Supporting Information

Pérez-Victoria et al. 10.1073/pnas.1004756107

SI Text

SI Materials and Methods. ***Cloning, protein expression, and purification.*** Based on secondary structure prediction and sequence similarity alignments using PredictProtein and BLAST programs (1, 2), respectively, we designed 32 constructs of the Vps54 C-terminal region. All constructs were generated by PCR amplification of complementary DNA coding for mouse Vps54 (GenBank: CAI24173.1) and inserted into the GST-parallel vector (3) using NcoI and XbaI restriction sites. All constructs were verified by DNA sequencing. After small-scale expression tests in the *Escherichia coli* BL21 (DE3) strain and subsequent solubility evaluation, a construct comprising residues 836–977 of Vps54 was selected for large-scale production. Cells were grown in LB medium to OD_{600} of 0.8 at which protein expression was induced with 1 mM IPTG at 20 °C overnight. Harvested cells were resuspended in PBS complemented with 0.2 mg/mL DNase, 1 mg/mL lysozyme, and 10 mM β-mercaptoethanol (β-ME), and lysed at 4 °C by high-pressure homogenization at 27 Kpsi (Constant System, Ltd). All subsequent purification steps were carried out at 4 °C. Insoluble material was removed by ultracentrifugation and supernatant was loaded onto a gravity column packed with glutathione-Sepharose™ 4B (GE Healthcare) and preequilibrated in PBS with 10 mM β-ME. After extensive washing, the protein was eluted by addition of 50 mM reduced glutathione to the buffer and GST was cleaved with tobacco etch virus protease (4) while dialyzed in PBS, 10 mM β-ME overnight. The sample pH was then adjusted to 5.6 using a stock solution of 1 M MES pH 5.6 and loaded onto an ion exchange chromatography column, HiTrap SP HP 5 mL (GE Healthcare). Vps54-CT was eluted by an isocratic gradient from 0.1 to 1 M NaCl in 30 column volumes. Fractions containing Vps54-CT identified by SDS-PAGE were concentrated and loaded onto a HiLoad 16/60 Superdex 75 column (GE Healthcare) equilibrated in 150 mM NaCl, 10 mM β-ME, 50 mM MES pH 5.6. Purified Vps54-CT was concentrated to 1.5 mg/mL, as measured by UV absorbance, and stored at −80 °C.

Selenomethionine (SeMet)-substituted protein was produced in *E. coli* B834 (Stratagene) a methionine auxotroph strain, using SelenoMet medium (Molecular Dimensions) following manufacturer's instructions. SeMet Vps54-CT purification was carried out under the same conditions as that of the native protein. The efficiency of SeMet incorporation was evaluated by mass differences between native (unlabeled) and SeMet-labeled protein samples using MALDI-TOF mass spectrometry. The observed mass differences confirmed 100% selenomethionine incorporation at seven expected sites.

Crystallization and data collection. Crystals used for data collection were grown by hanging drop vapor diffusion at 4 °C by mixing a 1:1 ratio of protein stock (10 mg/mL) to well solution. The three crystal forms used for Vps54-CT structure determination were obtained from three different reservoir solutions. SeMet Vps54-CT crystals in C2 space group were obtained in 20% PEG 3000, 0.2 M diammonium sulphate, and 5% glycerol. Native Vps54-CT crystals in C2 space group were obtained in 22–24% PEG 3350, 0.1 M BisTris (pH 8.5), and 0.1 M I3C, whereas Vps54-CT crystals in P212121 space group were obtained in 18% PEG 3350, 0.2 M $MgCl_2$, and 5% glycerol. All crystal forms appeared within one week and were flash-frozen under liquid nitrogen using paraffin oil as cryoprotectant.

A dataset was collected at 1.7-Å wavelength from a native crystal obtained in presence of the additive I3C (Hampton Research) at beamline ID23-1 of the European Synchrotron Radiation Facility. Additionally, two-wavelength (peak and inflection) multiwavelength anomalous dispersion datasets were collected from a crystal of SeMet Vps54-CT at the same beamline. Diffraction datasets were processed with the HKL2000 software package (5) without merging symmetry equivalents. Because no significant anomalous signal was observed from the I3C iodine atoms the dataset was further treated as native.

