

# Functional architecture of the retromer cargo-recognition complex

Aitor Hierro<sup>1\*</sup>, Adriana L. Rojas<sup>1\*</sup>, Raul Rojas<sup>2</sup>, Namita Murthy<sup>2</sup>, Grégory Effantin<sup>3</sup>, Andrey V. Kajava<sup>4</sup>, Alasdair C. Steven<sup>3</sup>, Juan S. Bonifacino<sup>2</sup> & James H. Hurley<sup>1</sup>

The retromer complex<sup>1,2</sup> is required for the sorting of acid hydrolases to lysosomes<sup>3–7</sup>, transcytosis of the polymeric immunoglobulin receptor<sup>8</sup>, Wnt gradient formation<sup>9,10</sup>, iron transporter recycling<sup>11</sup> and processing of the amyloid precursor protein<sup>12</sup>. Human retromer consists of two smaller complexes: the cargo recognition VPS26–VPS29–VPS35 heterotrimer and a membrane-targeting heterodimer or homodimer of SNX1 and/or SNX2 (ref. 13). Here we report the crystal structure of a VPS29–VPS35 subcomplex showing how the metallophosphoesterase-fold subunit VPS29 (refs 14, 15) acts as a scaffold for the carboxy-terminal half of VPS35. VPS35 forms a horseshoe-shaped, right-handed,  $\alpha$ -helical solenoid, the concave face of which completely covers the metal-binding site of VPS29, whereas the convex face exposes a series of hydrophobic interhelical grooves. Electron microscopy shows that the intact VPS26–VPS29–VPS35 complex is a stick-shaped, flexible structure, approximately 21 nm long. A hybrid structural model derived from crystal structures, electron microscopy, interaction studies and bioinformatics shows that the  $\alpha$ -solenoid fold extends the full length of VPS35, and that VPS26 is bound at the opposite end from VPS29. This extended structure presents multiple binding sites for the SNX complex and receptor cargo, and appears capable of flexing to conform to curved vesicular membranes.

The retromer cargo-recognition complex consists of the 38-kDa VPS26, 20-kDa VPS29 and 92-kDa VPS35 subunits. The structures of the two smaller subunits have been determined in isolation. VPS26 is a structural cousin of the arrestins<sup>16</sup>, a family of trafficking proteins that directly bind to cell surface receptors and direct their internalization. VPS29 has a metallophosphoesterase fold<sup>14,15</sup> that can bind two metal ions. Compared to functional metallophosphoesterases, a key histidine residue that serves as a catalytic base in the metallophosphoesterase active site is replaced by Phe 63. Thus, VPS29 is completely inactive with respect to generic phosphatase substrates<sup>15</sup>. However, metal-dependent activity *in vitro* against a phosphorylated peptide from a major retromer cargo, the cation-independent mannose 6-phosphate receptor (CI-MPR), has been reported<sup>17</sup>. Despite its centrality to multiple trafficking pathways, the precise function of retromer has been enigmatic. Various proposals have emphasized potential roles as a coat, adaptor, or cargo protein phosphatase. Here we take a structural approach to gain insight into the function of retromer.

The crystal structure of a VPS29–VPS35 subcomplex, containing the C-terminal 40% of the large VPS35 subunit, was determined at 2.8 Å resolution (Fig. 1a, Supplementary Fig. 1 and Supplementary Table 1). The C-terminal portion of VPS35 consists of a single

right-handed superhelix with a pitch of 12 Å and a total of 13 helices (Fig. 1a and Supplementary Fig. 2). VPS35 resembles many other helical solenoid proteins, including other important proteins in coated vesicle trafficking. These include the trunk domains of the large subunits of the AP adaptor protein complexes<sup>18,19</sup> (Supplementary Fig. 3), the function of which is to link clathrin to cargo and membranes, the helical repeat regions of the heavy chain of clathrin itself<sup>20,21</sup>, and the Sec13–Sec31 assembly unit of the COPII coat<sup>22</sup>. VPS35 wraps itself nearly halfway around the VPS29 subunit, burying 3,300 Å<sup>2</sup> of solvent-accessible surface area (Fig. 1b, c). The VPS35 binding site on VPS29 includes the entire metal-binding site, as well as flanking residues. Ile 91 of VPS29, previously shown to contact VPS35 (ref. 15), interacts extensively with VPS35 near the centre of the VPS29–VPS35 interface (Fig. 1a). VPS35 and its C-terminal fragment are both poorly stable when expressed in isolation, consistent with a fold that depends on scaffolding by VPS29 for its stability. The even-numbered helices of the convex face of VPS35 are splayed apart from each other because the solenoid is curved. This surface thus consists of a series of ridges separating hydrophobic grooves, suggestive of potential cargo-binding sites (Fig. 1d). These grooves are reminiscent of the CI-MPR binding site on the helical solenoidal VHS domains of the GGA clathrin adaptor proteins<sup>23,24</sup>. The binding site for Vps10 to yeast Vps35 (ref. 25) maps to part of this surface (Fig. 1d).

Two models have been proposed for the function of VPS29. In the first, VPS29 is a scaffold for interaction with VPS35 and SNX1 (ref. 15) with no phosphatase activity of functional significance. In the second, VPS29 is an active protein phosphatase for the CI-MPR and perhaps other receptors<sup>17</sup>. The complete burial of the putative VPS29 metallophosphoesterase active site prompted us to revisit the report of CI-MPR dephosphorylation activity by retromer<sup>17</sup>. Working with preparations purified to homogeneity, neither the complete VPS26–VPS29–VPS35 complex nor the VPS29–VPS35 subcomplex had detectable activity (Supplementary Fig. 4). Furthermore, VPS29 alone had no detectable activity.

The lack of *in vitro* phosphatase activity suggests, but does not prove, that the main function for the metallophosphoesterase site is to provide a scaffold for VPS35. VPS29 must associate with VPS35 to maintain its known biological functions in sorting<sup>15</sup>. We considered whether a conformational change in VPS35 could expose the VPS29 metallophosphoesterase site without completely disassembling the complex. One region of VPS35 covers the putative VPS29 active site and is thus potentially an inhibitory contact site, whereas another region is remote from the putative active site and unlikely to be involved in inhibition (Fig. 1b, c). Each of these two

<sup>1</sup>Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, <sup>2</sup>Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, and <sup>3</sup>Laboratory of Structural Biology, National Institute of Arthritis, Musculoskeletal, and Skin Diseases, National Institutes of Health, US Department of Health and Human Services, Bethesda, Maryland 20892, USA. <sup>4</sup>Centre de Recherches de Biochimie Macromoléculaire, CNRS, University of Montpellier, 1919 Route de Mende, 34293 Montpellier, France.

