



Figures and figure supplements

The architecture of EMC reveals a path for membrane protein insertion

John P O'Donnell et al

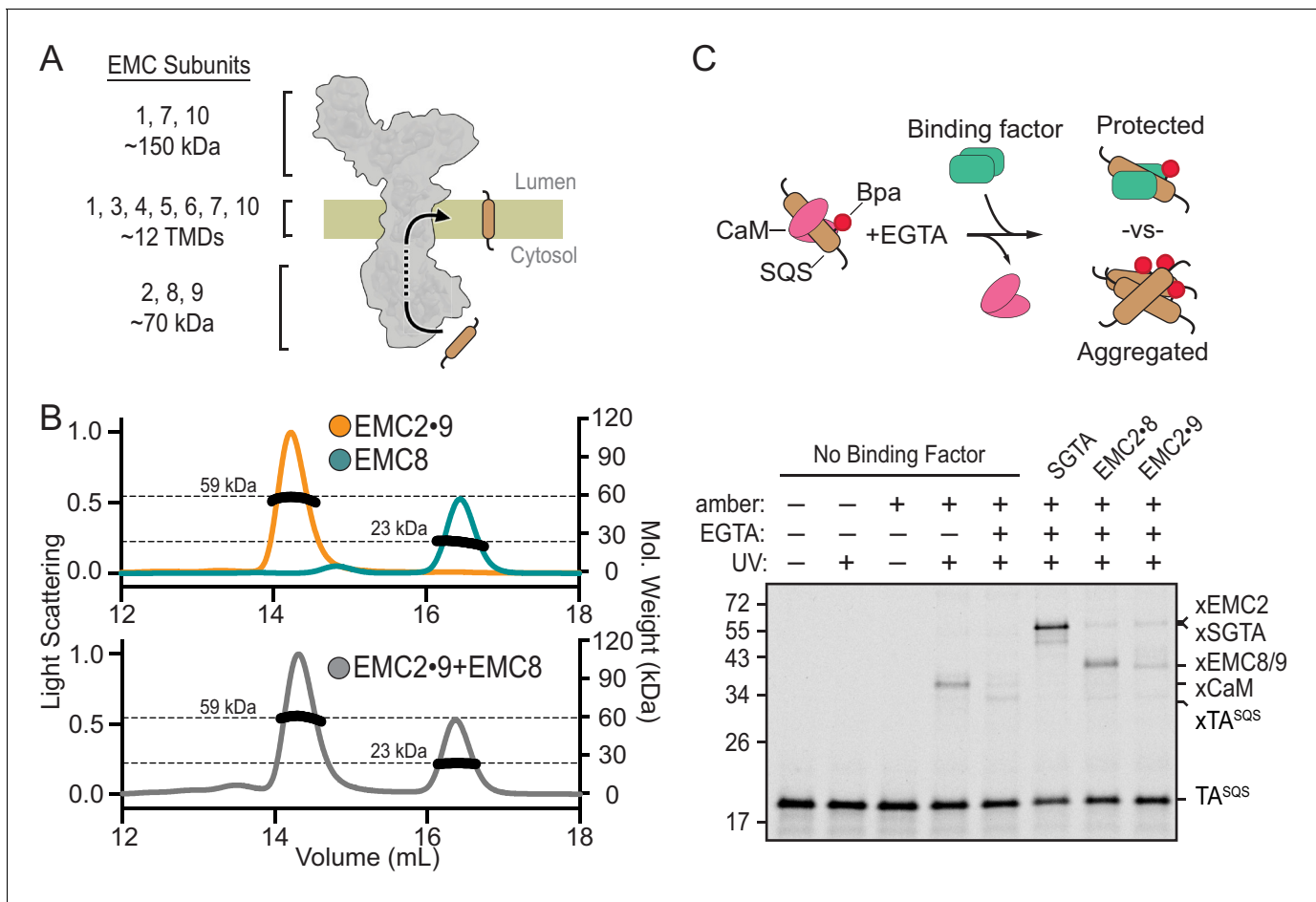


Figure 1. EMC2•EMC8 and EMC2•EMC9 form complexes that can bind a TMD. (A) Diagram of predicted EMC mass distribution and subunits. (B) Size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) analysis of recombinant EMC2•EMC9 reveals a stable complex in a 1:1 ratio at the expected molecular weight of 59 kDa (orange). Recombinantly expressed EMC8 analyzed independently has a Mw of 23 kDa consistent with a monomeric state (teal). Addition of EMC8 to a pre-formed EMC2•EMC9 complex does not result in the formation of a ternary complex (grey). (C) ³⁵S-methionine-labeled TA^{SQS} containing the benzoyl-phenylalanine (Bpa) photo-crosslinker within the TMD was produced as a defined complex with CaM using the PURE in vitro translation system. Bpa is incorporated into the TMD by amber suppression. Addition of EGTA releases CaM. The released TA^{SQS} will either aggregate or be protected from aggregation by a TMD-binding protein. The outcome can be monitored by UV-mediated crosslinking via Bpa to nearby proteins. The position of crosslinks between TA^{SQS} and various partners are indicated. No UV-mediated crosslinks are seen when TA^{SQS} does not contain an amber codon (lanes 1 and 2).

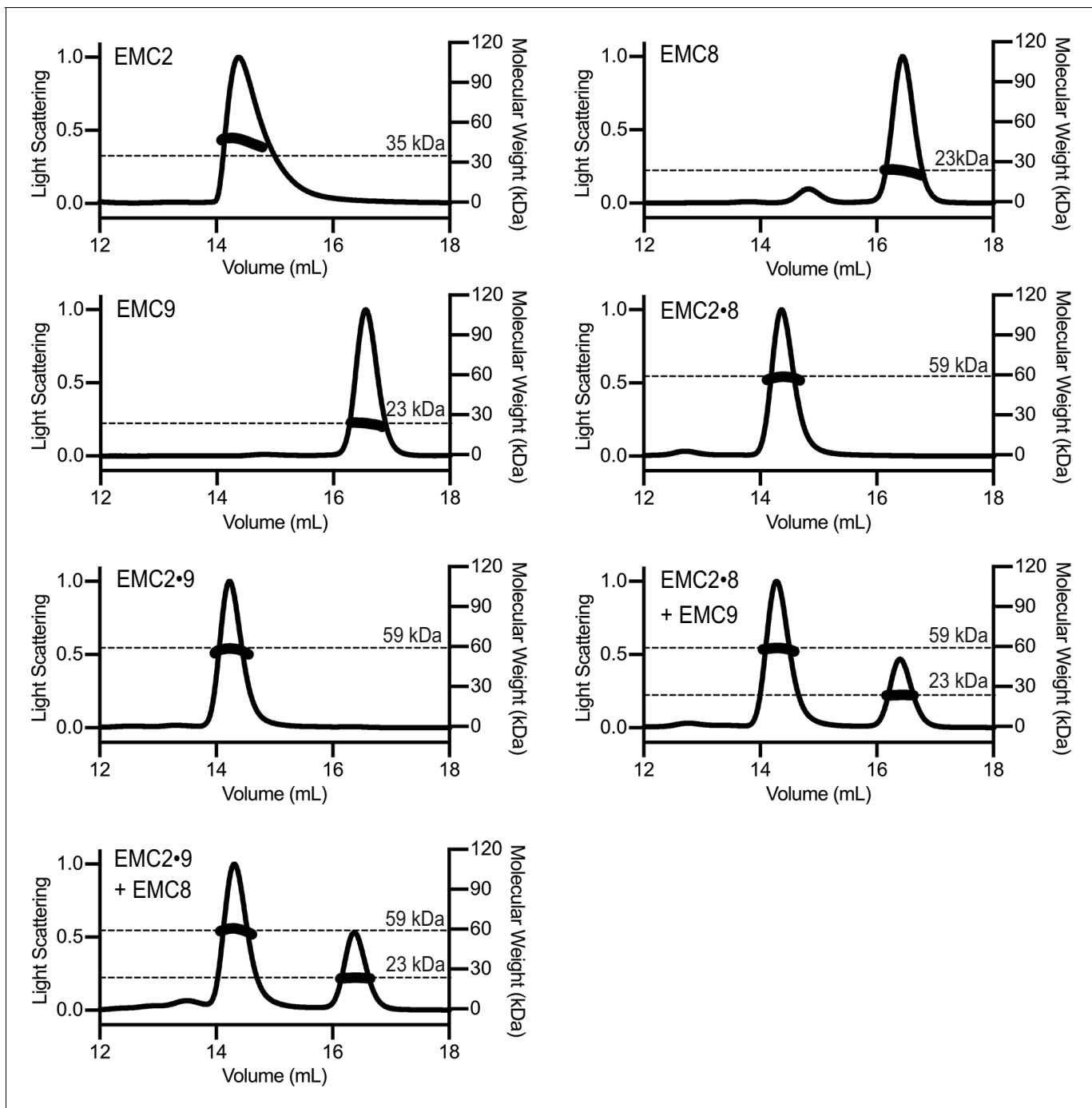


Figure 1—figure supplement 1. SEC-MALS of individual and complexed cytoplasmic EMC-subunits. Size Exclusion Multi Angle Light Scattering (SEC-MALS) analysis of recombinant EMC2, EMC8 and EMC9. The signal from the 90°-scattering detector is plotted on the left y-axis as a solid line and the calculated molecular weight is plotted on the right y-axis and depicted as points across the elution peak. A dotted line notated with text indicates the theoretical molecular weight for each protein. The calculated molecular weights are consistent with those predicted for monomeric subunits and complexed subunits. Neither addition of EMC8 to a preformed EMC2•EMC9 complex nor addition of EMC9 to a pre-formed EMC2•EMC8 complex resulted in the formation of a ternary complex.

