FEBS Letters

FEBS PRESS science publishing by scientists

Conformational switch of harmonin, a submembrane scaffold protein of the hair cell mechanoelectrical transduction machinery

Amel Bahloul[1,2,3], Elise Pepermans[1,2,3], Bertrand Raynal[4], Nicolas Wolff[5], Florence Cordier[5], Patrick England[4], Sylvie Nouaille[1,2,3], Bruno Baron[4], Aziz El-Amraoui[1,2,3], Jean-Pierre Hardelin[1,2,3], Dominique Durand[6] and Christine Petit[1,2,3,7]

1 Unité de Génétique et Physiologie de l'Audition, Institut Pasteur, Paris, France
2 UMRS1120, Institut National de la Santé et de la Recherche Médicale (INSERM), Paris, France
3 Sorbonne Universités, UPMC Université Paris 6, Paris, France
4 Plateforme de Biophysique Moléculaire, Institut Pasteur, Paris, France
5 Unité de RMN des Biomolécules, Institut Pasteur, Paris, France
6 Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Sud, Université Paris-Saclay, Gif-sur-Yvette, France
7 Collège de France, Paris, France

Correspondence
A. Bahloul and C. Petit, Unité de Génétique et Physiologie de l'Audition, Institut Pasteur, 75015 Paris, France
Tel: +33 1 45 68 88 93
E-mails: abahloul@pasteur.fr; christine.petit@pasteur.fr

(Received 1 April 2017, revised 15 May 2017, accepted 8 June 2017, available online 10 July 2017)

doi:10.1002/1873-3468.12729

Mutations in the gene encoding harmonin, a multi-PDZ domain-containing submembrane protein, cause Usher syndrome type 1 (congenital deafness and balance disorder, and early-onset sight loss). The structure of the protein and biological activities of its three different classes of splice isoforms (a, b, and c) remain poorly understood. Combining biochemical and biophysical analyses, we show that harmonin-a1 can switch between open and closed conformations through intramolecular binding of its C-terminal PDZ-binding motif to its N-terminal supramodule NTD-PDZ1 and through a flexible PDZ2-PDZ3 linker. This conformational switch presumably extends to most harmonin isoforms, and it is expected to have an impact on the interaction with some binding partners, as shown here for cadherin-related 23, another component of the hair cell mechanoelectrical transduction machinery.

Keywords: conformation switch; PDZ domain; Usher syndrome

The sensory cells of the auditory organ, the hair cells, convert sound pressure waves into receptor electrical potentials through mechanoelectrical transduction. This process takes place in the hair bundle of the hair cell, an ensemble of actin-filled, stiff microvilli called stereocilia, organized in three rows of graduated height. The mechanosensitive ion channels involved in this process are located at the tips of the stereocilia of the short and middle rows [1]. They are gated by the tip-links, oblique fibrous links connecting the tip of a stereocilium to the side of the adjacent taller stereocilium [2]. Functional characterization of the genes underlying the various genetic forms of Usher syndrome type I (USH1), which defined by profound congenital deafness, vestibular dysfunction, and early-onset sight loss, has shown that USH1 proteins are critical components of the mechanoelectrical transduction machinery [3]. Cadherin-related proteins 23 (cdhr23, USH1D) and 15 (cdhr15, USH1F) make up the upper and lower parts of the tip-links, respectively [4,5]. Myosin VIIa (USH1B), sans (USH1G), and harmonin (USH1C) [6–12], through their interactions with

Abbreviations
AH, α-helix; CC, coiled-coil; CD, circular dichroism; EOM, ensemble optimization method; HEK, Human embryonic kidney cells; NMR, nuclear magnetic resonance; PBM, PDZ-binding motif; SPR, surface plasmon resonance; TCEP, tris(2-carboxyethyl)phosphine.

cdhr23 and stereocilia actin filaments, are viewed as components of the upper part of the mechanoelectrical transduction machinery.

The PDZ (PSD-95/Dlg/ZO-1) domain-containing submembrane scaffold protein harmonin [13,14] interacts with cdhr23, myosin VIIa, and sans [6,8,9,15–18]. At least 10 different harmonin transcripts have been identified in the cochlea. They result from alternative splicing, and their predicted encoded proteins have been grouped into three subclasses (a, b, and c) on the basis of differences in their modular structures [14,19] (Fig. 1A). All contain an N-terminal domain (NTD) followed by two PDZ domains (PDZ1 and PDZ2) and a predicted coiled-coil domain (CC1). The proteins of a and b subclasses contain an additional PDZ domain (PDZ3) at the C-terminal end, which is preceded, in b isoforms, by a second predicted coiled-coil domain (CC2) and a proline–serine–threonine (PST)-rich sequence (Fig. 1A). PDZ domains mediate protein–protein interactions by binding to PDZ-binding motifs (PBMs), which are short amino-acid (aa) sequences usually located at the protein C terminus [20,21]. Harmonin isoforms of a and b, but not c, subclasses contain a C-terminal class I PBM (Fig. 1A). Deciphering the three-dimension structure of harmonin individual domains and their interactions within the entire protein is expected to improve our understanding of its scaffold function. The structures of the NTD, NTD-PDZ1 supramodule, and PDZ2 or PDZ3 individual domains of harmonin, alone or in complex with the class I PBMs of cdhr23 and sans, have been solved [12,22] (RIKEN Structural Genomics/Proteomics Initiative, Protein Data Bank 1V6B). We now show that harmonin can adopt two different structural states, 'open' and 'closed', as a result of the self-interaction between its PBM and NTD-PDZ1 supramodule, and that the switch between these conformations can affect the interaction with the tip-link component cdhr23.

Materials and methods

Production of harmonin fragments

Different fragments of harmonin-a1 (GenBank NM_023649 were produced and purified as previously described [9]: harmonin-a1 ΔLTFF (amino acids 1-544), NTD-PDZ1 (amino acids 1-203), PDZ2 (amino acids 197-308), PDZ2-Cter (amino acids 197-548), PDZ2-CterΔLTFF (amino acids 197-544), PDZ2-PDZ2/3 L (amino acids 197-380), PDZ2-PDZ2/3 L (amino acids 197-437), and CC (amino acids 298-388). The integrity of these proteins was verified by mass spectrometry (MALDI-TOF; Bruker Daltonics, Germany).

Recombinant expression vectors

A recombinant modified pcDNA3.1 expression vector encoding Flag-tagged full-length harmonin-a1 or flag-tagged harmonin-a1 ΔLTFF, and recombinant pCMV vectors encoding Myc-tagged sans (GenBank NM_176847), the Myc-tagged cytoplasmic fragment of cadherin-related 23 (GenBank AY563163.1), or the Myc-tagged myosin VIIa tail (GenBank AAB03679.1) were engineered for transient expression in transfected HEK293 cells. Cells were transfected with the DNA of interest (2 μg), using Lipofectamine Plus Reagent.

Surface plasmon resonance (SPR) experiments

SPR experiments were performed on a ProteOn XPR 36 system (Bio-Rad) equilibrated with a buffer containing 200 mM KCl, 50 mM Tris pH 8, 0.5 mM TCEP (Tris(2-carboxyethyl)phosphine), and 1 mM EDTA. Biotinylated peptides were captured on a Neutravidin-immobilized NLC sensor chip (Biorad) as ligands, over which the different harmonin fragments were allowed to flow as analytes at a rate of 20 $\mu L \cdot min^{-1}$ for 2 min. The steady-state SPR responses (*R*eq, experimental or extrapolated) were plotted against the concentration (*C*) of analyte and fitted with the following equation: $R_{eq} = (R_{max} * C)/(K_d + C)$, where K_d is the equilibrium dissociation constant and R_{max} is the maximal binding capacity of the surface (BIAevaluation 4.1 software; Biacore).

