Structural Studies of Membrane Proteins in Organellar & Cellular Homeostasis

Yifei (Abby) Lin

Department of Biochemistry,
University of Oxford
Supervisor: Prof. Simon Newstead

Jesus College

A thesis submitted for the degree of
Doctor of Philosophy in Cellular Structural Biology

48,190 words
Date of submission: 11th December 2025

Abstract

Cells rely on organelle-specific membrane proteins and lipids to maintain ionic and pH homeostasis and to oversee the biogenesis/quality control of multi-pass proteins. This thesis examines two such systems: the lysosomal chloride/proton antiporter CLC7 and a putative endoplasmic reticulum (ER) chaperone, RUSF1.

In the first half of the thesis, through cryo-electron microscopy, I solved three structures of human CLC7: wild-type transporter complex bound to a lysosomal signalling lipid PI(3,5)P2, wild-type apo transporter complex, as well as a hyper-active mutant Y715C. Through structural analysis, I could confirm that the lipid $PI(3,5)P_2$ acts as a structural "glue" that stabilizes contacts between the transmembrane domain and cytosolic CBS domains of CLC7, thereby inhibiting transport; loss of this lipid (or disruption at the interface) promotes CBS disengagement. The structure of a hyperactivating, disease-linked mutant (Y715C) also captured previously unseen TMD conformations, including rearrangements of helices that correlate with CBS dissociation and expose elements of the Cl^-/H^+ exchange pathway, consistent with an activated state. Importantly, two titratable residues E313 and E314 were exposed, and these residues has been shown to be important in the activation of *E. coli* CLC.

Therefore, our structural results suggested a potential mechanism of activation for human CLC7, where dissociation of the signalling lipid PI(3,5)P2 lead to detachment of the CBS domains, causing structural rearrangement in the transmembrane domains as well, exposing important glutamate residues which could be protonated to activate transport.

In the second half of the thesis, a 3.13 Å cryo-EM structure establishes RUSF1 as a homodimer with an MFS-like architecture, as well as possessing a large hydrophilic cavity, and extensive cytosolic domain–micelle engagement. Mutational analysis identifies conserved residues that control oligomeric state; notably, K230G drives monomerization of human RUSF1. Co-immunoprecipitation shows RUSF1 engages the transmembrane domain of GABAa receptor α1 (GABRA1), albeit more weakly than canonical chaperones, and knockout/perturbation phenotypes support a broader role in ER homeostasis rather than a client-specific chaperone. A SpyTag/SpyCatcher proteomics workflow nominates potential binding partners (e.g., TMEM126A and 14-3-3ε). In addition, immunofluorescence data from colleagues of RUSF1 knock-down cells showed hole-like structures present in the ER, suggesting a potentially more general role of RUSF1 in ER stress pathways. These results provided foundation on which more functional characterisations of RUSF can be built upon, such as co-localisation studies of RUSF1 and various ER-stress related proteins,

Finally, to aid further research on both projects, I selected and characterised multiple synthetic nanobodies against both CLC7 and RUSF1. These nanobodies can be used in immunofluorescence of native cells to probe the localisation of CLC7 and RUSF1, and in co-immunoprecipitation assays to identify interaction partners, furthering the functional understanding of CLC7 and RUSF1.

Acknowledgement

This thesis would not have been possible without the support and guidance I received throughout my PhD. I am deeply grateful to Prof. Simon Newstead, my primary supervisor, for his mentorship and for the many opportunities to present my work at conferences and meetings. I also thank Dr Joanne Parker for her steady support and valuable research advice.

I'm indebted to my colleagues in the Newstead Lab: Dr Oliver Adams, Dr Gabriel Kuteyi, Dr Chitra Joshi, Dr Mark Lobel, Dr Daniel Frey, Sigurbjörn Markússon, Anuja Brotha, and many others, for their friendship and for making Oxford feel like home over the past four years.

I am especially grateful to Dr Takafumi Kato and Dr Justin Deme (Susan Lea Lab) for mentorship in cryo-EM sample preparation and data processing, and for their contributions to this work. I also thank Dr Jacob Hilton (Joseph Mindell Lab) for outstanding electrophysiology data, and Dr Martyna Biadun for superb immunofluorescence images included in this thesis.

I am thankful to Prof. Joseph Mindell for the opportunity to collaborate on the CLC7 project, and to Dr Ramanujan Hegde for collaboration on the RUSF1 project.

Finally, my heartfelt gratitude to Carlos Riechmann for accompanying me throughout this journey.

Contents

List of Abbreviations

ABC transporter	ATP-Binding Cassette transporter
ADP	Adenosine diphosphate
AF	AlphaFold
AMP	Adenosine monophosphate
AMPA	Alpha-amino-3-hydroxy-5-methyl-4-isooxazole-propionic acid
ATP	Adenosine triphosphate
BCA	Bicinchoninic Acid assay
BLI	Biolayer interferometry
BOS	Back of Sec61
BSA	Bovine serum albumin
CBS	Cystathionine beta-synthase
cDNA	Complementary DNA
CDR	Complementarity-determining regions
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
CHO	Chinese Hamster Ovary (cells)
CHS	Cholesteryl Hemisuccinate
CLC	CLC Family Transporters and Chloride Channel
CLIC	Chloride Intracellular Channel
CLIMP-63	Cytoskeleton-Linking Membrane Protein
CLN7	Neuronal Ceroid Lipofuscinosis type 7
CLPTM1L	Cleft lip and palate transmembrane protein 1-like
cmc	Critical micelle concentration
COS-7	African green monkey kidney cells

CryoEM	Cryogenic-electron Microscopy
CryoET	Cryogenic-Electron Tomography
CryoSPARC	Cryogenic-electron Microscopy Single Particle Ab-Initio Reconstruction and Classification
CTF	Contrast transfer function
cv	Column volume
DAPI	4′,6-diamidino-2-phenylindole
DDM	n-Dodecyl β-D-maltoside
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
DSF	Differential scanning fluorimetry
DTT	Dithiothreitol
DUF	Domain of Unknown Function
EDTA	Ethylenediaminetetraacetic acid
EER	Electron Event Representation
ELISA	Enzyme-linked immunosorbent assay
EMC	ER membrane protein complex
ER	Endoplasmic Reticulum
ERAD	ER-associated degradation
ERES	ER exit sites
ERQC	ER quality control compartment
FL	Full length
FRET	Fluorescence Resonance Energy Transfer

FSC	Gold-standard Fourier Shell Correlation
FSEC	Fluorescence Size Exclusion Chromatography
GABA	Gamma aminobutyric acid
GABRA1	GABAa Receptor Subunit Alpha 1
GEL	GET- and EMC-like
GFP	Green Fluorescence Protein
HA	Hemagglutinin
HEK	Human embryonic kidney (cells)
HeLa	Cervical cancer cells of Henrietta Lacks
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horseradish peroxidase
IMAC	Immobilized metal affinity chromatography
IP	Immunoprecipitations
IPTM	Interface predicted template modelling
KD	Knock Down
Kd	Dissociation Constant
KO	Knock out
LB	Lysogeny broth
LC-MS	Liquid chromatography–mass spectrometry
LMNG	Lauryl Maltose Neopentyl Glycol
mAU	Milli absorbance unit
MES	2-Morpholinoethanesulfonic acid
MFS	Major Facilitator Superfamily
NMR	Nuclear Magnetic Resonance
OSTM1	Osteopetrosis-associated transmembrane protein 1

PAE	Predicted Aligned Error
PAT	Protein associated with translocon
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PI	Phosphatidylinositol
PI(2,5)P2	Phosphatidylinositol 2,5-bisphosphate
PI(3,5)P2	Phosphatidylinositol 3,5-bisphosphate
PI(4,5)P2	Phosphatidylinositol 4,5-bisphosphate
PI3P	Phosphatidylinositol 3-phosphate
PIKfyve	Phosphoinositide kinase
pLGIC	Pentameric ligand-gated ion channel
PSM	Peptide-spectrum match
PTM	Predicted template modelling
PVDF	Polyvinylidene fluoride
RELION	REgularized LIkelihood OptimizatioN
RMSD	Root Mean Square Deviation
RNA	Ribonucleic acid
RTN	Reticulon
RUS	Root UV-B Sensitive
RUSF1	Root UVB Sensitive Family 1
SA	Streptavidin
SDS-PAGE	Sodium Dodecyl Sulfate–PolyAcrylamide Gel Electrophoresis
SEM	Representing Standard Error of the Mean
SERCA	SARCO/endoplasmic reticulum Ca2+-ATPase
SLC	Solute Carrier Family

TA	Tail-Anchored
TBS	Tris-Buffered Saline
TEV	Tobacco Etch Virus protease
TG	Thapsigargin
TLR	Toll-like receptor
TM	Transmembrane helix
TMD	Transmembrane domain
TPC2	Two-pore channel 2
TRPML1	Transient Receptor Potential Mucolipin-1
TST	Twin strep tag
UPR	Unfolded protein response
UVB	Ultraviolet B radiation
WT	Wild type
XAR	Extreme acid resistance

1. Introduction

1.1 Membrane Proteins & Their Roles in Organellar Homeostasis

Eukaryotic cells are distinguished by their complex internal compartmentalization, wherein intracellular membranes demarcate specialized organelles—such as the nucleus, endoplasmic reticulum (ER), Golgi apparatus, endo-lysosomal system, mitochondria, and peroxisomes. These organelles occupy nearly half of the total cell volume and harbour unique ionic compositions, pH levels, protein make-up, and lipid compositions that facilitate their specific functions, such as protein synthesis and processing (the ER), trafficking (Golgi), degradation (endo-lysosomes), and ATP synthesis (mitochondria) (Alberts *et al.*, 2002).

The maintenance of each organelle's distinct identity and functionality is achieved, in part, through dedicated membrane proteins—pumps, channels, and transporters for ions and metabolites; receptors and enzymes for signalling and biosynthesis; and trafficking/tethering machines that preserve compartment identity and connectivity (Alberts *et al.*, 2002; Saftig *et al.,* 2009). For example, the acidification of endosomes and lysosomes depend on the vacuole type ATPase, facilitated by a host of ion channels and transporters such as TRPML1 and CLC7 for electrical shunt; nutrient sensors such as SLC38A9 that link amino-acid levels to mTORC1 signalling, positioning the lysosome as a metabolic hub (Collins & Forgac *et al*., 2020; Leray *et al*., 2022; Samie *et al.*, 2013; Wang *et al*., 2015; Lamming & Bar-Peled *et al.*, 2018). On the endoplasmic reticulum membrane, the DP1/Yop1p (and reticulon) family shapes tubular ER structures by “wedging” into the outer leaflet of the membrane and stabilizing curvature, maintaining ER architecture and connectivity (Voeltz *et al*., 2006; Shibata *et al*., 2008; Shibata *et al.*, 2010). Within the endoplasmic reticulum itself, the ER membrane protein complex (EMC) exemplifies a specialized assembly of transmembrane subunits that facilitate folding,

insertion, and quality control of multi-pass membrane proteins, crucial for proteostasis in the ER.

Dysregulation of these membrane proteins can lead to various human diseases. Disrupted ion flux through lysosomal channels and transporters such as CLC7, TRPML1, or TMEM175 underlies multiple lysosomal storage and neurodegenerative conditions (Zifarelli *et al.*, 2022; Wang *et al.,* 2014, Medina et al., 2015; Qu *et al.*, 2022). Similarly, mutations in essential ER resident proteins such as the Sec61 translocon, the ER membrane complex (EMC complex), and Sec61 translocon can cause a broad range of disease, including neurodevelopmental delay, obesity, immune deficiency, and cancers (Lang *et al.*, 2017; Zhu *et al.,* 2024), though the underlying molecular mechanism still remains largely elusive, due to their involvement in the biosynthesis of diverse groups of proteins.

A key challenge in this field is understanding the precise mechanisms through which these membrane proteins are trafficked to the correct cellular compartment and regulated within specialized organelles. What ensures they are synthesized correctly, shielded and inhibited until reaching the right environment? What factors fine-tune their activity and maintain ion, pH, and protein homeostasis in various organelles?

To advance understanding in this area, this work investigates two organellar proteins: the lysosomal chloride-proton antiporter CLC7 and a putative ER-localized chaperone RUSF1.

1.2 CLC7 & Lysosomal pH Homeostasis

1.2.1 Membrane-bound Channels & Transporters Facilitates Lysosomal pH Homeostasis

Lysosomes are essential organelles in the cellular degradation and recycling system, responsible for breaking down macromolecules and participating in nutrient sensing and signalling. Their degradative function depends critically on maintaining an acidic luminal pH of around 4.5 to 5.0, which is vital for optimal activity of hydrolytic enzymes (Freeman *et al*., 2023).

In addition to their degradative functions, lysosomes also contribute to cellular processes such as membrane repair, antimicrobial defence, autophagy, antigen presentation, and cell migration (Freeman *et al*., 2023; Steinberg *et al*., 2007). These functions are spatially and functionally coordinated through the bidirectional transport of lysosomes along microtubules by motor proteins such as dyneins and kinesins. Perinuclear lysosomes, which are typically more acidic, predominantly support catabolic degradation and metabolite recycling, whereas lysosomes located at the cell periphery tend to be less acidic and are involved in processes like antigen presentation and plasma membrane repair. The positioning of lysosomes in the cell is also dependent on pH. Experimental dissipation of the lysosomal proton gradient causes perinuclear lysosomes to relocalise to the periphery, while activation of kinesin-based transport pathways similarly increases lysosomal alkalinity.

Moreover, lysosomal pH influences membrane dynamics; acidic conditions are required for the remodelling of lysosomal membranes into highly curved structures such as vesicles and tubules (Freeman *et al*., 2023; Steinberg *et al*., 2007). These remodelling events are essential for proper cargo sorting by coat proteins and membrane-deforming complexes. Alkalinization of the

lysosomal lumen rapidly halts budding and vesiculation, resulting in defective cargo trafficking and mistargeting of ligands.

Thus, the pH of lysosomal lumen is tightly and intricately controlled, enabling functionality across diverse cellular contexts (Figure 1 A). The establishment and maintenance of this pH gradient is orchestrated by a cohort of membrane proteins, including proton pumps, ion channels, and transporters.

Among the most well-characterized components is the vacuolar-type H^+-ATPase (V-ATPase), a multi-subunit machinery that utilises the hydrolysis of ATP to actively pump protons into the lysosomal lumen (Figure 1 B). However, proton transport alone is not sufficient to sustain acidification, as the accumulation of positive charge would eventually inhibit further proton import. Therefore, effective lysosomal acidification by V-ATPases requires counterion movement to neutralize the transmembrane voltage generated during proton pumping. This may involve anion influx, cation efflux, or both (Mindell *et al*., 2012).

Several cation-selective lysosomal ion channels, including TMEM175, TRPMLs, and two-pore channels (TPCs), are candidates for providing a cation efflux to sustain lysosomal acidification, through transport of calcium, sodium, or protons themselves (Figure 1 B, Wang *et al*., 2014; Medina *et al*., 2015; Wu *et al*., 2023). Meanwhile, chloride has been shown to be the most abundant anion involved in many cellular processes, and multiple vesicular transporters and channels, such as CLC7 and CLN7, have been implicated in the regulation of lysosomal pH (Figure 1 B, Mindell *et al*., 2012; Stauber *et al*., 2013; Graves *et al*., 2008; Jentsch & Pusch, 2018; Wang *et al*., 2021).

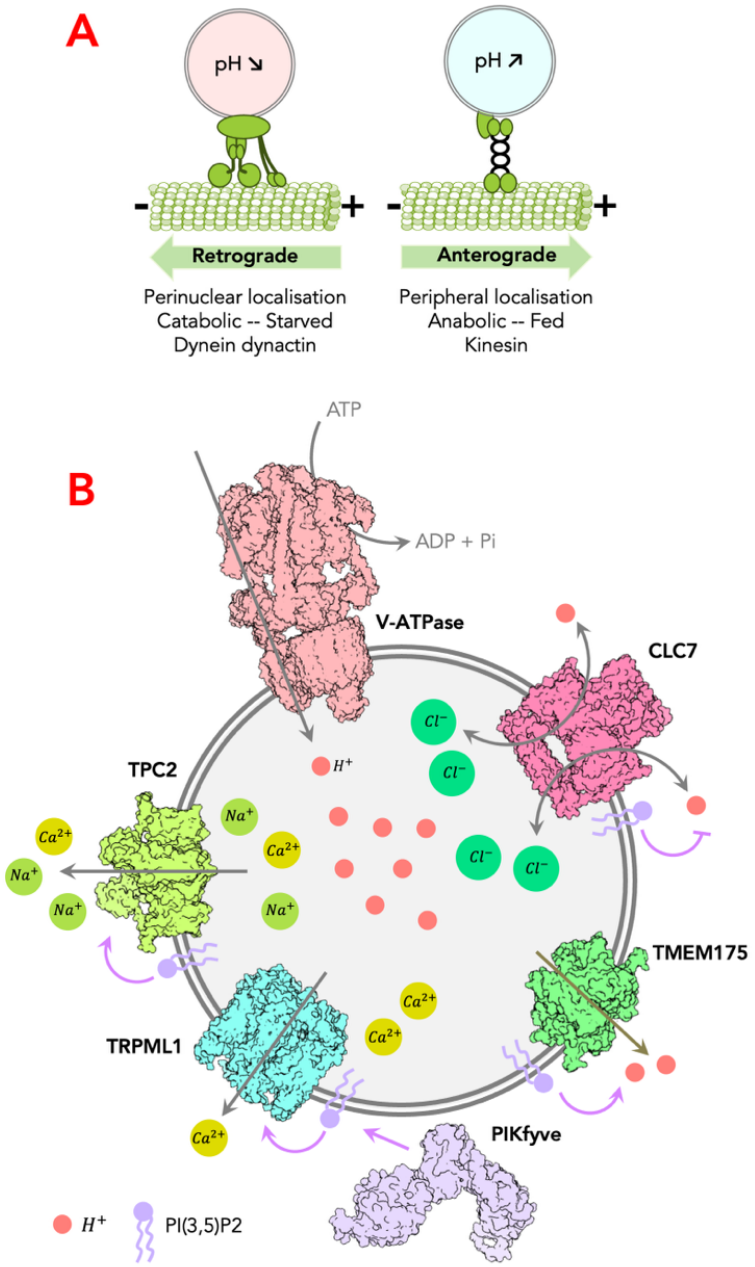


Figure 1. Lysosomal Channels and Transporters involved in pH and Ion Homeostasis

A. Cartoon illustration showing how varying vesicular pH affects the localisation and specialised function of lysosomes. B. Schematic diagram of the process of acidifying lysosomes by V-ATPase (PDB: 6WM2), facilitated by several well-characterised ion channels and transporters: sodium/calcium channel TPC2 (PDB: 6NQ1), calcium channel TRPML1 (PDB: 6E7P), proton channel TMEM175 (PDB: 6WC9), and chloride/proton antiporter CLC7 (PDB: 7JM7); the effect of PI(3,5)P2, synthesized by PIKfyve (PDB: 7K2V), was also indicated.

The maintenance of lysosomal pH homeostasis is a complex process, which also involves endo-lysosomal lipids, which function to regulate the ion channels and v-type ATPase pumps. Lysosomal function is tightly regulated by the internal luminal pH, and the internal

concentration of protons is critical for lysosomes to function in different cellular contexts and respond to cellular signals. More research is required to dissect in more detail regarding each player in this process, and the first part of my thesis will focus on understanding how the lipid PI(3,5)P2 regulates CLC7 from a structural perspective.

1.2.2 The Role of Phosphoinositides in Vesicular pH and Ion Homeostasis

In addition to various membrane proteins, phosphoinositides have also been shown to participate in the maintenance of lysosomal pH homeostasis, both through defining membrane identity, controlling membrane-associated processes (Posor *et al*., 2022; Lolicato *et al*., 2024), and through regulating the functions of multiple lysosomal channels and transporters (Figure 1 B, Hammond *et al*., 2020; Leray *et al*., 2022; She *et al*., 2019; Gan *et al*. 2022; Dong *et al*., 2010).

Chemically, phosphoinositides are phosphorylated derivatives of phosphatidylinositol (Figure 2 A, PtdIns, also called PI), a glycerophospholipid whose head group constitutes a myo-inositol ring (Balla *et al*., 2013). The myo-inositol adopts a chair conformation (Figure 2 B, likened to a turtle) in which five of the six hydroxyl groups are equatorial (flippers and tail) and the 2-position hydroxyl is axial (head); one hydroxyl (the D1 position) is esterified to diacylglycerol through a phosphodiester linkage to form the PtdIns backbone, while the remaining hydroxyls are available for reversible phosphorylation (Balla *et al*., 2013). In practice, only the 3, 4 and 5 positions of the inositol ring are phosphorylated in cells, producing a set of 7 distinct phosphoinositide species: PI3P, PI4P, PI5P, $PI(4,5)P_2$, $PI(3,4)P_2$, $PI(3,5)P_2$ and $PI(3,4,5)P_3$ (Balla *et al*., 2013). The synthesis and turnover of each species is governed by a complex and highly regulated repertoire of kinases and phosphatases (Figure 2 C, D).

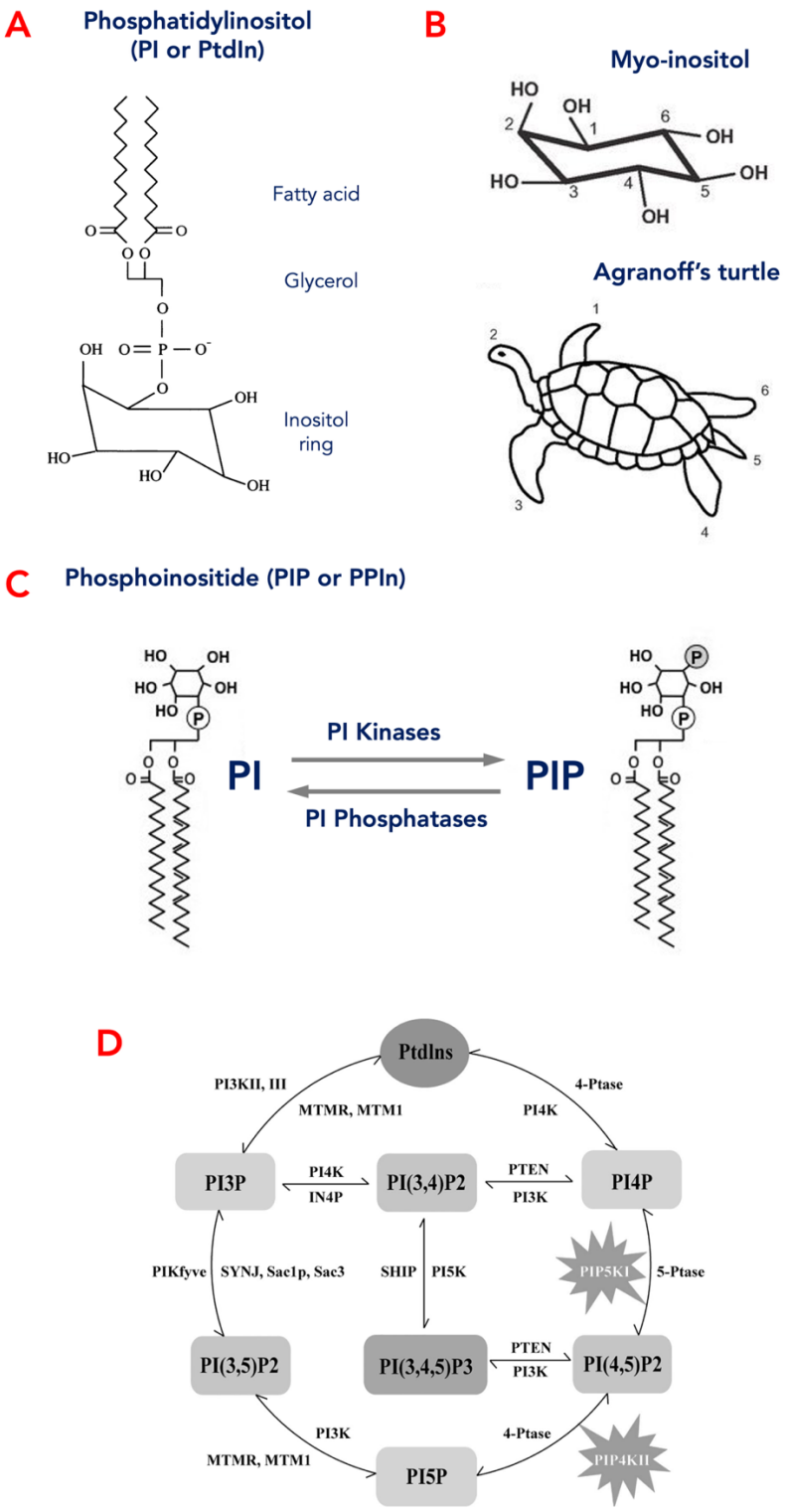


Figure 2. Chemical Structure of Phosphoinositides

A. Chemical structure of phosphatidylinositol. B. Chemical structure of the myo-inositol ring and the simplified analogy of Agranoff's turtle, showing the numbering of carbon atoms and hydroxyl group along the ring; other than hydroxyl group on carbon 2 (head of the turtle) being at axial position, the rest of the hydroxyl groups are at equatorial position (Agranoff BW., 1978). C. Simplified diagram of inter-conversion between phosphatidylinositol and phosphoinositides by various kinases and phosphatases. D. Kinases and phosphatases involved in the inter-conversion of phosphoinositides. Adapted from Jin *et al.*, 2023.

Although they constitute only a small fraction of total membrane lipids (< 1% of total phospholipids), phosphoinositides function as dynamic regulating factors rather than passive structural components (Balla *et al*., 2013). Because different organelles and membrane subdomains are enriched in different phosphoinositide species, these lipids serve as molecular signatures that help define organelle identity and spatially organize membrane traffic (Figure 3, Paolo *et al*., 2006; Posor *et al*., 2022). For example, the plasma membrane is characteristically enriched in PI4P and $PI(4,5)P_2$, the Golgi is enriched in PI4P, early endosomes are defined by PI3P, and late endosomes/lysosomes contain $PI(3,5)P_2$ (and mixed pools of other PIPs, Balla *et al*., 2013). These distinctive lipid fingerprints are read by effector proteins that bind specific phosphorylated inositol headgroups, thereby coupling lipid composition to downstream protein recruitment and activity.

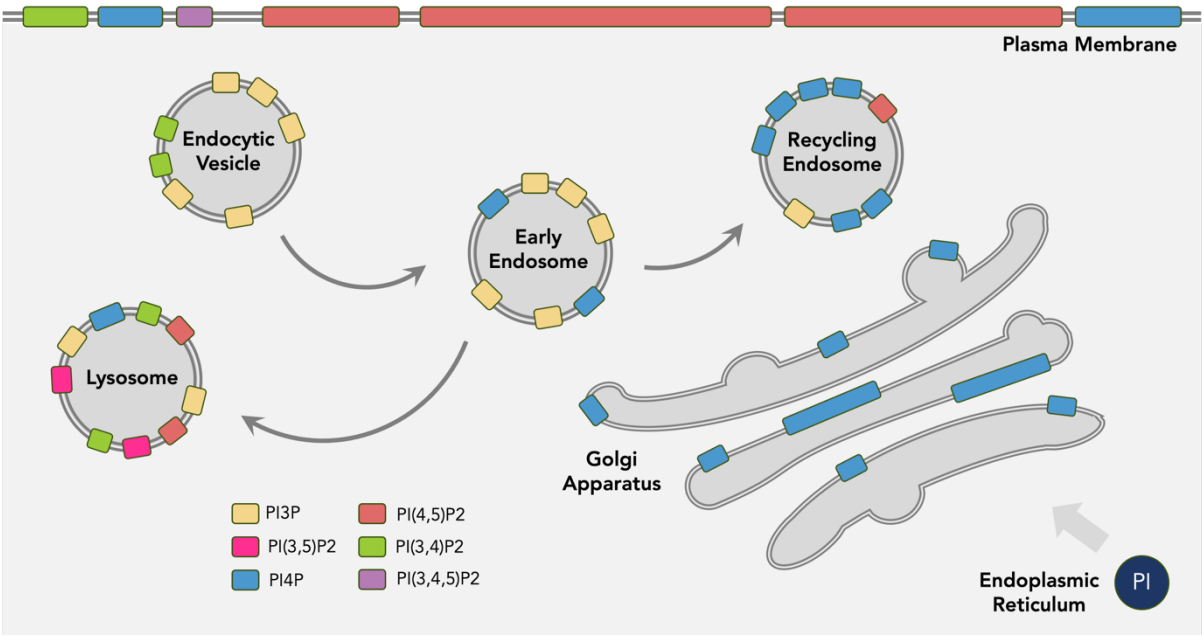


Figure 3. Phosphoinositides Function as Organelle Signatures

Schematic diagram showing the distribution of 6 phosphoinositide species across multiple membrane compartments in the cells: endoplasmic reticulum, Golgi apparatus, plasma membrane, endocytic vesicles, early endosomes, recycling endosomes, and lysosomes. The parent lipid PI is synthesised in the ER, then distributed throughout the endo-membrane systems and PM, and processed into other PIP lipids through kinases and phosphatases.

Phosphoinositides also have important roles directly influencing the activity of multiple ion channels and transporters, and provide an additional layer of regulation over the ion channels, transporters and exchangers responsible for vesicular ion homeostasis. Endo-lysosomal phosphoinositides — most notably $PI(3,5)P_2$, an endo-lysosome-enriched, low-abundance species — regulates the function of multiple lysosomal channels and transporters that fine tune luminal ion concentrations and acidification (Leray *et al*., 2022). Although steady-state levels of $PI(3,5)P_2$ are extremely low (~0.05-0.1% of total phosphatidylinositol pool), acute perturbation of its synthesis (for example by pharmacological inhibition of the kinase PIKfyve, Figure 2 D) produces dramatic defects in lysosomal acidification, indicating that mis-regulation of PI(3,5)P2 can have large physiological consequences.

A number of examples illustrate how specific phosphoinositide species directly regulate ion channels and transporters. $PI(4,5)P_2$ at the plasma membrane can activate inward-rectifier potassium channels such as Kir2.2 (Figure 4 A-C, Hansen *et al*., 2012); structural studies showed that $PI(4,5)P_2$ binding draws the channel's soluble domain closer to the transmembrane domain via a 6 Å movement (Figure 4 A), and electrophysiology experiments demonstrate increased current upon PIP_2 addition that is reversed by lipid washout (Figure 4 B).

Potential $PI(4,5)P_2$ binding sites have also been discovered in Solute carrier spinster homolog 2 (SPNS2), a lysolipid transporter (Tang *et al*., 2023). Mutagenesis of the proposed lipid pocket identified through molecular dynamics led to reduced transport activity (Tang *et al*., 2023).

Two-pore channel 2 (TPC2), a sodium-selective endo-lysosomal channel, is robustly and specifically activated by another PIP lipid, $PI(3,5)P_2$ (Figure 4 D-F); cryo-EM structures of $PI(3,5)P_2$-bound TPC2 reveal a widening of the ion selectivity filter (Figure 4 D, She *et al*., 2019).

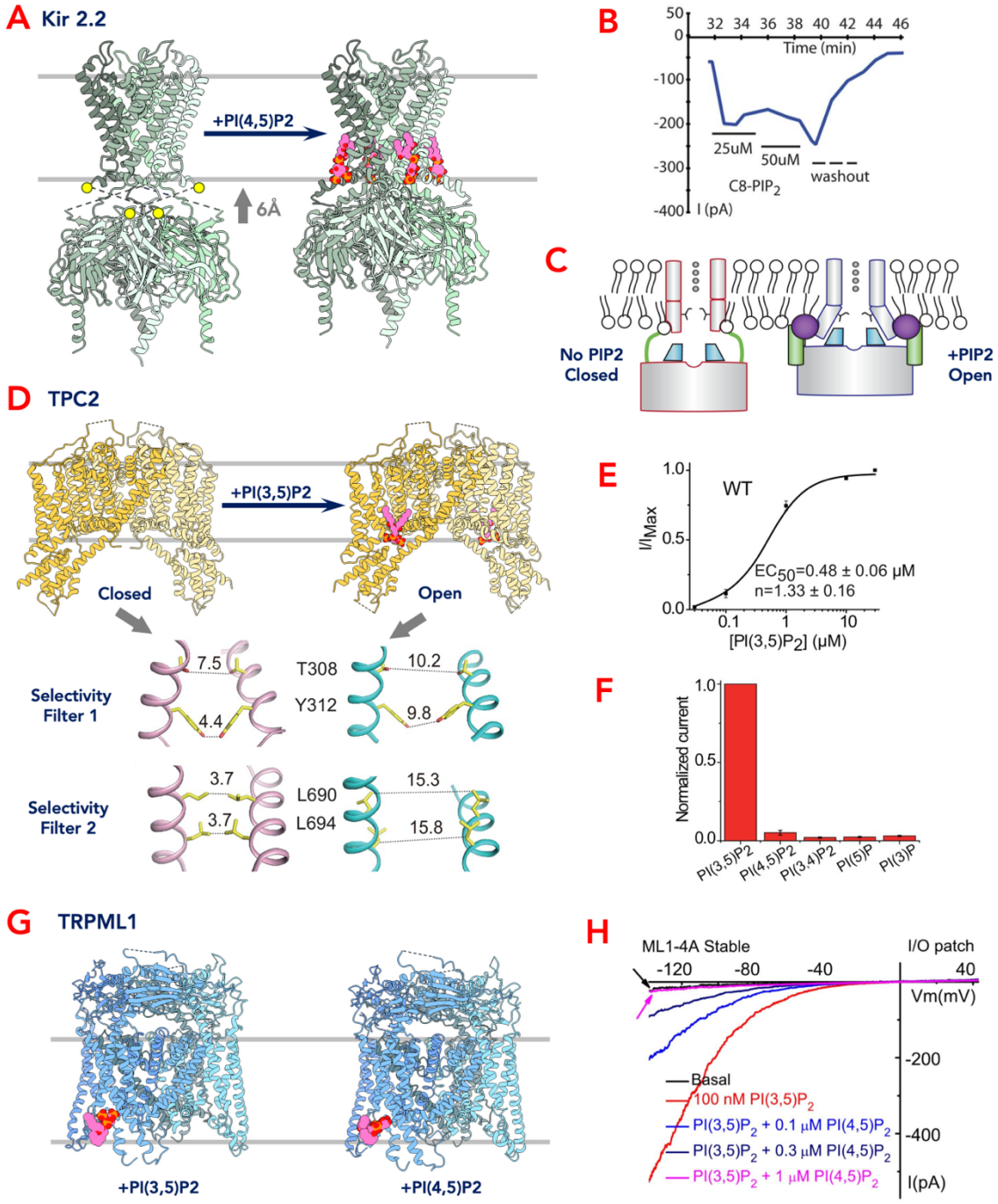


Figure 4. Phosphoinositides Regulate Activities of Membrane Channels

A. Structural data of PI(4,5)P2-dependent conformational changes in Kir 2.2 potassium channel (from left to right PDB 3JYC, 3SPY, Hansen *et al.*, 2012). B. Addition of PIP2 stimulated Kir 2.2 currents in *Xenopus* oocytes, which was diminished when PIP2 was washed out. C. Schematic diagram adapted from Hansen *et al.*, 2012, depicting the mechanism of lipid-linked activation of Kir 2.2 channel, through the binding of soluble domain and structural rearrangement in the transmembrane domain that widened the selectivity filter. D. Structural data of PI(3,5)P2-dependent activation of TPC2 channel, and widening of selectivity filter, adapted from She *et al.*, 2019 (from left to right PDB 6NQ1, 6NQ0). E,F. PI(3,5)P2-dependent activation of TPC2 and ligand specificity probed by whole-cell patch clamp electrophysiology, adapted from She *et al.*, 2019. G. Structural data of calcium channel TRPML1 bound to PI(3,5)P2 and PI(4,5)P2 (from left to right PDB 6E7P, 6E7Y, Fine *et al.*, 2018). H. TRPML1 current from whole-cell electrophysiology showed that PI(3,5)P2 addition led to strong signal, which was abolished by the addition of PI(4,5)P2.

TRPML1 is also directly activated by $PI(3,5)P_2$ with high specificity and potency (Figure 4 G, H, Dong *et al*., 2010; Zhang *et al*., 2012; Fine *et al*., 2018; Gan *et al*., 2022), but interestingly, currents recorded in electrophysiology are completely abolished by the addition of 1 µM $PI(4,5)P_2$ (Figure 4 H, Zhang *et al*., 2012). CryoEM structures showed that while $PI(4,5)P_2$ and $PI(3,5)P_2$ bound overlapping pockets in TRPML1, $PI(4,5)P_2$ preferentially stabilizes closed-state residues and thereby allosterically inhibits the channel (Figure 4 G, Fine *et al*., 2018).

Finally, the chloride/proton antiporter CLC7 is reported to be tonically inhibited by $PI(3,5)P_2$ (Leray *et al*., 2022), a regulatory relationship that has implications for lysosomal acidification and is discussed further in **Chapter 1.2.8**.

1.2.3 Chloride and Chloride Channels/Transporters

Chloride ions (Cl^-) are the most abundant anions present both extracellularly and intracellularly within animal cells. Chloride transport across the plasma membrane fulfils diverse physiological roles, including systemic electrolyte homeostasis, modulation of cellular excitability, and cell volume control (Stauber *et al*., 2013). These varied roles are facilitated by numerous chloride channels and transporters, which differ widely in structure, localization, and activity (Stauber *et al*., 2013).

However, the significance of chloride extends beyond these roles; Chloride ions play essential roles in regulating vesicular pH as an electrical shunt. Mathematical modelling shows that Cl^- pathways (CLC7 or other Cl^- channels) uniquely sustain robust, stable acidification compared with purely cation-based shunts (Ishida *et al.,* 2013). Mechanistically, prioritizing Cl^- avoids coupling acidification to cation pools that double as second messengers (e.g., Ca^{2+} for lysosomal release events; Na^+/K^+ shaping excitability and transport), thereby minimizing crosstalk with signalling networks (Mindell, 2012).

Across various cellular compartments within the secretory pathway, a noticeable gradient in luminal pH exists. As cargo progresses, vesicular pH decreases—from ~7.2 in the endoplasmic reticulum (ER) to ~6.4–6.7 in the Golgi and ~5.2–5.4 in secretory granules (Stauber *et al*., 2013). Similarly, endocytic compartments acidify progressively, with early endosomes at pH ~6.0 and lysosomes below pH 5.0. This acidification is accompanied by a rise in luminal Cl^- concentration ($[Cl^-]_{lumen}$), supporting its role as a counterion to protons. Following endocytosis, $[Cl^-]_{lumen}$ initially drops from ~140 mM to ~20 mM, likely due to repulsion by vesicle surface charges. It then rises to ~40 mM in early endosomes, ~50 mM in the Golgi, ~60 mM in late endosomes, and is estimated to exceed 80 mM in lysosomes (Stauber *et al*., 2013).

There have been many chloride channel and transporter families discovered throughout the years. These include the ligand gated chloride channels, such as the GABAA receptors, Volume-Regulated Anion Channels (VRACs), voltage dependent anion channels (VDACs), Chloride-Bicarbonate Exchangers from SLC family (SLC26), Chloride Intracellular Channel (CLIC) family, TMEM16A, CFTR, CLN7, and CLC family, among others (Kim *et al*., 2022, Wang et al., 2021).

Anion channels and transporters employ distinct strategies compared to cation channels to achieve ion selectivity. In many cation channels (such as potassium channels), selectivity arises from the precise coordination of dehydrated cations by backbone carbonyl oxygens in a narrow pore. This architecture effectively mimics the ion's hydration shell, allowing one cation (e.g., K^+) to permeate rapidly while excluding others of slightly different radius (e.g., Na^+) despite similar charge (Doyle *et al*., 1998; Zhou *et al*., 2001).

By contrast, anion channels are often less discriminating among monovalent anions of similar size, since the pool of available anions are less diverse in the cell, usually dominated by chloride ions (10-40 mM in the cytosol, even higher in endo-lysosomal vesicles, Lüscher *et al.,* 2020), followed by bicarbonates, phosphates, lactates, pyruvates and citrates (each less than 10 mM

in the cell. He *et al.,* 2021; Greiner *et al.,* 2021; Wu *et al.,* 2021; San Martín *et al.,* 2014; Zhao *et al.,* 2020). Therefore, anion channels rely more on electrostatics and hydrogen bonding than on size-exclusion, pores are typically lined with basic or polar residues (lysine, arginine, serine, threonine) that provide a positively charged environment stabilizing Cl^- or HCO_3^- passage. (Linsdell, 2001; Fahlke, 2001; Dutzler *et al*., 2002). For instance, studies of CFTR and CLC channels show that Cl^-, Br^-, and NO_3^- frequently permeate with comparable ease (Fahlke, 2001; Dutzler *et al*., 2002), while cation channels can exclude ions differing by only 0.2 Å in radius (Doyle *et al*., 1998; Zhou *et al*., 2001).

Furthermore, there are ambiguity between chloride channels and transporters in terms of molecular architecture (Kim *et al*., 2022). Both CFTR and CLN7 exemplify proteins that adopt transporter-like architectures but function as ion channels. High-resolution cryo-EM studies of human CFTR revealed that, although it preserves the canonical ABC transporter fold, distinctive structural features, including a helix-to-loop transition in transmembrane helix 8 and a lateral portal, repurpose the scaffold to form an anion-conducting pore, while gating is controlled by phosphorylation-dependent movement of the regulatory (R) domain and ATP-driven NBD dimerization (Liu et al., 2017; Zhang et al., 2018; Hwang and Kirk, 2018). In contrast, CLN7 (MFSD8) was initially predicted to belong to the major facilitator superfamily with a 12-transmembrane topology (Siintola et al., 2007), but electrophysiological recordings demonstrated that it acts as a lysosomal chloride channel (Wang et al., 2021).

The most extensively studied CLC family also includes both pH sensitive chloride channels, and secondary active chloride proton antiporters, while utilising very similar homo-dimeric fold with independent ion transport in each subunit.

1.2.4 CLC Family

The Chloride Channel (CLC) family encompasses a diverse group of transporters and channels conserved across both prokaryotic and eukaryotic organisms. Characterized either as chloride channels or chloride/proton antiporters, these proteins have evolved complex functionalities essential to cellular ion homeostasis and regulation of membrane potential (Jentsch & Pusch, 2018).

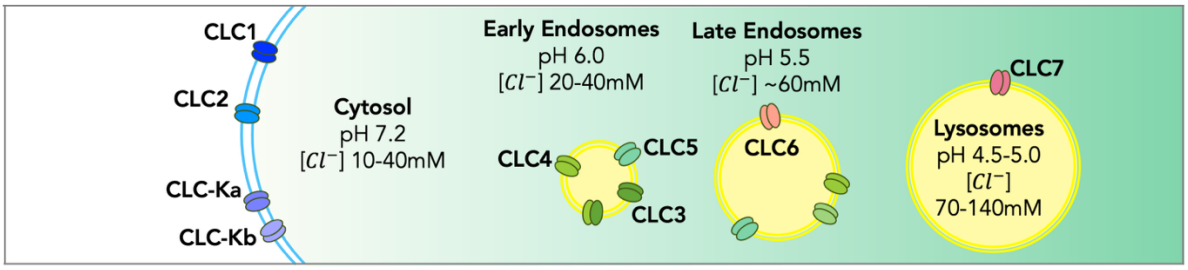
The first member of this protein family discovered was the *Torpedo* electroplax Cl− channel, CLC0 (Miller *et al.*, 1983; Miller *et al.*, 1984), by expression cloning from the *Torpedo* electric organ. Since then, genes corresponding to prokaryotic and eukaryotic members of the CLC family was rapidly identified (Jentsch *et al.,* 1999).

In prokaryotes such as *E. coli*, there are only two genes encoding CLC family proteins. The most studied member, CLCec1, is a chloride/proton antiporter localized to the outer membrane and contributes significantly to the organism's extreme acid resistance (XAR) response. This mechanism is crucial for enabling bacteria to survive harsh acidic environments (pH 2-3), which are conditions necessary for the successful colonization of the gut by benign microorganisms and for the virulence of pathogenic strains (Lyer *et al.*, 2003).

Plants, exemplified by *Arabidopsis thaliana*, possess a broader array of CLC transporters, with seven distinct isoforms (AtCLCa through AtCLCg). Of these, AtCLCa is the most thoroughly studied and is localized to the tonoplast membrane. Its primary role involves importing nitrate ($NO3^-$), and to a lesser extent chloride (Cl^-), from the cytoplasm into vacuoles for storage, harnessing the energy from the existing pH gradient across the membrane (Geelen *et al.*, 2001; Yang *et al*., 2023; He *et al*., 2023).

In humans, the complexity and diversity of the CLC family are further amplified, with nine distinct CLC proteins identified (CLC1 to CLC7, along with the isoforms CLC-Ka and CLC-Kb; Figure 5 A, B). Each member plays specialized roles, tailored to specific cellular contexts and physiological needs. This diversity underscores the crucial importance of precise regulation and functionality of CLC proteins in human health and disease (Jentsch & Pusch, 2018).

A



B

Human CLC Proteins		β-Subunit	Tissue	Function	Related Diseases
PM-localised Chloride Channels	CLC1	?	skeletal muscle	maintanence of membrane potential	myotonia
	CLC2	glialCAM	wide	transepithelial transport, regulation of excitability	leukodystrophy, aldosteronism
	CLC-Ka	Barttin	kidney, inner ear	transepithelial transport	Bartter IV
	CLC-Kb				Bartter III, Bartter IV
Vesicular Chloride-Proton Antiporters	CLC3	TMEM9	wide	endosomal ion and pH homeostasis	?
	CLC4	?	wide	endosomal ion and pH homeostasis	mental retardation, epilepsy
	CLC5	?	kidney	endosomal ion and pH homeostasis	Dent's disease
	CLC6	?	neuronal	endosomal ion homeostasis	?
	CLC7	OSTM1	wide	lysosomal ion and pH homeostasis, acidification of osteoclast	osteopetrosis, CNS or retina degeneration

Figure 5. Members of Human CLC Family Channels and Transporters

A. Schematic diagram of the localisation of human CLC family members in the cell; CLC1, 2, 3, Ka, and Kb are plasma-membrane localised chloride channels, while CLC4, 5, 6, and 7 localised to various endo-lysosomal vesicles. B. Table of tissue expression, function and disease related to human CLC family members; beta-subunits confirmed for several members also listed (adapted from Jentsch & Pusch, 2018).

Plasma membrane-localized members (CLC1, CLC2, CLC-Ka, and CLC-Kb) are chloride channels sensitive to pH fluctuations and are primarily involved in maintaining membrane potential and transepithelial transport. Conversely, CLC3 to CLC7 function as chloride/proton antiporters and localize to vesicular compartments within endo-lysosomal pathways, critically regulating vesicular ion composition and pH homeostasis (Figure 5 B; Jentsch & Pusch, 2018).

The evolutionary trajectory of the CLC family also reveals intriguing complexities in domain or subunit organisations. Prokaryotic CLC transporters exhibit a relatively simpler structural framework, comprising solely of a transmembrane domain (TMD) with 18 alpha helices (Figure 6 A). These transporters function as homodimers, with each subunit containing independent chloride permeation pathways (Accardi & Miller, 2004). In contrast, eukaryotic members have retained this fundamental TMD configuration but have evolved an additional cytoplasmic domain known as the cystathionine beta-synthase (CBS) domain, situated at the C-terminal region (Figure 6 B). Each CBS domain consists of two motifs that form dimers, subsequently dimerizing again to complete the final cytoplasmic domain structure. Though CBS domains do not directly participate in ion transport, they are hypothesized to regulate channel or transporter activity (Further discussed in **Chapter 1.2.6**; Jentsch & Pusch, 2018).

Furthermore, several human CLC proteins depend on auxiliary beta-subunits for proper trafficking and functionality (Figure 6 C, D; Jentsch & Pusch, 2018; Stölting *et al*, 2014). GlialCAM, a cell adhesion molecule expressed prominently within the central nervous system, is essential for targeting CLC2 to cell junctions, influencing its functional expression (Jeworutzki *et al*., 2012). Similarly, barttin, primarily expressed in the kidney and inner ear, is vital for the proper localization and function of CLC-Ka and CLC-Kb, integral to renal salt reabsorption (Krämer *et al*., 2008). OSTM1 serves as a mandatory beta-subunit for CLC7, with deficiencies in either protein leading to osteopetrosis symptoms (Figure 6 D; Lange *et al*., 2006). Most recently, studies have identified TMEM9 as an inhibitory beta-subunit for CLC3,

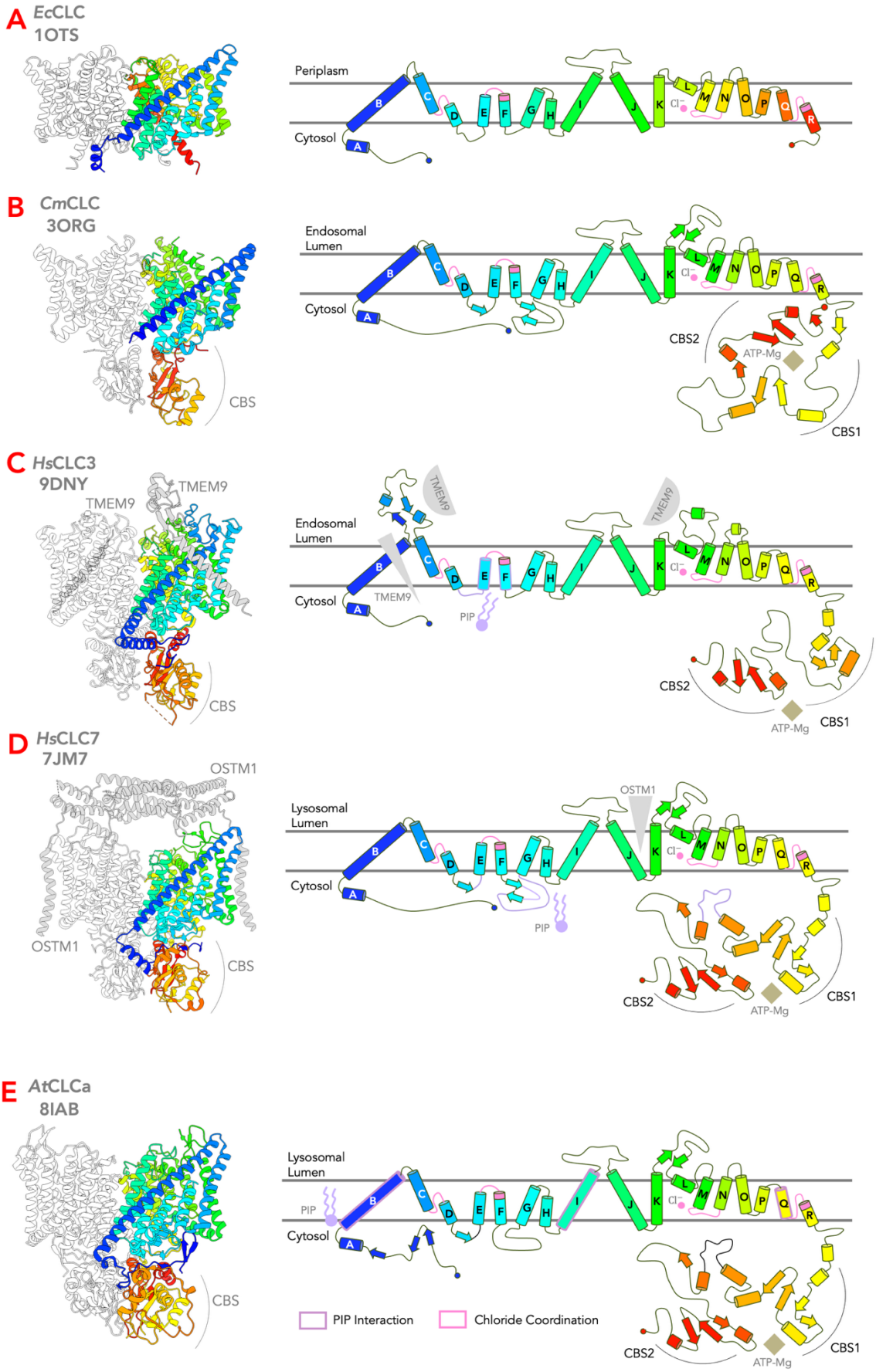

A EcCLC 1OTS
Periplasm
Cytosol
B CmCLC 3ORG
CBS
Endosomal Lumen
Cytosol
CBS2
ATP-Mg
CBS1
C HsCLC3 9DNY
TMEM9
TMEM9
CBS
Endosomal Lumen
Cytosol
TMEM9
TMEM9
PIP
CBS2
ATP-Mg
CBS1
D HsCLC7 7JM7
OSTM1
OSTM1
CBS
Lysosomal Lumen
Cytosol
OSTM1
PIP
CBS2
ATP-Mg
CBS1
E AtCLCa 8IAB
CBS
Lysosomal Lumen
PIP
Cytosol
PIP Interaction
Chloride Coordination
CBS2
ATP-Mg
CBS1

Figure 6. Topology of Representative CLC Transporters

A-E. Topology of representative CLC transporters, with conserved chloride binding sites coloured or highlighted in pink, beta-subunits indicated in grey, nucleotides in light brown, and lipid binding sites highlighted in purple. Structure of one protomer coloured by rainbow (N- to C- terminus), the other coloured in white. Topology diagram coloured by rainbow. A. Structure and topology of *E. coli* CLC transporter, containing only transmembrane domain (PDB 1OTS, Dutzler *et al.*, 2003). B. Structure and topology of *Cyanidioschyzon merolae* CLC transporter, with the additional CBS domain (PDB 3ORG, Feng *et al.*, 2010). C. Structure and topology of human CLC3 transporter in complex with TMEM9 (PDB 9DNY, Schrecker *et al.*, 2025). D. Structure and topology of human CLC7 transporter in complex with OSTM1 (PDB 7JM7, Schrecker *et al.*, 2020). E. Structure and topology of *Arabidopsis thaliana* CLCa transporter (PDB 8IAB, Yang *et al.*, 2023).

CLC4, and CLC5, which prevent toxic chloride accumulation and vacuole swelling (Figure 6 C; Schrecker *et al.*, 2025).

In addition to regulation mediated by beta-subunits, human CLC proteins also exhibit intricate lipid-mediated regulation. Notably, CLC3 and CLC7 are both found to be inhibited by the signalling lipid PI(3,5)P2, located on endo-lysosomal membranes (Leray *et al.*, 2022; Schrecker *et al.*, 2025). Interestingly, CLC3 and CLC7 interact with the inhibitory lipid with slightly different epitopes around helix E, F, and G (Figure 6 C, D). Furthermore, in the structure of *Arabidopsis thaliana* CLCa, PI(4,5)P2 was found binding at the dimer interface (Figure 6 E; Yang *et al.*, 2023), a completely distinct binding site from that of human CLC3 and 7, suggesting these specific mechanisms of regulation might have evolved independently, allowing precise control of CLC activity in specific cellular compartments and environmental conditions.

In the following chapter, I will introduce the structural features of CLC7 in more detail as an example of a mammalian CLC transporter.

1.2.5 Structure of CLC7

CLC7, consistent with other members of the CLC family, exhibits a highly conserved TMD fold. Each subunit of CLC7 adopts a triangular shape when viewed from either the extracellular

or cytoplasmic side (Figure 7 A, B; Schrecker *et al*., 2020; Zhang *et al*., 2020). Structurally, the TMD is characterized by 18 alpha helices (denoted A–R), of which 17 are partially embedded within the membrane, with several helices not traversing completely but forming small intra-membrane loops critical for chloride coordination (Figure 6 D, figure 7 A; Schrecker *et al*., 2020). These loops significantly influence ion transport specificity and efficacy (Dutzler *et al*., 2002; Zhang *et al*., 2020). Each CLC7 monomer reveals an inverted symmetry architecture within its TMD (Figure 6 D). This pseudo-symmetry is evident as helices A–I mirror the structural organization of helices J–R, albeit with an inverted membrane orientation (Dutzler *et al*., 2002). Such symmetry is essential for the functionality of the transporter, allowing it to adopt different conformational states crucial for ion translocation and gating mechanisms (Dutzler *et al*., 2002).

CLC7 functions as a dimer, characterized by notably an extensive, mostly hydrophobic interface of 3020 Å^2 (calculated with published structure PDB id 7JM7, Schrecker *et al*., 2020). However, unlike most ion channels, in which one single ion permeation pathway is formed in the centre of the complex, each monomer of CLC7 is capable of carrying out chloride/proton exchange (Jentsch & Pusch, 2018).

Central to the function of CLC7 are its chloride binding sites, formed by four highly conserved regions on helices D, F, N, and R (Figure 6 D; Jentsch & Pusch, 2018). The last membrane-embedded helix, R-helix, forms a significant part of the CLC7 ion conducting pore and acts as the selectivity filter for chloride ions (Jentsch & Pusch, 2018). A crucial cytoplasmic segment, termed the R-helix linker, connects the R-helix to the intracellular CBS domain (Figure 6 D).

The C-terminal region of CLC7 comprises two cystathionine-β-synthase (CBS) domains, conserved motifs commonly implicated in modulating protein function through nucleotide binding. The CBS domains form a sandwich-like structure, called a Bateman module, via their beta strands, typically binding adenine nucleotides (Figure 7 C; Jentsch & Pusch, 2018).

Comparative structural analyses highlight that in lower organisms like the red alga *Cyanidioschyzon merolae*, CBS1 resides distantly from the membrane domain, while CBS2 closely interfaces with the TMD (Figure 6 B; Feng *et al*., 2010). In contrast, in higher organisms like *Arabidopsis thaliana* and mammals (CLC3 and CLC7), both CBS motifs closely interact with the TMD, sometimes mediated by an N-terminal loop situated between the TMD and CBS domains (Figure 6 C-E; Yang *et al*., 2023; Schrecker *et al*., 2020; Schrecker *et al*., 2025; Zhang *et al*., 2025). In CLC7 specifically, an interaction surface of 1763 Å² is formed between the CBS domains and the N-terminal loop and TMD, with complexity resembling antibody-antigen interactions (Schrecker *et al*., 2020).

CLC7 associates with the β-subunit OSTM1, a heavily glycosylated protein positioned above and around the TMD (Lange *et al*., 2006; Schrecker *et al*., 2020). OSTM1 acts as a protective "lid," shielding CLC7 from harsh lysosomal conditions (Figure 7 A). The subunit extensively interacts with CLC7's luminal and TMD regions, forming an interface of 1190 Å² (and 2314 Å² between the OSTM1 monomers) (Zhang *et al*., 2020; Schrecker *et al*., 2020). Unlike TMEM9, which is known to inhibit the activity of CLC3, CLC4, and CLC5 (Planells-Cases *et al*., 2025), the effect of OSTM1 on CLC7 is not yet known, and binding of OSTM1 did not seem to change the conformation of the transporter (Schrecker *et al*., 2020).

1.2.6 Gating, Fast and Slow

Electrophysiological studies of CLC family proteins have identified two distinct gating mechanisms: a "fast gating" mechanism occurring on millisecond scales, independently controlling each pore, and a "slow gating" mechanism affecting both pores simultaneously over seconds (Figure 8 A; Miller *et al*., 1984; Jentsch & Pusch, 2018).

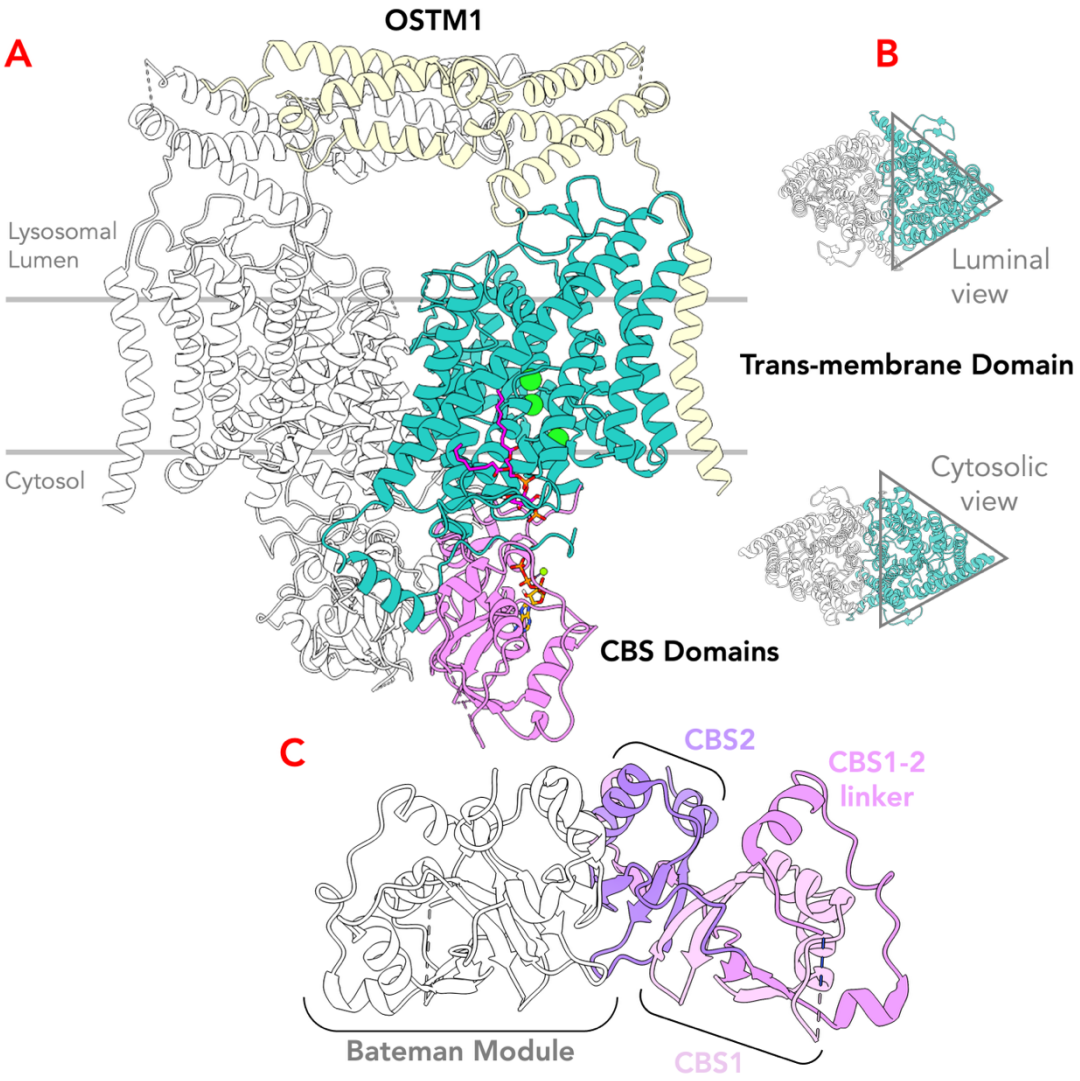


Figure 7. Architecture of Human CLC7

A. Structure of human CLC7 (PDB 7JM7, Schrecker *et al.*, 2020), transmembrane domain of one protomer coloured in teal, CBS domain coloured in pink, chloride coloured in green, lipid coloured in magenta; one of the beta-subunit OSTM1 coloured in light yellow. B. Transmembrane domain of CLC7 viewed from either lysosomal lumen, or cytosol, one protomer coloured teal, the other protomer coloured white. C. CBS domains of CLC7 composed of two Bateman module (one of which coloured in white); within the other Bateman module, CBS1 motif is coloured in pink, CBS2 motif in dark purple, and the linker coloured in light pink.

The fast gate in CLC proteins involves structural elements including the GSGIP sequence on helix D, the G(K/R)**E**GP motif on helix F, the GXFXP sequence on helix N, and a critical conserved tyrosine residue on helix R (Y602 on CLC7). Among these, the gating glutamate (Glu$_{gate}$, E247 in human CLC7) positioned on helix F is the most conserved residue critical for gating across the majority of CLC channels and transporters (Figure 8 A, C; Strange *et al.*,

2011). In humans, 7 out of 9 CLC genes encodes the gating glutamate residue (Supplementary Figure 1), only CLC-Ka and CLC-Kb encodes a valine in the equivalent position. The gating glutamate protrudes into the chloride permeation pathway and blocks chloride ions from communicating with lysosomal luminal bulk solvent (Zhang *et al*., 2020; Schrecker *et al*., 2020). In CLC7, mutating the gating glutamate to alanine completely abolishes chloride proton coupling, and led to strong, constant currents, turning the transporter into a “leak channel” (Leisle *et al*., 2011). This is also observed in other CLCs such as the Torpedo CLC0 channels, and CLC3, 4, and 5 transporters (Traverso *et al*., 2003; Picollo and Pusch, 2005; Scheel *et al*, 2005)

Extensive structural studies of homologs such as those from *E. coli* and *Cyanidioschyzon merolae* revealed four distinct states for the gating glutamate residue (Figure 8 B): “middle” (inward-open, PDB 1OTS, Dutzler *et al*., 2003), “up” (occluded, PDB 1OTU, Dutzler *et al*., 2003), “down” (apo, PDB 3ORG, Feng *et al*., 2010), and “out” (outward-open, PDB 6V2J, Chavan *et al*., 2020). Notably, all published structures of CLC7 show Glugate (E247) adopting the "up" conformation, indicative of an occluded state (Zhang *et al*., 2020; Schrecker *et al*., 2020).

Multiple other conserved residues surrounding the chloride binding site coordinate ions directly. Specifically, in CLC7, S204 creates a central, deep-pore site adjacent to the gating glutamate E247; and Y505 along with Y602 form an internal cytosolic site, stabilizing chloride ions prior to their release into the cytoplasm (Figure 8 C; Schrecker *et al*., 2020). Among these, tyrosine 602 is another highly conserved residue involved in chloride transport, found in all human CLCs. In *E. coli*, mutations to alanine and leucine led to uncoupling of chloride and proton transport and chloride leak (Walden *et al*., 2007).

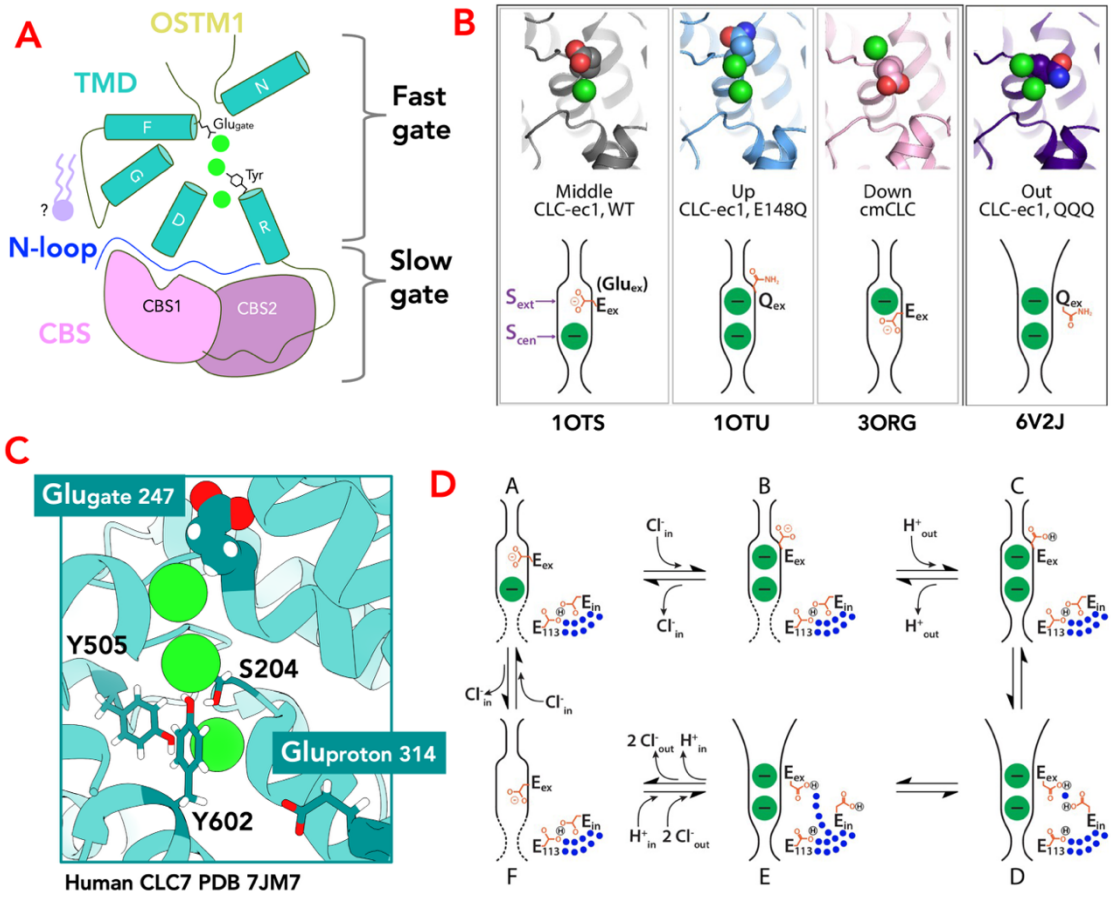


Figure 8. Fast gating of CLC Transporters

A. Schematic diagram of the fast gate and slow gate of CLC7, with the transmembrane domain coloured in teal, CBS domains in pink and dark purple, a stretch of N terminal loop in blue, and chloride ions in green. Gating glutamate E247 and Tyrosine 602 side chain shown on helix F and R. Lipid binding shown between helix F and G. B. 4 possible conformational states the gating glutamate could adopt, with the PDB of respective structures shown. Adapted from Chavan *et al.*, 2020. C. Chloride binding sites in human CLC7 (using PDB 7JM7); side chain of S204, E247, E314, Y505, and Y602 shown, chloride ions coloured in green. D. Schematic diagram of chloride/proton antiport in CLC transporters, adapted from Chavan *et al.*, 2020. Blue circles indicate water molecules.

Unique to CLC transporters is a titratable proton glutamate (Figure 8 C; Glu_{proton}, E314 in human), situated slightly off the chloride pathway, resulting in a Y-shaped divergence of the chloride and proton pathway on the cytosolic side (Dutzler *et al.,* 2003). In humans, the proton glutamate is only present in CLC3, 4, 5, 6, and 7. In CLC7, mutating the proton glutamate to alanine led to complete abolishment of all ion transport (Leisle *et al*., 2011). In *E. coli*, mutations of proton glutamate E203 to residues such as aspartic acid, histidine, lysine or arginine retain proton/chloride exchange; non-protonatable substitutions such as valine,

isoleucine or cysteine abolish proton coupling and strongly depress chloride flux (Lim *et al*., 2009). “Neutral” replacements such as glutamine, serine or alanine still allow chloride movement at ~30% the rate of wild type, but greatly impair H^+ transport, i.e., uncoupling (Lim *et al*., 2009).

Compared to chloride transport, the mechanism through which protons are transported in CLC transporters is less well understood. One hypothesis is that chloride ion binding and protonation of the gating glutamate can shift its conformation from "up" to "out," facilitating proton transfer to other residues or water networks, ultimately reaching the proton glutamate and the cytosolic bulk solvent (Figure 8 D; Chavan *et al*., 2020).

Despite extensive characterisation of fast gating, the precise structural composition and exact location of the "slow gate" remain unclear (Figure 8 A).

In *E. coli* CLC, the common gate is formed by the cooperative occlusion or exposure of the intracellular proton pathways at the dimer interface (Fortea *et al.,* 2024). In the inactive state, the N-terminal helix A and loops between helices H and I from each protomer cross over and block the H^+ vestibules of the opposite protomer. Activation occurs through pH-dependent rearrangements that disengage helix A and retract the H-I loop, thereby opening both H^+ pathways of the homo-dimer simultaneously. A key trigger is the protonation of E202, which lowers the energy barrier for helix A disengagement, leading to hydration of the H^+ vestibules and the formation of short water wires that connect the gating glutamate to the cytosol. This hydration dynamically destabilizes the Cl^- pathway gates, enabling Cl^-/H^+ exchange (Fortea *et al.,* 2024, discussed more in **Chapter 3.6**).

In eukaryotic homologs, the slow gate is widely attributed to interactions involving CBS domains, given the strong correlation of mutations in these domains with disease phenotypes across multiple CLC proteins (Zhang *et al*., 2020).

Proposed regulatory mechanisms involving CBS domains include but do not limit to nucleotide binding and post-translational modifications. Explicit demonstrations of nucleotide regulation are evident in CLC1, CLC3, CLC5, and CLC6 (Tseng *et al*., 2011; Wan *et al*., 2024; Zifarelli *et al*., 2009; Zhang *et al*., 2023). For instance, Zifarelli *et al*. (2009) showed that ATP, ADP, and AMP binds to CLC5 with similar affinity, and equally enhanced its activity. CryoEM studies by Wan *et al*. (2024) revealed ATP binding in CLC3 specifically alters chloride ion occupancy, enhancing chloride transport.

In CLC6 which is the closest to CLC7, ATP substantially increases currents, as recently demonstrated (Zhang *et al*., 2023). However, for CLC7, multiple studies suggest ATP binds with high affinity but does not directly modulate its transport activity (Schrecker *et al*. 2020, Leisle at al. 2011). Instead, ATP likely plays a structural role by bridging CBS domains with the transmembrane regions by uniquely interacting with a stretch of the N-terminal loop (Figure 8 A, figure 9 A).

Domain and dimer interfaces have also been implicated in slow gating, especially in CLC7. Zhang *et al*. (2020) demonstrated that stable packing interactions at multiple interfaces (N-terminus/TMD/CBS, CBS dimer, TMD dimer, and CLC7/Ostm1 complexes) create high kinetic barriers underlying the characteristic slow gating of CLC7. Mutations weakening these interactions consistently accelerate activation kinetics (Figure 9 B; Zhang et al., 2020).

Additionally, the CBS1–CBS2 linker, highly variable among CLC proteins, is suspected to mediate regulatory interactions via PTMs or lipid/ligand binding. For instance, the *C. elegans* homolog CLH-3b is regulated through kinase GCK3, which interacts within the CBS1-CBS2 linker (Strange *et al*., 2011). Similarly, phosphorylation is proposed as a regulatory mechanism in human CLC1, CLC2, and CLC3 via other kinases and phosphatases (Camerino *et al*., 2014; Conte *et al*., 2020; Furukawa *et al*., 2002; Zheng *et al*., 2002; Wang *et al*., 2013; Cuddapah *et al*., 2010). In human CLC6 and CLC7, part of the CBS1–CBS2 linker forms close contact with

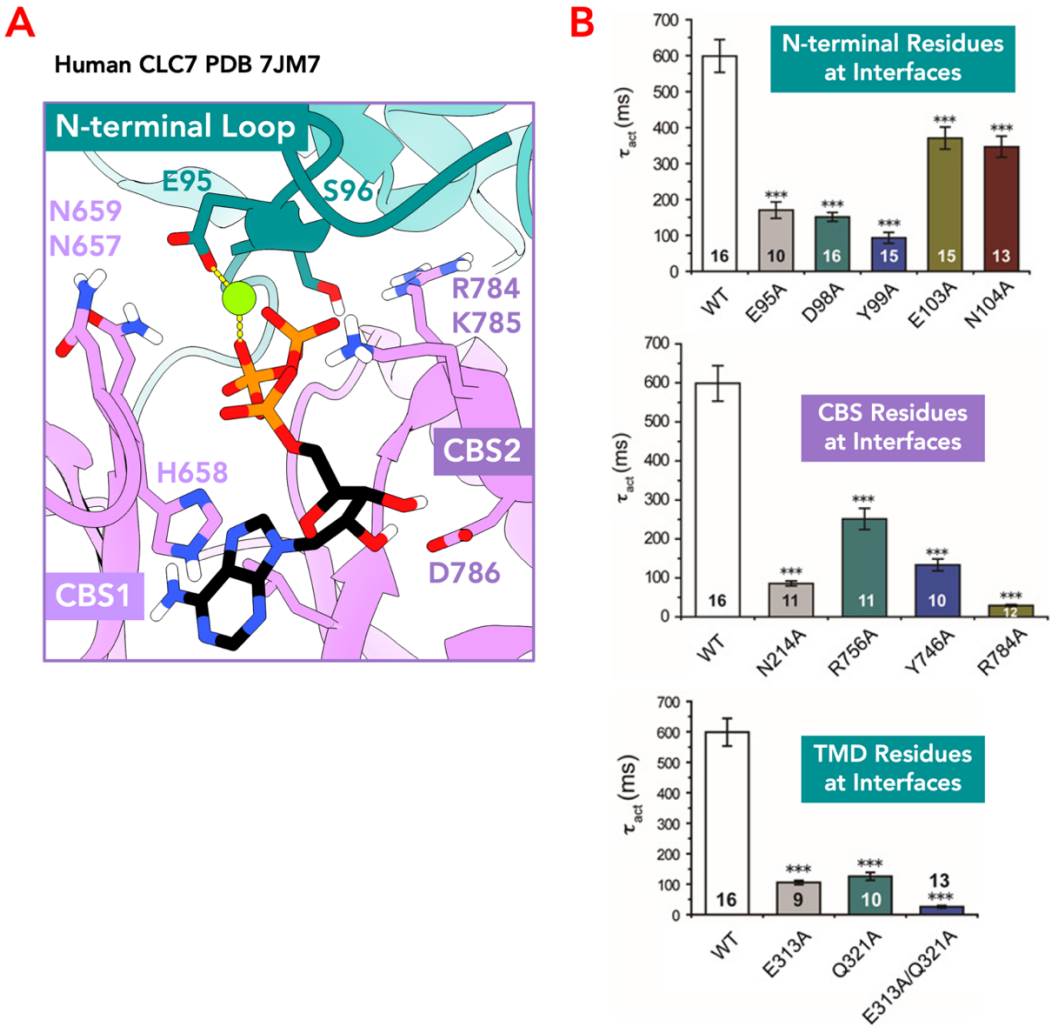


Figure 9. Slow Gating of CLC Transporters

A. Interaction of ATP-magnesium with the CBS domains (pink) and N-terminal loop (teal) of human CLC7 stabilised the domain interface. B. Mutations long the domain interfaces in human CLC7 lead to faster activation kinetics. The $\tau_{activation}$ values were measured at +140mV, with data presented with error bars representing standard error of the mean. Numbers in bars indicate repeats of experiments in each group.

the transmembrane domain (TMD) and forms a lipid-binding pocket (Schrecker *et al*., 2020; Zhang *et al*., 2023). In CLC7, it has been shown that lipid binding and interaction with that region of the CBS1-CBS2 linker inhibits transport (Leray *et al*., 2022; discussed in more detail in **Chapter 1.2.8**).

A critical gap remains in understanding precisely how conformational changes caused by nucleotide/lipid binding or PTMs within CBS domains translate to transport modulation in

CLC transporters. Some suggested that, the last transmembrane helix, helix R, might play a role in signal transduction (Strange *et al*., 2011). The R-helix–CBS linker, despite sequence variability among CLC homologs, often closely interacts with CBS2 domains, potentially able to translate CBS domain alterations into conformational changes affecting the transporter's fast gate.

