UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Structure-function and a genetic roadmap of the mammalian Endoplasmic Reticulum Membrane Protein Complex (EMC)

Permalink
https://escholarship.org/uc/item/21v5h18h

Author
Popova, Katerina Dimitrova

Publication Date
2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Structure-function and a genetic roadmap of the mammalian Endoplasmic Reticulum Membrane Protein Complex (EMC)

by

Katerina Dimitrova Popova

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of

DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Peter Walter
9C44B4D1D50740D...

Peter Walter

Chair

DocuSigned by:
Jonathan S Weissman
[illegible]4E2...

Jonathan S Weissman

DocuSigned by:
Carol Gross
7AFBE9046E0744C...

Carol Gross

Committee Members

Copyright 2021

by

Katerina Popova

Dedication

I dedicate this work to my parents, Diana and Dimitar, who love me unconditionally, and to Paul Baxter and Glenda Ross, my American parents, without whom I would not be here today.

Acknowledgments

I want to thank all the people who have supported me unconditionally throughout my life and my PhD.

First and foremost, I would like to thank my advisor, Jonathan S. Weissman, for his mentorship both in lab and in life. Thank you for allowing me to pursue my interests inside and outside of the lab and for helping me grow as a scientist and as a person. You were always there for me and my family. Thank you for that.

Thank you to my thesis committee members, Carol Gross and Peter Walter, for believing in me and for supporting me throughout the challenges and successes of the last five years. Carol, thank you for being the best role model a woman in the sciences can ask for and for inspiring me with your fearlessness.

I want to thank outstanding UCSF faculty for their mentorship and guidance with my research questions: Rushika Perera, Hani Goodarzi, and Luke Gilbert.

Huge thanks to the following people for welcoming me to the Weissman lab and making sure I had the best experience possible during my rotation and beyond: Dan Santos, Liz Boydston, Marco Jost, Jeff Hussmann, Jin Chen, James Nunez. Special thanks to my fellow Bulgarian, Kamena Kostova, for her scientific and life advice, communicated for the most part in Bulgarian. Thank you to the following kind humans for an incredibly stimulating experience during the writing of the EMC structure-function paper – Lakshmi Miller-Vedam, Bastian Brauning, and last but not least Jess Bonnar for her hard work and dedication to helping complete the project. Thank you to the following people who have supported the lab throughout my tenure here: Manny de Vera, Cristen Muresan, and Joan Kanter.

Thank you to amazing collaborators and scientists – Hana Retallack, Dan Itzhak, and Andreas Puschnik.

Special thanks to David Bartel and the whole Bartel lab for welcoming me into their lab – I loved the stimulating scientific environment and your kindness.

Thank you to my scientific mentors and friends at Smith College, who believed in me more than I believed in myself as a shy undergrad – Rob Dorit, Kalina Dimova, and Cristina Suarez.

Thank you to my friends across the world who were always there for me despite distance and time zones – Zlatka Madjarova, Tzvetelina Zhelyazkova, Dragomir Kichev, Nevena Eliseeva and Svilen Iskrov, Vesi Hristova, Ivan Vodenicharski and Anuta Vodenicharska, Galya Iliev, Sanda Soltan and Ken Wong, Brendan Smalec, Zak Doric, Silvan Wittwer. Thank you to my awesome and talented BMS classmates for an amazing five years.

Thank you to my American family – Glenda Ross and Paul Baxter, who passed away while I was in graduate school; Sharon Yarbrough and Gary Ubben – thank you for believing in that 12-year old kid who spoke broken English and for bringing her to the US and setting her up for success. I owe you everything, and I love you.

Thank you from the bottom of my heart to my parents, Diana and Dimitar, for always trusting me and supporting me in my endeavors. I love you so so much. To my brother, Dimitar, and the little ones, Dimitar and Zlatin, for always making me laugh.

Thank you to my love, Johnny Yu. Another reason why I'm so happy to have come to UCSF is that I met you. You are my best friend and the best partner I could ask for. Thank you for guiding me to become the best I can be and for helping me get over my fears. I love you.

Contributions

This dissertation represents the scientific work I contributed to in inception and execution. **Chapter I** is an introduction into the world of membrane protein folding. **Chapter II** of this dissertation constitutes a highly collaborative project published in 2020. I am grateful to all my co-authors: Lakshmi Miller-Vedam, Bastian Bräuning, Nicole Schirle Oakdale, Jessica Bonnar, Jesuraj Prabu, Elizabeth Boydston, Natalia Sevillano, Matthew Shurtleff, Robert Stroud, Charles Craik, Brenda Schulman, Adam Frost, Jonathan Weissman. **Chapter III** is represents exciting unpublished research, I have worked on in the past five years.

Miller-Vedam LE, Bräuning B, Popova KD, Schirle Oakdale NT, Bonnar JL, Prabu JR, Boydston EA, Sevillano N, Shurtleff MJ, Stroud RM, Craik CS, Schulman BA, Frost A, Weissman JS. Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients. Elife. 2020 Nov 25;9:e62611. doi: 10.7554/eLife.62611. PMID: 33236988.

Structure-function and a genetic roadmap of the mammalian Endoplasmic Reticulum Membrane Protein Complex (EMC)

Katerina Dimitrova Popova

Abstract

Cells are surrounded and contained by a plasma membrane consisting of a double layer of fats and proteins. These proteins monitor and facilitate the movement of food, oxygen and messages in and out of the cell, and help neighboring cells communicate. Membrane proteins are manufactured in a cell compartment called the endoplasmic reticulum. Cellular machines called ribosomes manufacture proteins that need to be secreted or embedded into the cell's membranes. As these proteins are made, they are pulled into the endoplasmic reticulum so they can be stabilized, folded correctly and inserted in the membrane. A cellular machine in this compartment's membrane that aids this process is the endoplasmic reticulum membrane protein complex (EMC). The EMC is a large protein complex made up of ten subunits in mammalian cells. EMC's intricate structure and large range of clients it stabilizes underscore its importance in maintaining protein homeostasis. Our studies illuminate the structural and mechanistic basis of EMC's multifunctionality and point to its role in differentially regulating the biogenesis of distinct client protein classes. Beyond the molecular role of EMC, we build a genetic roadmap of the EMC in mammalian cells and helping decipher its crucial role in maintaining cholesterol homeostasis. Together, this work lays the framework for understanding both the molecular and physiological role of the EMC in mammalian cells.

Table of Contents

Chapter 1: Membrane protein folding at the Endoplasmic Reticulum (ER)……..1

References……5

Chapter 2: Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients…. ……11

References……..67

Chapter 3: Building a genetic roadmap of the mammalian Endoplasmic Reticulum Membrane Protein Complex (EMC)……..……...85

References…….....108

List of Figures

Figure 2.1: Experimental strategy for the dissection of EMC function……..……..……...35

Figure 2.2: Overall structures of yeast and human EMC……..……36

Figure 2.3: The EMC cytoplasmic domain contains conserved functional Interfaces and may engage C-tail-anchored clients directly…….….38

Figure 2.4: The EMC houses two transmembrane cavities with conserved core structures and distinct accessibilities…….……..40

Figure 2.5: EMC houses an insertase module centered on EMC3 in the Gated membrane cavity……..……..41

Figure 2.6: A lipid-filled cavity in the EMC transmembrane domain stabilizes disparate client proteins……..……..42

Figure 2.7: The large EMC lumenal domain is the site for several annotated disease mutations……..43

Figure 2.8: Model of coordinated EMC functions……..……..45

Figure S2.1: Overview of strain creation and functional assays……..……...47

Figure S2.2: Western blots for EMC1 and EMC2……..……...49

Figure S2.3: Western blots for EMC3 and EMC5…….……51

Figure S2.4: Flow cytometry for mutations in the EMC cytoplasmic domain……..……53

Figure S2.5: Additional Flow cytometry for mutations in the EMC cytoplasmic domain.....55

Figure S2.6: Flow cytometry and western blot of gated cavity mutants……..…….57

Figure S2.7: Additional flow cytometry of gated cavity mutants……..……..59

Figure S2.8: Flow cytometry of lipid-filled cavity mutants…….……..60

Figure S2.9: Additional flow cytometry of lipid-filled cavity mutants……..……..62

Figure S2.10: Flow cytometry of lumenal domain mutants…….……64

Figure S2.11: Additional flow cytometry of lumenal domain mutants……..……66

Figure 3.1: Functional knockout and rescue of EMC function in K562 dCas9 KRAB cell line……..94

Figure 3.2: Genetic interaction screen identifies genes synthetic with EMC……..…….95

Figure 3.3: Perturb-Seq identifies functional mammalian EMC clusters…….……96

Figure 3.4: KO clusters superimposed on the rescue dataset show poor correlation……...97

Figure 3.5: Transcriptional signatures inform cluster identity……98

Figure 3.6: StringDB visualization of Cluster 1……99

Figure 3.7: StringDB visualization of Cluster 2……..100

Figure 3.8: Gene expression of cholesterol-related genes in Cluster 2 in comparison with rest of clusters……...101

Figure 3.9: StringDB visualization of Cluster 3…….….102

Figure 3.10: StringDB visualization of Cluster 4……103

Figure 3.11: Synthetic score from GI screen plotted against protein/mRNA ratio of EMC4 Rescue/EMC KO……..……104

Figure 3.12: Colorimetric-based quantification of total cholesterol and cholesteryl esters……..……...105

Figure 3.13: EMC4 KO and ECM4 Rescue cells stained for cholesterol distribution……....106

Chapter 1: Membrane protein folding at the Endoplasmic Reticulum (ER)

Integral membrane proteins make up a quarter of the human proteome and they serve diverse and critical cellular roles, including signal transduction, lipid biosynthesis, adhesion, and transport of molecules across the bilayer. In eukaryotic cells, the endoplasmic reticulum (ER) serves as the primary site of integral membrane protein synthesis, targeting (co- or post-translationally), insertion, folding, and quality control (Ellgaard et al., 2016; Costa et al., 2018). However, the features of membrane-spanning regions (e.g. low hydrophobicity, charged residues, non-optimal lengths, lipid- and ion-binding sites and hairpins or kinked transmembrane helices) that mediate important functions pose particular challenges for transmembrane protein biosynthesis and folding. Consequently, membrane protein biogenesis is prone to failure, and this can lead to cellular stress and disease (Marinko et al., 2019). Although some aspects of these processes are well understood, recent discoveries of new machineries that contribute to these events has shed new light on an increasingly complex and adaptable system that drives high-fidelity membrane protein biogenesis (Phillips et al., 2021).

The machinery necessary for membrane protein synthesis is located in the cytosol so initial membrane protein synthesis by ribosomes begins in the aqueous environment of the cytosol where membrane proteins are insoluble and prone to misfolding and aggregation leading to pathologies. As the hydrophobic signal sequence or the first transmembrane domain (TMD) of a membrane protein emerges from the cytosolic ribosome, the signal recognition particle (SRP) captures it and delivers the ribosome nascent chain complex to the Sec61 complex at the ER membrane. As the emerging nascent polypeptide enters the Sec61 translocon, it samples the lipid environment, and stretches of high hydrophobicity are thought to partition through the lateral gate to embed in the membrane (Grudnik et al., 2009; Park et al., 2012). The role of Sec61 in membrane protein insertion and secretory protein translocation has been dissected by a combination of genetics, biochemistry and structural approaches and was the only well characterized machinery for membrane protein handling at the ER for a long time (Phillips et al.,

2021). Here, I will focus on several recently characterized mammalian machineries that ensure folding, insertion, and stabilization of membrane proteins at the ER.

The ER membrane protein complex (EMC) emerged as a conserved player in membrane protein biogenesis when it was first identified in *Saccharomyces cerevisiae* as an abundant and stable multi-protein membrane complex whose disruption results in stress mirroring that caused by misfolded membrane proteins (Jonikas et al., 2009). Loss of the EMC in mammalian cells is associated with failed biogenesis and degradation of a subset of membrane proteins (Christianson et al., 2012). Accordingly, the EMC has been implicated in several mechanistically distinct steps of membrane protein biogenesis, stabilization, and quality control (Bircham et al., 2011; Richard et al., 2013; Satoh et al., 2015; Savidis et al., 2016; Shurtleff et al., 2018; Volkmar et al., 2019; Tian et al., 2019). One well-established EMC function is as an insertase for terminal transmembrane helices. EMC's insertase function has been demonstrated for two classes of clients: low hydrophobicity tail-anchored proteins (i.e. those that contain C-terminal membrane anchors) and a subset of polytopic transmembrane proteins in which the first helix is inserted with the N-terminus in the lumen (Guna et al., 2018; Chitwood et al., 2018).

However, many studies indicate EMC functions beyond initial insertion of N- or C-terminal helices. The EMC has been implicated in the biogenesis and stability of many membrane protein classes that do not require a terminal transmembrane insertase (Bircham et al., 2011; Louie et al., 2012; Richard et al., 2013; Shurtleff et al., 2018; Coelho et al., 2019; Luo et al., 2002; Volkmar et al., 2019; Talbot et al., 2019; Petkovic et al., 2020). Recent studies have shown that the EMC is required for stability of internal transmembrane helices of human and viral multi-pass membrane proteins (Hiramatsu et al., 2019; Lin et al., 2019; Ngo et al., 2019; Coelho et al., 2019; Xiong et al., 2020). Additionally, the human EMC (hEMC) physically interacts with the NS4A-B region of the Dengue Virus polyprotein following Sec61-dependent translocation and signal peptidase cleavage, suggesting roles in post-translational stabilization of polytopic membrane proteins (Ngo et al., 2019; Lin et al., 2019). Similarly, the *S. cerevisiae* EMC (yEMC) co-

immunoprecipitated with full-length polytopic transmembrane clients, including Pma1p (Luo et al., 2002), Mrh1p, and Fks1p (Shurtleff et al., 2018). In addition to varying types of transmembrane protein clients, the EMC also associates with a range of regulatory factors, including many general and substrate-specific chaperones in the cytoplasm and in the ER lumen (Bagchi et al., 2016; Coelho et al., 2019; Kudze et al., 2018; Richard et al., 2013; Shurtleff et al., 2018).

The complex architecture of the EMC provides additional support for multifunctionality in membrane protein biogenesis beyond the small subset of membrane proteins discussed above. The EMC is an eight (yeast) or nine (mammalian) component, 248–284 kDa complex with considerable mass in the ER lumen, membrane, and cytosol. The cytoplasmic domain contains conserved tetratricopeptide repeats (TPR) repeats in EMC2, and the human complex accommodates an additional subunit, EMC8/9, whose function is not yet understood. The ER lumenal domain in yeast does not contain an N-terminal EMC1 expansion seen in hEMC. Notably, the ER lumenal domain has been linked to a number of disease-associated phenotypes (Junes-Gill et al., 2011; Probert et al., 2015; Harel et al., 2016; Abu-Safieh et al., 2013; Diamantopoulou et al., 2017; Marquez et al., 2020), and presents the possibility of additional functions for the human lumenal domain. One EMC subunit (EMC3) shares limited sequence homology with a family of insertases that are evolutionarily related to the bacterial insertase YidC (Samuelson et al., 2000; Kumazaki et al., 2014; Borowska et al., 2015; Anghel et al., 2017), perhaps explaining the insertase function of the complex.

Beyond the characterization of EMC as a membrane protein chaperone, recent structural characterization of an expanded translocon apparatus comprising Sec61, TMCO1, CCDC47 and the Nicalin-TMEM147-NOMO complex suggests that an array of accessory molecules around the Sec61 translocon may help to stabilize or insert TMDs of multipass membrane proteins (McGilvray et al., 2020). In addition, a novel protein complex called the PAT10 complex, comprising PAT10 and Asterix, has been demonstrated to engage poorly hydrophobic TMDs to promote the biogenesis of multiple multipass membrane protein substrates (Chitwood et al.,

2020). The complex binds to unshielded hydrophilic or polar residues before releasing the substrate once all the TMDs have entered the bilayer and folding can be completed (Phillips et al, 2021).

The discoveries of these numerous membrane-bound molecular machineries engaged in stabilization and folding of membrane proteins in recent years underscores that simplified models of the co-translational biogenesis process have not captured the complex array of accessory factors required for fidelity in inserting and stabilizing membrane proteins in mammalian cells. Although we have been identifying important parts of the puzzle that is maintaining membrane protein homeostasis, we are still only scratching the surface. Key questions remain unanswered: What is the dependency of different types of membrane proteins on different insertion/stabilization machineries and factors? How are these different pathways coordinated to maximize efficient folding? What is the physiological role of these machineries beyond stabilization of a particular set of membrane proteins? Some of these questions are addressed in the following chapters. Our studies illuminate the structural and mechanistic basis of EMC's multifunctionality and point to its role in differentially regulating the biogenesis of distinct client protein classes.

REFERENCES

Abu-Safieh L, Alrashed M, Anazi S, Alkuraya H, Khan AO, Al-Owain M, Al-Zahrani J, Al-Abdi L, Hashem M, Al-Tarimi S, Sebai MA, Shamia A, Ray-Zack MD, Nassan M, Al-Hassnan ZN, Rahbeeni Z, Waheeb S, Alkharashi A, Abboud E, Al-Hazzaa SA, et al. 2013. Autozygome-guided exome sequencing in retinal dystrophy patients reveals pathogenetic mutations and novel candidate disease genes. *Genome Research* 23:236–247. DOI: https://doi.org/10.1101/gr.144105.112, PMID: 23105016

Anghel SA, McGilvray PT, Hegde RS, Keenan RJ. 2017. Identification of Oxa1 homologs operating in the eukaryotic endoplasmic reticulum. *Cell Reports* 21:3708–3716. DOI: https://doi.org/10.1016/j.celrep.2017.12. 006, PMID: 29281821

Bagchi P, Inoue T, Tsai B. 2016. EMC1-dependent stabilization drives membrane penetration of a partially destabilized non-enveloped virus. *eLife* 5:e21470. DOI: https://doi.org/10.7554/eLife.21470, PMID: 28012275

Bircham PW, Maass DR, Roberts CA, Kiew PY, Low YS, Yegambaram M, Matthews J, Jack CA, Atkinson PH. 2011. Secretory pathway genes assessed by high-throughput microscopy and synthetic genetic array analysis. *Molecular BioSystems* 7:2589. DOI: https://doi.org/10.1039/c1mb05175j, PMID: 21731954

Borowska MT, Dominik PK, Anghel SA, Kossiakoff AA, Keenan RJ. 2015. A YidC-like protein in the archaeal plasma membrane. *Structure* 23:1715–1724. DOI: https://doi.org/10.1016/j.str.2015.06.025, PMID: 26256539

Chitwood PJ, Juszkiewicz S, Guna A, Shao S, Hegde RS. 2018. EMC is required to initiate accurate membrane protein topogenesis. *Cell* 175:1507–1519. DOI: https://doi.org/10.1016/j.cell.2018.10.009, PMID: 30415835

Christianson JC, Olzmann JA, Shaler TA, Sowa ME, Bennett EJ, Richter CM, Tyler RE, Greenblatt EJ, Wade Harper J, Kopito RR. 2012. Defining human ERAD networks through

an integrative mapping strategy. *Nature Cell Biology* 14:93–105. DOI: https://doi.org/10.1038/ncb2383

Coelho JPL, Stahl M, Bloemeke N, Meighen-Berger K, Alvira CP, Zhang ZR, Sieber SA, Feige MJ. 2019. A network of chaperones prevents and detects failures in membrane protein lipid bilayer integration. *Nature Communications* 10:672. DOI: https://doi.org/10.1038/s41467-019-08632-0, PMID: 30737405

Costa EA, Subramanian K, Nunnari J, Weissman JS. 2018. Defining the physiological role of SRP in protein-targeting efficiency and specificity. *Science* 359:689–692. DOI: https://doi.org/10.1126/science.aar3607, PMID: 29348368

Diamantopoulou A, Sun Z, Mukai J, Xu B, Fenelon K, Karayiorgou M, Gogos JA. 2017. Loss-of-function mutation in Mirta22/Emc10 rescues specific schizophrenia-related phenotypes in a mouse model of the 22q11.2 deletion. *PNAS* 114:E6127–E6136. DOI: https://doi.org/10.1073/pnas.1615719114, PMID: 28696314

Ellgaard L, McCaul N, Chatsisvili A, Braakman I. 2016. Co- and Post-Translational protein folding in the ER: co-and Post-Translational protein folding in the ER. *Traffic* 17:615–638. DOI: https://doi.org/10.1111/tra.12392, PMID: 26947578

Grudnik, Przemyslaw, Bange, Gert and Sinning, Irmgard. "Protein targeting by the signal recognition particle", vol. 390, no. 8, 2009, pp. 775-782. https://doi.org/10.1515/BC.2009.102

Harel T, Yesil G, Bayram Y, Coban-Akdemir Z, Charng WL, Karaca E, Al Asmari A, Eldomery MK, Hunter JV, Jhangiani SN, Rosenfeld JA, Pehlivan D, El-Hattab AW, Saleh MA, LeDuc CA, Muzny D, Boerwinkle E, Gibbs RA, Chung WK, Yang Y, et al. 2016. Monoallelic and biallelic variants in EMC1 identified in individuals with global developmental delay, Hypotonia, scoliosis, and cerebellar atrophy. *The American Journal of Human Genetics* 98:562–570. DOI: https://doi.org/10.1016/j.ajhg.2016.01.011, PMID: 26942288

Hiramatsu N, Tago T, Satoh T, Satoh AK. 2019. ER membrane protein complex is required for the insertions of late-synthesized transmembrane helices of Rh1 in Drosophila photoreceptors. *Molecular Biology of the Cell* 30: 2890–2900. DOI: https://doi.org/10.1091/mbc.E19-08-0434, PMID: 31553680

Jonikas MC, Collins SR, Denic V, Oh E, Quan EM, Schmid V, Weibezahn J, Schwappach B, Walter P, Weissman JS, Schuldiner M. 2009. Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* 323:1693–1697. DOI: https://doi.org/10.1126/science.1167983, PMID: 19325107

Junes-Gill KS, Gallaher TK, Gluzman-Poltorak Z, Miller JD, Wheeler CJ, Fan X, Basile LA. 2011. hHSS1: a novel secreted factor and suppressor of glioma growth located at chromosome 19q13.33. *Journal of Neuro-Oncology* 102:197–211. DOI: https://doi.org/10.1007/s11060-010-0314-6, PMID: 20680400

Kudze T, Mendez-Dorantes C, Jalloh CS, McClellan AJ. 2018. Evidence for interaction between Hsp90 and the ER membrane complex. *Cell Stress and Chaperones* 23:1101–1115. DOI: https://doi.org/10.1007/s12192-018-0908-z, PMID: 29808299

Kumazaki K, Kishimoto T, Furukawa A, Mori H, Tanaka Y, Dohmae N, Ishitani R, Tsukazaki T, Nureki O. 2014. Crystal structure of Escherichia coli YidC, a membrane protein chaperone and insertase. *Scientific Reports* 4: 7299. DOI: https://doi.org/10.1038/srep07299, PMID: 25466392

Lin DL, Inoue T, Chen YJ, Chang A, Tsai B, Tai AW. 2019. The ER membrane protein complex promotes biogenesis of dengue and zika virus Non-structural Multi-pass transmembrane proteins to support infection. *Cell Reports* 27:1666–1674. DOI: https://doi.org/10.1016/j.celrep.2019.04.051, PMID: 31067454

Louie RJ, Guo J, Rodgers JW, White R, Shah N, Pagant S, Kim P, Livstone M, Dolinski K, McKinney BA, Hong J, Sorscher EJ, Bryan J, Miller EA, Hartman JL. 2012. A yeast phenomic model for the gene interaction network modulating CFTR-DF508 protein

biogenesis. *Genome Medicine* 4:103. DOI: https://doi.org/10.1186/gm404, PMID: 23270647

Luo WJ, Gong XH, Chang A. 2002. An ER membrane protein, Sop4, facilitates ER export of the yeast plasma membrane [H+] ATPase, Pma1. *Traffic* 3:730–739. DOI: https://doi.org/10.1034/j.1600-0854.2002.31005.x, PMID: 12230471

Marinko JT, Huang H, Penn WD, Capra JA, Schlebach JP, Sanders CR. 2019. Folding and misfolding of human membrane proteins in health and disease: from single molecules to cellular proteostasis. *Chemical Reviews* 119:5537–5606. DOI: https://doi.org/10.1021/acs.chemrev.8b00532, PMID: 30608666

Marquez J, Criscione J, Charney RM, Prasad MS, Hwang WY, Mis EK, Garcı´a-Castro MI, Khokha MK. 2020. Disrupted ER membrane protein complex-mediated topogenesis drives congenital neural crest defects. *Journal of Clinical Investigation* 130:813–826. DOI: https://doi.org/10.1172/JCI129308, PMID: 31904590

McGilvray PT, Anghel SA, Sundaram A, Zhong F, Trnka MJ, Fuller JR, Hu H, Burlingame AL, Keenan RJ. 2020. An ER translocon for multi-pass membrane protein biogenesis. *eLife* 9:e56889. DOI: https://doi.org/10.7554/eLife.56889, PMID: 32820719

Ngo AM, Shurtleff MJ, Popova KD, Kulsuptrakul J, Weissman JS, Puschnik AS. 2019. The ER membrane protein complex is required to ensure correct topology and stable expression of flavivirus polyproteins. *eLife* 8:e48469. DOI: https://doi.org/10.7554/eLife.48469, PMID: 31516121

Park E, Rapoport TA. Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annu Rev Biophys*. 2012;41:21-40. doi: 10.1146/annurev-biophys-050511-102312. Epub 2011 Dec 16. PMID: 22224601.

Petkovic M, Oses-Prieto J, Burlingame A, Jan LY, Jan YN. 2020. TMEM16K is an interorganelle regulator of endosomal sorting. *Nature Communications* 11:3298. DOI: https://doi.org/10.1038/s41467-020-17016-8, PMID: 32620747

Phillips, B. P., & Miller, E. A. (2021). Membrane protein folding and quality control. In *Current Opinion in Structural Biology* (Vol. 69, pp. 50–54). Elsevier Ltd. https://doi.org/10.1016/j.sbi.2021.03.003

Probert F, Rice P, Scudamore CL, Wells S, Williams R, Hough TA, Cox IJ. 2015. ÅH NMR metabolic profiling of plasma reveals additional phenotypes in knockout mouse models. *Journal of Proteome Research* 14:2036–2045. DOI: https://doi.org/10.1021/pr501039k, PMID: 25849460

Richard M., Boulin T., Robert V. J. P., Richmond J. E. and Bessereau J.-L. (2013). Biosynthesis of ionotropic acetylcholine receptors requires the evolutionarily conserved ER membrane complex. *Proc. Natl. Acad. Sci. USA* 110, E1055-E1063. 10.1073/pnas.1216154110

Samuelson JC, Chen M, Jiang F, Moller I, Wiedmann M, Kuhn A, Phillips GJ, Dalbey RE. 2000. YidC mediates membrane protein insertion in Bacteria. *Nature* 406:637–641. DOI: https://doi.org/10.1038/35020586, PMID: 10949305

Satoh T, Ohba A, Liu Z, Inagaki T, Satoh AK. 2015. dPob/EMC is essential for biosynthesis of rhodopsin and other multi-pass membrane proteins in Drosophila photoreceptors. *eLife* 4:e06306. DOI: https://doi.org/10.7554/eLife.06306

Savidis G, McDougall WM, Meraner P, Perreira JM, Portmann JM, Trincucci G, John SP, Aker AM, Renzette N, Robbins DR, Guo Z, Green S, Kowalik TF, Brass AL. 2016. Identification of zika virus and dengue virus dependency factors using functional genomics. *Cell Reports* 16:232–246. DOI: https://doi.org/10.1016/j.celrep.2016.06.028, PMID: 27342126

Shurtleff MJ, Itzhak DN, Hussmann JA, Schirle Oakdale NT, Costa EA, Jonikas M, Weibezahn J, Popova KD, Jan CH, Sinitcyn P, Vembar SS, Hernandez H, Cox J, Burlingame AL, Brodsky JL, Frost A, Borner GH, Weissman JS. 2018. The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins. *eLife* 7:e37018. DOI: https://doi.org/10.7554/eLife.37018, PMID: 29809151

Talbot BE, Vandorpe DH, Stotter BR, Alper SL, Schlondorff JS. 2019. Transmembrane insertases and N-glycosylation critically determine synthesis, trafficking, and activity of the nonselective cation channel TRPC6. *Journal of Biological Chemistry* 294:12655–12669. DOI: https://doi.org/10.1074/jbc.RA119.008299, PMID: 31266804

Tian S, Wu Q, Zhou B, Choi MY, Ding B, Yang W, Dong M. 2019. Proteomic analysis identifies membrane proteins dependent on the ER membrane protein complex. *Cell Reports* 28:2517–2526. DOI: https://doi.org/10.1016/j.celrep.2019.08.006, PMID: 31484065

Volkmar N, Thezenas ML, Louie SM, Juszkiewicz S, Nomura DK, Hegde RS, Kessler BM, Christianson JC. 2019. The ER membrane protein complex promotes biogenesis of sterol-related enzymes maintaining cholesterol homeostasis. *Journal of Cell Science* 132:jcs223453. DOI: https://doi.org/10.1242/jcs.223453, PMID: 30578317

Xiong L, Zhang L, Yang Y, Li N, Lai W, Wang F, Zhu X, Wang T. 2020. ER complex proteins are required for rhodopsin biosynthesis and photoreceptor survival in Drosophila and mice. *Cell Death & Differentiation* 27: 646–661. DOI: https://doi.org/10.1038/s41418-019-0378-6, PMID: 31263175

Chapter 2: Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients

A version of this chapter was published as: Lakshmi E Miller-Vedam[1,2,3,4†,] Bastian Brauning[5†], Katerina D Popova[3,4,6†], Nicole T Schirle Oakdale[4‡], Jessica L Bonnar[3,4], Jesuraj R Prabu[5], Elizabeth A Boydston[4§], Natalia Sevillano[7#], Matthew J Shurtleff[4], Robert M Stroud[2], Charles S Craik[7], Brenda A Schulman[5*], Adam Frost[2*], Jonathan S Weissman[3,4,8*]. Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients

Affiliations

[1]Molecular, Cellular, and Computational Biophysics Graduate Program, University of California, San Francisco, San Francisco, United States

[2]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, United States

[3]Department of Biology, Whitehead Institute, MIT, Cambridge, United States

[4]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, United States

[5]Department of Molecular Machines and Signaling, Max Planck Institute of Biochemistry, Martinsried, Germany

[6]Biomedical Sciences Graduate Program, University of California, San Francisco, San Francisco, United States

[7]Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, United States

[8]Howard Hughes Medical Institute, Chevy Chase, United States

*For correspondence: schulman@biochem.mpg.de (BAS); adam.frost@ucsf.edu (AF); weissman@wi.mit.edu (JSW)

†These authors contributed equally to this work

ABSTRACT

Membrane protein biogenesis in the endoplasmic reticulum (ER) is complex and failure-prone. The ER membrane protein complex (EMC), comprising eight conserved subunits, has emerged as a central player in this process. Yet, we have limited understanding of how EMC enables insertion and integrity of diverse clients, from tail-anchored to polytopic transmembrane proteins. Here, yeast and human EMC cryo-EM structures reveal conserved intricate assemblies and human-specific features associated with pathologies. Structure-based functional studies distinguish between two separable EMC activities, as an insertase regulating tail-anchored protein levels and a broader role in polytopic membrane protein biogenesis. These depend on mechanistically coupled yet spatially distinct regions including two lipid-accessible membrane cavities which confer client-specific regulation, and a non-insertase EMC function mediated by the EMC lumenal domain. Our studies illuminate the structural and mechanistic basis of EMC's multifunctionality and point to its role in differentially regulating the biogenesis of distinct client protein classes.

INTRODUCTION

Here, we determined high-resolution cryo-EM structures of yeast EMC bound to a Fab and two conformations of the human EMC structure. Furthermore, we characterized the phenotypes of three distinct classes of EMC clients associated with a series of structure-based EMC mutants. Both yEMC and hEMC structures reveal a path for transmembrane helix insertion from the cytoplasm into the membrane via a conserved cavity. Our structures and mutants also revealed a second lipid-filled cavity with regions of importance for all three client types probed. Analysis of human disease mutations in hEMC1 and our structure-informed mutations enabled us to decouple the EMC insertase function from non-insertase functions and reveal a potential role of the EMC in differentially controlling the biogenesis of distinct classes of client proteins. These structure-function studies collectively establish that the EMC adopts a modular architecture enabling its diverse functions in membrane protein biogenesis.

