Structure of Myo7b/USH1C complex suggests a general PDZ domain binding mode by MyTH4-FERM myosins

Jianchao Li1,†, Yunyun He2,†, Meredith L. Weck3, Qinig Lu4,5, Matthew J. Tyska6, and Mingjie Zhang7,8,9,10

1Division of Life Science, State Key Laboratory of Molecular Neuroscience, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong; 2Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, TN 37240-7935; and 3Center of Systems Biology and Human Health, School of Science and Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

Edited by James A. Spudich, Stanford University School of Medicine, Stanford, CA, and approved April 6, 2017 (received for review February 9, 2017)

Unconventional myosin 7a (Myo7a), myosin 7b (Myo7b), and myosin 15a (Myo15a) all contain MyTH4-FERM domains (myosin tail homology 4-band 4.1, ezrin, radixin, moesin; MF) in their cargo binding tails and are essential for the growth and function of microvilli and stereocilia. Numerous mutations have been identified in the MyTH4-FERM tandems of these myosins in patients suffering visual and hearing impairment. Although a number of MF domain binding partners have been identified, the molecular basis of interactions with the C-terminal MF domain (CMF) of these myosins remains poorly understood. Here we report the high-resolution crystal structure of Myo7b CMF in complex with the extended PDZ3 domain of USH1C (a.k.a., Harmonin), revealing a previously uncharacterized interaction mode both for MyTH4-FERM tandems and for PDZ domains. We predicted, based on the structure of the Myo7b CMF/USH1C PDZ3 complex, and verified that Myo7a CMF also binds to USH1C PDZ2 using a similar mode. The structure of the Myo7b CMF/USH1C PDZ complex provides mechanistic explanations for >20 deafness-causing mutations in Myo7a CMF. Taken together, these findings suggest that binding to PDZ domains, such as those from USH1C, PDZD7, and Whirlin, is a common property of CMFs of Myo7a, Myo7b, and Myo15a.

Microvilli and stereocilia are both actin bundle-based, finger-like protrusions found on apical surfaces of many epithelial cells (1). Microvilli line the apical surface of epithelial cells forming a densely packed structure known as the brush border in tube-like tissues, such as intestines, kidney, and lung (2–5). The best known function of microvilli is to massively increase membrane surface area of these tissues to promote solute exchange, although these protrusions may also allow cells to communicate with their extracellular surroundings (6). Stereocilia, on the other hand, are composed of several rows of protrusions with graded height forming a staircase-like structure on the apical surface of inner ear hair cells, which collectively function to sense sound waves (7, 8). Despite clear morphological and functional differences, microvilli and stereocilia share certain features at the molecular level. For example, the packing and organization of microvilli and stereocilia both rely on homologous cadherin-based tip-link complexes that target to the distal ends of these protrusions (9, 10). Additionally, a set of homologous unconventional myosins—including Myo1, Myo3, Myo6, and Myo7—are shared by and critical for the development, main-
tenance, and function of microvilli and stereocilia (1, 11, 12).

This study focuses on Myo7b, an unconventional myosin enriched in the microvilli of transporting epithelial cells (11, 13). Myo7b is characterized by a pair of MyTH4-FERM tandems in its tail cargo-binding domain. In addition to Myo7b, Myo7a, Myo10, and Myo15a also contain one or two MyTH4-FERM tandems in their tail domains (14). A common property of MyTH4-FERM myosins is their involvement in the formation or elongation/sta-
bilization of actin-bundle supported structures, including filopo-
dia, microvilli, and stereocilia (15). This property first appeared in ancient MyTH4-FERM myosins and is conserved over 1 billion years of evolution (14, 16, 17). Mutations of MyTH4-FERM myosins cause a broad spectrum of human diseases, including hearing loss, vision defects, digestive disorders, and cancers (18–21), often because of defects in actin bundle-supported protrusions (i.e., filo-
podia, microvilli, or stereocilia). The specific functions of Myo7a, Myo7b, Myo10, and Myo15a in regulating actin-bundle protrusion structures in various tissues must be a result of, at least in part, the unique cargo binding properties of their MyTH4-FERM domains. For example, the N-terminal MyTH4-FERM tandem (NMF) of Myo7b binds to the central domain (CEN) of USH1G (a.k.a., Sans) with very high affinity (22). Mutations of Myo7a or USH1G that weaken or disrupt the Myo7a/USH1G interaction can cause stereocilia deformation and hearing impairment (22–24). The Myo7b NMF binds to the central region of ANKS4B, which shares high homology with USH1G with a similar structural mechanism as the Myo7a/USH1G interaction (25, 26). The C-terminal MyTH4-FERM tandem (CMF) of Myo7b has been shown to bind to the C-terminal PDZ domains of USH1C (10, 25, 26). The direct binding of USH1C PDZ domains to Myo7b CMF is highly unexpected, as Myo7b CMF does not contain any recognizable C-terminal or internal PDZ-binding motif (PB) (27). Additionally, binding of one or a pair of PDZ domains to a MyTH4-
FERM tandem is also unique with respect to all previously character-
ized MyTH4-FERM target interactions (22, 26, 28–30). Elucidating the binding mechanism for the Myo7b CMF/USH1C PDZ interaction will not only be helpful for understanding the role of Myo7b in intermicrovillar adhesion complex (IMAC) as-
sembly but also provide valuable insights into possible interactions

