Supporting Information

Chartron et al. 10.1073/pnas.1006036107

SI Text

SI Methods. ***Cloning, expression, and purification of proteins.*** The sequences of Get4 (YOR164c) and Get5 (YOL111c) were obtained from the *Saccharomyces* Genome Database, codon optimized for expression in *Escherichia coli* using DNAWorks and synthesized by PCR (1). Residues 1–9 of Get4 were truncated to Met-Gly due to lack of conservation and predicted disorder (Fig. S1*A*). The genes were inserted sequentially into a pET33b(+)-derived vector (Novagen) and separated by an internal ribosome binding site. The C-terminus of Get4 was fused to a tobacco etch virus TEV protease cleavage site followed by a hexa-histidine tag. The proteins were overexpressed in BL21(DE3) (Novagen) grown in 2xYT media for 3 h at 37 °C after induction with 0.3 mM IPTG. Cells were lysed using an M-110L pneumatic microfluidizer (Microfluidics) and purified as a complex by Ni-affinity chromatography (Qiagen). The affinity tag was removed by a 3 h TEV protease digest at room temperature while dialyzing against 20 mM Tris pH 8.0, 30 mM NaCl and 12 mM β-mercaptoethanol. The sample separated into two peaks using a 6 mL Resource Q anion exchange column (GE Healthcare). Each peak was further purified using a Superdex 200 16/60 size exclusion column (GE Healthcare) equilibrated with 5 mM Tris pH 8.0, 300 mM NaCl and 12 mM β-mercaptoethanol and concentrated to 10 mg/mL.

The *Aspergillus fumigatus* homologs of Get4 (NCBI sequence XP_747572) and Get5 (NCBI sequence XP_748165) were cloned, expressed, and purified in the same way as the *Saccharomyces cerevisiae* proteins. Get5 started from the second methionine in the Genbank sequence, omitting the first 42 annotated residues, as they were not consistent with our multiple sequence alignments. The sequence used in this study is shown in Fig. S1*B*. Expression and purification of Get3 were performed as previously described (2).

Mutations and truncations of Get3, Get4, or Get5 were introduced using the Quikchange method (Stratagene). All mutants were expressed under identical conditions as the wild type Get4/5, purified by Ni-affinity chromatography and cleaved by TEV protease where indicated. Analytical size exclusion chromatography was performed using either a Superdex 200 10/300 column with a QuadTec UV detector (Biorad) or a Shodex KW 803 column with a Dawn Heleos MALS detector (Wyatt Technology).

Formation and detection of complexes between Get3 and Get4/5. Complexes were formed by incubating 800 pmol Get3 and 400 pmol Get4/5 for 2 h at room temperature in 200 μL of 50 mM HEPES pH 7.3, 100 mM NaCl, 2 mM $MgCl_2$, 30 mM imidazole and, where indicated, 2 mM ADP or 2 mM ATP. After incubation, the reaction was added to 10 μL of gravity settled Ni-NTA agarose beads and mixed. The beads were washed twice with 100 μL of the incubation buffer within 1 minute. Bound proteins were eluted with 25 μL of incubation buffer containing 300 mM imidazole.

Limited proteolysis and crystallization of Get4/5. Crystallization screening was performed using a TTP LabTech Mosquito robot and commercially purchased kits (Hampton Research, Qiagen, Molecular Dimensions Limited). Adding TLCK treated α-chymotrypsin (Sigma) at a 1 : 125 mass ratio to either oligomeric state immediately prior to screening, as described (3), resulted in crystals growing as hexaganol rods within a day at room temperature. Limited proteolysis was performed at room temperature for 15 minutes or 60 minutes on ice and stopped by the addition of 0.1 mM PMSF. Six products were separated by a Resource Q column, and only the major product, Get4/5-N, crystallized after concentration to 5 mg/mL. Sequences of the proteolytic fragments were determined using LC/MS with tryptic digestion and/or N-terminal sequencing. After refinement of the condition, crystals approximately 150 μm in length grew after 3 days at room temperature, although the majority had a twist approximately half way down their length and were unable to diffract (Fig. S2*D*).

A small molecule additive screen (Hampton Research) revealed that the addition of spermidine aggravated twisting while 10 mM L-proline, trimethylamine-HCL or sarcosine yielded three-sided single crystals (Fig. S2*D*). Diffracting crystals were grown using the hanging drop vapor diffusion method where 1 μL of Get4/5-N at 5 mg/mL in 20 mM Tris pH 8.0, 90 mM NaCl, and 12 mM β-mercaptoethanol was mixed with 1 μL of a reservoir of 17% PEG 6000, 0.14 M ammonium sulfate, 0.1 M Bis-tris pH 5.5, and 10 mM L-proline. Crystals were transferred into 20 μL of reservoir, cryoprotected by repeatedly removing and adding reservoir solution supplemented with 1% increments of glycerol to 10%, and flash frozen in liquid nitrogen. Selenomethionine derived protein was expressed following established methods (4) and purified, crystallized and frozen as described for the native protein.