Ultracentrifugation experiments

Harmonin samples (5–20 μM) in 300 mM NaCl, 50 mM Tris pH 8, 0.5 mM TCEP, 1 mM EDTA were centrifuged in a Beckman Coulter XLI analytical ultracentrifuge at 20 °C in an AN60-Ti rotor equipped with 12-mm double-sector Epon centerpieces. Protein concentration was determined as a function of radial position and time by measuring absorbance at 280 nm. Samples were spun for 8 h at 142 000 × ***g***. We calculated each c(s) distribution with Sednterp with a fitted frictional ratio f/f_0 .

Circular dichroism (CD) experiments

All CD measurements were acquired with an Aviv 215 spectropolarimeter. Far-UV (195–260 nm) spectra were recorded at 20 °C on 10 μM harmonin samples in a cylindrical cell with a 0.2-mm path-length cylindrical cell. Ellipticity was measured every 0.5 nm and averaged over 2 s. The final spectrum of the protein sample was obtained by averaging three successive scans and subtracting the baseline spectrum of the buffer recorded under the same conditions. The CONTIN program was used for quantitative decomposition of the far-UV CD spectrum.

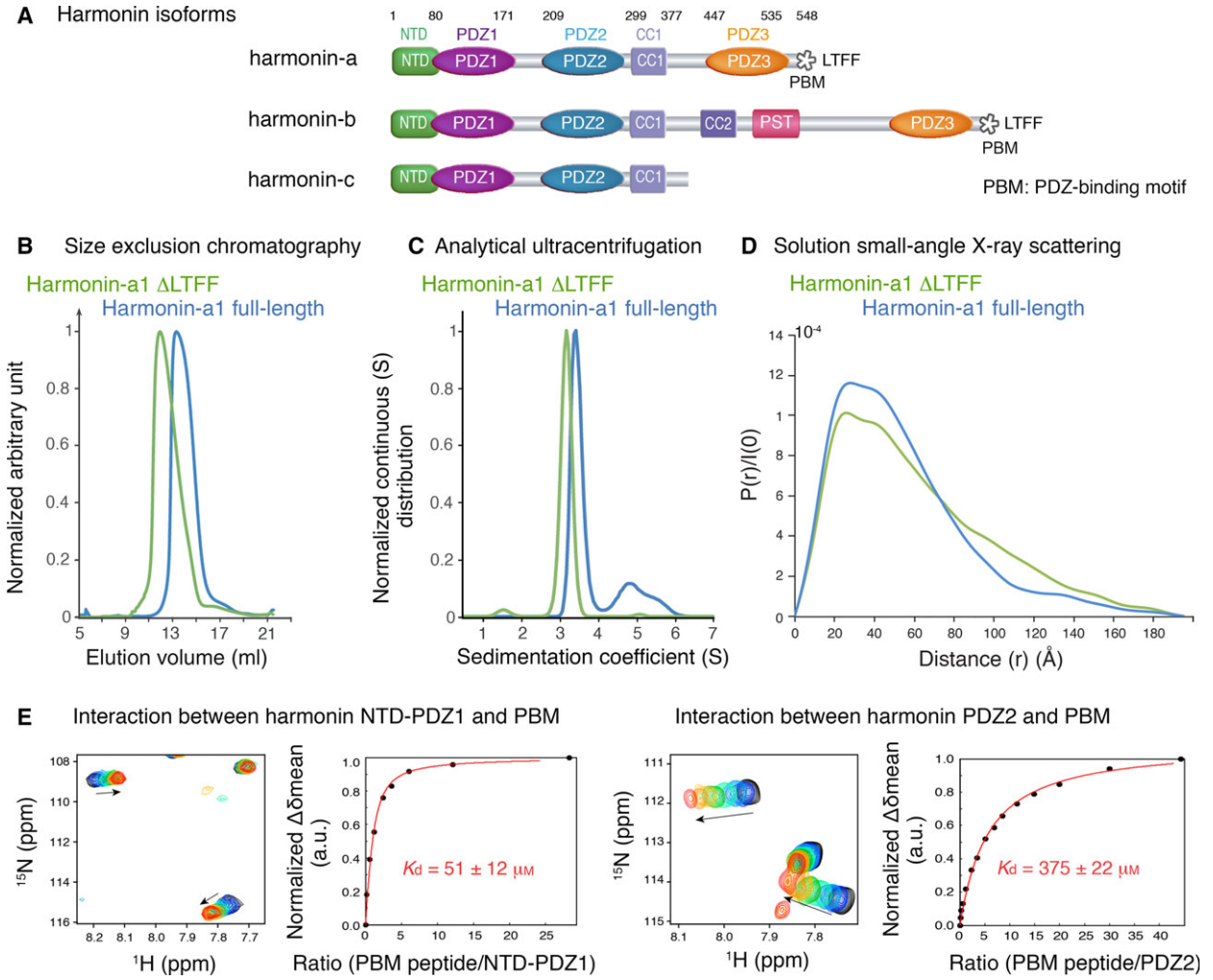

Fig. 1. Role of the C-terminal PDZ-binding motif of harmonin-a1 in the overall conformation of the protein. (A) Diagram of harmonin isoforms. Abbreviations: NTD: N-terminal domain; PDZ: postsynaptic density, disc large, zonula occludens; CC: coiled-coil; PST: proline–serine–threonine rich; PBM: PDZ-binding motif (LTFF). (B) Size-exclusion chromatography (Superdex 200 GL 10/300 column) traces for full-length harmonin-a1 (blue) and harmonin-a1 ΔLTFF (green). The elution volume of harmonin-a1 ΔLTFF is shifted toward higher molecular masses. (C) Analytical ultracentrifugation results for full-length harmonin-a1 (blue) and harmonin-a1 ΔLTFF (green) at a loading concentrations of 10 μM. Plot of the distribution of sedimentation coefficients (s) against normalized continuous s, where s, plotted in Svedberg units (S), is calculated from sedimentation velocity experiments. The single peak shows the highly monodisperse nature and monomeric state of full-length harmonin-a1 and harmonin-a1 ΔLTFF. (D) Small-angle X-ray scattering (SAXS) in solution. P(r) functions were deduced from the SAXS data with GNOM program. P(r) is the pair-distance distribution function as a function of the distance r. Full-length harmonin-a1 and harmonin-a1 ΔLTFF produce different P(r) curves, shown in blue and green, respectively. (E) Interaction between the PBM and NTD-PDZ1 (left panels) or PDZ2 (right panels) domains of harmonin-a1. Representative regions of ^{15}N HSQC (Heteronuclear Single-Quantum Correlation) NMR spectra of NTD-PDZ1 at 90 μM and PDZ2 at 70 μM, with increasing amounts of a synthetic peptide containing the C-terminal 16 amino-acid residues including the PBM (black to red spectra). Peaks displaying significant averaged chemical shift changes upon binding ($\Delta\delta_{av} = [(\Delta\delta_H)^2 + (0.159*\Delta\delta_N)^2]^{1/2}$, where $\Delta\delta_H$ and $\Delta\delta_N$ are the observed ^{1}H and ^{15}N chemical shift changes, respectively) are shown. For the determination of dissociation constants (*Kd*), the normalized $\Delta\delta_{av}$ was plotted as a function of the molar ratio (PBM-containing peptide: PDZ domain-containing harmonin fragment) for several peaks, and fitted using nonlinear regression and assuming a simple complex formation model. The average and standard deviation on six significantly shifting peaks was calculated for each complex.

