQTQQTRNPTQW
mTMEM63A 355 EKSMATYILKDFNACKCQGLRCKGEPQPSSYSRELCVSKW
mTMEM63B 368 NETITAIILKDFNVCKCQGCTCRGEPRASSCSEALHISNW
mTMEM63C 347 DARTVRRIYDDYKYIHC-----GRHPKQSSVTTIVKNYHW
dTMEM63 353 TVQEAQNI----------VTHFTPGTYRQW
K
130 kDa
106 kDa
mTMEM63A
mTMEM63B
dTMEM63
Normalized fluorescence
9.217 min
9.317 min
Retention time (min)
GFP-tagged:
OSCA1.2
mTMEM63A
mTMEM63B
dTMEM63
kDa
GLA
L
mTMEM63A
mTMEM63B
dTMEM63
-50 mmHg
5 pA
100 ms
M
Imax (pA)
Pressure (-mmHg)
P50 (-mmHg)

Figure A.8. Biochemical and electrophysiological analyses of TMEM63 and OSCAl.2 proteins, related to Figures A. 1, 4 and 5.

(*a*) Representative western blot of HEK293T cell lysate with overexpression of GFP-tagged OSCAl.2, hTMEM63A or hTMEM63B. The cell lysate was treated with or without 10 mM GLA. (*b*) Representative FSEC traces of HEK293T cell lysate containing expression of C-terminal GFP-tagged hTMEM63C or indicated co-expression. Retention time corresponding to a potential homo- or hetero-dimer (around 230 kDa without GFP) is marked. (*c*) Intensity distribution before (red) and after (gray) photobleaching for control proteins with a monomeric transmembrane (TM) domain, and either one or two GFP tags fused to their intracellular C-termini. The control proteins also contain an N-terminal SNAP tag. (*d*) *Left*, coomassie-blue staining of purified OSCAl.2, human TMEM63A and TMEM63B with a C-terminal GFP-tag after SDS- PAGE. Proteins were treated with or without 10 mM GLA; *Right*, representative SEC traces of indicated purified proteins. The UV absorbance is normalized to the peak amplitude for a better comparison. (*e*) *Left*, representative stretch-activated current traces from TMEM63A-expressing HEK293T cells in the absence or presence of 30 μM Gd^{3+} in the pipette solution; *Right*, quantification of I_{max} at various negative pressures in the absence (n = 7) or presence of Gd^{3+} (n = 7). $^{***}p < 0.001$. Two-way ANOVA analysis. (*f*) Representative stretch-activated current traces with prolonged application of negative pressures (1 s). (*g*) *Left*, representative traces of stretch-activated currents in HEK293T cells with overexpression of hTMEM63A, OSCAl.2 or mouse Piezol; *Right*, quantification of the maximal area under the current traces, which represents the number of conducted charges. (*h*) Biotinylation assay detecting both the total expression and cell surface expression of hTMEM63A WT and E571K mutant (*Left*) or hTMEM63B WT and D584K mutant (*Right*) in HEK293T cells. (*i*) Left, representative trace of single-channel recordings of OSCAl.2 reconstituted in liposomes. The voltage was held at -50 mV. Closed (C) and open (O) states are indicated; *Right*, current-voltage relationship of stretch-activated single-channel currents. The calculated single-channel conductance is shown. (*j*) Sequence alignment of the region within IL2 that mediates dimerization of OSCAl.2. Residues that directly mediate OSCAl.2 dimerization are highlighted in red. The inserted loop in TMEM63s is indicated by the red line. UniProt accession numbers of sequences used are: OSCAl.2: Q5XEZ5; mTMEM63A: Q91YT8; mTMEM63B: Q3TWI9; mTMEM63C: Q8CBX0; dTMEM63: Q6NP91. (*k*) Representative FSEC traces (*Left*) or western blot (*Right*) of HEK293T cell lysate with overexpression of GFP-tagged mouse TMEM63A (mTMEM63A), TMEM63B (mTMEM63B) or *Drosophila melanogaster* TMEM63 (dTMEM63). Cell lysates were treated without (-) or with (+) 10 mM GLA. (*l*) Representative traces of stretch-activated currents (-80 mV) from HEK293T cells with overexpression of mTMEM63A, mTMEM63B or dTMEM63. (*m*) Quantification of I_{max} (*Left*) and P_{50} (*Right*) from each individual patches; *Middle*, averaged pressure response current curves fitted with a Boltzmann equation (mTMEM63A, n = 8; mTMEM63B, n = 8; dTMEM63, n = 8).