Significance

MyTH4-FERM myosins (Myo7a, Myo7b, and Myo15a) regulate actin-bundle protrusion structures in various tissues, including brush border microvilli of intestines and stereocilia of inner ear hair cells. Mutations of the cargo binding MyTH4-FERM tandems of these myosins are frequently associated with human diseases, including hearing loss, vision defects, and digestive disorders, but with poorly understood mechanisms. In this work, we present the high-resolution crystal structure of Myo7b C-terminal MyTH4-
FERM tandem (CMF) in complex with the USH1C PDZ3 domain. The structure, together with biochemical studies, indicates that binding to PDZ domain scaffold proteins is a general property of these myosins. The complex structure reported here also helps to explain why numerous mutations identified in Myo7a CMF can cause deafness and blindness in humans.

Author contributions: J.L., M.J.T., and M.Z. designed research; J.L., Y.H., M.L.W., and Q.L. performed research; J.L., Y.H., M.L.W., M.J.T., and M.Z. analyzed data; and J.L., Y.H., M.L.W., M.J.T., and M.Z. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 5XBF).

1J.L. and Y.H. contributed equally to this work.

2To whom correspondence should be addressed. Email: mzhang@ust.hk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.
1073/pnas.1702251114/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1702251114

PNAS Early Edition | 1 of 10
between MyTH4-FERM tandems from Myo7a to PDZ domains of PDZD7 and USH1C or potential interactions between Myo15a MyTH4-FERM and the PDZ domains of Whirlin.

We performed a detailed biochemical characterization of the interaction between Myo7b CMF and USH1C, and solved the high-resolution crystal structure of Myo7b CMF in complex with the extended PDZ3 domain of USH1C. The complex structure reveals previously uncharacterized interaction modes both for MyTH4-FERM tandems and for PDZ domains. The Myo7b CMF/USH1C PDZ3 complex structure, together with direct biochemical binding experiments, also helps to explain why numerous mutations in the Myo7a CMF may impair its function and thus lead to hearing and vision impairments in patients. Our findings provide a structural framework for future studies of CMFs in Myo7a, Myo7b, and Myo15a.

**Results**

**The Overall Structure of the Myo7b CMF/USH1C PDZ3 Complex.** Our earlier study has shown that Myo7b CMF directly binds to USH1C and the interaction is mediated by Myo7b CMF and the extended PDZ3 domain of USH1C (26) (Fig. 1). To elucidate the molecular basis governing this unexpected interaction mode, we obtained high-quality crystals of the Myo7b CMF/USH1C PDZ3 complex that diffracted to 1.8 Å resolution and solved the complex structure by a single-wavelength anomalous dispersion method using gold derivatives (Table 1).

The overall structure of the Myo7b CMF/USH1C PDZ3 complex reveals previously uncharacterized target-binding modes for MyTH4-FERM tandems as well as for PDZ domains (Fig. 1B). The overall architecture of Myo7b CMF is similar to the structures of other MyTH4-FERM tandems (17, 22, 26, 28, 29), which all contain an N-terminal helical MyTH4 domain and a C-terminal clover-leaf–shaped FERM domain. The USH1C PDZ3 domain in the complex contains a canonical PDZ fold plus an N-terminal extension composed of two short α-helices and an extended C-terminal tail (Fig. 1B). Instead of using its canonical PBM binding αB/βB-groove, the USH1C PDZ3 uses an extended surface at the back of the PBM binding groove to interact with the F3 lobe of Myo7b CMF. The

![Structure of Myo7b CMF/USH1C PDZ3 complex](image-url)
Li et al.

extended C terminus of USH1C PDZ3, which contains a putative PBM sequence (L1-TFF), inserts into the central pocket formed by the three lobes of the FERM domain of Myo7b CMF with matching shape and charge complementation (Fig. 1 B and C, and see below). The USH1C PDZ3 binding mode of Myo7b CMF is significantly different from the target binding modes of other MYTH4-ERM tandems reported in the Myo7a/USH1G, Myo7b/ANKS4B, and Myo10/DCC complexes (Fig. 1 D and E) (22, 26, 28, 29).

There are noticeable differences between the Myo7b CMF structure and all of the rest of MYTH4-ERM structures solved to date. Such large differences presumably prevented us from solving the Myo7b CMF structure using the molecular replacement method with currently available models. We performed a structural analysis of more than a handful of the different MYTH4-ERM structures solved to date (Fig. S1). By aligning the F1 lobes of all of the MYTH4-ERM structures (17, 22, 26, 29), as well as the FERM domain of Moesin (which is a founding member of the FERM domain) (31), it is clear that the relative orientations of the three lobes of each FERM domain are all different from each other (see Fig. S1 and legend for details), indicating that the three lobes of FERM domains have considerable conformational freedoms. Such interlobe conformational differences are likely to be important for the target bindings of FERM domains, as the central interlobe interfaces are used as target binding sites of FERM domains from MYTH4-ERM tandems (22, 26).