Small-angle X-ray scattering (SAXS)

Data acquisition and primary reduction

Data for all constructs (full-length harmonin-a1, harmonin-a1 ΔLTFF, NTD-PDZ1, NTD-PDZ1-PDZ2, PDZ2-Cter, PDZ2-Cter-ΔLTFF, CC1 region) were collected on an in-house SAXS instrument (Brüker, Nanostar) with a Micro-star rotating anode (λ = 1.54 Å), a double multilayer focusing optics, scatterless slits, and a Vantec 2000 detector. We placed 30 μL of sample in a quartz capillary thermalized

cell (2.0 mm diameter) inserted into an evacuated sample chamber. A dozen 20-min frames were recorded for the sample and for the buffer. In some cases (full-length harmonin-a1, harmonin-a1 ΔLTFF, PDZ2-Cter, PDZ2-Cter-ΔLTFF), these experiments showed that the solutions contained a fraction of aggregates and/or oligomers. Aggregates/oligomers were separated from isolated proteins, by collecting the SAXS data at the SWING beamline of the SOLEIL synchrotron (Gif-sur-Yvette, France) with an size-exclusion HPLC column (Agilent Bio Sec-3) online with the SAXS measuring cell (a 1.5 mm diameter quartz capillary in an evacuated sample chamber) [23]. All protein constructs were diluted in the buffer, containing 25 mM Tris pH 8.0, 300 mM KCl, and 0.5 mM TCEP.

Data analysis

Data were first analyzed with Foxtrot, the Swing in-house software suite, and with PRIMUS software [24] (https://www.embl-hamburg.de/biosaxs/primus.html). Selected identical frames corresponding to the elution peak were averaged. The scattering of the elution buffer before void volume was recorded and subtracted from all protein scattering curves. The radius of gyration (R_g) was evaluated with the Guinier approximation [25] and from the distance distribution function P(r) calculated with GNOM [26]. Two independent determinations of the molecular mass were obtained with SaxsMow2 [27] and ScÅtter3 [28] available from http://saxs.ifsc.usp.br/ and https://bl1231.als.lbl.gov/scatter/, respectively. The atomic coordinates of the crystal structure of NTD-PDZ1 (PDB code 3K1R) were used to calculate its scattering pattern and to fit the calculated curve to the experimental one. We then used BUNCH [29] to model the conformation in solution of NTD-PDZ1-PDZ2 by a combination of rigid-body and *ab initio* modeling. This program repeated 100 times allowed us to identify the optimal positions and orientations of the NTD-PDZ1 (PDB code 3K1R) and PDZ2 (PDB code 2KBS) domains and to determine the probable conformations of the linker between them. CRYSOL [30] was then used to adjust the curve calculated according to the Bunch model to the experimental curve. An ideal modeling method for highly flexible, partly or fully unstructured proteins such as a1 ΔLTFF or the CC1 region involves the construction of a description from ensembles of conformations with the EOM (Ensemble Optimization Method) package [31,32]. The RANCH program within EOM creates a large (10 000) pool of random conformations from chains of dummy residues describing the unstructured parts of the protein. The atomic model previously obtained for NTD-PDZ1-PDZ2 and the crystal structure of PDZ3 (PDB code 1V6B) were used to describe the structured parts of the protein. The dummy residues were subsequently replaced by complete residues, with the programs PD2 [33] and SCWRL4 [34], before the calculation of all scattering patterns with CRYSOL. The last step was performed by the GAJOE program within EOM, which fits the experimental curve by averaging the calculated scattering patterns of an ensemble of conformations by a genetic algorithm protocol. The quality of fit was assessed by the χ^2 test. The GAJOE program was run 10 times and yielded equally good fits with different ensembles containing a few tens of conformations. Finally, in the case of the full-length harmonin-a1, GAJOE was applied to two pools of conformations, the 'unfolded' pool previously used to describe harmonin-a1 ΔLTFF, and a 'folded' pool consisting of conformations in which PBM interacts with the NTD-PDZ1 module. The volume fractions of the 'unfolded' and 'folded' conformations were further refined with OLIGOMER [24].

Nuclear magnetic resonance (NMR) experiments

NMR samples were prepared in a solution containing 250 mM KCl, 0.5 mM TCEP, 50 mM Tris pH 8, and 1 mM EDTA. For NMR titration experiments, the dried peptide was dissolved in this buffer to achieve a final peptide concentration of 1 mM. NMR experiments were carried out at 25 °C on a Varian Inova 600 MHz spectrometer (Agilent Technologies, Santa Clara, CA, USA).

Co-immunoprecipitation experiments

Human embryonic kidney cells (HEK293) were transfected in the presence of Lipofectamine Plus Reagent (Invitrogen), according to the manufacturer's instructions. The cell pellets were solubilized by sonication in 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.1% Triton, supplemented with an EDTA-free cocktail of protease inhibitors (Roche). Immunoprecipitation was performed with anti-Flag-M2 antibody-conjugated agarose beads (Sigma). The immunoprecipitates were electroblotted onto nitrocellulose sheets and probed with an anti-cMyc polyclonal antibody (Sigma, 1:500 dilution). Horseradish peroxidase-conjugated goat anti-rabbit antibody (Amersham) and the ECL chemiluminescence system (Pierce) were used for detection.

Determination of the proportion of PDZ domain-containing proteins containing a PBM

The Ensembl Gene IDs (ENSMUSG) of all PDZ domain-containing proteins (exhaustive list published by [20]) were retrieved with Biomart. We recovered the corresponding Ensembl protein ID (ENSMUSP) and protein sequence from Ensembl Gene ID. The four C-terminal residues of each protein were categorized as class I, class II, class III PBM, or non-PBM sequences.

Results and Discussion

Role of the C-terminal PBM of harmonin in the overall conformation of the protein

We analyzed the effect of the PBM on overall protein structure in a comparative analysis of the harmonin-a1 isoform (a1 full-length, amino acids 1-548, GenBank accession number NM_023649) and a truncated form of this protein lacking the C-terminal PBM (a1 ΔLTFF, amino acids 1-544) (Fig. 1A). We found that a1 full-length and a1 ΔLTFF had different elution volumes in a size-exclusion chromatography experiment, with a1 full-length having a higher elution volume than a1 ΔLTFF (Fig. 1B), potentially reflecting differences in molecular shape or oligomerization state. We performed analytical ultracentrifugation experiments to distinguish between these two possibilities. Sedimentation velocity experiments at protein concentrations of 1.5–13 μM resulted in curves that, when extrapolated to zero concentration, yielded a sedimentation coefficient in solution ($S_{20,w}$) and a frictional ratio (f/f_0) of 3.5 ± 0.1 S and 1.6, respectively, for a1 full-length, and of 3.1 ± 0.1 S and 1.7, respectively, for a1 ΔLTFF (Fig. 1C). These values indicate that both the full-length harmonin-a1 and a1 ΔLTFF are monomers in solution. Both have an estimated molecular mass of 63 kDa, consistent with the results of sedimentation equilibrium experiments. The lower sedimentation coefficient for a1 ΔLTFF than for the full-length protein indicates that these two proteins have different shapes. The calculated frictional ratio was characteristic of elongated molecules in both cases (typical values for globular proteins are 1.1-1.2), but a1 ΔLTFF was presumably more elongated than the full-length protein.

We, thus, investigated the possible existence of an intramolecular interaction between the PBM of harmonin and its own PDZ domains, by nuclear magnetic resonance (NMR) and surface plasmon resonance (SPR) techniques. A peptide consisting of the C-terminal 16 residues (amino acids 533-548) including the PBM was synthesized, and the NTD-PDZ1 supramodule (amino acids 1-192) and PDZ2 domain (amino acids 197-308, NM_023649) of harmonin-a1 were produced, ^{15}N-labeled, and purified. Analyses of ^{15}N-^{1}H-HSQC spectra and SPR experiments using the polypeptide and either NTD-PDZ1 or PDZ2 showed that the affinity of the interaction between the polypeptide and NTD-PDZ1 was higher (*K*d = 51 ± 12 μM) than the affinity of the interaction between the polypeptide and PDZ2 (*K*d = 375 ± 22 μM) (Fig. 1E and Fig. S1A). Together, these results suggested that the C-terminal PBM of harmonin-a1 can fold back to interact with its own NTD-PDZ1 supramodule.