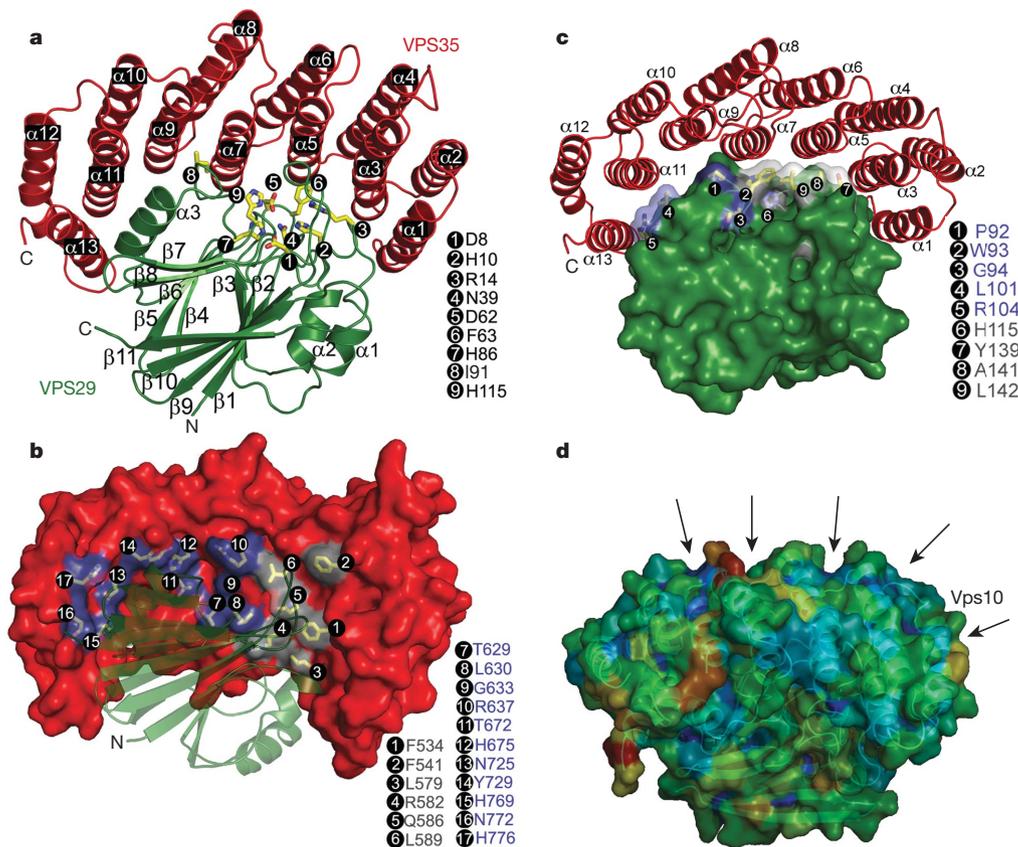
\*These authors contributed equally to this work.

regions comprises roughly half of the total contact area. The scaffold model predicts that mutations that selectively disrupt the putative inhibitory region should prevent assembly of the complex. The phosphatase model, on the other hand, predicts that these mutations should allow integration into the retromer but lead to constitutive exposure of the metallophosphoesterase site. HeLa cells were transfected with mammalian expression plasmids encoding haemagglutinin (HA)-tagged wild-type or mutant VPS35 constructs carrying multiple substitutions on residues (Fig. 1b, grey residues) covering the putative VPS29-phosphatase active site or residues (Fig. 1b, purple residues) distal to this region. After transfection, the proteins were isolated by immunoprecipitation and analysed by SDS-polyacrylamide gel electrophoresis (PAGE) and immunoblotting with antibodies to VPS26 and VPS29 (Fig. 2a, top). We found that wild-type HA-VPS35 can assemble with endogenous VPS29 and VPS26. In contrast, all of the HA-tagged VPS35 mutants in both regions prevented assembly with VPS29, consistent with the scaffolding model. All of the recombinant HA-tagged VPS35 constructs were expressed at similar levels in the transfected cells (Fig. 2a, bottom). In addition, these HA-tagged VPS35 mutants all assembled with VPS26 (Fig. 2a), indicating that they are capable of proper folding and that the VPS26 and VPS29 binding sites are independent of one another. We used double-labelling fluorescent microscopy to study the cellular distribution of these HA-tagged constructs (Supplementary Fig. 5a). Unlike wild-type HA-VPS35, which is mostly endosomal, all of the mutant HA-VPS35 constructs showed a diffuse cytosolic distribution.

Similarly, we transfected wild-type Myc-tagged VPS29 and versions harbouring mutations of VPS35-interacting residues surrounding the metallophosphoesterase binding site (Fig. 1c, grey residues) and

remote from it (Fig. 1c, purple residues). These constructs were immunoprecipitated and analysed by SDS-PAGE and immunoblotting using antibodies to VPS35 and VPS26 (Fig. 2b, top). All but one (L101D) of the characterized Myc-tagged VPS29 mutants failed to assemble with endogenous VPS35, re-emphasizing that all regions of the interface are important for assembly. We also tested nine mutants of the VPS29 metallophosphoesterase site and found that only three (D8N, H86A and H117A) assembled partially with VPS35, whereas the remaining mutants did not assemble (Supplementary Table 2). The levels of these mutants were much lower than wild-type Myc-tagged VPS29 and almost undetectable in lysates by immunoblot analysis (Supplementary Table 2). This led us to conclude that an intact metal-binding site is required for the stability of VPS29.

Indeed, the use of a metallophosphoesterase fold for non-catalytic purposes may be common. Of 5,307 putative metallophosphoesterase catalytic domains (see Methods), only 41% of the sequences conserve eight key residues directly involved in metal binding or catalysis. Even though catalytically inactive metallophosphoesterases seem to comprise most of this vast protein fold family, little is known about their functions. It will be interesting to learn whether the VPS29 example is a paradigm for other members of this class. How might so many such folds have evolved? The structure of full-length protein phosphatase 5 shows that its TPR domain, a helical solenoid, is inserted into and sterically blocks the active site of the catalytic domain<sup>26</sup>, and suggests the nature of a potential evolutionary precursor. We suggest that most of the metal ligands were retained during evolution for stability reasons, despite their lack of catalytic function.