Data collection, structure solution, and refinement. A single wavelength anomalous dataset to 2.8 Å resolution was collected on beam line 9-2 at the Stanford Synchrotron Radiation Laboratory (SSRL) (Table S1). Data were integrated and scaled using XDS (5), the space group was determined by POINTLESS and the data were merged and converted to amplitudes with SCALA/CTRUNCATE (6). The selenium substructure was determined using SHELXD (7) and initial phases, density modification, and model building were performed by PHASER (8) and RESOLVE (9) as implemented automatically by PHENIX (10). The asymmetric unit consisted of three copies of Get4/5-N arranged along a distorted 3-fold axis. Chain E was the most complete copy of Get4, with continuous electron density observed for residues 9–299. Similarly, chain F was the most complete copy of Get5-N consisting of residues 3–56. Manual rebuilding was performed using COOT (11). Difference densities in two copies near the N-terminal poles of α3 of Get4 (residues 49–52) that were too large for ordered waters were modeled as L-proline. Reciprocal space refinement with local NCS restraints was performed using REFMAC v5.6 (12) using the direct single wavelength anomalous dispersion (SAD) target (13). A native dataset of comparable quality was collected on beam line 11-1 at SSRL, however refinement against the SAD data resulted in a higher quality model (determined by R-factors and model geometry) and thus the selenomethionine structure was used for deposition into the PDB (ID code 3LKU). The final model had an R_{crys} of 18.2% and an R_{free} of 22.4% with residues in the Ramachandran plot in 97.9% preferred, 2.1% allowed, and 0.0% in the disallowed and restricted regions [COOT (11)]. Fig. 1 *A–C*, Fig. 2, Fig. 4 *D* and *E*, Fig. 5*B*, Fig. S3 and Fig. S8 were prepared using Pymol (14). Figs. S4 and S9 were prepared using UCSF Chimera (15, 16).

Yeast growth assays. Knockout strains BY4741 *YOR164C::KanMX* and BY4741 *YOL111c::KanMX* were purchased from America Type Culture Collection (ATCC) (17). The Get4 rescue plasmid was constructed by PCR amplifying the open reading frame with 339 bp upstream and 87 bp downstream flanking regions from

BY4741 genomic DNA and inserting into the YEp-352 shuttle vector (18). Similarly, the Get5 open reading frame was amplified with 506 bp upstream and 69 bp downstream flanking sequence. Mutations were then introduced with the Quikchange method. The parent BY4741 and deletion strains were transformed using the LiAc/single-stranded carrier DNA/PEG method (19).

Phenotypic rescue was determined by growing each transformant in SC-Ura media at 30 °C to an $OD_{600\,nm}$ between 1 and 2, diluting to 3.85×10^6 cells/mL and spotting 4 μL of serial dilutions onto SC -Ura agar plates in the presence or absence of 2 mM $CuSO_4$. Plates were then incubated at 37 °C for 20 h and photographed.

1. Hoover DM, Lubkowski J (2002) DNAWorks: An automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucl Acids Res* 30(10):e43.
2. Suloway CJ, Chartron JW, Zaslaver M, Clemons WM, Jr. (2009) Model for eukaryotic tail-anchored protein binding based on the structure of Get3. *Proc Natl Acad Sci U SA* 106(35):14849–14854.
3. Dong A, et al. (2007) In situ proteolysis for protein crystallization and structure determination. *Nat Methods* 4(12):1019–1021.
4. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* 229(1):105–124.
5. Kabsch W (2010) XDS. *Acta Crystallogr D* 66:125–132.
6. Collaborative Computation Project No. 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50(5):760–763.
7. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* 64(1):112–122.
8. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(4): 658–674.
9. Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D* 56(8):965–972.
10. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* 66:213–221.
11. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of COOT. *Acta Crystallogr D* 66:486–501.
12. Winn MD, Murshudov GN, Papiz MZ (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Method Enzymol* 374:300–321.
13. Skubak P, Murshudov GN, Pannu NS (2004) Direct incorporation of experimental phase information in model refinement. *Acta Crystallogr D* 60(12):2196–2201.
14. Delano WL (1998) *The PyMol Molecular Graphics System* (DeLano Scientific, Palo Alto, CA).
15. Pettersen EF, et al. (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
16. Sanner MF, Olson AJ, Spehner JC (1996) Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers* 38(3):305–320.
17. Winzeler EA, et al. (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285(5429):901–906.
18. Hill JE, Myers AM, Koerner TJ, Tzagoloff A (1986) Yeast/*E. coli* shuttle vectors with multiple unique restriction sites. *Yeast* 2(3):163–167.
19. Gietz RD, Schiestl RH (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2(1):31–34.
20. Larkin MA, et al. (2007) Clustal W and clustal X version 2.0. *Bioinformatics* 23(21): 2947–2948.