1.2.7 CLC7 Dysregulation Leads to Diverse Pathologies

Dysregulation or mutations in CLC7 or its accessory protein OSTM1 lead primarily to osteopetrosis, a severe bone disease characterised by impaired bone resorption and consequently dense, brittle bones (Figure 10 A-C; Kornak *et al*., 2001; Leisle *et al*., 2011). The primary mechanism underlying osteopetrosis involves malfunctioning osteoclasts unable to acidify the resorption lacuna, essential for dissolving bone minerals (Figure 10 B; Kornak *et al*., 2001; Leisle *et al*., 2011).

Besides osteopetrosis, loss-of-function mutations in CLC7 have also been associated with other severe disorders, notably retinal degeneration and neurodegenerative conditions (Figure 10 A, C; Kasper *et al*., 2005; Nicoli *et al*., 2019; Zifarelli *et al*., 2022). The precise role of CLC7 in these diseases was debated, especially given several controversial findings regarding its involvement in lysosomal pH regulation, with several studies indicating unchanged lysosomal pH despite CLC7 knock-out (Figure 10 D; Kasper *et al*., 2005; Leray *et al*., 2022).

Mutations in CLC7 exhibit notable complexity. Certain mutations which disrupt normal protein trafficking, leading to a reduction in functional CLC7 at lysosomal membranes or the resorption lacuna, thereby exhibiting classical loss-of-function characteristics as observed in osteopetrosis; even though in reconstituted condition, the mutants exhibited paradoxical features such as faster activation kinetics or increased ion currents on a molecular level (Leisle

et al., 2011). Specific alanine mutations disrupting the transmembrane domain (TMD) and CBS domain interfaces have also been reported to have faster activation kinetics (Figure 10 D; Zhang *et al.*, 2020).

An intriguing gain-of-function mutation identified in CLC7 is the Y715C variant, another mutation at the TMD CBS interface, causing hyperactivation of the protein (Figure 10 D; Nicoli *et al*., 2019). Clinically, this mutation leads to distinct pathologies such as albinism and lysosomal storage diseases in various organs, but not osteopetrosis (Figure 10 E). More recently, another gain-of-function disease variant, K285T, has also been discovered (Polovitskaya *et al*., 2024). Like tyrosine 715, lysine 285 is also located at the TMD CBS interface, and patients with the K285T variant exhibited the same disease phenotypes as those with the Y715C variant, namely, hypopigmentation, organomegaly, and delayed myelination and development -- the "HOD syndrome" (Polovitskaya *et al*., 2024).

The discovery of various functionally significant residues and disease-causing mutations at domain interfaces, especially at the TMD CBS interfaces suggested the importance this region in the regulation of CLC7 activity, and one of the hypothesised mechanism of regulation is through the signalling lipid PI(3,5)P2.

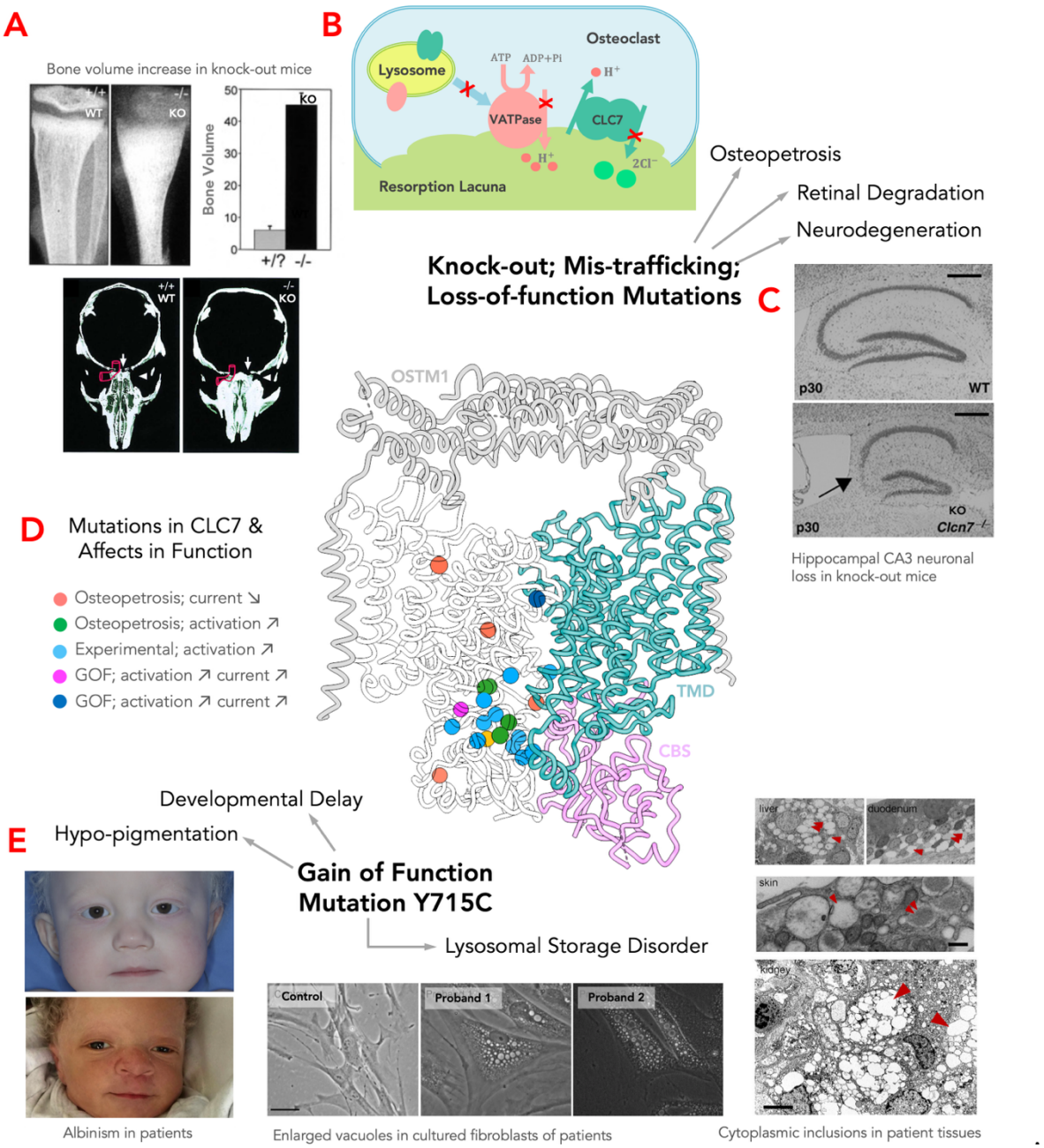


Figure 10. Disease Phenotypes Related to Human CLC7

A,C. CLC7 knock-out led to osteopetrosis, signified by high bone density (Kornak *et al.*, 2001), retinal degradation (Kornak *et al.*, 2001), and neurodegeneration (Kasper *et al.*, 2005). B. Schematic diagram of the role of CLC7 at the ruffle border of osteoclasts, and disruptions during osteopetrosis. D. Structure of human CLC7 (PDB 7JM7, licorice representation), one protomer coloured by domain, another protomer coloured in white, with various mutations mapped on the structure as spheres (Leisle *et al.*, 2011; Zhang *et al.*, 2020; Leray *et al.*, 2022). E. Clinical gain of function mutation Y715C led to hypopigmentation, developmental delay, lysosomal storage disorder, but not osteopetrosis.

1.2.8 Lipid-linked regulation of CLC7

Lipid-linked regulation of ion channels and transporters has emerged as a significant mechanism modulating cellular processes. A recent study by Leray *et al*. (2022) specifically explored the lipid regulation of the chloride transporter CLC7 by phosphatidylinositol 3,5-bisphosphate (PI(3,5)P2). This lipid is synthesized by the kinase PIKfyve, and its depletion through pharmacological inhibition with apilimod results in significant hyperacidification of lysosomes. Such hyperacidification corresponds to a reduction in lysosomal pH by approximately 0.31 units, effectively doubling the proton concentration within these organelles (Figure 11 A).

Remarkably, the role of CLC7 in this context became evident when knockout cells for this transporter exhibited significantly reduced hyperacidification upon apilimod treatment, compared to wild-type cells (Figure 11 B; Leray *et al*., 2022). These findings suggested a direct involvement of CLC7 in the pathway leading to lysosomal hyperacidification under conditions of reduced PI(3,5)P2 levels.

Further electrophysiological studies using patch-clamp techniques provided compelling evidence that PI(3,5)P2 specifically and directly inhibits CLC7 currents (Figure 11 C; Leray *et al*., 2022). Experimental introduction of PI(3,5)P2 into cells expressing CLC7 significantly reduced transporter currents, an effect not seen with the addition of other phosphoinositides.

Intriguingly, the Y715C mutation in CLC7, previously identified as a disease-causing variant, disrupts this lipid-mediated regulation (Figure 11 C). Cells expressing the Y715C mutant displayed enhanced transporter currents that remained unaffected by PI(3,5)P2, mirroring the lysosomal hyperacidification observed during PI(3,5)P2 depletion (Leray *et al*., 2022). This observation highlights the critical role of Y715 in the lipid-mediated regulatory mechanism.

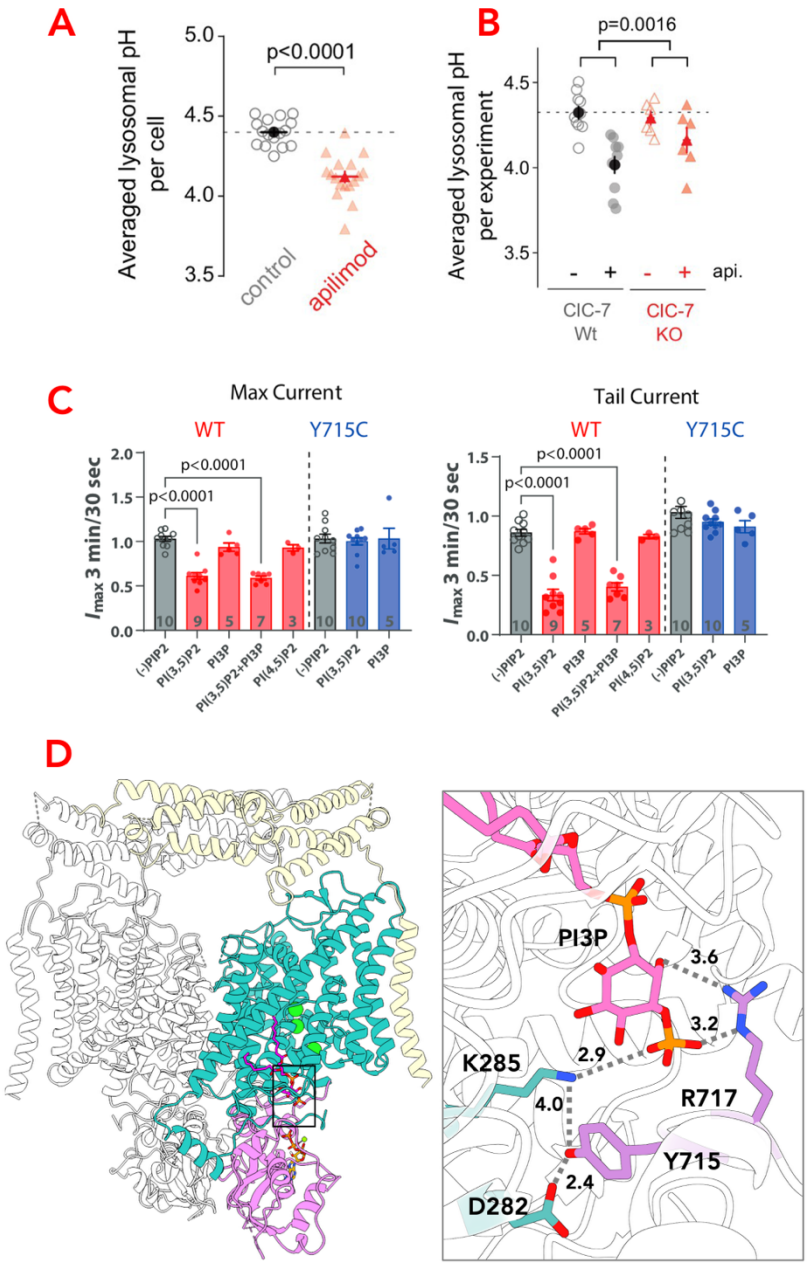


Figure 11. Lipid Linked Regulation of Human CLC7

A. Apilimod treatment led to a reduction of lysosomal pH through inhibition of PI(3,5)P2 synthesis by PIKfyve. B. CLC7 knock-out diminished the effect of apilimod treatment. C. PI(3,5)P2 specifically inhibited wtCLC7 currents, but did not affect the currents of Y715C mutant. D. Left: structure of human CLC7 (PDB 7JM7), one protomer coloured by domain. Right: zoomed in view of lipid binding pocket with PI3P bound (hot pink), side chains of several interacting residues shown and coloured by the domain (TMD: teal; CBS: purple). Hydrogen bonds shown in dotted line.

Based on these findings, Leray *et al*. proposed a model where CLC7 activity in lysosomes is tonically inhibited by PI(3,5)P2. According to this model, under certain physiological conditions, PI(3,5)P2 dissociates from CLC7, activating the transporter and thereby facilitating

lysosomal acidification. Mutations like Y715C hyperactivates the transporter by disrupting lipid binding.

Structurally, PI3P, another non-lysosomal phosphoinositide, has been identified bound in a pocket between the transmembrane domain and the CBS domains of CLC7 (PDB ID: 7JM7; Schrecker *et al*., 2020). The lipid engaged with a number of residues in a hydrogen bond network, including residue Y715. Notably, despite its structural binding, PI3P does not exert inhibitory effects on CLC7 currents (Figure 11 D; Leray *et al*., 2022).

Therefore, important questions remain regarding the detailed mechanism of PI(3,5)P2-mediated inhibition of CLC7. Specifically, it remains unclear how lipid binding translates into conformational changes in the transporter, how these changes affect the transport pathway, and the roles that domain interfaces and CBS domains play in this regulation.

1.2.9 Aims of CLC7 Study

The aims of my PhD thesis are focused on elucidating the unresolved aspects regarding the lipid-linked regulation of CLC7 described above. My research aims to:

1. determine the high-resolution structure of CLC7 bound to PI(3,5)P2, as well as the structure of the clinically relevant mutant Y715C. These structures are expected to clarify the nature of structural changes induced by the inhibitory lipid or the hyperactivating mutation, providing insights into their mechanistic implications for CLC7 function (Chapter 3).
2. In addition, I plan to produce high affinity binders specific to CLC7, allowing further research of its cellular role *in situ*, in more native environments (Chapter 4).

1.3 RUSF1 & the Biogenesis & Quality Control of Membrane Proteins

1.3.1 The Morphology & Compartmentalisation of the Endoplasmic Reticulum

The endoplasmic reticulum (ER) is central to cellular physiology, playing critical roles in protein biosynthesis, lipid metabolism, and calcium homeostasis. Approximately 25–30% of all eukaryotic proteins fold within the ER, which demands a highly coordinated protein synthesis and quality control system to safeguard fidelity during protein maturation. Misfolded proteins that escape surveillance can accumulate and aggregate, leading to disruption of cellular homeostasis and contributing ER storage diseases (ERSDs) as well as other disorders, including cancer and neurodegeneration (Chen *et al*., 2023).

Structurally, the ER is highly heterogeneous in its morphology and is organised into distinct domains that enable functional compartmentalisation. The rough ER, characterised by ribosome-studded membrane sheets, provides an environment optimised for co-translational insertion and folding of nascent proteins. The morphogenetic protein CLIMP-63 is essential for sheet architecture, as it spans the ER lumen and stabilises the spacing between membranes, thereby promoting sheet formation (Shibata *et al*., 2010). In contrast, the smooth ER is composed predominantly of highly curved tubules, which are often devoid of ribosomes and specialised for lipid metabolism, organelle contact sites, and intracellular trafficking (Shibata *et al*., 2010). Tubule morphology is stabilised by reticulon (RTN) family proteins, which insert wedge-like domains into the membrane, thereby maintaining curvature (Shibata *et al*., 2010).

Within this broader architecture, the ER is further partitioned into nano-domains that facilitate specialised functions. For example, sheet edges and tubule junctions generate local curvature that promotes the assembly of ER exit sites (ERES), where selected cargo is sorted for COPII-mediated trafficking (Hughes *et al*., 2009; Kurokawa *et al*., 2014). Other specialised

subdomains are dedicated to protein quality control. A distinct perinuclear ER quality control compartment (ERQC) serves as a hub where misfolded proteins, molecular chaperones, and degradation machinery converge to drive ER-associated degradation (ERAD) (Kamhi-Nesher *et al*., 2001; Bernasconi *et al*., 2010).

The biosynthesis of soluble proteins and the soluble domains of integral membrane proteins have been extensively studied and is relatively well understood (Brodsky *et al*., 2011; Braakman *et al*., 2013). Key factors in this process include signal recognition and cleavage by the signal peptidase complex (Shan and Walter, 2005; Jiang *et al*., 2008; Sun *et al*., 2022), translocation into the ER lumen through the Sec61 translocon (Rapoport *et al*., 2008), and maturation via post-translational modifications such as glycosylation and disulfide bond formation, mediated by oligosaccharyltransferases and protein disulfide isomerases (Brodsky *et al*., 2011; Braakman *et al*., 2013). Proper folding and stabilisation of these proteins are further facilitated by a network of molecular chaperones, including members of the *Hs*p70 and *Hs*p90 families, J-domain proteins, and the lectin chaperone calnexin, among others (Brodsky *et al*., 2011; Braakman *et al*., 2013).

By contrast, the biogenesis and quality control of membrane proteins remain comparatively elusive and present unique challenges. Roughly one quarter of the proteome in all organisms is composed of α-helical membrane proteins, whose insertion and assembly within the lipid bilayer require a fundamentally different set of mechanisms compared to soluble proteins (Hedge *et al*., 2022). The biosynthetic pathway of membrane proteins can be divided into several distinct steps: targeting to the ER, insertion of transmembrane domains (TMDs) with correct topology, folding into the final three-dimensional architecture, and, in many cases, assembly with obligatory binding partners (Hedge *et al*., 2022). The structural and physicochemical diversity of transmembrane helices demands the evolution of specialised

molecular machinery to facilitate each of these steps for different classes of client membrane proteins (Figure 12).

In the following sections, I will introduce the key molecular machineries that facilitate membrane protein biogenesis and quality control, highlighting their structural and functional features.

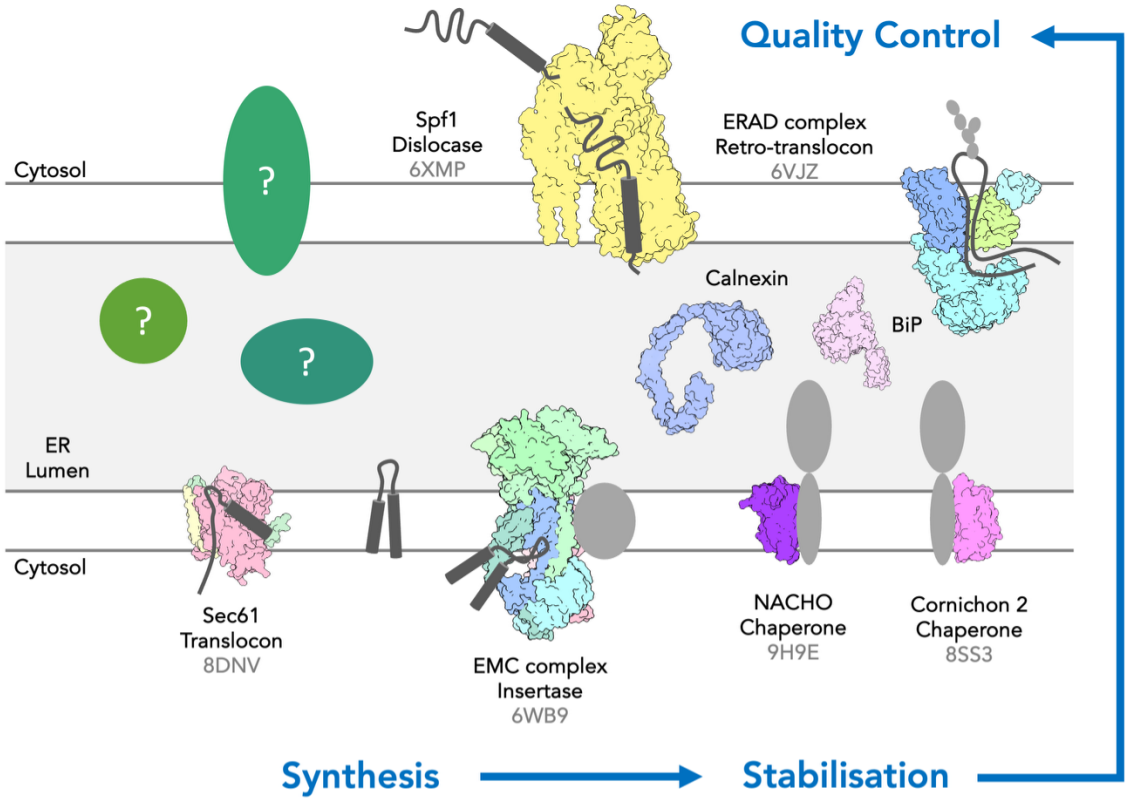


Figure 12. Machineries in Membrane Protein Biogenesis and Quality Control

Schematic diagram of various ER resident proteins involved in membrane protein synthesis, stabilisation, and quality control. Sec61 translocon (from human, PDB 8DNV, Itskanov *et al.*, 2023) and EMC complex (from *S. cerevisiae*, PDB 6WB9, Bai *et al.*, 2020) are machineries responsible for inserting membrane protein helices. NACHO (from human, PDB 9H9E, Hooda *et al.*, 2024, preprint) is a chaperone stabilising subunit of pentameric ligand-gated ion channels for proper assembly. Cornichon 2 (from human, PDB 8SS3, Gangwar *et al.*, 2023) is a chaperone regulating activities of AMPA receptors. Spf1 (from *S. cerevisiae*, PDB 6XMP, McKenna *et al.*, 2020) recognises mis-inserted mitochondrial proteins and remove them from ER membrane. ERAD-L complex (from *S. cerevisiae*, PDB 6VJZ, Wu *et al.*, 2020) retro-translocates ubiquitinated protein for degradation by the proteosome. Grey circles indicate binding partner/substrate of the respective resident ER proteins, green ovals indicate unknown factors involved in membrane protein synthesis, stabilization, and quality control.

1.3.2 Machineries Involved in Membrane Protein Biogenesis & Assembly

The insertion of transmembrane helices into the ER membrane is mediated primarily by two well-characterised protein families: the SecY/Sec61 translocon family and the Oxa1 family of insertases (Figure 12; Hedge and Keenan, 2024). Both families support the integration of α-helices into the membrane but operate through distinct architectural principles.

In mammals, the Sec61 translocon is the central component of co-translational protein translocation (Figure 13 A). It forms a heterotrimeric complex composed of the large membrane-spanning Sec61α subunit together with two smaller subunits, Sec61β and Sec61γ (Figure 13 B-C, Hedge *et al*., 2022; Braunger *et al*., 2018; Voorhees *et al*., 2016). Sec61α is organised into two pseudo-symmetric halves that together form a central channel across the membrane, which in its resting state is occluded by a short "plug" helix in the N-terminal bundle (PDB 6FTI, Figure 13 B, Braunger *et al*., 2018). Upon ribosome binding, conformational changes in Sec61 open this channel, while simultaneously creating a lateral gate through which hydrophobic TMDs can be released into the lipid bilayer. This dual architecture enables Sec61 to support both the translocation of hydrophilic domains through its aqueous channel and the partitioning of hydrophobic segments into the membrane (PDB 3JC2, Figure 13 C, Voorhees *et al*., 2016).

In contrast, the ER membrane complex (EMC) exemplifies a different mechanistic approach to TMD insertion (Figure 13 D-F, Hedge *et al*., 2022; Kumazaki *et al*., 2014; Bai *et al*., 2020). Its core subunit, EMC3, is homologous to the bacterial homolog YidC (Figure 13 E), and features a conserved three-TMD bundle that forms a hydrophilic vestibule extending partway into the lipid bilayer (Kumazaki *et al*., 2014). Rather than forming a complete channel, this "half-channel" locally distorts and thins the surrounding membrane, thereby reducing the energetic barrier for the insertion of shorter luminal loops (≤50 amino acids). In mammalian

cells, EMC3 forms a functional core together with EMC6 (Figure 13 F), and the complete nine-subunit EMC associates with ribosome-nascent chain complexes to mediate co-translational insertion (Bai *et al*., 2020). The EMC is particularly important for the biosynthesis of multi-pass membrane proteins with short inter-helical loops (Bai *et al*., 2020).

Once nascent proteins enter the endoplasmic reticulum (ER) or become integrated into the ER membrane, they rely on a wide range of chaperone factors to ensure their correct folding and maturation. These chaperones act on both the soluble domains and the transmembrane segments of client proteins. For soluble proteins/domains, molecular chaperones such as members of the *Hs*p70 family (e.g., BiP) and lectin-like chaperones such as calnexin are crucial (Figure 12; Brodsky *et al*., 2011; Braakman *et al*., 2013). They prevent premature aggregation by shielding exposed hydrophobic residues until these residues are properly buried within the protein core. By contrast, membrane proteins face a distinct folding challenge: they must achieve correct structural and functional conformations within a hydrophobic lipid bilayer (Hedge *et al*., 2022; Hedge and Keenan, 2024). While the surfaces of mature membrane proteins are typically hydrophobic, key residues critical for activity—such as those mediating ion coordination or transport—are often hydrophilic. As a result, specialized chaperones are needed to recognize and transiently stabilize these hydrophilic residues until the protein has folded correctly (Hedge *et al*., 2022; Hedge and Keenan, 2024). In addition to soluble ER factors, several families of membrane-bound chaperones have been characterized that specifically regulate the folding, assembly, and function of transmembrane proteins.

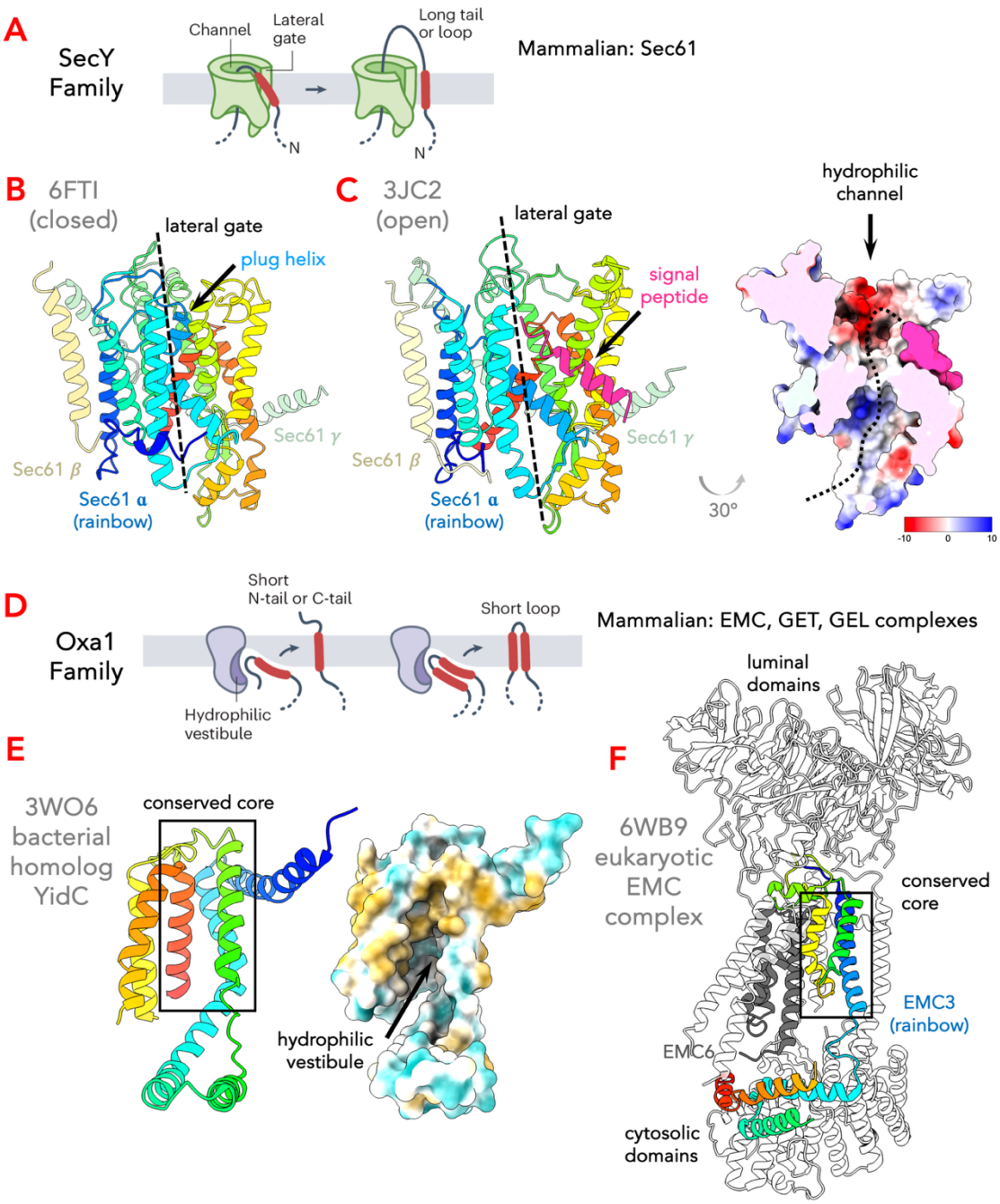


Figure 13. Sec61 Translocon and EMC Complex Facilitates Membrane Protein Insertion

A. Schematic diagram of SecY-family mediated transmembrane helix insertion (adapted from Hedge *et al.*, 2022). B. Structure of a mammalian Sec61 translocon in a closed state (PDB 6FTI, Braunger *et al.*, 2018), Sec61α coloured by rainbow (N- to C- terminus), Sec61β and Sec61γ coloured light yellow and light green respectively. C. Left: Structure of a mammalian Sec61 translocon in an open state bound to a signal peptide (PDB 3JC2, Voorhees *et al.*, 2016), Sec61α coloured by rainbow, Sec61β and Sec61γ coloured light yellow and light green respectively, signal peptide coloured in hot pink. Right: sliced view of the open Sec61 structure, surface coloured by electrostatic potential, surface of signal peptide coloured in hot pink. D. Schematic diagram of Oxa1-family mediated transmembrane helix insertion (adapted from Hedge *et al.*, 2022). E. Structure of a bacterial homolog of Oxa1 family insertase, YidC (PDB: 3WO6, Kumazaki *et al.*, 2014). Left: cartoon coloured by rainbow, box indicating the conserved 3-helix core. Right: surface coloured by hydrophobicity. F. Structure of *S. cerevisiae* EMC complex (PDB 6WB9, Bai *et al.*, 2020). EMC3 coloured by rainbow, black box indicating the conserved 3-helix core. EMC6 coloured in dark grey, rest of the subunits coloured in white.

Among the best-studied examples are the claudin family (such as TARPs) and the cornichon family (Figure 12), both of which serve as auxiliary subunits for AMPA-type glutamate receptors (Shanks *et al*., 2014; Twomey *et al*., 2016; Nakagawa *et al*., 2019; Gangwar *et al*., 2023). These 4-TM chaperones bind primarily to the outer helices of their client proteins, and in some cases also contribute extracellular domains—so-called "head" elements—that fine-tune receptor function. Notably, they do not merely assist folding but can influence the gating properties and trafficking of receptors, ultimately shaping their physiological activity (Twomey *et al*., 2016; Nakagawa *et al*., 2019; Gangwar *et al*., 2023).

Another example of a 4-TM ER-resident assembly factor is NACHO (TMEM35A). Unlike TARPs or cornichons, which remain associated with their mature client proteins, NACHO functions primarily at the stage of assembly. Since its discovery in 2016, NACHO has been shown to be indispensable for the biogenesis of pentameric ligand-gated ion channels such as nicotinic acetylcholine receptors (nAChRs). In the absence of NACHO, surface expression and current of the α7 nAChR are abolished (Figure 14 A), whereas overexpression of NACHO rescues receptor function (Gu *et al*., 2016).

More recently, NACHO has also been identified as an assembly factor for $GABA_a$ receptors. A cryo-EM structure of a NACHO–$GABA_a$ α1 complex (PDB 9H9E, Figure 14 B, Hooda *et al*., 2024, preprint) revealed that NACHO forms a compact four-helix bundle, directly shielding the interface of the α1 subunit that normally pairs with β or γ subunits (Figure 14 C). This prevents premature or incorrect assembly and ensures that α subunits remain competent for incorporation into functional pentamers. The author also showed that NACHO could interact with β1 subunit, and facilitate the correct assembly of $GABA_a$ hetero-pentamer (Figure 14 D). Unlike TARPs or cornichons, NACHO therefore does not fine-tune channel function but instead escorts client subunits until compatible partners become available.

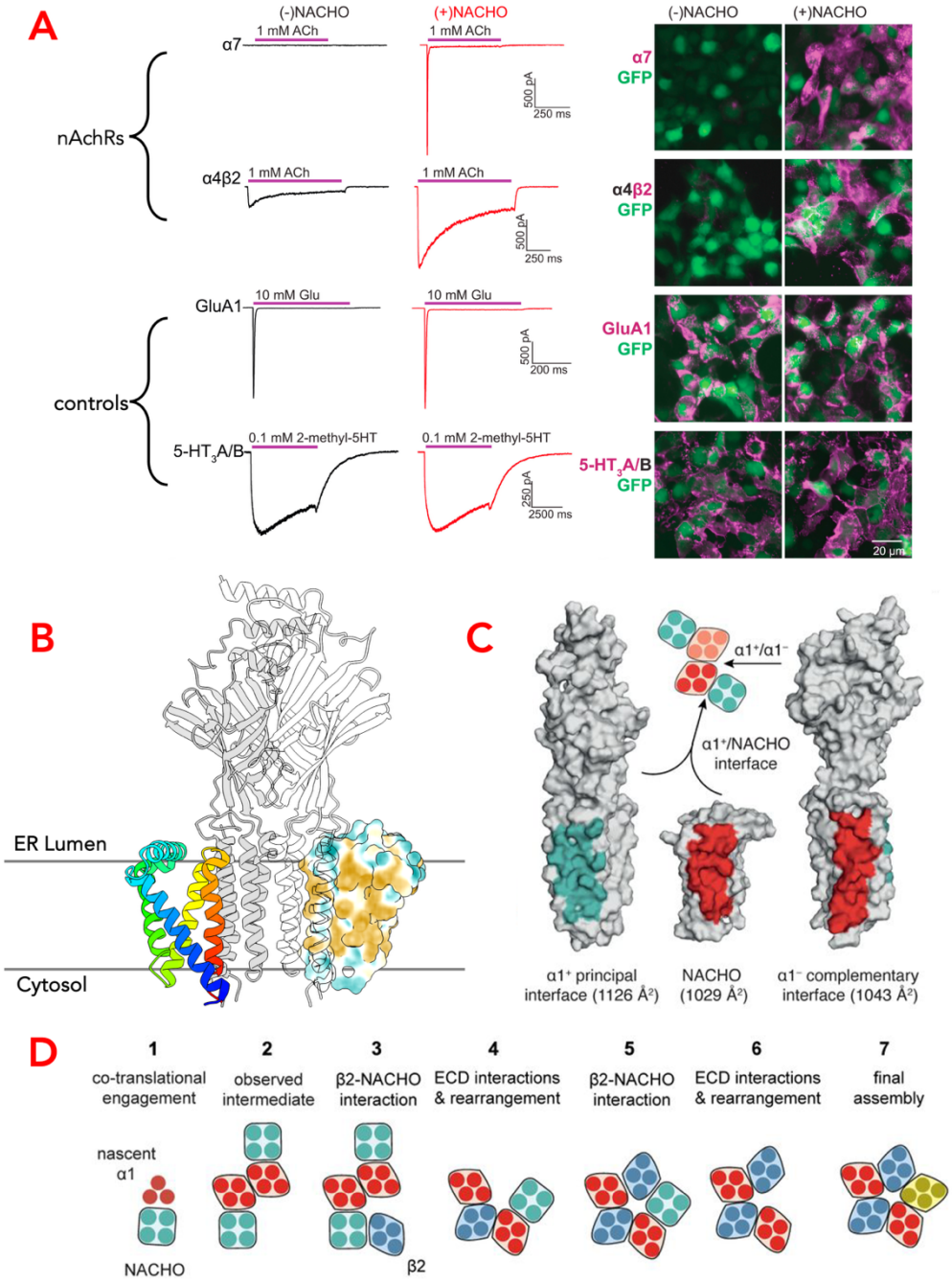


Figure 14. NACHO Facilitates Pentameric Ligand-gated Ion Channel Assembly

A. Left: whole-cell electrophysiology of HEK293Ts transfected with GFP-tagged ionotropic receptors, with or without NACHO; currents of nAchRs are significantly increased in the presence of NACHO, whereas currents of controls were not affected. Right: HEK293T cells transfected with GFP-tagged ionotropic receptors, with or without NACHO; co-expression of NACHO led to higher surface expression level of nAchRs, whereas the surface expression level of control ionotropic receptors was not affected. B. Structure of NACHO in complex with GABAa receptor subunit α1 in 2:2 stoichiometry (PDB 9H9E, Hooda *et al.*, 2024, preprint). α1 subunits are coloured in grey and white, NACHO molecules are either coloured by rainbow (cartoon representation), or by hydrophobicity (surface representation). C. From left to right: interaction surface of α1 to NACHO (teal), NACHO to α1 (red), α1 to α1 (red), adapted from Hooda *et al.*, 2024. D. Proposed mechanism of NACHO function in facilitating GABAa receptor assembly (NACHO in teal, α1 subunits in red, β2 subunits in blue, γ subunit in yellow), adapted from Hooda *et al.*, 2024.

A different class of membrane chaperones is represented by UNC93B1, which is specialized for endosomal Toll-like receptors (TLRs) that sense nucleic acids in endosomal compartments (Ishida *et al*., 2021). UNC93B1 is a multi-pass ER protein that adopts a fold resembling members of the major facilitator superfamily (MFS) of transporters, and multiple CryoEM structures were published in 2021 (PDB: 7C76, 7CYN, Ishida *et al*., 2021), showing how UNC93B1 utilised a transporter architecture for a different cellular function.

Structural studies have shown that UNC93B1 interacts directly with the single TM helix of TLR, as well as juxta-membrane segments, stabilizing them during trafficking from the ER (Figure 15 A-D, Ishida *et al*., 2021). For example, cryo-EM studies of UNC93B1 in complex with TLR3 (Figure 15 A) revealed the single transmembrane helix of TLR3 packed tightly against UNC93B1's TM3, stabilized by aromatic stacking interactions, while a juxta-membrane β-turn inserted into a complementary positively charged pocket (Figure 15 B). Mutagenesis of these UNC93B1 interface residues impaired binding, confirming that the observed contacts are essential for its chaperone activity (Figure 15 C).

Beyond folding, UNC93B1 also exerts control over receptor stoichiometry: TLR3 forms a 1:1 complex with UNC93B1, whereas TLR7 assembles as a 2:2 complex stabilized by both receptor–receptor and chaperone–chaperone interactions (Figure 15 D). Interestingly, structural modelling showed that UNC93B1 physically clashes with the active signalling conformation of its client TLRs, implying that dissociation of the chaperone is required before receptor activation can occur (Figure 15 E, Ishida *et al*., 2021). In this way, UNC93B1 acts not only as an escort for trafficking and folding but also as a gatekeeper, preventing premature signalling and ensuring that activation occurs only in the appropriate endosomal context. Within UNC93B1 itself, a cavity still retained, likely reminiscent of its evolutionary origin, although it was blocked by a long loop between TM helix 5 and 6, and inaccessible to solvent in all the published structures (Ishida *et al*., 2021).

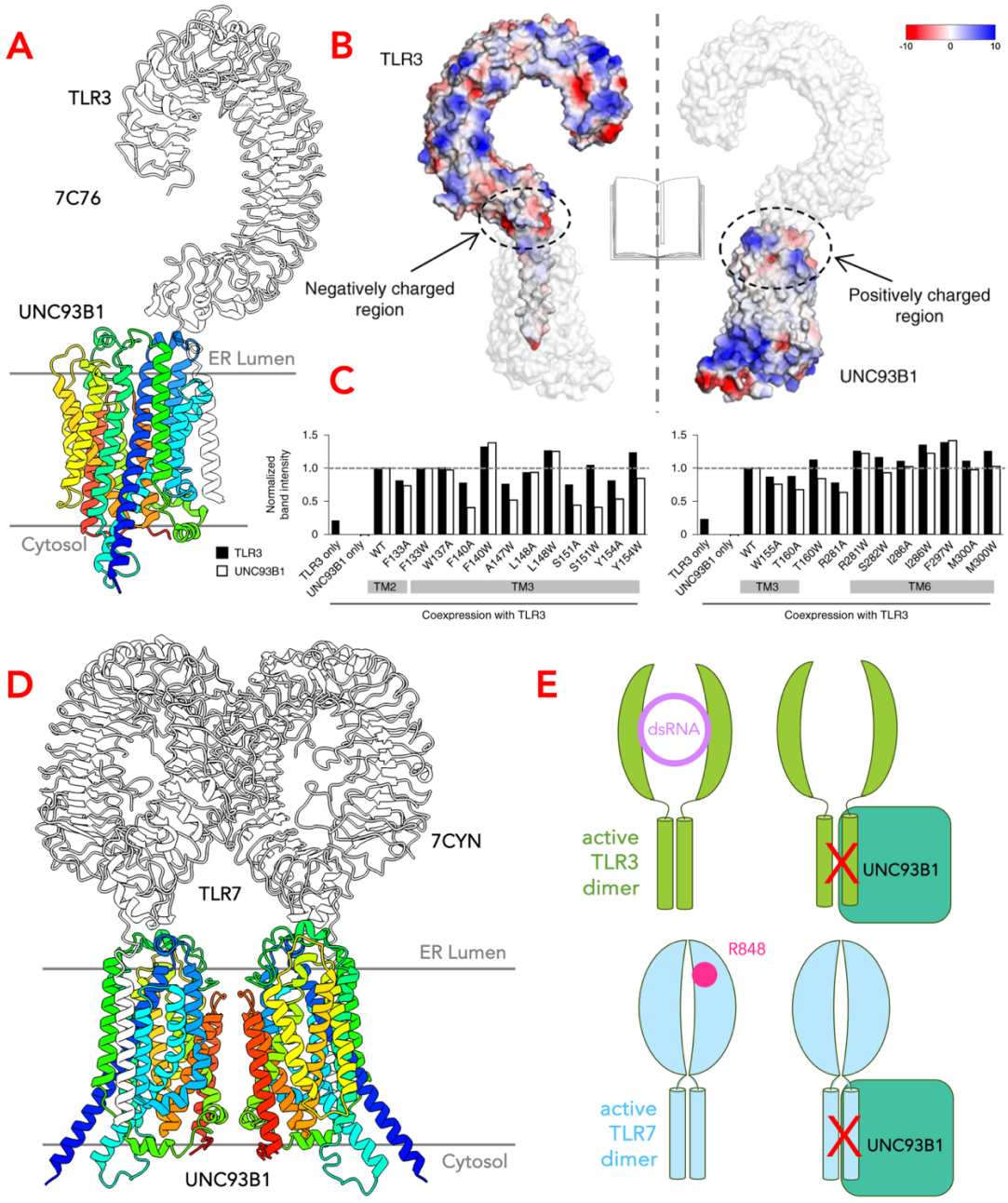


Figure 15. UNC83B1 is a Chaperone Utilising Transporter Architecture

A. Structure of human UNC93B1 in complex with TLR3 (PDB 7C76, Ishida *et al.*, 2021). UNC93B1 chaperone coloured by rainbow (N- to C- terminus), client protein TLR3 coloured white. B. Book view of UNC93B1/TLR3 complex in surface representation. Left: TLR3 coloured by electrostatic potential, UNC93B1 coloured white, transparency 50%. Right: UNC93B1 coloured by electrostatic potential, TLR3 coloured white, transparency 50%. C. Band intensity of co-immunoprecipitation of UNC93B1 mutants at the interface residues with FLAG-tagged TLR3. D. Structure of human UNC93B1 in complex with TLR7 in 2:2 stoichiometry (PDB 7CYN, Ishida *et al.*, 2021). UNC93B1 chaperones coloured by rainbow (N- to C- terminus), client protein TLR7s coloured white. E. Schematic representation of UNC93B1 function through preventing TLR dimerisation (TLR3: green; UNC93B1: teal; TLR7: light blue).

1.3.3 Machineries Involved in Membrane Protein Quality Control

While nascent membrane proteins are being inserted and folded within ER membrane, they are also surveyed by multiple quality control factors (Figure 12). The ER has evolved a highly sophisticated quality control system to safeguard the fidelity of protein synthesis, folding, and maturation. *At* its core are three major, interconnected pathways: the unfolded protein response (UPR), ER-associated degradation (ERAD), and ER-phagy. Together, these pathways operate in parallel to monitor nascent proteins, correct misfolded conformations, and eliminate irreparably damaged substrates (Chen *et al*., 2023; Christianson and Carvalho, 2022; Chino and Mizushima, 2023).

The UPR functions as the primary stress-sensing hub, detecting accumulation of misfolded proteins through three key sensors—PERK, IRE1α, and ATF6 (Hetz *et al*., 2020; Wiseman *et al*., 2022). Upon activation, these sensors attenuate protein translation, upregulate transcription of chaperones and folding factors, and, if stress remains unresolved, can trigger apoptotic signalling cascades (Wang and Kaufman, 2016; Hetz and Papa, 2018).

Proteins that cannot be refolded are instead targeted for removal via ERAD, where they are recognized by chaperones such as BiP or lectins like OS9, retro-translocated through ER membrane complexes, ubiquitinated, and subsequently degraded by the cytosolic proteasome (Christianson and Carvalho, 2022; Lemberg and Strisovsky, 2021).

Larger or aggregation-prone proteins that escape ERAD surveillance are cleared by ER-phagy, a selective autophagy pathway mediated by receptors such as FAM134B, RTN3L, and SEC62, which target aggregate-containing ER fragments for lysosomal degradation (Smith *et al*., 2018; Chino and Mizushima, 2023).

While soluble protein retro-translocation and UPR signalling have been studied extensively, relatively few dedicated quality control factors for membrane proteins have been identified.

Recent advances, however, have uncovered several specialized ER-resident factors that actively monitor and eliminate misfolded or mistargeted transmembrane proteins.

A striking example is the P5A-ATPase family, represented by Spf1 in yeast and ATP13A1 in humans, which functions as ATP-dependent "transmembrane helix dislocases." They remove mistargeted mitochondrial membrane proteins from the ER by extracting mis-inserted helices post-insertion (Figure 12; figure 16 A-C, McKenna *et al*., 2020).

In vitro reconstitution demonstrated that ATP13A1-deficient membranes accumulate mitochondrial TA proteins due to impaired TM extraction, a defect rescued by wild-type but not catalytically inactive ATP13A1 (Figure 16 A, McKenna *et al*., 2020). In human cells, ATP13A1 knockout caused ER mis-localization of mitochondrial TA proteins without affecting initial targeting, supporting a role in post-insertion removal (Figure 16 B, McKenna *et al*., 2020).

Cryo-EM structures of the yeast homolog Spf1 in inward-open and outward-open states (3.3–3.7 Å) was solved by the same authors (Figure 16 C, McKenna *et al*., 2020). The outward (ER lumen)-facing structure revealed an unusually large substrate-binding pocket accessible laterally from the membrane. This cavity, lined with hydrophilic, hydrophobic, and negatively charged residues, can alternately open to the cytosol during the transport cycle. In the outward-open state, Spf1 co-purified with an α-helical density spanning the bilayer, consistent with a bound TM substrate. These data support a model in which P5A-ATPases laterally capture mis-inserted, moderately hydrophobic helices, flip their short luminal segments across the membrane, and release them to the cytosol for degradation or correct targeting.

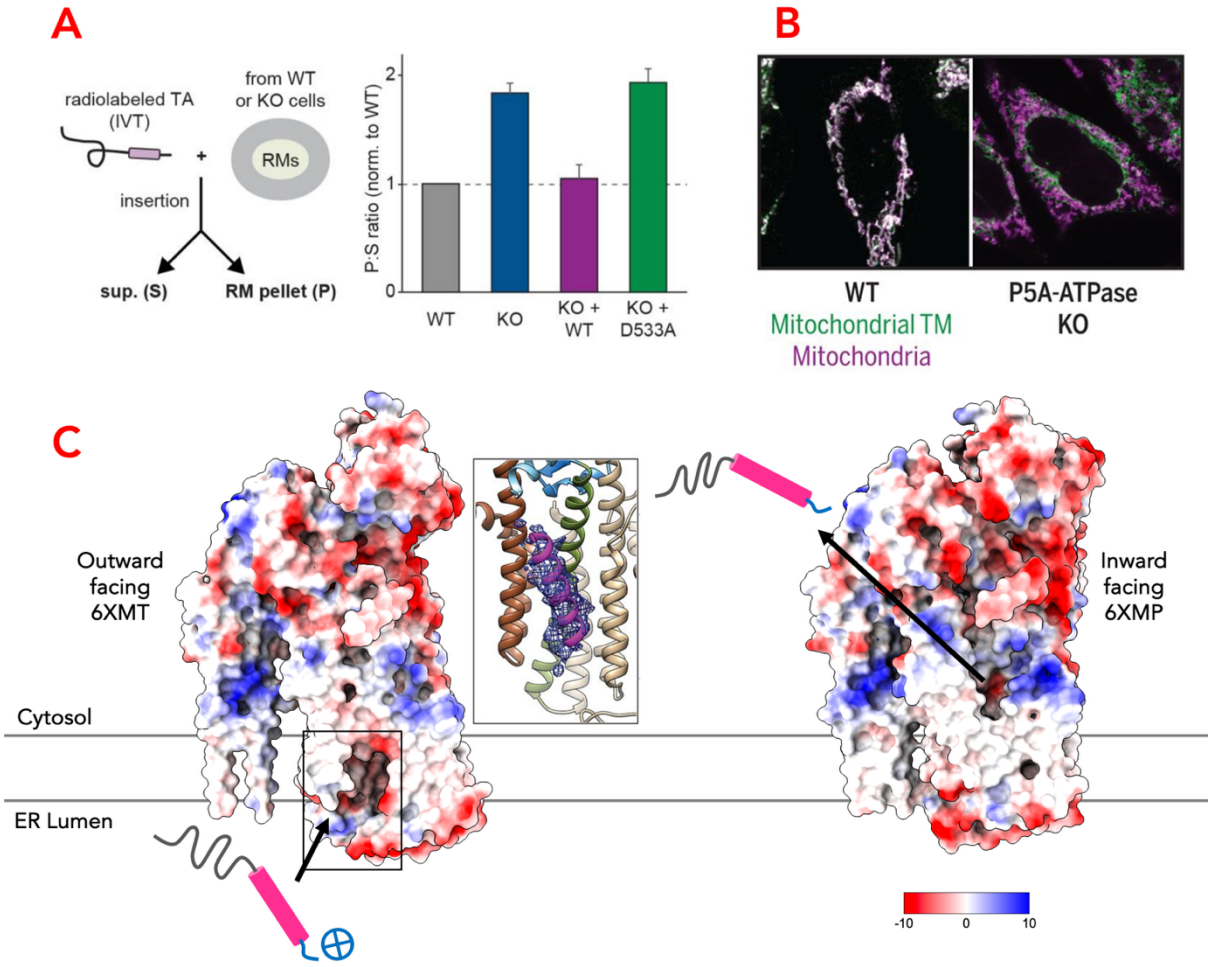


Figure 16. Spf1/ATP13A1 Facilitates Removal of Mislocated Transmembrane Helices

A. Spf1 was able to remove reconstituted tail-anchored (TA) proteins from rough ER microsomes (RMs), adapted from McKenna *et al.*, 2020. Left: Schematic diagram of experimental set up, where radiolabelled TA protein was translated into RMs derived from either WT Spf1 or ATP13A1 cells or knock-out cells followed by cross-linking. Then supernatant and pallet containing RMs were separated, and ratio of radiolabelled TA protein between pallet and supernatant calculated (P:S ratio). Right: WT Spf1 were efficient in removing radiolabelled Tas from RMs, and has significantly lower P:S ratio than KO cells; D533A mutant abolishes Spf1 activity and led to accumulation of TA proteins in RMs. B. Co-localisation of a mitochondrial transmembrane protein and mitochondria in the presence or absence of the P5A-ATPase ATP13A1 (McKenna *et al.*, 2020); knock-out of ATP13A1 led to mis-localisation of mitochondrial proteins. C. CryoEM structure of *S. cerevisiae* Spf1 in outward facing (left, PDB 6XMT) or inward facing conformation (right, PDB 6XMP), surface representation, coloured by electrostatic potential. TA protein in hot pink; arrows indicated the binding of positive residues and the release of TA protein on the other side of the membrane. Centre: density for putative substrate in CryoEM structure, adapted from McKenna *et al.*, 2020.

Another recently characterized machinery dedicated to membrane protein quality control is the ERAD-M complex (Rao *et al.*, 2021; Rao *et al.*, 2023), which mediates retro-translocation of misfolded membrane proteins (Figure 17 A-C). In 2021, single-particle cryo-EM studies revealed that the core of this machinery is a homo-tetrameric assembly of Derlin-1, which

forms a continuous membrane-spanning channel (Rao *et al*., 2021). The pore, reaching up to ~30 Å in width at its widest point and narrowing to 12–18 Å at a central constriction, is sufficient to accommodate transmembrane α-helices, allowing retro-translocation of partially unfolded clients. Lateral gates provide entry points for misfolded helices, while cytosolic and luminal loops act as regulatory covers for channel opening (Rao *et al*., 2021). These unique structural features suggest that, unlike soluble substrates, which use "half-channels" in ERAD-L pathways, membrane proteins are extracted through a complete channel (Wu & Rapoport, 2018; Schoebel *et al*. 2017).

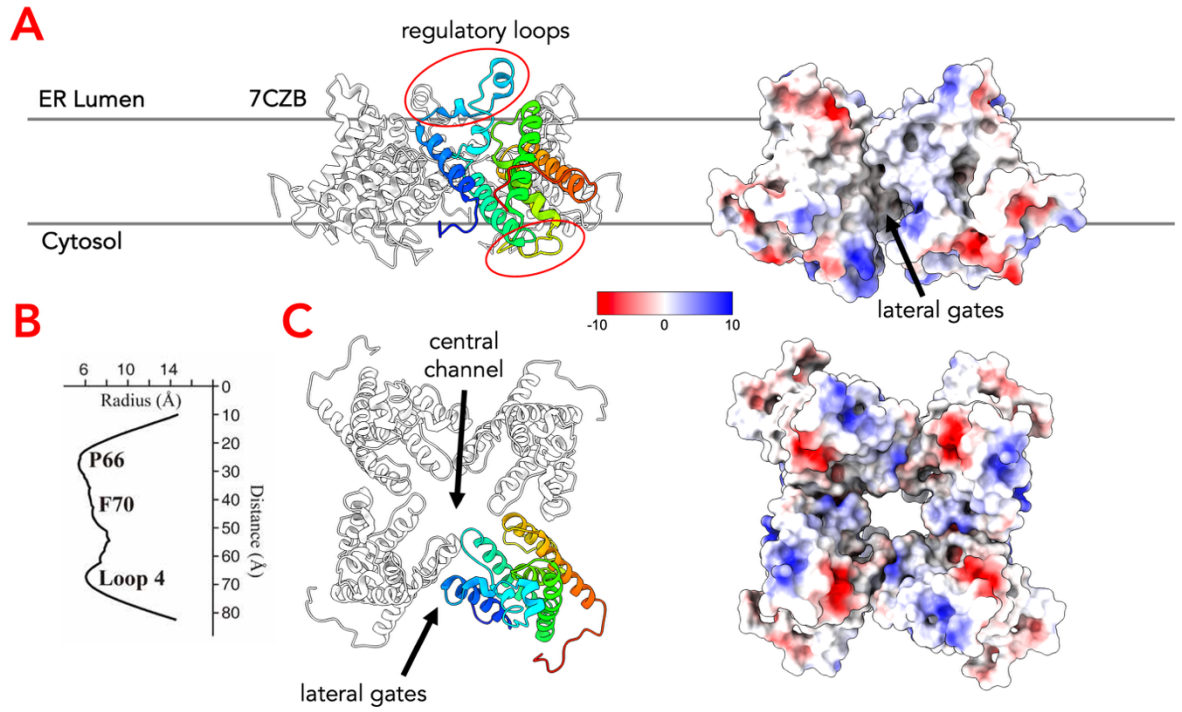


Figure 17. ERAD-M Channel Facilitates Retro-translocation of Membrane Proteins

A, C. Structure of human ERAD-M channel formed by tetrameric Derlin-1 (PDB 7CZB, Rao *et al.*, 2021). Left: one protomer coloured by rainbow (N- to C- terminus), other protomers coloured in white; red circles indicate regulatory loops, arrows indicate central channel and lateral gate. Right: surface representation, coloured by electrostatic potential. B. Channel radius of tetrameric Derlin-1, in Å, adapted from Rao *et al.*, 2021.

Beyond their individual activities, many of these ER factors act in larger networks of super-complexes that integrate protein synthesis, chaperoning, and quality control (Figure 18 A,D;

Sundaram *et al*., 2022; Yang *et al*., 2025; Haßdenteufel *et al*., 2018; Chen *et al*., 2023; Klose *et al*., 2025).

For example, the Sec61 translocon can associate with accessory complexes such as the GEL (GET- and EMC-like), PAT (protein associated with translocon), and BOS (back of Sec61) complexes, forming "multipass translocons" (Sundaram *et al*., 2022, Figure 18 B). These assemblies provide a semi-hydrophilic environment that accommodates nascent transmembrane helices during the synthesis of multipass membrane proteins. Sec61 has also been shown to interact with the dislocase ATP13A1, forming a localized synthesis–quality control pipeline (Figure 18 C). In this arrangement, atypical signal sequences—often unusually hydrophobic or lacking characteristic charges—are initially targeted to the ER by the signal recognition particle, sometimes in an incorrect orientation. ATP13A1 then dislocates these misoriented sequences and delivers them to Sec61 for proper insertion (Yang *et al*., 2025). In addition, Sec61 gating can be regulated by classical ER chaperones such as BiP and J-proteins (Haßdenteufel *et al*., 2018).

Similarly, the ER membrane protein complex (EMC) has functions that extend beyond its canonical role in transmembrane helix insertion (Figure 18 D). The "back" side of the EMC has been implicated in chaperone activity, acting as a holdase or stabilizing factor for client proteins. Indeed, structural studies in 2023 revealed that the EMC forms a complex with the voltage-gated calcium channel Cav1.2, where it plays a direct stabilizing role (Chen *et al*., 2023, Figure 18 E). Moreover, EMC has also been shown to form an evolutionarily conserved super-complex with ATP13A1 (Figure 18 F). A recent preprint (Klose *et al*., 2025) reported the structure of the human EMC–ATP13A1 complex, revealing a shared cavity that enables insertion and quality control to occur in close coordination. These findings emphasize that ER quality control factors rarely act in isolation, but instead operate as interconnected networks that integrate folding, surveillance, and degradation processes.

Given that many of these pathways and super-complexes have only been uncovered within the past 5-10 years, the landscape of membrane protein quality control is still far from complete. Numerous yet-undiscovered players are likely to exist. In the second part of this thesis, I will focus on one such newly identified factor, RUSF1 (Root UVB Sensitive Family 1), detailing its discovery and the efforts made to dissect its molecular function.

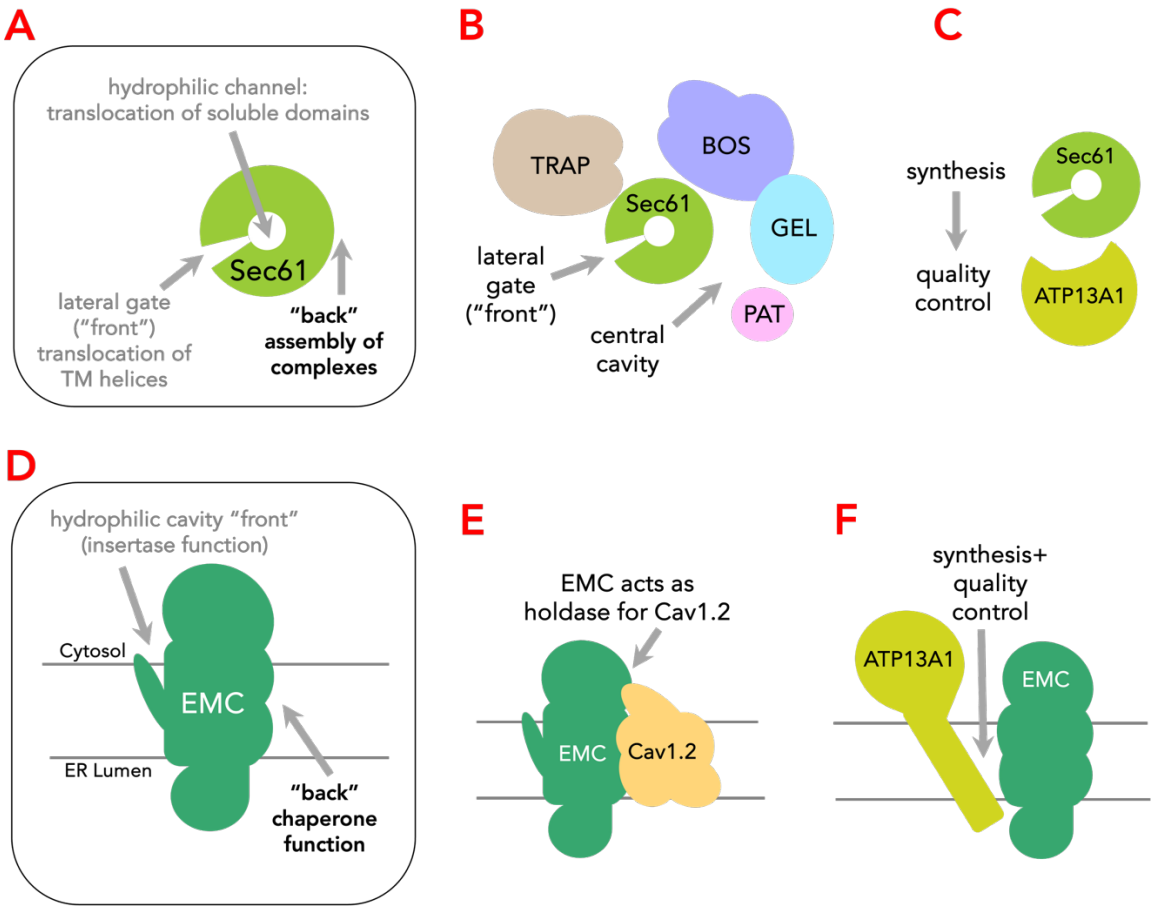


Figure 18. Sec61 and EMC are Involved in Various Super-complexes

A. Schematic diagram of Sec61 translocon viewed from ER lumen, with sites of translocation/insertion, and sites of multi-pass translocon assembly indicated by arrows. B. Schematic diagram of the multi-pass translocon, made up of multiple protein complexes around Sec61, providing a semi-hydrophilic environment for folding of multi-pass membrane protein. C. Sec61 translocon also interacts with dislocase ATP13A1, forming an insertion-quality control pipeline. D. Schematic diagram of EMC complex viewed from the lipid bilayer, with sites of insertion, and sites of client stabilisation indicated by arrows. E. Schematic diagram of Cav1.2 channel in complex with EMC. F. Schematic diagram of ATP13A1 in complex with EMC, also forming an insertion-quality control pipeline.

1.3.4 Investigating the Biogenesis and Assembly of $GABA_a$ Receptors

Pentameric ligand-gated ion channels (pLGICs), such as the γ-aminobutyric acid type A ($GABA_a$) and nicotinic acetylcholine receptors (nAChRs), play essential roles in neurotransmission throughout the nervous system (references). These receptors are composed of five subunits, often of different types, each containing a large extracellular domain, a transmembrane region with four transmembrane helices (TM1–TM4), and variable intracellular loops (Figure 19 A, Sarto-Jackson & Sieghart, 2008; Kim *et al*. 2021).

A central question in membrane biology has been how such structurally complex receptors are accurately inserted into the endoplasmic reticulum (ER) membrane and subsequently assembled into functional pentameric complexes, given the wide diversity of possible subunit combinations. Owing to their relatively small transmembrane domains, the presence of large extracellular regions, and the requirement for precise multimeric assembly, pLGICs have emerged as ideal model systems for dissecting the principles of multi-pass membrane protein biogenesis and assembly (Wu *et al*. 2024).

This area has been a major focus of the Hegde laboratory, with whom we collaborated for the present work. Their previous studies have illuminated the insertion pathway of the $GABA_a$ receptor α1 subunit, showing that the luminal domain and the first three transmembrane helices are co-translationally inserted into the ER membrane through the Sec61 translocon, whereas the final helix, TM4, which is preceded by an long cytosolic loop, requires the ER membrane protein complex (EMC) for insertion (Figure 19 B, Wu *et al*. 2024). These findings established a mechanistic framework for receptor subunit insertion and raised the question of how individual receptor subunits are stabilised within the ER membrane until they are correctly assembled into pentamers. Identifying the chaperones and assembly factors responsible for this process thus became a natural next step (Figure 19 C).

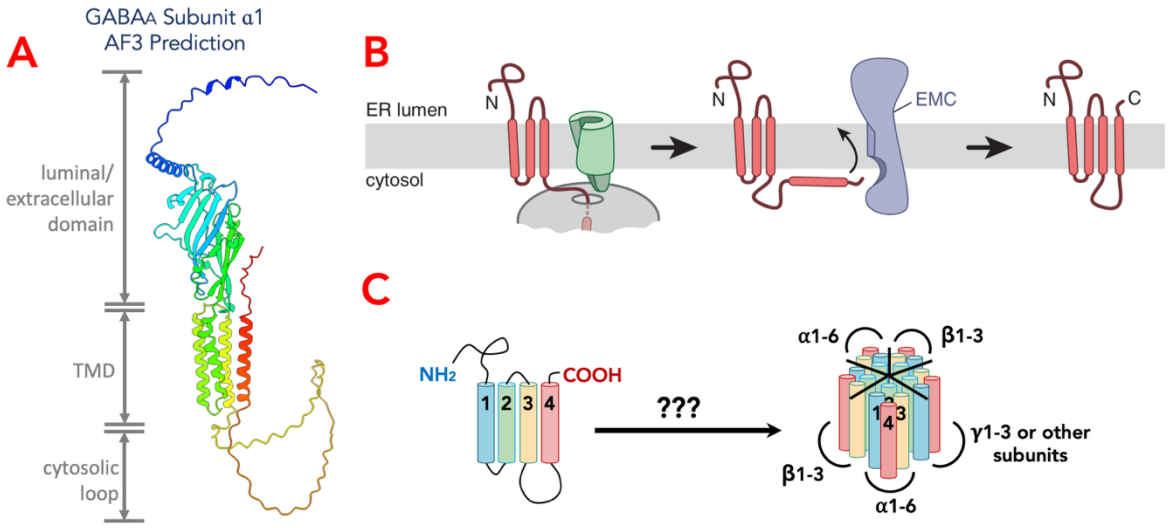


Figure 19. Biogenesis and Assembly of $GABA_a$ Receptors

A. Alphafold3 prediction of full length $GABA_a$ receptor α1 coloured by rainbow (N- to C- terminus), and domains labelled. B. Schematic diagram of the insertion of $GABA_a$ receptor α1 (red) cooperated by Sec61 translocon (green) and EMC (purple), adapted from Wu *et al.* 2024. C. Schematic diagram of pentameric $GABA_a$ receptor assembly from individual subunits.

To address this, a recent preprint (Hooda *et al.*, 2024) described the generation of a truncated $GABA_a$ receptor α1 construct, termed "mini-α1," which lacks the extracellular domain but retains the transmembrane region. This construct was expressed in ER-derived microsomes from both canine pancreas and HEK293 cells and subsequently affinity-purified (Figure 20 A). Mass spectrometry analysis of co-purifying proteins revealed the set of factors that specifically interact with the α1 subunit's membrane domain prior to receptor assembly. As a control for specificity, interaction partners of a thermostabilised β1-adrenergic receptor variant (β1ARΔCL3), a monomeric GPCR, were also analysed in parallel (Figure 20 B).

The mini-α1 proteomic screen identified a wide range of ER-associated proteins, many with established or putative roles in protein folding, quality control, and trafficking (Hooda *et al.*, 2024, preprint). These included subunits of the EMC complex, TMED proteins, NACHO and the putative lipid scramblase CLPTM1L, among others (Figure 20 B). Importantly, NACHO (also known as TMEM35A), an ER-resident membrane protein recently identified as a

biogenesis factor for nAChRs, was also found, reinforcing the validity of the approach. The author subsequently verified the interaction between NACHO and $GABA_a$ receptor α1 full-length construct, and a CryoEM structure of the complex was solved, which was discussed in earlier chapters (**Chapter 1.1.3**, McKenna *et al.*, 2020). However, the most prominent interactor in the dataset was a previously uncharacterised protein, RUSF1 (also known as C16orf58), whose cellular role had not yet been defined.

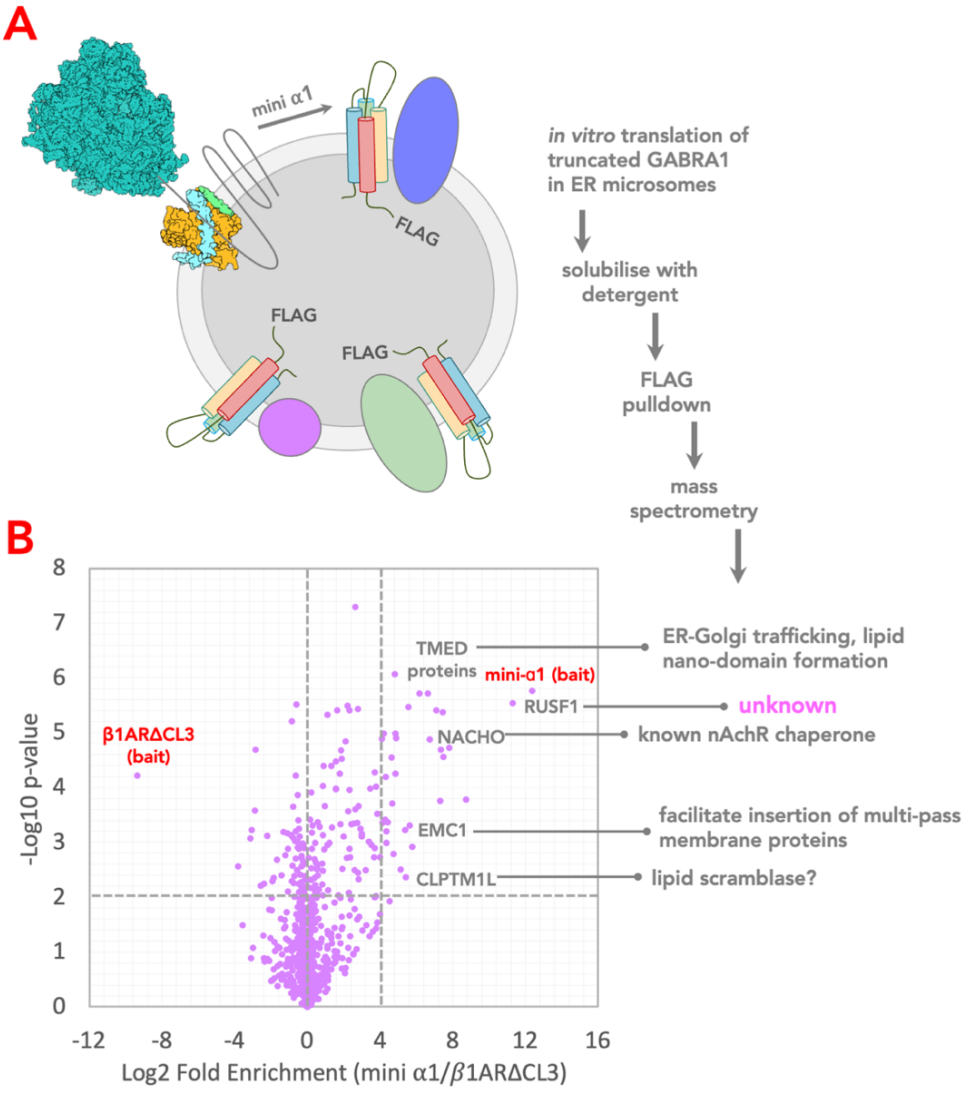


Figure 20. Identification of $GABA_a$ Chaperones

A. Schematic diagram of the identification of mini-α1 chaperones from ER microsomes. B. mini-α1 interactors identified through affinity capture mass spectroscopy, presented on a volcano plot, with various hits annotated.

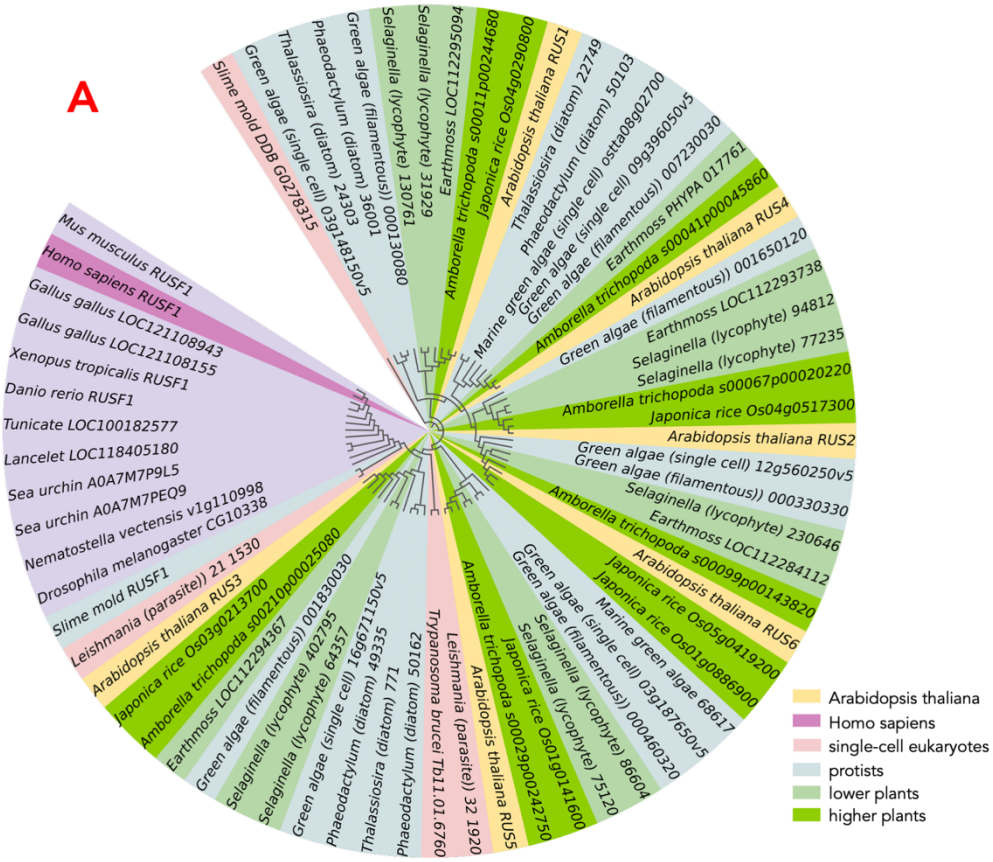
1.3.5 RUSF1 — A Molecular Chaperone with Transporter Architecture?

Bioinformatic analysis provided the first clues to RUSF1's identity. Multiple sequence alignments classified RUSF1 as a member of a family of proteins containing a conserved domain known as DUF647 (Pfam PF04884: "Vitamin B6 photo-protection and homoeostasis domain"). Notably, proteins harbouring this domain are exclusively found in eukaryotes. Phylogenetic analysis across ~30 representative eukaryotic proteomes, including human, mouse, chicken, zebrafish, Drosophila, yeast, green algae, and Arabidopsis, revealed clustering patterns among family members, although the tree was not rooted to infer a true evolutionary trajectory (Figure 21 A). Human RUSF1 grouped closely with other animal RUS proteins, consistent with evolutionary expectations. Generally, a single RUSF1 orthologue is encoded per animal genome, though alternative isoforms have been described in humans and mice. Interestingly, multiple gene entries exist for chicken (*Gallus gallus*) and sea urchins, although the validity of these annotations remains uncertain.

No RUSF1 homologues could be identified in budding or fission yeast (*S. cerevisiae* or *S. pombe*). In contrast, plants and protists were found to encode multiple genes containing the DUF647 domain. For instance, *Arabidopsis thaliana* possesses six distinct homologues (*At*RUS1–*At*RUS6), although only *At*RUS3 shows clear similarity to animal RUS proteins. Proteomes of other higher and lower plants, as well as green algae, displayed similar diversification, leading to plant-specific phylogenetic clusters.

Comparative sequence analysis reinforced these distinctions (Figure 21 B). Among the Arabidopsis paralogues, *At*RUS3 is the most similar to animal RUS proteins (the only plant member that clusters near the animal clade), showing ~35% identity to human RUSF1, whereas *At*RUS1 and *At*RUS2 are more divergent at ~27.1% and ~24.7% identity to human RUSF1, respectively. *At*RUS1 and *At*RUS2 themselves share only ~26% identity. By contrast,

conservation among animal orthologues is markedly higher: zebrafish RUSF1 is ~50% identical and mouse RUSF1 ~80% identical to the human protein.



B

	Human RUSF1	Mouse RUSF1	Zebrafish RUS3	*Arabidopsis* RUS3	*Arabidopsis* RUS1	*Arabidopsis* RUS2
Human RUSF1	100	80.26	49.77	35.14	27.11	24.69
Mouse RUSF1	80.26	100	50.23	36.02	26.79	26.05
Zebrafish RUSF1	49.77	50.23	100	36.34	25.06	25.59
Arabidopsis RUS3	35.14	36.02	36.34	100	27.67	25.50
Arabidopsis RUS1	27.11	26.79	25.06	27.67	100	26.06
Arabidopsis RUS2	24.69	26.05	25.59	25.50	26.06	100

Figure 21. Clustering and Sequence Identities of RUSF1 Homologs

A. Phylogenetic tree of DUF647 domain-containing proteins across 31 representative eukaryotic species, branching only representative of clustering, not evolutionary trajectory. Individual genes coloured by custom colour key. B. Sequence identity of various RUS family members: human RUSF1, mouse RUSF1, Zebrafish RUSF1, Arabidopsis RUS1, Arabidopsis RUS2, and Arabidopsis RUS3.

