Structural studies of ER-associated protein quality control systems

Jacob Robson-Tull

Sir William Dunn School of Pathology
Lincoln College
University of Oxford



A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy at the University of Oxford.

Trinity Term 2023

49,406 words

Acknowledgements

There are a great number of people who helped, encouraged and enriched my studies across the last four years, all of whom I am tremendously grateful for. Firstly, I would like to thank my supervisors, Susan Lea and Pedro Carvalho, for providing me with the opportunity to undertake such an exciting, but difficult project. They both provided constant support throughout and were great motivators through tough times. The opportunity to conduct research across two continents has been a tremendous experience that I am sure it will shape my outlook on life for a long time to come. For this, I am extremely grateful. Past and present members of both the Lea and the Carvalho labs were also invaluable in providing a fun and enriching atmosphere in both the UK and USA. They provided a constant sounding board for scientific ideas and introduced me to many of the techniques employed during my studies. It was truly fun to be a part of both labs. I am especially grateful to Justine Deme for his time and guidance with membrane protein purification and structure determination. Michael van de Weijer was also particularly helpful and welcoming during my start in the lab and gave up a lot of his time to get me started with the basics.

I want to thank the countless support across the Sir William Dunn School of Pathology, Oxford Biochemistry Department and the National Cancer Institute (NCI) for creating a great environment to perform science in. This particularly includes staff of the Oxford COSMIC facility (Errin Johnson, Adam Costin and Joseph Caeser) for training in electron microscopy and help with computational facilities. The biophysics facility at the NCI for taking the time to train and help me optimise samples for use on their instruments. And finally, Anna Valkov and her team in the Protein Technology Core at the NCI for their continued hard work in generating sybodies for the project. I also want to thank Christian Siebold for gifting us an important cell line early on in the project, to help the beginning of the project move faster. Thank you also to staff in the Dunn School and at the NCI

(particularly Lucinda Risius and Cindy Castle) for helping me with non-academic duties and facilitating an easy transition to work in the USA. I am also grateful for funding from the Medical Research Council during the first years of my studies.

Lastly, I would like to thank the support of my partner, Eunjin Kim, countless friends and my parents for their constant support in and outside of the lab. Eunjin and my friends were always there to see me through bad days in the lab and this experience has bonded us all, for better or worse. Both my parents also provided tremendous support throughout my studies, whether that be help with moving house year on year or occasionally providing a free laundry service. My studies would not have been as smooth without their continual support. The many members of the Lincoln College MCR made Oxford really enjoyable, even throughout the pandemic. Studying at Oxford was a truly special experience and I will be sad to finally bring my time as a student to an end.

Abstract

ER-associated protein degradation (ERAD) is an important degradation pathway for the removal of misfolded proteins from the endoplasmic reticulum (ER). This pathway encompasses a set of evolutionary conserved mechanisms for the selection, ubiquitination and retrotranslocation of membrane and lumenal proteins from the ER to the cytosol. Recent structural studies of the yeast Hrd1 E3 ligase complex identified a possible mechanism by which ERAD substrates are retrotranslocated across the ER membrane. Yeast Hrd1 forms a heterodimer with the rhomboid pseudoprotease Der1, each of which is proposed to contribute a half-channel in the membrane for retrotranslocation. Crosslinking studies also support this hypothesis. However, the mechanisms for selection of membrane substrates are still unknown and the identity of the translocon in other ERAD pathways remains elusive. This is further hindered by a lack of structural studies, especially in mammalian systems.

This thesis presents a bottom-up screen for the identification of mammalian ERAD factors amenable to structure determination by cryo-electron microscopy (cryo-EM). High-resolution structures of the mammalian E3 ligase Hrd1 were determined and two distinct dimeric states identified. These are termed the “open” and “closed” state due to rearrangements at the dimer interface that lead to repositioning of the E3 ligase RING domains within the dimer. These structures are compared with their yeast counterpart and analysed to understand key differences that may be important in ERAD. A mammalian expression system was developed for reconstitution of Hrd1 complexes containing additional binding partners. Complexes containing Hrd1 and its core interactors were isolated and imaged by cryo-EM. Other ERAD complexes were also purified and imaged by cryo-EM. This thesis provides a foundation for future structural studies of mammalian ERAD complexes.

Contents

Acknowledgements ..... *i*

Abstract ..... *iii*

Contents *1*

1. Introduction *6*

1.1. Membrane protein biosynthesis in the ER **6**

1.2. Quality control in the ER **7**

1.3. ERAD as a quality control pathway **9**

1.3.1. Ubiquitination cascade 9

1.3.2. ERAD principles 10

1.2.3. Differences between yeast and mammals 13

1.4. Substrate recognition in ERAD **14**

1.4.1. Lumenal surveillance 14

1.4.2. Membrane surveillance 15

1.5. Export of proteins from the ER **16**

1.5.1. The Hrd1-Der1 retrotranslocon 17

1.5.2. Rhomboid pseudoproteases 20

1.5.3. Other translocons 21

1.6. Project aims **22**

2. Materials and Methods *24*

2.1. Cloning & mutagenesis **24**

2.1.1. Mammalian expression vectors 24

2.1.2. Cloning genes 25

2.1.3. Mutagenesis 26
2.1.4. Restriction digestion 26
2.1.5. DNA assembly 26
2.1.6. Agarose gel purification 27
2.1.7. Bacterial transformation 27
2.1.8. Plasmid purification and sequencing 27
2.2. Cell culture 27
2.2.1. Transfection 28
2.2.2. Stable cell line generation 28
2.2.3. Suspension culture 30
2.2.4. C-terminal CRISPR tagging 31
2.3. Protein expression and purification 32
2.3.1. Mammalian protein expression 33
2.3.2. Mammalian membrane protein purification 33
2.3.3. Mammalian soluble protein purification 38
2.3.4. Bacterial protein expression 39
2.3.5. Bacterial protein purification 39
2.3.6. Detergent screening 40
2.3.7. Small-scale sybody expression and purification 41
2.4. Biochemical methods 42
2.4.1. Flow cytometry 42
2.4.2. SDS-PAGE and western blotting 42
2.4.3. Small-scale non-denaturing pulldowns 43
2.4.4. Analysis of sybody binding by pulldown 45
2.4.5. Analysis of sybody binding by SEC 45
2.4.6. Negative stain-EM 45
2.4.7. TEV protease protection assay 46
2.4.8. Circular dichroism (CD) 46

2.4.9. Nano-differential scanning fluorimetry (nanoDSF) 47
2.4.10. Size exclusion with multi-angle light scattering (SEC-MALS) 47

2.5. Cryo-EM 47

2.5.1. Vitrification 47
2.5.2. Screening 47

3. A bottom-up screen for candidate E3 ligase complexes for structural studies 48

3.1. Screening and characterisation 48

3.1.1. Identifying targets 48
3.1.3. Developing a mammalian expression system 57
3.1.4. Expression screening 60
3.1.5. Detergent screening 63

3.2. Small-scale purifications 69

3.3. Electron microscopy screening 71

3.4. Discussion 76

4. Cryo-EM and structural analysis of mammalian Hrd1 78

4.1. Cryo-EM of the Hrd1 E3 ligase 78

4.1.1. Sample preparation 78
4.1.2. Data collection 80
4.1.3. Data processing 82
4.1.4. Model building 93

4.2. Structural analysis of Hrd1 93

4.2.1. Prior structural work on yeast Hrd1 93
4.2.2. Hrd1 closed conformation 94
4.2.3. Hrd1 open conformation 99

4.2.4. Analysing in vitro assembly of purified Hrd1 107
4.2.5. Comparison with yeast Hrd1 structures 110

5. Discussion 114

5. Reconstituting higher-order Hrd1 E3 ligase complexes for cryo-EM 117

5.1. Endogenous partners of mammalian Hrd1 117

5.2. Purification of higher-order Hrd1 complexes 121

5.2.1. Purification of Hrd1-FAM8A1 121
5.2.2. Purification of Hrd1-Herp-FAM8A1 130
5.2.3. Purification of Hrd1-Herp-FAM8A1-DERL2 133
5.2.4. Attempts to reconstitute Sel1L with Hrd1 141
5.2.5. Purification of Hrd1-Herp-FAM8A1-DERL2-Sel1L 146
5.2.6. Purification of DERL1 with a Hrd1 complex 153
5.2.7. Attempts to reconstitute substrate with Hrd1 156

5.3. Discussion 157

6. Generation of high-affinity binders to Hrd1 complexes 161

6.1. Nanobodies as tools in cryo-EM 161

6.1.1. Synthetic nanobodies 163

6.2. Generation of nanobodies from llamas 165

6.2.1. Liposome reconstitution of Hrd1 complexes 165
6.2.2. Generating reagents for panning 170
6.2.3. Panning and assessment of hits after llama immunisation 175

6.3. Generating sybodies 176

6.3.1. Sybody strategy 176
6.3.2. Initial screening of sybody library against Hrd1$_{1-344}$-Avi 177

6.4. Discussion 182

7. Reconstitution of the RNF185 E3 ligase complex for cryo-EM..... 185

7.1. Composition of the RNF185 E3 ligase complex 185

7.2. Purification of H. sapiens MBRL..... 187

7.2.1. Screening other membrane mimetics 190

7.2.2. Screening the biochemical behaviour of MBRL homologues 196

7.3. Co-purification of partners with MBRL 202

7.3.1. Reconstitution of RNF185-MBRL 202

7.3.2. Reconstitution of RNF185-MBRL-TMUB complexes..... 205

7.3.3. Reconstitution of MBRL with substrate 208

7.4. Isolation of the endogenous RNF185-Membralin complex..... 210

7.5. Discussion 213

8. Conclusions and future perspectives..... 217

8.1. The Hrd1 E3 ligase complex 217

8.2. The RNF185 E3 ligase complex 220

8.3. Future directions for structural studies of mammalian E3 ligases 222

9. References..... 225

10. Supplementary Data..... 248

1. Introduction

Eukaryotic cells are characterised by the presence of membrane-bound intracellular compartments, each contributing a unique functional role. The physiochemical properties of these cellular organelles are diverse and are driven by organelle-specific proteomes. These proteomes are highly specialised and promote the retention of a unique set of proteins whilst clearing those with incorrect folding, assembly or localisation. This fidelity is needed to maintain a diverse set of cellular functions.