Harmonin-a can switch between 'open' and 'closed' conformations

To gain insight into the tertiary structure of harmonin-a1, we studied solutions containing the full-length protein, harmonin-a1 ΔLTFF, or a protein fragment lacking the NTD-PDZ1 supramodule (PDZ2-Cter or PDZ2-Cter ΔLTFF that also lacks the PBM; Fig. S2) by small-angle X-ray scattering (SAXS) (Fig. 2). The various parameter values deduced from the SAXS scattering curve I(q) are listed in Table S1. The deduced molar masses showed that all these molecules were monomeric. The distance distribution function p(r), which characterizes the shape of the molecule in real space, was similar for PDZ2-Cter and PDZ2-Cter ΔLTFF, and was typical of extended objects (Fig. 2A). By contrast, p(r) differed between a1 full-length and a1 ΔLTFF (Fig. 1D), and this difference was reflected in the radius of gyration (R_g) values (R_g(P(r)) = 44.2 Å versus 50.4 Å, respectively). Overall, SAXS experiments indicated that a1 ΔLTFF was more extended than the full-length protein, which is consistent with the results of the analytical ultracentrifugation experiments, again suggesting that the C-terminal PBM is able to fold back onto the N-terminal region of the protein.

The linker region between PDZ2 and PDZ3 (PDZ2/3 linker, amino acids 295-447) is the most likely candidate region capable of mediating the ability of harmonin-a1 to fold back. However, structural data concerning this region, which is predicted to contain a coiled-coil domain (CC1), are currently lacking. We used both circular dichroism (CD) and SAXS techniques coupled to secondary structure predictions to determine the structure of the PDZ2/3 linker region. We first studied two protein fragments containing the PDZ2 domain, together with the entire PDZ2/3 linker region, PDZ2-PDZ2/3 L (amino acids 197-437), or a C-terminally truncated PDZ2/3 linker region, PDZ2-PDZ2/3 L' (amino acids 197-380) (Fig. 3A and Fig. S2). CD spectroscopy analysis allowed us to examine the presence of α helices, β strands, and non-structured elements. A comparison between the CD spectra of the PDZ2-PDZ2/3 L and PDZ2-PDZ2/3 L' harmonin fragments showed that the C-terminal region of the PDZ2/3 linker was intrinsically unstructured (Fig. S1B). We then focused on the N-terminal region of the PDZ2/3 linker (amino acids 298-388), which all algorithms (LOGICOILS, COILS, MULTICOILS) have predicted to adopt a coiled-coil domain, either

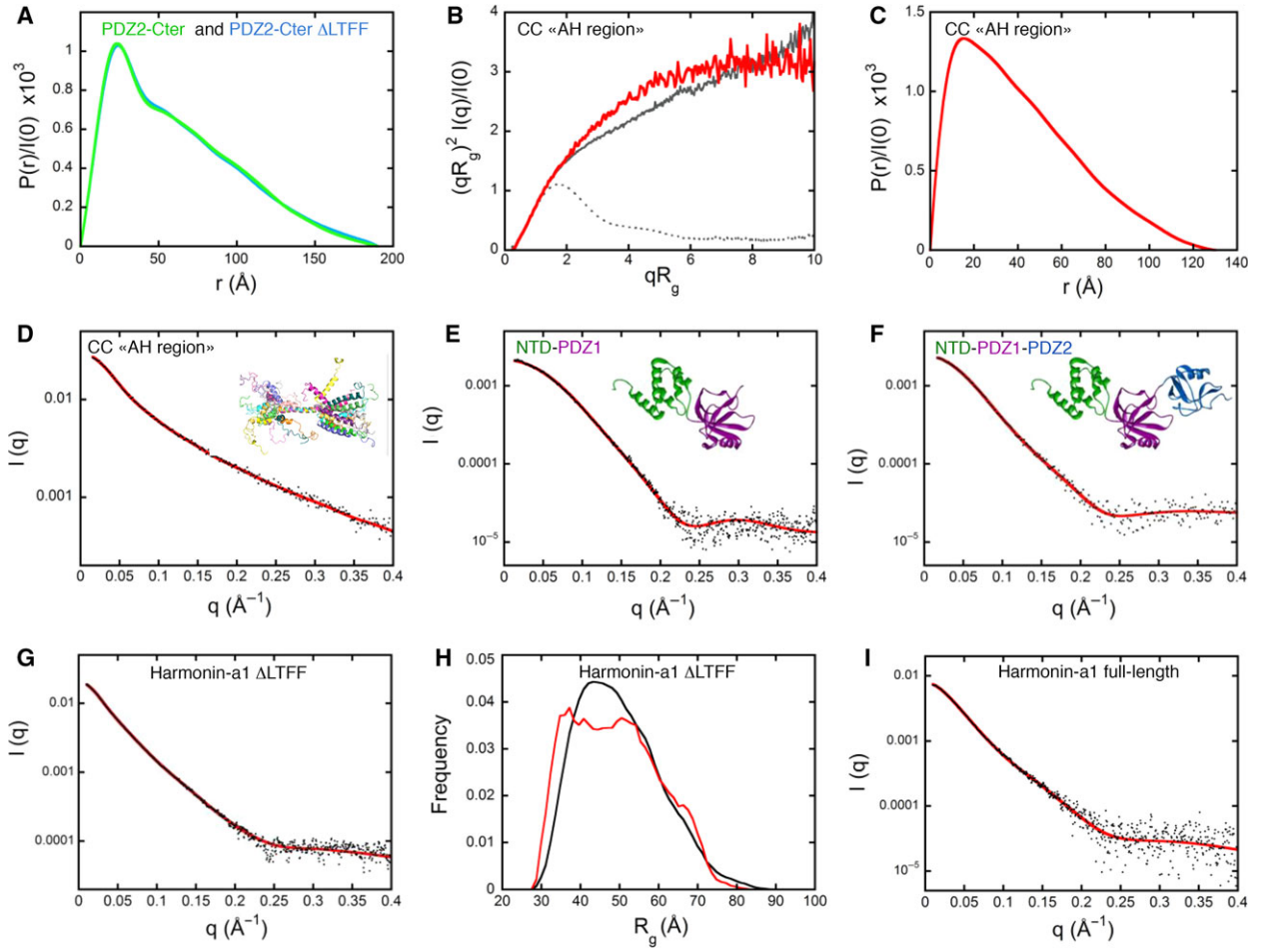

Fig. 2. SAXS studies of harmonin-a1 full-length or ΔLTFF and of different fragments in solution. (A) Distance distribution functions, P(r), scaled to I(0) and derived from PDZ2-Cter (blue line) and PDZ2-Cter-ΔLTFF (green line) scattering patterns. (B) Dimensionless Kratky plot of the CC1 scattering pattern (red line). The scattering patterns of PolX [41] (compact, fully structured, globular protein) and IB5 [42] (fully disordered and extended protein) are shown for a comparison (dashed and solid gray lines, respectively). (C) Distance distribution function P (r) derived from the CC1 scattering pattern scaled to I(0). (D) Fit of the CC1 SAXS pattern (black dots), by the ensemble optimization method (EOM). The fit (red line) is the average scattering pattern of the ensemble of 13 conformations shown in the insert (χ^2 = 0.78). (E) Comparison of the experimental SAXS curve (black dots) for NTD-PDZ1 with the curve calculated by the program CRYSOL (red line) from the structural model proposed by [22] (PDB ID: 3K1R) (χ^2 = 0.89). (F) Comparison of the experimental SAXS curve (black dots) for NTD-PDZ1-PDZ2 with the curve calculated by using CRYSOL (red line) from the structural model shown in the insert. This model was built by starting from the atomic models of NTD-PDZ1 and PDZ2 (PDB ID: 3K1R and PDB ID: 2KBS), respectively. The two structural models were then assembled with BUNCH software to fit the calculated curve for NTD-PDZ1-PDZ2 to the experimental curve (χ^2 = 0.77). (G) Experimental scattering curve for harmonin-a1 ΔLTFF (black dots) superimposed on the calculated curve derived from the best ensemble of conformations selected by the program EOM (red curve) (χ^2 = 0.86). This ensemble contains 30 'unstructured' conformations with PDZ3 exploring a large space (see Fig. 3B for an example of unstructured conformation). (H) Harmonin-a1 ΔLTFF: distribution of radius of gyration values for conformations from ensembles selected by EOM (red line) compared with the corresponding distribution for the complete pool of randomly generated conformations (black line). (I) Experimental SAXS pattern (black dots) for full-length harmonin-a1 fitted with an ensemble of 53% of unstructured conformations and 47% of 'folded' conformations (red curve) (χ^2 = 0.77) (see Fig. 3B for an example of folded conformation).