Human RUSF1 is a protein composed of 468 amino acids, with a predicted molecular weight of approximately 51 kDa (Uniprot Q96GQ5). Sequence analyses and domain annotations suggested that the protein contains both a transmembrane region (DUF647 domain) and a soluble domain, which is supported by AlphaFold structural predictions. Specifically, AlphaFold predicts the presence of six transmembrane helices, with the soluble portion of the protein formed predominantly by C-terminal residues. Interestingly, several β-strands from the N-terminal region are also predicted to contribute to the soluble domain (Figure 22 A, Supplementary Figure 8).

More intriguingly, both AlphaFold2 and AlphaFold3 confidently predict that RUSF1 forms homodimers (Supplementary Figure 8). Although the models differ slightly in their conformations, both identify a large and stable dimer interface within the transmembrane region. Notably, the N-terminal β-strands mentioned before now interact with the C-terminal domain of the opposing protomer in a domain-swapped arrangement, further reinforcing the rigidity of the dimeric structure (Figure 22 A).

Closer examination of the 12-transmembrane (TM) architecture of the predicted RUSF1 dimer reveals a striking similarity to the fold observed in the Major Facilitator Superfamily (MFS) of transporters (Figure 22 B). Like RUSF1, most MFS transporters are integral membrane proteins with a 12-TM organization. Structural comparison between RUSF1 and a monomeric MFS transporter PepT2 (PDB ID: 7NQK, Parker *et al*., 2021) indicates that the helical arrangement of each RUSF1 protomer corresponds closely to the N- and C-terminal bundles of PepT2.

This resemblance suggests that RUSF1 may represent one of the few cases where a protein with an MFS-like fold does not primarily serve as a transporter. Another example is UNC93B1 (mentioned in **Chapter 1.1.3**, McKenna *et al*., 2020), a molecular chaperone that was shown to adopt an MFS-like fold despite functioning in protein trafficking rather than substrate

transport. If RUSF1 is validated as a chaperone of $GABA_a$ receptors, it would therefore represent only the second known example of a molecular chaperone utilizing an MFS-type structural architecture.

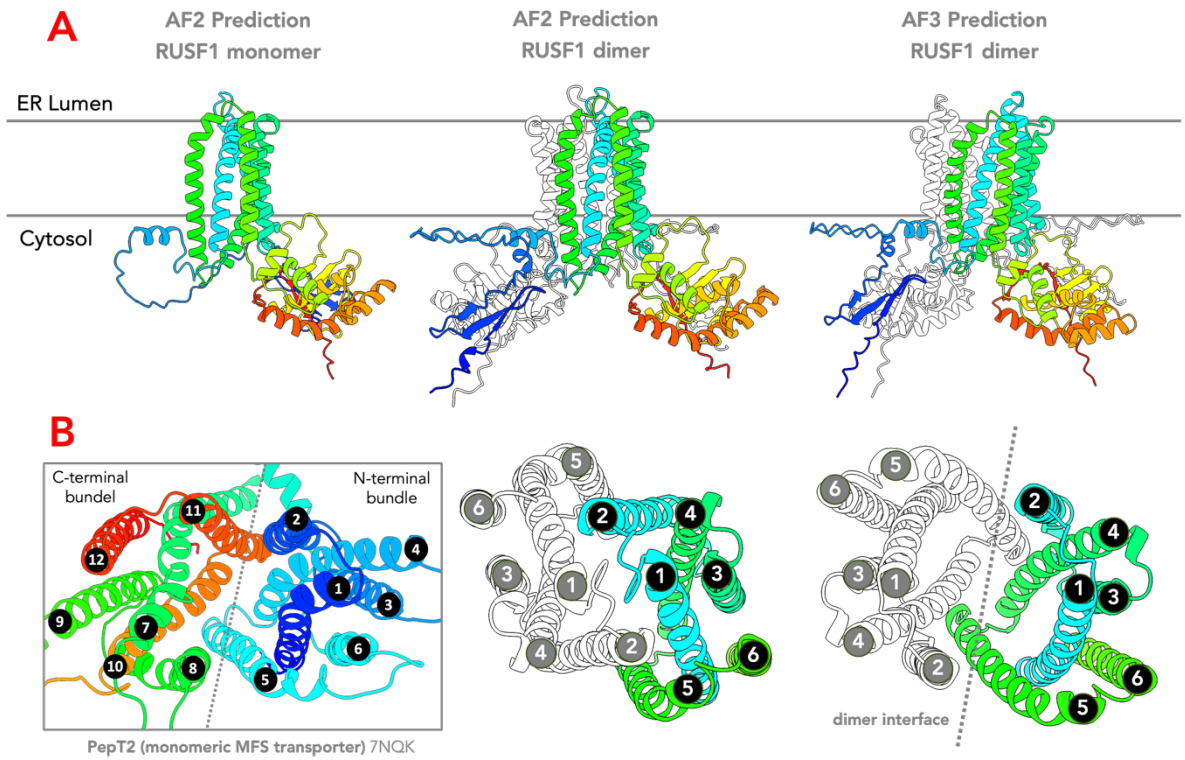


Figure 22. AlphaFold Predicted RUSF1 Homodimer to Adopt MFS Architecture

A. AlphaFold predictions of monomeric or dimeric RUSF1 structures. In dimeric RUSF1 predictions, one protomer is coloured by rainbow (N- to C- terminus), the other protomer coloured in white. B. Comparison of dimeric RUSF1 prediction with a structure of monomeric MFS transporter PepT2 (Parker *et al.*, 2021). Left: structure of PepT2 coloured by rainbow (N- to C- terminus), helix number annotated, dotted line indicates bundle interface. Middle & right: AlphaFold predictions of dimeric RUSF1 (soluble domains not shown), helix number annotated, dotted line indicates dimer interface.

The possibility that RUSF1 employs such a fold, while existing as a dimer, also raises intriguing questions both about its biological function and its evolutionary origins.

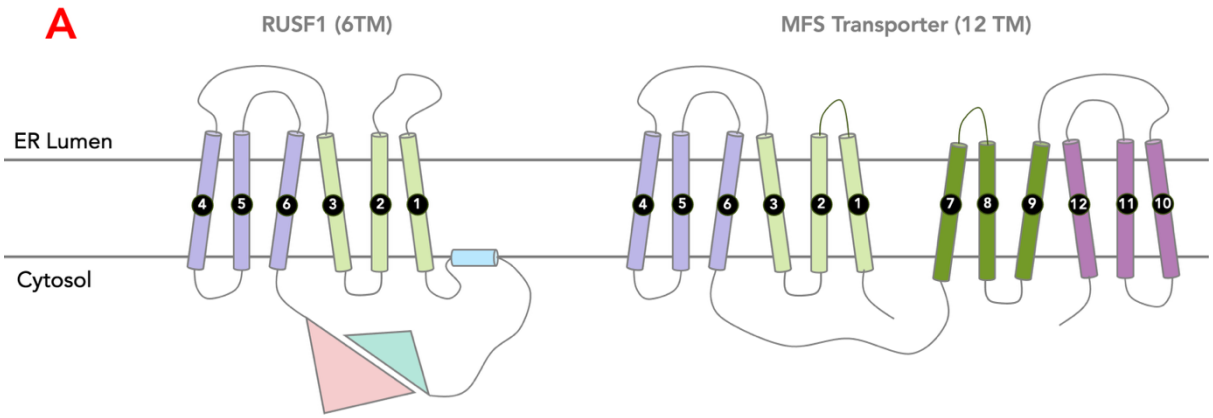
The Major Facilitator Superfamily is one of the largest families of membrane transporters, widely distributed across all kingdoms of life (Madej *et al.*, 2013; Varela *et al.*, 2023). These proteins typically catalyse proton-coupled symport or antiport and share a common 12-TM fold composed of two inverted six-helix bundles (Figure 23 A). This arrangement is thought to have

arisen through intragenic duplication of a primordial three-helix bundle, ultimately giving rise to pseudo-symmetric N- and C-terminal domains that enclose a central transport cavity (Madej *et al*., 2013). While MFS proteins display poor sequence similarity across subfamilies, they nevertheless preserve conserved three-dimensional arrangements of functional residues. For instance, despite having only ~10% sequence identity, the sugar/proton symporters LacY and FucP exhibit nearly identical positioning of key residues involved in sugar binding and proton coupling (Madej *et al*., 2013).

Despite the wide sequence divergence within the MFS, several motifs are consistently conserved (*Varela et al., 2023,* Figure 23 B). One of the most universal is Motif A, defined as G(X3)D(R/K)XGR(R/K), located in the loop between TM2 and TM3. This motif plays an important role in gating and bundle interface stability (Henderson *et al*., 1990; Griffith *et al*., 1992). Analysis of representative RUS family sequences reveals partial retention of this motif, as G(X3)D (Figure 23 C). In particular, animal RUS proteins show a highly conserved sequence, GSKLD, at this position, implying a functional or structural importance. AlphaFold predictions suggest that TM2 of RUSF1 contributes to the dimer interface, though this sequence itself is not directly at the dimer interface, rather, it is positioned at the TMD-soluble domain interface, so it might instead contribute to the overall stability of the protein.

A second point of resemblance is found with Motif D2, which has been implicated in determining substrate binding specificity in MFS transporters (Paulsen *et al*., 1996). In canonical MFS proteins, this motif is located at the C-terminal end of TM1, with a consensus sequence gIgl(X2)PvlP, with the residues in lower case to be less than 70% prevalent (Figure 23 B). Although less well-conserved in RUS family members, several features are retained. In human and mouse RUSF1, for example, the sequence GIGV is observed in human and mouse (on a flexible loop between TM1 and TM2 predicted by AlphaFold), while other homologues

preserve at least one of the conserved glycines and the overall hydrophobic character of the motif (Figure 23 C).



B

Motifs (Consensus Sequences)	Locations	Functions
Motif A G X3 D R/K X G R R/K	Loop between TM2 and TM3 (Loop2–3)	Gating, bundle interface stability, ion gradient sensor, conformational switching
Motif B L X3 R X2 q G X3 a R X2 Q G	TM4	Proton binding and transport, energization, conformational changes
Motif C G X8 G X3 G P X2 G G	TM5 of antiporters	Antiport, conformational changes, permeability barrier, substrate binding and specificity, molecular hinge, bundle interface forming
Motif D2 g I g l X2 P v l P	C-terminal end of TM1 in proteins with 12-TM	Postulated in substrate binding and specificity
Motif G G P L l g	TM11 of 12-TM proteins	Evolutionary duplication of motif C, postulated to play similar functional roles during transport

C

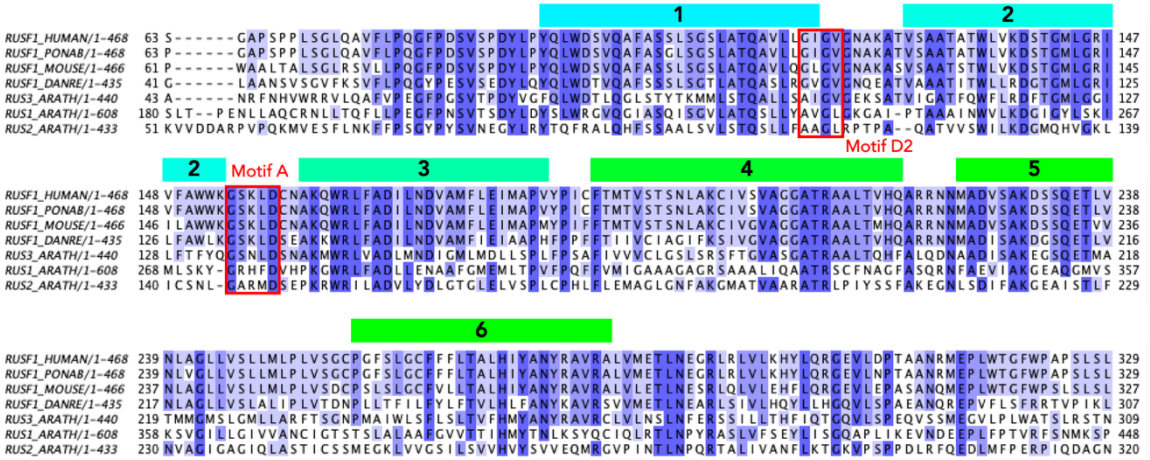


Figure 23. RUSF1 and the Evolution of MFS Transporters

A. Topology of RUSF1 (left) and an MFS transporter (right), helix number annotated. B. Conserved motifs of MFS transporters, their locations on transmembrane helices, and proposed functions of the motifs. Motifs found in RUSF1 in red. C. Sequence alignment of multiple RUS family members: human RUSF1, orangutan RUSF1, mouse RUSF1, Zebrafish RUSF1, Arabidopsis RUS1, Arabidopsis RUS2, and Arabidopsis RUS3. Residues coloured by sequence similarity. Transmembrane helices annotated above sequences, and red box indicates motifs conserved with MFS transporters.

In addition, the AlphaFold predictions of RUSF1 showed two distinct conformations. Namely, AF2 predicted an “inward-open” structure, where the ER lumen-facing side of the homodimer was tightly closed, through various hydrophobic interactions formed by TM1, TM2 and the loop between on both protomers forming the “extracellular gate”, reminiscent of an MFS transporter; meanwhile, AF3 predicted an “outward-open” structure, where the cytosolic-facing side of the transmembrane domain was closed, also mainly driven by hydrophobic interactions, between TM4 and TM5 of both protomers. Unlike MFS transporters, which usually possess salt bridges at the intracellular and extracellular gates which is responsible for substrate recognition or proton coupling (Parker *et al.,* 2017; Stelzl *et al.,* 2014), the lack of charged residues in the “gates” of RUSF1 suggests that the activation of the protein might require larger scale factors, such as protein-protein interaction, or membrane curvature/thickness change.

Taken together, the strong architectural similarity of RUSF1 to MFS transporters, along with partial retention of key MFS motifs, suggests that RUSF1 may have originated from the evolutionary lineage of the MFS superfamily. Unlike canonical transporters, however, which combine N- and C-terminal six-helix bundles within a single monomer to produce pseudo-symmetric folds, RUS family proteins appear to achieve symmetry through homo- or heterodimerization. This shift in structural organization could be central to their function, with monomer-dimer dynamics potentially playing a regulatory role.

1.3.6 Plant RUS Homologs Form Hetero-dimers and are Involved in Root UVB Sensing

As mentioned before, unlike most eukaryotes, which typically encode a single RUS protein, plants contain six distinct homologs of the RUS family. These paralogs have undergone functional diversification, with some members playing essential roles in development and

others contributing to specialized physiological processes. For example, Perry *et al*. (2021) systematically investigated multiple RUS paralogs in *Arabidopsis thaliana* and found that RUS6 is essential for embryogenesis. Homozygous loss-of-function mutants of RUS6 were inviable, displaying severe defects with embryos arresting at early stages of development. Complementation with a RUS6–GFP fusion protein restored viability, confirming the functional requirement of RUS6. Moreover, expression analysis revealed broad localization of RUS6 during embryogenesis. By contrast, other paralogs such as RUS3, RUS4, and RUS5 were dispensable under standard laboratory conditions, suggesting that different RUS proteins have evolved non-redundant, specialized functions within plants.

The first RUS genes to be identified, however, were RUS1 and RUS2, which gave the protein family its name — *Root UV-B Sensitive (RUS)*. The physiological relevance of these proteins was first uncovered by Tong *et al*. (2008), who demonstrated that *Arabidopsis* seedlings are capable of sensing extremely low intensities of UV-B light through their roots, adjusting growth patterns accordingly. A single recessive mutation in *RUS1: rus1-1 (or rus1-2),* resulted in severe developmental arrest under UV-B exposure, characterized by stunted root growth while leaf development remained unaffected (Figure 24 A). Importantly, shielding the roots from UV-B rescued this phenotype, directly linking RUS1 function to UV-B sensing. Subsequently, Leasure *et al*. (2009) cloned *RUS2*, discovered that it was co-expressed with *RUS1*, and demonstrated that mutations of either gene (rus1-2, rus2-1, or double mutant) produced identical developmental phenotypes (Figure 24 A). This finding strongly suggested that the two proteins function together within the same pathway.

Further molecular analyses revealed that RUS1 and RUS2 physically interact to form a heterodimeric complex (Leasure *et al*., 2009). Yeast two-hybrid assays confirmed that full-length RUS1 and RUS2 associate with each other, but neither protein self-associates. Truncation studies mapped the interaction interface to the conserved DUF647 domain,

implicating this region as critical for heterodimerization (Figure 24 B, upper two panels). This assay further highlighted the importance of several interaction: substitution of four highly conserved residues in RUS1 found in every single RUS family member (K281G, E298G, N342G, and K349G) disrupted binding to RUS2 to varying degrees (Figure 24 B, upper two panels). Specifically, K281G and K349G completely abolished interaction, whereas E298G and N342G markedly weakened it. When these mutant alleles were introduced into *rus1-1* mutant plants, surprisingly, partial phenotypic rescue was observed, suggesting the complexity of the system and the potential involvement other regulatory (Figure 24 B, lower panel).

Building on these findings, Tong *et al*. (2021, preprint) identified an additional regulatory layer. In a genetic knock-out screen, the authors discovered that the RUS1–RUS2 complex could be inhibited by certain aminotransferases involved in vitamin B6 metabolism. Vitamin B6 (especially the pyridoxal 5'-phosphate vitamer) is a vital cofactor for various metabolic activities during early seedling development, which RUS1 and RUS2 were also found have an effect on.

Among the aminotransferases discovered in the screen, ASP2 was found to physically interact with RUS1 (Figure 24 C), suggesting that the RUS1–RUS2 heterodimer acts not merely as a sensor but as part of a broader metabolic and signalling scaffold.

Collectively, studies in plants highlight the critical role of monomer-dimer dynamics in RUS protein function. The requirement for heterodimer formation, along with the demonstrated interactions of RUS proteins with metabolic enzymes, suggests that members of the RUS family might function primarily through protein–protein interactions rather than by transporting substrates directly, as is typical of their distant evolutionary relatives, the MFS transporters. These findings provide important context for the study of RUS proteins in other organisms, including humans, where the sole homolog RUSF1 may likewise rely on homo-dimerization and protein interactions to carry out its biological role.

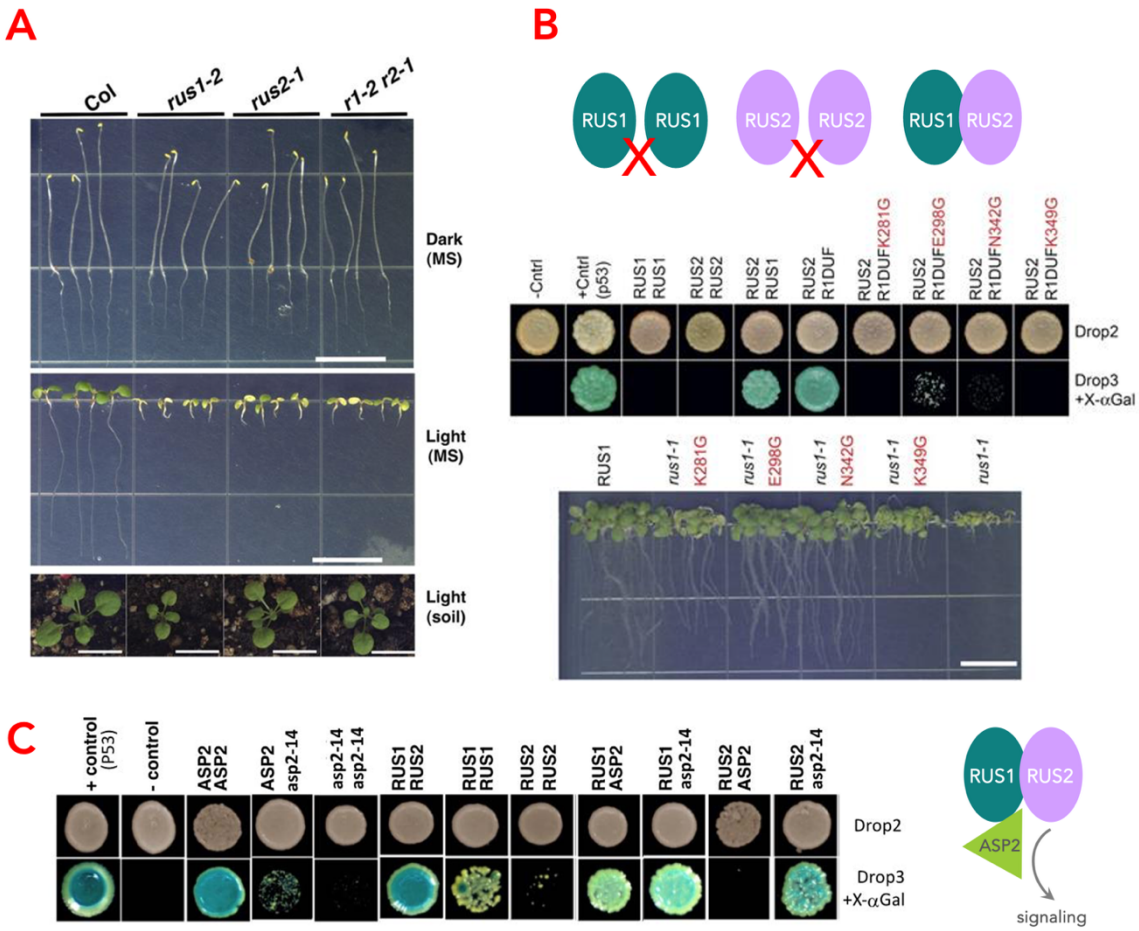


Figure 24. *Arabidopsis* RUS1 and RUS2 Forms Heterodimer

A. Root and leaf phenotypes of either wild type *Arabidopsis thaliana*, or mutants rus1-2, rus2-1, and double mutant r1-2 r2-1, under conditions with or without light. B. Top: schematic diagram of RUS1 RUS2 interaction. Middle: yeast two-hybrid assay of Arabidopsis RUS1 and RUS2. Multiple RUS1 mutants were also tested which disrupted the dimer interface. Down: introducing the same mutations to the rus1-1 mutant partially rescued root phenotypes under UVB exposure. C. Left: yeast two-hybrid assay of aminotransferase ASP2 with Arabidopsis RUS1 and RUS2. Right: Schematic diagram of ASP2 using RUS1/RUS2 as a docking platform for downstream signalling.

1.3.7 Aims of RUSF Study

The central aim of this part of the thesis is to advance our understanding of the structure and function of human RUSF1. Specifically, I aim to:

1. Determine the molecular architecture of RUSF1 using cryo-electron microscopy (Chapter 5.1-5.2).
2. Complementary biochemical approaches will be employed to probe the monomer–dimer dynamics of human RUSF1 (Chapter 5.3).
3. In addition to structural characterisation, I aim to investigate the functional interactions of RUSF1 with client proteins (Chapter 6.1-6.2). In particular, evidence points to a role for RUSF1 in regulating $GABA_a$ receptor biogenesis, and I will test this hypothesis by probing its physical interaction with the $GABA_a$ receptor α1 subunit. I'll also attempt to identify other potential client proteins of RUSF1.
4. I aim to generate RUSF1-specific synthetic nanobodies, which could be used in immunofluorescence studies of RUSF1 in cells (Chapter 6.3-6.4).

By integrating these structural and functional insights, I hope to contribute to a broader understanding of how RUSF1 participates in maintaining endoplasmic reticulum homeostasis, potentially revealing new principles by which MFS-like folds are repurposed as chaperones.

2. Materials and Methods

2.1 Molecular Biology

2.1.1 Construct Cloning

A variety of plasmid constructs were used across the CLC7 and RUSF1 projects, as summarized in the accompanying table (Table 1). RUSF1-FLAG in pcDNA3.1, HA-GABRA1, and NACHO-TST in pcDNA5.1 were generously provided by the Hedge lab (Laboratory of Molecular Biology, Cambridge).

Table 1. Constructs Used within the Scope of the Thesis

Vectors	Proteins	Expression System	Purpose
pLexM	CLC7-TEV-GFP-His	Mammalian Expression	Test expression and purification
	CLC7-His (and mutant Y715C)		Purification for structural studies
	CLC7-Avi-His		Purification for sybody selection
	OSTM1-Flag		Mammalian expression and purification
	RUSF1-Avi-Flag		Purification for sybody selection
	RUSF1-Myc (and mutants K162G, E179G, N223G, K230G)		Immunoprecipitations
	RUSF1-TST		Immunoprecipitations
	GABRA1-His		Immunoprecipitations
	GABRA1-TMD-His (Δ2-249)		Immunoprecipitations
	HA-GABRA1-lumD (Δ250-455)		Immunoprecipitations
	HA-GABRA1-TMD (Δ2-249)		Immunoprecipitations
PCDNA3.1	RUSF1-Flag (and mutants K162G, E179G, N223G, K230G)		Purification for structural studies or immunoprecipitations
	TMEM126A-Myc		Immunoprecipitations
PCDNA5.1	HA-GABRA1		Immunoprecipitations
	Flag-GABRA1		Immunoprecipitations
	NACHO-TST		Immunoprecipitations
pDx_init	Synthetic nanobodies	*E.coli* Expression	Phage Displays in the Sybody Pipeline
pSb_init	Synthetic nanobodies		Sybody expressions and purifications

Full-length constructs were subcloned into the pLexM vector using SacI and XhoI (FastDigest, Thermo Fisher Scientific) and ligated with QuickLigase (New England Biolabs) prior to transformation into *E. coli*. Synthetic nanobody sequences were inserted into pDx_init and pSb_init vectors using fragment exchange cloning, as described in the sybody protocol (Zimmermann *et al.*, 2020). Point mutations in both CLC7 and RUSF1 were introduced via

site-directed mutagenesis using KOD Hot Start Polymerase (Novagen), followed by DpnI digestion (FastDigest, Thermo Fisher Scientific). All plasmid constructs were verified by Sanger sequencing.

2.1.2 Transformation into E. coli

5 µl of ligation product (in cases of restriction cloning), DpnI digested materials (in cases of site-directed mutagenesis), or 0.5 µl of purified plasmids (in cases of plasmid propagation) were mixed with 100 µl of chemically competent *E. coli* cells (produced in house). For plasmid production, Omnimax cells were used; for sybody production, MC1061 cells were used.

The mixture was incubated on ice for 5 minutes, heat-shocked at 42 °C for 45 seconds, and returned to ice for 1 minute. Cells were then recovered in 400 µL of LB medium at 37 °C for 30 minutes before plating onto LB agar plates containing the appropriate antibiotic for selection.

2.1.3 Quantification of gene transcripts (qPCR)

Primers for quantitative PCR (qPCR) specific for human or mouse variants of RUSF1 and the housekeeping genes ACTB and HPRT1 were designed with annealing temperatures between 58 and 60 °C, guanine-cytosine contents between 40 and 60 %, low tendencies to form secondary structures and 3’ guanine-cytosine clamps were conceived. One primer within each primer pair was designed to span an exon-exon junction of the target gene, preventing the amplification of contaminating genomic DNA during qPCR.

For the verification of knock-down, HEK293T cells were seeded in 6-well plates 24 hours prior to transfection. siRNAs were resuspended in 1× siRNA buffer to a final concentration of 5 µM, aliquoted (30 µL/tube), and stored at –20 °C. For each transfection, 10 µL of siRNA (or 2 µg

of RUSF1 or pLexM DNA) was mixed with 3.2 µL or 1.6 µL of DharmaFECT reagent in DMEM to yield two 200 µL solutions. After 5 minutes of incubation, the solutions were combined and incubated for an additional 20 minutes before being diluted with complete media to 2 mL. Culture media was replaced with transfection media, and fresh media was exchanged 24 hours later, and harvested according to experimental design.

For control, or the quantification of gene transcript in various cell types, untreated cells (HeLa, HEK293T, LAT1 KO HCT, COS-7, N2A, and CHO cells) were seeded and harvested at more than 80 % confluency. All samples were snap-frozen in liquid nitrogen to preserve expression profiles.

For RNA extraction, 100 µL of ice-cold QuickExtract RNA Extraction Solution was added directly to each well (Biosearch Technologies), followed by mixing for 1-minute via vortexing. DNase I treatment was omitted as qPCR primers spanned exon–exon junctions. For reverse transcription, 7 µL of RNA was mixed with 1 µL Oligo(dT) (Invitrogen) and 2 µL dNTP, incubated at 70 °C for 10 minutes, then chilled on ice. The reaction was completed by adding 2 µL RT buffer, 2 µL DTT, 1.6 µL dNTP, 2.9 µL RNase-free water, 0.5 µL RNase inhibitor, and 1 µL reverse transcriptase (Agilent Technologies). cDNA synthesis was performed at 37 °C for 1 hour, followed by heat inactivation at 95 °C for 5 minutes. cDNA was purified using the NucleoSpin Gel & PCR Clean-up Kit (Macherey-Nagel).

For qPCR, 0.3 µL of 10 µM primers, 5 µL of 2× SYBR Select Master Mix, 0.5 µL DMSO, and up to 3.9 µL of cDNA were combined in 10 µL reactions, performed in duplicate. Primer sets included RUSF1 and house-keeping genes, with amplification performed using an annealing temperature of 57 °C.

2.2 Protein Expression Trials & Analyses

2.2.1 Protein Expression in HEK293F Cells

For small-scale transfections, 20 mL of HEK293 cells were seeded at a density of 0.7×10^6 cells/mL in FreeStyle media (Gibco) approximately 24 hours prior to transfection, shaking at 37 °C, 100rpm, allowing cells to reach $1.3–1.4 \times 10^6$ cells/mL at the time of transfection. On the day of transfection, 22 µg of plasmid DNA—or a combination of plasmids totalling 22 µg in the case of multi-component complexes—was diluted in unsupplemented DMEM (Gibco) to a final volume of 300 µL and incubated for 5 minutes. In parallel, 44 µL of PEI Max (1 mg/mL stock, made in house) was diluted in DMEM to 300 µL. The diluted PEI solution was added to the DNA mixture (total volume 600 µL), mixed gently, and incubated at room temperature for 10–15 minutes. The resulting transfection mixture was added dropwise to the cells with gentle swirling, and cells were returned to the incubator.

For large-scale transfections (2 L culture volume), cells were prepared using the same seeding conditions. On the day of transfection, 2.2 mg of plasmid DNA was diluted in DMEM to 30 mL and incubated for 5 minutes. Separately, 4.4 mL of PEI Max (1 mg/mL) was diluted to 30 mL. The PEI solution was added to the DNA, gently mixed, and incubated for 10–15 minutes at room temperature. The DNA-PEI complex was then slowly added to the 2 L cell culture with gentle swirling, followed by the addition of sodium butyrate to a final concentration of 8 mM. Cells were harvested 36 hours post-transfection by centrifugation, or at the optimal time point for the protein of interest.

2.2.2 Fluorescence Size Exclusion Chromatography (FSEC)

2 ml of cell culture were harvested via centrifugation and the pellet resuspended in cold PBS with benzonase (Sigma), and various detergents (usually DDM or LMNG, with or without cholesteryl hemisuccinate) were added to a final concentration of 1 %. The membrane was then rotated at 4 ºC for at least 1 hour to solubilise, before being subjected to ultracentrifuge at 200,000g for 20 minutes at 4 ºC. GFP Fluorescence of the sample before ultracentrifuge and the supernatant after were measured for the calculation of membrane solubilisation efficiency. 200 µL of the solubilised membrane were then injected into Superose 6 10/300GL column (Cytiva) in line with RF-20A Fluorescence Detector (Shimadzu), equilibrated in FSEC buffer (20 mM Tris buffer (pH 7.5), 150 mM NaCl and 0.03 % n-Dodecyl- β -D-Maltoside (DDM, Anatrace) or 0.003 % lauryl maltose neopentyl glycol (LMNG, Anatrace):0.0003 % cholesteryl hemisuccinate tris salt (CHS, Sigma-Aldrich). Fluorescence signal at 480nm was detected.

2.2.3 Immunoprecipitations (IP)

For IP experiments a 14–15 mL HEK293F cell pellet was resuspended in 600 µL PBS supplemented with 0.3 µL benzonase and incubated for approximately 30 minutes on ice. Membrane proteins were solubilized in a total volume of 1 mL containing 1xPBS, and 1 % detergent, with rotation in a cold room for 1 hour. The lysate was clarified by ultracentrifugation at 200,000 × g for 20 minutes.

For FLAG tag-based IP, the supernatant was incubated with 10 µL of pre-washed Pierce Anti-DYKDDDDK affinity resin (Thermo Fisher Scientific) for 1 hour at 4 °C with gentle rotation. Following binding, the unbound material was collected, and the beads were washed at least twice using wash buffer (20 mM Tris pH 7.5, 300 mM NaCl, 0.1 % detergent), 500 µL per

wash. After washing, bound proteins were eluted with SDS sample buffer (NuPAGE LDS sample buffer 4x, Thermo Fisher Scientific).

For His tag-based IP, supernatant was incubated with 10 µL of pre-washed HisPur Ni-NTA resin (Thermo Fisher Scientific) with 15 mM imidazole supplemented, rotating at 4 °C for 1 hour. Following binding, the unbound material was collected, and the beads were washed at least twice using wash buffer (20 mM Tris pH 7.5, 300 mM NaCl, 0.1% detergent) with 30mM Imidazole, 500µL per wash. After washing, bound proteins were eluted with SDS sample buffer (NuPAGE LDS sample buffer 4x, Thermo Fisher Scientific).

For Strep tag-based IP, the supernatant was incubated with 10 µL of pre-washed Strep-tactin Superflow resin (IBA Lifesciences) for 1 hours at 4 °C with gentle rotation. Following binding, the flow-through was collected, and the beads were washed at least twice using wash buffer (20 mM Tris pH 7.5, 300 mM NaCl, 0.1% detergent), 500µL per wash. After washing, bound proteins were eluted with SDS sample buffer (NuPAGE LDS sample buffer 4x, Thermo Fisher Scientific).

For Spy-tag based IP, pre-washed Hi Capacity Streptavidin Agarose Resin (Thermo Fisher Scientific) was incubated with biotinylated Spy-Catcher (produced in house) for half an hour, before being added to the ultracentrifuge supernatant. After 1 hour rotation at 4 °C, the flow-through was collected, and the beads were washed at least twice using wash buffer (20 mM Tris pH 7.5, 300 mM NaCl, 0.1% detergent), 500µL per wash. After washing, bound proteins were eluted with SDS sample buffer (NuPAGE LDS sample buffer 4x, Thermo Fisher Scientific).

For IP experiments utilising synthetic nanobodies (sybodies) with a Myc-His-tag, ultracentrifuge supernatant (3 mg aliquots, quantified with BCA assay) were incubated with 3 µg Myc-His-tagged sybody, 4 µg (or 27µg) anti-Myc antibody, and 20 µl 50% Pierce Protein

G Agarose resin (Thermo Fisher Scientific) for 1 hour at 4°C. After collecting the flow-through, beads were washed twice with lysis buffer and three times with wash buffer (20 mM Tris pH 7.5, 300 mM NaCl, 0.1% detergent). Elution was performed with 3× column volumes of 4× sample buffer (NuPAGE LDS sample buffer 4x, Thermo Fisher Scientific).

2.2.4 SDS-PAGE, Coomassie Staining, and Silver Staining

SDS-PAGE was performed using 8, 10, 13.5 or 15 % polyacrylamide gels (made in-house) at a constant voltage of 170 V for 60 minutes in an XCell SureLock Mini-Cell electrophoresis system (Invitrogen). Samples were prepared in NuPAGE LDS 4x Sample Buffer (non-reducing, Thermo Fisher Scientific) without heat denaturation prior to loading. Protein molecular weight standards included either the PageRuler Prestained Protein Ladder (Thermo Fisher Scientific) or the Eco Pre-stained Ladder (Protein Ark). Protein bands were either stained by Coomassie staining using Quick Coomassie Stain (Protein Ark), or by silver staining with the Pierce Silver Stain Kit (Thermo Fisher Scientific). For samples intended for mass spectrometry, the Pierce Silver Stain for Mass Spectrometry kit (Thermo Fisher Scientific) was used.

2.2.5 Western Blotting

A frozen cell pellet (2 mL) was thawed on ice, resuspended in 140 µL of 1× PBS with 0.5 µL benzonase, and incubated on ice for 30 minutes, then mixed with SDS sample buffer (NuPAGE LDS sample buffer 4x, Thermo Fisher Scientific) and loaded onto a 15-well gel (10 µL/well), with untreated HEK cells as a negative control. (Pulldown eluents were mixed with sample buffer and loaded directly onto the gels)

A PVDF membrane (6 cm × 9 cm) was activated in methanol for 5–10 minutes before rinsing in transfer buffer (3 g/L Tris, 14.4 g/L glycine, 200 mL/L methanol). The transfer sandwich consisted of two layers of filter paper, the PVDF membrane, the gel, and two additional layers of filter paper. Transfer was performed at 70 mA for 30 minutes.

The membrane was blocked for one hour at room temperature in PBST (PBS with 0.1% Tween-20) containing 3% milk powder, followed by three PBST washes (5 minutes each). The primary antibody (1:5000) was incubated for one hour at room temperature or overnight at 4°C, followed by three PBST washes. The secondary antibody (1:10,000) was then incubated for one hour at room temperature, followed by another three PBST washes.

For detection, the membrane was incubated in ECL Prime Western Blotting Detection Reagent (Amersham) for 1-2 minutes, then imaged using a gel documentation system (Syngene).

2.3 Protein Expression & Purification

2.3.1 Mammalian Cell Membrane Preparation

Cell pellets were thawed rapidly by immersion in lukewarm water and immediately resuspended in cold PBS (at least 8× pellet weight) supplemented with DNaseI (Sigma). Cells were homogenized thoroughly using a Dounce homogenizer, then lysed by sonication for 2 minutes total, on ice. Lysates were centrifuged at 12,000 × g for 10 minutes at 4 °C, and the supernatant was transferred to Ti45 ultracentrifuge tubes (Beckman Coulter). Samples were topped up with PBS if necessary and centrifuged at 45,000 rpm for 1 hour and 15 minutes in type 45 Ti Fixed-Angle Titanium Rotor (Beckman Coulter).

The resulting membrane pellet was resuspended in one-third to one-half of the original lysis volume using cold 15 mM HEPES (pH 7.5), 20 mM KCl, and homogenized again. The sample was subjected to a second ultracentrifugation (45,000 rpm, 1 hour, 4°C) in the type 45 Ti rotor.

Membranes were finally resuspended in cold PBS (~8 mL per gram of membrane), homogenized, transferred to Falcon tubes, frozen in liquid nitrogen, and stored at –80°C.

2.3.2 Purification of C-terminal TEV-GFP-His Tagged CLC7

Protein purification was carried out through solubilization, affinity purification, dialysis, reverse immobilized metal affinity chromatography (IMAC), and size-exclusion chromatography (SEC). The solubilization step involved incubating the sample (~3g of membrane per litre of cell culture) in 1% detergent for one hour, followed by ultracentrifugation. For Ni-NTA affinity purification, 1 ml of HisPur Ni-NTA resin (Thermo Fisher Scientific) was used, with a three-hour binding step. The resin was washed with 30 column volumes (CV) of wash buffer 1 (1x PBS, 150mM NaCl, 20 mM imidazole, 1 mM ATP, 10 mM $MgCl_2$, 0.1% detergent), followed by 15 CV of wash buffer 2 (1x PBS, 150mM NaCl, 30 mM imidazole, 0.1% detergent). Protein was eluted using 5 CV of elution buffer containing 150 mM imidazole and 0.1% detergent.

Dialysis was performed in 500 ml of TBS containing 0.03% DDM:CHS or 0.003% LMNG:CHS. Reverse IMAC was conducted using a HisTrap column, with the flow-through collected in the presence of 12 mM imidazole, followed by a wash with 30 mM imidazole and elution in 250 mM imidazole. Finally, the sample was subjected to SEC in TBS (20 mM Tris, 150 mM NaCl) supplemented with 0.03% DDM:CHS or 0.003% LMNG:CHS, with the Superdex 200 Increase 10/30 column (Cytiva) and flow rate of 0.5 mL/min and injection volume of 500 µL.

2.3.3 Purification of C-terminal FLAG Tagged Protein (or Complexes through a FLAG Tag)

Membrane was thawed and solubilised in 1% (w/v) detergent, 10% Glycerol, and 150 mM NaCl. Solubilised proteins were separated from other fractions by ultracentrifugation, followed by binding by gentle stirring to Pierce Anti-DYKDDDDK affinity resin (Thermo Fisher Scientific) for 2 hours in buffer equilibrated with PBS, 1% detergent, 10% glycerol, 150 mM NaCl, (Buffer A). FLAG affinity chromatography was then performed in an Econo-Column (Bio-Rad) by washing with 10 column volumes of Buffer A, 30 column volume of Buffer A with 1 mM ATP, 10 mM MgCl2, and finally eluting in 5 column volume of Buffer A with 0.5 mg/ml FLAG peptide (Merck Millipore). Flow through collected from the first FLAG purification was mixed with additional FLAG resin and subjected to another 1.5 hour of rotation, after which transferred to gravity column, washes and elutions were repeated.

For CLC7/OSTM1 samples supplemented with $PI(3,5)P_2$, an additional wash step was included, with 10 column volumes of Buffer A supplemented with 50 µM $PI(3,5)P_2$ (Avanti Polar Lipids or Echelon Biosciences) prior to elution.

Elutions containing target protein were concentrated to 500 ul with a 100 kDa MWCO concentrator (Amicon, Millipore) and subjected to size exclusion chromatography in Superdex 200 Increase 10/300 GL column (Cytiva) in TBS supplemented with 0.03% DDM:CHS or 0.003% LMNG:CHS. Main fractions containing the complex were pooled and concentrated.

2.4 Membrane Protein Characterisation

2.4.1 NanoDSF

Purified protein or complexes at around 0.2 mg/ml were loaded into nanoDSF grade capillaries (NanoTemper Technologies) and placed onto Prometheus NT.48 (NanoTemper Technologies).

The sample was subjected to a linear temperature ramp from 25 ºC to 90 ºC with the rate of 1 ºC /minute, and tryptophan fluorescence at 350nm and 330nm were measured for the calculation of unfolding transition midpoint of the protein of interest. Data was analysed by plotting either: 1) the ratio between fluorescence between 350nm and 330nm, or 2) the derivative of the ratio, against temperature.

2.4.2 Analysis of Oligomeric States — Crosslinking with Glutaraldehyde

A total of 6 µg of DDM:CHS-purified RUSF1 in 20 mM Tris, 150 nM NaCl buffer was incubated with 0.005%, 0.01%, or 0.1% glutaraldehyde in a 10 µL reaction volume. DDM:CHS-purified PepTSo was used as a control. Reactions were performed at room temperature for 20 minutes and quenched with 1 M Tris (pH 8.0). Samples were analyzed by SDS-PAGE on 12 % acrylamide gels or by blue native PAGE using Bis-Tris 4–16 % gradient gels (Invitrogen).

2.4.3 In-Gel Digestion of Coomassie or Silver-Stained Proteins for Mass Spectrometry

Protein bands were excised from Coomassie-stained SDS-PAGE gels after rinsing with Milli-Q water (15 min, room temperature), and cut into ~1 mm^3 cubes. Gel pieces were destained using 0.1 M ammonium bicarbonate: acetonitrile (1:1, v/v) for 30 min with intermittent vortexing, followed by dehydration with 100% acetonitrile.

Reduction was carried out with 10 mM TCEP in 0.1 M ammonium bicarbonate (30 min, RT), followed by acetonitrile dehydration. Alkylation was performed in the dark using 50 mM chloroacetamide in 0.1 M ammonium bicarbonate (30 min, RT), and gel pieces were again dehydrated with acetonitrile.

Proteins were digested overnight (16–20 hr, 37 °C, 650 rpm) with 10 ng/µL trypsin in 10 mM ammonium bicarbonate containing 10 % acetonitrile. Peptides were extracted in three steps using 5 % formic acid: acetonitrile (1:2, v/v), each with 15 min incubation at 37 °C with shaking. Supernatants were pooled and dried by SpeedVac prior to storage at –20 °C before being analysed using liquid chromatography-tandem mass spectrometry (LC-MS/MS). Data analysis was carried out by the proteomics facility at the Oxford University Biochemistry Department.

2.5 Structural Biology

2.5.1 CLC7 Structural Studies

CLC7/OSTM1 complexes were concentrated to 0.75–1 mg/mL. For PI(2,5)P_2-supplemented (Avanti Polar Lipids) samples, 0.1% LMNG and 2.5 mM ATP were added prior to vitrification. Samples were applied to glow-discharged holey carbon-coated grids (Quantifoil R1.2/1.3, 300 mesh, Au) for 10 seconds, blotted for 2 seconds at 100 % humidity and 10 °C, and plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific).

Cryo-EM data were collected on a Titan Krios G4 microscope (Thermo Fisher Scientific) equipped with a cold field emission gun (CFEG) and operating at 300 kV. Imaging was performed in counted mode using Electron Event Representation (EER) format with a Selectris X imaging filter (10 eV slit width) at 165,000× nominal magnification on a Falcon 4 direct electron detector (Thermo Fisher Scientific), corresponding to a physical pixel size of 0.693 Å. Movies were recorded at a total electron dose of 55.7–56.5 $e^-/Å^2$, fractionated to approximately 1 $e^-/Å^2$ per frame.

Initial micrograph processing was performed in real time using the SIMPLE pipeline for motion correction, CTF estimation, particle picking and particle extraction (Caesar *et al.*,

2020). All subsequent processing was performed in either cryoSPARC (Punjani *et al*., 2017) or RELION-3.1 (Zivanov *et al*., 2020).

For WT-CLC7/OSTM1 with PIP2 (Figure 25 A), processing begins with multiple rounds of 2D classification followed by multi-class *ab initio* (k=4) in C1 with a subset of particles, and the resulting volumes were used as references for heterogenous refinement (k=4) in C1 with the full set of particles. A suitable class was selected and particles corresponding to this class subjected to non-uniform refinement in cryoSPARC with C2 symmetry to yield an initial volume. Aligned particles were imported into RELION for Bayesian polishing. Polished particles were subjected to non-uniform refinement in cryoSPARC and the resulting volume used as a reference for a further multi-class ab initio (k=4) in C1 with a subset of particles, and the classes used for hertergenous refinement (C1) with the full particle set. Two classes containing CLC7 were processed further individually with CTF refinements (C2), per-particle defocus, tilt, trefoil, spherical aberration, and tetrafoil estimation/correction applied to generate the final volumes (Figure 25 B,C).

In the case of WT-CLC7/OSTM1 and CLC7-Y715C/OSTM1(Figure 25 D, G), multiple rounds of 2D classification followed by multi-class *ab initio* (k=5) in C1. A suitable class was selected and particles corresponding to this class subjected to non-uniform refinement in cryoSPARC with C2 symmetry to yield an initial volume. Aligned particles were imported into RELION for Bayesian polishing. Polished particles were subjected to non-uniform refinement in cryoSPARC with per-particle defocus, tilt, trefoil, spherical aberration, and tetrafoil estimation/correction applied to generate the final volume (Figure 25 E,F,H,I).

Grid preparation and cryo-EM sample processing steps were performed by Justin Deme.

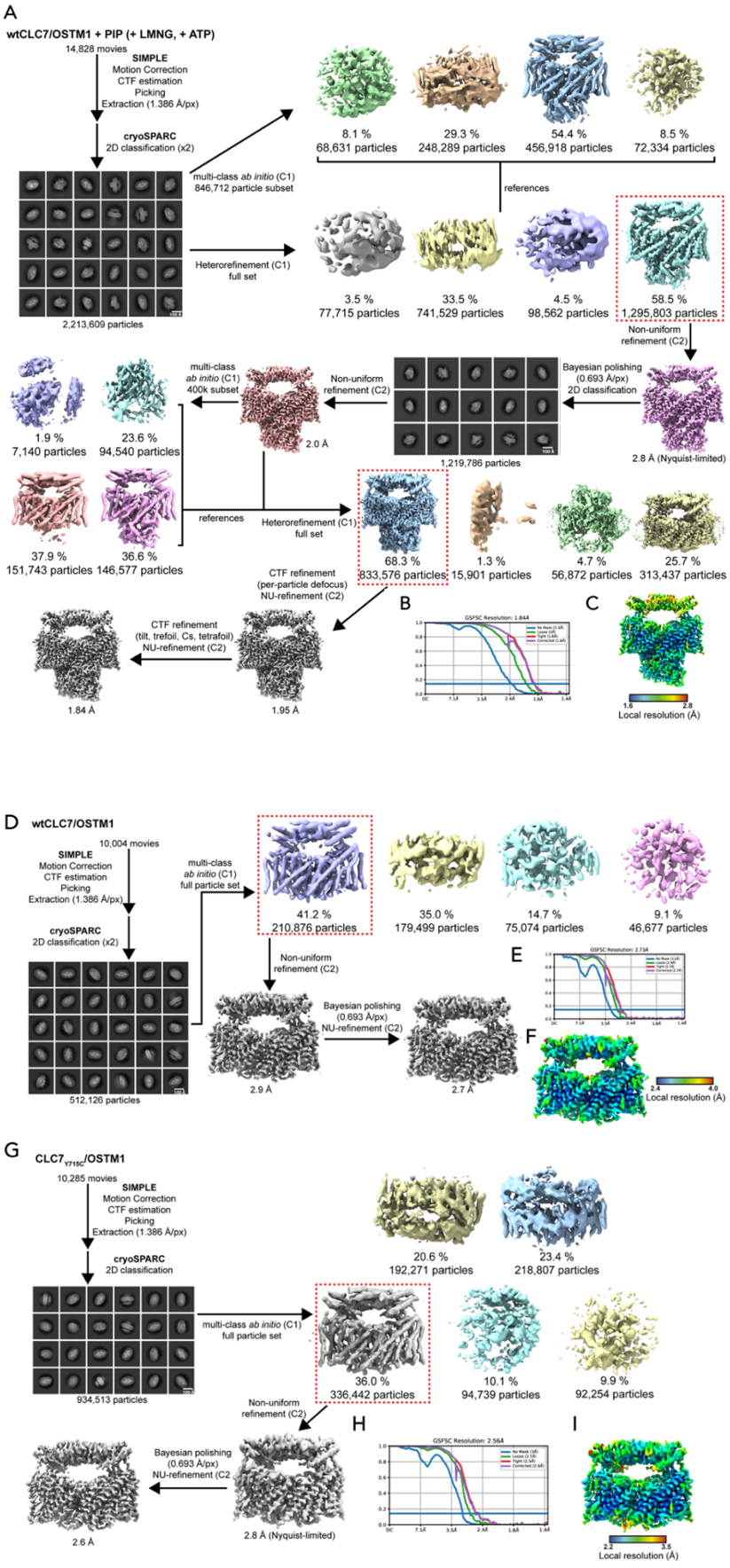


Figure 25. CryoEM processing pipeline for CLC7/OSTM1 structures (same as Figure 29-31)

2.5.2 CLC7 Model Building

The published structure of the wild-type CLC7/OSTM1 complex (PDB ID: 7JM7) was initially fitted into the WT-CLC7/OSTM1 + PI(3,5)P_2 cryo-EM map and manually adjusted in Coot (Emsley *et al*., 2010). The model underwent multiple rounds of real-space refinement in Phenix (Adams *et al*., 2010), interspersed with manual refinement in Coot, until no Ramachandran outliers remained. The refined model was then fitted into the WT-CLC7/OSTM1 (apo) and CLC7-Y715C/OSTM1 maps. Each was trimmed and manually adjusted in Coot, followed by iterative cycles of Phenix real-space refinement and Coot adjustment until the geometry was optimized and free of Ramachandran outliers.

2.5.3 RUSF1 Structural Studies

Protein purified in 50 mM citrate, 100 mM phosphate pH 6.0 with 0.02 % DDM:CHS was concentrated to 3.96 mg/ml with a 30 kDa MWCO concentrator (Amicon, Millipore). Samples were applied to glow-discharged holey carbon-coated grids (Quantifoil R1.2/1.3, 300 mesh, Cu), blotted for 5 seconds at 100 % humidity and 4 °C, and plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific).

Grids were screened on Talos Arctica Cryo-TEM microscope (Thermo Fisher Scientific), with EPU software (Thermo Fisher Scientific). Ice thickness, particle distribution, and particle homogeneity was considered when determining the optimal grids for data collection.

Data were collected in counted super-resolution mode on a Titan Krios G3 (FEI) operating at 300 kV with a BioQuantum imaging filter (Gatan) and K3 direct detection camera (Gatan) at 105,000× magnification (0.832 Å/pixel). We collected 14,967 movies for RUSF1 at a dose rate of 16.8 e− per Å2 per s, exposure time of 2.50 s, corresponding to a total dose of 42 e− per Å2 over 40 fractions.

Initial micrograph processing was performed in real time using the SIMPLE pipeline (Caesar *et al*., 2020), using SIMPLE-unblur for patched (15 × 10) motion correction, SIMPLE-CTFFIND for patched CTF estimation, SIMPLE-picker for particle picking and particle extraction. All subsequent processing was performed in either cryoSPARC (Punjani *et al*., 2017) or RELION-3.1 (Zivanov *et al*., 2020). Resolution estimates were derived from gold-standard Fourier shell correlations (FSCs) using the 0.143 criterion as calculated within cryoSPARC. Local resolution estimations were calculated within cryoSPARC.

4,945,927 particles were extracted to run 2D classification in cryoSPARC, and 1,909,800 particles (selection A) and 783,019 particles (selection B) are selected. Four classes *ab initio* reconstruction in cryoSPARC was performed by selection B. Selection A was used for heterogeneous refinement under two junk and one RUSF1 maps derived from *ab initio*. Particles belonging to the RUSF1 class were subsequently classified by the heterogeneous refinement with *C*2 symmetry expansion. The non-uniform refinement was performed by both selected 488,615 particles and the map of the *C*2 heterogeneous refinement (8 Å initial low-pass filter) and Bayesian polished in RELION. Polished particles were subjected to 2D classification in cryoSPARC and subsequent alignment-free classification with a whole protein mask in RELION. 284,046particles belonging to RUSF1 were used for non-uniform (8 Å initial low-pass) and local (8 Å initial low-pass) refinements with *C*2 symmetry expansion, and the final map was determined at 3.13 Å, based on the FSC = 0.143 criterion.

Cryo-EM sample processing steps were performed by Takafumi Kato (Figure 26).

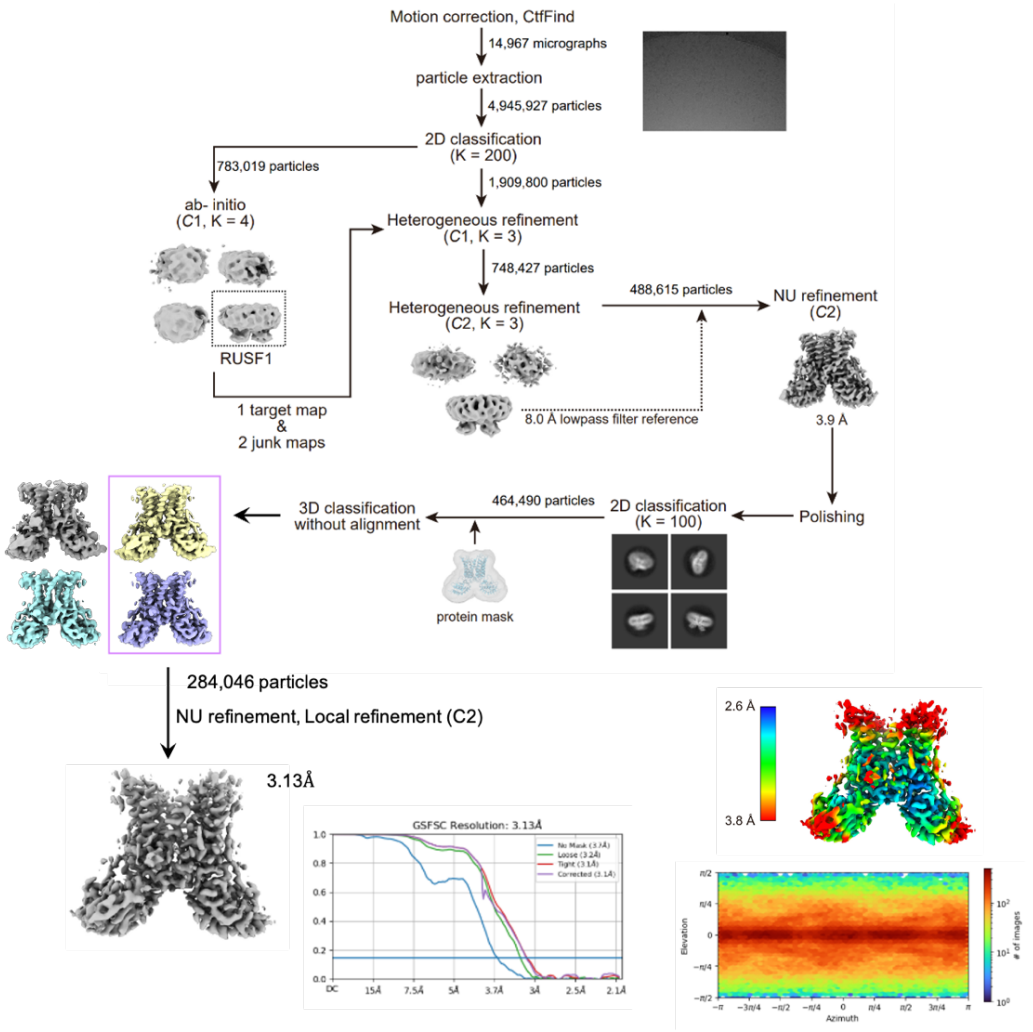


Figure 26. CryoEM processing pipeline for RUSF1 structure (same as Figure 54)

2.5.4 RUSF1 Model Building

AlphaFold prediction of RUSF1 homodimer was trimmed and manually adjusted in Coot (Emsley *et al*., 2010), followed by iterative cycles of Phenix real-space refinement and Coot adjustment until the geometry was optimized and free of Ramachandran outliers (Adams *et al*., 2010).

2.6 Synthetic Nanobody (Sybody) Pipeline

Sybody selection was carried out following the protocol described by Zimmermann *et al*. (2020), with minor modifications. Briefly, one round of ribosome display was followed by three rounds of phage display against the target protein, and a final round of ELISA screening was performed. Positive ELISA hits were identified by Sanger sequencing.

For the initial selection against the cytosolic domain of CLC7, ribosome and phage display were performed using purified, biotinylated CLC7 in LMNG:CHS detergent, followed by ELISA using biotinylated CLC7/OSTM1 to shield the luminal domain and enrich for cytosolic-binding sybodies.

In the optimized selection strategy, all steps were performed with purified, biotinylated CLC7/OSTM1 in LMNG:CHS supplemented with 50 µM $PI(3,5)P_2$ to stabilize the complex. During the second round of phage display, 5 µM purified OSTM1 was used as a competitor to enhance specificity for CLC7.

For RUSF1, all selection steps—including ribosome display, phage display, and ELISA—were carried out using purified, biotinylated RUSF1 in LMNG:CHS.

2.6.1 Protein Biotinylation

For the biotinylation reaction, 500 µL of total reaction volume was prepared using the following components: 175 µL of concentrated protein-Avi to achieve a final concentration of 40 µM, 50 µL each of 10× Biomix A (0.5M bicine buffer, pH 8.3) and Biomix B (100mM ATP, 100mM MgOAc, 500µM d-biotin), and 9.7 µL of GST-BirA (12.5 µg per 10 nmol protein). The reaction was completed with 165.3 µL of water to reach the final volume. Care was taken to maintain a final NaCl concentration of ~50 mM to avoid inhibition of BirA activity. Detergent concentration was adjusted as needed to remain above the critical micelle

concentration to maintain protein stability. The reaction was incubated overnight at 4 °C, followed by size-exclusion chromatography in 20 mM Tris, 150 mM NaCl, and detergent same as protein purification at 3 times cmc concentration. Desired fractions were collected, and biotinylation efficiency was assessed via band shift using streptavidin on SDS-PAGE.

2.6.2 Purification of Sybodies Through Periplasmic Extraction

Selected sybody plasmids were transformed into chemically competent *E. coli* MC1061 cells (produced in house) by heat shock at 42 °C for 45 seconds, followed by recovery in 200 µL LB at 37 °C for 30 minutes. Transformants were plated on LB agar containing 25 µg/mL chloramphenicol and incubated overnight at 37 °C.

Individual colonies were used to inoculate 1 mL of Terrific Broth (TB) with 25 µg/mL chloramphenicol, which was expanded into 50 mL TB cultures. Cultures were grown at 37 °C for 1 hour, then shifted to 22 °C until an OD_{600} of 0.4–0.8 was reached. Protein expression was induced with 0.02% L-(+)-arabinose, and sybodies were expressed overnight at 22 °C.

Cells were harvested by centrifugation at 4,000 × g for 20 minutes and resuspended in 5 mL cold periplasmic extraction buffer (100 mM Tris, pH 8.0, 20 % sucrose, 0.5 mM EDTA), incubated at 4 °C for 1 hour. Cells were then extracted with 20 mL of a second cold buffer (25 mM Tris, pH 8.0, 5% sucrose, 5 mM $MgCl_2$) for 20 minutes at 4 °C. After centrifugation (4,000 × g, 20 minutes), the supernatant was adjusted to 200 mM NaCl, 20 mM Tris (pH 8.0), and 15 mM imidazole.

Approximately 100 µL of pre-washed nickel resin (HisPur Ni-NTA resin, Thermo Fisher Scientific) was added, and the mixture incubated for 1 hour at room temperature with gentle agitation. Resin was collected by centrifugation (500 × g, 2 minutes), transferred to a spin column, and washed once with buffer containing 15 mM imidazole, followed by three washes

with buffer containing 25 mM imidazole. Sybodies were eluted in buffer containing 300 mM imidazole.

Eluted sybodies were dialysed overnight in 20mM Tris (pH7.5), 150mM NaCl, before being used for pulldown analyses.

2.6.3 Sybody Binding Analysis — Mini Pulldowns

To assess sybody binding, 25 µL of target protein (0.2 mg/mL in 20 mM Tris, 200 mM NaCl and 3 times cmc detergent buffer) was incubated with 25 µL of sybody (~2.5 µg) or control buffer/sybody for 1 hour on ice. HisPur Ni-NTA resin (Thermo Fisher Scientific, 10 µL, pre-washed) and 15 mM imidazole were added, followed by rotation at 4 °C for 30 minutes. Samples were centrifuged gently, and the supernatant was removed. The resin was washed three times with 0.5 mL of wash buffer (Tris/NaCl/detergent + 15 mM imidazole), with the second wash optionally extended to 30–60 minutes to displace weak binders. After the final wash, beads were resuspended in 5 µL of 1 M imidazole and 12 µL of 2× loading dye, incubated for 5 minutes, and analysed on a 15 % SDS-PAGE gel.

2.6.4 ELISA-Based Affinity Comparison

For affinity comparison, 96-well Nunc Maxisorp plates were coated overnight at 4 °C with 100 µL per well of Protein A (1 mg/ml) diluted 1:1,000 in PBS. The next day, wells were washed once with 250 µL TBS and blocked with TBS containing 0.5 % BSA for 30 minutes at room temperature. Plates were then washed three times with 250 µL TBS.

A 1:2,000 dilution of anti-c-Myc monoclonal antibody (company) in TBS-BSA with detergent (2× critical micelle concentration, CMC) was added (100 µL per well) and incubated for 45 minutes. After three washes with TBS + detergent (TBS-D), binders were applied in TBS-D-

BSA at the desired concentration or as a dilution series (range), and incubated for 30 minutes. Plates were washed again three times with TBS-D.

Biotinylated target protein (typically 50 nM in TBS-BSA-D) was added and incubated for 40 minutes, followed by three additional washes with TBS-D. A 1:5,000 dilution of streptavidin–peroxidase polymer in TBS-BSA-D (100 µL) was added and incubated for 40 minutes, then washed three times with TBS-D.

Colorimetric detection was performed using ELISA developing buffer (components) (100 µL per well), and absorbance was measured at 650 nm once sufficient signal developed (typically within 1–8 minutes).

To ensure accurate comparison of binder affinities, all samples were analysed on a single plate where possible. Binder performance correlated with BLI-derived affinities, with stronger binders showing earlier saturation and higher signal at lower concentrations. Dilution series ranging from 1 µM to sub-nanomolar levels (e.g., 1 µM, 200 nM, 100 nM, 40 nM, 8 nM, 1.6 nM, 0.32 nM, 0 nM) were used depending on required resolution.

2.6.5 Bio-layer Interferometry

Bio-layer interferometry was performed using an Octet Red385 system (Sartorius) with Streptavidin (SA) biosensors, as previously described (Parker *et al*., 2021). Biosensors were equilibrated for 10 minutes in 200 µL of assay buffer containing 20 mM Tris (pH 7.4), 150 mM NaCl, and 0.003% LMNG, 0.0003% CHS. Sensors were then loaded with biotinylated CLC7/OSTM1 (75 nM) in 50 µL of the same buffer. Sybodies were introduced during the association step at concentrations ranging from 1 µM to sub-nanomolar levels (e.g., 1 µM, 200 nM, 100 nM, 40 nM, 8 nM, 1.6 nM, 0.32 nM, 0 nM), specified in the Results section. Each

experiment included the following steps: equilibration, loading, second equilibration, association, and dissociation, each for 5 minutes.

2.7 Structural Prediction

2.7.1 AlphaFold3 Structural Prediction

To generate structural predictions from AlphaFold3, amino acid sequences of target proteins were taken from corresponding Uniprot entries, and uploaded to the AlphaFold3 server (Abramson *et al.,* 2024). Structural prediction results were analysed on the web server and in ChimeraX (Elaine *et al.,* 2023), as well as examining confident scores and PAE plots.

3. Structural Studies of CLC7 Complexes

3.1 Optimisation of CLC7/OSTM1 Purification

To determine structures of the CLC7 transporter complexed with OSTM1, particularly focusing on interactions involving the lipid PI(3,5)P2 and the Y715C mutant, it was essential to optimize purification strategies to obtain complexes of sufficient yield, purity, and stoichiometry suitable for cryo-electron microscopy (cryo-EM).

Initially, a purification strategy involving a CLC7 construct tagged with C-terminal TEV-cleavable GFP-His8, co-expressed with a C-terminal FLAG-tagged OSTM1, was adopted (Figure 27 A). The purification protocol comprised of sequential steps: Ni-NTA affinity purification, overnight dialysis accompanied by TEV protease cleavage, reverse immobilized metal affinity chromatography (IMAC) to remove His-tagged TEV protease and GFP, followed by size exclusion chromatography (SEC). Although this method resulted in relatively good yields (153 µg/L), SDS-PAGE analysis revealed partial loss of OSTM1 during purification steps, leading to suboptimal stoichiometry of the final complex (based solely on relative band intensity visually).

In an attempt to improve stoichiometry, an alternative tandem purification protocol was evaluated (Figure 27 B). This strategy employed FLAG affinity purification subsequent to overnight dialysis and TEV protease cleavage. Although the resulting CLC7/OSTM1 complex exhibited improved stoichiometry, impurities persisted due to inefficient removal of TEV protease, as demonstrated in Figure 27 B (final yield: 97 µg/L).

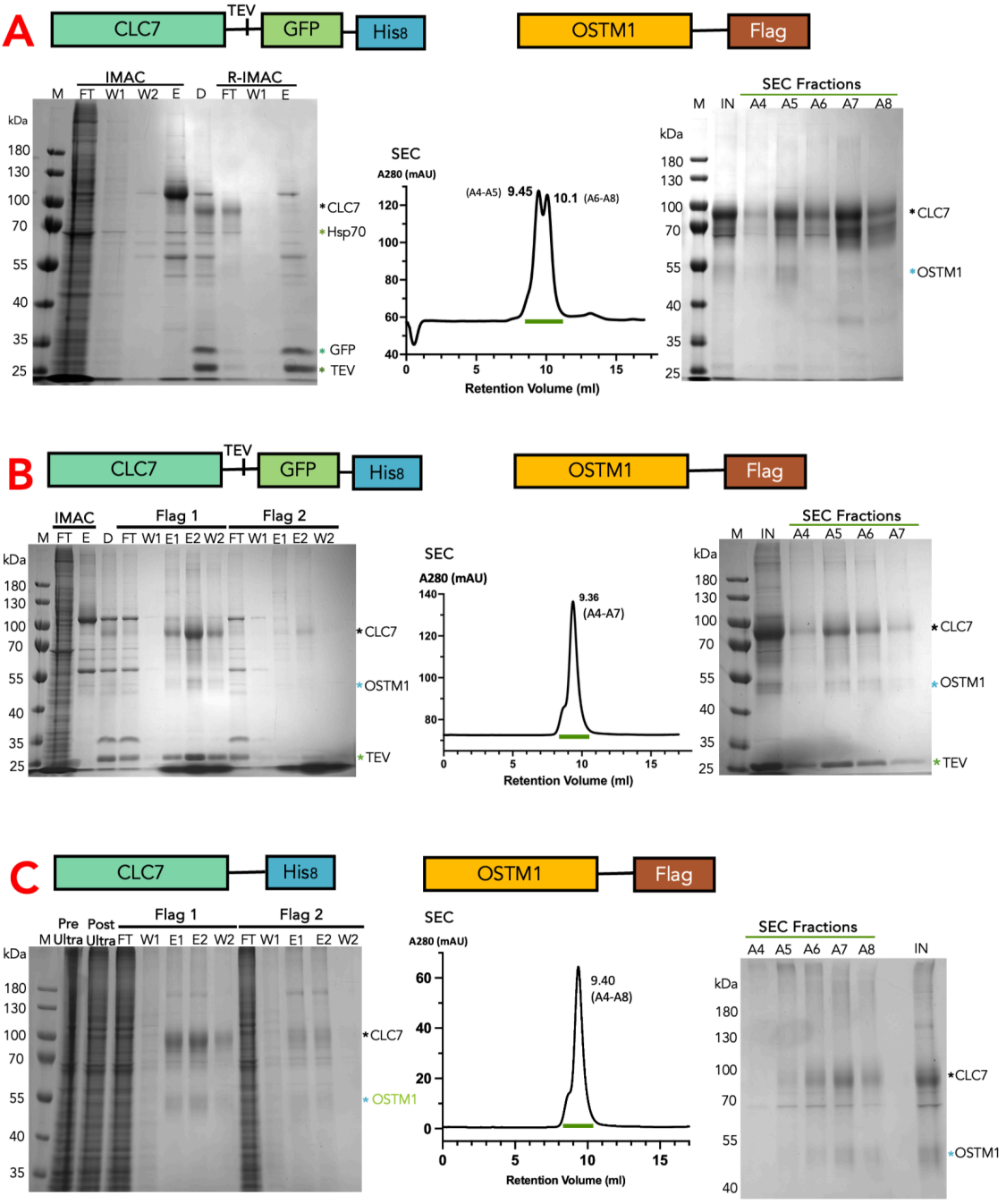


Figure 27. Optimising CLC7/OSTM1 Complex Purification

A. Purification trial of CLC7-TEV-GFP-His/OSTM1-FLAG complex through IMAC and reverse IMAC purification. Top: construct diagrams. Bottom left: SDS-PAGE results of IMAC and reverse-IMAC purification, complex components and impurities annotated. Centre: size-exclusion chromatogram of complex, retention volume annotated. Right: SDS-PAGE results of size-exclusion chromatography, complex components annotated. B. Purification trial of CLC7-TEV-GFP-His/OSTM1-FLAG complex through tandem affinity purification. Top: construct diagrams. Bottom left: SDS-PAGE results of Ni-NTA and FLAG affinity purification, complex components and impurities annotated. Centre: size-exclusion chromatogram of complex, retention volume annotated. Right: SDS-PAGE results of size-exclusion chromatography, complex components and impurities annotated. C. Purification trial of CLC7- His/OSTM1-FLAG complex through FLAG affinity purification. Top: construct diagrams. Bottom left: SDS-PAGE results of FLAG purification, complex components annotated. Centre: size-exclusion chromatogram of complex, retention volume annotated. Right: SDS-PAGE results of size-exclusion chromatography, complex components annotated.

Ultimately, the purification strategy yielding optimal results involved a streamlined one-day purification protocol combining FLAG affinity chromatography directly followed by size exclusion chromatography (Figure 27 C). While this method had comparatively lower yield (91 µg/L), it significantly improved the purity and stoichiometry of the CLC7/OSTM1 complex, making it ideally suited for structural studies using cryo-EM.

Using the refined protocol, several samples were purified for detailed structural analysis via cryo-EM, including wild-type CLC7/OSTM1 complexes and Y715C mutant complexes, both in the presence (WT+PIP2, Y715C+PIP2) and absence (Apo) of PI(3,5)P2 (Avanti Polar Lipids) lipid (Figure 28 A-D). All complexes consistently eluted at approximately 9.5 mL during size exclusion chromatography, in largely mono-dispersed peaks. Coomassie-stained SDS-PAGE confirmed the purity of these samples, clearly indicating both CLC7 (~100 kDa) and glycosylation-smeared OSTM1 (~55 kDa) bands (A-D).

Thermostability assays conducted on purified samples indicated uniform melting profiles around 70°C for all tested conditions, confirming that the structural integrity of the transporter complex remained largely unaffected by the Y715C mutation (Figure 28 E). Notably, an intriguing additional unfolding event appeared between 20-40°C exclusively in the wild-type complex with PI(3,5)P2 (Avanti Polar Lipids) lipid (WT+PIP2). This anomaly may correspond to domain dissociation or partial unfolding events distinct to lipid-bound complexes, warranting further investigation (discussed later).

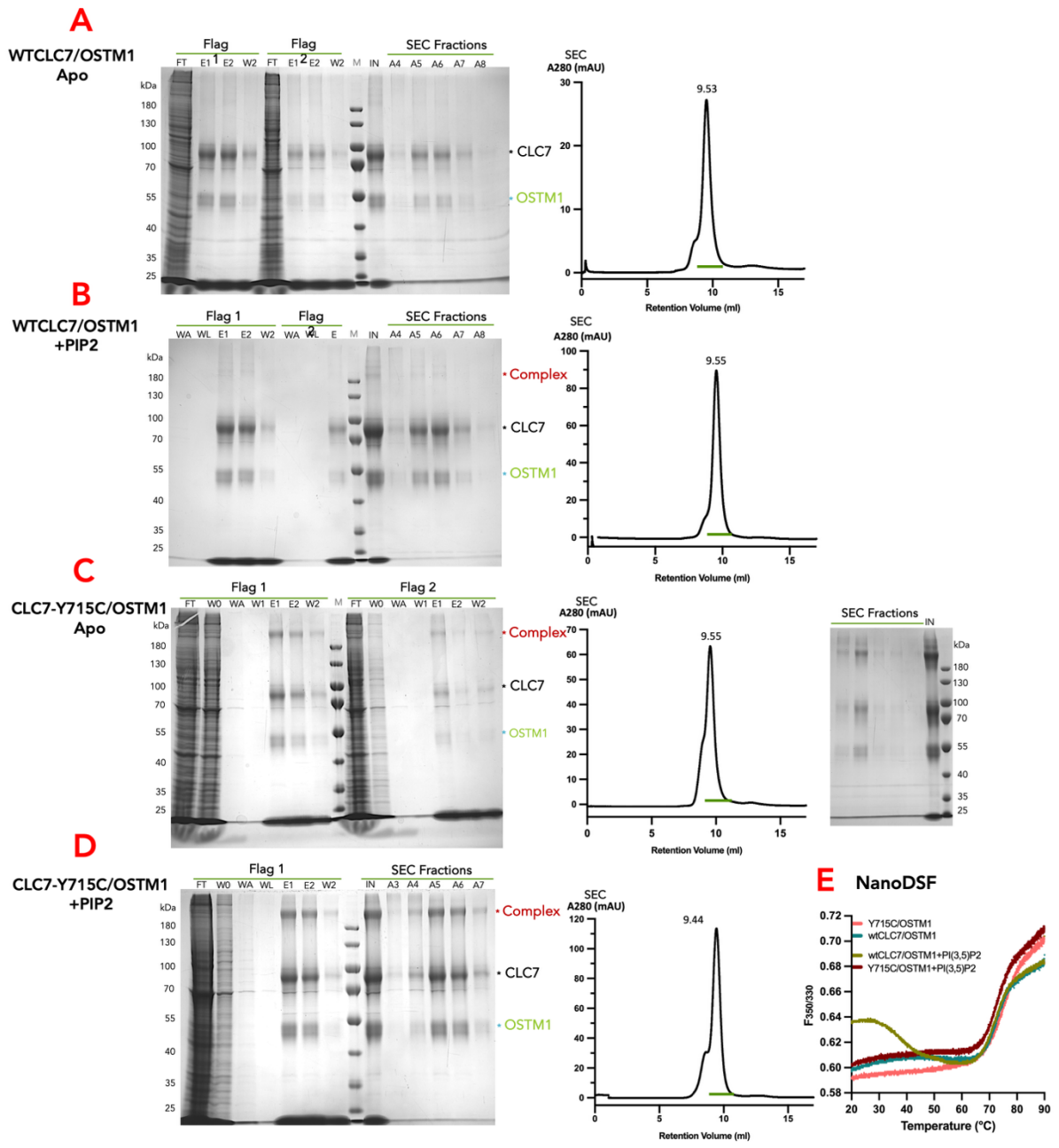


Figure 28. CLC7/OSTM1 Complex Purification for CryoEM

A. Purification of CLC7-His/OSTM1-FLAG complex. Left: SDS-PAGE results of FLAG purification and size-exclusion fractions, complex components annotated. Right: size-exclusion chromatogram of complex, retention volume annotated. B. Purification of CLC7-His/OSTM1-FLAG complex supplemented with PIP2. Left: SDS-PAGE results of FLAG purification and size-exclusion fractions, complex components annotated. Right: size-exclusion chromatogram of complex, retention volume annotated. C. Purification of Y715C-His/OSTM1-FLAG complex. Left: SDS-PAGE results of FLAG purification, complex components annotated. Centre: size-exclusion chromatogram of complex, retention volume annotated. Right: SDS-PAGE results of size-exclusion chromatography, complex components annotated. D. Purification of Y715C-His/OSTM1-FLAG complex supplemented with PIP2. Left: SDS-PAGE results of FLAG purification and size-exclusion fractions, complex components annotated. Right: size-exclusion chromatogram of complex, retention volume annotated. E. Nano-DSF based thermal stability assay of purified CLC7/OSTM1 complexes. Duplicate experiments were performed for each sample, and results of one repeat are shown. Y715C/OSTM1 in salmon, wtCLC7/OSTM1 in teal, wtCLC7/OSTM1+PIP2 in olive, Y715C/OSTM1+PIP2 in magenta.

3.2 PIP2 is Essential for CBS-TMD Interaction in CLC7

The initial attempts to elucidate the structure of the CLC7 mutant in complex with OSTM1(CLC7–Y715C/OSTM1) using cryo-electron microscopy (CryoEM) yielded unexpected results (Figure 29 A-C). After 2D classification, there were clear densities corresponding to OSTM1, forming a cap-like shape over the detergent micelle density (Figure 29 A). However, on the other side of the micelle, signals corresponding to the CBS domains were blurred. Following *ab initio* model generation, it was clear that all the transmembrane helices in the homo-dimeric CLC7 complex were resolved, as well as the OSTM1 dimer, though densities for the CBS domains were lost, potentially due to high flexibility. This dataset was solved to 2.6 Å global resolution after non-uniform refinement and polishing (Figure 29 B, C).