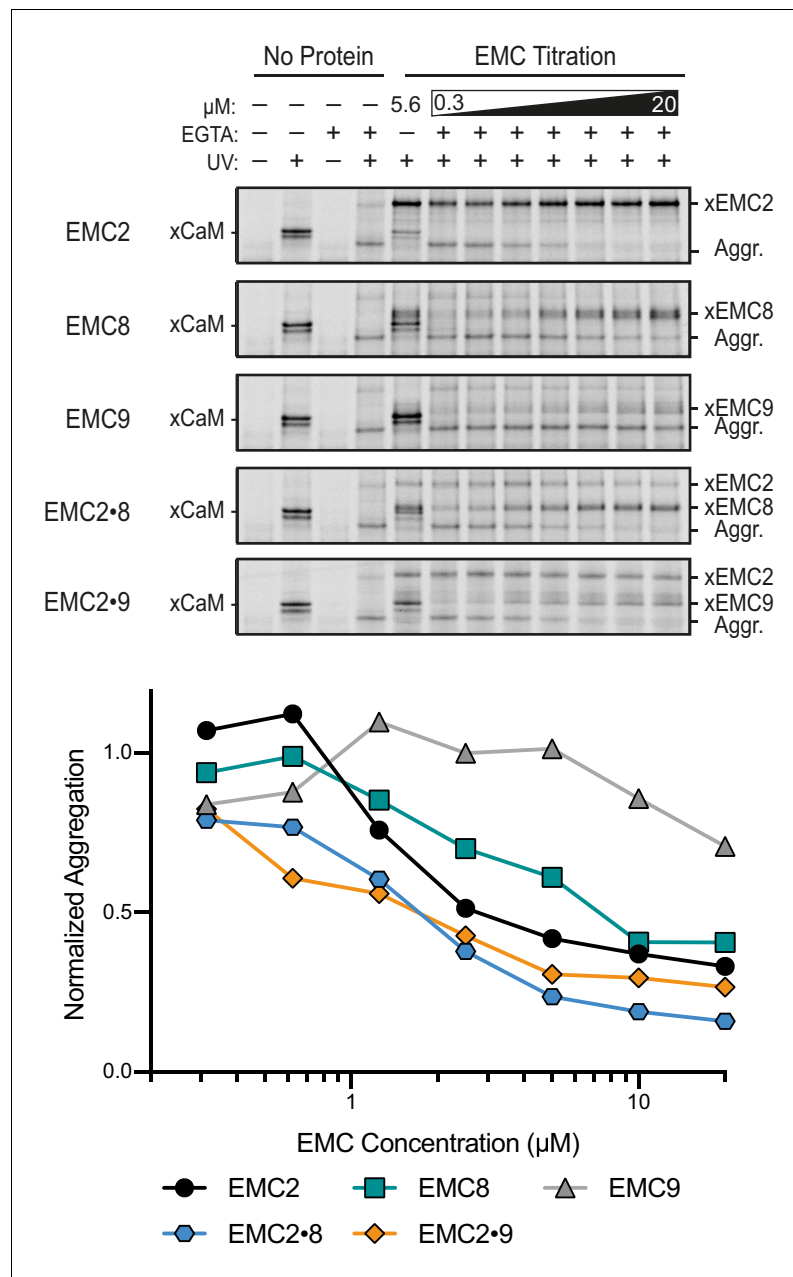


Figure 1—figure supplement 2. Activity of EMC cytosolic subunits in preventing TMD aggregation. ³⁵S-methionine labeled SQS containing the benzoyl-phenylalanine (Bpa) photo-crosslinker within the TMD was prepared in complex with CaM as characterized in **Figure 1C**. The complex was mixed with various concentrations of different proteins (indicated on the left), treated with EGTA to release the substrate from CaM, and substrate interactions monitored by UV-mediated crosslink formation. Release from CaM in the absence of any additional protein results in UV-dependent self crosslinks indicating of aggregation (Aggr.). This aggregate product is progressively reduced to a comparable extent by the inclusion of EMC2, EMC2•EMC8, or EMC2•EMC9. By contrast, EMC9 alone is not able to reduce aggregation, while EMC8 alone has a partial effect on aggregation. Note that the crosslinking observed to EMC9 when it is added in isolation is probably due to its co-aggregation with substrate and not a protective interaction because the substrate aggregate product is not diminished appreciably. The plot below the gels quantifies the relative intensity of the aggregate band normalized to the sample without added protein.

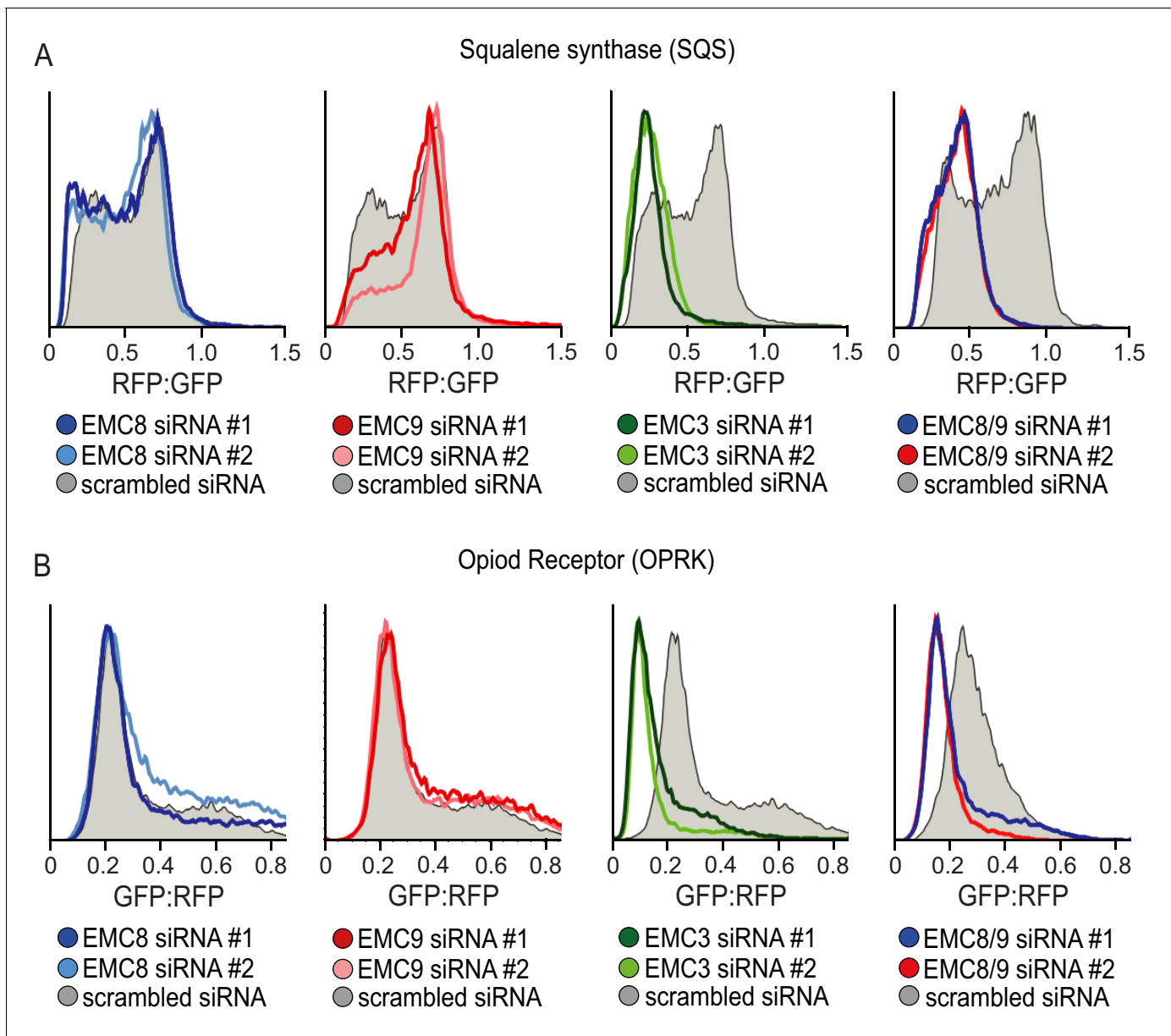


Figure 1—figure supplement 3. EMC8 and EMC9 necessary for membrane protein biogenesis but are functionally redundant. (A) Flow cytometry analysis of a tail-anchored RFP-squalene synthase (RFP-SQS) substrate, relative to an internal GFP expression control. SQS is an established EMC substrate (Guna et al., 2018; Volkmar et al., 2019). Reduced biogenesis of the TMD substrate results in a shift of the histogram ratio to the left. siRNA mediated knockdown of either EMC8 or EMC9 had minimal impact on RFP-squalene synthase biogenesis compared to knockdown of the integral transmembrane component EMC3. Combined knockdown of EMC8 and EMC9 resulted in a significant defect in RFP-SQS biogenesis. (B) The same experiment as in panel A but using the mu opioid receptor (GFP-OPRK), a multi-pass membrane protein, as the substrate. Here, OPRK is GFP-tagged and RFP serves as the internal expression control. Previous studies have shown that OPRK is an EMC substrate (Chitwood et al., 2018).

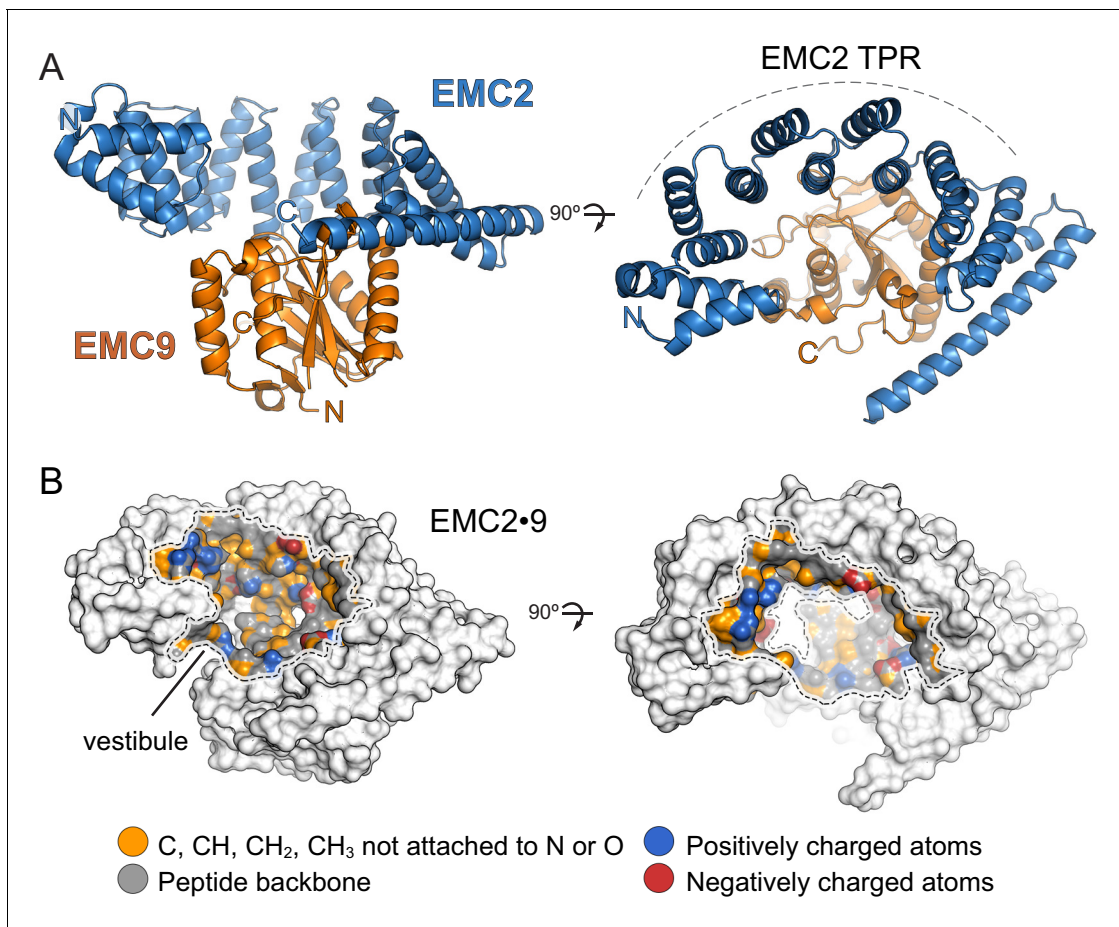


Figure 2. Structure of the EMC2•EMC9 complex. (A) Crystal structure of the EMC2•EMC9 heterodimer (PDB: 6Y4L). The heterodimer consists of EMC2 (residues 11–274), depicted in blue, and EMC9 (residues 1–200), depicted in orange. The TPR-repeat motif of EMC2 is indicated. (B) Physicochemical properties of a vestibule in the EMC2•EMC9 complex. Surface rendering of crystal structures coloured according to chemical properties (Hagemans et al., 2015).