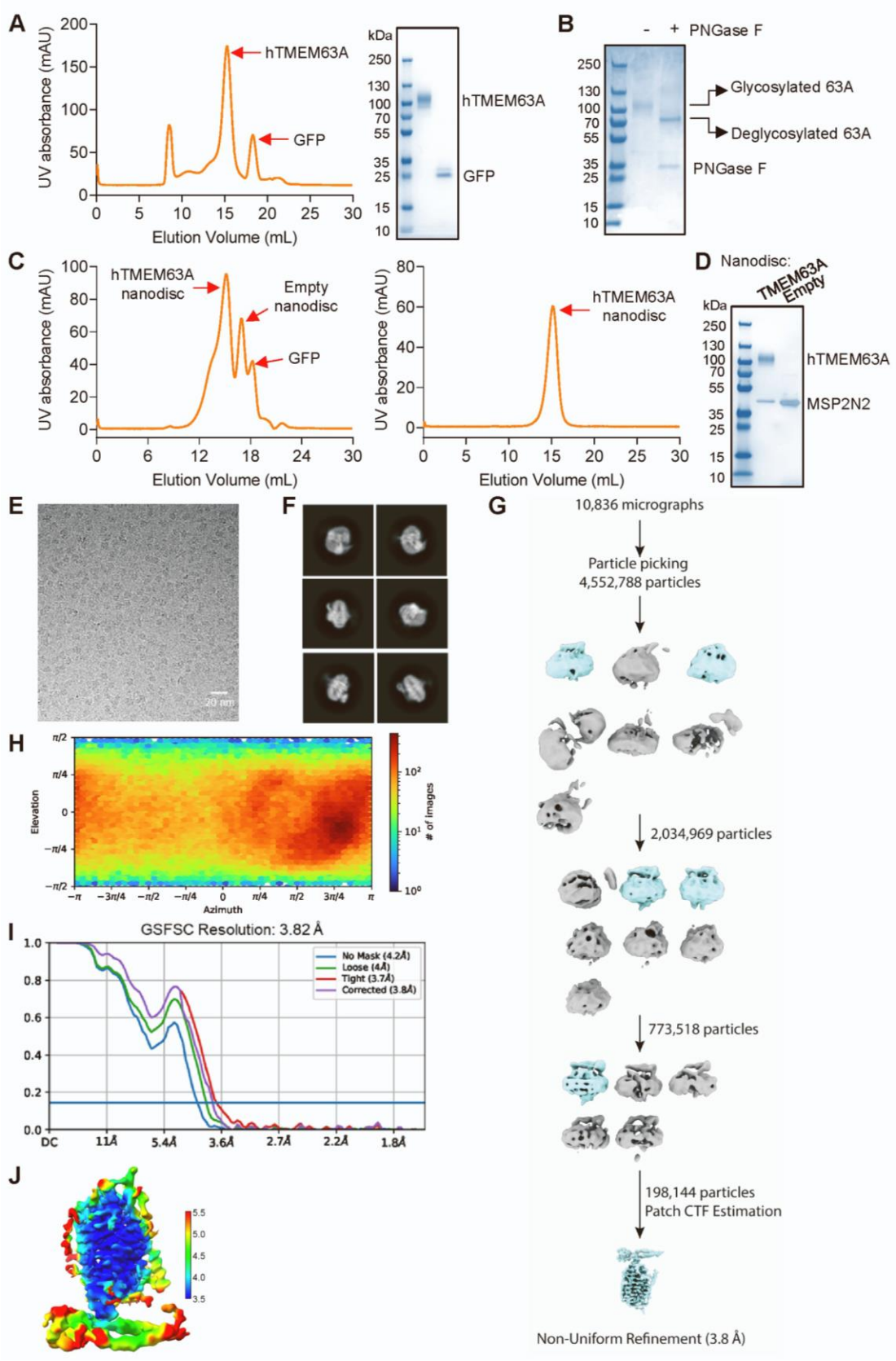

A
UV absorbance (mAU)
Elution Volume (mL)
hTMEM63A
GFP
kDa
hTMEM63A
GFP
B
- + PNGase F
Glycosylated 63A
Deglycosylated 63A
PNGase F
C
hTMEM63A nanodisc
Empty nanodisc
GFP
hTMEM63A nanodisc
D Nanodisc:
TMEM63A Empty
hTMEM63A
MSP2N2
E
F
G
10,836 micrographs
Particle picking
4,552,788 particles
2,034,969 particles
773,518 particles
198,144 particles
Patch CTF Estimation
Non-Uniform Refinement (3.8 Å)
H
Elevation
Azimuth
# of Images
I
GSFSC Resolution: 3.82 Å
No Mask (4.2Å)
Loose (4Å)
Tight (3.7Å)
Corrected (3.8Å)
J

Figure A.9. Purification and reconstruction of human TMEM63A, related to Fig. A. 2.

(*a*) *Left*, representative SEC trace of purified human TMEM63A protein in LMNG / CHS, treated with 3C protease to cleave GFP. The peak fractions corresponding to hTMEM63A and GFP are indicated; *Right*, SDS-PAGE gel staining of hTMEM63A and GFP peak fractions. (*b*) SDS-PAGE gel staining of hTMEM63A peak fraction treated without or with PNGase F. Glycosylated and deglycosylated hTMEM63A are indicated. (*c*) *Left*, SEC trace of hTMEM63A reconstituted in lipid nanodiscs with MSP2N2; *Right*, SEC trace of hTMEM63A nanodisc peak fraction collected from the *left* panel. (*d*) SDS-PAGE gel staining of peak fractions of hTMEM63A nanodiscs and empty nanodiscs. (e and f) Representative cryo-EM raw image (*e*) and 2D-class averages (*f*) of hTMEM63A in nanodiscs. (*g*) Cryo-EM data processing flowchart for nanodisc-embedded hTMEM63A density map. (*h*) Angular distribution of reconstructed particles in the final map of hTMEM63A. (*i*) FSC plots of unmasked and masked maps. (*j*) Final reconstruction map colored by local resolution.

