

Supplemental Information

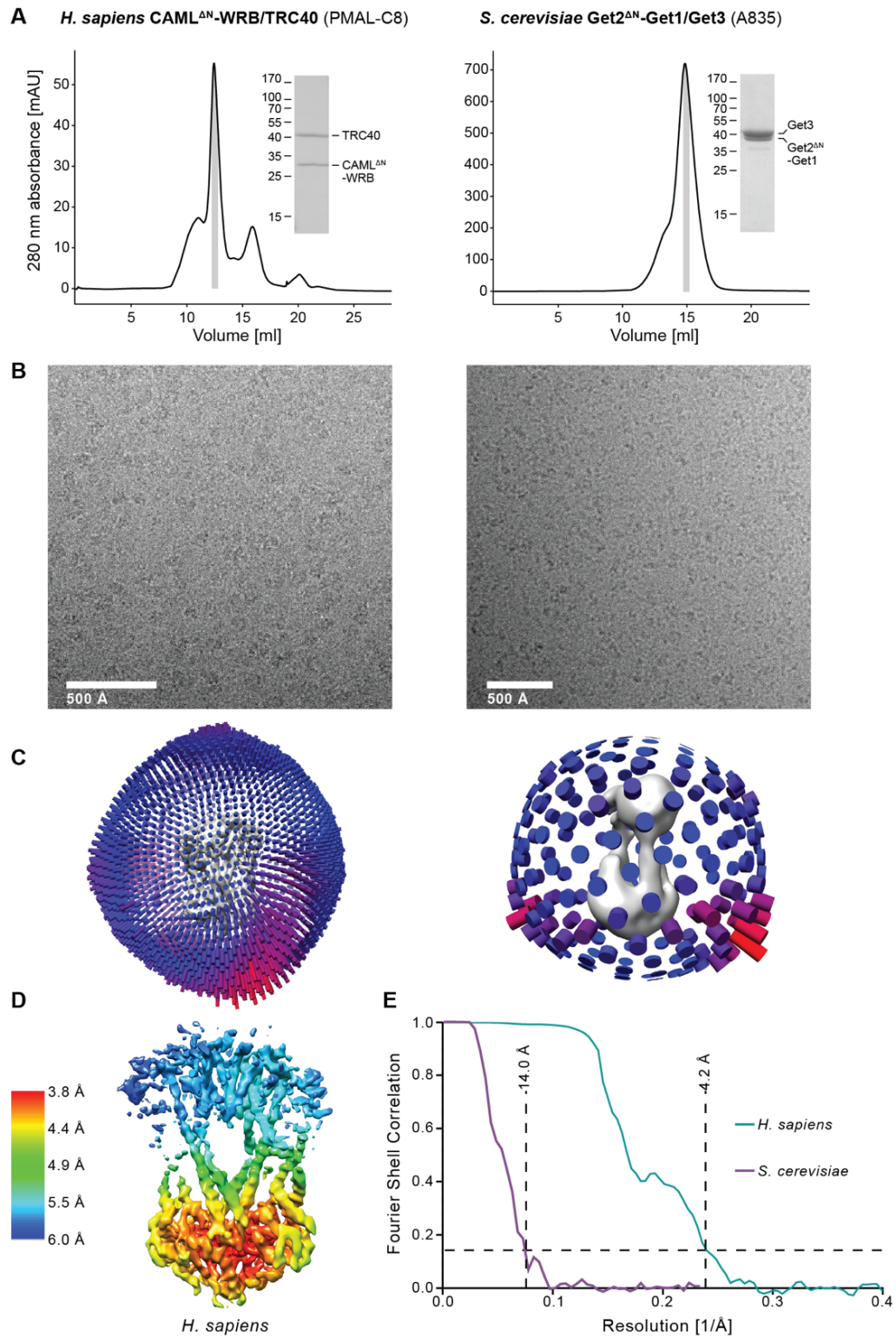
Structural Basis of Tail-Anchored Membrane

Protein Biogenesis by the GET Insertase Complex

Melanie A. McDowell, Michael Heimes, Francesco Fiorentino, Shahid Mehmood, Ákos Farkas, Javier Coy-Vergara, Di Wu, Jani Reddy Bolla, Volker Schmid, Roger Heinze, Klemens Wild, Dirk Flemming, Stefan Pfeffer, Blanche Schwappach, Carol V. Robinson, and Irmgard Sinning

