nature cell biology

ARTICLES

https://doi.org/10.1038/s41556-019-0393-3

Molecular identification of a BAR domain-containing coat complex for endosomal recycling of transmembrane proteins

Boris Simonetti[1,6], Blessy Paul[2,6], Karina Chaudhari[3], Saroja Weeratunga[2], Florian Steinberg[4], Madhavi Gorla[3], Kate J. Heesom[5], Greg J. Bashaw[3], Brett M. Collins[2,7*] and Peter J. Cullen[1,7*]

Protein trafficking requires coat complexes that couple recognition of sorting motifs in transmembrane cargoes with biogenesis of transport carriers. The mechanisms of cargo transport through the endosomal network are poorly understood. Here, we identify a sorting motif for endosomal recycling of cargoes, including the cation-independent mannose-6-phosphate receptor and semaphorin 4C, by the membrane tubulating BAR domain-containing sorting nexins SNX5 and SNX6. Crystal structures establish that this motif folds into a β-hairpin, which binds a site in the SNX5/SNX6 phox homology domains. Over sixty cargoes share this motif and require SNX5/SNX6 for their recycling. These include cargoes involved in neuronal migration and a *Drosophila snx6* mutant displays defects in axonal guidance. These studies identify a sorting motif and provide molecular insight into an evolutionary conserved coat complex, the 'Endosomal SNX–BAR sorting complex for promoting exit 1' (ESCPE-1), which couples sorting motif recognition to the BAR-domain-mediated biogenesis of cargo-enriched tubulo-vesicular transport carriers.

Thousands of transmembrane cargo proteins routinely enter the endosomal network where they transit between two fates: retention within the network for degradation in the lysosome or export from the network for recycling and reuse at the cell surface, the *trans*-Golgi network (TGN) and the autophagosome[1,2]. Although the mechanistic details of degradative sorting are well established, the events that define the export pathways remain poorly described[2]. The identification of sorting motifs in the cytosolic tails of transmembrane proteins that provide the signature for export as well as the membrane remodelling complexes that recognise these signatures and drive packaging into transport carriers is central to the understanding of endosomal recycling[3].

Although the sorting motifs for endocytic and Golgi-directed transport are well known, those that control export from the endosomal network are poorly understood[2–4]. The characterization of AP-1 and the CCC (CCDC22, CCDC93, COMMD) complex[5–7], Retromer and Retriever, and their associated sorting nexin (SNX) cargo adaptors[8–16] has identified some export motifs for the sorting of hundreds of integral proteins. However, additional sorting motifs are likely to be required and how their recognition is coordinated with membrane remodelling to form cargo-enriched transport carriers remains largely unknown for the majority of the identified sorting motifs[17,18].

Here, we define the role of a SNX containing a carboxy-terminal Bin, Amphiphysin, Rvs (BAR) domain (SNX–BAR) membrane-remodelling complex in the sequence-dependent sorting of integral proteins[19–24]. Through structural and functional analyses, we identify a sorting motif that is required for endosomal recycling through direct interaction with these SNX–BARs. Our study provides insight into an endosomal coat complex that couples sequence-dependent cargo recognition with the BAR domain-mediated biogenesis of tubulo-vesicular transport carriers.

Results

Structure of SNX5 bound to the CI-MPR. We recently observed a role for SNX5, SNX6 and SNX32—a subset of SNX–BAR proteins—in the endosome sorting of the cation-independent mannose-6-phosphate receptor (CI-MPR) to the TGN[23,24]. To provide molecular insight into these observations (Fig. 1a), we mapped the region of SNX5 necessary for CI-MPR binding. As noted by others[25], protein alignment revealed a 38-amino-acid insert in the PX domains of these SNX–BAR proteins, which is absent in the PX domain of SNX1 and SNX2—functionally related SNX–BAR proteins that form a heterodimeric complex with SNX5 (ref. [20]) but fail to associate with the CI-MPR (Fig. 1a,b). CI-MPR binding could indeed be transferred from SNX5 to SNX1 by engineering a SNX1–SNX5 switch chimera (Fig. 1c).

A green fluorescent protein (GFP)-nanotrap screen of deletion mutants across the CI-MPR cytosolic tail (covering residues 2328–2491) established that residues 2347–2375 were essential for the SNX5 and SNX6 association (Fig. 1d). A peptide corresponding to these residues bound directly to the isolated SNX5 PX domain (K_d of approximately 25 μM; Fig. 1e) and the corresponding domains of SNX6 and SNX32 (K_d of 24 μM and 12 μM, respectively; Fig. 1e and Supplementary Table 1).

Next, we designed a fusion protein constituting the SNX5 PX domain coupled through a flexible linker to the CI-MPR sequence, crystallized the protein in two different crystal forms and determined

---

[1]School of Biochemistry, University of Bristol, Bristol, UK. [2]Institute for Molecular Bioscience, The University of Queensland, St. Lucia, Queensland, Australia. [3]Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [4]Center for Biological Systems Analysis (ZBSA), Faculty of Biology, Albert Ludwigs Universitaet Freiburg, Freiburg, Germany. [5]Proteomic Facility, School of Biochemistry, University of Bristol, Bristol, UK. [6]The authors contributed equally: Boris Simonetti, Blessy Paul. [7]These authors jointly supervised this work: Brett M. Collins, Peter J. Cullen. *e-mail: b.collins@imb.uq.edu.au; pete.cullen@bristol.ac.uk

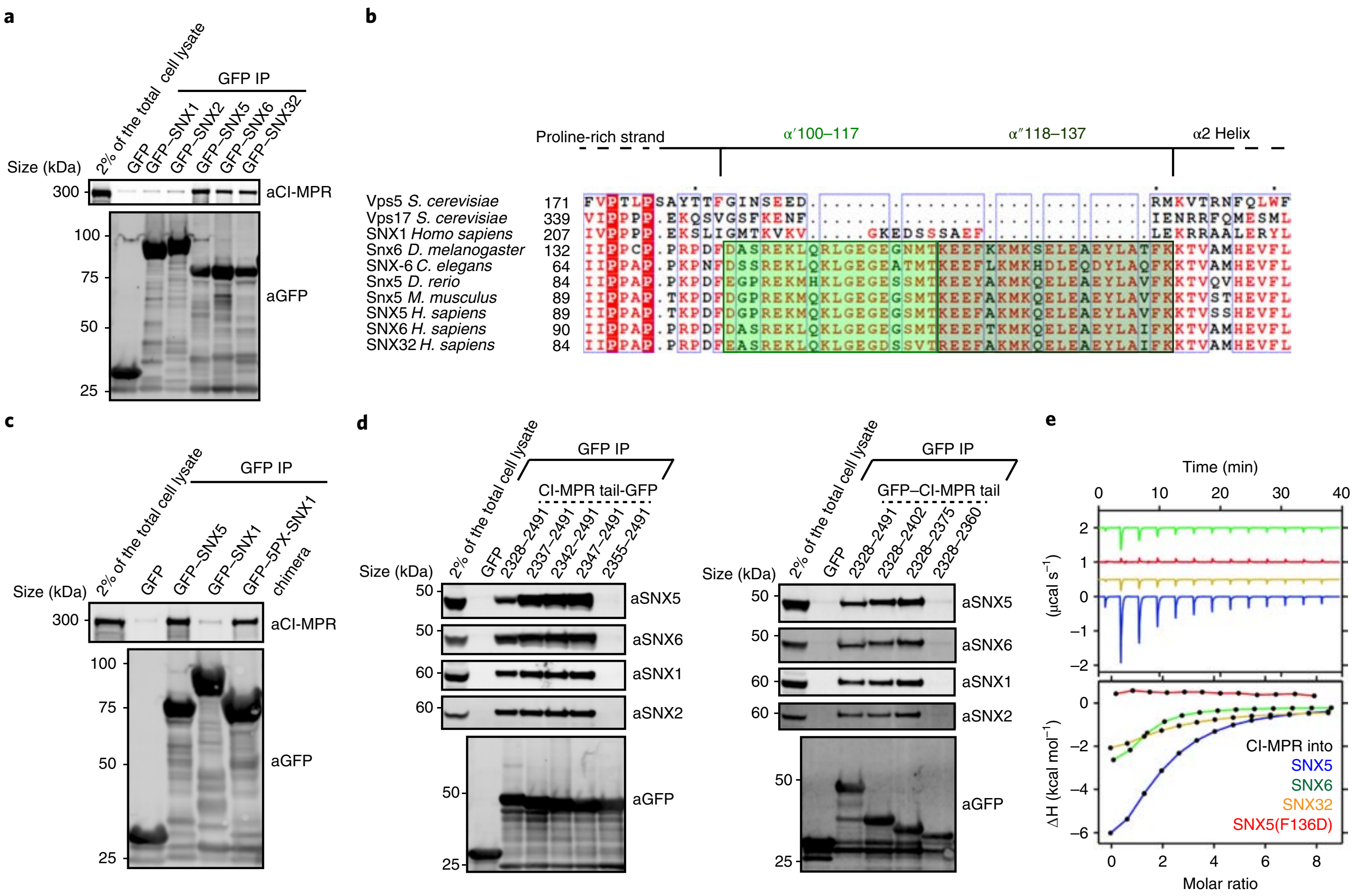


Fig. 1 | Mapping the interaction between SNX5 and CI-MPR. a, Co-immunoprecipitation of GFP-tagged SNX–BARs transiently transfected in HEK293T cells. **b**, Alignment of yeast Vps17 and Vps5 with SNX1, SNX5, SNX6, SNX32 and SNX5 homologues from different species. The α′ and α″ helices that compose the unique helix-turn-helix extension are highlighted in green. **c**, Co-immunoprecipitation of GFP-tagged SNX1 and SNX5, and a SNX1 chimera generated by the replacement of the SNX1 PX domain with that of SNX5. **d**, Co-immunoprecipitation of GFP-tagged CI-MPR tail-deletion mutants transiently transfected in HEK293T cells. **e**, The CI-MPR peptide was injected into SNX5, SNX5(F136D), SNX6 or SNX32 PX domains and binding was measured by ITC. The raw data (top) and the integrated and normalized data fit with a 1:1 binding model (bottom) are shown. The binding of CI-MPR with SNX5 and SNX6 was measured over $n=3$ independent experiments; the binding of CI-MPR with SNX32 was measured once. The ITC binding parameters, including s.d. (where calculated), are provided in Supplementary Table 1. The blots in **a**,**c**,**d** are representative of three independent GFP traps. The unprocessed original scans of the immunoblots are shown in Supplementary Fig. 3. IP, immunoprecipitation.

their structures by X-ray crystallography (Fig. 2a,b and Supplementary Table 2). The CI-MPR peptide interacted with the SNX5 PX domain via a β-sheet augmentation, forming two antiparallel β-strands (βA and βB) connected by a long flexible linker. Although the sequences and specific side-chain interactions are different, the bound CI-MPR structures resemble that of the IncE protein from *Chlamydia trachomatis*[26–28] (Supplementary Fig. 1a,b).

In both structures, the first β-strand consists of the residues ^{2349}VSYKYSK^{2355} (Fig. 2a,b). The key side chains of SNX5 that mediate binding include Arg103, Met106, Tyr132 and Phe136. The Val2349, Tyr2351 and Tyr2353 side chains in the βA-strand are in close contact with SNX5, with Tyr2351 forming a stacking interaction with SNX5 Phe136 and Val2349 forming a hydrogen bond with SNX5 Arg103. This mirrors a similar stacking interaction of Phe116 from IncE (Supplementary Fig. 1a). Arg103 from SNX5 forms a hydrogen bond with the backbone carbonyl of Val2349, causing it to adopt a substantially different orientation to its pose in the IncE-bound SNX5 structure. The Y2351N and Y2351D mutations completely disrupted the interaction, thus confirming the importance of the ^{2349}VSYKYSK^{2355} residues (Fig. 2c).

Interestingly, the second βB-strand included a WLM sequence previously identified as necessary for CI-MPR endosome-to-TGN sorting[29]. There was a major difference in the register of this second βB-strand in the two crystal forms of SNX5–CI-MPR (Fig. 2a,b), which we confirmed by anomalous difference maps of the SeMet-labelled proteins to unambiguously place the CI-MPR Met2371 side chain (Fig. 2d). The βB strand in Form 1 includes the residues ^{2369}WLMEEI^{2374}, whereas in Form 2 it consists of ^{2368}EWLMEE^{2373}. Therefore, Leu2370 in Form 1 engages a hydrophobic pocket composed of the CI-MPR side chains Tyr2351 and Tyr2353, and the SNX5 side chains Tyr132, Leu133 and Phe136; in Form 2, the flip in the βB strand orientation results in CI-MPR Met2371 occupying this pocket (Fig. 2a,b). Which of these peptide configurations is preferred? We theorized that mutation of either Leu2370 or Met2371 to an aspartate side chain would block the interaction with SNX5, depending on which residue was most favoured. Neither mutation had a major effect on the binding affinity—determined by isothermal titration calorimetry (ITC) experiments—although there was a reduced enthalpy of interaction, whereas mutation of the entire WLM sequence resulted in a larger reduction in binding affinity (Fig. 2e and Supplementary Table 1). This suggests that the two configurations of the βB strand are both equally able to sustain binding; what is important is not the sequence per se but that a correctly positioned hydrophobic side chain is the main requirement (either

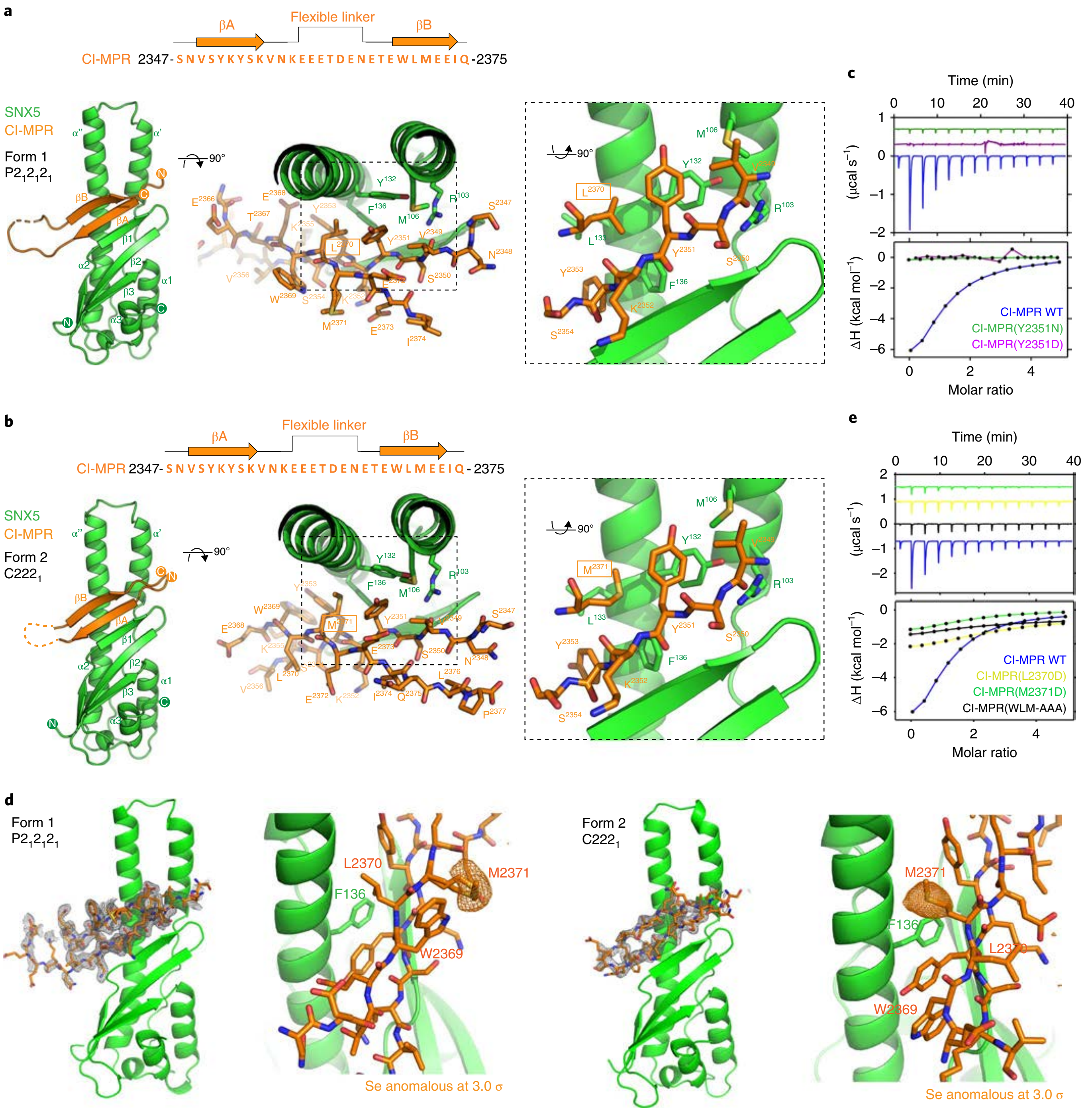


Fig. 2 | Crystal structure of the SNX5–CI-MPR complex. **a**,**b**, The crystal structures of SNX5–CI-MPR Form 1 (spacegroup $P2_12_12_1$; **a**) and Form 2 (spacegroup $C222_1$; **b**) are shown as ribbon diagrams (left), with a close-up of the bound CI-MPR peptide shown as a stick representation (right). **c**, Binding of SNX5 to the CI-MPR peptide and CI-MPR βA-strand mutants Y2351N and Y2351D. The raw data (top) and the integrated and normalized data fit with a 1:1 binding model (bottom) are shown. **d**, The crystal structures of SNX5–CI-MPR Form 1 and Form 2 are shown with the corresponding CI-MPR peptide Fo–Fc omit maps (grey; 3 σ level; left); the anomalous difference maps for the Se atoms in the SeMet-labelled protein (orange; 3 σ level) are shown for both forms (right). This identifies the Met2371 side chain unambiguously in each structure. **e**, Binding of SNX5 to the CI-MPR βB-strand mutants L2370D, M2371D and WLM-AAA, as measured by ITC. The raw data (top) and the integrated and normalized data fits with a 1:1 binding model (bottom) are shown. Binding of SNX5 to the CI-MPR wild type (WT) was measured over $n = 3$ independent experiments; the binding of SNX5 with the CI-MPR βA/βB-strand mutants was measured once. The ITC binding parameters, including the s.d. (where calculated), are provided in Supplementary Table 1.