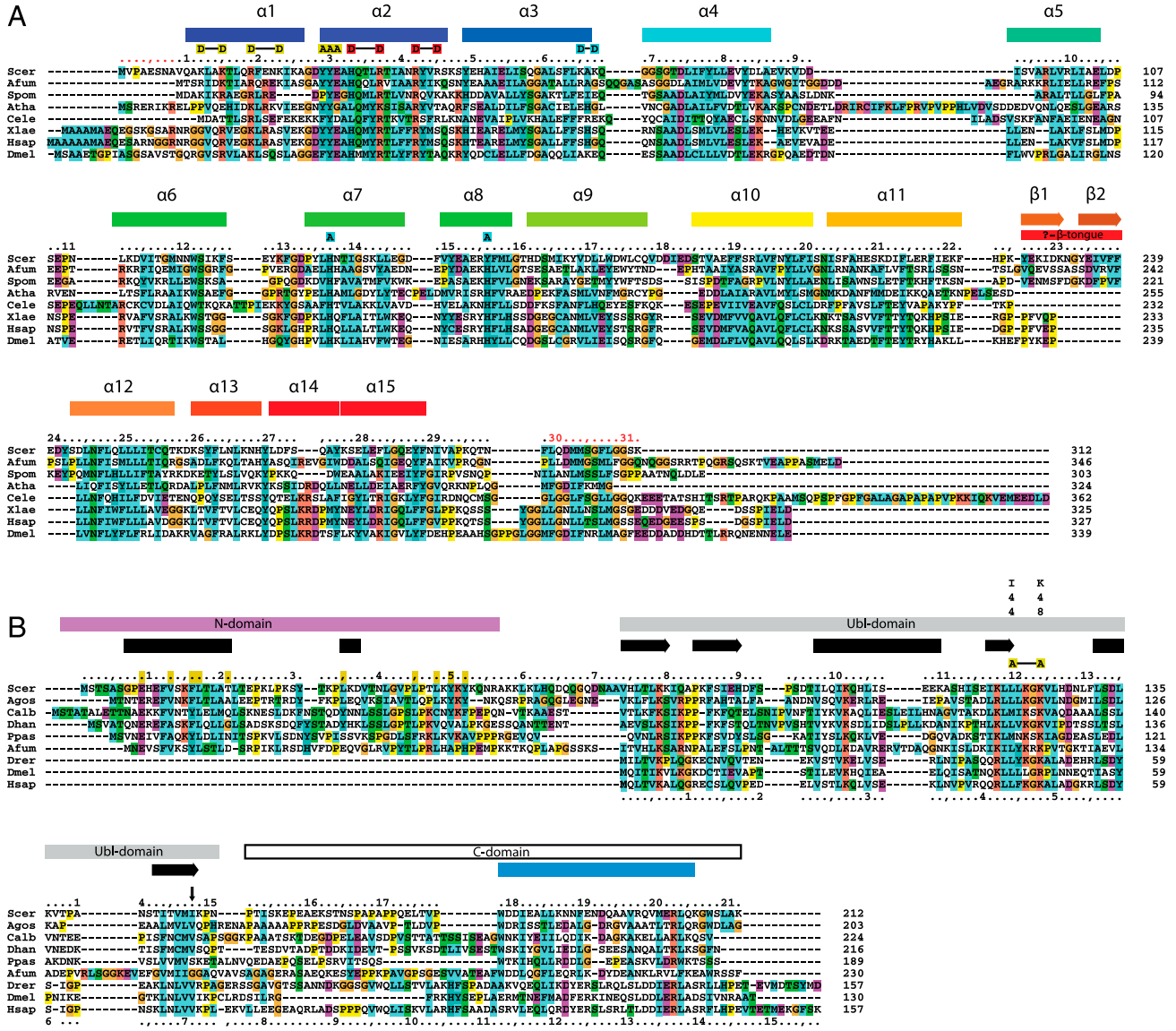

Fig. S1. Sequence alignments of Get4 and Get5. (*A*) Sequences of Get4 were aligned using ClustalX (20). Residue coloring is based on the program output (amino acid type). Species in order are Scer (*Saccharomyces cerevisiae*), Afum (*Aspergillus fumigatus*), Spom (*Schizosaccharomyces pombe*), Atha (*Arabidopsis thaliana*), Cele (*Caenorhabditis elegans*), Xlae (*Xenopus laevis*), Hsap (*Homo sapiens*), and Dmel (*Drosophila melanogaster*). Numbering above the sequences is based on *S. cerevisiae*, with residues observed in the crystal structure in black and residues absent in red. Mutants generated for this study are indicated above the numbering, with double mutants connected with bars, and are highlighted based on wild type (cyan), weak (yellow) or weakest (red) interactions with Get3. Secondary structure is shown on top with helices indicated by boxes and β-strands by arrows color ramped to match Fig. 1*B*. (*B*) Alignment of Get5. Species in order are Scer (*Saccharomyces cerevisiae*), Agos (*Ashbya gossypii*), Calb (*Candida albicans*), Dhan (*Debaryomyces hansenii*), Ppas (*Pichia pastoris*), Afum (*Aspergillus fumigatus*), Drer (*Danio rerio*), Dmel (*Drosophila melanogaster*), and Hsap (*Homo sapiens*). Secondary structure in the N-domain is from the crystal structure, whereas secondary structure of the Ubl-domain is based on the NMR structure of human Ubl4A (PDB ID code 2DZI). Residues 179–205 are predicted to be helical and are colored cyan. The arrow above residue 148 indicates the start of the C-domain peptide observed during purications. I44 and K48 labeled vertically above the Ubl-domain designation refer to residues in ubiquitin specically addressed in the text. The L120A/K120A mutant is indicated in yellow representing a moderate growth defect.

