Supporting Information

Wei et al. 10.1073/pnas.1306768110

SI Materials and Methods

Protein Expression and Purification. The coding sequences of myosin V paralog a (MyoVa)-globular tail domain (GTD) (residues 1,469–1,853), MyoV paralog b (Vb)-GTD (residues 1,434–1,818), Rab interacting lysosomal protein-like 2 (RILPL2), RILP, melanophilin (MLPH)-MyoVa-GTD binding domain (GTBD; residues 170–208), and granuphilin were PCR amplified from a mouse cDNA library. The N-terminal His_6-tagged MyoVa-GTD was expressed in *Escherichia coli* BL21 (DE3). The Se-Met–labeled MyoVa-GTD was expressed in *E. coli* B834 host strain supplemented with selenomethionine. Fusion proteins in their native or derivative forms were purified by Ni^{2+}-nitrilotriacetic acid (NTA) affinity chromatography followed by size-exclusion chromatography in the buffer containing 50 mM Tris, 100 mM NaCl, 1 mM EDTA, and 1 mM DTT at pH 7.5. Mouse MyoVb-GTD, RILPL2, RILPL2-RH1 (residues 1–97), RILPL2-RH1_ΔC (residues 1–83), RILP-N (residues 1–103), MLPH-GTBD, and granuphilin_230-350, each fused with an N-terminal Trx-His_6-tag, were expressed in *E. coli* BL21 (DE3) and purified by Ni^{2+}-NTA affinity chromatography followed by size-exclusion chromatography. All point mutants described in this study were prepared using PCR-based methods and purified using essentially the same methods as used for the corresponding WT proteins.

Analytical Gel Filtration Chromatography. Analytical gel filtration chromatography was carried out on an AKTA FPLC system (GE Healthcare). Protein samples with concentrations of 30–50 μM were loaded onto a Superose12 10/300 GL column (GE Healthcare) equilibrated with a buffer containing 50 mM Tris·HCl, 100 mM NaCl, 1 mM EDTA, and 1 mM DTT at pH 7.5.

Isothermal Titration Calorimetry Assay. Titration experiments were carried out on a VP-ITC calorimeter at 25 °C. All protein samples were in 50 mM Tris buffer at pH 7.5 plus 100 mM NaCl and 1 mM DTT. The titration processes were performed by injecting 5- to 10-μL aliquots of protein samples in a syringe (concentration of 100 μM) into protein samples in cells (concentration of 10 μM) at time intervals of 2 min to ensure that the titration peak returned to the baseline. The titration data were analyzed using the program Origin7.0 and fitted by the one-site binding model.

Crystallization. Crystals of native MyoVa-GTD and its selenomethionine derivative were obtained by the hanging drop vapor diffusion method at 16 °C. To set up a hanging drop, 1 μL of concentrated protein solution was mixed with 1 μL of crystallization solution with 20–30% vol/vol ethylene glycol; 40% vol/vol ethylene glycol was used for cryoprotection.

To prepare the MyoVa-GTD/RILPL2-RH1/MLPH-GTBD complex, freshly purified MyoVa-GTD, Trx-RILPL2-RH1, and Trx-MLPH-GTBD proteins were mixed at a ratio of 1:1:1, and followed by protease cleavage to remove the Trx-tag from the two MyoVa cargos. The protease-digested protein mixture was further purified by size-exclusion chromatography to obtain the MyoVa-GTD/RILPL2-RH1/MLPH-GTBD ternary complex. Crystallization of the triple complex was performed using the hanging drop vapor diffusion method at 16 °C. Crystals from initial screens only diffracted to ∼4 Å. Better crystals were obtained by introducing a Cys to Ser mutation at residue 1674 of MyoVa-GTD (Fig. S1). The optimized crystallization buffer contains 12–16% wt/vol PEG3350 and 2–4% vol/vol tacsimate pH 7.0. Before diffraction experiments, crystals were soaked in crystallization solution containing 30% vol/vol ethylene glycol for cryoprotection.

Structure Determination. The native and SeMet datasets of MyoVa-GTD and the dataset of the MyoVa-GTD/RILPL2-RH1/MLPH-GTBD complex were collected at Shanghai Synchrotron Radiation Facility beamline BL17U. The diffraction data were processed and scaled using HKL2000 (1).

By using the SeMet derivative dataset, the single-wavelength anomalous diffraction phase was determined and a partial structural model was traced in AutoSol (2). The structure model was built manually based on the experimental phase. The model was refined again the native dataset of MyoVa-GTD in PHENIX (3). COOT was used for model rebuilding and adjustments (4). In the final stage, an additional TLS refinement was performed in PHENIX. The model quality was check by MolProbity (5). In the final model, 97.4% and 2.6% of the residues were in favored and additional allowed regions of the Ramachandran plot, respectively.