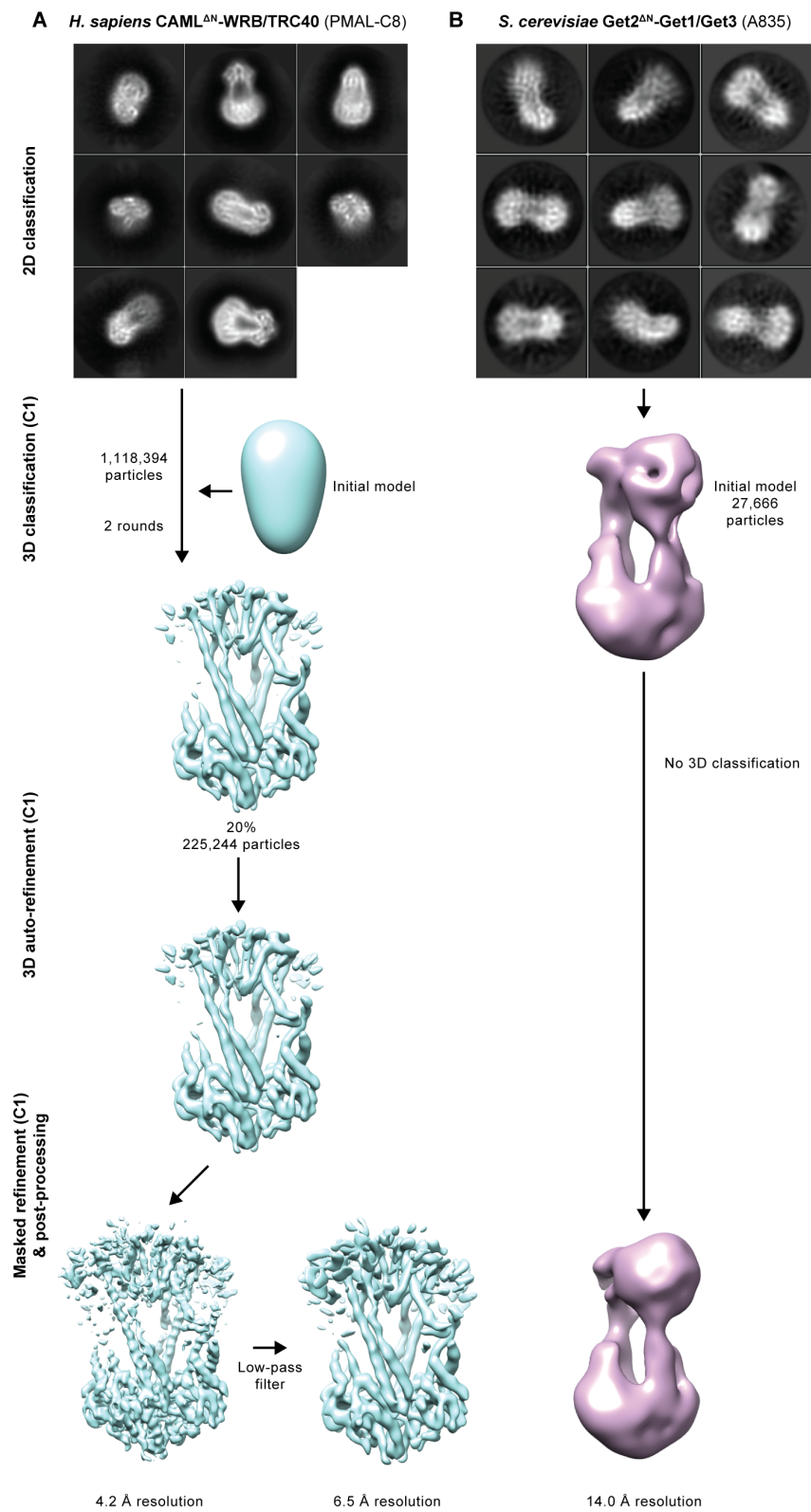
Supplemental Figures

Figure S1. Sample preparation and experimental quality for cryo-EM. Related to Figures 1 and 3.