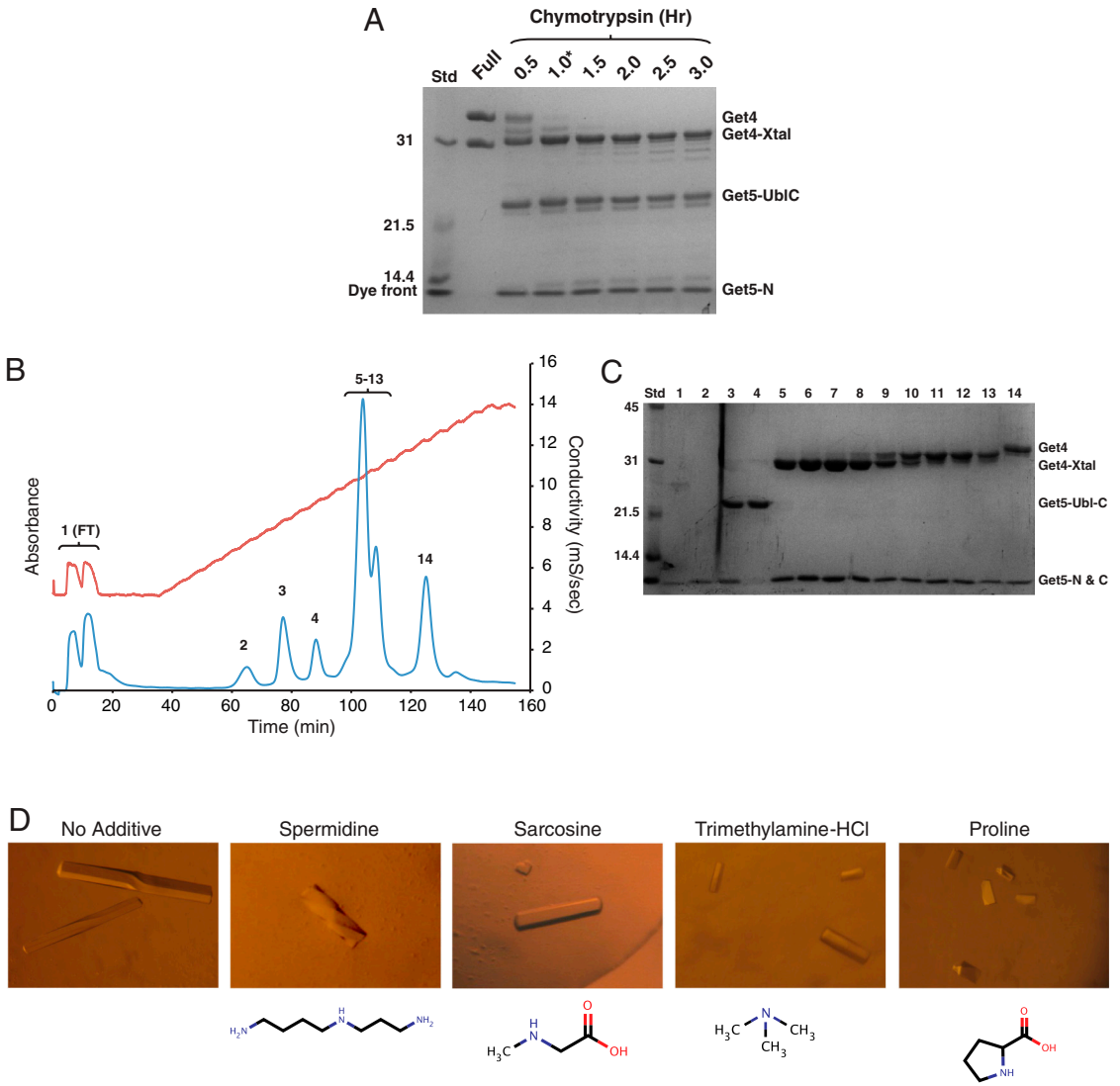

Fig. S2. Proteolysis, purication, and crystallization. (*A*) Chymotrypsin is added to Ni-NTA anity puried Get4/5 on ice with aliquots removed at the indicated time points in hours and separated by SDS-PAGE. "Full" represents dimeric Get4/5 prior to the addition of enzyme. The identities of the Get4-Xtal, Get5-Ubl-C and Get5-N bands were confirmed by tryptic digest followed by LC/MS. The 1 h time point, indicated with an asterisk, represents the degree of digestion used for further purication (*B*) Anion exchange chromatogram of Get4/5 after digest by chymotrypsin using a 30–130 mM NaCl gradient. Conductivity is plotted as the red trace and A280 as the blue trace. The numbers above peaks represent lanes in *C*. (*D*) Representative crystals of Get4/5-N. Crystals are grown in 17% PEG 6000, 0.14 M ammonium sulfate, 0.1 M Bis-tris pH 5.5 with the following additives: None, 10 mM spermidine, 10 mM sarcosine, 10 mM Trimethylamine-HCL, or 10 mM L-proline.