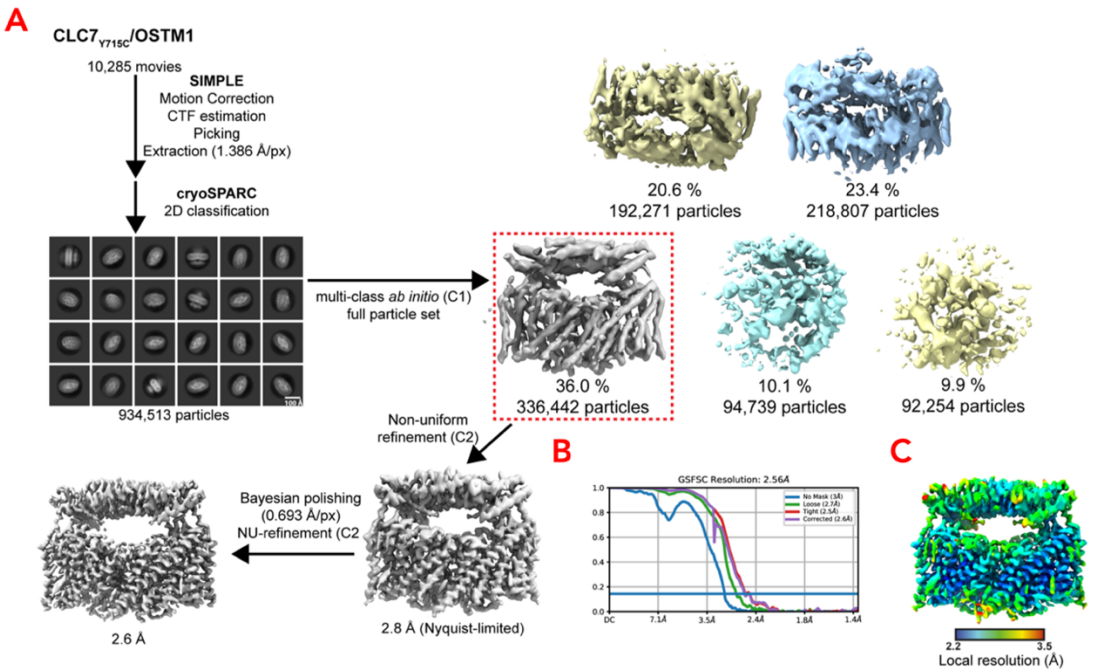


Figure 29. CryoEM Processing Pipeline for CLC7-Y715C/OSTM1 Complex

A. Image processing workflow for CLC7-Y715C/OSTM1. B. Gold-standard Fourier Shell Correlation (FSC) curves used for global resolution estimation. C. Local resolution estimate of the final volume. Data processing done by Justin Deme.

To assess if this observation was due to the instability of Y715C mutant, or human errors during purification, we subsequently performed structural studies using wild-type CLC7 (wtCLC7)/OSTM1. Despite purification results confirmed that the protein was intact (Figure 33 C), the CBS domain remained unresolved both in 2D classification, and the 3D reconstruction (Figure 30 A). The final reconstruction yielded a 2.7 Å map with clear density for the transmembrane domain of the CLC7 dimer and density for OSTM1 (Figure 30 B, C). The results were perplexing, since multiple structures of CLC7 have been published prior to our attempts (Schrecker *et al*., 2020; Zhang *et al*., 2020), all of which resulted in maps with clear CBS domain density.

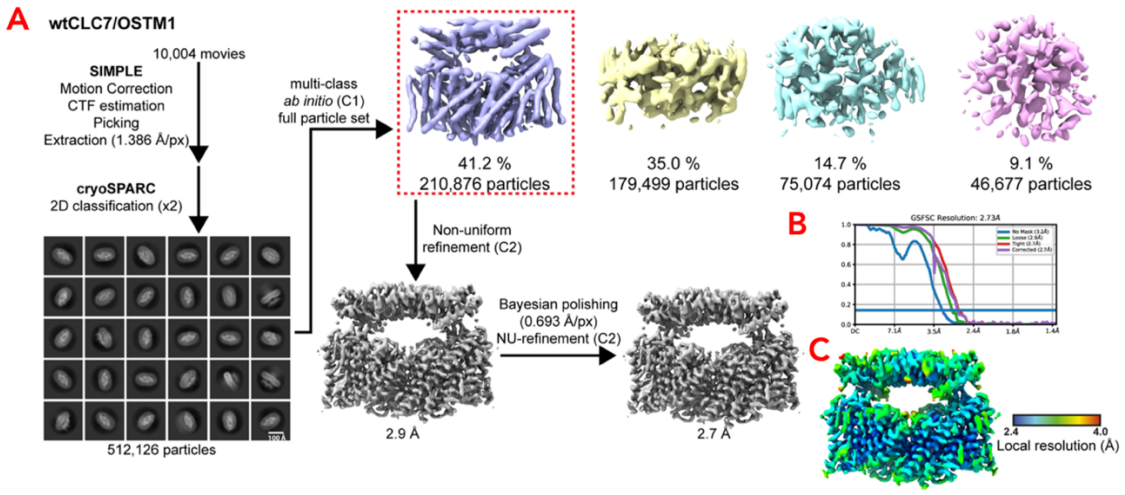


Figure 30. CryoEM Processing Pipeline for wtCLC7-Apo /OSTM1 Complex

A. Image processing workflow for wtCLC7 /OSTM1. B. Gold-standard Fourier Shell Correlation (FSC) curves used for global resolution estimation. C. Local resolution estimate of the final volume. Data processing done by Justin Deme.

A significant breakthrough occurred when 50 µM of phosphatidylinositol 3,5-bisphosphate (PI(3,5)P2, Avanti Polar Lipids) was introduced during the purification of the wtCLC7/OSTM1 complex. Under these conditions, distinct densities corresponding to the CBS domain appeared clearly in the 2D class averages and in the 3D reconstructions, suggesting that the addition of

PI(3,5)P2 (Avanti Polar Lipids) directly facilitated the association of CBS domains to the transmembrane domains in CLC7 (Figure 31 A).

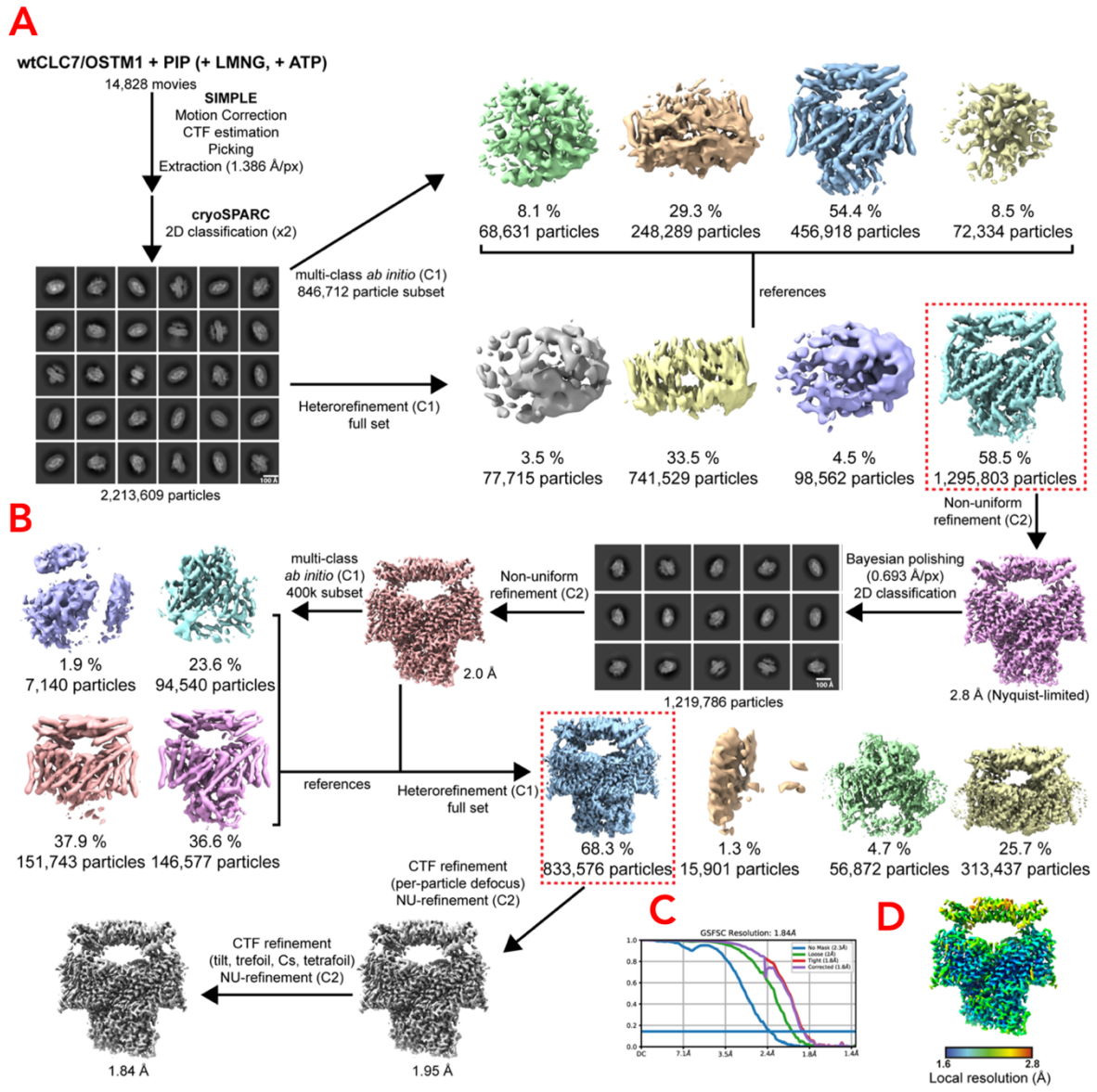


Figure 31. CryoEM Processing Pipeline for wtCLC7+PIP2/OSTM1 Complex

A. Image processing workflow for wtCLC7/OSTM1 supplemented with PIP2. B. 3D classes showing populations of CLC7 with or without dissociated CBS domains. C. Gold-standard Fourier Shell Correlation (FSC) curves used for global resolution estimation. D. Local resolution estimate of the final volume. Data processing done by Justin Deme.

Interestingly, in one of the multi-class *ab initio* model generation steps intended to separate heterogenous particles, two classes of particles were separated which corresponded to CLC7: one class with resolved CBS domains comprising 36.6% of the total particles, and another class

with unresolved CBS domains, comprising 37.9% of the total particles (Figure 31 B). This suggested that even in the presence of PI(3,5)P2 (Avanti Polar Lipids), the CBS domains of a significant portion of particles remained flexible. This dataset was solved to a final resolution of 1.8 Å, resulting in a map with clear density for both OSTM1 dimer and CLC7 dimer (Figure 31 C, D).

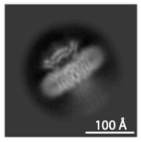
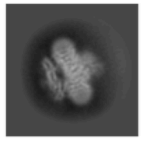
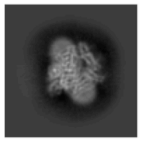
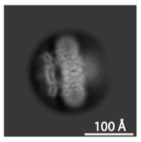
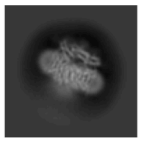
To rigorously confirm the specific effect of PI(3,5)P2 (Avanti Polar Lipids) observed in the previous datasets, we examined multiple smaller CryoEM datasets across various conditions, systematically altering components such as ATP and detergent concentration before CryoEM grid preparation (Table 2). Consistently, only datasets incorporating PI(3,5)P2 (Avanti Polar Lipids) demonstrated stable and discernible CBS domain densities, suggesting PI(3,5)P2's potential role as a molecular glue, effectively bridging the CBS and TMD regions.

The maps used for further analysis were the aforementioned 2.7 Å map for the wtCLC7(Apo)/OSTM1 complex, 1.8 Å map for the wtCLC7/OSTM1 complex with PI(3,5)P2 (Avanti Polar Lipids), and 2.6 Å map for the Y715C/OSTM1 complex (Supplementary Table 1).

Notably, all of the maps revealed highly detailed features, including the distinct densities of cholesterol, lipids, and detergents (Figure 33 A). Importantly, a single molecule of PI(3,5)P2 (Avanti Polar Lipids) was resolved on each monomer of CLC7 in the WT PIP2 map, binding at a previously recognised lipid-binding pocket situated at the interface between the TMD and CBS domains (Schrecker *et al*., 2020), forming hydrogen bonding network with residues D282, K285, Y715 and R717 (Figure 33 B). The lipid's position was additionally stabilised by a critical beta hairpin structure between helix F and helix G, designated as the FG loop, through interactions involving residues R265, S266, and T267 (Figure 36 A-D).

Table 2. Representative 2D Classes of CLC7/OSTM1 Complex Under Various Conditions

Summary of the conditions (PI(3,5)P2 from Avanti Polar Lipids, detergent, and ATP) in different WT-CLC7 or Y715C EM datasets, the resolution of final maps, and whether the CBS domains are resolved; maps with resolution in bold were analysed further. Examples of 2D Classes from the collected EM datasets shown, with side view of CLC7/OSTM1 complex for visualization of resolved / not resolved CBS domains.

Sample	PIP2	Excess LMNG:CHS	Excess ATP	CBS Domain	Representative 2D Class	Resolution (Å)
wtCLC7 /OSTM1	-	-	-	Not resolved		**2.7 WT Apo PDB: 9G6D**
	+	+	+	Resolved		**1.8 WT+PIP2 PDB: 9G6C**
	+	+	-	Resolved		2.8
	+	-	+	Resolved		2.8
	-	+	-	Not resolved		2.9
CLC7-Y715C /OSTM1	-	-	-	Not resolved		**2.6 Y715C Apo PDB: 9G6E**
	+	+	+	Not resolved		2.8
	+	-	-	Not resolved		2.7

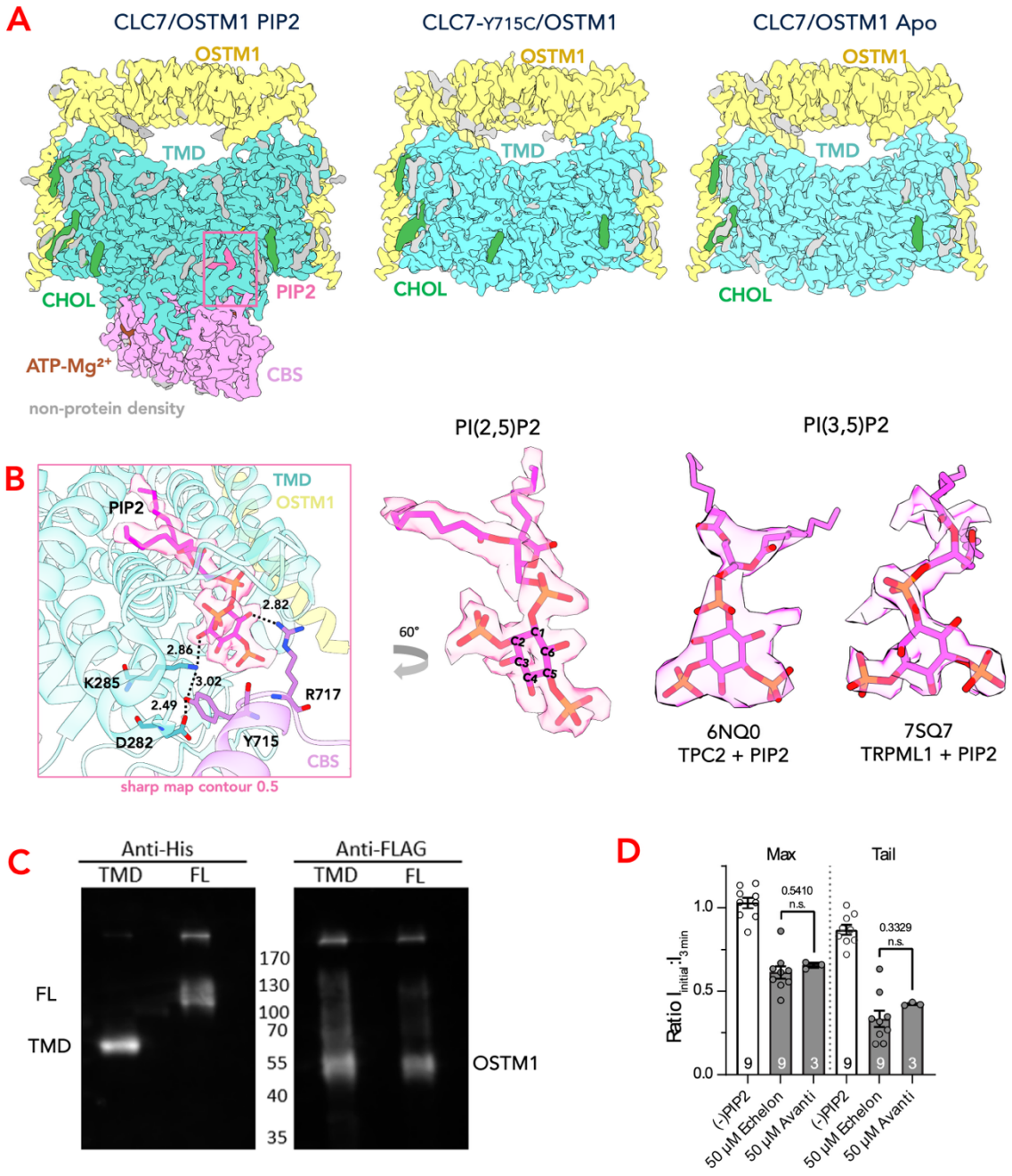


Figure 33. Structure of CLC7/OSTM1 Complexes

A. Density maps are shown for wtCLC7/OSTM1 + PI(3,5)P2 from Avanti Polar Lipids (left), Y715C/OSTM1 (centre) and WT-apo/OSTM1 (right). The TM domain is cyan, CBS domains are light purple and OSTM1 is yellow. Additional lipid and detergent densities are shown in green and grey. PIP2 lipid shown in hot pink. B. Left: density of the bound phosphoinositide (hot pink) with the outline of the atomic model for CLC7. Note the location of phophate moieties on the 2 and 5 positions of the inositol ring. Centre: close-up view of PI(2,5)P2 lipid density (lipid sourced from Avanti Polar Lipids), carbon on the inositol ring numbered. Right: PI(3,5)P2 density from two published structures (6NQ0 and 7SQ7), both sourced from Echelon Biosciences. C. Western blots of purified wtCLC7/OSTM1 without added PI(3,5)P2 (Avanti Polar Lipids), compared to CLC7(TMD)/OSTM1. Full-length CLC7: between 100 kDa and 130 kDa. TMD only: between 55 kDa and 70 kDa. D. (Done by Jacob Hilton) Whole-cell patch clamp recordings compare inhibition by 50 µM PI(3,5)P2 of CLC7PM currents by nominally PI(3,5)P2 diC8 lipids purchased from either Avanti Polar Lipids or Echelon Biosciences. Mean and Standard Error are shown, with symbols representing measurements from individual cells.

A surprising revelation emerged upon closer examination of the bound lipid structure: the phosphate group was unexpectedly attached to carbon 2 instead of carbon 3, indicating that the CryoEM studies inadvertently utilised non-physiological PI(2,5)P2 sourced from Avanti Polar Lipids (Figure 33 B). This was in contrast to electrophysiology experiments conducted by collaborators, which correctly employed physiological PI(3,5)P2 from Echelon Biosciences. The chemical make-up of the lipid from different sources were further confirmed via NMR analysis (Supplementary Figure 3). Remarkably, both lipid variants exhibited comparable inhibitory effects on CLC7 transporter activity (Figure 33 D).

Based on the structural insights, mechanistic hypothesis could be proposed, wherein physiological PI(3,5)P2 (or the non-physiological PI(2,5)P2) serves as a molecular glue facilitating robust interaction between the TMD and CBS domains. This interaction presumably stabilises the transporter in an inactive state, whereas the absence of the lipid or mutations disrupting lipid binding would increase CBS domain flexibility, activating the transporter (Figure 34 A).

To validate this hypothesis, electrophysiological analyses (patch clamp on Xenopus oocytes) by collaborators were conducted on mutants designed to disrupt lipid binding. Mutants designed to disrupt TMD-CBS interactions included D282A, K285E, Y715C, and R717E (Figure 34 B).

These mutants indeed demonstrated hyperactivity relative to wild-type, and importantly, their activity was less effectively inhibited upon addition of PI(3,5)P2 from Echelon Biosciences (Figure 34 B). Further experiments included a construct comprising only the CLC7 TMD domain, which similarly exhibited increased activity and lacked proper inhibition (Figure 34 C, D). Remarkably, reintroducing the CBS domains at the cytosolic terminus of OSTM1 partially restored regulation and activity patterns akin to the wild-type complex.

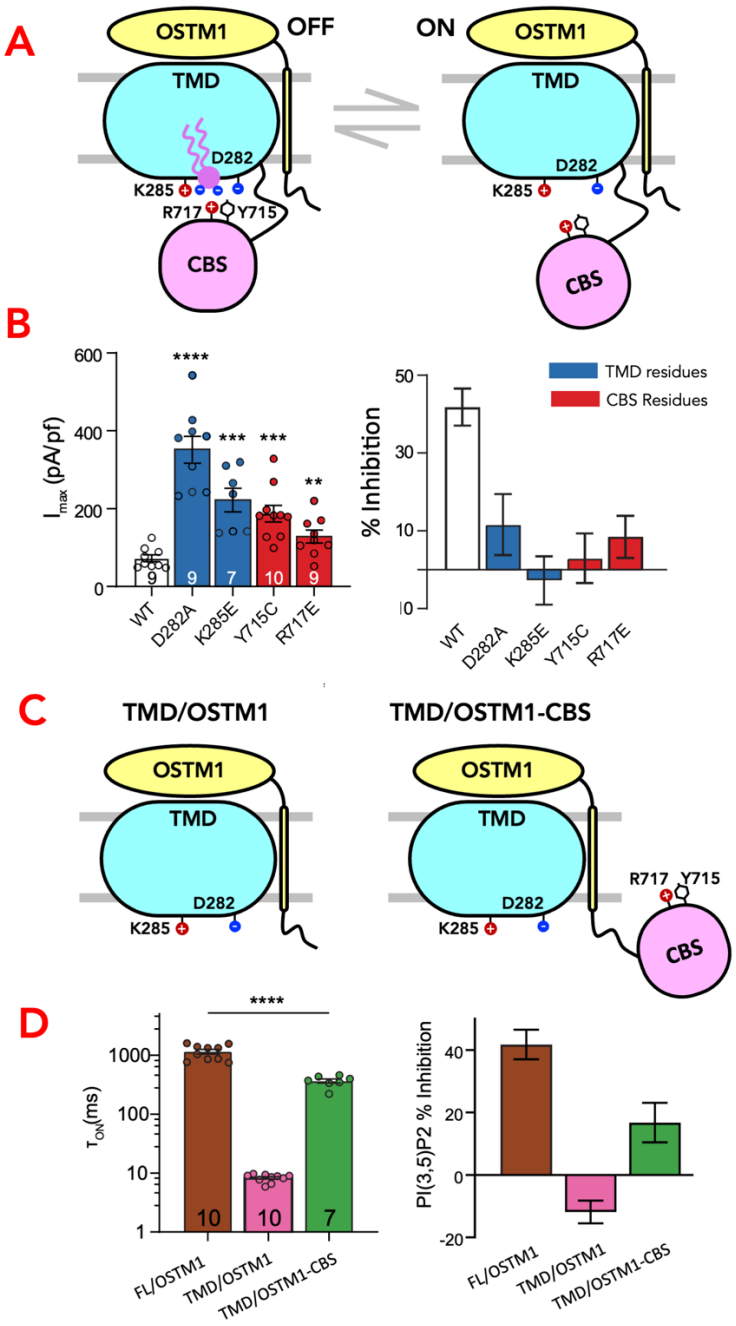


Figure 34. Effect of Weakening Lipid Binding Sites & Removal of CBS Domains

A. Schematic illustration of the proposed model of PI(3,5)P2-facilitated association of CBS domains and TM domain of CLC7. The PI(3,5)P2-bound state with associated CBS domains represents the inhibited state, whereas the apo, CBS dissociated state represents the active state. B. (Done by Jacob Hilton) Left: mean Peak currents of various CLC7 constructs with point mutation at the CBS-PI(3,5)P2-TMD interfaces. Right: percent inhibition by 50 μM PI(3,5)P2 (Echelon Biosciences) of maximum currents measured at +140 mV for the same mutants, error bars representing standard error of the mean (SEM). C. Illustrations depicting TMD-only, and split transporter constructs. D. (Done by Jacob Hilton) Left: activation time constants of full-length, TMD only and split CLC7. Right: percentage inhibition of maximum currents measured at +140 mV by 50 μM PI(3,5)P2 (Echelon Biosciences), error bars representing standard error of the mean (SEM). Experiments performed by Jacob Hilton.

3.3 Effects of PIP2 Binding

Previously, Schrecker *et al*. (2020) reported the structure of WT CLC7 bound to PI3P, a phosphoinositide that notably does not inhibit CLC7 activity. To investigate the inhibitory mechanism associated specifically with PI(3,5)P2 (or PI(2,5)P2), we superposed our PI(2,5)P2-bound structure with the previously published PI3P-bound structure (PDB ID: 7JM7, Schrecker *et al*., 2020). The alignment yielded an RMSD of 0.86 Å (between 1773 pruned atom pairs), suggesting structural congruence across most protein regions, particularly backbone of helices. Variability was predominantly localized to peripheral loop regions (Figure 35 A-C).

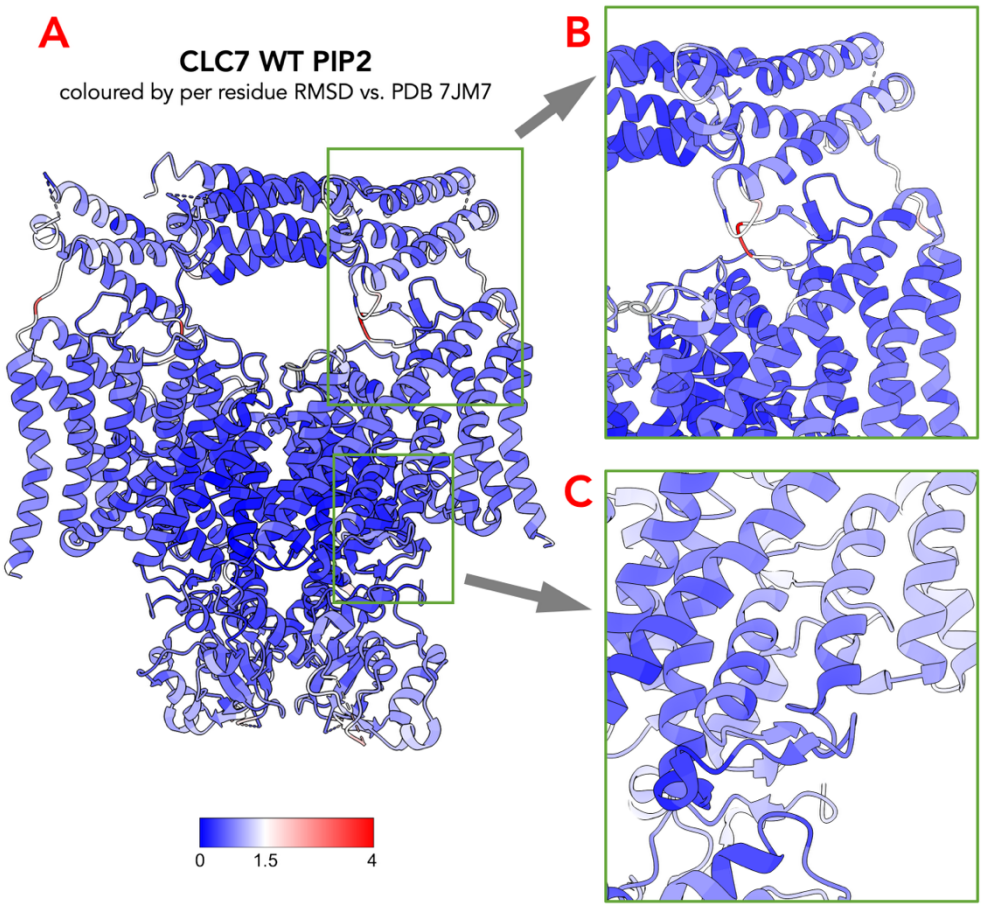


Figure 35. Comparison of WT PIP2 Structure with Published Structure

A. Structure of WT PI(2,5)P2 structure shown, coloured by per residue RMSD from a published structure (PDB 7JM7), colour key shown below, blue indicated regions with high structural congruence, red indicated regions with low structural congruence (lowest ~ 4 Å), white indicated RMSD of around 1.5 Å. B, C. Zoomed-in view of regions with low structural congruence (B), between CLC7 and OSTM1; and F-G loop (C).

Focused examination of the lipid-binding pocket—comprised of transmembrane helices E, F, the F-G loop, and the CBS loop near Y715C—revealed key structural differences between the two phosphoinositides. The 5' phosphate of our PI(2,5)P2 lipid (Avanti Polar Lipids) occupied a similar pocket as the 3' phosphate from PI3P, at the groove between helix E and F, the additional phosphate group at carbon 2 interacted with R265 on the F-G loop. This interaction caused the inositol ring to sit approximately 1 Å deeper into the pocket, packing the fatty acid tail base closer to the transmembrane domain, and also pulling the residues of the F-G loop closer to the TMD by about 1 Å (Figure 36 A-C).

Overall, this moderate structural shift suggested that the extra negatively charged phosphate facilitated stronger interactions with positively charged residues on the F-G loop, and might potentially enhance affinity between the lipid and CBS domains. Such tighter domain interactions presumably underpin the inhibition observed physiologically.

Supporting this hypothesis, mutations introduced around the lipid-binding site (primarily on the F-G loop) by our collaborators resulted in reduced lipid-mediated inhibition, consistent with the structural role of the F-G loop in lipid binding (Figure 36 D). Meanwhile, the mutations only affected activation kinetics minimally, also consistent with the current hypothesis that slow gating is largely mediated by CBS domains, and F-G loop did not directly form any interactions with the CBS domains.

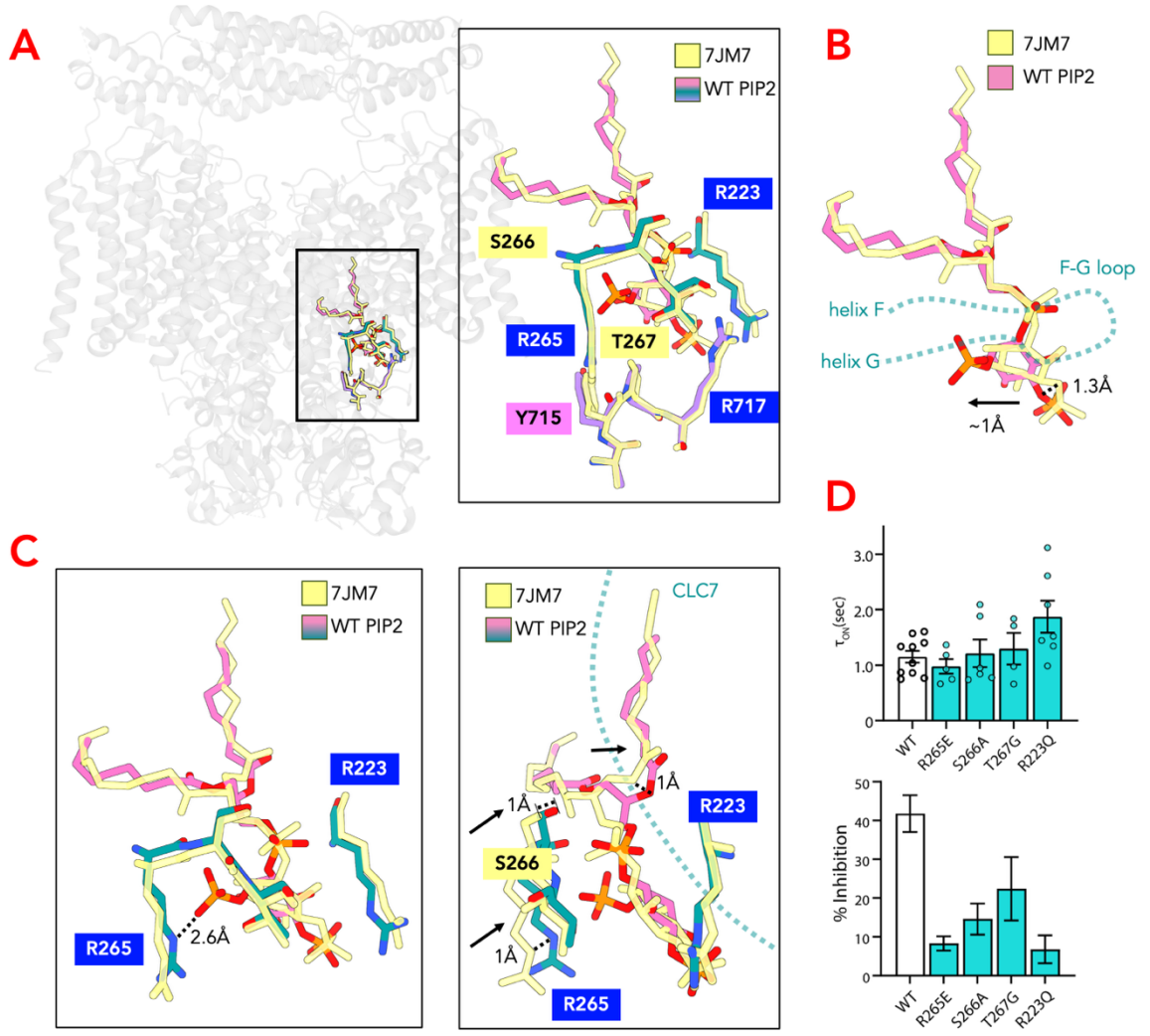


Figure 36. Specificity of PIP2 Inhibition of CLC7

A. Overlay of PIP2 lipid (hot pink), key interacting residues on F-G loop (teal) and CBS domain (purple) onto WT PIP2 structure (light grey); corresponding PI3P lipid and interacting residues from the published structure 7JM7 superimposed (yellow). B. Comparison of binding pose of PI(2,5)P2 and PI3P, the inositol ring of PI(2,5)P2 shifted ~ 1 Å deeper into the binding pocket. Teal dotted line indicated F-G loop. C. Left: PI(2,5)P2 formed an additional interaction with R265 with the phosphate group on carbon 2. Right: PI(2,5)P2 binding lead to a tightening of the entire binding pocket towards the transporter, compared to PI3P-bound structure; several residues (e.g., R265, S266) shifted ~ 1 Å closer to the TM domain of CLC7 indicated by teal dotted line. Corresponding residues and PI3P from structure 7JM7 superimposed (yellow). D. (Done by Jacob Hilton) Left: activation time constants of various F-G loop mutants compared to wild-type; mutants behaved similarly to wild-type, indicating that activation kinetics was governed by CBS domains alone. Right: percentage inhibition by 50 μM PI(3,5)P2 of maximum currents measured at +140 mV of the same mutants, which were less inhibited than wild-type, indicating lower affinity to PIP2.

Further comparative analysis across the human CLC chloride transporters showed that the F-G loop structure is also present in CLC6, the closest homolog of CLC7, but absent in CLC3-5 (Figure 38 A-C). This implies that CLC6 might similarly bind phosphoinositides or other negatively charged lipids. Indeed, inspection of a recently published human CLC6 structure (PDB ID: 8JPO, Zhang *et al.,* 2023) revealed an unmodelled density consistent with lipid binding at the F-G loop, although lipid identity remained ambiguous due to limited EM map resolution (Figure 37 B). Moreover, CLC6's F-G loop possesses fewer positively charged residues compared to CLC7 (Figure 38 A), which might contribute to weaker lipid binding affinity, potentially explaining the less defined lipid density in the EM map. Since no evidence of lipid-linked regulation has been found in CLC6, it remains unclear whether the F-G loop in CLC6 serves the same function as in CLC7.

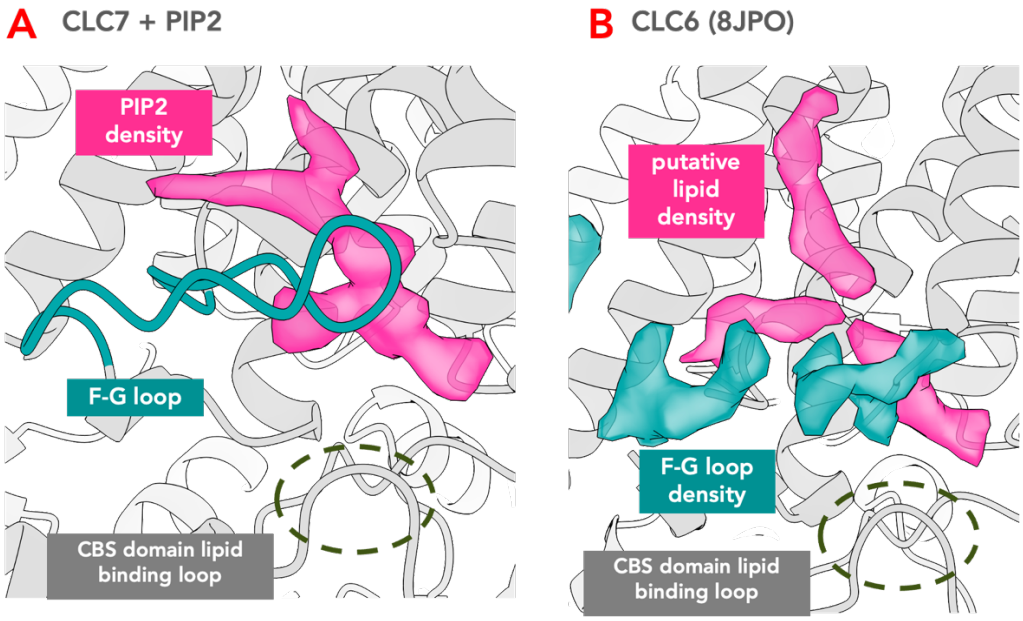


Figure 37. F-G Loop and Lipid Density in CLC6

A. WT PIP2 structure (grey) with F-G loop coloured in teal, PI(2,5)P2 density coloured in hot pink. B. Human CLC6 structure (grey, PDB 8JPO) with putative F-G loop density coloured in teal, PIP lipid density coloured in hot pink. CBS domain lipid binding loop (where Y715 in human is located) circled with dotted line.

On the other hand, in CLC3-5, the F-G loop was replaced by short linkers, even though literature suggest that they might also bind and be regulated by PI(3,5)P2 lipid (Figure 39 A-

B, Schrecker *et al*., 2025). A recently published study showed that in CLC3, PI(3,5)P2 facilitates the binding of its beta subunit TMEM9, leading to inhibition of the protein (Schrecker *et al*., 2025). Surprisingly, the deposited CLC3 structure (PDB ID: 9DNY) showed a PIP2 lipid binding also adjacent to helices E and F despite the absence of a pronounced F-G loop. Here, one of the phosphate groups (on carbon 3) engages primarily with helices E, F, while the lipid tail packed against helix C as well as the beta subunit TMEM9 (Figure 39 A-B).

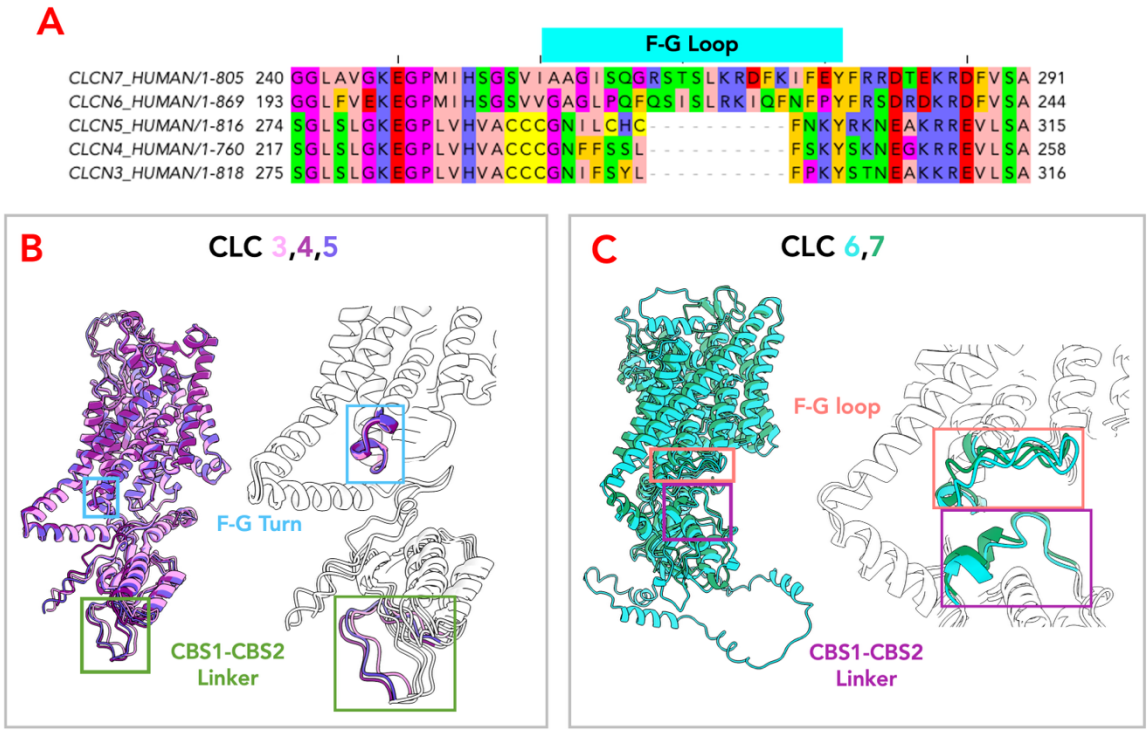


Figure 38. F-G Loop in Human CLC Transporters

A. Sequence alignment of human CLC transporters between helix F and G. CLC7 and CLC6 have long F-G loops, whereas CLC3, CLC4, and CLC5 only have short turns connecting the helices. B. AlphaFold 2 prediction of CLC3, 4, and 5 (in light pink, magenta and purple) with F-G turn. C. AlphaFold 2 prediction of CLC6, and CLC7 (in light cyan and teal), and zoomed in view of F-G turn region and CBS1-CBS2 linker in proximity.

In summary, it was intriguing that in both CLC3 and CLC7, PI(3,5)P2 acts essentially as a "molecular glue"—by bringing distinct domains or subunits into close proximity to regulate transport. The recurring theme of lipid-induced domain interactions within the CLC transporter family prompts speculation regarding whether such regulatory mechanisms arose independently among family members, and if it extends broadly across other membrane protein families. Literature on Kir 2.2 potassium channels (Hansen *et al*., 2012) revealed that it is another example of PIP2, PI(4,5)P2 specifically, being a molecular glue connecting soluble domains and transmembrane domains of the channel, though the binding of PIP2 led to activation of Kir channels rather than inhibition (**Chapter 1.2.2**; Hansen *et al*., 2012).

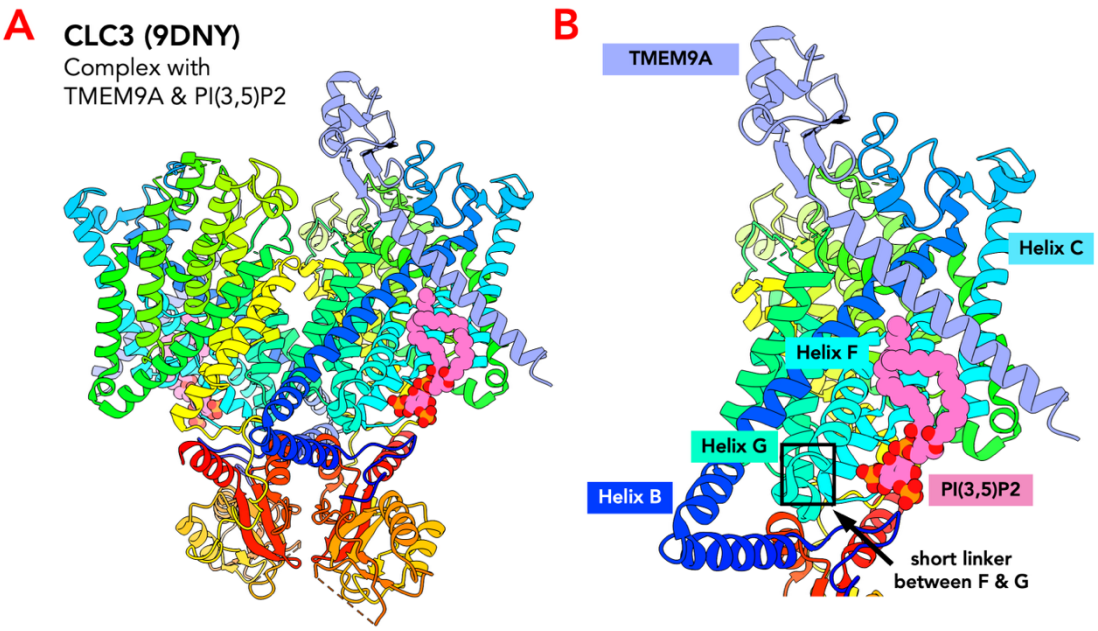


Figure 39. Binding of PIP2 to CLC3

A. Structure of homo-dimeric CLC3 (PDB 9DNY), both protomer coloured by rainbow (N- to C- terminus), bound PI(3,5)P2 shown in spheres, coloured in hot pink. Beta-subunit TMEM9A coloured in lavender. B. Zoomed in view of A (showing only one monomer on the right)., helix B, C, F, G labelled, black box outline indicated the short F-G turn. PI(3,5)P2 bind in proximity of F-G turn and TMEM9A.

3.4 Structure of CLC7-Y715C Revealed New Conformational States

Superimposition of the structure of the CLC7-Y715C/OSTM1 complex onto the WT PIP2 structure revealed significant global conformational differences in the transmembrane domain, in addition to the dissociation of CBS domains (Figure 40 A-J). Notably, a pronounced disorder-order transition was observed at the onset of helix B. Specifically, residue R113, initially positioned in the N-terminal loop, underwent a substantial shift of approximately 7.8 Å, aligning closer to the transporter's side and consequently forming an extended segment of helix B (Figure 40 B). Unlike helix B, helices C through G displayed relatively stable conformations, with minor fluctuations observed predominantly in helices H and I, which experienced modest shifts averaging 1.5 Å (Figure 40 C, D).

Remarkably, the I–J loop, a critical long loop spanning the luminal interface of the transporter, displayed substantial structural flexibility. The most pronounced shift occurred at residue D364, moving by approximately 4 Å (Figure 40 E). Additionally, significant positional adjustments were observed at the beginning of helix J, with residue I372 relocating approximately 3.1 Å toward the luminal side (Figure 40 F).

Helix L, composed of two short helices linked by an extended loop incorporating two beta-sheets, also underwent moderate positional adjustments. Specifically, residues G443 and H481 shifted by about 3.2 Å and 2.6 Å respectively, along with residues around it (Figure 40 G).

Helices M and N showed minimal conformational variability (Figure 40 H). Conversely, helices Q and R displayed significant disorder-order transitions, notably extending helix Q, facilitating interactions with the newly extended helix B as indicated above (Figure 40 J; discussed in more detail in **Chapter 3.5**). Such rearrangements prominently influenced preceding helices O and P (Figure 40 I).

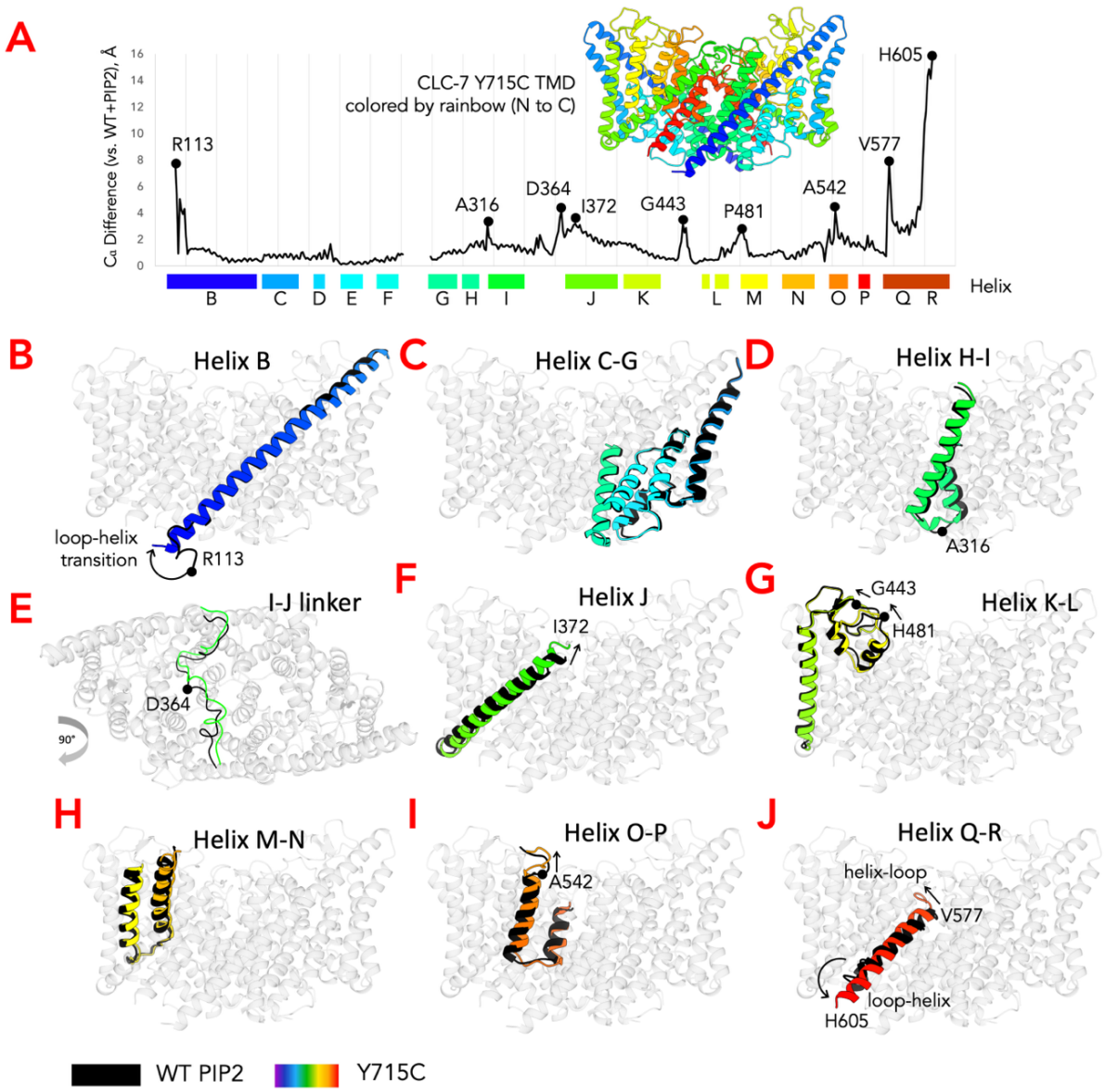


Figure 40. Global Conformational Changes in CLC7-Y715C

A. Per residue Cα RMSD between the transmembrane domain of WT PIP2 structure and Y715C structure, plotted against residue number. Structure of CLC7 Y715C shown, coloured by rainbow (N- to C- terminus), helix name and colour also annotated below the plot. Representative residues with significant structural shift indicated by black circle, and residue number annotated. B-J. Overlay of isolated helices onto CLC7 transmembrane domain (light grey), WT PIP2 structure in black, and Y715C structure coloured by rainbow (N- to C- terminus); representative residues with significant shift shown in spheres, arrows indicating direction of shift from WT PIP2 to Y715C: helix B (B), helix C-G (C), helix H, I (D), I-J linker (E), helix J (F), helix K-L (G), helix M, N (H), helix O, P (I), and helix Q, R (J).

Collectively, these local helix-specific structural modifications propagated into comprehensive global conformational shifts within the TMD (Figure 40 A). Notably, helices located

peripherally at the dimeric interface interacting with OSTM1—specifically, the C-terminus of helix B and J, along with helices C, K, and the N-terminus of helix L—remained relatively fixed. In contrast, helices crucial to the dimer interface (N-terminus of helix B and J, helices H, I, P, Q, R, and the I–J linker), alongside neighbouring helices L and O, demonstrated considerable conformational variability.

Importantly, these conformational adjustments significantly expanded the dimer interface's buried surface area (Figure 41 A, Supplementary Table 2). Specifically, the buried area increased nearly 1000 Å², changing from 2219 Å² in the wild-type structure with PIP2 to 3107 Å² in the Y715C mutant. The expansion primarily arose from the newly extended helices, particularly between helices B, Q, and R (Supplementary Table 2).

Furthermore, the conformational shifts within the dimer interface altered the residue interaction profiles. A critical example includes the disruption of the salt bridge between residues E313 and Q321 (distance 2.6 Å), previously observed in the wild-type PIP2 structure (Figure 41 B). The disruption of this salt bridge in the Y715C mutant (distance 5.3 Å) holds significant functional implications, as highlighted by Zhang *et al*. (2020). Their research demonstrated that mutational disruption of either E313 or Q321 to alanine markedly accelerated transporter activation, and the double mutation led to a twenty-fold increase in activation kinetics. Consequently, the absence of the salt bridge in the Y715C mutant aligns well with the hyperactivation phenotype previously documented, suggesting potential correlations between structural changes seen in this structure and functional dynamics of CLC7 slow gating.

Looking at the fast gate within each protomer, the Y715C mutation also influenced the permeation pathway compared to the WT. Specifically, the extracellular vestibule's bottleneck was partially widened, with the distance between residues E247 (the gating glutamate) and I515 expanding by approximately half an angstrom (Figure 42 A, B). Notably, the distance between I515 and K246 remained unchanged, and the gating glutamate conformation persisted

in the up position, suggesting that the mutation mostly affects the slow gate rather than the fast gate.

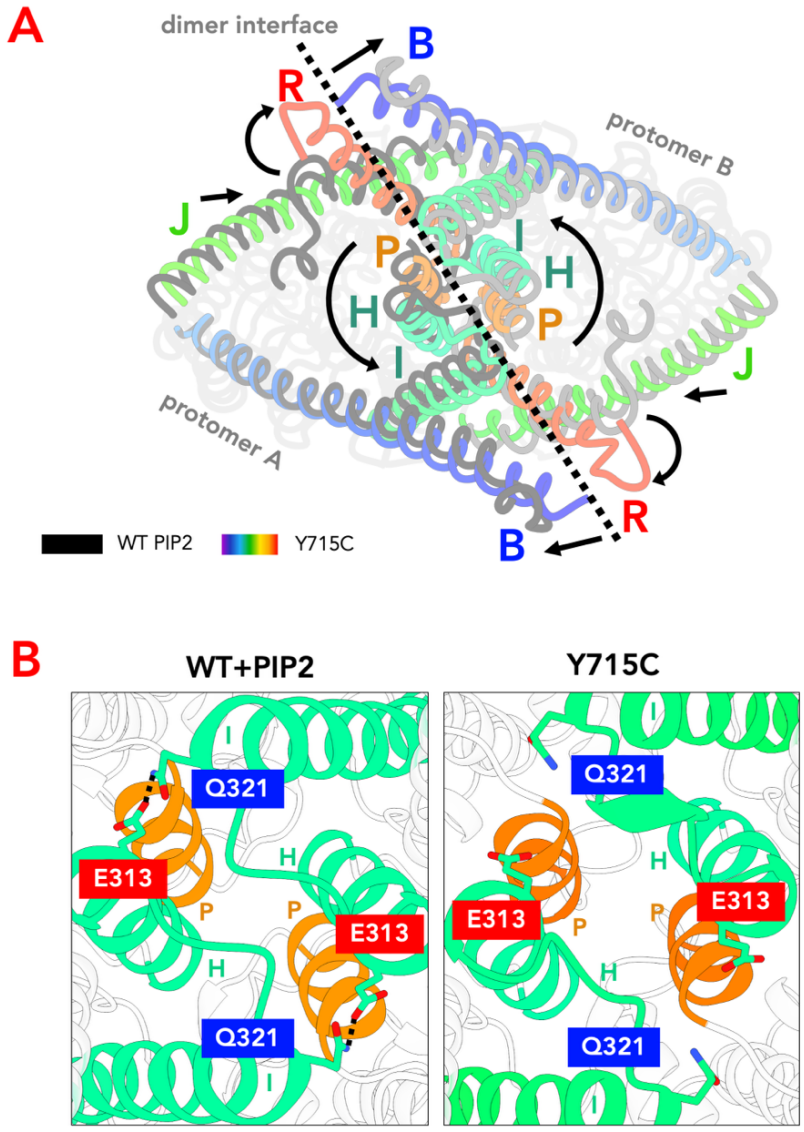


Figure 41. Conformational Changes in CLC7-Y715C Dimer Interface

A. Super-imposition of helices at the dimer interface of CLC7 transmembrane domain (helix B, H, I, J, P, and R), WT PIP2 in black and dark grey, Y715C coloured by rainbow. View from the cytosol, CBS domains in WT PIP2 structure hidden. Arrows indicate direction of shift from WT PIP2 structure to Y715C structure. Black dotted line indicates dimer interface at the transmembrane domain. B. Zoomed in view (from the cytosol) of helices at the dimer interface: helix H (green), I (green), P (orange), rest of the transmembrane domain in white. Residues E313 and Q321 shown (labelled in red and blue). Black dotted lines indicate ionic bonds between the residues in WT PIP2 structure (left), which is broken in the Y715C structure (right).

Substantial alterations were also observed in chloride coordination sites, predominantly toward the cytosolic part of the transport pathway (Figure 42 C). While the WT structure resolved three distinct chloride densities, the Y715C mutant primarily exhibited a prominent density only at the external site. A particularly remarkable shift involved the highly conserved residue Y602, located at the start of helix R, which typically contributes to cytosolic constriction. In the Y715C mutant, Y602 relocated approximately 15 Å outward, integrating into the extended Q helix, opening the cytosolic pathways to the cytosol. This residue, crucially implicated in the slow gating mechanism and proton coupling in related CLC transporters (specifically in CLCec1), was now absent from the chloride permeation pathway. Consequently, chloride densities at the central and internal binding sites in the Y715C structure were notably weaker. Moreover, proton glutamate E314, previously linked to proton coupling, was exposed to the same cytosol-facing vestibule (Figure 42 C).

Lacking the ability to form an occluded state, the Y715C mutant structurally resembles an open channel more closely than a transporter, supporting previous studies on the functional importance of Y602 and E314 (Dutzler *et al.,* 2003, Walden *et al*., 2007). Further investigations are required to determine whether the Y715C mutation explicitly leads to decoupled chloride transport.

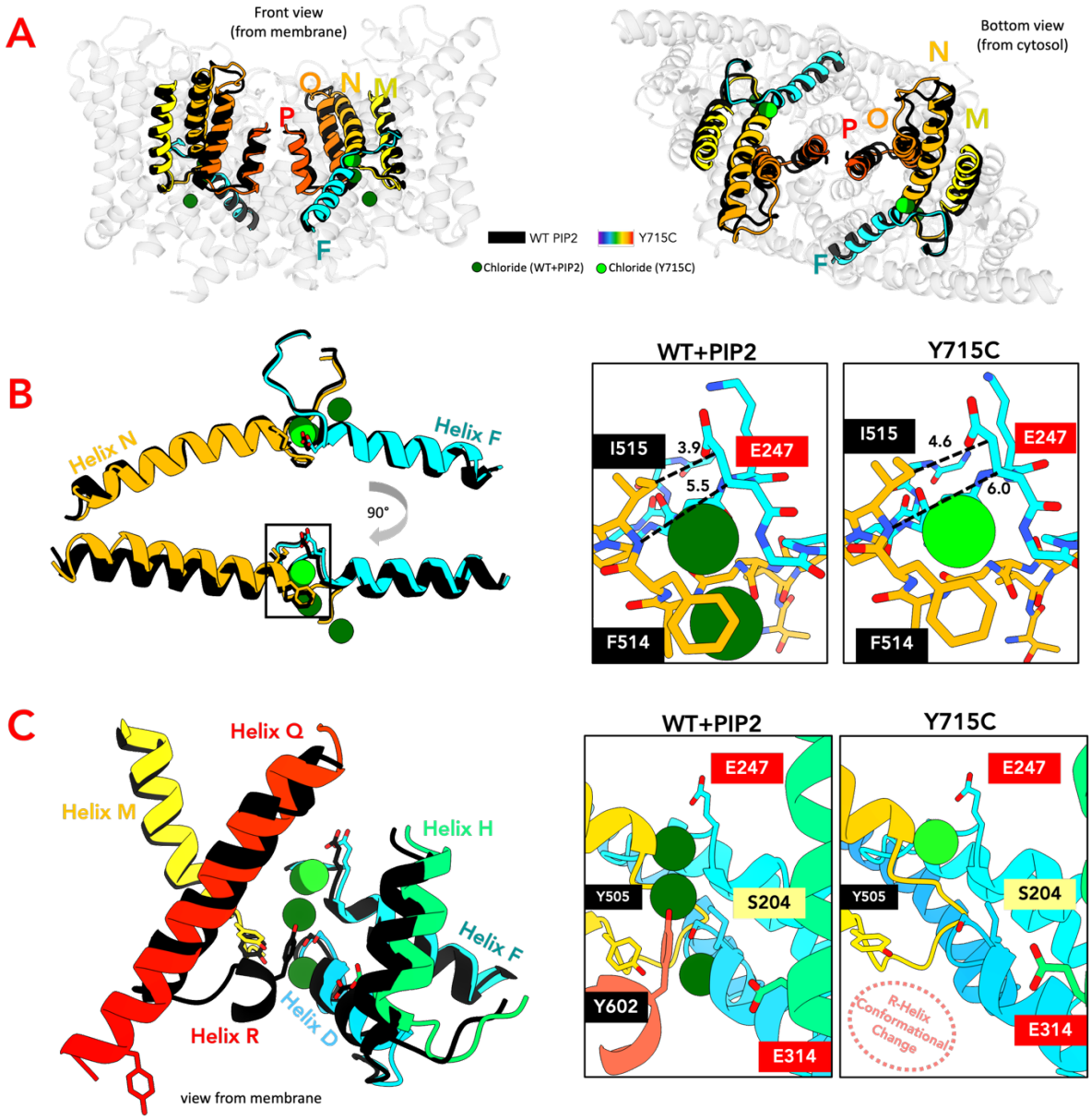


Figure 42. Shifts in Fast Gate of CLC7-Y715C

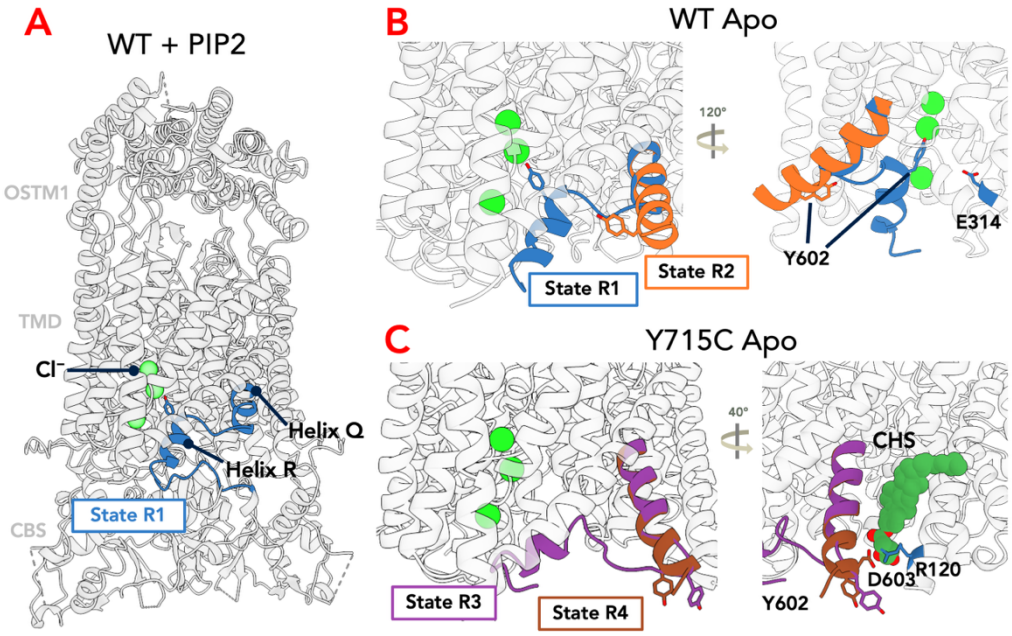
A. Super-imposition of helices involved in fast gating in CLC7, on transmembrane domain of WT PIP2 structure (light grey), helices F, M, N, O, P in WT PIP2 structure coloured black, corresponding helices in Y715C structure coloured by rainbow. Chloride ions coloured dark green in WT PIP2 structure, bright green in Y715C structure. Left: view from the lipid bilayer; right: view from the cytosol. B. Left: isolated super-imposition of helices F and N, coloured the same as A. Right: comparison of luminal chloride constrictions between WT PIP2 structure and Y715C. Distance between I515-Cγ1 and E247-Cγ, I515-N and Cβ shown. C. Left: isolated super-imposition of helices involved in cytosolic chloride constriction: D, F, H, M, Q, and R, coloured the same as A. Right: comparison of cytosolic chloride constrictions between WT PIP2 structure and Y715C.

3.5 Conformational Changes in R Helix Corresponded to CBS Dissociation

As previously described, the opening of the cytosolic vestibule within the chloride permeation pathway is predominantly associated with conformational changes in the R helix. Existing literature has proposed that the R helix plays a pivotal role in translating cytosolic signals and conformational adjustments within the CBS domains into the transmembrane domain (TMD), thus regulating transport functions. However, until now, there were no structural confirmations of significant alterations in the R helix within any CLC members.

Our findings, including the wild type (WT) apo structure and the Y715C mutant structure, represent the first instances in which substantial conformational shifts in the R helix have been identified structurally. Traditionally, in all known structures of CLC family proteins, including previously published CLC7 structures (PDB IDs: 7JM7, 7CQ5, 7CQ6, 7CQ7, 7BXU; Supplementary Figure 6), a characteristic "zigzag" configuration of helices Q and R has consistently been observed. Namely, helix Q resides at the periphery of the transmembrane domain and is connected via a linker looping back towards the centre of the transporter. This linker leads to a critical tyrosine residue (Y602 in humans) situated at the tip of helix R, forming the cytosolic constriction site for chloride ions. From there, helix R extends towards the cytosol, leading to a linker and then the CBS domains (Figure 43 A).

In contrast, our WT apo and Y715C structures revealed multiple conformational states of helix R, indicated by overlapping electron microscopy (EM) densities. These alternative states have been designated as states R2, R3, and R4, with the canonical state in the structure of WT PIP2 termed R1 (Figure 43 A-D). In the WT apo structure specifically, clear EM densities supported both the canonical R1 state and an equally populated alternative state R2 (Figure 43 B). In state R2, the linker between helices Q and R, along with the initial residues of helix R, swung dramatically outward from the central transporter core towards the periphery.



D

Structure	CLC7+PIP2	CLC7 Apo	CLC7$_{Y715C}$
State Helix Q & R	R1	R1 OR R2	R3 OR R4
Chloride Pathway	Closed	Closed or Cytosolic Open	Cytosolic Open
CBS Domain	Resolved (Rigid)	Not Resolved (Disordered/Flexible)	Not Resolved (Disordered/Flexible)

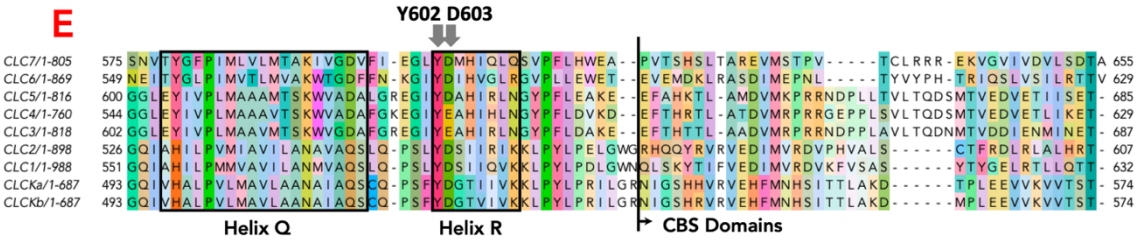


Figure 43. Novel Conformation of R Helix Identified in WT Apo and Y715C

A. Structure of WT PIP2 structure, with helix Q and R of one protomer coloured blue ("state R1"); chloride ions coloured green, note that central chloride is restricted by Y602 (side chain shown). B. Views of helix Q and R in WT apo structure, with two alternative conformations shown in blue ("state R1") and orange ("state R2"); chloride ions coloured green. In state R2, Y602 is removed from chloride permeation pathway. C. Views of helix Q and R in Y715C structure, with two alternative conformations shown in purple ("state R3") and brown ("state R4"); chloride ions coloured green. In both states, Y602 is removed from chloride permeation pathway. In addition, D603 in state R4 (brown) is brought into proximity to R120 (coloured blue), stabilized by a cholesteryl hemisuccinate molecule (coloured green). D. Table summarising the states of Q-R helix, openness of the chloride pathway, and the state of CBS domains in the three CryoEM structures. E. Sequence alignment of human CLC members around helix Q and R, arrows indicating the conserved Y602 and D603.

This outward movement effectively extended helix Q by two helical turns, displacing the conserved residue Y602 by approximately 15.5 Å. This movement led to significant opening of the cytosolic vestibule.

The Y715C structure displayed two additional novel states, R3 and R4 (Figure 43 C). Unlike the WT apo structure, both R3 and R4 showed an extended helix Q conformation and a pronounced outward shift of residue Y602. Importantly, due to the notable global conformational changes within the TMD described in previous chapters, the conformational changes in R helix in the Y715C structure specifically led to new interactions in the dimer interface. Specifically, residue D603, which also became part of the extended helix structure, was observed forming a new salt bridge with residue R120 located on helix B (in state R4). This interaction notably enlarged the dimer interface, a phenomenon further stabilized by the presence of a cholesterol hemisuccinate molecule observed in the EM density. The newly formed ionic interaction between D603 and R120 was presumably a major driving force behind the stabilisation of this "open" conformation (Figure 43 C).

In examining sequence conservation, Y602 is universally conserved across human CLC proteins, explained by its critical role in chloride coordination at the proto gate (Figure 43 E). Aspartic acid 603 is conserved in CLC1, CLC2, CLC5, CLC6, CLC-Ka, and CLC-Kb, while being replaced by a glutamic acid residue in CLC3 and CLC4. In contrast, residue R120 on helix B is only conserved within CLC6, the closest relative to CLC7. Although positively charged residues are common in similar regions across various CLC proteins, their capacity to form ionic interactions with the conserved acidic residues on helix R remains unclear. Nonetheless, it seems plausible that CLC6 could adopt this same open conformation observed in the Y715C structure, stabilised by a similar ionic bridge between helices R and B.

Collectively, our findings demonstrate that loss of PIP2 binding correlates with the dissociation of the CBS domains, an extended conformation of helix Q incorporating residue Y602, and

subsequent opening of the cytosolic vestibule, as evidenced in both WT apo and Y715C structures. Moreover, the Y715C mutation introduces an additional global conformational shift, resulting in further widening of the chloride transport pathway towards the luminal side, alongside enhanced structural stability facilitated by novel ionic interactions and cholesterol binding. These observations pose intriguing questions for further study, such as how exactly the Y715C mutation induces these global structural changes, whether similar conformations could be adopted by wild-type proteins, and what conformational dynamics the CBS domains might undergo upon their dissociation from the TMD.

3.6 Discussion & Chapter Conclusions

In this chapter I have presented structural studies of human CLC7 in two distinct states: the PIP2-bound wild-type (WT) structure and the hyperactivating Y715C mutant, together with a WT apo structure in which the lipid is absent. The structures converge on a consistent message: PIP2 is a key stabiliser of interactions between the regulatory cytosolic CBS domains and the transmembrane domain that houses the chloride/proton transport pathway. When PIP2 is present the CBS domains form defined, ordered contacts with the TM domain, but when PIP2 is lost (WT apo) or when the protein is unable to bind PIP2 (Y715C), the CBS domains become detached, mobile and are no longer resolvable by cryo-EM. These observations argue that lipid binding at the cytosolic interface is central to maintaining the quaternary architecture of CLC7.

The Y715C mutation produces more than local perturbations: it induces global conformational rearrangements within the TM domain. The most striking change is in helix R, which swings outward from the TM core and thereby opens the cytosolic entry to the transport pathway. This outward movement is accompanied by large-scale rearrangements throughout the TM helices. Together with the observed loss of ordered CBS density, these structural changes provide a plausible structural explanation for the hyperactive phenotype of the mutant.

When placed in the context of previously published CLC7 structures (for example PDB entries 7JM7 and 7BXU), the WT PIP2-bound conformation shares key features with an occluded state of the transporter. Specifically, both luminal and cytosolic constrictions for chloride are present and the gating glutamate adopts an 'up' conformation, consistent with a closed/occluded ion conduction path. Because these hallmarks are conserved across multiple WT structures, we interpret the PIP2-bound map as representing an inactive, occluded state.

By contrast, the WT apo and Y715C structures are notable as the first human CLC family structures in which the CBS domains are completely dissociated from the TM domain. The Y715C structure is also among the first CLC structures to display global TM rearrangements as a principal change between states. This is a behaviour previously less commonly ascribed to CLC transport-cycle transitions, which are often interpreted primarily through local movements of the gating glutamate. These observations raise the possibility that CLC7 can access alternative, more global modes of conformational change during activation or gating; however, additional comparative structures and functional data will be required to establish the physiological relevance and the sequence of events.

Insight into plausible activation pathways can be drawn by comparison to recent work on the bacterial CLC homolog ec-CLC by Fortea *et al*., where cryo-EM structures were solved under a range of pH and chloride conditions. In their work, at 100 mM chloride and acidic pH (pH 4.5) they observed three distinct conformations — termed "swap," "turn," and "twist" (Figure 44 A-C, figure 45 D), whereas at neutral pH, or in the absence of chloride, the most particles populated the "swap" conformation. Under low pH and no chloride, the population split roughly half in "swap" and half in "turn."

The author suggested that the 'swap' conformation resembles an occluded state in which a domain swap between protomers maintains a tight dimer interface (3,630 Å^2, Figure 44 A), whereas the 'turn' and 'twist' conformations feature displacement of a transmembrane helix

(helix A) from the dimer interface and progressive reductions in TM dimer contact area, from 2,148 Å² in "turn" and 1,945 Å² in "twist" conformations (Figure 44 B, C).

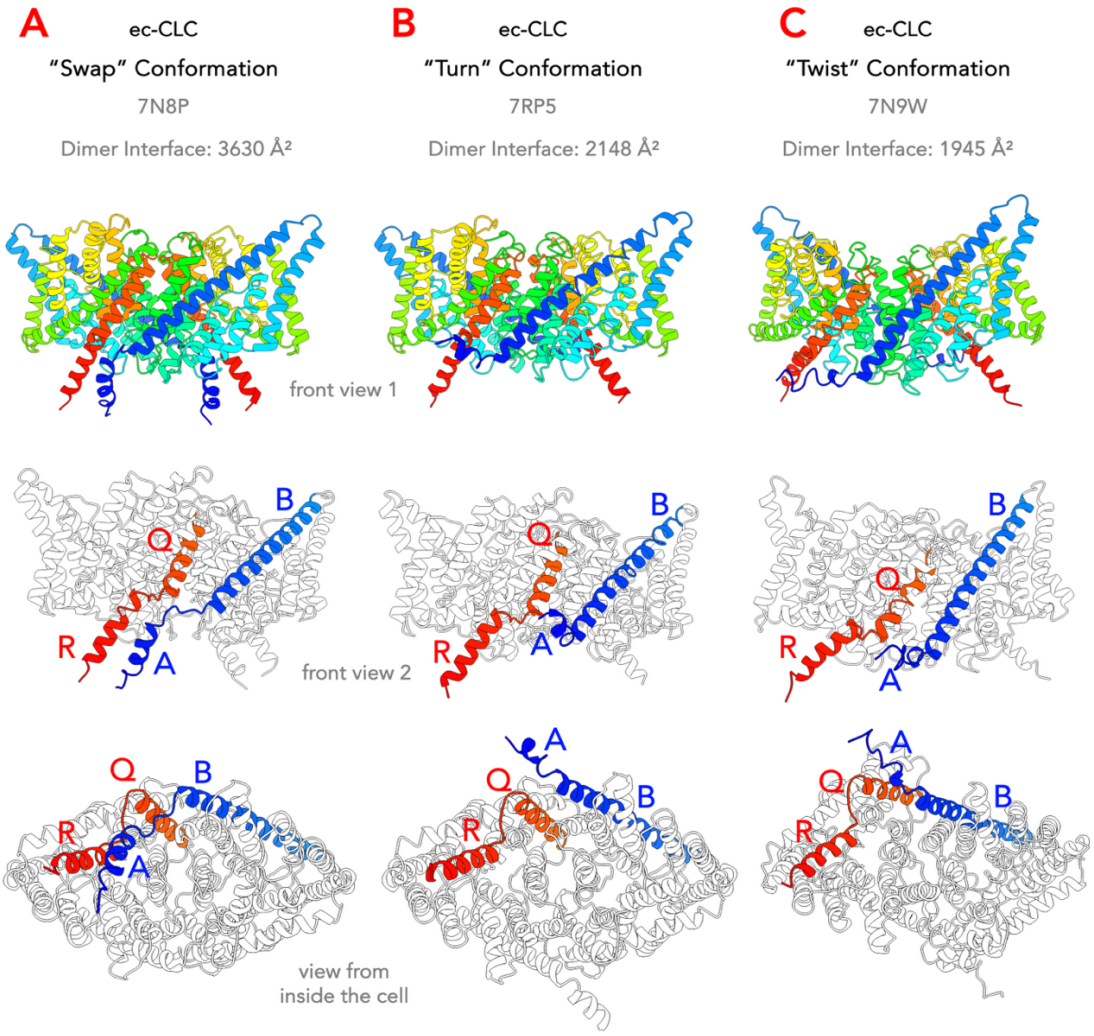


Figure 44. Conformational Changes in ec-CLC Upon Activation

A-C. Structures of E. coli CLC transporter in various stages of activation (Fortea *et al.*, 2024). Top: structures coloured by rainbow in each chain (N- to C- terminus), view from lipid bilayer/micelle ("front view"). Middle: structures in white except helix A, B (blue), and Q, R (red), view from lipid bilayer/micelle. Bottom: same colouring as middle, view from inside the cell ("bottom view"). A. "Swap" conformation (PDB: 7N8P). B. "Turn" conformation (PDB: 7RP5). C. "Twist" conformation (PDB: 7N9W).