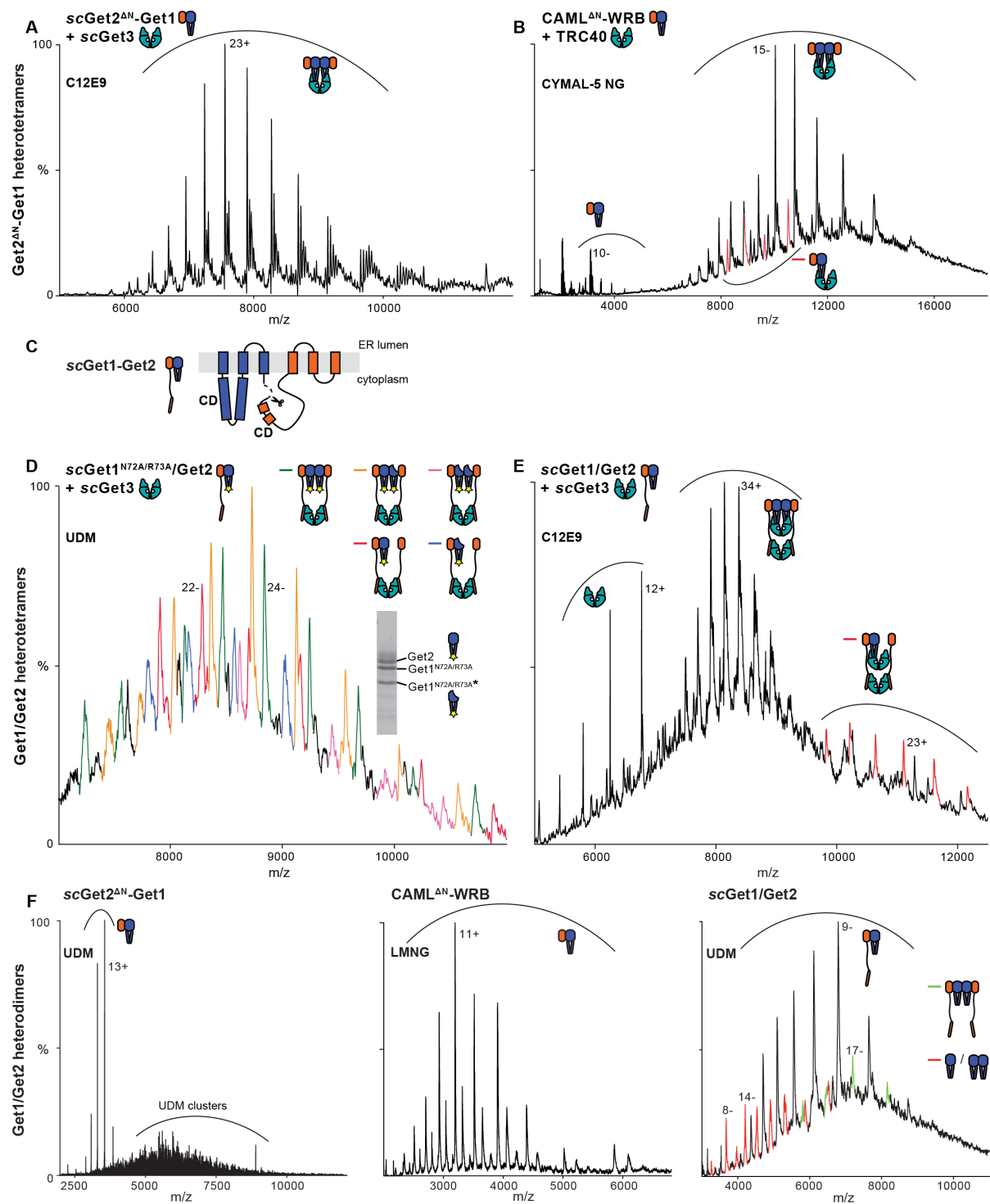
(A) Size exclusion chromatography profiles for the complexes in amphipol, with the single fraction taken for cryo-EM analysis highlighted in grey. Insets, Coomassie-stained 12.5% SDS-PAGE gels showing the purity of the final complexes. **(B)** Representative micrographs for each complex. **(C)** Angular distribution plots following 3D classification (*H. sapiens*) or 3D auto-refinement (*S. cerevisiae*). **(D)** Local resolution estimates of the final reconstruction obtained from the *H. sapiens* complex. **(E)** Gold-standard FSC curves of the final reconstructions calculated using a soft-edge mask.

Figure S2. Workflow for cryo-EM single-particle data processing of CAML^{ΔN}-WRB/TRC40 and scGet2^{ΔN}-Get1/Get3 complexes. Related to Figures 1 and 3 and STAR Methods.



(A) Autopicked particles for *H. sapiens* CAML^{ΔN}-WRB/TRC40 in PMAL-C8 were subjected to 2D classification followed by 3D classification against a low-resolution *de novo* model. Particle images were then refined using a soft mask, resulting in a reconstruction with an average resolution of 4.2 Å. More complete density for low resolution regions was present in a reconstruction low-pass filtered to 6.5 Å. **(B)** For *S. cerevisiae* Get2^{ΔN}-Get1/Get3 in A835, 3D classification did not produce sensible particle subsets. Therefore, the initial model was refined directly, resulting in a reconstruction at 14.0 Å resolution.