**Figure 1 | Structure of the VPS29-VPS35 subcomplex.** **a**, VPS29 is green and VPS35 red. **b**, The surface of VPS35 is shown, with the residues blocking the metallophosphoesterase site of VPS29 in grey, and other residues that contact VPS29 in purple. **c**, The surface of VPS29 is shown, with residues surrounding the metallophosphoesterase site in light blue, and other VPS35-contacting residues in purple. **d**, Hydrophobic grooves on the outer surface

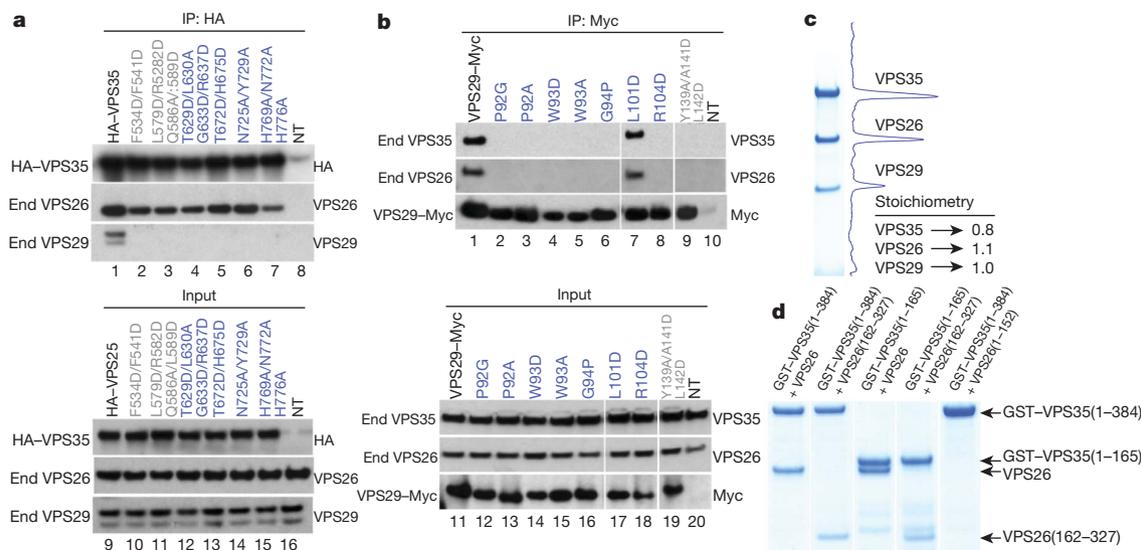
of VPS35 are formed between even-numbered helices. The probability of the surface to participate in ligand binding was coloured from lowest to highest in a blue to red gradient using the hotpatch server (<http://hotpatch.mbi.ucla.edu/>). Structural figures were generated with pymol (<http://www.pymol.org/>).

The intact recombinant VPS26–VPS29–VPS35 complex was found to contain an equimolar ratio of the three subunits by Coomassie-stained SDS–PAGE (Fig. 2c) and by quantitative amino acid analysis (data not shown). The elution of the complex on gel filtration chromatography (data not shown) is potentially consistent with either an elongated 1:1:1 complex or a globular 2:2:2 complex. Given that the electron microscopy described below indicates that the structure is elongated and occupies a volume consistent with a 1:1:1 complex, we favour this interpretation. Through co-expression of truncated recombinant proteins, we mapped the interaction of VPS26 and VPS35 to the C-terminal lobe of VPS26 and the N-terminal 165 amino acids<sup>4</sup> of VPS35 (Fig. 2d). The organization of the large subunit, VPS35, is central to understanding the structure of retromer. Analysis of the VPS35 sequence showed that it consists of regular repeats of a two-helix motif (Fig. 3a), strongly suggesting that the  $\alpha$ -solenoid visualized in the crystallized fragment extends throughout essentially the entire structure. The total number of helices is anticipated to be 34, including the 13 in VPS35-C. This corresponds to a contour length of  $\sim 204$  Å for the complete molecule, given a pitch of 12 Å per two-helix turn of the solenoid in the crystallized portion. The conformations of other  $\alpha$ -solenoids vary from ring-like to quasi-linear<sup>27,28</sup>. How is the putative  $\alpha$ -solenoid arranged in this case? To investigate this question, we performed negative staining electron microscopy. The images (for example, see Supplementary Fig. 6) exhibit some variability but many molecules are filamentous and about 210 Å long. To enhance detail, we performed image averaging after alignment and classification. The class averages exhibit some variation in local structure but are rather consistent in length, between 210 and 220 Å (Fig. 3b). Each has five or six stain-excluding densities. These data can be accounted for by a model consisting of the VPS29–VPS35-C heterodimer at one end, extended by  $\sim 135$  Å of  $\alpha$ -solenoid with VPS26 attached near its distal end (Fig. 3c, d). This model was

tested by generating resolution-limited projections from the model for comparison with the negatively stained class averages (Fig. 3c). At one end, the structure is consistent with projections of the VPS29–VPS35-C subcomplex. The respective class averages correspond to different side views of this assembly, with some flexing about its mid-section (Supplementary Fig. 7). These data strongly support the inferred continuation of the VPS35  $\alpha$ -solenoid in an approximately linear conformation, with VPS26 bound near its end.