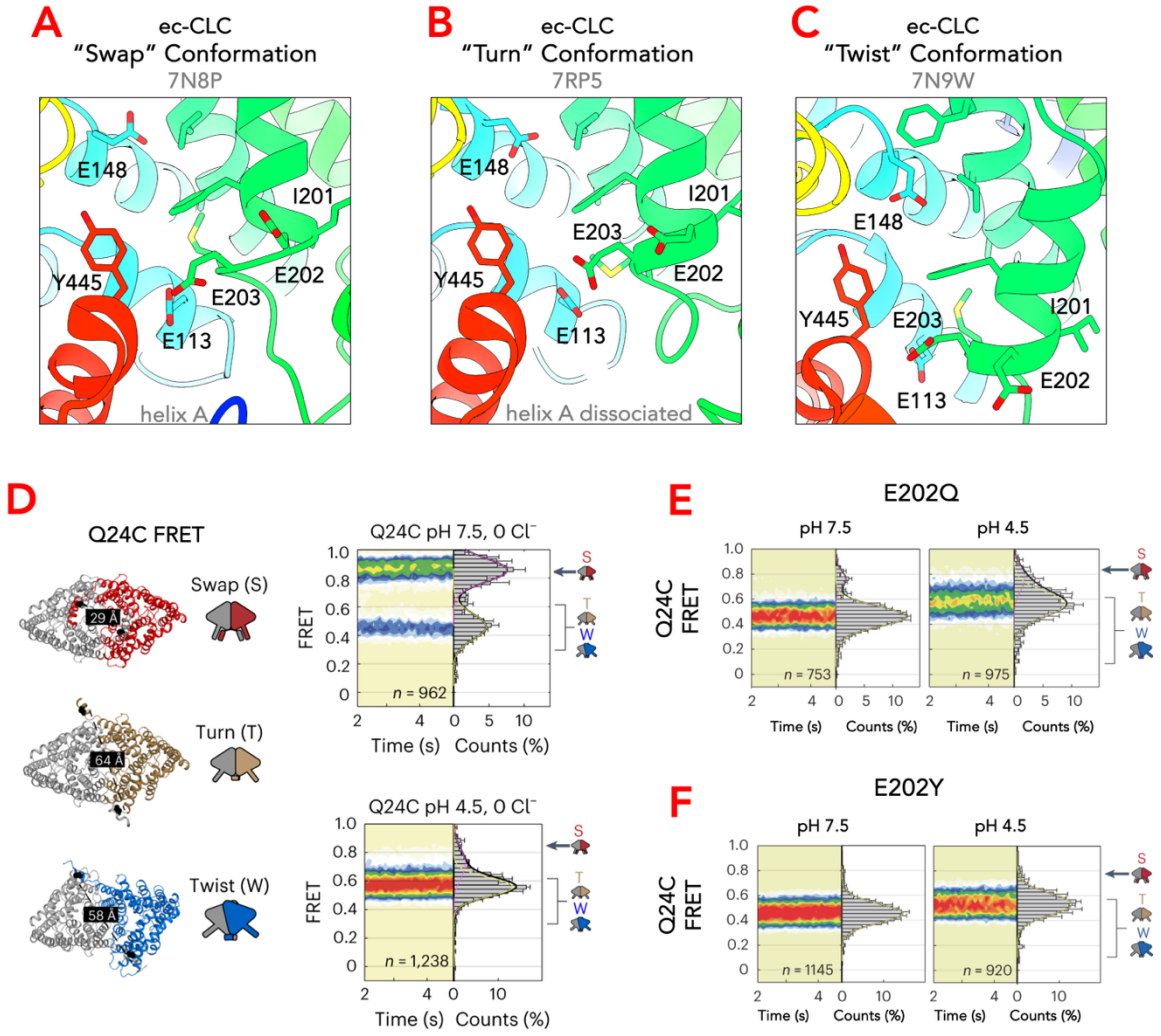


Figure 45. Single-molecule FRET of ec-CLC

A-C. View of chloride permeation pathway within the ec-CLC structures under various conformations. Helices coloured by rainbow (N- to C- terminus), side chains of important residues shown: E148 (gating glutamate), E203 (proton glutamate), Y445 (cytosolic chloride constriction). A. "Swap" conformation. B. "Turn" conformation. C. "Twist" conformation. D-F. Single-molecule FRET data adapted from Fortea *et al.*, 2024. D. Schematic diagrams of smFRET experimental design. Left: labelled residue Q24 on helix A shown as black spheres on CryoEM structures, distance between Q24 on opposite protomers in each conformational states annotated. Right: Population contour plots of labelled ec-CLC at pH 7.5 in 0 mM chloride (top), and pH 4.5 in 0 mM chloride (bottom). At pH 7.5, ec-CLC is split between Swap (S), Turn (T), and Twist (W) states, whereas at pH 4.5, T and W states are dominant. E. Population contour plots of labelled ec-CLC E202Q mutant at pH 7.5 or pH 4.5. F. Population contour plots of labelled ec-CLC E202Y mutant at pH 7.5 or pH 4.5.

To support this hypothesis, the authors performed single-molecule FRET experiments on Q24C-labeled mutants of ec-CLC. These experiments allowed monitoring of helix A movements through the distance between C24 residues (residing on helix A) on opposing protomers (Figure 45 D). At neutral pH, ec-CLC primarily exhibited two populations: one with

a 29 Å inter-residue distance corresponding to the inactive “swap” conformation, and another with distances of 55–65 Å, consistent with a mixture of “turn” and “twist” states. At pH 4.5, the protein population shifted predominantly toward the low-FRET configurations of “turn” and “twist” states (Figure 45 D), aligning with the cryo-EM observations.

Furthermore, mutagenesis of a conserved acidic residue, E202 (Figure 45 A-C)—located near the proton-coupling residue E203 and the only titratable residue exposed to cytosolic solvent in the inactive “swap” conformation, provided additional insight. The authors hypothesized that protonation of E202 initiates the activation sequence in ec-CLC, by promoting the displacement of domain-swapped helix A. This was supported by FRET results showing that the E202Q mutation shifted the protein population to the “twist” and “turn” conformations even at neutral pH, confirming that protonation of E202 alone can trigger dissociation of helix A and initiate the inter-protomer rearrangements associated with activation, which the author suggested to be the mechanism of slow gating in ec-CLC (Figure 45 E).

Another mutant, E202Y, which disrupts the dimer interface, also yielded a FRET profile consistent with the “turn” and “twist” states across both pH conditions. This further emphasizes the role of the dimer interface and E202 in slow gating of ec-CLC (Figure 45 F).

Interestingly, E202 in ec-CLC is conserved as E313 in human CLC7, directly preceding the proton glutamate E314 (Figure 46 A, B). In the WT PIP2-bound human structure (presumed to be in an occluded state), E313 is buried and shielded from the cytosol by a loop extending from helix R into the CBS domain. If a similar protonation-based activation mechanism operates in CLC7, this loop, and possibly the entire CBS domain, would need to dissociate first to expose E313 to solvent. This aligns with observations from both our WT apo and Y715C structures, where the CBS domains are detached (Figure 46 C, D). These differences suggest that while ec-CLC and human CLC7 share conserved residues and structural themes, the mechanisms of slow gating may diverge. In ec-CLC, protonation of E202 appears to induce helix A

dissociation and gating. In contrast, CLC7 may rely on CBS domain detachment—possibly triggered by PIP2 dissociation—to expose key titratable residues such as E313 and E314, thereby initiating activation.

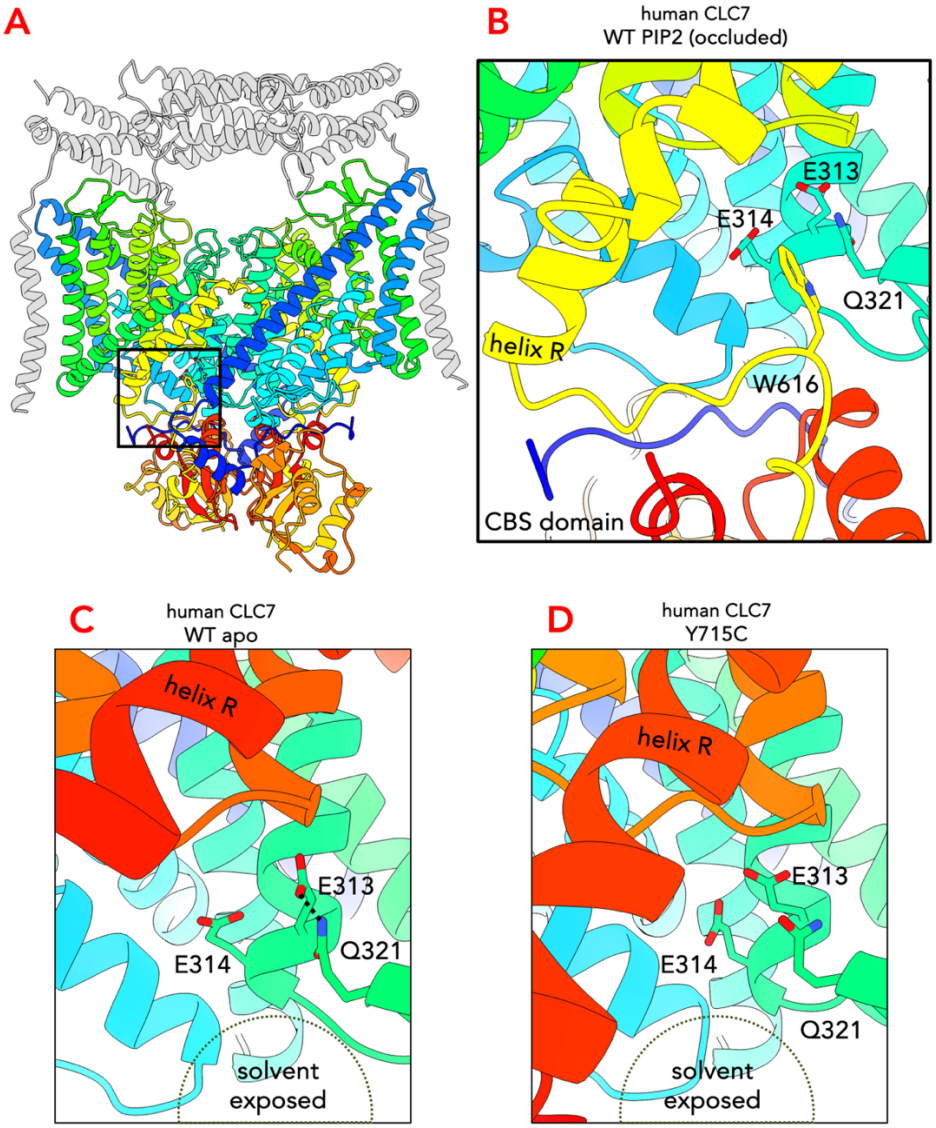


Figure 46. Dissociation of CBS Domain May be Required for Protonation of Human CLC7

A, B. Structure of human CLC7 bound to PIP2 (WT PIP2). Beta-subunit OSTM1 coloured in light grey, CLC7 coloured by rainbow in each protomer (N- to C- terminus), helix R labelled. Black box indicates cytosolic chloride permeation pathway. E314 (proton glutamate, E203 in ec-CLC), E313 (E202 in ec-CLC), Q321, and W616 side chains are shown, CBS domains attached to TMD. C, D. Cytosolic chloride pathway of human CLC7 apo (WT apo, C) or Y715C (D). Helices coloured by rainbow (N- to C- terminus), helix R labelled. Side chains of E314, E313, Q321 shown. CBS domains detached from TMD.

Moreover, WT PIP2 and WT apo structures revealed that E313 forms an ionic interaction with Q321 on the opposing protomer. This interaction is disrupted in the Y715C mutant, accompanied by global rearrangements (Figure 46 C, D). Thus, the WT apo structure might correspond to the "swap" state in ec-CLC, in which titratable acidic residues are solvent-exposed but not yet protonated. Meanwhile, the Y715C structure could represent an active "turn" or "twist"-like state, in which protonation (presumably of E313) breaks inter-protomer salt bridges, leading to activation.

In summary, through structural comparisons between our human CLC7 cryo-EM structures and the ec-CLC structures under various states, we can propose a tentative activation mechanism for CLC7 (Figure 47 B).

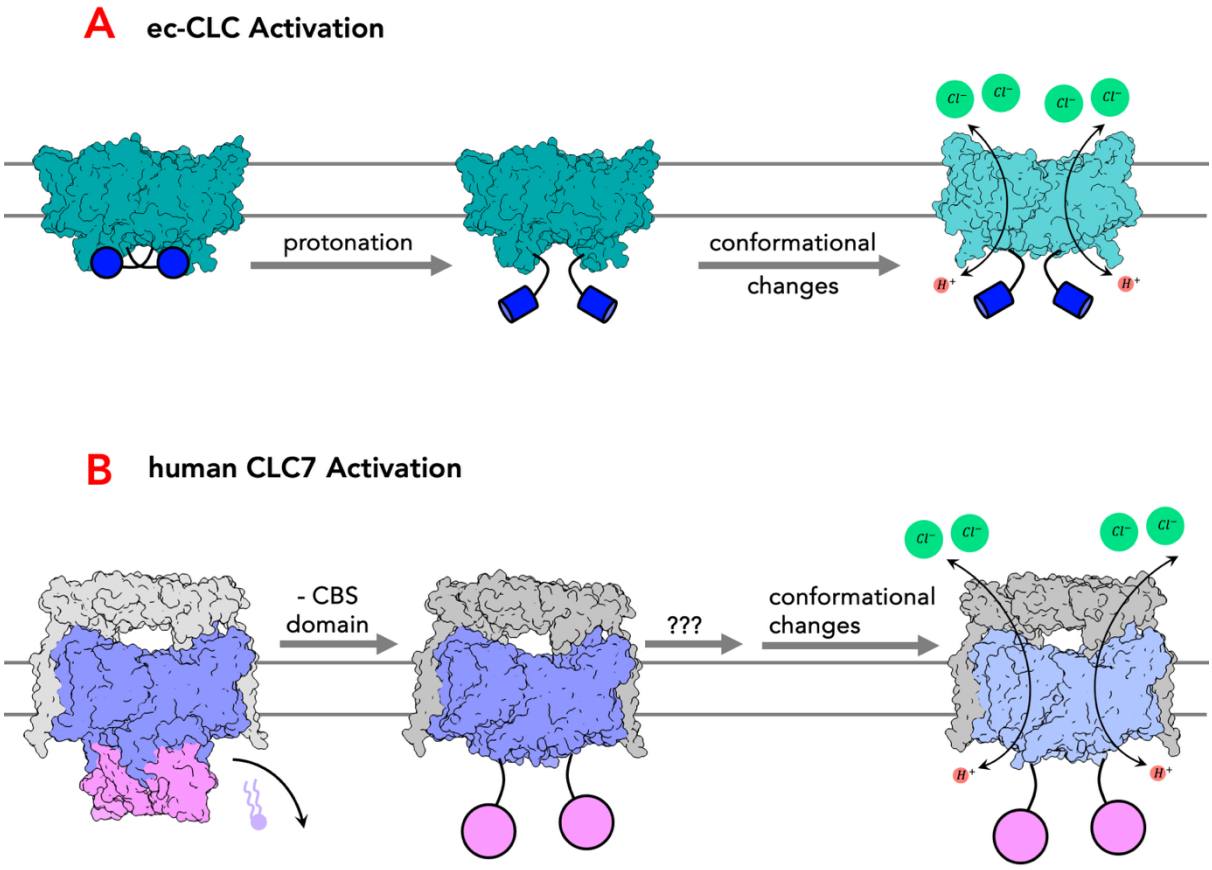


Figure 47. Schematic Diagram of Hypothetical Activation Mechanism of ec-CLC and Human CLC7

A. ec-CLC (teal) has domain-swapped helix A (dark blue) in the inactive state. Upon protonation of E202, helix A dissociate, and the TMD can undergo conformational changes which results in an active transporter. B. In human CLC7, CBS domains (pink) represent another layer of regulation. Dissociation of PIP2 (light purple), and subsequently CBS domains, exposes titratable residues such as E313 and E314, and potentially through protonation and breaking of ionic interactions between E313 and Q321, the TMD (slate blue) can undergo conformational changes which activate transport.

Dissociation of PIP2 destabilizes CBS–transmembrane interactions, leading to detachment of the CBS domains and exposure of titratable residues like E313. Subsequent protonation of E313 may disrupt ionic interactions (e.g., with Q321), initiating conformational rearrangements in the transmembrane domain that activate the transporter.

However, further functional validation is required to determine whether E313 protonation is indeed a critical step in CLC7 activation, or whether other residues, such as E314, are responsible. Given that the CBS domain detachment in CLC7 results in a significantly more solvent-exposed cytosolic environment than in ec-CLC, it is also possible that CLC7 has evolved a distinct biochemical activation pathway, potentially enabling regulation by additional cellular signals or under different physiological conditions.

4. Synthetic Nanobodies Against CLC7

4.1 Selection & Analysis of Binding

To gain deeper insights into the functional roles of the CBS domains and to facilitate the study of endogenous CLC7 using advanced techniques such as super-resolution microscopy, the generation of high-affinity, specific binders became essential. Given that the luminal side of the transporter complex, particularly OSTM1, is extensively glycosylated, targeting the CBS domains on the cytosolic side was more desirable. Cytosolic binders could also provide greater flexibility in experimental applications, including live-cell imaging, immunoprecipitation of endogenous protein complexes, and cryo-electron microscopy or tomography studies without the destruction of lysosomes, allowing understanding of CLC7 in more native context.

To achieve the goal of producing specific, high affinity nanobodies, a synthetic nanobody pipeline developed by Zimmerman *et al*. was employed. This method required minimal amounts of protein, advantageous given the typically low yield (less than 100 µg/L) of CLC7 during purification. The approach facilitated selection under detergent conditions and enabled ligand supplementation, enhancing the specificity and affinity of the resulting binders to a specific state of the target protein.

Initially, due to the absence of structural details on how phosphatidylinositol-3,5-bisphosphate (PIP2) stabilizes the CBS domains against the transmembrane domain (TMD), sybody selections were performed without additional PIP2 (Figure 48 A). To selectively enrich cytosolic binders, purified and biotinylated CLC7 without OSTM1 was utilized in the early stages, including ribosome and phage display steps (Figure 49 A). The final ELISA screening employed biotinylated CLC7/OSTM1 complexes to exclude binders recognizing the luminal side, aiming primarily for CBS domain binders. This approach yielded eleven unique synthetic nanobodies (sybodies) from three libraries (Figure 48 B).

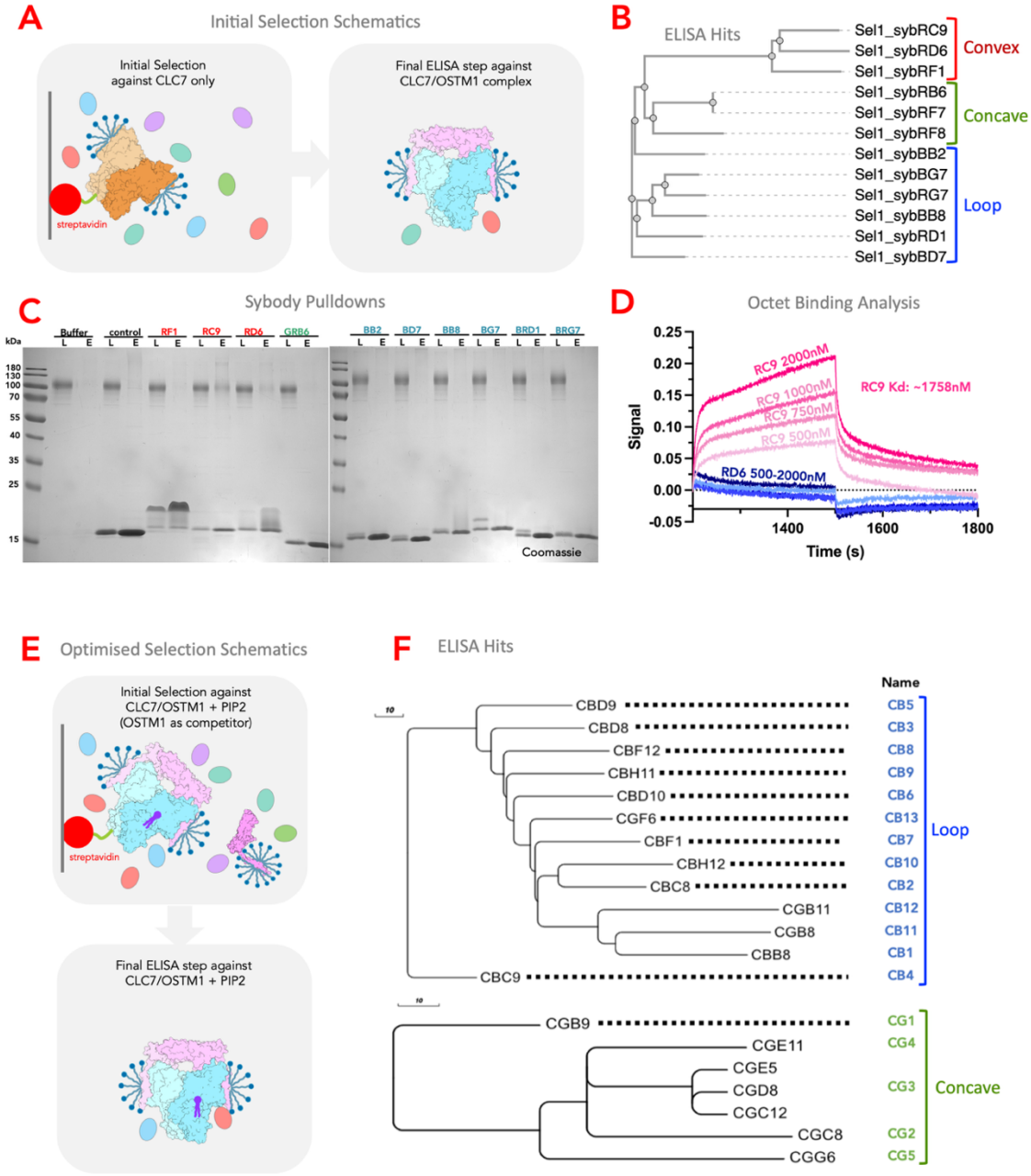


Figure 48. Synthetic Nanobody Selection Against CLC7

A. Schematic diagram of the initial attempt to select sybody against the CBS domains of CLC7, where ribosome display and phage display steps were carried out with purified and biotinylated CLC7 alone (left), and the final ELISA step was carried out against purified and biotinylated CLC7/OSTM1 complex. B. Phylogenetic tree of hits identified from the initial sybody selection, from convex, concave, and loop library. C. Ni-NTA pulldowns of identified sybodies from the initial selection. L: "load", samples prior to application on Ni-NTA resin. E: "eluent", samples eluted from Ni-NTA resin. Buffer control and control sybody was included. Test sybodies coloured by the library they are from (red for convex library, green for concave library, blue for loop library). CLC7 was enriched in the eluent of RC9 and RD6. D. Octet measurements of sybody binding to immobilised CLC7. 75 nM of biotinylated CLC7 was immobilised on streptavidin tips, which were submerged in buffers containing RC9 and RD6 at 500, 750, 1000, and 2000 nM for 300s, then buffers without sybodies for 300s. E. Schematic diagram of the optimised attempt to select sybody against the CBS domains of CLC7, where biotinylated CLC7/OSTM1 complexes supplemented with PIP2 were used throughout the selection, and un-biotinylated OSTM1 was used in one step of phage display as competitor. F. Phylogenetic tree of hits identified from the optimised sybody selection, from concave and loop library.

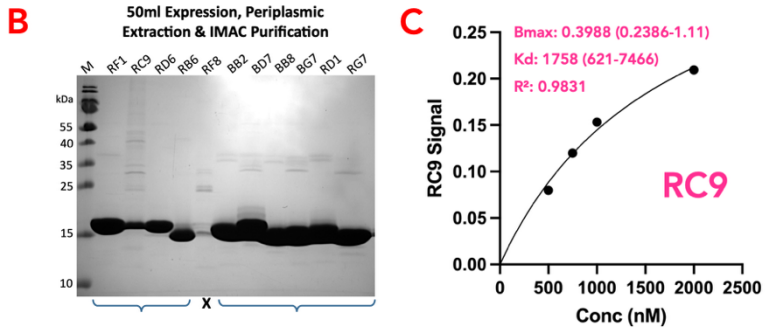
The purified sybodies were expressed in *E. coli* MC1061, periplasmically extracted, and purified using Ni-NTA resin followed by overnight dialysis (Figure 49 B). Subsequent pulldown assays indicated that only two sybodies, RC9 and RD6, showed weak binding to CLC7 after rigorous washing (Figure 48 C).

To further investigate their binding kinetics, I performed biolayer-interferometry on the Octet system. In brief, biotinylated CLC7 was immobilised on streptavidin tips at 75 nM, and dipped sequentially into buffers with or without sybodies at particular concentrations. Binding was measured as changes in interference patterns of reflected white light, resulting from size changes of immobilised molecules. Octet experiments demonstrated a dose-dependent interaction only for sybody RC9, which displayed a rapid dissociation rate and an estimated dissociation constant (K_d) of approximately 2 μM (Figure 48 D). This limited success suggested that most selected sybodies may have targeted the luminal side or failed to enrich effectively due to dynamic conformational changes of CLC7 without PIP2 during the selection process.

With additional structural insights, a second optimized sybody selection strategy was devised. This involved using the CLC7/OSTM1 complex supplemented with PIP2 throughout the entire selection process to maintain the transporter in its most stable state (Figure 48 E, Figure 49 D, Supplementary Figure 4). Purified OSTM1 was also used in one of the phage display stages as a competitor to favour cytosolic binder. This refined method successfully produced eighteen unique sybodies from only two libraries, all of which purified well (Figure 48 F, figure 49 E).

A Initial Selection

Sybody Library	Phage Display Library Enrichment		
	Round One	Round Two	Round Three
Small (Concave)	2.10	0.35	1.03
Medium (Loop)	0.66	1.75	17.85
Large (Convex)	0.64	0.77	35.81



D Optimised Selection

Sybody Library	Phage Display Library Enrichment		
	Round One	Round Two	Round Three
Small (Concave)	6.36	0.18	12.71
Medium (Loop)	4.58	0.7	6.56

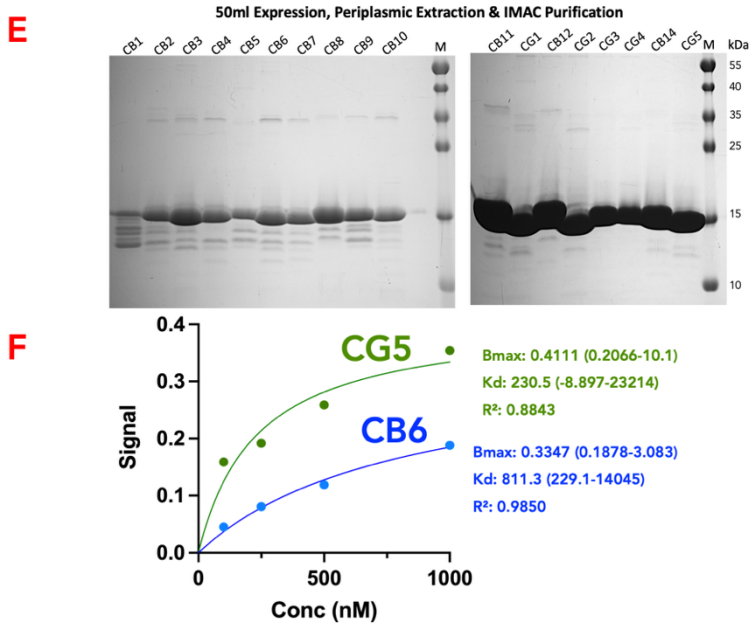


Figure 49. Results of Initial and Optimised Sybody Selection Against CLC7

A. Results of initial selection against CLC7. Phage display enrichment for Convex, Concave, and Loop library against negative control PepTso. Enrichment was calculated as the ratio between isolated phages against CLC7 and that of the negative control, quantified by qPCR. An enrichment larger than 2 usually indicates successful enrichment of target-specific sybodies. B. SDS-PAGE analysis of purified sybodies from the initial selection. Lane labels indicate different sybodies tested; molecular weight marker (M) is shown on the left. A strong band around 15 kDa corresponds to the expected size of sybodies. C. Dose–response binding curve of sybody RC9 measured against its target. Data were fitted to a one-site specific binding model. D. Enrichment factors from the phage display for Concave and Loop libraries following an optimized selection strategy. E. SDS-PAGE analysis of purified sybodies from the optimized selection. F. Dose–response binding curves of sybodies CG5 (green) and CB6 (blue). Nonlinear regression fits were applied to determine K_d and B_{max} values.

Pulldown assays using immobilized CLC7/OSTM1 complexes revealed that ten out of eighteen sybodies were effectively enriched, with eight demonstrating significantly stronger binding compared to the initial screen's sybodies (Figure 50 A). Additional experiments using CLC7 alone, both with and without PIP2, showed that two sybodies, CG4 and CG5, consistently associated with CLC7, indicating a likely interaction with the CBS domains (Figure 50 B). Other sybodies appeared to target either OSTM1 or epitopes present only on the complex.

To further elucidate the binding epitopes, epitope binning experiments using biolayer interferometry were conducted on four selected sybodies: CG4 and CG5, which were hypothesized to bind CBS domains, and CB6 and CG3, which potentially bound elsewhere Biotinylated CLC7/OSTM1 complexes supplemented with PIP2 were immobilized at a concentration of 75 nM on streptavidin biosensor tips, followed by sequential exposure to sybodies at a concentration of 500 nM each. Additive increases in the binding signal indicated binding to distinct epitopes. The first experimental setup, where CG4, CG5, CG3, and CB6 were added sequentially, revealed a modest additive effect upon CG5 addition and a more significant increase with CG3, followed by another small increase upon CB6 addition (Figure 50 C). These observations suggested that CG4 and CG5 bound similar epitopes, with CG5 showing higher affinity, evidenced by its ability to saturate shared epitopes more effectively. Reversing the order of sybody addition reinforced these conclusions, as CG4 failed to generate additional binding signals following CG5 saturation.

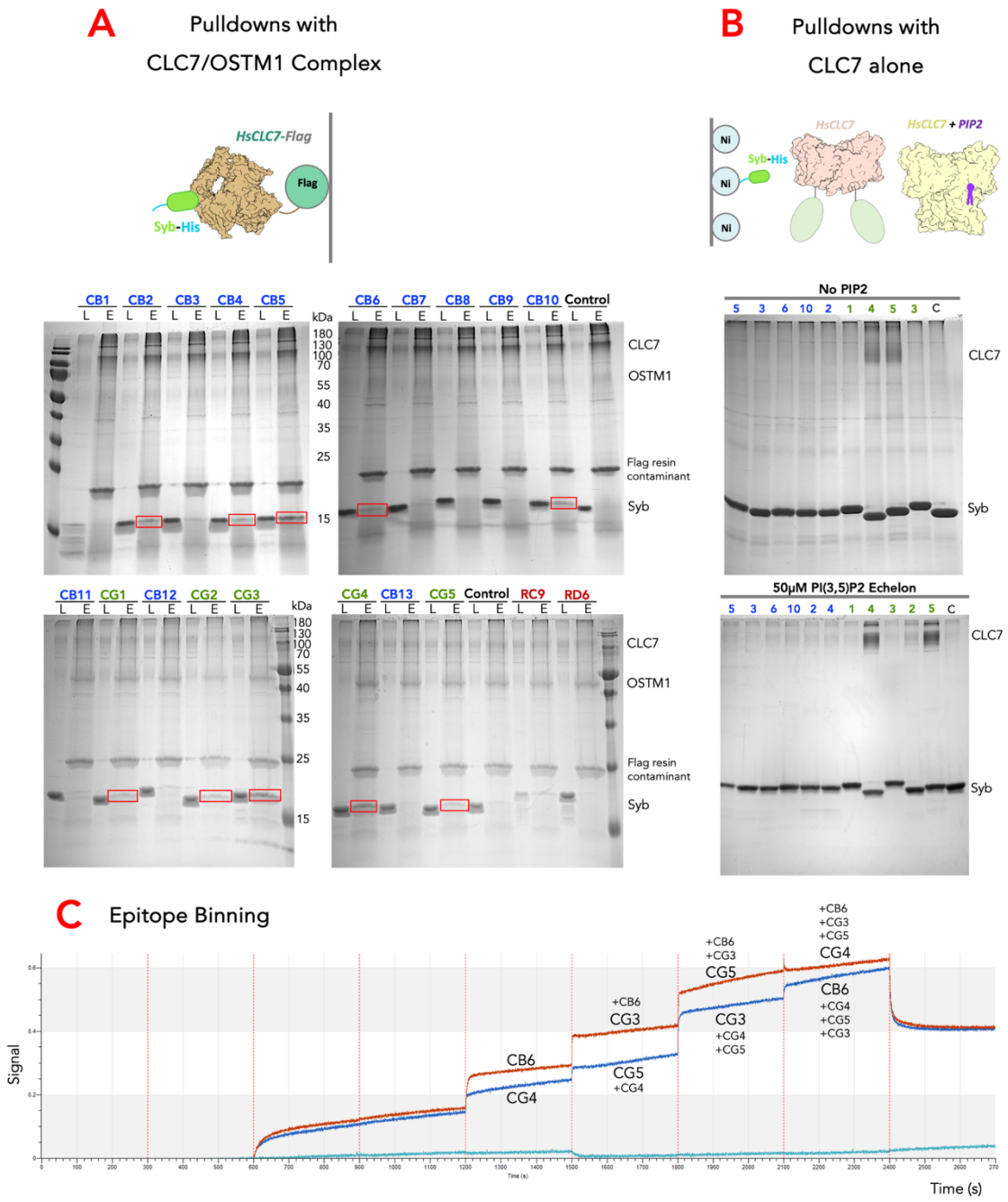


Figure 50. Analyses of Sybody Binding to CLC7

A. SDS-PAGE analysis of sybody pulldowns using the CLC7-FLAG/OSTM1 complex. Sybodies (Syb-His) were incubated with the complex and captured via FLAG resin. Lane labels indicate load (L) and elution (E) fractions. Bands corresponding to CLC7, OSTM1, and sybodies are indicated. Several sybodies (highlighted with red boxes) successfully co-eluted with the complex. B. SDS-PAGE analysis of pulldowns using CLC7 in the absence (top) or presence (bottom) of PI(3,5)P2 lipid. Sybodies were incubated with CLC7, and eluted fractions were analysed. Bands corresponding to CLC7 and sybodies are indicated. C. Biolayer interferometry (BLI) analysis of binding competition among sybodies CB6, CG3, CG4, and CG5. Red and blue lines indicate two separate experiments where sybodies were added in opposite orders. Sequential binding assays revealed overlapping and distinct binding epitopes.

However, results concerning CG3 were less definitive due to inconsistent signal jumps observed in both experimental setups. These inconsistencies raised several possibilities: a buffer composition discrepancy due to handling errors, non-specific interactions of CG3 with either CLC7/OSTM1 complex, detergent micelle or the biosensor surface, or intrinsic variability in CG3's binding characteristics. Meanwhile, CB6 consistently bound a distinct epitope to CG4 and 5, likely on OSTM1 or elsewhere on the complex, although its exact positional relationship to CG3 and potential for non-specific interactions remained unresolved.

Subsequent kinetic analyses of individual sybodies revealed substantial improvements, with CG5 exhibiting a dissociation constant of approximately 230 nM and CB6 around 811 nM, marking a significant advancement over the previous screening efforts (Figure 49 F).

4.3 Sequence Analysis of CLC7 Sybodies

Between the two attempts at selecting sybodies against CLC7, no redundant hits were observed (Table 3). This indicates that employing the full CLC7/OSTM1 complex and supplementing with PIP2 significantly altered the epitopes presented on CLC7, underscoring the influence of experimental conditions the outcome of sybody selection.

In the initial selection round, sybodies that successfully pulled down CLC7 originated exclusively from the convex library. However, this library presented challenges due to a high prevalence of aromatic residues within the CDR3 loops, which often resulted in issues such as non-specific binding or reduced biophysical quality. Notably, sybodies RC9 and RD6 exemplified these challenges. RD6 was capable of pulling down CLC7 but failed to exhibit a specific binding signal in subsequent Octet analyses, indicating potential non-specific interactions. Similarly, RC9 displayed only a minimal binding signal during Octet analysis

(0.05 when 500nM binder was added, and 0.2 when 2uM of binder added), with significantly lower affinity compared to sybodies identified in the subsequent selection round (~0.2 signal when only 500nM binders were added). Consequently, due to these limitations, the convex library was omitted from the second selection pipeline.

The concave library, characterized by shorter CDR3 loops, yielded only two unique sybodies in the initial screening, neither of which demonstrated binding to CLC7. In contrast, the second screening significantly improved outcomes, with five unique sybodies enriched from this library. All five successfully pulled down the CLC7/OSTM1 complex, and notably, two bound specifically to the CBS domain.

Among the five sybodies enriched from the second screening using the concave library, four (CG2, CG3, CG4, and CG5) demonstrated distinctive sequence characteristics within their CDR3 loops. Each of these sybodies exhibited a variable alanine residue at the start of their CDR3 loops, and a positively charged residue immediately after the conserved valine residue. Additionally, multiple consecutive aromatic residues were consistently found near the end of these CDR3 loops. A further shared sequence feature included an alanine residue located as the final variable residue in their CDR1 loops. Due to these similarities, it would be tempting to inferred from sequence analyses that these four sybodies likely recognize overlapping or similar epitopes on CLC7. Nevertheless, functional assays showed differential binding behaviours. Specifically, CG4 and CG5 could successfully pull down CLC7 independently of the presence of OSTM1, while CG2 and CG3 could not. Furthermore, epitope binning suggested that CG3 interacts with a distinct epitope compared to CG4 and CG5. These observations support a model in which, although the conserved CDR3 loop may interact directly with epitopes on CLC7, stable and high-affinity binding of certain sybodies could require additional interactions provided by their CDR1 and CDR2 loops with specific regions on OSTM1.

Historically, the loop library has proven most successful within the lab for selecting sybodies against membrane protein targets. However, in the first screening for CLC7, all 6 unique binders from this library failed to pull down CLC7. The second screening drastically improved outcome again, with 13 unique hits from ELISA, and 8 that pulled down CLC7/OSTM1 complex. Comparative analysis of the loop library sybodies from the first and second screenings revealed notable differences. Particularly, sybodies from the second screening prominently featured CDR3 loops starting with residues tryptophan (W) or arginine (R), despite these residues typically being rare in the library (each with 5% propensity at that specific position). This enrichment implies selective pressure favouring these residues, potentially reflecting their roles in binding interactions with CLC7. Contrarily, acidic residues such as aspartic acid dominated the CDR3 loops in the first screen, whereas the second screen favoured sequences rich in alternating hydrophobic and polar uncharged residues.

The high similarity among loop library sybodies suggests they might recognize a common epitope, presumably on OSTM1, also backed up by their selective ability to pull down the complex but not CLC7 independently.

Examination of CDR1 and CDR2 loops across the two screens highlighted consistent yet minor variations. CDR1 loops remained histidine-rich across both screens, while CDR2 loops contained numerous charged residues, a feature especially pronounced in the second screening.

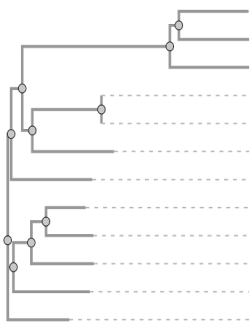
Most binders from this latter group exhibited multiple charged residues within their CDR2 loops, often incorporating both positively and negatively charged amino acids simultaneously. However, the presence of these charged residues did not appear to strongly correlate with binding success. For instance, sybodies CB2, CB3, and CB8 all contained a mixture of positively and negatively charged residues within their CDR2 loops. Despite these similarities, only CB2 and CB3 effectively pulled down the CLC7/OSTM1 complex, whereas CB8 failed

to do so. This suggests that rather than directly mediating target binding, the primary functional role of the CDR1 and CDR2 loops in these sybodies may be structural.

Table 3. Sequence Analysis of CLC7 Sybodies

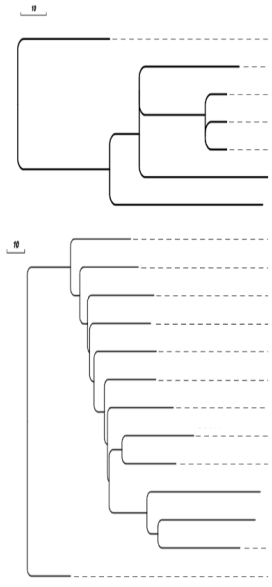
Top: Phylogenetic tree and CDR1–3 sequence alignments of sybodies isolated from the initial selection rounds. Sybodies are grouped by library origin (Convex, Concave, or Loop). Pulldown outcomes are indicated. Bottom: Phylogenetic tree and CDR1–3 sequence alignments of sybodies isolated after optimized selection rounds. Sybodies are grouped by library origin (Concave, or Loop). Pulldown outcomes are indicated. And Sybodies with predicted epitopes annotated.

Initial Selection



Name	CDR1	CDR2	CDR3	Epitope?	Pulldown?	
RC9	SINNIYYLG	NTNWGT	WQGWDWPLYRRWYY	untested	Yes	Convex
RD6	NITGIKYLG	STDSGE	NYGYWWPLSRRDYE	untested	Yes	
RF1	NIAHIKYLG	ATTTGR	KRGNKEPLNSSRYT	untested	no	
RB6	ATFRMF	QSNGLHT	YVNVGQSYT	untested	no	Concave
RF7				untested	no	
RF8	DQTGMW	ESFGDRT	HVNVGTHYH	untested	no	
BB2	WNNTMY	TSTGNETT	YGNHYGD	untested	no	
BG7	NAHYMH	YSWGYGTA	AGINATQ	untested	no	
RG7	NAYYMH	YSSGYGTA	VGADHWA	untested	no	Loop
BB8	ASHYMH	YSNGYGTK	VGDLYNT	untested	no	
RD1	SSSTMT	ASDGRHTK	VGHPQNS	untested	no	
BD7	YIEMA	RSTGYHTH	SGMDGTE	untested	no	

Optimised Selection



Name	CDR1	CDR2	CDR3	Epitope?	Pulldown?	
CG1	NANSMF	SSYGLWTF	HVNVGARYE	untested	Yes	
CG4	RTSHMA	RSSGKTTW	AVKVGYYYT	CBS	Yes	
CG3	KVDQMA	KSHGKWTW	AVKVGRYYK	Complex?	Yes	Concave
CG2	SSAGMA	KSAGQ-TW	AVKVGSYYT	untested	Yes	
CG5	HEWHMA	QSWGEYTH	AVHVGNWYK	CBS	Yes	
CB5	KTLNMY	TSFGYETT	WGAYMYW	untested	Yes	
CB3	AHLNMH	VSAGHRTE	WGSYMNW	untested	Yes	
CB8	NHSKMR	DSKGASTK	WGAYMHW	untested	no	
CB9	QYYKMH	HSEGNETY	WGNYMHW	untested	no	
CB6	MYNSMH	VSEGTGTR	WGQYMTW	OSTM1	Yes	
CB13	EYTNMF	MSEGDGTE	WGQYMHW	untested	no	Loop
CB7	SNTHMY	YSFGDGTT	WGEYMTW	untested	no	
CB10	YSSRME	GSNGTWTK	WGAYMTW	untested	Yes	
CB2	SSSTMT	ESHGDETI	WGAYMTW	untested	Yes	
CB12	AYNWMR	NSWGDYTE	RGHQTNT	untested	Yes	
CB11	NNAKML	QSHGLKTR	RGHPQFI	untested	Yes	
CB1	HLYEMY	ASTGHTTY	YGHPQNM	untested	no	
CB4	NTASMY	YSIGDVME	WGAYMSW	untested	Yes	

4.3 Chapter Conclusions

In summary, this chapter describes the development of synthetic nanobodies against CLC7. The initial pipeline, performed without PIP2, yielded only weak and non-specific sybodies; by contrast, an optimized strategy employing the full CLC7/OSTM1 complex with PIP2 supplementation produced a diverse panel of higher-affinity binders, including CG5 and CB6, which reached sub-micromolar affinities. Sequence and binding analyses revealed clear library-dependent differences, with concave and loop libraries yielding functional sybodies only under optimized conditions. Collectively, the results presented here could provide molecular tools, such as nanobody CG5, that will be valuable for further structural and functional studies of lysosomal pH regulation. Importantly, these binders now can be used in cryo-EM or tomography as fiducial markers, or used in live-cell imaging, where they may facilitate visualization of endogenous complexes and enable mechanistic insights into CLC7 regulation and function in a native cellular context.

5. Structural Studies of RUSF1

5.1 Expression & Characterisation of RUSF1

To structurally characterize the protein RUSF1 using cryo-electron microscopy (CryoEM), study its function *in vitro*, or generate high-affinity binders, the protein must first be produced in large quantities and purified effectively. To achieve this, RUSF1 was expressed with a C-terminal FLAG tag in mammalian cells. Subsequently, purification was tested using two different detergents: DDM and LMNG, each with or without the addition of cholesteryl hemisuccinate (CHS).

From 2-liter cultures of HEK293F cells, cell pellets were lysed, washed, and membranes were separated. These membranes were solubilized in each detergent under the previously described conditions (Chapter 2.3.3). After solubilization, FLAG resin was added to the ultracentrifuge supernatant. The FLAG resin eluents showed clear bands corresponding to RUSF1 on Coomassie Blue-stained SDS-PAGE gels (13.5% or 15% acrylamide) in all conditions tested (Figure 51 A). Size-exclusion chromatography (SEC) revealed mono-dispersed peaks eluting between 10.2 and 12.7 mL, corresponding to RUSF1. Differences in retention volumes across conditions likely reflect variations in detergent micelle sizes and the presence or absence of cholesterol.

Purified proteins were subjected to thermal melt analysis using nanoscale differential scanning fluorimetry to assess stability. Proteins purified in DDM-CHS, LMNG, and LMNG-CHS unfolded within the temperature range of 40–50 °C (Figure 51 B). Protein purified in LMNG-CHS exhibited the highest stability, with a melting temperature (Tm) of 47.2 °C, followed by protein purified in DDM-CHS (Tm 46 °C), and then LMNG alone (Tm 43.6 °C).

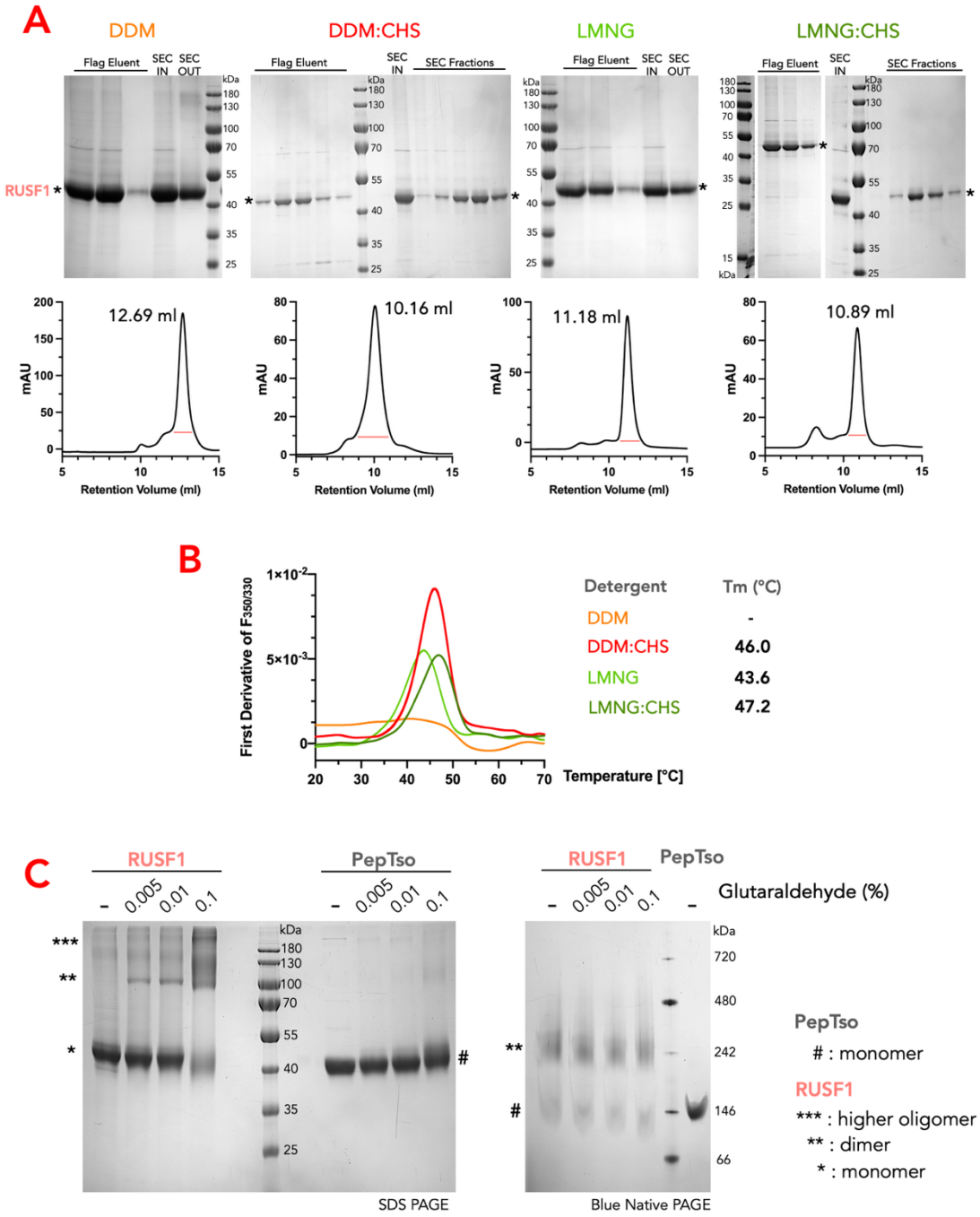


Figure 51. Purification of RUSF1 in Detergent and Investigation of Oligomeric States

A. Purification of RUSF1 using different detergents (DDM, DDM:CHS, LMNG, LMNG:CHS). SDS-PAGE gels show FLAG elution and SEC fractions. SEC chromatograms (bottom) reveal distinct elution volumes. B. Thermal denaturation profiles of RUSF1 in different detergents as measured by differential scanning fluorimetry (DSF). Duplicate experiments were performed for each sample, and results of one repeat are shown. First derivative plots of fluorescence are shown. Melting temperatures (Tm) were determined as follows: DDM:CHS (46.0 °C), LMNG (43.6 °C), LMNG:CHS (47.2 °C). No defined Tm was observed for DDM. C. SDS-PAGE (left) and Blue Native PAGE (right) analysis of RUSF1 and PepTso after glutaraldehyde cross-linking at indicated concentrations. Monomers, dimers, and higher-order oligomers are labelled. RUSF1 forms dimers and higher oligomers upon cross-linking, while PepTso serves as a monomeric control.

Interestingly, protein purified in DDM alone did not exhibit detectable melting within the experiment's duration, as indicated by the lack of intrinsic fluorescence change typically associated with unfolding. This observation suggested that the protein purified solely in DDM may have unfolded during the purification process itself, or there were abnormally strong interactions between the tryptophan/tyrosine residues in RUSF1 and the DDM micelle, leading to them never being exposed to the solvent during the heating process, thus the lack of signal. Therefore, despite DDM providing the highest apparent yield in protein purification trials (approximately 175 mAU absorption peak on SEC), it was deemed unsuitable for structural or functional studies due to protein instability.

Additionally, SEC retention volumes suggested RUSF1 might purify as a homodimer or possibly higher-order oligomers. For comparison, a known dimeric CLC7 complex purified in LMNG-CHS eluted at approximately 10.1 mL (molecular weight ~180 kDa), closely matching the elution profile of RUSF1 purified in LMNG-CHS at 10.89 mL (estimated dimeric size ~110 kDa).

To confirm biochemically whether RUSF1 forms oligomers, crosslinking experiments were performed using glutaraldehyde at low concentrations on DDM-CHS-purified RUSF1 (Figure 51 C). The goal was to crosslink potential oligomers within the detergent micelle without linking multiple micelles together. Cross-linked proteins were then analysed by SDS-PAGE and blue native-PAGE gels. If there were more than one copies of RUSF1 within a micelle, an upward shift should be observed in the SDS-PAGE, but not the native-PAGE, as SDS-PAGE unfolds oligomers in the absence of crosslinking reagents, whereas native-PAGE does not.

As expected, in the presence of 0.005%, 0.01%, or 0.1% glutaraldehyde, SDS-PAGE analysis revealed an additional band at approximately 100 kDa, consistent with the molecular weight of dimeric RUSF1. At 0.1% glutaraldehyde, additional higher molecular weight bands also appeared, possibly indicative of crosslinking between micelles, as similar bands appeared in

the negative control PepTso, a monomeric transporter of about 50 kDa. No band shift was observed on native PAGE gels, confirming that the band shift observed in SDS-PAGE corresponded to crosslinking events that primarily occurred within individual micelles rather than across multiple micelles.

5.2 Structural Elucidation of RUSF1 Homodimer in Detergent

The structural elucidation of the RUSF1 homodimer in detergent commenced with the observation that detergent-purified RUSF1 appeared to exist predominantly as a dimer, exhibiting a total molecular size of approximately 110 kDa. AlphaFold predictions indicated that a significant portion of the RUSF1 complex comprised a soluble domain tightly associated with the transmembrane region. This encouraged the decision to proceed with structural determination efforts without the use of binder molecules such as nanobodies, commonly employed to increase particle size and facilitate structural analysis.

Initial attempts at sample preparation involved grids prepared with 1.5 mg/ml LMNG-CHS-purified RUSF1, chosen due to its superior thermal stability in initial purification trials. However, preliminary screening using the Arctica electron microscope revealed heterogeneous particle shapes and many species lacking the characteristic oval shapes typically observed for membrane proteins enveloped by detergent micelles (Figure 52 A). Thus, further optimisation of sample preparation was necessary.

In the following attempt, the detergent was switched from LMNG-CHS to DDM-CHS, and a comprehensive thermal stability screen was performed (Figure 52 B). RUSF1 proteins were tested across various buffer systems, including citrate phosphate at pH 5.5, 6, and 6.5; MES at pH 5.5 and 6.5; HEPES at pH 6.8 and 7.2; and Tris NaCl at pH 7.5. Differential scanning fluorimetry analysis revealed the greatest stability of RUSF1 at pH 6 in citrate phosphate buffer, exhibiting a melting temperature (Tm) of 49°C, a notable increase of four degrees compared

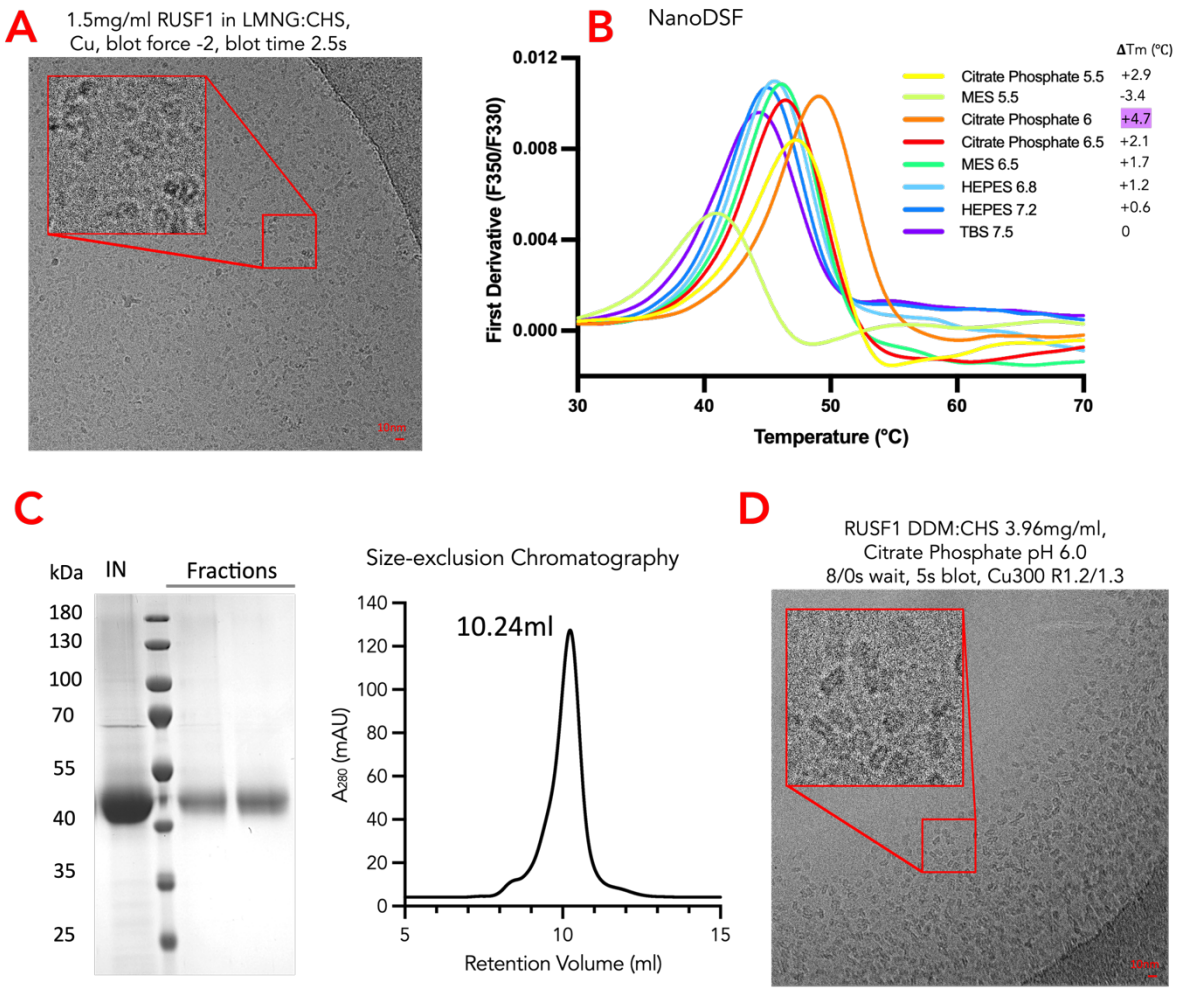


Figure 52. RUSF1 Sample Optimisation for CryoEM

A. Representative cryo-EM micrograph of RUSF1 purified in LMNG:CHS. Grids were prepared using copper support. B. NanoDSF analysis of RUSF1 in various buffer conditions, thermo shifts calculated against sample in TBS pH 7.5 buffer. Duplicate experiments were performed for each sample, and results of one repeat are shown. Citrate phosphate pH 6.0 conferred the highest stability (ΔTm = +4.7 °C). C. Purification of RUSF1 in citrate phosphate buffer, SDS-PAGE analysis (left) of RUSF1 after affinity purification and size-exclusion chromatography (SEC). D. Representative cryo-EM micrograph of RUSF1 in DDM:CHS micelles under optimized buffer conditions (Citrate phosphate pH 6.0).

to the Tm recorded in the routinely used Tris buffer at pH 7.5. Citrate phosphate buffers at pH 5.5 and pH 6.5 resulted in slightly lower Tm values of 47.2°C and 46.4°C, respectively. Conversely, MES buffers at similar pH values provided comparatively less stabilisation, with Tm values around 46°C at pH 6.5 and even lower at 41°C at pH 5.5. Similarly, HEPES buffer systems did not significantly enhance stability relative to Tris, though slight stabilization (Tm = 46°C) was observed at pH 6.8.

Given these insights, large-scale purification was repeated with solubilization in DDM-CHS and subsequent FLAG affinity purification in PBS under standard protocols (Figure 52 C). The eluted protein was then subjected to size exclusion chromatography (SEC) using citrate phosphate buffer at pH 6, resulting in a monodisperse protein peak eluting at 10.24 ml, slightly later than previously observed with Tris-based buffers. SDS-PAGE confirmed the purity of the SEC fractions, which were pooled, concentrated to 3.96 mg/ml, and subsequently applied onto copper grids (Quantifoil R1.2/1.3, 300 mesh, Cu). These grids were blotted and rapidly vitrified by plunge-freezing in liquid ethane.

Subsequent screening of the optimised grids revealed significantly improved particle quality, displaying greater homogeneity and the characteristic oval shape expected of detergent micelle-embedded membrane proteins (Figure 52 D). Although particle distribution was predominantly observed near grid hole edges with slightly thicker ice, particle density was found to be optimal for data collection. Consequently, a comprehensive data collection comprising 14,967 movies was conducted from one optimized grid.

Initially, data processing followed the standard pipeline in the group: motion correction, CTF refinement, and particle picking done in SIMPLE, followed by multiple rounds of 2D classification in cryoSPARC. This resulted in multiple classes showing side view of the DDM-CHS micelle, some signal for transmembrane helices, and very clear soluble domains consistent with AlphaFold prediction of RUSF1 homodimer (Figure 53 A). Tilted side views as well as top/bottom views were also present, although slightly under-represented. However, after *ab initio* model generation and non-uniform refinements, the resolution was limited to 6.75 Å (Figure 53 B). The map revealed some densities for transmembrane helices, as well as two globular domains outside the micelle, which the AlphaFold model roughly fitted, indicating the presence of RUSF1 dimer (Figure 53 C). However, resolution needed to be improved for secondary structural details to be revealed.

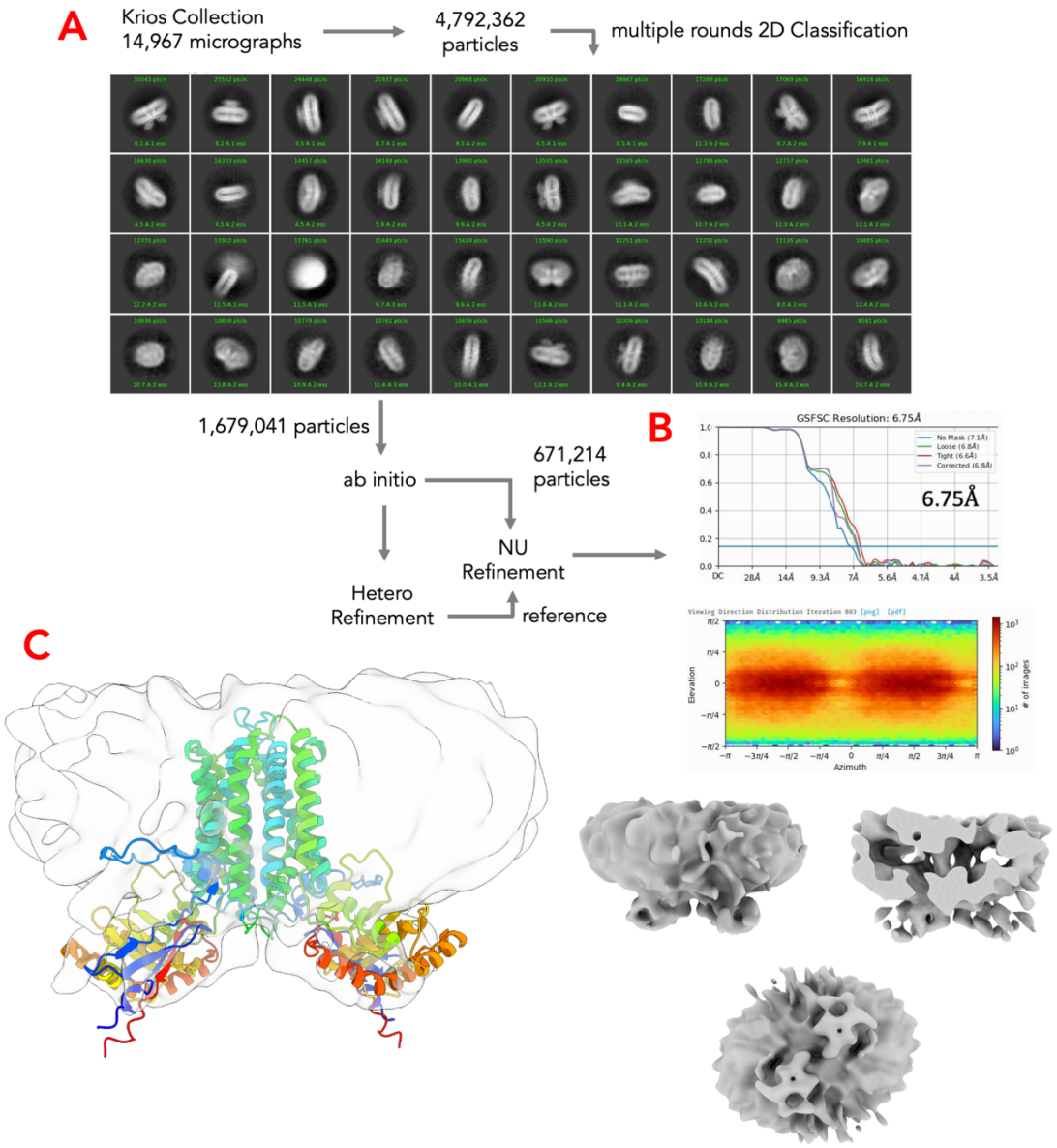


Figure 53. Initial Attempts of CryoEM Processing for RUSF1

A. Cryo-EM data processing workflow, resulting in a low-resolution map. B. Gold-standard Fourier shell correlation determined a resolution of 6.75 Å. C. Left: cryo-EM density map with fitted AlphaFold prediction of RUSF1 dimer, showing overall goodness of fit in the soluble domains. Right: multiple views of the reconstructed cryo-EM density. Sliced view showed density for transmembrane helices. Data processing done by Takafumi Kato.

An optimised pipeline which by-passed *ab initio* model generation on the whole dataset proved to be successful in improving resolution. In brief, a subset of particles from the final 2D classification was used for *ab initio* model generation, yielding multiple maps, which were used as references (1 map as the reference for target protein, 2 maps as references for junk

particles). These reference maps were used in heterogeneous refinements, for the whole dataset. After two rounds of heterogeneous refinement, the protein-containing map was used as a reference for non-uniform refinement, yielding a density map of 3.9 Å. Further polishing and 3D classification in RELION improved the resolution even further, and a final density map of global resolution 3.13 Å was obtained, with the core of the protein reaching 2.6 Å (Figure 54, Supplementary Table 3).

The structural analysis revealed that two copies of RUSF1 are present within the micelle, forming a symmetrical homodimer. The transmembrane domain was predominantly buried within the micelle, while the cytosolic domains protruded outward approximately 22 Å, adopting a characteristic "V" shaped conformation (Figure 55 A).

On the luminal side, the transmembrane helices were partially buried within the micelle, consistent with their predominantly hydrophobic character (Figure 55 A). The membrane-facing portions of the helices were largely hydrophobic, whereas centrally, a significant cavity was apparent, created by transmembrane helices 1, 2, 4, and 6 from each RUSF1 protomer. This cavity, oriented toward the luminal side, primarily featured polar residues (Figure 55 B); notably, no lipid or detergent density was detected within this space.

On the cytosolic side of the transmembrane domain, patches of charged residues began to emerge, partially interacting with detergent head groups (Figure 55 B). A notable basic patch was identified at the junction where transmembrane helix 6 transitions into the soluble domains, accompanied by an acidic groove located at the dimer interface.

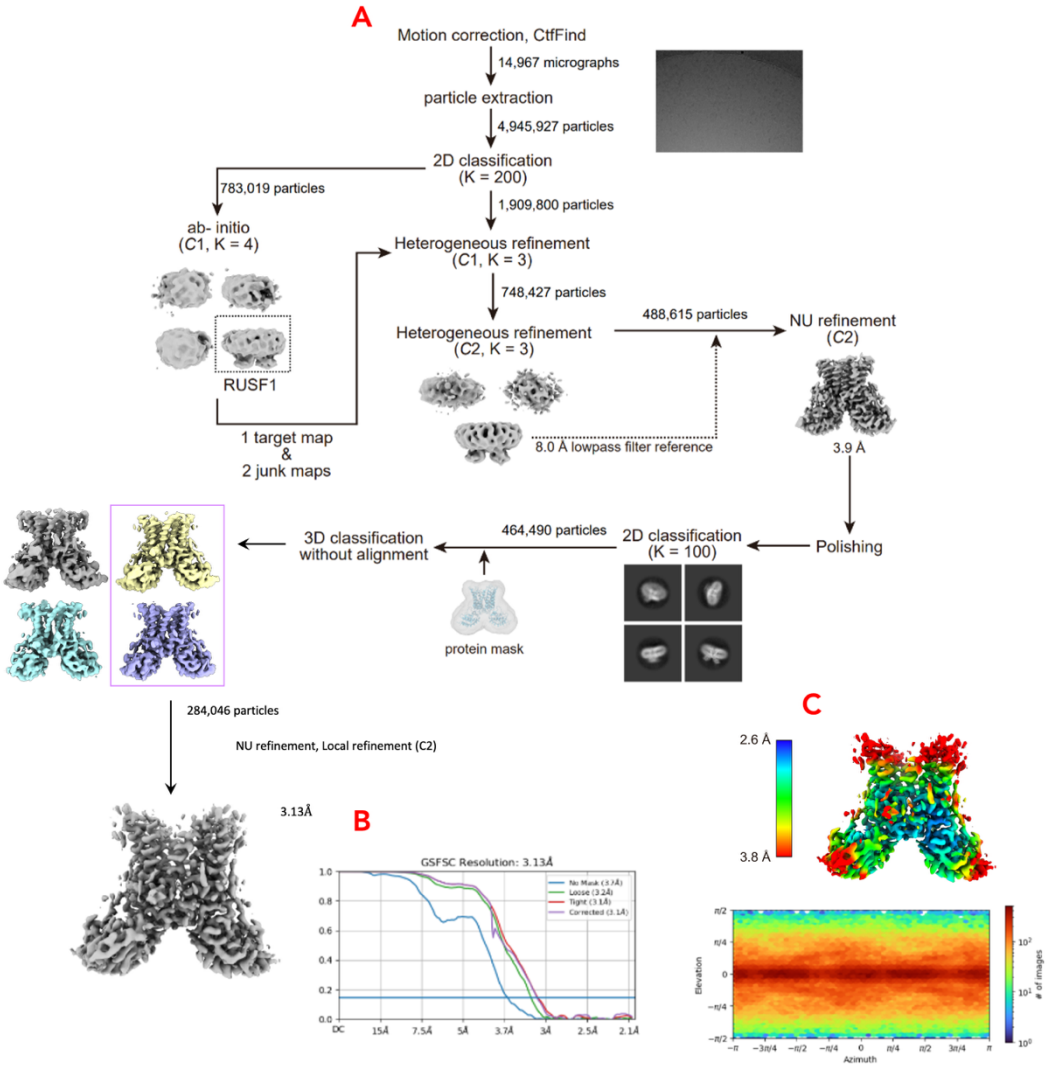


Figure 54. Optimised CryoEM Processing Pipeline of RUSF1

A. Optimised processing steps for RUSF1 dataset, which bypassed *ab initio* model generation steps using all particles. B. Gold-standard Fourier shell correlation curve of the final reconstruction, indicating a global resolution of 3.13 Å at the 0.143 FSC cutoff. C. Local resolution estimation of the final cryo-EM map. The map is colored according to local resolution (2.6–3.8 Å), with the angular distribution plot showing uniform sampling of particle orientations. Data processing done by Takafumi Kato.

The cytosolic side of the RUSF1 complex revealed two symmetrical structural modules composed of domain-swapped N- and C-terminal regions from each RUSF1 protomer (Figure 55 B). The surface of these domains exhibited predominantly polar characteristics, becoming increasingly acidic towards the periphery and more basic towards the central region.

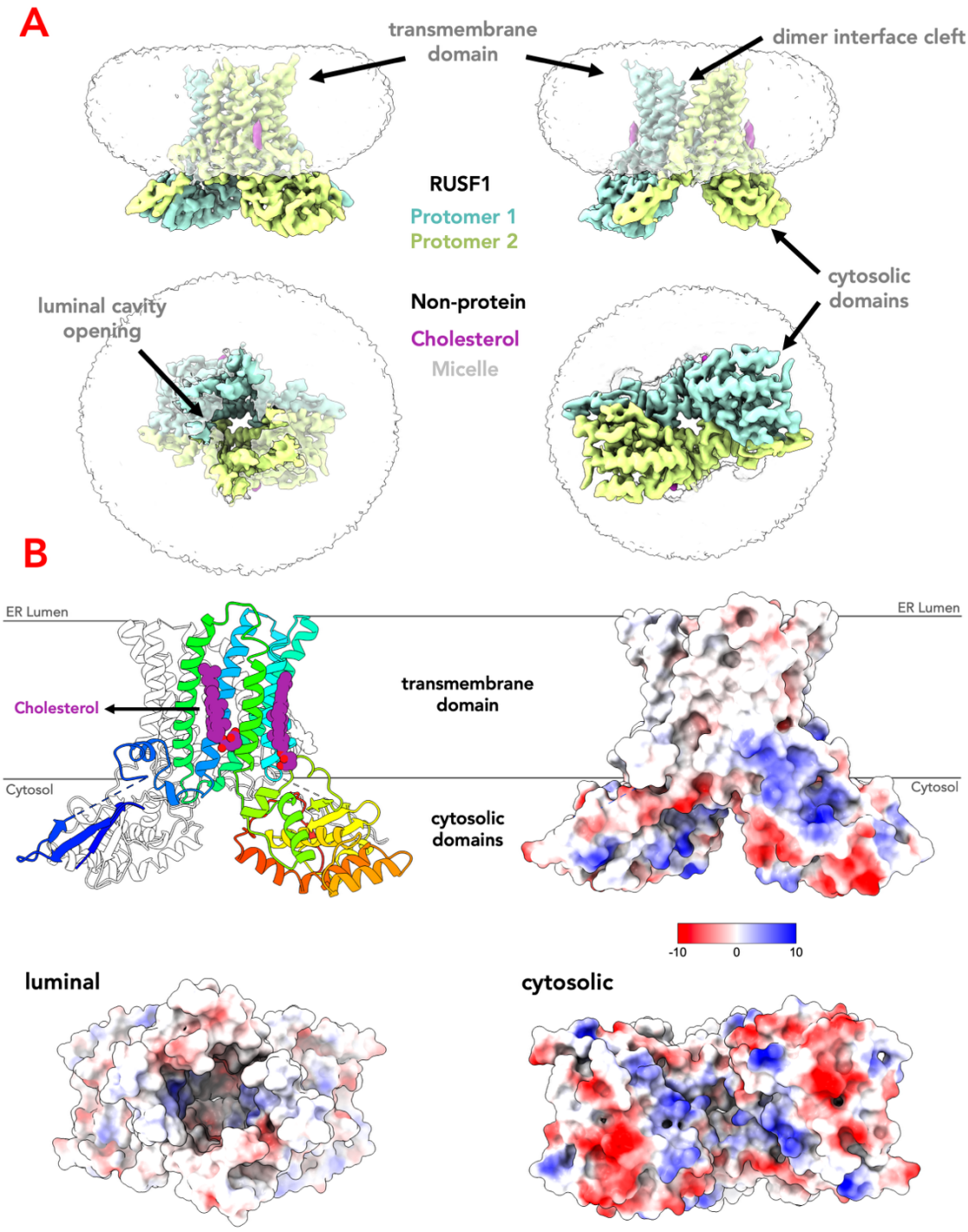


Figure 55. CryoEM Structure of RUSF1 Homodimer in Detergent Micelle

A. Cryo-EM density map of RUSF1 displayed in different orientations. The protein forms a dimer composed of two protomers (coloured cyan and yellow), with distinct transmembrane and cytosolic domains. The dimer interface cleft and luminal cavity opening are indicated. Non-protein densities corresponding to cholesterol molecules (magenta) and micelles are also observed. B. Structural features of RUSF1. Left: cartoon representation of the RUSF1 homodimer, showing the transmembrane helices and the cytosolic domains positioned below. One protomer is coloured white, the opposite protomer is coloured by rainbow (N- to C- terminus). Two cholesterol molecule (magenta) is bound within the transmembrane region. Right: RUSF1 surface representation, coloured by electrostatic potential, with positively charged regions in blue and negatively charged regions in red. Bottom: surface representations of the luminal (left) and cytosolic (right) views, coloured by electrostatic potential.

In line with previous AlphaFold predictions, the homo-dimeric structure of RUSF1 closely resembled the architecture of the Major Facilitator Superfamily (MFS) transporters (Figure 56 A, B). Analysis using FoldSeek (Kim *et al.,* 2025) identified 197 structurally similar proteins from the Protein Data Bank (PDB), with 195 belonging to the MFS (Supplementary Table 4). The top two matches were bacterial homologs, specifically a lactate transporter from the SLC16 family and an *E. coli* D-galactonate: proton symporter from the SLC17 family, each were predicted to have a 99% probability to be structurally homologous to RUSF1, despite sequence similarities of only 12.9% and 13.6%, respectively (Supplementary Table 4). Further comparative analysis in this chapter will focus primarily on the top match, the lactate transporter SfMCT from *Syntrophobacter fumaroxidans* (PDB ID 6ZGR, Bosshart *et al.*, 2021).

Examination of the arrangement of transmembrane helices further underscored the MFS fold similarity in the RUSF1 dimer, with protomers corresponding structurally to the N- and C-terminal bundles of monomeric MFS transporters (Figure 56 A, B). Both the N- and C-termini of RUSF1 and SfMCT were cytosolic, and the orientation of TM helices 1–6 in RUSF1 mirrored the orientation of TM helices 1–6 and 7–12 in SfMCT.

Observing from the cytosolic side, the helical arrangements appeared markedly similar; helices 2 and 5 of both RUSF1 protomers formed the dimer interface, corresponding to helices 2, 5, 7, and 11 in SfMCT. However, significant differences were apparent: while SfMCT helices were densely packed centrally, excluding helices 3, 6, 9, and 12 from the core, RUSF1 helices assembled in a semi-circular fashion, creating a substantially wider cavity. Notably, aside from helix 6 in each protomer, the remaining helices directly faced the cavity at specific regions.