The initial phase of the MyoVa-GTD/RILPL2-RH1/MLPH-GTBD complex was determined by molecular replacement using the MyoVa-GTD structure as the search model. The RILPL2-RH1 and MLPH-GTBD molecules were also built into the model. The model was refined using the same strategy as that used for the apo-MyoVa-GTD. In the final model, 98.2% and 1.7% of the residues were in favored and additional allowed regions of the Ramachandran plot, respectively. The final refinement statistics are listed in Table S1. All structure figures were prepared by PyMOL (www.pymol.org).

Analytical Ultracentrifugation. Sedimentation velocity experiments were performed on a Beckman XL-I analytical ultracentrifuge equipped with an eight-cell rotor at 25 °C. The final sedimentation velocity data were analyzed and fitted to a continuous sedimentation coefficient distribution model with the fitting result shown as solid lines using the program SEDFIT (www.analyticalultracentrifugation.com/default.htm).

GST Pull-Down Assay. The interactions between MyoVa-GTD and various granuphilin fragments were assayed in PBS (pH 7.4). GST-GTD (∼0.6 nmol)-loaded GSH-Sepharose beads were incubated with HEK293T-expressed GFP-tagged full-length granuphilin and its various fragments. GTD-bound proteins were separated by SDS/PAGE. The GFP-granuphilin proteins were visualized by immunodetection using anti-GFP antibody.

1. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
2. Terwilliger TC, et al. (2009) Decision-making in structure solution using Bayesian estimates of map quality: The PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* 65(Pt 6):582–601.
3. Adams PD, et al. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58(Pt 11):1948–1954.
4. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
5. Davis IW, et al. (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35(Web Server issue):W375–W383.