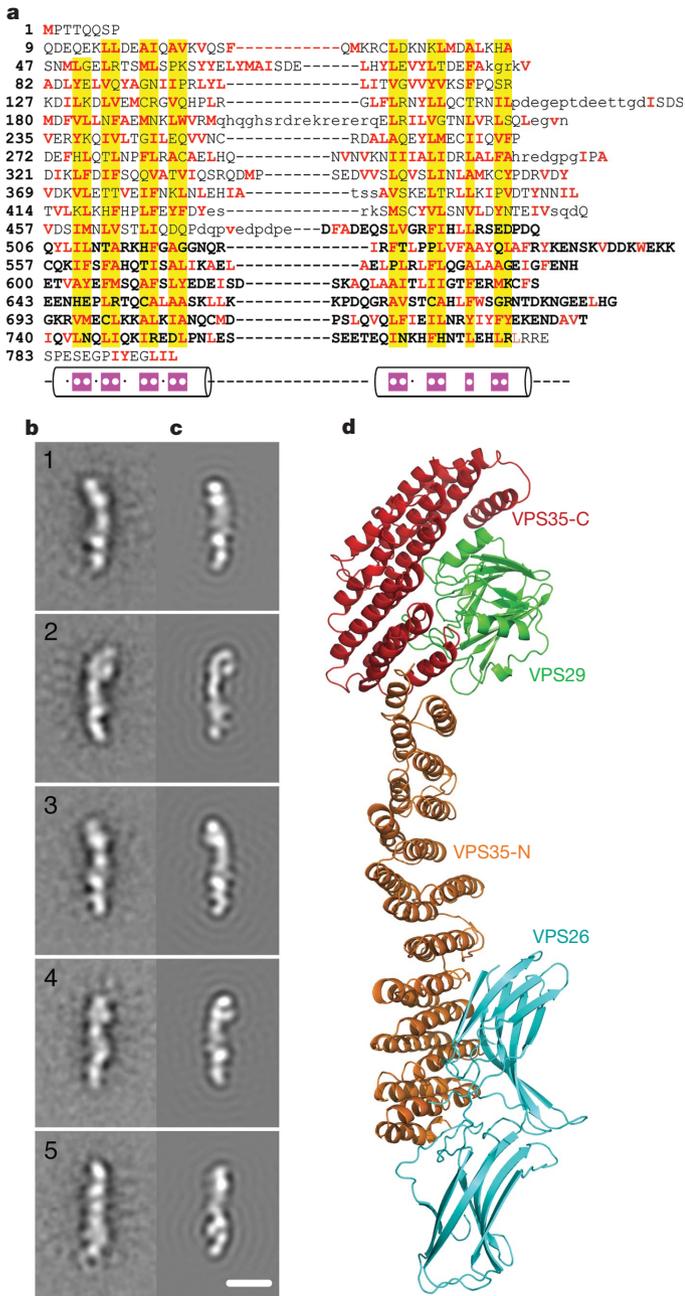
The overall structural model suggests a mechanism for cooperative binding to multiple membrane-anchored molecules (Fig. 4a). The membrane-binding SNX1 subunit interacts with a groove on VPS29 (at Leu 25 and Leu 152)<sup>15</sup> on the face distal to VPS35, and with two different sites on the N- and C-terminal halves of VPS35 (ref. 4). The N- and C-terminal halves of yeast Vps35 have been shown to interact with the distinct cargoes A-ALP and Vps10, respectively<sup>25</sup>. The Vps35 binding site on the A-ALP cytosolic domain is within 20 residues of the A-ALP transmembrane region<sup>29</sup>, dictating that the N-terminal half of Vps35 must be close to the membrane. The flexibility of the cargo-recognition complex suggests that it could adapt to the shape of curved tubulovesicular membranes and so bind multiple membrane-embedded cargo molecules. The curvature and contour length of the SNX dimer can be estimated from the related structure of the amphiphysin BAR domain<sup>30</sup>, which has a contour length of 150 Å along its inner-membrane binding surface. The cargo recognition complex has at least three distinct points of interaction with the SNX dimer. Their contour lengths are similar, suggesting that they could form curved side-by-side arrays at a 1:1 ratio (Fig. 4b). Although speculative, this conceptual model helps lay a foundation for further analysis of the assembly of the full retromer complex on cargo-bearing vesicles.

The clathrin and COPII coats, which direct the budding of roughly spherical vesicles, have been the principal models for the structural



**Figure 2 | Assembly of the cargo-recognition complex.** **a**, Mutations in the VPS29-interaction surface of VPS35 block assembly *in vivo*. Lysates from HeLa cells transfected with cDNAs encoding wild-type or mutant forms of HA-tagged VPS35, or lysates from non-transfected cells (lanes 8 and 16, NT), were subjected to immunoprecipitation (IP) using mouse monoclonal antibody to the HA epitope. Lysates (2% of total, lanes 9–16) and immunoprecipitates (lanes 1–8) were analysed by SDS–PAGE and immunoblotting with antibodies to the HA tag, VPS26 and VPS29. Coloured text matches the colour scheme in Fig. 1. **b**, Mutations in the VPS35-interaction surface of VPS29 abrogate assembly, with the exception of L101D. Lysates from HeLa cells transfected with cDNAs encoding wild-type or mutant VPS29–Myc, or lysates from non-transfected cells (lanes 10 and 20, NT), were subjected to immunoprecipitation using a mouse monoclonal antibody to the Myc epitope. Lysates (5% of total, lanes 11–20) and

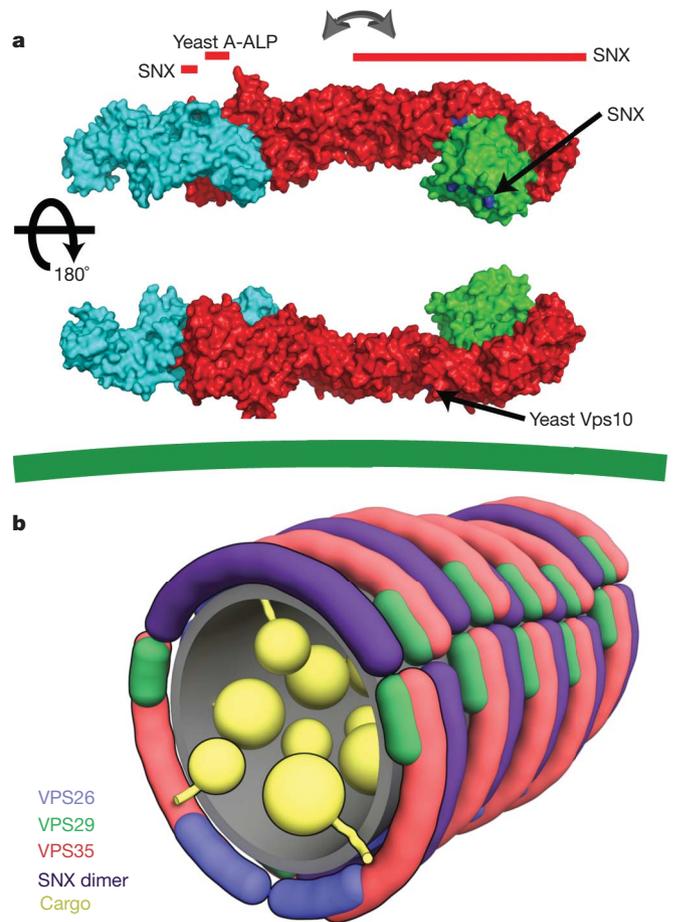
immunoprecipitates (lanes 1–10) were analysed by SDS–PAGE and immunoblotting with antibodies to VPS35, VPS26 and the Myc tag. **c**, The cargo-recognition complex is an equimolar complex of VPS26, VPS29 and VPS35. Gels were stained with Coomassie blue (Simply Blue SafeStain; Invitrogen). Image acquisition and analysis was done in an Epichemi Darkroom (UVP BioImaging Systems) using Labworks 4.5 software. The integrated area from each peak was normalized to the calculated molecular mass for each protein, and the value was determined relative to VPS29. **d**, Coomassie-stained gel showing that the N terminus of VPS35 forms a stable subcomplex with the C-terminal lobe of VPS26. For each lane, GST-tagged VPS35 fragments and untagged VPS26 or its fragments were co-expressed in *Escherichia coli* and purified by glutathione-Sepharose chromatography.