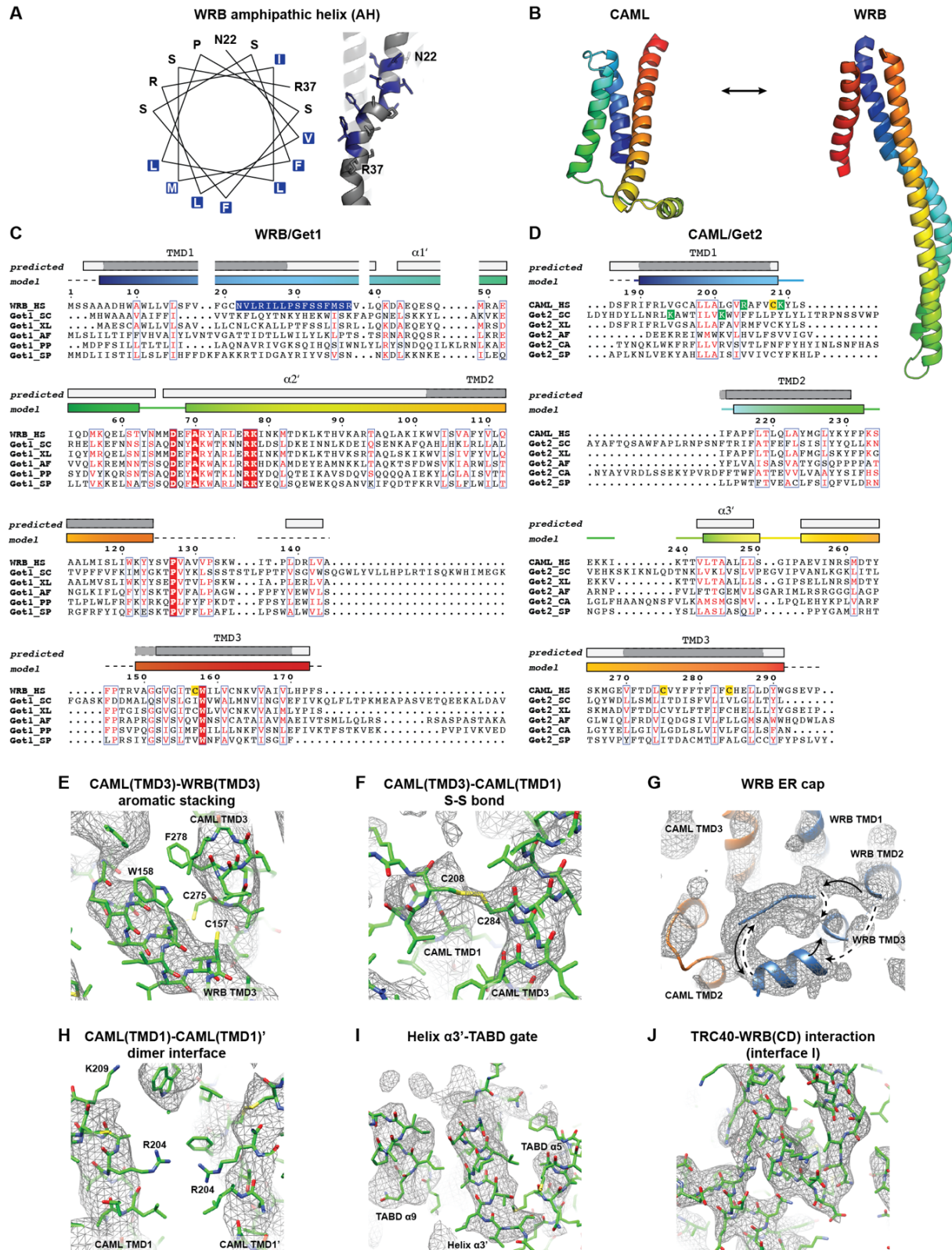
Figure S3. Characterisation of *scGet1/Get2/Get3* and *CAML^{ΔN}-WRB/TRC40* by native-MS. Related to Figure 2.



Mass spectra of **(A)** *scGet2*^{ΔN}-Get1/Get3 and **(B)** CAML^{ΔN}-WRB/TRC40 complexes liberated from detergent micelles collected under non-denaturing conditions. **(C)** Topology of the full-length *scGet1*-Get2 single-chain construct used in this study showing the Get1 and Get2 CDs. The dashed line represents the non-native linker sequence that is cleaved by specific proteases during purification to yield *scGet1*/Get2. Mass spectra of full-length *scGet1*/Get2, **(D)** with and **(E)** without N72A/R73A mutations in the Get1-CD, liberated from detergent micelles collected under non-denaturing conditions. **(D)** Inset, Coomassie-stained 12.5% SDS-PAGE gel showing the degradation product of *scGet1*^{N72A/R73A} (*scGet1*^{N72A/R73A*}) observed during native-MS. **(E)** The Get3₂ species represents excess protein added to solution during complex formation. The *scGet1*/Get2₂/Get3₂ complex arises from collision induced dissociation and is not present in solution. **(F)** Mass spectra of *scGet2*^{ΔN}-Get1, CAML^{ΔN}-WRB and *scGet1*/Get2 heterodimers liberated from detergent micelles collected under non-denaturing conditions. For each assigned complex in all mass spectra the charge state for the most intense peak is shown, with the sign denoting data collection in either positive or negative mode.