Leu2370 or Met2371 in CI-MPR). Notably, this site in the SNX5–IncE complex is occupied by a valine side chain (Supplementary Fig. 1a). Therefore, the model is that SNX5 (and SNX6) can be engaged by a 'promiscuous' β-hairpin structure, where essential features include aromatic residues in the third and fifth positions of the βA strand, a loop region and a hydrophobic side chain in the third position of the βB strand.

The CI-MPR sorting motif is required for retrograde transport. Validation of the ^{2349}VSYKYSK^{2355} and ^{2369}WLM^{2371} sequences as

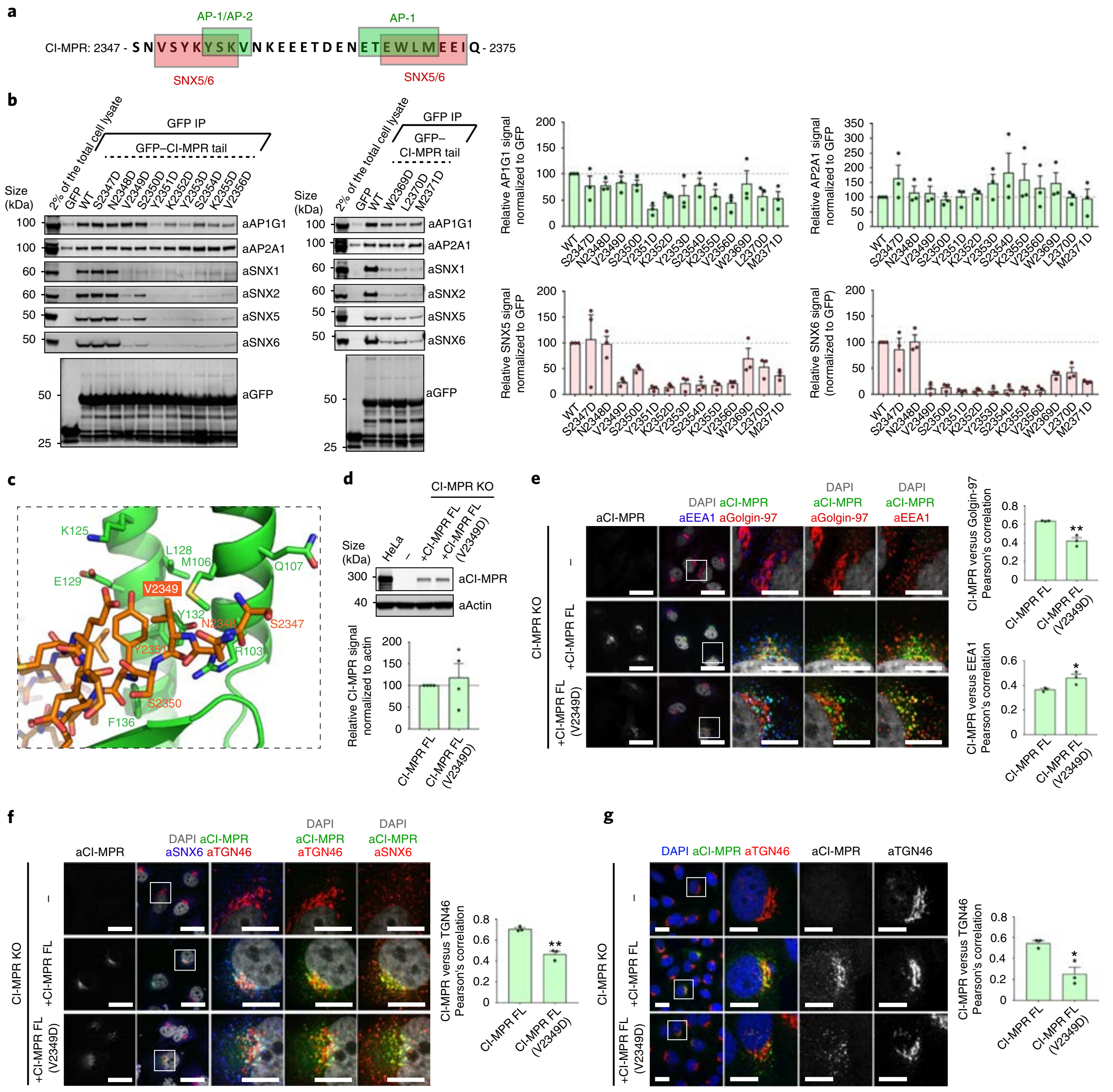


Fig. 3 | The interaction between SNX5 and the CI-MPR β-hairpin structure is required for CI-MPR retrograde trafficking. a, Schematic of the interactions reported to overlap with the CI-MPR β-hairpin structure. **b**, Co-immunoprecipitation of the GFP-tagged CI-MPR-tail constructs transiently transfected in HEK293T cells (left). Summary of the relative levels of binding to the indicated proteins (right). The band intensities were measured from $n = 3$ independent experiments using Odyssey software. The band intensities, normalized to GFP expression, are presented as the average fraction of the GFP–CI-MPR WT control. **c**, Bound CI-MPR peptide as a stick representation highlighting the position of Val2349. **d**, Re-expression of full-length CI-MPR and the CI-MPR(V2349D) mutant in a HeLa CI-MPR knockout (KO) clonal line. CI-MPR levels were analysed by quantitative fluorescence-based western blotting. The band intensities were measured using Odyssey software and normalized to β-actin before calculating the percentage of protein compared with the full-length CI-MPR control. The bars represent the mean of $n = 4$ independent experiments. **e**,**f**, HeLa CI-MPR KO clonal line transiently transfected with full-length CI-MPR or the CI-MPR(V2349D) mutant (left). Magnified views of the white boxes are shown in the images to their right. The steady-state colocalization of CI-MPR with EEA1, Golgin-97 and TGN46 was analysed using Pearson's correlation (right). The Pearson's coefficient values were compared using a two-tailed unpaired *t*-test. **e**, Cell numbers analysed for CI-MPR versus Golgin-97 colocalization: 64 CI-MPR KO + CI-MPR and 56 CI-MPR KO + CI-MPR(V2349D) cells across $n = 3$ independent experiments; $P = 0.0036$. Cell numbers analysed for the Pearson's correlation of CI-MPR versus EEA1: 69 CI-MPR KO + CI-MPR and 49 CI-MPR KO + CI-MPR(V2349D) cells across $n = 3$ independent experiments; $P = 0.0385$. **f**, Cell numbers analysed for CI-MPR versus TGN46 colocalization: 70 CI-MPR KO + CI-MPR and 66 CI-MPR KO + CI-MPR(V2349D) cells across $n = 3$ independent experiments; $P = 0.0019$. **g**, The HeLa CI-MPR KO clonal line was transfected with full-length CI-MPR or the CI-MPR(V2349) mutant and CI-MPR colocalization with TGN46 analysed after 40 min chase of surface CI-MPR. Magnified views of the white squares are shown in the images to their right. Cell numbers analysed: 45 CI-MPR KO + CI-MPR and 53 CI-MPR KO + CI-MPR(V2349D) across $n = 3$ independent experiments. The Pearson's coefficient values were compared using an unpaired *t*-test (right); $P = 0.0133$. Scale bars, 20 μm (micrographs) and 10 μm (magnified images). Bars, error bars and circles represent the mean, s.e.m. and individual data points, respectively. The unprocessed original scans of the immunoblots are shown in Supplementary Fig. 3. The statistics source data are in Supplementary Table 4. FL, full-length; $^*P < 0.05$, $^{**}P < 0.01$.

sorting motifs is complicated by their overlap with sequences recognized by other cargo adaptors, most notably AP-1 and AP-2, which mediate TGN-to-endosome transport and CI-MPR endocytosis, respectively[30] (Fig. 3a). We performed aspartate mutagenesis across the ^{2349}VSYKYSK^{2355} and ^{2369}WLM^{2371} sequences to identify a mutant that negated SNX5/SNX6 binding while retaining AP-1/AP-2 association (Fig. 3b). Quantitative western blot analysis revealed that the V2349D mutant conveyed a loss of SNX5/SNX6 binding without a discernible effect on AP-2/AP-1 binding (Fig. 3b). Mutations of the βB-strand caused a minor loss of SNX5/SNX6 binding, thus confirming that the ^{2349}VSYKYSK^{2355} motif constitutes the major contributor to the interaction (Fig. 3b).

To establish the ^{2349}VSYKYSK^{2355} motif as the dominant element of the sorting signal, we generated a HeLa cell line knocked out for CI-MPR. Next, we incorporated a Val2349Asp mutation (Fig. 3c) into an expression construct encoding full-length CI-MPR (Fig. 3d). In contrast to the expression of full-length wild-type CI-MPR, which displayed a steady-state enrichment at the TGN46- and Golgin-97-labelled TGN, full-length CI-MPR(V2349D) failed to enrich in the TGN and displayed a more prominent endosomal distribution (Fig. 3e,f). This is consistent with the V2349D mutant retaining the ability to undergo endocytosis and transport from the TGN despite having a defect in retrograde transport. Confirming this, antibody uptake experiments established that both wild-type CI-MPR and the V2349D mutant underwent endocytosis, but only the V2349D mutant failed to undergo endosomal export, as defined by enrichment in the TGN (Fig. 3g). Thus, the ^{2349}VSYKYSK^{2355} sorting motif, alongside the ^{2369}WLM^{2371} motif[29], is required for the SNX5/SNX6-dependent endosome-to-TGN transport of the CI-MPR.

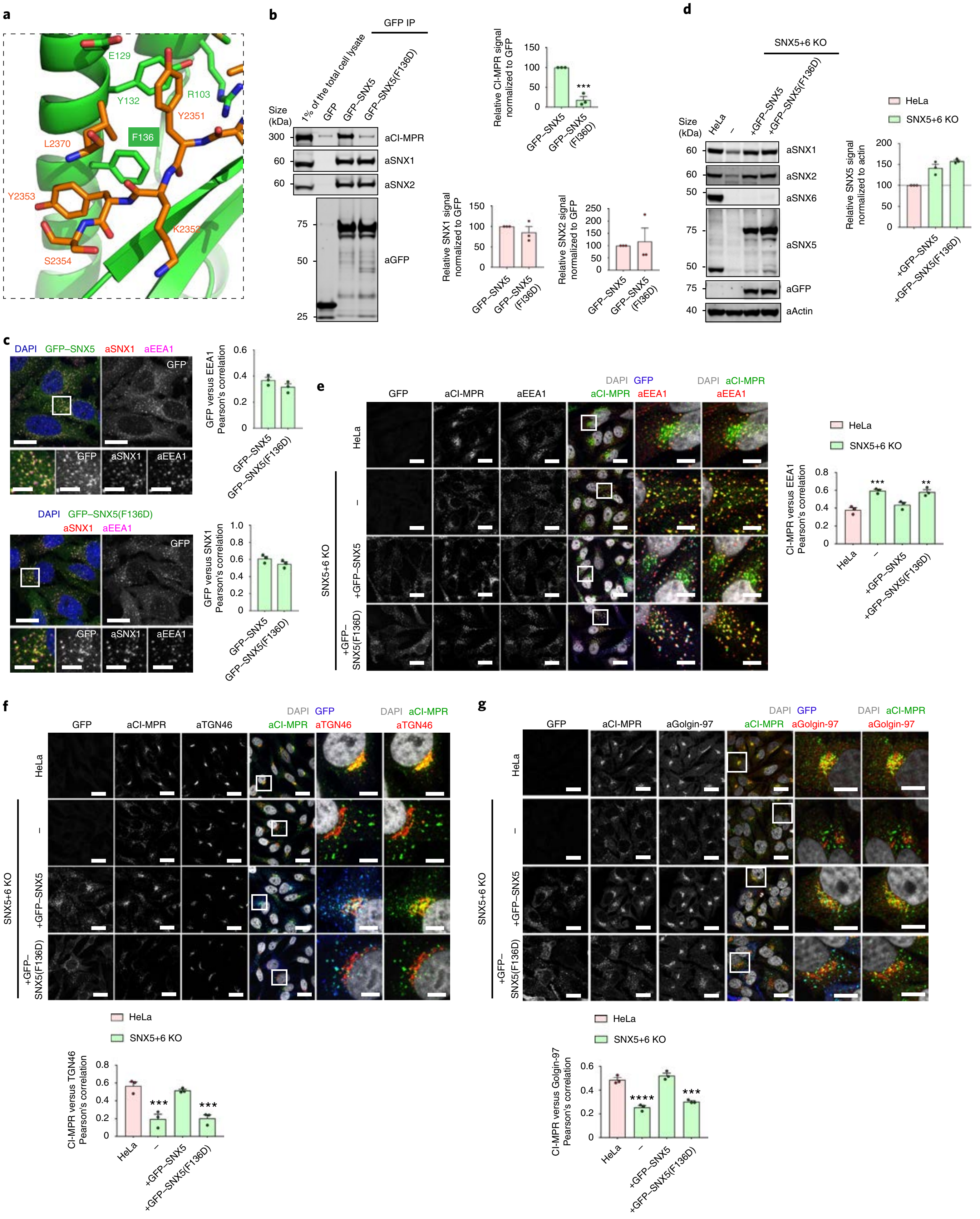
SNX5 recognition of the CI-MPR sorting motif is required for endosomal export. We next examined the interaction of GFP-tagged SNX5 and an aspartate mutant of the key Phe136 residue (F136D; Fig. 4a) with full-length CI-MPR. Although SNX5 displayed CI-MPR binding, the GFP–SNX5(F136D) mutant failed to interact in co-immunoprecipitation experiments (Fig. 4b). Importantly, the GFP–SNX5(F136D) mutant retained endosomal association and formed functional BAR-domain-mediated heterodimers with SNX1 and SNX2[20,21] (Fig. 4b,c). The loss of SNX5 and SNX6 expression in a knockout cell line results in the steady-state localization of the CI-MPR shifting from the TGN to peripherally dispersed endosomes[23,24]. In this background, transduction of wild-type GFP–SNX5 rescued CI-MPR sorting, whereas GFP–SNX5(F136D) expression failed to do so (Fig. 4d–g). Overall, by identifying a sorting motif and revealing its mode of recognition by the SNX5/SNX6 components of the SNX–BAR membrane tubulating complex, these data establish the mechanistic basis of tubular-based endosome-to-TGN export of the CI-MPR.

