Supporting Information

Shi and Blobel 10.1073/pnas.1013038107

SI Materials and Methods

Protein Expression, Purification, and Crystallization. The cDNA encoding full-length, mutant, and various C-terminal truncated She4p was cloned into modified pGEX-KG vector between NdeI and XhoI sites. *Escherichia coli* BL21 (DE3) Gold carrying expression plasmid was grown in LB media at 37 °C to OD_{600} = 0.5 and was induced by adding IPTG to 1 mM final concentration at 23 °C for overnight incubation. GST fusion protein was first purified by glutathione sepharose and cleaved by thrombin in buffer containing 10 mM Hepes pH 8.0, 200 mM NaCl, 10 mM DTT. Eluted protein was further purified by ion exchange (5 mL HiTrap Q FF) and gel filtration (Superdex 75 16/60) (GE Healthcare). Fractions were pooled and concentrated to 8 mg/mL in 10 mM Tris pH 8.0, 300 mM NaCl, 10 mM DTT and 1 mM $CaCl_2$ for crystallization. Diffraction quality crystals used for native data collection were obtained using buffer containing 200 mM Na citrate, 20% (wt/vol) PEG 3350, 10 mM DTT and crystallization was carried out at 12 °C. Se-Met substituted protein was produced as previously described (1), purified, and crystallized as above.

Data Collection and Structure Determination. Data were collected at 98 K at the beamline X-12C and X-29, National Synchrotron Light Source and processed with HKL2000 (HKL Research, Inc.). The structure was determined using three wavelengths multiwavelength anomalous dispersion (MAD) datasets from an Se-Met crystal. Two molecules were found in one asymmetric unit, and 18 ordered Se sites were identified using SHELXD (2). No noncrystallographic symmetry was used for structural determination. Subsequent phasing and density modification were carried out using SHELXE (2) and RESOLVE (3). Iterative cycles of density improvement, model building, and refinement were carried out using the Coot (4), CNS (5), and CCP4 (6) program packages. In CNS, 8.8% of diffractions were used for cross-validation. The final model includes a segment that can only be modeled as poly Ala due to poor density quality (residue 410–415). Additional unmodeled residues are shown in Fig. S1. Only 145 water molecules were modeled into the density due to low resolution (note: the average B factor of water molecules is much lower than that of the protein atoms). High B-factors across the entire protein and slightly high R_{free}, R_{work}, and their difference (6.5%) are likely due to the high flexibility typical of helical repeat proteins (7) and a significant percentage (4%) of the protein in the asymmetric unit that remains unmodeled. Sixteen outliers in Ramachandran plot locate to regions with weak density. Figures were prepared by PyMOL and the electrostatic potentials of the molecular surface were calculated by APBS (8).

Pull-Down Assays. Fragments of motor domain of Myo4p were cloned into modified pGEX-KG vector and expressed in BL21 (DE3) Gold strain. Expression condition is described above. GST-Myo4p fragments were immobilized on 50 μL of glutathione sepharose beads and washed with binding buffer containing 10 mM Hepes, pH 8.0, 150 mM NaCl, 1 mM DTT. One hundred micrograms of purified She4p or truncated She4p 1–544 were incubated with beads and washed twice with 500 μL binding buffer. Proteins were eluted by SDS sample buffer and analyzed by SDS-PAGE and subsequent staining with Coomassie blue.

Multiangle Light Scattering. Purified She4p WT was characterized by multiangle light scattering following size-exclusion chromatography. Fifty microliters of 2 mg/mL protein was injected onto a Superdex 200 10/300 GL size-exclusion chromatography column (GE Healthcare) equilibrated with buffer containing 10 mM Hepes pH 8.0, 150 mM NaCl, and 1 mM DTT. The chromatography system was connected to an 18-angle light scattering detector (DAWN HELEOS) and refractive index detector (Optilab rEX, Wyatt Technology). Data were collected every 1 s at a flow rate of 0.25 mL/ min at 25 °C. Data analysis was carried out using the program ASTRA.