parallel or antiparallel [35–37]; hence the name 'CC1' attributed to the entire region. CD spectra revealed that about 70% of the CC1 region consisted of α helices (Fig. 1B). The SAXS pattern of the CC1 region alone in solution was used to characterize the conformations adopted by this region. Fig 2B and 2C shows the scattering pattern I(q) obtained with a dimensionless Kratky plot and the distance distribution function P(r) derived from I(q), respectively. These two plots are characteristic of a fully unfolded

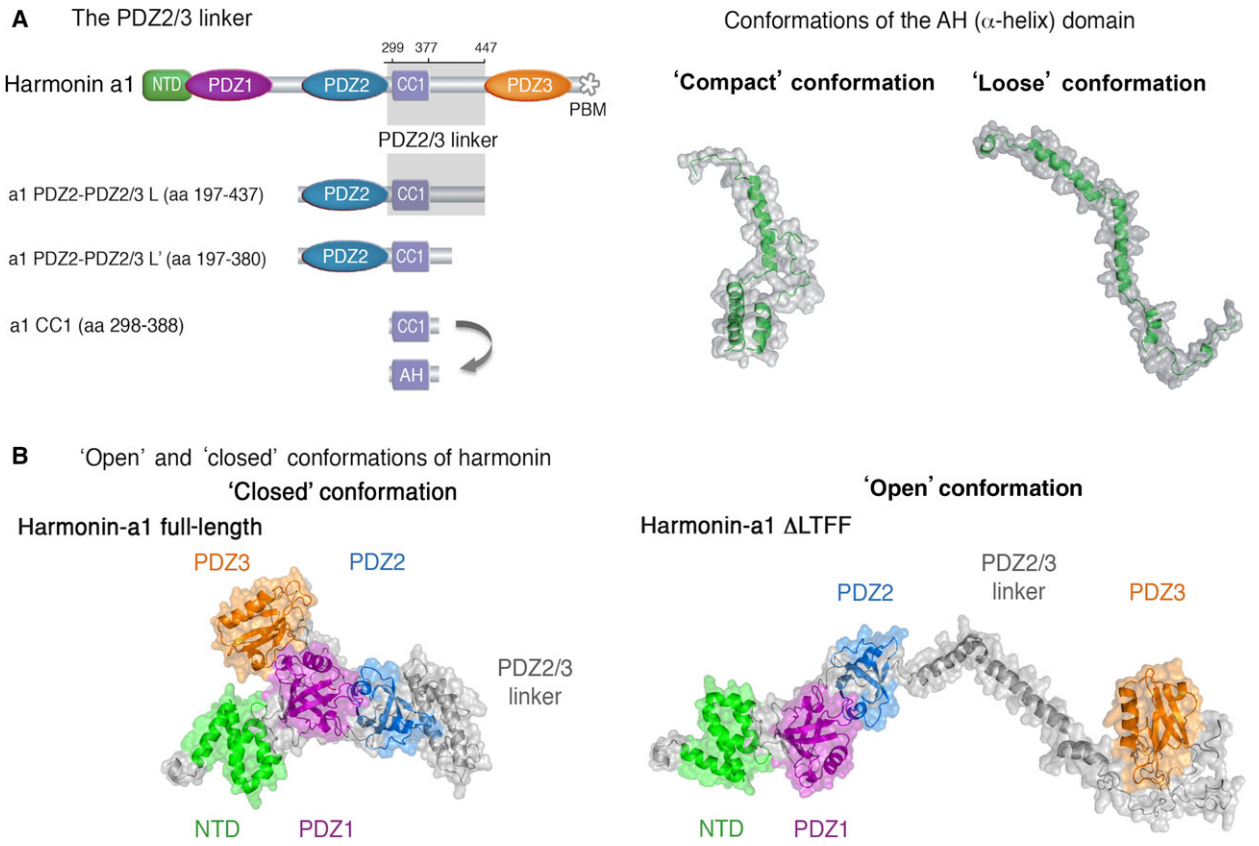

Fig. 3. Model of the open and closed conformations of harmonin-a1. (A) Model of CC1 harmonin with a compact and loose conformations of the alpha helix domain (AH). Left panel: schematic representation of the full-length harmonin-a1 isoform and various fragments of the PDZ2/3 linker. Right panel: compact ('closed') and loose ('open') conformations of the α-helix domain (AH). These two conformations were extracted from the ensemble of conformations shown in the insert of Fig. 2D. (B) Structural models of the full-length harmonin-a1 and harmonin-a1 ΔLTFF. The models derived from SAXS data analysis show the compact, closed conformation of the full-length protein and the extended, open conformation of the protein lacking the PBM motif (harmonin-a1 ΔLTFF). NTD: N-terminal domain (green); PDZ1, PDZ2, PDZ3: postsynaptic density, disc large, zonula occludens domains 1, 2, and 3 (purple, blue, orange). The AH domain and unstructured regions of the protein are represented in gray.

polypeptide. We pursued this analysis by first building 10 000 conformations of the predicted CC1 region, taking into account the results of secondary structure prediction programs, which suggested that three stretches of amino acids at positions 302-324, 329-352, and 356-569 were probably α-helical. Atomic models of these three α-helices were obtained with the tertiary structure prediction program Phyre2 [38]. In all 10 000 conformations, the interhelical linkers were constructed as short chains of residues in random conformations that just satisfied the polypeptide chain constraints. In the last stage of the analysis, we used the GAJOE program to select from among the pool of 10 000 conformations the ensembles of conformations (typically comprising a few tens of conformations) for which the average scattering pattern of the ensemble provided a good fit to the experimental curve (Fig. 2D). Together, these results indicate that the CC1 region in solution does not have the predicted typical structure of a rigid coiled-coil domain. Instead, it seems to consist of a flexible necklace of α-helices oriented randomly relative to each other, potentially capable of accommodating the folding back of the C-terminal PBM of harmonin onto its NTD-PDZ1 supramodule. We will, therefore, refer to this domain hereafter as the α-helix (AH) domain (Fig. 3A).