1.1. Membrane protein biosynthesis in the ER

Cells require a never-ending supply of protein to maintain their diverse functions, up to one third of which is estimated to be membrane and secretory proteins [1,2]. Membrane proteins carry out diverse functions and are involved in cell signalling, intracellular trafficking, inter- or intra-cellular communications, organelle biogenesis and more [3]. The endoplasmic reticulum (ER) is the largest membrane-bound organelle in eukaryotes and is the entry point for membrane and secretory protein synthesis [4]. This organelle serves as the staging ground for initial protein folding, modification and assembly [4]. The minimal requirement for a membrane protein is the presence of a single transmembrane domain (TMD) that consists of ~15-25 mostly hydrophobic amino acid residues, arranged in an α-helical segment spanning the lipid bilayer [5]. All integral membrane proteins are translated by cytosolic ribosomes and thus they present a challenge for synthesis given their hydrophobic nature and the presence of soluble domains whose premature folding can preclude membrane insertion [4]. This challenge is overcome by coupling synthesis to ER membrane insertion [6]. Co-translational insertion minimises exposure of hydrophobic TMDs to the aqueous cytosol and bypasses premature folding [7]. Protein activity and trafficking is dependent on correct folding and attachment of post-translational modifications. With the aid of a myriad of folding and

assembly chaperones in the ER lumen and membrane, unfolded nascent chains entering the ER fold towards their native conformations. Glycosylation and formation of disulphide bonds in the oxidising ER lumen also facilitate this [8]. After preliminary folding, assembly of subunits into multiprotein complexes with defined stoichiometries takes place [4].

Despite a host of chaperones available to guide protein folding in the ER, a large proportion of newly synthesised proteins fail to achieve their native conformation. 12-15% of newly synthesis proteins in mammalian cells are co-translationally eliminated by the ubiquitin-proteasome system [9,10]. This number is possibly higher in the ER where proteins may fail to translocate efficiently, fail to be post-translationally modified or for membrane proteins, fail to achieve the correct topology [11]. This problem is more pronounced in disease states that compromise protein biogenesis, such as mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) that impair domain assembly [12]. Failure to achieve the correctly folded state can be caused by limited chaperone availability, non-stoichiometric quantities of binding partner(s), mutations, oxidation, heat or even heavy metal stress [13]. Moreover, proteins fold with different efficiencies and consequently some are more prone to misfolding such as CTFR, of which 50% fails to mature [14]. An inability of proteins to reach their final state or to clear aberrant proteins is implicated in numerous human diseases, such as Parkinson's and Alzheimer's disease [15]. The ER must therefore be constantly surveilled and cleared of terminally misfolded proteins. Quality control systems play an integral part in this maintenance program and constantly surveille newly synthesised and existing protein, ensuring that proteins reach their mature states [16].

1.2. Quality control in the ER

Protein homeostasis in the ER is directly maintained by three distinct pathways: the unfolded protein response (UPR), ER-associated protein degradation (ERAD) and

autophagy [11]. UPR is reactive and responds to an inability to process unfolded/misfolded protein in the ER, leading to large scale cellular changes to cope with folding demand [17]. This involves adjustments to protein and lipid synthesis, chaperone levels and the activity of degradation pathways [18]. High ER folding stress leads to activation of UPR and prolonged stress can lead to apoptosis [19]. ERAD and autophagy remove protein substrates from the ER for degradation. ERAD is the most understood protein quality control pathway and was historically known for its role in degradation of misfolded and unassembled/orphan substrates, but recent work has also shown it to be critical for regulated proteolysis [20]. This pathway provides a level of fine post-translational control to a large proportion of the proteome, as well as a safety net for times of high folding demand.

Protein fidelity and localisation is also maintained by a set of conserved chaperones at the ER membrane. The ER membrane protein complex (EMC) is implicated in post-translational insertion and reinsertion of low hydrophobicity TMDs [21,22]. This 10-component transmembrane complex spans the ER membrane and possess a cytosolic envelope for TMD insertion [23,24]. The EMC has been shown to act as both an insertase and holdase, and recent cryo-EM structures reveal that it can engage both cytosolic and transmembrane segments of substrates, acting as a holdase during membrane protein assembly [25,26]. A selectivity filter within the cytosolic leaflet also reduces mislocalisation by preventing mistargeting of mitochondrial tail-anchored (TA) proteins [27]. TA proteins are characterised by a single TMD at their C-terminus [28]. Conversely, the P5A-ATPase Spf1 (ATP13A1 in mammals) works to dislocate mislocalised and misinserted TMDs from the ER membrane [29,30].

1.3. ERAD as a quality control pathway

ERAD encompasses a set of evolutionary conserved pathways in eukaryotes for the export of misfolded proteins from the ER into the cytosol [20]. No proteases are present in the ER lumen that can degrade terminally misfolded proteins and so the cytosolic ubiquitin-proteasome system is utilised during ERAD for degradation of all substrates [31]. The ER membrane acts as a physical barrier for separation of the folding and degradation machineries, thereby protecting folding intermediates from premature degradation. Therefore, removal of proteins from the ER is a multi-step process that requires cooperation of factors on both sides of the ER membrane.

1.3.1. Ubiquitination cascade

ERAD substrates are degraded by the ubiquitin-proteasome system upon delivery to the cytoplasm. This requires covalent modification of substrates with the 76-residue polypeptide ubiquitin via conjugation to substrate lysines, and in some cases, serine, threonine or cystine residues [32–34]. Ubiquitin is highly conserved amongst eukaryotes and serves as a platform for interaction with proteins containing ubiquitin-binding domains [35]. Ubiquitination requires an enzyme cascade whereby a ubiquitin-activating enzyme (E1) activates a ubiquitin molecule by formation of a thioester linkage with the C-terminus of ubiquitin (figure 1.1) [36]. This ubiquitin is then transferred to a ubiquitin-conjugating enzyme (E2) [37]. An E3 ubiquitin ligase (hereafter referred to as E3 ligase) recruits both substrate and an E2-ubiquitin conjugate to mediate the transfer of ubiquitin to substrate acceptor sites [38]. There are three types of E3 ligases: HECT (Homologous to the E6-AP carboxy terminus), RING (Really Interesting New Gene) and U-box types [39]. RING E3 ligases are the largest class and E3s involved in ER-resident E3s implicated in ERAD are all RING-type ligases [40]. RING domains adopt a conserved fold for the recruitment of a charged E2-ubiquitin and mediate ubiquitin transfer to

substrate by bringing substrate in close proximity to the E2 enzyme [41]. Ubiquitin is added as either a monomer or is extended to form chains through conjugation to one of its seven lysines or its N-terminus [42]. The length and linkage type determines the fate of the substrate but in ERAD, polyubiquitination usually results in extraction from the membrane and degradation by the 26S proteasome [16].

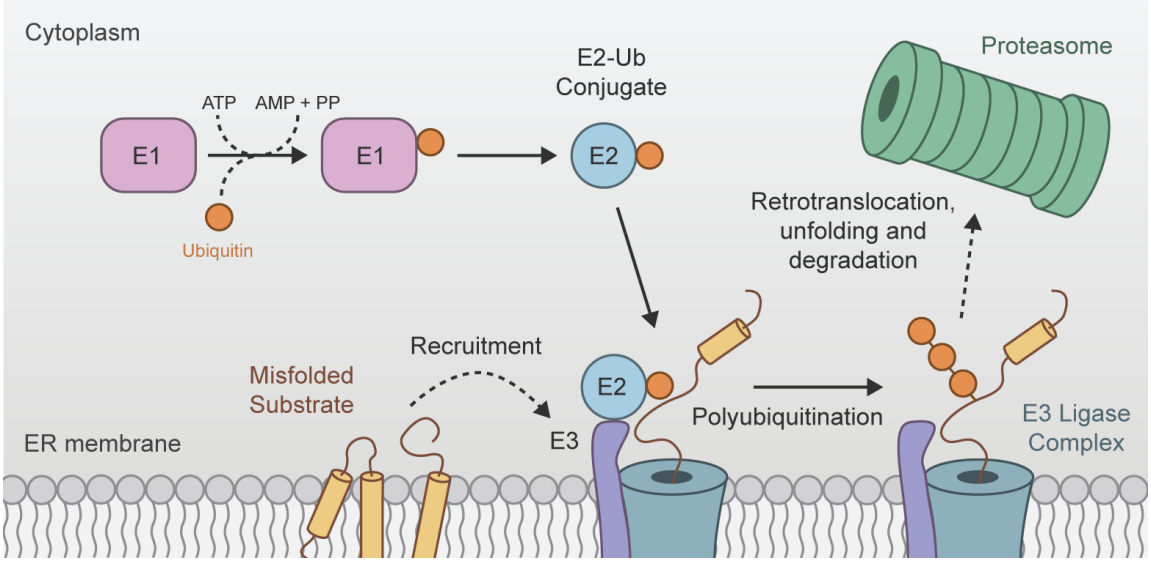


Figure 1.1. The ubiquitination cascade in ERAD. A ubiquitin-E1 conjugate is formed by ATP hydrolysis. Ubiquitin is then transferred to the catalytic site of an E2 enzyme. The E3 ligase recruits both protein substrate and a ubiquitin-loaded E2, mediating transfer of ubiquitin to substrate. Polyubiquitinated substrates then undergoes retrotranslocation and unfolding before delivery to the proteasome for degradation.

1.3.2. ERAD principles

The ERAD machinery must distinguish folding intermediates from misfolding and retrotranslocate substrates from the ER lumen and membrane into the cytosol. This process is made more complex by the fact that it also recognises substrates for regulated proteolysis [43,44]. These substrates are unlikely to be truly misfolded and instead are

thought to undergo conformational changes that expose surfaces for recognition by ERAD factors. ERAD is subdivided into branches governed by individual E3s, each processing a distinct set of substrates whilst operating in parallel with other ERAD pathways [16]. These E3s also interact with a discrete set of partners and together they constitute a system capable of degrading a diverse range of substrates [20].

ERAD begins with the delivery of misfolded substrate to an E3 ligase complex (figure 1.2). Substrates are engaged in either the lumen, membrane or cytosol by ERAD factors, termed ERAD-L, ERAD-M and ERAD-C, respectively [45–47]. In yeast, the recognition location drives E3 ligase specificity [45]. Substrates for retrotranslocation are either soluble lumenal proteins or ER-resident membrane proteins. Since the ubiquitination machinery of these complexes resides in the cytosol, soluble substrates must at least be partially unfolded and threaded through the ER membrane before ubiquitination [35]. Membrane substrates may already expose ubiquitination sites upon recruitment to E3 ligase complexes but this process is still unknown. Once suitable cytosolic substrate sites are polyubiquitinated by the E3 ligase complex, all ERAD branches converge on the cytosolic AAA-ATPase cdc48 in yeast, or p97/VCP in mammals, for retrotranslocation and unfolding [48–50]. Cdc48/p97/VCP monomers assemble into a pair of double-stacked hexameric rings, termed D1 and D2, with a central pore [50,51]. The co-factors Npl4/Ufd1 are recruited to the D1 side and these mediate initial interaction with substrate [50]. This complex is tethered to the ER membrane by VCP-interacting motifs present in ER-resident membrane proteins, thus facilitating continuous processing of substrates emerging from the ER membrane [52,53]. Npl4-Ufd1 bind polyubiquitin chains and binding of Npl4 to ubiquitin, surprisingly, causes its unfolding which then allows for threading of the substrate polypeptide into the pore of Cdc48/VCP/p97 [50]. ATPase cycles then move substrate through the pore in a hand-over-hand mechanism, similar to other AAA-ATPases, using aromatic residues to pull the polypeptide through [50,54]. This process unfolds substrates as they move through the pore. Membrane extraction must

be coupled with either a chaperone or translocon at the ER membrane to overcome the energetic barrier for protein dislocation from the membrane [55]. The mechanisms by which substrates are retrotranslocated are still unclear and it is an active area of research. Substrate traps have been developed and numerous crosslinking studies performed to track the path of substrates through E3 ligase complexes [55,56]. This, together with structural approaches for ERAD-substrate complexes, will undoubtedly shed more light on this process.