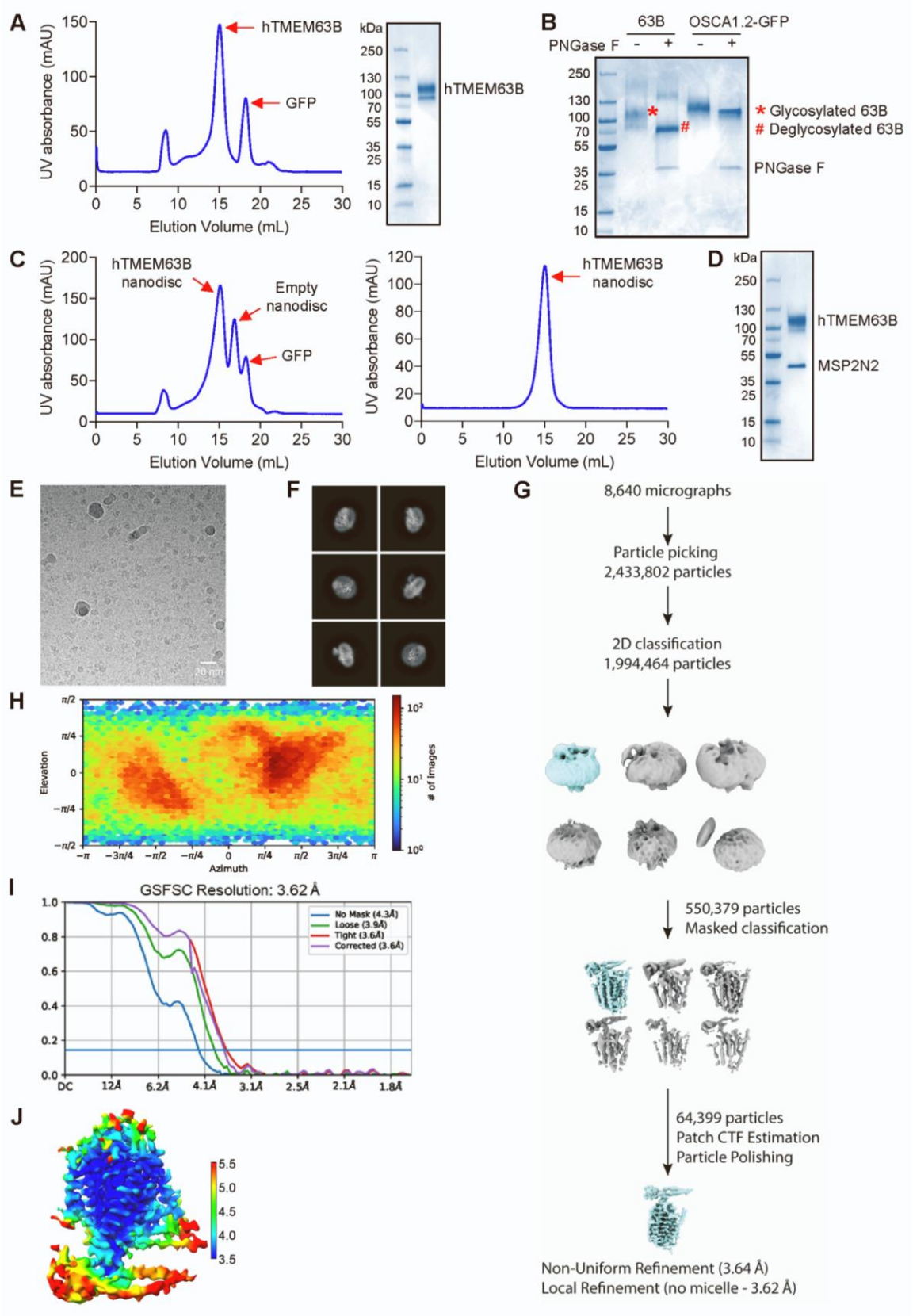

A
UV absorbance (mAU)
Elution Volume (mL)
hTMEM63B
GFP
kDa
hTMEM63B
B
63B
OSCA1.2-GFP
PNGase F
* Glycosylated 63B
# Deglycosylated 63B
PNGase F
C
hTMEM63B nanodisc
Empty nanodisc
GFP
hTMEM63B nanodisc
D
hTMEM63B
MSP2N2
E
F
G
8,640 micrographs
Particle picking
2,433,802 particles
2D classification
1,994,464 particles
550,379 particles
Masked classification
64,399 particles
Patch CTF Estimation
Particle Polishing
Non-Uniform Refinement (3.64 Å)
Local Refinement (no micelle - 3.62 Å)
H
Elevation
Azimuth
# of Images
I
GSFSC Resolution: 3.62 Å
No Mask (4.3Å)
Loose (3.9Å)
Tight (3.6Å)
Corrected (3.6Å)
J

Figure A.10. Purification and reconstruction of human TMEM63B, related to Fig. A. 2.

(*a*) SEC trace of purified human TMEM63B protein in LMNG / CHS. GFP was cleaved by 3Cprotease treatment. The hTMEM63B peak fraction was examined by SDS-PAGE. (*b*) SDS-PAGE gel staining of purified hTMEM63B and OSCAI.2-GFP treated without or withPNGase F. Glycosylated and deglycosylated hTMEM63B are indicated. (*c*) *Left,* SEC trace of hTMEM63B reconstituted in lipid nanodiscs with MSP2N2; *Right,* SECtrace of hTMEM63B nanodisc peak fraction collected from the *left* panel. (*d*) SDS-PAGE gel staining of the hTMEM63B nanodisc peak fraction. (*e* and *f*) Representative cryo-EM raw image (e) and 2D-class averages (f) of hTMEM63B inLMNG / CHS. (*g*) Cryo-EM data processing flowchart for the map ofhTMEM63B in LMNG / CHS.(*h*) Angular distribution of reconstructed particles in the final map of hTMEM63B. (*i*) FSC plots of unmasked and masked maps. (*j*) Final reconstruction map colored by local resolution.

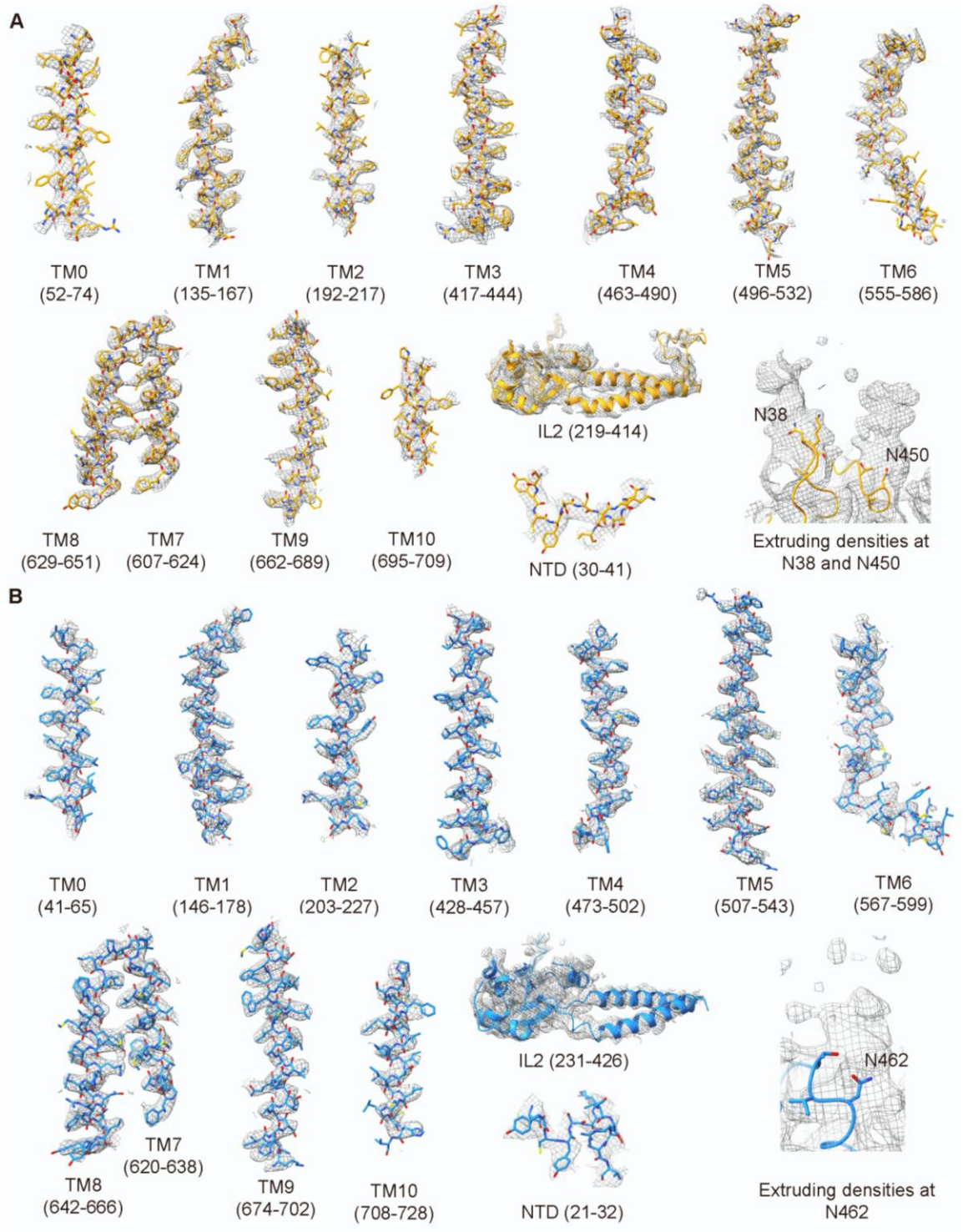


Figure A.11. Cryo-EM density and fit with models of TMEM63s, related to Fig. A. 2.