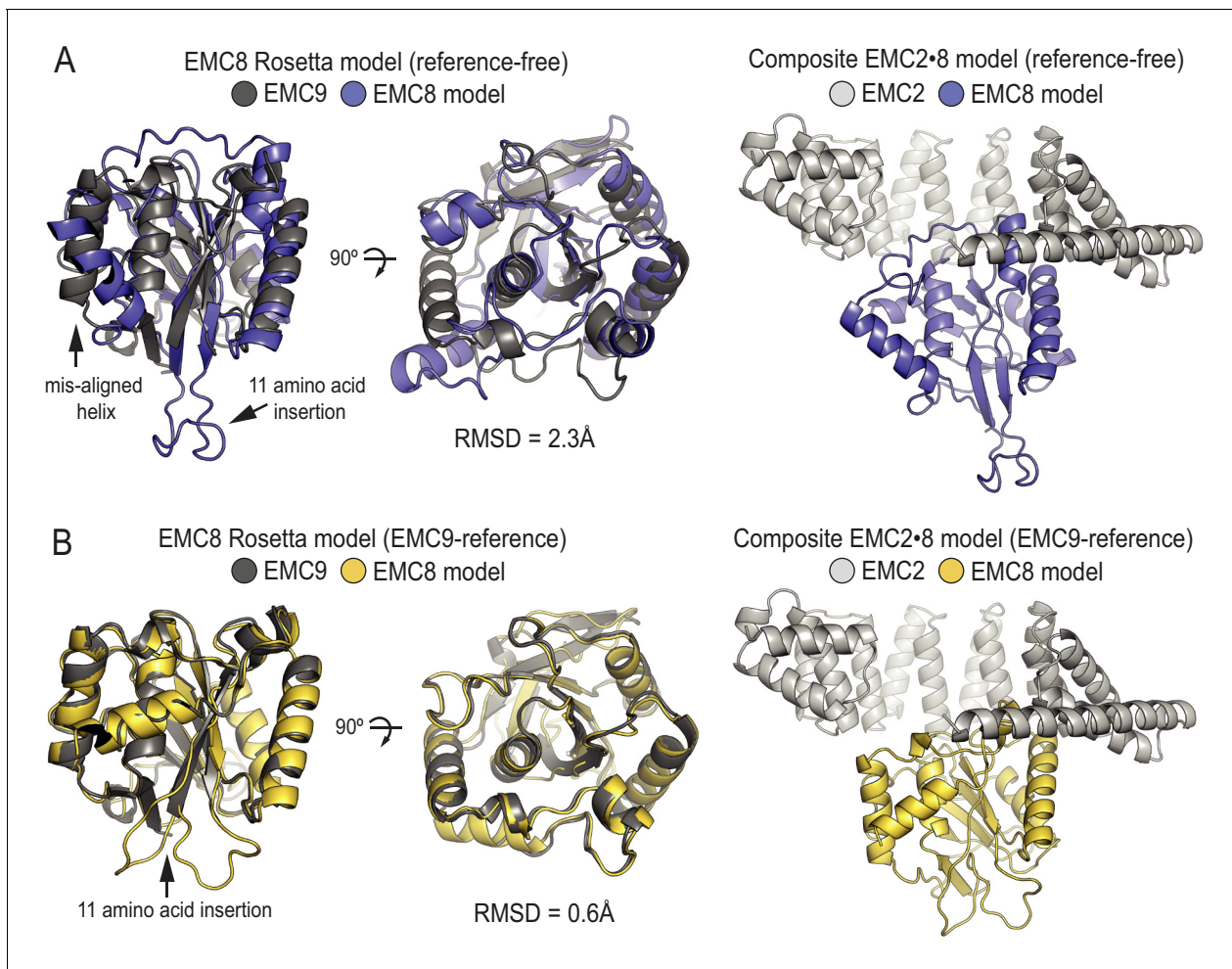


Figure 2—figure supplement 1. Protein structure prediction confirms structural homology between EMC9 and EMC8. (A) Reference free ab initio modelling of EMC8 (blue) using Robetta results in a model with an overall RMSD of 2.3 Å when aligned to the EMC9 crystal structure (grey). The majority of this disparity is the result of a single misaligned helix and the location of an 11 residue insertion in a loop in EMC8. Docking of the predicted structure into the EMC2•EMC9 structure in the place of EMC9 results in a good fit with no steric violations. (B) Ab initio modelling of EMC8 (yellow) using Robetta with the EMC9 crystal structure (grey) as a template results in an improved model with an overall RMSD of 0.6 Å. As with the reference-free model, the template-based model of EMC8 yields the same core fold and fits well when substituted into the EMC2•EMC9 crystal structure.

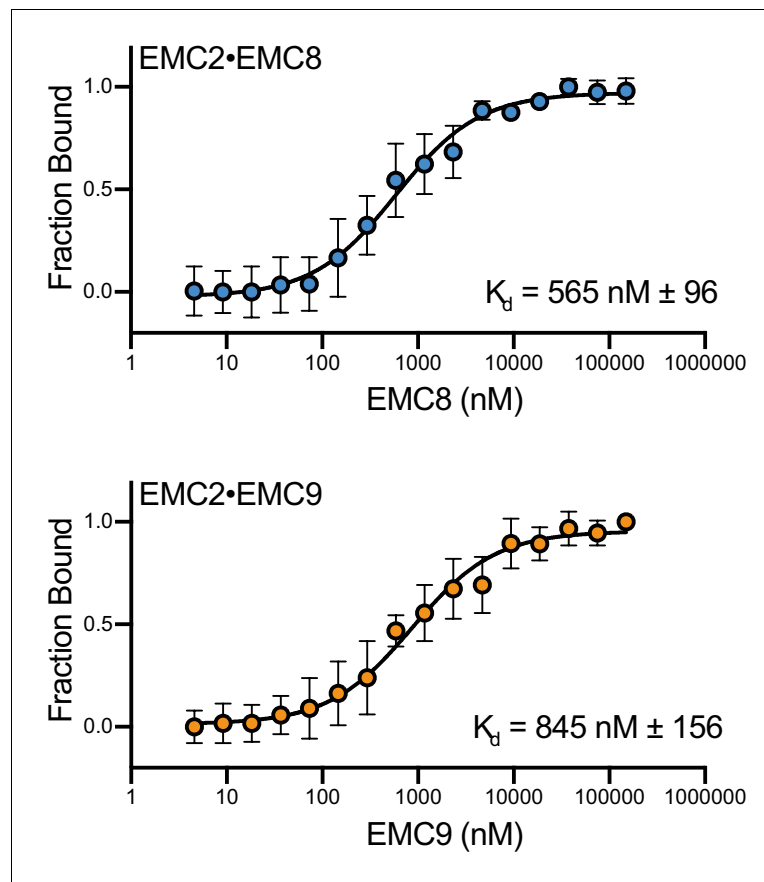


Figure 2—figure supplement 2. EMC8 and EMC9 have similar affinities for EMC2. Microscale thermophoresis was used to determine affinities of the heterodimer. An endogenous solvent accessible cysteine in EMC2 was labeled with maleimide-OG488 and held at a fixed concentration of 50 nM. Either EMC8 or EMC9 were titrated from 4 nM - 150 μ M. Affinities calculated for EMC2•EMC8 (blue) and EMC2•EMC9 (orange) were ~550 nM and ~850 nM respectively. Thermophoresis data were fit using NanoTemper's in software quadratic model, and equivalent results were attained using nonlinear fit model in Prism GraphPad. Data consist of two biological replicates with two technical replicates each. Data points represent the mean and error bars represent standard deviation.

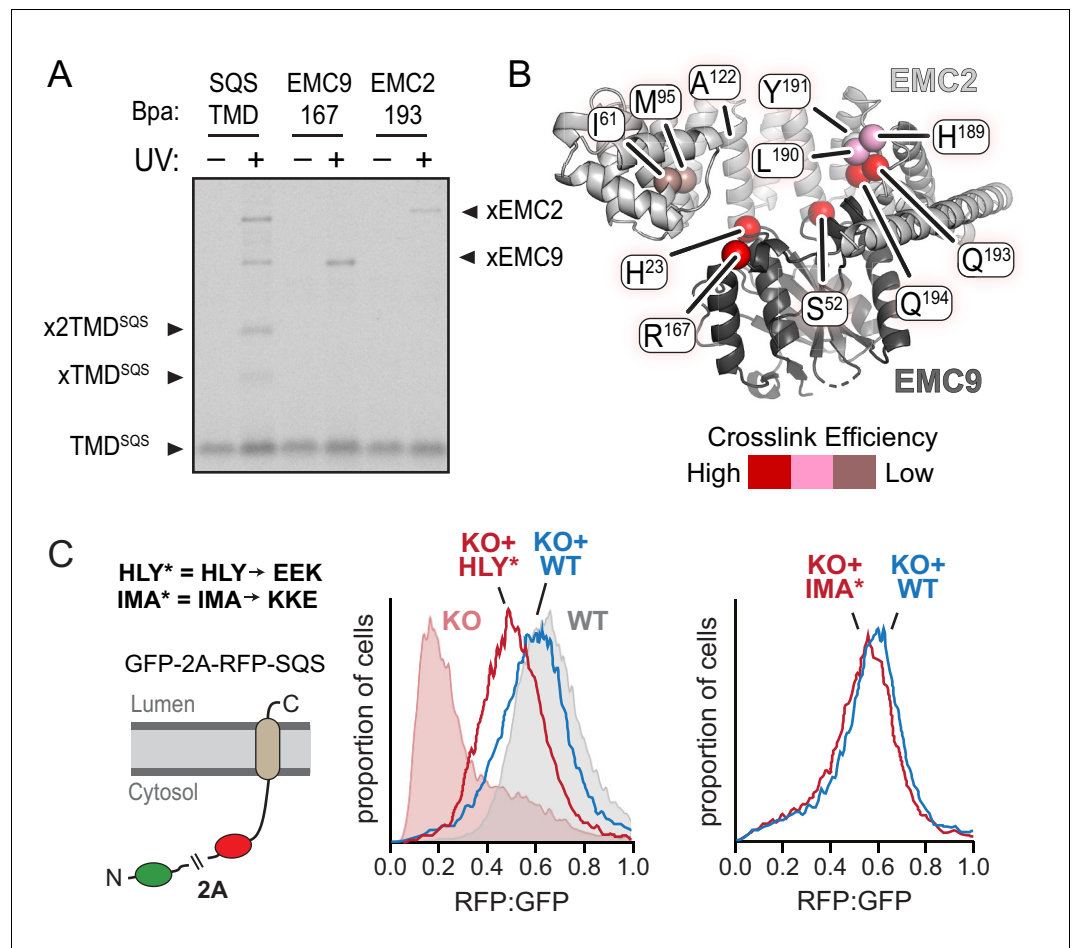


Figure 3. Functional analysis of the EMC2•EMC9 cytosolic vestibule. (A) ³⁵S-methionine-labeled TMD of SQS was mixed with recombinant purified EMC2•EMC9 complex as in **Figure 1** and analyzed directly or after UV irradiation as indicated. The photocrosslinking amino acid benzoyl-phenylalanine (Bpa) was incorporated into either the SQS substrate, EMC9 (at codon 167), or EMC2 (at codon 193) as indicated. Both R167 in EMC9 and Q193 in EMC2 line the vestibule. (B) Bpa was incorporated at different positions within the vestibule of the EMC2•EMC9 complex and crosslinking efficiency to SQS was determined as in panel A. Locations of the Bpa are annotated as spheres on the EMC2•EMC9 heterodimer. The sphere colors correspond to the intensity of the resulting crosslink. Position 191 (obscured behind L190 in this view) showed no crosslinking to substrate, consistent with its rearward facing location. (C) Shown on the left is a diagram of the dual color reporter for insertion of the TMD of SQS. Expression of this reporter results in a free GFP protein and an RFP-tagged SQS protein due to ribosomal skipping at the a viral 2A sequence. The left graph shows flow cytometry analysis of the SQS reporter in WT cells (grey), EMC2 knockout (KO) cells (shaded pink), KO cells complemented with WT EMC2 (blue line), and KO cells complemented with the HLY* EMC2 mutant (red line). The data are represented as histograms of the RFP to GFP ratio. The right graph shows a comparison of the SQS reporter in KO cells complemented with either WT EMC2 or the IMA* EMC2 mutant. The mutated amino acids, whose positions are shown in panel B, are: H189E, L190E, Y191K, I61K, M95K, and A122E.