RESULTS

Overview of strategy to comprehensively reveal EMC structure and function

To comprehensively dissect both conserved and species-specific functions of the EMC, we developed approaches to produce EMC for structure determination and broad mutational analysis (Figure 2.1). We developed systems to produce robust quantities of pure intact yEMC and hEMC to determine structures for the two organisms in which different facets of EMC function have been described in detail (Jonikas et al., 2009; Christianson et al., 2012; Guna et al., 2018; Shurtleff et al., 2018). Parallel efforts converged on an approach involving FLAG affinity-tagging of the EMC5 C-terminus, which was performed for endogenous yEMC and recombinant hEMC in human embryonic kidney (HEK) cells. In parallel, to enable testing of hypotheses based on structures, we created a suite of human (K562) knockout cell lines deleted for individual hEMC subunits - hEMC1 (lumen), hEMC2 (cytoplasm), hEMC3, and hEMC5 (transmembrane) - and a series of reporters of EMC-dependent transmembrane protein biogenesis (Figure S2.1). Reintroduction of the wild-type hEMC subunits in the respective knockout cells fully rescued the knockout phenotype (Figure S2.2, S2.3). This allowed for introduction of structure-based mutations in hEMC subunits into the respective knockout cells to determine features supporting biogenesis of fluorescently tagged versions of three different types of EMC clients: the transmembrane domain of a C-terminal tail-anchored transmembrane protein (squalene synthase, SQS378-410) (Guna et al., 2018), a polytopic transmembrane protein that depends on the EMC N-terminal insertase activity (Beta 1 adrenergic receptor, B1AR) (Chitwood et al., 2018), and a polytopic transmembrane protein (Sigma intracellular receptor 2, TMEM97) whose biogenesis requires the Sec61 translocon but does not require a terminal helix insertase (Figure S2.1). Three individual EMC clients were fused to mCherry fluorescent protein and GFP separated by a P2A ribosomal skipping sequence. Translation of the described mRNA generates two products due to peptide bond skipping at the P2A sequence. For each molecule of the client-mCherry fusion, there is one GFP molecule. Reduction in mCherry levels relative to GFP reflects

post-translational degradation of the client fused to mCherry. Each of the client reporters were introduced into five separate cell lines: wild-type K562 cells, hEMC1 knockout K562 cells, hEMC2 knockout K562 cells, hEMC3 knockout K562 cells, and hEMC5 knockout K562 cells. Monitoring the effect of an hEMC mutation on fluorescent reporter levels provided a quantitative measure of its impact on EMC-dependent biogenesis of each class of client protein. A number of mutations of varying severity, varying conservation between yeast and human, were designed and tested spanning the hEMC structure. Subsequently, these 49 mutations were mapped onto the structure grouped by reporter phenotype. To allow for direct comparison of our structure-guided mutant phenotypes with those published recently by others (Pleiner et al., 2020; Bai et al., 2020; O'Donnell et al., 2020), we summarized all mutant data (see full publication for videos). A subset of the mutant cell lines was validated by genotyping. Western blots against the endogenous hEMC subunits allowed us to control for mutational effects on the production and stability of the hEMC complex itself. We concurrently blotted for three clients, SQS, TMEM97, and BCAP31, to assay changes in endogenous protein levels for each of the mutations (Figure S2.2, S2.3). This strategy thus distinguishes effects resulting from a global disruption of the EMC complex from those caused by specific disruption of EMC function. These functional assays of the hEMC show a broad dependence of all these clients on the EMC, consistent with previous work (Shurtleff et al., 2018; Guna et al., 2018; Chitwood et al., 2018; Volkmar et al., 2019; Tian et al., 2019). In order to understand the mechanism of action, we will now go in more detail through several of the mutants with the strongest functional phenotypes in differing regions of the three-dimensional structure.

The EMC is an intricate molecular machine spanning the ER membrane and exhibits a conserved core architecture

We determined structures of yEMC and hEMC — all showing overall compositional similarity, with regional conformational differences between the yeast and human complexes (see publication). We obtained reconstructions of yEMC bound to an antigen binding fragment (Fab)

and hEMC reconstituted both in detergent micelles and lipid nanodiscs, with the latter strategy yielding the most isotropic and highest resolution data. For yEMC+FabDH4 and hEMC, the global map resolutions reached 3.2 A° and 3.4 A°, respectively. The cryo-EM maps allowed for de novo model building of both human and yeast complexes (see publication). As described in the following sections, our multiple EMC structures enable a broad survey of its conserved architecture, with variations between the structures pointing to conformational and compositional differences (see publication). We note that our maps and models are consistent with recent cryo-EM data from yeast EMC (Bai et al., 2020), human EMC (O'Donnell et al., 2020; Pleiner et al., 2020), and a crystal structure of human EMC2-EMC9 (O'Donnell et al., 2020).

The EMC is comprised of cytoplasmic, transmembrane, and lumenal domains arranged similarly for yeast and human, despite significant evolutionary separation (Figure 2.2A-D). For both species, subunits encompassing EMC2 to EMC7 form an interconnected core complex, while there is additional density capping both the cytoplasmic and lumenal domains of hEMC, occupied by an hEMC8/9 and an hEMC1 N-terminal expansion, respectively (Figure 2.2C–D). hEMC8 and hEMC9 are paralogs of each other, which have not been identified in yeast (Wideman, 2015). We modeled and depict only hEMC8 for clarity, but due to the 44% sequence identity with hEMC9 and both being present in the recombinant system we refer to this as hEMC8/9. The large hEMC1 insertion in hEMC constitutes the majority of a membrane distal beta-propeller domain protruding into the lumen, a feature missing from *S. cerevisiae.* Compared to other ER-resident proteins implicated in membrane protein biogenesis (Suloway et al., 2009; Pfeffer et al., 2017; Ramirez et al., 2019; McDowell et al., 2020; McGilvray et al., 2020), the arrangement of domains of the EMC is unusual with the transmembrane domain connecting prominent cytoplasmic and lumenal domains (Figure 2.2E). On a global level, the structure suggests complexities beyond those of some other ER machineries fulfilling select functions in transmembrane protein biogenesis.

The cytoplasmic domain provides a platform for protein-protein interactions

The exterior interface of the cytoplasmic domain is formed by EMC2, EMC3, EMC4, and parts of hEMC8/9 (in human), while parts of EMC5, EMC2, and EMC8/9 are shielded from the cytoplasm (Figure 2.3A–B). The helical fold of EMC2 constitutes the central organizer of this platform, established by five or six TPR motifs in human versus yeast, respectively (Figure 2.3C). TPR domains are commonly found mediating protein-protein interactions and are present in numerous well-characterized chaperone-protein and other interaction networks (Blatch and La¨ssle, 1999; Scheufler et al., 2000; Schlegel et al., 2007; Assimon et al., 2015; Krysztofinska et al., 2017; Graham et al., 2019). Yeast EMC2 features a more curved helical arrangement with N- and C-terminal domains in closer proximity to each other than seen in hEMC2. Notably, the canonical peptide-binding TPR groove is occupied by the partially helical C-terminus of EMC5, which forms a large interaction surface with EMC2. To test the functional roles of this interaction, we mutated three residues within the hEMC2 TPR motif (hEMC2$^{K125E + R126D + K127E}$) or a single hEMC5 residue buried in the TPR-binding groove (hEMC5^{F90A}). The mutations on both sides of the interface decreased hEMC integrity by western blot, with a modest decrease of hEMC subunits for hEMC5^{F90A} and a strong reduction in the levels of several hEMC subunits for hEMC2$^{K125E + R126D + K127E}$ (Figure 2.3C, S2.2). This suggests that this interface might be critical for EMC complex assembly rather than EMC function.

The multi-protein cytoplasmic cap has distinct elements between hEMC and yEMC. Capping the cytoplasmic domain in hEMC is hEMC8/9; the functional roles of this cap-like structure are not yet clear. An hEMC8-9 heterodimer is not observed and our cryo-EM permits tracing with both the hEMC8 or hEMC9 amino acid sequence (see publication). Mass spectrometric analysis of our hEMC preparations reveals slightly higher abundance of hEMC8 to hEMC9 (see publication), so we modeled the cytoplasmic cap structure with the hEMC8 sequence. A groove on hEMC8/9 cradles an N-terminal peptide of hEMC4, which proceeds into the EMC4 segment that traverses over hEMC2 and the three-helix bundle of hEMC3 (Figure

2.3D). Although yEMC lacks EMC8/9, yEMC4 follows a similar binding trajectory along cytoplasmic yEMC2 and yEMC3 surfaces. A stretch of 20 hEMC4 amino acids (residues 23–42) after the hEMC8/9 binding site is only poorly resolved in our cryo-EM maps and predicted to be disordered (40% glycine content). This loop contains primarily polar amino acids, and traverses the top of the hEMC2 TPR domain. To see whether this dynamic hEMC2-hEMC4 interface played a role in client stabilization, we mutated two charged patches on hEMC2 to alanines (hEMC2$^{E146A+E149A+Q150A}$, hEMC2$^{E168A+D170A+K173A}$), lying in close vicinity to hEMC4$^{23-42}$ (Figure 2.3E–F). These mutants lead to a modest accumulation of the tail-anchored client (SQS$^{378-410}$) but did not affect polytopic client abundance or decrease of hEMC subunits (Figure 2.3E–F, S2.2). Several mutants across the cytoplasmic domain showed similar phenotypes, supporting a key role in tail anchor protein biogenesis (Figure S2.3).

Two distinct cavities are present in the transmembrane domain

The transmembrane core of EMC is predicted to include contributions from each subunit except for EMC2 and, in humans, hEMC8/9 (Figure 2.2C–D). The EMC presents two distinct and structurally conserved cavities on opposite sides of the transmembrane core that differ in size, shape, subunit compositions and apparent function (Figure 2.4A–B). One cavity, which we refer to as the lipid-filled cavity, appears contiguous with the ER lipid environment (Figure 2.4A). The second cavity, which we refer to as the gated cavity, appears to open toward the cytoplasm in our structures and is more occluded by a transmembrane helix gate from the lipid environment (Figure 2.4B). Notable structural hallmarks present in both species include a superimposable core of nine transmembrane helices, a set of flexible gate helices, and an amphipathic EMC1 brace helix (Figure 2.4C).

The gated cavity serves as a conduit for terminal helix insertion

Evaluating potential client paths from the cytoplasm into the transmembrane domain revealed a cavernous opening at the membrane-cytoplasmic interface of the gated cavity, wide enough to allow passage of a client helix, and tapering toward the lumen (Figure 2.4D). Consistent

with its potential role as a cytoplasmic conduit into the EMC, the EMC3 portion of the cytoplasmic domain, which delineates this opening, sits approximately 45 A° from the lumenal side of the gated cavity. This dimension exceeds the thickness of the ER membrane (Mitra et al., 2004; Heberle et al., 2020; Cornell et al., 2020; Figure 2.4D). This cavity is lined primarily by EMC3, EMC4, and EMC6 (Figure 2.5A). Simulating the dimension of the first transmembrane helix of a known terminal insertase-client (B1AR - Chitwood et al., 2018) suggests that there is sufficient space for a client helix even in the client-free state of the EMC (Figure 2.5B). The gated cavity is hydrophilic on the cytoplasmic side and becomes increasingly hydrophobic toward the lumenal side (Figure 2.5C).

The entrance into the gated cavity interior (Figure 2.5A) is formed primarily by the EMC3 cytoplasmic domain. To test its function, charge swap mutations were introduced along the rim of this opening (hEMC3$^{E63K + D213K + E223K}$, hEMC3$^{R59E + R62E + K216E}$) (Figure 2.5D). These mutants resulted in loss of the tail-anchored client (SQS$^{378-410}$) and partial loss of the N-terminal insertase-dependent polytopic client (B1AR), reflecting a failure to support insertase activity. These mutants had no appreciable effect on the abundance of the polytopic transmembrane client (TMEM97) reporter (Figure 2.5E, S2.7). A similar phenotype was observed with alanine substitutions for a pair of lysines at the periphery of this cytoplasmic rim (hEMC3$^{K42A + K43A}$) (Figure S2.7).

Having identified a functionally important entry route for terminal helix insertase clients, we next considered potential surfaces inside the cavity that might accommodate a client helix. A polar patch close to the membrane interior of this cavity was conspicuous, even though the specific amino acid residues are not strictly conserved. Mutating a pair of adjacent asparagine residues to equivalently sized but negatively charged aspartates (hEMC3$^{N114D+N117D}$) resulted in a dramatic decrease in SQS$^{378-410}$ reporter levels and no significant decrease in the other two client reporter levels (Figure 2.5F). Western blot analysis for this mutant showed wild-type rescue levels of hEMC subunits and a decrease in endogenous SQS levels (Figure S2.3). Meanwhile, mutating a neighboring positively charged residue to an alanine (hEMC3^{R180A}), a residue that is conserved

in some of the YidC-superfamily insertase proteins (Anghel et al., 2017), resulted in partial loss of only the tail-anchored insertase client (SQS378-410) (Figure S2.3, S2.6).

Lastly, we surveyed residues closer to the hydrophobic lumenal side of the gated cavity. Lipid density was resolved at positions along the cavity in hEMC and yEMC cryo-EM maps (Figure 2.4B) and the properties of this hydrophobic seal to the lumen are conserved (Figure S2.6). The importance of this hydrophobic seal is suggested by the strong effect of a structurally mild mutation of a conserved methionine to a leucine (hEMC3^{M151L}), which caused significant decrease in both SQS$^{378-410}$ and B1AR abundance (Figure 2.5G). Mutation of a neighboring aromatic residue (hEMC3^{F148L}), contacting both a lipid and a hEMC4 C-terminal transmembrane helix, caused a decrease in all three client types without altering the levels of hEMC subunits (Figure S2.6). Together these results indicate that proper EMC insertase function depends on the exact composition of the cavity and not simply on its hydrophobic nature.

Structural heterogeneity suggests a role for the gate in regulating access to the insertase transmembrane cavity

While the core transmembrane helices of the gated cavity are superimposable in all four of our EMC structures, the adjacent gate helices appear in different relative orientations. The structural variability likely reflects dynamics of the gate (Figure 2.4C). Comparing detergent and nanodisc maps for both species identified two major gate conformations (see publication). One of the conformations, referred to as the closed-gate conformation, results in a more occluded membrane cavity. The other conformation, referred to as the open-gate conformation, would provide space for client accommodation.

The C-terminal transmembrane helix of EMC4 and ensuing lumenal segment are well resolved in all four structures; however, other regions of EMC4, including the segment connecting the cytoplasmic domain to the transmembrane gate helices, were poorly resolved, perhaps owing to mobility. The yEMC detergent map, yEMC nanodisc map, and hEMC detergent map all show the unassigned helices in the closed conformation, preventing client residence in the gated cavity.

By contrast, the hEMC nanodisc map reveals an open-gate conformation with the unassigned helices shifted away from the transmembrane core to provide space for a client (Figure 2.5B). Consistent with our observations, the closed transmembrane gate conformation can also be seen in recently published cryo-EM maps of hEMC (O'Donnell et al., 2020) and yEMC (Bai et al., 2020), which studied LMNG and digitonin-solubilized complexes, respectively. We note that the conformational heterogeneity and concomitant lower resolution of the gate likely accounts for the challenges in making unambiguous subunit assignments, reflected by the three different interpretations reported in recent structures (Pleiner et al., 2020; O'Donnell et al., 2020; Bai et al., 2020).

Considering the apparent flexibility of the gate, we sought to mutate the hEMC4 interfaces resolved in the cytoplasm versus the membrane. As described above, mutating residues that together form a composite-binding surface for the cytoplasmic domain of hEMC4 (hEMC2$^{E146A+E149A+Q150A}$, hEMC2$^{E168A\ +\ D170A\ +\ K173A}$, Figure 2.3E–F), we observed a modest accumulation of the tail-anchored insertase client (SQS$^{378-410}$). Likewise, mutating residues in the center of the gated cavity, close to one of the unassigned helices in the closed-gate conformation (hEMC3$^{V118A\ +\ I122A}$) led to an increase of SQS$^{378-410}$. This SQS$^{378-410}$ accumulation effect stands in contrast to mutating a residue that contacts the lumenal anchor of hEMC4 (hEMC3^{F148L}), which caused a reduction of SQS$^{378-410}$ levels (Figure S2.6).

The lipid-filled cavity is critical for both insertase-dependent and insertase-independent EMC functions

In addition to the gated cavity, the EMC harbors another membrane-accessible cavity. The surface of the lipid-filled cavity includes contributions from EMC1, EMC3, EMC5, and EMC6 (Figure 2.6A). In our structures, the EMC2 N-terminus occludes cytoplasmic accessibility to this cavity (Figure 2.4D, Figure 2.6A–B). However, this cavity may be accessible from the membrane or the ER lumen. The respective distance from the cytoplasmic EMC2 N-terminus to the lumenal

side of the lipid-filled cavity is approximately 35 A° across, which is close to the average ER membrane thickness (Mitra et al., 2004).

The lipid-filled cavity features a uniformly hydrophobic surface (Figure 2.6C) and superimposes across our ensemble of EMC structures. As noted, we resolved several lipids in our cryo-EM maps lining the cavity wall and modeled four POPC (1-palmitoyl-2-oleoyl-sn-glycero-3 phosphatidylcholine) molecules in the hEMC nanodisc map (Figure 2.6C). The residues near these lipids are moderately conserved (Figure S2.8). To characterize the functional role of the lipid-filled cavity, we mutated cavity-lining and lipid-proximal residues (Figure 2.6D, S2.8). Most of these mutations resulted in an increased abundance of the tail-anchored reporter ($SQS^{378-410}$) and wild-type rescue levels for the other two reporters (B1AR, TMEM97). However, one lipid-proximal mutant showed decreased levels of all three client reporter types with varying severity ($hEMC3^{R13E}$) without altering overall EMC levels (Figure 2.6E). Western blotting for the endogenous SQS and TMEM97 revealed a decrease in endogenous SQS and TMEM97 levels for this mutant (Figure S2.3). An analogous mutation in Drosophila EMC3 was recently reported to cause reduced levels of Rh1 in this mutant background (Xiong et al., 2020). The amphipathic EMC1 brace helix, which packs against the transmembrane helices of EMC5, is a structural hallmark of the lipid-filled cavity (Figure 2.6D). Here, mutating interfacial residues from hEMC5 ($hEMC5^{H19L+S23A+Q26L}$) caused a marked decrease in the N-lumenal polytopic reporter (B1AR) and no effect on either the tail-anchored client ($SQS^{378-410}$) or the polytopic client reporter (TMEM97) (Figure 2.6F). Unexpectedly, mutating interfacial residues from hEMC1 ($hEMC1^{F473Y+R487K}$) showed a diametrically opposed phenotype in which B1AR was unaffected, increased $SQS^{378-410}$ levels, and TMEM97 levels markedly decreased (Figure 2.6G). Another mutation in this brace ($hEMC1^{M483A+R487H+Q491N)}$ resulted in a decrease in TMEM97 and no significant effect on the other two client reporters. An adjacent $hEMC5^{D44K}$ mutations in the interfacial brace had yet different resulting client flow cytometry profiles, with an increase in $SQS^{378-410}$ and no effect on

either of the polytopic client reporters (Figure S2.9). The pleiotropic client phenotypes across the panel of interfacial brace mutants suggest that this feature is critical for multiple EMC functions.

The EMC lumenal domain is crucial for the biogenesis of multi-pass transmembrane proteins

Composed primarily of EMC1, EMC7, and EMC10, the extensive EMC lumenal domain (Figure 2.7A) is important for polytopic client biogenesis and interactions with lumenal chaperones (Luo et al., 2002; Shurtleff et al., 2018; Hiramatsu et al., 2019; Coelho et al., 2019). EMC7 and EMC10 are scaffolded on two beta-propellers of EMC1, one distal and the other proximal to the membrane. The lumenal cap differs between hEMC and yEMC, with a four-bladed distal beta-propeller in yeast and eight-bladed distal propeller the human complex (Figure 2.7B). All three lumenal EMC subunits have structural folds known to participate in protein-protein interactions (Reinisch and De Camilli, 2016). Mutations in this lumenal domain have been linked to loss of the EMC complex (Bircham et al., 2011), a trafficking delay for membrane protein Pma1 (Luo et al., 2002), and male infertility (Zhou et al., 2018).

Several regions of the lumenal domain form stabilizing interactions with the membrane cavities. The gate helices of the gated cavity are anchored via the embedding of EMC4's C-terminus within the membrane-proximal EMC1 propeller. The lipid-filled cavity is connected to the ER lumenal domain via the amphipathic EMC1 brace helix, which is tethered to the membrane-proximal EMC1 beta-propeller. The connections between the lumenal domain and the transmembrane cavities could allow for conformational coupling during client handling. Indeed, superimposing the two conformations presented above, the open- and closed-gate states, revealed not only differences in the transmembrane domain but also a rotation of the lumenal domain relative to the membrane cavities (see publication). The lumenal positioning is consistent for all three of our closed-gate conformation reconstructions (hEMC detergent, yEMC nanodisc, yEMC detergent). By contrast, the one map with an open gated cavity displayed a luminal rotation and concomitant shifts in position of the hEMC1 brace helix (Figure S2.10). Indeed, our set of interfacial hEMC1 brace mutants described above (Figure 2.6F–G, S2.10), showed differing client

phenotypes when mutated from either the hEMC1 or the hEMC5 side. This suggests a complex conformational interplay between lumenal and transmembrane domains during the engagement of diverse client types.

We investigated several known disease mutations in both conserved and human-specific regions of hEMC1 (Figure 2.7C–D, S2.10; Harel et al., 2016; Abu-Safieh et al., 2013; Amberger et al., 2019). One of these disease-associated residues sits near the anchor point for the lumenal hEMC4 transmembrane gate helix (hEMC1^{R881C}), while the majority are found farther from the membrane (hEMC1^{G868R}, hEMC1^{A144T}, hEMC1^{T82M}) (Figure 2.7C–D, S2.10). Incorporating each of these disease mutations into our EMC functional assay resulted in lower levels of the N-cytoplasmic polytopic client (TMEM97) and an increase in the level of the tail-anchored client (SQS$^{378-410}$), discussed in more detail below.

Two different hEMC1 mutants associated with cerebellar atrophy, visual impairment, and psycho-motor retardation (hEMC1^{T82M}, hEMC1^{G868R}) map to the hinge region between the hEMC1 beta propellers where hEMC7 binds (Figure 2.7D). Both the mutants at this protein-protein interface resulted in depletion of the N-cytoplasmic polytopic client (TMEM97). EMC7 and EMC10 form beta-sandwich domains on either side of the membrane-proximal beta-propeller of EMC1 and contact each other across the EMC1 surface. Consistent with our structures, coupling of these subunits is supported by the prior finding that in the absence of EMC7, EMC10 is also lost from the complex while the other EMC components appear unaffected (Shurtleff et al., 2018). EMC7 and EMC10 have been proposed to be auxiliary components with weaker phenotypes compared to core EMC subunits (Jonikas et al., 2009; Shurtleff et al., 2018; Dickinson et al., 2016). Upon deleting yEMC7, multi-pass transmembrane clients are retained in the ER but tail-anchored clients, including SQS-homolog Erg9, decrease in abundance (Shurtleff et al., 2018).

Several features of our data suggest dynamic association of hEMC7. Density for the hEMC7 beta-sandwich at the hinge between the two hEMC1 beta propellers was relatively weak in the consensus hEMC nanodisc map (see publication). Additional rounds of 3D classification

revealed two distinct classes, one with clear density for hEMC7, and one with weak density in this region. Mass spectrometric analysis of purified hEMC, however, revealed that the abundance of hEMC7 was similar to that of the other hEMC components. Both reconstructions, with and without density for the hEMC7 lumenal domain, displayed well-resolved density for hEMC10. Together, we conclude that hEMC7 is associated with hEMC1 in two different conformational states of hEMC7 with potentially distinct functions.

The OMIM database (Amberger et al., 2019) lists a mutation of unknown significance linked to retinitis pigmentosa (hEMC1^{A144T}) residing in the EMC1 distal propeller (Figure S2.11). Additionally, we also generated mutations in two surface exposed patches of the membrane-distal EMC1 beta-propeller projecting into the lumen (hEMC1^{D31K,} hEMC1^{R69D}, hEMC1^{G71S}, hEMC1$^{H93D + E138D + N282K}$, Figure 2.7E, S2.10). Overall, these mutations displayed the same client effect: a decrease in the N-cytoplasmic polytopic client reporter (TMEM97), no change in the N-lumenal polytopic client reporter (B1AR), and accumulation of the tail-anchored client reporter (SQS$^{378-410}$). Upon identifying antibodies against yEMC, we observed that the top two antibodies bind to a similar extended loop in the distal yEMC1 beta-propeller, perhaps suggesting that this site is accessible for co-factor binding in the ER. Intriguingly, this region of the lumenal domain corresponds to the region where hEMC1 has an expanded distal beta-propeller. Taken together, the data provide evidence that the lumenal domain is functionally coupled to the broader EMC role in transmembrane client stabilization. Moreover, these data support that the EMC is acting as a holdase chaperone to shield polytopic clients from degradation while they are folding to their functional form.

DISCUSSION

Our collection of yeast and human EMC structures revealed the intricate and dynamic architecture of this multifunctional transmembrane molecular machine. The structures served as the starting point for our systematic dissection of EMC's multifaceted functions by exploring the impact of structure-based mutations on the ability of the EMC to support the biogenesis of

representative members of three classes of membrane proteins: SQS, a tail-anchored protein, which exploits EMC's C-terminal insertase activity; B1AR, which relies on EMC's N-terminal insertase activity; and TMEM97, a polytopic membrane protein, which depends on the EMC for its biogenesis but does not rely on either of EMC's terminal insertase activities. Our data revealed that a conserved dual membrane cavity architecture supports the biogenesis of this diverse panel of transmembrane clients.

Overall, our studies present a nuanced picture of EMC's multifunctionality, revealing structural regions that differentially impact production of the three distinct client types. Unexpectedly, we also find that alterations to either the cytoplasmic or lumenal domain of EMC lead to enhanced abundance of the TA substrate. Moreover, our work provides a foundational framework for understanding how discrete yet allosterically coupled regions of the complex enable the multiple functions of the EMC to support membrane protein biogenesis. Taken together, these studies suggest a model in which the EMC differentially regulates the biogenesis of distinct membrane proteins, thereby contributing to cellular coordination of membrane protein abundance in accordance with physiological needs. We propose a model of the EMC functioning both as a terminal insertase as well as a holdase chaperone that is potentially modulated by post-translational modifications, lipid interactions, and protein-protein interactions (Figure 2.8). Here, we summarize our findings into a proposed model of EMC function for these three clients.

Terminal insertase clients require an embedded insertase module within the EMC

EMC3's fold at the interface between the cytoplasm and membrane forms the core of the gated cavity and is reminiscent of proteins from the YidC family of insertases (Borowska et al., 2015; Dalbey and Kuhn, 2015; Anghel et al., 2017). Indeed, mutations in either the cytoplasmic or transmembrane domains of EMC3 establish that these features are critical for terminal helix insertase activity. In light of our observation of multiple gate conformations, we speculate that these conformations modulate insertion and release into the ER membrane.

Notably, mutating the surface of the cytoplasmic cap, which extends beyond the EMC3 cytoplasmic helices toward EMC8/9, resulted in an unexpected increase in C-tail anchor client ($SQS^{378-410}$) abundance. Of the three clients analyzed, SQS was the only one to show enhanced levels. It is unclear if this enhancement is SQS-specific or representative more broadly of all post-translationally targeted EMC tail-anchored clients. Future studies will be required to address if this is due to regulated insertion of SQS by the EMC, parallel pathways for inserting SQS into the membrane (i.e. mediated by TRC40/GET), and/or slower cytoplasmic clearance of chaperone-bound SQS.

Post-translational insertase clients have previously been shown to be targeted to the ER by cytoplasmic chaperones (Guna et al., 2018). Structural analysis and coupled mutagenesis, from our and recent studies (O'Donnell et al., 2020; Pleiner et al., 2020; Bai et al., 2020), suggest that clients then engage the cytoplasmic domain of the EMC, the transmembrane gate opens, the terminal helix is inserted into the EMC-gated cavity, and then another conformational change would allow for release into the lipid bilayer (Figure 2.8A–B). Further studies are needed to establish the precise C-terminal client range, as most tail anchor clients have been shown to be inserted by the GET (in yeast) or WRB (in human) complexes (Denic et al., 2013; Mateja and Keenan, 2018).

Both EMC cavities have resolved lipids and are critical for client biogenesis

Both the N-terminal (B1AR) and C-terminal insertase (SQS) clients depend on the EMC-gated cavity. Indeed, both the SQS tail-anchored helix and the first transmembrane helix of B1AR are moderately hydrophobic, with polar residues near the cytoplasmic end of the transmembrane helix, and both showed a strong dependence on the gated cavity. Nevertheless, our panel of mutants revealed some notable differences in the handling of these two client types. B1AR showed more dependence than SQS on the lipid-filled cavity in contrast to mutants elsewhere in the complex. Consistent with this, a number of mutations, primarily in the gated cavity, show residues of importance to both SQS and B1AR. However, there are also a number of mutations

that appear to only affect SQS. One possible reason could be due to differences in the mechanism of initial engagement: SQS is targeted to the ER by cytoplasmic chaperones, while B1AR is targeted by SRP. Another key difference is that B1AR is polytopic and needs to overcome the additional challenge of tertiary transmembrane packing to reach its folded state. This work provides support for a model where the EMC inserts both types of terminal transmembrane helices into the gated cavity with differences in initial targeting and perhaps release into the lipid environment (Figure 2.8C–D). Future work will address the interplay between B1AR synthesis and its co-translational engagement with the translocon to ascertain whether there is a direct handoff between the translocon and the EMC or the EMC acts post-translationally to insert the N-terminal helix of B1AR.

The EMC lumenal domain orchestrates holdase chaperone function important for polytopic clients

Unlike the two terminal insertase clients we investigated, TMEM97 biogenesis was negatively impacted by mutation of the lumenal EMC1. The depletion of TMEM97 observed in these mutant backgrounds is consistent with the lumenal domain contributing to a holdase chaperone function, passively shielding its client while it is being synthesized and/or folded (Zhang et al., 2017). Interestingly, the diametrically opposed phenotype of mutants in the EMC lumenal domain on SQS raises the possibility that occupancy by one type of client can support an EMC conformation that is unfavorable for receiving the other. Alternative conformations could establish competition between client types for EMC occupancy. One explanation for this observation is that there is a conformational change between the insertase-active versus the holdase-active states. Interestingly, we identified at least two EMC conformations in our collection of structures, and EMC may adopt different conformations in various client and cofactor-engaged states.

In yeast, the polytopic clients co-purifying with the EMC are also glycosylated. One possible model is that the putative carbohydrate-binding domains in EMC7 or EMC10 directly contribute to engagement with client proteins. We speculate post-translational modifications on

clients and the EMC could modulate function including client binding, chaperone binding, or regulating signaling in response to cellular cues.

Multi-pass transmembrane proteins require membrane factors to assist after insertion into the membrane to pack transmembrane helices in the correct order and topology. We propose that the EMC may act as a chaperone holdase to facilitate one of the following: helix and lipid packing, shielding from degradation while synthesis is in progress, or assisting in the assembly of multi-protein transmembrane complex formation. This is consistent with observations that in the absence of the EMC numerous integral membrane proteins are degraded (Shurtleff et al., 2018; Volkmar et al., 2019; Tian et al., 2019). Direct interactions with multi-pass transmembrane proteins have been shown previously (Shurtleff et al., 2018; Coelho et al., 2019). Furthermore, EMC dependence of internal transmembrane domain segments has also been established (Ngo et al., 2019; Hiramatsu et al., 2019). In the absence of yEMC7, a primarily lumenal subunit, a polytopic membrane protein was retained for longer in the ER, suggesting the possibility that yEMC7 may be involved in client release from the EMC. We propose a model where the EMC engages polytopic clients either during or directly after translation and remains bound until the client is released either to the membrane environment directly or handed off to client-specific and general ER chaperones (Figure 2.8E–F). It remains to be seen whether these polytopic clients directly engage with the lipid-filled cavity or the gated cavity or the lumen domain, what the determinants for engaging with a client or release into the membrane are, and how the EMC fits into the broader ER lumenal chaperone network

Potential role of the EMC as a master regulator of membrane protein biogenesis as the basis for its pleiotropic phenotypes

Why does the cell use a multifunctional EMC molecular machinery rather than specialized machinery for each of the functions encompassed by the EMC? Considering that the cell already has general machinery (Sec61 translocon) and tail-anchor insertase machinery (GET/TRC complex), we speculate that the EMC coordinates biogenesis of diverse membrane proteins.

Several observations suggest broader roles of the EMC as an integrator of information sensing the protein and lipid environment and coordinating its multiple activities, including the regulating the biogenesis of membrane proteins. For example, the initial identification of the EMC included numerous genetic interactions with both protein and lipid synthesis factors in yeast (Jonikas et al., 2009) and these disparate interdependencies have been subsequently observed in numerous species including human EMC (Lahiri et al., 2014; Tang et al., 2017; Guna et al., 2018; Volkmar et al., 2019; Volkmar and Christianson, 2020). Also, several client proteins are enzymes or cofactors involved in multiple stages of lipid synthesis or trafficking, and this may provide a unifying explanation for the range of genetic interactions and co-essentiality observations reported to date (Guna et al., 2018; Shurtleff et al., 2018; Volkmar et al., 2019; Tian et al., 2019; Wainberg et al., 2019; Corradi et al., 2019; Volkmar and Christianson, 2020). Perhaps by facilitating the insertion of sterol synthesis protein SQS, the EMC allows for modulation of local membrane thickness and lipid composition to accommodate differences within the broad range of membrane proteins being synthesized. In this regard, one structural feature of particular interest is the EMC1 amphipathic brace, which resides adjacent to the lipid-filled cavity. This conserved feature sits within the interfacial membrane boundary, raising the possibility that it can modulate the lipid or protein composition of this cavity. Notably, several other membrane proteins involved in ER homeostasis, including Opi1 and Ire1, also contain amphipathic helices that have been proposed to sense the properties of the lipid bilayer (Volmer et al., 2013; Jacquemyn et al., 2017; Halbleib et al., 2017; Hofbauer et al., 2018; Cho et al., 2019). Future work will explore how the EMC overall, and the EMC1 brace helix in particular, govern client release into the membrane, interface with the local structure of the lipid bilayer, and play roles in specific client-lipid interactions.