CI-MPR retrograde trafficking requires multiple features of SNX5-containing heterodimers. SNX5 (SNX6 and SNX32) form functional heterodimers with SNX1 (and SNX2) through their BAR domains[20,22]. The resulting SNX1–SNX5 heterodimers are associated with the cytosolic leaflet of endosomes through multiple features of these organelles, including specific phosphoinositides and the sensing/induction of membrane curvature. To explore the relationship between membrane association and CI-MPR sorting, we performed rescue experiments with previously validated mutants that perturb aspects of the membrane binding activity of the SNX1–SNX5 heterodimer: SNX5(S226E) displays a reduced ability to form heterodimers with SNX1 (and SNX2)[31], SNX1(K214A) prevents association with phosphoinositides[19] (the SNX5 PX domain does not bind phosphoinositides[32]), the BAR-domain-targeted SNX1(K429E:K430E:R431E) triple mutant (SNX1(KKR–EEE)) is unable to sense or induce membrane curvature[19] and the SNX1(M287E:F288E) double mutant (SNX1(ΔAH))[22] has a defective amphipathic helix required for membrane insertion (Fig. 5a).

Co-immunoprecipitations confirmed that SNX5(S226E) displayed a reduced interaction with SNX1 (and SNX2) but, importantly, had little-to-no effect on the CI-MPR association (Fig. 5b). All of the SNX1 mutants retained SNX5 heterodimer formation and CI-MPR association (Fig. 5c). When expressed in the SNX5+SNX6 (SNX5+6) knockout HeLa cell line, SNX5(S226E) displayed a

Fig. 4 | Interaction between CI-MPR and the hydrophobic groove of the SNX5 PX domain is required for CI-MPR retrograde trafficking. a, CI-MPR peptide bound to the SNX5 hydrophobic groove highlighting Phe136 position. **b**, Co-immunoprecipitation of GFP-tagged SNX5 and SNX5(F136D) transiently transfected in HEK293T cells (left) and a summary of the relative binding levels (right). The band intensities were quantified from $n=3$ independent experiments using Odyssey software. The band intensities, normalized to GFP expression, are presented as the average fraction of the GFP–SNX5 control. The GFP–SNX5(F136D) mutant was compared with GFP–SNX5 using a two-tailed unpaired *t*-test; $P=0.3655$ (SNX1), 0.7646 (SNX2) and 0.0006 (CI-MPR). **c**, The F136D mutation in the SNX5 PX domain does not affect SNX5 endosomal localization. HeLa cells were transduced with GFP–SNX5 or GFP–SNX5(F136D) lentivirus particles. The transduced cells were fixed and stained for EEA1 and SNX1 (left). Magnified views of the white boxes are shown below. Cell numbers analysed for colocalization between GFP and EEA1: 50 GFP–SNX5 and 57 GFP–SNX5(F136D) cells across $n=3$ independent experiments. The Pearson's coefficient values were compared using a two-tailed unpaired *t*-test (right); $P=0.1862$. Cell numbers analysed for colocalization between GFP and SNX1: 51 GFP–SNX5 and 59 GFP–SNX5(F136D) cells across $n=3$ independent experiments. The Pearson's coefficient values were compared using a two-tailed unpaired *t*-test; $P=0.1771$. Scale bars, 20 µm (micrographs) and 10 µm (insets). **d**, GFP–SNX5 and GFP–SNX5(F136D) was reintroduced in SNX5 + SNX6 (SNX5+6) KO cells at levels comparable to endogenous. The HeLa SNX5+6 KO clonal line was transduced with GFP–SNX5 or GFP–SNX5(F136D) lentivirus particles and the SNX5 levels were analysed by quantitative fluorescence-based western blotting. The band intensities were measured using Odyssey software and normalized to β-actin before calculating the percentage of protein compared with the parental HeLa control. The bars represent the mean of $n=3$ independent experiments. **e–g**, Distribution of endogenous CI-MPR in HeLa cells, the HeLa SNX5+6 KO clonal line and the HeLa SNX5+6 KO clonal line transduced with GFP–SNX5 or GFP–SNX5(F136D). Magnified views of the white boxes are shown in the images to their right. Colocalization analysis between CI-MPR and EEA1, CI-MPR and TGN46, and CI-MPR and Golgin-97. The Pearson's coefficient values were compared with the HeLa control using a one-way analysis of variance (ANOVA) and Dunnett's test. **e**, Cell numbers analysed across $n=3$ independent experiments for CI-MPR and EEA1 colocalization: 129 HeLa, 144 SNX5 + 6 KO, 159 SNX5+6 KO + GFP–SNX5 and 122 SNX5+6 KO + GFP–SNX5(F136D) cells; $P=0.0007$ (SNX5+6 KO), 0.2886 (SNX5+6 KO + GFP–SNX5) and 0.0010 (SNX5+6 KO + GFP–SNX5(F136D)). **f**, Cell numbers analysed across $n=3$ independent experiments for CI-MPR and TGN46 colocalization: 126 HeLa, 111 SNX5+6 KO, 121 SNX5+6 KO + GFP–SNX5 and 110 SNX5+6 KO + GFP–SNX5(F136D) cells; $P=0.0005$ (SNX5+6 KO), 0.6897 (SNX5+6 KO + GFP–SNX5) and 0.0006 (SNX5+6 KO + GFP–SNX5(F136D)). **g**, Cell numbers analysed across $n=3$ independent experiments for CI-MPR and Golgin-97 colocalization: 105 HeLa, 127 SNX5+6 KO, 122 SNX5+6 KO + GFP–SNX5 and 131 SNX5+6 KO + GFP–SNX5(F136D) cells; $P=0.3482$ (SNX5+6 KO + GFP–SNX5), 0.0002 (SNX5+6 KO + GFP–SNX5(F136D)) and ≤ 0.0001 (SNX5+6 KO). Scale bars, 20 µm (micrographs) and 5 µm (magnified images). The bars, error bars and circles represent the mean, s.e.m. and individual data points, respectively. The unprocessed original scans of the immunoblots are shown in Supplementary Fig. 3. The statistics source data are in Supplementary Table 4. $^{**}P<0.01$, $^{***}P<0.001$, $^{****}P<0.0001$.

cytosolic localization and failed to rescue the CI-MPR mis-sorting phenotype (Fig. 5d). Furthermore, in rescue experiments performed in a double SNX1 and SNX2 knockout HeLa cell line[23,24], all of the individual SNX1 mutants localized to the cytosol and failed to rescue the defect in CI-MPR retrograde transport (Fig. 5e). Consistent with these data, all SNX1 mutants failed to drive the enrichment of



SNX6 to endosomes (an antibody to detect endogenous SNX5 is lacking at present; Fig. 5f). The SNX5 (SNX6)-dependent CI-MPR retrograde transport therefore requires the formation of heterodimers with SNX1 (or SNX2) and association of the resulting SNX1–SNX5 complex to the endosomal membrane through sensing of this organelle. With the SNX5(F136D) mutant retaining heterodimer formation and maintaining a localization to endosomes (Fig. 4c), cargo engagement does not seem to be the major driver for endosomal association of the SNX1–SNX5 complex. We speculate that cargo recognition may aid cargo clustering and the SNX1–SNX5 complex as a pre-requisite for the biogenesis of tubular profiles and transport carriers (Fig. 5g).

SEMA4C and IGF1R bind to SNX5 through a similar mechanism. Do other cargo proteins utilize the same molecular mechanism for SNX5/SNX6-mediated endosomal sorting? In addition to CI-MPR, IGF1R (a receptor critical in malignant transformation) and SEMA4C (a plexin B2 receptor involved in axon guidance) are two cell surface transmembrane proteins that interact with SNX5[24]. Co-immunoprecipitations of GFP–SNX5 and GFP–SNX5(F136D) established that SNX5 binds to these proteins through the conserved CI-MPR binding site within the PX domain (Fig. 6a). Truncation mutagenesis identified residues 731–755 of SEMA4C and residues 1275–1285 of IGF1R in SNX5 binding (Fig. 6b,c). Using the corresponding peptides, binding of SEMA4C and IGF1R to the PX domain of SNX5 was direct and with micromolar affinities (K_d of 26 and 15 µM, respectively), which is comparable to those observed for CI-MPR (Fig. 6d,e and Supplementary Table 1).

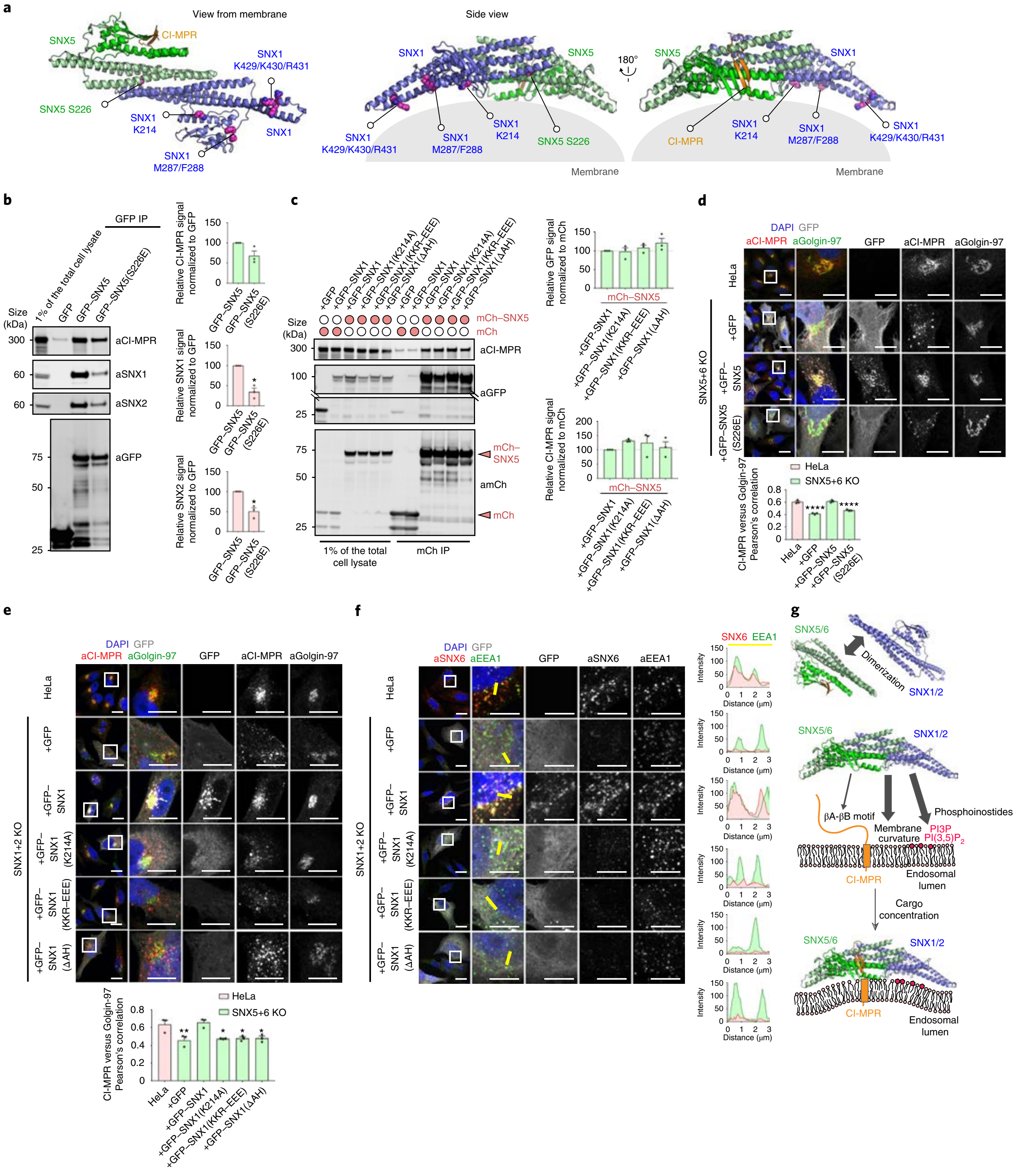
Although we were unable to obtain diffracting crystals for the SNX5–IGF1R complex, crystals of the SNX5 PX domain with the SEMA4C sequence were grown and the structure was determined at a resolution of 2.45 Å (Fig. 6f and Supplementary Table 2). SEMA4C interacts with SNX5 via a similar anti-parallel βA- and βB-sheet augmentation, with the ^{734}VGYYYS^{739} peptide making the first β-strand, followed by a tight turn (residues ^{740}DGS^{742}) and then the Leu743 side chain engaging the hydrophobic pocket where Leu2370 or Met2371 are bound in the CI-MPR structures (Fig. 6f). Immunoprecipitations of GFP-tagged SEMA4C cytosolic-tail mutants, probing for endogenous SNX5 association, confirmed that the ^{734}VGYYYS^{739} sequence was required for the SNX5–SEMA4C association (Fig. 6g).

In examining the sequence of the IGF1R peptide (residues 1275–1285), we noted a ^{1277}VSFYYS^{1282} sequence very similar to the βA strand of SEMA4C and CI-MPR (^{734}VGYYYS^{739} and ^{2349}VSYKYS^{2355}, respectively). Immunoprecipitation using the wild-type GFP-tagged IGF1R tail and mutants targeting the ^{1277}VSFYYS^{1282} sequence confirmed that this is required for the SNX5 interaction (Fig. 6h).

In contrast to the role of SNX5/SNX6 in the endosome-to-TGN transport of the CI-MPR, the role of these SNX–BARs in the sorting of the IGF1R is distinct[24]. Following ligand activation at the cell surface, the IGF1R undergoes internalization and enters the endosomal network for its SNX5/SNX6-dependent recycling to the cell surface. Thus, the internalized IGF1R in SNX5 and SNX6 double-knockout cells fails to engage the recycling itinerary and instead enters the lysosomal degradative pathway[24]. To extend this analysis into its functional role during endosome-to-plasma membrane recycling, we introduced the F1279D mutant into full-length IGF1R and confirmed that this displayed a clear loss of endogenous SNX5 binding (Fig. 7a). We introduced wild-type IGF1R or IGF1R(F1279D) into an IGF1R-knockout HeLa cell line. Importantly, the mutation did not affect receptor function, as defined by the activation of ERK–MAPK following IGF-1 stimulation (Fig. 7b). Consistent with the ^{1277}VSFYYS^{1282} sequence being a sorting motif, long-term IGF-1 stimulation of IGF1R(F1279D) resulted in mis-sorting into the lysosomal degradative pathway (Fig. 7c). In addition, IGF1R was degraded in the SNX5 and SNX6 double-knockout cells, and re-expression of SNX5, but not of the SNX5(F136D) mutant, rescued the IGF1R degradative fate and promoted cell surface recycling (Fig. 7d,e). These data establish that SNX5 (as well as SNX6/SNX32) engage the cytosolic tail of cargoes through the recognition of a ‘promiscuous’ β-hairpin structure.