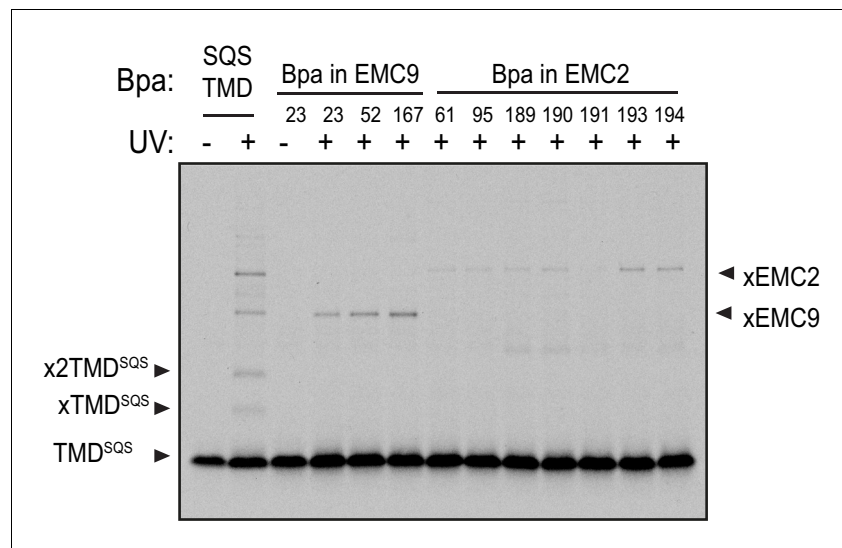


Figure 3—figure supplement 1. Crosslinking analysis of the EMC2•EMC9 cytosolic vestibule. ³⁵S-methionine-labeled TMD of SQS was mixed with recombinant purified EMC2•EMC9 complex as in **Figure 1** and analyzed directly or after UV irradiation as indicated. The photocrosslinking amino acid benzoyl-phenylalanine (Bpa) was incorporated into either the SQS substrate, EMC9, or EMC2 as indicated.

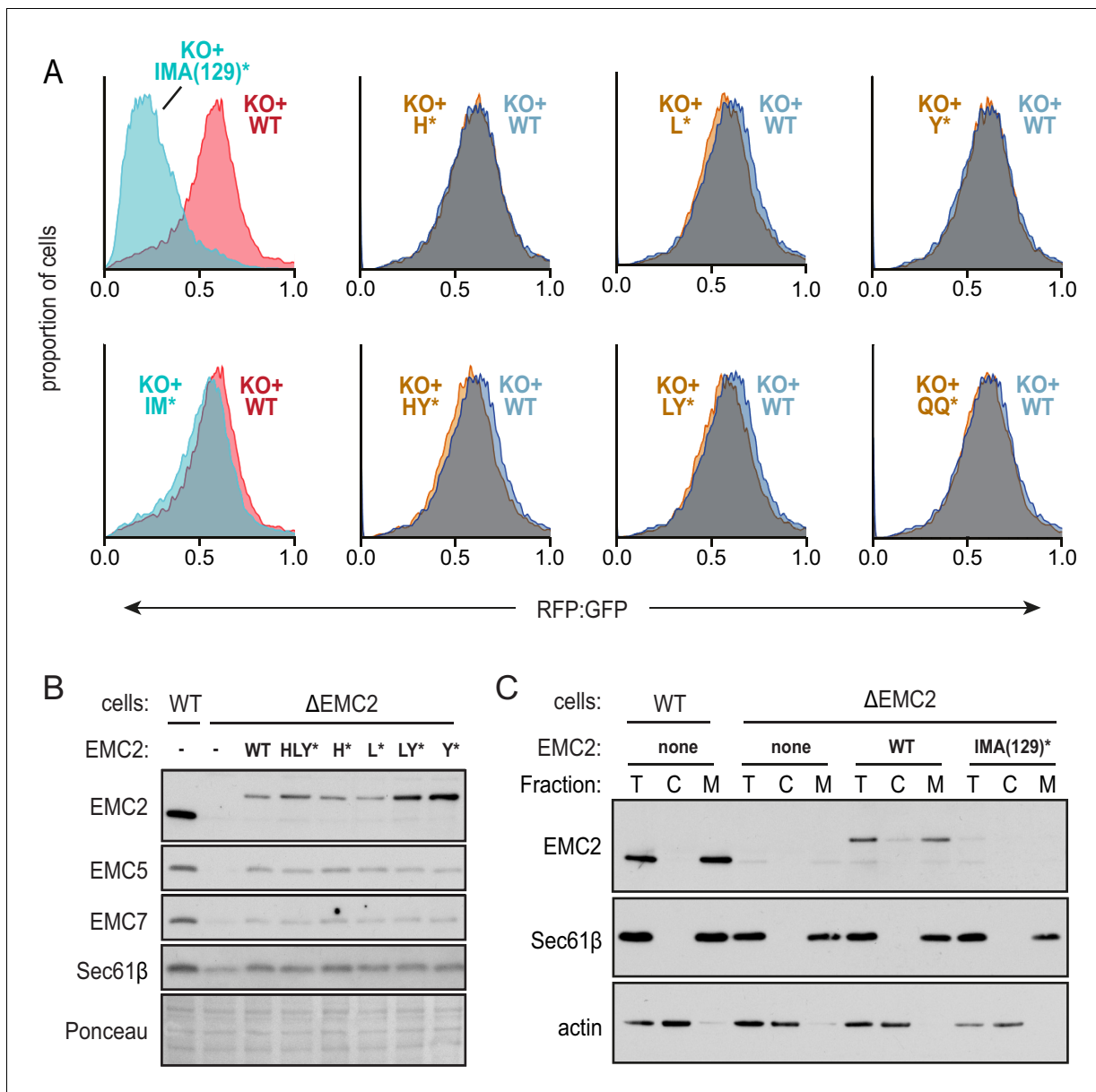


Figure 3—figure supplement 2. Functional analysis of the EMC2•EMC9 cytosolic vestibule. (A) Expression of the dual color SQS reporter in EMC2 knockout cells complemented with either WT EMC2 or the indicated EMC2 mutants. The data are represented as histograms of the RFP to GFP ratio. The two left graphs are from a different experiment from the remainder of graphs (which were analyzed together) and are colored differently to indicate this. The knockout-like phenotype of the IMA(129)^{*} mutation is apparently due to the A129K mutation since the IM^{*} mutation shows almost complete rescue. The mutated amino acids are: I61K, M95K, A129K, H189E, L190E, Y191K, Q193E, and Q194K. (B) Immunoblotting of whole cell extracts from WT or ΔEMC2 cells transfected with the indicated EMC2 constructs. A section of the Ponceau-stained blot shows equal loading. (C) WT and ΔEMC2 cells transfected with the indicated EMC2 constructs were fractionated into cytosol (C) and membrane (M) fractions and analyzed by immunoblotting for EMC2 relative to an equivalent amount of total (T) cell lysate. The membrane marker Sec61β and cytosolic marker actin were included as controls. Note that relative to WT EMC2 expressed in ΔEMC2 cells or endogenous EMC2 in WT cells, the IMA(129)^{*} mutant is very poorly expressed. This is presumably due to its degradation secondary to inefficient assembly with other EMC subunits, explaining its strong phenotype in the flow cytometry assay.

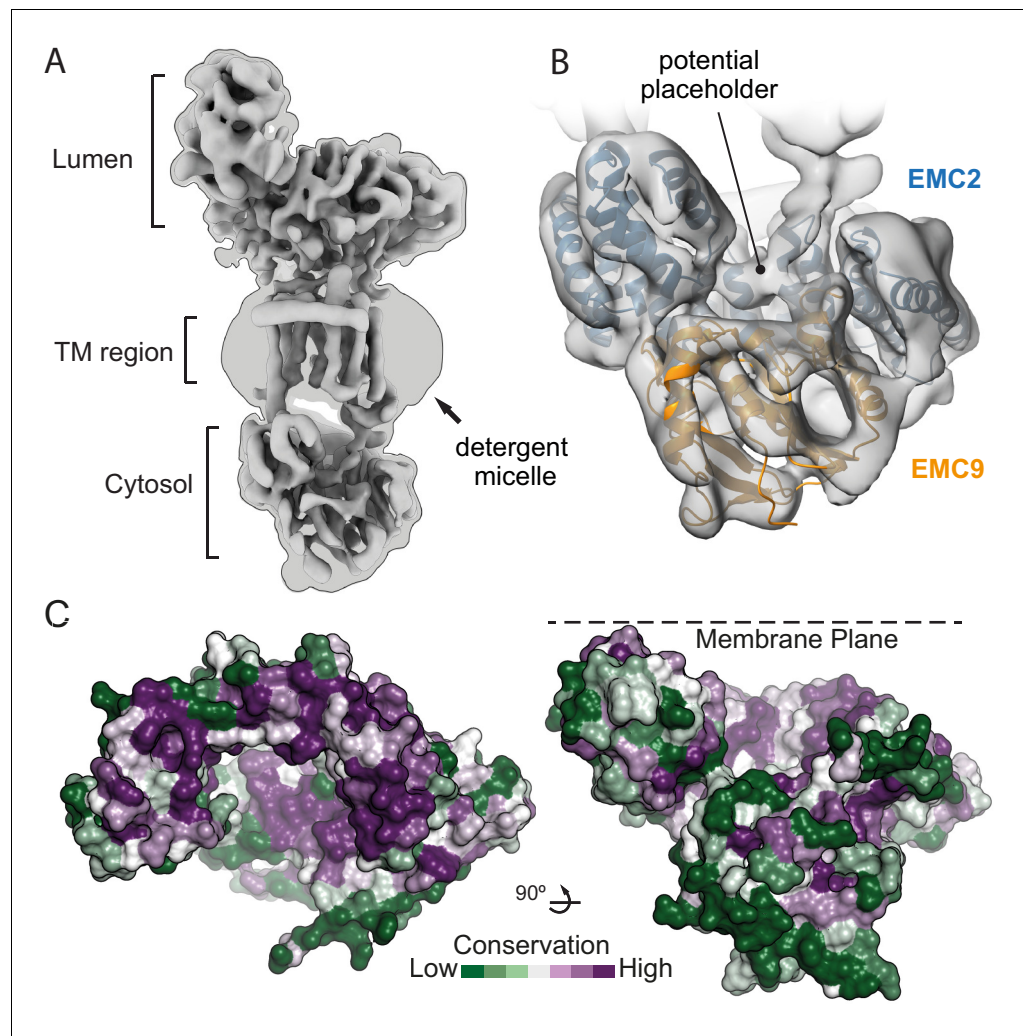


Figure 4. The position of EMC2•EMC9 within native EMC. (A) Cryo-EM map of the EMC at 6.4 Å resolution reveals the architecture of the complex. The map is shown at two contour levels: a stringent contour that illustrates secondary structure features (0.21) superimposed with a liberal contour that shows the detergent micelle (0.15). (B) Refinement of the EMC2•EMC9 crystal structure (blue and orange) into the cytosolic density (grey) using Flex-EM, Coot and PHENIX. (C) Surface rendering of the EMC2•EMC9 crystal structure coloured by residue conservation (Ashkenazy et al., 2016) from highly conserved (purple) to weakly conserved (green). The top rim of EMC2's TPR is highly conserved. This surface faces the membrane and regions of it interact with other EMC-subunits. The substrate binding vestibule also exhibits high conservation in comparison to the remaining solvent exposed surface of the EMC2•EMC9 heterodimer.