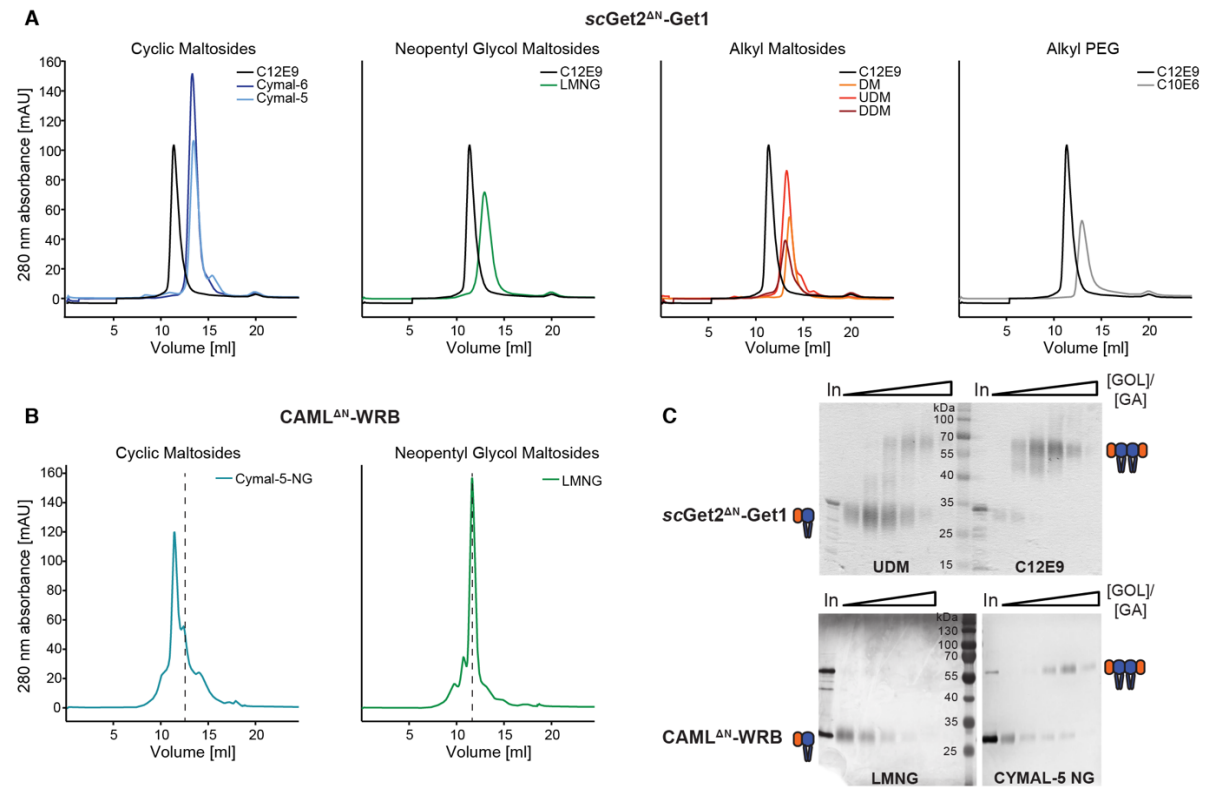
Figure S4. Model building of CAML^{AN}-WRB/TRC40 in the cryo-EM density. Related to

Figure 3.



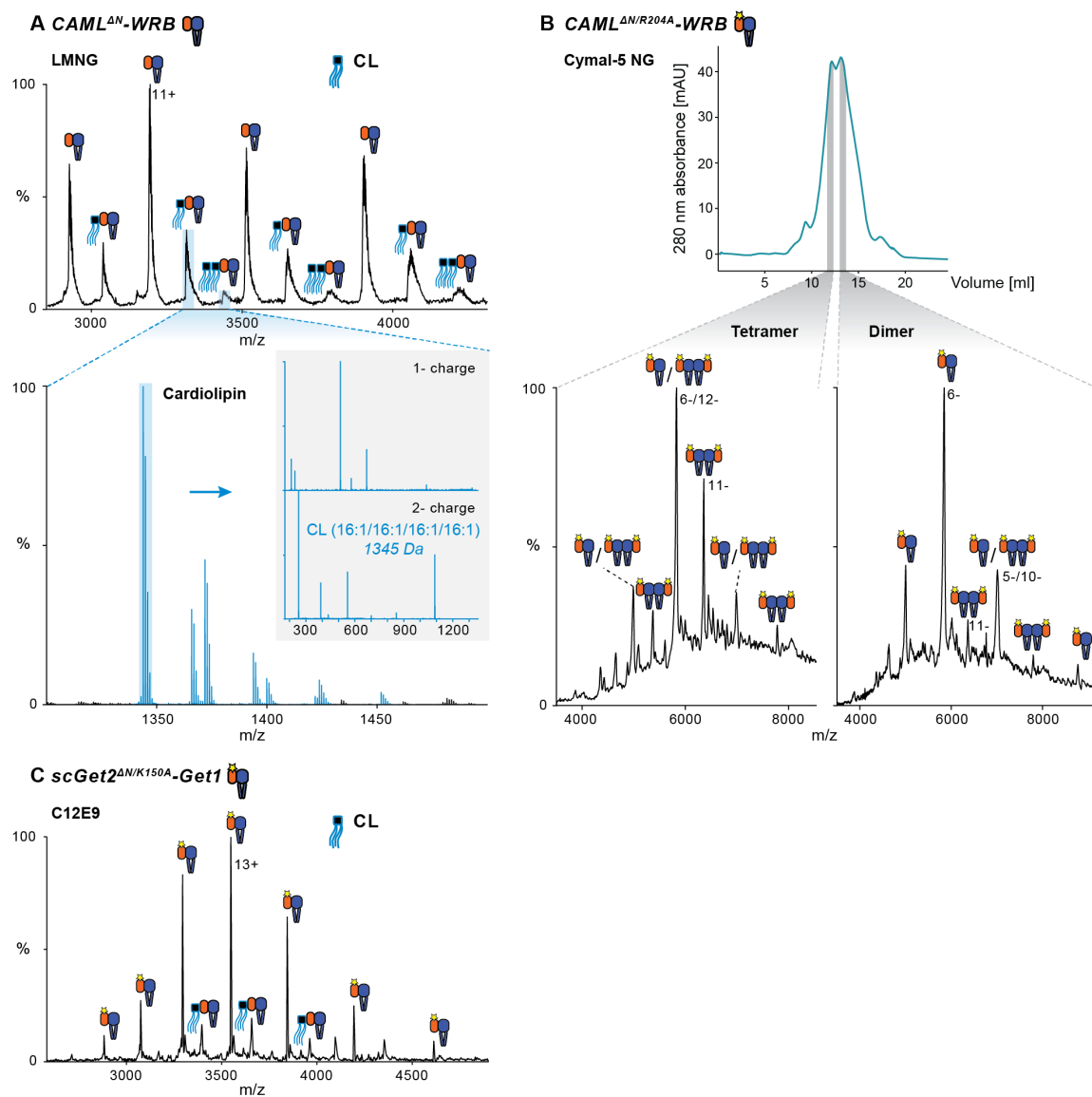
(A) The amphipathic helix of WRB (residues 22-37) is shown in both a helical wheel and structural representation. Hydrophobic residues are coloured blue. **(B)** The structures of CAML and WRB (artificially separated in the membrane plane for clarity) coloured from the N- (blue) to the C- (red) terminus. **(C, D)** Sequence alignments for **(C)** WRB/Get1 and **(D)** CAML/Get2 generated using simultaneous transmembrane prediction and secondary structure alignment in PRALINE (Pirovano et al., 2008) and visualised using Esript (Gouet et al., 2003). Sequences are shown for *H. sapiens* (HS), *S. cerevisiae* (SC), *Xenopus laevis* (XL), *Aspergillus fumigatus* (AF), *Candida albicans* (CA), *Schizosaccharomyces pombe* (SP) and *Pichia pastoris* (PP). Conserved and similar residues are shown in red, CAML/*sc*Get2 TMD1 positively charged residues that may point into the heterotetramer interface are highlighted in green and CAML/WRB cysteines that may be involved in disulphide bond formation are highlighted in yellow. The amphipathic helix present in the structure of WRB is highlighted in blue. Residues missing in our CAML^{ΔN}-WRB model are indicated by dashed lines. Helices predicted for CAML/WRB by PSIPRED (Buchan and Jones, 2019) are shown in grey, with dark shading showing transmembrane domains predicted by Phobius (Käll et al., 2007). Helices present within our model for CAML^{ΔN}-WRB are coloured according to their representation in **(B)**. **(E-J)** Representative density for CAML^{ΔN}-WRB/TRC40 showing key structural features. For the WRB ER cap **(G)**, two possible connectivities are indicated with arrows on the cartoon representation, with the solid arrows indicating the most likely arrangement based on density shape and size.