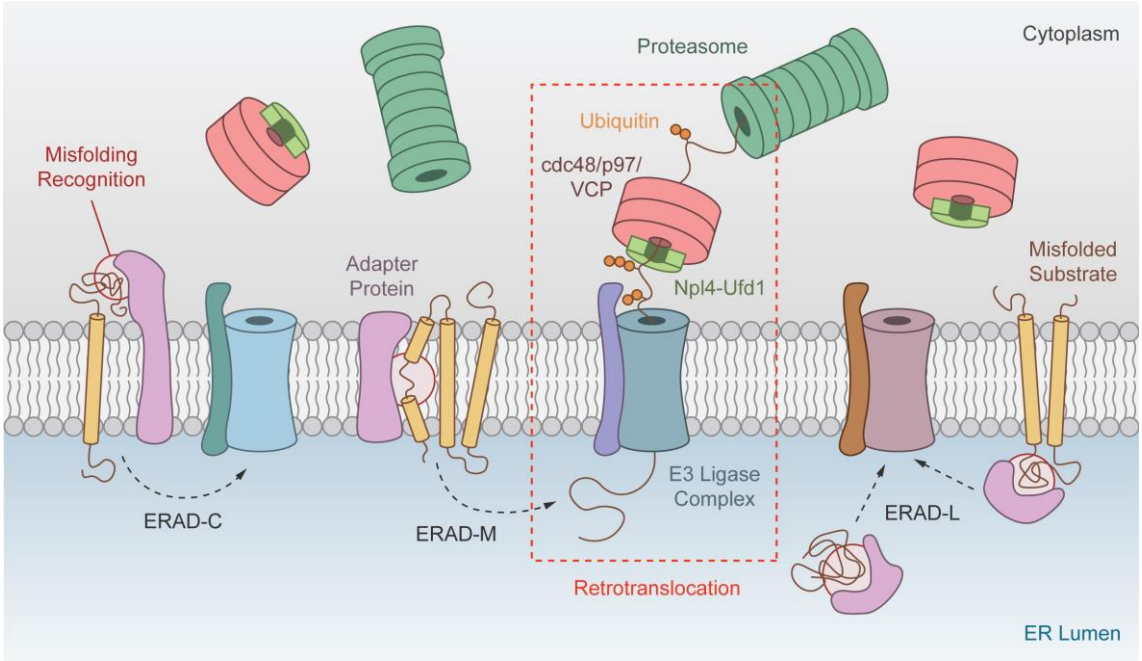


Figure 1.2. Key processing steps in ERAD. Substrates are recognised in the cytosol (ERAD-C), membrane (ERAD-M) or lumen (ERAD-L) by ERAD factors and recruited to ER-resident E3 ligase complexes. Polyubiquitination by the E3 ligase triggers extraction of substrate from the ER membrane by Cdc48/p97/VCP and its cofactors Npl4-Ufd1. Unfolding by Cdc48/p97/VCP is a probable prerequisite to degradation by the proteasome.

1.2.3. Differences between yeast and mammals

Our understanding of ERAD builds primarily on studies in *Saccharomyces cerevisiae* where there are three ERAD branches. Hrd1 and Doa10 process substrates in the ER whilst the Asi complex surveilles the nucleus for orphan subunits [45,57]. Each of these E3 ligases was initially thought to process a distinct class of substrates but this is only true to some extent [45]. Hrd1 is the only E3 ligases capable of processing soluble, lumenal substrates (ERAD-L) whilst Doa10 processes all substrates with cytosolic legions (ERAD-C) [45]. However, all yeast ERAD branches are capable of processing membrane-associated substrates (ERAD-M) [57–59]. The inability of other branches to process lumenal substrates is probably derived from the interactions that Hrd1 makes with Hrd3 (Sel1L in mammals). Hrd3/Sel1L drives recruitment of glycosylated substrates to Hrd1 in what is thought to be a mutually exclusive interaction [24,60]. There is no evidence of other E3 ligases interacting with Hrd3/Sel1L. The lack of redundancy within ERAD-L may be compensated for by other ER homeostasis pathways such as autophagy. Cells may also be less tolerant to accumulation of ERAD-M substrates and so have evolved redundancy in ERAD branches for processing of these substrates.

In mammals, there are more than 25 different ER-resident E3 ligases, 11 of which have already been implicated in ERAD [20]. Presumably, this expansion is attributed to larger genomes, multicellular organisation and longer lifespans of mammals, all of which increase the misfolding risk. Mammalian Hrd1 is also the only E3 ligase capable of processing lumenal substrates but redundancy between E3 ligases is very prevalent amongst other substrate classes [20]. For membrane substrates in mammalian ERAD, there is high redundancy between branches, especially for regulated proteolysis in sterol metabolism [43]. Mammalian Hrd1 associates with similar partners to its yeast counterpart, however some have no yeast homologues [61,62]. All substrates in eukaryotes converge on Cdc48/p97/VCP for membrane dislocation and this interaction

is enhanced by recruitment of Cdc48/p97/VCP to the ER membrane. In yeast, this is driven by the UBX domain of the membrane protein Ubx2 but in mammals, this is driven by the Ubx2 homolog UBXD8, other domains within ERAD factors like Derlin-1/2 and domains within E3 ligases [52,63–65]. Tomography studies in the algae *Chlamydomonas reinhardtii* have also shown proteasomes to localise with Cdc48 in specific patches around ER cristae, thereby suggesting the entire ERAD process is directly and locally coupled for high efficiency [66]. Whether this is also true for mammalian ERAD is unknown.

1.4. Substrate recognition in ERAD

1.4.1. Lumenal surveillance

All ERAD-L substrates are recruited to the Hrd1 E3 ligase complex by lumenal interacting partners [45]. A misfolded version of yeast carboxypeptidase (CPY*) was used to decipher the factors and mechanisms by which misfolded glycoproteins are recognised and transported across the ER membrane by the Hrd1 E3 ligase complex [31]. Hrd1 associates with ERAD factors on both sides of the membrane and in the ER lumen, it associates with Hrd3 and Yos9. The lectin domains of Yos9 and its mammalian orthologs OS-9 and XTP3-B recruit misfolded glycoproteins by recognition of N-linked glycans with a terminal α1,6-linked mannose residue [67]. Glycoproteins with prolonged residence times in the ER lumen are subject to slow-acting mannosidase enzymes [68]. Long residence time and a failure to mature from the ER is indicative of unsuccessful cycles of folding/refolding and serves as a signal for degradation. Trimming exposes a terminal α1,6-linked mannose residue that is recognised by Yos9/OS-9/XTP3-B [69]. Substrates are then delivered to Hrd3, or its mammalian homologue Sel1L [55]. Hrd3/Sel1L cooperate with Kar2/BiP to bind to unstructured polypeptides for delivery to Hrd1 and eventual ubiquitination [70,71]. However, certain mammalian substrates can

bypass this pathway and interact with Hrd3/Sel1L directly, suggesting a different mode of recognition [72]. Yos9/OS-9/XTP3-B appear to play no role in this pathway and instead in mammals, BiP, ERDJ5 and the Hrd1 partner HERPUD1 are involved [73].

1.4.2. Membrane surveillance

All membrane proteins in the ER theoretically have access to the TMDs of E3 ligases and so membrane protein quality control could occur within the membrane at E3s. From early ERAD studies, membranous “adapter proteins” were shown to drive recruitment of substrates for ERAD [74,75]. The first evidence for membrane degradation initiators came from studies on human cytomegalovirus (hCMV). US2 and US11 are hCMV gene products that induce degradation of the major histocompatibility complex class I (MHC-I) heavy chain, aiding host immune system evasion [74,75]. Both these proteins are type I membrane proteins that co-translationally insert into the ER, where they bind newly synthesised MHC-I and induce its degradation [76]. US2 and US11 are degraded by distinct ERAD pathways and they have been useful in dissecting mammalian ERAD pathways [77]. US2 interacts with MHC-I through its transmembrane domain (TMD) and recruits the E3 ligase TRC8 via its cytosolic tail [78,79]. The lumenal domain of MHC-I is then rapidly ubiquitinated implying that MHC-I must be partially retrotranslocated to expose its lumenal domains to the cytosolic ubiquitination machinery. Conversely, US-11 recognition of MHC-I occurs via its lumenal domain and degradation proceeds via an E3 ligase complex consisting of TMEM129, Derlin-1 and 2, VIMP and p97 [80].

ERAD is not exclusive to the degradation of misfolded proteins but it also plays a role in regulated proteolysis. Sterol metabolism is tightly controlled by regulated proteolysis of the multi-pass membrane proteins 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase and squalene monooxygenase (SM) [43,44]. Both these enzymes are key control points in sterol synthesis and are conserved across eukaryotes [81]. High sterol

synthesis results in binding of HMG-CoA with sterol intermediates via its sterol-sensing domain [82,83]. This causes conformational changes that recruit INSIG whose association with several E3s leads to degradation [84]. INSIG also binds sterol intermediates and in low sterol conditions, is subject to degradation by the E3 ligase Gp78 [85]. This also likely causes intramembrane conformational changes that serve as a signal for degradation. SM is degraded in a similar manner to HMG-CoA and its degradation by MARCHF6 is stimulated by high cholesterol concentrations [86,87]. A combination of SM association with cholesterol and increased stability of cholesterol-bound MARCHF6 trigger degradation [86,88]. All these pathways suggest a mechanism by which E3 ligases detect conformational changes in membrane proteins for ERAD initiation.