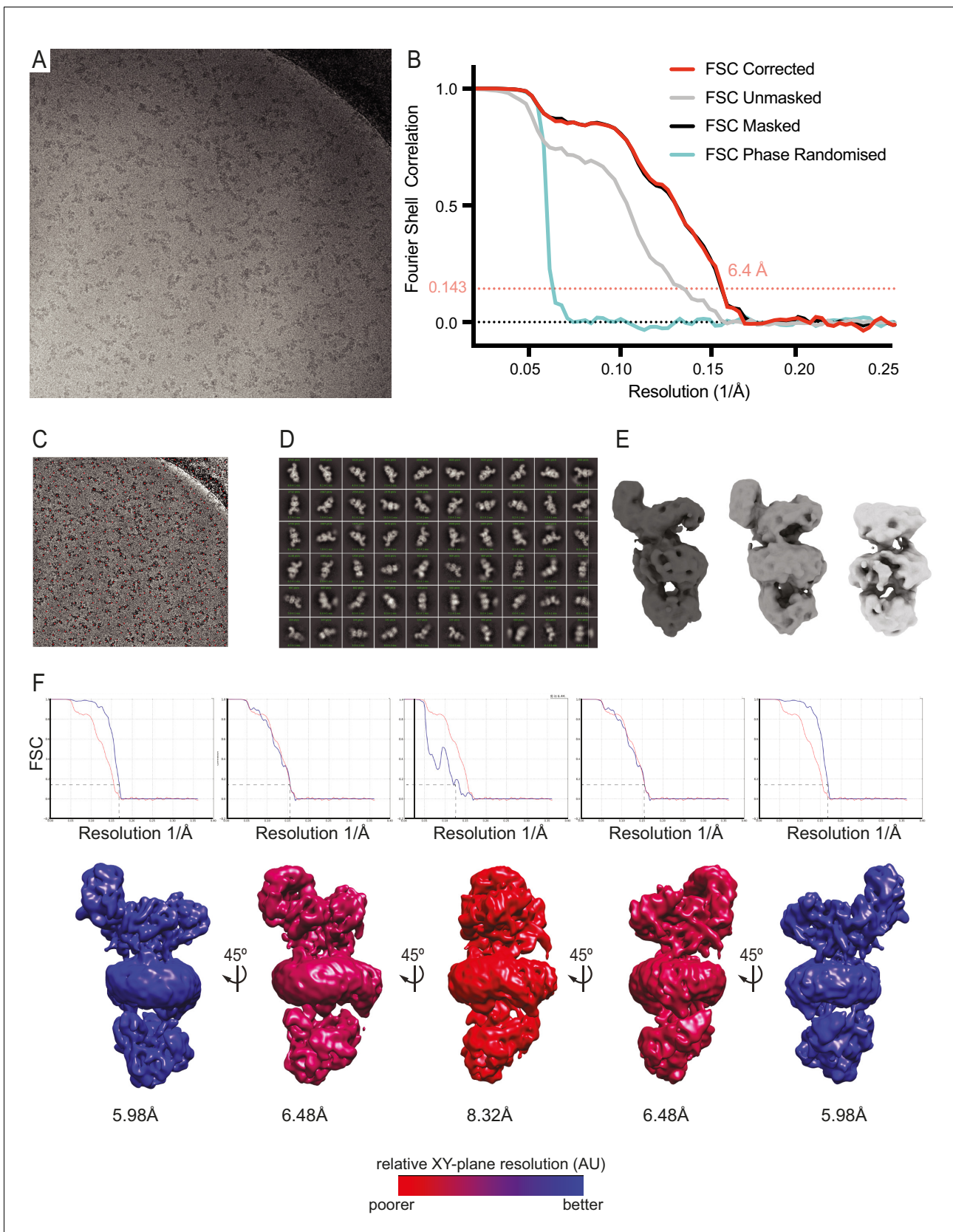


Figure 4—figure supplement 1. Cryo-electron microscopy data processing. (A) Micrograph showing EMC particles imaged at $-0.5 \mu\text{m}$ defocus with volta phase plate at a nominal pixel size of 1.38 \AA using a K2 summit detector. (B) Fourier Shell Correlation (FSC) curves for phase randomised, Figure 4—figure supplement 1 continued on next page

Figure 4—figure supplement 1 continued

unmasked, masked and corrected 3D reconstructions. Representative images of template-based autopicking (C), 2D-classification (D) and 3D classification (E) stages during the processing of dataset 1. (F) 3D FSC curves demonstrating resolution anisotropy during reconstruction of a map of the human EMC. 5 views of the same map are shown in rotational increments of 45°. Each view is accompanied by an FSC plot comparing the FSC in the respective viewing orientation (blue) to the global FSC (red). Resolution values given below the maps are calculated from the FSC plots above the maps using the 0.143 criterion. Maps are coloured from blue (high) to red (low) to indicate reconstruction resolution from different angles. Maps generated using 3DFSC implemented in CryoSPARC (Tan et al., 2017).

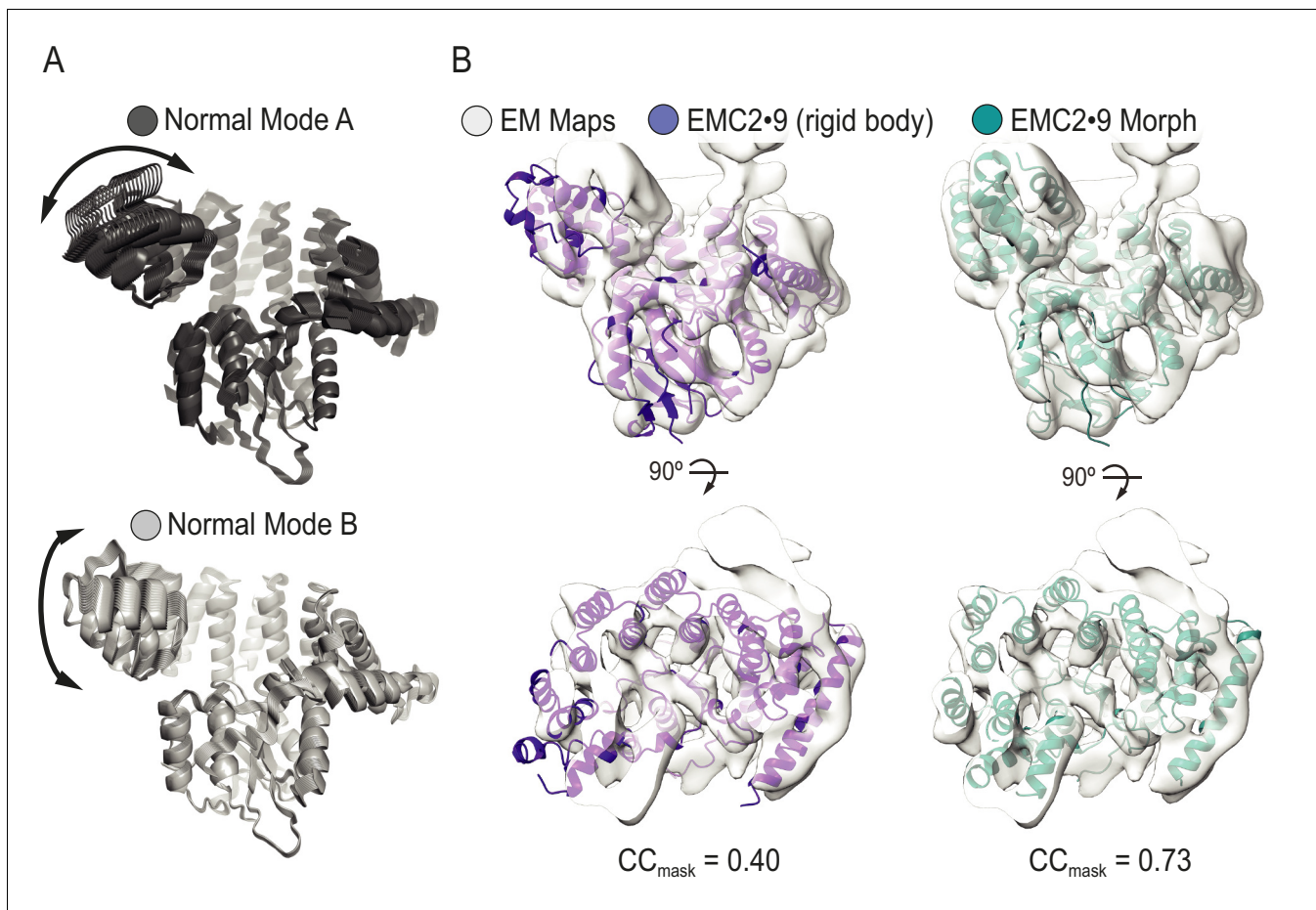


Figure 4—figure supplement 2. Normal mode analysis and flexible fitting of the EMC2•EMC9 crystal structure into the full EMC cryo-EM map. (A) Normal mode analysis using the el Némo server (Suhre and Sanejouand, 2004) reveals conformational flexibility in the N-terminal region of EMC2•EMC9. All calculated states are shown for modes A and B in dark and light grey, respectively. Arrows indicate direction of motion. (B) Rigid body docking of the EMC2•EMC9 crystal structure (purple) into the experimental cryo-EM density (grey) is in overall good agreement with the exception of EMC2's N-terminal region. The fit is improved (teal) following flexible fitting conducted with Flex-EM, and real space refinement in Coot and PHENIX. The map to model CC_{mask} calculated in PHENIX real space refine increases from 0.40 for the rigid body fit to 0.73 for the flexibly fitted structure. During flexible fitting the largest movements occurred in the N-terminal helical bundle of EMC2, consistent with the normal mode analysis.