Fig. 5 | CI-MPR retrograde trafficking requires functional SNX5 heterodimers and co-incidence detection of multiple membrane features. a, Model of the SNX5–SNX1 heterodimer. **b**, Co-immunoprecipitation of GFP-tagged SNX5 and SNX5(S226E) transiently transfected in HEK293T cells (left) and a summary of the relative binding levels (right). The band intensities were quantified from $n = 3$ independent experiments using Odyssey software. The band intensities, normalized to GFP expression, are presented as the average fraction of the GFP–SNX5 control. The GFP–SNX5(S226E) mutant was compared with GFP–SNX5 using a two-tailed unpaired t-test; $P = 0.0193$ (SNX1), 0.0324 (SNX2) and 0.1160 (CI-MPR). **c**, Co-immunoprecipitation of mCherry (mCh) and mCh–SNX5 transiently transfected in HEK293T cells alongside GFP, GFP–SNX1, GFP–SNX1(K214A), GFP–SNX1(KKR-EEE) and GFP–SNX1(ΔAH) (left) and a summary of the relative binding levels. The band intensities of GFP and CI-MPR were quantified from $n = 3$ independent experiments using Odyssey software. The band intensities, normalized to mCh expression, are presented as the average fraction of the mCh–SNX5 + GFP–SNX1 control. The interactions of the mCh-SNX5 + GFP-SNX1 mutants were compared with mCh-SNX5 + GFP-SNX1 WT using a one-way ANOVA and Dunnett’s test; GFP, $P = 0.9969$ (SNX5 + GFP-SNX1(K214A)), 0.8473 (SNX5 + GFP-SNX1(KKR-EEE)) and 0.2983 (SNX5 + GFP-SNX1(ΔAH)); CI-MPR, $P = 0.4109$ (SNX5 + GFP-SNX1(K214A)), 0.6042 (SNX5 + GFP-SNX1(KKR-EEE)) and 0.9648 (SNX5 + GFP-SNX1(ΔAH)). **d**, Distribution of endogenous CI-MPR in HeLa cells and the HeLa SNX5+6 KO clonal line transfected with GFP, GFP–SNX5 or GFP–SNX5(S226E) (top). Colocalization analysis between CI-MPR and the TGN marker Golgin-97 (bottom). Cell numbers analysed: 55 HeLa, 62 SNX5+6 KO + GFP, 63 SNX5+6 KO + GFP–SNX5 and 68 SNX5+6 KO + GFP–SNX5(S226E) cells across $n = 3$ independent experiments; $P \leq 0.0001$ (SNX5+6 KO + GFP and SNX5+6 KO + GFP–SNX5(S226E)) and $P = 0.9173$ (SNX5+6 KO + GFP–SNX5). **e**, Distribution of endogenous CI-MPR in HeLa cells, the HeLa SNX1 + SNX2 (SNX1+2) KO clonal line and the HeLa SNX1+2 KO clonal line transfected with GFP, GFP–SNX1, GFP–SNX1(K214A), GFP–SNX1(KKR-EEE) and GFP–SNX1(ΔAH) (top). Colocalization analysis between CI-MPR or Golgin-97 (bottom). Cell numbers analysed: 60 HeLa, 53 SNX1+2 KO + GFP, 58 SNX1+2 KO + GFP–SNX1, 72 SNX1+2 KO + GFP–SNX1(K214A), 77 SNX1+2 KO + GFP–SNX1(KKR-EEE) and 73 SNX1+2 KO + GFP–SNX1(ΔAH) cells across $n = 3$ independent experiments; $P = 0.0055$ (SNX1+SNX2 KO + GFP), 0.9752 (SNX1+2 KO + GFP–SNX1), 0.0106 (SNX1+2 KO + GFP–SNX1(K214A)), 0.0126 (SNX1+2 KO + GFP–SNX1(KKR-EEE)) and 0.0132 (SNX1+2 KO + GFP–SNX1(ΔAH)). **f**, Distribution of endogenous SNX6 in HeLa cells, the HeLa SNX1+2 KO clonal line and the HeLa SNX1+2 KO clonal line transfected with GFP, GFP–SNX1, GFP–SNX1(K214A), GFP–SNX1(KKR-EEE) or GFP–SNX1(ΔAH) (left). SNX6 colocalization with EEA1 (right). The figure is representative of $n = 3$ independent experiments with similar results. **d–f**, Magnified views of the white boxes are shown in the images to their right. Micrograph scale bars, 20 µm; magnified image scale bars, 10 µm. The Pearson’s coefficient values in **d**,**e** were compared with the HeLa control using a one-way ANOVA and Dunnett’s test. **g**, Model for how SNX5/6–SNX1/2 heterodimers sense multiple features of endosomal membranes, including the presence of specific phosphoinositides, local membrane curvature and cytosolic tails of cargoes. By co-incident sensing, these feature SNX5/6–SNX1/2 heterodimers assemble into functional membrane-associated complexes that couple cargo recognition with membrane remodelling for the formation of cargo-enrich transport carriers. The bars, error bars and circles represent the mean, s.e.m. and individual data points, respectively. The statistics source data are in Supplementary Table 4. The unprocessed original scans of the immunoblots are shown in Supplementary Fig. 3. $^*P < 0.5$, $^{**}P < 0.01$, $^{****}P < 0.0001$.

Numerous cargoes contain a ΦxΩxΦ(x)$_n$Φ motif necessary for SNX5-mediated endosomal sorting. We next considered whether other cargoes utilize the same mechanism of binding to SNX5/SNX6 to mediate their endosomal sorting. To analyse this, we designed a comparative proteomic screen, which involved the transient transfection of HEK293T cells with GFP, GFP–SNX5 or GFP–SNX5(F136D), followed by GFP-nanotrap immunoprecipitation. The interacting proteins were identified by liquid chromatography with tandem mass spectrometry (LC–MS/MS) and enrichment quantified through isobaric tandem mass tagging (TMT; Supplementary Fig. 2a). This identified 48 proteins, several of which were transmembrane proteins, whose loss was significant in



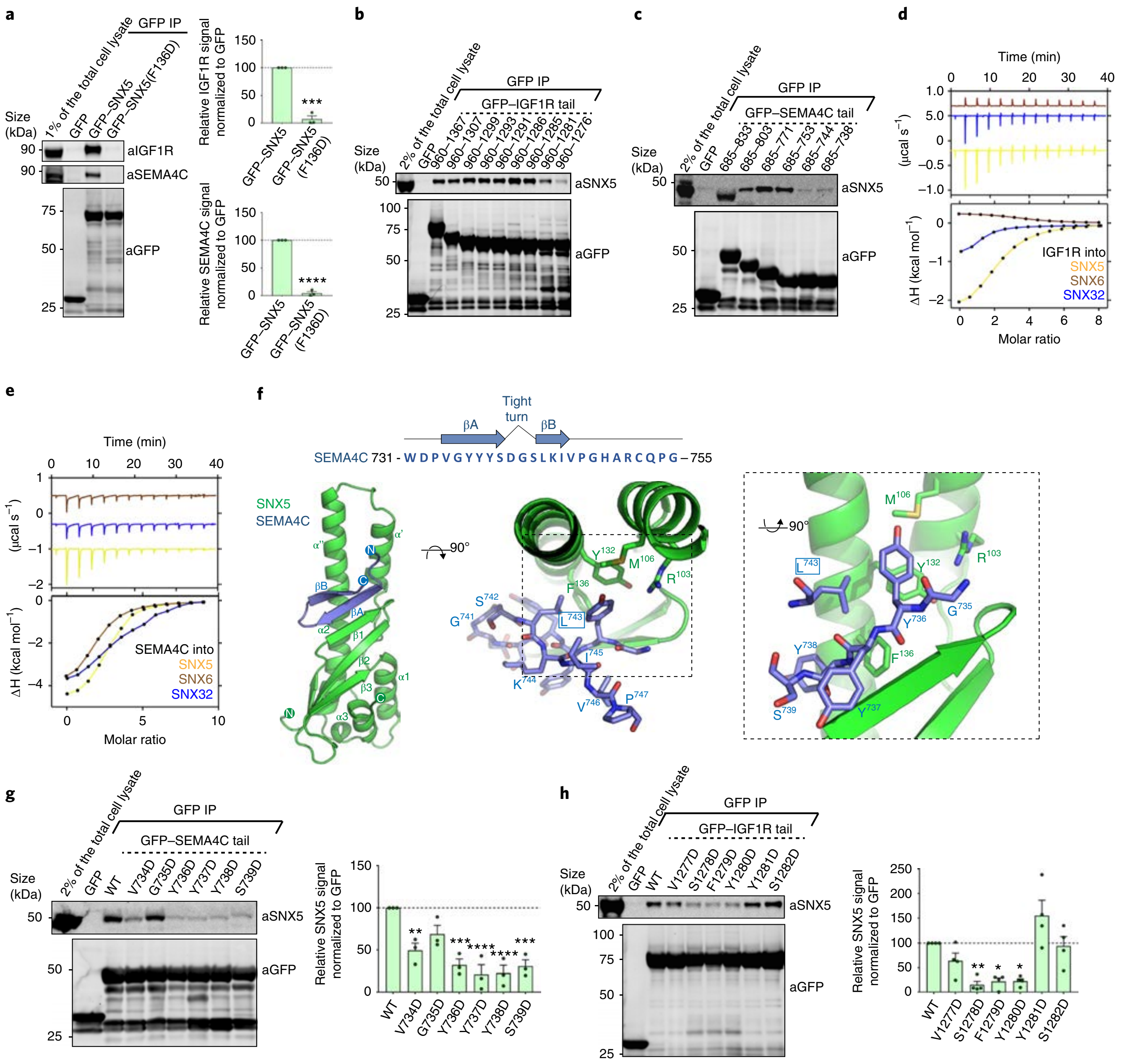


Fig. 6 | Mechanism of SEMA4C and IGF1R cargo binding to SNX5. a, Co-immunoprecipitation of GFP-tagged SNX5 and SNX5(F136D) transiently transfected in HEK293T cells (left) and a summary of the relative binding (right). The band intensities were quantified from $n = 3$ independent experiments using Odyssey software. The band intensities, normalized to GFP expression, are presented as the average fraction of the GFP–SNX5 control. The GFP–SNX5(F136D) mutant was compared with GFP–SNX5 using a two-tailed unpaired *t*-test; $P = 0.0001$ (IGF1R) and ≤ 0.0001 (SEMA4C). **b**, Co-immunoprecipitation of GFP-tagged IGF1R-tail-truncation mutants transiently transfected in HEK293T cells. The blot is representative of three independent GFP traps. **c**, Co-immunoprecipitation of GFP-tagged SEMA4C-tail-truncation mutants transiently transfected in HEK293T cells. The blot is representative of three independent GFP traps. **d**,**e**, Binding of the SNX5, SNX6 and SNX32 PX domains to the IGF1R peptide (**d**) and SEMA4C peptide (**e**), as measured by ITC. The raw data (top) and integrated and normalized data fits with a 1:1 binding model are shown (bottom). Binding of IGF1R and SEMA4C with SNX5 measured over $n = 3$ independent experiments; the binding of IGF1R and SEMA4C with SNX6 or SNX32 was measured once. The ITC binding parameters, including s.d. (where calculated) are provided in Supplementary Table 1. **f**, SNX5–SEMA4C crystal structure shown as a ribbon diagram (left), with bound SEMA4C peptide shown as a stick representation (middle and right). **g**, Co-immunoprecipitation of GFP-tagged SEMA4C-tail constructs transiently transfected in HEK293T cells (left). The band intensities were quantified from $n = 3$ independent experiments using Odyssey software (right). The band intensities, normalized to GFP expression, are presented as the average fraction of the GFP–SEMA4C wild-type control. The GFP–SEMA4C mutants were compared with the GFP–SEMA4C wild-type control using a one-way ANOVA and Dunnett's test; $P = 0.0045$ (V734), 0.0929 (G735D), 0.0003 (Y736D and S739D) and ≤ 0.0001 (Y737D and Y738D). **h**, Co-immunoprecipitation of GFP-tagged IGF1R-tail constructs transiently transfected in HEK293T cells (left). The band intensities were quantified from $n = 4$ independent experiments using Odyssey software (right). The band intensities, normalized to GFP expression, are presented as the average fraction of the GFP–IGF1R wild-type control. The GFP–IGF1R mutants were compared with the GFP–IGF1R wild-type control using a one-way ANOVA and Dunnett's test; $P = 0.4031$ (V1277D), 0.0049 (S1278D), 0.0101 (F1279D), 0.0107 (Y1280D), 0.0941 (Y1281D) and 0.9996 (S1282D). The bars, error bars and circles represent the mean, s.e.m. and individual data points, respectively. The unprocessed original scans of the immunoblots are shown in Supplementary Fig. 3. The statistics source data are in Supplementary Table 4. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

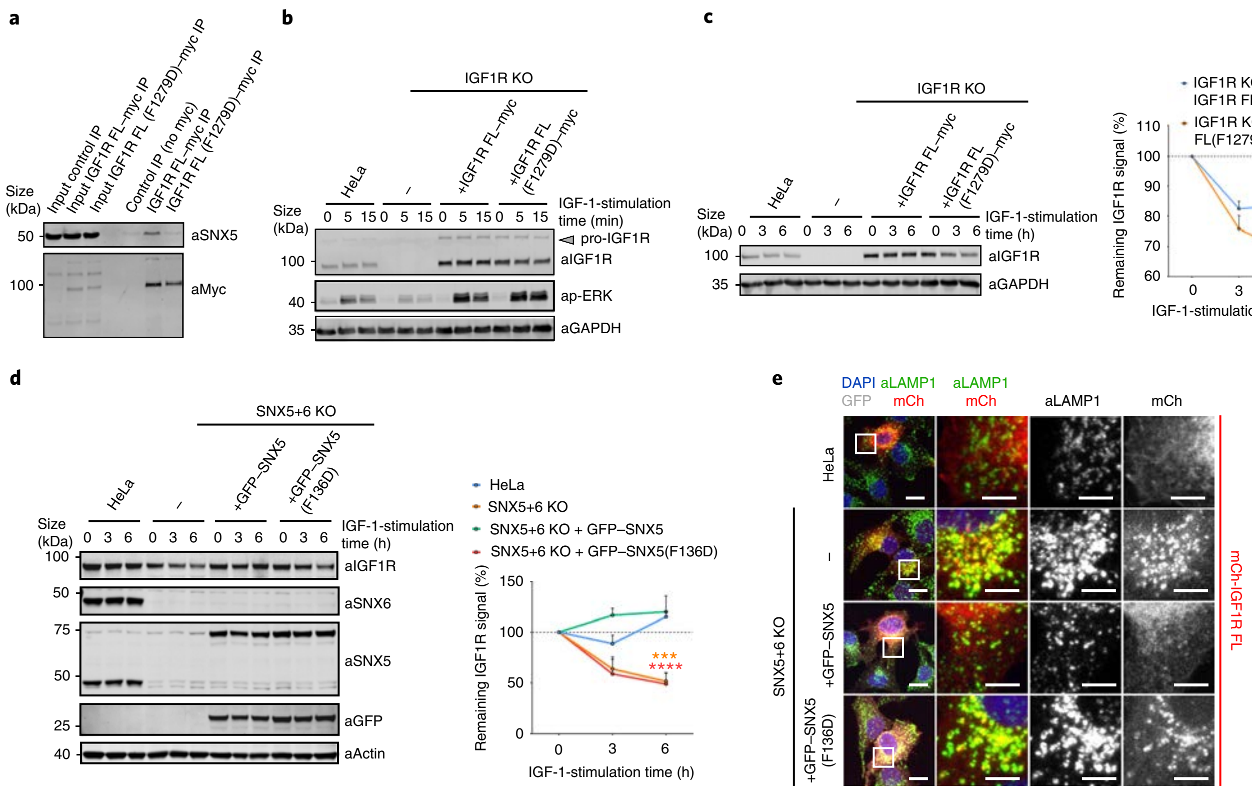


Fig. 7 | Interaction between the IGF1R cytosolic tail and the SNX5 PX-domain hydrophobic groove is required for the endosome-to-plasma membrane recycling of the receptor. **a**, Co-immunoprecipitation of myc-tagged full-length IGF1R and IGF1R(F1279D) in HEK293T cells. Blot is representative blot of $n = 3$ independent myc immunoprecipitation experiments. **b**, HeLa cells and the HeLa IGF1R KO cell line with or without transduction with full-length IGF1R or IGF1R(F1279D) were serum starved and treated with 10 nM IGF-1 for the indicated periods. The levels of IGF1R and p-ERK were analysed by immunoblotting. Blot is representative of $n = 3$ independent experiments. **c**, HeLa cells and the IGF1R KO cell line with or without transfection with full-length IGF1R or IGF1R(F1279D) were serum starved and treated with 10 nM IGF-1 for the indicated periods. The levels of IGF1R were analysed by quantitative fluorescence-based western blotting (left). The band intensities of IGF1R were quantified from $n = 3$ independent experiments using Odyssey software (right). The band intensities, normalized to GAPDH expression, are presented as the average fraction of the IGF1R signal relative to time-point 0. The IGF1R(F1279D) levels were compared with the IGF1R levels using a two-way ANOVA and Sidak test; $P = 0.4742$ (3 h) and 0.0219 (6 h). **d**, HeLa SNX5+6 KO clonal lines with or without transduction with GFP–SNX5 or GFP–SNX5(F136D) were serum starved and treated with 10 nM IGF-1 for the indicated periods. The levels of endogenous IGF1R were analysed by quantitative fluorescence-based western blotting. The band intensities of endogenous IGF1R were quantified from $n = 3$ independent experiments using Odyssey software. The band intensities, normalized to the β-actin expression, are expressed as a relative fraction of the IGF1R signal to time-point 0. The levels of IGF1R in the different cell lines were compared with the levels of IGF1R in HeLa controls using a two-way ANOVA and Dunnett's test; $P = 0.1418$ (SNX5+6 KO), 0.0919 (SNX5+6 KO + GFP-SNX5) and 0.0663 (SNX5+6 KO + GFP-SNX5(F136D)) at 3 h; $P = 0.0001$ (SNX5+6 KO), 0.9622 (SNX5+6 KO + GFP-SNX5) and ≤0.0001 (SNX5+6 KO + GFP-SNX5(F136D)) at 6 h. **e**, Distribution of mCh–IGF1R in the HeLa SNX5+6 KO clonal line with or without transduction with GFP–SNX5 or GFP–SNX5(F136D). The cells were serum starved and treated with 10 nM IGF-1 for 2 h before fixation and immunostaining. Magnified views of the white squares are shown in the images to their right. Micrograph scale bars, 20 μm; magnified image scale bars, 10 μm. The circles in **c**,**d** represent the mean and the error bars the s.e.m. Images are representative of $n = 3$ independent experiments. The unprocessed original scans of the immunoblots are shown in Supplementary Fig. 3. The statistics source data are in Supplementary Table 4. $^{*}P < 0.05$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

the SNX5(F136D) interactome compared with the wild-type SNX5 interactome (Fig. 8a and Supplementary Fig. 2a,b). We selected the following transmembrane proteins for further validation: ROBO1, a receptor in axon guidance and neuronal precursor cell migration[33]; TMEM230, a protein linked to Parkinson's disease[34] and GPR50, a regulator of melatonin and TGFβ signalling[35]. Immunoprecipitation of GFP-tagged SNX5 and the F136D mutant validated the Phe136-dependent association to SNX5 (Fig. 8b). Moreover, alignment of their cytosolic tails with CI-MPR, SEMA4C and IGF1R revealed the presence of the ^{1309}HTYGY^{1313}, ^{40}IPYKA^{44} and ^{317}IFFSG^{321} sequences in ROBO1, TMEM230 and GPR50, respectively, which conform with a general ΦxΩxΦ consensus motif for the βA strand (Φ and Ω, hydrophobic and aromatic side chains, respectively; Fig. 8c).