**Figure 3 | Structural analysis of the complete cargo-recognition complex.** **a**, Alignment of human VPS35 repeats. Columns of residues that are likely to form the hydrophobic core of the structure have a yellow background. Apolar residues are in red. Cylinders show the  $\alpha$ -helical regions. Magenta boxes within the  $\alpha$ -helices indicate the positions of apolar residues in the consensus sequences. The C-terminal region of the crystal structure is in bold. Loop regions predicted on the basis of the multiple sequence alignment of VPS35 proteins are shown by lower-case letters. **b**, Averaged images of the cargo-recognition complex from negative stain electron microscopy (Supplementary Fig. 6). The images were obtained by multivariate statistical analysis with reference-free alignment. The number of images per class 1 to 5 is respectively 288, 362, 367, 206 and 327. **c**, Corresponding projections, limited to 25 Å resolution, of the cargo-recognition complex model shown in **d** (C-terminal crystallized region of VPS35, red; VPS29, green; VPS26, cyan; and N-terminal modelled region of VPS35, orange). Each image in **c** is oriented such that VPS35–VPS29 corresponds always to the top part of the image and VPS26 to the bottom part, as in **d**. The correlation coefficient between the electron microscopy class average and the corresponding model projection for class 1 to 5 is respectively 0.79, 0.70, 0.86, 0.75 and 0.73. Scale bar in **c**, 100 Å.

1066

dissection of coat complexes<sup>20–22</sup>. These two coats consist of an inner layer (the clathrin adaptors and the Sec23–Sec24 complex, respectively) that binds to cargo proteins and lipids, and an outer layer that forms an organizing cage (composed of clathrin and the Sec13–Sec31 complex, respectively) but does not directly contact cargo or lipids. Retromer has extensive structural similarities to these classical systems: VPS26 looks like the clathrin adaptor  $\beta$ -arrestin<sup>16</sup>, whereas VPS35 has an  $\alpha$ -solenoid fold similar to the AP trunk domains and to the clathrin heavy chain itself. On the other hand, retromer contrasts sharply in that the inner layer functions of the classical coats—binding to cargo and membrane lipids—are divided between the cargo recognition and SNX complexes, and are not shared by one inner complex. Neither the cargo recognition nor the SNX complex appears to serve as an outer coat, and it is unclear if an outer coat exists for this system. The apparent lack of an outer coat in this system is consistent with the absence of a classical electron-dense coat on vesicles of retrograde trafficking<sup>2</sup>. The putative single-layer ‘cage free’



**Figure 4 | Integration of cargo and targeting signals by the cargo-recognition complex.** **a**, The VPS26–VPS29–VPS35 complex is predicted to align roughly parallel to the membrane (green line at bottom), such that its multiple SNX<sup>4,15</sup> and cargo-binding sites<sup>25</sup> cooperatively interact. The arrows mark the central region about which VPS35 is proposed to flex so as to interact with cargo embedded in curved membranes. Binding sites that have been mapped to individual residues within crystallized components are coloured dark blue. Binding sites that have been mapped to regions of VPS35 or to as yet non-crystallized portions of VPS35 are marked by red bars aligned with the region of interest. Binding sites for yeast cargo proteins are not necessarily conserved in human VPS35; however, the overall architecture of the yeast and other orthologous complexes is proposed to be very similar to the human complex. **b**, Schematic rendering of a speculative model for the retromer coat on a tubular vesicle, coloured as above, with the SNX dimer in purple.

organization of the system seems to be ideally designed for the coating of tubular vesicles of variable length. The rough cylindrical symmetry and ill-defined endpoints present very different geometric constraints compared with the quasi-spherical vesicles of endocytosis and the secretory pathway, and thus demand different solutions.

## METHODS SUMMARY

The crystal structure of the VPS29–VPS35 C-terminal subcomplex was determined by single anomalous dispersion at the Se edge and refined to an *R*-free value of 0.278 at 2.8 Å resolution.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 2 July; accepted 4 September 2007.**