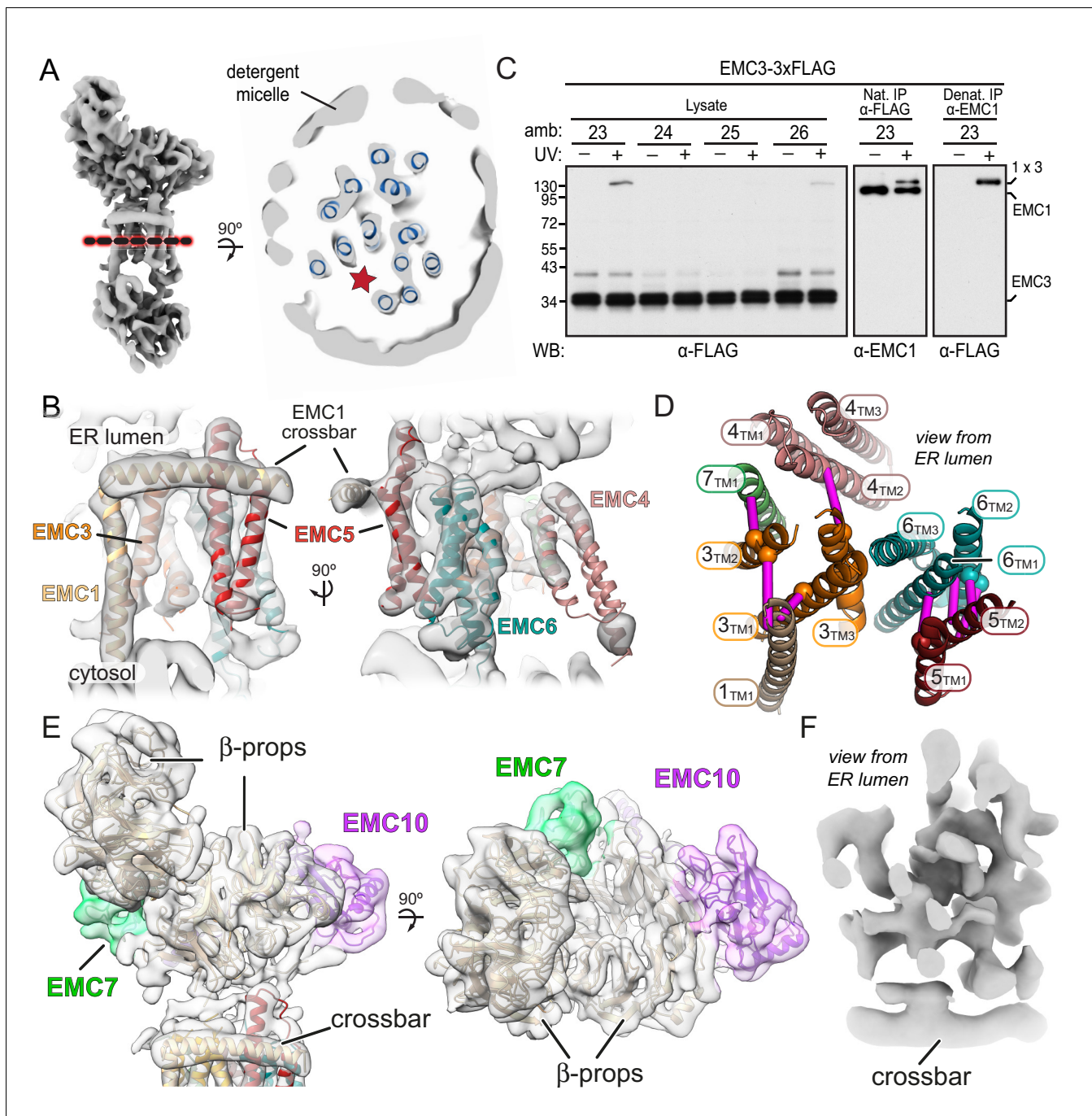


Figure 5. A composite model of EMC's membrane and luminal domains. (A) Cross section through the TMD region of the cryo-EM map of EMC (light grey) identifies 13 helix-like densities (dark blue) that define an intramembrane groove (red star) open to the lipid bilayer. Density corresponding to the annular detergent micelle is indicated. (B) Ab initio trRosetta models fitted into cryo-EM density for the EMC membrane domain are shown for EMC1 (wheat), EMC3 (orange), EMC4 (light pink), EMC5 (red), EMC6 (teal), and EMC7 (green). The horizontal density at the membrane-lumen interface termed the 'crossbar' is assigned to an amphipathic helix in EMC1. (C) EMC3-3xFLAG constructs containing amber codons (amb.) at the indicated positions were expressed together with an amber suppressor tRNA and cognate aminoacyl-tRNA synthetase that accepts the UV-activated crosslinking amino acid 3'-azibutyl-N-carbamoyl-lysine (AbK). Cells were left untreated or irradiated with UV and analyzed by immunoblotting for EMC3-3xFLAG. A prominent UV-dependent crosslink is seen from position 23 and to a lesser extent, position 26. Native FLAG immunoprecipitation (IP) recovers EMC1 (indicating that EMC3-3xFLAG is incorporated into EMC), which shifts with UV. Denaturing EMC1 IP confirms the crosslinked product contains both EMC1 and EMC3. (D) TMD helices positioned based on docking of ab initio models overlaid with AbK-mediated crosslinks (see **Figure 5—figure supplement 3**). The positions where AbK was incorporated are shown as spheres, with magenta lines showing the closest point of the target protein in the model. (E) Composite model of the EMC luminal domain generated by ab initio modelling in trRosetta and real-space refinement in PHENIX. Cryo-EM density is shown in grey. (F) Cryo-EM density map of the luminal domain. *Figure 5 continued on next page*

Figure 5 continued

EM density has been colored according to subunit identity and the composite lumenal domain model accounts for almost all the lumenal EM density. (F) Cross-section through EMC at the plane of the membrane-lumen interface illustrating that a pore is not evident across the membrane. All EM data visualized in UCSF ChimeraX with EM maps contoured at 0.15 (panel A) and 0.21 (panels B, E, and F) with hide dust setting of 10. Panel D was generated in PyMOL.

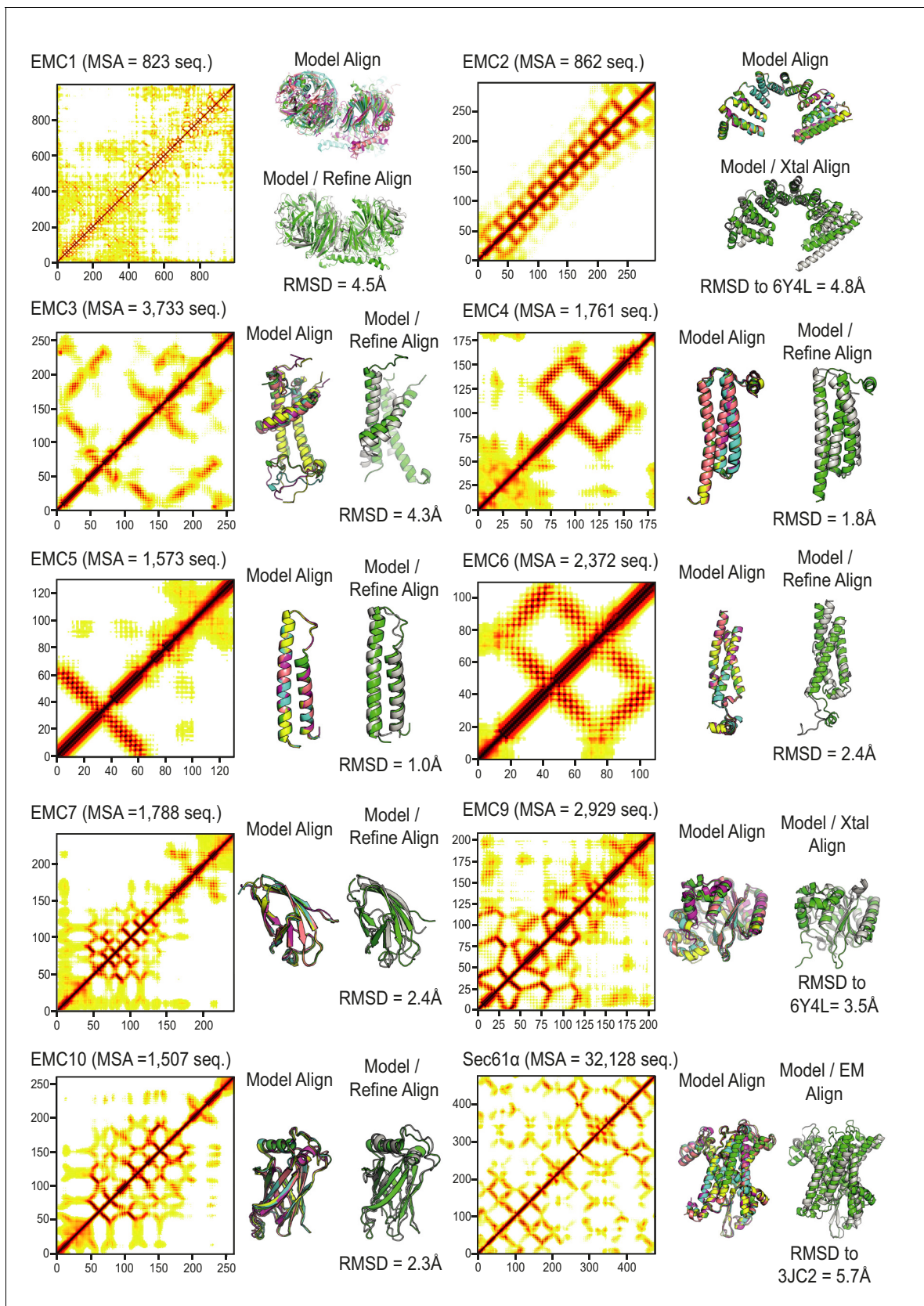


Figure 5—figure supplement 1. Ab initio prediction of EMC subunit structure and flexible fitting using real-space refinement. Structures were predicted for 9 EMC subunits using the trRosetta server. The best ranked solution for each subunit was then docked into the cryo-EM density and *Figure 5—figure supplement 1 continued on next page*

Figure 5—figure supplement 1 continued

subject to real-space refinement in PHENIX. Number of alignments in the multiple sequence alignment (MSA) is indicated along with the resulting distance contact heat map. The top five ranked models were aligned for the core folds used in EM docking to show modelling confidence. A comparison of the model before and after real space refinement illustrates only subtle changes (green vs grey, RMSD is indicated). Soluble and membrane proteins (EMC2, EMC9, and Sec61a) were modelled and compared to their known structures.