A limitation of the interactome-based analysis is the reliance on sufficient interacting protein being retained through the immunoprecipitation methodology for identification and quantification by LC–MS/MS, key determinants of which are the abundance of the interacting protein and its SNX5 binding affinity. To expand the cargoes and identify low-abundance and/or low-affinity cargoes, we designed a functionally orientated proteomic analysis to quantify how the altered cell surface proteome in SNX5+6 knockout cells was rescued by wild-type SNX5 or SNX5(F136D). SNX5+6 knockout cells and knockout cells transduced with wild-type SNX5 or SNX5(F136D) were subjected to surface biotinylation and TMT labelling to quantify the proteomic data (Supplementary Fig. 2c). Of the transmembrane proteins that displayed decreased levels in the SNX5+6 knockout cells compared with the cells rescued with SNX5, 61 were significantly depleted from the plasma membrane of cells expressing SNX5(F136D) compared with the wild-type (Fig. 8d and Supplementary Fig. 2d). This is consistent with these proteins requiring the SNX5 hydrophobic groove for their endosomal escape and recycling from the lysosomal degradative fate. A Panther gene-ontology analysis revealed enrichment for proteins involved in 'axonal guidance', 'neuronal morphogenesis' (for example, L1CAM, ROBO1, SEMA4C, PLXNA3 and CELSR2) and 'cell migration' (for example, ITGA6 and ADAM17; Fig. 8e), and, importantly, sequence analysis revealed that 20 proteins contain a putative ΦxΩxΦ motif in their cytosolic tails (Fig. 8f).

In vivo studies establish a role for *Drosophila snx-6* in axonal guidance. Finally, we considered the evolutionary conserved helix-turn-helix extension in the PX domain (Fig. 1b) and the aforementioned gene-ontology analysis by examining axon guidance in the *Drosophila* embryonic central nervous system. We initially tested whether loss-of-function mutations in either *snx1* or *snx6* result in defects in axon guidance at the midline by examining late-stage embryos with markers to label all axons (horseradish peroxidase) or subsets of ipsilaterally projecting interneurons (Fasciclin II, FasII; *Drosophila* encode a single SNX5/SNX6 ortholog, *snx6*). Neither *snx1* nor *snx6* zygotic mutants resulted in dramatic defects in this coarse-level analysis; however, Snx1 and Snx6 are maternally deposited (https://flybase.org/), and this could potentially prevent the detection of phenotypes[36]. We turned to sensitized genetic backgrounds to analyse the roles for the Snx1 and Snx6 proteins more carefully. First, we analysed reducing *snx1* and *snx6* genes in embryos heterozygous for mutations in *slit*—the major midline repellent—and its neuronal receptor *robo*[37] (Fig. 8g,i). In this background, a small percentage of FasII-positive axons abnormally cross the midline and mutations in other genes involved in midline guidance can modify this phenotype to reveal roles in promoting or inhibiting midline crossing[38]. Removing one copy of either *snx1* or *snx6* significantly attenuated the midline-crossing defects in this background, with a simultaneous reduction of both *snx1* and *snx6* leading to a greater reduction in ectopic crossing (Fig. 8g,i). Next, we made use of a second genetic background where normal midline crossing of commissural axons is partially disrupted. In this case, expression of a truncated Frazzled receptor (FraΔC) in a subset of commissural axons resulted in the failure of a small percentage of axons to cross the midline. Similar to the *slit*, *robo* background, mutations in other genes involved in midline guidance can modify this phenotype to reveal roles in either promoting or inhibiting midline crossing[39]. Midline crossing defects were not significantly altered for embryos in this background, which are heterozygous for either *snx1* or *snx6*, whereas a simultaneous reduction of both *snx1* and *snx6* led to a striking enhancement in the percentage of axons that failed to cross the midline (Fig. 8h,j). Our preliminary genetic analysis is suggestive of a role of these *snx* genes in promoting the normal axonal growth across the midline.

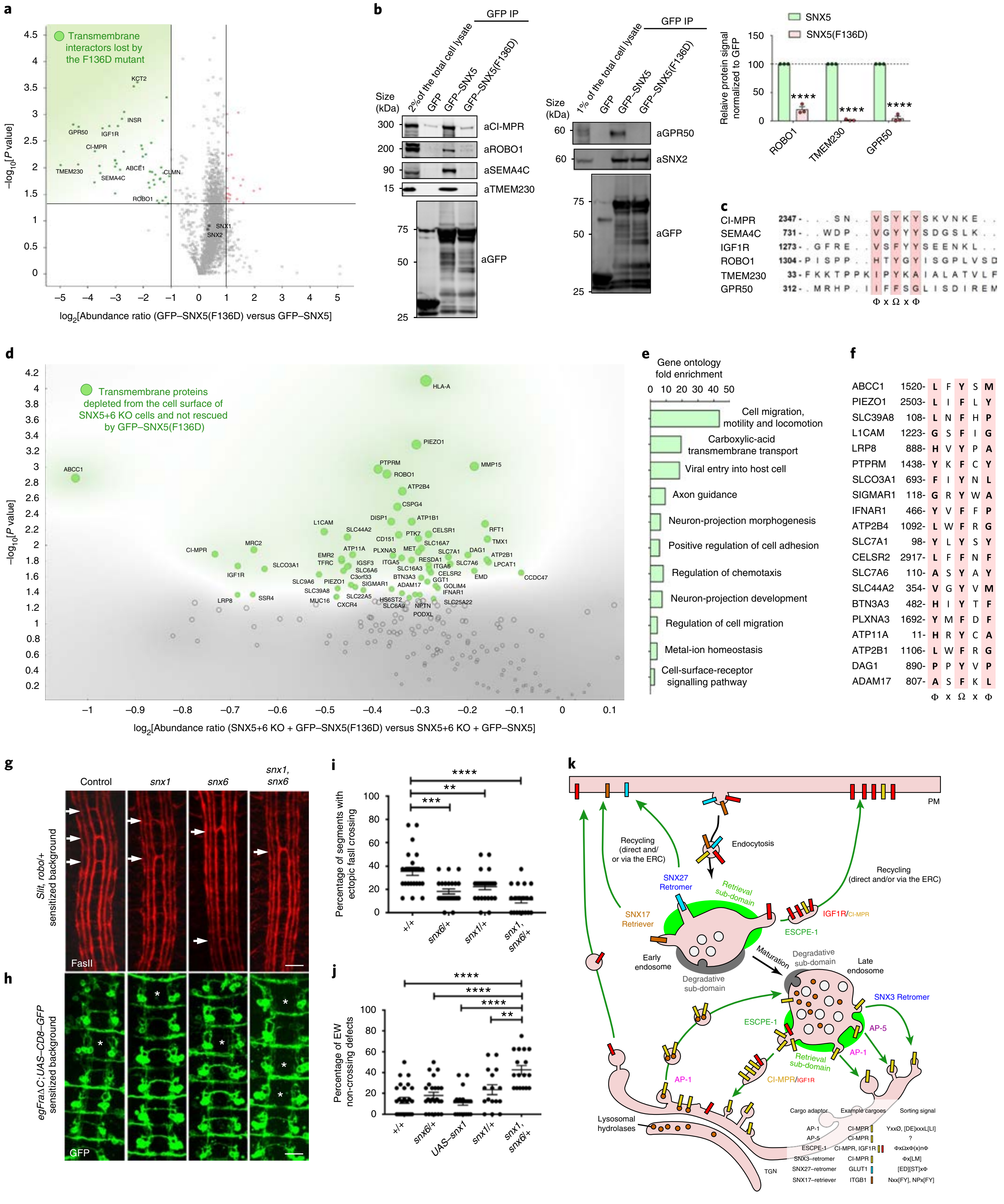
Discussion

The BAR-domain-containing SNX5 and SNX6 drive the biogenesis of tubular profiles and tubulo-vesicular transport carriers through the formation of heterodimeric assemblies with SNX1 or SNX2[19,22]. Here, we establish that a conserved hydrophobic groove in the PX domains of SNX5 and SNX6 (and presumably the related neuronal SNX32) associates with a ΦxΩxΦ$(x)_n$Φ motif present in the cytosolic tails of numerous integral proteins (x, any amino acid; n, variable linker region). Our study therefore establishes that SNX1/SNX2–SNX5/SNX6 heterodimers co-ordinate sequence-dependent cargo recognition with the biogenesis of tubulo-vesicular transport carriers for cargo export from the endosomal network.

The degradative capacity of lysosomes relies on the delivery of over 60 acid hydrolases from the biosynthetic pathway by means of membrane transport[40]. Their lysosomal delivery requires association with the CI-MPR at the TGN[30]. Transport of the CI-MPR–hydrolase complex from the TGN is mediated by clathrin-coated

Fig. 8 | Identification of a ΦxΩxΦ consensus motif for SNX5-mediated cargo recruitment. a, Comparative interactome analysis of wild-type SNX5 versus the SNX5(F136D) mutant across $n = 3$ independent experiments using a one-sample *t*-test and Benjamini–Hochberg false-discovery rate. Vertical black lines represent the threshold for interactions altered in the SNX5(F136D) interactome. Horizontal black line is the threshold for statistical significance. Green circles are interactions whose loss is statistically significant between GFP-SNX5(F136D) and GFP-SNX5. Red circles are interactions whose increase is statistically significant between GFP-SNX5(F136D) and GFP-SNX5. Grey circles are interactions whose change is not statistically significant. **b**, Co-immunoprecipitation of GFP-tagged SNX5 and SNX5(F136D) transiently transfected in HEK293T cells (left and middle) and a summary of the relative binding (right). The band intensities were quantified from $n = 3$ independent experiments using Odyssey software. The band intensities, normalized to GFP expression, are presented as the relative fraction of the GFP–SNX5 control. The unprocessed scans of immunoblots are shown in Supplementary Fig. 3. **c**, Alignment of the cytosolic tail of the SNX5 cargoes identified through the comparative interactome analysis of wild-type SNX5 versus the SNX5(F136D) mutant. **d**, Comparative proteomic analysis of transmembrane proteins that rely on the ability to engage the SNX5 hydrophobic groove for their plasma-membrane localization. Analysis was performed across $n = 3$ independent experiments using a one-sample *t*-test and Benjamini–Hochberg false-discovery rate. **e**, List of the most represented gene ontologies of the SNX5 cargoes, established by comparative surfaceosome analysis. **f**, Alignment of the cytosolic tail of the SNX5 cargoes that require the SNX5 hydrophobic groove for their endosomal retrieval and recycling, as established by comparative surfaceosome analysis. **g**, *Drosophila* Snx1 and Snx6 promote axon growth across the midline. Representative images of stage-17 *Drosophila* embryos stained with anti-FasII antibodies to reveal a subset of ipsilaterally projecting interneurons in the central nervous system. All embryos were heterozygous for mutations in *slit* and *robo*. The arrows indicate the segments in which the axon bundles are abnormally crossing the midline. The percentage of segments showing ectopic midline crossing was reduced in embryos that were also heterozygous for *snx1*, *snx6* or both *snx1* and *snx6* compared with control embryos. **h**, Representative images of stage-16 embryos stained with anti-GFP to reveal the Eagle subset of commissural interneurons, comprised of Eg axons that cross the midline in the anterior commissure and Eg axons that cross in the posterior commissure. The asterisks indicate the segments where the EW axons have failed to cross the midline. All embryos selectively express the FraΔC transgene, which results in the failure of a portion of EW axons to cross the midline. There was a significant reduction in the percentage of EW axons that crossed the midline in embryos heterozygous for mutations in both *snx1* and *snx6*. **i**, Quantification of ectopic midline crossing in the indicated genotypes. Eleven segments were scored in each embryo; $n = 26$ (wild-type control and *snx6/+*), 24 (*snx1/+*) and 23 (*snx1* and *snx6/+*) embryos. **j**, Quantification of failed midline crossing in the indicated genotypes. The eight abdominal segments were scored in each embryo; $n = 26$ (wild-type control), 24 (*snx6/+*), 16 (*snx1/+*), 17 (*snx1* and *snx6/+*) and 22 (UASSnx1) embryos. **i**,**j**, Each embryo (in which segments were quantified) was considered an independent trial. Statistical significance was determined using a one-way ANOVA, followed by a Tukey's test. **k**, Model of the role of the ESCPE-1 complex, which consists of heterodimeric combinations of either SNX5 or SNX6 dimerized to either SNX1 or SNX2, in retrieving and recycling transmembrane cargo protein on the cytosolic-facing surface of endosomes. The cartoon also represents the other machineries involved in the retrograde transport of cargoes to the TGN or in their recycling to the plasma membrane. The known sorting motifs in the cytosolic domain of the prototypical cargoes are listed. ERC, endocytic recycling compartment. The bars, error bars and circles in **b**,**i**,**j** represent the mean, s.e.m. and individual data points, respectively. The statistics source data are in Supplementary Table 4. The proteomic datasets are listed in Supplementary Table 3. Scale bars, 10 µm. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

transport vesicles[30]. Following fusion with the endosome, exposure to the acidic environment leads to hydrolase dissociation of the CI-MPR–hydrolase complex, which moves into the lysosome and becomes enzymatically active following endosome–lysosome fusion[41]. For continued hydrolase delivery, the unbound CI-MPR is sorted from the endosome into tubulo-vesicular transport carriers for transport back to the TGN[42,43]. The details of this step in CI-MPR sorting have remained elusive[44]. A number of molecules have been implicated, including TIP47 and the Rab9 GTPase[45], PACS-1 (ref. 46), AP-1 (ref. 47), AP-5 (ref. 48) and the Retromer complex[44], but their



precise role remains controversial for many of these molecules. Our structural identification and validation of the sorting motif in the CI-MPR and its recognition by the SNX5 and SNX6 components of the SNX1/SNX2–SNX5/SNX6 membrane tubulating complex has provided a detailed mechanistic framework on which to elucidate the early sorting steps of this important pathway.