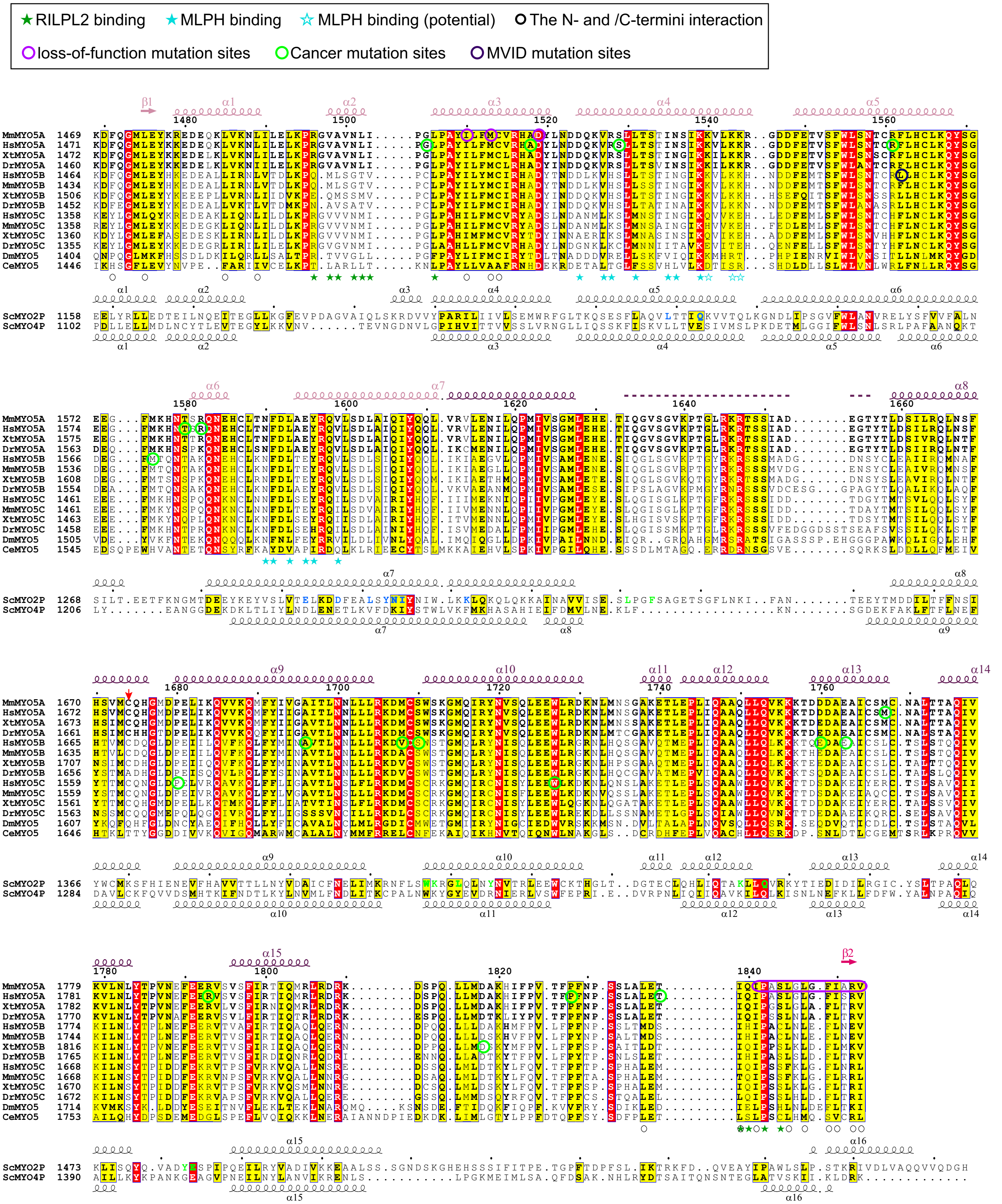

Fig. S1. Structural-based sequence alignments of the GTD regions of type V myosins from different species (related to Fig. 1). The species are represented by two letters. Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; Dr, *Danio rerio*; Hs, human; Mm, mouse; Sc, *Saccharomyces cerevisiae*; Xt, *Xenopus tropicalis*. The sequences were aligned twice with different combinations to present more comprehensive information. First, the sequences from vertebrates, *Drosophila* and *C. elegans*, were aligned. Then, the sequences of yeast myo2p and myo4p were aligned to the first alignment. The secondary structure elements of MyoVa-GTD are labeled above the first alignment using the same color coding as used in Fig. 1. The secondary structure elements of myo2p and myo4p are labeled above and below the second alignment, respectively. Residues that are identical and highly similar in each alignment are shown in red and yellow boxes, respectively. The disordered region is indicated by a dashed line. The cysteine residue, which was mutated to a serine for the complex sample

Legend continued on following page

preparation, is marked by a red arrow. The residues, of which missense or truncation mutations lead to the dilute phenotypes (purple) or human diseases [microvillus inclusion disease (MVID), blue; cancers, green], are circled. The residues involved in the formation of the N terminus and the C-terminal half of the C-terminal loop interface are indicated by open circles. The residues forming the MyoVa-GTD/RILPL2-RH1 interface and the MyoVa-GTD/MLPH-GTBD interface are indicated by solid green and cyan stars, respectively. The three lysine residues, which contribute to the positively charged, MLPH-GTBD binding surface on MyoVa-GTD, are indicated by open cyan stars. The myo2p residues, which were identified as the binding sites for vacuole/mitochondrion-specific and vesicle-specific adaptors (1), are in blue and green, respectively.

1. Eves PT, Jin Y, Brunner M, Weisman LS (2012) Overlap of cargo binding sites on myosin V coordinates the inheritance of diverse cargoes. *J Cell Biol* 198(1):69–85.

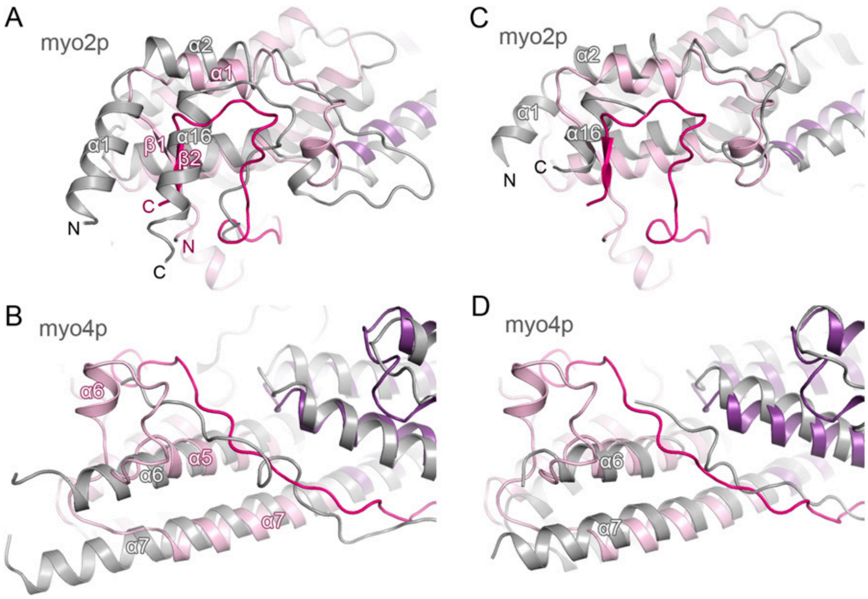

Fig. S2. Structural differences of the GTDs between MyoVa and two type V myosins from yeast (related to Fig. 1). Structural superimpositions of MyoVa with myo2p (*A* and *B*) and of MyoVa with myo4p (*C* and *D*) reveal the differences in the N terminus and the C-terminal half of the C-terminal loop interface (*A* and *C*) and the α5/α7 connecting region (*B* and *D*). MyoVa ribbons are colored with the same coding as in Fig. 1; myo2p and myo4p ribbons are gray.