Structural figures were generated using PyMOL, LIGPLOT, and NACCESS programs (6, 7).

Multiple sequence alignment. Initial multiple sequence alignment of the murine Vps54 (residues 656–977) and C and D domains from structures of representative multisubunit tethering complexes was generated by using the MUSCLE program (8). This alignment was edited manually based on the structural alignments between the corresponding protein structures reported by the Dali server (9). The location and alignment of the C domain of Vps54 was obtained by applying a generalized sequence profile method and *pftools* package (10). We constructed the profile from structure-based alignments of the C domains from the known structures and used it to identify the C domain in Vps54. The best match produced by this profile (positions 685–831) had E value <0.0002. The E values of the matches were calculated as described by Hofmann and Bucher (1995) (11). Subsequent examination of this match found mostly hydrophobic residues at positions that should be apolar because the corresponding residues of known structures have interior locations.

Circular dichroism (CD). CD data were recorded by a Jasco J-815 CD spectrometer equipped with thermoelectric temperature control. Several dilutions of Vps54-CT and L967V mutant were prepared and placed in a 2-mm cell. CD wavelength scan was performed at a constant temperature of 30 °C between 200 and 280 nm with a scanning speed of 50 nm/min and a data pitch of 1 nm until the decay reached a plateau. Data analysis was done using Origin and JASCO Spectra Manager® integrated software.

Mammalian expression constructs and reagents. Complementary DNAs of full-length murine Vps54 and deletion mutants were cloned in-frame with a C-terminal V5 tag in pEF6-V5-His TOPO (Invitrogen) vector and verified by sequencing. The wobbler-V5 and Vps54 C-terminal-V5 (Vps54 535–977-V5) constructs were described elsewhere (12).

HeLa cells were transfected with oligofectamine or Lipofectamine 2000 (Invitrogen) as described (12). Protein lysates were quantified with the BCA kit (Pierce), resolved by electrophoresis on 4–12% NuPage gels (Invitrogen) and immunoblotted following standard conditions.

Coimmunoprecipitations. HeLa cells were transfected for 16 h in 100 mm plates with 5–10 μg cDNA for the different Vps54 alleles containing a C-terminal V5 tag, lysed under native conditions in lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 2 mM EDTA, 10% glycerol, 0.5% NP-40, pH 7.4), and immunoprecipitated with anti-V5 monoclonal antibody and True-Blot™ anti-mouse Ig IP beads (eBioscience). Immunoprecipitates were analyzed by immunoblotting for V5 or endogenous Vps53.

Cycloheximide chase. HeLa cells were transfected in six-well plates for 18 h with the different Vps54 cDNA constructs. The medium was then aspirated and replaced with medium containing

40 μ/mL of cycloheximide (Sigma) to stop protein synthesis. Aliquots were taken at different times, lysates prepared in lysis buffer, quantified, and 10 μg of proteins loaded per lane for immunoblotting.

Quantification of Vps54 in tissues from wobbler and control mice. Pieces from liver, cervical spinal cord, lumbar spinal cord, and brain cortex were ground in a mortar under liquid nitrogen. Proteins were extracted under native conditions in the above mentioned lysis buffer, quantified, and 10–30 μg were analyzed by SDS-PAGE and immunoblotting for endogenous Vps54 and Vps53. Quantification was performed by densitometry. Statistical significance was assessed on the average data from all the blots for each individual mouse analyzed.

Antibodies. Antibodies to the V5 epitope, p230 and TGN46 for immunofluorescence microscopy, to Vps53 and the V5 epitope for immunoblotting, and the corresponding secondary antibodies were described before (12). A new polyclonal antibody to murine Vps54 was raised in rabbits using the whole crystallized protein domain (mVps54 835–977) as antigen (Covance). This antibody recognized endogenous Vps54 as a specific band of ~120 kDa by immunoblotting at a 1:1,000 serum dilution, and was able to immunoprecipitate Vps54-V5 and wobbler-V5 with similar efficiencies.