How this insight interfaces with existing models of CI-MPR sorting, most notably with regard to Retromer, remains controversial[44]. The yeast Retromer, identified in yeast as a coat complex involved in endosome-to-Golgi transport[8], is a stable pentamer of two sub-complexes: a Vps26, Vps35 and Vps29 heterotrimer, and a heterodimer of the SNX–BAR proteins Vps5 and Vps17. In higher metazoans, the VPS26–VPS35–VPS29 heterotrimer[9,10] does not form a stable pentameric assembly with the corresponding SNX1/SNX2–SNX5/SNX6 heterodimer (SNX1 and SNX2 being Vps5 orthologs, with SNX5 and SNX6 being Vps17 orthologs)[19,20]. Moreover, the human VPS26–VPS35–VPS29 heterotrimer can function independently of the SNX1/SNX2–SNX5/SNX6 heterodimer in cargo sorting[23,24], a conclusion recently extended in vivo through the endosomal sorting of planar cell polarity proteins in the *Drosophila* wing[49]. Although it cannot be excluded that a metazoan yeast-like pentameric Retromer complex may assemble under certain specific condition(s), the divergence in the association and functional interaction between the two complexes would argue that such a pentameric complex may form only a minor functional population. In metazoans, the divergence away from the yeast pentameric complex has led to a refined nomenclature where the VPS26–VPS35–VPS29 heterotrimer is referred to as Retromer[50].

The evolution of Retromer and the SNX1/SNX2–SNX5/SNX6 complex into separate functional complexes is paralleled by Retromer acquiring a number of interactions with proteins not found in yeast, including sorting nexin-27 (SNX27)[13,14,51], the actin polymerizing WASH complex[52–54] and ANKRD50 (ref. [55,56]). Equally, the functional role of Retromer has expanded to include the sorting of hundreds of cargoes to the cell surface[2]. The evolution in diversity can now be extended to the SNX1/SNX2–SNX5/SNX6 heterodimer. The extended helix-loop-helix cargo-binding groove in SNX5 and SNX6 is absent from yeast Vps17 (Fig. 1b). This mode of sequence-dependent cargo recognition is therefore not a feature shared with the yeast pentameric Retromer. However, the conservation of the hydrophobic groove in *Caenorhabditis elegans*, *Drosophila melanogaster* and *Danio rerio* (Fig. 1b), indicates that an equivalent mode of sequence-dependent endosomal sorting is likely to exist across these organisms, a conclusion supported by our preliminary analysis of axonal guidance in *Drosophila*.

A number of studies have highlighted the mammalian Retromer in the endosome-to-TGN transport of the CI-MPR[44]. Here, cargo selection is proposed to occur through the recognition of the ^{2369}WLM^{2371} motif at an interface between VPS26 and the non-BAR-domain-containing sorting nexin SNX3 (ref. [15]). Although structural evidence remains to be established, both Retromer and SNX3 can mediate CI-MPR transport in carriers that depend on the Golgin GCC88 for tethering at the TGN[57]. In contrast, SNX1 and SNX2 (and by extension SNX5 and SNX6) mediate CI-MPR trafficking in carriers dependent on the Golgin-245 tether, leading to the proposal of two independent pathways for CI-MPR sorting to the TGN: a SNX3–Retromer pathway[57] and a separate pathway dependent on SNX1/SNX2–SNX5/SNX6 (ref. [23,24]). In addition, AP-1 and AP-5 adaptors mediate endosome-to-TGN CI-MPR transport[43,58,59] (Fig. 8k). Disentangling the sequential versus concurrent relationship of these pathways will provide invaluable insight into the global organization of the endosome-to-TGN transport pathways.

For the SNX1/SNX2–SNX5/SNX6 coat complex, cargoes that are exported from the endosomal network through engaging SNX5/SNX6 can be sorted to (at least) two distinct destinations: for IGF1R (and numerous other cargoes) a steady-state cell surface enrichment and a TGN-enrichment for the CI-MPR[23,24] (Fig. 8k). Which principles govern how recognition of similar motifs by the same sorting complex ultimately lead to cargo sorting to the plasma membrane versus the TGN? At present, we have no direct experimental data to address this fundamental question. However, von Zastrow and colleagues have documented that endosomal exit of two cargo proteins, β2-adrenergic receptor and Wntless, occurs through a shared tubular profile, although their steady-state enrichments are ultimately defined as the cell surface and TGN, respectively[60]. Although the mechanistic basis of these sorting events are distinct, sorting of the β2-adrenergic receptor and Wntless is Retromer-dependent[23,24] and the central problem is conceptually the same. Similar to the mechanism suggested for the β2-adrenergic receptor and Wntless[60], we believe that the relative enrichment of the two cargoes is an important element in their sorting itineraries. Following activation, IGF1R enters the endosomal network from the cell surface while the majority of the CI-MPR undergoes continual rounds of TGN-to-endosome and endosome-to-TGN transport from maturing endosomes. The relative enrichment of the two cargoes therefore reflects the maturation status of the endosome: endocytosed IGF1R (and other cell surface cargoes) will be enriched in newly formed endosomes, whereas CI-MPR will be enriched in the 'late' endosomes. Tubular exit across the endosomal maturation pathway promoted by SNX1/SNX2–SNX5/SNX6 will therefore display distinct cargo enrichments: early exit will favour IGF1R over CI-MPR, whereas later exit favours CI-MPR over IGF1R. If early exit promotes entry into cell surface recycling pathways (either directly or via the endocytic recycling compartment) and late exit promotes entry into retrograde transport pathways, then such a sorting 'refinery' would achieve differential sorting. Such sorting may not immediately result in the correct steady-state localization for all cargoes—for example, some IGF1R may reach the TGN—but the inter-connected nature of membrane trafficking pathways provides a route for cell surface delivery through re-secretion (Fig. 8k). There is indeed evidence for cell surface cargoes, such as β1-integrin, being recycled through the TGN[61]. The presence of such 'sub-refineries', including the TGN and the endocytic recycling compartment, would provide further platforms for differential sorting.

Finally, the interaction between *C. trachomatis* effector IncE and SNX5/SNX6 serves to manipulate host mechanisms to promote pathogen survival and replication in the intracellular environment of the host[26,62]. The secreted IncE is associated with the cytosolic face of the *C. trachomatis* inclusion body from where it recruits SNX5 and SNX6, and overrides the normal endosome association of these SNX–BARs[2,62]. This leads to a defect in the CI-MPR endosome-to-TGN transport[26–28,62]. Our structural description of CI-MPR binding to SNX5/SNX6 and its resemblance to that of IncE establishes a molecular explanation for this phenotype. With the importance of CI-MPR recycling for hydrolase delivery and lysosomal degradative health, *C. trachomatis* may induce perturbation in this pathway to reduce the restrictive role of this organelle. Our analysis has identified a number of signalling receptors that require SNX5/SNX6 for their endosomal sorting, which are likely to provide insight into the significance of IncE-perturbed SNX5/SNX6 trafficking. One way by which *C. trachomatis* secures survival of the host cell is by modulating the MEK–ERK and PI 3-kinase survival pathway downstream of EPHA2 receptor activation. This receptor is internalized with, and retained on, the maturing inclusion from where it triggers long-lasting anti-apoptotic signalling[63]. It is tempting to speculate that the secreted IncE, by preventing the SNX5/SNX6-mediated transport of active signalling receptors away from the inclusion, may ensure retention of pro-survival signalling receptors during replication.

Overall, we have identified a mechanism through which the endosome-associated SNX1/SNX2–SNX5/SNX6 coat complex coordinates sequence-dependent cargo recognition with membrane remodelling to generate cargo-enriched tubulo-vesicular transport carriers. We have provided evidence that this mechanism is required for the export of numerous cargoes from the endosomal network for onward transport to the TGN and the cell surface. To highlight this function and to reflect that the mammalian VPS26–VPS35–VPS29 Retromer and the SNX1/SNX2–SNX5/SNX6 coat complex have functionally diversified, we propose to refer to the SNX1/SNX2–SNX5/SNX6 coat complex as the 'Endosomal SNX–BAR sorting complex for promoting exit-1' (ESCPE-1; Fig. 8k). We speculate that additional ESCPE complexes may function in sequence-mediated tubulo-vesicular-based endosomal sorting.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41556-019-0393-3.

Received: 30 January 2019; Accepted: 19 August 2019;
Published online: 1 October 2019

References

1. Goldenring, J. R. Recycling endosomes. *Curr. Opin. Cell Biol.* **35**, 117–122 (2015).
2. Cullen, P. J. & Steinberg, F. To degrade or not to degrade: mechanisms and significance of endocytic recycling. *Nat. Rev. Mol. Cell Biol.* **19**, 679–696 (2018).
3. Behnia, R. & Munro, S. Organelle identity and the signposts for membrane traffic *Nature* **438**, 597–604 (2005).
4. McNally, K. E. & Cullen, P. J. Endosomal retrieval of cargo: retromer is not alone. *Trends Cell Biol.* **28**, 807–822 (2018).
5. Phillips-Krawczak, C. A. et al. COMMD1 is linked to the WASH complex and regulates endosomal trafficking of the copper transporter ATP7A. *Mol. Biol. Cell* **26**, 91–103 (2015).
6. Bartuzi, P. et al. CCC- and WASH-mediated endosomal sorting of the LDLR is required for normal clearance of circulating LDL. *Nat. Commun.* **7**, 10961 (2016).
7. Meyer, C. et al. μ1A-adaptin-deficient mice: lethality, loss of AP-1 binding and rerouting of mannose 6-phosphate receptors. *EMBO J.* **19**, 2193–2203 (2000).
8. 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).
9. Seaman, M. N. Cargo-selective endosomal sorting for retrieval to the Golgi requires retromer. *J. Cell Biol.* **165**, 111–122 (2004).
10. 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).
11. Strochlic, T. I., Setty, T. G., Sitaram, A. & Burd, C. G. Grd19/Sns3p functions as a cargo-specific adapter for retromer-dependent endocytic recycling. *J. Cell Biol.* **177**, 115–125 (2007).
12. Harterink, M. et al. A SNX3-dependent retromer pathway mediates retrograde transport of the Wnt sorting receptor Wntless and is required for Wnt secretion. *Nat. Cell Biol.* **13**, 914–923 (2011).
13. Temkin, P. et al. SNX27 mediates retromer tubule entry and endosome-to-plasma membrane trafficking of signalling receptors. *Nat. Cell Biol.* **13**, 715–721 (2011).
14. Steinberg, F. et al. A global analysis of SNX27-retromer assembly and cargo specificity reveals a function in glucose and metal ion transport. *Nat. Cell Biol.* **15**, 461–471 (2013).
15. Lucas, M. et al. Structural mechanism for cargo recognition by the retromer complex. *Cell* **167**, 1623–1635 (2016).
16. McNally, K. E. et al. Retriever is a multiprotein complex for retromer-independent endosomal cargo recycling. *Nat. Cell Biol.* **19**, 1214–1225 (2017).
17. Purushothaman, L. K. & Ungermann, C. Cargo induces retromer-mediated membrane remodelling on membranes. *Mol. Biol. Cell* **29**, 2709–2719 (2018).
18. Kovtun, O. et al. Structure of the membrane-assembled retromer coat determined by cryo-electron tomography. *Nature* **561**, 561–564 (2018).
19. 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).
20. Wassmer, T. et al. A loss-of-function screen reveals SNX5 and SNX6 as potential components of the mammalian retromer. *J. Cell Sci.* **120**, 45–54 (2007).
21. Wassmer, T. et al. The retromer coat complex coordinates endosomal sorting and dynein-mediated transport with carrier recognition by the trans-Golgi network. *Dev. Cell* **17**, 110–122 (2009).
22. van Weering, J. R. et al. Molecular basis for SNX–BAR-mediated assembly of distinct endosomal sorting tubules. *EMBO J.* **31**, 4466–4480 (2012).
23. Simonetti, B., Danson, C. M., Heesom, K. J. & Cullen, P. J. Sequence-dependent cargo recognition by SNX-BARs mediates retromer-independent transport of CI-MPR. *J. Cell Biol.* **216**, 3695–3712 (2017).
24. Kvainickas, A. et al. Cargo-selective SNX–BAR proteins mediate retromer trimer independent retrograde transport. *J. Cell Biol.* **216**, 3677–3693 (2017).
25. Teasdale, R. D., Loci, D., Houghton, F., Karlsson, L. & Gleeson, P. A. A large family of endosome-localised proteins related to sorting nexin 1. *Biochem. J.* **358**, 7–16 (2001).
26. Elwell, C. A. et al. *Chlamydia* interfere with an interaction between the mannose-6-phosphate receptor and sorting nexins to counteract host restriction. *eLife* **6**, e22709 (2017).
27. Paul, B. et al. Structural basis for the hijacking of endosomal sorting nexin proteins by *Chlamydia trachomatis*. *eLife* **6**, e22311 (2017).
28. Sun, Q. et al. Structural and functional insights into sorting nexin 5/6 interaction with bacterial effector IncE. *Signal Transduct. Target. Ther.* **2**, 17030 (2017).
29. Seaman, M. N. Identification of a novel conserved sorting motif required for retromer-mediated endosome-to-TGN retrieval. *J. Cell Sci.* **120**, 2378–2389 (2007).
30. Coutinho, M. F., Prata, M. J. & Alves, S. Mannose-6-phosphate pathway: a review on its role in lysosomal function and dysfunction. *Mol. Genet. Metab.* **105**, 542–550 (2012).
31. Itai, N. et al. The phosphorylation of sorting nexin 5 at serine 226 regulates retrograde transport and macropinocytosis. *PLoS One* **13**, e0207205 (2018).
32. Chandra, M. et al. Classification of the human phox homology (PX) domains based on their phosphoinositide binding specificities. *Nat. Commun.* **10**, 1528 (2019).
33. Blockus, H. & Chedotal, A. Slit-Robo signaling. *Development* **143**, 3037–3044 (2016).
34. Deng, H. X. et al. Identification of *TMEM230* mutations in familial Parkinson's disease. *Nat. Genet.* **48**, 733–739 (2016).
35. Wojciech, S. et al. The orphan GPR50 receptor promotes constitutive TGFβ receptor signaling and protects against cancer development. *Nat. Commun.* **9**, 1216 (2018).
36. Fan, X., Labrador, J. P., Hing, H. & Bashaw, G. J. Slit stimulation recruits Dock and Pak to the roundabout receptor and increases rac activity to regulate axon repulsion at the CNS midline. *Neuron* **40**, 113–127 (2003).
37. Kidd, T., Bland, K. S. & Goodman, C. S. Slit is the midline repellent for the Robo receptor in *Drosophila*. *Cell* **96**, 785–794 (1999).
38. Coleman, H. A., Labrador, J. P., Chance, R. K. & Bashaw, G. J. The Adam family metalloprotease Kuzbanian regulates the cleavage of the roundabout receptor to control axon repulsion at the midline. *Development* **137**, 2417–2426 (2010).
39. Hernandez-Fleming, M., Rohrbach, E. & Bashaw, G. J. Sema-1a reverse signaling promotes midline crossing in response to secreted Semaphorins. *Cell Rep.* **18**, 174–184 (2017).
40. Klumperman, J. & Raposo, G. The complex ultrastrutcure of the endolysosomal system. *Cold Spring Harb. Perspect. Biol.* **6**, a016857 (2014).
41. Bright, N. A., Davis, L. J. & Luzio, J. P. Endolysosomes are the principal intracellular sites of acid hydrolase activity. *Curr. Biol.* **26**, 2233–2245 (2016).
42. Mari, M. et al. SNX1 defines an early endosomal recycling exit for sortilin and mannose 6-phosphate receptors. *Traffic* **9**, 380–393 (2008).
43. Johannes, L. & Wunder, C. Retrograde transport: two (or more) roads diverged in an endosomal tree? *Traffic* **12**, 956–962 (2011).
44. Seaman, M. N. J. Retromer and the cation-independent mannose 6-phosphate receptor—time for a trial separation? *Traffic* **19**, 150–152 (2018).
45. Carroll, K. S. et al. Role of Rab9 GTPase in facilitating receptor recruitment by TIP47. *Science* **292**, 1373–1376 (2001).
46. Wan, L. et al. PACS-1 defines a novel gene family of cytosolic sorting proteins required for trans-Golgi network localisation. *Cell* **94**, 205–216 (1998).
47. Robinson, M. S., Sahlender, D. A. & Foster, S. D. Rapid inactivation of proteins by rapamycin-induced rerouting to mitochondria. *Dev. Cell* **18**, 324–331 (2010).
48. Hirst, J. et al. Interaction between AP-5 and the hereditary spastic paraplegia proteins SPG11 and SPG15. *Mol. Biol. Cell* **24**, 2558–2569 (2013).
49. Strutt, H. et al. Retromer controls planar polarity protein levels and asymmetric localization at intercellular junctions. *Curr. Biol.* **29**, 484–491 (2019).
50. Burd, C. & Cullen, P. J. Retromer: a master conductor of endosome sorting. *Cold Spring Harb. Perspect. Biol.* **6**, a016774 (2014).
51. Gallon, M.et al. A unique PDZ domain and arrestin-like fold interaction reveals mechanistic details of endocytic recycling by SNX27-retromer. *Proc. Natl Acad. Sci. USA* **111**, E3604–13 (2014).
52. Gomez, T. S. & Billadeau, D. D. A FAM21-containing WASH complex regulates retromer-dependent sorting. *Dev. Cell* **17**, 699–711 (2009).
53. Harbour, M. E., Breusegem, S. Y. & Seaman, M. N. Recruitment of the endosomal WASH complex is mediated by the extended 'tail' of Fam21 binding to the retromer protein Vps35. *Biochem. J.* **442**, 209–220 (2012).
54. Jia, D., Gomez, T. S., Billadeau, D. D. & Rosen, M. K. Multiple repeat elements within the FAM21 tail link the WASH actin regulatory complex to the retromer. *Mol. Biol. Cell* **23**, 2352–2361 (2012).