The Detailed Interactions Between Myo7b CMF and USH1C PDZ3. In the Myo7b CMF/USH1C PDZ3 complex, the binding interface can be divided into two parts (Fig. 2A). The C-terminal extension of USH1C PDZ3 inserts into the central pocket of the FERM cloverleaf formed by the F1_β1/F5, F2_α1, F3_β1/α1, and F1/F2 linker (Fig. 2B). The USH1C carboxyl group on the C-terminal residue Phe552 forms charge-charge interaction with Lys1918 in as well as hydrogen bonds with Gln1914 and Trp1895 in Trp1895F1/F2_linker (Fig. 2 B and C). Isothermal titration calorimetry (ITC) assays were used to evaluate the mutational effect of the critical residues (Fig. 2E). Substitution of the Lys1918 by Glu totally disrupted the binding (Fig. 2E). The two Phe residues upstream also make extensive hydrophobic interactions with Phe1843 and Lys2086 and Lys1818 and Phe1923 (Fig. 2B). Mutation of the C-terminal Phe552 to Ala also abolished the binding (Fig. 2E). Additionally, a pair of salt bridges between Arg1921 and Asp546 further contributes to the interaction (Fig. 2B). Reverse-charged mutations of each of two residues, namely Arg1921 and Asp546Arg, significantly weakened the binding (Fig. 2E). The extensive charge-charge and hydrophobic interactions between USH1C tail and the central pocket of the FERM domain of Myo7b CMF are consistent with our previous result that the C-terminal extension of USH1C PDZ3 is essential for the Myo7b CMF and USH1C interaction (26). The canonical PDZ fold of USH1C PDZ3 uses its ββ/βC-loop to interact with Myo7b CMF F3 lobe (Fig. 2A). In particular, Glu469 in the PDZ3 ββ/βC-loop forms several hydrogen bonds with the sidechain of Ser2061 and the sidechain and backbone amide of Thr2062 (Fig. 2D). Fittingly, mutation of the Glu469 of PDZ3 to Ala also abolishes the binding (Fig. 2E). Additionally, hydrophobic interactions formed by Pro475 and Ile476 and Tyr2063 further stabilize the PDZ3/CMF interface (Fig. 2D). The N-terminal extension of USH1C PDZ3 was also involved in this interaction (Fig. 2E).

Table 1. Statistics of X-ray crystallographic data collection and model refinement

<table>
<thead>
<tr>
<th>Data collection and refinement</th>
<th>Native</th>
<th>Au-derivative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9787</td>
<td>0.9783</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
<td>P2₁</td>
</tr>
<tr>
<td>Unit cell parameters (Å)</td>
<td>a = 69.6, b = 42.5, c = 118.4; a = 70.2, b = 41.9, c = 119.4; a = γ = 90°, β = 97.5°, γ = 98.3°</td>
<td></td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>50–1.8 (1.83–1.80)</td>
<td>50–2.3 (2.34–2.30)</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>63,640 (3,196)</td>
<td>30,149 (1,513)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.0 (3.0)</td>
<td>12.2 (12.5)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.1 (99.8)</td>
<td>97.7 (98.4)</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>5.9 (49.2)</td>
<td>9.5 (75.3)</td>
</tr>
<tr>
<td><strong>Structure refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50–1.8 (1.83–1.80)</td>
<td>18.71/21.07 (27.15/30.03)</td>
</tr>
<tr>
<td>Rmerge/Rfree (%)</td>
<td>0.011/0.997</td>
<td>37.5</td>
</tr>
<tr>
<td>No. of atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein atoms</td>
<td>4,617</td>
<td></td>
</tr>
<tr>
<td>Ligands</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>265</td>
<td></td>
</tr>
<tr>
<td>No. of reflections</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working set</td>
<td>60,479 (2,956)</td>
<td>3,147 (174)</td>
</tr>
<tr>
<td>Test set</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot regions†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favored (%)</td>
<td>98.1</td>
<td></td>
</tr>
<tr>
<td>Allowed (%)</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Outliers (%)</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Numbers in parentheses represent the value for the highest resolution shell.

*Rmerge = Σ[Ioobs − <Io>|/ΣIoobs], where Io is the intensity of measured reflection and <Io> is the mean intensity of all symmetry-related reflections.

1Rmerge = Σ[Fobs] − |Fobs|/ΣFobs, where Fobs and Fcalc are observed and calculated structure factors.

‡Rmerge = Σ[Fobs] − |Fcalc|/ΣFobs, where T is a test dataset of about 5% of the total unique reflections randomly chosen and set aside before refinement.

§B factors and Ramachandran plot statistics are calculated using MOLPROBITY (54).
PDZ3 is not involved in the Myo7b CMF binding. The weakening of the USH1C PDZ3 and Myo7b CMF interaction induced by truncating the PDZ3 N-terminal extension is likely a result of overall structural stability change of PDZ3. Therefore, the USH1C PDZ3 represents another example of an extended PDZ domain, the function of which requires additional sequences outside the canonical PDZ fold (32). Taken together, the above structural and biochemical results suggest that the two interfaces between USH1C PDZ3 and CMF Myo7b act synergistically to form the stable complex. The interaction between PDZ3 and F3 may help to position the PDZ3 C-terminal extension for inserting into the central pocket of the FERM domain, which in turn stabilizes the PDZ3/ F3 interaction mediated by only a few hydrogen bonds and a small hydrophobic interface. This analysis is consistent with our previous biochemical characterization that neither the C-terminal extension alone nor the PDZ3 without the C-terminal extension can efficiently interact with Myo7b CMF (26).