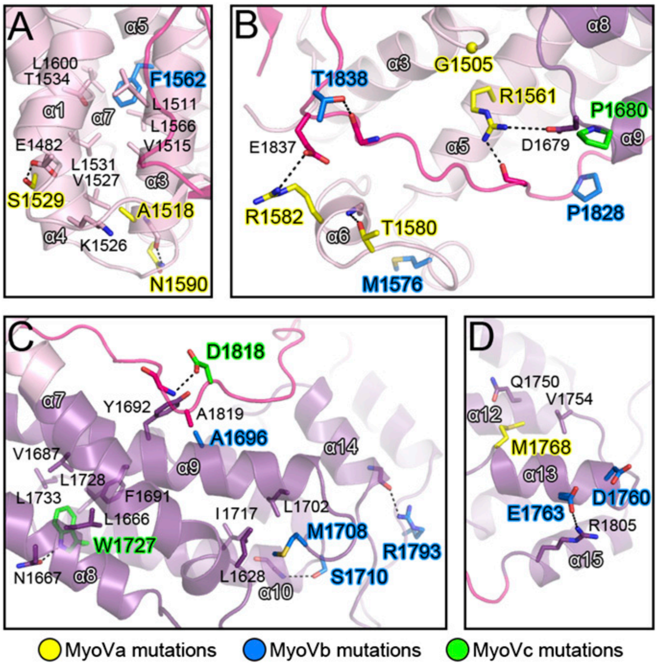

Fig. S3. Structural details of residues in disease-causing mutations (related to Fig. 2). Panels *A*–*D* represent enlarged regions of the apo-MyoVa structure showing mutations described in Fig. 2. The mutations found in MyoVa, Vb, and Vc are colored in yellow, blue, and green, respectively. Hydrogen bonds and salt bridges are indicated by dashed lines.

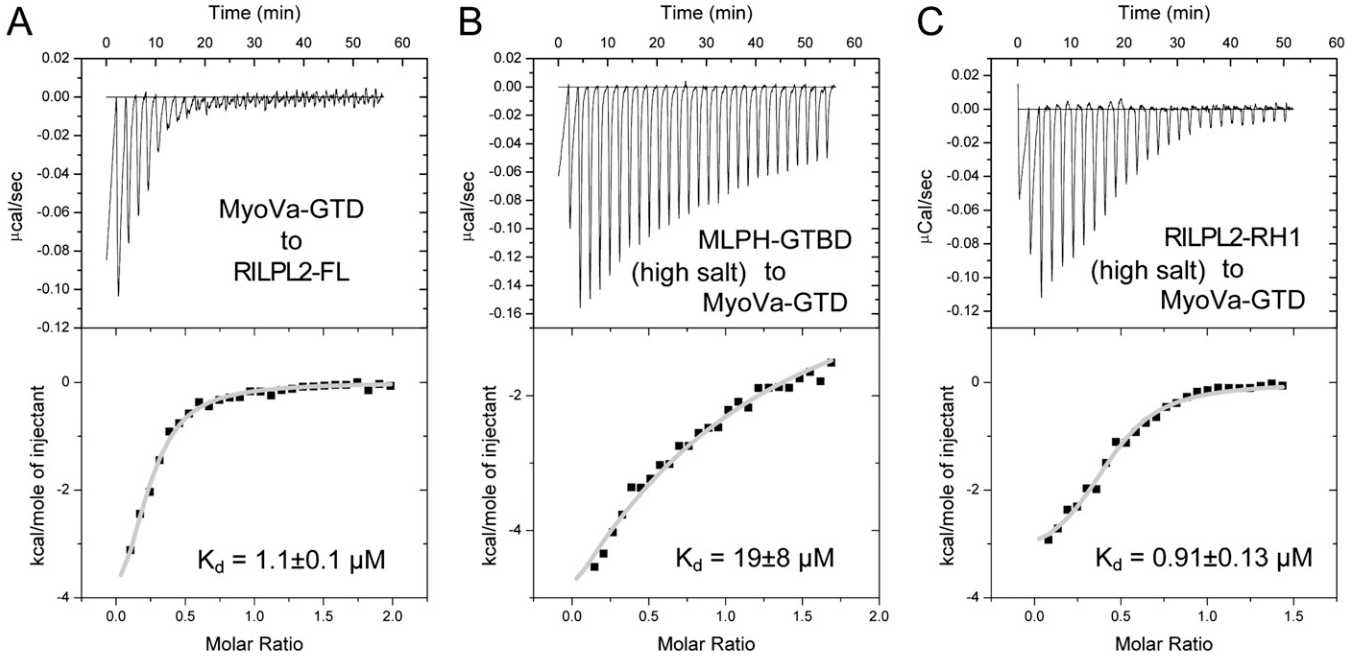

Fig. S4. ITC-based analysis of the MyoVa-GTD/RILPL2 and MyoVa-GTD/MLPH interactions (related to Fig. 3). (*A*) Because the RILPL2 full-length protein (RILPL2-FL) is easy to form aggregate in a high concentration, it was placed in cell instead of syringe during the ITC titration. (*B* and *C*) The sample buffers contained 500 mM NaCl.