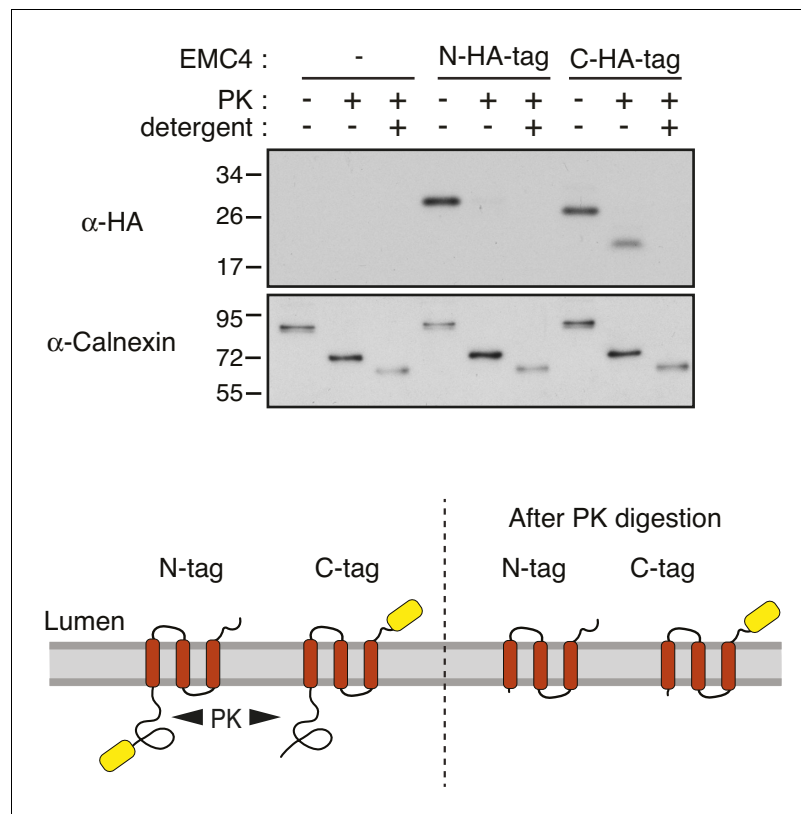


Figure 5—figure supplement 2. Protease-protection analysis of EMC4 topology. Microsomes isolated from HEK293T cells transfected with an empty vector, EMC4 with an N-terminal 3xHA tag (N-HA-tag), or EMC4 with a C-terminal 3xHA tag (C-HA-tag) were analyzed by a protease-protection assay. Equal aliquots were left untreated, treated with proteinase K (PK), or treated with PK in the presence of detergent (Triton X-100) and analyzed by immunoblotting for the HA tag and the N terminus of calnexin. The N-terminal 3xHA tag was accessible to PK in the absence of detergent, suggesting the N-terminus is exposed to the cytosol. The C-terminal 3xHA tag was protected from PK digestion in the absence, but not presence, of detergent, suggesting that the C-terminus is in the ER lumen. This supports a three TMD model in which the loop between TMD2 and TMD3 is too short to be accessible to PK. A schematic of the results and proposed topology of EMC4 is also shown. In principle, the protease-digestion results are also consistent with a single-spanning topology with the N-terminus facing the cytosol. However, this is unlikely given that all topology prediction algorithms predict at least two TMDs and trRosetta predicts a three helix bundle. Note that the core region in the N-terminal luminal domain of calnexin is resistant to complete PK digestion.

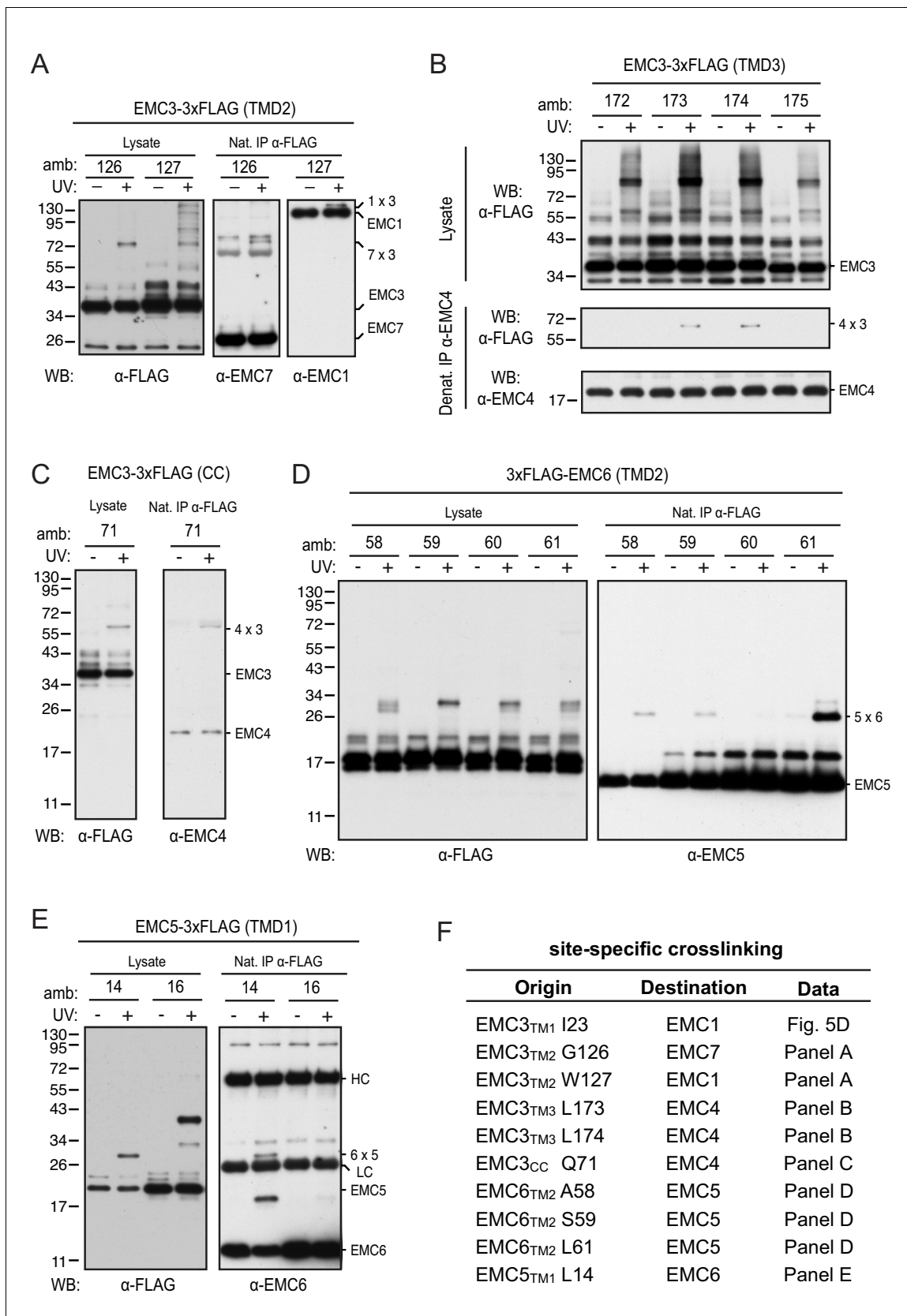


Figure 5—figure supplement 3. Site-specific photocrosslinking between EMC subunits. (A–E) Experiments similar to that shown in **Figure 5C** are displayed for the key crosslinks depicted in **Figure 5D**. In each case, the FLAG-tagged construct indicated on top and containing amber codons (amb.) *Figure 5—figure supplement 3 continued on next page*

Figure 5—figure supplement 3 continued

at the indicated positions was expressed together with an amber suppressor tRNA and cognate aminoacyl-tRNA synthetase that accepts the UV-activated crosslinking amino acid 3'-azibutyl-N-carbamoyl-lysine (AbK). Cells were left untreated or irradiated with UV and analyzed by immunoblotting for the FLAG epitope in total lysate as indicated. In panels A, C, D, and E, lysates were subjected to native FLAG immunoprecipitation (IP) and immunoblotted for the indicated targets. In panel B, lysates were subjected to denaturing EMC4 IP and blotted for the FLAG tag and EMC4 as indicated. (F) Summary of the observed crosslinks and where the primary data are displayed in the figure. Note that various positions from which crosslinks to a particular target was not observed are not shown for brevity.

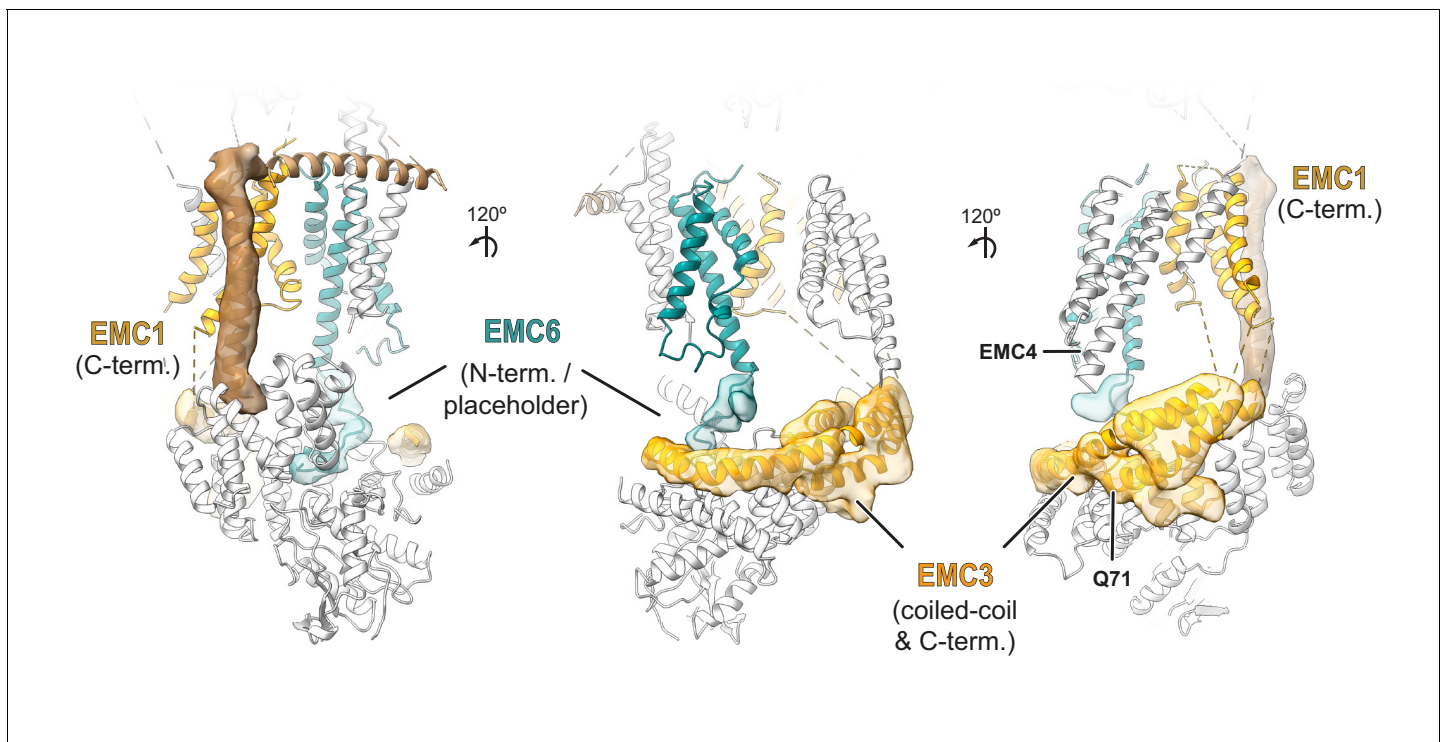


Figure 5—figure supplement 4. Provisional assignment of non-EMC2•EMC9 cytosolic density. Semi-automated segmentation in UCSF chimera revealed regions of density in the cytosolic domain of the cryo-EM map of the EMC which are not accounted for by the EMC2•EMC9 heterodimer (translucent densities). Based on the position and connectivity to the respective TMDs in the membrane portions of this density were provisionally assigned to EMC1 (brown) and the N-terminus of EMC6 (teal) which is proposed to form a ‘placeholder’ in the cytosolic substrate binding pocket. A predicted three-helix bundle formed from a coiled-coil between TMD1 and TMD2 of EMC3 interacting with a helix in EMC3’s C-terminal tail was fit into a similarly shaped and sized density. In this fitted model, Q71 of EMC3 is surface-exposed and near where the N-terminal tail of EMC4 might reside. AbK incorporated at Q71 shows a UV-dependent crosslink to EMC4 (**Figure 5—figure supplement 3**), consistent with our assignments. The remaining density is continuous with the C-terminal helix in this three-helix bundle and was therefore provisionally assigned to the rest of EMC3’s C-terminal tail.