Figure S5. *scGet2^{ΔN}-Get1* and *CAML^{ΔN}-WRB* form heterotetramers in C12E9 and Cymal5-NG respectively. Related to Figure 4.



(A) Size exclusion chromatography for *scGet2^{ΔN}-Get1* heterodimers in different detergents using a Superdex 200 increase 10/30 GL column shown relative to the elution profile for the *scGet2^{ΔN}-Get1* heterotetramer in C12E9 (black). **(B)** Size exclusion chromatography for *CAML^{ΔN}-WRB* in different detergents using a Superdex 200 increase 10/30 GL column. The dashed line denotes the elution volume for the heterodimeric fusion. **(C)** GraFix cross-linking of *scGet2^{ΔN}-Get1* and *CAML^{ΔN}-WRB* oligomers in the given detergents. Silver-stained 12.5% SDS-PAGE gels show the input (In) and fractions taken from an increasing gradient of glycerol (GOL) and glutaraldehyde (GA).

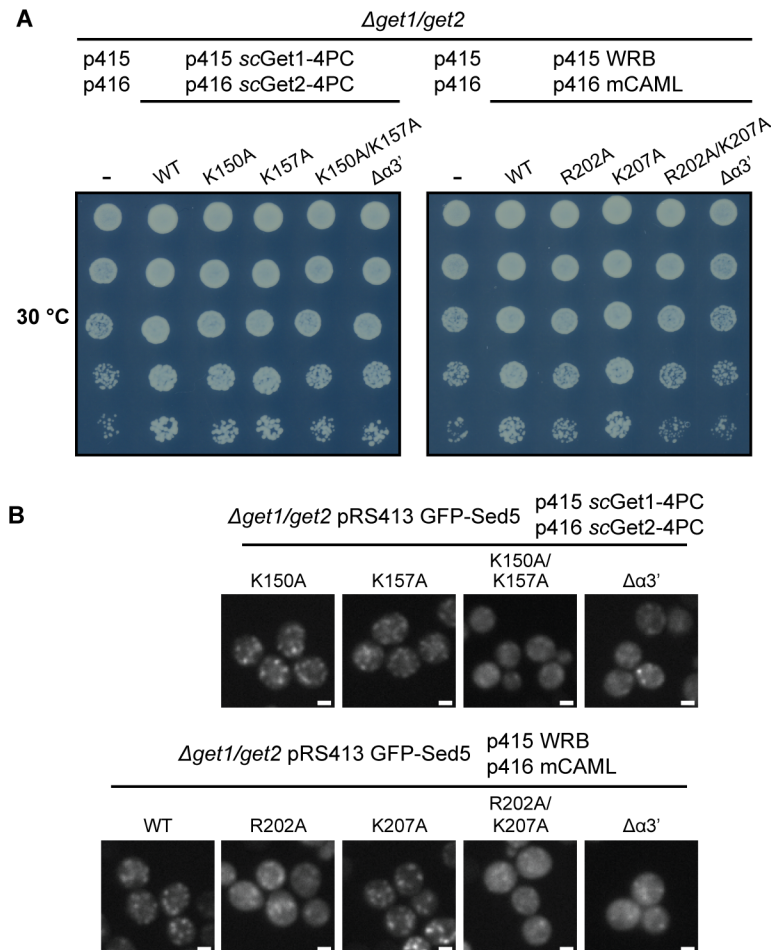
Figure S6. CAML^{ΔN}-WRB and *scGet2*^{ΔN/K150A}-Get1 heterodimers do not bind PI. Related to Figure 4 and 5.