(*a* and *b*) Selected regions of cryo-EM map of TMEM63A in nanodiscs (*a*) and TMEM63B in LMNG (*b*) superposed with the atomic models.

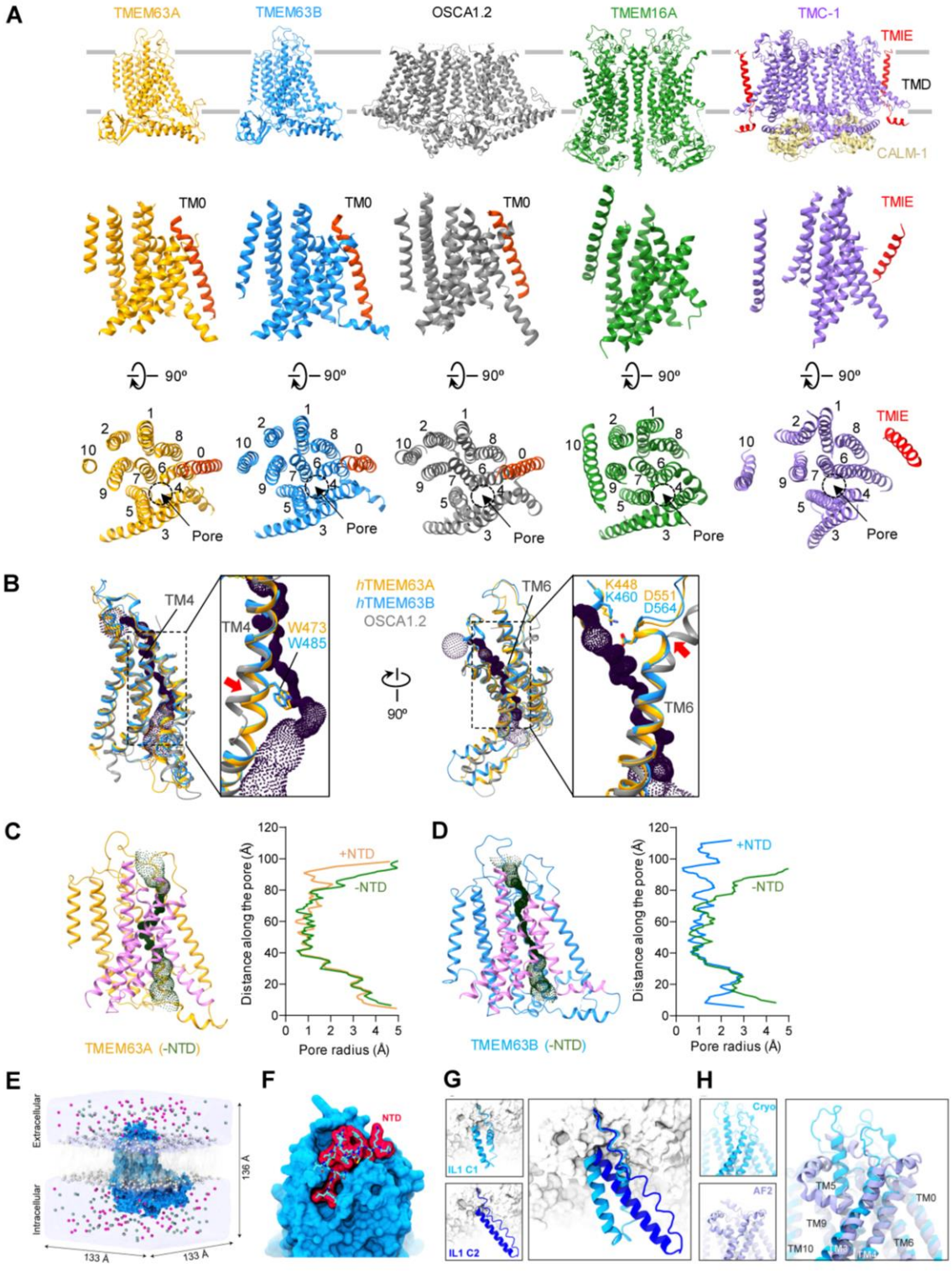

A
TMEM63A
TMEM63B
OSCA1.2
TMEM16A
TMC-1
TMIE
TMD
CALM-1
TM0
90°
Pore
B
TM4
W473
W485
hTMEM63A
hTMEM63B
OSCA1.2
TM6
K448
K460
D551
D564
C
TMEM63A (-NTD)
+NTD
-NTD
Distance along the pore (Å)
Pore radius (Å)
D
TMEM63B (-NTD)
E
Extracellular
Intracellular
136 Å
133 Å
F
NTD
G
IL1 C1
IL1 C2
H
Cryo
AF2
TM5
TM0
TM9
TM10
TM6

Figure A.12. Structural comparisons and ion permeation pore analysis, related to Fig. A. 3.

(*a*) Comparison of TMDs. *Upper,* side view of ribbon structures of human TMEM63A, human TMEM63B, *Arabidopsis thaliana* OSCAI.2 (PDB: 6MGV), mouse TMEM16A (PDB: 5OYG) and C. *elegans* TMC-1 bound with TMIE and CALM-I (PDB: 7USX). *Middle,* side-view of ribbon structures of TMD isolated from TMEM63A, TMEM63B, OSCAI.2, TMEM16A and TMC-1. The TM0 in TMEM63A, TMEM63B and OSCAI.2 and TMIE is highlighted in red. *Lower,* top view of the TMD with the location of putative ion permeation pore indicated. (*b*) Structural comparisons of the pore domain among TMEM63A, TMEM63B and OSCAI.2 (PDB: 6MGV). The noticeable differences in TM4 and TM6 were highlighted. (*c* and *d*) Effects of NTD on the ion permeation pore profile in TMEM63A and TMEM63B. *Left,* ribbon structures of TMEM63A (*c*) and TMEM63B (*d*) without the NTD (S30-V45 in TMEM63A; Pl9-L36 in TMEM63B). The calculated pore profile is shown as green dots. Pore lining TMs are colored in magenta; *Right,* pore radii of TMEM63A (*c*) and TMEM63B (*d*) with(+NTD) and without NTD (-NTD) against distance along the pore. (*e*) All-atom MD simulation setup for TMEM63B in a POPC membrane (Sim1a). Protein is shown in surface representation. Chloride and potassium ions are in dark green and pink, respectively. Lipid phosphates are white spheres and water is in transparent blue surface. (*f*) Highlight of TMEM63B NTD, removed in some simulations. (*g*) Two predicted conformations for ILI. (*h*) Experimentally derived (Cryo) and AF2 predicted conformations of extracellular regions of TM3-TM4 and TM5-TM6. Other helices (e.g., TM0, TM9, TMI0) have matching conformations.