In addition to the three client classes, we investigate here, it is clear that EMC has a broader range of clients including multi-protein assemblies (Richard et al., 2013; Talbot et al., 2019), lipid-modulating proteins (Volkmar et al., 2019), lipid-binding proteins (Salas-Estrada et al., 2018; Sejdiu and Tieleman, 2020), and those with helices that do not span the bilayer (Lin et

al., 2019; Ngo et al., 2019). The compartmentalization and interdependence that we observe for effects of mutations on client handling provide a foundation for understanding this multifunctionality. We propose that the complexity of the EMC machine, combining insertase and holdase chaperone functions within one molecular machine, has arisen to mitigate the error prone biogenesis of a diverse range of membrane spanning proteins in the dynamic environment of the ER.

ACKNOWLEDGEMENTS

We thank J Weibazahn, P Walter, R Irannejad, J Gestwicki, R Scheltema, O¨ Karayel, H Nguyen, I Johnson, N Talledge, L Kenner, E Thompson, K Hickey, J Kellermann, S von Gronau, M Feige, K Swain, M Liao, C-W Lee, F Wilfling, and members of the Weissman, Frost, and Schulman laboratories for assistance and helpful discussion; L Metzger, Z Roe-Zurz, M Tessema, DW Chester, and S Aller for assisting with fermentation; M Sun, H Autzen, and E Green for advice on nanodisc reconstitution; F Hagn and I Goba for the gift of the MSP1D1 vector and advice on nanodisc reconstitution; P Thomas and D Asarnow for computational support; M Braunfeld, G Gilbert, E Tse, D Bulkley, M Harrington, A Myasnikov and Z Yu of the UCSF Center for Advanced CryoEM for microscopy support and funded by NIH grants S10OD020054 and 1S10OD021741; J Baker-LePain and the QB3 shared cluster (NIH grant 1S10OD021596-01) for computational support; and the Howard Hughes Medical Institute (HHMI); Z Yu and H Chou of the CryoEM Facility at the HHMI Janelia Research Campus (NIH grant 1S10OD021596-01). D Bollschweiler, T Scha¨ fer and the cryo-EM facility at the Max Planck Institute of Biochemistry; B Steigenberger, the mass spectrometry core facility at the Max Planck Institute of Biochemistry; E Gouaux for the gift of the pEG vector; A Titan X Pascal used for this research was donated by the NVIDIA Corporation. This study was supported in part by the HDFCCC Laboratory for Cell Analysis Shared Resource Facility through a grant from NIH (P30CA082103). Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from

National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases. UCSF Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

MATERIALS AND METHODS

Cell line maintenance

K562 dCas9 KRAB cells were grown in RPMI 1640 (GIBCO) with 25 mM HEPES, 2 mM l-glutamine, 2 g/L NaHCO3 and supplemented with 10% (v/v) fetal bovine serum (FBS), 100 units/mL penicillin, 100 mg/mL streptomycin, 2 mM l-glutamine. HEK293T cells were grown in Dulbecco's modified eagle medium (DMEM, GIBCO) with 25 mM d-glucose, 3.7 g/L NaHCO3, 4 mM l-glutamine and supplemented with 10% (v/v) FBS, 100 units/mL penicillin, 100 mg/mL streptomycin. All cell lines were grown at 37°C. All cell lines were periodically tested for Mycoplasma contamination using the MycoAlert Plus Mycoplasma detection kit (Lonza).

DNA transfections and virus production

Lentivirus was generated by transfecting HEK39T cells with standard fourth-generation packaging vectors using TransIT-LT1 Transfection Reagent (Mirus Bio). Media was changed 10 hr post-transfection. Viral supernatant was harvested 60 hr after transfection, filtered through 0.45 mm PVDF filters and frozen prior to transduction.

Knockout hEMC cell lines

A single and dual knockout guide system was developed in the pX458 backbone (Addgene plasmid #48138) with guides targeting hEMC1, hEMC2, hEMC3, or hEMC5. Targeting guides were selected using the Broad's guide selection tool (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design). For the single hEMC5 knockout system, an hEMC5 targeting guide was cloned into pX458 by digesting with BbsI and ligating to annealed oligos for the hEMC5 sgRNA. For the dual knockout system, a four-step cloning process generated the final knockout plasmid: (1) Each of the two guides targeting the

same locus were individually cloned into pX458. (2) Then pX458_sgRNA1 was digested with XbaI. (3) SgRNA2 cassette from pX458_sgRNA2 was PCR amplified with oligos containing overhangs spanning the XbaI cloning site and purified. (4) Finally, the final dual guide vector was generated by Gibson cloning (NEBuilder). To generate the hEMC knockout cell lines, K562 dCas9 KRAB cells were nucleofected with the respective hEMC knockout plasmids using Lonza SF Cell Line 96-well Nucleofector Kit (V4SC-2096). Two days post-nucleofection, GFP-positive cells were single cell sorted into 96-well plates using BD FACS AriaII. After colonies from single cells grew out, genomic DNA was isolated using QuickExtract (Lucigen), the sgRNA-targeted sites were PCR amplified and then NGS-sequenced via Genewiz's EZ-Amplicon service. Sequencing data was analyzed and aligned to the respective reference alleles in the human genome. Clones whose alleles harbored only indel mutations for hEMC1, hEMC2, hEMC3, and hEMC5 (full knockouts) respectively were further validated on the protein level.

Dual fluorescent EMC client reporter cell lines

Dual client reporters for TMEM97, ADRB1 (protein name: B1AR), and FDFT1 (protein name: SQS) were introduced lentivirally into each of the EMC1, EMC2, EMC3, and EMC5 knockout cell lines. TMEM97 and ADRB1 full-length sequences were used with a C-terminal tag -mCherry-P2A-GFP. The sequence for FDFT1 transmembrane domain ($SQS^{378-410}$) was tagged N-terminally with GFP-P2A-mCherry- and an opsin tag on the C-terminus as used in a prior study (Guna et al., 2018). Three days post-transduction, GFP/mCherry-positive cells were sorted on BDAriaII.

Mutant EMC cell lines

The EMC mutant genes were synthesized and cloned by Twist into pKDP119-SFFV-[insert site]-IRES-Puro-P2A-BFP. Mutant hEMC cell lines were generated by lentiviral introduction of the respective hEMC mutant subunit into the respective knockout cell lines (hEMC1, hEMC2, hEMC3, or hEMC5) containing the dual fluorescent reporters for each EMC client (pKDP110_ADRB1_mCherry_P2A_GFP, pKDP111_TMEM97_mCherry_P2A_GFP, or

GFP_P2A_m-Cherry_FDFT1_TMD_opsintag). The expression of each fluorescent reporter was read out 6 days after puromycin selection in each of the hEMC mutant cell lines.

Flow analysis

For each hEMC mutant cell line, 20,000 live cells were recorded on Attune NxT flow cytometer. FlowCal flow analysis package was used for analysis in Python. First, live cells were gated based on FSC/SSC. Then GFP (BL1-A) and mChery (YL2-A) were plotted for each mutant and control cell line. mCherry:GFP intensity ratios were calculated for individual cells in each cell line. Fluorescence ratios for each substrate in an hEMC mutant cell line were normalized to the mCherry:GFP ratio of the same substrate in the hEMC wild-type rescue cell line. Distributions of fluorescence ratios were plotted as histograms in Python using seaborn.

Fluorescent reporter statistical analysis

We performed bootstrap estimates of the mean of normalized mCherry/GFP ratio from the FACS data. For bootstrapping, we performed 1000 iterations with 50 cells/iteration to fit normal distributions. We performed two separate one-sided T-tests at a p-value cutoff of 0.01 between each mutant and the respective subunit WT to test for significant decreases or increases in ratios based on bootstrapped estimates of the mean.

Western blotting

Cell pellets were lysed using lysis buffer (20 mM Tris pH 7.5, 150 mM NaCl, 5 MgCl2, 1% Triton x-100, 1 mM DTT, 24 U/ml) Turbo DNase (Ambion). Clarified lysate was quantified and samples were boiled with 4x LDS sample (Thermo Fisher, NP0007) buffer for 5 min at 95°C. Samples were separated on 4–12% or 12% Bolt Bis-Tris Plus Gels (Invitrogen, NP0322PK2). Proteins were transferred onto nitrocellulose membranes using Bio-Rad Trans-Blot Turbo transfer system. Membranes were blocked in Odyssey Blocking Buffer (LI-COR, 927–50000) for an hour at room temperature. Blocked membranes were incubated with primary antibody diluted in TBST and incubated overnight at 4°C on a shaker. Primary antibodies were detected by incubating membranes with 1:10,000 dilution of IRDye-conjugated (LI-COR) secondary anti-mouse and anti-

rabbit antibodies for 1 hr at room temperature. Blots were visualized using LI-COR imaging system.

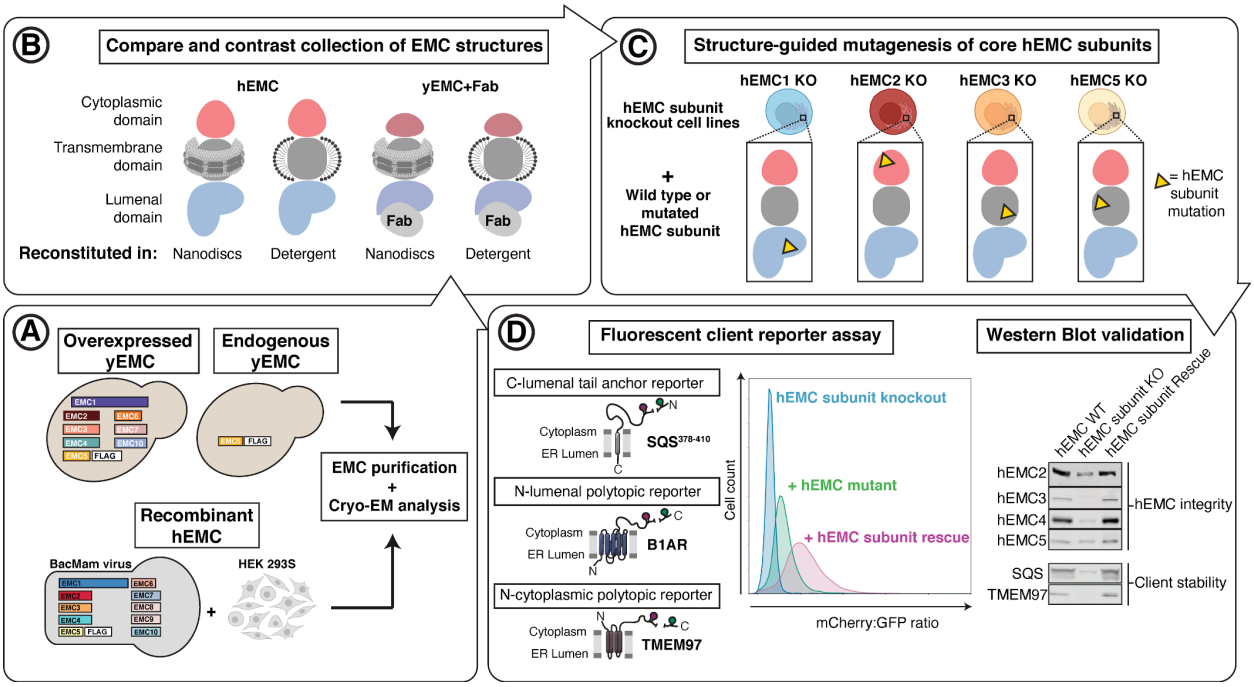


Figure 2.1. Experimental strategy for the dissection of EMC function. Schematic representation of the combined structural and mutational approach to dissect EMC function. (A) yEMC was purified either by overexpression of all subunits together and affinity pulldown with 3xFlag-tagged yEMC5 or by pulldown of endogenous yEMC proteins using an affinity pulldown with 3xFlag-tagged yEMC5. For hEMC, all subunits were overexpressed together with Flag-tagged EMC5 via a single recombinant BacMam virus. Both yEMC and hEMC were purified by column chromatography and subjected to cryo-EM analysis. (B) The obtained collection of cryo-EM structures of yEMC and hEMC in lipid nanodiscs or detergent micelles were compared to identify similarities and differences. (C) Structure-guided mutagenesis was performed across four core hEMC subunits: hEMC1, hEMC2, hEMC3, and hEMC5 in mammalian K562 cells. (D) Each hEMC subunit knockout (KO) cell line was individually transduced with three different fluorescent client reporters: SQS378-410, full-length B1AR, and full-length TMEM97. Mutant hEMC subunits were then introduced into the corresponding subunit KO cell lines carrying each of the three fluorescent hEMC client reporters. hEMC client stability in each mutant hEMC subunit cell line was assessed by quantifying the mCherry-to-GFP ratio. Western blotting was performed for each mutant-transduced cell line to assess EMC integrity (by immunoblotting for hEMC subunits) as well as client stability (by immunoblotting for hEMC clients) compared against both wild-type (WT) and KO cell lines.

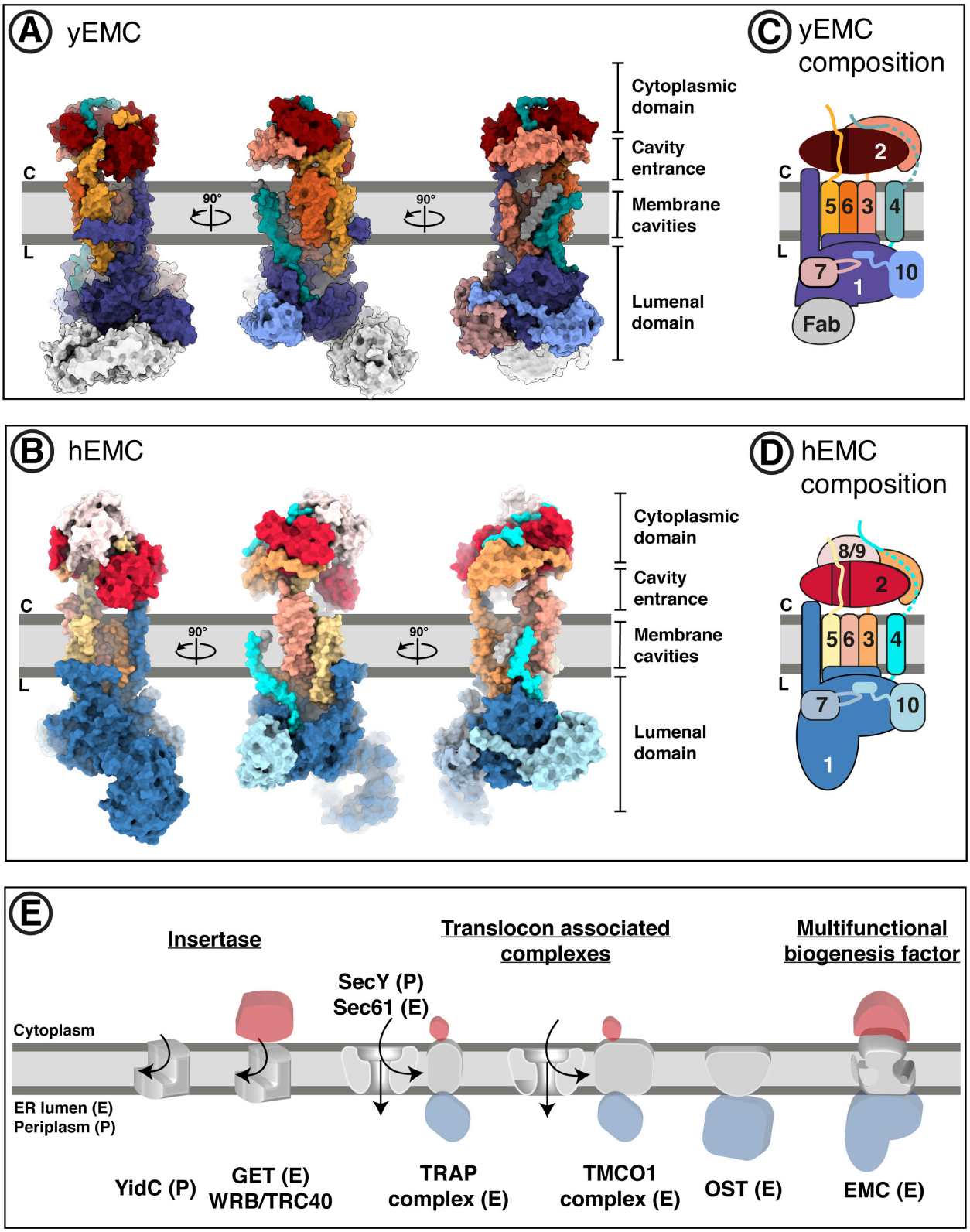


Figure 2.2. Overall structures of yeast and human EMC. (A) Cryo-EM structure of yEMC in nanodiscs. Three orthogonal views of the yEMC cryo-EM structure shown as surface rendering. Gray bars delineate the approximate ER membrane boundaries with the cytoplasmic (C) and

lumenal (L) sides indicated. The FAb molecule bound to the yEMC1 lumenal domain is colored in gray. (B) Cryo-EM structure of hEMC in nanodiscs. Labeling as in (A). (C) Subunit composition and color scheme of yEMC used throughout the manuscript. Dotted line indicates a portion of yEMC4 unresolved in the cryo- EM map and left unmodeled. (D) Subunit composition and color scheme of hEMC used throughout the manuscript. (E) Schematic depiction and comparison of the EMC architecture to known transmembrane protein biogenesis factors in the ER and the bacterial plasma membrane. Cytoplasmic, transmembrane and lumenal domains are depicted as cartoons colored red, gray and blue, respectively. E, eukaryotic; P, prokaryotic.

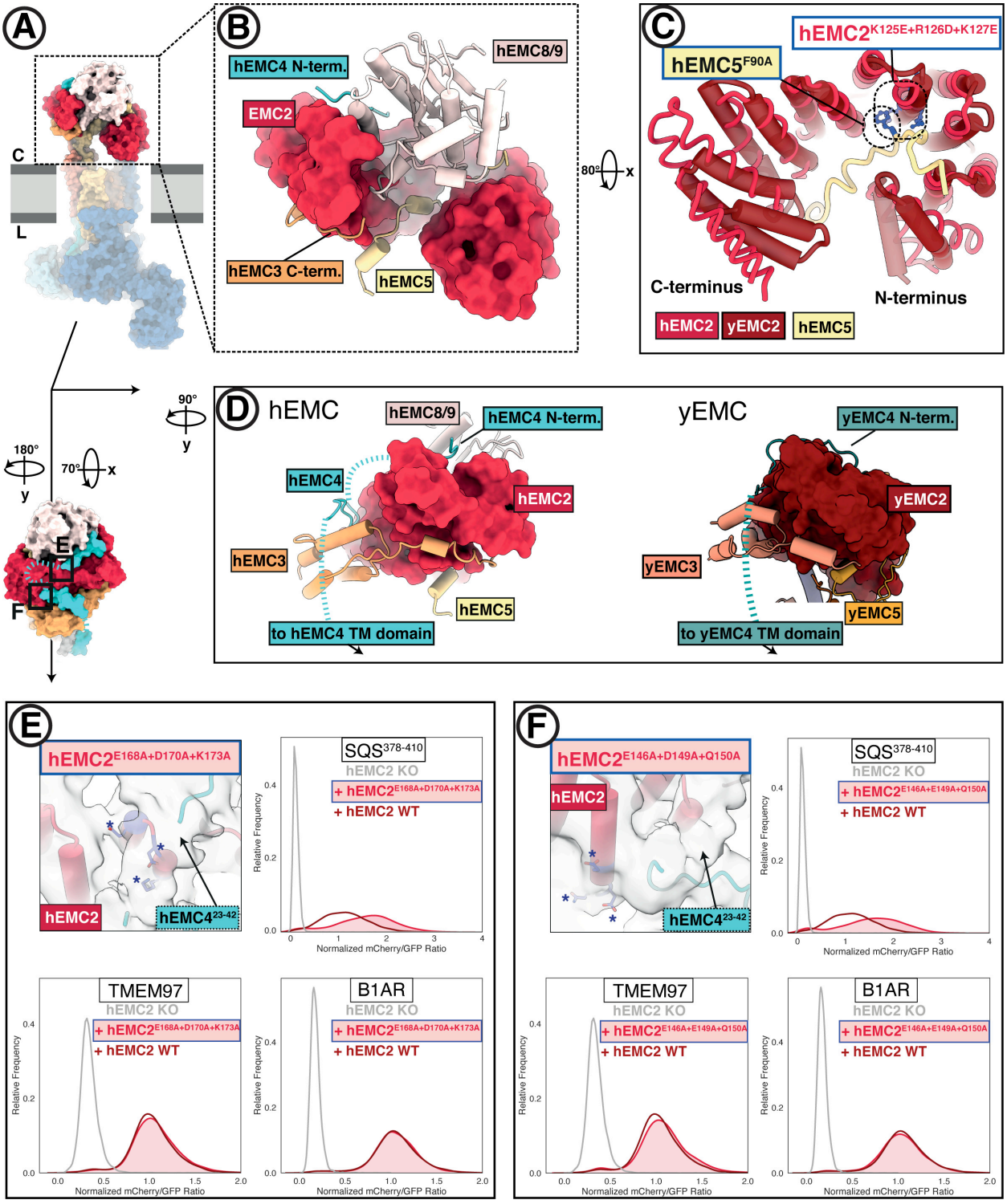


Figure 2.3. The EMC cytoplasmic domain contains conserved functional interfaces and may engage C-tail-anchored clients directly. (A) Position of the hEMC cytoplasmic domain relative to the membrane and the rest of the complex. Shown is the surface rendered hEMC structure reconstituted in nanodiscs. (B) EMC2 nucleates a protein-protein interaction hub in the cytoplasm. Zoomed-in view of the cytoplasmic domain from (A). EMC2 is shown as surface rendering while interacting EMC subunits are shown as cartoon cylinders. (C) EMC2 forms a TPR domain which

binds EMC5. Overlaid are hEMC2 (red) and yEMC2 (dark red), illustrating the more tightly wound yEMC2 TPR solenoid. Two mutants, one in EMC5 and three in EMC2, are colored in blue, and show destabilizing phenotypes for EMC integrity. (D) A cytoplasmic cap structure involving EMC4 is conserved in yEMC and hEMC. Shown is a side-by-side comparison between the cytoplasmic domains of hEMC (left) and yEMC (right), highlighting the similar path EMC4 takes from the cytoplasmic domain toward the transmembrane domain. While an interaction surface between EMC8/9 and the EMC4 N-terminus is absent in yeast, yEMC4 binds at the top of the EMC2 TPR domain and assumes as similar position across the EMC3 cytoplasmic domain at the cytoplasm-membrane interface. (E) Fluorescent client reporter stability assay for TMEM97 (N-cytoplasmic polytopic client), B1AR (N-lumenal polytopic client) and $SQS^{378-410}$ (C- lumenal tail-anchored client) in EMC2 KO cells expressing mutant $hEMC2^{E168A+D170A+K173A}$ (shaded) or WT hEMC2 rescue (unshaded). Shown is the model of hEMC in nanodiscs superposed with the unsharpened cryo-EM map, where the weaker density for EMC4 (23–42) becomes apparent. Mutated residues are colored blue and marked with asterisks for clarity. (F) Fluorescent client reporter stability assay, as in E, for the $hEMC2^{E146A+E149A+Q150A}$ mutant.

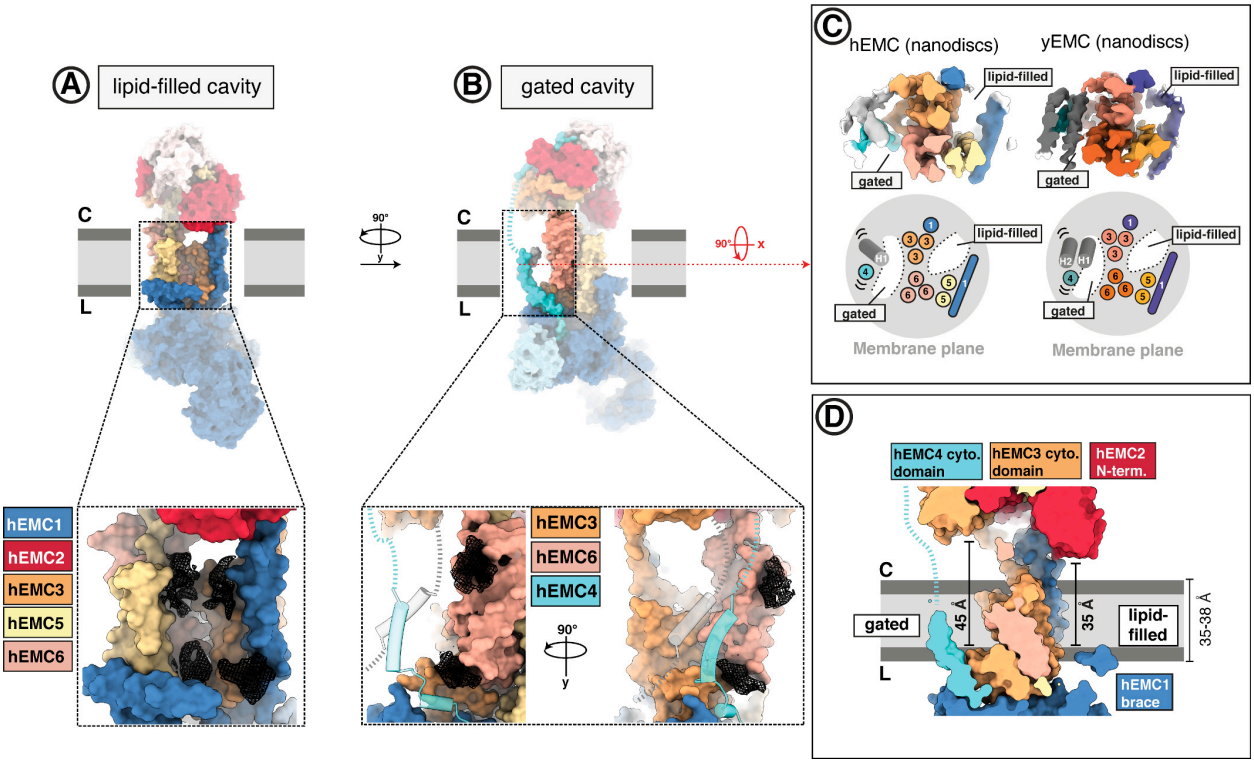


Figure 2.4. The EMC houses two transmembrane cavities with conserved core structures and distinct accessibilities. (A) Location and composition of the lipid-filled cavity. A zoom-in view on the cavity is shown below, which is composed of EMC1, EMC3, EMC5, and EMC6. Resolved lipid densities from the cryo-EM map of hEMC in POPC nanodiscs are shown as black mesh zoned within 3 A° of modeled POPC molecules. (B) Location and composition of the gated cavity. Two orthogonal zoom-in views of the cavity are shown below, which is composed of EMC3 and EMC6. A transmembrane gate opposite the cavity wall is depicted as transparent cartoon cylinders and has contributions from the C-terminal EMC4 transmembrane helix along with up to two additional, unassigned helices. Resolved lipid densities are shown as in (A). (C) The dual-cavity architecture of the EMC transmembrane domain is conserved between yEMC and hEMC. Unsharpened cryo-EM maps of hEMC and yEMC in nanodiscs (top) are shown along with corresponding schematic representations of the spatial organization of all transmembrane helices (bottom). The gate helices of the gated cavity represent the region of highest conformational heterogeneity across our collection of EMC structures. (D) The two EMC transmembrane cavities feature distinct accessibilities. Shown is a central slice through the surface rendered hEMC nanodisc structure with the two membrane cavities on opposite sides. Measuring from the lumenal to the cytoplasmic side, gated and lipid-filled cavities measure 45 A° and 35 A° across, respectively. This suggests that the gated cavity has accessibility from the cytoplasm while the lipid-filled cavity does not.

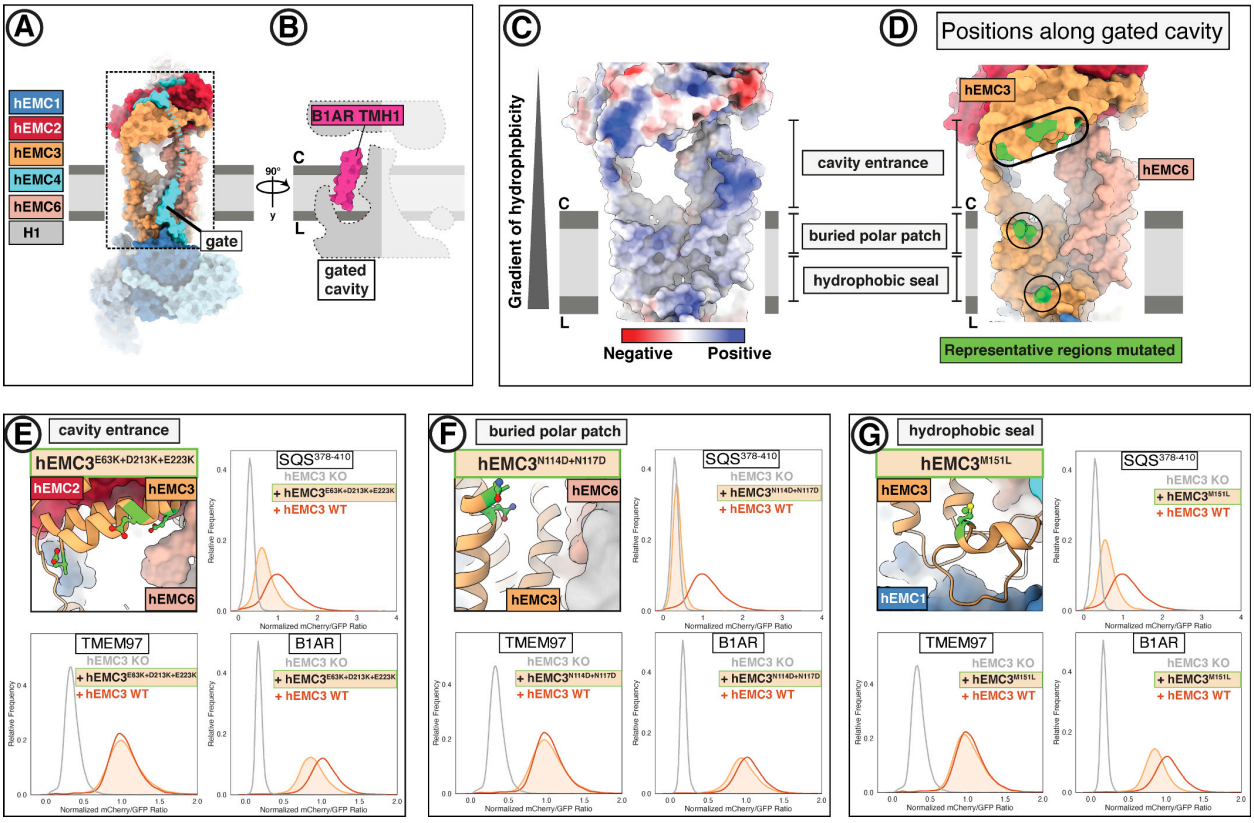


Figure 2.5. EMC houses an insertase module centered on EMC3 in the gated membrane cavity. (A) A transmembrane gate anchored in the cytosol and the lumen is a structural hallmark of the EMC gated cavity. Shown is a surface rendering of the hEMC model in lipid nanodiscs with an unresolved EMC4 connection between the cytoplasm and the membrane depicted as a dashed line. An unassigned helix of the gate is shown in gray (H1). (B) The gated cavity in the hEMC nanodisc structure has sufficient space to accommodate a client transmembrane helix. The space-filling model of the first transmembrane helix of B1AR (B1AR TMH1) is shown placed inside an outline of the EMC gated cavity. (C) A hydrophobic gradient characterizes the surface of the EMC gated cavity from the cytoplasmic to the lumenal side. Gate helices have been omitted for clarity. The surface of the hEMC nanodisc structure is colored by electrostatic surface potential ranging from -15 (red) to +15 (blue) kcal/(mol e). (D) Distinct EMC3 regions along the gated cavity hydrophobic gradient targeted for mutagenesis. Mutated residues are colored in lime. (E) Fluorescent client reporter stability assay for the EMC3 cavity entrance mutant, hEMC3$^{E63K+D213K+E223K}$. (F) As in (E) for the EMC3 buried polar patch mutant, hEMC3$^{N114D+N117D}$. (G) As in (E) for the EMC3 hydrophobic seal mutant, hEMC3^{M151L}.