Recent work on yeast Hrd1 deciphered the role of different structural regions for ERAD-L and ERAD-M substrate degradation [89]. Site-directed mutagenesis of TMD1 identified mutants capable of ERAD-M but not ERAD-L. These mutations were not at interfaces of known Hrd1 interactors and so this suggests an ability of Hrd1 to surveille for different substrate classes. The transmembrane region of Hrd1 also contains a large number of hydrophilic amino acids [58]. Mutation of these Hrd1 sites resulted in ERAD-M defects, particularly in residues that clustered near the entrance of the lateral gate [58]. Together, this suggests a role of Hrd1 in direct recognition of substrates. Whether other E3 ligases possess this ability is unknown and will require precise biochemical and structural studies to confirm this.

1.5. Export of proteins from the ER

To reach the ubiquitination machinery, all soluble substrates in the ER lumen must be at least partially retrotranslocated to expose suitable sites for ubiquitination. Both soluble and membrane substrates need a mechanism to assist dislocation from the ER

membrane. These two processes are likely coupled, and recent biochemical and structural studies have shed light on this process.

1.5.1. The Hrd1-Der1 retrotranslocon

A path for substrate retrotranslocation was revealed by structures of the yeast Hrd1-Der1 heterodimer and monomeric Hrd1-Hrd3 [55]. For degradation of ERAD-L substrates, the yeast E3 ligase Hrd1 assembles a membrane protein complex containing Usa1, Der1 and Hrd3. Usa1 bridges the interaction between the Hrd1-Der1 heterodimer, where Hrd1 and Der1 each contribute a “half-channel” for substrate retrotranslocation (figure 1.3a) [55]. On the lumenal side of this complex, Hrd1 and Der1 form a hydrophilic cavity for substrate binding where the membrane is thinned at its centre, possibly to reduce the energetic barrier for polypeptide movement across the membrane [55]. This heterodimer is supported by site-specific crosslinks as well as previously established crosslinks between substrate and Hrd1 [90,91]. However, endogenous Hrd1 in both yeast and mammals is known to form higher order oligomers, some of which correspond to dimeric Hrd1 [62,91]. Prior to the Hrd1-Der1 structure, a dimeric structure of Hrd1-Hrd3 reconstituted in amphipol was solved [60]. This structure was later refuted since the assigned membrane boundaries in the later monomeric Hrd1-Hrd3 structure would result in soluble portions of Hrd3 in the dimeric structure entering the membrane. A steric clash would also be imposed between the two Hrd3 molecules if they were arranged the same as in the monomeric structure [55]. The Hrd1-Hrd3 dimer was suggested to be a purification artefact caused by the choice of detergent/amphipol but the propensity for Hrd1 to form dimeric complexes is consistent with other biochemical data [60]. Site-specific crosslinking of the endogenous yeast Hrd1-Der1 complex corroborated the heterodimeric structure, with and without a lumenal substrate [56,92]. However, Hrd1-Hrd1 crosslinks were also identified and these were shown to be independent of Der1

and Usa1 [56]. These crosslinks were introduced at the lateral gate of Hrd1 and could be observed in both the presence and absence of substrate.

Auto-ubiquitination of yeast Hrd1 is known to be a stimulator of retrotranslocation activity as well as a modulator of substrate recruitment to Hrd1 [93]. Lysine residues near the RING domain are auto-ubiquitinated (particularly the strongly conserved K373) and this was shown to drive pore opening when Hrd1 was reconstituted in lipid bilayers, suggesting some level of channel activity with Hrd1 alone [93]. Hrd1 structural studies have failed to answer how autoubiquitination might regulate activity. Cross-linking analysis of Hrd1 also revealed that Hrd1-Hrd1 crosslinks increased in the absence of UBC7, an E2 enzyme involved in ERAD, suggesting a role for ubiquitination in dimer dissolution [56]. This was also supported by an inactive Hrd1 RING mutant that increased Hrd1-Hrd1 crosslinking [56]. Together, the data imply a role for auto-ubiquitination in Hrd1 dimer dissolution and downstream assembly of other Hrd1 complexes. Conversely, mammalian Hrd1 does not have many lysine residues near its RING domain for potential auto-ubiquitination, nor has it been shown to auto-ubiquitinate. However, it shares the same strongly conserved lysine residue that is auto-ubiquitinated in yeast. This leaves unanswered questions as to how mammalian Hrd1 activity is regulated and whether mammalian ERAD mechanisms are similar to yeast.

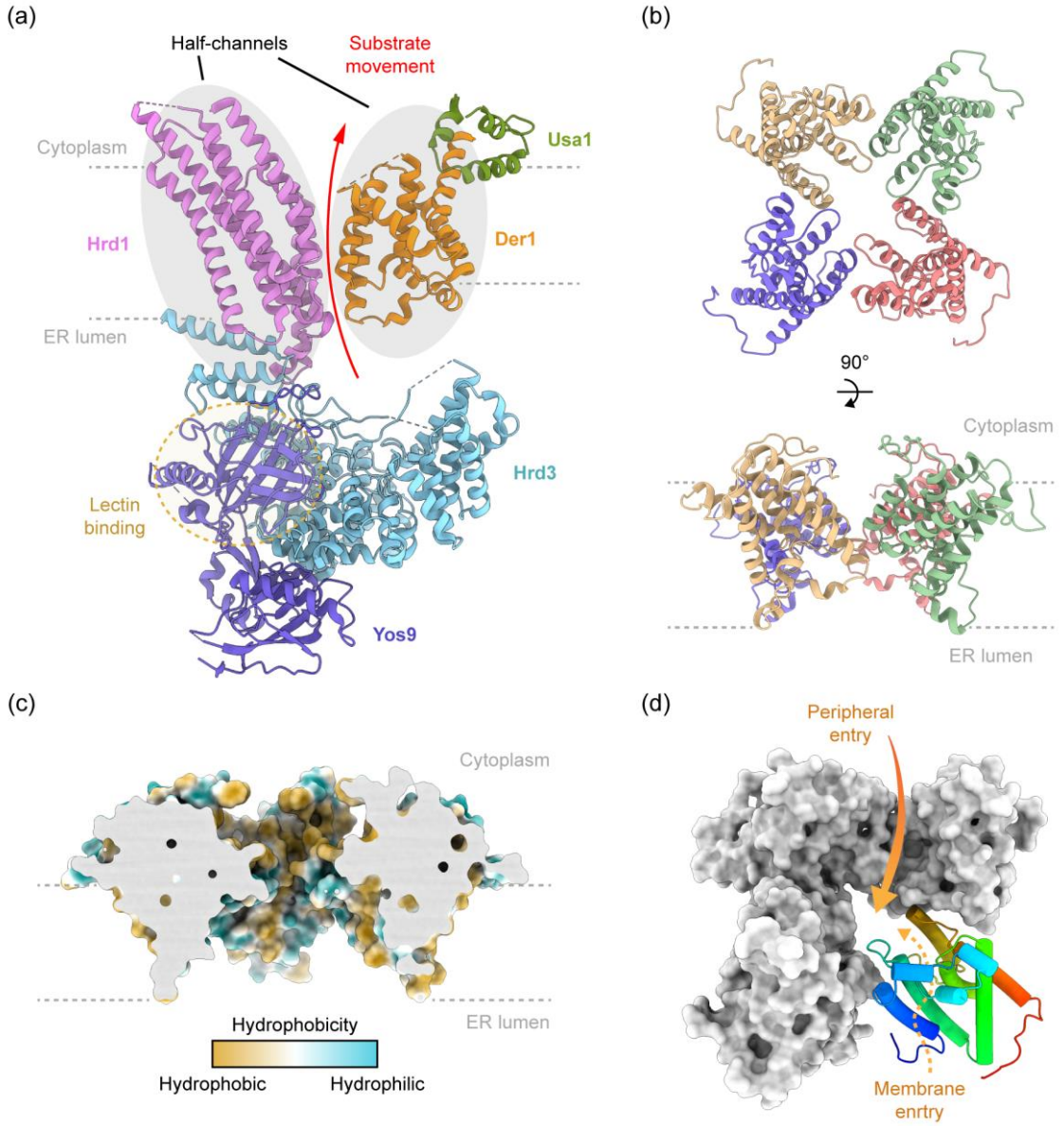


Figure 1.3. ER membrane export mechanisms in ERAD. (a) Cryo-EM structure of the Hrd1-Der1 retrotranslocon [55]. The position of Yos9 was fitted by alignment of the Hrd3-Yos9 and Hrd1-Usa1-Der1-Hrd3 models [55]. Misfolded substrates are recruited to the complex by Hrd3 and a lectin-binding domain in Yos9. Hrd1 and Der1 each contribute a proposed half-channel for substrate retrotranslocation. (b) Cryo-EM structure of the mammalian Derlin-1 tetramer [94]. Each protomer is coloured differently. Membrane boundaries were approximated based on the hydrophobicity of the TM helices. (c) Hydrophobicity surface plot for a cross-sectional view of the Derlin-1 tetramer, showing

a central hydrophobic pore. (d) Proposed substrate entry paths into the central pore of the Derlin-1 tetramer, adapted from Rao et al. [94]. Peripheral membrane protein entry at the monomer interfaces is proposed as this may be exposed to the lipid bilayer. A route for membrane substrate entry through the TMDs is also proposed based on the presence of “gating” helices [94].

1.5.2. Rhomboid pseudoproteases

Rhomboids are multi-pass membrane proteins, some of which are resident in the ER, and are capable of cleaving TMD substrates within the membrane [95]. Within the rhomboid superfamily is a class of rhomboids which are pseudoproteases and these lack the catalytic residues needed for proteolysis but retain the same overall structure [96]. Yeast Der1 is a rhomboid pseudoprotease that assembles with Hrd1 to assist retrotranslocation of ERAD-L substrates [55]. Rhomboids cleave scissile peptide bonds within α-helical TMDs and so their prevalence in ERAD may be related to their ability to partially unfold TMDs whilst accessing scissile bonds [65]. Degradation of yeast ERAD-M substrates occurs independent of Der1 and instead, Dfm1 associates with Hrd1 in a mutually exclusive but similar complex to drive ERAD-M [97,98]. Dfm1 is also a rhomboid pseudoprotease and is related to Der1 but it is unknown whether it also assembles as a heterodimer with Hrd1. Interestingly, Dfm1 was also shown to drive Hrd1-independent ERAD-M, by assisting other ERAD pathways such as Doa10 [98,99]. In mammals, the Der1 and Dfm1 homologues Derlin-2 and Derlin-1, respectively, are also known to associate with Hrd1 [100]. Derlin-2 is indispensable for ERAD-L whilst Derlin-1 is thought to be required for other substrate classes [65]. Recent structural work on the Dfm1 homologue Derlin-1 revealed that Derlin-1 can assemble into tetramers, which is consistent with previous data that Der1 and Dfm1 can self-associate (figure 1.3b-d) [94,97]. This tetramer has a central hydrophobic pore that is likely thinned at the

membrane (figure 1.3c). Peripheral membrane protein substrates are proposed to enter at the protomer interface whilst other membrane substrates are proposed to move through “gating” helices within a single protomer (figure 1.3d) [94]. There is no evidence to support these proposals yet. The significance of the tetramer is unknown but Dfm1 also contains a SHP box domain for recruitment of Cdc48 for substrate dislocation and so likely plays an integral part in this process [98,101]. More recent work has also shown Dfm1 and Derlin-1 to contain residues important for ERAD-M substrate recruitment [101]. Mutation of certain residues within loop 1 abolish ERAD-M and are required for substrate binding [101]. These residues are also conserved in catalytic rhomboids where they are used for substrate recognition, indicating conserved mechanisms between pseudoproteases and their catalytic ancestors [101]. The authors suggest that this region recognises integral membrane proteins and is poised to attract unstable TM helices, similarly to the bacterial rhomboid GlpG [102]. Association of rhomboid pseudoproteases with Hrd1 and other E3 ligases is a common theme amongst ERAD and could be a way to drive substrate selectivity. Curiously, rhomboids may also be involved in other quality control pathways, as evidenced by RHBDL4’s ability to rescue a substrate from ERAD and promote its maturation to the Golgi [103]. This active rhomboid was shown to not cleave its substrate and instead promotes ER-to-Golgi trafficking by an unknown mechanism, perhaps via a novel quality control pathway.