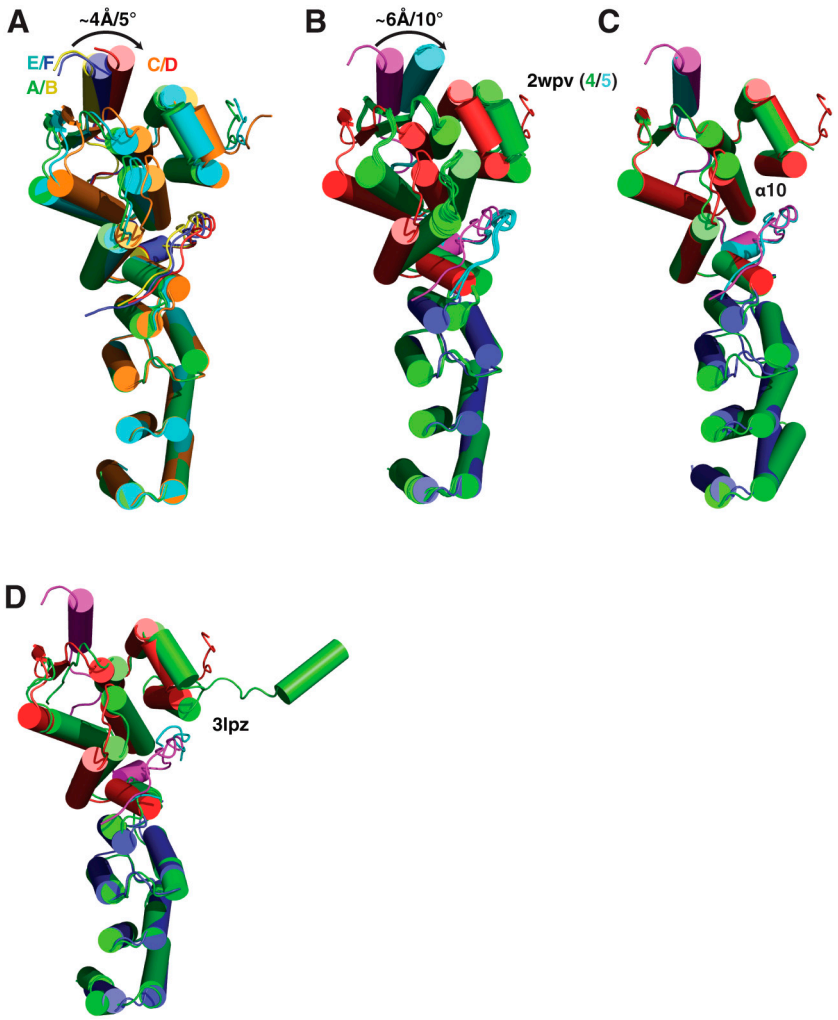

Fig. S3. Comparison to other crystal structures. All views are from the bottom of Get4/5-N relative to Fig. 1*A*. (*A*) Aligning the three molecules in the asymmetric unit relative to the first six helices in molecule A. The molecules are colored the same as in Fig. 1*C*. The rotation relative to the domains is indicated. (*B*) Similar to (*A*) with the A/B molecule colored N-domain (blue), C-domain (red), and Get5-N (magenta). The four molecules in the asymmetric unit of Chang et al (1 wpv) are aligned to the N-domain as in (*A*) colored by Get4 (green) and Get5-N (cyan). (*C*) Similar to (*B*) but molecules 1 wpv-A/B are aligned to the C-domain. The shifted α10 is indicated. (*D*) The Bozkurt et al. (1 lpz) Get4 (green) aligned to the A molecule N-domain. Included are a C-terminal peptide seen in that crystal structure that docks into the Get5-N binding groove (cyan).

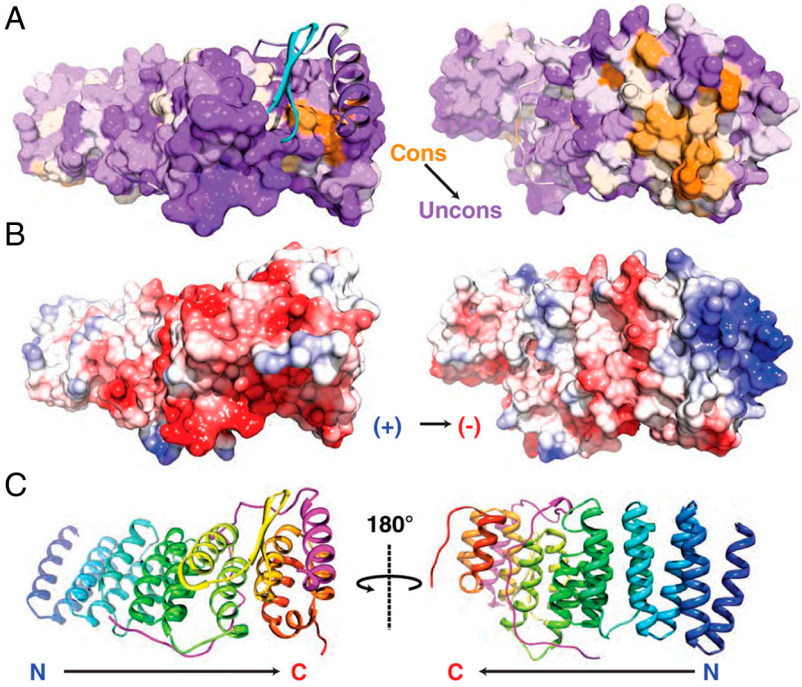

Fig. S4. Surface conservation and charge. (*A*) Get4 shown as an accessible surface and Get5-N as a ribbon colored from most (orange) to least (purple) conserved. The fungal specific β-tongue is shown as a cyan ribbon. (*B*) As in *A* with the total accessible surface colored by Coulombic charge. (*C*) Ribbons diagram in the same orientations for reference.