**Published online 23 September 2007.**

- Seaman, M. N. Recycle your receptors with retromer. *Trends Cell Biol.* **15**, 68–75 (2005).
- Bonifacino, J. S. & Rojas, R. Retrograde transport from endosomes to the trans-Golgi network. *Nature Rev. Mol. Cell Biol.* **7**, 568–579 (2006).
- Seaman, M. N. J., McCaffery, J. M. & Emr, S. D. A membrane coat complex essential for endosome-to-Golgi retrograde transport in yeast. *J. Cell Biol.* **142**, 665–681 (1998).
- Haft, C. R. *et al.* Human orthologs of yeast vacuolar protein sorting proteins Vps26, 29, and 35: Assembly into multimeric complexes. *Mol. Biol. Cell* **11**, 4105–4116 (2000).
- Arighi, C. N., Hartnell, L. M., Aguilar, R. C., Haft, C. R. & Bonifacino, J. S. Role of the mammalian retromer in sorting of the cation-independent mannose 6-phosphate receptor. *J. Cell Biol.* **165**, 123–133 (2004).
- Carlton, J. *et al.* Sorting nexin-1 mediates tubular endosome-to-TGN transport through coincidence sensing of high-curvature membranes and 3-phosphoinositides. *Curr. Biol.* **14**, 1791–1800 (2004).
- Seaman, M. N. Cargo-selective endosomal sorting for retrieval to the Golgi requires retromer. *J. Cell Biol.* **165**, 111–122 (2004).
- Verges, M. *et al.* The mammalian retromer regulates transcytosis of the polymeric immunoglobulin receptor. *Nature Cell Biol.* **6**, 763–769 (2004).
- Coudreuse, D. Y. M., Roel, G., Betist, M. C., Destree, O. & Korswagen, H. C. Wnt gradient formation requires retromer function in Wnt-producing cells. *Science* **312**, 921–924 (2006).
- Prasad, B. C. & Clark, S. G. Wnt signaling establishes anteroposterior neuronal polarity and requires retromer in *C. elegans*. *Development* **133**, 1757–1766 (2006).
- Strochlic, T. I., Setty, T. G., Sitaram, A. & Burd, C. G. Grd19/Snx3p functions as a cargo-specific adapter for retromer-dependent endocytic recycling. *J. Cell Biol.* **177**, 115–125 (2007).
- Small, S. A. *et al.* Model-guided microarray implicates the retromer complex in Alzheimer's disease. *Ann. Neurol.* **58**, 909–919 (2005).
- Rojas, R., Kametaka, S., Haft, C. R. & Bonifacino, J. S. Interchangeable but essential functions of SNX1 and SNX2 in the association of retromer with endosomes and the trafficking of the mannose 6-phosphate receptors. *Mol. Cell Biol.* **27**, 1112–1124 (2007).
- Wang, D. *et al.* Crystal structure of human vacuolar protein sorting protein 29 reveals a phosphodiesterase/nuclease-like fold and two protein-protein interaction sites. *J. Biol. Chem.* **280**, 22962–22967 (2005).
- Collins, B. M., Skinner, C. F., Watson, P. J., Seaman, M. N. & Owen, D. J. Vps29 has a phosphoesterase fold that acts as a protein interaction scaffold for retromer assembly. *Nature Struct. Mol. Biol.* **12**, 594–602 (2005).
- Shi, H., Rojas, R., Bonifacino, J. S. & Hurley, J. H. The retromer subunit Vps26 has an arrestin fold and binds Vps35 through its C-terminal domain. *Nature Struct. Mol. Biol.* **13**, 540–548 (2006).
- Damen, E., Krieger, E., Nielsen, J. E., Eygensteyn, J. & van Leeuwen, J. E. M. The human Vps29 retromer component is a metallo-phosphoesterase for a cation-independent mannose 6-phosphate receptor substrate peptide. *Biochem. J.* **398**, 399–409 (2006).
- Collins, B. M., McCoy, A. J., Kent, H. M., Evans, P. R. & Owen, D. J. Molecular architecture and functional model of the endocytic AP2 complex. *Cell* **109**, 523–535 (2002).
- Heldwein, E. E. *et al.* Crystal structure of the clathrin adaptor protein 1 core. *Proc. Natl Acad. Sci. USA* **101**, 14108–14113 (2004).
- Ybe, J. A. *et al.* Clathrin self-assembly is mediated by a tandemly repeated superhelix. *Nature* **399**, 371–375 (1999).
- Fotin, A. *et al.* Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* **432**, 573–579 (2004).
- Fath, S., Mancias, J. D., Bi, X. & Goldberg, J. Structure and organization of coat proteins in the COPII cage. *Cell* **129**, 1251–1252 (2007).
- Misra, S., Puertollano, R., Kato, Y., Bonifacino, J. S. & Hurley, J. H. Structural basis for acidic-cluster-dileucine sorting-signal recognition by VHS domains. *Nature* **415**, 933–937 (2002).
- Shiba, T. *et al.* Structural basis for recognition of acidic-cluster dileucine sequence by GGA1. *Nature* **415**, 937–941 (2002).
- Nothwehr, S. F., Bruinsma, P. & Strawn, L. A. Distinct domains within Vps35p mediate the retrieval of two different cargo proteins from the yeast prevacuolar/endosomal compartment. *Mol. Biol. Cell* **10**, 875–890 (1999).
- Yang, J. *et al.* Molecular basis for TPR domain-mediated regulation of protein phosphatase 5. *EMBO J.* **24**, 1–10 (2005).
- Groves, M. R. & Barford, D. Topological characteristics of helical repeat proteins. *Curr. Opin. Struct. Biol.* **9**, 383–389 (1999).
- Conti, E. & Izaurralde, E. Nucleocytoplasmic transport enters the atomic age. *Curr. Opin. Cell Biol.* **13**, 310–319 (2001).
- Nothwehr, S. F., Ha, S. A. & Bruinsma, P. Sorting of yeast membrane proteins into an endosome-to-Golgi pathway involves direct interaction of their cytosolic domains with Vps35p. *J. Cell Biol.* **151**, 297–310 (2000).
- Peter, B. J. *et al.* BAR domains as sensors of membrane curvature: The amphiphysin BAR structure. *Science* **303**, 495–499 (2004).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank H. Shi for contributing to the early stages of this project; X. Zhu and H. T. Tsai for technical assistance; B. Canagarajah for bioinformatics and computational crystallography support; E. Tyler for artwork; Z.-Q. Fu and all the staff of APS SER-CAT beamline 22 for assistance with data collection; and C. Haft for antibodies to retromer. Use of the APS was supported by the US DOE, Basic Energy Sciences, Office of Science. This project was funded by the Intramural Programs of NIDDK, NICHD and NIAMS, NIH.

**Author Contributions** A.L.R. and A.H. expressed and purified protein complexes, crystallized the VPS29–VPS35 C-terminal subcomplex, collected crystallographic data, and determined and refined the crystal structure; A.H. carried out phosphatase assays; R.R. and N.M. carried out immunoprecipitation and optical microscopy studies; G.E. and A.C.S. carried out and interpreted electron microscopy studies; A.V.K. carried out sequence analysis; and J.H.H., J.S.B. and A.C.S. designed the study. A.H. and A.L.R. contributed equally to this study.

**Author Information** Crystallographic coordinates have been deposited with the Protein Data Bank with accession number 2R17. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to J.H.H. ([hurley@helix.nih.gov](mailto:hurley@helix.nih.gov)).

## METHODS

**Protein expression and purification.** DNA encoding full-length human VPS29 and residues 476–780 of human VPS35 (C terminus) was amplified using the polymerase chain reaction (PCR). The products were inserted into pmr101A (modified pmr101; American Type Culture Collection) and pGEX-KG, respectively. VPS35 was fused to an N-terminal GST tag followed by a thrombin cleavage site, and VPS29 was untagged. The plasmids were co-transformed into *Escherichia coli* BL21 (DE3) star (Invitrogen) and protein was expressed at 15 °C for 12 h by induction with 1 mM isopropylthiogalactoside (IPTG) when cells reached an  $A_{600}$  of 0.8. Cells were lysed by high-pressure homogenization in 50 mM Tris (pH 8.0), 1 M NaCl and 10 mM dithiothreitol (DTT). The complex was isolated using glutathione-Sepharose TM 4B (Amersham Biosciences) affinity chromatography. The fusion protein was cleaved with thrombin while bound to the matrix in 50 mM Tris (pH 8.0), 50 mM NaCl and 10 mM DTT. The subcomplex was further purified by ion exchange (5 ml HiTrap Q FF) and gel filtration (Superdex 200 16/60) (GE Healthcare) chromatography. Fractions were pooled, concentrated to 2 mg ml<sup>-1</sup> in 50 mM Tris (pH 8.0), 1 M NaCl, 10 mM DTT, and stored at -80 °C. Full-length human VPS35, VPS26A and VPS29 cDNAs were amplified by PCR, and the products were subcloned into pGST-parallel2 vector (VPS35) and pmr101A vectors (VPS26A and VPS29), respectively. The VPS29 and VPS35 expression plasmids were co-transformed into BL21 (DE3) star whereas VPS26A was transformed separately. Expression was induced as above except with 0.3 mM IPTG. Cells from both lines were harvested together and lysed by high-pressure homogenization in 50 mM Tris (pH 8.0), 150 mM NaCl and 10 mM DTT. The full complex was then purified as described above. Human VPS29 was expressed and purified as above.