Mouse colonies and tissue dissection. All experimental procedures met the ethical guidelines of the Cleveland Clinic Foundation Animal Research Committee and were approved by the Research Program Committee. We used the B6NZB hybrid wobbler mouse and performed PCR amplification of microsatellite DNA containing variable numbers of GT repeats located in the 5′ noncoding region of the glutamine synthetase pseudogene 1 (*glns-ps1*) to identify mutant and wild-type mice. Deeply anesthetized mice had their brains, spinal cords, and liver rapidly removed, snap frozen, and stored at −80 °C until used.

1. Rost B, Yachdav G, Liu J (2004) The PredictProtein server. *Nucleic Acids Res* 32:W321–326 (Web server issue).
2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410.
3. Sheffield P, Garrard S, Derewenda Z (1999) Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expres Purif* 15(1):34–39.
4. Kapust RB, Waugh DS (2000) Controlled intracellular processing of fusion proteins by TEV protease. *Protein Expres Purif* 19(2):312–318.
5. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol*, eds Carter CW, Sweet RM (Academic Press, New York), Vol 276, pp 307–326.
6. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 8(2):127–134.
7. Hubbard SJ, Thorton JM (1993) NACCESS Computer Program (Dept of Biochemistry and Molecular Biology, University College London, London).
8. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797.
9. Holm L, Kaariainen S, Rosenstrom P, Schenkel A (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics* 24(23):2780–2781.
10. Bucher P, Karplus K, Moeri N, Hofmann K (1996) A flexible motif search technique based on generalized profiles. *Comput Chem* 20(1):3–23.
11. Hofmann K, Bucher P (1995) The FHA domain: A putative nuclear signalling domain found in protein kinases and transcription factors. *Trends Biochem Sci* 20(9):347–349.
12. Perez-Victoria FJ, Mardones GA, Bonifacino JS (2008) Requirement of the human GARP complex for mannose 6-phosphate-receptor-dependent sorting of cathepsin D to lysosomes. *Mol Biol Cell* 19(6):2350–2362.

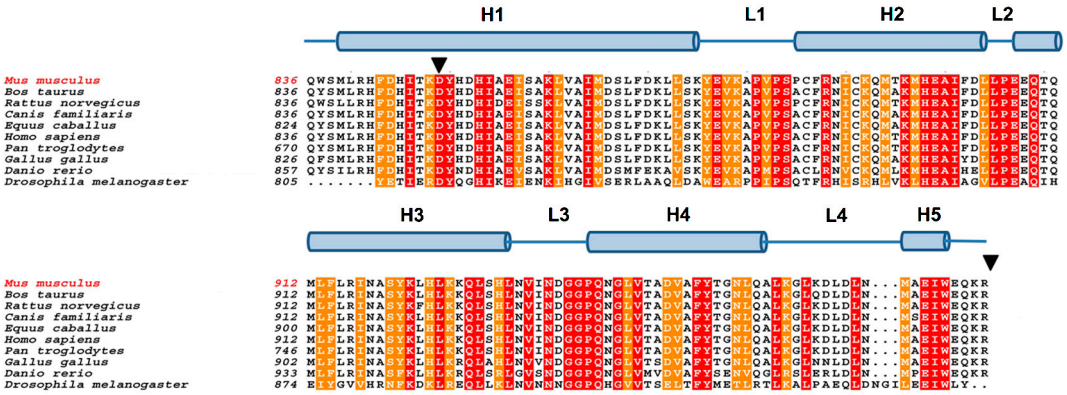

Fig. S1. Sequence alignment based on residues 836–977 from *Mus musculus* Vps54 (GenBank: CAI24173.1) and secondary structure elements from the crystal structure. The D domain of Vps54 (residues 848–977) is delimited by two black triangles. Completely conserved residues and residues with 90% similarity are colored red and orange, respectively. The alignment and graphical representation were done using the ClustalW2 and ESPript programs respectively.

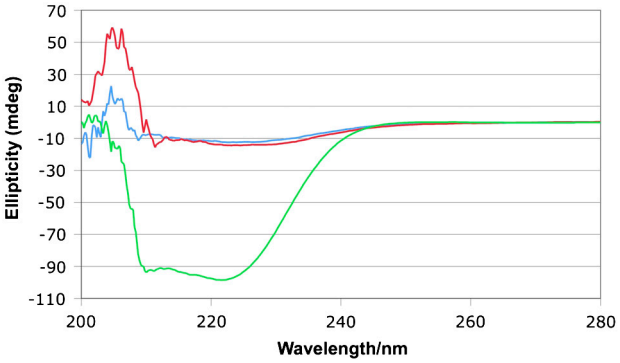

Fig. S2. Irreversible thermal denaturation of Vps54 D domain. Far UV CD spectra at 4 °C (green), after thermal denaturation at 46 °C (red), and after subsequent cooling to 4 °C (blue).