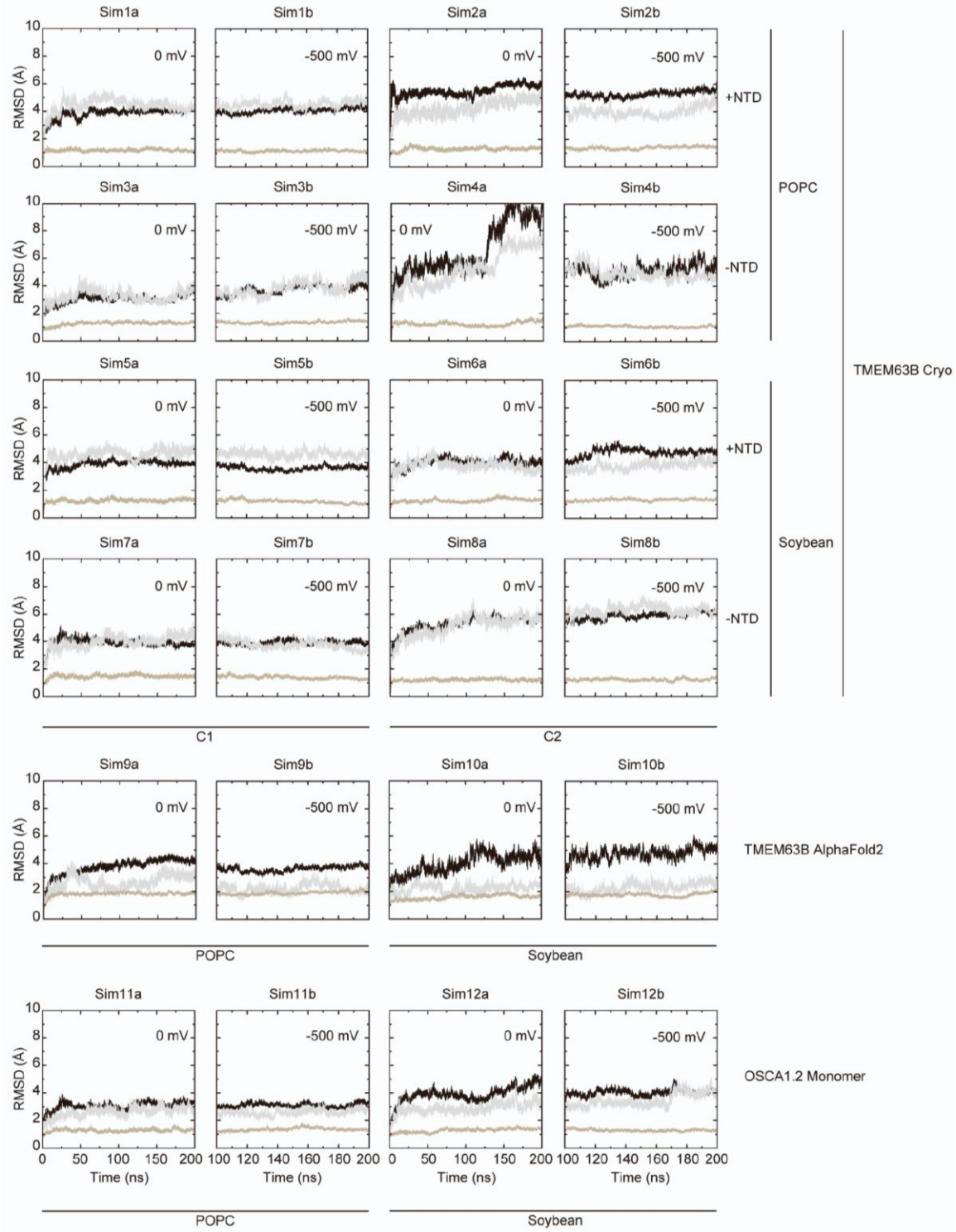


Figure A.13. Domain stability in simulations, related to Fig. A. 3.

RMSD values for all simulations in Table A. 2 computed using C_α atoms of the entire protein(black), of the IL2 (gray), and of the TMD (tan).