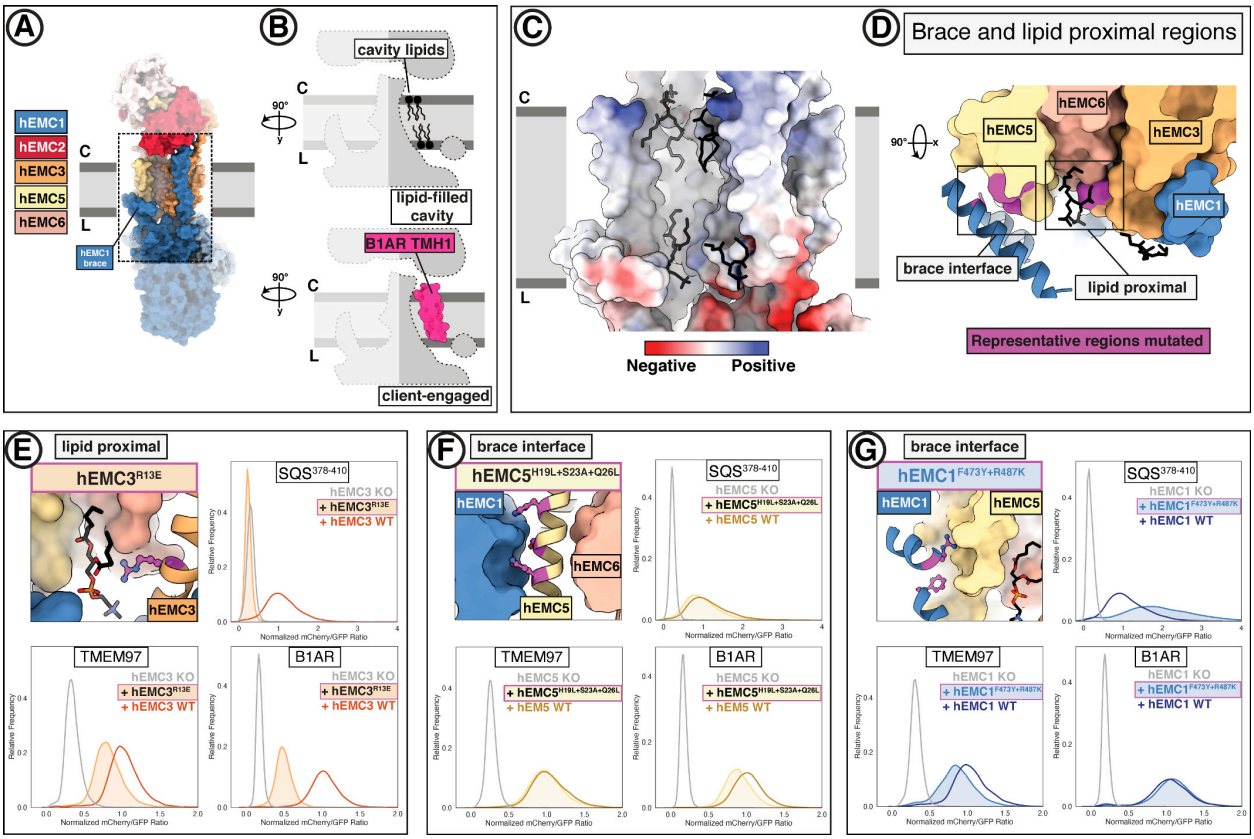


Figure 2.6. A lipid-filled cavity in the EMC transmembrane domain stabilizes disparate client proteins. (A) An EMC1 amphipathic brace helix delineates the boundary of the lipid-filled transmembrane cavity and packs against EMC5. Shown is a surface rendering of the hEMC model in nanodiscs. EMC4, EMC5, EMC6, and EMC1 subunits all contribute to the cavity lining. (B) The lipid-filled cavity in the hEMC nanodisc is occupied by several lipid molecules. Cartoon outlines of the gated cavity illustrate that the cavity could in principle allow for occupancy of a client helix (B1AR TMH1), possibly by lipid displacement or movement of the EMC1 brace helix. (C) The lipid-filled cavity has a uniform hydrophobic lining. Shown is an electrostatic surface rendering of the hEMC nanodisc structure colored as in Figure 5C. The cytoplasm-membrane interface contains positively charged residues and the lumenal interface contains negatively charged residues. Modeled phospholipid molecules are displayed in black. (D) Lipid-proximal and brace interface residues targeted for mutagenesis. Selected regions targeted for mutagenesis are colored in magenta and include brace interface mutations both in EMC1 and EMC5, as well as a lipid-proximal residue in EMC3. (E) Fluorescent client reporter stability assay for the hEMC3^{R13E} mutant, which is in close proximity to a modeled POPC molecule. (F) As in (E) for the $^{hEMC5H19L+S23A+Q26L}$ mutant, which sits at the interface to the EMC1 amphipathic brace helix. (G) As in (E) for the hEMC1$^{F473Y+R487K}$ mutant, which sits at the interface to the EMC5 transmembrane helices.

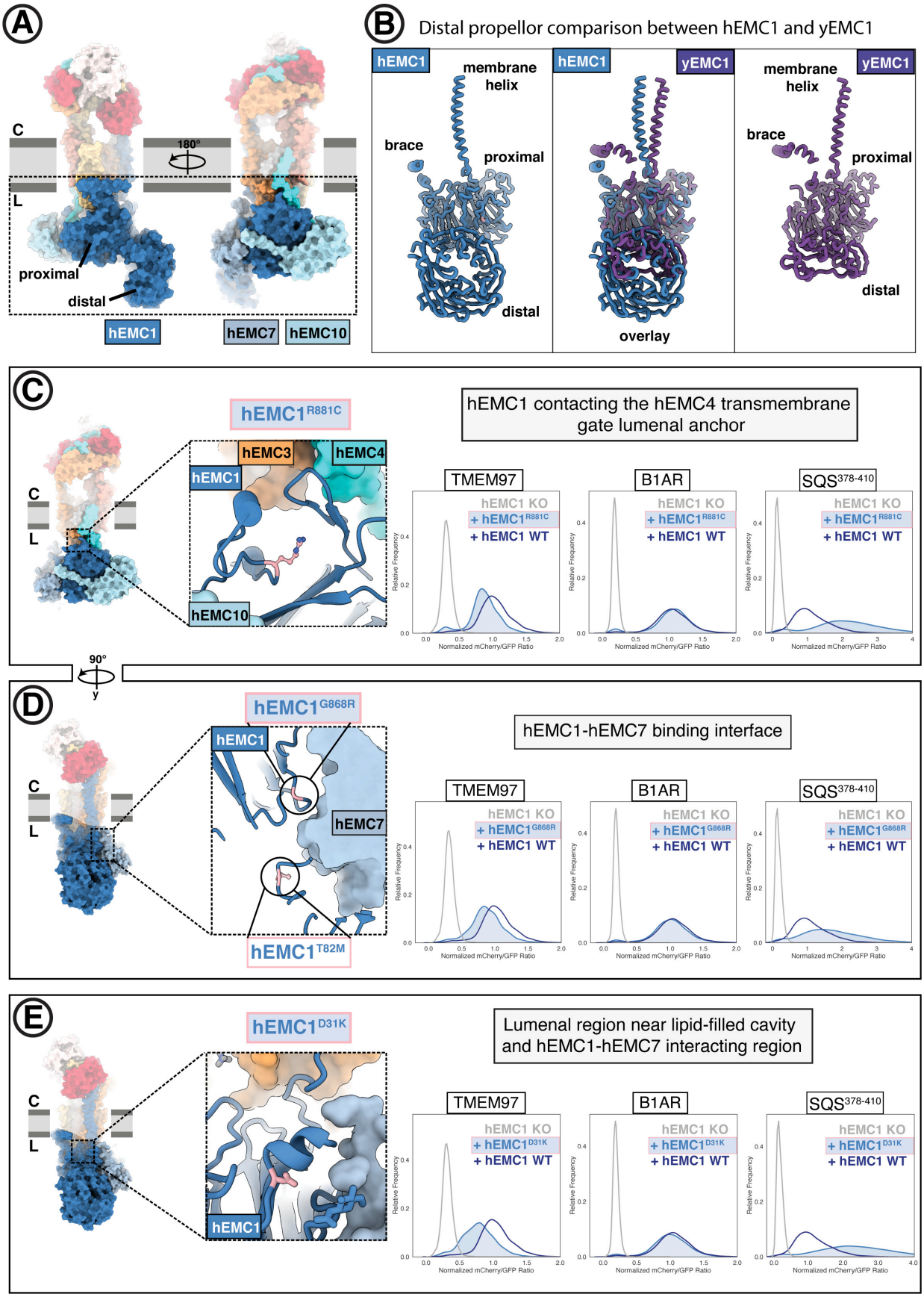

A
C
180°
L
proximal
distal
hEMC1
hEMC7
hEMC10
B
Distal propellor comparison between hEMC1 and yEMC1
hEMC1
membrane helix
brace
proximal
distal
hEMC1
yEMC1
overlay
yEMC1
membrane helix
brace
proximal
distal
C
hEMC1R881C
hEMC1 contacting the hEMC4 transmembrane gate lumenal anchor
hEMC3
hEMC4
hEMC1
hEMC10
TMEM97
B1AR
SQS378-410
hEMC1 KO
+ hEMC1R881C
+ hEMC1 WT
Relative Frequency
Normalized mCherry/GFP Ratio
90°
y
D
hEMC1G868R
hEMC1-hEMC7 binding interface
hEMC1
hEMC7
hEMC1T82M
+ hEMC1G868R
E
hEMC1D31K
Lumenal region near lipid-filled cavity and hEMC1-hEMC7 interacting region
+ hEMC1D31K

Figure 2.7. The large EMC lumenal domain is the site for several annotated disease mutations. (A) Two views of the hEMC nanodisc structure. Two beta propellers are present in EMC1, one proximal to the membrane and one distal. (B) EMC1 is the largest EMC subunit and differs in size between yeast and human. Shown are human EMC1 (nanodisc), an overlay of human and yeast EMC1 (both nanodisc), and yeast EMC1 (nanodisc). (C) The hEMC1^{R881C} mutant sits near the EMC4 lumenal gate anchor. Left: Location of the mutation (colored pink). Right: Fluorescent client R881C reporter stability assay for hEMC1. (D) As in (C) for for the hEMC1^{G868R} mutant. (E) As in (C) for the hEMC1^{D31K} mutant.

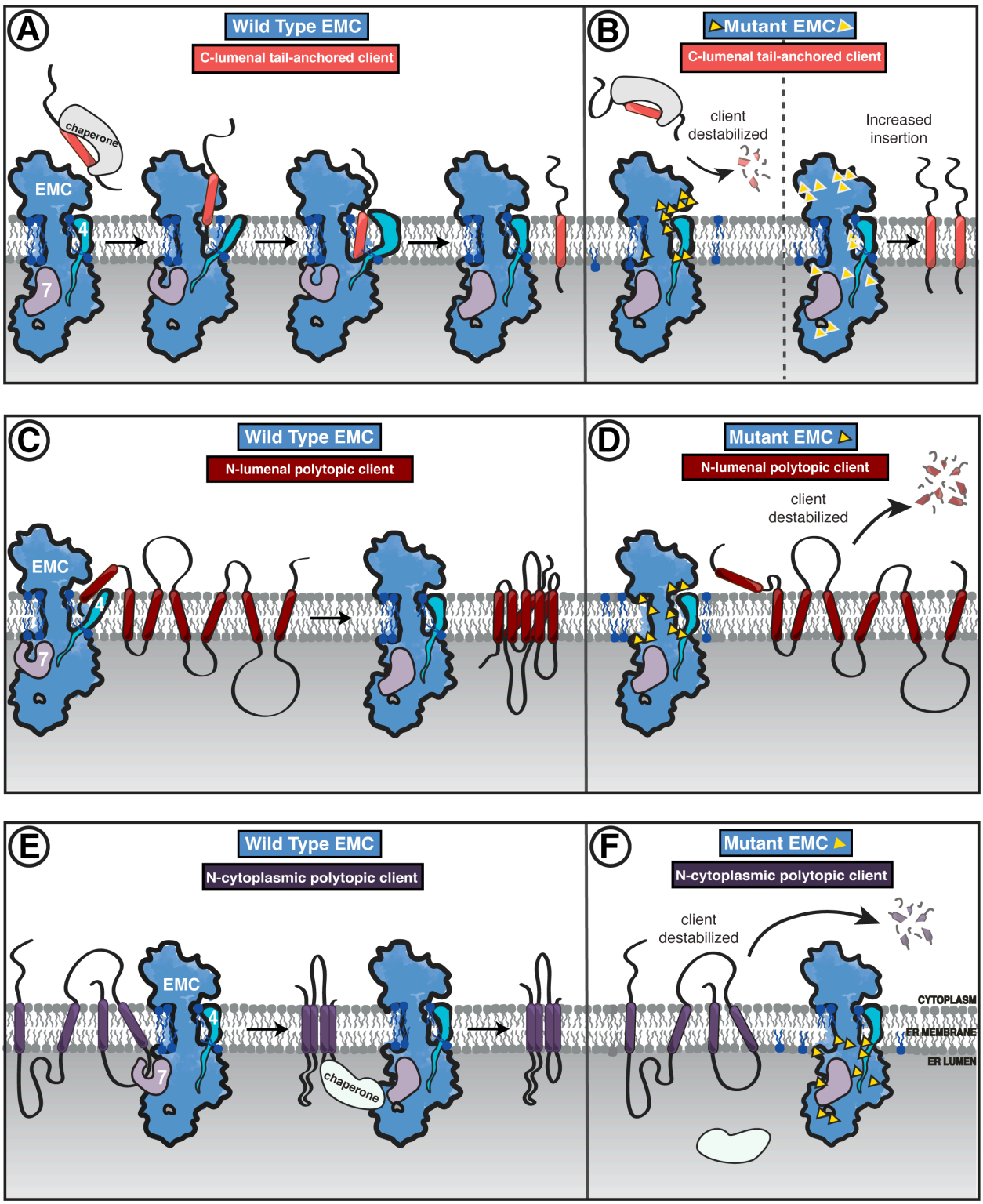


Figure 2.8. Model of coordinated EMC functions. (A) Model of EMC insertase function for a C-lumenal tail-anchored client. Cytosolic factors bring post-translationally localized clients to the ER. Then the client engages the EMC cytoplasmic domain. The polar roof modulates entry into the gated cavity. A hydrophobic slide facilitates the client helix fully entering the cavity. A lateral movement of the gate releases the client helix into the membrane and the EMC gate closes. (B)

Our mutagenesis data provide the following insights into EMC regions of functional importance for each of the three client types we tested. Mutants are depicted by yellow triangles. Tail-anchored client (coral) abundance was depleted upon mutagenesis of the cytoplasmic domain entrance to the gated cavity, polar and charged residues at the cytoplasm-membrane boundary, residues along the length of the gated cavity, in the hydrophobic seal to the lumen, and lipid interacting residues in both cavities (left). We also observed a subset of mutants that resulted in higher levels of the C-lumenal tail-anchored client (right) that are positioned in the cytoplasmic domain cap, throughout the ER lumenal domain, and one mutation at the center of the gated cavity. (C) The EMC facilitates biogenesis of N-lumenal polytopic client protein B1AR (dark red). (D) Regions important for B1AR stability primarily map to the transmembrane region of the EMC structure, with depletion observed for lipid proximal residues on both sides of the cavity, the polar entrance roof of the gated cavity, and the EMC1 brace helix. (E) The EMC facilitates biogenesis of N-cytoplasmic polytopic client protein TMEM97 (dark purple). (F) Regions important for TMEM97 stability were primarily located in the lumenal domain spanning both propellers, in EMC1. In addition to these lumenal regions, there was a depletion of TMEM97 at the lipid-interacting positions at the lumenal interface of both membrane cavities of the EMC

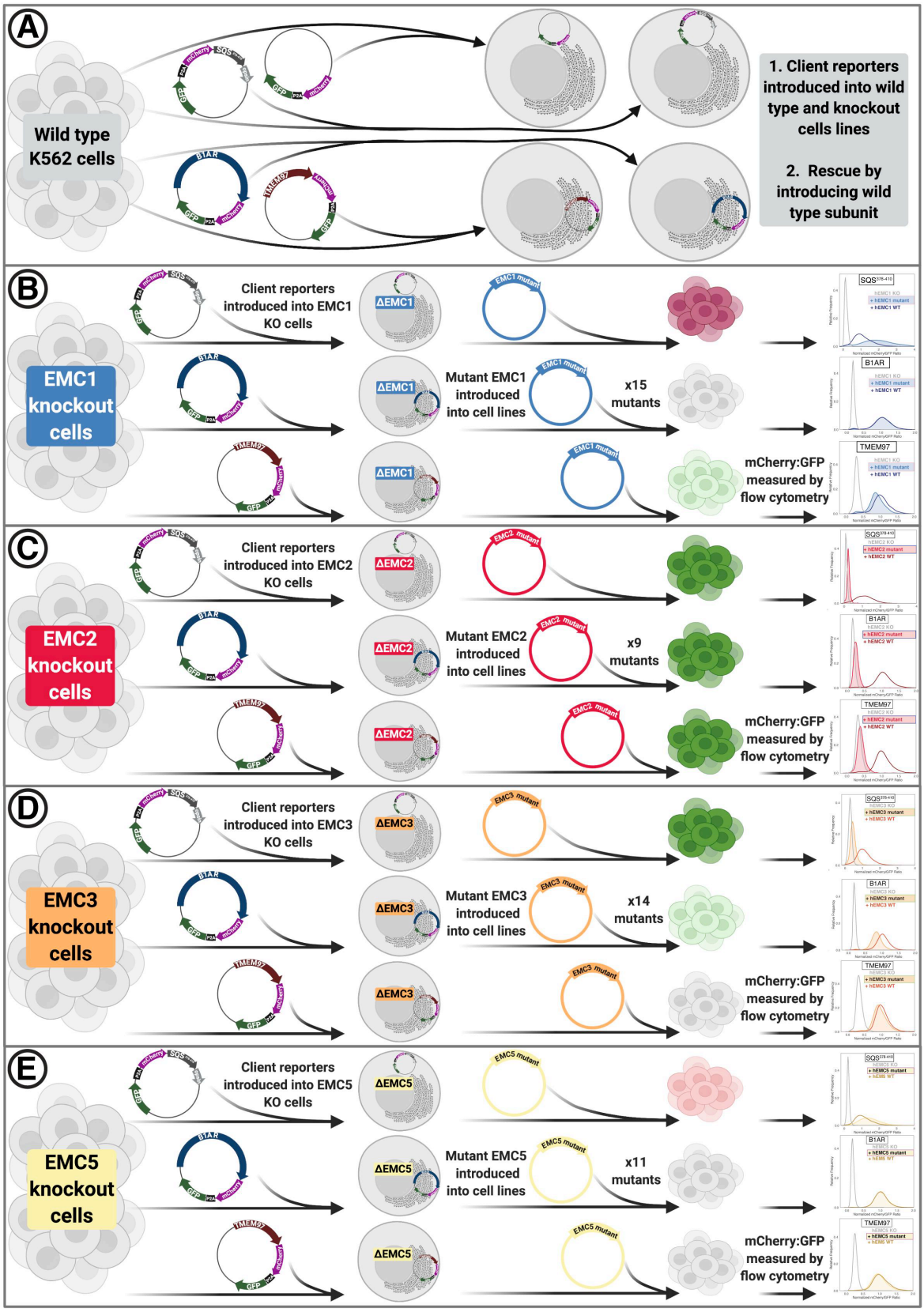

A
Wild type K562 cells
1. Client reporters introduced into wild type and knockout cells lines
2. Rescue by introducing wild type subunit
B
EMC1 knockout cells
Client reporters introduced into EMC1 KO cells
ΔEMC1
EMC1 mutant
Mutant EMC1 introduced into cell lines
x15 mutants
mCherry:GFP measured by flow cytometry
C
EMC2 knockout cells
Client reporters introduced into EMC2 KO cells
ΔEMC2
EMC2 mutant
Mutant EMC2 introduced into cell lines
x9 mutants
mCherry:GFP measured by flow cytometry
D
EMC3 knockout cells
Client reporters introduced into EMC3 KO cells
ΔEMC3
EMC3 mutant
Mutant EMC3 introduced into cell lines
x14 mutants
mCherry:GFP measured by flow cytometry
E
EMC5 knockout cells
Client reporters introduced into EMC5 KO cells
ΔEMC5
EMC5 mutant
Mutant EMC5 introduced into cell lines
x11 mutants
mCherry:GFP measured by flow cytometry
SQS
B1AR
TMEM97

Figure S2.1. Overview of strain creation and functional assays. (A) Into each K562 cell line (WT or knockout of respective subunit), fluorescent client reporters were introduced lentivirally. Subsequently, knockout phenotypes were rescued by re-introducing the wild-type hEMC subunit. (B) After introducing client reporters into hEMC1 knockout cell lines cells were sorted to obtain a pure population. Then mutant hEMC subunits were re-introduced lentivirally and selected with puromycin. Upon reaching a pure population, cells lines were subjected to flow cytometry to measure abundance of mCherry and GFP for each client reporter in each mutant background. Cell pellets were collected for subsequent western blot analysis for each mutant in WT and knockout cell line in a cell line with an mCherry-P2A-GFP reporter. (C) Same as (B) for hEMC2. (D) Same as (B) for hEMC3. (E) Same as (B) for hEMC5.

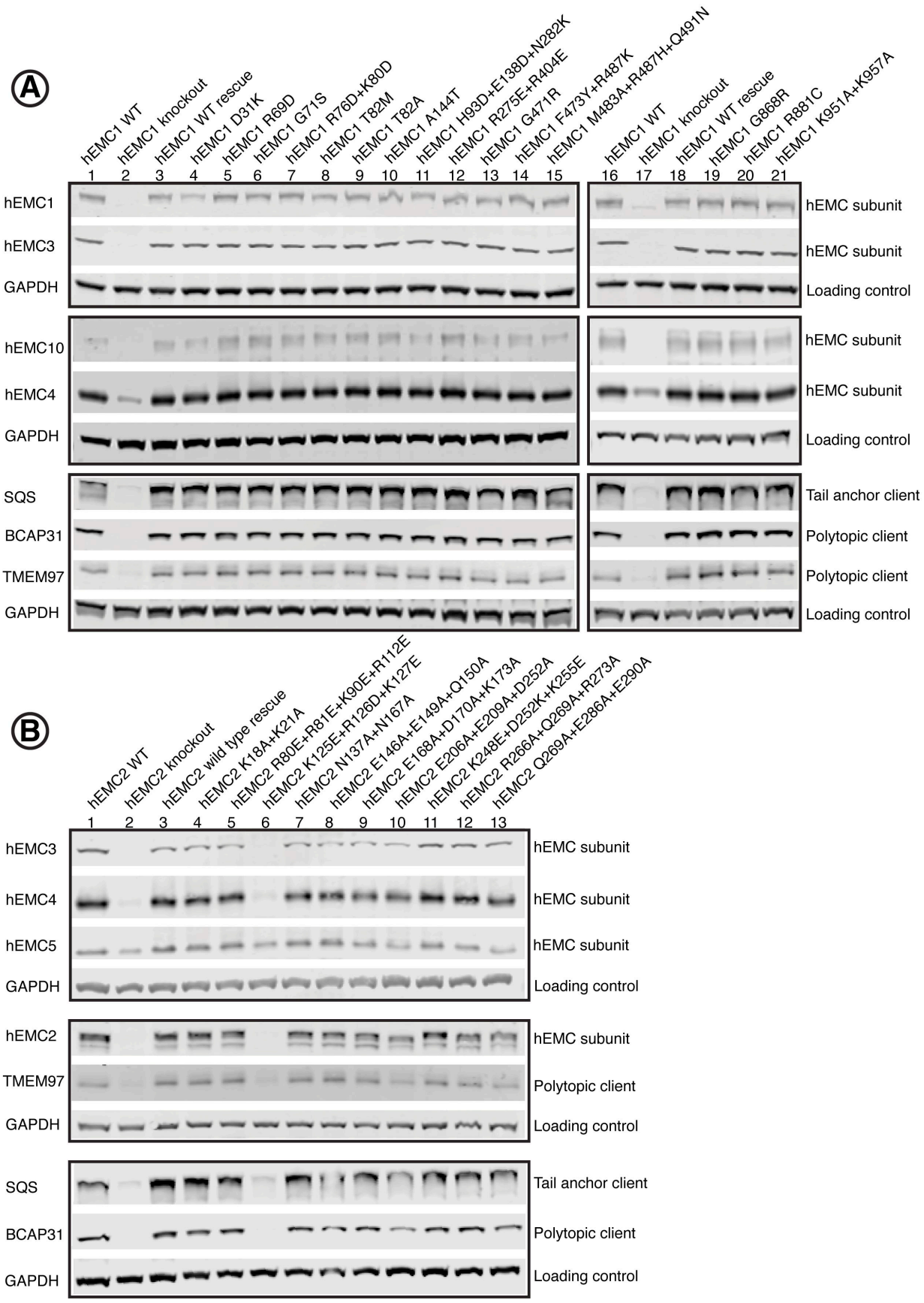

A
hEMC1 WT
hEMC1 knockout
hEMC1 WT rescue
hEMC1 D31K
hEMC1 R69D
hEMC1 G71S
hEMC1 R76D+K80D
hEMC1 T82M
hEMC1 T82A
hEMC1 A144T
hEMC1 H93D+E138D+N282K
hEMC1 R275E+R404E
hEMC1 G471R
hEMC1 F473Y+R487K
hEMC1 M483A+R487H+Q491N
hEMC1 WT
hEMC1 knockout
hEMC1 WT rescue
hEMC1 G868R
hEMC1 R881C
hEMC1 K951A+K957A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
hEMC1
hEMC3
GAPDH
hEMC subunit
hEMC subunit
Loading control
hEMC10
hEMC4
GAPDH
hEMC subunit
hEMC subunit
Loading control
SQS
BCAP31
TMEM97
GAPDH
Tail anchor client
Polytopic client
Polytopic client
Loading control
B
hEMC2 WT
hEMC2 knockout
hEMC2 wild type rescue
hEMC2 K18A+K21A
hEMC2 R80E+R81E+K90E+R112E
hEMC2 K125E+R126D+K127E
hEMC2 N137A+N167A
hEMC2 E146A+E149A+Q150A
hEMC2 E168A+D170A+K173A
hEMC2 E206A+E209A+D252A
hEMC2 K248E+D252K+K255E
hEMC2 R266A+Q269A+R273A
hEMC2 Q269A+E286A+E290A
1 2 3 4 5 6 7 8 9 10 11 12 13
hEMC3
hEMC4
hEMC5
GAPDH
hEMC subunit
hEMC subunit
hEMC subunit
Loading control
hEMC2
TMEM97
GAPDH
hEMC subunit
Polytopic client
Loading control
SQS
BCAP31
GAPDH
Tail anchor client
Polytopic client
Loading control

Figure S2.2. Western blots for EMC1 and EMC2. Western blots of endogenous human EMC subunits and client proteins functional assay to check for complex stability and endogenous protein levels. For each mutant, abundance of several hEMC subunits as well as several representative client proteins were tested. (A) Western blots for hEMC1 mutant cell lines, blotting for hEMC subunits hEMC1, hEMC3, hEMC4, and hEMC10. In addition, levels of three endogenous client proteins (SQS, BCAP31, and TMEM97) were blotted for. Wild-type cells with the fluorescent reporter displayed in Lanes 1 and 16. hEMC1 knockout cells displayed in Lanes 2 and 17. hEMC1 knockout cells with reintroduction of wild-type hEMC1 shown in Lanes 3 and 18. (B) Western blots for hEMC2 mutant cell lines, blotting for hEMC subunits hEMC2, hEMC3, hEMC4, and hEMC5. In addition, levels of three endogenous client proteins (SQS, BCAP31, and TMEM97) were blotted for. Wild-type cells with the fluorescent reporter displayed in Lane 1. hEMC2 knockout cells displayed in Lane 2. hEMC2 knockout cells with reintroduction of wild-type hEMC2 shown in Lane 3. Mutant hEMC2 E206A+E209A+D252A is in Lane 10, flow cytometry of this mutant is not included. For the remaining mutants, both western blot and flow cytometry were conducted.

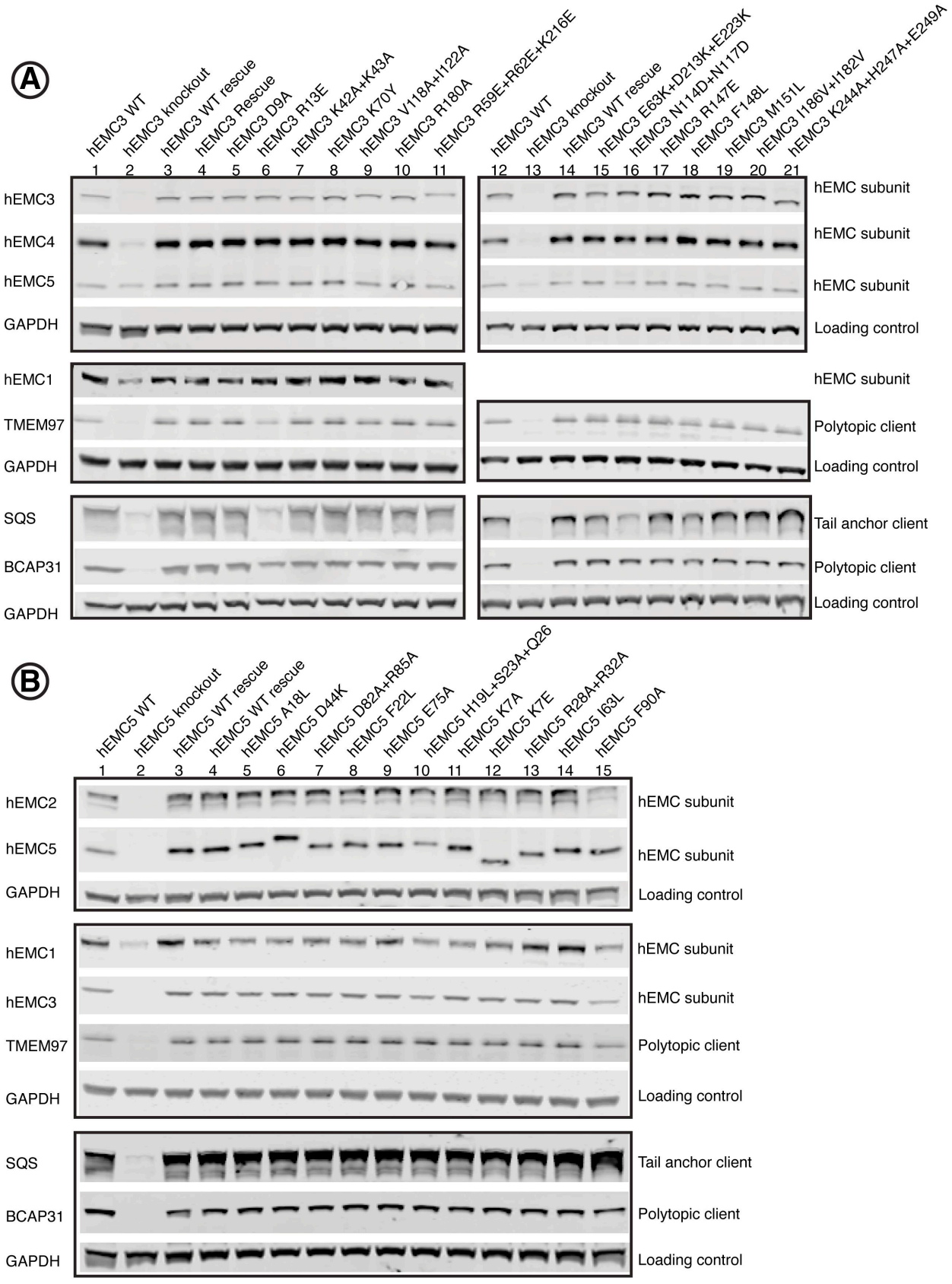

A
hEMC3 WT
hEMC3 knockout
hEMC3 WT rescue
hEMC3 Rescue
hEMC3 D9A
hEMC3 R13E
hEMC3 K42A+K43A
hEMC3 K70Y
hEMC3 V118A+I122A
hEMC3 R180A
hEMC3 R59E+R62E+K216E
hEMC3 WT
hEMC3 knockout
hEMC3 WT rescue
hEMC3 E63K+D213K+E223K
hEMC3 N114D+N117D
hEMC3 R147E
hEMC3 F148L
hEMC3 M151L
hEMC3 I186V+I182V
hEMC3 K244A+H247A+E249A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
hEMC3
hEMC4
hEMC5
GAPDH
hEMC subunit
hEMC subunit
hEMC subunit
Loading control
hEMC1
TMEM97
GAPDH
hEMC subunit
Polytopic client
Loading control
SQS
BCAP31
GAPDH
Tail anchor client
Polytopic client
Loading control
B
hEMC5 WT
hEMC5 knockout
hEMC5 WT rescue
hEMC5 WT rescue
hEMC5 A18L
hEMC5 D44K
hEMC5 D82A+R85A
hEMC5 F22L
hEMC5 E75A
hEMC5 H19L+S23A+Q26
hEMC5 K7A
hEMC5 K7E
hEMC5 R28A+R32A
hEMC5 I63L
hEMC5 F90A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
hEMC2
hEMC5
GAPDH
hEMC subunit
hEMC subunit
Loading control
hEMC1
hEMC3
TMEM97
GAPDH
hEMC subunit
hEMC subunit
Polytopic client
Loading control
SQS
BCAP31
GAPDH
Tail anchor client
Polytopic client
Loading control

Figure S2.3. Western blots for EMC3 and EMC5. Human functional assay to check for complex stability. For each mutant, abundance of several EMC subunits as well as several representative client proteins was tested. (**A**) Western blots for hEMC3 mutant cell lines, blotting for hEMC subunits hEMC1, hEMC3, hEMC4, and hEMC5. In addition, levels of three endogenous client proteins (SQS, BCAP31, and TMEM97) were blotted for. Wild-type cells with the fluorescent reporter displayed in Lanes 1 and 12. hEMC3 knockout cells displayed in Lanes 2 and 13. hEMC3 knockout cells with reintroduction of wild-type hEMC1 shown in Lanes 3 and 14. (**B**) Western blots for hEMC5 mutant cell lines, blotting for hEMC subunits hEMC1, hEMC2, hEMC3, and hEMC5. In addition, levels of three endogenous client proteins (SQS, BCAP31, and TMEM97) were blotted for. Wild-type cells with the fluorescent reporter displayed in Lane 1. hEMC5 knockout cells displayed in Lane 2. hEMC5 knockout cells with reintroduction of wild-type hEMC5 shown in Lanes 3 and 4.