USH1C is known to undergo extensive alternative splicing (33, 34). These spliced isoforms can be divided into three classes (a-c in Fig. 2F). The c isoforms do not contain PDZ3 and thus are not expected to bind to Myo7b CMF (as well as Myo7a CMF, see below). The isoforms a1, a2, and b1 contain the extended C-terminal tail as characterized in this study, and these isoforms are expected to bind to Myo7b CMF (Fig. S2). Although the isoforms b2–b4 contain a PDZ3 with C-terminal extensions, the C-terminal extensions of these isoforms do not have the characteristic “LTFF” sequence required for inserting into the FERM central pocket. Therefore, these isoforms are also not expected to bind to Myo7b CMF. Indeed, replacing the C-terminal extension of isoform a1 with that of isoform b2 resulted in complete loss of binding of USH1C PDZ3 to Myo7b CMF (Fig. 2E). USH1C isoform a1 is known to be expressed in the eye, cochlea, vestibule, heart, kidney, small intestine, and testis, and isoform b1 is expressed in cochlea and vestibule. The tissues expressing the USH1C isoforms a1 and b1 are all known to express either Myo7a or Myo7b (13, 35, 36).

**Fig. 2.** Detailed interaction between Myo7b CMF and USH1C PDZ3. (A) Ribbon representation of the Myo7b CMF/USH1C PDZ3 complex structure with the two major interfaces highlighted with the black boxes. (B) Detailed interaction between USH1C C-terminal tail and Myo7b CMF central pocket. Residues involved in binding are highlighted with stick models. Salt bridges and hydrogen bonds are indicated with dashed lines. (C) Surface electrostatic potential of the Myo7b CMF showing the positively charged central pocket. (D) Detailed interaction between USH1C PDZ3 βββC loop and Myo7b CMF F3 lobe. (E) Summary of dissociation constants showing that mutations of the critical residues in the interface either weakened or even abolished the binding. (F) Schematic diagrams showing domains organizations of different classes of USH1C spliced isoforms.

The Myo7b/USH1C Interaction Is Critical for IMAC Formation and Function in CACO-2 

Next, we used CACO-2 

BBE cells to evaluate the role of Myo7b/USH1C binding in IMAC function and brush border assembly. Our previous studies showed that knockdown (KD) of Myo7b in CACO-2 

BBE impairs IMAC formation, as shown by loss of IMAC component (USH1C, CDHR2 and CDHR5) enrichment at microvillar tips; these phenotypes were rescued by reexpression of WT Myo7b (37). We used a similar KD/rescue approach to evaluate the function of several mutants of Myo7b defective in USH1C PDZ3 binding (K1918E and R1921E in Myo7b CMF as shown in Fig. 2 and a variant of Myo7b lacking the CMF, denoted ΔCMF). CACO-2 

BBE cells were subjected to Myo7b KD and then infected with lentivirus encoding Myo7b rescue constructs. After allowing cells to differentiate for 14 d, they were fixed and stained for USH1C and CDHR2 and then imaged using confocal microscopy. The resulting images were used to calculate Pearson’s correlation coefficient to examine the extent of colocalization between Myo7b constructs and USH1C or CDHR2 in microvilli. As a readout for IMAC function, we also visually scored the fraction of cells that exhibit robust microvillar clustering, as previously described (37). As a positive control, we reexpressed EGFP-tagged WT full-length Myo7b, which strongly colocalized with USH1C (Fig. 3A, J, and B) and CDHR2 (Fig. 3C and Fig. S3) in microvilli, and rescued microvillar clustering in KD cells (Fig. 3A, J, and D). In contrast, neither the point mutants nor ΔCMF fully rescued Myo7b KD-induced loss of IMAC component localization or microvillar clustering (Fig. 3 and Fig. S3). The K1918E mutant and the ΔCMF mutant, which are both unable to bind to USH1C PDZ3, completely failed to rescue the normal microvillar enrichment of USH1C and clustering of microvilli (Fig. 3A, 2 and 3). The R1921E mutant, which is still able to bind weakly to USH1C PDZ3, showed a slightly less severe phenotype in both assays (Fig. 3), suggesting that strong Myo7b CMF/USH1C PDZ3 binding is important for stable localization and function of USH1C in microvilli.
Myo7a CMF Also Binds to USH1C PDZ3. Myo7a and Myo7b are homologous and their CMFs share 57% sequence identity (Fig. 4A). Moreover, mapping of Myo7a CMF amino acid residues onto the structure of Myo7b CMF revealed that the residues responsible for USH1C PDZ3 binding (i.e., both the PDZ3 domain binding surface on F3 and the C-terminal tail binding pocket at the interface of the three FERM lobes) are nearly identical in Myo7a (Fig. 4B and Fig. S4), suggesting that USH1C PDZ3 may also interact with Myo7a CMF. We tried to test the possible interaction between Myo7a CMF and USH1C PDZ3 using purified recombinant proteins. Although isolated Myo7a CMF could not be expressed in its soluble forms in Escherichia coli, coexpression of USH1C PDZ3 with Myo7a CMF greatly enhanced the solubility of Myo7a CMF (Fig. 4C). The resulting Myo7a CMF/USH1C PDZ3 complex can be purified by immobilized metal ion-affinity chromatography followed by a size-exclusion column (Fig. 4D), confirming the direct interaction of the two proteins. We managed to separate Myo7a CMF and USH1C PDZ3 by treating the complex with buffer containing 300 mM NaCl (Fig. 4D3). ITC-based titration experiments using separated proteins showed that USH1C PDZ3 binds to Myo7a CMF with a $K_d \sim 11\mu M$ (Fig. 4E), which is a few fold weaker than that between Myo7b and USH1C ($K_d \sim 1.7\mu M$) and is consistent with the result shown in our earlier pull-down assay (26).