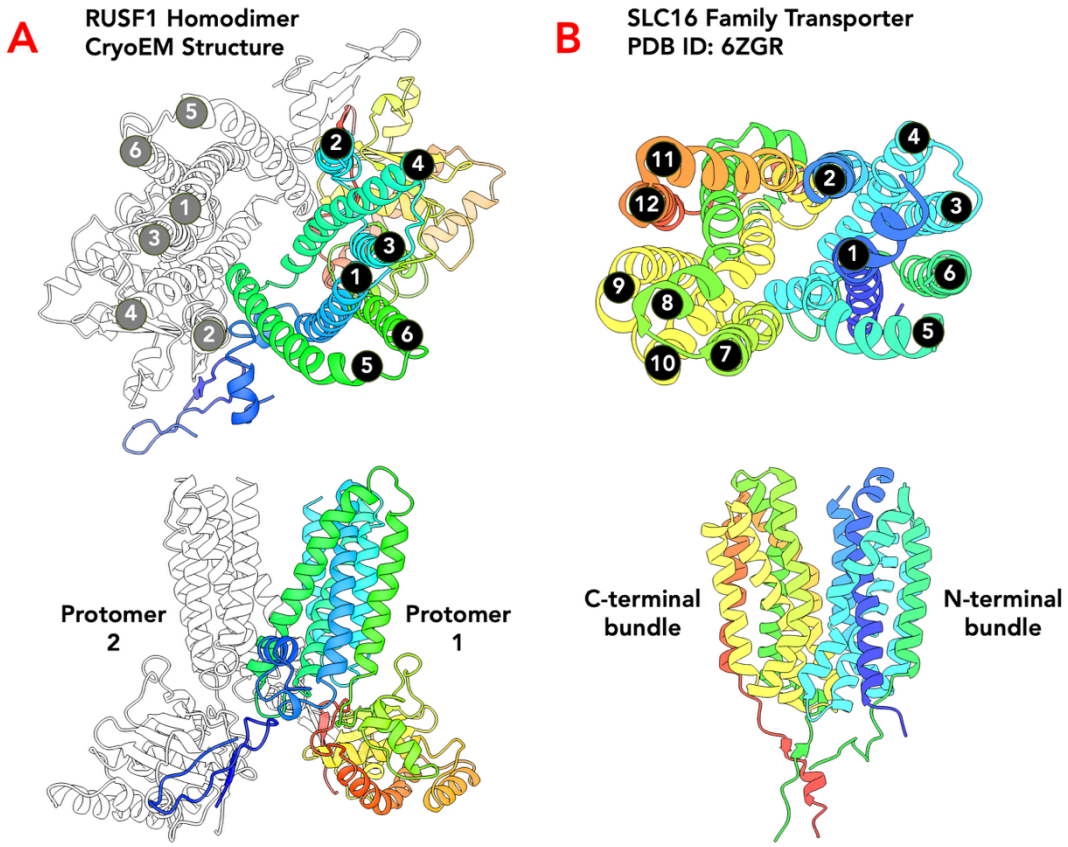


Figure 56. Comparison of RUSF1 EM Structure with MFS Transporters

A. Cryo-EM structure of the RUSF1 homodimer. Top: top view of the dimer showing protomer 1 (coloured rainbow from N- to C-terminus) and protomer 2 (white). Transmembrane helix number annotated on black circles for protomer 1, and grey circles for protomer 2. Bottom: side view of RUSF1 homodimer. B. Crystal structure of a bacterial SLC16 family transporter sfMST (PDB 6ZGR, Bosshart *et al.*, 2021). Top: top view showing the monomeric transporter coloured by rainbow (N- to C- terminus). Transmembrane helix number annotated on black circles for. Note that helices arrangement on the N-terminal bundle is the same as RUSF1 protomer 1. Bottom: side view of sfMST.

Remarkably, many of the MFS-related structures identified to be structurally similar to RUSF1 by FoldSeek, such as SfMCT, were in outward-open conformations, a characteristic clearly mirrored by the expansive luminal-facing cavity observed in RUSF1 (Figure 57 A, C). This cavity, with a volume of approximately 1450 Å^3, had a luminal opening radius of 8 Å, narrowing to a cytosolic constriction of about 0.6 Å radius, reaching a maximum internal radius of 9.5 Å (Figure 57 A, B). The dimensions suggest capacity sufficient to accommodate larger molecules, potentially polypeptides or structured helices, a significant departure from the smaller cavity of SfMCT (volume approximately 301.6 Å^3, maximum radius 5.0 Å, bottleneck

0.3 Å radius; Figure 57 C, D). Such a profound disparity implies a functional divergence between RUSF1 and typical MFS transporters, indicative of evolved specificity for larger substrates.

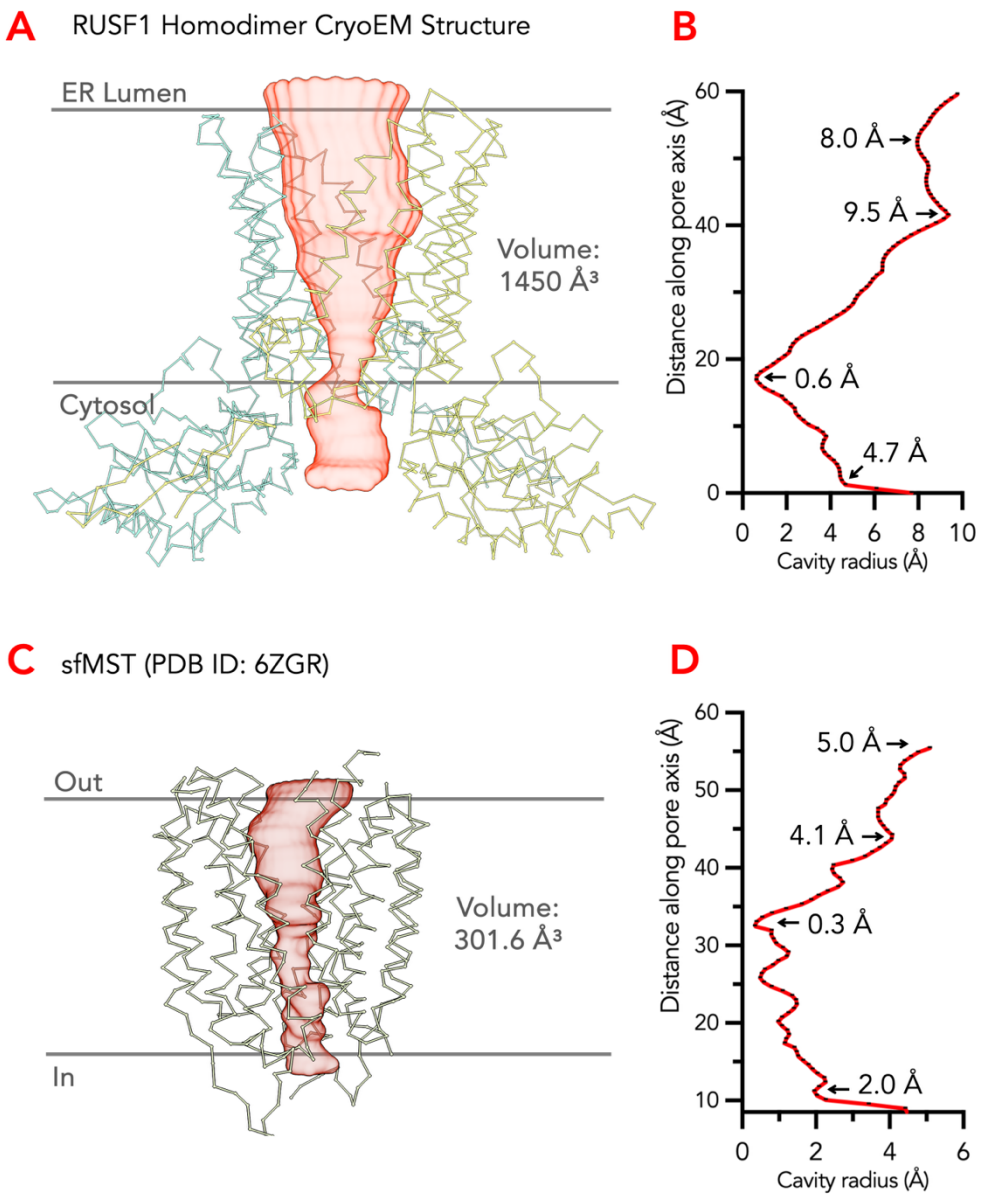


Figure 57. Comparison of RUSF1 Luminal Cavity to Outward-open SLC16 Family Transporter

A. Cryo-EM structure of the RUSF1 homodimer, showing the central cavity (red surface) extending from the ER lumen to the cytosol. The cavity has an estimated volume of 1450 Å³. B. Cavity radius profile of RUSF1 along the pore axis, revealing multiple constriction sites with radii of 0.6 Å and 4.7 Å, and wider regions up to 9.5 Å. C. Structure of the SLC16 family transporter sfMST (PDB ID: 6ZGR), with the internal cavity shown in red. The cavity is smaller in volume (301.6 Å³) and spans the membrane from extracellular to intracellular space. D. Radius profile of the sfMST pore, indicating narrower constrictions compared to RUSF1. Images of cavity surface made by PoreAnalyser (Seiferth & Biggin, 2024).

An additional distinction between RUSF1 and SfMCT cavities involved hydrophilicity. RUSF1's cavity exhibited predominantly hydrophilic residues forming symmetrical, charged patches: acidic residues dominated the cytosolic constriction base, basic patches characterized the central portion, and the luminal opening reverted to acidic residues (Figure 58 A). Intriguingly, the very luminal tips of the helices remained hydrophobic, consistent with partial micelle burial.

At the juncture of transmembrane and soluble domains on the cytosolic side, RUSF1 featured an amphipathic helix partially buried in the micelle, with the exposed cytosolic side negatively charged (Figure 58 A). Adjacent soluble domain regions were also negatively charged, positioned near the cavity bottleneck, opposite an acidic patch within the cavity. The possibility of conformational shifts connecting these regions through solvent exposure remains speculative.

Conversely, SfMCT presented a more uniformly hydrophobic transmembrane domain, with charged residues mainly localized around the ligand-binding site and peripheral entry/exit points, as exemplified by ligand interactions observed in its crystal structure, where an inhibitor was bound, with its negatively charged functional groups fitting precisely in the positively charged ligand binding pocket (Figure 58 B).

In summary, though the structure of RUSF1 exhibited similarities to MFS transporters in terms of arrangement of transmembrane helices, the properties of these helices and the resultant environment created at the core of the protein were very different, both in terms of size, and hydrophobicity, indicating that the role of RUSF1 might be quite distinct from any MFS transporter.

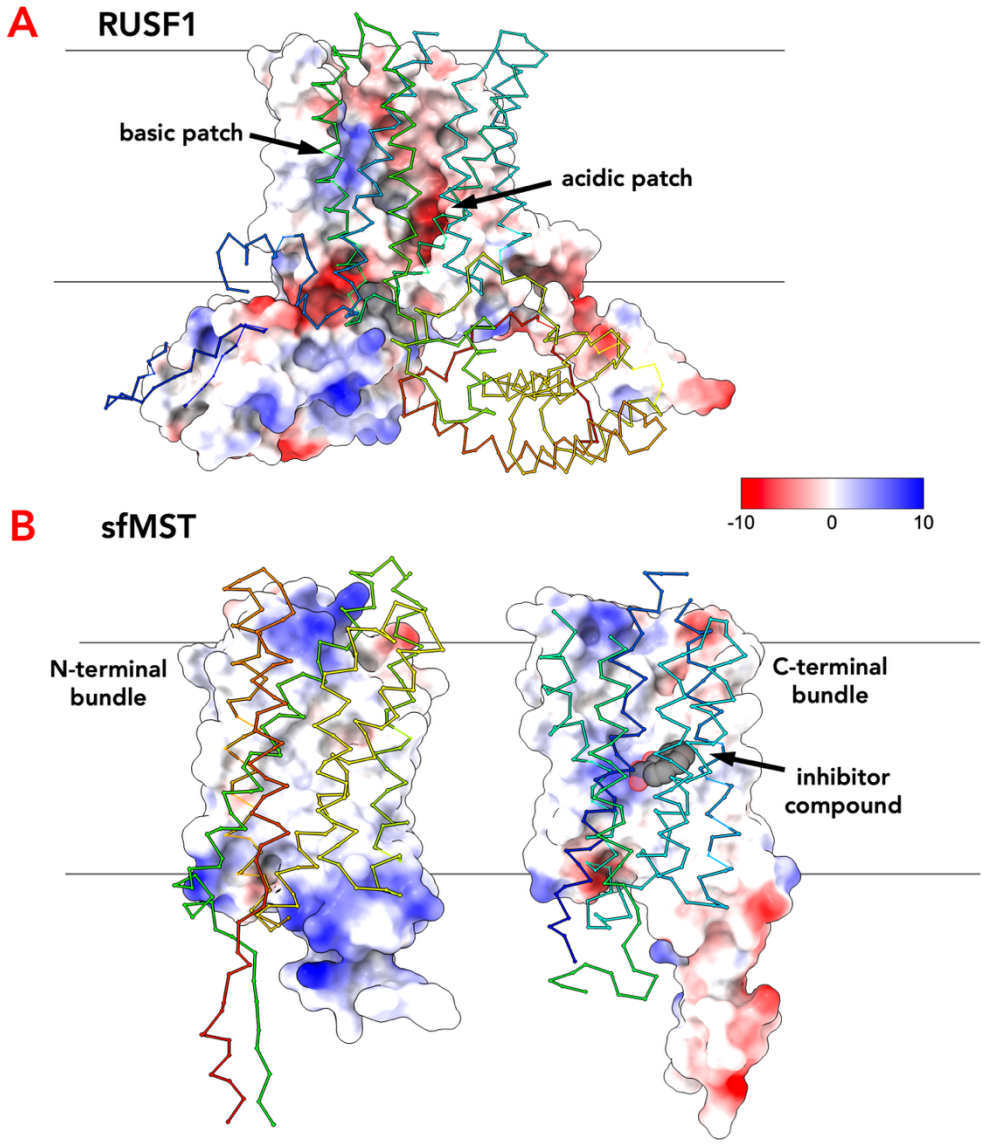


Figure 58. Electrostatics of RUSF1 Cavity Differs from MFS Transporters

A. Electrostatic surface potential of RUSF1 mapped onto one protomer (the opposite protomer in wire representation, coloured by rainbow from N- to C- terminus). Distinct charged regions are observed in the cavity, and cytosolic soluble domains. B. Electrostatic surface potential of sfMST mapped onto the N-terminal bundle (left) or the C-terminal bundle (right), and the rest of the protein in wire representation, coloured by rainbow from N- to C- terminus. The bundle interface is largely hydrophobic, except for a positively charged patch on the C-terminal bundle, where ligand and inhibitor compounds bind (Bosshart *et al.*, 2021).

5.3 Mutations on Conserved RUSF1 Residues Drive Monomerization

Previous studies have demonstrated that AtRUSF1 and AtRUSF2 form heterodimers, and mutations at four conserved positions in AtRUSF1 induce monomerization of the complex (**Chapter 1.3.6,** Leasure *et al*., 2009). To explore whether a similar monomerization mechanism applies to human RUSF1, I introduced glycine substitutions at the analogous conserved residues in human RUSF1, replicating the approach of earlier research (Figure 59 A).

Mapping these mutations onto the existing cryo-EM structure of RUSF1 revealed that residue E179 is positioned away from the dimer interface, oriented instead towards the luminal cavity. Furthermore, residues K162 and N223 were not directly situated at the dimer interface; rather, they are involved in interactions connecting the transmembrane domain and the cytosolic soluble domain. Only the residue K230 was located precisely at the dimer interface (Figure 59 A).

After expressing these mutants in suspension HEK cells, the harvested cells underwent lysis and solubilisation using DDM-CHS. Following solubilisation, protein extraction was carried out using FLAG affinity pulldown. Subsequent analysis through acrylamide gel electrophoresis and western blotting confirmed protein expression. Interestingly, expression was only successful for the E179G mutant, which faces the luminal cavity, and the K230G mutant at the dimer interface. Mutants K162G and N223G, located at critical domain contact sites, failed to express, indicating that the rigidity of this domain interface is critical for either synthesis or structural stability of RUSF1 (Figure 59 B).

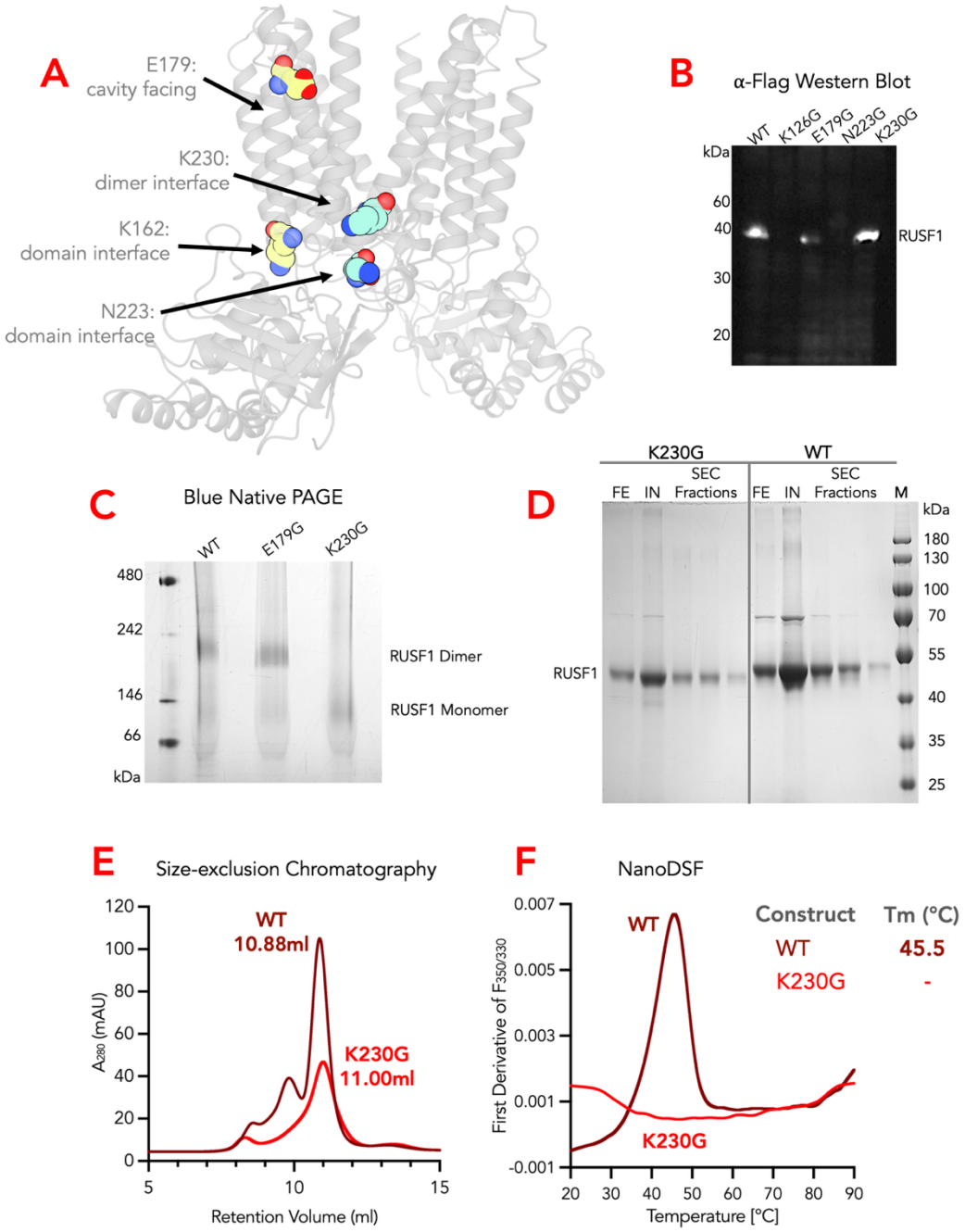


Figure 59. Investigation of Conserved Residues on RUSF1 Dimer Interface

A. Structural model of RUSF1 highlighting conserved residues targeted for mutagenesis in order to generate monomeric protein. E179 (cavity facing), K162 and N223 (domain interface), and K230 (dimer interface) are shown as spheres. B. Western blot analysis of wild-type (WT) and mutant RUSF1 constructs (K162G, E179G, N223G, K230G) expressed with a C-terminal FLAG tag. K126G and N223G did not express. C. Blue native PAGE of WT, E179G, and K230G mutants. WT RUSF1 migrates predominantly as a dimer, whereas the K230G mutant disrupts dimer formation, with a shift toward monomeric species. D. SDS-PAGE of size-exclusion chromatography (SEC) fractions for WT and K230G purified from 2L HEK cells. Both migrated at ~55 kDa. E. SEC chromatograms of WT and K230G RUSF1. WT elutes at 10.88 mL, while K230G elutes slightly later at 11.00 mL. F. Thermal stability analysis using nano differential scanning fluorimetry (NanoDSF). Duplicate experiments were performed for each sample, and results of one repeat are shown. WT RUSF1 displays a melting temperature (Tm) of 45.5 °C, while the K230G mutant shows no defined unfolding transition, indicating structural instability.

To assess the oligomeric state of the expressed mutants, I performed blue native-PAGE on the pulldown eluents (Figure 59 C). The E179G mutant migrated similarly to wild-type (WT) RUSF1, indicating retention of its dimeric state. In contrast, the K230G mutant displayed significantly reduced molecular weight, suggesting monomerization.

To further characterize the K230G mutant, I attempted large-scale purification using the protocol optimized for WT protein (Figure 59 D-F). When analysed by size exclusion chromatography, the K230G mutant eluted only slightly later (~0.1 ml) than WT RUSF1 but yielded roughly one-quarter the amount of protein (Figure 59 E). SDS-PAGE revealed both the WT and K230G mutants as a single predominant band near the 55 kDa marker (Figure 59 D). Despite mono-dispersity of SEC profile, thermal stability assays showed negligible changes in intrinsic fluorescence signals for K230G mutant as compared to WT, indicating either unfolding events during purification, or abnormal interactions between protein and detergent micelle. Consequently, further structural analysis of this mutant was not pursued (Figure 59 F).

Although it was remarkable that a single amino acid mutation (two in total across the dimer) could disrupt a dimer interface spanning approximately 3000 Å^2, questions remained about whether a single K230G mutation was sufficient to prevent dimer formation with a WT protomer. To investigate this, WT RUSF1 tagged with either a C-terminal Myc or FLAG tag was co-expressed with the oppositely tagged K230G mutant (Figure 60 A). Following FLAG pulldown, western blot analysis showed that FLAG-tagged WT RUSF1 interacted exclusively with Myc-tagged WT protein and not with K230G mutants (Figure 60 B). The reverse experiment yielded similar results, confirming that a single K230G mutation at the dimer interface was sufficient to induce monomerization.

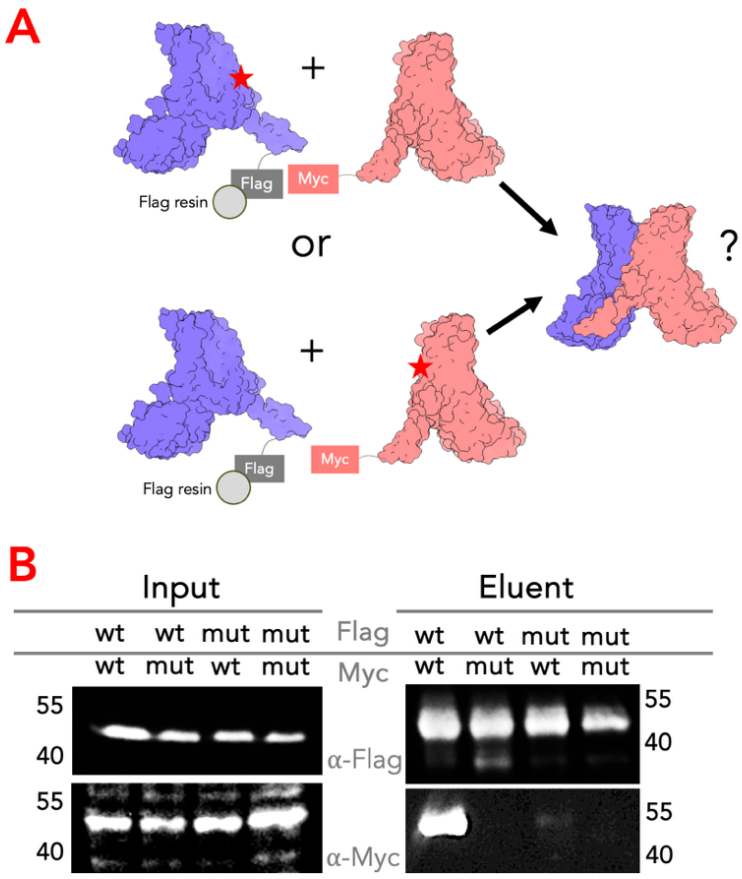


Figure 60. A Single K230G Mutation Abolishes Dimer Interaction

A. Schematic of co-immunoprecipitation assay design. RUSF1 constructs were either FLAG- or Myc-tagged. Wild-type (WT) and mutant (K230G) proteins were co-expressed, and the ability of FLAG-tagged RUSF1 to pull down Myc-tagged partners was assessed. B. Co-immunoprecipitation of WT and mutant RUSF1 constructs. Input lanes confirm comparable expression of FLAG- and Myc-tagged proteins. Elution from FLAG resin shows robust co-purification of Myc-tagged WT RUSF1 with WT FLAG-RUSF1. In contrast, the K230G mutant fails to co-immunoprecipitate Myc-tagged protein, demonstrating that the single K230G mutation disrupts dimer formation.

5.4 Chapter Conclusions

In this chapter, I presented the first structure of RUS family proteins, human RUSF1, solved to 3.1Å by cryo-EM. The structure revealed a dimeric complex with both membrane-embedded domains and soluble domains. The transmembrane domain of the dimeric RUSF1 resembled the architecture of transporters in the Major Facilitator Superfamily, composing of two 6-helix bundles. The structure of RUSF1 differed from MFS transporters, though, by exhibiting a large

cavity facing the ER lumen, almost 5 times the size of the cavity found in an outward facing MFS transporter. In addition, the cavity in RUSF1 is largely charged or polar, in contrast to that of MFS transporters, which are more hydrophobic.

The RUSF1 dimer structure had an extensive dimer interface of around 3000 Å^2, and one residue at the dimer interface, K230, is conserved in *Arabidopsis* RUS1-RUS2 interface. Mutating K230 into glycine in human RUSF1 disrupted the dimer interface, consistent with previous results shown in *Arabidopsis* RUS1-RUS2 (Leasure *et al*., 2009).

The structural and biochemical data presented in this chapter suggested that there is a functional significance in the luminal cavity and dimer interface of RUSF1, providing a foundation for functional characterisation of human RUSF1, and RUS family proteins at large.

6. Towards the Understanding of RUSF1 Function in Cell

6.1 RUSF1 Interacted with a Small Population of GABARA1

The interaction between RUSF1 and GABRA1 was initially suggested by results from a mass-spectrometry screen aimed at identifying chaperones for the GABA receptor subunit alpha 1 (GABRA1, **Chapter 1.3.4**). Consequently, one of the primary objectives was to verify this interaction and further characterize its nature. In particular, I aimed to determine whether RUSF1 interacts with GABRA1 co-translationally or whether it can also recognize fully folded and assembled GABRA1. Additionally, the impact of the K230G mutation in RUSF1 on its ability to bind GABRA1 was investigated, due to the potential physiological significance of this mutant variant.

For initial immunoprecipitation experiments, NACHO was employed as a positive control, as it emerged from the same chaperone screen as another GABRA1 interactor. Structurally, NACHO is a small membrane protein containing four transmembrane helices and lacking a soluble domain. Its interaction with GABRA1 had previously been verified, with recent cryo-EM structural analysis showing a 2:2 heterotetrameric assembly (**Chapter 1.3.5**). In this complex, two peripheral copies of NACHO sandwich two central copies of GABRA1. Given that RUSF1 is significantly larger than NACHO, possesses a sizeable soluble domain, and forms stable homodimers, it could be speculated that its mode of interaction with GABRA1 might differ substantially.

To assess the interaction between RUSF1 and GABRA1, I co-expressed bait proteins, each tagged at the C-terminus with twin strep tags (NACHO-TST or RUSF1-TST), alongside prey protein, N-terminally HA-tagged full-length GABRA1, in HEK293F cells. Following expression, cells were harvested, lysed, and solubilized in LMNG detergent—in line with

previous structural analyses of the NACHO-GABRA1 complex (Hooda *et al*., 2024, preprint). Strep pulldowns were performed, and subsequent western blot analyses clearly demonstrated robust enrichment of GABRA1 in the NACHO pulldown fraction (Figure 61 A, top panel, left). In contrast, RUSF1 pulldown fractions initially appeared to lack enrichment of GABRA1 (Figure 61 A, top panel, right). However, upon enhancing the brightness of the blots, faint bands corresponding to GABRA1 could indeed be observed (Figure 61 A, middle panel, right). Despite comparable expression levels and binding efficiencies to the strep-tactin resin for both NACHO and RUSF1 (Figure 61 A, lower panel), it was evident that NACHO more effectively enriched GABRA1.

Further experiments explored the questions of co-translational versus post-translational interaction and the effect of the K230G mutation on RUSF1's interaction capabilities. Wild-type RUSF1 or its K230G mutant, each bearing a C-terminal FLAG tag, were expressed alongside N-terminally HA-tagged full-length GABRA1, either within the same culture or separately. Cultures expressing bait and prey separately were harvested and combined prior to solubilization to allow post-translational interaction *in vitro*. FLAG pulldowns were conducted, and western blot analyses revealed that under co-expression conditions ("CO" lanes), both wild-type and K230G mutant RUSF1 successfully pulled down comparable amounts of GABRA1 (Figure 61 B). This suggests the monomeric K230G mutant may still be functionally competent in cellular contexts for binding GABRA1.

In contrast, when bait and prey proteins were expressed separately and then combined post-expression ("MIX" lanes), neither wild type nor K230G mutant RUSF1 enriched GABRA1 (Figure 61 B). These findings indicate that RUSF1 likely interacts co-translationally with GABRA1 and that its recognition epitope may be inaccessible after GABRA1 folds and assembles fully.

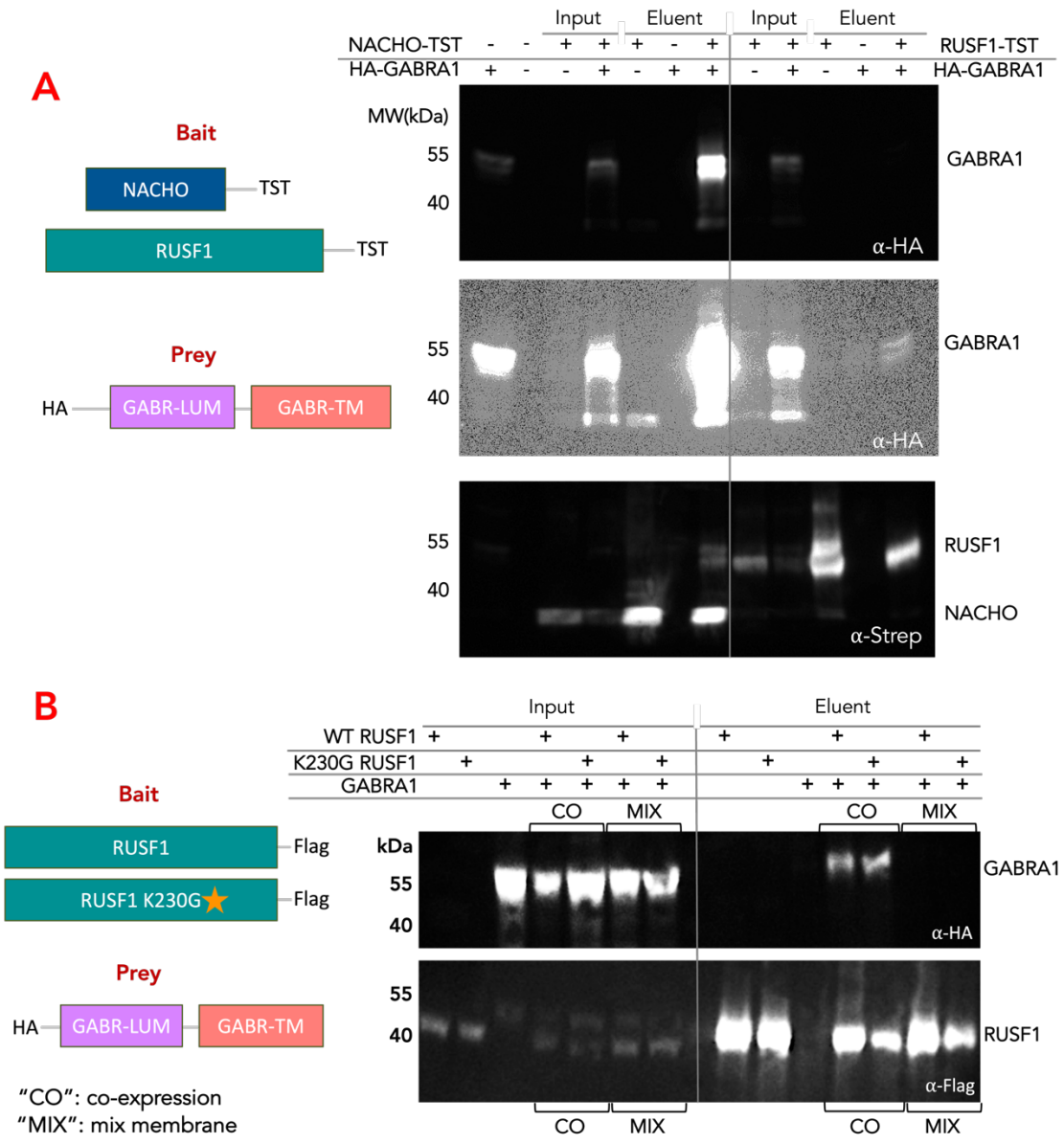


Figure 61. Investigation of RUSF1/GABRA1 Interaction

A. Co-immunoprecipitation assay testing interaction between NACHO or RUSF1 and GABRA1. Schematic (left) shows the bait constructs (NACHO-TST or RUSF1-TST) and prey (HA-tagged GABRA1). Input and elution fractions were probed with anti-HA (to detect GABRA1) and anti-Strep (to detect bait). GABRA1 co-purified with NACHO robustly, but barely interacted with RUSF1. B. Effect of the K230G mutation or co-translational expression on RUSF1–GABRA1 interaction. Schematic (left) shows WT RUSF1 or K230G mutant (FLAG-tagged) as bait and HA-tagged GABRA1 as prey. Co-immunoprecipitation was performed under two conditions: co-expression (CO) and mixed membranes (MIX). Both WT RUSF1 and K230G mutant pulled down comparable amount of GABRA1, but only when they are co-expressed.

Lastly, identifying the domain of GABRA1 with which RUSF1 interacts was crucial. Full-length GABRA1 is approximately a 55kDa protein with a sizeable luminal domain and four

transmembrane helices. Constructs containing either only the luminal domain or only the transmembrane domain (TMD), each with N-terminal HA tags were made, and expressed as prey, along with full-length constructs, in combination with twin-strep-tagged RUSF1 bait (Figure 62). Following strep pulldowns, western blot analyses confirmed the successful expression of all constructs: full-length ("FL", ~55kDa), luminal domain ("LU", ~35kDa), and TMD ("TM", <20kDa).

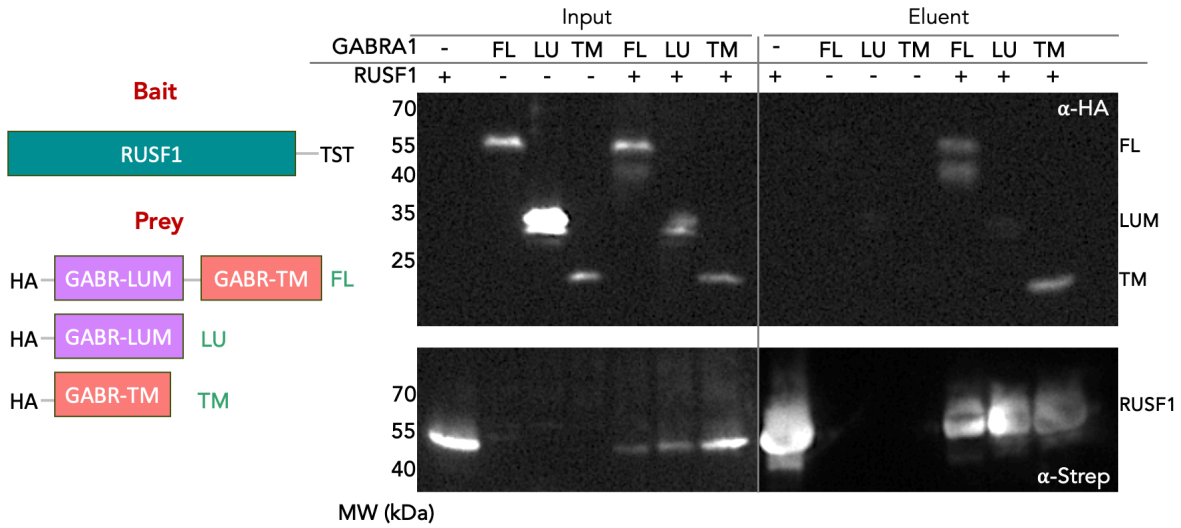


Figure 62. Epitope Binning of RUSF1/GABRA1 Interaction

Co-immunoprecipitation assay testing which domains of GABRA1 mediate interaction with RUSF1. Left: schematic of bait (RUSF1-TST) and prey constructs, including full-length GABRA1 (FL), the luminal domain (LU), and the transmembrane domain (TM), all tagged with HA. Right: input and elution fractions probed with anti-HA (top) to detect GABRA1 constructs and anti-Strep (bottom) to detect RUSF1. Full-length GABRA1 and the transmembrane domain co-purify with RUSF1, whereas the luminal domain alone does not, indicating that the interaction is mediated primarily through the transmembrane region of GABRA1.

Notably, only the full-length and TMD constructs were enriched in RUSF1 pulldowns, while the luminal domain was not. This indicated that RUSF1 specifically interacts with the transmembrane domain of GABRA1, aligning with the interaction mode seen with NACHO. This observation was consistent with previous mass-spectrometry screening setup, where the GABRA1 construct used contained only the transmembrane domain. In addition, the soluble domain of GABRA1 is luminal facing, while the RUSF1 soluble domains were cytosolic

facing, and its TM domain was likely largely buried in membrane. Given the orientation and structural constraints of both proteins, it is logical that RUSF1 would not robustly interact with the luminal domain of GABRA1.

6.2 Identification of Potential Clients of RUSF1

To further investigate potential interactors of RUSF1 beyond GABRA1, which had only demonstrated weak or partial interaction, I aimed to identify additional cellular binding partners using a proteomics approach. This investigation necessitated a highly specific and reliable system to isolate true interactors from nonspecific proteins. Consequently, I employed the SpyCatcher-SpyTag technology (Figure 63 A), previously developed by the Howarth lab (Keeble *et al*., 2019), based on the bacterial protein FbaB. Within FbaB, the CnaB2 domain forms a highly stable internal isopeptide bond between the ε-amine group of lysine residue K31 and the carboxyl group of aspartic acid residue D117. The Howarth lab split this domain into two complementary parts: SpyCatcher, the incomplete immunoglobulin-like domain containing the reactive lysine, and SpyTag, the complementary peptide containing the reactive aspartate. Upon mixing, these fragments spontaneously and covalently link, providing a robust and irreversible interaction that greatly facilitates selective protein capture and analysis.

To utilize this system effectively for identifying RUSF1 interactors, I expressed RUSF1 in HEK293F cells, engineering it to include a C-terminal FLAG-SpyTag. Concurrently, SpyCatcher was produced recombinantly in *E. coli*, purified, and biotinylated via the BirA enzyme. Immunoprecipitation assays were subsequently performed using DDM-CHS solubilized RUSF1-FLAG-Spy samples, applying them to Streptavidin resin immobilized with SpyCatcher (Figure 63 B, C). To confirm the specificity and validate the approach, I incorporated appropriate controls: a positive control consisting of the RUSF1 sample applied

to FLAG resin, and a negative control with RUSF1 sample applied to Streptavidin resin without immobilized SpyCatcher (Figure 63 D, E).

Following binding, the resins were washed rigorously at least four times to eliminate nonspecific binders, and bound complexes were eluted with SDS sample buffer. The eluents were resolved on acrylamide gels and visualized using silver staining. In the gel analysis, a prominent band corresponding to RUSF1 was readily observable in the FLAG resin eluents, as anticipated, yet absent in both negative controls and SpyCatcher eluents (Figure 63 D, E). This is because in the SpyCatcher-based samples, RUSF1 was irreversibly and covalently bound to the streptavidin-immobilized SpyCatcher, preventing its elution even in harsh conditions, such as the presence of SDS. This feature could enhance the visualization of potential RUSF1 binding proteins, particularly those near the molecular weight range of RUSF1 itself. The co-IP was also repeated, and eluents were run on an 8% acrylamide gel in addition to a 15% gel to better separate higher MW proteins (Figure 63 E).

To systematically identify potential interacting proteins, I examined bands present in both FLAG and SpyCatcher elution lanes but absent from the negative control lane (Figure 63 D, E). Notably, the SpyCatcher eluents revealed two faint bands within the 15–25 kDa molecular weight range, a distinct band just below 70 kDa, and additional bands above 130 kDa. These specific regions, along with corresponding negative control regions, were excised from the gels, subjected to trypsin digestion, and prepared for subsequent mass spectrometry analysis.

Following peptide identification, protein species that appeared with equal or greater abundance in the control eluate compared with the RUSF1 samples were removed from analyses, since these were likely non-specific contaminants. Remaining candidates were scored based on their peptide-spectrum match (PSM), providing a confidence metric for each potential interactor.

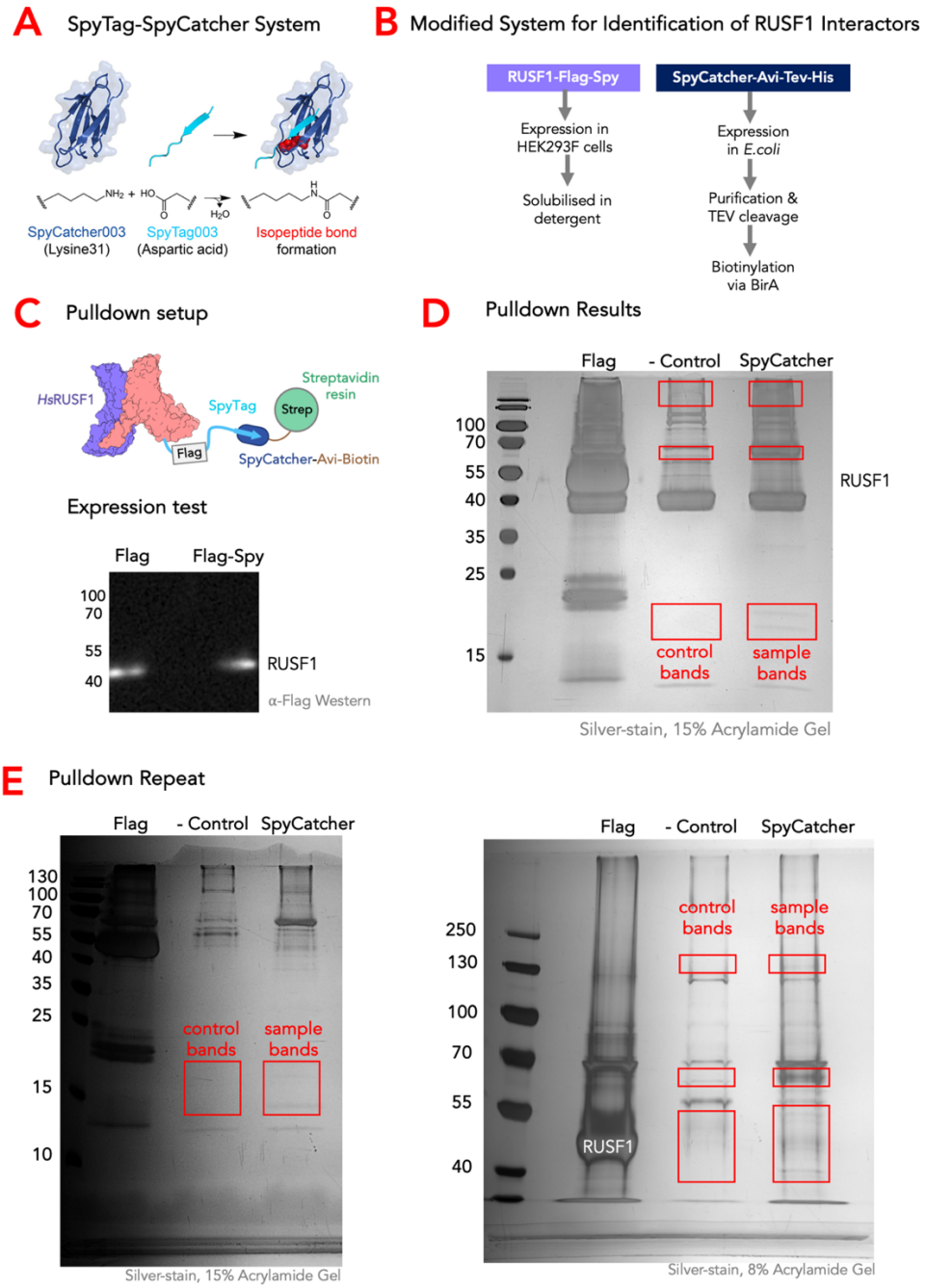


Figure 63. Identification of RUSF1 Interactors Through the SpyTag-SpyCatcher System

A. Schematic of the SpyTag–SpyCatcher system. SpyCatcher contains a lysine residue that forms an isopeptide bond with the aspartic acid in SpyTag, enabling covalent complex formation (Keeble *et al.*, 2019). Image adapted from Rahikainen *et al.*, 2023. B. Workflow for identifying RUSF1 interactors. RUSF1-FLAG-Spy was expressed in HEK293F cells and solubilized in detergent, while SpyCatcher-Avi-TEV-His was expressed and biotinylated in *E. coli* and biotinylated to be immobilised on Streptavidin resin. C. Pulldown assay setup. RUSF1-FLAG-Spy was immobilized via streptavidin resin through SpyCatcher, and potential interactors is pulled down and eluted by SDS sample buffer. Expression of RUSF1-FLAG-Spy construct was verified by anti-FLAG western blot. D. Pulldown results. Silver-stained SDS-PAGE shows protein bands specifically enriched in the SpyCatcher pulldown compared with controls. Red boxes highlight candidate bands corresponding to potential RUSF1 interactors. E. Independent repeat pulldown assays. Silver-stained gels (15% and 8% acrylamide) confirm reproducibility of RUSF1-specific interactors.

To increase robustness, we repeated the co-IP and mass spectrometry workflow twice, and only proteins detected in both independent pull-downs were carried forward for further analysis (Table 4).

The final list of reproducible hits was ranked by their average PSM-derived confidence scores across the two pulldowns. To gain structural insight into each putative interaction, I used AlphaFold3 to model RUSF1 as a dimer in complex with each candidate. We recorded both the pTM score, reflecting overall model reliability, and the iPTM score, reflecting interface confidence. Finally, for each prediction we examined the Predicted Aligned Error (PAE) plots—specifically looking for cross-peak signals indicating regions of high local confidence in the inter-protein interface.

Overall, the results were convoluted. First of all, the confidence scores for the same protein were sometimes inconsistent between two repeats of pulldown experiments. Protein species such as ribosomal subunits are generally contaminants in IP experiments, especially in over-expression systems, however, they could dominate mass spec results if resin was not thoroughly washed. Indeed, although large ribosomal subunit protein uL14 scored very highly in mass spectrometry, the corresponding AlphaFold3 model showed no confident interface with RUSF1, confirming its status as a false positive (Figure 64 A, B).

Table 4. Identification of Potential Interactors of RUSF1

List of candidate RUSF1 interactors identified through the SpyTag–SpyCatcher pulldown system and subsequent proteomic analysis. The table includes UniProt ID, protein name, and confidence score from replicate samples (S1, S2) from proteomics analysis. AlphaFold predictions were also analysed, pTM and iPTM scores are presented for each prediction. Predicted epitope localization (cytosolic, luminal, transmembrane, or both) is also shown, along with analysis of PAE (Predicted Aligned Error) plots. Proteins highlighted in red are discussed further below.

Uniprot ID	Name	S1	S2	iPTM	pTM	Epitope Predicted	PAE Plot
P62829	**Large ribosomal subunit protein uL14**	15.05	396.42	0.44	0.52	cytosolic	no cross peak
Q8N163	**Cell cycle and apoptosis regulator protein 2**	48.11	162.80	0.35	0.45	luminal	cross peak
P23396	**Small ribosomal subunit protein uS3**	81.86	99.67	0.44	0.53	cytosolic	no cross peak
Q9HCN8	**Stromal cell-derived factor 2-like protein 1**	111.79	2.27	0.42	0.55	cytosolic	**cross peaks**
P22670	**MHC class II regulatory factor RFX1**	12.62	93.31	0.35	0.43	both/luminal	**cross peaks**
Q96RE7	**Nucleus accumbens-associated protein 1**	25.83	58.73	0.38	0.48	cytosolic	no cross peak
Q96QF0	**Rab-3A-interacting protein**	2.15	80.70	0.42	0.52	luminal	cross peak
P51610	**Host cell factor 1**	17.27	61.63	0.34	0.41	both/luminal	**cross peaks**
P16220	**Cyclic AMP-responsive element-binding protein 1**	3.39	63.88	0.48	0.58	luminal	cross peak
Q9HCK8	**Chromodomain-helicase-DNA-binding protein 8**	43.47	23.67	0.31	0.37	cytosolic	no cross peak
P82979	**SAP domain-containing ribonucleoprotein**	41.14	23.32	0.51	0.6	cytosolic	no cross peak
Q2M2I8	**AP2-associated protein kinase 1**	26.62	34.96	0.36	0.46	luminal	**cross peaks**
Q9P2N5	**RNA-binding protein 27**	18.41	42.00	0.42	0.47	both	cross peak
Q9H061	**Transmembrane protein 126A**	50.66	8.40	0.44	0.55	TM	**overall confident**
Q15750	**TGF-beta-activated kinase 1**	22.27	31.99	0.37	0.47	cytosolic	no cross peak
Q96TA2	**ATP-dependent zinc metalloprotease YME1L1**	29.78	23.78	0.36	0.45	both/luminal	**cross peaks**
P62258	**14-3-3 protein epsilon**	32.66	18.17	0.47	0.56	cytosolic	**overall confident**
Q5T8P6	**RNA-binding protein 26**	18.99	30.53	0.44	0.48	both/luminal	cross peak
P35232	**Prohibitin 1**	43.92	1.79	0.43	0.52	luminal/TM	**cross peak**
Q15723	**ETS-related transcription factor Elf-2**	23.67	21.94	0.31	0.45	cytosolic	no cross peak
Q9Y6X9	**ATPase MORC2**	6.88	34.66	0.33	0.43	both/luminal	cross peak
O15075	**Serine/threonine-protein kinase DCLK1**	11.05	28.48	0.34	0.44	cytosolic	no cross peak
P19105	**Myosin regulatory light chain 12A**	21.45	17.27	0.48	0.58	cytosolic	**cross peaks**
Q92974	**Rho guanine nucleotide exchange factor 2**	16.73	21.98	0.33	0.43	luminal	cross peak
Q6ZSR9	**Uncharacterized protein FLJ45252**	18.03	18.36	0.45	0.57	cytosolic/cavity	**cross peaks**
Q07021	**C1QBP**	30.50	2.13	0.4	0.51	cytosolic	no cross peak
P06748	**Nucleophmin**	22.24	9.23	0.45	0.55	luminal	no cross peak
Q9BY77	**Polymerase delta-interacting protein 3**	2.38	29.08	0.45	0.55	luminal	cross peak
Q99569	**Plakophilin-4**	6.38	22.59	0.33	0.43	both/luminal	**cross peaks**
Q14739	**Delta(14)-sterol reductase LBR**	10.71	15.94	0.36	0.47	luminal/TM	cross peak
Q86V48	**Leucine zipper protein 1**	1.87	23.78	0.35	0.45	luminal	no cross peak
Q8WYP5	**Protein ELYS**	3.51	21.92	0.31	0.4	both/luminal	cross peak
P60709	**Actin, cytoplasmic 1**	13.68	9.21	0.39	0.5	cytosolic	no cross peak
P18846	**Cyclic AMP-dependent transcription factor ATF-1**	6.66	14.64	0.5	0.6	luminal	cross peak
Q5SW79	**Centromal protein of 170 kDa**	13.65	7.38	0.33	0.42	luminal	**cross peaks**
P98175	**RNA-binding protein 10**	2.31	18.49	0.36	0.47	luminal	no cross peak
Q9NQY0	**Bridging integrator 3**	18.07	2.43	0.44	0.52	cytosolic	no cross peak
Q9NXV6	**CDKN2A-interacting protein**	3.89	16.53	0.39	0.5	luminal	cross peak
P78364	**Polyhomeotic-like protein 1**	1.63	18.54	0.38	0.48	luminal	**cross peaks**
Q16594	**Transcription initiation factor TFIID subunit 9**	14.60	4.41	0.49	0.58	luminal	not inside cavity
Q9ULW0	**Targeting protein for Xklp2**	11.63	6.35	0.36	0.47	luminal	cross peak
P25208	**Nuclear transcription factor Y subunit beta**	9.98	6.33	0.51	0.59	luminal	cross peak
Q9Y2Q9	**Small ribosomal subunit protein bS1m**	8.44	7.75	0.42	0.53	cytosolic/cavity	**cross peaks**
Q8WWM7	**Ataxin-2-like protein**	13.35	2.57	0.38	0.48	luminal	cross peak
Q96GA3	**Protein LTV1 homolog**	3.18	12.74	0.4	0.5	luminal	no cross peak
P07477	**Serine protease 1**	7.33	7.83	0.38	0.51	cytosolic	no cross peak
Q6ZTW0	**Tubulin polyglutamylase complex subunit 1**	10.94	3.19	0.4	0.52	cytosolic	**cross peaks**
O95835	**Serine/threonine-protein kinase LATS1**	6.02	7.81	0.33	0.43	luminal	cross peak
Q9NZI8	**Insulin-like growth factor 2 mRNA-binding protein 1**	10.55	3.13	0.38	0.48	luminal	cross peak
Q8N9M5	**Transmembrane protein 102**	2.37	11.17	0.34	0.46	cytosolic/cavity	not confident
Q03060	**cAMP-responsive element modulator**	7.10	5.88	0.49	0.59	luminal	**cross peaks**
Q96EV2	**RNA-binding protein 33**	6.56	5.99	0.37	0.46	both/luminal	**cross peaks**
Q6NW34	**Nucleolus and neural progenitor protein**	1.65	10.88	0.37	0.47	luminal	no cross peak
Q9Y520	**Protein PRRC2C**	4.13	7.33	-	-	-	prediction failed
Q14671	**Pumilio homolog 1**	1.83	9.09	0.37	0.46	both/cytosoic	**cross peaks**
Q96KK5	**Histone H2A type 1-H**	2.42	7.64	0.53	0.59	luminal	**cross peaks**
O75928	**E3 SUMO-protein ligase PIAS2**	5.09	4.90	0.37	0.47	both/luminal	cross peak
Q9GZU7	**CTDSP1**	7.73	2.02	0.43	0.55	cytosolic/cavity	no cross peak
Q9P270	**SLAIN motif-containing protein 2**	7.06	2.27	0.43	0.54	luminal	**cross peaks**
Q6ZU65	**Ubinuclein-2**	3.78	5.10	0.35	0.45	luminal	**cross peaks**
Q8N5L8	**Ribonuclease P protein subunit p25-like protein**	5.22	2.19	0.49	0.56	luminal	**cross peaks**
Q15054	**DNA polymerase delta subunit 3**	2.10	4.46	0.42	0.52	luminal	no cross peak
Q9ULH1	**Arf-GAP with SH3 domain**	2.57	3.71	0.33	0.42	both/luminal	**cross peaks**
Q92665	**Small ribosomal subunit protein mS31**	3.20	2.64	0.41	0.51	luminal	**cross peaks**
P23469	**Receptor-type tyrine-protein phosphatase epsilon**	3.83	1.82	0.36	0.46	both	**cross peaks**
P38398	**Breast cancer type 1 susceptibility protein**	2.22	2.00	0.32	0.4	both	no cross peak
Q12802	**A-kinase anchor protein 13**	1.74	2.45	0.33	0.4	luminal	cross peak
Q9Y6X8	**Zinc fingers and homeobes protein 2**	1.92	2.24	0.36	0.47	cytosolic	no cross peak
Q9UQR0	**Sex comb on midleg-like protein 2**	2.00	2.15	0.35	0.45	cytosolic	no cross peak
P04406	**Glyceraldehyde-3-phphate dehydrogenase**	2.22	1.69	0.4	0.5	cytosolic	no cross peak
Q9NVJ2	**ADP-ribosylation factor-like protein 8B**	2.02	1.73	0.43	0.55	cytosolic/cavity	**cross peaks**
Q9UHY1	**Nuclear receptor-binding protein**	2.00	1.67	0.39	0.49	cytosolic	cross peak
P62280	**Small ribosomal subunit protein uS17**	1.85	1.72	0.45	0.57	cytosolic	no cross peak
P62847	**Small ribosomal subunit protein eS24**	0.34	2.28	0.49	0.56	no	no cross peak
P46776	**Large ribosomal subunit protein uL15**	0.17	2.03	0.47	0.54	luminal	no cross peak
P30626	**Sorcin**	0.09	1.66	0.46	0.58	cytosolic/cavity	**cross peaks**

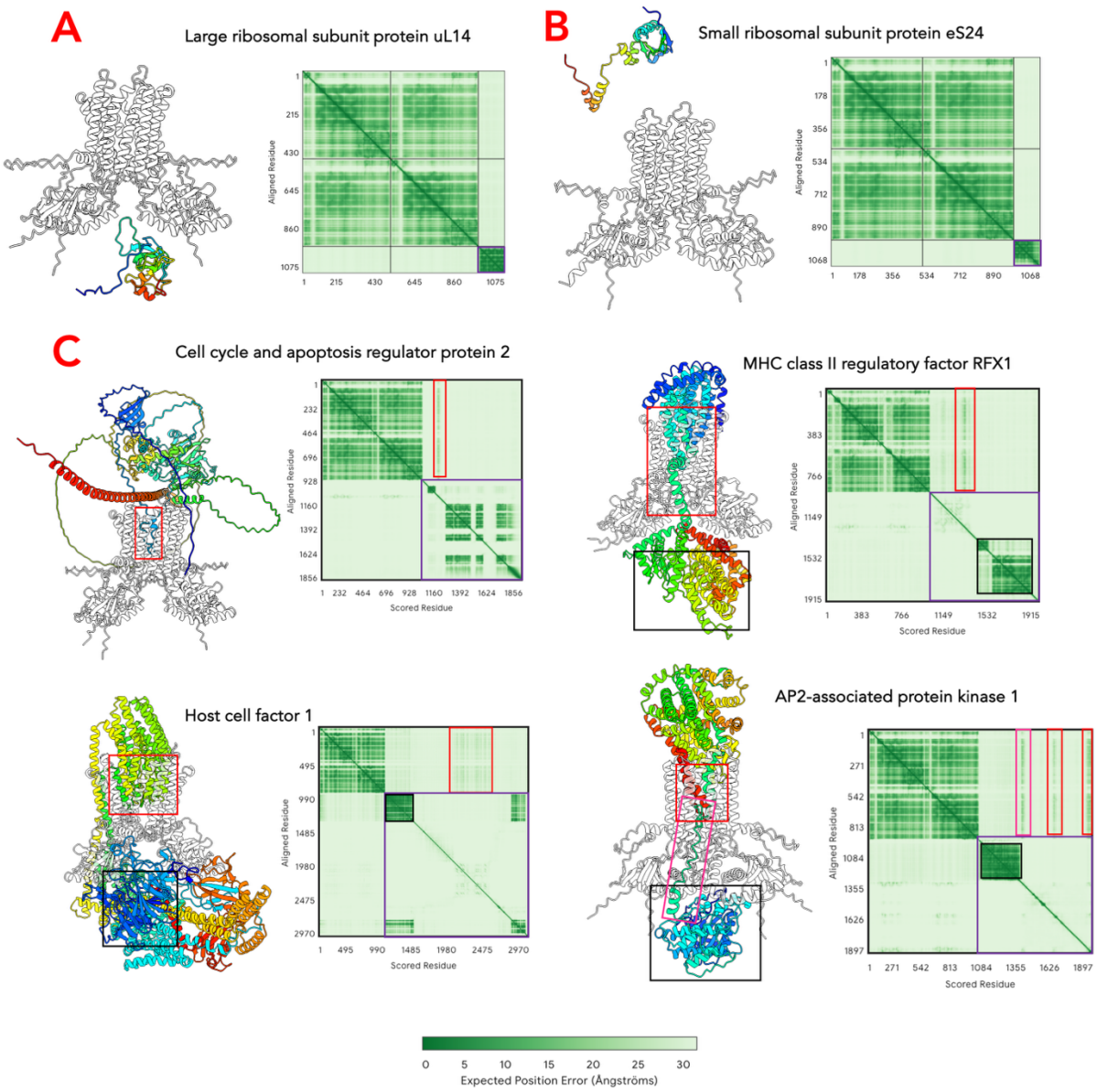


Figure 64. Structural predictions and PAE (Predicted Aligned Error) plots of selected RUSF1 interactors

A, B. AlphaFold 3 predictions of RUSF1 in complex with ribosomal subunits uL14 or eS24. Despite appearing as top hits of proteomic analysis, AlphaFold did not predict any confident interactions between uL14 or eS24 and RUSF1 (as cross peak in PAE plot), indicating these hits are likely artefacts. C. AlphaFold 3 predictions of RUSF1 in complex with various hits (names indicated above each prediction) identified in mass spectroscopy. These predictions all exhibit some confidence in interactions in relation to RUSF1 cavity, but not to the rest of the protein.

Secondly, although PTM scores and iPTM scores should provide some insights into the confidence of structural prediction, it could be observed that proteins with existing high-resolution structures tended to yield artificially elevated pTM and iPTM values, even when

their predicted interfaces lacked physical plausibility. For example, the prediction of the small ribosomal subunit protein eS24 in complex with RUSF1 generated pTM score of 0.56 and iPTM of 0.49, despite showing no significant PAE cross-peaks. Moreover, across the list shown, the PTM and iPTM score did not correlate with the confidence score from mass spectroscopy results whatsoever, only adding complications to analysis.

Instead, the most helpful indicator for analysis was the presence of localized PAE cross-peaks.

In several top-ranked hits, the predicted interaction "hotspots" (visible as red-boxed cross-peaks) mapped to disordered segments (that AlphaFold3 sometimes falsely models as helices), indicated by the high expected error in the purple squares (Figure 64 A, B). Interestingly, all of these regions bind to the charged cavity of RUSF1. The lack of overall confidence combined with confidence in local regions suggested these are promiscuous electrostatic contacts which could occur during immunoprecipitation experiments, but does not represent genuine and stable binding interfaces (Figure 64 C).

Nevertheless, there were still promising candidates (Figure 65-67, Supplementary Figure 11). TMEM126A also appeared in the top one third of the hits, and AF3 prediction of 2:2 TMEM126A:RUSF1 showed cross peaks across all transmembrane region (Figure 65 A-C). Visual inspection of the model revealed that TMEM126A adopts a fold reminiscent of cornichon membrane chaperones and positions several positively charged residues through a negatively charged cleft formed between the soluble domain of RUSF1 and one of its amphipathic helices.

To test this prediction experimentally, RUSF1 bearing a C-TST tag was co-expressed with TMEM126A bearing a C-Myc tag. Immunoprecipitation was carried out in LMNG:CHS detergent, and Western blotting showed that RUSF1 enriched a small fraction of TMEM126A (Figure 65 D). Although the interaction was detectable by western blot, it was too weak to

visualize on Coomassie–stained gels. Further optimization of detergent composition and buffer conditions may strengthen the interaction and allow for capture of this complex.

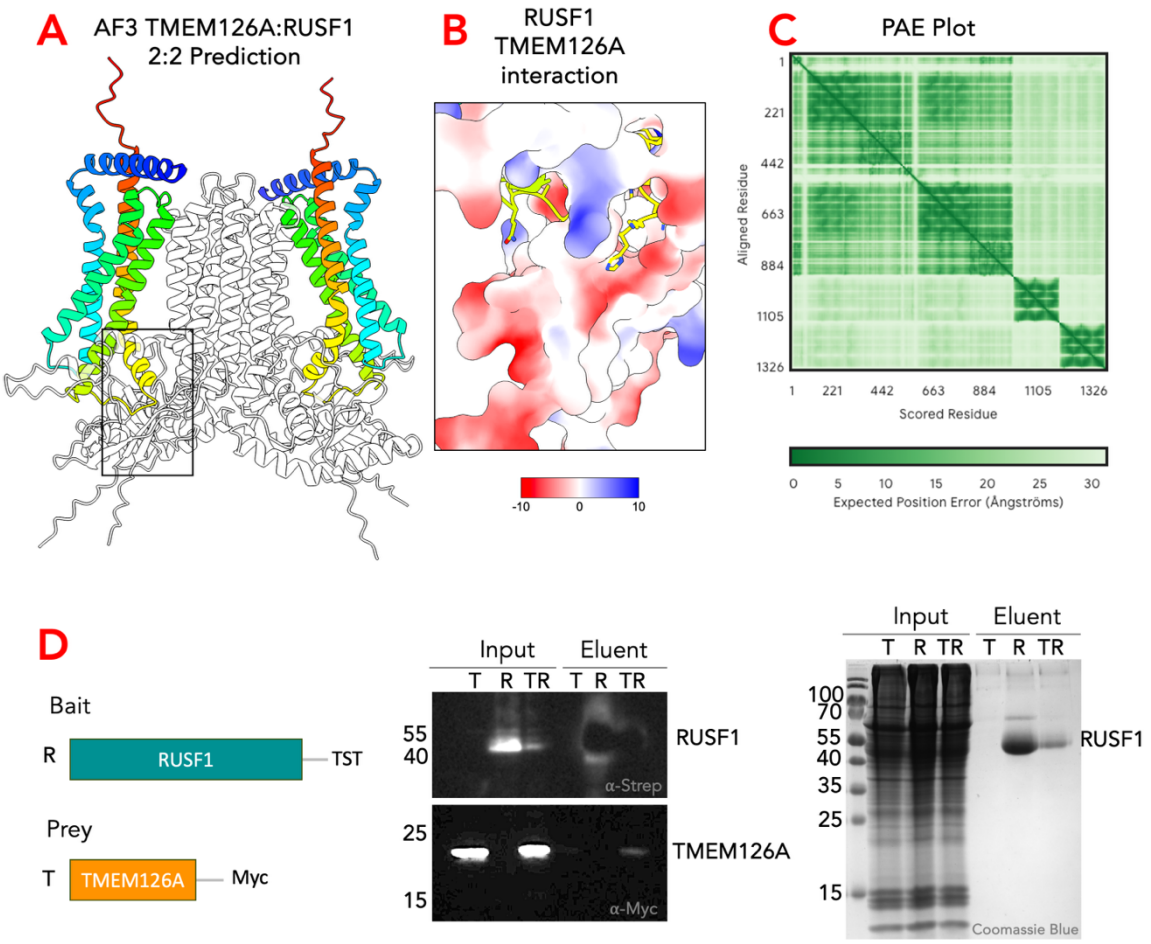


Figure 65. Verifying Interactions Between RUSF1 and TMEM126A

A. AF3 prediction of a RUSF1:TMEM126A 2:2 complex. Predicted models show RUSF1 (white) and TMEM126A (rainbow-colored N- to C- terminus). B. Electrostatic surface potential of RUSF1 at the TMEM126A interface, showing complementary charge distribution that may stabilize binding. C. PAE (Predicted Aligned Error) plot of the RUSF1–TMEM126A complex, showing cross peak (confidence) across both transmembrane domain and soluble domain. D. Experimental validation of the interaction by co-immunoprecipitation. Left: schematic of the pulldown assay with RUSF1-TST (bait) and TMEM126A-Myc (prey). Right: Input and elution fractions probed by western blot (anti-Strep for RUSF1, anti-Myc for TMEM126A) and Coomassie staining. TMEM126A co-purifies with RUSF1, which is visible on western blot, but not Coomassie stained gels.

TMEM126A has been implicated in co-translational quality control within the mitochondrial inner membrane, though its potential localization to other organelles (e.g., ER or Golgi)

remains uncharacterized. Should RUSF1 and TMEM126A truly complex *in vivo*, it could suggest a broader role for RUSF1 in membrane protein quality control or organelle crosstalk.

Interestingly, TMEM126A was not the only mitochondrial protein that appeared in the screen. Prohibitin 1, part of a large chaperone complex on the mitochondrial inner membrane, was also pulled down. AF3 predicted its transmembrane helix to be parallel to RUSF1's transmembrane domain, though no interaction was formed upon visual inspection; however, a basic C-terminal region of Prohibitin 1 docked into RUSF1's cavity with moderate confidence (Figure 66 A-C). Compared to the predicted interactions shown above in figure 64, this seemed more physiologically possible.

In March 2025, the first structure of the prohibitin1-prohibitin2 complex was solved *in situ* by cryoET to ~16 Å (Figure 66 D, Lange *et al.,* 2025, PDB 8RRH), which revealed a structural arrangement of prohibitin 1 distinct from the AlphaFold prediction shown above. A more compact N-terminal globular domain was observed, and C-terminal helices were extended to interact with other protomers of the ring.

Although the localisation of prohibitin complex in places other than mitochondria is still debated, and the localisation of RUSF1 is also under active investigation (recent data from the lab suggested mostly ER localisation), if RUSF1 and Prohibitin 1 do interact in cells, two models emerge (Figure 66 D): (1) a cis interaction within the same organelle membrane, leveraging Prohibitin's flexible C-terminal helices to dock into RUSF1's cavity; or (2) a trans-interaction across different organelles where the extended C-terminal helices of prohibitin 1 interact with the cavity of RUSF1 on a different membrane, allowing for molecular signals to be transmitted across different cellular compartments.

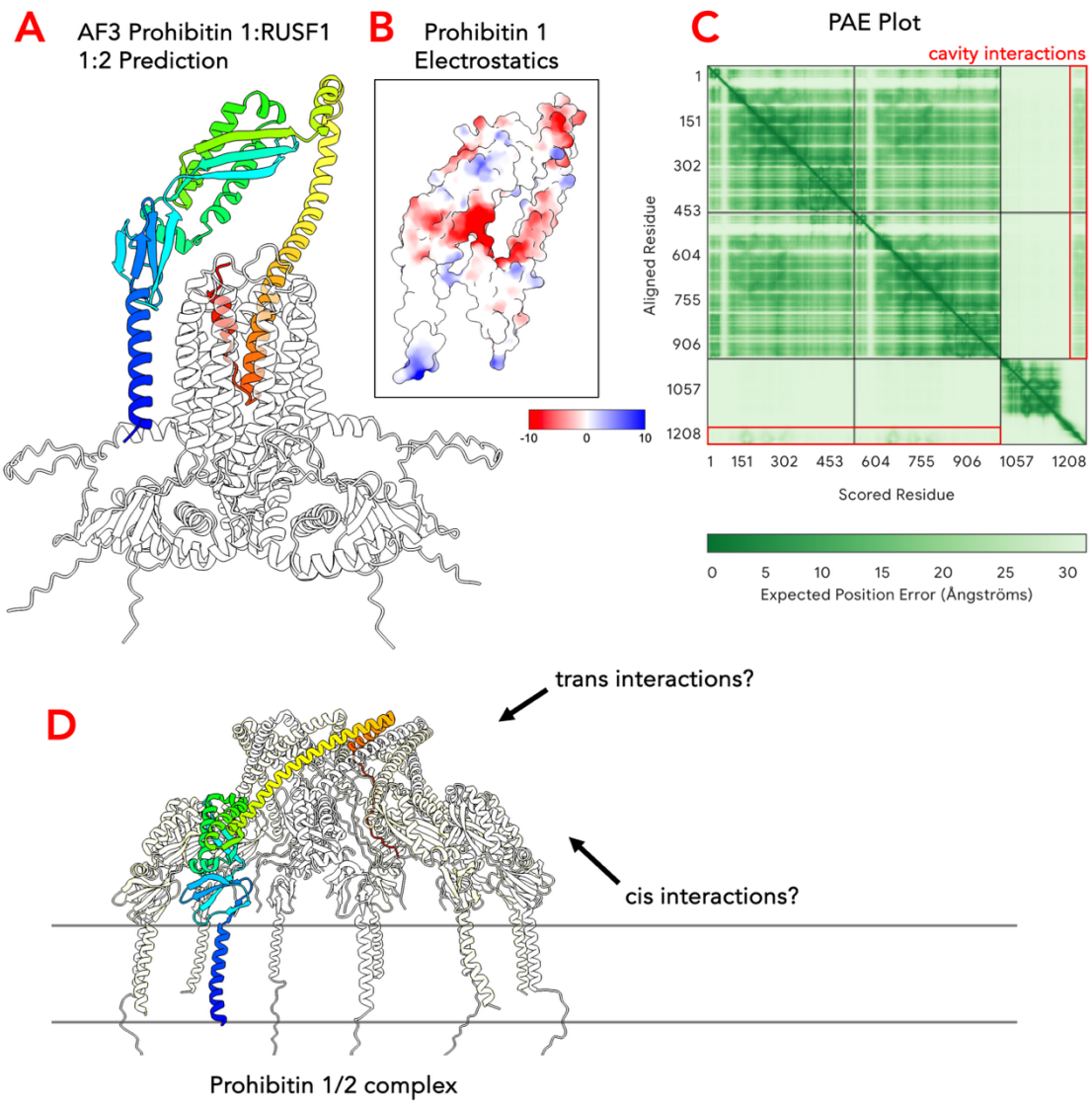


Figure 66. AlphaFold Prediction of RUSF1/Prohibitin Complex

A. AF3 prediction of a 1:2 Prohibitin–RUSF1 complex. Prohibitin (rainbow-colored N- to C-terminus) is modelled interacting with the cavity of RUSF1 (white). B. Electrostatic surface potential of Prohibitin. C. Predicted Aligned Error plot of the Prohibitin–RUSF1 model. Low PAE values (green) in the cavity-interaction regions (highlighted in red) support the confidence of the predicted binding interface. D. CryoET structure of prohibitin complex (PDB 8RRH), where the C-terminus is extended, facilitating oligomeric interactions. Arrows indicate possible sites of interaction with RUSF1.

Another hit worth noting was 14-3-3 epsilon. 14-3-3 proteins are adaptor proteins involved in various cellular signalling pathways, they bind to clients through recognising phosphorylated residues. 14-3-3 epsilon has been shown to function as either monomer or dimer. AlphaFold3 confidently modelled a 14-3-3ε dimer engaging RUSF1's cytosolic domain, with PAE cross-

peaks distributed throughout the predicted interface (Figure 67 A, B). Remarkably, RUSF1's C-terminal tail—containing a potential phospho-Ser site—appears to insert into the canonical substrate-binding groove of 14-3-3ε, mimicking the published crystal structure of 14-3-3ε bound to a phosphorylated PEAK3 peptide (Figure 67 C, D). This structural mimicry strongly suggests that RUSF1 could be a bona fide 14-3-3 client, pending confirmation of its phosphorylation status and binding in cells.

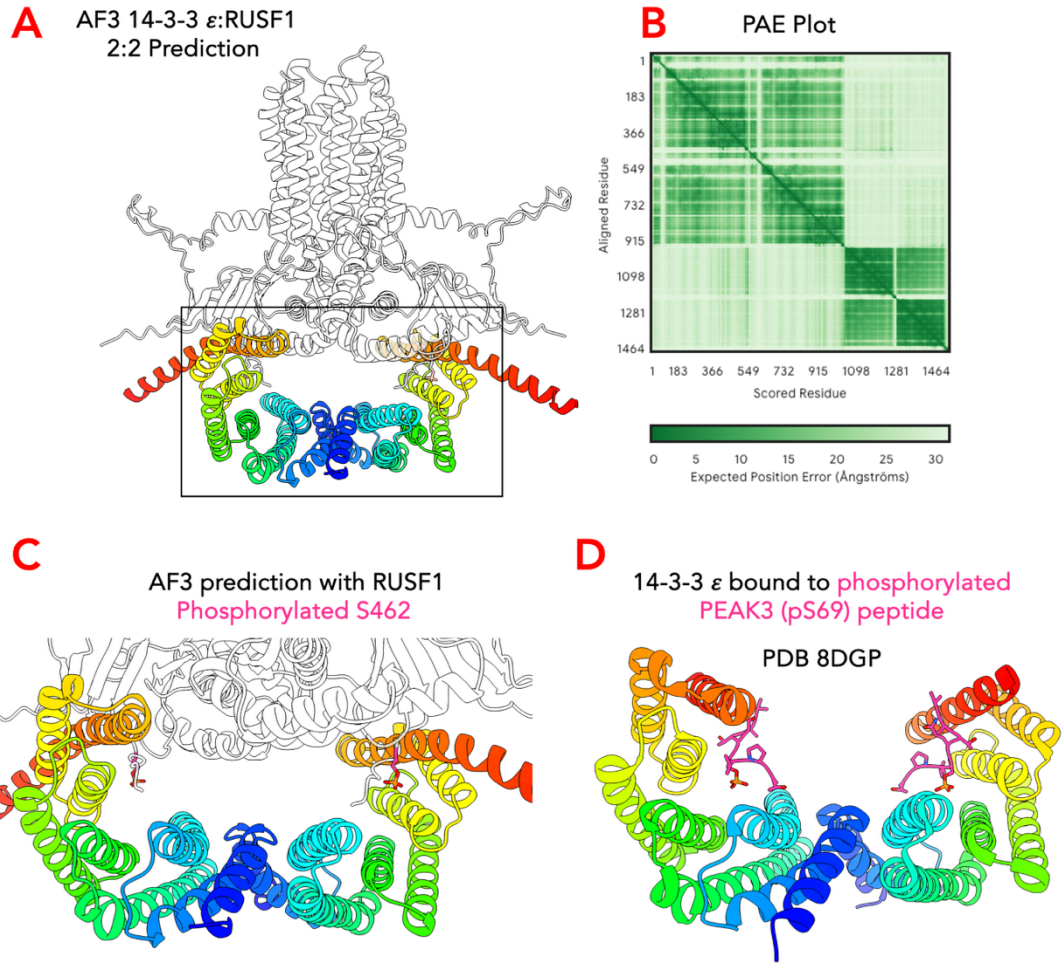


Figure 67. AlphaFold Prediction of RUSF1/14-3-3 Epsilon Complex

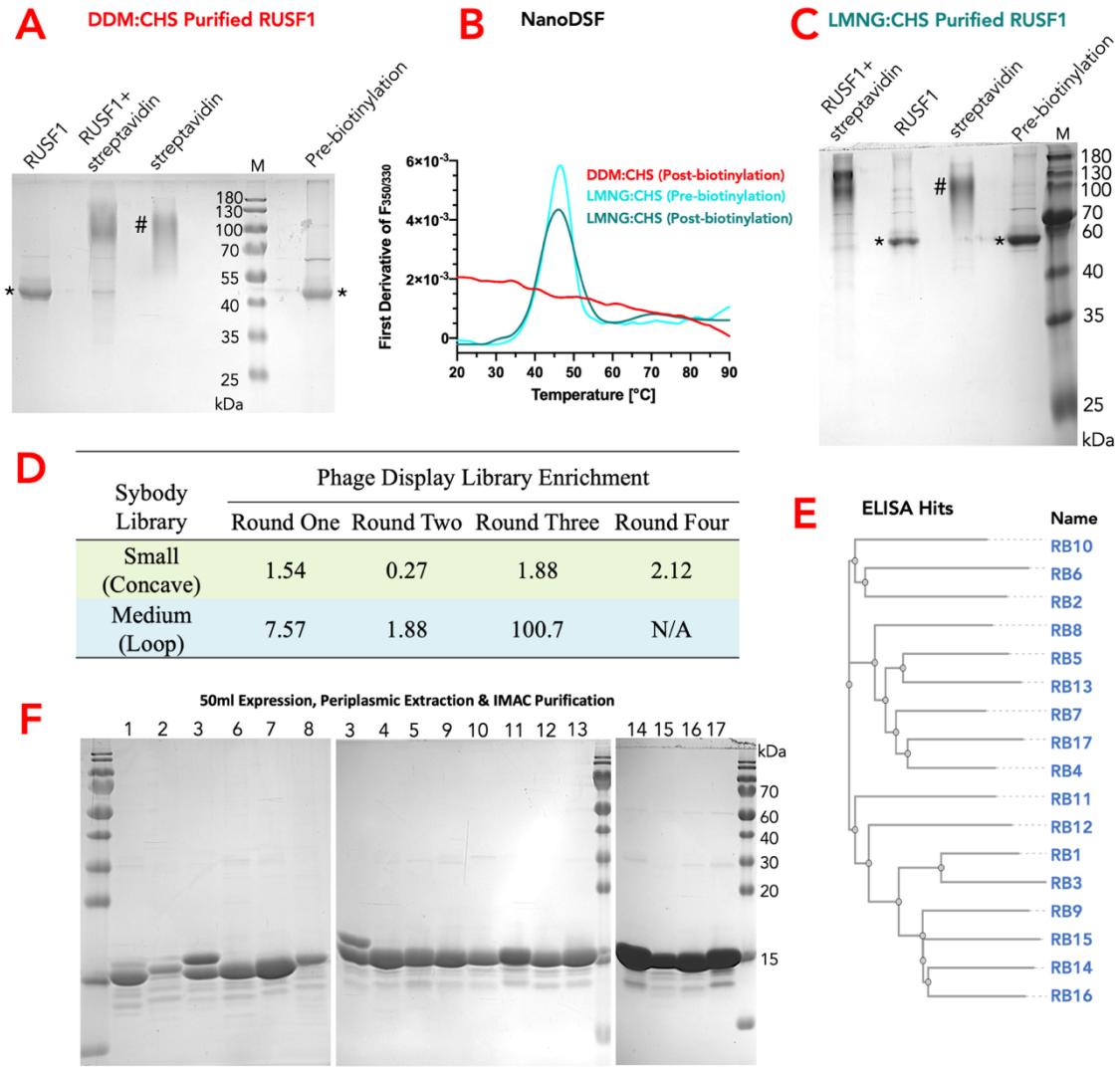
A. AlphaFold-Multimer (AF3) prediction of a 2:2 RUSF1–14-3-3ε complex. The 14-3-3ε dimer (rainbow-colored) is modelled interacting with RUSF1 (white), with binding localized to the cytosolic domain. B. Predicted Aligned Error plot for the RUSF1/14-3-3ε complex. C. AF3 prediction incorporating a phosphorylated RUSF1 residue (S462). Phosphorylated S462 (shown in stick representation) is positioned within the putative 14-3-3ε binding groove. D. Structure of 14-3-3ε bound to a phosphorylated PEAK3 peptide (pS69) (PDB ID: 8DGP). The similarity of binding modes between the predicted RUSF1–14-3-3ε complex and the experimentally determined peptide–14-3-3ε structure supports the plausibility of phosphorylation-dependent recognition of RUSF1 by 14-3-3ε.