1.5.3. Other translocons

Other candidates with retrotranslocation activity are slowly emerging and many of these are associated with E3 ligases. Recent work on the transmembrane and ubiquitin-like domain-containing (TMUB) proteins has highlighted their role in ERAD [104–106]. TMUBs are small transmembrane proteins with a short cytosolic domain [107]. They contain ubiquitin-like domains (UBLs) and are known to associate with E3 ligase

complexes containing Gp78 or RNF185 [105,106]. Recent work showed that TMUB1 can engage membrane substrates and act as an escortase during retrotranslocation, prior to engagement of substrates by p97/VCP [104]. Through pulldowns and crosslinking, TMUB1 was shown to engage short TMD segments independently of other ERAD components [104]. Depletion of TMUB1 significantly reduced cytosolic substrate abundance. The UBL domain of TMUB1 stably associates with p97/VCP and was critical for substrate degradation. Interestingly, the authors show that TMUB1 may exist as a higher-order oligomer in membranes [104]. They propose that TMUB1 chaperones TMD segments immediately as they emerge from the membrane since p97/VCP is too large to engage close to thinned membranes [104]. Whether TMUBs act alone or in concert with other factors for retrotranslocation is unknown. The RNF185 E3 ligase complex is known to bind Membralin and either TMUB1 or TMUB2 [105]. This complex degrades membrane substrates but interestingly it contains the 8 TMD-containing protein Membralin, of unknown function. Membralin knockout is perinatal lethal in mice but it robustly scaffolds the other components of the complex and interacts with substrate [105,108]. Perhaps Membralin contributes retrotranslocation activity in concert with TMUBs. Other E3 ligases are also candidates for substrate retrotranslocation such as mammalian MARCHF6 and TRC8. Both these E3 ligases process TA proteins and the data suggest some level of substrate recognition within the membrane [109]. These E3s possess 14 and 13 predicted TMDs, respectively, and so may facilitate retrotranslocation of substrate together with p97/VCP [110].

1.6. Project aims

Structural work on ER-associated yeast E3 ligases is limited to Hrd1 and so the mechanisms by which other E3s operate is unknown. Furthermore, the precise mechanisms of substrate recognition and retrotranslocation remain elusive. Given that

no structures exist of any mammalian E3s with their associated membrane regions and partners, our knowledge of mammalian ERAD is further limited. There are still many unanswered questions regarding ERAD mechanisms generally and especially the mechanisms of mammalian E3s:

- What drives selection and recruitment of membrane substrates to E3 ligases during ERAD? Can E3 ligases directly recognise substrates or are there always intermediate steps?
- Are mammalian E3 ligases mechanistically equivalent to their yeast counterparts during ERAD?
- How do mammalian E3 ligases cooperate with their partners to facilitate retrotranslocation of substrates? Do they function similarly to the yeast Hrd1-Der1 translocon?
- What is the significance of higher-order oligomers for many ERAD factors such as TMUBs, Derlins and some E3 ligases? Do these represent regulatory states or other routes for substrate retrotranslocation?
- Do E3 ligases always act as translocons, as with the Hrd1 E3 ligase?

To explore these questions further, this thesis will present a combination of structural and biochemical approaches to study various mammalian ERAD complexes. A bottom-up screen for ERAD factors amenable to cryo-EM structure determination was initially performed and structures of the mammalian Hrd1 E3 ligase determined. A mammalian membrane protein expression system was developed and used to isolate complexes of Hrd1 with its endogenous membrane partners, for cryo-EM analysis. Attempts were also made to develop high-affinity binders to these complexes for high-resolution cryo-EM studies. The RNF185 E3 ligase complex was also isolated and analysed by cryo-EM.

2. Materials and Methods

2.1. Cloning & mutagenesis

Manufacturer's protocols were followed for DNA purification and DNA gel extraction kits, PCR amplifications, Gibson Assembly and restriction digestions, unless otherwise specified.

2.1.1. Mammalian expression vectors

Proteins of interest were expressed in mammalian cell lines. To generate these cell lines, lentiviral transfer plasmids were constructed for producing the gene of interest under stable expression [111]. Modifications were made to an existing mammalian expression system for stable expression of target genes [111]. Lentivirus are used to transduce HEK293S cells, a cell line adapted to tolerate the rapid transition from adherent to suspension culture [112]. For expression from stable cell lines, HEK293S TetR GnTi$^-$ cells were used. These cells constitutively express the repressor of the tetracycline response element (TetR), thereby preventing expression of sequences downstream of tetracycline operators (TetO) [113]. This repression is relieved by addition of tetracycline and allows for inducible expression of the downstream gene. N-acetylglucosaminyltransferase I (GnTI) activity is also removed from these cells by knockdown to restrict N-glycan formation, thereby increasing recombinant protein homogeneity [113].

The pHR-CMV-TetO_2-3C-mVenus-TwinStrep lentiviral transfer plasmid was modified to replace the 3C site with a TEV site, remove the RPTPσ signal sequence and increase the length of the domain-linking regions. Variations of this backbone were generated to encode for different N- and C-terminal tags (table 2.1). For isolation of multi-component complexes from stably expressing cell lines, this backbone was also modified to include

different antibiotic selection cassettes. With these, cell lines could be enriched for multiple transgenes. To clone these backbones, pHR-CMV-$TetO_2$-TEV-mVenus-TwinStrep was cut with the restriction enzyme SalI, and PCR fragments containing the human phosphoglycerate kinase (hPGK) promoter and different mammalian resistance cassettes were ligated in place with Gibson Assembly (NEB). This promoter is constitutive and so will provide a basal level of expression of the antibiotic resistance cassette for selection [114]. Transfer plasmids with puromycin, zeocin, geneticin and hygromycin selection cassettes were generated (table 2.1).

pHR-CMV-$TetO_2$	Tag Type	Resistance Cassette
TEV-mVenus-TwinStrep	C-terminal	N/A
6xHis-mCherry-TEV	N-terminal	N/A
TEV-TwinStrep	C-terminal	hPGK-Puromycin
TEV-6xHis	C-terminal	hPGK-Puromycin
TwinStrep-TEV	N-terminal	hPGK-Puromycin
TEV-TwinStrep	C-terminal	hPGK-Zeocin
TEV-6xHis	C-terminal	hPGK-Zeocin
TwinStrep-TEV	N-terminal	hPGK-Zeocin
3C-mVenus-TwinStrep	C-terminal	hPGK-Geneticin
3C-mVenus-TwinStrep	C-terminal	hPGK-Hygromycin

Table 2.1. List of lentiviral transfer plasmids constructed for inducible overexpression of proteins in HEK293S TetR GnTi$^-$ cells.

2.1.2. Cloning genes

Double-stranded DNA fragments of genes, codon optimised for expression in *Homo sapiens*, were ordered with 20-40 bp overhangs homologous to the pHR-CMV-$TetO_2$ transfer plasmid for insertion by Gibson Assembly (gBlocks, Integrated DNA Technologies) [115]. The lentiviral transfer vectors were usually prepared by restriction digestion with EcoRI and XhoI, and DNA fragments ligated in place with Gibson Assembly.

2.1.3. Mutagenesis

Mutations were introduced by Gibson Assembly in the pHR-CMV-TetO_2 vector. Primers were designed to amplify regions either side of the mutation site and 20-40 bp overhangs introduced to complement with the PCR fragments and backbone. The wildtype transgene was cut out of its backbone with EcoRI and XhoI, and PCR fragments encoding the mutant transgene were ligated in place with Gibson Assembly.

2.1.4. Restriction digestion

All restriction digestions were performed with FastDigest restriction enzymes (ThermoFisher), similarly to the manufacturer's instructions. Briefly, at least 1 µg of template DNA was mixed in a 20 µl reaction with 2 µl of 10x restriction buffer and no more than 10% (v/v) of restriction enzyme. This was incubated at 37°C for 1 hour with agitation before separating the cleavage products by agarose gel electrophoresis (as in 2.1.6.). DNA fragments were then purified with a gel extraction kit (GeneJet, ThermoFisher) before further use.

2.1.5. DNA assembly

DNA fragments were always assembled by Gibson assembly to form plasmids or longer DNA fragments. Backbones for DNA assembly were always restriction digested and these were assembled with PCR amplified sequences with Gibson Assembly Master Mix or NEBuilder HiFi DNA Assembly Master Mix (NEB). Usually, 1 µl of each DNA fragment was mixed together with an equal volume of assembly master mix. This was incubated at 50°C for 1 hour before transformation into *Escherichia coli* (as in 2.1.7.).

2.1.6. Agarose gel purification

DNA was mixed 5:1 with 6X loading dye (NEB) and loaded on either 1% or 2% (w/v) agarose dissolved in TAE buffer containing 40 mM Tris-HCl (pH 8.0), 1 mM EDTA, 20 mM glacial acetic acid and 1X CYBR Safe (Invitrogen). Gels were run for 20-40 minutes at 120 V in gel tanks containing TAE buffer without CYBR safe.

2.1.7. Bacterial transformation

Frozen aliquots of *E. coli* Turbo or DH5a cells were thawed on ice for 10 minutes. 20-50 µl of cells were added to ~100 ng of plasmid DNA and incubated on ice for 5 minutes. Cells were heat shocked at 42°C for 50 seconds and incubated on ice for a further 5 minutes. 150 µl of Luria broth (LB) was added and the mix incubated at 37°C for 45-60 minutes. Cells were plated on antibiotic resistance plates and grown overnight at 37°C.