Disease Mutations Impact Myo7 CMF/USH1C PDZ3 Interactions. Mutations in myo7a (USH1B) are estimated to be responsible for $\sim50\%$ of all type I Usher syndrome cases, and the other half are caused by the rest of the four USH1 genes (38). A total of
25 missense mutations in 23 residues are located in Myo7a CMF based on the data extracted from ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/) and HGMD database (www.hgmd.cf.ac.uk/). The high sequence identity between Myo7a and Myo7b allowed us to confidently build the Myo7a CMF structure by homology modeling (Fig. S4) (39). We then mapped the 25 mutations to the modeled structure (Fig. 5 A and B). We can divide the mutations into four categories (Fig. 5 A and B and Fig. S4). The most interesting category (category 1), which is also directly relevant to this study, contains four mutations (K2021R, G2163S, L2186P, and L2193F) located at the PDZ3 binding interface. Category 2 includes mutations within the folding core of each of the subdomains. These variants are likely to affect the folding of each individual subdomain. In category 3 are mutations at the subdomain–subdomain interfaces, and these mutations may destabilize the MyTH4-FERM supramodule formation. Interestingly, the residues in Myo7a NMF, which correspond to E1812 and P1887 in MyTH4 domain core and to R1883 at MyTH4-F1 interface of Myo7b CMF, are also found to be mutated in USH1B patients, highlighting the importance of the correct folding of the MyTH4 domain and the formation of MyTH4-FERM supramodule in the functions of Myo7a MyTH4-FERM tandems. The fourth category of mutations cannot be well explained by the current structural model, and residues in this category are located on the surface of Myo7a CMF (Y1719C) or in the flexible loops that cannot be visualized in our Myo7b CMF crystal structure (R1743W, R1873Q/W, and A2204P). It is noted that impact of the Y1719C mutation is ambiguous, because this mutation has been reported to be a common polymorphism in the Moroccan population (40). There is also a de novo mutation (W1642C)
found in Myo7b CMF, which is proposed to be linked with autism (41). W1642 is partially exposed and the mutation may destabilize MyTH4α1. However, the potential impact of the W1642C mutation in nervous systems requires further investigation.

Given the extremely high structural similarity between Myo7a and Myo7b and the easy availability of Myo7b CMF, we first tested some of these mutations using Myo7b as the surrogate. We focused on the four Myo7a CMF mutations, of which the corresponding sites are located at the PDZ3 binding interface of Myo7b CMF. Substitutions of K1918 with Arg or L2193 with Phe (corresponding to K2021R and L2193F in Myo7a CMF indicated in Fig. 5B, the two residues intimately involved in USH1C C-terminal interaction) (Fig. 2B), either abolished or weakened Myo7b CMF’s binding to PDZ3 (Fig. 5C). Although the two mutations do not dramatically change the sidechain properties (maintaining the charged or hydrophobic properties), substitutions with larger sidechains will decrease the size of the central pocket and thus impair the USH1C PDZ3 C-terminal tail binding. L2083 (corresponding to L2186 on Myo7a CMF) is also located at FERM-domain central pocket and participates the hydrophobic interaction with USH1C C-terminal extension. The substitution of L2083 with Pro led to formation of inclusion bodies in the process of protein expressions in E. coli (Fig. 5C), which may indicate that such mutation can impair the folding of the F3 lobe of Myo7b CMF and thus prevent USH1C PDZ3 from binding. At the PDZ3/FERM F3 lobe interface, the G2060S mutation (corresponding to G2163 on Myo7a CMF) also totally abolished the interaction between Myo7b CMF and USH1C PDZ3 (Fig. 5C). We also tested another disease related mutation G2084D (corresponding to G2187 on Myo7a CMF), which is immediately following L2083 and is in the vicinity of the PDZ3 tail binding pocket. We found that the G2084D mutation also totally disrupted the interaction between Myo7b CMF and USH1C PDZ3 (Fig. 5C).