55. McGough, I. J. et al. Identification of molecular heterogeneity in SNX27-retromer-mediated endosome-to-plasma membrane recycling. *J. Cell Sci.* **127**, 4940–4953 (2014).
56. Kvainickas, A. et al. Retromer- and WASH-dependent sorting of nutrient transporters requires a multivalent interaction network with ANKRD50. *J. Cell Sci.* **130**, 382–395 (2017).
57. Cui, Y. et al. Retromer has a selective function in cargo sorting via endosome transport carriers. *J. Cell Biol.* **218**, 615–631 (2019).
58. Hirst, J. et al. Distinct and overlapping roles for AP-1 and GGAs revealed by the 'knocksideways' system. *Curr. Biol.* **22**, 1711–1716 (2012).
59. Hirst, J., Itzhak, D. N., Antrobus, R., Borner, G. H. H. & Robinson, M. S. Role of the AP-5 adaptor protein complex in late endosome-to-Golgi retrieval. *PLoS Biol.* **16**, e2004411 (2018).
60. Varandas, K. C., Irannejad, R. & von Zastrow, M. Retromer endosome exit domains serve multiple trafficking destinations and regulate local G protein activation by GPCRs. *Curr. Biol.* **26**, 3129–3142 (2016).
61. Shafaq-Zadah, M. et al. Presistent cell migration and adhesion rely on retrograde transport of β1-integrin. *Nat. Cell Biol.* **18**, 54–64 (2016).
62. Mirrashidi, K. M. et al. Global mapping of the Inc-human interactome reveals that retromer restricts *Chlamydia* infection. *Cell Host Microbe* **18**, 109–121 (2015).
63. Subbarayal, P. et al. EphrinA2 receptor (EphA2) is an invasion and intracellular signalling receptor for *Chlamydia trachomatis*. *PLoS Pathog.* **11**, e1004846 (2015).

Acknowledgements

This work was supported by the MRC (grant nos MR/L007363/1 and MR/P018807/1), the Wellcome Trust (grant no. 104568/Z/14/2), the Lister Institute of Preventive Medicine to P.J.C. and the Wellcome Trust PhD Studentship for the Dynamic Cell Biology programme (grant no. 083474) to B.S. B.M.C. is supported by an NHMRC Senior Research Fellowship (grant no. APP1136021) and the NHMRC (grant no. APP1099114), the Australian Research Council (grant no. DP160101743) and the Bright Focus Foundation (grant no. A2018627S). G.J.B. is supported by the National Institutes of Health (grant no. R35 NS097340). We thank the University of Queensland Remote Operation Crystallisation and X-ray facility and the staff for their support with the crystallization experiments, the staff of the Australian Synchrotron for assistance with X-ray diffraction data collection and the University of Bristol Wolfson Bioimaging Facility.

Author contributions

Initial conceptualization: B.S., F.S., B.M.C. and P.J.C. Evolution of conceptualization: B.S., B.P., S.W., F.S., G.J.B., B.M.C. and P.J.C. Formal analysis: K.J.H. Investigation: B.S., B.P., K.C., S.W., M.G., F.S. and K.J.H. Writing of the original draft: B.S., B.M.C. and P.J.C. Writing, review and editing: all authors. Funding acquisition: F.S., G.J.B., B.M.C. and P.J.C. Supervision: F.S., G.J.B., B.M.C. and P.J.C.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41556-019-0393-3.

Correspondence and requests for materials should be addressed to B.M.C. or P.J.C.

Methods

Antibodies. The antibodies used in this study were: mouse monoclonal antibodies to SNX1 (clone 51/SNX1; cat. no. 611482, BD; 1:1,000 for western blots, 1:200 for immunofluorescence), SNX2 (clone 13/SNX2; cat. no. 5345661, BD; 1:1,000 for western blots, 1:200 for immunofluorescence), SNX6 (clone d-5; cat. no. 365965, Santa Cruz Biotechnology, Inc.; 1:1,000 for western blots, 1:200 for immunofluorescence), β-actin (A1978, Sigma-Aldrich; 1:2,000 for western blots), GFP (clones 7.1 and 13.1; cat. no. 11814460001, Roche; 1:2,000 for western blots, 1:400 for immunofluorescence), CI-MPR (clone 2G11; MA1-066, Thermo Fisher Scientific; 1:1,000 or western blots, 1:100 for immunofluorescence); Golgin-97 (CDF4; A-21270, Thermo Fisher Scientific; 1:200 or immunofluorescence), rabbit polyclonal antibodies to mCherry (ab167453, Abcam; 1:2,000 for western blots, 1:400 for immunofluorescence), SNX1 (ab995, Abcam; 1:1,000 for western blots), SNX5 (17918-1-AP, Proteintech; 1:1,000 for western blots), GAPDH (G9545, Sigma-Aldrich; 1:2,000 for western blots), AP2A1 (11401-1-AP, Proteintech; 1:1,000 for western blots), AP1G1 (13258-1-AP, Proteintech; 1:1,000 for western blots), SEMA4C (ab171559, Abcam; 1:1,000 for western blots), TMEM230 (21466-1-AP, Proteintech; 1:1,000 for western blots), ROBO1 (20219-1-AP, Proteintech; 1:1,000 for western blots), GPR50 (19762-1-AP, Proteintech; 1:1,000 for western blots); rabbit monoclonal antibodies to SNX5 (EPR14358; cat. no. 180520, Abcam; 1:1,000 for western blots), CI-MPR (EPR6599; cat. no. 124767, Abcam; 1:1,000 for western blots, 1:200 for immunofluorescence), IGF1R (clone D23H3; cat. no. 111A9, Cell Signaling Technologies; 1:1,000 for western blots), EEA1 (clone c45b10; cat. no. 3288, Cell Signaling Technology; 1:200 for immunofluorescence); goat polyclonal antibodies to EEA1 (N-19; sc-6415, Santa Cruz Biotechnology, Inc.; 1:200 for immunofluorescence) and sheep polyclonal antibody to TGN46 (AHP500G; Bio-Rad Laboratories; 1:200 for immunofluorescence).

Cell culture and transfection. The HeLa, HEK293T and RPE1 cell lines were originally sourced from the American Type Culture Collection. Authentication was from the American Type Culture Collection. We did not independently authenticate the cell lines. The parental HeLa cells used for the generation of all CRISPR-edited cell lines were mycoplasma tested, which revealed no contamination. Our HEK293T cell line also tested negative for mycoplasma contamination. No commonly misidentified cell lines were used. Cells were grown in DMEM medium (Sigma-Aldrich) supplemented with 10% (vol/vol) FCS (Sigma-Aldrich) and penicillin/streptomycin (Gibco), and grown under standard conditions. FuGENE HD (Promega) was used for transient transfection of DNA according to the manufacturer's instructions. For the isolation of CRISPR–Cas9 clonal lines, HeLa cells were transfected with pX330 plasmid coding for the genomic RNA (gRNA) of the gene of interest together with a puromycin-resistance-expressing plasmid. The cells were then subjected to puromycin selection for 24 h. The cells were subsequently resuspended using accutase (BioLegend) and diluted to 2.5 cells ml^{-1} in Iscov's modified Dulbecco's medium (Sigma-Aldrich) supplemented with 10% (vol/vol) FBS (Sigma-Aldrich). Next, 200 µl was plated in the wells of 10× 96-well plates and the plates were screened for the presence of cell colonies after three weeks. The cell colonies were expanded and subjected to lysis and western blotting to determine the levels of the target proteins. The SNX5+6 double-knockout cell line used in this study was characterized previously (clone13 in [23]); the SNX1+2 double-knockout cell line used in this study was also characterized previously (clone4 in [23]). The CI-MPR-knockout cell line was generated for this study using the following three gRNAs: 5′-GCTTGTCCTGAGTTACGTGA-3′, 5′-GTGTGCACTACTTTGAGTGG-3′ and 5′-GAGAAGGAAGACCTCCTCTG-3′. The IGF1R-knockout cell line was generated for this study using the following three gRNAs: 5′-GATGATGCGATTCTTCGACG-3′, 5′-GGTTGGGGAAGAGGTCTCCG-3′ and 5′-GGAGAACGACCATATCCGTG-3′.

Generation of stable lentiviral RPE1 cell lines. The gene of interest was subcloned into the lentiviral vector pXLG3 for the generation of lentiviral particles. Lentiviral particles were produced and harvested in HEK293T cells. HeLa cells were transduced with near-to-endogenous levels of lentiviral particles to produce stably expressing cell lines. Transduced HeLa cells were grown in DMEM supplemented with 10% (vol/vol) FCS and penicillin/streptomycin and grown at 37 °C in a 5% CO$_2$ incubator.

Immunoprecipitation and quantitative western blot analysis. Cells were lysed in PBS with 1% Triton X-100 and protease inhibitor cocktail for western blotting. The protein concentration was determined using a BCA assay kit (Thermo Fisher Scientific) and equal amounts were resolved on NuPAGE 4–12% precast gels (Invitrogen). Blotting was performed onto polyvinylidene fluoride membranes (Immobilon-FL, EMD Millipore), followed by detection using the Odyssey infrared scanning system (LI-COR Biosciences). When using the Odyssey, we routinely performed western blot analysis where a single blot was simultaneously probed with antibodies against two proteins of interest (distinct antibody species), followed by visualization with the corresponding secondary antibodies conjugated to distinct spectral dyes. For the GFP-based immunoprecipitations, HEK293T cells were transfected with GFP constructs using polyethylenimine (Sigma-Aldrich). The cells were lysed in immunoprecipitation buffer (50 mM Tris–HCl, 0.5% NP-40 and Roche protease inhibitor cocktail) 48 h after transfection and subjected to GFP-trap (ChromoTek) or myc-tap (ChromoTek) beads. Immunoblotting was performed using standard procedures. Detection was performed on an Odyssey infrared scanning system (LI-COR Biosciences) using fluorescently labelled secondary antibodies.

Biotinylation of cell surface proteins. For the surface biotinylation experiments, fresh Sulfo NHS-SS-Biotin (cat. no. 21217, Thermo Scientific) was dissolved in ice-cold PBS at pH 7.8 at a final concentration of 0.2 mg ml^{-1}. Cells were washed twice in ice-cold PBS and placed on ice to slow down the endocytic pathway. Next, the cells were incubated with the biotinylation reagent for 30 min at 4 °C, followed by incubation in TBS for 10 min to quench the unbound biotin. The cells were then lysed in lysis buffer and subjected to Streptavidin-bead-based affinity isolation (GE Healthcare).

TMT labelling and high pH reversed-phase chromatography. Samples were digested overnight with 2.5 µg trypsin at 37 °C, labelled with TMT ten-plex reagents according to the manufacturer's protocol (Thermo Fisher Scientific) and the labelled samples were pooled. The pooled sample was then dried by evaporation, resuspended in 5% formic acid and desalted using a SepPak cartridge according to the manufacturer's instructions (Waters). The eluate from the SepPak cartridge was again dried by evaporation and resuspended in buffer A (20 mM ammonium hydroxide, pH 10) before fractionation by high pH reversed-phase chromatography using an UltiMate 3000 liquid chromatography system (Thermo Scientific). Briefly, the sample was loaded onto an XBridge BEH C18 column (130 Å, 3.5 µm, 2.1 mm × 150 mm; Waters) in buffer A and the peptides were eluted with an increasing gradient of buffer B (20 mM ammonium hydroxide in acetonitrile, pH 10) from 0–95% over 60 min. The resulting fractions were dried by evaporation and resuspended in 1% formic acid before analysis by nano-LC–MS/MS using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific).

Nano-LC–MS/MS. The high pH reversed-phase fractions were further fractionated using an UltiMate 3000 nano-LC system in line with an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific). Briefly, peptides in 1% (vol/vol) formic acid were injected onto an Acclaim PepMap C18 nanotrap column (Thermo Scientific). After washing with 0.5% (vol/vol) acetonitrile in 0.1% (vol/vol) formic acid, the peptides were resolved on a 250 mm × 75 µm Acclaim PepMap C18 reverse-phase analytical column (Thermo Scientific) over a 150 min organic gradient, using seven gradient segments (1–6% solvent B over 1 min, 6–15% solvent B over 58 min, 15–32% solvent B over 58 min, 32–40% solvent B over 5 min, 40–90% solvent B over 1 min, held at 90% solvent B for 6 min and then reduced to 1% solvent B over 1 min) with a flow rate of 300 nl min^{-1}. Solvent A was 0.1% formic acid and solvent B was aqueous 80% acetonitrile in 0.1% formic acid. The peptides were ionized by nano-electrospray ionization at 2.0 kV using a stainless steel emitter with an internal diameter of 30 µm (Thermo Scientific) and a capillary temperature of 275 °C.

All spectra were acquired using an Orbitrap Fusion Tribrid mass spectrometer controlled by Xcalibur 2.0 software (Thermo Scientific) and operated in data-dependent acquisition mode using an SPS-MS3 workflow. FTMS1 spectra were collected at a resolution of 120,000, with an automatic-gain-control target of 200,000 and a maximum injection time of 50 ms. Precursors were filtered with an intensity threshold of 5,000, according to the charge state (to include charge states 2–7) and with monoisotopic precursor selection. Previously interrogated precursors were excluded using a dynamic window (60 s ± 10 ppm). The MS2 precursors were isolated with a quadrupole mass filter set to a width of 1.2 *m/z*. ITMS2 spectra were collected with an automatic-gain-control target of 10,000, maximum injection time of 70 ms and CID collision energy of 35%.

For the FTMS3 analysis, the Orbitrap was operated at a resolution of 50,000 with an automatic-gain-control target of 50,000 and a maximum injection time of 105 ms. The precursors were fragmented by high-energy collision dissociation at a normalized collision energy of 60% to ensure maximal TMT reporter-ion yield. Synchronous precursor selection was enabled to include up to five MS2 fragment ions in the FTMS3 scan.

Data analysis. The raw data files were processed and quantified using Proteome Discoverer software v2.1 (Thermo Scientific) and searched against the UniProt human database (downloaded 14 September 2017; 140,000 sequences) using the SEQUEST algorithm. The peptide precursor mass tolerance was set at 10 ppm and MS/MS tolerance was set at 0.6 Da. Search criteria included oxidation of methionine (+15.9949) as a variable modification, and carbamidomethylation of cysteine (+57.0214) and the addition of the TMT (+229.163) to the peptide N termini and lysine as fixed modifications. The searches were performed with full tryptic digestion and a maximum of two missed cleavages were allowed. The reverse database search option was enabled and the data were filtered to satisfy a false-discovery rate of 5%. The Perseus software platform was used for the statistical analysis of the proteomics datasets[64].

Immunofluorescence staining. Cells were fixed in 4% paraformaldehyde for 20 min, washed in PBS and then permeabilized with 0.1% Triton X-100. The fixed cells were blocked in 1% BSA and incubated in primary antibody and the respective secondary antibody (Alexa Fluor; Thermo Fisher Scientific) in 1% BSA.