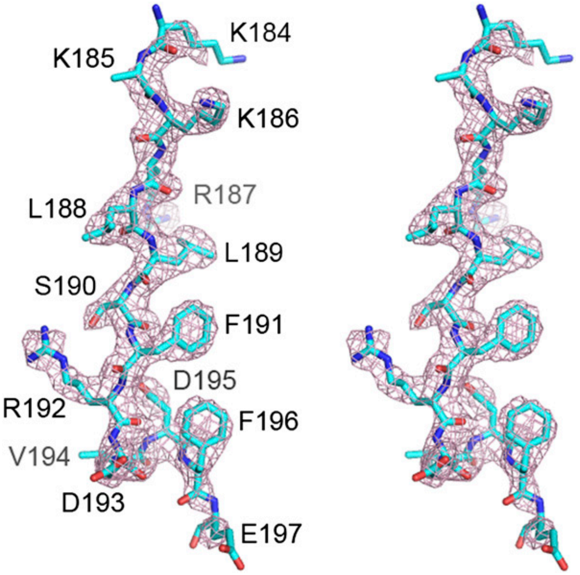

Fig. S5. The stereo view of the electron densities of MLPH-GTBD in the MyoVa-GTD/RILPL2-RH1/MLPH-GTBD complex (related to Fig. 4). The F_o-F_c map was calculated by omitting the MLPH-GTBD part from the final model and contoured at 3σ. The MLPH-GTBD structure is displayed in the explicit stick model. The side chain of K185 was not assigned because of lack of electron density information.