Isothermal Titration Calorimetry. Affinity of a synthetic peptide of Myo4p (residues 561–587) for She4p was determined by isothermal titration calorimetry. Proteins were diluted into 10 mM Hepes, pH 8.0, 150 mM NaCl. The chamber contained 1.7 mL 10 μM She4p and was kept at 25 °C. The concentration of the synthetic Myo4p peptide was 200 μM. First one 5-μL aliquot was injected followed by the 29 consecutive injections of 10-μL aliquots in about 3 min intervals. Data were plotted and analyzed with a single site binding model using MicroCal Origin software.

Yeast Plasmids Construction. To generate the yeast expression vector p425Gal1-2FLAG, the original p425Gal1 vector (9) was digested with BamHI and XhoI and ligated with an annealed primer set of double-FLAG-fwd and double-FLAG-rev (Table S1). This procedure yielded two consecutive N-terminal FLAG tags and new cloning sites in the vector p425Gal1-2FLAG. All WT or mutant*SHE4* was cloned into p425Gal1-2FLAG using NdeI and XhoI sites. All constructs were verified by sequencing.

Yeast Strain Construction. *Saccharomyces cerevisiae* strain BY4741 was used in the analysis of She4Mp function. Yeast strains were maintained in yeast-peptone-dextrose plates or grown in liquid medium before transformation. Yeast strain BY4741 *she4*Δ::*KanMX* from the deletion library (Invitrogen Corporation) was transformed with *EGFP*::*HIS5* cassette amplified from pFA6a vector with primers M4-C and M4-C-rev to create a C-terminal EGFP fusion *MYO4-EGFP*::*HIS5* cassette. Transformation was performed by the lithium acetate procedure, as previously described (10), and yielded strain BY4741 *she4*Δ::*KanMX, MYO4-EGFP*::*HIS5*. Flag-tagged WT and mutant *SHE4* was amplified from plasmids described above using primers S4-FLAG-LEU and S4-LEU-rev, carrying LEU selection marker and integrated into the strain above for imaging analysis. Colonies grew on LEU selection medium but not on KAN selection medium. All integrations were confirmed by sequencing. All primer sequences used for plasmid construction and integration are listed in Table S1.

Imaging. Yeast strains were grown at 30 °C in synthetic complete medium containing 2% glucose until stationary phase. These cultures were aliquoted into the same medium to OD_{600} of 0.1 and grown at 30 °C until OD_{600} reached approximately 0.6. Cells from 1 mL culture were twice centrifuged and resuspended in 1 mL medium. Finally, cells were resuspended into 50 μL medium. Two microliters of each resuspension solution was used to prepare slides for imaging analysis. Images were taken using a Zeiss AXZ1 microscope and figures were prepared with Adobe Photoshop.

1. Shi H, Rojas R, Bonifacino JS, Hurley JH (2006) The retromer subunit Vps26 has an arrestin fold and binds Vps35 through its C-terminal domain. *Nat Struct Mol Biol* 13:540–548.

2. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr, Sect A: Found Crystallogr* 64:112–122.

3. Terwilliger T (2004) SOLVE and RESOLVE: Automated structure solution, density modification, and model building. *J Synchrotron Radiat* 11:49–52.
4. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr, Sect D: Biol Crystallogr* 60:2126–2132.
5. Brunger AT, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr, Sect D: Biol Crystallogr* 54:905–921.
6. Anonymous (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr, Sect D: Biol Crystallogr* 50:760–763.
7. Conti E, Muller CW, Stewart M (2006) Karyopherin flexibility in nucleocytoplasmic transport. *Curr Opin Struct Biol* 16:237–244.
8. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037–10041.
9. Mumberg D, Muller R, Funk M (1994) Regulatable promoters of *Saccharomyces cerevisiae*: Comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res* 22:5767–5768.
10. Gietz RD, Schiestl RH (2007) Quick and easy yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2:35–37.
11. Kippert F, Gerloff DL (2009) Highly sensitive detection of individual HEAT and ARM repeats with HHpred and COACH. *PLoS One* 4:e7148.
12. Wong KC, Naqvi NI, Iino Y, Yamamoto M, Balasubramanian MK (2000) Fission yeast Rng3p: An UCS-domain protein that mediates myosin II assembly during cytokinesis. *J Cell Sci* 113:2421–2432.
13. Barral JM, Bauer CC, Ortiz I, Epstein HF (1998) Unc-45 mutations in *Caenorhabditis elegans* implicate a CRO1/She4p-like domain in myosin assembly. *J Cell Biol* 143:1215–1225.
14. Hodge T , Cope MJ (2000) A myosin family tree. *J Cell Sci* 113:3353–3354.