In summary, while many top mass-spec hits prove to be artefacts (notably ribosomal proteins and disorder-driven “stickiness”), through combining MS ranking and PAE analysis, multiple candidates including TMEM126A, Prohibitin 1, and 14-3-3ε could be identified for further biochemical and cellular validation.

6.3 Sybody Selection Against RUSF1

Although the molecular structure of RUSF1 was obtained, its interaction with GABRA1 verified, and other potential binding partners investigated in an over-expression system, the identification of RUSF1’s functional role in cell required additional tools, specifically, high affinity binders, that could allow for imaging endogenous RUSF *in vivo*, or quantification of RUSF1 levels in cells in various ER stress conditions either through western blotting, or immunoprecipitation. To that end, I decided to generate synthetic nanobodies specific for RUSF1, utilising the pipeline described in **Chapter 4**.

Initially, RUSF1 was purified with DDM:CHS detergent before biotinylation by BirA, as this was the detergent used for structural studies. The efficiency of biotinylation was tested by the addition of streptavidin, which would lead to a band shift of biotinylated protein on SDS-PAGE. This was observed for DDM:CHS purified RUSF1 (Figure 68 A). However, the biotinylated protein failed to give a thermal shift in NanoDSF (Figure 68 B), suggesting protein unfolding (or abnormal interactions between protein and DDM:CHS micelle) during purification or biotinylation process, possibly due to the highly basic environment during biotinylation process (pH 8.5- 9). Luckily, repeating the purification and biotinylation with LMNG:CHS detergent resulted in thermally stable (Tm ~47 °C), and fully biotinylated RUSF1, indicated by the band shift (Figure 68 B, C).



Sybody Library	Phage Display Library Enrichment			
	Round One	Round Two	Round Three	Round Four
Small (Concave)	1.54	0.27	1.88	2.12
Medium (Loop)	7.57	1.88	100.7	N/A

Figure 68. Synthetic Nanobody Selection Against RUSF1

A. SDS-PAGE of DDM:CHS-purified RUSF1 before and after streptavidin incubation and pre-biotinylation. Bands corresponding to RUSF1 are indicated (*), with excess streptavidin labelled (#). B. Nano differential scanning fluorimetry (NanoDSF) analysis comparing thermal stability of RUSF1 purified in different detergents (DDM:CHS or LMNG:CHS) before and after biotinylation. Duplicate experiments were performed for each sample, and results of one repeat are shown. LMNG:CHS provides superior stability post-biotinylation. C. SDS-PAGE of LMNG:CHS-purified RUSF1 before and after streptavidin incubation and pre-biotinylation. Bands corresponding to RUSF1 are indicated (*), with excess streptavidin labelled (#). D. Phage display enrichment from concave (small) and loop (medium) sybody libraries. Loop library enriched significantly after three rounds of phage displays, while concave library reached an enrichment of 2.12 after four rounds. E. Phylogenetic tree of the 17 sybodies identified through ELISA. F. Expression and purification of sybody clones. Periplasmic extracts from *E. coli* expressing individual sybodies were purified via IMAC and analysed by SDS-PAGE, confirming robust expression for all clones.

Biotinylated RUSF1 was used in ribosome display, phage displays and the final ELISA steps, while additional un-biotinylated RUSF1 also purified in LMNG:CHS was use as competitor in the second phage display, aimed to remove binders that were non-specific or fast off-rate. Of

the two libraries, the loop library enriched after three rounds of phage displays, with an enrichment value of 100.7, while the concave library did not meet the target enrichment value of 2 (enrichment of target RUSF1 smaller than 2-fold of enrichment of control PepTso). Therefore, an additional phage display step was performed, which led to increase of enrichment to 2.12 (Figure 68 D). The enriched sybodies were then identified through ELISA.

The sybody selection pipeline utilizing LMNG-CHS purified RUSF1 successfully generated a total of 18 unique sybodies (Figure 68 E). All identified sybodies originated exclusively from the loop library. Based on sequence similarity, these sybodies could be categorized into three distinct clades: the first clade comprising RB10, 6, and 2; the second including RB8, 5, 13, 7, 17, and 4; the third consisting of RB11, 12, 1, 3, 9, 15, 14, and 16.

Each of these sybodies was able to be expressed and purified from *E. coli*, enabling subsequent analyses to determine their binding affinities to RUSF1 (Figure 68 F). Two experimental methodologies were employed to assess affinity: Ni-NTA pulldowns and ELISA-based affinity assay.

In the Ni-NTA pulldown assays (Figure 69 A), His-tagged sybodies were immobilized onto Ni-NTA resin, followed by incubation with purified RUSF1 for a minimum duration of one hour on ice. After incubation, the resin underwent a series of three rigorous washes, including an extended incubation wash step of 30 minutes at room temperature, designed specifically to eliminate any low-affinity binders or those exhibiting rapid dissociation rates. Following this thorough washing protocol, the sybodies were eluted from the resin using SDS sample buffer. The pulldown assay results indicated that several sybodies successfully enriched RUSF1, notably RB2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 15, and 17 (Figure 69 B). However, the majority demonstrated only weak enrichment of RUSF1, suggesting that many sybodies, such as RB3, 8, 9, 10, 12, 13, 15, and 17, likely possessed relatively fast off-rates or low affinity. In contrast, RB2, 6, 7, and 11 displayed comparatively stronger binding.

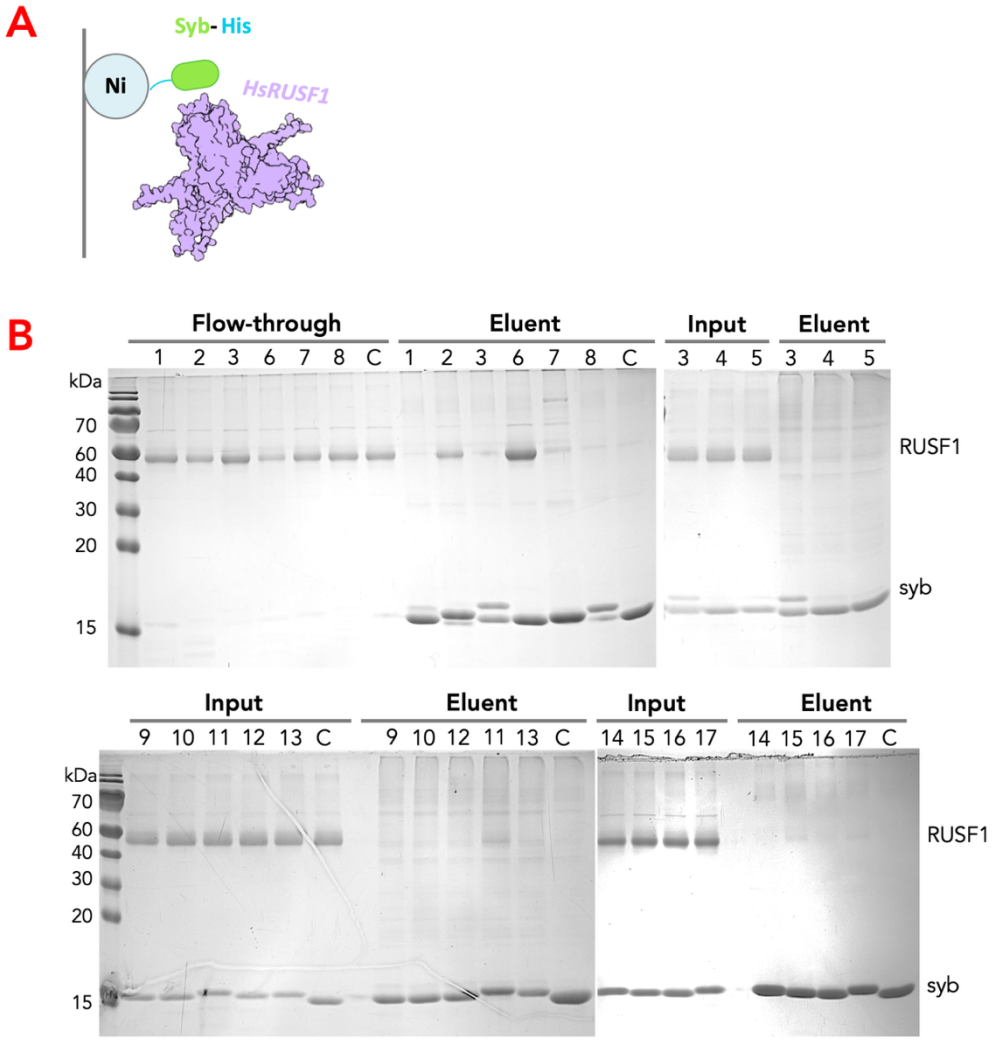


Figure 69. Analysis of Binding of RUSF1 Sybodies

A. Schematic of the pulldown setup. His-tagged sybodies (Syb-His) were immobilized on Ni resin and incubated with purified RUSF1 to test binding. B. SDS-PAGE analysis of pulldown assays with different sybody clones. Flow-through or input fractions, and elution fractions are shown. Several sybodies co-elute with RUSF1, confirming specific binding. The sybody (~15 kDa) and RUSF1 (~55 kDa) bands are visible in eluates.

To further refine the assessment of binding affinities, affinity ELISA was utilized (Figure 70 A, B). This assay involved immobilizing Myc-His tagged sybodies (at multiple concentration points) onto wells via protein A and an anti-Myc antibody, with biotinylated RUSF1 subsequently introduced at a constant concentration (50nM). Signal detection was achieved using streptavidin-conjugated HRP, with the final readout obtained via TMB substrate development (Figure 70 A).

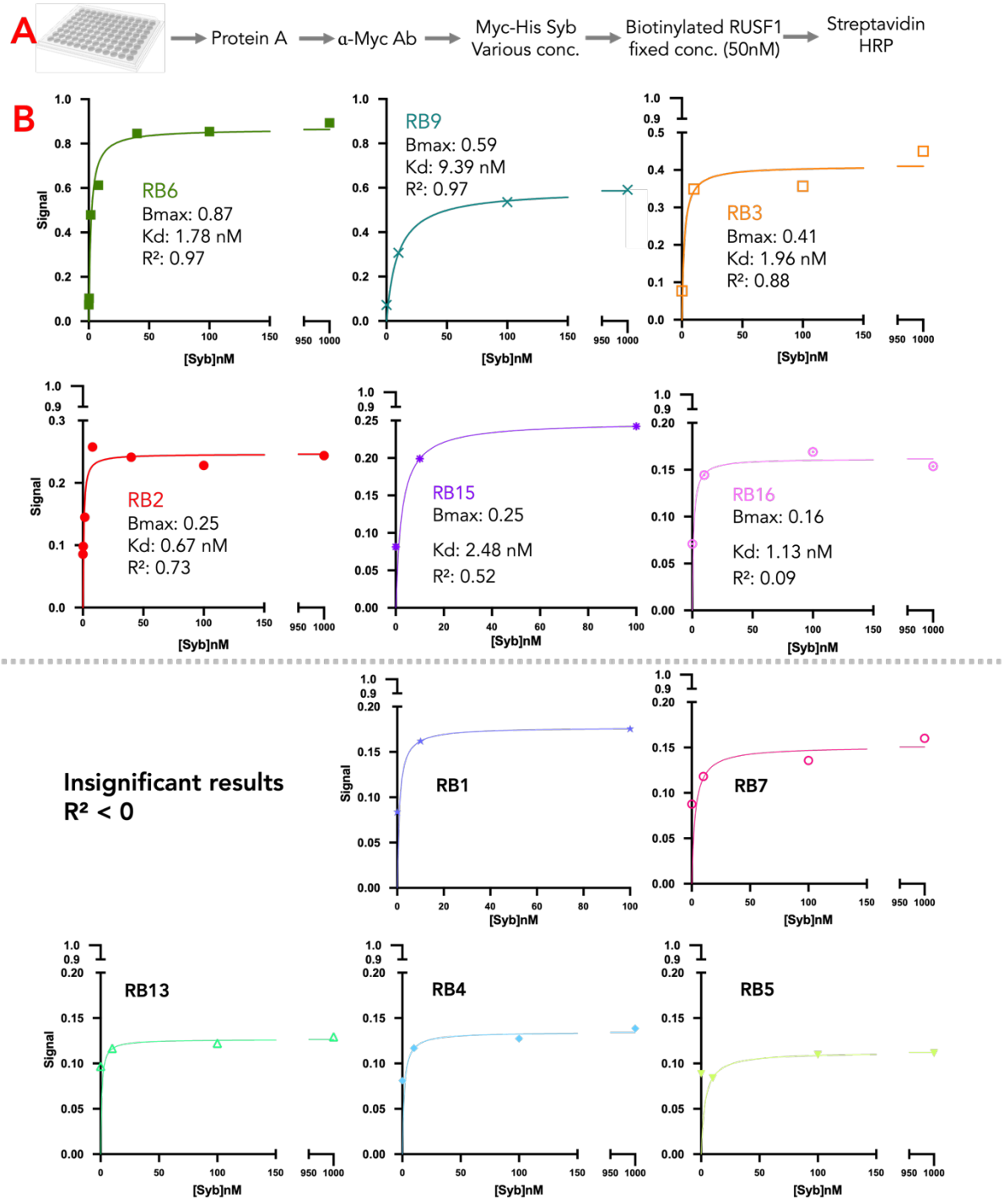


Figure 70. Affinity-based ELISA Assays of RUSF1 Sybodies

A. Schematic of ELISA-based binding assay workflow. Myc-His-tagged sybodies were captured on Protein A plates coated with anti-Myc antibody, followed by incubation with biotinylated RUSF1 (50 nM). Binding was detected using streptavidin-HRP. B. Binding curves for selected sybodies. Nonlinear regression fitting was used to determine binding parameters (B_{max}, K_d, R^2). Only one repeat was done. Plots under the dotted line were insignificant results ($R^2<0$).

Based on initial pulldown data, 12 sybodies were selected for detailed ELISA analysis (RB2, 3, 4, 5, 6, 7, 9, 11, 13, 15, and 16). Most sybodies were tested across four concentrations (0, 10, 100, and 1000 nM), while RB2 and RB6, predicted as high-affinity binders from pulldown assays, were examined with additional lower concentration points (2.5, 5, and 10 nM). Unfortunately, due to handling errors, wells immobilized with RB11 did not yield measurable signals, thus precluding its analysis (Figure 70 B).

Of the remaining sybodies, signal corresponding to various concentrations were plotted, and an apparent K_d could be calculated. though this technique is not as accurate as biolayer interferometry or SPR, nor does the results represent an objective dissociation constant (it's generally significantly over-optimistic), it provides a rough ranking of association kinetics and binding affinity of the sybodies (Figure 70 B).

Five sybodies showed minimal signal (B_{max} approximately or below 0.15, near background levels of 0.08), making the resulting K_d values unreliable (R squared value smaller than 0) and therefore omitted (Figure 70 B). For the six sybodies demonstrating dose-dependent responses, RB16 displayed a poor fit (R^2=0.09), whereas RB2, 3, 6, 9, and 15 exhibited much better fits with R^2 values above 0.5. Among these, RB6 emerged as the highest affinity candidate, evidenced by a steep binding curve and the highest observed signal (B_{max} of 0.87). RB9 also exhibited a strong affinity, with a notable B_{max} of 0.59, though the binding curve was comparatively less steep. RB3, RB2, and RB15 followed, indicating moderate affinity levels suitable for further characterization.

6.4 Utilising Sybodies for Functional Characterisation of RUSF1: Preliminary Trials

To assess the specificity and functionality of sybodies selected against the RUSF1 protein, eight candidate sybodies that demonstrated either pull-down capabilities or positive responses in affinity ELISA assays were further evaluated. The chosen sybodies were RB6, RB2, RB9, RB3, RB15, RB16, RB11, and RB7.

Firstly, these sybodies were employed in fluorescence microscopy experiments using fixed COS-7 cells (Figure 71, done by Martyna Biadun). Cells expressing endogenous levels of RUSF1 was tested, whereas RUSF1 knock-down cells were used as negative controls. Detection was facilitated by anti-Myc antibodies and fluorophore-labelled secondary antibodies. Analysis of fluorescence signal intensities revealed non-specific binding for sybodies RB2, RB6, RB7, RB11, and RB16, as indicated by their binding to both native and knock-down cells. In contrast, RB3 and RB15 exhibited specificity by staining only the native RUSF1-expressing cells. Sybody RB9, however, did not yield detectable signals under the conditions tested (4 µM sybody concentration).

Subsequently, these same eight sybodies were also evaluated for their ability to detect purified RUSF1 protein through Western blotting (Figure 72). Purified RUSF1 (0.5 µg) was separated by acrylamide gel electrophoresis and transferred onto a PVDF membrane, which was then incubated with each sybody individually. Detection involved anti-Myc antibodies, secondary antibodies, and chemiluminescent exposure.

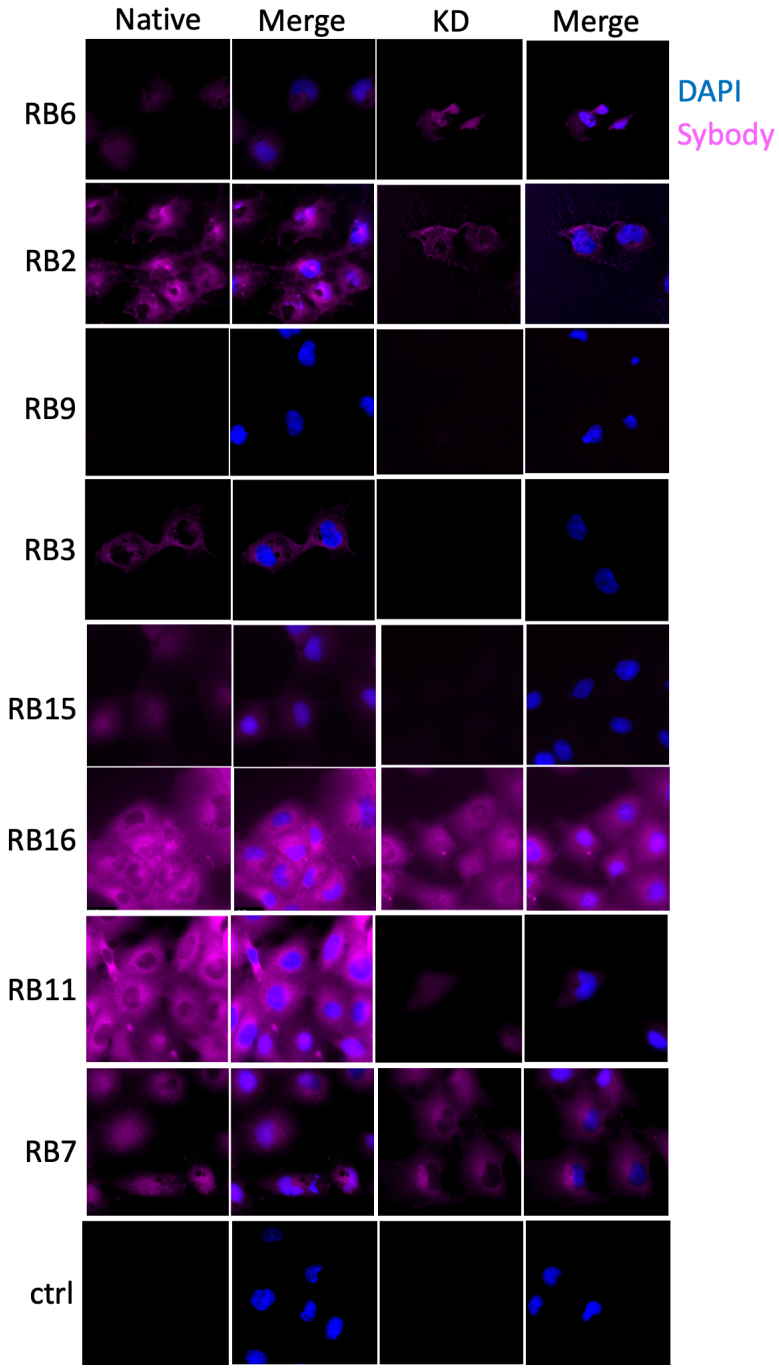


Figure 71. Utilising RUSF1 Sybodies in Immunofluorescence

Representative immunofluorescence images of cells stained with selected anti-RUSF1 sybodies (magenta) and DAPI (blue, nuclei). "Native" panels show staining in untreated cells, while "KD" panels show cells after RUSF1 knockdown to assess specificity. RB3 and RB15 showed specific staining of RUSF1. Experiment done by Martyna Biadun.

Consistent with microscopy results, RB6 displayed non-specific binding to unrelated proteins, including molecular weight markers. Sybodies RB2, RB7, RB9, RB11, and RB16 did not detect the purified RUSF1 under these conditions. However, RB3 and RB15, which recognised RUSF1 specifically in microscopy, also specifically bound to RUSF1 on the membrane, suggesting these sybodies potentially interact with epitopes resilient enough to denaturation by SDS sample buffer. Interestingly, RB11, which recognized RUSF1 specifically in cell imaging, failed to bind RUSF1 under the denaturing conditions of Western blotting, indicating that its epitope likely unfolded with the addition of SDS buffer.

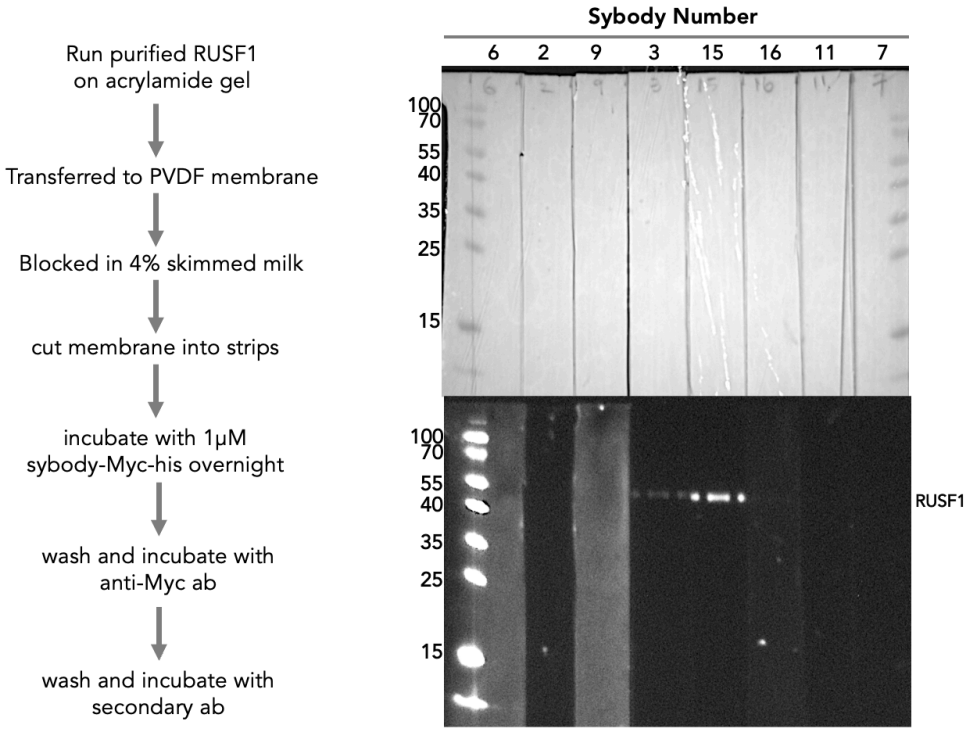


Figure 72. Utilising RUSF1 Sybodies in Western Blot

Left: workflow of using sybodies in western blot detection of purified RUSF1. Right: results of western blotting; each strip was incubated with a different sybody (number at the top of the strip, "6" indicated incubation with sybody RB6).

Taken together, these data demonstrate that although some sybodies such as RB2 and RB6 initially appeared promising due to their high affinity in pull-down and ELISA assays, they

ultimately lacked specificity to be used in techniques such as fluorescence microscopy, Western blotting (to quantify the amount of endogenous RUSF1 under various cellular conditions), or co-immunoprecipitation of endogenous binding partners of RUSF1.

Detailed sequence analysis of the sybodies could offer insights into their binding specificity, or the lack thereof. RB2 and RB6, both characterized by abundant hydrophobic and aromatic residues in their CDR1 and CDR3 loops, demonstrated strong yet non-specific interactions, potentially mediated by interactions with detergent molecules or unfolded proteins (Table 5).

Similarly, sybodies RB4, RB5, RB7, RB8, RB13, and RB17 featured consecutive aromatic residues within their CDR3 loops, likely contributing to their non-specific binding behaviour. Despite having alternating charged residues within CDR1 loops that might theoretically facilitate specific interactions with charged regions of RUSF1, these sybodies did not demonstrate sufficient specificity or affinity for practical use, RB7 in specific was proved to be non-specific in imaging.

In contrast, RB3 and RB15 emerged as specifically binding candidates with distinct sequence features.

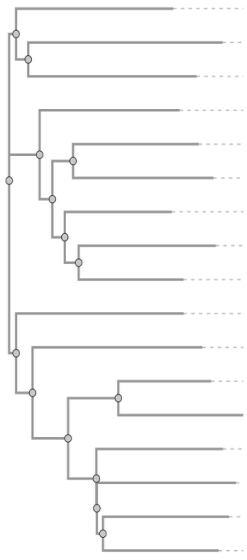
RB3 uniquely contained two cysteine residues at each end of its CDR3 loop, potentially enabling intramolecular disulfide bond formation and stabilizing loop conformations, as indicated by double bands in purification and pull-down assays. Interestingly, cysteine residues were generally excluded when the randomised synthetic nanobody library was generated (Zimmermann et al. 2018), so this specific nanobody probably arose from further mutagenesis.

RB15, alongside RB9 and RB16, possessed two tryptophan residues at the equivalent positions, potentially stabilizing CDR3 loops via pi-pi stacking interactions. Stabilisation of CDR3 loop might be beneficial in increasing binding affinity and specificity. However, other sybodies with

similar residues in CDR3, such as RB1 and RB14, failed to pull down RUSF1, suggesting cooperativity from other loops.

Table 5. Sequence Analysis of RUSF1 Sybodies

Phylogenetic tree and CDR sequences of RUSF1-binding sybodies (RB1–RB17). Functional properties are summarized on the right, including performance in pulldown assays, affinity ELISA assay results, immunofluorescence imaging, and western blotting.



Name	CDR1	CDR2	CDR3	Pulldown	ELISA	Imaging	Western Blot
RB10	YANRMA	QSSGYSTW	HGYMYQY	yes	N/A	N/A	N/A
RB6	FITKMA	WSFGHWTT	NGIFAGY	yes	yes	non specific	non specific
RB2	YMLRMR	SSWGEWTH	SGYGAAK	yes	yes	non specific	no signal
RB8	DNKKMR	QSTGNNTY	LGYWYYK	no	N/A	N/A	N/A
RB5	EKFRMY	YSAGFETI	LGYWYFH	yes	no	N/A	N/A
RB13	ENHEMY	YSEGFEKR	VGIWYYR	yes	no	N/A	N/A
RB7	ERYHMA	QSSGSQTH	LGIWYYR	yes	no	non specific	no signal
RB17	ESRQMA	YSQGANTF	AGYWYYR	yes	N/A	N/A	N/A
RB4	EAHHMT	YSNGSSTA	IGYWYYR	yes	no	N/A	N/A
RB11	DK-HMT	ESEGHYTW	KGSNKSR	yes	N/A	specific	no signal
RB12	YQYYMH	QSDGQTTH	VGTYNWN	yes	N/A	N/A	N/A
RB1	SWYNMY	LSTGAGTE	WGSYMNW	no	no	N/A	N/A
RB3	SRYHMY	SSTGDGTE	CGSYLNC	yes	yes	specific	signal
RB9	KYMSMH	YSYGISTE	WGGYMHW	yes	yes	no signal	no signal
RB15	AYMNMY	YSTGTQTY	WGEYMSW	yes	yes	specific	signal
RB14	NAYSMY	YSTGSHTK	WGQYMHW	no	N/A	N/A	N/A
RB16	NHMSMY	QSVGNITE	WGAYMLW	no	yes	non specific	no signal

Looking at CDR1, RB3 and RB9 were distinct from RB15, with two positively charged residues, RB15 was mostly hydrophobic/aromatic. However, in CDR2, both RB9 and RB15 had multiple aromatic tyrosine, while RB3 had a poplar/negatively charged CDR2. Since RB9 seemed to be less specific to RUSF1 than RB3 and RB15, while the latter behaved similarly across all the analyses, one could speculate that the hydrophobic CDR1 (unique to RB15 among the three), as well as the polar/negative CDR2 (unique to RB3 among the three) were beneficial features facilitating specific binding to RUSF1. This might provide guidance for future engineering aimed at achieving enhanced affinity and specificity.

6.5 Discussions

In the preceding chapters (Chapter 5, 6), I presented structural and biochemical studies of human RUSF1, including structural characterisation, analysis of monomer–dimer dynamics, and investigation of its interaction with the putative client protein GABRA1.

I also explored functional aspects of RUSF1, identifying additional candidate clients or co-factors (e.g., TMEM126A) and generating synthetic nanobodies against RUSF1 to perform both immunofluorescence and western blot analyses. In this chapter, I will attempt to integrate these findings into a broader cellular and physiological context, proposing potential roles for RUSF1 in endoplasmic reticulum homeostasis and reflecting on the significance of its conformational flexibility.

RUSF1 was initially identified as the top candidate in a mass-spectrometry screen for chaperones of the GABAA receptor subunit alpha1 (Hooda *et al*., 2024, preprint). While biochemical assays confirmed that RUSF1 interacts with GABRA1, the interaction was considerably weaker compared with that of NACHO, a known and robust GABRA1 chaperone (Chapter 6.1). This difference in binding affinity suggests that RUSF1 is unlikely to function exclusively as a dedicated chaperone for GABRA1. Instead, its role appears to be broader, encompassing additional clients and potentially extending beyond chaperoning.

Immunofluorescence studies done by Martyna Biadun in the lab (unpublished) further support the idea that RUSF1 functions as a general ER-resident protein. Under over-expressed conditions in COS-7 cells, RUSF1-GFP colocalises extensively with both CLIMP63, a marker of rough ER sheets, and reticulon, a marker of tubular ER (Figure 73 A). Using synthetic nanobody RB3 to stain RUSF1 under endogenous condition, colocalization with CLIMP63 and reticulon was also observed, indicating that RUSF1 is widely distributed across distinct ER

subdomains (Figure 73 B). Such a pattern implies that RUSF1 may contribute to fundamental aspects of ER homeostasis rather than serving a highly specialized role.

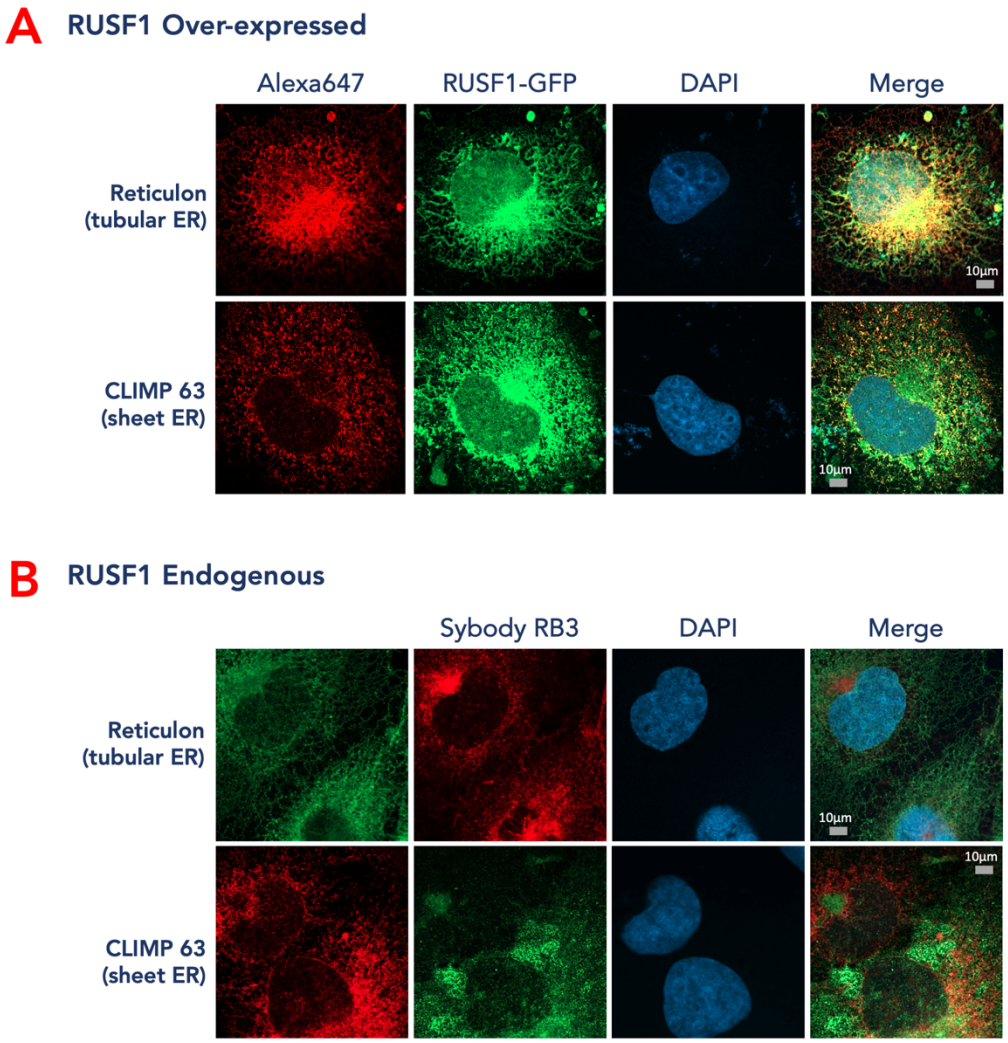


Figure 73. Localisation of RUSF1 in COS-7 Cells

A. Overexpressed RUSF1-GFP co-localizes with ER markers. Immunofluorescence staining using Alexa647-conjugated antibodies against reticulon (tubular ER marker, top row) and CLIMP-63 (sheet ER marker, bottom row) was performed in COS-7 cells overexpressing RUSF1-GFP. Merged images show RUSF1 signal strongly overlapping with both ER markers, indicating ER localization. Nuclei were stained with DAPI (blue). B. Endogenous RUSF1 detected using sybody RB3. COS-7 cells were stained with sybody RB3, and co-localisation was tested against reticulon, or CLIMP-63. Merged images confirm endogenous RUSF1 localization within ER membranes, showing overlap with both tubular and sheet ER markers. Microscopy done by Martyna Biadun.

Loss-of-function studies further underscore RUSF1's significance in ER maintenance. Knockout of RUSF1 in COS-7 cells resulted in striking morphological defects, characterized by the formation of prominent "holes" within the ER, as visualized by CLIMP63 staining

(Figure 74). Interestingly, these disruptions were not apparent under reticulon staining, suggesting that the phenotype was specific to ER sheets. Expression of the monomeric K230G mutant of RUSF1 in COS-7 cells also produced a similar phenotype, although the holes were smaller and less abundant. The concordance of these results indicates that RUSF1 integrity and its oligomeric state are essential for maintaining ER structure. Given that COS-7 cells do not endogenously express GABA receptors, these findings argue strongly for a general role of RUSF1 in ER quality control and homeostasis, rather than client-specific stabilization.

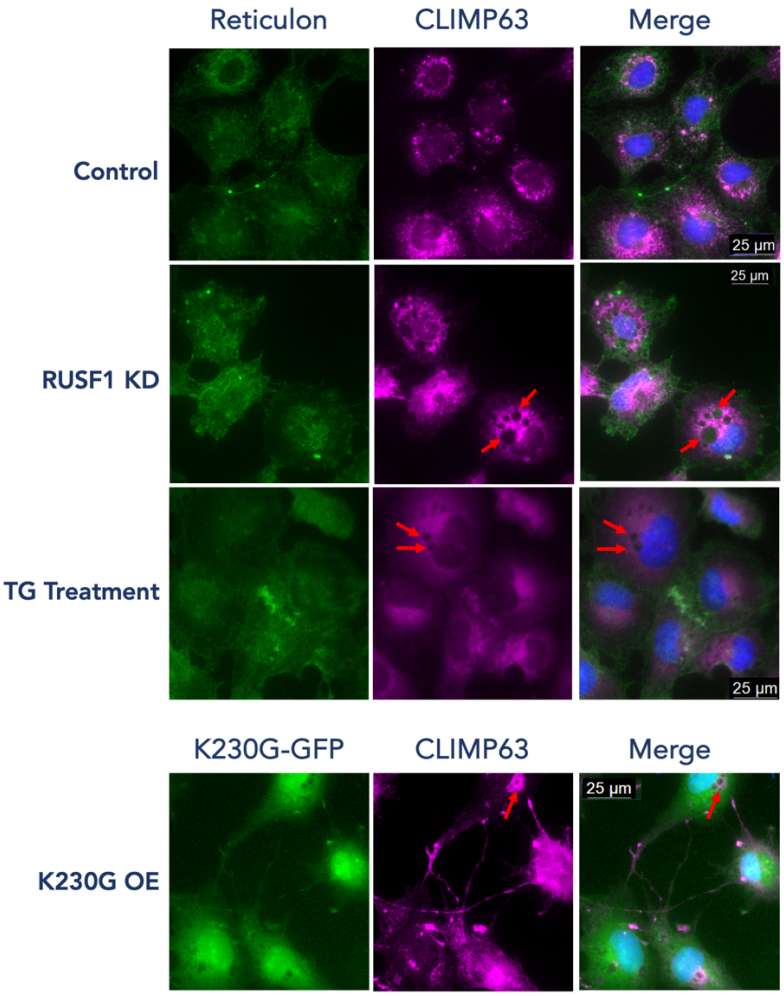


Figure 74. Phenotype of RUSF1 KD COS-7 Cells Resembling ER-Stress Conditions

Top panels: representative immunofluorescence images of COS-7 cells stained for reticulon (green, tubular ER marker) and CLIMP-63 (magenta, sheet ER marker). Nuclei are stained with DAPI (blue). Knock-down of RUSF1 or Thapsigargin treatment (1 µM overnight prior to fixing) led to hole formation in the ER. Bottom panel: COS-7 cells over-expressing RUSF1-K230G-GFP and stained with CLIMP-63 (magenta, sheet ER marker). Over-expression of K230G mutant also lead to formation of holes. Microscopy done by Martyna Biadun.

The nature of the “holes” observed in RUSF1-deficient cells raises intriguing parallels with ER stress-induced remodelling (Xu *et al*., 2021). Notably, treatment of COS-7 cells with thapsigargin, a SERCA (sarco/endoplasmic reticulum Ca^{2+}-ATPase) inhibitor that induces ER stress, produced a similar phenotype under CLIMP63 staining (Figure 74).

Thapsigargin (TG) treatment of eukaryotic cells usually leads to depletion of ER calcium levels, which disrupts protein folding, activates all branches of the unfolded protein response (PERK, IRE1, ATF6), and drives downstream outcomes ranging from adaptation to apoptosis via calcium dysregulation, oxidative stress, and mitochondrial crosstalk. Previous work by Xu *et al*. (2021) demonstrated that thapsigargin also leads to remodelling of ER morphology, triggering the formation of multilayered ER whorls in normal rat kidney (NRK) cells and HEK293T cells (Figure 75 A, B). These structures were thought to emerge via COPII-dependent budding and requiring PERK activity. These whorls sequester Sec61 and concentrate PERK, thereby achieving translational inhibition and mitigating ER protein load. While the structural features of the thapsigargin-induced whorls differ somewhat from the “holes” observed upon RUSF1 loss, the resemblance suggests that the latter may also reflect an aberrant stress adaptation response.

Determining whether RUSF1-associated holes correspond to *bona fide* ER whorls would be an important next step in understanding the cellular function of RUSF1. Co-localisation experiments with Sec61, COPII components, or PERK could clarify whether these structures represent ER stress-induced remodelling or a different morphological phenomenon. If confirmed as stress-derived whorls, this would suggest that RUSF1 functions in detecting or resolving protein misfolding stress, potentially acting as a quality control factor, mediator of misfolded protein extraction, or sensor linking protein misfolding to downstream UPR signalling.

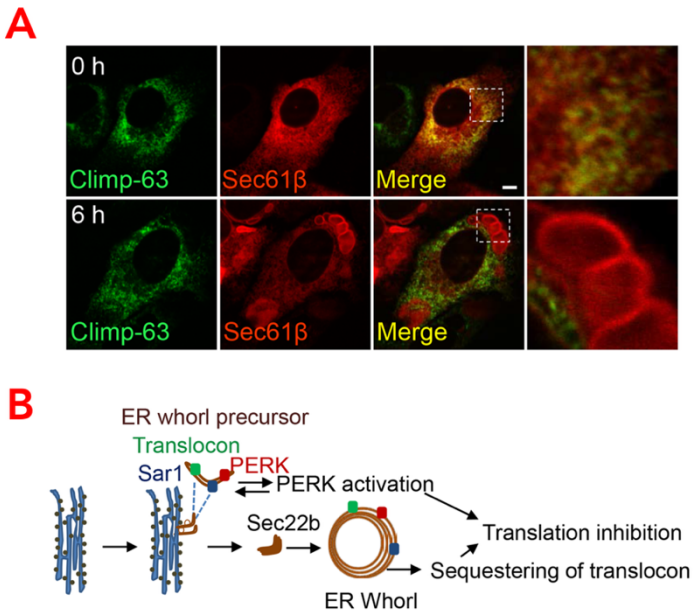


Figure 75. ER Whorl Formation Under Thapsigargin Treatment

A. Immunofluorescence of COS-7 cells treated with thapsigargin (TG) for 0 h or 6 h. Cells were stained for CLIMP-63 (green, ER sheets) and Sec61β (red, translocon). After 6 h TG treatment, Sec61β redistributes and co-localizes with CLIMP-63 into compact circular structures resembling ER whorls (right panels show magnified views). Adapted from Xu et al., 2021. B. Schematic model of ER whorl formation. ER whorls are proposed to arise from precursor membranes containing translocon and ER-shaping proteins (e.g., Sar1, Sec22b, CLIMP-63). Upon TG treatment, PERK is activated, leading to translation inhibition and sequestration of the translocon into ER whorls. This process reflects a stress-induced remodelling mechanism that alters ER architecture.

Structural studies of RUSF1 also provided additional clues to its potential function. RUSF1 harbours a large, charged cavity reminiscent of ER quality control proteins such as ATP13A1, which removes mis-inserted mitochondrial proteins (McKenna *et al.*, 2020), and Derlin-1, which mediates retro-translocation of misfolded membrane proteins (Rao *et al.*, 2021). The conformational dynamics of RUSF1 dimers may permit transient exposure of this cavity, enabling interactions with unfolded helices or partially misfolded membrane proteins. Although peripheral charged surfaces of the dimer could, in principle, mediate stable client interactions, our co-immunoprecipitation and mass spectrometry analyses failed to identify high-affinity binding partners. This lack of stable interactions suggests that RUSF1 does not operate as a classical chaperone, but is more likely a dynamic quality control factor.

The question then arises: how are RUSF1's conformational states regulated? In the CryoEM structure of RUSF1 dimer, the complex adopted a state reminiscent of an "outward open" conformation of an MFS transporter, with a large luminal-facing cavity. However, multiple AlphaFold predictions, especially that of RUSF1 dimer in complex of TMEM126A, a interactor identified through mass spectroscopy, revealed a different state similar to an "inward-open" conformation, where the luminal transmembrane interface closes, while a cavity opens facing the cytosol. The conformational change reduced the buried dimer interface from 3100 Å^2 to 2700 Å^2. More importantly, the shift disengaged residue K230.

Experimentally, it has been shown that K230G mutation destabilises the dimer, producing monomeric, thermally unstable species. However, research on *Arabidopsis* RUS1 and RUS2 suggested that monomer-dimer transitions was vital for proper cellular function (Leasure *et al*., 2009). Therefore, if the AlphaFold predicted "inward open" state is physiologically relevant, it implies that under certain physiological stimuli, RUSF1 can transition from "outward open" to "inward open", then transiently sample monomeric states before dimerizing again (Figure 76 A).

Several mechanisms may drive this transition. Firstly, client or co-factor binding at the periphery of the dimer could induce conformational changes leading to monomerization. Secondly, post-translational modifications may play a role: mass spectrometry identified 14-3-3 epsilon as a potential RUSF1 interactor, predicted to bind the C-terminal tail near serine 462. Preliminary proteomic evidence verified phosphorylation at this residue in overexpression systems (Supplementary Figure 12), though further studies are required to confirm its occurrence endogenously, and establish whether 14-3-3 epsilon mediates this modification.

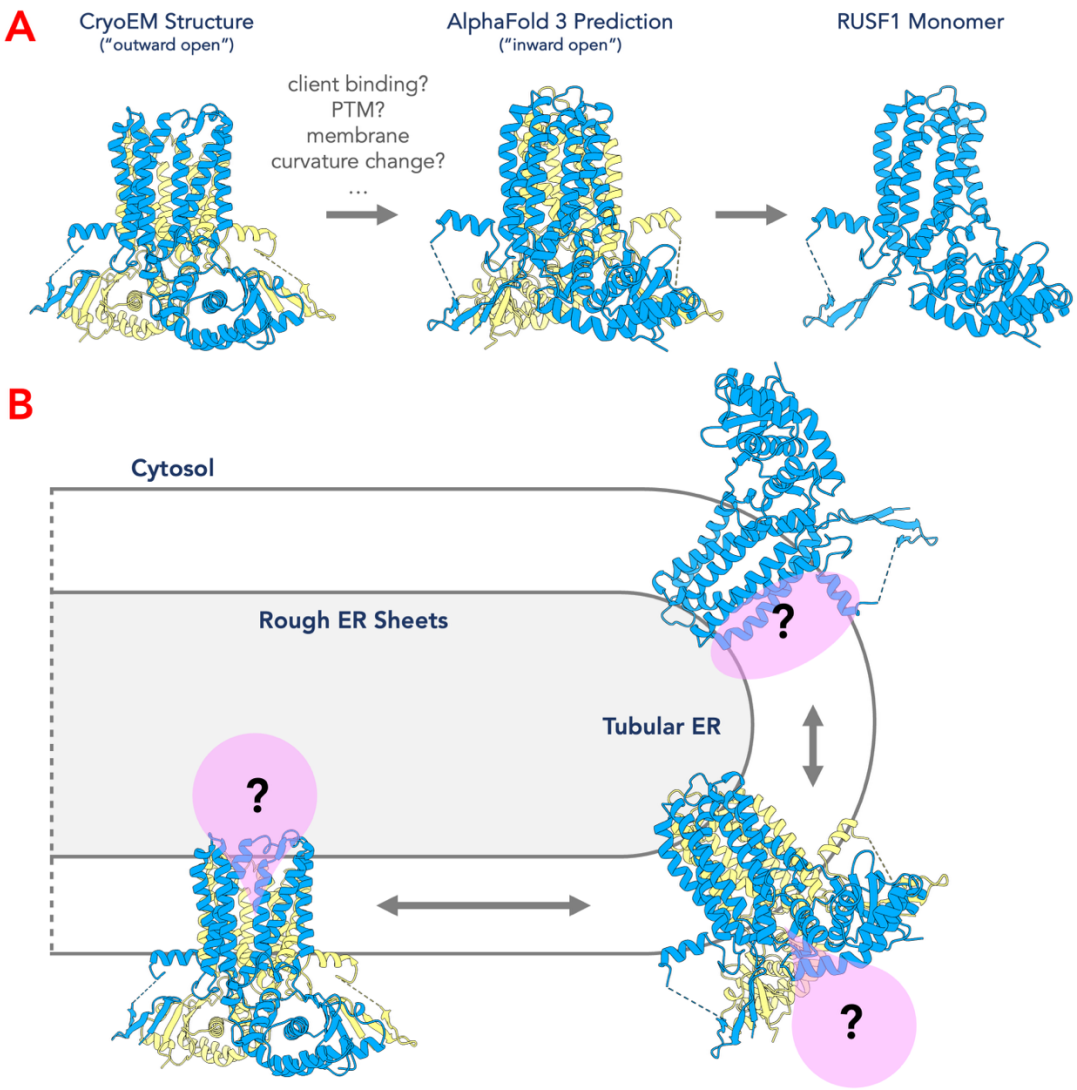


Figure 76. Proposed Mechanisms of Function of RUSF1 in Cells

A. Comparison of the cryo-EM structure ("outward open"), AlphaFold 3 prediction ("inward open"), and the monomeric state of RUSF1, suggesting that RUSF1 may undergo conformational changes (from various factors) which could drive monomerization. B. Hypothetical cellular roles of RUSF1. In the ER, RUSF1 may localize to both rough ER sheets and tubular ER membranes. Its conformational plasticity could enable interactions with unknown partners, trafficking clients between sub-compartments of the ER.

Finally, membrane curvature may regulate oligomeric state (Figure 76 B). The short amphipathic helix preceding the first transmembrane segment of RUSF1 appeared perpendicular to the transmembrane domain in the CryoEM structure, and underwent a shift of ~9° tilted in AlphaFold "inward open" models, suggesting that changes in bilayer tension might be able promote conformational rearrangements of this helix and drive global conformational

changes. This is especially relevant, since functionalities in the ER are highly compartmentalised, and RUSF1 might adopt distinct conformational states, e.g., dimeric in the flat ER sheets, and monomeric at the curved tubular ER, to carry out its intended function.

Taken together, our findings point toward a model in which RUSF1 is an important ER-resident protein that contributes to protein quality control and ER homeostasis. Its role may involve client stabilization, extraction of misfolded proteins, or initiation of stress responses, potentially integrating several of these functions. Like other ER-resident complexes such as EMC, RUSF1 may perform multiple activities within the broader process of membrane protein biogenesis and surveillance. Central to these roles is its conformational plasticity, mediated by monomer–dimer transitions and modulated potentially by client interactions, post-translational modifications, or membrane environment.

Future experiments will be required to resolve the precise role of RUSF1. Structural studies of RUSF1 reconstituted into liposomes could directly test the influence of membrane curvature on conformational states. Similarly, immunofluorescence experiments probing the co-localisation of ER-stress markers with the holes observed in knockout cells could confirm whether these structures correspond to ER whorls. Ultimately, a combination of biochemical, structural, and cell-based studies will be necessary to define how RUSF1 contributes to ER protein homeostasis and to elucidate the physiological significance of its conformational dynamics.

7. Conclusions

This thesis aimed to advance the understanding of how membrane proteins contribute to the maintenance of homeostasis within distinct organellar compartments. To address this, two proteins with divergent functions were investigated: the lysosomal chloride/proton antiporter CLC7 and the putative endoplasmic reticulum (ER) chaperone RUSF1.

The first part of the thesis focused on the regulation of CLC7 by the lipid PI(3,5)P2. Structural analyses of CLC7 in complex with PI(3,5)P2, as well as a hyperactivating mutant (Y715C), revealed that PI(3,5)P2 acts as an inhibitor by stabilizing interactions between the transmembrane and soluble CBS domains. In the absence of PI(3,5)P2, the CBS domains dissociate from the transporter, leading to a previously uncharacterized conformational rearrangement of helix R. This movement exposes a segment of the chloride/proton transport pathway, including several titratable residues essential for exchange activity (e.g., E313 and E314), while simultaneously removing Y602 as a steric barrier for chloride ions. Functionally, this conformational shift likely represents an activated state of the transporter. Future studies, such as single-molecule FRET experiments, with a pair of residues labelled to measure the distance between the C-terminus of R helix and the rest of the transmembrane domain, could provide insights into the proportion of transporters adopting these distinct conformations under varying pH or chloride conditions. Additionally, mutational disruption of helix R stability may clarify its role in transporter activation. To further characterize CLC7 in its native context, I also generated and validated multiple nanobodies targeting CLC7, which can be applied in immunofluorescence and proteomics experiments. Collectively, these findings enhance the mechanistic understanding of how phosphoinositides fine-tune the activity of lysosomal membrane proteins to maintain ion and pH homeostasis within intracellular vesicles; and how the disease mutant Y715C, through disrupting the intricate interaction between the lipid and

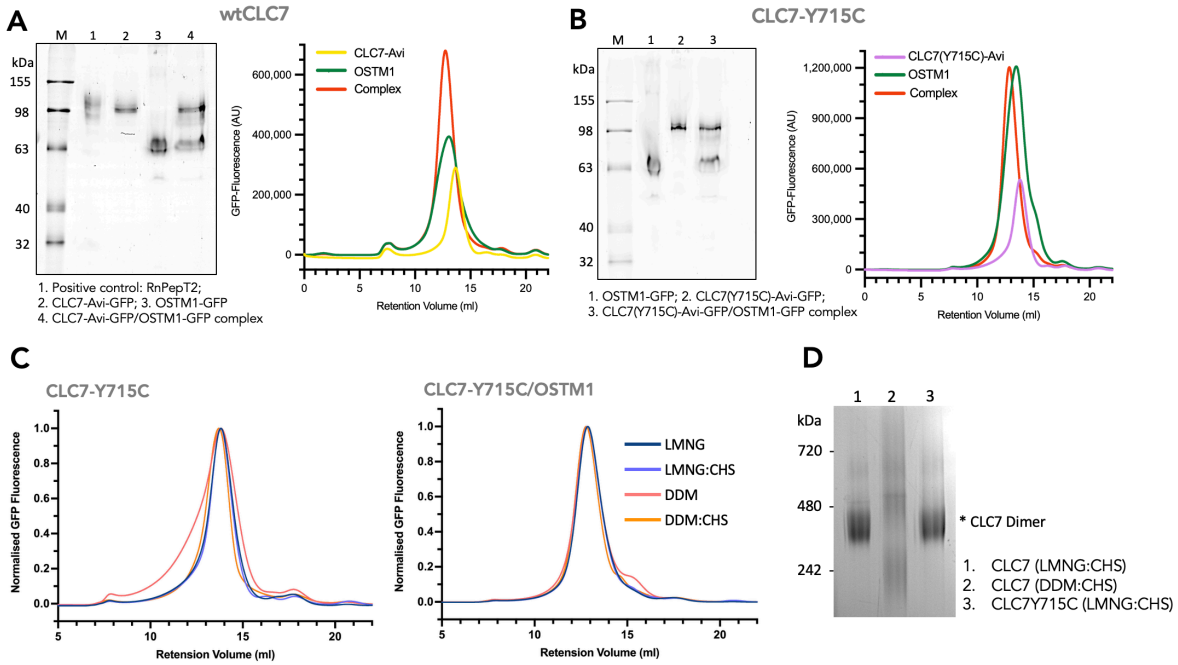
CLC7, can cause such a drastic structural rearrangement, resulting in significant hyper-activating phenotypes.

The second part of the thesis investigated the function of the putative ER chaperone RUSF1, which exhibits transporter-like architecture. Structural studies revealed the first structure of this protein family, at 3.1 Å, showing RUSF1 as a homodimer with several notable features: a hydrophilic cavity oriented toward the ER lumen, an amphipathic helix perpendicular to the transmembrane domain, and extensive interactions of the soluble domains with the detergent micelle. Protein–protein interaction studies confirmed RUSF1 binding to two potential client or partner proteins, GABRA1 and TMEM126A, and additional candidates such as 14-3-3 epsilon were identified via co-immunoprecipitation of overexpressed RUSF1. These findings provided a foundation for future studies, suggesting several potential RUSF client proteins to be tested with pulldowns and co-localisation assays. These data also provided a hypothesis that RUSF1 activity is regulated by monomer-dimer transitions, which could be tested once client proteins are discovered, and an activity assay established. In addition, I developed specific nanobodies against RUSF1 that enabled the detection of endogenous protein by immunofluorescence *in vivo*. These nanobodies will facilitate further investigations into the role of RUSF1 in ER stress, the unfolded protein response (UPR), ER-associated degradation (ERAD), and ultimately ER proteostasis.

Appendix


CLC7/1-805
CLC6/1-869
CLC5/1-816
CLC4/1-760
CLC3/1-818
CLC2/1-898
CLC1/1-988
CLCKa/1-687
CLCKb/1-687

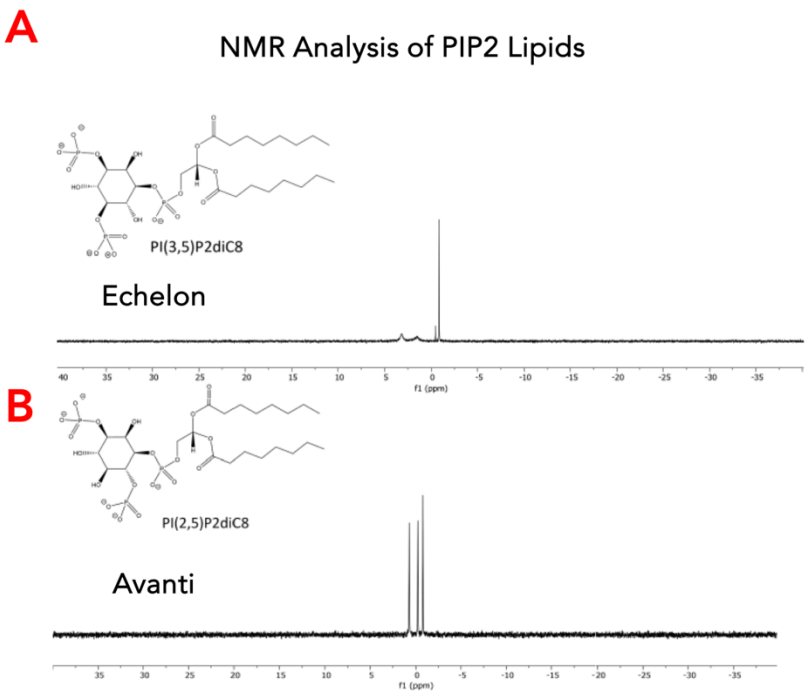
Supplementary Figure 1. Sequence alignment of human CLC channels and transporters.



Supplementary Figure 2. Test expression of various CLC7 and OSTM1 GFP-tagged constructs. A. Left: In-gel fluorescence of 1. Positive control RnPepT2-GFP; 2. CLC7-Avi-GFP; 3. OSTM1-GFP; and 4. Co-expression of CLC7-Avi-GFP and OSTM1-GFP. CLC7 appears as a band at ~98 kDa, and OSTM1 appears as a band at ~63 kDa. Right: FSEC chromatogram of the samples shown on the left, color keys included, note that complex exhibited lower retention volume than individual components. B. Left: In-gel fluorescence of 1. OSTM1-GFP; 2. CLC7(Y715C)-Avi-GFP; 3. Co-expression of CLC7(Y715C)-Avi-GFP and OSTM1-GFP. FSEC chromatogram of the samples shown on the left, color keys included. C. FSEC chromatogram of CLC7-Y715C (left) and CLC7-Y715C/OSTM1 complex (right) solubilised in different detergents (colour keys shown). D. Blue Native PAGE of various purified CLC7 samples. CLC7 (wild type and Y715C) in LMNG:CHS appeared as one condensed band, whereas wild-type CLC7 purified in DDM:CHS appeared spread out, indicating protein degradation or unfolding.

	wtCLC7/OSTM1 (EMD-51098) (PDB: 9G6D)	CLC7-Y715C/OSTM1 (EMD-51099) (PDB: 9G6E)	CLC7/OSTM1+PIP2 (EMD-51097) (PDB: 9G6C)
Microscope	Krios G4	Krios G4	Krios G4
Detector	Falcon 4	Falcon 4	Falcon 4
Slit width (eV)	10	10	10
Total Dose ($e^{-}Å^{-2}$)	55.7	56.5	55.7
Defocus range (μm)	-2.0 to -0.5	-2.5 to -0.5	-2.5 to -0.5
Pixel size ($Åpx^{-1}$)	0.693	0.693	0.693
Micrographs	10,004	10,285	14,828
Picking software	SIMPLE	SIMPLE	SIMPLE
Particles (initial)	2,455,968	2,723,428	4,258,989
Particles (final)	210,876	336,442	833,576
Symmetry	C2	C2	C2
Resolution (Å)	2.73	2.56	1.84
FSC cut-off	0.143	0.143	0.143
Map-sharpening B-factor ($Å^2$)	78.9	83.5	46.0
Mask CC	0.82	0.87	0.88
Volume CC	0.82	0.86	0.87
Mean Protein B-factor ($Å^2$)	126.36	94.14	71.22
Mean Ligand B-factor ($Å^2$)	162.03	127.79	87.78
R.M.S deviations			
Bond lengths (Å)	0.003	0.004	0.002
Bond angles (°)	0.512	0.567	0.478
MolProbity score	1.77	2.27	1.4
Clash score	7.26	10.39	6.44
Poor rotamers (%)	2.92	4.48	1.16
Ramachandran Plot			
Favoured (%)	97.94	96.41	98.03
Allowed (%)	2.06	3.59	1.97
Disallowed (%)	0.00	0.00	0.00

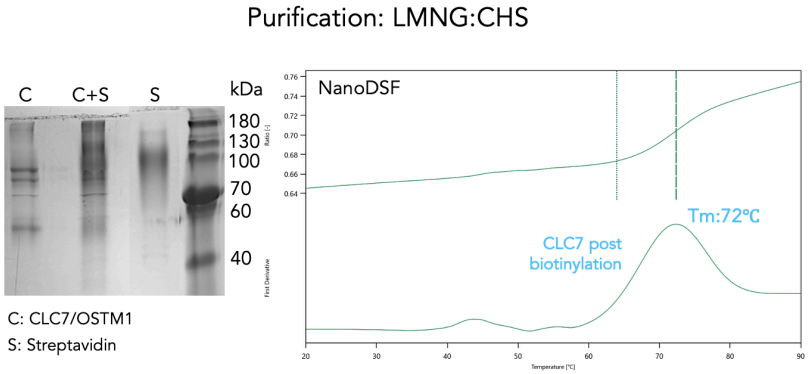
Supplementary Table 1. CryoEM data collection, refinement and validation statistics for CLC7/OSTM1 structures



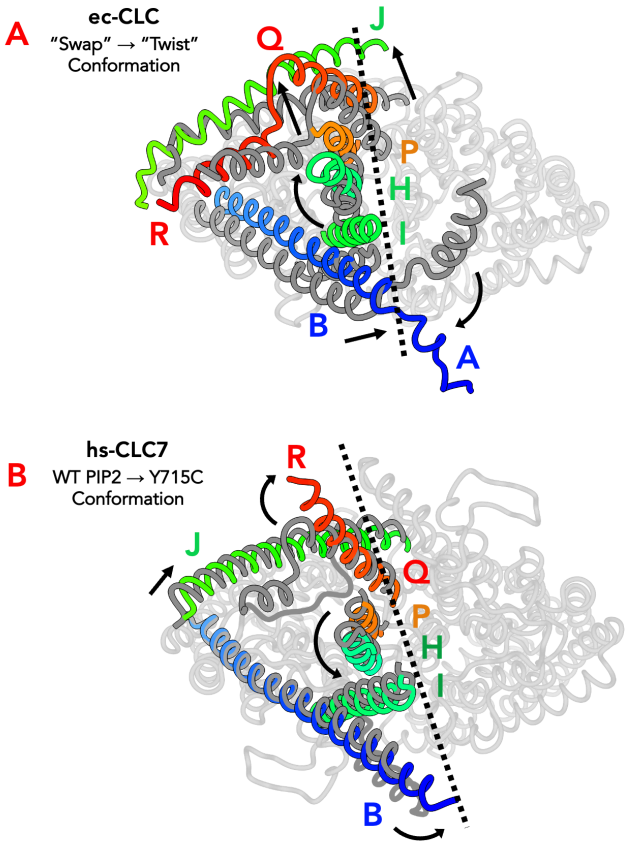
Supplementary Figure 3. 31P NMR of lipid samples from each supplier. Top spectrum is from Echelon; note the single prominent peak consistent with nearly equivalent environments for each phosphate, equally spaced around the inositol ring, i.e. lipid being PI(3,5)P. Bottom spectrum is PI(3,5)P2 sourced from Avanti Polar Lipids, where three peaks were present, suggesting 3 different phosphate environments. This is incompatible with PI(3,5)P2 but is consistent with PI(2,5)P2 as suggested by the structure.

Y715C		Helix No.	WT	
residue no.	a.a.		residue no.	a.a.
113	ARG	Helix B		
116	HIS			
119	PHE			
120	ARG			
123	GLU			
124	ILE			
127	TRP		127	TRP
128	VAL			
131	ALA			
305	VAL	Helix H	305	VAL
309	LEU		309	LEU
			312	LEU
313	GLU		313	GLU
			316	ALA
317	SER		317	SER
			318	PHE
319	TRP		319	TRP
321	GLN	Helix I	321	GLN
322	PHE			
324	THR			
325	TRP		325	TRP
328	PHE		328	PHE
329	PHE		329	PHE
332	MET		332	MET
			333	ILE
336	PHE		336	PHE
340	PHE		340	PHE
349	MET	I-J Loop	349	MET
350	TRP		350	TRP
351	ASP		351	ASP
352	LEU		352	LEU
			355	PRO
362	ARG		362	ARG
			364	ASP
365	SER			
			366	GLU
			369	ALA
370	TYR		370	TYR
371	THR		371	THR
372	ILE		372	ILE
373	HIS	Helix J		
375	ILE			
		Helix N	540	ILE
562	MET	Helix P		
563	THR		563	THR
564	LEU		564	LEU
567	THR			
568	VAL		568	VAL
571	MET		571	MET
572	GLU		572	GLU
575	SER		575	SER
576	ASN			
			577	VAL
			578	THR
579	TYR	Helix Q		
580	GLY			
581	PHE		581	PHE
584	MET		584	MET
588	MET		588	MET
591	LYS	Helix R		
599	GLU			
602	TYR			
603	ASP			
605	HIS			

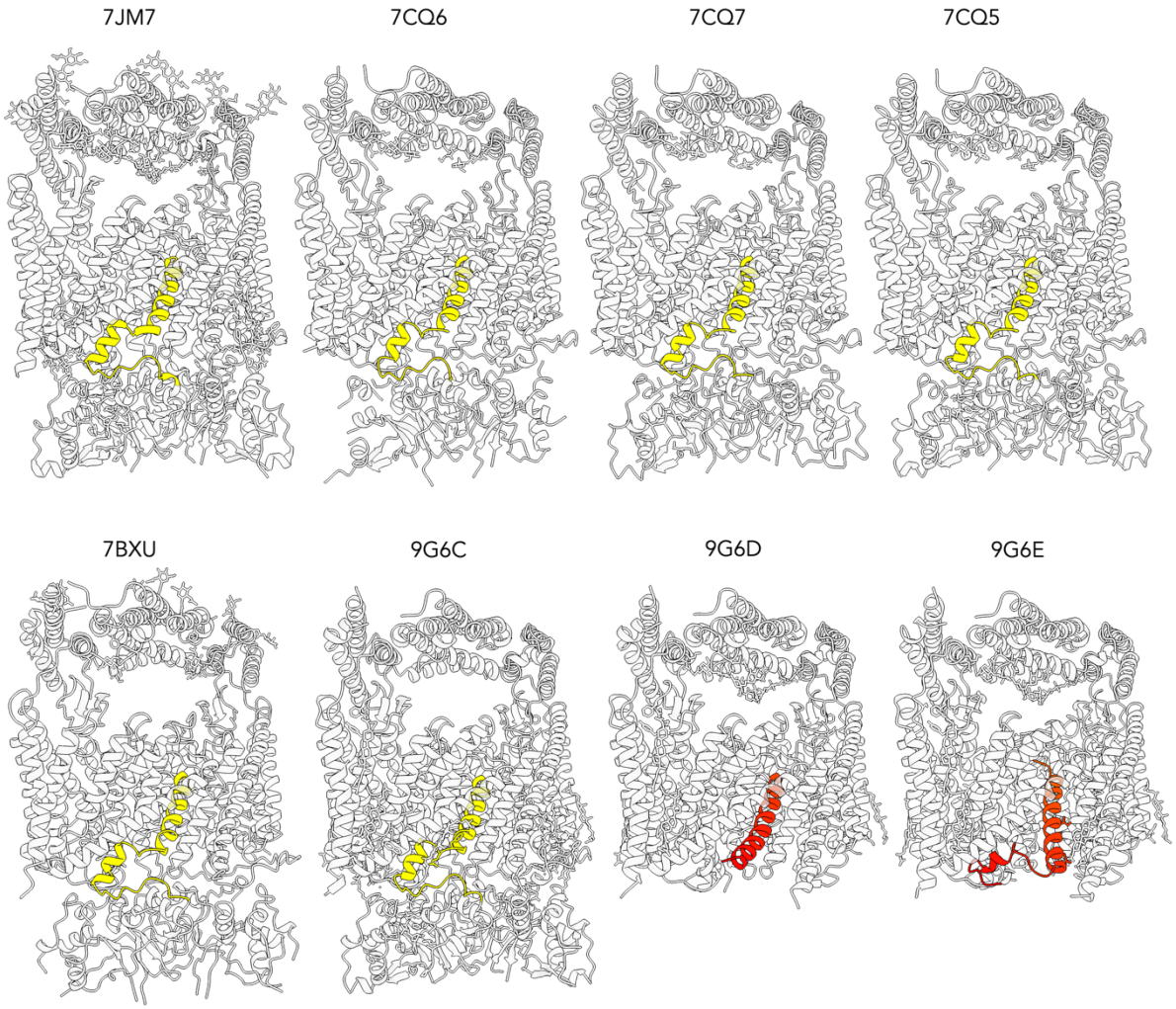
Supplementary Table 2. Residues involved in the dimer interface of CLC7 dimer interface in WT PIP2 structure and Y715C structure. Number of residue, type of amino acid, and location on transmembrane helix annotated.



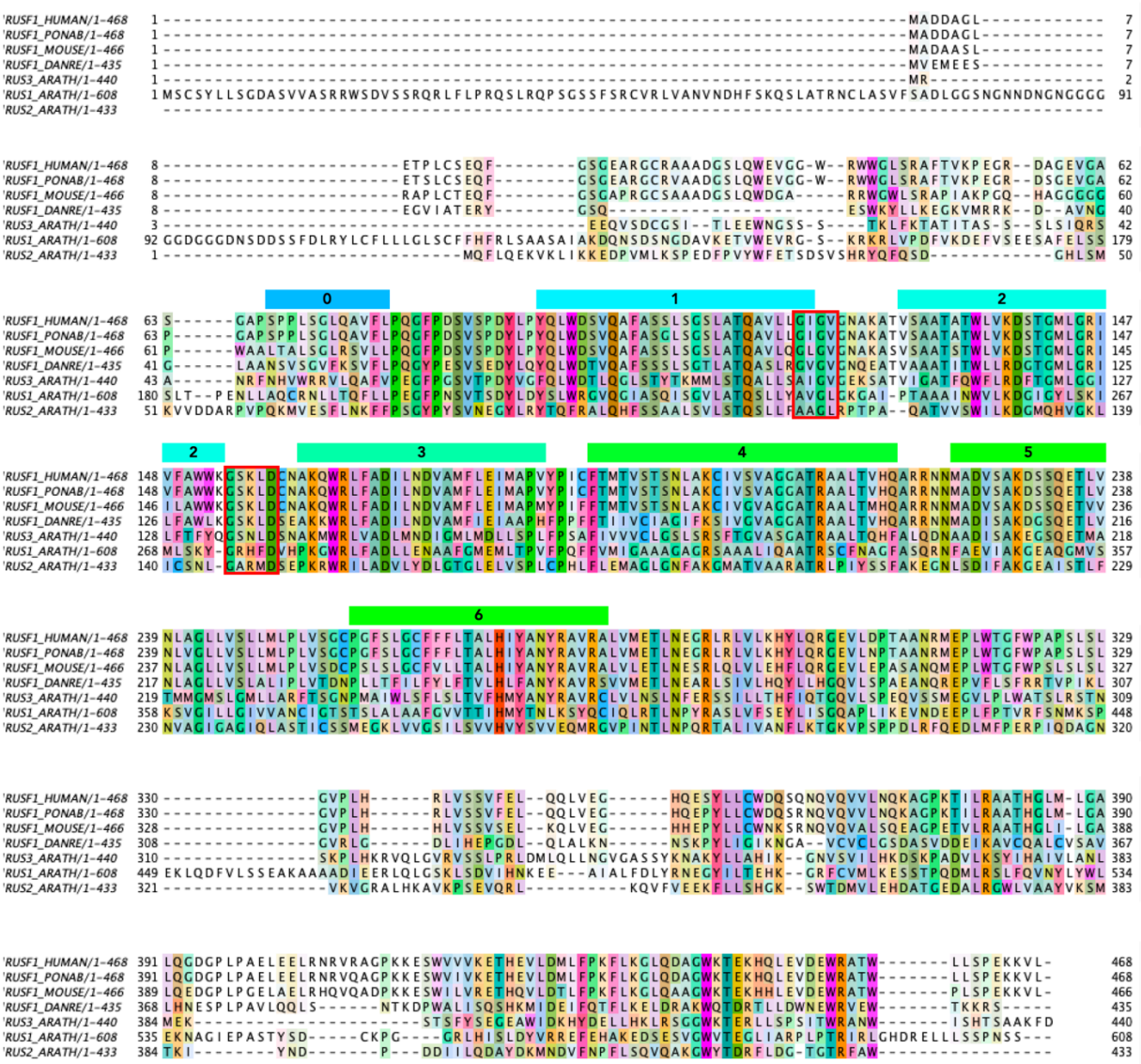
Supplementary Figure 4. Purification of CLC7/OSTM1 complex as well as OSTM1 for Sybody Selection. Top left: purified CLC7/OSTM1 complex (C), Streptavidin (S), and the complex (C+S), run on acrylamide gel. Addition of streptavidin to CLC7 led to an upshift of bands, indicating successful biotinylation. Top right: NanoDSF analysis of CLC7 post-biotinylation, with a Tm of ~72 degree Celsius, confirming protein stability. Bottom: SDS-PAGE analysis of purified OSTM1. "C" indicating CLC7/OSTM1 complex, "O" indicating OSTM1 only. OSTM1 appeared as a band around 40-70 kDa, note that in sample of the complex, additional bands at ~100 kDa correspond to CLC7, which are mostly absent from the OSTM1 only samples.