2.1.8. Plasmid purification and sequencing

Most plasmids were purified from E. coli with miniprep kits following the manufacturer's instructions (GeneJet; ThermoScientific). Lentiviral packaging and envelope plasmids were purified from midiprep kits (GeneJet; ThermoScientific) to guarantee higher yield and purity.

2.2. Cell culture

HEK293S TetR GnTi$^-$ cells in adherent culture were usually maintained in Dulbecco's Modified Eagle Medium (DMEM; Gibco) supplemented with 10% FBS (Gibco), 1% glutamine (Gibco) and 1% penicillin-streptomycin (Gibco), unless otherwise specified. Adherent cell cultures were incubated at 37°C with 5% CO_2. Suspension cells were maintained in Freestyle 293 Expression Medium (Thermo Fisher) supplemented with 1%

FBS and 1% non-essential amino acids (NEAA; Gibco). These were grown at 37°C with 8% CO_2 and shaking between 90-130 RPM. Trypsinisations were performed with TrypLE Express Enzyme (Gibco) and inhibited by addition of an excess of DMEM or Freestyle media.

2.2.1. Transfection

Plasmid transfections were performed using Trans-LT1 (Mirus), according to the manufacturer's instructions. Briefly, the appropriate amount of plasmid was incubated with Trans-LT1 in Opti-MEM reduced serum media (Gibco) for 30 minutes. This was then added dropwise to the cell media and incubated for at least 24 hours with cells before cell passage.

2.2.2. Stable cell line generation

Stable cell lines were generated similarly to Elegheert et al. [111]. For consistent generation of high expression cell lines, this protocol was modified to guarantee greater than 1 transgene copy per cell. Transfection efficiency was improved by using TransIT-LT1 instead of polyethylenimine, which was known to have an increased transfection efficiency in mammalian cells [116].. The density of the cells to be transduced was also reduced. Both these changes increased the effective viral titre and the quantity of virus presented to each cell during transduction, thereby increasing transgene copy number. This usually removed the need to enrich expressing cells later by flow cytometry and resulted in higher protein yields of the same quality. The protocol was also adapted to T25 flasks to minimise waste and save time.

$2.5x10^6$ Lenti-X 293T cells (Takara Bio) were initially seeded in a T25 flask with 5 ml DMEM and grown to ~70-80% confluency over 24 hours. The next day, a transfection mix containing 2nd generation lentiviral packaging and envelope plasmids (psPAX2 and

pMD2.G, respectively), and transfer plasmid was prepared (table 2.2). 640 µl Opti-MEM was placed in a 1.5 ml tube, 2.13 µg of each plasmid added and the solution mixed thoroughly. 19.2 µl of TransIT-LT1 was added dropwise, taking care to avoid contact with the walls of the tube. This solution was mixed carefully by pipetting and incubated at room temperature for 20-30 minutes. Before transfection, the Lenti-X 293T media was refreshed with 5 ml DMEM supplemented with 2% FBS, 1% glutamine and 1% penicillin-streptomycin. Transfection mix was applied dropwise to the media and the flask incubated for 72 hours. 1 day prior to maturation of the lentivirus flask, $3x10^6$ HEK293S TetR GnTi$^-$ cells (kindly provided by Christian Siebold, Oxford; hereafter referred to as HEK293S cells) were seeded in a T25 flask with 5 ml DMEM. After 24 hours, the HEK293S cells were ~50% confluent and ready for transduction. The lentiviral culture media was harvested, filtered through a 0.45 µm cellulose acetate filter and topped up with 2 ml fresh DMEM. 10 µg/ml of polybrene (Santa Cruz Biotechnology) was added and the HEK293S cell media replaced with the lentiviral media. After 72 hours incubation with lentivirus, the media was aspirated and the cells passaged. The cells were left to recover for a further 2 days and were maintained in DMEM hereafter.

Reagent	Final Concentration	Volume
psPAX2	2.13 µg	Variable
pMD2.G	2.13 µg	Variable
pHR-CMV-TetO2	2.13 µg	Variable
OptiMEM	N/A	640 µl
TransIT-LT1	3 µl per 1 µg DNA	19.2 µl

Table 2.2. Transfection mix for generating stable cell lines.

For vectors encoding a fluorescent recombinant protein, expression was induced with 10 µg/ml of doxycycline for ~16 hours and cell homogeneity assessed by flow cytometry (see 2.4.1 for preparation of cells). If less than 90% of cells were GFP positive,

fluorescent cells were enriched by FACS before downstream expression and purification. However, protein yields were generally better when the cell line was instead recreated.

For vectors encoding an antibiotic resistance cassette, transduced cells were enriched by antibiotic selection following lentiviral transduction. For puromycin-based selection, 2 μg/ml of puromycin was incubated with cells for 2 days. For zeocin-based selection, 100 μg/ml of zeocin was incubated with cells for at least 5 days. For geneticin-based selection, 400 μg/ml of geneticin was incubated with cells for at least 10 days. If cells reached confluency during selection, cells were passaged into fresh media with the appropriate antibiotic. A selection control with non-transduced HEK293S cells was always performed and selection time adjusted if the selection was slower than anticipated.

For expression of multi-component complexes from stable cell lines, HEK293S cells were first transduced with a transfer vector containing a fluorescently tagged transgene without an antibiotic resistance cassette. If the transduction event was good quality (>1 copy per cell as assessed by flow cytometry), these cells were further transduced with lentiviruses containing transfer vectors of the other complex components. Subsequent vectors for lentiviral transduction encoded antibiotic selection markers and positive transducants were enriched by consecutive antibiotic selection. Sometimes mixed lentiviral populations were generated by mixing different transfer plasmids during transfection. This was done when the transfer plasmids were of similar sizes and so would have similar transfection efficiencies.

2.2.3. Suspension culture

HEK293S suspension cultures were maintained in Freestyle media and cells handled similarly to Elegheert et al. [111]. For increased protein yields, protein expression was induced when the cells reached higher densities than those stated in the paper. Briefly,

HEK293S cells were transitioned to suspension culture by expansion to 1-2x T175 flasks and grown to ~80% confluency in DMEM. Cells were trypsinised with 5 ml TrypLE Express per T175 flask for 10 minutes and 25 ml of Freestyle media added to resuspend the cells and inhibit trypsin activity. Cells were then diluted to $0.5x10^6$ viable cells/ml with Freestyle media. These were transferred to polycarbonate erlenmeyer flasks with vent lids (Corning) and incubated in a shaking incubator. Once cells reached ~$2x10^6$ viable cells/ml, cells were diluted to $0.5x10^6$ cells/ml and the process repeated until the required culture scale was reached. Protein expression was induced with 10 µg/ml of doxycycline once the cell density exceeded $3x10^6$ viable cells/ml.

2.2.4. C-terminal CRISPR tagging

CRISPR tagging was performed by co-transfection of a Cas12a-encoding plasmid and a PCR cassette, following an already published protocol but with some adjustments [117]. The PCR tagging vector pMaCTag-Z07 was modified to include a TEV-mVenus-TwinStrep tag. The antibiotic resistance cassette was also replaced with a geneticin resistance cassette so that the vector could be used to tag genes within Flp-In T-REx cells. These cells already contain a zeocin resistance cassette and cell lines were to be validated using puromycin-resistant CIRSPR-Cas9 knockout plasmids. The homology fragment for CRISPR was generated by PCR with M1 and M2 primers designed according to Fueller et al. [117]. 2x 50 ul PCR reactions were set up to amplify the PCR cassette (Table 2.3) and the PCR products gel purified.

Component	Volume (µl)
5X Q5 Reaction Buffer	10
10 mM dNTPs	1
10 µM M1	2.5
10 µM M2	2.5
Template DNA	1
Q5 polymerase	0.5
Betaine	5
Nuclease-free water	27.5

Table 2.3. PCR reaction mixture for amplification of the tagging cassette.

150,000 Flp-In T-REx were seeded in a single well of a 24-well plate in 900 µl of DMEM. After ~24 hours incubation and when the cells reached greater than 50% confluency, cells were transfected with pY230 and the PCR tagging cassette with homology arms (according to 2.2.1). After 48 hours incubation with the transfection mix, cells were passaged and refreshed into DMEM supplemented with 400 µg/ml of geneticin. Cells were selected for 10 days and GFP positive cells were single cell sorted into 2x 96-well plates containing 200 µl of DMEM with 20% FBS. Single cell clones were left to grow out for ~2 weeks or until the media began changing colour. Viable clone populations were validated by CRISPR-Cas9 mediated knockdown of the respective gene with an appropriate guide RNA [80]. Clones were then assessed for knockdown of the knocked-in fluorescent tag by flow cytometry and western blotting.

2.3. Protein expression and purification

Proteins were generally purified according to the following protocols. Any adaptations are stated in the proceeding chapters. All purification steps were carried at 4°C unless otherwise stated.

2.3.1. Mammalian protein expression

Recombinant mammalian proteins were expressed from HEK293S cells stably expressing protein under control of a tetracycline operator. Once the density of the suspension culture exceeded $3x10^6$ viable cells/ml, expression was usually induced for 24-48 hours by addition of 10 µg/ml of doxycycline. Cells were harvested by centrifugation at 1,500 g for 10 minutes at 4°C. Pellets were washed with Dulbecco's phosphate-buffered saline (DPBS; Gibco) and stored at -80°C until further use.

2.3.2. Mammalian membrane protein purification

Recombinant mammalian membrane proteins were generally purified as follows. These protocols were optimised over the course of the project.

2.3.2.1. Lysis and membrane preparation

Frozen cells were thawed at room temperature and resuspended in ~20 ml lysis buffer per 1 L equivalent of cell pellet, containing 100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA and complete protease cocktail inhibitor (Roche). Lysis was performed by passing the cells 5 times through an EmulsiFlex-C5 (Avestin) with at least 12,000 PSI of backpressure. Cellular debris was pelleted by centrifugation at 12,000 g for 20 minutes at 4°C. The supernatant was then spun at 100,000 g for 1 hour at 4 °C to pellet membranes.

2.3.2.2. Membrane solubilisation

Membranes were resuspended in resuspension buffer containing 100 mM Tris-HCl (pH 8.0), 150 mM NaCl and 1 mM EDTA supplemented with high concentrations of detergent (table 2.4). The ratio of detergent to protein was determined by membrane

mass. For mammalian membranes, 10 ml of 1% detergent buffer per 1 g of membranes was used for solubilisation. Solubilisation was performed at 4°C for 2 hours with agitation. Unsolubilised material was pelleted by centrifugation at 100,000 g for 30 minutes at 4°C. The supernatant was collected for downstream purification.