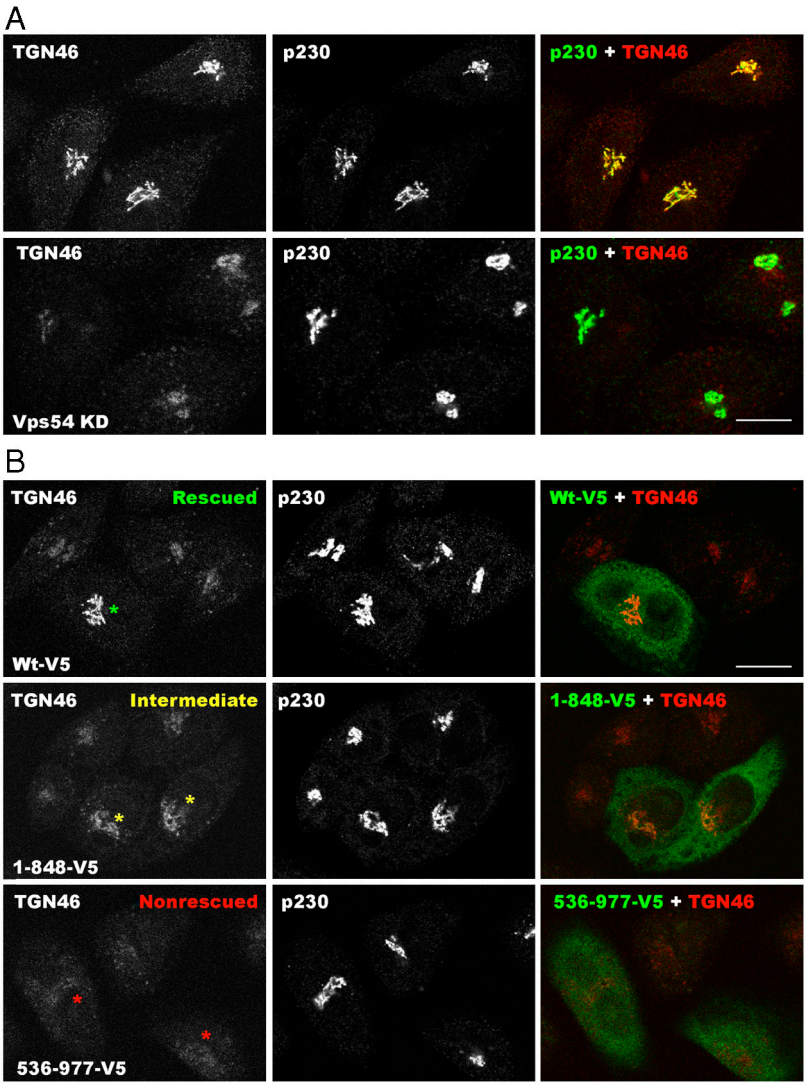

Fig. S3. Effect of Vps54 KD and rescue on the localization of TGN46 to the *trans*-Golgi network (TGN). (*A*) Depletion of Vps54 in HeLa cells leads to mislocalization of TGN46 (red) but not the TGN-resident protein, p230 (green). (*B*) Examples of different rescue phenotypes. HeLa cells depleted of Vps54 were transfected with the indicated siRNA-resistant Vps54 alleles containing a C-terminal V5 tag, fixed and triple-labeled with antibodies to TGN46, V5, and p230. The staining pattern for TGN46 in each transfected cell (denoted by asterisks) was scored visually, distinguishing three different phenotypic groups, as shown in the images: (*i*) rescued cells, displaying intense TGN46 staining in the TGN area (similar to nondepleted cells, green asterisk); (*ii*) nonrescued cells, in which TGN46 remained mislocalized and not concentrated at the TGN (similar to depleted cells, red asterisks); and (*iii*) intermediate phenotype, with TGN46 delineating the TGN but without reaching the normal intensity at this organelle (yellow asterisks). (Scale bar, 10 μm.)