We next sought to test these disease mutations directly on Myo7a CMF. With the help of USH1C PDZ3 coexpression, we can purify four of the five mutants tested and used the ITC-based assay to assess the impact of each of the mutation on USH1C PDZ3 binding, and the result is summarized in Fig. 5C. As we have anticipated, the impact of each of the Myo7a CMF mutation on USH1C PDZ3 binding is essentially the same as what we have observed for the corresponding mutations in Myo7b CMF. The K2021R, G2163S and G2187D mutations totally abolished the interaction between Myo7a CMF and USH1C PDZ3, and the L2193F mutation weakened the interaction (Fig. 5C). As in Myo7b CMF, we were not able to express the L2186P mutant of Myo7a CMF in soluble form either.

**Fig. 5.** Missense mutations in Myo7a CMF found in USH1B patients. (A) Ribbon representation of the Myo7a CMF structural model. The 20 missense mutation sites in USH1B patients are highlighted with spheres and colored in magenta, green, yellow, and pink, corresponding to the four categories. A de novo mutation found in Myo7b (W1642C) is shown in red. (B) Schematic cartoon diagram showing the distributions of the USH1B mutations mapped on to the structural model of Myo7a CMF. (C) Summary of impacts of some USH1B mutations on the Myo7a CMF/USH1C PDZ3 interaction. The matching mutations were also made in Myo7b CMF corresponding to each of the USH1B mutations, and their impacts on USH1C PDZ3 binding are measured.

Formation of the Myo7b/ANKS4B/USH1C Tripartite Complexes Requires all Three Binary Interactions. Up to this stage, we had quantitatively studied all binary interactions between pairs of proteins within the
Myo7b/ANKS4B/USH1C tripartite complexes that are expected to form in cells (Fig. 6A). We next tried to reconstitute the tripartite complexes in vitro. After numerous trials, we successfully purified large quantities of Myo7b tail containing the two MyTH4-FERM tandems (denoted as Myo7b NCMF) with the help of coexpression of USH1C PDZ3. When loaded on a size-exclusion column, the 1:1:1 mixture of Myo7b NCMF, the full-length ANKS4B, and the full-length USH1C were eluted significantly earlier than each individual component (Fig. 6B). SDS/PAGE analysis showed that the complex consisted of all three proteins (Fig. 6C), indicating the formation of the Myo7b/ANKS4B/USH1C tripartite complex.

The Myo7b/ANKS4B/USH1C tripartite complex contains three pairs of binary interactions (Fig. 6A). Myo7b uses its NMF to interact with ANKS4B CEN domain and its CMF to interact with USH1C extended PDZ3 domain, each with a micromolar dissociation constant. ANKS4B also uses its SAM-PBM to interact with USH1C NPDZ1 with a $K_d$ at nanomolar range. Next, we sought to determine how each pair of binary interactions contributes to the tripartite complex formation. Our current and earlier structural studies allowed us to design mutations to specifically disrupt or weaken each pair of binary interaction, and we used pull-down assays to evaluate each of the mutations on the Myo7b/ANKS4B/USH1C tripartite complex.

We have demonstrated in this study, by solving the atomic structure and performing detailed biochemical characterization, that Myo7b CMF can specifically bind to USH1C PDZ3 using a previously uncharacterized mode for MyTH4-FERM tandems, as well as for PDZ domains. We have also shown that the highly homologous Myo7a CMF can bind to USH1C PDZ3 as well. Given the domain organization similarities among USH1C, PDZD7, and Whirlin, we speculate that the recently reported Myo7a/PDZD7 complex (42, 43) and the Myo15a/Whirlin

**Discussion**

We have demonstrated in this study, by solving the atomic structure and performing detailed biochemical characterization, that Myo7b CMF can specifically bind to USH1C PDZ3 using a previously uncharacterized mode for MyTH4-FERM tandems, as well as for PDZ domains. We have also shown that the highly homologous Myo7a CMF can bind to USH1C PDZ3 as well. Given the domain organization similarities among USH1C, PDZD7, and Whirlin, we speculate that the recently reported Myo7a/PDZD7 complex (42, 43) and the Myo15a/Whirlin...
complex (44), existing in stereocilia of hair cells, may also form via a similar MyoTH4-FERM/PDZ interaction observed in the Myo7b/USH1C and Myo7a/USH1C complexes. As we have illustrated in the present study, the high-resolution structure of the Myo7b CMF/USH1C PDZ3 complex structure may help explain why many of the USH1B mutations in the Myo7a CMF cause deafness and blindness in human patients. The Myo7b CMF/USH1C PDZ3 structure can also serve as a template for understanding the potential impacts of mutations found in Myo15a CMF in nonsyndromic deaf patients.