We then built a structural model of the full-length harmonin-a1 and a1 ΔLTFF, using the previously reported crystal structures of NTD-PDZ1, PDZ2, and PDZ3 [12,22], RIKEN Structural Genomics/Proteomics Initiative, Protein Data Bank 3K1R, 2KBS, 1V6B, respectively. First, we determined the SAXS patterns of the NTD-PDZ1 and NTD-PDZ1-PDZ2 protein fragments experimentally (Fig. 2E,F). We found that the conformation of NTD-PDZ1 in

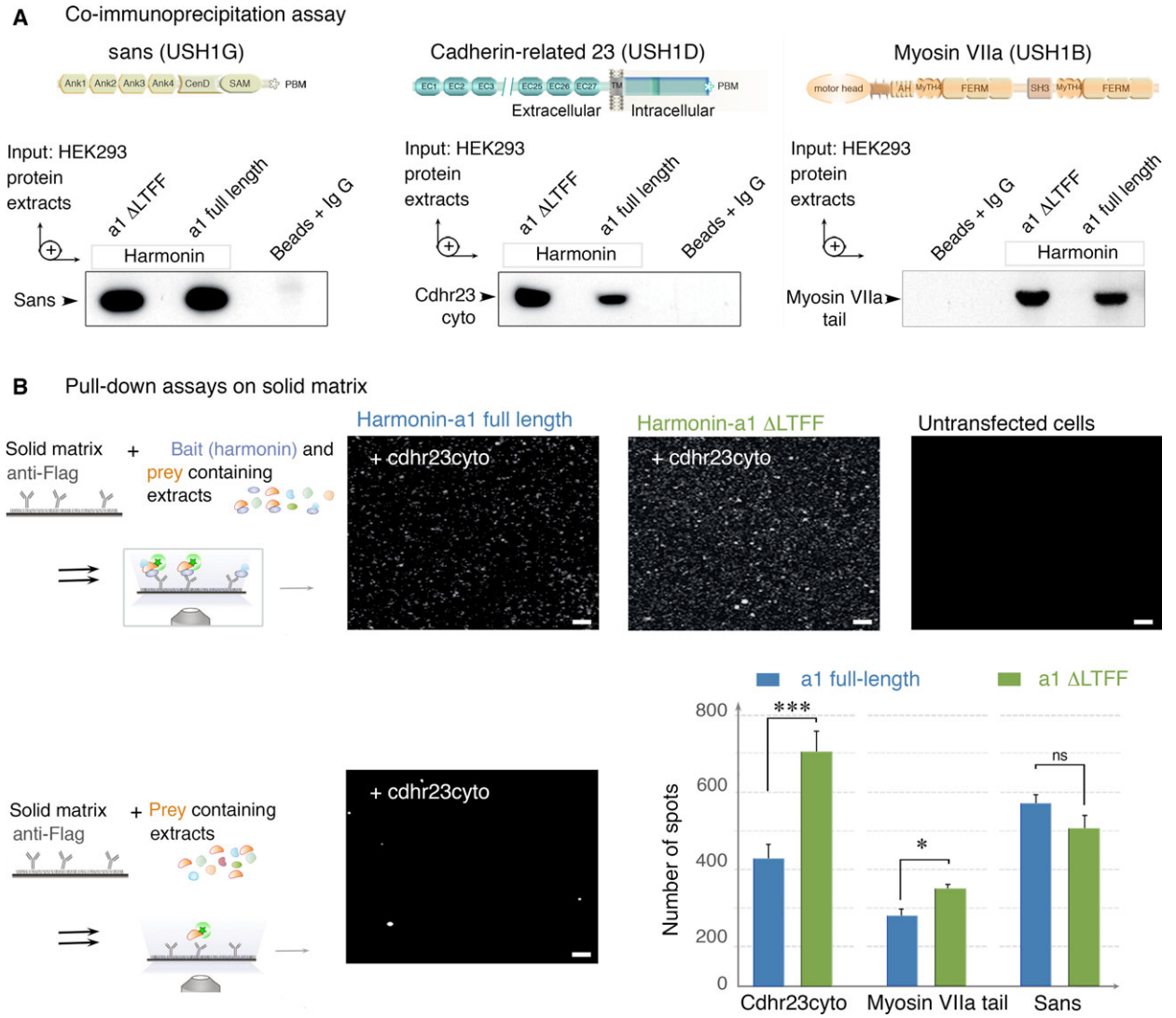

Fig. 4. Interactions of the putative open and closed conformations of harmonin-a1 with its binding partners. (A) Co-immunoprecipitation experiments. Left panel: Schematic representation of three binding partners of harmonin; the horizontal bars indicate the protein fragments used in the co-immunoprecipitation experiments. Right panel: western-blot analysis. Protein extracts from co-transfected HEK293 cells producing full-length harmonin-a1 (a1 full length) or harmonin-a1 lacking the C-terminal PBM motif (a1 ΔLTFF) in the presence of sans, the entire myosin VIIa tail, or the cytoplasmic region of cadherin-related 23 (cdhr23cyto) were used. Similar amounts of sans or the myosin VIIa tail were co-immunoprecipitated with a1 full length and a1 ΔLTFF, whereas larger amounts of cdhr23cyto are co-immunoprecipitated with a1 ΔLTFF than with a1 full length. Dividing line has been used to separate lanes that were not contiguous in the original gel. (B) Single-molecule pull-down assays on a solid matrix. The anti-Flag antibody is immobilized on a NHS-activated surface. The supernatant extract from co-transfected HEK293 cells producing Flag-tagged a1 full length or a1 ΔLTFF and Myc-tagged sans, myosin VIIa tail, or cdhr23cyto is added to the surface (upper diagram). Extracts from cells producing sans, the myosin VIIa tail, or cdhr23cyto only are used as negative controls (lower diagram). The pulled-down Flag-tagged full-length harmonin-a1/Myc-tagged cdhr23cyto or Flag-tagged harmonin-a1 ΔLTFF/Myc-tagged cdhr23cyto molecular complex is probed with an anti-Myc primary antibody and an Atto-conjugated secondary antibody. Co-immunoprecipitated prey proteins are visualized as isolated fluorescent spots (scale bars in the pictures = 5 μm). The bar graph shows the numbers of fluorescent spots (mean ± s.d.) per image (10 000 μm^2) Asterisks indicate significant differences (*$P < 0.05$; ***$P < 0.001$; ns, not statistically significant; Student's *t*-test).

solution was identical to its crystal conformation, and we were able to identify the likely position of the PDZ2 domain with respect to the NTD-PDZ1 supramodule. We built models of 10 000 different conformations of harmonin-a1 ΔLTFF, by connecting PDZ3 to PDZ2 with a flexible linker containing the

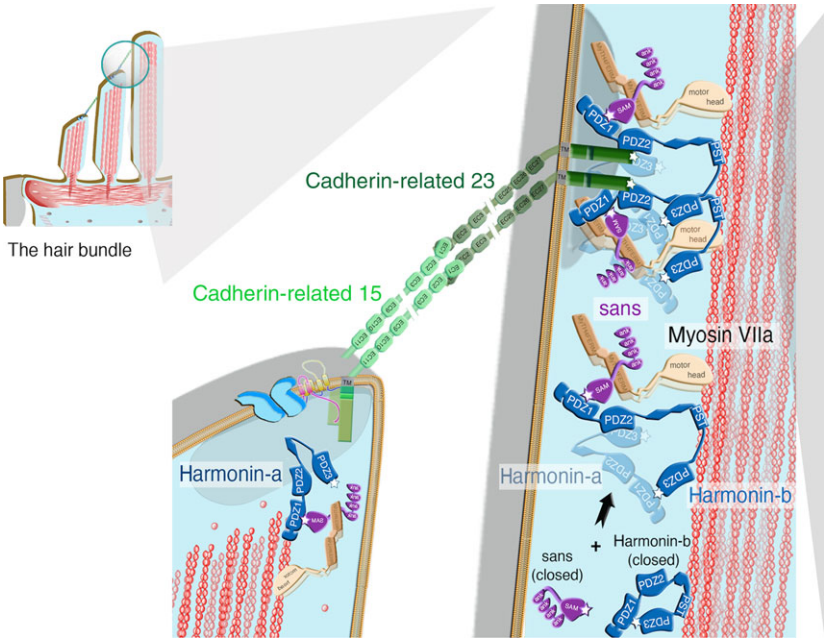

Fig. 5. Diagram showing harmonin isoforms at the tip-link upper and lower densities**.** Diagram of the molecular complexes associated with the tip-link upper and lower densities, showing harmonin-a and harmonin-b isoforms in the 'open' and 'closed' conformations, and the predicted interactions with their binding partners: cadherin-related 23, myosin VIIa, and sans.