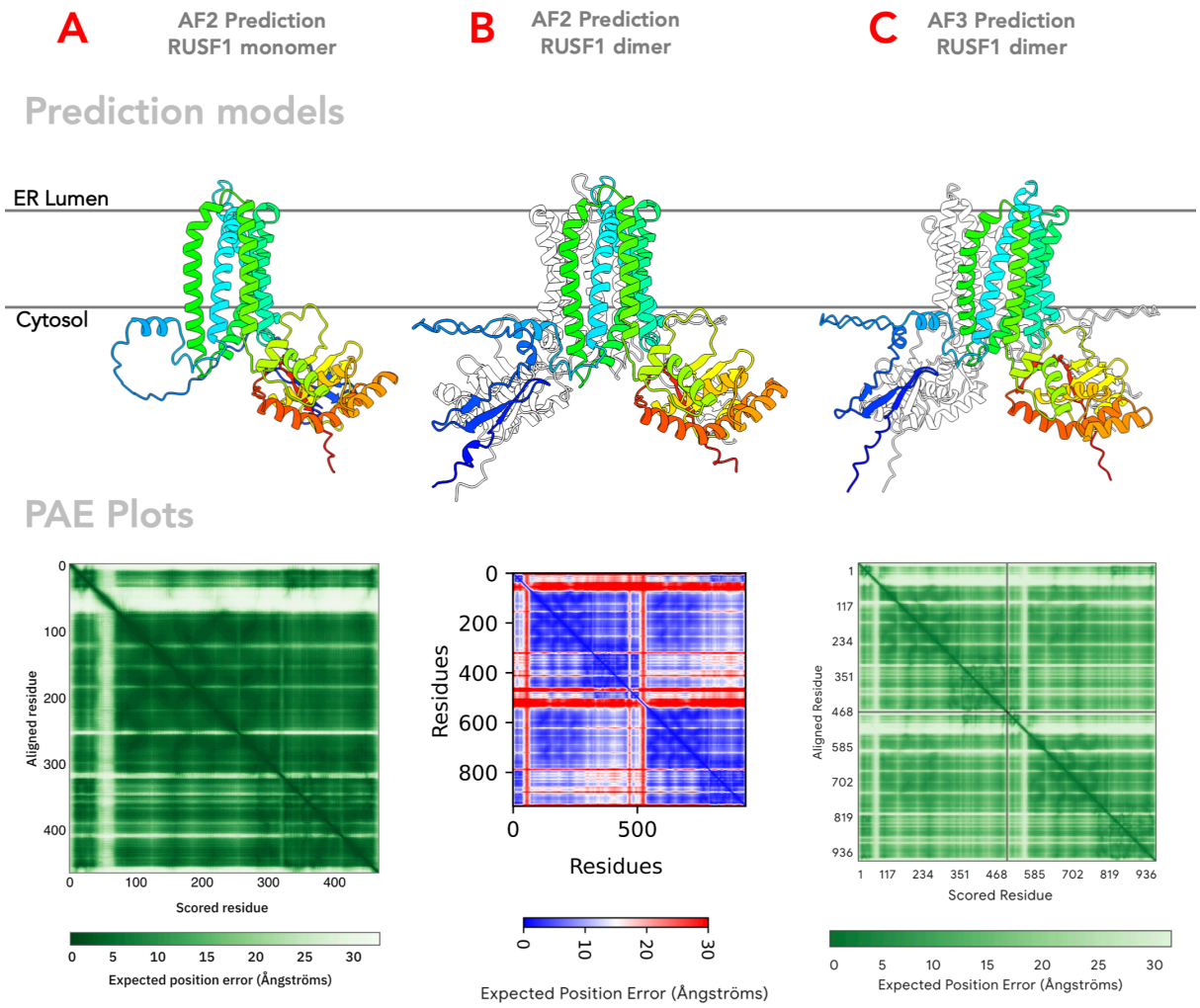
Supplementary Figure 5. Comparison of global conformational changes in ec-CLC and hs-CLC between closed/occluded states and active states. A. Overlay of the ecCLC "Swap" (grey) and "Twist" (rainbow) states (PDB 7N8P, 7N9W), showing one protomer in colour indicated, presented on top of shadow of the dimeric structure (light grey). Major conformational changes indicated with arrows. B. Overlay of the human CLC occluded state (grey, WT PIP2) and active state (rainbow, Y715C), showing one protomer in colour indicated, presented on top of shadow of the dimeric structure (light grey). CBS domains not shown. Major conformational changes indicated with arrows.



Supplementary Figure 6. Q and R helix (residue 589-620) conformations in deposited human CLC7 structures (PDB IDs shown above structures). All structures with CBS domains in close contact with TMD exhibited the canonical "zig-zag" Q-R helix conformation, whereas 9G6D (WT apo) and 9G6E (Y715C) showed an extended QR helix



Supplementary Figure 7. Sequence alignment of Various Eukaryotic RUSF1 Members. Transmembrane helices 1-6 annotated with boxes above sequence. Helix 0 refers to the helix perpendicular to the TMD, prior to helix 1.



Supplementary Figure 8. AlphaFold prediction models (top, one protomer coloured by rainbow from N- to C-terminus) and PAE plots (bottom) of RUSF1. A. AlphaFold 2 prediction of RUSF1 monomer. B. AlphaFold 2 prediction of RUSF1 dimer. C. AlphaFold 3 prediction of RUSF1 dimer.

Supplementary Table 3. CryoEM data collection, refinement and validation statistics for RUSF1 homodimer

	RUSF1 **EMD-54786 PDB: 9SDH**
Microscope	TFS Krios
Detector	*Gatan K3 Bioquantum (6k x 4k)*
Total Dose ($e^{-}Å^{-2}$)	42
Defocus range (μm)	-2.0 to -0.8
Pixel size ($Åpx^{-1}$)	1.04
Micrographs	14,967
Picking software	SIMPLE
Particles (initial)	4,945,927
Particles (final)	284,046
Symmetry	C2
Resolution (Å)	3.13
FSC cut-off	0.143
Map-sharpening B-factor ($Å^2$)	96.2
Mask CC	0.86
Volume CC	0.85
Mean Protein B-factor ($Å^2$)	111.87
Mean Ligand B-factor ($Å^2$)	140.22
R.M.S deviations	
Bond lengths (Å)	0.003
Bond angles (º)	0.594
MolProbity score	1.52
Clash score	6.70
Poor rotamers (%)	0.14
Ramachandran Plot	
Favoured (%)	97.14
Allowed (%)	2.86
Disallowed (%)	0.00

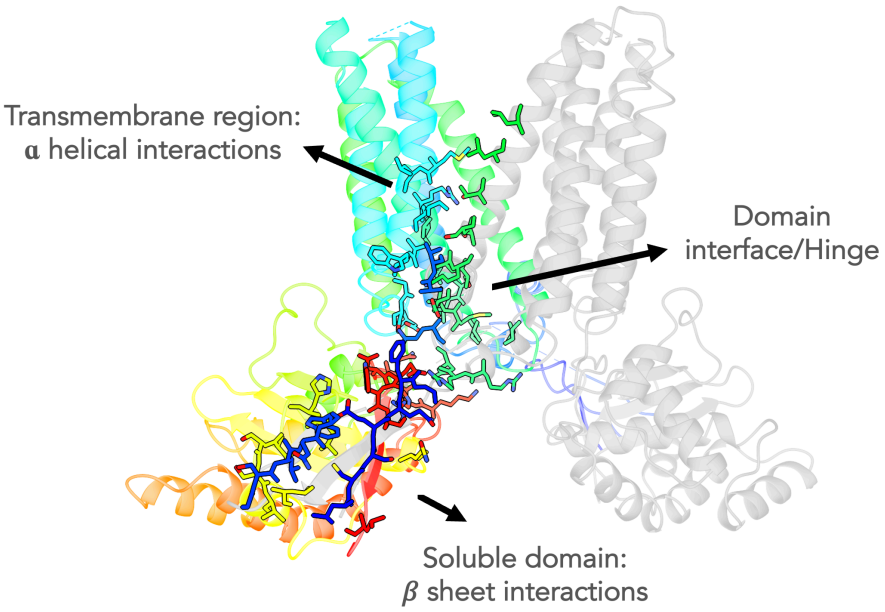
Supplementary Table 4. FoldSeek hits (from PDB100 database) against RUSF1 homodimer.

PDB-assambly_chain	Description	Species	Probability	Seq. Id.	E-Value
6zgr-assembly1_A	Crystal structure of a MFS transporter with bound 1-hydroxynaphthalene-2-carboxylic acid at 2.67 Angstroem resolution	Syntrophobacter fumaroxidans MPOB	0.99	12.90	1.29E-01
6e9o-assembly2_A	E. coli D-galactonate:proton symporter mutant E133Q in the outward substrate-bound form	Escherichia coli	0.99	13.60	1.29E-01
8dl8-assembly1_A	Cryo-EM structure of human ferroportin/slc40 bound to Co2+ in nanodisc	Homo sapiens	0.98	16.90	1.23E-01
6zgu-assembly2_B	Crystal structure of a MFS transporter with bound 3-(2-methylphenyl)propanoic acid at 2.41 Angstroem resolution	Syntrophobacter fumaroxidans MPOB	0.97	12.90	2.78E-01
6hcl-assembly1_A	Crystal structure of a MFS transporter with Ligand at 2.69 Angstroem resolution	Syntrophobacter fumaroxidans MPOB	0.97	13.50	3.37E-01
8dwi-assembly1_A	Molecular Mechanism of Sialic Acid Transport Mediated by Sialin	Homo sapiens	0.97	11.80	3.71E-01
6ob6-assembly2_B	Human equilibrative nucleoside transporter-1, S-(4-nitrobenzyl)-6-thioinosine bound, merohedrally twinned	Homo sapiens	0.96	14.80	3.06E-01
7bp3-assembly1_B	Cryo-EM structure of the human MCT2	Homo sapiens	0.96	17.70	5.45E-01
8dl7-assembly1_A	Cryo-EM structure of human ferroportin/slc40 bound to minihepcidin PR73 in nanodisc	Homo sapiens	0.96	16.30	2.78E-01
8et9-assembly1_A	Cryo-EM structure of the organic cation transporter 2 in complex with 1-methyl-4-phenylpyridinium	Homo sapiens	0.94	12.60	4.28E-01
6g9x-assembly2_B	Crystal structure of a MFS transporter at 2.54 Angstroem resolution	Syntrophobacter fumaroxidans MPOB	0.93	12.40	4.50E-01
7wn1-assembly1_A	Structure of PfNT1(Y190A) in complex with nanobody 48 and inosine	Plasmodium falciparum	0.93	15.40	1.42E-01
5aym-assembly1_A	Crystal structure of a bacterial homologue of iron transporter ferroportin in outward-facing state with soaked iron	Bdellovibrio bacteriovorus HD100	0.93	7.40	5.99E-01
7pn1-assembly1_A	Apo HsPepT1 in the outward facing open conformation	Homo sapiens	0.93	15.50	3.37E-01
6g9x-assembly1_A	Crystal structure of a MFS transporter at 2.54 Angstroem resolution	Syntrophobacter fumaroxidans MPOB	0.92	13.30	3.71E-01
7d5p-assembly2_B	Structure of NorC transporter in an outward-open conformation in complex with a single-chain Indian camelid antibody	Staphylococcus aureus subsp. aureus COL	0.92	13.70	5.99E-01
6w4s-assembly1_F	Structure of apo human ferroportin in lipid nanodisc	Homo sapiens	0.92	16.90	3.89E-01
6zgs-assembly1_A	Crystal structure of a MFS transporter with bound 3-phenylpropanoic acid at 2.39 Angstroem resolution	Syntrophobacter fumaroxidans MPOB	0.91	13.90	2.30E-01
6ob6-assembly1_A	Human equilibrative nucleoside transporter-1, S-(4-nitrobenzyl)-6-thioinosine bound, merohedrally twinned	Homo sapiens	0.90	14.90	5.99E-01
5ayn-assembly1_A	Crystal structure of a bacterial homologue of iron transporter ferroportin in outward-facing state	Bdellovibrio bacteriovorus HD100	0.90	9.30	4.28E-01
7d5p-assembly1_A	Structure of NorC transporter in an outward-open conformation in complex with a single-chain Indian camelid antibody	Staphylococcus aureus subsp. aureus COL	0.90	13.20	7.62E-01
7l17-assembly1_A	Crystal structure of sugar-bound melibiose permease MelB	Salmonella enterica subsp. enterica serovar Typhimurium str. LT2	0.87	12.90	1.07E+00
8sc4-assembly1_A	Human OCT1 bound to metformin in inward-open conformation	Homo sapiens	0.60	9.20	1.42E+00
7ydq-assembly1_A	Structure of PfNT1(Y190A)-GFP in complex with GSK4	Plasmodium falciparum 3D7	0.60	13.70	6.92E-01
7lo8-assembly1_Z	NorA in complex with Fab36	Staphylococcus aureus	0.57	12.70	7.99E-01
7ckr-assembly1_A	Cryo-EM structure of the human MCT1/Basigin-2 complex in the presence of anti-cancer drug candidate BAY-8002 in the outward-open conformation.	Homo sapiens	0.57	10.80	2.53E+00
1pw4-assembly1_A	Crystal Structure of the Glycerol-3-Phosphate Transporter from E.Coli	Escherichia coli	0.57	11.80	2.78E+00
4ja4-assembly2_B	Inward open conformation of the xylose transporter XylE from E. coli	Escherichia coli K-12	0.57	8.80	1.81E+00
8sc6-assembly1_A	Human OCT1 bound to thiamine in inward-open conformation	Homo sapiens	0.57	9.70	2.19E+00
6m2l-assembly2_B	Crystal structure of Plasmodium falciparum hexose transporter PfHT1 bound with C3361	Plasmodium falciparum	0.57	13.40	1.42E+00
8jzu-assembly1_A	SLC15A4_TASL complex	Homo sapiens	0.57	10.80	1.07E+00
7eeb-assembly1_L	Structure of the CatSpermasome	Mus musculus	0.57	11.80	1.12E+00
4iu9-assembly2_A	Crystal structure of a membrane transporter	Escherichia coli K-12	0.54	15.40	1.07E+00
6vbg-assembly2_B	Lactose permease complex with thiodigalactoside and nanobody 9043	Escherichia coli	0.54	13.40	1.42E+00
4xni-assembly1_A	X-ray structure of PepTst1	Streptococcus thermophilus LMG 18311	0.54	8.10	1.90E+00
8jzr-assembly1_B	Outward_facing SLC15A4 monomer	Homo sapiens	0.54	14.50	1.90E+00
6eia-assembly1_A	PepTSt in complex with HEPES (100 mM)	Streptococcus thermophilus	0.54	8.60	2.41E+00
4uvm-assembly1_A	In meso crystal structure of the POT family transporter PepTSo	Shewanella oneidensis MR-1	0.54	11.20	1.99E+00
7xpz-assembly1_A	Structure of Apo-hSLC19A1	Homo sapiens	0.51	10.90	3.54E+00
6e9n-assembly2_B	E. coli D-galactonate:proton symporter in the inward open form	Escherichia coli	0.51	13.50	2.41E+00
4aps-assembly1_B	Crystal structure of a POT family peptide transporter in an inward open conformation.	Streptococcus thermophilus LMG 18311	0.51	15.40	1.02E+00
8sc2-assembly1_A	Human OCT1 bound to diltiazem in inward-open conformation	Homo sapiens	0.51	11.60	1.90E+00
5oxo-assembly1_A	PepTSt apo structure	Streptococcus thermophilus LMG 18311	0.47	8.70	2.92E+00
5a2o-assembly1_B	Crystal structure of the nitrate transporter NRT1.1 from Arabidopsis thaliana in complex with nitrate.	Arabidopsis thaliana	0.47	11.70	2.53E+00
7pmy-assembly1_A	HsPepT2 bound to Ala-Phe in the inward facing partially occluded conformation	Homo sapiens	0.47	17.20	1.49E+00
8jta-assembly1_A	Human VMAT2 complex with tetrabenazine	Homo sapiens	0.44	12.50	2.30E+00
8wx2-assembly1_B	Cryo-EM structure of human SLC15A3 (outward-facing partially occluded)	Homo sapiens	0.41	12.70	1.12E+00
8dl6-assembly1_A	Cryo-EM structure of human ferroportin/slc40 bound to Ca2+ in nanodisc	Homo sapiens	0.41	15.60	1.56E+00
4ikx-assembly1_A	Crystal structure of peptide transporter POT (E310Q mutant)	Geobacillus kaustophilus	0.41	14.20	1.72E+00
7pmw-assembly1_A	HsPepT1 bound to Ala-Phe in the outward facing occluded conformation	Homo sapiens	0.41	15.40	1.49E+00
4zow-assembly1_A	Crystal structure of E. coli multidrug transporter MdfA in complex with chloramphenicol	Escherichia coli K-12	0.38	15.80	2.41E+00
6rw3-assembly2_B	The molecular basis for sugar import in malaria parasites.	Plasmodium falciparum	0.38	10.00	2.53E+00
4ja4-assembly1_A	Inward open conformation of the xylose transporter XylE from E. coli	Escherichia coli K-12	0.38	10.10	2.78E+00
6n3i-assembly1_A	Crystal structure of a double Trp XylE mutants (G58W/L315W)	Escherichia coli	0.38	10.10	1.42E+00
8pnl-assembly1_A	Outward-open conformation of a Major Facilitator Superfamily (MFS) transporter MHAS2168, a homologue of Rv1410 from M. tuberculosis, in complex with an alpaca nanobody	Mycolicibacterium hassiacum DSM 44199	0.38	12.00	1.81E+00

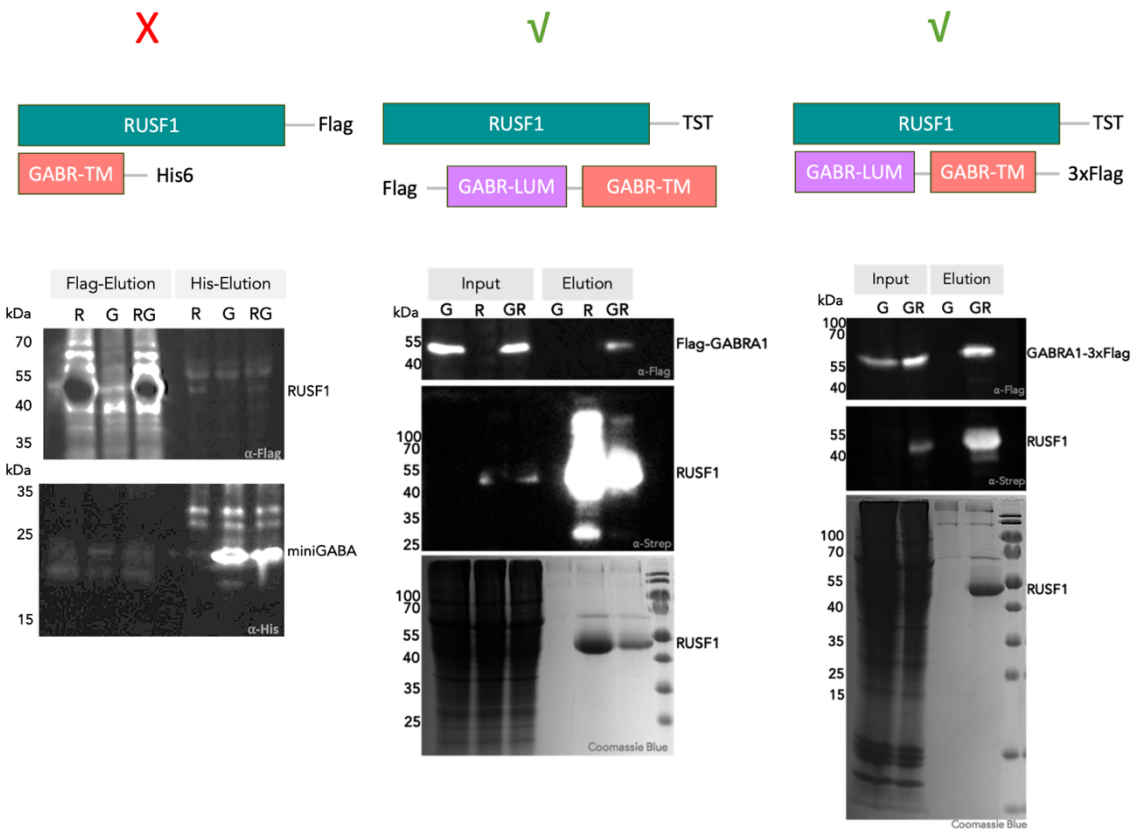
7d5q-assembly2_B	Structure of NorC transporter (K398A mutant) in an outward-open conformation in complex with a single-chain Indian camelid antibody	Staphylococcus aureus subsp. aureus COL	0.85	11.40	9.69E-01
8et8-assembly1_A	Cryo-EM structure of the organic cation transporter 1 in complex with verapamil	Homo sapiens	0.85	9.50	3.37E-01
7nqk-assembly1_A	Cryo-EM structure of the mammalian peptide transporter PepT2	Rattus norvegicus	0.85	16.00	6.92E-01
2y5y-assembly1_B	Crystal structure of LacY in complex with an affinity inactivator	Escherichia coli K-12	0.84	13.30	4.95E-01
7d5q-assembly1_A	Structure of NorC transporter (K398A mutant) in an outward-open conformation in complex with a single-chain Indian camelid antibody	Staphylococcus aureus subsp. aureus COL	0.84	12.50	9.69E-01
6hcl-assembly2_B	Crystal structure of a MFS transporter with Ligand at 2.69 Angstroem resolution	Syntrophobacter fumaroxidans MPOB	0.82	13.50	6.29E-01
8c03-assembly1_A	Structure of SLC40/ferroportin in complex with vamifeport and synthetic nanobody Sy12 in outward-facing conformation	Homo sapiens	0.82	16.10	6.92E-01
8sdz-assembly1_A	Structure of rat organic anion transporter 1 (OAT1) in complex with probenecid	Rattus norvegicus	0.82	11.30	1.29E+00
8bvt-assembly1_A	Cryo-EM structure of rat SLC22A6 bound to probenecid	Rattus nativitatis	0.82	13.00	9.23E-01
6lz0-assembly1_A	Cryo-EM structure of human MCT1 in complex with Basigin-2 in the presence of lactate	Homo sapiens	0.77	10.80	1.42E+00
6oom-assembly2_A-2	PROTEIN A	Escherichia coli	0.77	17.80	8.39E-01
6vbg-assembly1_A	Lactose permease complex with thiodigalactoside and nanobody 9043	Escherichia coli	0.75	12.60	1.99E+00
7xq2-assembly1_A	Structure of hSLC19A1+2'3'-cGAMP	Homo sapiens	0.75	13.00	1.35E+00
6rw3-assembly1_A	The molecular basis for sugar import in malaria parasites.	Plasmodium falciparum	0.75	10.00	7.99E-01
8sdy-assembly1_A	Structure of rat organic anion transporter 1 (OAT1) in complex with para-aminohippuric acid (PAH)	Rattus norvegicus	0.75	12.20	1.49E+00
6zgr-assembly2_B	Crystal structure of a MFS transporter with bound 1-hydroxynaphthalene-2-carboxylic acid at 2.67 Angstroem resolution	Syntrophobacter fumaroxidans MPOB	0.72	12.50	1.07E+00
8t6b-assembly1_A	Human VMAT2 in complex with serotonin	Homo sapiens	0.72	11.30	1.49E+00
6ob7-assembly1_A	Human equilibrative nucleoside transporter-1, dilazep bound	Homo sapiens	0.69	14.70	1.49E+00
7t3n-assembly1_A	R184Q/E191Q mutant of rat vesicular glutamate transporter 2 (VGLUT2)	Rattus norvegicus	0.69	10.30	8.39E-01
6kkl-assembly1_A	Crystal structure of Drug:Proton Antiporter-1 (DHA1) Family SotB, in the inward conformation (H115N mutant)	Escherichia coli K-12	0.66	16.50	1.90E+00
8thr-assembly1_A	Structure of the human vesicular monoamine transporter 2 (VMAT2) bound to tetrabenazine in an occluded conformation	Aequorea victoria	0.66	13.40	7.62E-01
8jzx-assembly1_B	SLC15A4 inhibitor complex	Homo sapiens	0.66	13.50	1.17E+00
6vs2-assembly1_A	protein D	Escherichia coli	0.63	15.50	1.29E+00
6wbv-assembly1_A	Structure of human ferroportin bound to hepcidin and cobalt in lipid nanodisc	Homo sapiens	0.63	17.20	9.69E-01
6m2l-assembly1_A	Crystal structure of Plasmodium falciparum hexose transporter PfHT1 bound with C3361	Plasmodium falciparum	0.63	12.60	1.12E+00
7spt-assembly1_A	Crystal structure of exofacial state human glucose transporter GLUT3	Homo sapiens	0.63	11.80	8.39E-01
6kkj-assembly1_B	Crystal structure of Drug:Proton Antiporter-1 (DHA1) Family SotB, in the inward open conformation	Escherichia coli K-12	0.60	16.80	3.54E+00
6lyy-assembly1_A	Cryo-EM structure of the human MCT1/Basigin-2 complex in the presence of anti-cancer drug candidate AZD3965 in the outward-open conformation.	Homo sapiens	0.60	11.00	1.07E+00
6euq-assembly1_A	MdfA(Q131R/L339E)	Escherichia coli K-12	0.60	17.30	1.49E+00
2y5y-assembly1_A	Crystal structure of LacY in complex with an affinity inactivator	Escherichia coli K-12	0.60	14.10	1.17E+00
6e9o-assembly1_B	E. coli D-galactonate:proton symporter mutant E133Q in the outward substrate-bound form	Escherichia coli	0.60	13.80	1.29E+00
7aar-assembly1_A	sugar/H+ symporter STP10 in inward open conformation	Arabidopsis thaliana	0.38	11.40	1.56E+00
8sbe-assembly1_A	Structure of the rat vesicular glutamate transporter 2 determined by single-particle Cryo-EM	Rattus norvegicus	0.35	10.60	5.99E-01
6wik-assembly1_A	Cryo-EM structure of SLC40/ferroportin with Fab in the presence of hepcidin	Carlito syrichta	0.35	15.70	3.21E+00
4ja4-assembly3_C	Inward open conformation of the xylose transporter XylE from E. coli	Escherichia coli K-12	0.35	9.30	3.21E+00
4qiq-assembly1_A	Crystal structure of D-xylose-proton symporter	Escherichia coli K-12	0.35	12.60	1.90E+00
8sc1-assembly1_A	Human OCT1 (Apo) in inward-open conformation	Homo sapiens	0.35	9.60	3.37E+00
4xnj-assembly1_A	X-ray structure of PepTst2	Streptococcus thermophilus LMG 18311	0.35	7.80	4.50E+00
8jzs-assembly1_A	Outward-facing SLC15A4 dimer	Homo sapiens	0.35	15.10	2.30E+00
6h7d-assembly1_A	Crystal Structure of A. thaliana Sugar Transport Protein 10 in complex with glucose in the outward occluded state	Arabidopsis thaliana	0.35	16.00	1.49E+00
4ikw-assembly1_A	Crystal structure of peptide transporter POT in complex with sulfate	Geobacillus kaustophilus	0.35	13.20	3.37E+00
8jt9-assembly1_A	Human VMAT2 complex with ketanserin	Homo sapiens	0.33	14.90	1.99E+00
8t69-assembly1_A	Human VMAT2 in complex with tetrabenazine	Homo sapiens	0.33	10.20	3.89E+00
6t1z-assembly1_A	LmrP from L. lactis, in an outward-open conformation, bound to Hoechst 33342	Lactococcus lactis	0.33	12.50	2.30E+00
4iu9-assembly1_B	Crystal structure of a membrane transporter	Escherichia coli K-12	0.33	12.80	2.30E+00
4ybq-assembly1_A	Rat GLUT5 with Fv in the outward-open form	Rattus norvegicus	0.33	14.20	9.69E-01
6m20-assembly3_C	Crystal structure of Plasmodium falciparum hexose transporter PfHT1 bound with glucose	Plasmodium falciparum	0.33	13.80	2.53E+00
8bvs-assembly1_A	Cryo-EM structure of rat SLC22A6 bound to tenofovir	Rattus norvegicus	0.33	15.20	3.89E+00
4oaa-assembly1_B	Crystal structure of E. coli lactose permease G46W,G262W bound to sugar	Escherichia coli str. K-12 substr. DH10B	0.30	12.80	4.95E+00
8b1j-assembly1_A	DtpB-Nb132-SL	Escherichia coli	0.30	12.60	2.65E+00
4ybq-assembly2_B	Rat GLUT5 with Fv in the outward-open form	Rattus norvegicus	0.30	14.70	1.35E+00
7zh0-assembly1_A	Structure of human OCT3 in lipid nanodisc	Homo sapiens	0.30	11.30	3.54E+00
7xtk-assembly1_A	Cryo-EM structure of SLC19A1	Homo sapiens	0.28	11.10	4.50E+00
7xq1-assembly1_A	Structure of hSLC19A1+2'3'-CDAS	Homo sapiens	0.28	13.10	5.72E+00
4jre-assembly1_D	Crystal structure of nitrate/nitrite exchanger NarK with nitrite bound	Escherichia coli K-12	0.28	15.20	3.54E+00
4u4w-assembly1_A	Structure of a nitrate/nitrite antiporter NarK in nitrate-bound occluded state	Escherichia coli str. K-12 substr. MG1655	0.28	10.70	1.99E+00
4u4v-assembly1_A	Structure of a nitrate/nitrite antiporter NarK in apo inward-open state	Escherichia coli str. K-12 substr. MG1655	0.28	11.40	2.09E+00
7ac6-assembly1_A	Structure of sponge-phase grown PepTst2 collected by rotation serial crystallography on a COC membrane at a synchrotron source	Streptococcus thermophilus LMG 18311	0.28	8.70	3.54E+00
5oxq-assembly1_A	PepTSt in complex with HEPES (300 mM)	Streptococcus thermophilus	0.28	8.40	5.45E+00
6ghj-assembly1_A	PepTSt in complex with tripeptide Phe-Ala-Gln	Streptococcus thermophilus LMG 18311	0.28	8.20	2.53E+00
4ikz-assembly1_A	Crystal structure of peptide transporter POT (E310Q mutant) in complex with alafosfalin	Geobacillus kaustophilus	0.28	10.50	3.54E+00

8hpj-assembly1_A	Crystal structure of the bacterial oxalate transporter OxlT in a ligand-free outward-facing form	Oxalobacter formigenes	0.25	10.50	2.78E+00
7bc7-assembly1_A	Cryo-EM structure of the proton coupled folate transporter at pH 6.0 bound to pemetrexed	Gallus gallus	0.25	14.50	5.45E+00
4d2b-assembly1_A	Structure of a ligand free POT family peptide transporter	Streptococcus thermophilus LMG 18311	0.25	7.90	6.60E+00
4pyp-assembly1_A	Crystal structure of the human glucose transporter GLUT1	Homo sapiens	0.25	15.30	1.99E+00
6rw3-assembly3_C	The molecular basis for sugar import in malaria parasites.	Plasmodium falciparum	0.25	11.10	3.37E+00
7wsn-assembly1_A	Cryo-EM structure of human glucose transporter GLUT4 bound to cytochalasin B in detergent micelles	Homo sapiens	0.25	14.60	2.41E+00
8wx4-assembly1_A	Cryo-EM structure of human SLC15A4 in complex with Lys-Leu (outward-facing open)	Homo sapiens	0.25	10.90	6.60E+00
4iky-assembly1_A	Crystal structure of peptide transporter POT (E310Q mutant) in complex with sulfate	Geobacillus kaustophilus	0.25	15.00	3.71E+00
8hik-assembly1_A	The TPP-bound BRIL-SLC19A1/Fab/Nb ternary complex	Escherichia coli	0.25	14.20	5.72E+00
7xq0-assembly1_A	Structure of hSLC19A1+3'3'-CDA	Homo sapiens	0.23	14.20	6.60E+00
2v8n-assembly2_B	Wild-type Structure of Lactose Permease	Escherichia coli	0.23	11.70	5.45E+00
7zh6-assembly1_A	Structure of human OCT3 in complex with inhibitor Corticosterone	Homo sapiens	0.23	10.20	5.45E+00
7q0m-assembly1_A	Crystal structure of the peptide transporter YePEPT-K314A in complex with LZNV at 2.66 A	Yersinia enterocolitica subsp. palearctica YE-P4	0.23	11.60	3.89E+00
8hnd-assembly1_A	Cryo-EM structure of human OATP1B1 in complex with estrone-3-sulfate	Homo sapiens	0.23	10.40	3.71E+00
5oge-assembly1_A	Crystal structure of a nucleotide sugar transporter	Saccharomyces cerevisiae S288C	0.21	9.80	8.80E-01
6c9w-assembly1_A	Crystal Structure of a ligand bound LacY/Nanobody Complex	Escherichia coli K-12	0.21	12.40	2.92E+00
8ex4-assembly1_A	Human S1P transporter Spns2 in an inward-facing open conformation (state 1)	Homo sapiens	0.21	12.90	4.29E+00
6s4m-assembly1_A	Crystal structure of the human organic anion transporter MFSD10 (TETRAN)	Homo sapiens	0.21	12.50	2.65E+00
7sp5-assembly2_B	Crystal Structure of a Eukaryotic Phosphate Transporter	Serendipita indica	0.21	13.30	4.95E+00
8sc3-assembly1_A	Human OCT1 bound to fenoterol in inward-open conformation	Homo sapiens	0.21	9.20	6.00E+00
5c65-assembly1_A	Structure of the human glucose transporter GLUT3 / SLC2A3	Homo sapiens	0.21	13.90	3.37E+00
6gs7-assembly1_A	Crystal structure of peptide transporter DtpA-nanobody in glycine buffer	Escherichia coli K-12	0.21	16.30	5.45E+00
8omu-assembly1_A	Cryo-EM structure of rat SLC22A6 bound to alpha-ketoglutaric acid in a low occupancy state	Rattus norvegicus	0.20	13.10	8.80E+00
4zyr-assembly2_B	Crystal structure of E. coli Lactose permease G46W/G262W bound to p-nitrophenyl alpha-D-galactopyranoside (alpha-NPG)	Escherichia coli K-12	0.18	13.60	7.27E+00
5gxb-assembly1_A	crystal structure of a LacY/Nanobody complex	Escherichia coli K-12	0.18	10.90	8.00E+00
3o7p-assembly1_A	Crystal structure of the E.coli Fucose:proton symporter, FucP (N162A)	Escherichia coli K-12	0.18	12.20	5.72E+00
4u4t-assembly1_A	Structure of a nitrate/nitrite antiporter NarK in nitrate-bound inward-open state	Escherichia coli str. K-12 substr. MG1655	0.18	11.50	1.90E+00
6tha-assembly1_A	Crystal structure of human sugar transporter GLUT1 (SLC2A1) in the inward conformation	Homo sapiens	0.18	16.00	3.54E+00
4ja3-assembly2_B	Partially occluded inward open conformation of the xylose transporter XylE from E. coli	Escherichia coli K-12	0.13	12.00	4.72E+00
5oxn-assembly1_A	PepTSt in complex with dipeptide Phe-Ala	Streptococcus thermophilus LMG 18311	0.13	7.80	8.39E+00
6i1z-assembly1_B	Outward facing structure of apo CST	Zea mays	0.12	12.50	2.09E+00
5oxk-assembly1_A	PepTSt in complex with dipeptide Ala-Gln	Streptococcus thermophilus LMG 18311	0.12	7.60	6.93E+00
5oxm-assembly1_A	PepTSt in complex with dipeptide Asp-Glu	Streptococcus thermophilus LMG 18311	0.12	8.90	6.93E+00
8pg0-assembly1_A	Human OATP1B3	Homo sapiens	0.12	11.00	5.72E+00
5oge-assembly4_D	Crystal structure of a nucleotide sugar transporter	Saccharomyces cerevisiae S288C	0.11	10.10	2.19E+00
8hpk-assembly1_A	Crystal structure of the bacterial oxalate transporter OxlT in an oxalate-bound occluded form	Oxalobacter formigenes	0.11	13.90	4.29E+00
8jtc-assembly1_A	Human VMAT2 complex with reserpine	Homo sapiens	0.11	11.60	8.39E+00
7dl9-assembly2_B	Crystal structure of nucleoside transporter NupG	Escherichia coli K-12	0.11	13.10	5.72E+00
4gby-assembly1_A	The structure of the MFS (major facilitator superfamily) proton:xylose symporter XylE bound to D-xylose	Escherichia coli K-12	0.11	11.70	9.69E+00
8pnl-assembly2_C	Outward-open conformation of a Major Facilitator Superfamily (MFS) transporter MHAS2168, a homologue of Rv1410 from M. tuberculosis, in complex with an alpaca nanobody	Mycolicibacterium hassiacum DSM 44199	0.11	12.80	6.60E+00
6hzp-assembly1_A	Crystal structure of a POT family transporter in complex with 5-aminolevulinic acid	Staphylococcus hominis	0.11	12.50	8.00E+00
5oge-assembly3_H-2	Crystal structure of a nucleotide sugar transporter	Saccharomyces cerevisiae S288C	0.10	10.40	2.92E+00
5oge-assembly4_F	Crystal structure of a nucleotide sugar transporter	Saccharomyces cerevisiae S288C	0.10	10.60	3.06E+00
6fmy-assembly1_A	IMISX-EP of S-PepTSt	Streptococcus thermophilus LMG 18311	0.10	14.80	8.39E+00
7t3o-assembly1_A	Rat vesicular glutamate transporter 2 (VGLUT2) in low Cl condition	Rattus norvegicus	0.10	8.10	2.65E+00

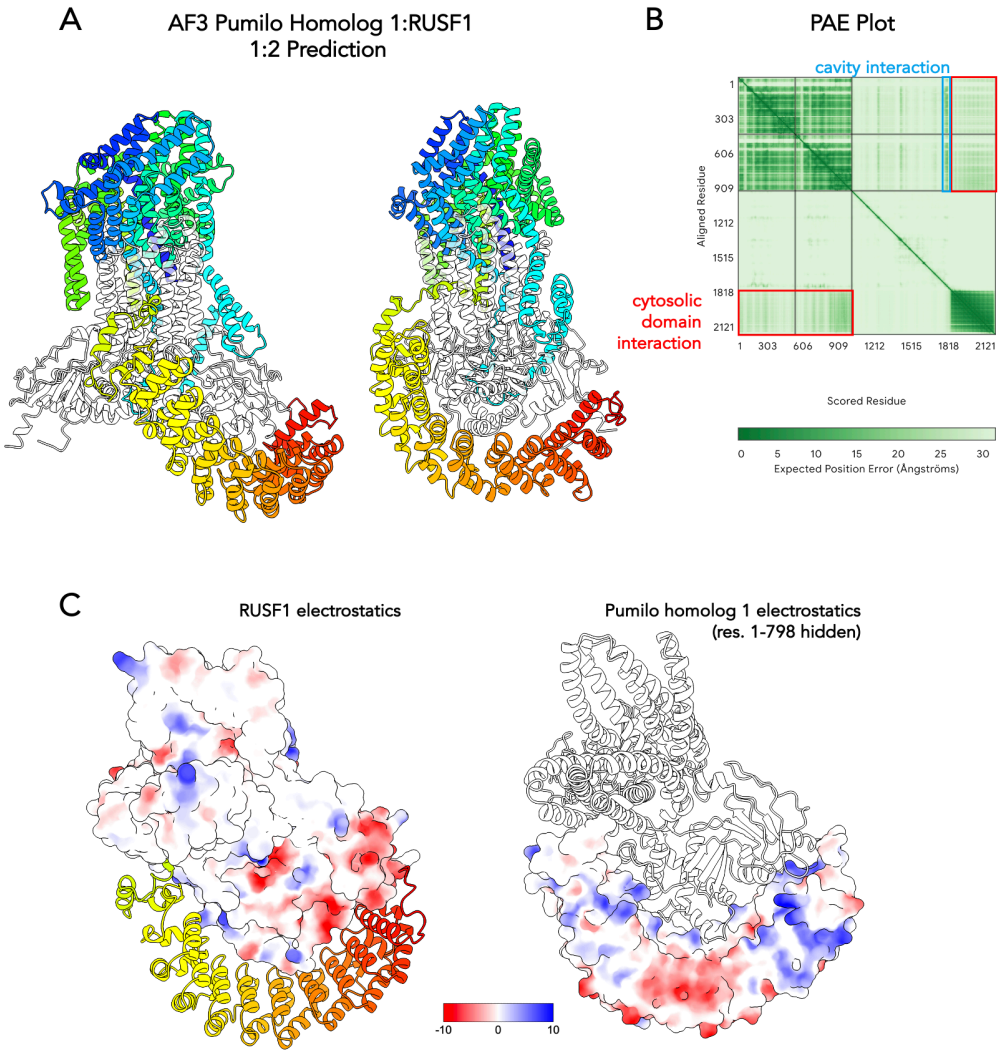
6yof-assembly1_A	Structure of PepTSt from COC IMISX setup collected by rotation serial crystallography on crystals prelocated by 2D X-ray phase-contrast imaging	Streptococcus thermophilus LMG 18311	0.18	16.00	2.92E+00
7q0l-assembly1_A	Crystal structure of the peptide transporter YePEPT-K314A at 2.93 A	Yersinia enterocolitica subsp. palearctica YE-P4	0.18	15.10	3.89E+00
6exs-assembly1_A	Crystal structure of a POT family transporter in complex with thioalcohol conjugated peptide.	Staphylococcus hominis	0.18	9.70	6.29E+00
7zha-assembly1_A	Structure of human OCT3 in complex with inhibitor decynium-22	Homo sapiens	0.18	9.90	5.45E+00
6ei3-assembly1_A	Crystal structure of auto inhibited POT family peptide transporter	Xanthomonas campestris	0.18	13.00	4.50E+00
7lo7-assembly1_Z	NorA in complex with Fab25	Staphylococcus aureus	0.16	15.80	6.00E+00
6kki-assembly1_A	Crystal structure of Drug:Proton Antiporter-1 (DHA1) Family SotB, in the inward-occluded conformation	Escherichia coli K-12	0.16	17.60	6.93E+00
6gv1-assembly1_A	Crystal structure of E.coli Multidrug/H+ antiporter MdfA in outward open conformation with bound Fab fragment	Escherichia coli K-12	0.16	15.20	8.80E+00
5oxl-assembly1_A	PepTSt in complex with dipeptide Ala-Leu	Streptococcus thermophilus LMG 18311	0.16	16.40	4.09E+00
3wdo-assembly1_A	Structure of E. coli YajR transporter	Escherichia coli	0.16	12.20	6.93E+00
2gfp-assembly1_B	Structure of the Multidrug Transporter EmrD from Escherichia coli	Escherichia coli	0.15	12.80	4.29E+00
5ayo-assembly1_A	Crystal structure of a bacterial homologue of iron transporter ferroportin in inward-facing state	Bdellovibrio bacteriovorus HD100	0.15	10.30	4.50E+00
8ex5-assembly1_A	Human S1P transporter Spns2 in an outward-facing open conformation (state 4)	Homo sapiens	0.15	15.00	4.09E+00
6rw3-assembly4_D	The molecular basis for sugar import in malaria parasites.	Plasmodium falciparum	0.15	11.10	4.72E+00
4u4w-assembly2_B	Structure of a nitrate/nitrite antiporter NarK in nitrate-bound occluded state	Escherichia coli str. K-12 substr. MG1655	0.15	11.20	4.09E+00
6yog-assembly1_A	Structure of PepTSt from COC IMISX setup collected by still serial crystallography on crystals prelocated by 2D X-ray phase-contrast imaging	Streptococcus thermophilus LMG 18311	0.15	6.70	5.19E+00
7sps-assembly2_B	Crystal structure of human glucose transporter GLUT3 bound with exofacial inhibitor SA47	Homo sapiens	0.15	14.10	4.95E+00
8qsl-assembly1_C	Cryo-EM structure of human SLC15A4 dimer in outward open state in LMNG	Homo sapiens	0.15	10.70	6.60E+00
6i1z-assembly1_A	Outward facing structure of apo CST	Zea mays	0.14	14.80	1.49E+00
4jre-assembly1_A	Crystal structure of nitrate/nitrite exchanger NarK with nitrite bound	Escherichia coli K-12	0.14	11.40	1.81E+00
5c65-assembly2_B	Structure of the human glucose transporter GLUT3 / SLC2A3	Homo sapiens	0.14	13.00	6.93E+00
7zc2-assembly1_A	Dipeptide and tripeptide Permease C (DtpC)	Escherichia coli	0.14	11.80	8.39E+00
6fmr-assembly1_A	IMISX-EP of Se-PepTSt	Streptococcus thermophilus LMG 18311	0.14	8.90	5.45E+00
6btx-assembly1_A	Structure of a bacterial metal transporter	Bdellovibrio bacteriovorus HD100	0.13	9.90	6.60E+00
4iu8-assembly1_A	Crystal structure of a membrane transporter (selenomethionine derivative)	Escherichia coli K-12	0.13	16.80	4.95E+00
4ja3-assembly1_A	Partially occluded inward open conformation of the xylose transporter XylE from E. coli	Escherichia coli K-12	0.13	10.00	8.80E+00
4yb9-assembly1_D	Crystal structure of the Bovine Fructose transporter GLUT5 in an open inward-facing conformation	Bos taurus	0.13	11.10	5.19E+00
4d2d-assembly1_A	Structure of a tri peptide bound POT family peptide transporter	Streptococcus thermophilus LMG 18311	0.10	9.00	6.93E+00
7y58-assembly1_A	CryoEM structure of QacA (D411N), an antibacterial efflux transporter from Staphylococcus aureus	Staphylococcus aureus	0.09	12.30	9.69E+00
8g92-assembly1_A	Structure of inhibitor 16d-bound SPNS2	Homo sapiens	0.08	10.60	6.93E+00
4d2c-assembly1_A	Structure of a di peptide bound POT family peptide transporter	Streptococcus thermophilus LMG 18311	0.08	5.90	8.80E+00
1cnt-assembly2_3	CILIARY NEUROTROPHIC FACTOR	Homo sapiens	0.08	13.60	3.21E+00
6qsk-assembly2_G-2	Crystal structure of a nucleotide sugar transporter with bound nucleotide monophosphate.	Saccharomyces cerevisiae S288C	0.08	10.90	4.29E+00
3fyq-assembly1_A	Structure of Drosophila melanogaster talin IBS2 domain (residues 1981-2168)	Drosophila melanogaster	0.06	12.80	8.00E+00
5i20-assembly5_E	Crystal structure of protein	Ancylobacter novellus DSM 506	0.04	15.30	7.62E+00
3q84-assembly2_G	Crystal structure of human PACSIN 1 F-BAR domain	Homo sapiens	0.04	13.90	9.24E+00



Supplementary Figure 9. Dimer interface of RUSF1 homodimer. One protomer coloured by rainbow (N- to C-terminus), the opposite protomer coloured light grey. Side chains of residues involved in the interface shown. Main interaction types annotated by arrows.

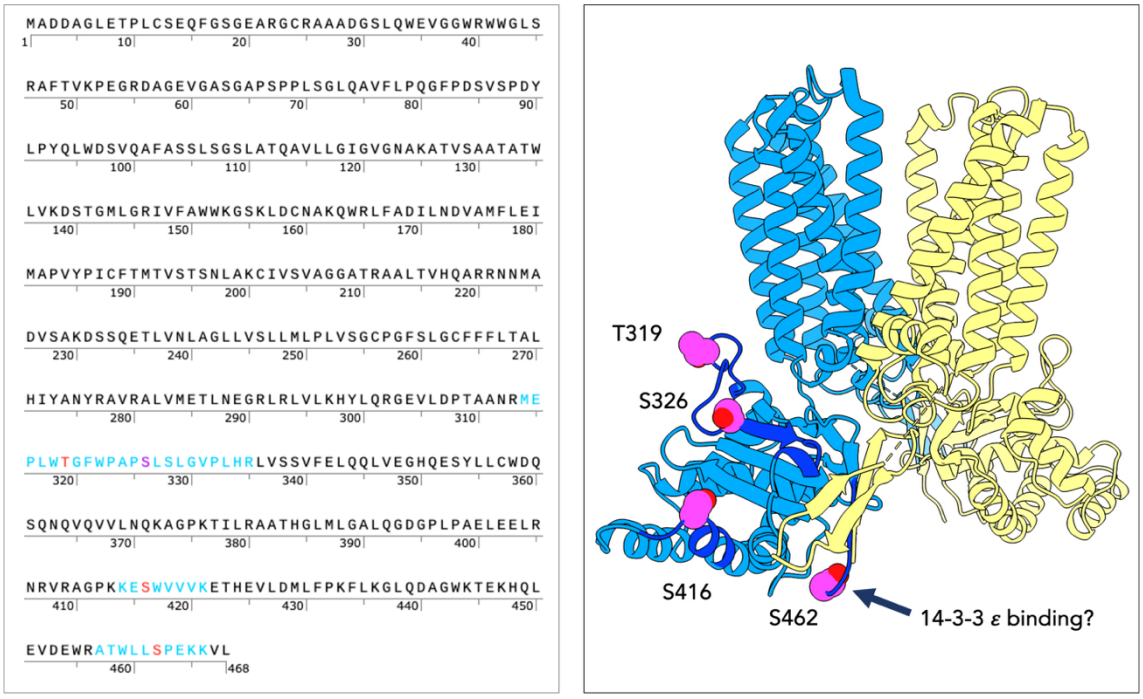


Supplementary Figure 10. Co-immunoprecipitation of various RUSF1 and GABRA1 constructs, showing that RUSF1-FLAG and GABRA1 TMD-His constructs did not interact (or not detected by western blot), while RUSF1-TST interacted with both N- and C- terminally tagged full-length GABRA1.



Supplementary Figure 11. A-B. AlphaFold prediction of RUSF dimer in complex with Pumilo homolog 1, showing extensive interactions in both luminal cavity and cytosolic soluble domain, PAE plot also showed confidence in these predicted interactions. C. Soluble domain interactions showed some degree of charge complementarity.

RUSF1 Phosphorylated Residues by Mass Spectroscopy



Supplementary Figure 12. Phosphorylated peptides in RUSF1 identified through mass spectroscopy. Left: phosphorylated peptide sequences coloured in blue, while phosphorylated residues coloured in red or purple. Right: phosphorylated peptides (dark blue) and phosphorylated residues (pink spheres) shown on the dimeric structure (one protomer coloured blue and another coloured yellow). S462, which AlphaFold predicted to bind 14-3-3 epsilon, was phosphorylated.

Bibliography

Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).

Accardi, A. & Miller, C. Secondary active transport mediated by a prokaryotic homologue of ClC Cl- channels. *Nature* **427**, 803–807 (2004).

Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221 (2010).

Alberts, B. *et al.* The Compartmentalization of Cells. in *Molecular Biology of the Cell. 4th edition* (Garland Science, 2002).

Bai, L., You, Q., Feng, X., Kovach, A. & Li, H. Structure of the ER membrane complex, a transmembrane-domain insertase. *Nature* **584**, 475–478 (2020).

Balla, T. Phosphoinositides: tiny lipids with giant impact on cell regulation. *Physiol Rev* **93**, 1019–1137 (2013).

Bernasconi, R., Galli, C., Calanca, V., Nakajima, T. & Molinari, M. Stringent requirement for HRD1, SEL1L, and OS-9/XTP3-B for disposal of ERAD-LS substrates. *J Cell Biol* **188**, 223–235 (2010).

Bosshart, P. D., Kalbermatter, D., Bonetti, S. & Fotiadis, D. The making of a potent L-lactate transport inhibitor. *Commun Chem* **4**, 128 (2021).

Braakman, I. & Hebert, D. N. Protein Folding in the Endoplasmic Reticulum. *Cold Spring Harb Perspect Biol* **5**, a013201 (2013).

Braunger, K. *et al.* Structural basis for coupling protein transport and N-glycosylation at the mammalian endoplasmic reticulum. *Science* **360**, 215–219 (2018).

Brodsky, J. L. & Skach, W. R. Protein Folding and Quality Control in the Endoplasmic Reticulum: Recent Lessons from Yeast and Mammalian Cell Systems. *Curr Opin Cell Biol* **23**, 464–475 (2011).

Caesar, J. *et al.* SIMPLE 3.0. Stream single-particle cryo-EM analysis in real time. *J Struct Biol X* **4**, 100040 (2020).

Camerino, G. M. *et al.* Protein kinase C theta (PKCθ) modulates the ClC-1 chloride channel activity and skeletal muscle phenotype: a biophysical and gene expression study in mouse models lacking the PKCθ. *Pflugers Arch* **466**, 2215–2228 (2014).

Chakraborty, K., Leung, K. & Krishnan, Y. High lumenal chloride in the lysosome is critical for lysosome function. *eLife* **6**, e28862 (2017).

Chavan, T. S. *et al.* A CLC-ec1 mutant reveals global conformational change and suggests a unifying mechanism for the CLC Cl–/H+ transport cycle. *eLife* **9**, e53479 (2020).

Chen, G., Wei, T., Ju, F. & Li, H. Protein quality control and aggregation in the endoplasmic reticulum: From basic to bedside. *Front Cell Dev Biol* **11**, 1156152 (2023).

Chen, G., Wei, T., Ju, F. & Li, H. Protein quality control and aggregation in the endoplasmic reticulum: From basic to bedside. *Front. Cell Dev. Biol.* **11**, (2023).

Chen, Z. *et al.* EMC chaperone–CaV structure reveals an ion channel assembly intermediate. *Nature* **619**, 410–419 (2023).

Chino, H. & Mizushima, N. ER-Phagy: Quality and Quantity Control of the Endoplasmic Reticulum by Autophagy. *Cold Spring Harb Perspect Biol* **15**, a041256 (2023).

Christianson, J. C. & Carvalho, P. Order through destruction: how ER-associated protein degradation contributes to organelle homeostasis. *EMBO J* **41**, e109845 (2022).

Collins, M. P. & Forgac, M. Regulation and function of V-ATPases in physiology and disease. *Biochim Biophys Acta Biomembr* **1862**, 183341 (2020).

Conte, E. *et al.* Changes in Expression and Cellular Localization of Rat Skeletal Muscle ClC-1 Chloride Channel in Relation to Age, Myofiber Phenotype and PKC Modulation. *Front Pharmacol* **11**, 714 (2020).

Cuddapah, V. A. & Sontheimer, H. Molecular Interaction and Functional Regulation of ClC-3 by Ca2+/Calmodulin-dependent Protein Kinase II (CaMKII) in Human Malignant Glioma. *J Biol Chem* **285**, 11188–11196 (2010).

Di Paolo, G. & De Camilli, P. Phosphoinositides in cell regulation and membrane dynamics. *Nature* **443**, 651–657 (2006).

Dong, X. *et al.* PI(3,5)P(2) controls membrane trafficking by direct activation of mucolipin Ca(2+) release channels in the endolysosome. *Nat Commun* **1**, 38 (2010).

Doyle, D. A. *et al.* The structure of the potassium channel: molecular basis of K+ conduction and selectivity. *Science* **280**, 69–77 (1998).

Dutzler, R., Campbell, E. B., Cadene, M., Chait, B. T. & MacKinnon, R. X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature* **415**, 287–294 (2002).

Dutzler, R., Campbell, E. B. & MacKinnon, R. Gating the Selectivity Filter in ClC Chloride Channels. *Science* **300**, 108–112 (2003).

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Cryst D* **66**, 486–501 (2010).

Fahlke, C. Ion permeation and selectivity in ClC-type chloride channels. *American Journal of Physiology-Renal Physiology* **280**, F748–F757 (2001).

Feng, L., Campbell, E. B., Hsiung, Y. & MacKinnon, R. Structure of a eukaryotic CLC transporter defines an intermediate state in the transport cycle. *Science* **330**, 635–641 (2010).

Fine, M., Schmiege, P. & Li, X. Structural basis for PtdInsP2-mediated human TRPML1 regulation. *Nat Commun* **9**, 4192 (2018).

Fortea, E. *et al.* Structural basis of pH-dependent activation in a CLC transporter. *Nat Struct Mol Biol* **31**, 644–656 (2024).

Freeman, S. A., Grinstein, S. & Orlowski, J. Determinants, maintenance, and function of organellar pH. *Physiological Reviews* **103**, 515–606 (2023).

Furukawa, T. *et al.* Phosphorylation and functional regulation of CIC-2 chloride channels expressed in Xenopus oocytes by M cyclin-dependent protein kinase. *J Physiol* **540**, 883–893 (2002).

Gan, N. *et al.* Structural mechanism of allosteric activation of TRPML1 by PI(3,5)P2 and rapamycin. *Proc Natl Acad Sci U S A* **119**, e2120404119 (2022).

Gangwar, S. P. *et al.* Modulation of GluA2–γ5 synaptic complex desensitization, polyamine block and antiepileptic perampanel inhibition by auxiliary subunit cornichon-2. *Nat Struct Mol Biol* **30**, 1481–1494 (2023).

Geelen, D. *et al.* Disruption of putative anion channel gene AtCLC-a in Arabidopsis suggests a role in the regulation of nitrate content. *The Plant Journal* **21**, 259–267 (2000).

Graves, A. R., Curran, P. K., Smith, C. L. & Mindell, J. A. The Cl-/H+ antiporter ClC-7 is the primary chloride permeation pathway in lysosomes. *Nature* **453**, 788–792 (2008).

Greiner, J. V. & Glonek, T. Intracellular ATP Concentration and Implication for Cellular Evolution. *Biology (Basel)* **10**, 1166 (2021).

Griffith, J. K. *et al.* Membrane transport proteins: implications of sequence comparisons. *Curr Opin Cell Biol* **4**, 684–695 (1992).

Gu, S. *et al.* Brain α7 Nicotinic Acetylcholine Receptor Assembly Requires NACHO. *Neuron* **89**, 948–955 (2016).

Hammond, G. R. & Burke, J. E. Novel roles of phosphoinositides in signaling, lipid transport, and disease. *Curr Opin Cell Biol* **63**, 57–67 (2020).

Hansen, S. B., Tao, X. & MacKinnon, R. Structural basis of PIP2 activation of the classical inward rectifier K+ channel Kir2.2. *Nature* **477**, 495–498 (2011).

Haßdenteufel, S. *et al.* Chaperone-Mediated Sec61 Channel Gating during ER Import of Small Precursor Proteins Overcomes Sec61 Inhibitor-Reinforced Energy Barrier. *Cell Rep* **23**, 1373–1386 (2018).

He, J. *et al.* Cryo-EM structure of the plant nitrate transporter AtCLCa reveals characteristics of the anion-binding site and the ATP-binding pocket. *Journal of Biological Chemistry* **299**, (2023).

He, P. *et al.* High phosphate actively induces cytotoxicity by rewiring pro-survival and pro-apoptotic signaling networks in HEK293 and HeLa cells. *The FASEB Journal* **35**, e20997 (2021).

Hegde, R. S. & Keenan, R. J. The mechanisms of integral membrane protein biogenesis. *Nat Rev Mol Cell Biol* **23**, 107–124 (2022).

Hegde, R. S. & Keenan, R. J. A unifying model for membrane protein biogenesis. *Nat Struct Mol Biol* **31**, 1009–1017 (2024).

Henderson, P. J. Proton-linked sugar transport systems in bacteria. *J Bioenerg Biomembr* **22**, 525–569 (1990).

Hetz, C. & Papa, F. R. The Unfolded Protein Response and Cell Fate Control. *Mol Cell* **69**, 169–181 (2018).

Hetz, C., Zhang, K. & Kaufman, R. J. Mechanisms, regulation and functions of the unfolded protein response. *Nat Rev Mol Cell Biol* **21**, 421–438 (2020).

Hooda, Y. *et al.* Mechanism of NACHO-mediated assembly of pentameric ligand-gated ion channels. *bioRxiv* 2024.10.28.620708 (2024) doi:10.1101/2024.10.28.620708.

Hughes, H. *et al.* Organisation of human ER-exit sites: requirements for the localisation of Sec16 to transitional ER. *J Cell Sci* **122**, 2924–2934 (2009).

Hwang, T.-C. *et al.* Structural mechanisms of CFTR function and dysfunction. *J Gen Physiol* **150**, 539–570 (2018).

Ishida, H. *et al.* Cryo-EM structures of Toll-like receptors in complex with UNC93B1. *Nat Struct Mol Biol* **28**, 173–180 (2021).

Ishida, Y., Nayak, S., Mindell, J. A. & Grabe, M. A model of lysosomal pH regulation. *J Gen Physiol* **141**, 705–720 (2013).

Itskanov, S. *et al.* A common mechanism of Sec61 translocon inhibition by small molecules. *Nat Chem Biol* **19**, 1063–1071 (2023).

Iyer, R., Iverson, T. M., Accardi, A. & Miller, C. A biological role for prokaryotic ClC chloride channels. *Nature* **419**, 715–718 (2002).

Jentsch, T. J. & Pusch, M. CLC Chloride Channels and Transporters: Structure, Function, Physiology, and Disease. *Physiol Rev* **98**, 1493–1590 (2018).

Jeworutzki, E., López-Hernández, T., Estévez, R. & Pusch, M. Glialcam Affects CLC-Chloride Channels by Activating the Slow Gate. *Biophysical Journal* **104**, 628a (2013).

Jiang, Y., Cheng, Z., Mandon, E. C. & Gilmore, R. An interaction between the SRP receptor and the translocon is critical during cotranslational protein translocation. *J Cell Biol* **180**, 1149–1161 (2008).

Jin, Y. & Xue, J. Lipid kinases PIP5Ks and PIP4Ks: potential drug targets for breast cancer. *Front. Oncol.* **13**, (2023).

Kamhi-Nesher, S. *et al.* A Novel Quality Control Compartment Derived from the Endoplasmic Reticulum. *Mol Biol Cell* **12**, 1711–1723 (2001).

Kasper, D. *et al.* Loss of the chloride channel ClC-7 leads to lysosomal storage disease and neurodegeneration. *EMBO J* **24**, 1079–1091 (2005).

Keeble, A. H. *et al.* Approaching infinite affinity through engineering of peptide–protein interaction. *Proceedings of the National Academy of Sciences* **116**, 26523–26533 (2019).

Kim, H. J., Lee, P. C.-W. & Hong, J. H. Chloride Channels and Transporters: Roles beyond Classical Cellular Homeostatic pH or Ion Balance in Cancers. *Cancers* **14**, 856 (2022).

Kim, J. J. & Hibbs, R. E. Direct structural insights into GABAA receptor pharmacology. *Trends Biochem Sci* **46**, 502–517 (2021).

Kim, W. *et al.* Rapid and sensitive protein complex alignment with Foldseek-Multimer. *Nat Methods* **22**, 469–472 (2025).

Klose, C. J. *et al.* The EMC acts as a chaperone for membrane proteins. *Nat Commun* **16**, 7097 (2025).

Kornak, U. *et al.* Loss of the ClC-7 Chloride Channel Leads to Osteopetrosis in Mice and Man. *Cell* **104**, 205–215 (2001).

Krämer, B. K., Bergler, T., Stoelcker, B. & Waldegger, S. Mechanisms of Disease: the kidney-specific chloride channels ClCKA and ClCKB, the Barttin subunit, and their clinical relevance. *Nat Rev Nephrol* **4**, 38–46 (2008).

Kumazaki, K. *et al.* Structural basis of Sec-independent membrane protein insertion by YidC. *Nature* **509**, 516–520 (2014).

Kurokawa, K., Okamoto, M. & Nakano, A. Contact of cis-Golgi with ER exit sites executes cargo capture and delivery from the ER. *Nat Commun* **5**, 3653 (2014).

Lamming, D. W. & Bar-Peled, L. Lysosome: The Metabolic Signaling Hub. *Traffic* **20**, 27–38 (2019).

Lang, S. *et al.* An Update on Sec61 Channel Functions, Mechanisms, and Related Diseases. *Front. Physiol.* **8**, (2017).

Lange, F. *et al.* In situ architecture of the human prohibitin complex. *Nat Cell Biol* **27**, 633–640 (2025).

Lange, P. F., Wartosch, L., Jentsch, T. J. & Fuhrmann, J. C. ClC-7 requires Ostm1 as a β-subunit to support bone resorption and lysosomal function. *Nature* **440**, 220–223 (2006).

Leasure, C. D. *et al.* ROOT UV-B SENSITIVE2 Acts with ROOT UV-B SENSITIVE1 in a Root Ultraviolet B-Sensing Pathway. *Plant Physiol* **150**, 1902–1915 (2009).

Leisle, L., Ludwig, C. F., Wagner, F. A., Jentsch, T. J. & Stauber, T. ClC-7 is a slowly voltage-gated 2Cl−/1H + -exchanger and requires Ostm1 for transport activity: Slowly voltage-gated 2Cl/H-exchange by ClC-7/Ostm1. *The EMBO Journal* **30**, 2140–2152 (2011).

Lemberg, M. K. & Strisovsky, K. Maintenance of organellar protein homeostasis by ER-associated degradation and related mechanisms. *Mol Cell* **81**, 2507–2519 (2021).

Leray, X. *et al.* Tonic inhibition of the chloride/proton antiporter ClC-7 by PI(3,5)P2 is crucial for lysosomal pH maintenance. *Elife* **11**, e74136 (2022).

Lim, H.-H. & Miller, C. Intracellular Proton-Transfer Mutants in a CLC Cl−/H+ Exchanger. *J Gen Physiol* **133**, 131–138 (2009).

Linsdell, P. Relationship between anion binding and anion permeability revealed by mutagenesis within the cystic fibrosis transmembrane conductance regulator chloride channel pore. *J Physiol* **531**, 51–66 (2001).

Liu, F., Zhang, Z., Csanády, L., Gadsby, D. C. & Chen, J. Molecular Structure of the Human CFTR Ion Channel. *Cell* **169**, 85-95.e8 (2017).

Lloyd-Evans, E. & Waller-Evans, H. Lysosomal Ca2+ Homeostasis and Signaling in Health and Disease. *Cold Spring Harb Perspect Biol* **12**, a035311 (2020).

Lolicato, F., Nickel, W., Haucke, V. & Ebner, M. Phosphoinositide switches in cell physiology - From molecular mechanisms to disease. *Journal of Biological Chemistry* **300**, (2024).

Lüscher, B. P., Vachel, L., Ohana, E. & Muallem, S. Cl− as a bona fide signaling ion. *Am J Physiol Cell Physiol* **318**, C125–C136 (2020).

Madej, M. G., Dang, S., Yan, N. & Kaback, H. R. Evolutionary mix-and-match with MFS transporters. *Proceedings of the National Academy of Sciences* **110**, 5870–5874 (2013).

Medina, D. L. & Ballabio, A. Lysosomal calcium regulates autophagy. *Autophagy* **11**, 970–971 (2015).

Meng, E. C. *et al.* UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* **32**, e4792 (2023).

Miller, C. & White, M. M. Dimeric structure of single chloride channels from Torpedo electroplax. *Proc Natl Acad Sci U S A* **81**, 2772–2775 (1984).

Mindell, J. A. Lysosomal Acidification Mechanisms*. *Annual Review of Physiology* **74**, 69–86 (2012).

Nakagawa, T. Structures of AMPA receptor in complex with its auxiliary subunit cornichon. *Science* **366**, 1259–1263 (2019).

Nicoli, E.-R. *et al.* Lysosomal Storage and Albinism Due to Effects of a De Novo CLCN7 Variant on Lysosomal Acidification. *The American Journal of Human Genetics* **104**, 1127–1138 (2019).

Parker, J. L. *et al.* Cryo-EM structure of PepT2 reveals structural basis for proton-coupled peptide and prodrug transport in mammals. *Science Advances* **7**, eabh3355 (2021).

Parker, J. L. *et al.* Proton movement and coupling in the POT family of peptide transporters. *Proc Natl Acad Sci U S A* **114**, 13182–13187 (2017).

Paulsen, I. T., Brown, M. H. & Skurray, R. A. Proton-dependent multidrug efflux systems. *Microbiol Rev* **60**, 575–608 (1996).

Perry, N., Leasure, C. D., Tong, H., Duarte, E. M. & He, Z.-H. RUS6, a DUF647-containing protein, is essential for early embryonic development in Arabidopsis thaliana. *BMC Plant Biology* **21**, 232 (2021).

Picollo, A. & Pusch, M. Chloride/proton antiporter activity of mammalian CLC proteins ClC-4 and ClC-5. *Nature* **436**, 420–423 (2005).

Planells-Cases, R. *et al.* Endosomal chloride/proton exchangers need inhibitory TMEM9 β-subunits for regulation and prevention of disease-causing overactivity. *Nat Commun* **16**, 3117 (2025).

Polovitskaya, M. M. *et al.* Gain-of-function variants in CLCN7 cause hypopigmentation and lysosomal storage disease. *J Biol Chem* **300**, 107437 (2024).

Posor, Y., Jang, W. & Haucke, V. Phosphoinositides as membrane organizers. *Nat Rev Mol Cell Biol* **23**, 797–816 (2022).

Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290–296 (2017).

Qu, L. *et al.* Lysosomal K+ channel TMEM175 promotes apoptosis and aggravates symptoms of Parkinson's disease. *EMBO reports* **23**, e53234 (2022).

Rahikainen, R. *et al.* Visible Light-Induced Specific Protein Reaction Delineates Early Stages of Cell Adhesion. *J. Am. Chem. Soc.* **145**, 24459–24465 (2023).

Rao, B. *et al.* The cryo-EM structure of an ERAD protein channel formed by tetrameric human Derlin-1. *Science Advances* **7**, eabe8591 (2021).

Rao, B. *et al.* The cryo-EM structure of the human ERAD retrotranslocation complex. *Science Advances* **9**, eadi5656 (2023).

Rapoport, T. A. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* **450**, 663–669 (2007).

Saftig, P. & Klumperman, J. Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. *Nat Rev Mol Cell Biol* **10**, 623–635 (2009).

Samie, M. *et al.* A TRP Channel in the Lysosome Regulates Large Particle Phagocytosis via Focal Exocytosis. *Dev Cell* **26**, 511–524 (2013).

San Martín, A. *et al.* Imaging Mitochondrial Flux in Single Cells with a FRET Sensor for Pyruvate. *PLoS One* **9**, e85780 (2014).

Sarto-Jackson, I. & Sieghart, W. Assembly of GABA(A) receptors (Review). *Mol Membr Biol* **25**, 302–310 (2008).

Scheel, O., Zdebik, A. A., Lourdel, S. & Jentsch, T. J. Voltage-dependent electrogenic chloride/proton exchange by endosomal CLC proteins. *Nature* **436**, 424–427 (2005).

Schoebel, S. *et al.* Cryo-EM structure of the protein-conducting ERAD channel Hrd1 in complex with Hrd3. *Nature* **548**, 352–355 (2017).

Schrecker, M., Korobenko, J. & Hite, R. K. Cryo-EM structure of the lysosomal chloride-proton exchanger CLC-7 in complex with OSTM1. *eLife* **9**, e59555 (2020).

Schrecker, M. *et al.* Structural basis of ClC-3 transporter inhibition by TMEM9 and PtdIns(3,5)P2. *Nat Struct Mol Biol* https://doi.org/10.1038/s41594-025-01617-2 (2025)

Shan, S. & Walter, P. Co-translational protein targeting by the signal recognition particle. *FEBS Letters* **579**, 921–926 (2005).

Shanks, N. F. *et al.* Molecular Dissection of the Interaction between the AMPA Receptor and Cornichon Homolog-3. *J. Neurosci.* **34**, 12104–12120 (2014).

She, J. *et al.* Structural mechanisms of phospholipid activation of the human TPC2 channel. *Elife* **8**, e45222 (2019).

Shibata, Y. *et al.* Mechanisms determining the morphology of the peripheral ER. *Cell* **143**, 774–788 (2010).

Shibata, Y. *et al.* The Reticulon and Dp1/Yop1p Proteins Form Immobile Oligomers in the Tubular Endoplasmic Reticulum. *J Biol Chem* **283**, 18892–18904 (2008).

Siintola, E. *et al.* The Novel Neuronal Ceroid Lipofuscinosis Gene MFSD8 Encodes a Putative Lysosomal Transporter. *Am J Hum Genet* **81**, 136–146 (2007).

Smith, M. D. *et al.* CCPG1 Is a Non-canonical Autophagy Cargo Receptor Essential for ER-Phagy and Pancreatic ER Proteostasis. *Dev Cell* **44**, 217-232.e11 (2018).

Steinberg, B. E., Huynh, K. K. & Grinstein, S. Phagosomal acidification: measurement, manipulation and functional consequences. *Biochem Soc Trans* **35**, 1083–1087 (2007).

Stelzl, L. S., Fowler, P. W., Sansom, M. S. P. & Beckstein, O. Flexible Gates Generate Occluded Intermediates in the Transport Cycle of LacY. *J Mol Biol* **426**, 735–751 (2014).

Stölting, G., Fischer, M. & Fahlke, C. CLC channel function and dysfunction in health and disease. *Front Physiol* **5**, 378 (2014).

Strange, K. Putting the pieces together: A crystal clear window into CLC anion channel regulation. *Channels* **5**, 101–105 (2011).

Sun, S., Li, X. & Mariappan, M. Signal sequences encode information for protein folding in the endoplasmic reticulum. *J Cell Biol* **222**, e202203070 (2022).

Sundaram, A. *et al.* Substrate-driven assembly of a translocon for multipass membrane proteins. *Nature* **611**, 167–172 (2022).

Tang, H. *et al.* The solute carrier SPNS2 recruits PI(4,5)P2 to synergistically regulate transport of sphingosine-1-phosphate. *Molecular Cell* **83**, 2739-2752.e5 (2023).

Tj, J., T, F., A, S. & H, Y. The CLC chloride channel family. *Pflugers Archiv : European journal of physiology* **437**, (1999).

Tong, H. *et al.* Role of root UV-B sensing in Arabidopsis early seedling development. *Proc Natl Acad Sci U S A* **105**, 21039–21044 (2008).

Tong, H. *et al.* Arabidopsis ROOT UV-B SENSITIVE 1 and 2 Interact with Aminotransferases to Regulate Vitamin B6 Homeostasis. 2021.03.01.433438 Preprint at https://doi.org/10.1101/2021.03.01.433438 (2021).

Traverso, S., Elia, L. & Pusch, M. Gating Competence of Constitutively Open CLC-0 Mutants Revealed by the Interaction with a Small Organic Inhibitor. *J Gen Physiol* **122**, 295–306 (2003).

Tseng, P.-Y. *et al.* Binding of ATP to the CBS domains in the C-terminal region of CLC-1. *J Gen Physiol* **137**, 357–368 (2011).

Twomey, E. C., Yelshanskaya, M. V., Grassucci, R. A., Frank, J. & Sobolevsky, A. I. Elucidation of AMPA receptor-stargazin complexes by cryo-electron microscopy. *Science* **353**, 83–86 (2016).

Varela, M. F., Ortiz-Alegria, A., Lekshmi, M., Stephen, J. & Kumar, S. Functional Roles of the Conserved Amino Acid Sequence Motif C, the Antiporter Motif, in Membrane Transporters of the Major Facilitator Superfamily. *Biology (Basel)* **12**, 1336 (2023).

Voeltz, G. K., Prinz, W. A., Shibata, Y., Rist, J. M. & Rapoport, T. A. A class of membrane proteins shaping the tubular endoplasmic reticulum. *Cell* **124**, 573–586 (2006).

Walden, M. *et al.* Uncoupling and Turnover in a Cl−/H+ Exchange Transporter. *J Gen Physiol* **129**, 317–329 (2007).

Wan, Y. *et al.* Structural basis of adenine nucleotides regulation and neurodegenerative pathology in ClC-3 exchanger. *Nat Commun* **15**, 6654 (2024).

Wang, M. & Kaufman, R. J. Protein misfolding in the endoplasmic reticulum as a conduit to human disease. *Nature* **529**, 326–335 (2016).

Wang, S. *et al.* The amino acid transporter SLC38A9 is a key component of a lysosomal membrane complex that signals arginine sufficiency to mTORC1. *Science* **347**, 188–194 (2015).

Wang, W., Zhang, X., Gao, Q. & Xu, H. TRPML1: an ion channel in the lysosome. *Handb Exp Pharmacol* **222**, 631–645 (2014).

Wang, X.-G. *et al.* Tyrosine 284 phosphorylation is required for ClC-3 chloride channel activation in vascular smooth muscle cells. *Cardiovasc Res* **98**, 469–478 (2013).

Wang, Y. *et al.* CLN7 is an organellar chloride channel regulating lysosomal function. *Sci Adv* **7**, eabj9608 (2021).

Wiseman, R. L., Mesgarzadeh, J. S. & Hendershot, L. M. Reshaping endoplasmic reticulum quality control through the unfolded protein response. *Mol Cell* **82**, 1477–1491 (2022).

Wolfgang Hanke & Christopher Miller. Single chloride channels from Torpedo electroplax. Activation by protons. *J Gen Physiol* **82**, 25–45 (1983).

Wu, D. *et al.* Single-cell metabolic imaging reveals a SLC2A3-dependent glycolytic burst in motile endothelial cells. *Nat Metab* **3**, 714–727 (2021).

Wu, H., Smalinskaitė, L. & Hegde, R. S. EMC rectifies the topology of multipass membrane proteins. *Nat Struct Mol Biol* **31**, 32–41 (2024).

Wu, X. & Rapoport, T. A. Mechanistic insights into ER-associated protein degradation. *Curr Opin Cell Biol* **53**, 22–28 (2018).

Wu, X. *et al.* Structural basis of ER-associated protein degradation mediated by the Hrd1 ubiquitin ligase complex. *Science* **368**, eaaz2449 (2020).

Yang, X. *et al.* ATP13A1 engages SEC61 to facilitate substrate-specific translocation. *Science Advances* **11**, eadt1346 (2025).

Yang, Z. *et al.* Molecular mechanism underlying regulation of Arabidopsis CLCa transporter by nucleotides and phospholipids. *Nat Commun* **14**, 4879 (2023).

Zhang, B. *et al.* Molecular basis of ClC-6 function and its impairment in human disease. *Science Advances* **9**, eadg4479 (2023).

Zhang, S. *et al.* Molecular insights into the human CLC-7/Ostm1 transporter. *Sci. Adv.* **6**, eabb4747 (2020).

Zhang, X., Li, X. & Xu, H. Phosphoinositide isoforms determine compartment-specific ion channel activity. *Proceedings of the National Academy of Sciences* **109**, 11384–11389 (2012).

Zhang, Z., Liu, F. & Chen, J. Molecular structure of the ATP-bound, phosphorylated human CFTR. *Proc Natl Acad Sci U S A* **115**, 12757–12762 (2018).

Zhao, Y., Shen, Y., Wen, Y. & Campbell, R. E. High-Performance Intensiometric Direct- and Inverse-Response Genetically Encoded Biosensors for Citrate. *ACS Cent Sci* **6**, 1441–1450 (2020).

Zheng, Y.-J., Furukawa, T., Ogura, T., Tajimi, K. & Inagaki, N. M phase-specific expression and phosphorylation-dependent ubiquitination of the ClC-2 channel. *J Biol Chem* **277**, 32268–32273 (2002).

Zhou, Y., Morais-Cabral, J. H., Kaufman, A. & MacKinnon, R. Chemistry of ion coordination and hydration revealed by a K+ channel–Fab complex at 2.0 Å resolution. *Nature* **414**, 43–48 (2001).

Zhu, Q., Zhu, X. & Zhang, L. ER membrane complex (EMC): Structure, functions, and roles in diseases. *FASEB J* **38**, e23539 (2024).

Zifarelli, G. The Role of the Lysosomal Cl−/H+ Antiporter ClC-7 in Osteopetrosis and Neurodegeneration. *Cells* **11**, 366 (2022).

Zifarelli, G. & Pusch, M. Intracellular regulation of human ClC-5 by adenine nucleotides. *EMBO Rep* **10**, 1111–1116 (2009).

Zimmermann, I. *et al.* Generation of synthetic nanobodies against delicate proteins. *Nat Protoc* **15**, 1707–1741 (2020).

Zimmermann, I. *et al.* Synthetic single domain antibodies for the conformational trapping of membrane proteins. *eLife* **7**, e34317 (2018).

Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, 253–267 (2020).

List of figures

List of tables