**VPS26–VPS35 fragment interactions.** For mapping VPS35–VPS26 interaction, the appropriate regions of human VPS35 were cloned into parallel GST2 vectors<sup>31</sup>, using a standard PCR-based cloning strategy, from cDNA encoding VPS35. Subsequently, VPS26A and GST–VPS35 constructs were amplified by PCR to generate cassettes containing the Shine–Delgarno translational start signal<sup>32</sup> and cloned directly into the polycistronic pST39 vector<sup>33</sup>. Proteins were expressed as for the intact complex and purified by glutathione-Sepharose affinity chromatography as described above.

**Crystallization and data collection.** Crystals of the retromer VPS29–VPS35 subcomplex were grown at 18 °C using hanging-drop vapour diffusion by mixing 2 µl of 16 mg ml<sup>-1</sup> protein solution (50 mM Tris pH 8.0, 1 M NaCl, 15 µM glycyl-glycyl-glycine) with 4 µl of 150 mM potassium sodium tartrate and 20% PEG 3350. Crystals were cryo-protected with 15% glycerol and flash frozen in liquid nitrogen. SAD experimental data at  $\lambda = 0.9793$  Å were collected at 95 K at APS SER-CAT Beamline 22-ID at the Se edge, and processed using HKL2000 (HKL Research).

**Structure determination and refinement.** Twelve of fourteen expected selenium sites were located and refined using SOLVE<sup>34</sup>, and initial phases were calculated at 3.1 Å. An interpretable electron density map was obtained after density modification and phase extension to 2.8 Å with RESOLVE. This map was used to place VPS29 in the subcomplex structure and the VPS35 C terminus was built by hand in COOT<sup>35</sup>. Subsequent rounds of model adjustment, simulated annealing and thermal factor refinement were performed with CNS<sup>36</sup>. During the final refinement stages, water molecules were inserted into the protein model. Statistics are summarized in Supplementary Table 1. The final atomic model is a dimer of the retromer subcomplex in the asymmetric unit, where chains A and B correspond to VPS29 and chains C and D correspond to VPS35 C terminus. This model includes two glycerol molecules and 53 water molecules. 88.4 % of the residues are in the most allowed region of the Ramachandran plot. Residues 41 and 115 of VPS29 are the only residues in the disallowed region of the Ramachandran plot, and are in the same conformations as in the structure of VPS29 previously determined in isolation. No electron density was observed for residues 679–694 or 734–738 in chain D. Figure 1d was coloured using the HotPatch server<sup>37</sup>.

**In vitro phosphatase assays.** Purified proteins were dialysed against 150 mM NaCl, 50 mM Tris/HCl pH 7.0 buffer. Each protein or protein complex was assayed at a final concentration of 17.5 µM in the presence of 1,500 µM of serine-phosphorylated peptide SFHDDpSDEDLLHI (>98% purity; New England Peptide LLC) and 50 µM ZnCl<sub>2</sub>. The volume for each reaction was adjusted to 100 µl with serine/threonine phosphatase assay buffer (50 mM Tris/HCl, pH 7.0, and 100 µM CaCl<sub>2</sub>) (Upstate Biotechnology). EDTA was added to duplicate reactions at a final concentration of 5 mM. As a positive control, 0.25 µM of shrimp alkaline phosphatase (Roche) was used under the same conditions. The reactions were incubated at 37 °C for 1 h or 16 h. From each sample, 30 µl was added to 60 µl of Biomol green (Biomol) and the amount of free phosphate was determined from the absorbance at 630 nm.

**Bioinformatics of metallophosphoesterases.** The alignment of 5,307 putative metallophosphoesterases was downloaded from the Pfam database<sup>38</sup> and residue

conservation at the designated positions was analysed with a locally written script.

**Recombinant DNA constructs for cell biology studies.** An *EcoRI*–*Bam*HI fragment (including a newly introduced start codon in position 1, and the replacement of a stop codon in position 183 for a tyrosine residue) encoding human VPS29 was obtained by PCR amplification from the VPS29–pGADT7 construct<sup>16</sup>. This fragment was later cloned into the corresponding sites of the pcDNA3.1+/Myc–His vector (Invitrogen). An *MfeI*–*NotI* fragment encoding VPS35 (2–796) was amplified by PCR from the VPS35–pGBKT7 construct and cloned into the respective sites of the pCI–HA<sub>3</sub> vector<sup>39</sup>. This construct was used as a template for the PCR amplification of the (HA)<sub>3</sub>–VPS35 fragment that was subsequently cloned into the pEF6/V5–His TOPO TA vector (Invitrogen). Mutagenesis of both the VPS29–Myc–pcDNA3.1 and (HA)<sub>3</sub>–VPS35–pEF6/V5 constructs was performed using the QuickChange site-directed mutagenesis kit (Stratagene).

**Immunoprecipitation.** HeLa cells (American Type Culture Collection) were cultured on 35-mm dishes as described previously<sup>13</sup>. Once the cells reached 80–90% confluency, they were transfected with 4 µg of wild-type or mutant (HA)<sub>3</sub>–VPS35–pEF6/V5 plasmids using Lipofectamine 2000 (Invitrogen). After 36 h, cells lysates were prepared and pre-cleared as previously described<sup>16</sup>. Pre-cleared lysates were then incubated for 2 h at 4 °C with 30 µl of protein–G-Sepharose beads bound to mouse monoclonal antibody to HA (clone HA. 11; Covance). Immunoprecipitates were washed and then subjected to SDS–PAGE and immunoblot analysis using antibodies to VPS29, VPS26 and HA, as described previously<sup>13,16</sup>. Immunoprecipitation of HeLa cells expressing either wild-type or mutant VPS29–Myc–pcDNA3.1 was performed as described above but with the difference that cell lysates were incubated with protein–G-Sepharose beads conjugated to mouse monoclonal antibody to Myc (9E10; Covance). Antibodies to VPS35 and Myc used for immunoblotting have been described previously<sup>13</sup>.

**Immunofluorescence labelling and microscopy.** HeLa cells grown on coverslips were transfected with wild-type and mutant VPS29–Myc–pcDNA3.1 or wild-type and mutant (HA)<sub>3</sub>–VPS35–pEF6/V5 constructs with Lipofectamine 2000 as indicated above. At 36 h after transfection, cells were washed, fixed and labelled with mouse monoclonal antibody to Myc (9E10), rat monoclonal antibody to HA (3F10) and rabbit polyclonal antibody to VPS26, and the respective fluorescently labelled secondary antibodies. Imaging was performed using an epifluorescence microscope (AX10; Carl Zeiss Microimaging, Inc) with a Plan–Apochromatic ×63 oil objective and the appropriate filter combination.