The structure of the Myo7b CMF/USH1C PDZ3 complex also extends our current knowledge about the target recognition modes of FERM and PDZ domains, both of which are highly versatile protein–protein interaction modules abundantly present in mammalian proteomes. We summarized all reported FERM/target bindings with structure coordinates deposited in PDB without considering Talin FERM because of its very different F1 lobe organization (Fig. S5 and Table S1). A total of 10 different target binding sites can be found in all three lobes of FERM domains. Additionally, the interlobe interfaces of FERM domains can also function as target binding sites. For example, the F1/F3 interface (site 2 in Fig. S5) can be protein or lipid binding sites of FERM domains (45, 46). The F1/F2/F3 interface in FERM domains of myosins binds to different protein targets with distinct mechanisms (22, 26; and present study). We also note that some targets simultaneously bind to more than one site on FERM domains, resulting in higher binding affinity and specificity. For example, we show in this study that USH1C binds to site 5 in F3 and site 6 formed by the F1/F2/F3 interface of Myo7b C-terminal FERM domain. In a number of FERM domain proteins with auto-inhibitory tails, each tail is known to simultaneously bind to sites 3 and 4 in the F3 and F2 lobes, respectively (31, 47, 48). The Crumbs tail is shown to simultaneously bind to site 1 in the F3-lobe and site 2 in the F1/F3 interface (49). Among all of the sites identified, site 1 in the α1/β5-groove of F3 is the most frequently used target binding site of FERM domains. The summary in Fig. S5 highlights the remarkable capacities of FERM domains in binding to very diverse target molecules, each with their own specificities.

Materials and Methods

Cloning and Constructs for Recombinant Protein Expression. The human cDNA construct of Myo7b used in this study was GI: 122937511. The protein coding sequences of Myo7b CMF (residues 1701–2215) and NCMF (residues 960–2116) were amplified by PCR from human Myo7b and cloned into a pET vector. The coding sequences of Myo7a CMF (residues 1704–2215) and NCMF (residues 965–2215) were amplified by PCR from a mouse USH1B gene (GI: U81453.1). The coding sequences of USH1C isoform b2 PDZ3 (residues 728–859) was amplified by PCR from mouse USH1C isoform b2 cDNA (a kind gift from Christine Pettit, Pasteur Institute, Paris). Other constructs used have been described in our previous study (26). Point mutations were generated by PCR-based site-directed mutagenesis.

Cloning and Constructs for CACO-2mac Transfection. Full-length (residues 1–2116) and ΔCMF (residues 1–1587) Myo7b constructs were generated by PCR and TOPO cloned into the pCR8 Gateway entry vector (Invitrogen). Point mutations and refractory silent mutations were introduced using QuickChange site-directed mutagenesis (Agilent). All entry vectors were verified by DNA sequencing. The entry vectors were then shuttled into the destination vector pDONR207-EF1-GFP-C1 (25) that was Gateway-adapted using the Gateway vector conversion kit (Invitrogen). A nontargeting scramble control shRNA (Addgene; plasmid 1864) and Myo7b KD shRNA clones were expressed in pLKO.1, corresponding to TRC clones TRCN0000247713 and TRCN0000247715 (Sigma). TRC clone TRCN0000247713 was used to KD Myo7b in the ΔCMF rescue cell line as this construct lacks the shRNA target sequence. All other Myo7b KD rescue constructs contain refractory silent mutations to TRC clone TRCN0000247715.

Protein Expression and Purification. Recombinant proteins were expressed in BL21 (DE3) E. coli cells. The N-terminal His<sub>6</sub>-tagged proteins were purified using Ni<sup>2+</sup>-nitrilotriacetic acid agarose column followed by another step of size-exclusion chromatography (Superdex 200 column from GE Healthcare) in the final buffer of 50 mM Tris-HCl, 1 mM EDTA, pH 7.8, and 100 mM NaCl. GST-fused proteins were purified by GSH-Sepharose affinity chromatography, followed by a step of size-exclusion chromatography, same as the one described above. For the separation of Myo7a CMF and USH1C PDZ2 from the copurified complex, an additional step of size-exclusion chromatography (Superdex 200 column from GE Healthcare) in the buffer of 50 mM Tris HCl, 1 mM EDTA, 1 mM EDDA, pH 7.8, and 300 mM NaCl was performed, and the fraction containing each of the individual components was collected.

Crystallography. The crystals of Myo7b CMF/USH1C PDZ3 complex (~10 mg/mL in 50 mM Tris, pH 7.8, 100 mM sodium citrate, 1 mM EDTA, and 1 mM DTT buffer) were grown in reservoir solution containing 8% tascimate, pH 7.0, and 20% (vol/vol) PEG3350 by sitting-drop vapor diffusion methods at 16 °C. To prepare Au-derivatives, crystals were soaked in crystallization solution containing additional 2–5 mM KAu(CN)₄ for 1 to 2 d. Crystals were then soaked in reservoir solution containing extra 10% (vol/vol) glycerol for cryoprotection. Diffraction data were collected at the Shanghai Synchrotron Radiation Facility (BL17U and BL19U1) at 100 K. Data were processed and scaled using HKL2000 or HKL3000 (50).

Six gold sites were found by SHELXD (51). Subsequence sites refinement, phase calculation, density modification and initial model building were carried out with AutoCoot (52). Additional manual model building and refinement were completed iteratively using COOT (53) and PHENIX (52). The final model was validated by MolProbity (54). The final refinement statistics are summarized in Table 1. All structure figures were prepared by PyMOL (www.pymol.org).