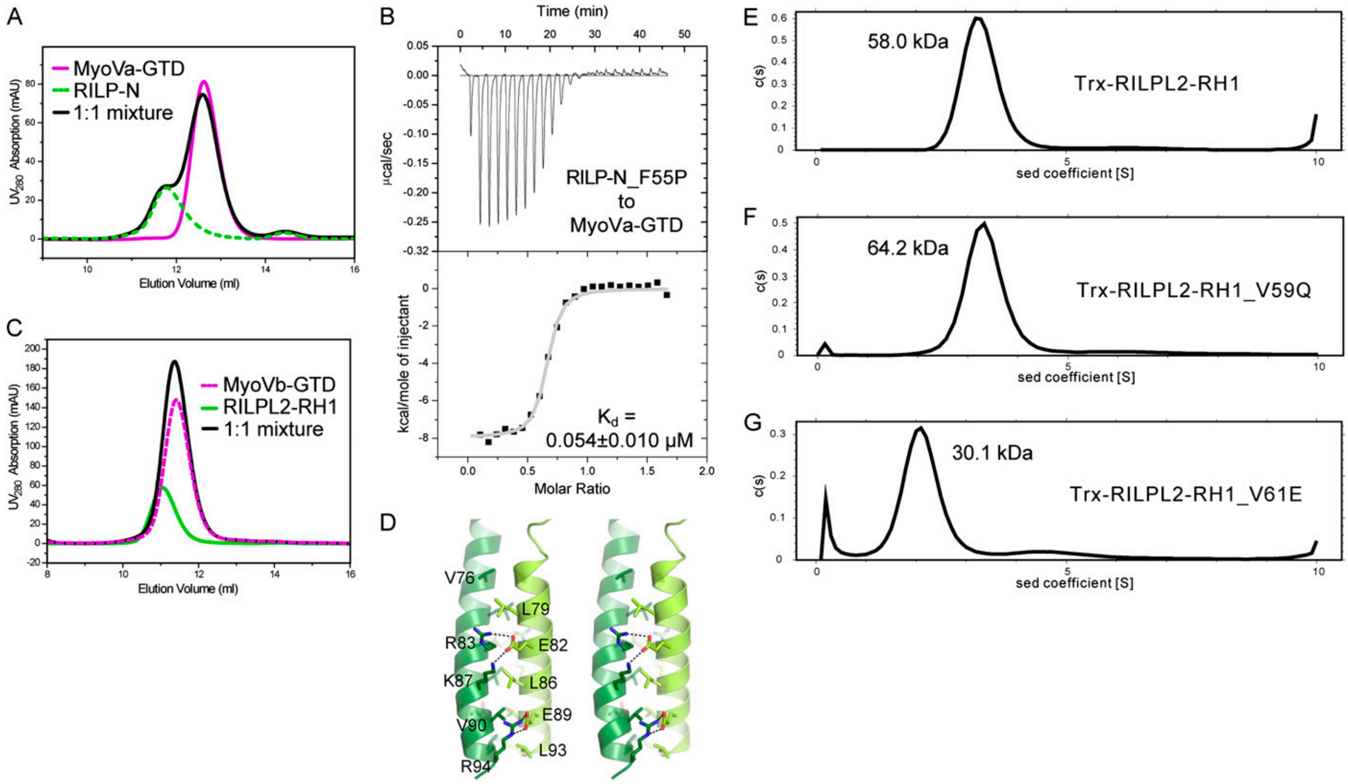

Fig. S6. The binding specificity of RILPL2 to MyoVa-GTD (related to Fig. 5). (*A*) The N-terminal region of RILP (RILP-N) containing the RH1 region did not show detectable binding to MyoVb-GTD. (*B*) A single point mutation (F55P) renders RILP to bind strongly to MyoVa-GTD. (*C*) The analytical gel filtration analysis showed that RILPL2-RH1 is not a MyoVb-GTD binder. (*D*) The molecular details of the coiled-coil formation in RILPL2. Hydrogen bonds and salt bridges are indicated by dashed lines. (*E–G*) Sedimentation velocity analysis of the RILPL2-RH1 dimerization. The derived molecular weights of the main peaks are indicated. The theoretical molecular weight is 27 kDa or 54 kDa for Trx-tagged RILPL2-RH1 in its monomeric or dimeric form, respectively. The figure shows that the V61E mutant of RILPL2-RH1 becomes a monomer, whereas the V59Q mutant of RILPL2-RH1 remains a dimer.

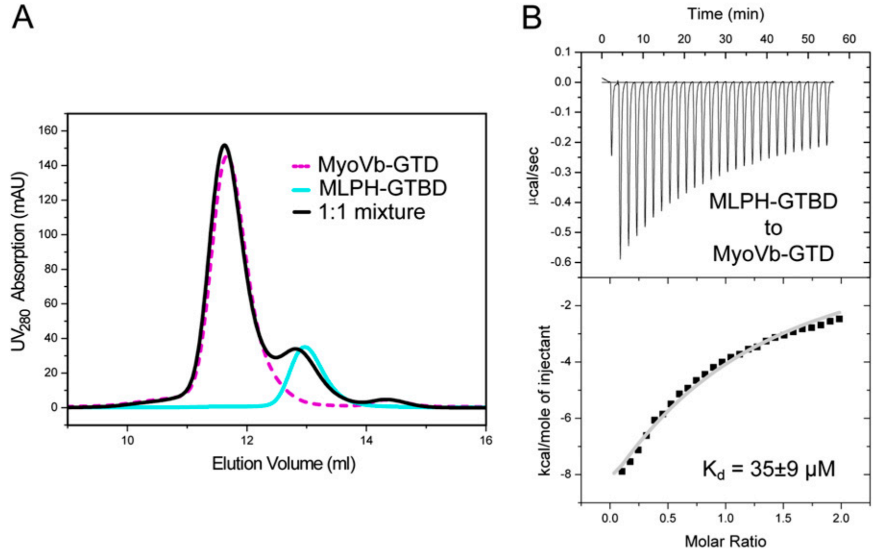

Fig. S7. Analytical gel filtration analysis (*A*) and ITC titration-based assay (*B*) showing that the interaction between MyoVb-GTD and MLPH-GTBD is very weak (related to Fig. 6).