three helices of the AH domain, linked by chains of residues in random conformations. The ensembles of conformations best accounting for the experimental curve (Fig. 2G) each included several tens of unfolded conformations exploring the entire space, and the distribution of gyration radii within these ensembles was similar to that of the complete pool (Fig. 2H), suggesting that the PDZ2/3 linker is indeed flexible in the absence of the C-terminal PBM. An example of representative conformations is shown in Fig. 3A. None of the ensembles from this pool of 10 000 unstructured conformations was able to account for the SAXS scattering curve of full-length harmonin-a1. We needed to create an additional pool of conformations in which the PDZ2/3 linker folded back such that the harmonin PBM came into contact with the NTD-PDZ1 module and interacted with it. The interaction between the PBM and NTD-PDZ1 was modeled on the basis of the solved structure for sans-PBM and harmonin (Protein Data bank: 3K1R) described by Yan *et al.* [12]. The SAXS pattern of full-length harmonin-a1 was then perfectly fitted (Fig. 2I) by ensembles containing roughly equal numbers of conformations from the 'extended' and 'folded' pools. Fig. 2B presents an example of back-folded conformations, in which the PBM interacts with the NTD-PDZ1 supramodule. In conclusion, the SAXS data indicate that the full-length harmonin-a1 exists in solution as an equilibrium between 'closed' and 'open' conformational states, probably due to the flexible nature of the AH domain in the central region of the protein.

Impact of harmonin topology on the interaction with binding partners

The closed conformation of harmonin-a1 resembles the auto-inhibited conformation of another PDZ domain-containing scaffold protein of the hair bundle, NHERF-1, in which the C-terminal PBM interacts with the first and second PDZ domains of the protein [39]. We, thus, investigated the impact of the open and closed conformations of harmonin-a on its interactions with other USH1 proteins. Co-immunoprecipitation experiments were carried out on co-transfected HEK293 cells producing the cytoplasmic region of cdhr23 (USH1D), the entire myosin VIIa (USH1B) tail, or sans (USH1G), and either full-length harmonin-a1 or harmonin-a1 ΔLTFF. Western-blot analysis showed no significant difference in binding between sans or the entire myosin VIIa tail and full-length harmonin-a1 or harmonin-a1 ΔLTFF. By contrast, a smaller amount of the cdhr23 cytoplasmic fragment was coprecipitated with the full-length harmonin-a1 than with harmonin-a1 ΔLTFF (Fig. 4A). We also showed, in single-molecule pull-down experiments that the cytodomain of cdhr23 and, to a lesser extent, the myosin VIIa tail bound preferentially to harmonin-a1 ΔLTFF, whereas sans invariably bound to full-length harmonin-a1 and harmonin-a1 ΔLTFF (Fig. 4B). The sans C-terminal PBM peptide has been shown to bind to the harmonin NTD-PDZ1 supramodule with high affinity (Kd = 1 μM) [12], stronger than that for the interaction between the C-terminal PBM of harmonin and its NTD-PDZ1 domain shown

here. Assuming similar local cell concentrations of sans and harmonin, we suggest that sans binding can disrupt the C-terminal PBM-PDZ1 interaction of harmonin, thereby triggering the conformational switch from a putative auto-inhibitory closed state of the protein to a putative active open state, allowing the PDZ1 and/or PDZ2 domain to interact with cdhr23.

A similar conformational switch probably also applies to the harmonin-b isoforms, which differ from harmonin-a1 solely by the additional presence of the CC2 and PST regions after the common AH region (Fig. 1A). On the basis of their interaction with the cytoplasmic region of cdhr23, their binding to F-actin, and their subcellular localization at the tip-link upper insertion point in the hair cells, harmonin-b isoforms presumably anchor the tip links to the actin filaments of the stereocilia [6,7]. Since the regions of harmonin previously predicted to be coiled coils (CC1 and presumably CC2 also) are in fact flexible domains rich in α helices, these isoforms are qualified to modulate the elasticity of the upper tip-link complex (Fig. 5). Finally, sequence analysis of all the PDZ domain-containing proteins in the Ensembl database (see Materials and Methods; Table S2) revealed that 43 proteins (i.e., about 28%) also contain a predicted C-terminal PBM, suggesting that intramolecular interactions similar to that reported here for harmonin might occur in many, if not all of them. Likewise, these self-interactions might be removed by competitive binding or by phosphorylation of the protein as reported for X11α/Mint1 and EBP50/NHERF1 [40].

Acknowledgements

We thank Sébastien Brulé from the biophysical platform for technical support. This work was supported by *Institut Pasteur* PTR program No.483, by the European Union Seventh Framework Programme under the grant agreement HEALTH-F2-2010-242013 (TREATRUSH), the European Commission (Hairbundle ERC-2011-ADG_294570), by *Agence Nationale de la Recherche* (ANR) within the framework of the *Investissements d'Avenir* program (ANR-15-RHUS-0001), *Laboratoire d'excellence* (LabEx) Lifesenses (ANR-10-LABX-65), ANR-11-IDEX-0004-02, ANR-11-BSV5-0011, and grants from *Fondation Agir pour l'Audition*, the BNP Paribas Foundation, the FAUN Stiftung, the LHW-Stiftung and Mrs. Errera Hoechstetter. We thank the staff of the SWING beamline for help during the SAXS measurements.

Author contributions

AB, DD and CP designed the research project. AB, EP and SN purified the recombinant proteins. AB, BR, FC, PE, BB and DD did the biophysical experiments. AB, PE, BR and DD analyzed the results. AB, AE, JPH, DD and CP wrote the article.

References

1 Beurg M, Fettiplace R, Nam JH and Ricci AJ (2009) Localization of inner hair cell mechanotransducer channels using high-speed calcium imaging. *Nat Neurosci* **12**, 553–558.

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

3 Petit C and Richardson GP (2009) Linking genes underlying deafness to hair-bundle development and function. *Nat Neurosci* **12**, 703–710.

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

5 Pepermans E, Michel V, Goodyear R, Bonnet C, Abdi S, Dupont T, Gherbi S, Holder M, Makrelouf M, Hardelin J-P *et al.* (2014) The CD2 isoform of protocadherin-15 is an essential component of the tip-link complex in mature auditory hair cells. *EMBO Mol Med* **6**, 984–992.

6 Lefevre G, Michel V, Weil D, Lepelletier L, Bizard E, Wolfrum U, Hardelin J-P and Petit C (2008) A core cochlear phenotype in USH1 mouse mutants implicates fibrous links of the hair bundle in its cohesion, orientation and differential growth. *Development* **135**, 1427–1437.

7 Michalski N, Michel V, Caberlotto E, Lefevre GM, van Aken AF, Tinevez JY, Bizard E, Houbron C, Weil D, Hardelin J-P *et al.* (2009) Harmonin-b, an actin-binding scaffold protein, is involved in the adaptation of mechanoelectrical transduction by sensory hair cells. *Pflugers Arch* **459**, 115–130.

8 Grillet N, Xiong W, Reynolds A, Kazmierczak P, Sato T, Lillo C, Dumont RA, Hintermann E, Sczaniecka A, Schwander M *et al.* (2009) Harmonin mutations cause mechanotransduction defects in cochlear hair cells. *Neuron* **62**, 375–387.