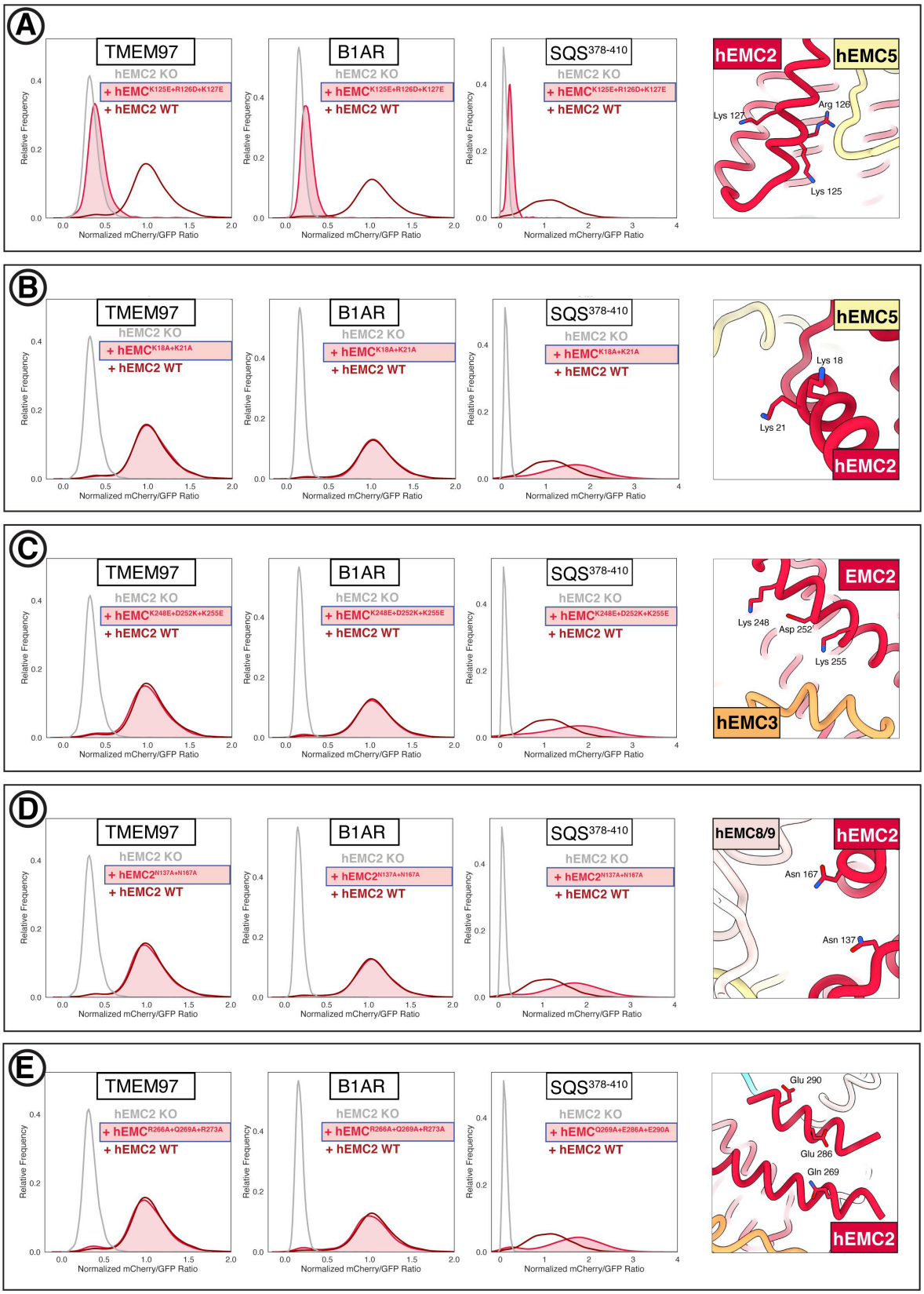


Figure S2.4. Flow cytometry for mutations in the EMC cytoplasmic domain. (A) Mutant hEMC2$^{K125E+R126D+K127E}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell

lines. Image of hEMC ND model displaying the residues mutated. (B) Mutant hEMC2$^{K18A+K21A}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (C) Mutant hEMC2$^{K248E+D252K+K255E}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (D) Mutant hEMC2$^{N137A+N167A}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (E) Mutant hEMC2$^{Q269A+E286A+E290A}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated.

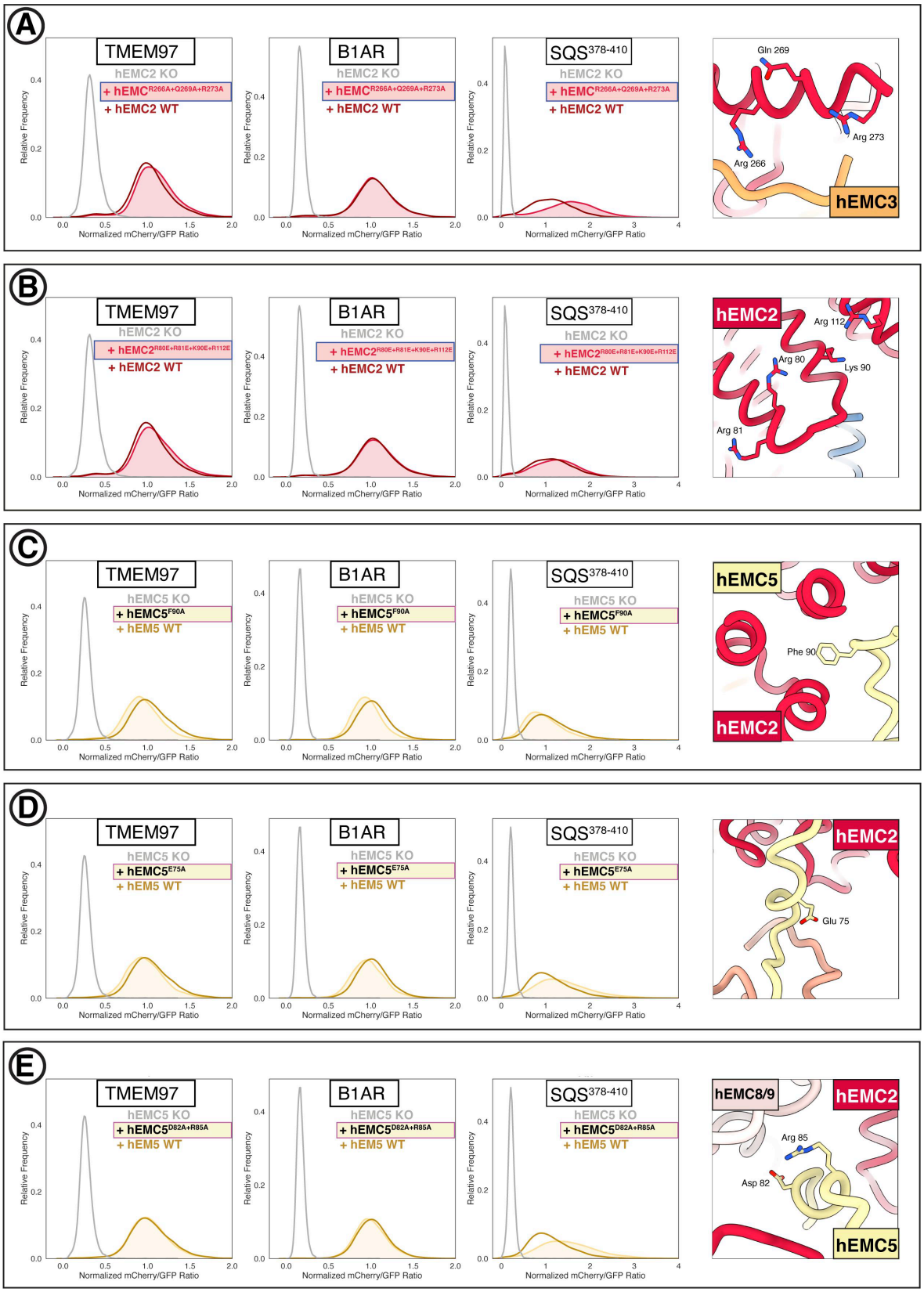


Figure S2.5. Additional Flow cytometry for mutations in the EMC cytoplasmic domain. (A) Mutant hEMC2$^{R266A+Q269A+R273A}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell

lines. Snapshot of hEMC ND model displaying the residues mutated. (B) Mutant hEMC2$^{R80E+R81E+K90E+R112E}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (C) Mutant hEMC5^{F90A} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (D) Mutant hEMC5^{E75A} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (E) Mutant hEMC5$^{D82A+R85A}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated.

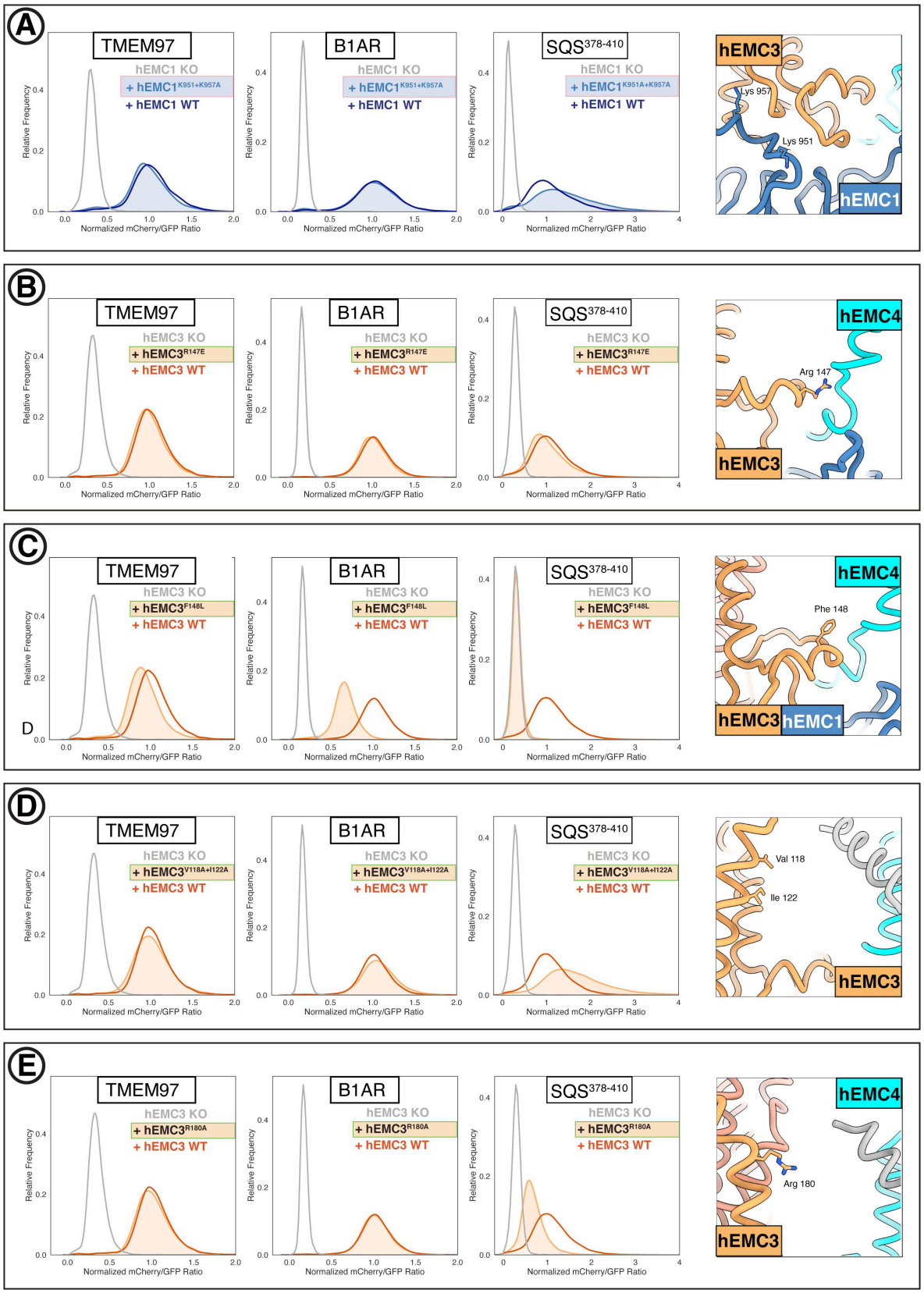


Figure S2.6. Flow cytometry and western blot of gated cavity mutants (A) Mutant hEMC1K$^{951A+K957A}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines.

Snapshot of hEMC ND model displaying the residues mutated. (B) Mutant hEMC3^{R147E} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (C) Mutant hEMC3^{F148L} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (D) Mutant hEMC3$^{V118A+I122A}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (E) Mutant hEMC3^{R180A} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated.

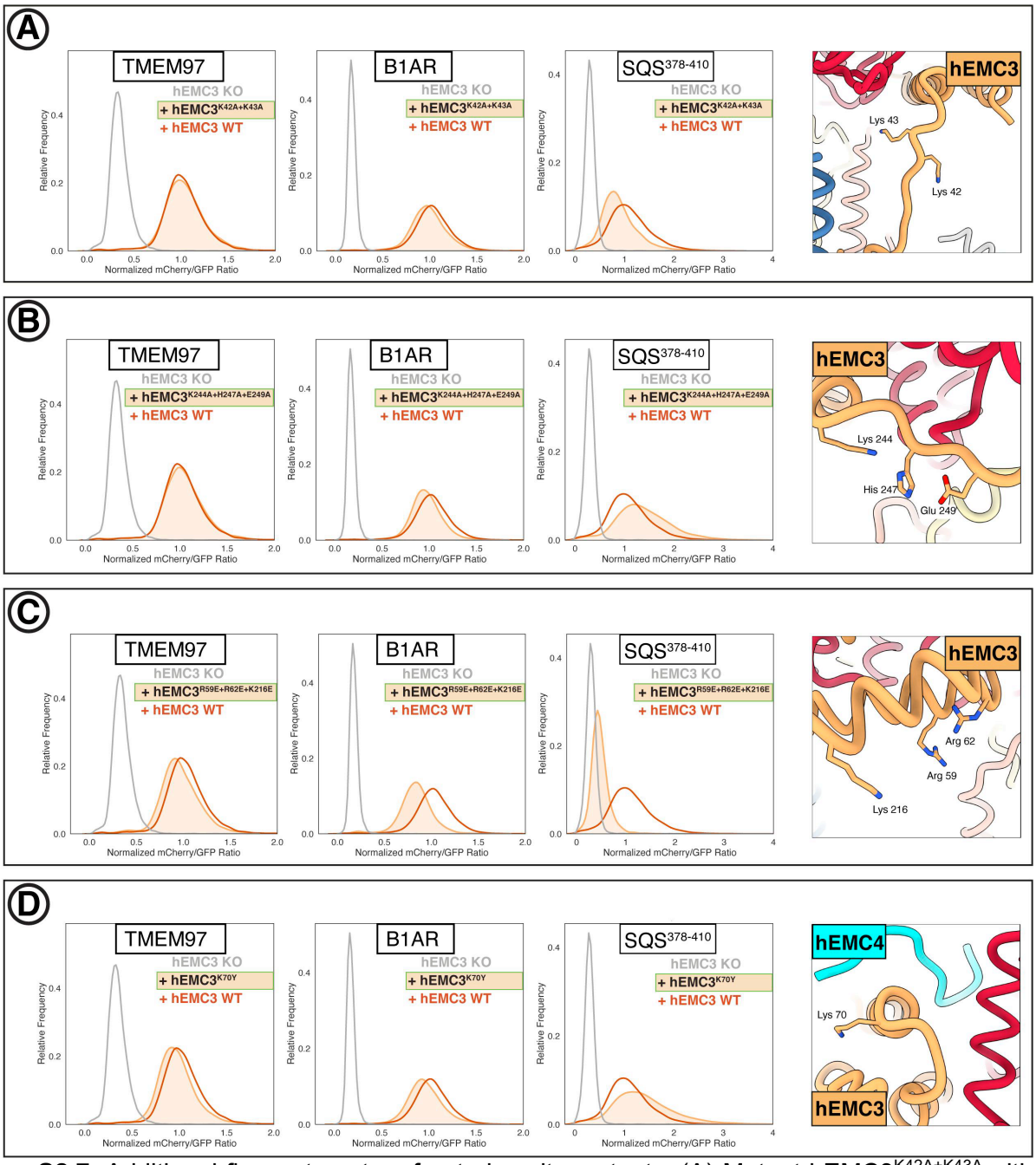


Figure S2.7. Additional flow cytometry of gated cavity mutants. (A) Mutant hEMC3$^{K42A+K43A}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (B) Mutant hEMC3$^{K244A+H247A+E249A}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (C) Mutant hEMC3$^{R59E+R62E+K216E}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (D) Mutant hEMC3^{K70Y} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated.

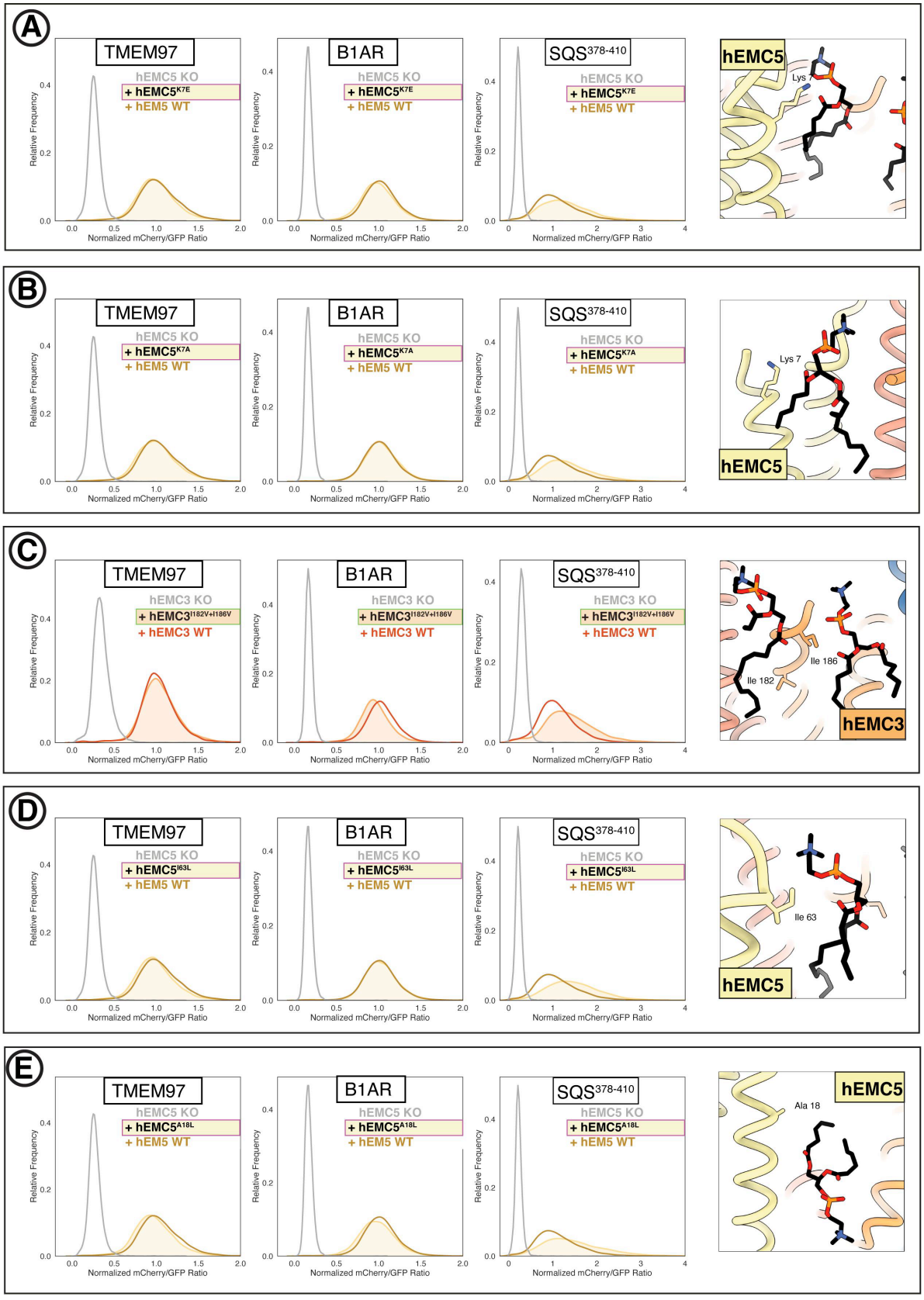

A
TMEM97
hEMC5 KO
+ hEMC5K7E
+ hEM5 WT
B1AR
hEMC5 KO
+ hEMC5K7E
+ hEM5 WT
SQS378-410
hEMC5 KO
+ hEMC5K7E
+ hEM5 WT
Relative Frequency
Normalized mCherry/GFP Ratio
hEMC5
Lys 7
B
TMEM97
hEMC5 KO
+ hEMC5K7A
+ hEM5 WT
B1AR
hEMC5 KO
+ hEMC5K7A
+ hEM5 WT
SQS378-410
hEMC5 KO
+ hEMC5K7A
+ hEM5 WT
Relative Frequency
Normalized mCherry/GFP Ratio
Lys 7
hEMC5
C
TMEM97
hEMC3 KO
+ hEMC3I182V+I186V
+ hEMC3 WT
B1AR
hEMC3 KO
+ hEMC3I182V+I186V
+ hEMC3 WT
SQS378-410
hEMC3 KO
+ hEMC3I182V+I186V
+ hEMC3 WT
Relative Frequency
Normalized mCherry/GFP Ratio
Ile 186
Ile 182
hEMC3
D
TMEM97
hEMC5 KO
+ hEMC5I63L
+ hEM5 WT
B1AR
hEMC5 KO
+ hEMC5I63L
+ hEM5 WT
SQS378-410
hEMC5 KO
+ hEMC5I63L
+ hEM5 WT
Relative Frequency
Normalized mCherry/GFP Ratio
Ile 63
hEMC5
E
TMEM97
hEMC5 KO
+ hEMC5A18L
+ hEM5 WT
B1AR
hEMC5 KO
+ hEMC5A18L
+ hEM5 WT
SQS378-410
hEMC5 KO
+ hEMC5A18L
+ hEM5 WT
Relative Frequency
Normalized mCherry/GFP Ratio
Ala 18
hEMC5

Figure S2.8. Flow cytometry of lipid-filled cavity mutants. (A) Mutant hEMC5^{K7E} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (B) Mutant hEMC5K7A with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (C) Mutant hEMC3$^{I182V+I186V}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (D) Mutant hEMC5^{I63L} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (E) Mutant hEMC5^{A18L} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated.

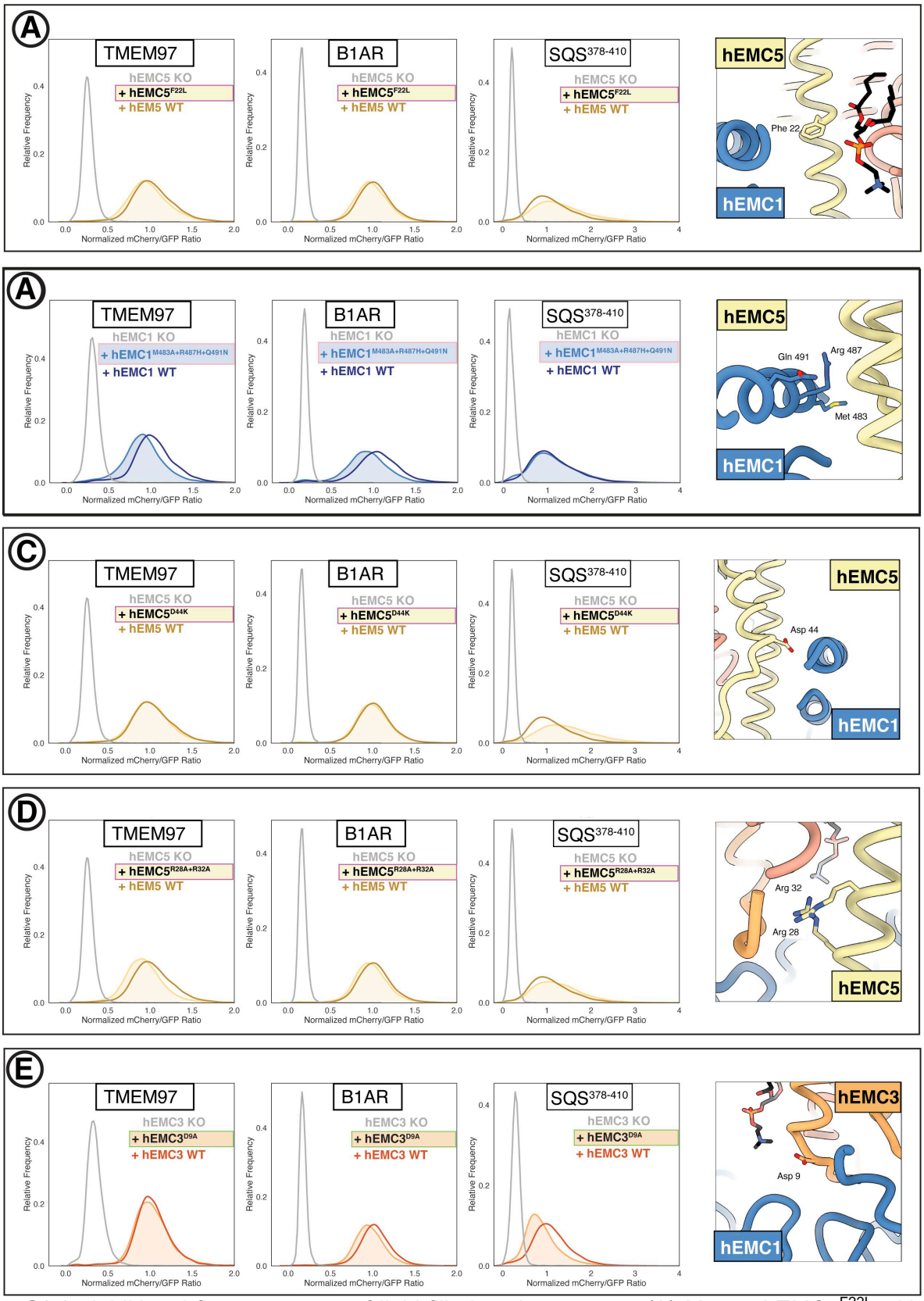


Figure S2.9. Additional flow cytometry of lipid-filled cavity mutants. (A) Mutant hEMC5^{F22L} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND

model displaying the residues mutated. (B) Mutant hEMC1$^{M483A+R487H+Q491N}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (C) Mutant hEMC5$^{R28A+R32A}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (D) Mutant hEMC3^{D9A} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (E) Mutant hEMC5^{D44K} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated.

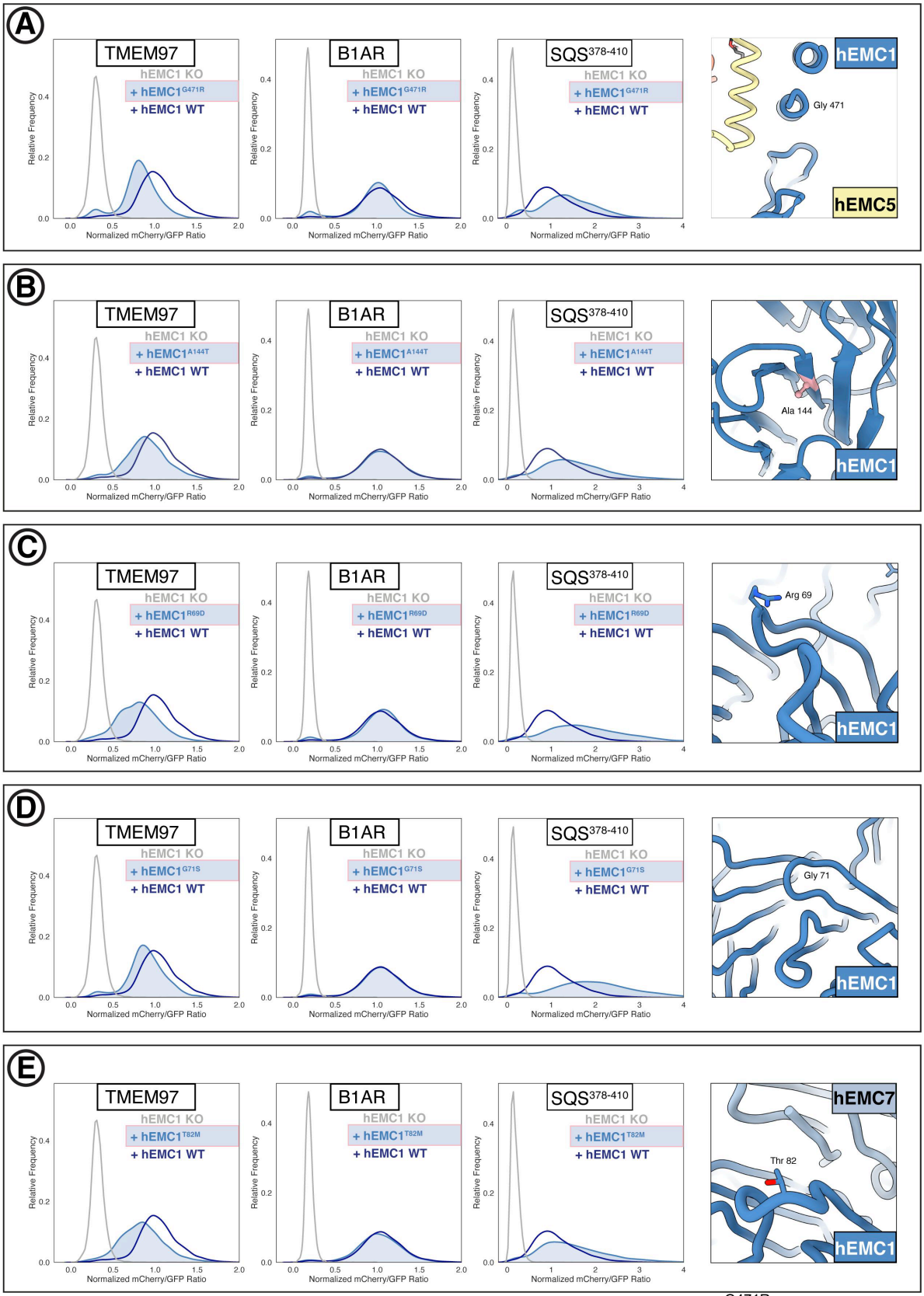


Figure S2.10. Flow cytometry of lumenal domain mutants. (A) Mutant hEMC1^{G471R} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model

displaying the residues mutated. (B) Mutant hEMC1^{D31K} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (C) Mutant hEMC1^{R69D} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (D) Mutant hEMC1^{G71S} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (E) Mutant hEMC1^{T82M} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated.

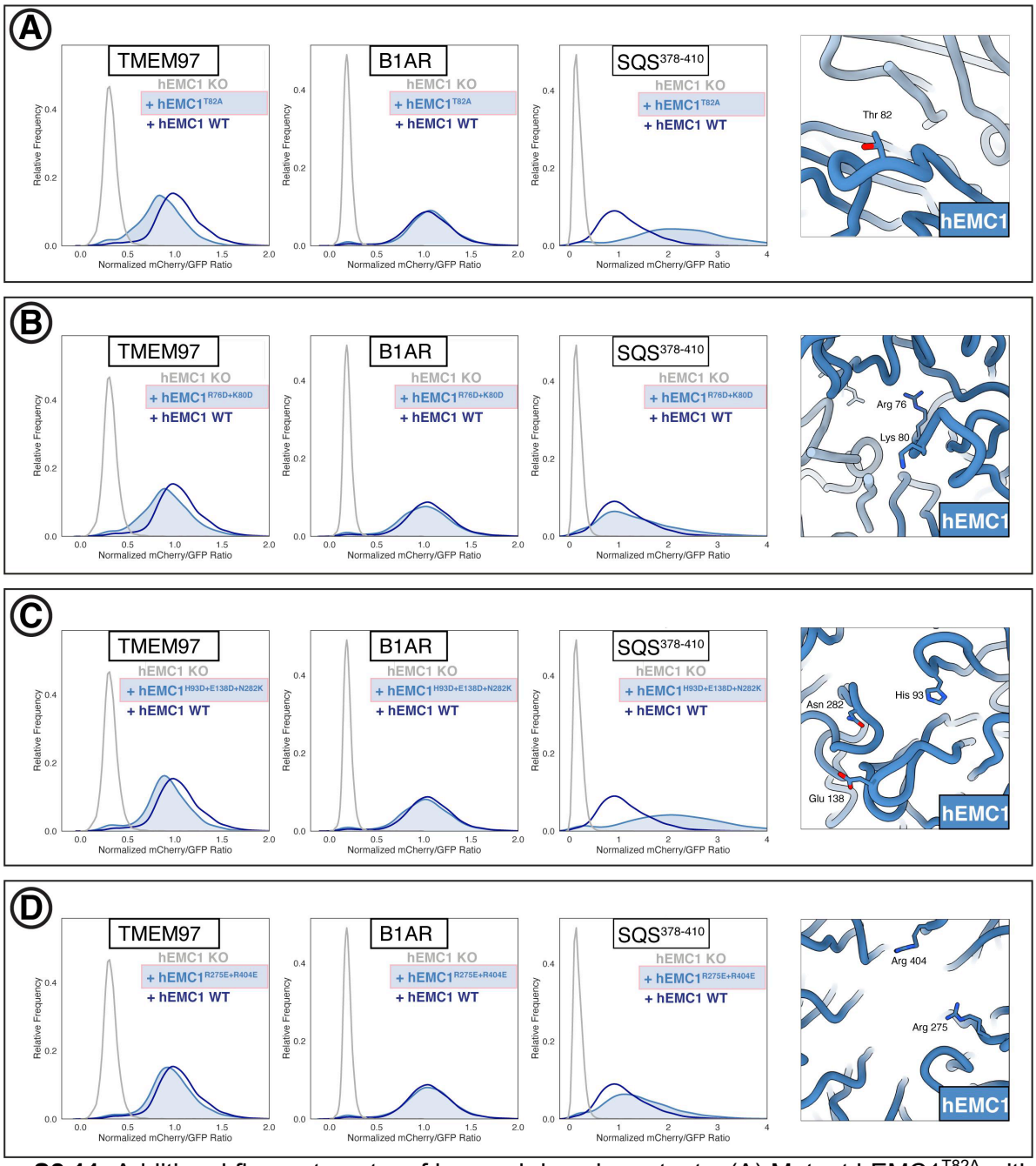


Figure S2.11. Additional flow cytometry of lumenal domain mutants. (A) Mutant hEMC1^{T82A} with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (B) Mutant hEMC1$^{R76D+K80D}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (C) Mutant hEMC1$^{H93D+E138D+N282K}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated. (D) Mutant hEMC1$^{R275E+R404E}$ with TMEM97-mCherry, B1AR-mCherry, and mCherry-SQScterm cell lines. Snapshot of hEMC ND model displaying the residues mutated.