For the aCI-MPR antibody uptake assay, HeLa cells were transferred to a well containing 3 ml ice-cold DHB (DMEM supplemented with 20 mM HEPES and 1% (vol/vol) FCS) for 15 min to stop the trafficking processes. Surface CI-MPR was labelled by incubation for 30 min in cold DHB containing an antibody targeting the extracellular domain of the receptor (clone 2G11; MA1-066, Thermo Fisher Scientific). Coverslips were washed twice in ice-cold PBS, transferred to pre-warmed growth media and returned to the incubator for 40 min to allow uptake of the anti-CI-MPR to occur. The cells were then washed, fixed and immunostained.

Image acquisition and analysis. Microscopy images were collected with a confocal laser scanning microscope (SP5 AOBS, Leica Microsystems) attached to an inverted epifluorescence microscope (DMI6000, Thermo Fisher Scientific). A ×63 1.4 NA (numerical-aperture) oil-immersion objective (Plan Apochromat BL, Leica Biosystems) and the standard SP5 system acquisition software and detector were used. Images were captured at room temperature as z stacks with photomultiplier tube detectors with a photocathode made of gallium-arsenide-phosphide (Leica Microsystems) for collecting light emission. The images were captured using Application Suite AF software (version 2.7.3.9723; Leica Microsystems) and then analysed with Volocity 6.3 software (PerkinElmer). For colocalization studies, Pearson's correlation colocalization coefficient was measured using the method of Costes to set automatic thresholds. For the quantification of CI-MPR tubules, an individual tubular profile was scored based on a length that was $\geq 3\,\mu\text{m}$.

Protein expression and purification. The constructs for bacterial expression of glutathione *S*-transferase (GST)-tagged PX domains of human SNX5 (residues 22–170), SNX6 (residues 26–170) and SNX32 (residues 17–166) were described previously[27]. The genes encoding the human SNX5 PX domain (residues 22–170) with C-terminal fusions of the CI-MPR sequence (residues 2347–2377) or SEMA4C sequence (residues 731–755) were codon optimized and synthesized by Genscript and cloned into pGEX4T-2. The GST-tagged proteins were expressed in *Escherichia coli* BL21 codon plus supplemented with the appropriate antibiotics. Single colonies from cultures grown on lysogeny-broth (LB) agar plates were inoculated into 50 ml LB media supplemented with ampicillin (0.1 mg ml^{-1}) and chloramphenicol (0.034 mg ml^{-1}), and grown overnight at 37 °C with shaking. The following day, 30 ml from the overnight culture was used to inoculate 1 l LB media containing ampicillin (0.1 mg ml^{-1}) and chloramphenicol (0.034 mg ml^{-1}), and incubated at 37 °C. The cells were cultured to an optical density of 0.7–0.8 at 600 nm and induced with 0.5 mM isopropyl-β-d-thiogalactopyranoside. The cultures were incubated with shaking overnight at 18 °C. The cells were harvested using a Beckman rotor JLA 8.1000 at 4,000 r.p.m. for 15 min at 4 °C. The pellets were resuspended in 10 ml lysis buffer (50 mM Tris pH 8.0, 100 mM NaCl, 5% glycerol, 1 mM dithiothreitol, 0.1 mg ml^{-1} benzamidine and 0.1 mg ml^{-1} DNase) per litre of culture. The cells were subjected to cell disruption and further centrifugation using a Beckman rotor JA25.50 at 18,000 r.p.m. for 30 min at 4 °C. The soluble fractions were first purified using affinity chromatography with glutathione-sepharose and the GST tags were cleaved by thrombin while still bound to the column. The proteins were eluted in 50 mM Tris pH 8.0, 100 mM NaCl, 5% glycerol and 1 mM dithiothreitol, and then further polished using gel-filtration chromatography (Superdex 75, GE Healthcare) in ITC buffer (50 mM HEPES pH 8.0 and 100 mM NaCl). The fractions corresponding to the respective proteins were then pooled and concentrated for ITC experiments or crystallization.

ITC. All of the synthetic peptides used for the ITC experiments were purchased from Genscript. For the ITC experiments, the peptides were weighed and dissolved in ITC buffer to make a stock peptide concentration of 2 mM. The ITC experiments were performed at 25 °C on a MicroCal PEAQ-ITC instrument. The proteins were buffer exchanged into ITC buffer by gel filtration before the ITC experiments. Transmembrane cargo peptides at 750 µM were titrated into 50 µM PX domain samples. The binding data were processed using the MicroCal PEAQ-ITC analysis software with a single site-binding model to determine the stoichiometry (n), binding constant (K_d), estimated heat of binding (ΔH), Gibbs free energy (ΔG) and binding entropy (ΔS). Three experiments were performed for each set of samples to determine the average ± s.e.m. for the thermodynamic quantities, except for the peptide mutations, SNX6 and SNX32 interaction experiments, where only single experiments were performed. For these single experiments, all experiments were performed using the same batch of protein to allow direct comparisons to be made.

Protein crystallization and structure determination. The SeMet-labelled SNX5–CI-MPR fusion protein was concentrated to 12 mg ml^{-1} for crystallization. Four 96-well crystallization hanging-drop screens were set up using a mosquito liquid handling robot (TTP, LabTech) at 20 °C in the UQ ROCX facility. After these 96-well screens had been set up, the trays were incubated at 20 °C in a Rockimager storage hotel (Formulatrix), where the drops were imaged at different time points for 21 d. Multiple hits were obtained and the best optimized conditions yielded two different crystal forms. Crystal form 1 (spacegroup $P2_12_12_1$) was produced using 0.1 M MES pH 6.5, 0.2 M ammonium sulphate, 30% (w/v) PEG5000 MME, 15% glycerol and 4% (v/v) polypropylene glycol P400. Crystal form 2 (spacegroup $C222_1$) was produced using 0.1 M MES pH 6.5, 0.2 M ammonium sulphate, 30% (w/v) PEG5000 MME, 15% glycerol and 3% (v/v) (±)-2-methyl-2,4-pentanediol. Both crystals were cryoprotected in mother liquor plus 10% glycerol before freezing in liquid nitrogen. Data were collected at the Australian Synchrotron MX2 Beamline, integrated with iMOSFLM[65] and scaled with AIMLESS[66] in the CCP4 suite[67]. The structures were initially solved by molecular replacement with PHASER[68] using the SNX5–IncE crystal structure as the input model (Protein Data Bank (PDB) accession code 5TGI)[27]. The anomalous signals from the Se atoms were used to unambiguously assign the CI-MPR Met2372 side chain. The resulting models were rebuilt with COOT[69], followed by repeated rounds of refinement with PHENIX[70].

The SNX5–SEMA4C fusion protein was concentrated to 12 mg ml^{-1} and crystals were grown in 25% PEG3350, 0.1 M Tris pH 8.0, 20 mM ammonium sulphate and 10 mM barium chloride dihydrate, cryoprotected in 25% glycerol and data were collected at the Australian Synchrotron MX1 Beamline. The structure was determined by molecular replacement using PHASER[68] and the resulting models were rebuilt with COOT[69], followed by repeated rounds of refinement with PHENIX[70]. All structural figures were generated using PyMOL (DeLano scientific).

Genetics. The following *Drosophila* mutant alleles were used: *slit*^{1}, *robo*^{5}, *snx1*$^{\Delta 2}$ and *snx5/6*$^{\Delta}$. The following transgenes were used: *UAS–FraΔC–HA* and *UAS–A5CD8–GFP*. The GAL4 driver used was eg-GAL4. All crosses were carried out at 25 °C. Embryos were genotyped using balancer chromosomes carrying lacZ markers or by the presence of epitope-tagged transgenes.

***Drosophila* embryo immunofluorescence and imaging.** Dechorionated, formaldehyde-fixed, methanol-devitellinized embryos were fluorescently stained as previously described[39]. The following primary antibodies were used: mouse anti-1D4/FasII (Developmental Studies Hybridoma Bank; 1:100), mouse anti-β-galactosidase (Developmental Studies Hybridoma Bank; 1:150), rabbit anti-GFP (Invitrogen, A11122; 1:500) and mouse anti-haemagglutinin (Covance, 16B12; 1:250). The following secondary antibodies were used: Alexa647-conjugated goat anti-horseradish peroxidase (Jackson Immunoresearch, 123-605-021; 1:500), cyanine 3-conjugated goat anti-mouse (Jackson; 1:1,000) and Alexa488-conjugated goat anti-rabbit (Molecular Probes; 1:500). The embryos were mounted in 70% glycerol in PBS. The phenotypes were analysed and images were acquired using a spinning disk confocal system (PerkinElmer) built on a Nikon Ti-U inverted microscope using a Nikon OFN25 ×60, ×40 or ×10 objective with a Hamamatsu C10600-10B CCD camera and Yokogawa CSU-10 scanner head with Volocity imaging software. The images were processed using ImageJ and Adobe Photoshop software.

Statistics and reproducibility. All of the quantified western-blot and confocal-colocalization data are the mean of at least three independent experiments. Statistical analyses were performed using Prism 7 (GraphPad Software). The graphs represent the means and s.e.m. For all statistical tests, $P < 0.05$ was considered significant and is indicated by asterisks.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The coordinates and structure factors for the SNX5 PX domain in complex with the CI-MPR sorting signal have been deposited at the PDB with accession codes 6N5X (form 1) and 6N5Y (form 2). The coordinates and structure factors for the SNX5 PX domain in complex with the SEMA4C sorting signal have been deposited at the PDB with the accession code 6N5Z. Supplementary Table 3 contains all of the raw data from the proteomic analysis shown in Fig. 7b,f. All of the other relevant raw data related to this study are available from the corresponding authors on request. The mass spectrometry data have been deposited in ProteomeXchange with the primary accession code PXD014927. The source data for Fig. 8a,d have been provided as Supplementary Table 3. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

References

64. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
65. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271–281 (2011).
66. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
67. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
68. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
69. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
70. Adams, P. D. et al. The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94–106 (2011).

nature research

Corresponding author(s): Peter Cullen / Brett Collins

Last updated by author(s): 5th August 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Software for confocal microscopy - LAS AF 2012.1.1. Quantification of western analysis - Odyssey Sa Software version 1.1 Xcalibur - version 2.0 Application Suite AF software - version 2.7.3.9723.
Data analysis	For Pearson's correlation coefficients - Volocity Version 6.3.0 Statistical analysis of data from westerns and confocal microscopy - GraphPad Prism 7 Statistical analysis of proteomic data sets - Perseus Version 1.6.1.3 Drosophila imaging - Image J and Adobe Photoshop Structural figures - PyMOL (DeLano scientific) MicroCal PEAQ-ITC analysis software For the proteomic analysis the raw data files were processed and quantified using Proteome Discoverer software v2.1 (Thermo Scientific) and searched against the UniProt Human database (downloaded 14-09-17: 140000 sequences) using the SEQUEST algorithm.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

X-ray structures have been deposited to the PDB with deposition IDs as follows: 6N5X; 6N5Y; 6N5Z. Raw proteomic data sets are included in an Excel spreadsheet as Supplementary Table 3 (raw proteomic data). Source data for Figs 3b, 3d, 3e, 3f, 3g, 4b, 4c, 4d, 4e, 4f, 4g, 5b, 5c, 5d, 5e, 6a, 6g, 6h 7c, 7d, 8b, 8i and 8l have been provided as Supplementary Table 4 (statistics source data). All unprocessed immunoblots are collated in Supplementary Figure 3. All other data supporting the findings of this study are available from the corresponding authors on reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size calculations were not preformed. All cell biology experiments were performed using at least 3 independent biological repeats. Analysis of cell-based phenotypes by confocal microscope was performed on in excess of 49 cells per experimental conditions. Quantification of Drosophila phenotypes was achieved from between 16 and 26 embryos (exact numbers are defined in the appropriate Figure Legend).
Data exclusions	No data was excluded from any of the analyses reported in this study.
Replication	All experiments were performed using at least 3 independent biological repeats. All experiments shown were reproducible. For ITC experiments three experiments were performed for each set of samples to determine the average ± standard error of the mean (SEM) for thermodynamic quantities, except for the peptide mutations, SNX6 and SNX32 interaction experiments where only single experiments were performed. For these single experiments, all experiments were performed using the same batch of protein to allow direct comparisons to be made.
Randomization	Experimental groups were assembled such that controls and positive/negative experimental conditions were generated, processed and analysed in parallel. For examples, the data shown in Figure 4D were collected from the same experimental group.
Blinding	Blinding was not used in this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Antibodies for mammalian experiments used in this study were: mouse monoclonal antibodies to SNX1 (clone 51/SNX1; 611482; BD) (1:1000 for WB, 1:200 for IF), SNX2 (clone 13/SNX2; 5345661; BD) (1:1000 for WB, 1:200 for IF), SNX6 (clone d-5,365965; Santa Cruz Biotechnology, Inc.) (1:1000 for WB, 1:200 for IF), β-actin (A1978; Sigma-Aldrich) (1:2000 for WB), GFP (clones 7.1 and 13.1; 11814460001; Roche) (1:2000 for WB, 1:400 for IF), CI-MPR (clone 2G11; MA1-066; Thermo Fisher Scientific) (1:1000 for WB, 1:100 for IF); Golgin-97 (CDF4) (A-21270, Thermo Fisher Scientifics) (1:200 for IF), rabbit polyclonal antibodies to

mCherry (ab167453, Abcam) (1:2000 for WB, 1:400 for IF), SNX1 (ab995; Abcam) (1:1000 for WB), SNX5 (17918-1-AP; Proteintech) (1:1000 for WB), GAPDH (G9545; Sigma-Aldrich) (1:1000 for WB), AP-2 (11401-1-AP; Proteintech) (1:1000 for WB), AP-1 (13258-1-AP; Proteintech) (1:1000 for WB), SEMA4C (ab171559; Abcam) (1:1000 for WB), TMEM230 (21466-1-AP; Proteintech) (1:1000 for WB), ROBO1 (20219-1-AP; Proteintech) (1:1000 for WB), GPR50 (19762-1-AP; Proteintech) (1:1000 for WB); rabbit monoclonal antibodies to SNX5 (EPR14358; 180520; Abcam) (1:1000 for WB), CI-MPR (EPR6599; 124767; Abcam) (1:1000 for WB, 1:200 for IF), IGF1R (clone D23H3; 111A9; Cell Signalling Technologies) (1:1000 for WB), EEA1 (clone c45b10; 3288; Cell Signaling Technology) (1:200 for IF); goat polyclonal antibodies to EEA1 (N- 19; sc-6415; Santa Cruz Biotechnology, Inc.); sheep polyclonal antibody to TGN46 (AHP500G; Bio-Rad Laboratories) (1:200 for IF).

Antibodies for Drosophila experiments used in this study were: mouse anti-1D4/FasII [Developmental Studies Hybridoma Bank (DSHB); 1:100], mouse anti-Beta gal [DSHB; 1:150], rabbit anti-GFP [Invitrogen (#A11122); 1:500], Mouse anti-HA [Covance (16B12) 1:250]. The following secondary antibodies were used: Alexa647- conjugated goat anti- HRP [Jackson Immunoresearch (#123-605-021); 1:500]. Cyanine 3-conjugated goat anti-mouse [Jackson; 1:1000], Alexa488-conjugated goat anti-rabbit [Molecular Probes; 1:500].

Validation

Besides the technical validation from the manufacturers we independently validated the key antibodies used in this study namely those against SNX1, SNX2, SNX5, SNX6, CI-MPR and IGF1-R antibodies. This validation was achieved through comparative and quantitative western analysis and immunofluorescent analysis of wild-type and CRISPR Cas9 gene-edited cells targeting the corresponding protein.

For the other antibodies validation was either through the manufacturers validation sheet (see detailed information above for the precise manufacturer and individual antibody ID) or published validation by other research groups (see reference list associated with the manuscript).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

HeLa, HEK-293T and RPE1 cell lines were originally sourced from ATCC.

Authentication

Authentication was from the ATCC. We have not independently authenticated the cell lines.

Mycoplasma contamination

The parental HeLa cells used for generation of all CRISPR edit cell lines were mycoplasma tested, which revealed no contamination. Our HEK-293T cell line has also been negatively tested for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

No laboratory animals were used in this study.

Wild animals

This study did not use wild animals.

Field-collected samples

This study did not use field-collected samples.

Ethics oversight

No ethical approval or guidance was required for this study as it did not involve the use of animals or human tissue samples.

Note that full information on the approval of the study protocol must also be provided in the manuscript.