(A) Zoomed region of the CAML^{ΔN}-WRB mass spectrum from Figure S3F showing the main charge state for the major species. Adduct peaks with a defined mass greater than the amino acid sequence mass correspond to bound CL. MS profile for CL species identified through in-solution lipidomics is shown beneath the spectrum. Inset, MS/MS of the highlighted ions identified the major CL species associated with the protein. **(B)** Size exclusion chromatography for CAML^{ΔN/R204A}-WRB in Cymal-5 NG using a Superdex 200 increase 10/30 GL column.

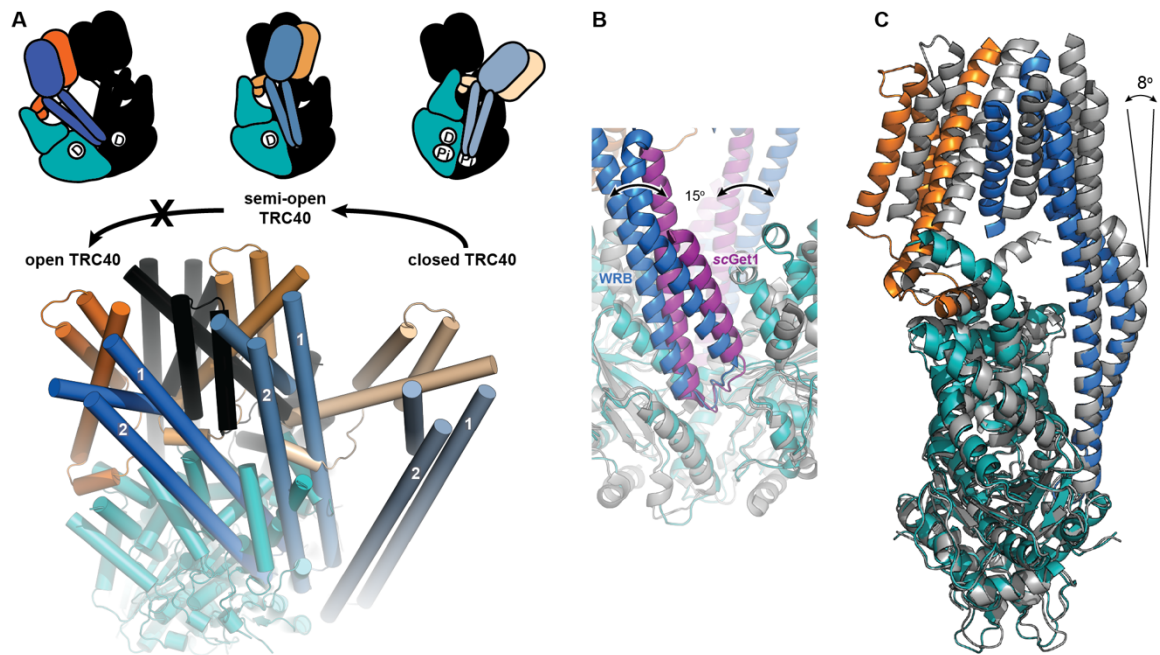
Shown beneath are mass spectra collected under non-denaturing conditions for the indicated fractions (corresponding to tetrameric and dimeric species) liberated from detergent micelles. **(C)** Mass spectrum of *scGet2*^{ΔN/K150A}-Get1 liberated from C12E9 micelles under non-denaturing conditions. **(B, C)** For each assigned complex the charge state for the most intense peak is shown, with the sign denoting data collection in either positive or negative mode.

Figure S7. Helix $\alpha 3'$ and positively charged residues in the heterotetramer interface are important for *scGet1/Get2* and WRB/mCAML function *in vivo*. Related to Figure 4, 5 and 6.



(A), (B) *scGet1-4PC/Get2-4PC* and WRB/mCAML were co-expressed from the indicated plasmids in *Δget1/get2* yeast strains for wild type (WT) and mutant *scGet2/mCAML* sequences. Transformants with the empty vector were taken as a negative control (-). (A) Control plates from the spotting assay shown in Figure 5B incubated at 30°C, demonstrating comparable growth of the strains under normal growth conditions. (B) Representative fluorescence microscopy images of *Δget1/get2* cells expressing GFP-Sed5 and *scGet1-4PC/Get2-4PC* or WRB/mCAML constructs from separate plasmids. Scale bar: 2 μ M.