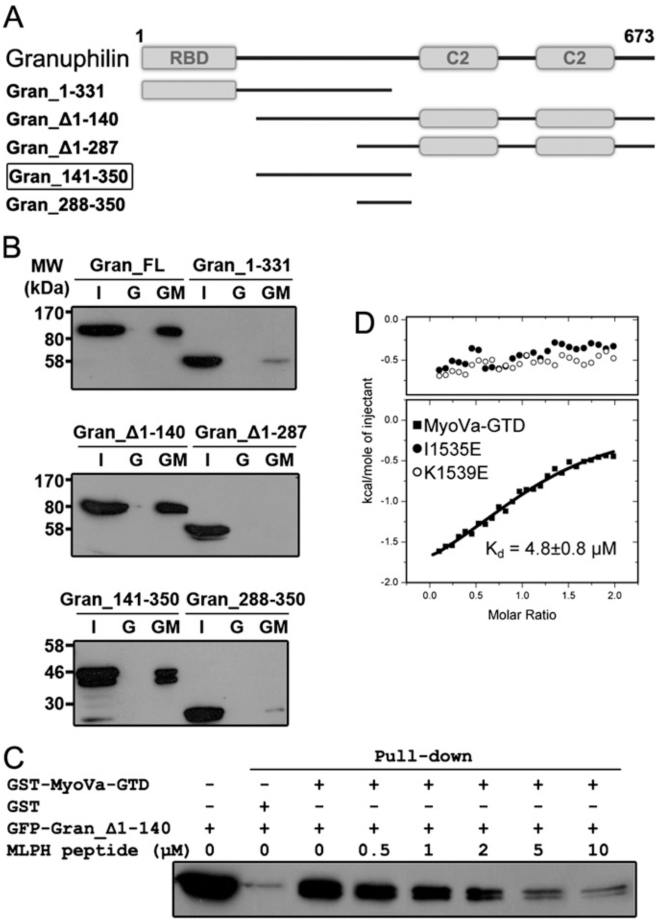

Fig. S8. Granuphilin (Gran) and MLPH-GTBD share overlapping binding sites on MyoVa-GTD. (*A*) The domain organization diagram of Gran and its derivatives tested in this study. The mapped minimal MyoVa-GTD–binding region is boxed. (*B*) GST pull-down analysis of the bindings of various forms of Gran to MyoVa-GTD. The input (I), the GST control (G), and GST-tagged MyoVa-GTD proteins (GM) are shown. (*C*). The competition experiment showing that MLPH can effectively compete with Gran_Δ1-140 for binding to GST-tagged MyoVa-GTD. The protein concentration of GST-tagged MyoVa-GTD is ~2 μM. The MLPH peptide contains residues 176–201, which has been shown to retain the high-affinity binding to MyoVa-GTD (1). (*D*) ITC-based analysis of the bindings of MyoVa-GTD (*Lower*) and its two MLPH-binding defective mutants (*Upper*) to Gran_230-350.

1. Geething NC, Spudich JA (2007) Identification of a minimal myosin Va binding site within an intrinsically unstructured domain of melanophilin. *J Biol Chem* 282(29):21518--21528.

Table S1. Statistics of data collection and model refinement

	Apo-Native	Apo-SeMet	Complex
Data collection			
Space group	$P2_1$	$P2_12_12_1$	$P2_1$
Unit cell parameters, Å	a = 69.0, b = 165.4, c = 173.0 β = 90.1	a = 69.0, b = 88.0, c = 166.2	a = 81.9, b = 107.9, c = 83.5 β = 96.1
Resolution range, Å	50–2.5 (2.54–2.5)	50–3.1 (3.15–3.1)	50–2.4 (2.44–2.4)
No. of unique reflections	131,888 (6,046)	36,139 (1,759)	54,968 (2,754)
Redundancy	4.9 (3.6)	14.5 (14.8)	3.6 (3.6)
I/σ	22.6 (2.3)	43.2 (7.1)	24.7 (2.3)
Completeness, %	98.9 (90.6)	99.9 (100)	98.4 (99.4)
R_{merge}, %*	6.7 (62.5)	11.7 (51.9)	6.7 (61.2)
Structure refinement			
Resolution (Å)	50–2.5 (2.53–2.5)		50–2.4 (2.45–2.4)
R_{cryst}/R_{free}, %†	17.3 (25.6)/22.5 (32.1)		19.2 (27.8)/24.1 (33.2)
rmsd bonds, Å/angles, °	0.008/1.1		0.004/0.8
Average B factor	63.1		53.5
No. of atoms			
Protein atoms	22,833		7,302
Water molecules	356		57
Other molecules	88		
Ramachandran plot‡			
Favored regions, %	97.4		98.2
Allowed regions, %	2.6		1.7
Outliners, %	0.0		0.1

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

*$R_{merge} = \sum|I_i - I_m|/\sum I_i$, where I_i is the intensity of the measured reflection and I_m is the mean intensity of all symmetry related reflections.

†$R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure factors. $R_{free} = \Sigma_T||F_{obs}| - |F_{calc}||/\Sigma_T|F_{obs}|$, where T is a test data set of about 5% of the total reflections randomly chosen and set aside before refinement.

‡Defined by MolProbity (1).

1. Davis IW, et al. (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35(Web Server issue):W375–W383.

Table S2. Summary of reported disease mutations located in the GTD region of three MyoV paralogs