Detergent	Solubilisation Concentration (w/v)
DDM	1%
DM	1%
LMNG	1%
DMNG	1%
GDN	1.50%

Table 2.4. Detergent concentrations during solubilisation of membrane proteins.

2.3.2.3. TwinStrep purification

Solubilised membrane proteins were applied to either StrepTactin XT or StrepTactin XT 4Flow pre-packed columns (IBA) equilibrated with wash buffer containing 100 mM Tris-HCl (pH 8.0), 150 mM NaCl and detergent at ~3x its CMC. Sample was recycled through the resin 2x before washing the column with 6-8 column volumes (CVs) of wash buffer. Protein was eluted with 6 CVs of elution buffer containing 100 mM Tris-HCl (pH 7.4), 150 mM NaCl, 55 mM D-biotin and detergent at ~3x its CMC.

2.3.2.4. Dual affinity purification

For enrichment of multi-component protein complexes, TwinStrep purified material was subject to a second round of purification with Ni-NTA resin. StrepTactin eluted material was incubated with 0.5-1 ml Ni-NTA resin (HisPur; Thermo Fisher) per 1 L equivalent of cell culture. This resin was pre-equilibrated in Ni-NTA wash buffer containing 100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 20 mM imidazole and detergent at

~3x its CMC. Protein was incubated with the resin overnight at 4°C with agitation. The following day, the resin was washed with 10 CVs of Ni-NTA wash buffer and eluted with 5 CVs of 100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 300 mM imidazole and detergent at ~3x its CMC.

2.3.2.5. Tag cleavage

Purified protein was incubated overnight at 4°C with purified TEV protease at an A280 ratio of 1:10 TEV to protein. If the cleavage was performed following Ni-NTA purification, the material was additionally dialysed during the TEV reaction into 2 L of 100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 20 mM imidazole and detergent at ~3x its CMC, inside of 10 kDa MWCO snakeskin dialysis tubing (Sigma). TEV protease was removed by reverse His-tag purification by incubation with a small quantity of Ni-NTA resin.

2.3.2.6. Size-exclusion chromatography

Protein complexes were further polished by size-exclusion chromatography (SEC). Purified material was concentrated in a 100 kDa MWCO vivaspin (Sartorius) to less than 500 µl and loaded on a Superose 6 Increase 10/300 GL equilibrated with 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and detergent at ~3x its CMC. This was carried out at room temperature with room temperature buffers. Elution fractions were assessed for purity by SDS-PAGE.

2.3.2.7. Fluorescence-based size-exclusion chromatography

To detect protein fluorescence during SEC, fluorescence readings were recorded using an inline fluorimeter. For mVenus fluorescence, 515 nm and 527 nm were used for the excitation and emission wavelengths, respectively [118]. For mCherry fluorescence,

587 nm and 610 nm were used for the excitation and emission wavelengths, respectively [119]. The sensitivity and gain were adjusted accordingly.

2.3.2.8. Liposome reconstitution

Proteoliposomes were prepared similarly to Newstead et. al. [120]. Hrd1 complexes for liposome reconstitution were purified in phosphate buffer (pH 7.4) containing 37.7 mM Na_2HPO_4, 12.3 mM NaH_2PO_4 and 100 mM NaCl. All complexes were reconstituted into liposomes containing a 3:1 mix of POPE:POPG (Avanti). Chloroform was evaporated from the lipids with a rotary evaporator and the lipids washed 2 times with pentane. Dried lipids were resuspended to 10 mg/ml in phosphate buffer. Lipids were then subject to 2 rounds of freeze-thaw with liquid nitrogen. Prior to reconstitution, lipids were extruded through a 0.8 µm filter and then through a 0.4 µm filter. This was carried out above the phase transition temperature of the lipids. To reconstitute Hrd1 and Hrd1 complexes into liposomes, purified protein was added to liposomes at a ratio of 1:50. Additional detergent stock was added to maintain a detergent concentration of ~3x CMC. This was incubated at room temperature for 1 hour and then on ice for an additional 1 hour. To remove the detergent, Bio-Beads (Bio-Rad) were added to 75% of the total reaction volume. Bio-Beads were incubated with the sample for 2 hours at 4°C with agitation. Bio-Beads were then removed, and fresh beads added to 75% of the total reaction volume. These were incubated overnight at 4°C with agitation. The following day, the supernatant was collected and diluted to a larger volume with phosphate buffer. Liposomes were harvested by centrifugation at 120,000 g for 40 minutes at 4°C. These were resuspended to a final protein concentration of 1 mg/ml and subject to 3 rounds of freeze-thaw cycles with liquid nitrogen, before storage at -80°C.

2.3.2.9 Biotinylation of mammalian membrane proteins

Purified proteins and complexes either before or after SEC were biotinylated with the biotinylation reaction kit (Avidity), according to the manufacturer's instructions. Briefly, purified membrane proteins were concentrated to < 500 µl in a 100 kDa MWCO vivaspin (Sartorius). Buffer A and Buffer B were diluted 10x with the purified protein, and 2.5 µg of biotin added per 10 nmol of recombinant protein. Detergent stock was added to any buffers without detergent, to maintain at least 3x CMC. The reaction was incubated overnight at 4°C with agitation. Excess biotin, ATP and BirA were removed by separation on a Superose 6 Increase 10/300 GL equilibrated in 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and detergent at ~3x CMC.

2.3.2.10. Amphipol reconstitution

Protein was always purified by affinity purification prior to amphipol reconstitution. In some cases, SEC was also performed. Reconstitution took place by broadly following other reconstitution protocols for small membrane proteins [60,121]. Purified protein was usually concentrated to > 3 mg/ml in a 100 kDa MWCO vivaspin. Stocks of 100 mg/ml A8-35 (Anatrace) or 10% PMAL C8 (Anatrace) were made with ultrapure water. Amphipol exchange reactions were set up with a range of amphipol to protein mass ratios. This was brought to a final volume of 400 µl and additional detergent added to maintain at least 3x the detergent CMC. After incubation for 2 hours with agitation at 4°C, 70% (w/v) Bio-Beads were added and incubated for an additional 2 hours. The supernatant was collected and incubated with fresh Bio-Beads at 70% (w/v) overnight at 4°C. The following day, the Bio-Beads were removed and insoluble material pelleted by centrifugation at 15,000 g for 30 minutes at 4°C. The supernatant was then injected immediately over a SEC column as in 2.3.2.6.

2.3.2.11. Styrene maleic-acid lipid particle (SMALP) reconstitution

Mammalian cells were lysed and membranes prepared as in 2.3.2.1 and 2.3.2.2. SMALP reconstitution was carried out using the guidelines published by Cube Biotech [122]. Membranes were resuspended to 30 mg/ml with resuspension buffer containing 100 mM Tris-HCl (pH 7.5), 150 mM NaCl and 1 mM EDTA. 1.5 ml of resuspended membranes was added directly to tubes containing 50 mg of lyophilised SMALP. Solubilisation occurred overnight by incubation at 4°C with agitation. The following day, the solubilisation mix was diluted by a factor of 10 with resuspension buffer. Unsolublised material was cleared by centrifugation at 70,000 g for 1 hour at 4°C. The supernatant was then used in pulldowns or injected immediately over a SEC column.

2.3.3. Mammalian soluble protein purification

Soluble proteins expressed in HEK293S cells were always TwinStrep tagged. Frozen cells were thawed at room temperature and resuspended in ~20 ml lysis buffer per 1 L equivalent of cell pellet, containing 100 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA and complete protease cocktail inhibitor (Roche). Lysis was performed by passing the cells 5 times through an EmulsiFlex-C5 (Avestin) with at least 12,000 PSI of backpressure. Cellular debris was pelleted by centrifugation at 35,000 g for 45 minutes at 4°C. The supernatant was filtered through a 0.4 µm filter before being passed 2x through either StrepTactin XT or StrepTactin XT 4Flow pre-packed columns equilibrated with wash buffer containing 100 mM Tris-HCl (pH 8.0) and 150 mM NaCl. Columns were washed with 6-8 CVs of wash buffer before elution with 6 CVs of elution buffer containing 100 mM Tris-HCl (pH 7.0), 150 mM NaCl, 55 mM biotin and 20 mM imidazole. Tags were then cleaved overnight by incubation with purified TEV protease at a 1:10 protein to TEV ratio. TEV protease was then removed by reverse-His tag purification before polishing the sample further by SEC. Protein was concentrated in either a 10 kDa or 100 kDa

MWCO vivaspin before being injected on either a Superose 6 Increase 10/300 GL or Superdex 200 Increase 10/300 GL equilibrated in 50 mM Tris-HCl (pH 7.4) and 150 mM NaCl. Clean elution fractions were pooled, concentrated in a vivaspin and flash frozen in liquid nitrogen. Frozen protein was stored at -80°C.

2.3.4. Bacterial protein expression

For bacterial expression, genes of interest were cloned into pET-28b backbones (Addgene), unless otherwise stated. Plasmids were generally transformed into *E. coli* BL21 or *E. coli* Shuffle (NEB). A scoop of transformants were used to inoculate 10 ml of LB containing 30 ug/ml of kanamycin, and this was incubated overnight at 37°C with agitation. The following day, 1 L of TB supplemented with 30 ug/ml of kanamycin was inoculated with the overnight preculture and grown at 37°C until OD600 0.5-0.8 reached. The culture was removed from the incubator and left on the bench for 30 minutes to cool down. Protein expression was induced with 0.5-1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and cultures incubated at 18°C overnight with agitation. Cells were harvested the following day by centrifugation at 4,000 g for 20 minutes at 4°C. Pellets were washed with distilled water and frozen at -20°C.

2.3.5. Bacterial protein purification

Cell pellets were resuspended in ~20 ml of lysis buffer containing 100 mM Tris-HCl (pH 7.0), 300 mM NaCl, 20 mM imidazole and complete protease inhibitor, per 1 L equivalent of TB culture. Lysis was performed by passing the suspension 5 times through an EmulsiFlex-C5 with at least 12,000 PSI of backpressure. Debris was pelleted by centrifugation at 35,000 g for 45 minutes at 4°C. The supernatant was filtered through a 0.4 um filter and applied to 2 ml Ni-NTA resin per 1 L equivalent of TB culture, equilibrated with lysis buffer. This was incubated for 1 hour at 4°C with agitation before washing the

resin with 10 CVs of lysis buffer. Protein was eluted with 5 CVs of lysis buffer containing 400 mM imidazole and concentrated to < 5 ml in a 10 kDa MWCO vivaspin (Sartorius). This was further polished on a Superdex 200 16/60 HiLoad equilibrated in 50 mM Tris-HCl (pH 7.0) and 150 mM NaCl. Clean elution fractions were pooled, concentrated in a vivaspin and flash frozen in liquid nitrogen before storage at -80°C.