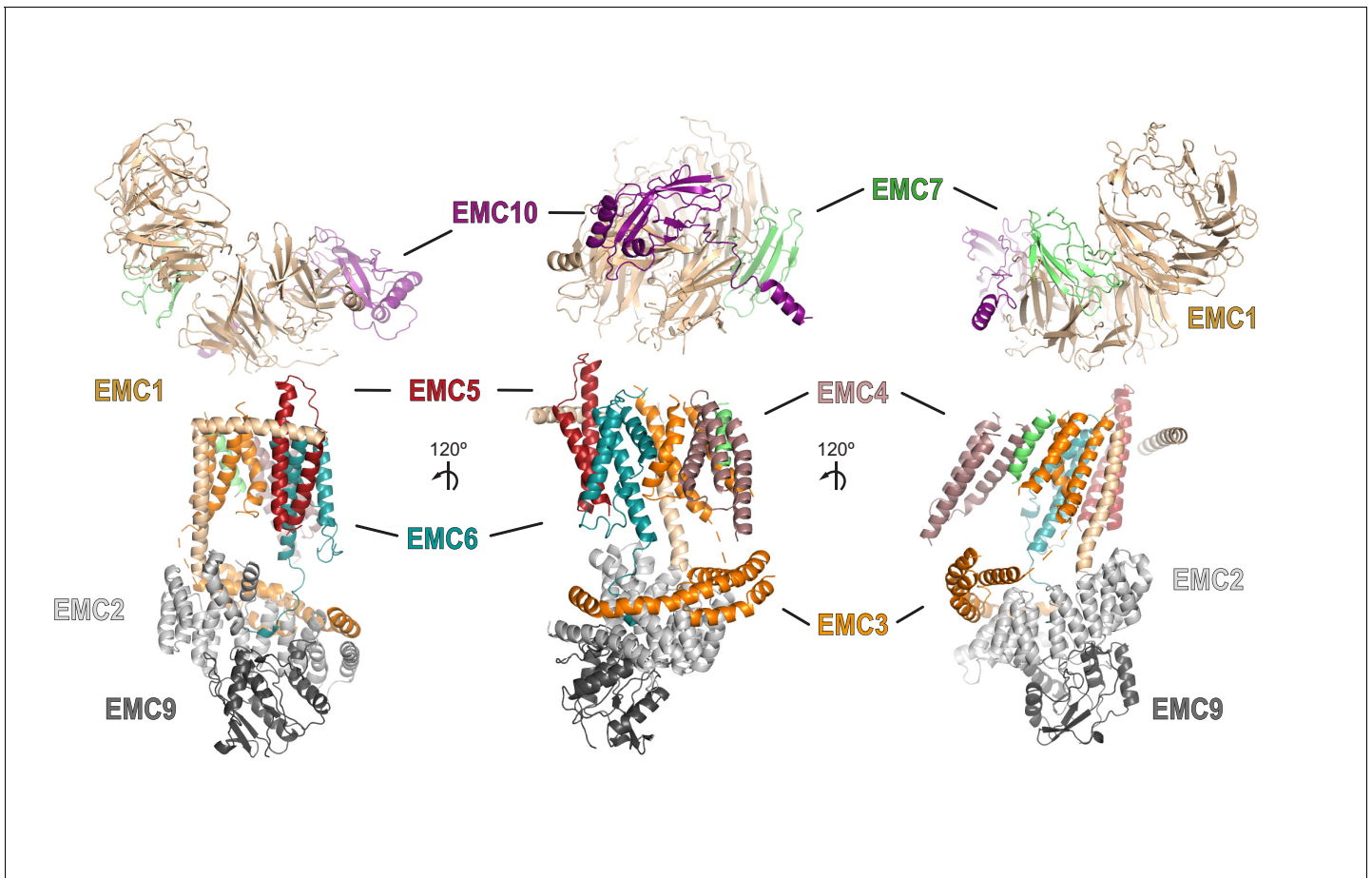


Figure 5—figure supplement 5. Views of the composite EMC model. Shown are multiple views of the composite EMC model which contains all nine EMC subunits. Where possible, the same color scheme as used throughout the main figures was employed with EMC1 (tan), EMC2 (light grey), EMC3 (orange), EMC4 (dark pink), EMC5 (red), EMC6 (teal), EMC7 (lime), EMC9 (dark grey), and EMC10 (purple).

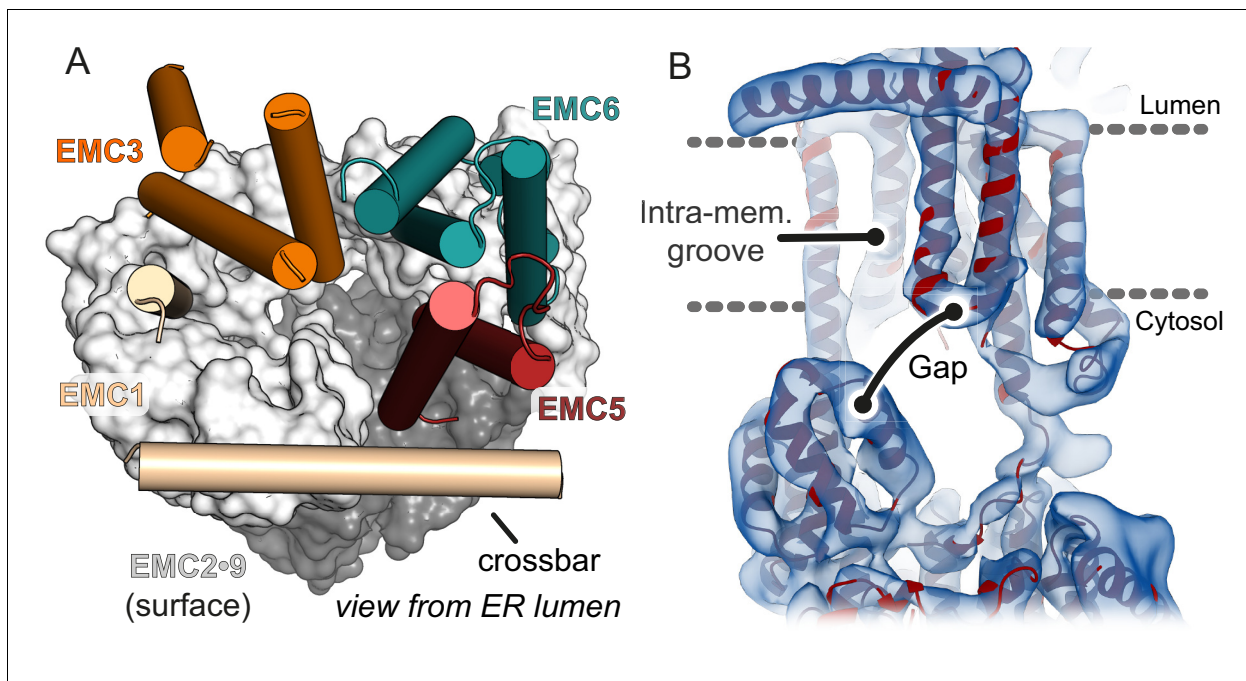


Figure 6. The cytosolic vestibule leads into an intramembrane groove. **(A)** The cytosolic vestibule in EMC2•EMC9 (grey surface) is contiguous with a horseshoe-shaped intramembrane groove lined by the TMDs of EMC1, EMC3, EMC5, and EMC6. The luminal domain of EMC and the TMDs of EMC7 and EMC4 are not shown for clarity. **(B)** Fitting the composite EMC model (red) into cryo-EM density (blue) reveals a ~ 6–12 Å gap between the cytosolic and membrane domains of EMC. A substrate with a TMD in the intramembrane groove can be released from EMC with the cytosolic flanking polypeptide passing through this gap.

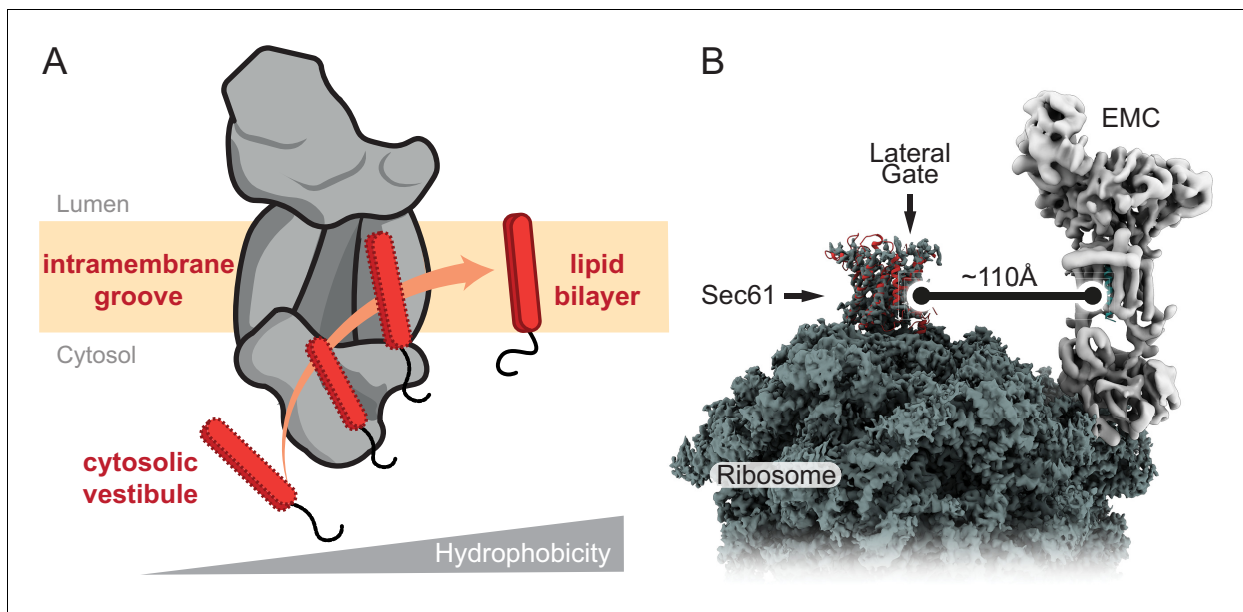


Figure 7. Model for EMC-mediated TMD insertion. (A) TMDs follow a gradient of increasing hydrophobicity from the cytosol through EMC and into the lipid bilayer. TMDs first enter the cytosolic vestibule formed by either the EMC2•EMC8 or EMC2•EMC9 complex. This vestibule leads into an intramembrane groove that is laterally open toward the lipid bilayer. (B) The ribosome-Sec61 complex is depicted with EMC at the same scale in the same membrane plane. The minimum distance from the Sec61 lateral gate to the intramembrane groove of EMC is ~110 Å. This implies that EMC can only mediate TMD insertion before ribosomes bind to Sec61 or if the polypeptide downstream of a TMD can span this distance. EMC cannot capture TMDs at the Sec61 lateral gate.