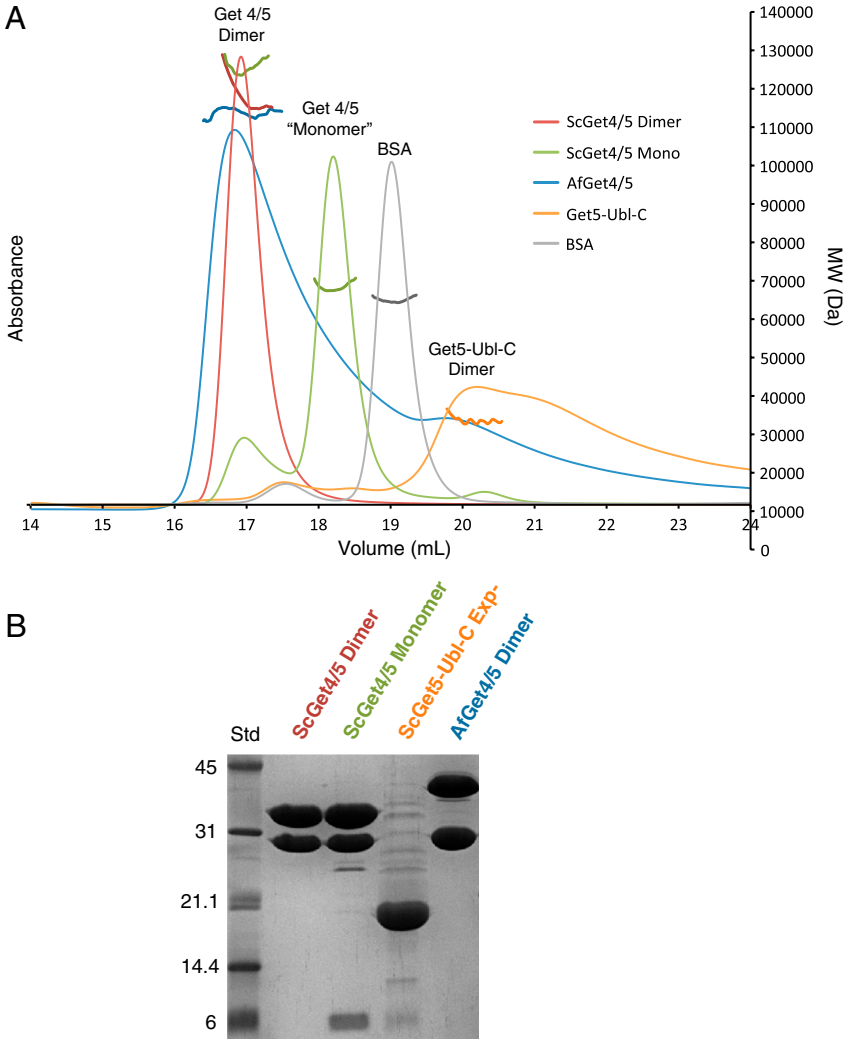

Fig. S5. Multiangle light scattering. (*A*) Overlay of size exclusion chromatograms using a Shodex KW 803 column. Traces represent UV absorbance (left axis). For each peak, molecular weight as determined using multiangle light scattering and refractive index is plotted (right axis). (*B*) Coomassie-stained SDS-PAGE of samples used in *A*.

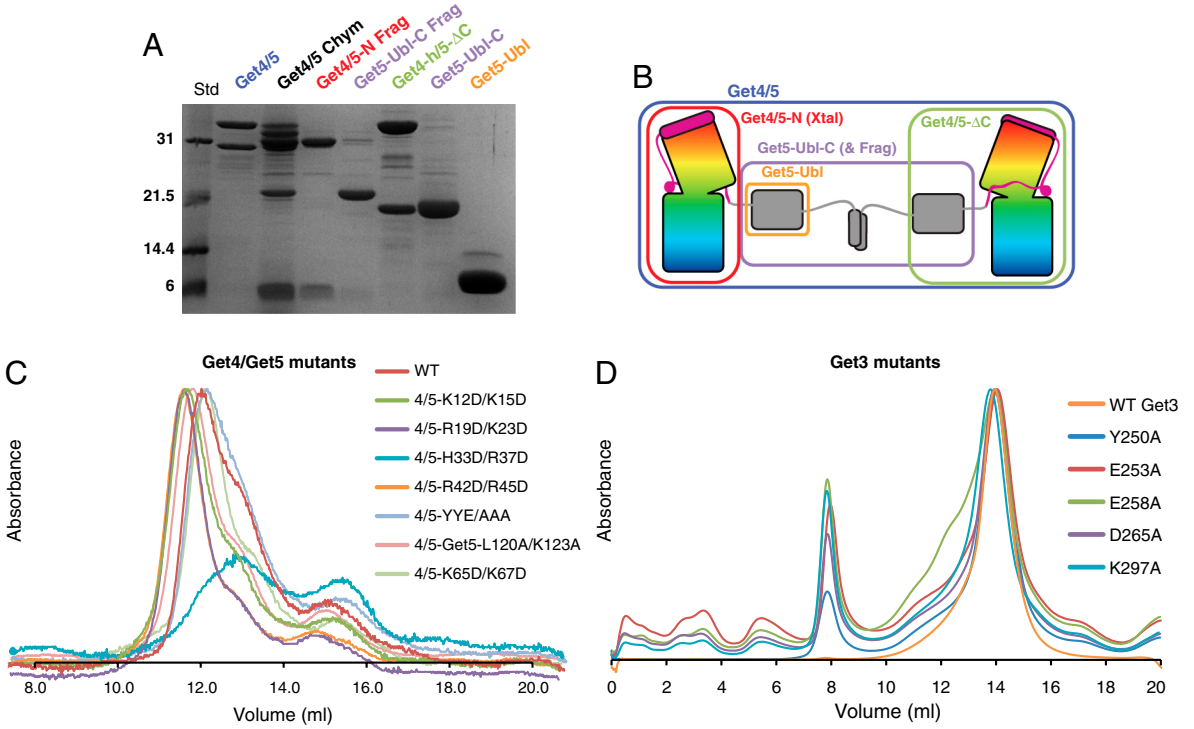

Fig. S6. Mutant purification. (*A*) Coomassie-stained SDS-PAGE of various puried Get4/5 constructs. Get4/5 Chym is the result of a 15 minute room temperature chymotrypsin digest. Get4/5-N Frag and Get5-UblC Frag are purified from the chymotrypsin digest using anion exchange chromatography as in Fig. S2*B*. Get4-h/5-C, Get5-UblC, and Get5-Ubl are generated at the genetic level. (*B*) Cartoon diagram of Get4/5 dimer with constructs boxed and colored as in *A*. (*C*) Size exclusion chromatograms of mutants of Get4/5 using a Superdex 200 10/300 column. The absorbences are normalized to the highest recorded value. (*D*) As in *C*, for mutants of Get3.