Gene	Mutation	Protein prediction	Related disease	Source	Corresponding residue in mouse MyoVa	Expected effect on protein folding
Human MYO5A	c.2332C > T	p.R778X	GS	(1, 2)	R778	Lacks the whole GTD
	c.4634ins47	p.K1545fs	GS	(1, 2)	K1543	Lacks most of GTD
	c.4453C > T	p.Q1485X	Lung cancer	COSMIC*	Q1483	Lacks most of GTD
	c.4520G > A	p.G1507E	Melanoma	(3)	G1505	Alters a turn conformation (Fig. S3*B*)
	c.4559C > T	p.A1520V	Lung cancer	COSMIC*	A1518	Mildly affects a hydrophobic core (Fig. S3*A*)
	c.4592C > T	p.S1531L	Skin cancer	(4)	S1529	Breaks a hydrogen bond (Fig. S3*A*)
	c.4687C > G	p.R1563G	Lung cancer	COSMIC*	R1561	Breaks a salt bridge and a hydrogen bond (Fig. S3*B*)
	c.4744A > G	p.T1582A	Colon cancer	Cancer Genome Atlas Network	T1580	Breaks a hydrogen bond (Fig. S3*B*)
	c.4750C > T	p.R1584C	Skin cancer	(4)	R1582	Breaks a salt bridge (Fig. S3*B*)
	c.4775A > G	p.N1592S	Lung cancer	(5)	N1590	Breaks a hydrogen bond (Fig. S3*A*)
	c.4938G > C	p.L1646F	Lung cancer	(5)	L1644	Unknown[†]
	c.4970G > T	p.G1657V	Lung cancer	(5)	G1655	Unknown[†]
	c.5308A > T	p.M1770L	Colon cancer	COSMIC*	M1768	Mildly affects a hydrophobic core (Fig. S3*D*)
Horse MYO5A	c.4183del1	p.R1395fs	LFS	(6)	R1369	Lacks the whole GTD
Human MYO5B	c.4667_4668TT > GC	p.L1556R	MVID	(7)	F1562	Disrupts a central hydrophobic core (Fig. S3*A*)
	c.4755_4756dupT	p.D1586X	MVID	(7)	D1592	Lacks most of GTD
	c.4840C > T	p.Q1614X	MVID	(7)	Q1620	Lacks the whole subdomain II
	c.5392C > T	p.R1795X	MVID	(8)	R1800	Lacks most of GTD
	c.4710G > A	p.M1570I	Breast cancer	(9)	M1576	No (Fig. S3*C*)
	c.5072C > A	p.A1691E	Renal cell carcinoma	(10)	A1696	Disrupts a few hydrophobic interactions (Fig. S3*C*)
	c.5108T > C	p.V1703A	Endometrial, kidney, and lung cancers	COSMIC*	M1708	Disrupts a hydrophobic core (Fig. S3*C*)
	c.5114C > T	p.S1705F	Skin cancer	(4)	S1710	Breaks a hydrogen bond (Fig. S3*C*)
	c.5263G > A	p.E1755K	Skin cancer	(4)	D1760	None (Fig. S3*D*)
	c.5272G > C	p.E1758Q	Urinary tract cancer	COSMIC*	E1763	Breaks a salt bridge (Fig. S3*D*)
	c.5362C > T	p.R1788W	Breast cancer	(9)	R1793	Breaks a hydrogen bond (Fig. S3*C*)
	c.5468C > A	p.P1823Q	Colon cancer	Cancer Genome Atlas Network	P1828	Alters a loop conformation (Fig. S3*B*)
	c.5497T > G	p.S1833A	Breast cancer	COSMIC*	T1838	Breaks a hydrogen bond (Fig. S3*B*)
Human MYO5C	c.4706C > T	p.P1569L	Skin cancer	(4)	P1680	Alters a turn conformation (Fig. S3*B*)
	c.4848G > T	p.W1616C	Colon cancer	(11)	W1727	Disrupts a central hydrophobic core (Fig. S3*C*)
	c.5119G > A	p.D1707N	Lung cancer	COSMIC*	D1818	None (Fig. S3*C*)

GS, Griscelli syndrome.

*The mutation data were obtained directly from COSMIC (http://cancer.sanger.ac.uk/cancergenome/projects/cosmic).

[†] The mutation sites are located on the disordered region in MyoVa-GTD.

1. Pastural E, et al. (2000) Two genes are responsible for Griscelli syndrome at the same 15q21 locus. *Genomics* 63(3):299–306.
2. Sanal O, et al. (2002) Griscelli disease: Genotype-phenotype correlation in an array of clinical heterogeneity. *J Clin Immunol* 22(4):237–243.
3. Berger MF, et al. (2012) Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature* 485(7399):502–506.
4. Durinck S, et al. (2011) Temporal dissection of tumorigenesis in primary cancers. *Cancer Discov* 1(2):137–143.
5. Imielinski M, et al. (2012) Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* 150(6):1107–1120.
6. Brooks SA, et al. (2010) Whole-genome SNP association in the horse: identification of a deletion in myosin Va responsible for Lavender Foal Syndrome. *PLoS Genet* 6(4):e1000909.
7. Ruemmele FM, et al. (2010) Loss-of-function of MYO5B is the main cause of microvillus inclusion disease: 15 novel mutations and a CaCo-2 RNAi cell model. *Hum Mutat* 31(5):544–551.
8. Müller T, et al. (2008) MYO5B mutations cause microvillus inclusion disease and disrupt epithelial cell polarity. *Nat Genet* 40(10):1163–1165.
9. Shah SP, et al. (2012) The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* 486(7403):395–399.
10. Peña-Llopis S, et al. (2012) BAP1 loss defines a new class of renal cell carcinoma. *Nat Genet* 44(7):751–759.
11. Wood LD, et al. (2007) The genomic landscapes of human breast and colorectal cancers. *Science* 318(5853):1108–1113.