GST Pull-Down Assay. Formation of the tripartite complex Myo7b/AN548B/USH1C was assayed in 50 mM Tris buffer, pH 7.5, containing 100 mM NaCl, 1 mM EDTA, and 1 mM DTT. In each assay, purified Trx tagged Myo7b NCMF and AN548 FL or its mutant were first mixed with GST-tagged USH1C or its mutant at 1:1:1 ratio (~0.28 nmol each) and then pelleted by adding 30 μL of fresh GSH-Sepharose beads. The pellet, after washing twice with the assay buffer, was analyzed by SDS/PAGE. The Coomassie Brilliant Blue-stained gel was analyzed by densitometry to determine the amount of pulled down Myo7b NCMF.

Analytical Gel-Filtration Chromatography. Protein samples (typically 100 μL at a concentration of 20 μM pre-equilibrated with the column buffer) were injected into an AKTA FPLC system with a Superose 12 10/300 GL column (GE Healthcare) using the column buffer of 50 mM Tris, pH 7.8, 1 mM EDTA, and 100 mM NaCl.

ITC Assay. ITC measurements were carried out on a VP-ITC Microcal calorimeter (Microcal) at 25 °C. Titration buffer contained 50 mM Tris-HCl, pH 7.8, 1 mM EDTA, 1 mM DTT, and 100 mM NaCl. Each titration point was performed by injecting a 10-μL aliquot of a protein sample from a syringe into a protein sample in the cell at a time interval of 120 s to ensure that the titration peak returned to the baseline. The titration data were analyzed by Origin7.0 (Microcal).

Lentivirus Production and Stable Cell Line Generation. CACO-2mac and HEK293FT cells were cultured at 37 °C and 5% CO<sub>2</sub> in DMEM with high glucose and 2 mM l-glutamine supplemented with 20% FBS or 10% FBS, respectively. Lentiviral particles were generated by cotransfecting HEK293FT cells with overexpression (pINDUCER20-EGFP-C1) or KD (pLKO.1) plasmids along with the packaging pSPAX2 and envelope pMD2.G plasmids using FuGENE 6 (Promega). Cells were incubated for 2 d to allow for lentivirus production before collecting the media containing lentiviral particles. Media were centrifuged at 500 x g for 10 min at 4 °C to remove cells and concentrated using Lenti-X Concentrator (Clontech). For generation of stable cell lines, CACO-2mac cells were grown to 90% confluency in T25 flasks. For lentivirus transduction of CACO-2mac cells, the media were supplemented with 8 μg/mL polybrene and incubated with lentivirus overnight. The following day, cells were resuspended into T75 flasks and grown for 3 d. Cells were then resuspended into T182 flasks with media containing 10 μg/mL puromycin or 1 mg/mL G418 and grown to saturation for stable integration.

Light Microscopy and Image Analysis. CACO-2mac cells were washed once with warm PBS and fixed with 4% paraformaldehyde in PBS for 15 min at 37 °C. After fixation, cells were washed three times with PBS and permeabilized with...
0.1% Triton X-100 in PBS for 15 min at room temperature. Cells were then washed three times with PBS and blocked overnight with 5% BSA in PBS at 4°C. Cells incubated with PBS and immunostaining was performed using anti-Ush1C (1:70; Sigma cat#HPA027398), and anti-CDHR2 (1:75; Sigma cat#HPA012569), at room temperature for 2 h. Coverslips were then washed three times with PBS and incubated with Alexa Fluor-647 goat anti-rabbit and Alexa Fluor-568 phalloidin (1:200; Life Technologies) in PBS for 1 h at 37°C.

Coverslips were then washed three times with PBS and blocked overnight with 5% BSA in PBS at 0.1% Triton X-100 in PBS for 15 min at room temperature. Cells were then washed three times with PBS and coverslips were mounted using ProLong Gold Antifade Mountant (Life Technologies). Cells were imaged using a Nikon A1R microscope using confocal microscopy equipped with 488-, 561-, and 647-nm excitation lasers and a 100×/1.49 numerical aperture (NA) APO total internal reflection fluorescence objective. Image acquisition was controlled with Nikon Elements software. Images were contrast enhanced and analyzed using ImageJ software (NIH). Pearson’s correlation coefficients were calculated using JACOP (Just Another Colocalization Plugin). Individual cells were blindly scored for microvillar clustering using the phalloidin signal. For all figures, error bars indicate that SD and n values are reported in the figure legend. All graphs were generated, and statistical analyses performed in Prism v.7 (GraphPad Software, Inc.). Unpaired t tests were used to determine statistical significance between reported values.

ACKNOWLEDGMENTS. We thank the Shanghai Synchrotron Radiation Facility BL19U1 and BL17U1 for X-ray beam time. This work was supported by the Research Grants Council of Hong Kong Grants 641113, 16103614, AoE-M09-12, and 113-60712/R, and a 973 program grant from the Minister of Science and Technology of China (Grant 2014CB910204; to M.Z.); Research Grants Council of Hong Kong Grant 16149516 (to O.L.); an American Heart Association predoctoral fellowship (to M.L.W.); National Research Service Award Predoctoral Fellowship F31DK108528 (to M.L.W.); and NIH Grants R01DK057555 and R01DK0959811 (to M.J.T.). M.Z. is a Kerry Holdings Professor in Science and a Senior Fellow of the Institute for Advanced Study at the Hong Kong University of Science and Technology.