REFERENCES

Abu-Safieh L, Alrashed M, Anazi S, Alkuraya H, Khan AO, Al-Owain M, Al-Zahrani J, Al-Abdi L, Hashem M, Al-Tarimi S, Sebai MA, Shamia A, Ray-Zack MD, Nassan M, Al-Hassnan ZN, Rahbeeni Z, Waheeb S, Alkharashi A, Abboud E, Al-Hazzaa SA, et al. 2013. Autozygome-guided exome sequencing in retinal dystrophy patients reveals pathogenetic mutations and novel candidate disease genes. *Genome Research* 23:236–247. DOI: https://doi.org/10.1101/gr.144105.112, PMID: 23105016

Adams PD, Afonine PV, Bunko´ czi G, Chen VB, Echols N, Headd JJ, Hung LW, Jain S, Kapral GJ, Grosse Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner RD, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2011. The Phenix software for automated determination of macromolecular structures. *Methods* 55: 94–106. DOI: https://doi.org/10.1016/j.ymeth.2011.07.005, PMID: 21821126

Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, Adams PD. 2018. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallographica. Section D, *Structural Biology.* 74:531–544. DOI: https://doi.org/10.1107/S2059798318006551, PMID: 29872004

Amberger JS, Bocchini CA, Scott AF, Hamosh A. 2019. OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Research* 47:D1038–D1043. DOI: https://doi.org/10.1093/nar/gky1151, PMID: 30445645

Anghel SA, McGilvray PT, Hegde RS, Keenan RJ. 2017. Identification of Oxa1 homologs operating in the eukaryotic endoplasmic reticulum. *Cell Reports* 21:3708–3716. DOI: https://doi.org/10.1016/j.celrep.2017.12. 006, PMID: 29281821

Asarnow D, Palovcak E, Cheng Y. 2019. UCSF pyem. *Zenodo*. v0.5.https://doi.org/10.5281/zenodo.3576630

Assimon VA, Southworth DR, Gestwicki JE. 2015. Specific binding of tetratricopeptide repeat proteins to heat shock protein 70 (Hsp70) and heat shock protein 90 (Hsp90) Is regulated

by affinity and phosphorylation. *Biochemistry* 54:7120–7131. DOI: https://doi.org/10.1021/acs.biochem.5b00801, PMID: 26565746

Bagchi P, Inoue T, Tsai B. 2016. EMC1-dependent stabilization drives membrane penetration of a partially destabilized non-enveloped virus. *eLife* 5:e21470. DOI: https://doi.org/10.7554/eLife.21470, PMID: 28012275

Bai L, You Q, Feng X, Kovach A, Li H. 2020. Structure of the ER membrane complex, a transmembrane-domain insertase. *Nature* 584:475–478. DOI: https://doi.org/10.1038/s41586-020-2389-3, PMID: 32494008

Bircham PW, Maass DR, Roberts CA, Kiew PY, Low YS, Yegambaram M, Matthews J, Jack CA, Atkinson PH. 2011. Secretory pathway genes assessed by high-throughput microscopy and synthetic genetic array analysis. Molecular BioSystems 7:2589. DOI: https://doi.org/10.1039/c1mb05175j, PMID: 21731954

Biyani N, Righetto RD, McLeod R, Caujolle-Bert D, Castano-Diez D, Goldie KN, Stahlberg H. 2017. Focus: the interface between data collection and data processing in cryo-EM. *Journal of Structural Biology* 198:124–133. DOI: https://doi.org/10.1016/j.jsb.2017.03.007, PMID: 28344036

Blatch GL, La¨ ssle M. 1999. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *BioEssays* 21:932–939. DOI: https://doi.org/10.1002/(SICI)1521-1878(199911)21:11<932::AID-BIES5>3.0.CO;2-N, PMID: 10517866

Borowska MT, Dominik PK, Anghel SA, Kossiakoff AA, Keenan RJ. 2015. A YidC-like protein in the archaeal plasma membrane. *Structure* 23:1715–1724. DOI: https://doi.org/10.1016/j.str.2015.06.025, PMID: 26256539

Chaudhary S, Pak JE, Gruswitz F, Sharma V, Stroud RM. 2012. Overexpressing human membrane proteins in stably transfected and clonal human embryonic kidney 293S cells. *Nature Protocols* 7:453–466. DOI: https://doi.org/10.1038/nprot.2011.453, PMID: 22322218

Chitwood PJ, Juszkiewicz S, Guna A, Shao S, Hegde RS. 2018. EMC is required to initiate accurate membrane protein topogenesis. *Cell* 175:1507–1519. DOI: https://doi.org/10.1016/j.cell.2018.10.009, PMID: 30415835

Cho H, Stanzione F, Oak A, Kim GH, Yerneni S, Qi L, Sum AK, Chan C. 2019. Intrinsic structural features of the human IRE1a transmembrane domain sense membrane lipid saturation. *Cell Reports* 27:307–320. DOI: https://doi.org/10.1016/j.celrep.2019.03.017, PMID: 30943411

Christianson JC, Olzmann JA, Shaler TA, Sowa ME, Bennett EJ, Richter CM, Tyler RE, Greenblatt EJ, Wade Harper J, Kopito RR. 2012. Defining human ERAD networks through an integrative mapping strategy. *Nature Cell Biology* 14:93–105. DOI: https://doi.org/10.1038/ncb2383

Coelho JPL, Stahl M, Bloemeke N, Meighen-Berger K, Alvira CP, Zhang ZR, Sieber SA, Feige MJ. 2019. A network of chaperones prevents and detects failures in membrane protein lipid bilayer integration. *Nature Communications* 10:672. DOI: https://doi.org/10.1038/s41467-019-08632-0, PMID: 30737405

Cornell CE, Mileant A, Thakkar N, Lee KK, Keller SL. 2020. Direct imaging of liquid domains in membranes by cryo-electron tomography. *PNAS* 117:19713–19719. DOI: https://doi.org/10.1073/pnas.2002245117, PMID: 32759217

Corradi V, Sejdiu BI, Mesa-Galloso H, Abdizadeh H, Noskov SY, Marrink SJ, Tieleman DP. 2019. Emerging diversity in Lipid-Protein interactions. *Chemical Reviews* 119:5775–5848. DOI: https://doi.org/10.1021/acs.chemrev.8b00451, PMID: 30758191

Costa EA, Subramanian K, Nunnari J, Weissman JS. 2018. Defining the physiological role of SRP in protein-targeting efficiency and specificity. *Science* 359:689–692. DOI: https://doi.org/10.1126/science.aar3607, PMID: 29348368

Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* 26:1367–1372. DOI: https://doi.org/10.1038/nbt.1511, PMID: 19029910

Cuff JA, Barton GJ. 2000. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins: Structure, Function, and Genetics* 40:502–511. DOI: https://doi.org/10.1002/1097-0134(20000815)40:3<502::AID-PROT170>3.0.CO;2-Q, PMID: 10861942

Dalbey RE, Kuhn A. 2015. Membrane insertases are present in all three domains of life. *Structure* 23:1559–1560. DOI: https://doi.org/10.1016/j.str.2015.08.002, PMID: 26331454

Denic V, Do¨ tsch V, Sinning I. 2013. Endoplasmic reticulum targeting and insertion of tail-anchored membrane proteins by the GET pathway. *Cold Spring Harbor Perspectives in Biology* 5:a013334. DOI: https://doi.org/10.1101/cshperspect.a013334, PMID: 23906715

Diamantopoulou A, Sun Z, Mukai J, Xu B, Fenelon K, Karayiorgou M, Gogos JA. 2017. Loss-of-function mutation in Mirta22/Emc10 rescues specific schizophrenia-related phenotypes in a mouse model of the 22q11.2 deletion. *PNAS* 114:E6127–E6136. DOI: https://doi.org/10.1073/pnas.1615719114, PMID: 28696314

Dickinson ME, Flenniken AM, Ji X, Teboul L, Wong MD, White JK, Meehan TF, Weninger WJ, Westerberg H, Adissu H, Baker CN, Bower L, Brown JM, Caddle LB, Chiani F, Clary D, Cleak J, Daly MJ, Denegre JM, Doe B, et al. 2016. High-throughput discovery of novel developmental phenotypes. *Nature* 537:508–514. DOI: https://doi.org/10.1038/nature19356, PMID: 27626380

Ellgaard L, McCaul N, Chatsisvili A, Braakman I. 2016. Co- and Post-Translational protein folding in the ER: co-and Post-Translational protein folding in the ER. *Traffic* 17:615–638. DOI: https://doi.org/10.1111/tra.12392, PMID: 26947578

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of coot. *Acta Crystallographica*. Section D, Biological Crystallography 66:486–501. DOI: https://doi.org/10.1107/S0907444910007493, PMID: 20383002

Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallographica*. Section D, Biological Crystallography 60:2126–2132. DOI: https://doi.org/10.1107/S0907444904019158, PMID: 15572765

Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, Qi LS, Kampmann M, Weissman JS. 2014. Genome-Scale CRISPR-Mediated control of gene repression and activation. *Cell* 159:647–661. DOI: https://doi.org/10.1016/j.cell.2014.09.029, PMID: 25307932

Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Science* 27:14–25. DOI: https://doi.org/10.1002/pro.3235, PMID: 28710774

Goehring A, Lee CH, Wang KH, Michel JC, Claxton DP, Baconguis I, Althoff T, Fischer S, Garcia KC, Gouaux E. 2014. Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nature Protocols* 9:2574–2585. DOI: https://doi.org/10.1038/nprot.2014.173, PMID: 25299155

Graham JB, Canniff NP, Hebert DN. 2019. TPR-containing proteins control protein organization and homeostasis for the endoplasmic reticulum. *Critical Reviews in Biochemistry and Molecular Biology* 54:103–118. DOI: https://doi.org/10.1080/10409238.2019.1590305, PMID: 31023093

Guna A, Volkmar N, Christianson JC, Hegde RS. 2018. The ER membrane protein complex is a transmembrane domain insertase. *Science* 359:470–473. DOI: https://doi.org/10.1126/science.aao3099, PMID: 29242231

Hagn F, Nasr ML, Wagner G. 2018. Assembly of phospholipid nanodiscs of controlled size for structural studies of membrane proteins by NMR. *Nature Protocols* 13:79–98. DOI: https://doi.org/10.1038/nprot.2017.094, PMID: 29215632

Halbleib K, Pesek K, Covino R, Hofbauer HF, Wunnicke D, Ha¨ nelt I, Hummer G, Ernst R. 2017. Activation of the unfolded protein response by lipid bilayer stress. *Molecular Cell* 67:673–684. DOI: https://doi.org/10.1016/j.molcel.2017.06.012, PMID: 28689662

Harel T, Yesil G, Bayram Y, Coban-Akdemir Z, Charng WL, Karaca E, Al Asmari A, Eldomery MK, Hunter JV, Jhangiani SN, Rosenfeld JA, Pehlivan D, El-Hattab AW, Saleh MA, LeDuc CA, Muzny D, Boerwinkle E, Gibbs RA, Chung WK, Yang Y, et al. 2016. Monoallelic and biallelic variants in EMC1 identified in individuals with global developmental delay, Hypotonia, scoliosis, and cerebellar atrophy. *The American Journal of Human Genetics* 98:562–570. DOI: https://doi.org/10.1016/j.ajhg.2016.01.011, PMID: 26942288

Heberle FA, Doktorova M, Scott HL, Skinkle AD, Waxham MN, Levental I. 2020. Direct label-free imaging of nanodomains in biomimetic and biological membranes by cryogenic electron microscopy. *PNAS* 117:19943–19952. DOI: https://doi.org/10.1073/pnas.2002200117, PMID: 32759206

Hiramatsu N, Tago T, Satoh T, Satoh AK. 2019. ER membrane protein complex is required for the insertions of late-synthesized transmembrane helices of Rh1 in Drosophila photoreceptors. *Molecular Biology of the Cell* 30: 2890–2900. DOI: https://doi.org/10.1091/mbc.E19-08-0434, PMID: 31553680

Hofbauer HF, Gecht M, Fischer SC, Seybert A, Frangakis AS, Stelzer EHK, Covino R, Hummer G, Ernst R. 2018. The molecular recognition of phosphatidic acid by an amphipathic helix in Opi1. *Journal of Cell Biology* 217: 3109–3126. DOI: https://doi.org/10.1083/jcb.201802027, PMID: 29941475

Hu M, Yu H, Gu K, Wang Z, Ruan H, Wang K, Ren S, Li B, Gan L, Xu S, Yang G, Shen Y, Li X. 2018. A particle-filter framework for robust cryo-EM 3D reconstruction. Nature Methods 15:1083–1089. DOI: https://doi.org/10.1038/s41592-018-0223-8, PMID: 30504871

Jacquemyn J, Cascalho A, Goodchild RE. 2017. The ins and outs of endoplasmic reticulum-controlled lipid biosynthesis. *EMBO Reports* 18:1905–1921. DOI: https://doi.org/10.15252/embr.201643426, PMID: 29074503

Jones DT, Taylor WR, Thornton JM. 1994. A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* 33:3038–3049. DOI: https://doi.org/10.1021/bi00176a037, PMID: 8130217

Jones DT. 1999. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology* 292:195–202. DOI: https://doi.org/10.1006/jmbi.1999.3091, PMID: 10493868

Jonikas MC, Collins SR, Denic V, Oh E, Quan EM, Schmid V, Weibezahn J, Schwappach B, Walter P, Weissman JS, Schuldiner M. 2009. Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* 323:1693–1697. DOI: https://doi.org/10.1126/science.1167983, PMID: 19325107

Junes-Gill KS, Gallaher TK, Gluzman-Poltorak Z, Miller JD, Wheeler CJ, Fan X, Basile LA. 2011. hHSS1: a novel secreted factor and suppressor of glioma growth located at chromosome 19q13.33. *Journal of Neuro-Oncology* 102:197–211. DOI: https://doi.org/10.1007/s11060-010-0314-6, PMID: 20680400

Kallberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu J. 2012. Template-based protein structure modeling using the RaptorX web server. *Nature Protocols* 7:1511–1522. DOI: https://doi.org/10.1038/nprot.2012.085,PMID: 22814390

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 10:845–858. DOI: https://doi.org/10.1038/nprot.2015.053, PMID: 25950237

Kim J, Stroud RM, Craik CS. 2011. Rapid identification of recombinant fabs that bind to membrane proteins. Methods 55:303–309. DOI: https://doi.org/10.1016/j.ymeth.2011.09.012, PMID: 21958987

Kimanius D, Forsberg BO, Scheres SH, Lindahl E. 2016. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* 5:e18722. DOI: https://doi.org/10.7554/eLife.18722, PMID: 27845625

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology* 305:567–580. DOI: https://doi.org/10.1006/jmbi.2000.4315, PMID: 11152613

Krysztofinska EM, Evans NJ, Thapaliya A, Murray JW, Morgan RML, Martinez-Lumbreras S, Isaacson RL. 2017. Structure and interactions of the TPR domain of Sgt2 with yeast chaperones and Ybr137wp. *Frontiers in Molecular Biosciences* 4:68. DOI: https://doi.org/10.3389/fmolb.2017.00068, PMID: 29075633

Kudze T, Mendez-Dorantes C, Jalloh CS, McClellan AJ. 2018. Evidence for interaction between Hsp90 and the ER membrane complex. *Cell Stress and Chaperones* 23:1101–1115. DOI: https://doi.org/10.1007/s12192-018-0908-z, PMID: 29808299

Kumazaki K, Kishimoto T, Furukawa A, Mori H, Tanaka Y, Dohmae N, Ishitani R, Tsukazaki T, Nureki O. 2014. Crystal structure of Escherichia coli YidC, a membrane protein chaperone and insertase. Scientific Reports 4: 7299. DOI: https://doi.org/10.1038/srep07299, PMID: 25466392

Lahiri S, Chao JT, Tavassoli S, Wong AK, Choudhary V, Young BP, Loewen CJ, Prinz WA. 2014. A conserved endoplasmic reticulum membrane protein complex (EMC) facilitates phospholipid transfer from the ER to mitochondria. *PLOS Biology* 12:e1001969. DOI: https://doi.org/10.1371/journal.pbio.1001969, PMID: 25313861

Laverty D, Desai R, Uchan´ ski T, Masiulis S, Stec WJ, Malinauskas T, Zivanov J, Pardon E, Steyaert J, Miller KW, Aricescu AR. 2019. Cryo-EM structure of the human a1b3g2 GABAA receptor in a lipid bilayer. *Nature* 565: 516–520. DOI: https://doi.org/10.1038/s41586-018-0833-4, PMID: 30602789

Li X, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, Agard DA, Cheng Y. 2013. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature Methods* 10:584–590. DOI: https://doi.org/10.1038/nmeth.2472, PMID: 23644547

Liebschner D, Afonine PV, Baker ML, Bunko´ czi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, Stockwell DH, et al. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in phenix. *Acta Crystallographica* Section D Structural Biology 75:861–877. DOI: https://doi.org/10.1107/S2059798319011471, PMID: 31588918

Lin DL, Inoue T, Chen YJ, Chang A, Tsai B, Tai AW. 2019. The ER membrane protein complex promotes biogenesis of dengue and zika virus Non-structural Multi-pass transmembrane proteins to support infection. *Cell Reports* 27:1666–1674. DOI: https://doi.org/10.1016/j.celrep.2019.04.051, PMID: 31067454

Louie RJ, Guo J, Rodgers JW, White R, Shah N, Pagant S, Kim P, Livstone M, Dolinski K, McKinney BA, Hong J, Sorscher EJ, Bryan J, Miller EA, Hartman JL. 2012. A yeast phenomic model for the gene interaction network modulating CFTR-DF508 protein biogenesis. *Genome Medicine* 4:103. DOI: https://doi.org/10.1186/gm404, PMID: 23270647

Luo WJ, Gong XH, Chang A. 2002. An ER membrane protein, Sop4, facilitates ER export of the yeast plasma membrane [H+] ATPase, Pma1. *Traffic* 3:730–739. DOI: https://doi.org/10.1034/j.1600-0854.2002.31005.x, PMID: 12230471

Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. *Science* 252:1162–1164. DOI: https://doi.org/10.1126/science.252.5009.1162, PMID: 2031185

Marinko JT, Huang H, Penn WD, Capra JA, Schlebach JP, Sanders CR. 2019. Folding and misfolding of human membrane proteins in health and disease: from single molecules to cellular proteostasis. *Chemical Reviews* 119:5537–5606. DOI: https://doi.org/10.1021/acs.chemrev.8b00532, PMID: 30608666

Marquez J, Criscione J, Charney RM, Prasad MS, Hwang WY, Mis EK, Garcia-Castro MI, Khokha MK. 2020. Disrupted ER membrane protein complex-mediated topogenesis drives congenital neural crest defects. *Journal of Clinical Investigation* 130:813–826. DOI: https://doi.org/10.1172/JCI129308, PMID: 31904590

Mastronarde DN. 2005. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology* 152:36–51. DOI: https://doi.org/10.1016/j.jsb.2005.07.007, PMID: 16182563

Mateja A, Keenan RJ. 2018. A structural perspective on tail-anchored protein biogenesis by the GET pathway. *Current Opinion in Structural Biology* 51:195–202. DOI: https://doi.org/10.1016/j.sbi.2018.07.009, PMID: 30173121

McDowell MA, Heimes M, Fiorentino F, Mehmood S, Farkas A, Coy-Vergara J, Wu D, Bolla JR, Schmid V, Heinze R, Wild K, Flemming D, Pfeffer S, Schwappach B, Robinson CV, Sinning I. 2020. Structural basis of Tail-Anchored membrane protein biogenesis by the GET insertase complex. *Molecular Cell* 80:72–86. DOI: https://doi.org/10.1016/j.molcel.2020.08.012, PMID: 32910895

McGilvray PT, Anghel SA, Sundaram A, Zhong F, Trnka MJ, Fuller JR, Hu H, Burlingame AL, Keenan RJ. 2020. An ER translocon for multi-pass membrane protein biogenesis. *eLife* 9:e56889. DOI: https://doi.org/10.7554/eLife.56889, PMID: 32820719

Mitra K, Ubarretxena-Belandia I, Taguchi T, Warren G, Engelman DM. 2004. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than

cholesterol. *PNAS* 101:4083–4088. DOI: https://doi.org/10.1073/pnas.0307332101, PMID: 15016920

Morin A, Eisenbraun B, Key J, Sanschagrin PC, Timony MA, Ottaviano M, Sliz P. 2013. Collaboration gets the most out of software. *eLife* 2:e01456. DOI: https://doi.org/10.7554/eLife.01456, PMID: 24040512

Ngo AM, Shurtleff MJ, Popova KD, Kulsuptrakul J, Weissman JS, Puschnik AS. 2019. The ER membrane protein complex is required to ensure correct topology and stable expression of flavivirus polyproteins. *eLife* 8:e48469. DOI: https://doi.org/10.7554/eLife.48469, PMID: 31516121

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* 302:205–217. DOI: https://doi.org/10.1006/jmbi.2000.4042, PMID: 10964570

O'Donnell JP, Phillips BP, Yagita Y, Juszkiewicz S, Wagner A, Malinverni D, Keenan RJ, Miller EA, Hegde RS. 2020. The architecture of EMC reveals a path for membrane protein insertion. *eLife* 9:e57887. DOI: https://doi.org/10.7554/eLife.57887, PMID: 32459176

Obradovic Z, Peng K, Vucetic S, Radivojac P, Dunker AK. 2005. Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins: Structure, Function, and Bioinformatics* 61 Suppl 7:176–182. DOI: https://doi.org/10.1002/prot.20735, PMID: 16187360

Ouali M, King RD. 2000. Cascaded multiple classifiers for secondary structure prediction. Protein *Science* 9:1162–1176. DOI: https://doi.org/10.1110/ps.9.6.1162, PMID: 10892809

Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z. 2006. Length-dependent prediction of protein intrinsic disorder. *BMC Bioinformatics* 7:208. DOI: https://doi.org/10.1186/1471-2105-7-208, PMID: 16618368

Petkovic M, Oses-Prieto J, Burlingame A, Jan LY, Jan YN. 2020. TMEM16K is an interorganelle regulator of endosomal sorting. *Nature Communications* 11:3298. DOI: https://doi.org/10.1038/s41467-020-17016-8, PMID: 32620747

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF chimera–a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25:1605–1612. DOI: https://doi.org/10.1002/jcc.20084, PMID: 15264254

Pfeffer S, Dudek J, Schaffer M, Ng BG, Albert S, Plitzko JM, Baumeister W, Zimmermann R, Freeze HH, Engel BD, Forster F. 2017. Dissecting the molecular organization of the translocon-associated protein complex. *Nature Communications* 8:14516. DOI: https://doi.org/10.1038/ncomms14516, PMID: 28218252

Pleiner T, Tomaleri GP, Januszyk K, Inglis AJ, Hazu M, Voorhees RM. 2020. Structural basis for membrane insertion by the human ER membrane protein complex. *Science* 369:433–436. DOI: https://doi.org/10.1126/science.abb5008, PMID: 32439656

Probert F, Rice P, Scudamore CL, Wells S, Williams R, Hough TA, Cox IJ. 2015. ÅH NMR metabolic profiling of plasma reveals additional phenotypes in knockout mouse models. *Journal of Proteome Research* 14:2036–2045. DOI: https://doi.org/10.1021/pr501039k, PMID: 25849460

Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* 14:290–296. DOI: https://doi.org/10.1038/nmeth.4169, PMID: 28165473

Punjani A, Zhang H, Fleet DJ. 2019. Non-uniform refinement: adaptive regularization improves single particle cryo-EM reconstruction. *bioRxiv*. DOI: https://doi.org/10.1101/2019.12.15.877092

Punjani A, Fleet DJ. 2020. 3D variability analysis: directly resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM images. *bioRxiv*. DOI: https://doi.org/10.1101/2020.04.08.032466

Raman S, Vernon R, Thompson J, Tyka M, Sadreyev R, Pei J, Kim D, Kellogg E, DiMaio F, Lange O, Kinch L, Sheffler W, Kim BH, Das R, Grishin NV, Baker D. 2009. Structure prediction for CASP8 with all-atom refinement using rosetta. *Proteins: Structure, Function, and Bioinformatics* 77 Suppl 9:89–99. DOI: https://doi.org/10.1002/prot.22540, PMID: 19701941

Ramı´rez AS, Kowal J, Locher KP. 2019. Cryo-electron microscopy structures of human oligosaccharyltransferase complexes OST-A and OST-B. *Science* 366:1372–1375. DOI: https://doi.org/10.1126/science.aaz3505, PMID: 31831667

Ramlaul K, Palmer CM, Nakane T, Aylett CHS. 2020. Mitigating local over-fitting during single particle reconstruction with SIDESPLITTER. *Journal of Structural Biology* 211:107545. DOI: https://doi.org/10.1016/j.jsb.2020.107545, PMID: 32534144

Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols* 2:1896–1906. DOI: https://doi.org/10.1038/nprot.2007.261, PMID: 17703201

Reinisch KM, De Camilli P. 2016. SMP-domain proteins at membrane contact sites: structure and function. *Biochimica Et Biophysica Acta* (BBA) - Molecular and Cell Biology of Lipids 1861:924–927. DOI: https://doi.org/10.1016/j.bbalip.2015.12.003, PMID: 26686281

Richard M, Boulin T, Robert VJ, Richmond JE, Bessereau JL. 2013. Biosynthesis of ionotropic acetylcholine receptors requires the evolutionarily conserved ER membrane complex. *PNAS* 110:E1055–E1063. DOI: https://doi.org/10.1073/pnas.1216154110, PMID: 23431131

Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, Sligar SG. 2009. Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs.

Methods in Enzymology 464:211–231. DOI: https://doi.org/10.1016/S0076-6879(09)64011-8, PMID: 19903557

Rost B. 2001. Review: protein secondary structure prediction continues to rise. *Journal of Structural Biology* 134: 204–218. DOI: https://doi.org/10.1006/jsbi.2001.4336, PMID: 11551180

Rothstein R. 1991. Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. *Methods in Enzymology* 194:281–301. DOI: https://doi.org/10.1016/0076-6879(91)94022-5, PMID: 2005793

Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* 5:725–738. DOI: https://doi.org/10.1038/nprot.2010.5, PMID: 20360767

Salas-Estrada LA, Leioatts N, Romo TD, Grossfield A. 2018. Lipids alter rhodopsin function via Ligand-like and Solvent-like interactions. *Biophysical Journal* 114:355–367. DOI: https://doi.org/10.1016/j.bpj.2017.11.021, PMID: 29401433

Samuelson JC, Chen M, Jiang F, Moller I, Wiedmann M, Kuhn A, Phillips GJ, Dalbey RE. 2000. YidC mediates membrane protein insertion in Bacteria. *Nature* 406:637–641. DOI: https://doi.org/10.1038/35020586, PMID: 10949305

Satoh T, Ohba A, Liu Z, Inagaki T, Satoh AK. 2015. dPob/EMC is essential for biosynthesis of rhodopsin and other multi-pass membrane proteins in Drosophila photoreceptors. *eLife* 4:e06306. DOI: https://doi.org/10.7554/eLife.06306

Savidis G, McDougall WM, Meraner P, Perreira JM, Portmann JM, Trincucci G, John SP, Aker AM, Renzette N, Robbins DR, Guo Z, Green S, Kowalik TF, Brass AL. 2016. Identification of zika virus and dengue virus dependency factors using functional genomics. *Cell Reports* 16:232–246. DOI: https://doi.org/10.1016/j.celrep.2016.06.028, PMID: 27342126

Scheufler C, Brinker A, Bourenkov G, Pegoraro S, Moroder L, Bartunik H, Hartl FU, Moarefi I. 2000. Structure of TPR domain-peptide complexes: critical elements in the assembly of

the Hsp70-Hsp90 multichaperone machine. *Cell* 101:199–210. DOI: https://doi.org/10.1016/S0092-8674(00)80830-2, PMID: 10786835

Schlegel T, Mirus O, von Haeseler A, Schleiff E. 2007. The tetratricopeptide repeats of receptors involved in protein translocation across membranes. *Molecular Biology and Evolution* 24:2763–2774. DOI: https://doi.org/10.1093/molbev/msm211, PMID: 17905998

Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2011. Global quantification of mammalian gene expression control. Nature 473:337–342. DOI: https://doi.org/10.1038/nature10098, PMID: 21593866

Sejdiu BI, Tieleman DP. 2020. Lipid-Protein interactions are a unique property and defining feature of G Protein-Coupled receptors. *Biophysical Journal* 118:1887–1900. DOI: https://doi.org/10.1016/j.bpj.2020.03.008, PMID: 32272057

Shaikh TR, Gao H, Baxter WT, Asturias FJ, Boisset N, Leith A, Frank J. 2008. SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nature Protocols* 3:1941–1974. DOI: https://doi.org/10.1038/nprot.2008.156, PMID: 19180078

Shurtleff MJ, Itzhak DN, Hussmann JA, Schirle Oakdale NT, Costa EA, Jonikas M, Weibezahn J, Popova KD, Jan CH, Sinitcyn P, Vembar SS, Hernandez H, Cox J, Burlingame AL, Brodsky JL, Frost A, Borner GH, Weissman JS. 2018. The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins. *eLife* 7:e37018. DOI: https://doi.org/10.7554/eLife.37018, PMID: 29809151

Song Y, DiMaio F, Wang RY, Kim D, Miles C, Brunette T, Thompson J, Baker D. 2013. High-resolution comparative modeling with RosettaCM. *Structure* 21:1735–1742. DOI: https://doi.org/10.1016/j.str.2013.08. 005, PMID: 24035711

Stagg SM, Noble AJ, Spilman M, Chapman MS. 2014. ResLog plots as an empirical metric of the quality of cryo-EM reconstructions. *Journal of Structural Biology* 185:418–426. DOI: https://doi.org/10.1016/j.jsb.2013.12.010, PMID: 24384117

Suloway CJ, Chartron JW, Zaslaver M, Clemons WM. 2009. Model for eukaryotic tail-anchored protein binding based on the structure of Get3. *PNAS* 106:14849–14854. DOI: https://doi.org/10.1073/pnas.0907522106, PMID: 19706470

Talbot BE, Vandorpe DH, Stotter BR, Alper SL, Schlondorff JS. 2019. Transmembrane insertases and N-glycosylation critically determine synthesis, trafficking, and activity of the nonselective cation channel TRPC6. *Journal of Biological Chemistry* 294:12655–12669. DOI: https://doi.org/10.1074/jbc.RA119.008299, PMID: 31266804

Tang X, Snowball JM, Xu Y, Na CL, Weaver TE, Clair G, Kyle JE, Zink EM, Ansong C, Wei W, Huang M, Lin X, Whitsett JA. 2017. EMC3 coordinates surfactant protein and lipid homeostasis required for respiration. *Journal of Clinical Investigation* 127:4314–4325. DOI: https://doi.org/10.1172/JCI94152, PMID: 29083321

Tian S, Wu Q, Zhou B, Choi MY, Ding B, Yang W, Dong M. 2019. Proteomic analysis identifies membrane proteins dependent on the ER membrane protein complex. *Cell Reports* 28:2517–2526. DOI: https://doi.org/10.1016/j.celrep.2019.08.006, PMID: 31484065

Volkmar N, Thezenas ML, Louie SM, Juszkiewicz S, Nomura DK, Hegde RS, Kessler BM, Christianson JC. 2019. The ER membrane protein complex promotes biogenesis of sterol-related enzymes maintaining cholesterol homeostasis. *Journal of Cell Science* 132:jcs223453. DOI: https://doi.org/10.1242/jcs.223453, PMID: 30578317

Volkmar N, Christianson JC. 2020. Squaring the EMC - how promoting membrane protein biogenesis impacts cellular functions and organismal homeostasis. *Journal of Cell Science* 133:jcs243519. DOI: https://doi.org/10.1242/jcs.243519, PMID: 32332093

Volmer R, van der Ploeg K, Ron D. 2013. Membrane lipid saturation activates endoplasmic reticulum unfolded protein response transducers through their transmembrane domains. *PNAS* 110:4628–4633. DOI: https://doi.org/10.1073/pnas.1217611110, PMID: 23487760

Vyas NK, Vyas MN, Chervenak MC, Johnson MA, Pinto BM, Bundle DR, Quiocho FA. 2002. Molecular recognition of oligosaccharide epitopes by a monoclonal fab specific for Shigella

flexneri Y lipopolysaccharide: x-ray structures and thermodynamics. *Biochemistry* 41:13575–13586. DOI: https://doi.org/10.1021/bi0261387, PMID: 12427018

Wainberg M, Kamber RA, Balsubramani A, Meyers RM, Sinnott-Armstrong N, Hornburg D, Jiang L, Chan J, Jian R, Gu M, Shcherbina A, Dubreuil MM, Spees K, Snyder MP, Kundaje A, Bassik MC. 2019. A genome-wide almanac of co-essential modules assigns function to uncharacterized genes. *bioRxiv*. DOI: https://doi.org/10.1101/827071

Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *Journal of Molecular Biology* 337:635–645. DOI: https://doi.org/10.1016/j.jmb.2004.02.002, PMID: 15019783

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189–1191. DOI: https://doi.org/10.1093/bioinformatics/btp033, PMID: 19151095

Weissmann F, Petzold G, VanderLinden R, Huis In 't Veld PJ, Brown NG, Lampert F, Westermann S, Stark H, Schulman BA, Peters JM. 2016. biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. *PNAS* 113:E2564–E2569. DOI: https://doi.org/10.1073/pnas.1604935113, PMID: 27114506

Wideman JG. 2015. The ubiquitous and ancient ER membrane protein complex (EMC): tether or not? *F1000Research* 4:624. DOI: https://doi.org/10.12688/f1000research.6944.1, PMID: 26512320

Wu S, Avila-Sakar A, Kim J, Booth DS, Greenberg CH, Rossi A, Liao M, Li X, Alian A, Griner SL, Juge N, Yu Y, Mergel CM, Chaparro-Riggers J, Strop P, Tampe´ R, Edwards RH, Stroud RM, Craik CS, Cheng Y. 2012. Fabs enable single particle cryoEM studies of small proteins. *Structure* 20:582–592. DOI: https://doi.org/10.1016/j.str.2012.02.017, PMID: 22483106

Xiong L, Zhang L, Yang Y, Li N, Lai W, Wang F, Zhu X, Wang T. 2020. ER complex proteins are required for rhodopsin biosynthesis and photoreceptor survival in Drosophila and mice.