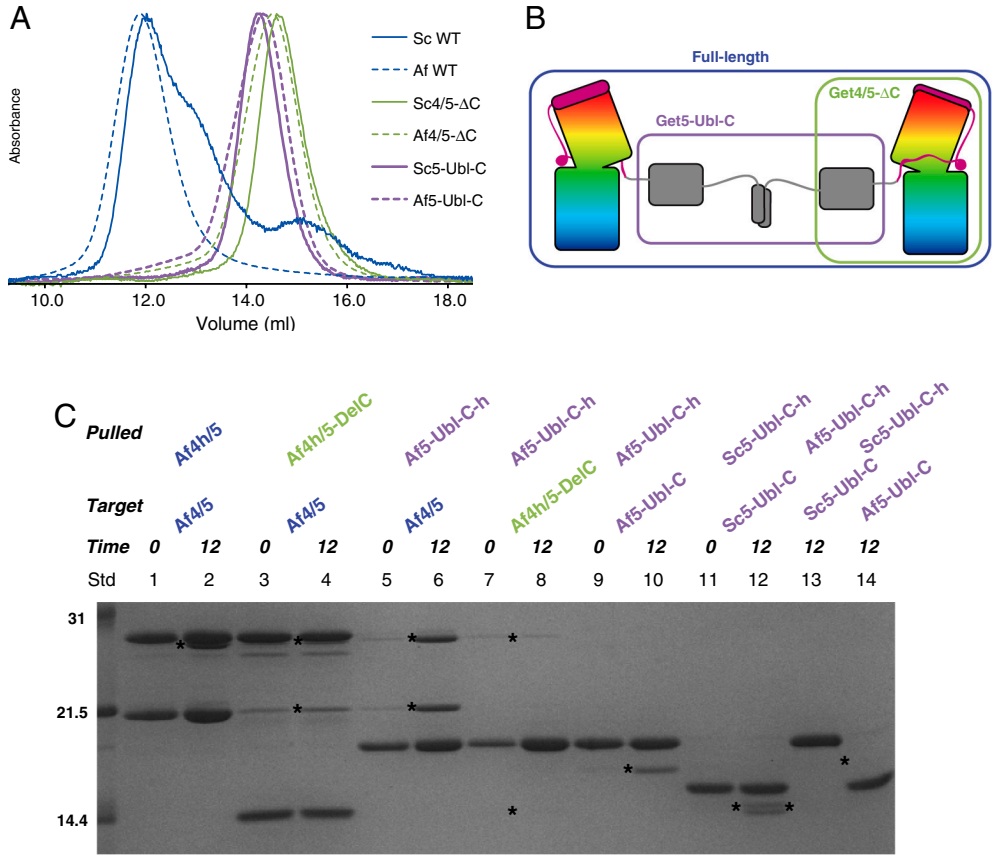

Fig. S7. Dimerization of *Af*Get4/5 by Get5-C. (*A*) Size exclusion chromatograms of constructs of Get4/5 from *Sc*, as in Fig. 2*A*, and *Af* (dashed lines). The absorbances are normalized to the highest recorded value. (*B*) Cartoon of constructs used in (*A*) indicated by colored boxes. (*C*) Domain swapping by dimerization. His-tagged proteins, indicated by "h," with and without Get5-C are mixed with untagged protein containing Get5-C for 0 and 12 h. Respective *Af* and *Sc* constructs are indicated. Asterisks (*) represent expected bands for successful capture.