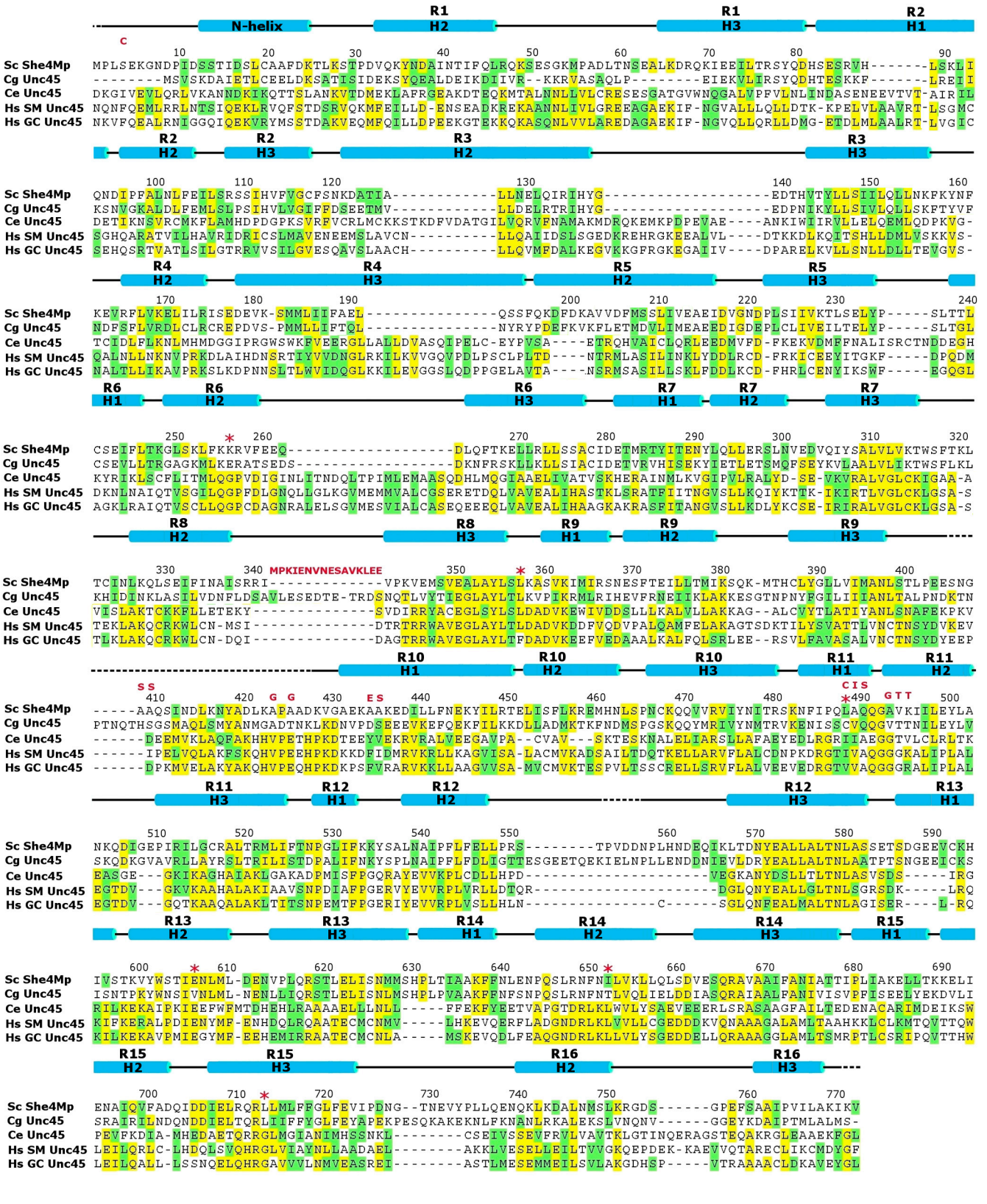

Fig. S1. Sequence comparison between UCS proteins. The residues of WT She4p that have been deleted or altered with respect to a mutant, She4Mp, are indicated by red letters. Aligned with She4Mp were selected fungal and animal homologues using the program Clustal W. Sc, *Saccharomyces cerevisiae*; Cg, *Candida glabrata*; Ce, *Caenorhabditis elegans*; Hs, *Homo sapiens*, with the two isoforms in smooth muscle (SM) cells and in general cells (GC) and manually edited using structural information. Identical residues are colored in yellow, similar residues are colored in green. Secondary structures, based on the coordinates of molecule *A*, are shown above the alignment as alternating loops (shown in straight lines) and helices (shown in cylinders colored blue). Helices are arranged in helical repeats numbered R1–R16 except N helix. Repeats R9–R16 carry conserved residues for classical Armadillo (ARM) repeats. Repeats R1–R8 share similar hydrophobic pattern to ARM/Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and the yeast PI3-kinase TOR1 (HEAT) repeats but are less conserved among UCS proteins. We therefore group them into diverged ARM repeats (11). Several disordered regions are shown in dashed lines.