Cell Death & Differentiation 27: 646–661. DOI: https://doi.org/10.1038/s41418-019-0378-6, PMID: 31263175

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2015. The I-TASSER suite: protein structure and function prediction. *Nature Methods* 12:7–8. DOI: https://doi.org/10.1038/nmeth.3213, PMID: 25549265

Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9:40. DOI: https://doi.org/10.1186/1471-2105-9-40, PMID: 18215316

Zhang K. 2016. Gctf: real-time CTF determination and correction. *Journal of Structural Biology* 193:1–12. DOI: https://doi.org/10.1016/j.jsb.2015.11.003, PMID: 26592709

Zhang S, Xu C, Larrimore KE, Ng DTW. 2017. Slp1-Emp65: a guardian factor that protects folding polypeptides from promiscuous degradation. *Cell* 171:346–357. DOI: https://doi.org/10.1016/j.cell.2017.08.036, PMID: 28919078

Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nature Methods 14:331–332. DOI: https://doi.org/10.1038/nmeth.4193, PMID: 28250466

Zhou Y, Wu F, Zhang M, Xiong Z, Yin Q, Ru Y, Shi H, Li J, Mao S, Li Y, Cao X, Hu R, Liew CW, Ding Q, Wang X, Zhang Y. 2018. EMC10 governs male fertility via maintaining sperm ion balance. *Journal of Molecular Cell Biology* 10:503–514. DOI: https://doi.org/10.1093/jmcb/mjy024, PMID: 29659949

Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, Gabler F, So¨ ding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *Journal of Molecular Biology* 430:2237–2243. DOI: https://doi.org/10.1016/j.jmb.2017.12.007, PMID: 29258817

Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* 7:e42166. DOI: https://doi.org/10.7554/eLife.42166, PMID: 30412051

Chapter 3: Building a genetic roadmap of the mammalian Endoplasmic Reticulum Membrane Protein Complex (EMC)

Abstract

As discussed in Chapter I and Chapter II, we are just beginning to understand the role of newly characterized molecular machines and accessory factors involved in membrane protein stabilization and folding. However, studying membrane protein folding processes has been limited to using mainly biochemical and structural approaches. Although high-throughput unbiased approaches have proven useful in the identification of membrane protein homeostasis factors such as the EMC back in 2009 (Jonikas et al, 2009), only recently is screening emerging as a tool to ask questions about membrane protein stabilization and trafficking processes (Kotov et al., 2019; Coukos et al., 2021; Cecchetti et al., 2021). Additionally, with the advent tracking transcriptional changes upon gene perturbation at the single cell level (Adamson et al., 2016; Dixit et al., 2016), we now have the ability to understand the role of membrane protein folding machineries not just at the molecular level – which protein gets stabilized by which chaperone complex – but also at the overall cellular level – what happens in each cell when given genes are turned off one at a time in combination with perturbing membrane protein folding complexes. Given the detailed molecular characterization of the each of EMC's functional subunits described in Chapter II and their role in stabilization of a range of membrane proteins, EMC must play a crucial role in maintaining a healthy cell. As a result, we set out to understand the broader physiological role of the EMC by building a genetic roadmap of the complex. Here, we perform the first unbiased screen and a follow up Perturb-Seq experiment in the context of a perturbed mammalian EMC.

A genetic screen identifies pathways genetically interacting with the EMC

We wanted to identify cellular pathways that the EMC interacts with genetically in a genome-wide and unbiased manner. To do so, we first generated a genetic and functional knockout of a core EMC subunit, EMC4. Knocking out EMC4 led to the destabilization and

depletion of the whole EMC complex and depletion of its client membrane proteins. We show that upon reintroducing EMC4 lentivirally, we rescue the knockout phenotype (Fig 3.1). We leveraged our CRISPRi functional genomics platform to screen for genes whose knockdown in EMC4 knockout cells affected their viability compared to WT EMC4 cells (Gilbert et al., 2013; Horlbeck et al., 2016). Briefly, we infected chronic myeloid leukemia (K562) EMC4 KO, EMC4 WT or EMC4 Rescue cells expressing dCas9-KRAB (CRISPRi) with our genome-scale CRISPRi five-guide library (Fig 3.2A). After performing a 10-day screen, the resulting gene-level phenotypes revealed genes for which knockdown strongly affects viability in EMC4 knockout cells. These genes can be characterized as buffering or synthetic with EMC (Fig 3.2B). Buffering arises when the combined perturbation of two genes results in a less severe phenotype than expected indicating the genes act in a linear pathway or encode subunits of a functional protein complex. Synthetic interactions are those in which the combined perturbation of two genes leads to a more severe phenotype than perturbing either of them alone. Synthetic interactions classically reflect genes acting in parallel pathways.

The EMC4 genetic interaction screen resulted in ~200 genes that were synthetic with EMC4 (Fig 3.2C). Before starting to investigate new biology, we wanted to confirm that the EMC4 genetic interaction screen replicated known membrane protein biology. Firstly, genetic knockdown of all the other EMC subunit genes (except for EMC1) resulted in mostly buffering interactions with EMC4 confirming their belonging to the same functional complex (Fig 3.2D). We believe that guides targeting EMC1 promoter also target the promoter of a nearby gene, MRTO4, leading to a its knockdown and phenotype different from the rest of the EMC genes (data not shown). Secondly, a number of genes responsible for tail-anchor protein insertion through the canonical TRC40 pathway (ASNA1, WRB1, CAMLG) and membrane protein folding (TMEM208, TMCO1) were strongly synthetic with EMC confirming the redundant role of EMC in tail-anchor insertion and multipass membrane protein stabilization with other characterized pathways Fig 3.2E). The EMC GI screen validated known EMC genetic interactions which confirmed our

confidence in the biological relevance of the screen. We next set out to dissect newly identified genetic interactions with EMC. The ~200 EMC synthetic genes fell in the following general pathway/functional categories: cholesterol synthesis, endosomal trafficking, membrane biogenesis, ERAD, translation, and uncharacterized. In order to elucidate the functional relationship of each synthetic gene with EMC and understand the transcriptional consequence of perturbing synthetic gene in EMC4 depleted cells, we performed a single-cell RNA seq screening experiment in EMC4 KO and Rescue cells.

Perturb-Seq in EMC KO mammalian cells defines functional clusters of genetic interactors

We used Perturb-seq, a platform for multiplexed profiling of perturbations with single-cell resolution, and used it to systematically dissect the interaction between the mammalian EMC and the synthetic pathways identified in the genome-wide screen. Briefly, we designed a single-guide sgRNA library targeting all (~200) strongly synthetic genes with EMC and cloned it into pJR85 Perturb-Seq vector including 5 non-targeting control guides (Joseph et al). Our sequence-verified library was then packaged into lentiviruses transfection, and lentiviral preparations was pooled for transduction into K562 dCas9-KRAB EMC4 KO and K562 dCas9-KRAB EMC4 Rescue cells. Three days post infection, transduced cells were sorted to near purity. Seven days post infection, cells were separated into droplet emulsions using the 10x Genomics Chromium controller and 10x sequencing library was prepared according to the 10x protocol (Fig 3.3A, 10x protocol reference). We used Cell Ranger 3.0 software (10x Genomics) for alignment of scRNA-seq reads, collapsing reads to unique molecular identifier (UMI) counts, cell calling, and depth normalization of transcriptome library. We sequenced it using the standard format for scRNA-seq from 10x Genomics (28 bp Read 1, 98 bp Read 2, and 8 bp Index Read 1) on a NovaSeq 6000 System (Illumina) with NovaSeq 6000 S4 Reagent kits (Illumina). Index reads were aligned to expected sequences using bowtie for GBC Perturb-seq.

To explore these data, we first constructed synthetic bulk expression (metacell) profiles by averaging normalized expression across cells containing each sgRNA. Heatmap displays

correlations between hierarchically clustered average bulk expression profiles for each of ~200 perturbations in the EMC4 KO background (Fig 3.3B, C). This hierarchical clustering revealed four distinct fictional clusters driven by EMC4 KO background. Each cluster is defined by a group of genes whose knockdown leads to a similar transcriptional response in cells. The brighter a cluster is the strongly these genes are correlated to each other based on gene expression. First, to confirm the validity of the transcriptional signatures in each cluster in the EMC KO experiment, we superimposed the name of the genes making up the clusters onto the EMC Rescue dataset. They showed no self-correlation suggesting that this unique transcriptional signature is specific to EMC KO phenotype (Fig 3.4). Second, we wanted to characterize the effect of control non-targeting guides in this experiment – we expect them to cluster with genes whose perturbation does not change the transcriptional profile of EMC KO cells. Reassuringly, (1) the non-targeting guides clustered tightly together in and (2) they belonged to the most strongly correlated cluster of genes defined by membrane protein biogenesis and ERAD (Fig 3.3). One possibility explaining this is that EMC KO cells have a high level of protein misfolding and cellular stress and perturbing other genes exacerbating protein folding and inhibiting the ability to recycle misfolded proteins drives the cell towards a similar but more exacerbated transcriptional response.

We next proceeded to characterize each cluster. For each cluster, we are showing the group of genes making up that cluster (visualized in StringDB) and the transcriptional response defining that cluster. We will start with the most highly correlated cluster, Cluster 2, we named “Protein Homeostasis”. Firstly, as mentioned earlier, Cluster 2 is defined by a few pathways consistent with protein folding, insertion, and degradation (Fig 3.6). Aside from canonical TA insertion pathway cytosolic genes (GET4, BAG6, SGTA, UBL4A), the cluster contains recently identified membrane protein biogenesis genes physically associating with Sec61 channel - TMCO1, CCDC47, and NOMO-NCLN-TMEM147 complex, a relationship that was not fully captured in StringDB association networks (ref; Fig 3.6). Another large group making up this cluster are E3 ubiquitin ligase genes and ERAD component genes ensuring proper handling of

misfolded proteins. The gene expression signature defining this cluster is defined by upregulation of stress response genes (Fig 3.5). The unifying thesis for the functionality of this cluster is that EMC4 KO cells are stressed at baseline level so perturbing any other parallel pathways or quality control pathways stress those KO cells even more. This result motivates follow up work that we are doing on parallel pathways for membrane protein biogenesis, range of substrates for each pathway, and redundancy between these pathways.

The next most highly correlated cluster is Cluster 3, we named “Cholesterol Synthesis” cluster. There were a few groups of genes defining this cluster: 1) in the middle of them all were the master transcriptional cholesterol regulators (SREBF1, SREBF2, MBTPS2, and the sensor SCAP); 2) the ER bound TA insertion machinery components (WRB, ASNA1, and CAMLG) together with peroxisomal protein import genes (PEX5 and PEX13); and 3) lobe A of the conserved oligomeric Golgi (COG) complex (COG2, COG3, COG4 including COG8 belonging to lobe B and connecting lobe A to lobe B) together with the RINT1-NBAS-ZW10 (NRZ) complex (Fig 3.7). The defining feature of this cluster is upregulation of immune response genes (Fig 3.5). Let us break down each group of genes and their potential resulting effect on EMC knockout cells. The transcriptional regulators in this group were one of the most strongly synthetic genes from the GI screen. This is most likely because EMC knockout cells have a lowered capacity for cholesterol production due to depletion of EMC client proteins such as: SQS, an enzyme important for cholesterol synthesis; TMEM97, a protein at the lysosome regulating cholesterol trafficking; and SOAT1, which uses long chain fatty acids to modify surplus free cholesterol for inert storage in lipid droplets (Pol et al., 2004). Any further abrogation of the cholesterol synthesis pathway causes further cholesterol depletion and synthetic lethality. The second strongly related genes in this group, according to StringDB are the TA protein insertion and peroxisomal import machinery genes. Interestingly, the TA insertion canonical pathway genes were split between this cluster and Cluster 2 – ER components in this cluster versus cytosolic components in Cluster 2. This may be because there are compensatory chaperones to bring TA proteins to the ER from

the cytosol but limited machinery at the ER for insertion. SQS is mainly an EMC client but could also be inserted less efficiently through the TRC40 pathway (Guna et al., 2018). It has been shown that functional peroxisomes are essential for efficient cholesterol sensing and synthesis, which could explain the functional grouping of PEX5 and PEX13 in this cluster (Charles et al., 2020). The last group, made up of COG genes, the NRZ complex, and COPE is crucial for intracellular trafficking (Willett et al., 2013; Raote et al., 2018). The COG complex, both physically and functionally, interacts with all classes of molecules maintaining intra-Golgi trafficking, namely SNAREs, and mediates intra-Golgi vesicle tethering. The evolutionarily conserved NRZ tether is a multi-subunit tether complex (MTC) that assembles at the surface of the ER and is required for retrograde capture of membranes, partially localizes to ER exit sites and interacts with SNAREs mediating Golgi-to-ER vesicle tethering (Aoki et al., 2009; Arasaki et al., 2006; Hirose et al., 2004; Liu et al., 2020; Schroter et al., 2016; Ren et al., 2009). These genes may be in the “Cholesterol Synthesis” functional cluster because they secure an alternative source of cholesterol in an EMC knockout background. As mentioned earlier, the “Cholesterol Synthesis” cluster is defined by upregulation of immune response genes. Pro-inflammatory molecules like IL32 have been found to induce cholesterol accumulation through inhibition of cholesterol efflux, thus, increasing the cholesterol capacity of cholesterol depleted EMC4 KO cells (Xu et al., 2017). Furthermore, cholesterol biosynthesis genes in cluster 3 are down regulated, possibly due to lack of master cholesterol regulators (Fig 3.8). The characterization of this cluster underscores the many ways the cell can make cholesterol and the importance of the EMC as a key player in cholesterol homeostasis.

The next cluster of interest, albeit not as strongly correlated as Clusters 2 and 3, is Cluster 4 we named “Cholesterol Uptake” (Fig 3.9). This cluster is characterized by upregulation of cholesterol and lipid metabolism genes (Fig 3.5). The groups of genes making up this cluster are: 1) cholesterol import and synthesis enzymes (LDLR – crucial role for cholesterol uptake, HMGCR, ACLY); 2) COMMD/CCDC22/CCDC93 (CCC) complex genes crucial for membrane trafficking

(COMMD2, COMMD9, COMMD9, COMMD10, CCDC22, VPS29, C16orf62, ZNF207, SPSB3); 3) V-type ATPase genes crucial for endosomal acidification (ATP6AP1, ATP6V0A, ATP6V0C, ATP6V0D1, ATP6V0E1, ATP6V0G1). Aside from *de novo* cholesterol biosynthesis, uptake of cholesterol via LDL Receptor is another way to maintain cholesterol homeostasis. This happens as follows: extracellular LDL is captured by **LDLR** on the cell surface and the LDL–LDLR complex is internalized via endocytosis. LDLR dissociates from LDL and is recycled by the **CCC (COMMD–CCDC22–CCDC93)** protein complex back to the surface. LDL is further delivered towards **lysosomes**, and the carried cholesteryl esters are hydrolyzed to cholesterol. The genes in this cluster, that are also strongly synthetic with EMC, all belong to this pathway for internalization of extracellular LDL its processing into cholesterol. When this process is perturbed in EMC4 knockout cells, it drives cholesterol depletion further. Not surprisingly, the cell tries to compensate by upregulating genes for *de novo* cholesterol synthesis (Fig 3.5). Yet another cluster from this Perturb-Seq experiment underscores the key role of EMC in cholesterol homeostasis validating our approach to understand the physiological role of the EMC in the cell beyond simply its molecular role.

The last cluster we will address is Cluster 4 and we will call “Translation” (Fig 3.10). It includes RNA binding proteins associated with splicing and translational regulation (HNRNPM, PQBP1, DNAJC8, ILF3, IGF2BP1, CSDE1); a translational initiation group (EIF4A1, EIF3H, RPS8, DDX3X), and a group of DNA damage genes (part of the CCC complex together w/ RBBP5, DPY30, and LEO1). This cluster is characterized by upregulation of ribosomal protein genes (Fig 3.5). One possible explanation is that in EMC KO cells, there is an increase in misfolded proteins which induces the UPR and leads to decrease in total protein translation. As a results, there is an overabundance of RPLs and RPSs that are not being translated. Given the presence of multiple small subclusters in Cluster 1, there are alternative hypotheses that could be tested.

Validation of cholesterol defect in EMC KO cells

Given the 1) strong synthetic interaction of EMC with cholesterol homeostasis genes from the screen 2) the functional validation of cholesterol synthesis and cholesterol uptake clusters as the most synthetically lethal with EMC and 3) reports that EMC deficiency limits the cellular boundaries defining cholesterol tolerance, reflected by diminished viability with limiting or excessive extracellular cholesterol, we wanted to further investigate the role EMC in cholesterol homeostasis (Volkmar et al., 2019). Our hypothesis was that EMC may be acting to maintain cholesterol homeostasis through sets of genes 1) whose protein products are EMC client proteins 2) that are epistatic to EMC and 3) that are post-translationally regulated. We took advantage of the different sets of high-throughput data we have generated and asked the question what those genes might be. Using the synthetic score for each gene from the GI screen, the transcriptional information from the Perturb Seq data, and the protein level expression data from a Mass-Spectrometry experiment all in EMC4 KO and EMC4 Rescue cells, we identified a set of genes that have epistatic relationship with EMC and are highly post-translationally downregulated in EMC4 KO cells meaning they are not transcriptionally inhibited but rather they get degraded after protein has been made consistent with misfolding of these proteins in the absence of the chaperone EMC (Fig 3.11). There was a group of proteins that are strongly depleted post-translationally and are epistatic with EMC. We have grouped them in a few major categories: cholesterol synthesis (FDFT1 (SQS), EBP, MSMO1), cholesterol storage (SOAT1, PLIN2), cholesterol trafficking (TMEM97, SOAT1), and fatty acid-related (ANO6, ELOVL6), and others well-characterized EMC clients (BCAP29, BCAP31, TRPV2). Clearly, EMC clients are crucial players in maintaining cholesterol homeostasis on every level – *de novo* synthesis, uptake and trafficking, and storage.

We next quantified the cholesterol defect in EMC4 KO cells vs EMC4 Rescue cells. Knockout cells had ~2-fold less total cholesterol and ~6-fold less stored cholesterol (Fig 3.11). This quantitative phenotype was validated in a microscopy experiment where in EMC KO cells

we clearly see the accumulation of cholesterol 1) at the plasma membrane and 2) intracellularly due to perturbed trafficking (Fig 3.12).

Conclusion

The research described in Chapter III is proof for the utility of using high-throughput and unbiased genetic approaches to answer fundamental physiological questions about protein homeostasis. Starting with a phenotype-first approach, we identified a set of ~200 genes that were synthetic interactors with the EMC. We then took these genes whose depletion in EMC KO background made cells sick and characterized the synthetic phenotype on transcriptional level using Perturb Seq. We identified four distinct functional clusters of genes of which the most synthetic were associated with cholesterol synthesis and cholesterol uptake. In fact, the most highly correlated cluster of genes was the Protein Homeostasis one and was relatively less synthetic than the cholesterol clusters (Fig 3.3C). This underscores the fact that EMC KO cells are not that sick when we perturb protein homeostasis further possibly due to compensatory mechanisms for maintaining protein balance in the cell. However, if we further perturb cholesterol balance, cells cannot cope, due to the central role of cholesterol in so many cellular processes. Here, we present the first complete view of the EMC cellular roadmap, an unbiased genetic-based approach to characterize the role of this conserved complex in mammalian cells. In two large scale experiments, we identified relationships of the EMC with other cellular pathways that have been studied in the past decade using single-track biochemical approaches. The EMC roadmap opens a wealth of exciting follow-up opportunities for the fields of membrane protein homeostasis and cholesterol homeostasis.

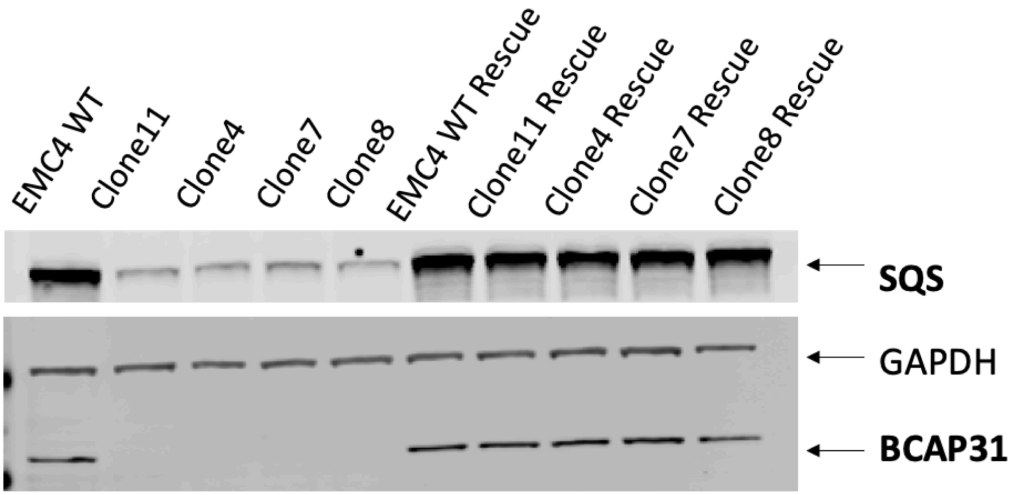


Figure 3.1. Functional knockout and rescue of EMC function in K562 dCas9 KRAB cell line. Four independent genetic EMC4 knockout clones were generated using CRISPR-Cas9. In all four EMC knockouts, clients BCAP31 and SQS are severely depleted, and upon lentiviral reintroduction of EMC4, functional defect is rescued. GAPDH – loading control.

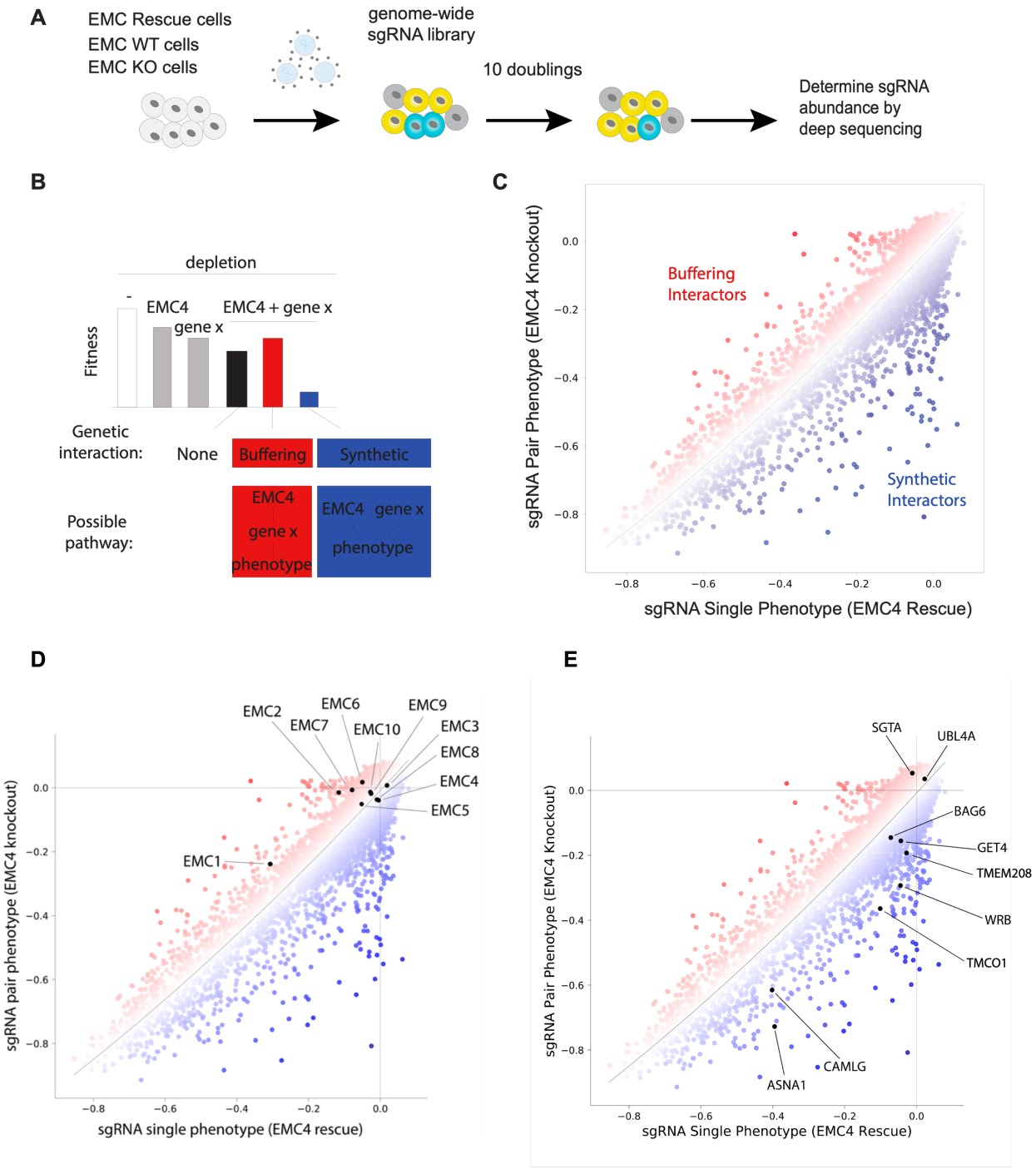


Figure 3.2. Genetic interaction screen identifies genes synthetic with EMC. A. Flowchart of a genetic interactions (GI) screen. B. GI screens differentiate between synthetic, buffering, or no interactions. C. Results from EMC4 genetic interactions screen. D. EMC complex subunits are slightly buffering. E. Parallel membrane protein biogenesis pathways are synthetic with EMC.

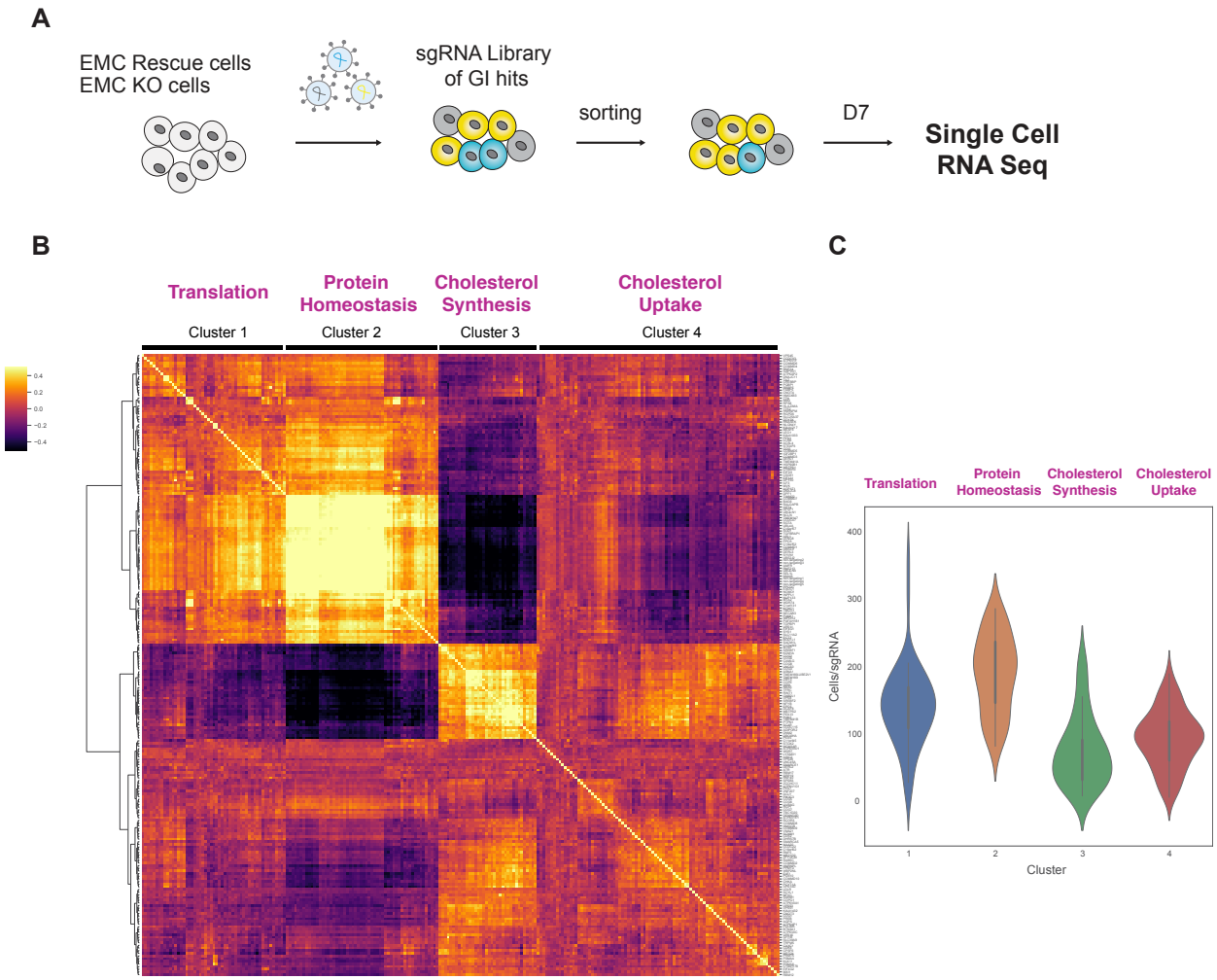


Figure 3.3. Perturb-Seq identifies functional mammalian EMC clusters. A. Flowchart of Perturb Seq experiment with a GI sublibrary in EMC4 KO or EMC4 Rescue cells. B. Heatmap displays correlations between hierarchically clustered average bulk expression profiles for each of ~200 perturbations in the EMC4 KO background. C. Clusters have different level of synthetic lethality with EMC4.

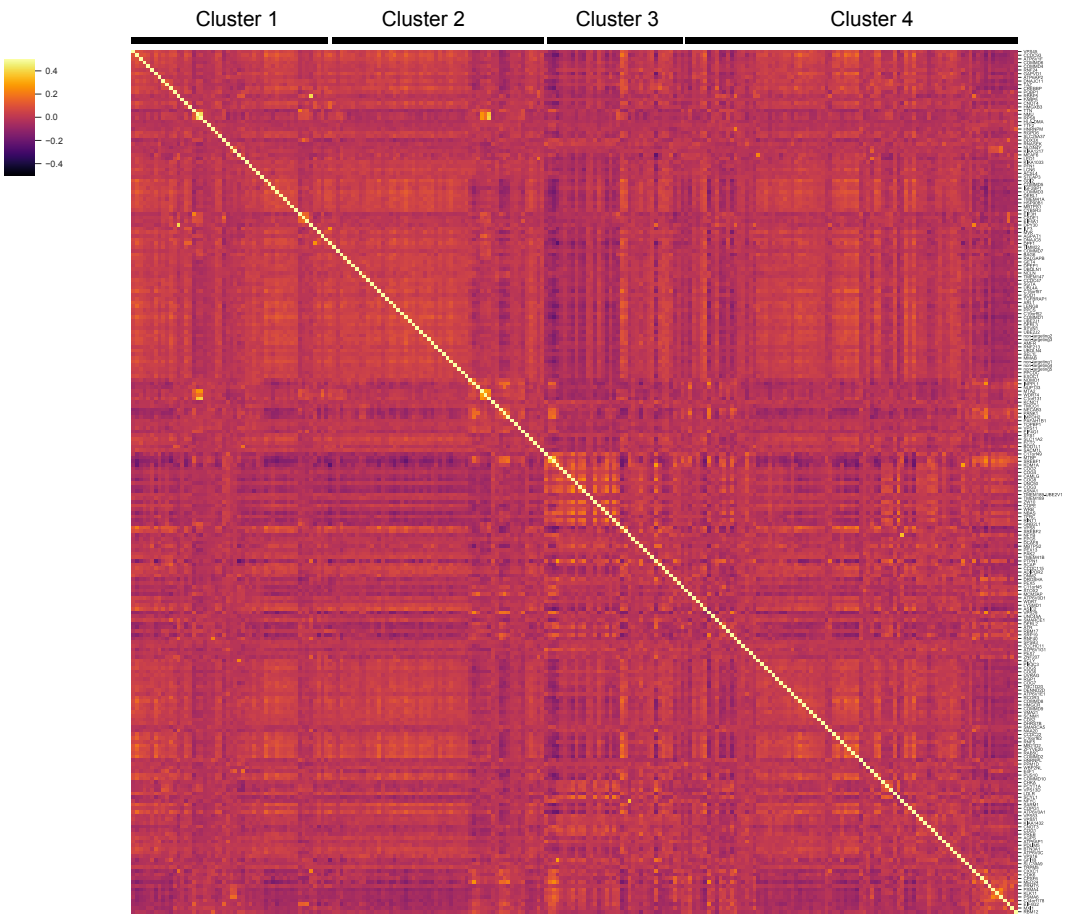


Figure 3.4. KO clusters superimposed on the rescue dataset show poor correlation.