Figure S8. Structural analysis of the WRB-CD interaction with TRC40. Related to Figures 3 and 7.



(A) Model for CAML^{ΔN}-WRB coupled to the closed to semi-open to open transition of TRC40. One CAML^{ΔN}-WRB/TRC40 (black) is kept in a fixed position by first aligning semi-open and closed *scGet3* templates (PDB accession codes 3SJC and 4XTR) via chain A (RMSD 1.32 Å over 234 Cα atoms), then superimposing chain A of TRC40 (RMSD 1.72 Å over 230 Cα atoms) with the semi-open state. Chain B of TRC40 (RMSD 1.60 Å over 226 Cα atoms) from our model is then superimposed on semi-open and closed *scGet3*. The adjoining CAML^{ΔN}-WRB maintains interactions via interface I (Stefer et al., 2011). Arrows show the expected movement of the mobile CAML^{ΔN}-WRB half (shown in Video S1). **(B)** Superimposition of the open *scGet3* dimer (grey; PDB accession 3SJB) with TRC40 (cyan; RMSD 2.31 Å over 456 Cα atoms), showing the tilt of bound *scGet1*-CD relative to WRB. **(C)** Superimposition of one TRC40 protomer on the other within the CAML^{ΔN}-WRB/TRC40 complex (RMSD 1.01 Å over 271 Cα atoms). The superimposed TRC40 protomer and the CAML^{ΔN}-WRB heterodimer interacting via interface I (Stefer et al., 2011) are shown in grey for clarity. The given angle represents the difference in tilt of the WRB-CD between each dimer half.

Supplemental Table

Table S1. Molecular masses determined by native mass spectrometry. Related to Figures 2 and 4.

Figure	Complex	Experimental molecular mass $\pm$ s.d. (Da)	Theoretical molecular mass (Da)
	<i>scGet2^{ΔN}-Get1</i>		
Figure 4B	(<i>scGet2^{ΔN}-Get1</i>) ₂	92,493 $\pm$ 2	92,328
Figure 4C	(<i>scGet2^{ΔN}-Get1</i>) ₂ + phospholipid	93,332 $\pm$ 11	N/A
Figure 4C	(<i>scGet2^{ΔN}-Get1</i>) ₂ + cardiolipin	93,883 $\pm$ 8	N/A
Figure S3F	<i>scGet2^{ΔN}-Get1</i>	46,163 $\pm$ 9	46,164
Figure S3A	(<i>scGet2^{ΔN}-Get1</i>) ₂ /Get3 ₂	173,745 $\pm$ 10	173,600
Figure 2A	(<i>scGet2^{ΔN}-Get1</i>) ₂ /Get3 ₂	173,702 $\pm$ 3	173,600
	<i>scGet2^{ΔN/K150A}-Get1</i>		
Figure S6C	<i>scGet2^{ΔN/K150A}-Get1</i>	46,105 $\pm$ 1	46,164
Figure S6C	<i>scGet2^{ΔN/K150A}-Get1</i> + cardiolipin	47,521 $\pm$ 18	N/A
	<i>scGet1/Get2</i>		
Figure S3F	<i>scGet1/Get2</i>	61,152 $\pm$ 12	61,136
Figure S3E	(<i>scGet1/Get2</i>) ₂ /Get3 ₂	284,851 $\pm$ 37	284,816
Figure 2C	(<i>scGet1/Get2</i>) ₂ /Get3 ₂	284,835 $\pm$ 12	284,816
	<i>scGet1/Get2/Get3₂</i>	142,346 $\pm$ 6	142,408
	<i>scGet1^{N72A/R73A}/Get2</i>		
Figure S3D	<i>scGet1^{N72A/R73A}/Get2₂/Get3₂</i>	174,006 $\pm$ 20	173,917
	(<i>scGet1^{N72A/R73A}/Get2</i>) ₂ /Get3 ₂	203,302 $\pm$ 48	203,287
	<i>CAML^{ΔN}-WRB</i>		
Figure 4B	(<i>CAML^{ΔN}-WRB</i>) ₂	70,067 $\pm$ 26	70,095
Figure 4D	(<i>CAML^{ΔN}-WRB</i>) ₂ + phospholipid	70,893 $\pm$ 16	N/A
Figure S3F	<i>CAML^{ΔN}-WRB</i>	35,090 $\pm$ 0.5	35,047
Figure S6A	<i>CAML^{ΔN}-WRB</i> + cardiolipin	36,435 $\pm$ 0.2	N/A
Figure S3B	(<i>CAML^{ΔN}-WRB</i>) ₂ /TRC40 ₂	150,548 $\pm$ 18	150,305
Figure 2B	(<i>CAML^{ΔN}-WRB</i>) ₂ /TRC40 ₂	150,375 $\pm$ 46	150,305
	<i>CAML^{ΔN/R204A}-WRB</i>		
Figure S6B	(<i>CAML^{ΔN/R204A}-WRB</i>) ₂	69,952 $\pm$ 77	70,095
Figure S6B	<i>CAML^{ΔN/R204A}-WRB</i>	35,077 $\pm$ 33	35,047