Previously identified temperature sensitive mutations (12, 13) are labeled with red asterisks (*). In order to improve the diffraction resolution for crystals, we consecutively introduced a series of mutations into She4p. First, we deleted a flexible loop (residues 343–358 in WT She4p). Second, we replaced Cys_4 with Ser. Third, we replaced several flexible residues (Gly, Ser, and Glu) in the region of residues 420–450 in WT She4p with Ala. Fourth, we replaced two tripeptide regions between residues 505 and 512 in WT She4p. These two regions were identified as sites of degradation by mass spectrometry and were replaced by residues that occur in these positions in the UCS protein of *Aspergillus fumigates* (gene ID: 159126781).

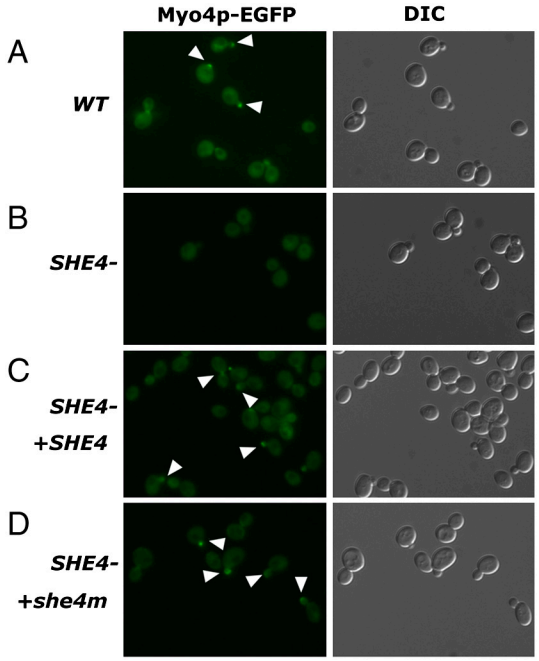

Fig. S2. She4Mp functionally substitutes for She4p in *S. cerevisiae*. (*A*) In WT cells , a C-terminal EGFP fusion to Myo4p (Myo4p-EGFP) localizes to the tip and the neck of the bud where it forms a bright dot. (*B*) Deletion of She4p results in the loss of spot-like localization of Myo4p-EGFP localization. (*C*) Genomic replacement by *SHE4* restores spot-like localization of Myo4p-EGFP. (*D*) Genomic replacement by *she4m* restores the localization of Myo4p-EGFP indicating that it can functionally substitute for WT She4p.

Fig. S3. The She4p binding site is highly conserved in myosin subfamilies. Alignment of the She4p binding sites among the various myosin subfamilies (I–XVII) (14). Identical and similar residues are shown in yellow and green, respectively. The characteristic beta strands and the helix-loop motif are shown at the top in orange.