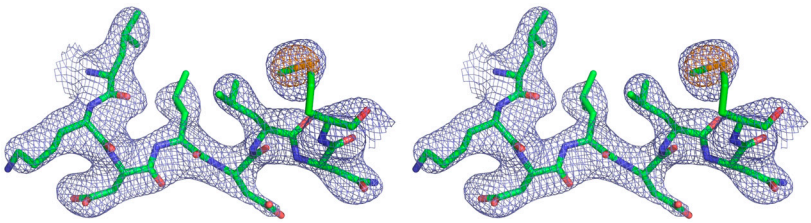

Fig. S4. Structure determination of the C-terminal domain of Vps54. Stereo view of the electron density map contoured at 1.1σ from the 2WMADN solvent-flattened dataset at 2.4-Å resolution. The map view corresponds to L962-A970 region. Anomalous Fourier map contoured at 10σ corresponding to selenium position of SeMet969 is shown in orange.

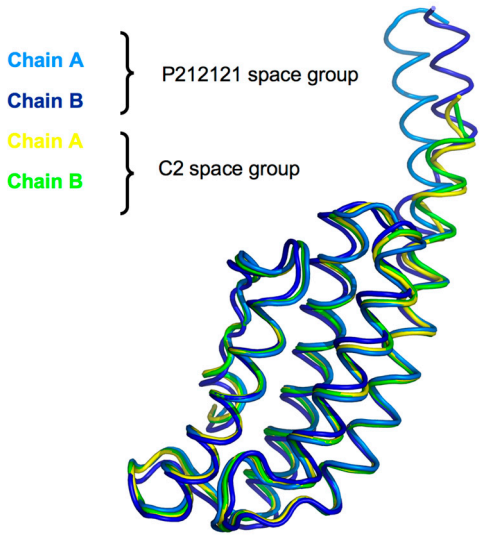

Fig. S5. Comparison of the four Vps54 C-terminal monomers. Dark- and light-blue models correspond to the two monomers from the crystallographic asymmetric unit corresponding to the P212121 space group, whereas green and yellow models correspond to the two monomers from the crystallographic asymmetric unit corresponding to the C2 space group.

Table S1. Data collection and refinement statistics

			SeMet MAD	
	Native (1)	Native (2)	Peak	Inflection
Data collection				
Resolution, Å	50-1.7 (1.76-1.70)	50-2.6 (2.7-2.6)	50-2.4 (2.5-2.4)	50-2.6 (2.7-2.6)
Wavelength, Å	0.87260	1.70000	0.97910	0.97940
Completeness, %	99.0 (96.0)	99.3 (94.6)	98.4 (90.3)	97.5 (85.8)
R_{merge}, %	6.3 (45.0)	7.0 (18.8)	6.6 (20.0)	5.6 (11.9)
Redundancy	7.3 (6.8)	6.3 (3.9)	3.3 (2.3)	3.1 (2.2)
$I/\sigma I$	26.5 (4.0)	24.3 (5.6)	12.3 (3.9)	15.7 (6.2)
Space group	$P2_12_12_1$	C2	C2	
Unit cell				
a, b, c, Å	29.8, 77.1, 120.1	126.3, 30.1, 88.8	125.9, 30.3, 89.0	
α, β, γ, °	90.0, 90.0, 90.0	90.0, 120.9, 90.0	90.0, 120.5, 90.0	
Refinement				
R_{work}/R_{free}, %	17.6/21.1		23.8/28.7	
No. reflections	29964		11296	
% reflections for R_{free}	5		5	
Number of atoms				
Protein	2355		2103	
Waters	330		63	
Average B factors, Å^2				
Protein	19.5		50.8	
Water	31.6		41.6	
rmsd				
Bond lengths, Å	0.012		0.002	
Bond angles, °	1.284		0.547	
Ramachandran statistics				
Most favored, %	98.1		94.1	
Additional allowed, %	1.9		5.4	
Generously allowed, %	0		0.4	

Data in parentheses are for highest-resolution shells.