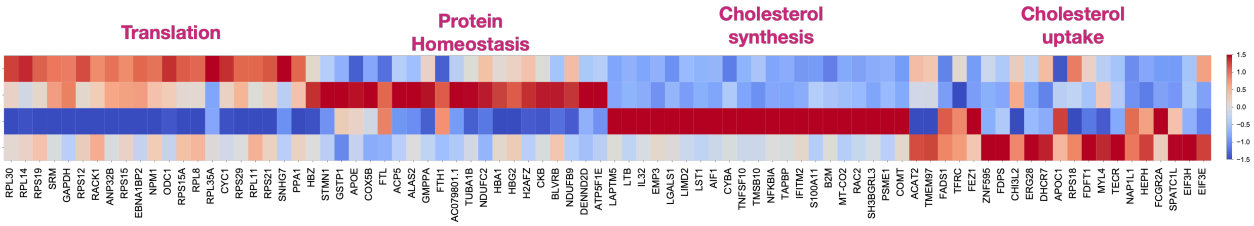


Figure 3.5. Transcriptional Signatures Inform Cluster Identity

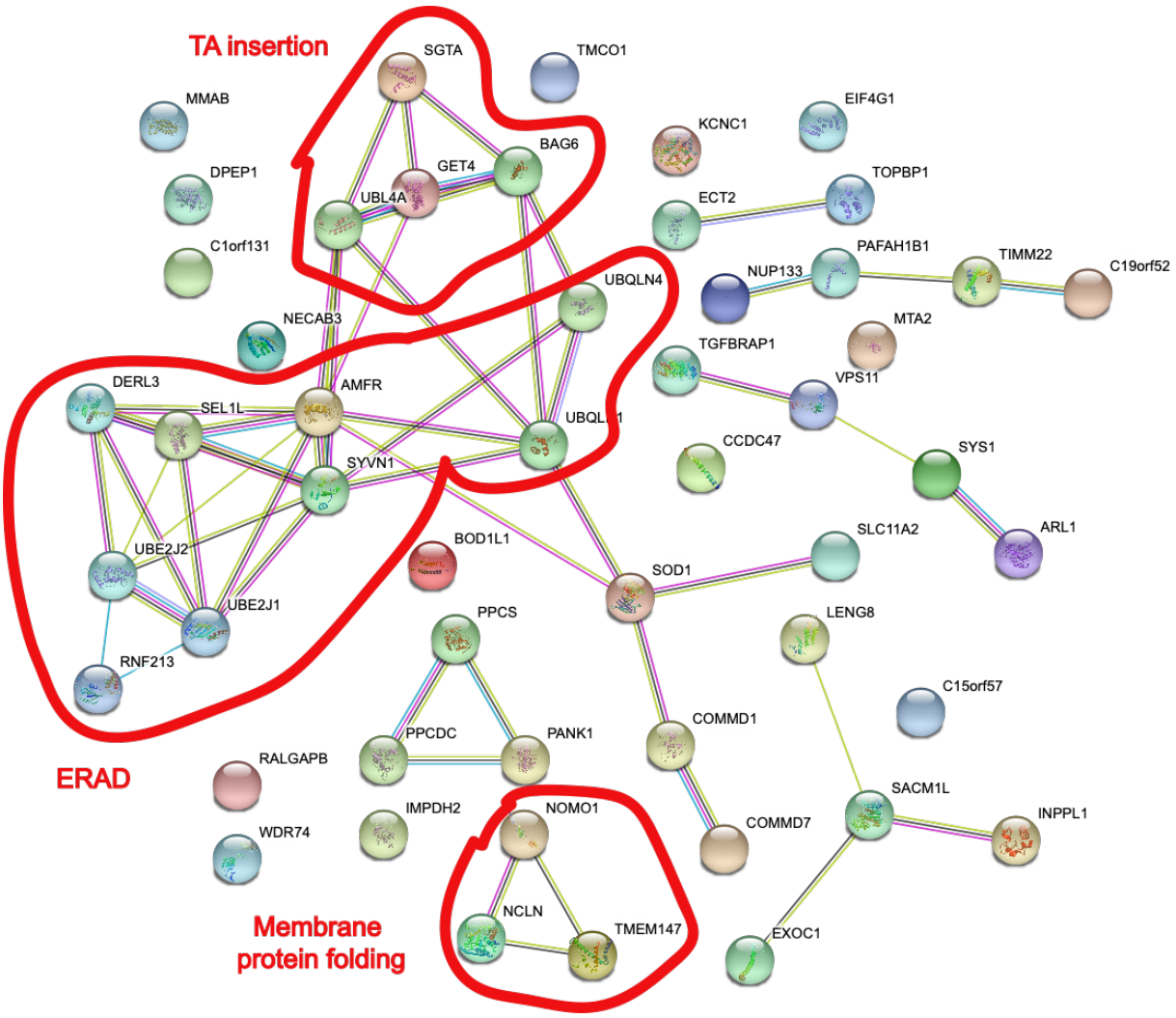


Figure 3.6. StringDB visualization of Cluster 1.

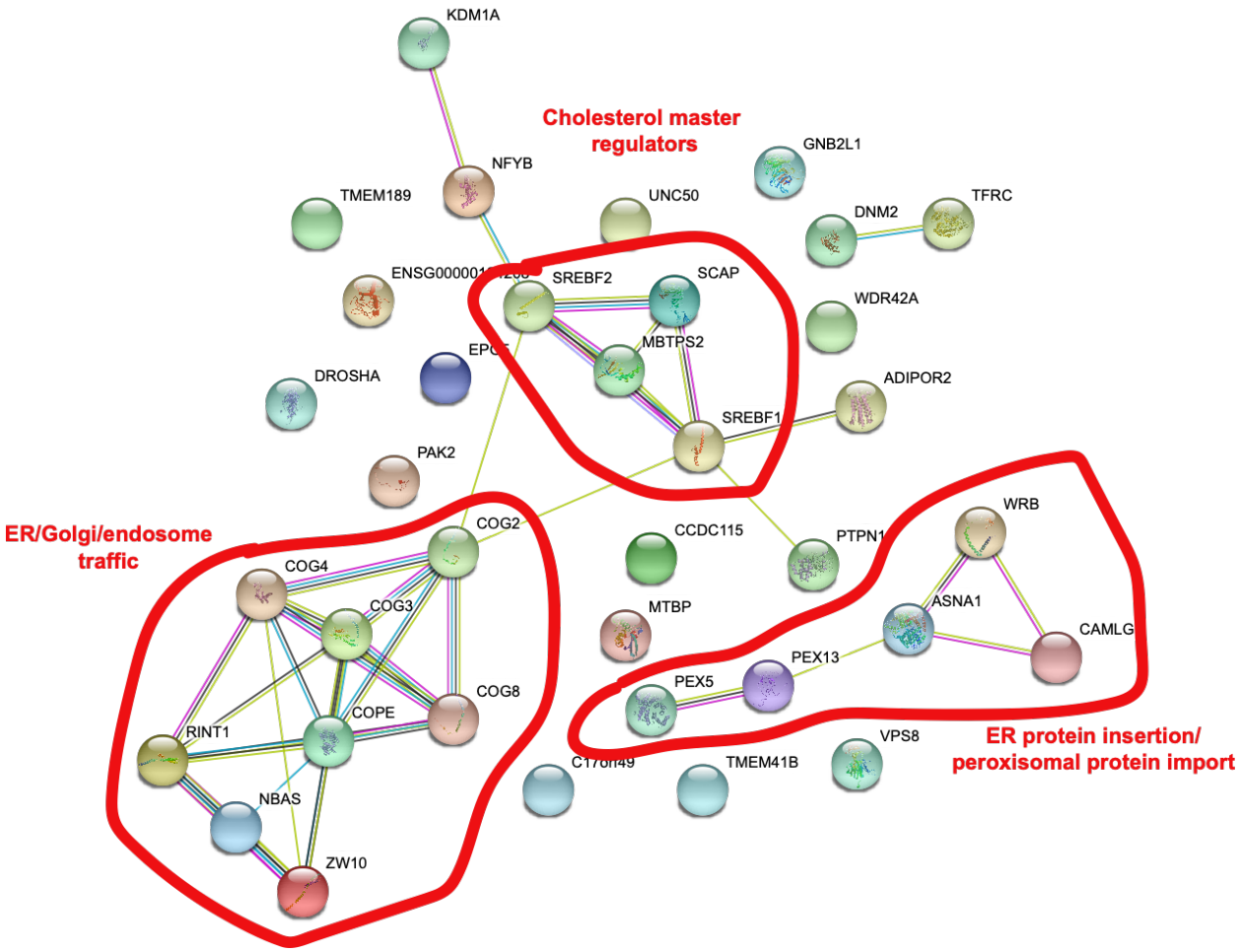


Figure 3.7. StringDB visualization of Cluster 2.

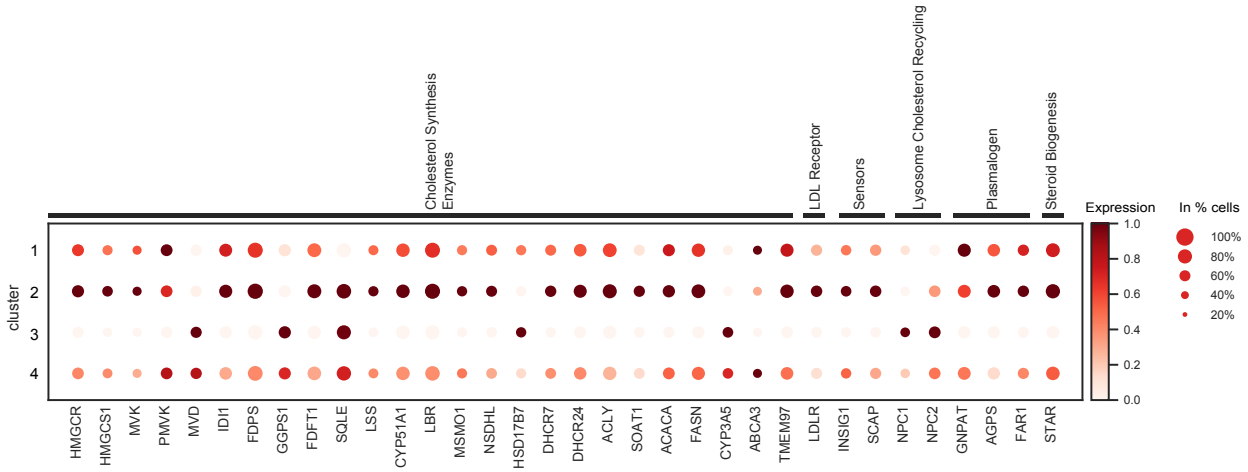


Figure 3.8. Gene expression of cholesterol-related genes in Cluster 2 in comparison with rest of clusters.

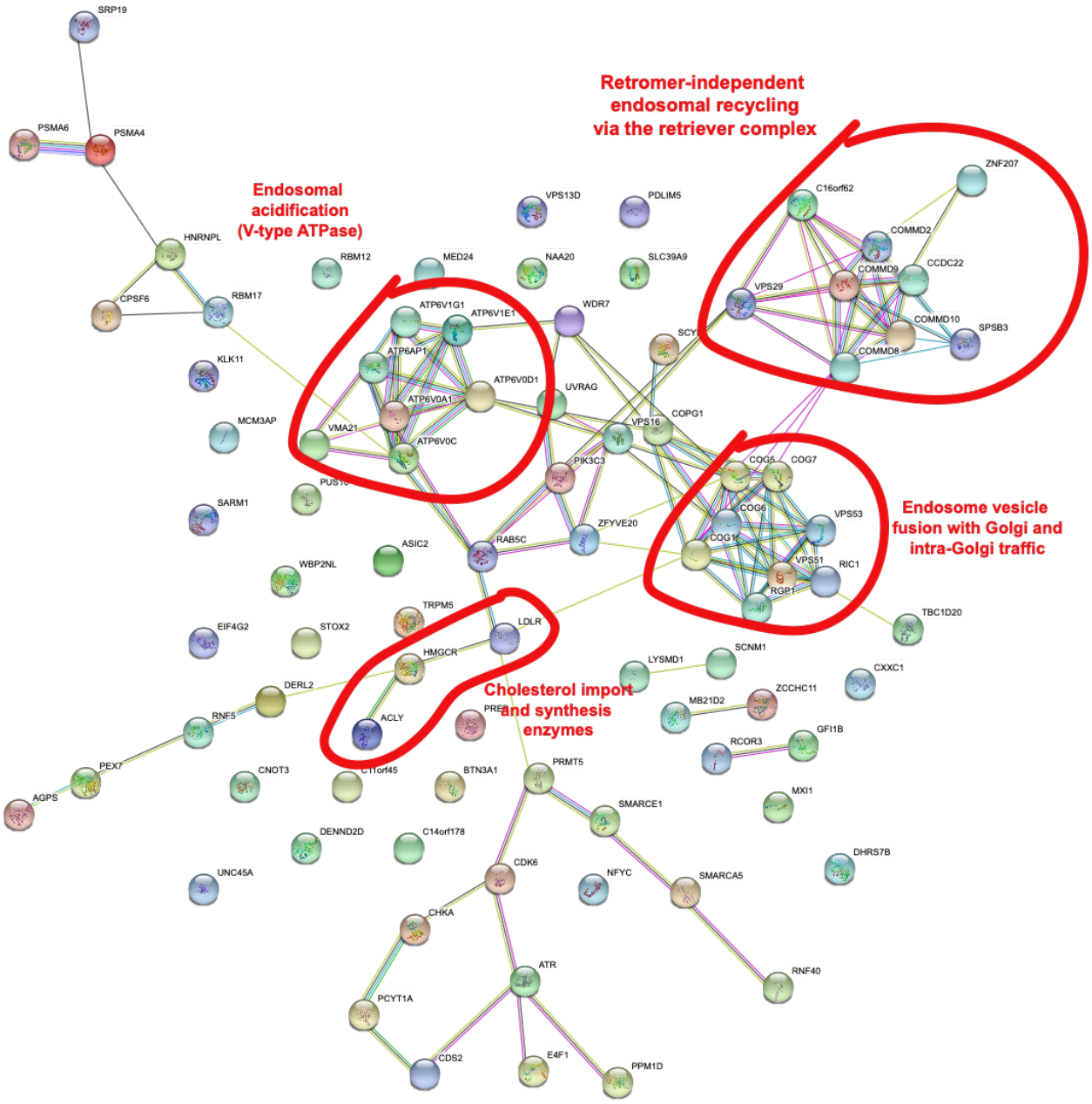


Figure 3.9. StringDB visualization of Cluster 3.

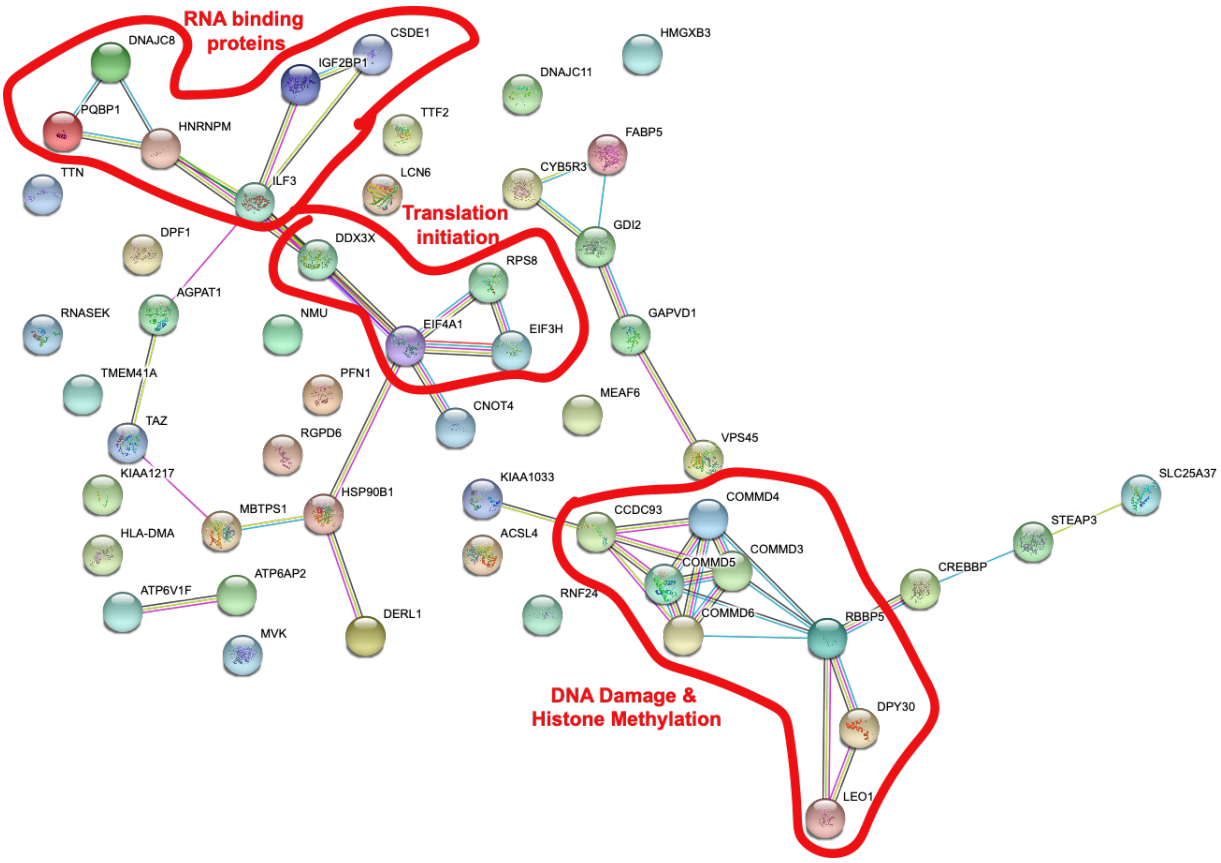


Figure 3.10. StringDB visualization of Cluster 4.

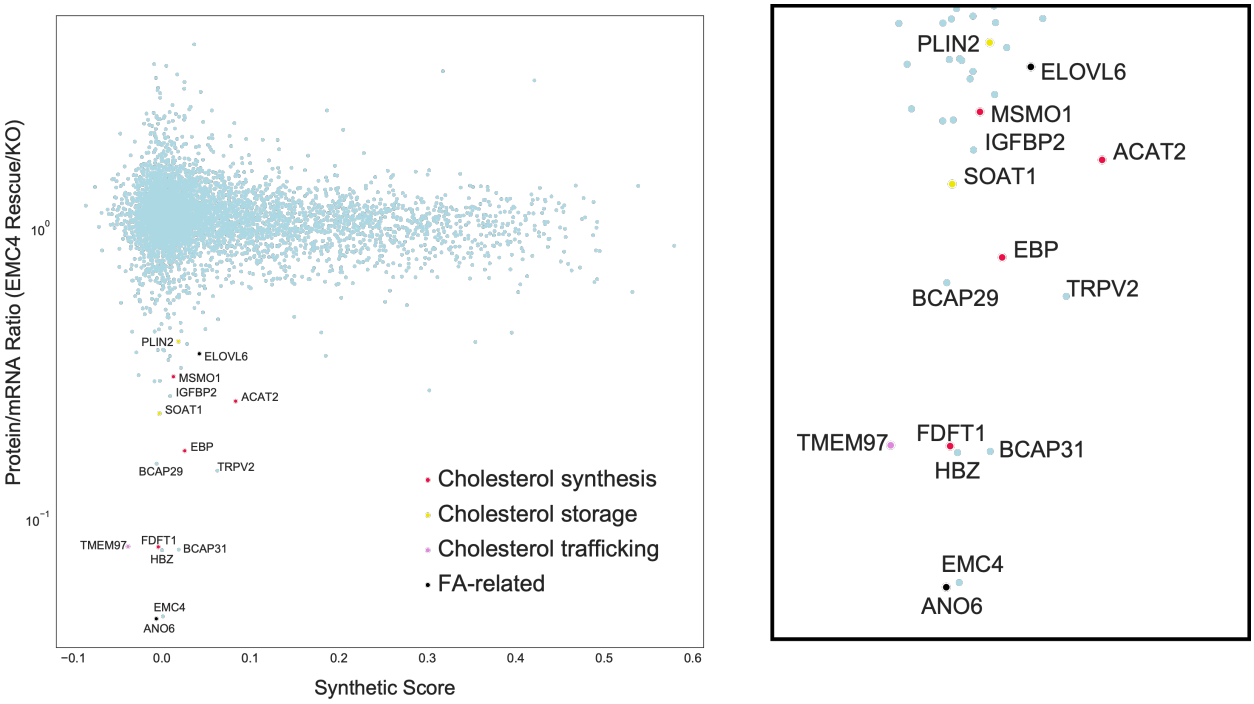


Figure 3.11. Synthetic score from GI screen plotted against protein/mRNA ratio of EMC4 Rescue/EMC KO.

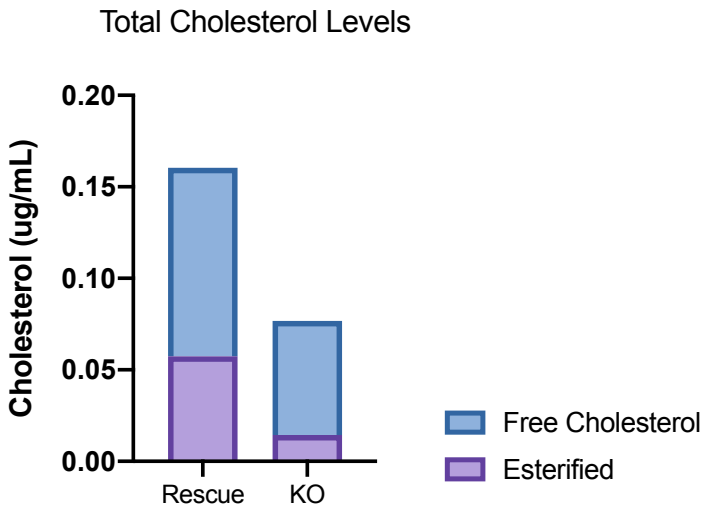


Figure 3.12. Colorimetric-based quantification of total cholesterol and cholesteryl esters.

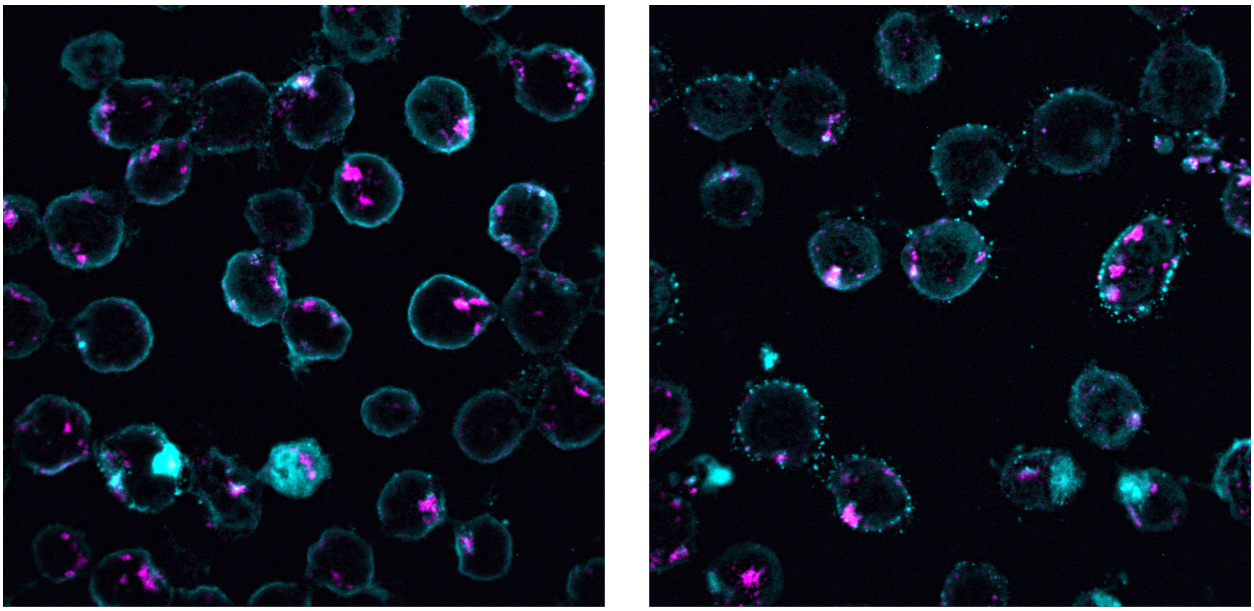


Figure 3.13. EMC4 KO and ECM4 Rescue cells stained for cholesterol distribution. Cyan: Filipin (binds free cholesterol). Magenta: LAMP1 (lysosomal protein)

REFERENCES

Adamson, B., Norman, T. M., Jost, M., Cho, M. Y., Nuñez, J. K., Chen, Y., Villalta, J. E., Gilbert, L. A., Horlbeck, M. A., Hein, M. Y., Pak, R. A., Gray, A. N., Gross, C. A., Dixit, A., Parnas, O., Regev, A., & Weissman, J. S. (2016). A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell*, *167*(7), 1867-1882.e21. https://doi.org/10.1016/j.cell.2016.11.048

Aoki, T., Ichimura, S., Itoh, A., Kuramoto, M., Shinkawa, T., Isobe, T., & Tagaya, M. (2009). Identification of the Neuroblastoma-amplified Gene Product as a Component of the Syntaxin 18 Complex Implicated in Golgi-to-Endoplasmic Reticulum Retrograde Transport. *Molecular Biology of the Cell*, *20*, 2639–2649. https://doi.org/10.1091/mbc.E08

Arasaki, K., Taniguchi, M., Tani, K., & Tagaya, M. (2006). RINT-1 Regulates the Localization and Entry of ZW10 to the Syntaxin 18 Complex. D. *Molecular Biology of the Cell*, *17*, 2780–2788. https://doi.org/10.1091/mbc.E05-10

Cecchetti, C., Strauss, J., Stohrer, C., Naylor, C., Pryor, E., Hobbs, J., Tanley, S., Goldman, A., & Byrne, B. (2021). A novel high-throughput screen for identifying lipids that stabilise membrane proteins in detergent-based solution. *PLoS ONE*, *16*(7 July). https://doi.org/10.1371/journal.pone.0254118

Charles, K. N., Shackelford, J. E., Faust, P. L., Fliesler, S. J., Stangl, H., & Kovacs, W. J. (2020). Functional Peroxisomes Are Essential for Efficient Cholesterol Sensing and Synthesis. *Frontiers in Cell and Developmental Biology*, *8*. https://doi.org/10.3389/fcell.2020.560266

Coukos, R., Yao, D., Sanchez, M. I., Strand, E. T., Olive, M. E., Udeshi, N. D., Weissman, J. S., Carr, S. A., Bassik, M. C., & Ting, A. Y. (2021). An engineered transcriptional reporter of protein localization identifies regulators of mitochondrial and ER membrane protein trafficking in high-throughput CRISPRi screens. https://doi.org/10.7554/eLife

Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C. P., Jerby-Arnon, L., Marjanovic, N. D., Dionne, D., Burks, T., Raychowdhury, R., Adamson, B., Norman, T. M., Lander, E. S., Weissman, J.

S., Friedman, N., & Regev, A. (2016). Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell*, *167*(7), 1853-1866.e17. https://doi.org/10.1016/j.cell.2016.11.038

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S., & Qi, L. S. (2013). XCRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, *154*(2), 442. https://doi.org/10.1016/j.cell.2013.06.044

Guna A, Volkmar N, Christianson JC, Hegde RS. 2018. The ER membrane protein complex is a transmembrane domain insertase. *Science* 359:470–473. DOI: https://doi.org/10.1126/science.aao3099, PMID: 29242231

Hirose, H., Arasaki, K., Dohmae, N., Takio, K., Hatsuzawa, K., Nagahama, M., Tani, K., Yamamoto, A., Tohyama, M., & Tagaya, M. (2004). Implication of ZW10 in membrane trafficking between the endoplasmic reticulum and Golgi. *EMBO Journal*, *23*(6), 1267–1278. https://doi.org/10.1038/sj.emboj.7600135

Horlbeck, M. A., Witkowsky, L. B., Guglielmi, B., Replogle, J. M., Gilbert, L. A., Villalta, J. E., Torigoe, S. E., Tjian, R., & Weissman, J. S. (2016). *Nucleosomes impede Cas9 access to DNA in vivo and in vitro*. https://doi.org/10.7554/eLife.12677.001

Jonikas MC, Collins SR, Denic V, Oh E, Quan EM, Schmid V, Weibezahn J, Schwappach B, Walter P, Weissman JS, Schuldiner M. 2009. Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. Science 323:1693–1697. DOI: https://doi.org/10.1126/science.1167983, PMID: 19325107

Kotov, V., Bartels, K., Veith, K., Josts, I., Subhramanyam, U. K. T., Günther, C., Labahn, J., Marlovits, T. C., Moraes, I., Tidow, H., Löw, C., & Garcia-Alai, M. M. (2019). High-throughput stability screening for detergent-solubilized membrane proteins. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-46686-8

Liu, L., Zhang, M., & Ge, L. (2020). Protein translocation into the ERGIC: an upstream event of secretory autophagy. *Autophagy*, *16*(7), 1358–1360. https://doi.org/10.1080/15548627.2020.1768668

McGilvray, P. T., Anghel, S. A., Sundaram, A., Zhong, F., Trnka, M. J., Fuller, J. R., Hu, H., Burlingame, A. L., & Keenan, R. J. (2020). An ER translocon for multi-pass membrane protein biogenesis. *ELife*, *9*, 1–43. https://doi.org/10.7554/ELIFE.56889

Pol, A., Martin, S., Fernández, M. A., Ingelmo-Torres, M., Ferguson, C., Enrich, C., & Parton, R. G. (2005). Cholesterol and Fatty Acids Regulate Dynamic Caveolin Trafficking through the Golgi Complex and between the Cell Surface and Lipid Bodies/ V. *Molecular Biology of the Cell*, *16*, 2091–2105. https://doi.org/10.1091/mbc.E04

Raote, I., Ortega-Bellido, M., nio Santos, A. J., Foresti, O., Zhang, C., Garcia-Parajo, M. F., Campelo, F., & Malhotra, V. (2018). *eLife*. TANGO1 builds a machine for collagen export by recruiting and spatially organizing COPII, tethers and membranes. https://doi.org/10.7554/eLife.32723.001

Ren, Y., Yip, C. K., Tripathi, A., Huie, D., Jeffrey, P. D., Walz, T., & Hughson, F. M. (2009). A Structure-Based Mechanism for Vesicle Capture by the Multisubunit Tethering Complex Dsl1. *Cell*, *139*(6), 1119–1129. https://doi.org/10.1016/j.cell.2009.11.002

Replogle, J. M., Norman, T. M., Xu, A., Hussmann, J. A., Chen, J., Cogan, J. Z., Meer, E. J., Terry, J. M., Riordan, D. P., Srinivas, N., Fiddes, I. T., Arthur, J. G., Alvarado, L. J., Pfeiffer, K. A., Mikkelsen, T. S., Weissman, J. S., & Adamson, B. (2020). Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nature Biotechnology*, *38*(8), 954–961. https://doi.org/10.1038/s41587-020-0470-y

Schröter S, Beckmann S and Schmitt HD (2016). Coat/Tether Interactions—Exception or Rule? *Front. Cell Dev. Biol*. 4:44. doi: 10.3389/fcell.2016.00044

Volkmar N, Thezenas ML, Louie SM, Juszkiewicz S, Nomura DK, Hegde RS, Kessler BM, Christianson JC. The ER membrane protein complex promotes biogenesis of sterol-

related enzymes maintaining cholesterol homeostasis. J Cell Sci. 2019 Jan 16;132(2):jcs223453. doi: 10.1242/jcs.223453. PMID: 30578317

Willett, R., Kudlyk, T., Pokrovskaya, I., Schönherr, R., Ungar, D., Duden, R., & Lupashin, V. (2013). COG complexes form spatial landmarks for distinct SNARE complexes. *Nature Communications*, *4*. https://doi.org/10.1038/ncomms2535

Xu, Z., Dong, A., Feng, Z., & Li, J. (2017). Interleukin-32 promotes lipid accumulation through inhibition of cholesterol efflux. *Experimental and Therapeutic Medicine*, *14*(2), 947–952. https://doi.org/10.3892/etm.2017.4596

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Katerina Popova
7371473E1C1F47A...
Author Signature

12/12/2021
Date