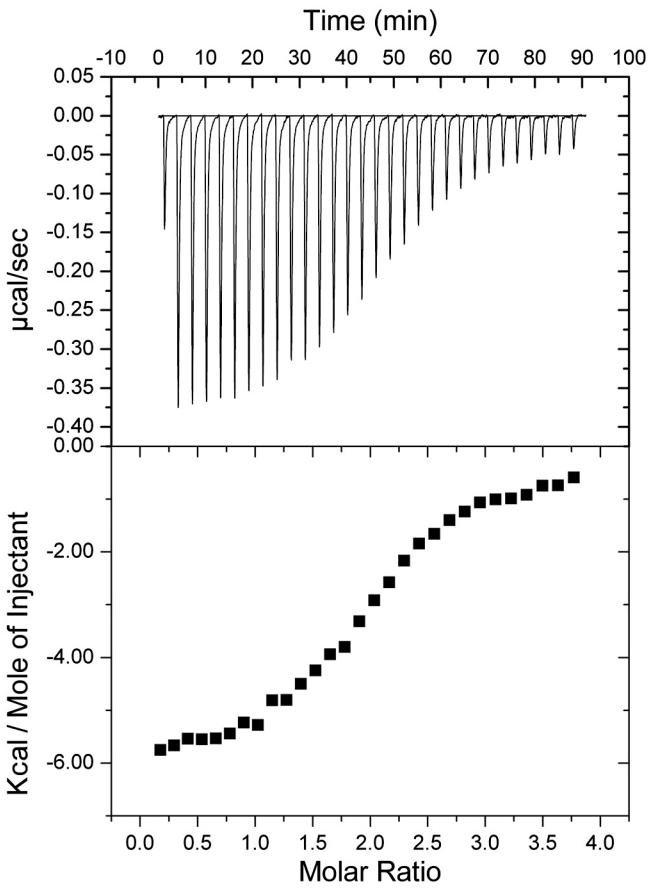

Fig. S4. Isothermal titration calorimetry. She4p and a synthetic peptide of *S. cerevisiae* Myo4p (residues 561–587) were incubated in a calorimeter to determine their affinity by fitting these data to a single binding site model.

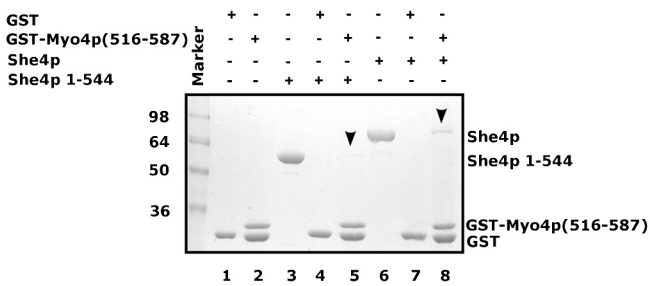

Fig. S5. C-terminal region of She4p mediates the interaction with Myo4p. GST and GST fusion Myo4p epitope (residue 561–587) were immobilized on glutathione sepharose beads (lanes 1 and 2), and incubated with purified N-terminal fragment of She4p (residues 1–544) and WT She4p, respectively (lanes 3 and 6). Neither form of She4p interacts with the GST alone (lanes 4 and 7). The binding of N-terminal She4p 1–544 is significantly reduced compared to that of full-length She4p (lanes 5 vs. 8).

Table S1. Primers used in yeast vector construction and integration

Name	Sequence
double-FLAG-fwd	GATCCCCGGGAATGCGCGATTACAAGGATGACGACGATAAGCGC GATTACAAGGATGACGACGATAAGCATATGGCCATGGGTCGAC
double-FLAG-rev	TCGAGTCGACCCATGGCCATATGCTTATCGTCGTCATCCTTGTA ATCGCGCTTATCGTCGTCATCCTTGTAATCGCGCATTCCCGGG
M4-C	AGCAATACAGAGGGCTTAGCTACTGTCAGTAAAATTATAAAATTAG ACAGAAAAGGCGGTGGCGGTGGCGGTGAAGCTCAAAAACTTAAT
M4-C-rev	TAGTAATCATCATATTGCTTGTTGTAATCATCGTCCCCATTGCTGTT CACTTTAGTCGATGCTGACGGTATCGATAAGCTT
S4-FLAG-LEU	TCTTCTACTTAATAGTGAAGACGTTTTATCTAAACTACTCACATG AAAAGATTACTAAAAAATTAGAATCACGGAACTAGTGGATCCCCCGGGA
S4-LEU-rev	TCTTGTGTGTTCATTATCAGGGTGCTGTTGTATTTCTTAGTGAAA ATTAAGAAGAGTTAAACATGGCGATTCTTTCCATATTAAGCAA GGATTTTCTTAACTTCTTCG