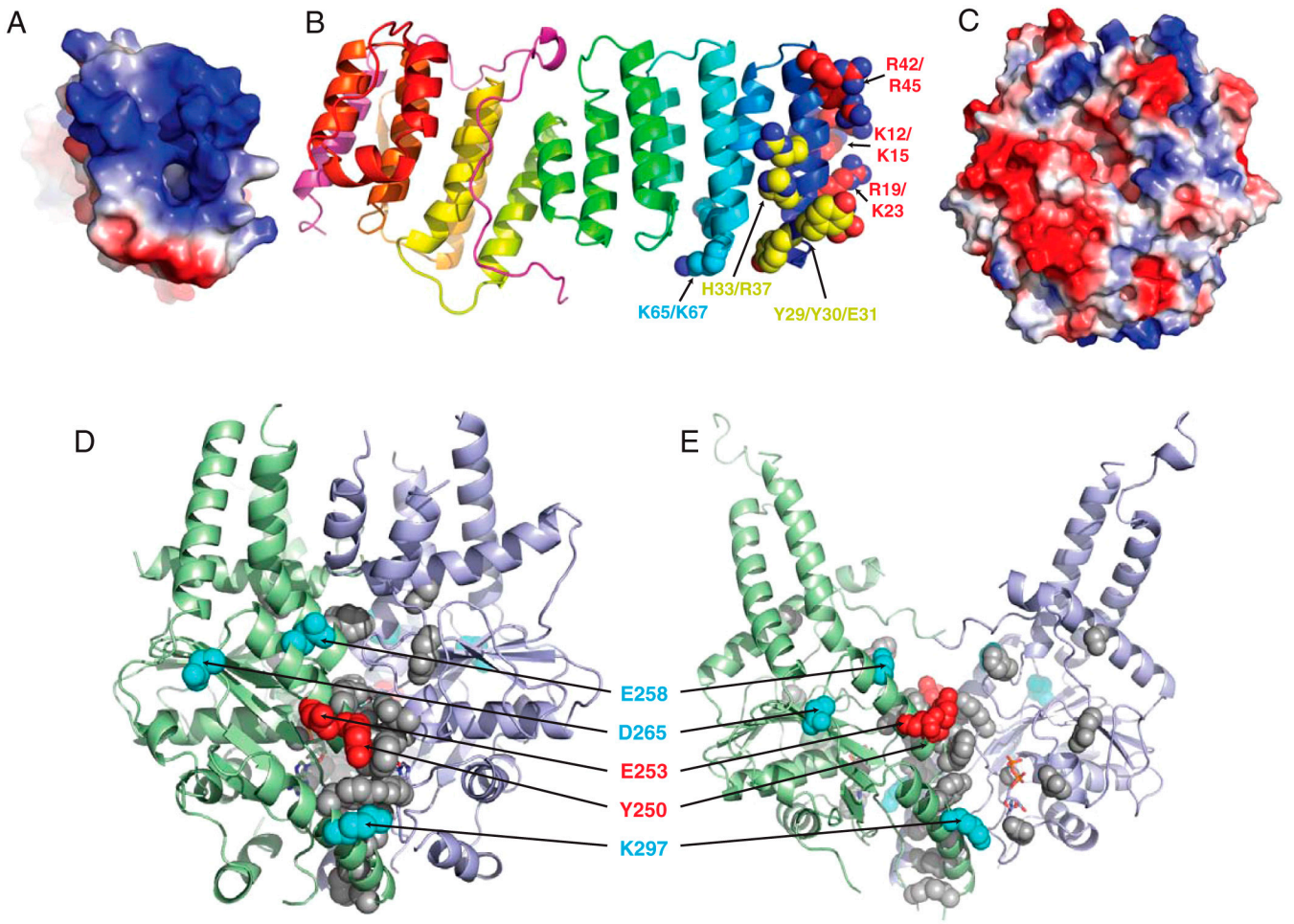

Fig. S8. Mutants on Get4 and Get3. (*A*) View from the N-terminal face of Get4 as an accessible surface colored from positive (blue) to negative (red) Coulombic charge. (*B*) Side view of Get4/5 as in Fig. 1*B*. The mutants used in Fig. 4 *B* and *C* are displayed as spheres as colored based on wild type (cyan), weak (yellow), or weakest (red) binding to Get3 as in Fig. 4*D*. (*C*) Get3 similar to (*A*) in the closed state (PDB ID code 2WOJ) (*D*) Ribbon diagram of the closed state of Get3 with mutants that had previously been shown to have strong phenotypes as spheres. Tested mutants are colored as in Fig. 4*E*. (*E*) Ribbon diagram of the open state of Get3 colored as in (*D*) (PDB ID code 3IBG).

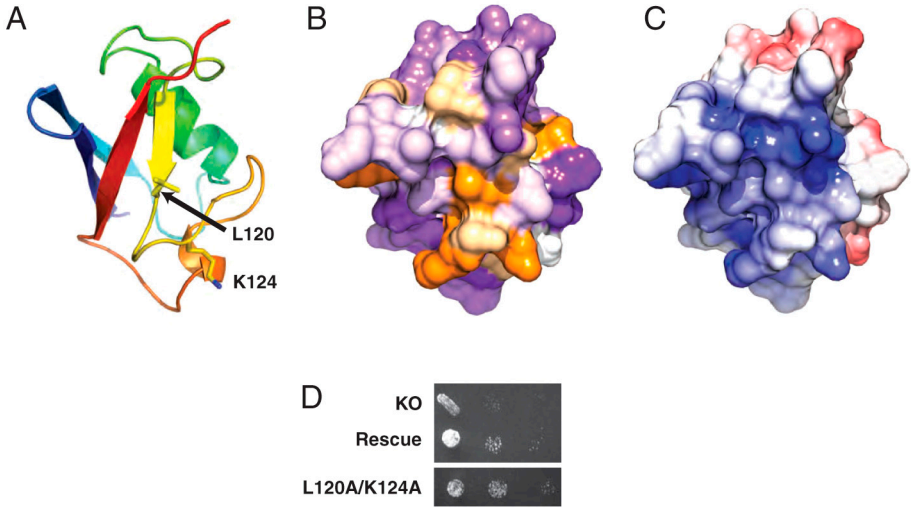

Fig. S9. Get5 model with phenotypic rescue. (*A*) Ribbons diagram of a homology model of the Get5-Ubl-domain generated by sequence threading through the NMR structure of human UBL4A (PDB ID code 2DZI) using SWISS-MODEL, color ramped from N- (blue) to C-terminus (red). Residues L120 and K124 are displayed as sticks. (*B*) Get5-Ubl oriented as in *A* and shown as an accessible surface colored from least (purple) to most (orange) conserved. (*C*) As in (*B*), colored from positive (blue) to negative (red) Coulombic charge. (*D*) Spot plate growth assays in the BY4741 Get5::KanMX background, performed as in Fig. 5.

Table S1. Crystallographic statistics

Data collection	Get4/5 Se peak
Beamline	SSRL 9-2
Wavelength (Å)	0.97862
Space group	P 3_1 2 1
Cell dimensions	
a, *b*, *c* (Å)	108.7, 108.7 169.8
α, β, γ (°)	90, 90, 120
Resolution range (Å)	30-2.8 (2.95–2.80)*
R_{merge} †	0.104 (0.500)
I/σI	20.3 (4.7)
Completeness (%)	99.9 (100.0)
Multiplicity	12.1 (11.7)
Anom. compl. (%)	99.9 (99.9)
Anom. multiplicity	6.3 (6.1)
Phasing	
Se sites	9
PHASER figure of merit (8)	0.29
RESOLVE figure of merit (9)	0.65
Refinement	
No. reflections, working	27722
No. reflections, free	1464
No. atoms	
Protein	8439
Water	84
Ligands	16
Avg. B-factor (Å^2)	52.3
Working R-factor‡	0.18152
Free R-factor	0.22437
rmsd bonds (Å)	0.011
rmsd angles (°)	1.2

*Values in parentheses represent highest resolution shell.

†$R_{merge} = \Sigma hkl \Sigma i |I(hkl)i - \langle I(hkl)\rangle| / \Sigma hkl \Sigma i \langle I(hkl)i\rangle$.

‡R-factor $= \Sigma hkl |F_o(hkl) - F_c(hkl)| / \Sigma hkl |F_o(hkl)|$, where F_o and F_c are observed and calculated structure factors, respectively.