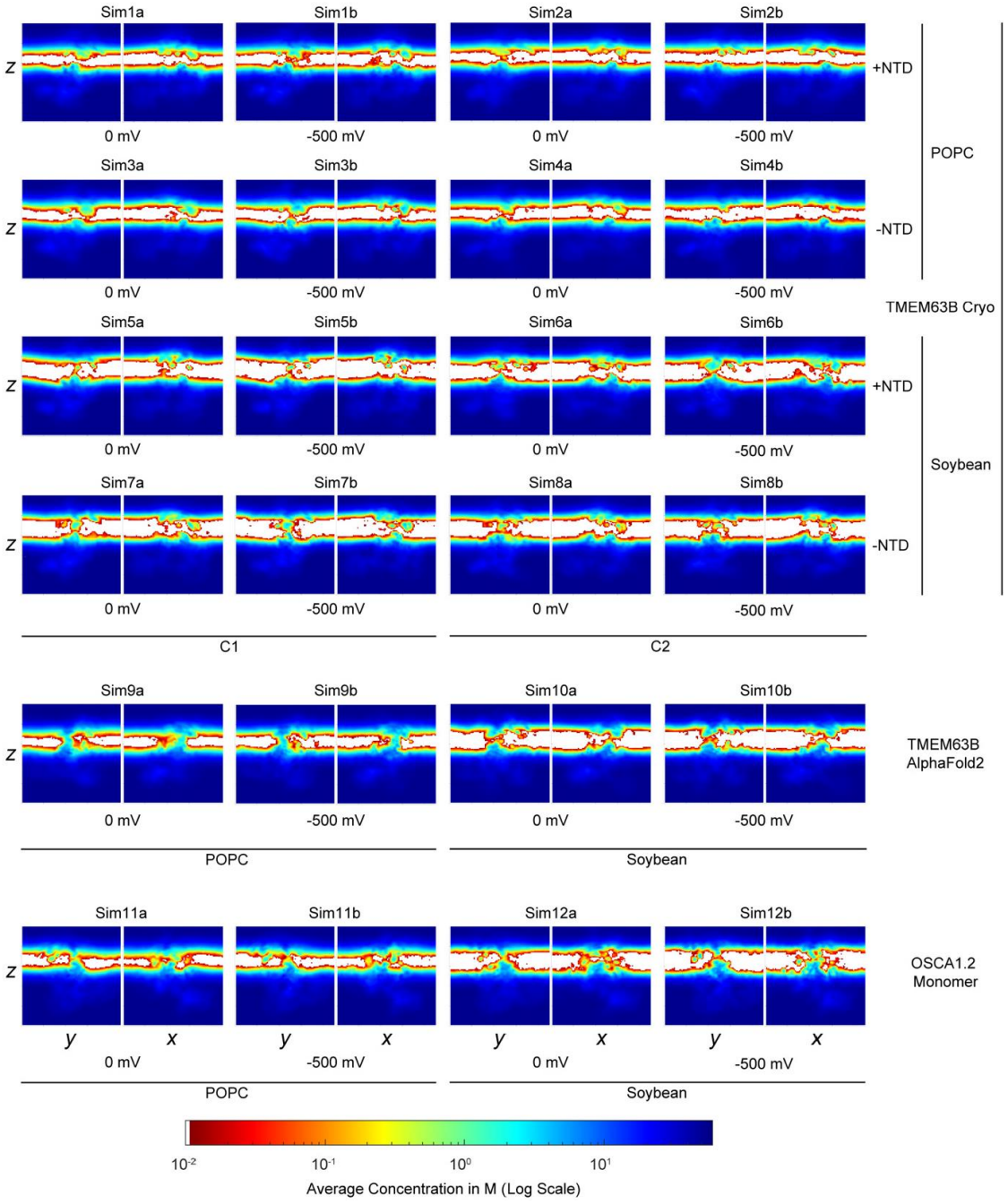


Figure A.14. Pore hydration in simulations of TMEM63B and OSCA1.2, related to Fig. A. 3.

Two dimensional slices showing time- and space-averaged water concentrations for all MD simulations listed in Table A.2. Data shown in a logarithmic color scale to highlight transient hydration of putative pores.

Figure A.15. Protein sequence alignment of TMEM63 proteins and OSCA1.2, related to Fig. A. 2 and 7.

Multiple sequence alignment was performed using ClustalW for human TMEM63A (UniProt Accession #: 094886), TMEM63B (Q5T3F8), TMEM63C (Q9PlW3) and *Arabidopsis thaliana* OSCA1.2 (Q5XEZ5). Secondary elements of human TMEM63A based on cryo-EM model are shown on the top, with the same color code as in Fig. A. 2*c*. Unresolved regions in the cryo-EM model are indicated with dashed line. Two putative N-linked glycosylation sites (N38 and N450) in hTMEM63A are indicated with yellow triangles. The π-helix in TM5 and TM6 is labeled with a black oval.