9 Bahloul A, Michel V, Hardelin J-P, Nouaille S, Hoos S, Houdusse A, England P and Petit C (2010) Cadherin-23, myosin VIIa and harmonin, encoded by Usher syndrome type I genes, form a ternary complex and

interact with membrane phospholipids. *Hum Mol Genet* **19**, 3557–3565.
10 Caberlotto E, Michel V, Foucher I, Bahloul A, Goodyear RJ, Pepermans E, Michalski N, Perfettini I, Alegria-Prevot O, Chardenoux S *et al.* (2011) Usher type 1G protein sans is a critical component of the tip-link complex, a structure controlling actin polymerization in stereocilia. *Proc Natl Acad Sci USA* **108**, 5825–5830.
11 Grati M and Kachar B (2011) Myosin VIIa and sans localization at stereocilia upper tip-link density implicates these Usher syndrome proteins in mechanotransduction. *Proc Natl Acad Sci USA* **108**, 11476–11481.
12 Yan J, Pan L, Chen X, Wu L and Zhang M (2010) The structure of the harmonin/sans complex reveals an unexpected interaction mode of the two Usher syndrome proteins. *Proc Natl Acad Sci USA* **107**, 4040–4045.
13 Bitner-Glindzicz M, Lindley KJ, Rutland P, Blaydon D, Smith VV, Milla PJ, Hussain K, Furth-Lavi J, Cosgrove KE, Shepherd RM *et al.* (2000) A recessive contiguous gene deletion causing infantile hyperinsulinism, enteropathy and deafness identifies the Usher type 1C gene. *Nat Genet* **26**, 56–60.
14 Verpy E, Leibovici M, Zwaenepoel I, Liu XZ, Gal A, Salem N, Mansour A, Blanchard S, Kobayashi I, Keats BJ *et al.* (2000) A defect in harmonin, a PDZ domain-containing protein expressed in the inner ear sensory hair cells, underlies Usher syndrome type 1C. *Nat Genet* **26**, 51–55.
15 Boeda B, El-Amraoui A, Bahloul A, Goodyear R, Daviet L, Blanchard S, Perfettini I, Fath KR, Shorte S, Reiners J *et al.* (2002) Myosin VIIa, harmonin and cadherin 23, three Usher I gene products that cooperate to shape the sensory hair cell bundle. *EMBO J* **21**, 6689–6699.
16 Siemens J, Kazmierczak P, Reynolds A, Sticker M, Littlewood-Evans A and Muller U (2002) The Usher syndrome proteins cadherin 23 and harmonin form a complex by means of PDZ-domain interactions. *Proc Natl Acad Sci USA* **99**, 14946–14951.
17 Adato A, Michel V, Kikkawa Y, Reiners J, Alagramam KN, Weil D, Yonekawa H, Wolfrum U, El-Amraoui A and Petit C (2005) Interactions in the network of Usher syndrome type 1 proteins. *Hum Mol Genet* **14**, 347–356.
18 Wu L, Pan L, Zhang C and Zhang M (2012) Large protein assemblies formed by multivalent interactions between cadherin23 and harmonin suggest a stable anchorage structure at the tip link of stereocilia. *J Biol Chem* **287**, 33460–33471.
19 Reiners J, Reidel B, El-Amraoui A, Boeda B, Huber I, Petit C and Wolfrum U (2003) Differential distribution of harmonin isoforms and their possible role in Usher-1 protein complexes in mammalian photoreceptor cells. *Invest Ophthalmol Vis Sci* **44**, 5006–5015.
20 Luck K, Charbonnier S and Trave G (2012) The emerging contribution of sequence context to the specificity of protein interactions mediated by PDZ domains. *FEBS Lett* **586**, 2648–2661.
21 Ye F and Zhang M (2013) Structures and target recognition modes of PDZ domains: recurring themes and emerging pictures. *Biochem J* **455**, 1–14.
22 Pan L, Yan J, Wu L and Zhang M (2009) Assembling stable hair cell tip link complex via multidentate interactions between harmonin and cadherin 23. *Proc Natl Acad Sci USA* **106**, 5575–5580.
23 David G and Perez J (2009) Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J Appl Crystallogr* **42**, 892–900.
24 Konarev PV, Volkov VV, Sokolova AV, Koch MHJ and Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* **36**, 1277–1282.
25 Guinier A (1939) La diffraction des rayons X aux très petits angles: application à l'étude de phénomènes ultramicroscopiques. Univ. de Paris, Paris.
26 Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* **25**, 495–503.
27 Fischer H, Neto MD, Napolitano HB, Polikarpov I and Craievich AF (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J Appl Crystallogr* **43**, 101–109.
28 Rambo RP and Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481.
29 Petoukhov MV and Svergun DI (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J* **89**, 1237–1250.
30 Svergun D, Barberato C and Koch MHJ (1995) CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28**, 768–773.
31 Bernado P, Mylonas E, Petoukhov MV, Blackledge M and Svergun DI (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc* **129**, 5656–5664.
32 Tria G, Mertens HDT, Kachala M and Svergun DI (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCr J* **2**, 207–217.
33 Moore BL, Kelley LA, Barber J, Murray JW and MacDonald JT (2013) High-quality protein backbone reconstruction from alpha carbons using gaussian mixture models. *J Comput Chem* **34**, 1881–1889.

34 Krivov GG, Shapovalov MV and Dunbrack RL (2009) Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* **77**, 778–795.
35 Berger B, Wilson DB, Wolf E, Tonchev T, Milla M and Kim PS (1995) Predicting coiled coils by use of pairwise residue correlations. *Proc Natl Acad Sci USA* **92**, 8259–8263.
36 Lupas A, Van Dyke M and Stock J (1991) Predicting coiled coils from protein sequences. *Science* **252**, 1162–1164.
37 Wolf E, Kim PS and Berger B (1997) MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci* **6**, 1179–1189.
38 Kelley LA, Mezulis S, Yates CM, Wass MN and Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845–858.
39 Morales FC, Takahashi Y, Momin S, Adams H, Chen X and Georgescu MM (2007) NHERF1/EBP50 head-to-tail intramolecular interaction masks association with PDZ domain ligands. *Mol Cell Biol* **27**, 2527–2537.
40 Lee HJ and Zheng JJ (2010) PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun Signal* **8**, 8.
41 Leulliot N, Cladiere L, Lecointe F, Durand D, Hubscher U and van Tilbeurgh H (2009) The family X DNA polymerase from Deinococcus radiodurans adopts a non-standard extended conformation. *J Biol Chem* **284**, 11992–11999.
42 Boze H, Marlin T, Durand D, Perez J, Vernhet A, Canon F, Sarni-Manchado P, Cheynier V and Cabane B (2010) Proline-rich salivary proteins have extended conformations. *Biophys J* **99**, 656–665.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Fig. S1. (A) Intramolecular interaction between the PBM and NTD-PDZ1 domains of harmonin. The interaction of the NTD-PDZ1 domain of harmonin-a1 with an immobilized biotinylated synthetic peptide containing the C-terminal 16 amino-acid residues of harmonin-a1 including the PBM was analyzed by SPR. Different concentrations of NTD-PDZ1 (0.6–19 μM) were injected in a randomized order. The binding affinity was fitted using the maximum response values. The K_d value was determined as 61 μM. RU, resonance unit. (B) Circular dichroism (CD) spectroscopy analysis of the secondary structure of the harmonin-a1 fragment connecting PDZ2 and PDZ3. CD spectra were recorded between 180 nm and 250 nm on a Jasco 600 CD spectropolarimeter, as described in the Materials and Methods. Protein secondary structures were predicted by the deconvolution of CD spectra. The coiled-coil (CC) fragment (amino acids 298-388) contains more helical structure and less random structure than the PDZ2-PDZ2/3 L' (amino acids 197-380) and PDZ2-PDZ2/3 L (amino acids 197-437) fragments.
Fig. S2. Diagram and SDS/PAGE analysis of the harmonin fragments used for biophysical studies. The domains of harmonin (NTD: N-terminal domain; PDZ: postsynaptic density, disc large, zonula occludens; CC: coiled-coil; PST: proline–serine–threonine rich; PBM: PDZ-binding motif) and the corresponding amino-acid positions are indicated. The apparent molecular masses (kDa) of the protein fragments analyzed by SDS/PAGE are consistent with the values calculated from their amino-acid sequences.
Table S1. SAXS data and scattering-derived parameters: Rg, Dmax, molecular mass (MM).
Table S2. List of all PDZ domain-containing proteins in man.