**Electron microscopy and image processing.** VPS35–VPS29–VPS26 retromer complexes (10 µg ml<sup>-1</sup>) were negatively stained with 1% uranyl formate and observed on a Philips CM120 microscope with LaB6 illumination. Micrographs, recorded at 60,000 magnification on Kodak SO-163 film, were digitized on a Nikon 9000 scanner and binned twofold, giving a pixel size of 2.1 Å. Particle picking, CTF correction by phase flipping, and other pre-processing operations, were carried out with Bsoft<sup>40</sup>. 4,331 particles, low-pass-filtered at 15 Å, were sorted by reference-free alignment as implemented in EMAN<sup>41</sup>, then classified in SPIDER<sup>42</sup> with multiple rounds of correspondence analysis, hierarchical clustering by the Ward criteria, and multi-reference alignment until stable classes were obtained. In the last iteration, thirteen classes, representing the major trends within the data set (3,815 particles), were individually averaged. Their resolutions, determined by Fourier ring correlation at 0.5 cutoff, were between 24 and 26 Å.

**Identification of VPS35 repeats.** We probed the N-terminal part of VPS35 (residues 1–480) for additional copies of repeats by applying a sensitive sequence profile method and pftools package<sup>43</sup>. First, we constructed the profile from structure-based alignments of the repeats from the C-terminal part of VPS35. The profile spanned three VPS35 repeats to increase the selectivity of the search. This is in line with the observations that a single repeat would be unlikely to form a stable structure on its own and that  $\alpha$ -helical solenoid proteins have tandem arrays of repeats. The best match produced by this profile (positions 459–592) had an *E*-value < 0.0001. The *E*-values of the matches were calculated as described<sup>44</sup>. Subsequent examination of this match found mostly hydrophobic residues in positions that should be apolar because the corresponding residues of known VPS35 structure have interior locations. The newly identified repeats were included in the initial structure-based alignment and the second profile was constructed. The profile identified three additional repeats in the N-terminal region (positions 272–413) with an *E*-value < 0.000001. For the remainder of VPS35, the repeat profiles produced a few weaker matches with *E*-values between 0.01 and 1. At the same time, several homologous VPS35 sequences were aligned and this allowed us to determine whether the residues conserved in the aligned repeats are also conserved between homologues. The alignment of VPS35 homologues also identified sites of insertions that probably correspond to loops in the

three-dimensional structures. This information was used to generate the alignment of the remaining repeats with the weak matches.

**Structural interpretation of electron microscopy.** A model of residues 1–482 of VPS35 as a right-handed  $\alpha$ -solenoid of 21 helices was constructed using the crystal structure of residues 483–780 as a template. The model was built manually, taking two copies of the 483–780 fragment containing a total of 26 helices as the starting point. The two fragments were manually aligned end to end with each other and with the VPS29–VPS35 subcomplex in a linear fashion as suggested by the processed electron microscopy images. Helical pairs at the N and C termini of each fragment were overlapped in order to maintain the pitch of the solenoid. The overlapping helices were then removed from the model, leaving a total of 21 modelled helices. Because the electron microscopy images offer no guidance for side-chain placement, no effort was made to assign the correct amino acid residues to the model. VPS26 was docked onto the N-terminal 165 residues of VPS35 such that its C-terminal lobe and in particular the loop from residues 235–246 (ref. 16) were in contact with each other. The curvature of the model was manually adjusted such that the overall cargo recognition complex model had a 210 Å end-to-end distance. The mean contour length of 0.26 Å per residue observed in the crystallized portion of VPS35 was maintained in the model.

For comparison with the electron microscopy class averages, the molecular model was band-limited to 25 Å resolution with Bsoft and two-dimensional projections were calculated, corresponding to a set of equally spaced side views. Finally, a CTF corresponding to the defocus condition of the micrographs was applied. The coordinates of the molecular model were then modified manually as follows to improve the agreement with the electron microscopy data. No modifications were made to the crystallographically determined structures of VPS35–C–VPS29 and VPS26. The relative position of the VPS35–VPS29 subcomplex versus VPS26 interacting with the N-terminal 165 residues of VPS35 was progressively adjusted by small increments and the modelled portion of VPS35 shifted to maintain continuity. Then, a new set of two-dimensional model projections was generated and compared with the electron microscopy data, both visually and computationally (by calculating correlation coefficients). Finally, the placement of VPS26 on the N-terminal region of VPS35 was also modified manually in a way that kept the total length of the model at 210 Å (the value consistently observed in the electron microscopy analysis) and maintained contact between VPS35–N and the VPS26 loop from residues 235–246, and evaluated in the same way. After these adjustments, disconnected loops were manually

reconnected in O<sup>45</sup>, and the structure was energy-minimized with CNS<sup>36</sup> to remove gaps and steric clashes. The consistency between the final model and the electron microscopy class average was estimated by computing the average Fourier ring cross-correlation curve as a function of resolution, giving a 29 Å limit for a 0.3 cutoff (Supplementary Fig. 8).

31. Sheffield, P., Garrard, S. & Derewenda, Z. Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expr. Purif.* **15**, 34–39 (1999).
32. Hierro, A., Kim, J. & Hurley, J. H. Polycistronic expression and purification of the ESCRT-II endosomal trafficking complex. *Methods Enzymol.* **403**, 322–332 (2005).
33. Tan, S. A modular polycistronic expression system for overexpressing protein complexes in *Escherichia coli*. *Protein Expr. Purif.* **21**, 224–234 (2001).
34. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55**, 849–861 (1999).
35. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
36. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
37. Pettit, F. K., Bare, E., Tsai, A. & Bowie, J. U. HotPatch: a statistical approach to finding biologically relevant features on protein surfaces. *J. Mol. Biol.* **369**, 863–879 (2007).
38. Bateman, A. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **30**, 276–280 (2002).
39. Martina, J. A., Moriyama, K. & Bonifacino, J. S. BLOC-3, a protein complex containing the Hermansky-Pudlak syndrome gene products HPS1 and HPS4. *J. Biol. Chem.* **278**, 29376–29384 (2003).
40. Heymann, J. B. & Belnap, D. M. Bsoft: image processing and molecular modeling for electron microscopy. *J. Struct. Biol.* **157**, 3–18 (2007).
41. Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97 (1999).
42. Frank, J. *et al.* SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**, 190–199 (1996).
43. Bucher, P., Karplus, K., Moeri, N. & Hofmann, K. A flexible motif search technique based on generalized profiles. *Comput. Chem.* **20**, 3–23 (1996).
44. Hofmann, K. & Bucher, P. The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors. *Trends Biochem. Sci.* **20**, 347–349 (1995).
45. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).