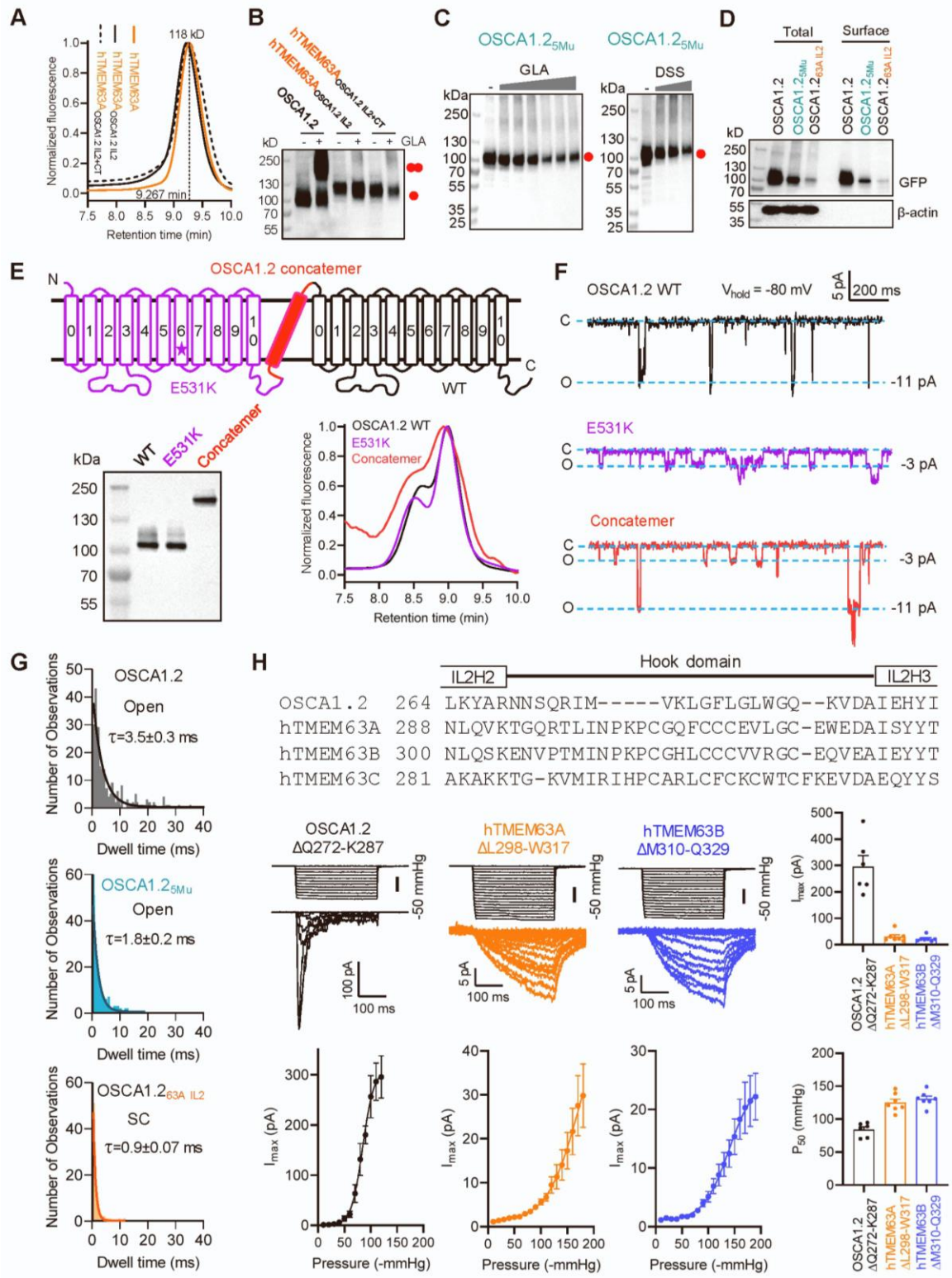

A
118 kD
hTMEM63A
hTMEM63A OSCA1.2 IL2
hTMEM63A OSCA1.2 IL2+CT
9.267 min
Normalized fluorescence
Retention time (min)
B
hTMEM63A OSCA1.2 IL2+CT
hTMEM63A OSCA1.2 IL2
OSCA1.2
kD
GLA
C
OSCA1.2 5Mu
GLA
DSS
kDa
D
Total
Surface
OSCA1.2
OSCA1.2 5Mu
OSCA1.2 63A IL2
GFP
β-actin
E
OSCA1.2 concatemer
N
C
E531K
WT
kDa
WT
E531K
Concatemer
OSCA1.2 WT
E531K
Concatemer
Normalized fluorescence
Retention time (min)
F
OSCA1.2 WT
V hold = -80 mV
5 pA
200 ms
-11 pA
E531K
-3 pA
Concatemer
-3 pA
-11 pA
G
OSCA1.2
Open
τ=3.5±0.3 ms
OSCA1.2 5Mu
Open
τ=1.8±0.2 ms
OSCA1.2 63A IL2
SC
τ=0.9±0.07 ms
Number of Observations
Dwell time (ms)
H
IL2H2
Hook domain
IL2H3
OSCA1.2 264 LKYARNNSQRIM-----VKLGFLGLWGQ--KVDAIEHYI
hTMEM63A 288 NLQVKTGQRTLINPKPCGQFCCCEVLGC-EWEDAISYYT
hTMEM63B 300 NLQSKENVPTMINPKPCGHLCCCVVRGC-EQVEAIEYYT
hTMEM63C 281 AKAKKTG-KVMIRIHPCARLCFCKCWTCFKEVDAEQYYS
OSCA1.2 ΔQ272-K287
hTMEM63A ΔL298-W317
hTMEM63B ΔM310-Q329
-50 mmHg
100 pA
100 ms
5 pA
I max (pA)
P 50 (mmHg)
Pressure (-mmHg)

Figure A.16. Characterization of engineered OSCAl.2, TMEM63A and TMEM63B channels, related to Fig. A. 5 and 6.

(*a* and *b*) Representative FSEC traces (*a*) and western blot (*b*) of HEK293T cell lysate with overexpression of GFP-tagged constructs, as indicated. hTMEM63A$_{OSCA1.2\ IL2}$, IL2 of hTMEM63A were replaced with OSCAI.2 IL2; hTMEM63A$_{OSCA1.2\ IL2+CT}$, both IL2 and C-terminus of hTMEM63A were replaced with those of OSCA1.2. Cell lysates were treated without (-) or with (+) 10 mM GLA. (*c*) Representative western blots of GFP-tagged OSCA1.2$_{5Mu}$ overexpressed in HEK293T cells. *Left*, cell lysates were treated with GLA as in Fig. A. 5*f*; *Right*, intact cells were treated with DSS as in Fig. A. 1*d*. (*d*) Biotinylation assay detecting both the total expression and cell surface expression in HEK293T cells overexpressing GFP-tagged OSCAI.2, OSCA1.2$_{5Mu}$ and OSCA1.2$_{63A\ IL2}$. (*e*) *Upper*, diagram showing the design of OSCAI.2 concatemer. E531K mutant (purple, the mutation site is indicated with a star) and WT (black) protomers are connected via a single transmembrane helix of glycophorin A (red); *Lower*, representative western blot (*Left*) and FSEC traces (*Right*) of HEK293T cell lysates with overexpression of GFP-tagged OSCA1.2 WT, E531K mutant or the concatemer. (*f*) Representative traces of stretch-activated single-channel currents recorded at -80 mV. The closed (C) and open (O) states, and amplitude of unitary currents are indicated. (*g*) Dwell time histograms for the fully open state of OSCA1 .2 and OSCA1.2$_{5Mu}$, and the assumed sub-conductance state of OSCA1.2$_{63A\ IL2}$. (*h*) *Upper*, sequence alignment of the hook domain; Middle, representative traces of stretch activated currents (-80 mV) from HEK293T cells with overexpression of Hook-deleted OSCA1.2 (ΔQ272-K287), hTMEM63A (ΔL298-W3l7) or hTMEM63B (ΔM310-Q329), and quantification of I_{max}; *Lower*, averaged pressure-response current curves fitted with a Boltzmann equation (OSCA1.2 ΔQ272-K287, n = 6; hTMEM63A ΔL298-W3l7, n = 7; hTMEM63B ΔM310-Q329, n = 7) and quantification of P_{50} from each individual patches.

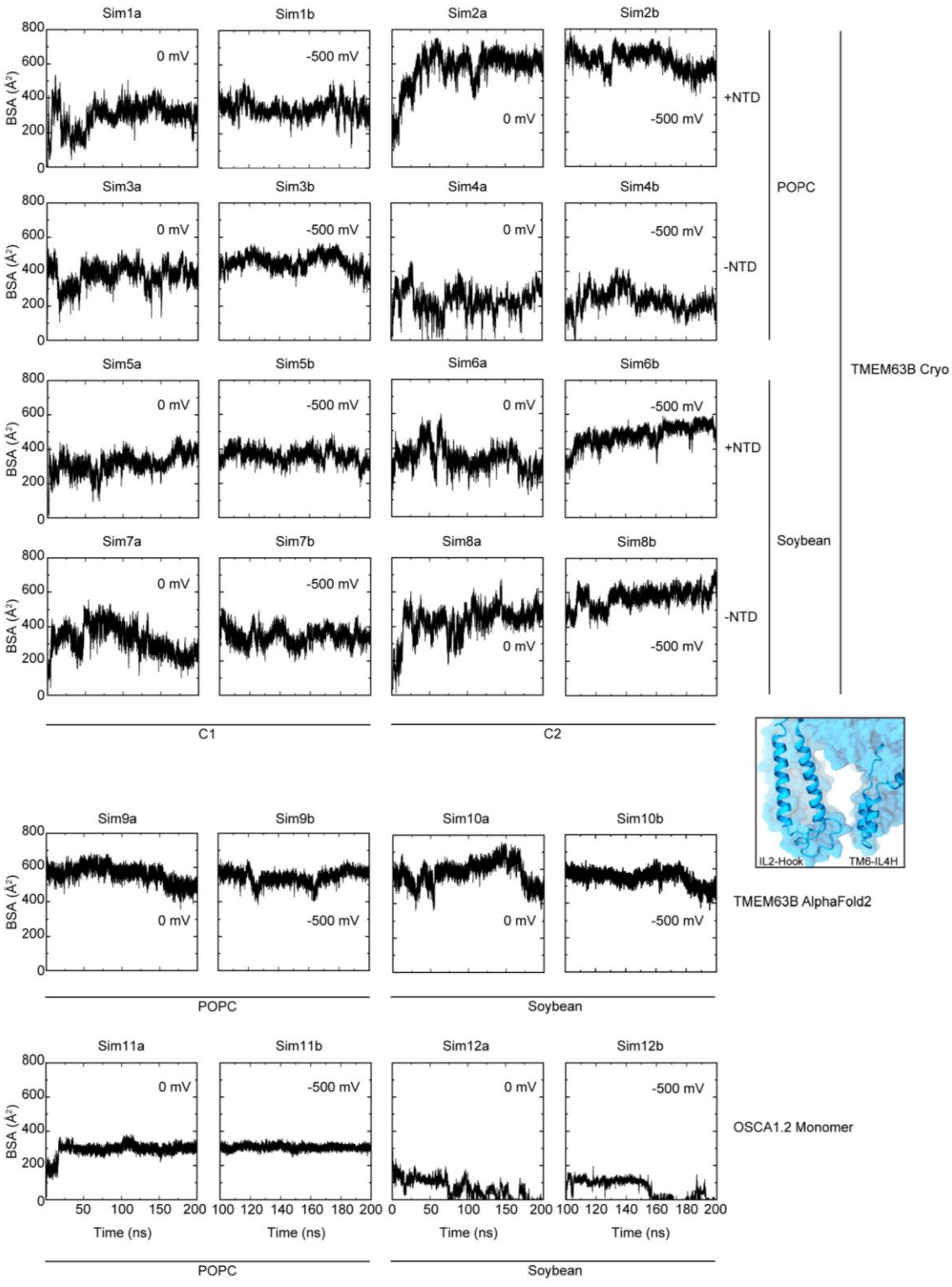

Sim1a
Sim1b
Sim2a
Sim2b
Sim3a
Sim3b
Sim4a
Sim4b
Sim5a
Sim5b
Sim6a
Sim6b
Sim7a
Sim7b
Sim8a
Sim8b
Sim9a
Sim9b
Sim10a
Sim10b
Sim11a
Sim11b
Sim12a
Sim12b
0 mV
-500 mV
BSA (Å²)
0
200
400
600
800
+NTD
-NTD
POPC
Soybean
TMEM63B Cryo
C1
C2
IL2-Hook
TM6-IL4H
TMEM63B AlphaFold2
OSCA1.2 Monomer
50
100
150
200
100
120
140
160
180
200
Time (ns)

Figure A.17. Predicted contacts between the IL2 domain and helices TM6b and ILH4, related to Fig. A. 6.

Buried surface area (BSA) is used to quantify the interaction between the IL2 hook and the TM6b and IL4H region near the bottom of the pore for TMEM63B and OSCA1.2 models. Inset image shows top view of this interaction in the TMEM63B cryo-EM model used for simulations (*middle right*). Initial BSA values were ~ 20 $Å^2$ for the TMEM63B cryo-EM model, ~ 600 $Å^2$ for the TMEM63B AF2 model, and ~ 260 $Å^2$ for OSCA1.2. Values shown for all simulations listed in Table A.2.

Table A.1. Cryo-EM data collection, refinement, and validation statistics

Data collection/processing	**TMEM63A in nanodiscs (EMD-28153) (PDB 8EHW)**	**TMEM63B in LMNG**	**TMEM63B in nanodiscs**	**TMEM6B in LMNG (EMD-28154) (PDB 8EHX)**
Magnification	105,000	36,000	105,000	105,000
Voltage (kV)	300	200	300	300
Electron exposure (e-/Å^2)	50.705	52.563	53.048	53.048
Defocus range (μm)	-1 to -2	-1.4 to -2.5	-0.8 to -2.1	-0.8 to -2.1
Number of frames	50	50	50	50
Pixel size (Å)	0.825	1.1	0.825	0.825
Micrographs (no.)	10,836			8,640
Initial particles (no.)	4,552,788			2,433,802
Final particles (no.)	198,144			64,399
Symmetry imposed	C1			C1
Map resolution (Å)	3.82			3.62
FSC threshold	0.143			0.143
Map sharpening B factor (Å^2)	-198.6			-131.2
Refinement				
Initial model	AF2			AF2
Model resolution (Å)	4.03			3.87
FSC threshold	0.5			0.5

Continued

Continued Table A.1

Model composition			
Non-hydrogen atoms	5,150		5,153
Protein residues	630		631
R.m.s deviations			
Bond lengths (Å)	0.003		0.003
Bond angles (°)	0.769		0.648
Validation			
MolProbity score	2.19		2.13
Clashscore	15.33		13.01
Poor rotamers (%)	0.18%		0
Ramachandran plot			
Favored (%)	91.67%		91.33%
Allowed (%)	8.33%		8.51%
Disallowed (%)	0		0.16%

Table A.2. Simulations summary for each system

Label	Protein	Membrane	NTD	IL1	Voltage (mV)	Length (ns)	Start	# atoms	Dimensions ($Å^3$)
Sim1a	TMEM63B (Cryo)	POPC	+	C1	0	200	-	205,840	133 × 133 × 136
Sim1b					-500	100	Sim1a (100 ns)		
Sim2a				C2	0	200	-	211,763	135 × 135 × 142
Sim2b					-500	100	Sim2a (100 ns)		
Sim3a			-	C1	0	200	-	209,874	135 × 135 × 138
Sim3b					-500	100	Sim3a (100 ns)		
Sim4a				C2	0	200	-	209,954	133 × 133 × 136
Sim4b					-500	100	Sim4a (100 ns)		
Sim5a		Soybean	+	C1	0	200	-	203,823	133 × 133 × 136
Sim5b					-500	100	Sim5a (100 ns)		
Sim6a				C2	0	200	-	205,514	135 × 135 × 142
Sim6b					-500	100	Sim6a (100 ns)		
Sim7a			-	C1	0	200	-	203,993	135 × 135 × 142
Sim7b					-500	100	Sim7a (100 ns)		
Sim8a				C2	0	200	-	204,024	133 × 133 × 136
Sim8b					-500	100	Sim8a (100 ns)		

Continued

Continued Table A.2

Sim9a	TMEM63B (AF2)	POPC	+	C2	0	200	-	200,163	135 × 135 × 135
Sim9b					-500	100	Sim9a (100 ns)		
Sim10a		Soybean			0	200	-	209,874	135 × 135 × 135
Sim10b					-500	100	Sim10a (100 ns)		
Sim11a	OSCA1.2 Monomer (PDB: 6MGV)	POPC		-	0	200	-	187,941	130 × 133 × 132
Sim11b					-500	100	Sim11a (100 ns)		
Sim12a		Soybean			0	200	-	194,793	135 × 135 × 132
Sim12b					-500	100	Sim12a (100 ns)		
Total Simulation Time						3,600			

C1 and C2 indicate the conformation of IL1 predicted by AF2 multimer prediction and monomer prediction, respectively.