2.3.6. Detergent screening

200 ml of HEK293S cells stably expressing recombinant protein were grown to greater than $3x10^6$ viable cells/ml in Freestyle media. Expression was induced with 10 µg/ml of doxycycline for 48 hours. Cells were harvested and membranes isolated (as in 2.3.1 and 2.3.2.1.). Membranes were resuspended in 5 ml of resuspension buffer containing 100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA and complete protease inhibitor. This was then flash frozen in liquid nitrogen and stored at -80°C until further processing. The membrane resuspension was thawed at room temperature, homogenised in a tissue grinder and aliquoted by 0.5 ml. For each detergent condition, 55 µl of the following detergent stocks was added to bring the detergent concentration to the final solubilisation concentration (table 2.5).

Detergent	Stock Concentration (w/v)	Solubilisation Concentration (w/v)
DDM	10% DDM	1% DDM
DDM-CHS (10:1)	10% DDM, 1% CHS	1% DDM, 0.1% CHS
DDM-CHS (5:1)	10% DDM, 2% CHS	10% DDM, 0.2% CHS
LMNG	10% LMNG	1% LMNG
LMNG-CHS (10:1)	10% LMNG, 1% CHS	1% LMNG, 0.1% CHS
LMNG-CHS (5:1)	10% LMNG, 2% CHS	1% LMNG, 0.2% CHS
DMNG	10% DMNG	1% DMNG
DMNG-CHS (10:1)	10% DMNG, 1% CHS	1% DMNG, 0.1% CHS
GDN	15% GDN	1.5% GDN

Table 2.5. Solubilisation conditions for detergent screening.

Membranes were solubilised for 1 hour at 4°C with agitation. Unsolubilised material was pelleted by centrifugation at 100,000 g for 30 minutes at 4°C. 100 µl of supernatant was then loaded over a Superose 6 Increase 10/300 GL equilibrated with 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.03% DDM and 0.006% CHS. Fluorescence readings were recorded to visualise recombinant protein.

2.3.7. Small-scale sybody expression and purification

Sybodies (synthetic nanobodies) were expressed and purified from small scale cultures according to Seeger et al. [123]. *E. coli* MC1061 (NEB) were transformed with pBXNPH3 or pBXNPHM3 backbones containing sybodies. These were plated on LB-agar plates containing 25 µg/ml chloramphenicol and left to grow overnight at 37°C. 1 ml of TB, supplemented with 25 µg/ml chloramphenicol, was inoculated with a scoop of colonies from the overnight transformants. This was grown for a few hours at 37°C with agitation. 50 ml of TB supplemented with 25 µg/ml chloramphenicol was then inoculated with this culture. This was grown at 37°C for an hour before dropping the temperature to 22°C. Cultures were grown until an OD600 0.4-0.8 was reached, before overnight induction with L-(+)-arabinose at a concentration of 0.02% (w/v). Cells were harvested by centrifugation at 4,000 g for 20 minutes.

To extract the sybodies, pellets were resuspended in ~5 ml of periplasmic extraction buffer containing 100 mM Tris-HCl (pH 8.0), 20% sucrose and 0.5 mM EDTA. This was incubated at 4°C for 1 hour with agitation. 20 ml of a second periplasmic extraction buffer containing 25 mM Tris-HCl (pH 8.0), 5% sucrose and 5 mM $MgCl_2$, was added and incubated with agitation at 4°C for 20 minutes. Cells were pelleted by centrifugation at 4,000 g for 20 minutes at 4°C. The supernatant was topped up with buffer stocks to bring the final buffer to 25 mM Tris-HCl (pH 8.0), 200 mM NaCl and 15 mM imidazole. 100 µl

Ni-NTA resin was added and incubated for 1 hour at room temperature with agitation. The resin was washed with 500 µl wash buffer containing 20 mM Tris-HCl (pH 8.0), 200 mM NaCl and 15 mM imidazole. This was further washed with 1500 µl wash buffer containing 25 mM imidazole. Protein was eluted with 500 µl wash buffer containing 300 mM imidazole. The 500 µl eluate was further polished by SEC. 500 µl Ni-ETA eluate was loaded onto a Sepax SRT-10C SEC100 column equilibrated in 20 mM Tris-HCl (pH 8.0), 150 mM NaCl. High purity SEC fractions were pooled and concentrated in a 3 kDa MWCO vivaspin before immediate binding with the target protein.

2.4. Biochemical methods

2.4.1. Flow cytometry

Cells were trypsinised and the trypsin inhibited with flow cytometry buffer (DPBS supplemented with 1 mM EDTA and 1% FBS). Cells were pelleted at 500 g for 5 minutes at room temperature and resuspended in 2 ml flow cytometry buffer. Data were collected on an LSRFortessa X20 equipped with 561 and 640 nm lasers (BD Biosciences).

2.4.2. SDS-PAGE and western blotting

3x loading buffer, containing 188 mM Tris-HCl (pH 6.8), 3% SDS, 30% glycerol, 0.01% bromophenol blue and 150 mM DTT, was added to samples. Membrane protein samples were incubated at 37°C for 15 minutes whilst soluble protein samples were incubated at 65°C for 15 minutes. All SDS-PAGE samples were run with 4-20% Criterion TGX Precast Gels (Bio-Rad) in running buffer containing 25 mM Tris base, 200 mM glycine and 3.5 mM SDS. Gels were either western blotted or stained with Coomassie blue (Expedeon or Lubio).

For western blotting, samples were run by SDS-PAGE and transferred onto polyvinylidene difluoride membranes (Bio-Rad). Membranes were blocked for 1 hour in 3% (w/v) BSA or 3% (w/v) milk in PBS-T (0.1% Tween20 in 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl and 2.7 mM KCl). Primary antibodies were diluted into PBS-T containing 1% (w/v) BSA or 1% (w/v) milk and incubated with membranes overnight at 4°C. Membranes were washed three times with PBS-T. Secondary antibody was diluted into 1% (w/v) BSA or 1% (w/v) milk and incubated with the membranes for 1 hour at room temperature. Blots were washed three times with PBS-T and developed with Western Lightning Plus (PerkinElmer). Primary and secondary antibodies used in this thesis are shown in table 2.6.

Type	Antigen	Organism	Source	Catalog Number
Primary	Hrd1	Rabbit	Cell Signalling	14773
Primary	HERPUD1	Rabbit	Abcam	ab150424
Primary	Sel1L	Rabbit	Abcam	ab78298
Primary	Derlin-2	Rabbit	MyBioSource	MBS9612470
Primary	His	N/A - IMAC-based	Thermo Fisher Scientific	15165
Primary	mNeonGreen	Mouse	Chromotek	32F6-20
Primary	GFP	Mouse	Chromotek	3H9-100
Primary	FLAG	Mouse	Sigma	F1804
Primary	Membralin	Rabbit	Atlas Antibodies	HPA042669
Primary	RNF185	Rabbit	Abcam	ab181999
Primary	TMUB1	Rabbit	Abcam	ab180586
Primary	TMUB2	Rabbit	ProteinTech	28044-1-AP
Secondary	Rabbit IgG	Mouse	Jackson ImmunoResearch	211-032-171
Secondary	Rat IgG	Goat	Jackson ImmunoResearch	112-035-175
Secondary	Mouse IgG	Goat	Jackson ImmunoResearch	115-035-174

Table 2.6. Primary and secondary antibodies used in western blotting.

2.4.3. Small-scale non-denaturing pulldowns

Cells were seeded in 6-well plates and brought to 80% confluency in DMEM media. For cell lines made from pHR transfer vectors, protein expression was induced with 10 µg/ml doxycycline for 48 hours. For cell lines made from pUC19 vectors, protein

expression was induced with 0.1 µg/ml tetracycline overnight. Membrane proteins for FLAG pulldowns were solubilised by addition of 1.5 ml of solubilisation buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, protease inhibitor, 1% DMNG and 250 units/ml of benzonase (Sigma). Membrane proteins for strep-tag pulldowns were solubilised with the same buffer similarly but at pH 8.0. Solubilisation was carried out for 1 hour at 4°C with agitation. Insoluble material was pelleted by centrifugation at 20,000 g for 20 minutes at 4°C.

For FLAG pulldowns, the supernatant was incubated with 20 µl anti-FLAG M2 beads (Sigma) equilibrated in wash buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA and 0.02% DMNG. Binding took place for 30 minutes at 4°C with agitation. Beads were washed three times with 10 bead volumes of wash buffer before protein elution with 5 bead volumes of elution buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 0.02% DMNG and 500 ng/ml 3xFLAG peptide (Sigma). 3x SDS-PAGE loading buffer was then added and samples analysed by SDS-PAGE and western blotting.

For strep-tag pulldowns, the supernatant was incubated with 60 µl MagStrep "type3" XT beads (IBA) equilibrated in wash buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA and 0.02% DMNG. Binding took place for 30 minutes at 4°C with agitation. Beads were washed three times with 5 bead volumes of wash buffer before protein elution with elution buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, 0.02% DMNG and 60 mM biotin. 3x SDS-PAGE loading buffer was then added and samples analysed by SDS-PAGE and western blotting.

GFP pulldowns were performed similarly to the FLAG pulldowns except that GFP-trap slurry (Chromotek) was used instead of anti-FLAG M2 beads. 50 µl beads were used per well of a 6-well plate and protein was eluted with 1x SDS-PAGE loading buffer.

2.4.4. Analysis of sybody binding by pulldown

This pulldown protocol was adapted from a protocol established by Simon Newstead and his group, based at the University of Oxford (unpublished). 25 µl of target protein at 0.2 mg/ml was mixed with 25 µl of purified sybody at 0.1 mg/ml in 20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.03% DDM and 0.003% CHS. After incubation at 4°C for 1 hour, 12 µl of HisPur Ni-NTA resin was added and the imidazole concentration brought to 15 mM. This was incubated for 30 minutes with agitation and washed twice with 500 µl of the same buffer containing 15 mM imidazole. A third wash was then performed but this was incubated with the resin for 45 minutes before removal, in an attempt to dissociate weaker binders. Protein was eluted by incubation of the beads with 5 µl of 1 M imidazole and 10 µl of 3x SDS-PAGE loading dye.

2.4.5. Analysis of sybody binding by SEC

This protocol for assessment of sybody binding by SEC was also adapted from one established by Simon Newstead's group. 40 µg of the target protein was mixed with sybody in a 1:1.5 molar ratio in a 100 µl reaction. Detergent stock was added to bring the CMC to at least 3x. Control reactions were also set up at the same concentration. After incubation at 4°C for 30 minutes, the reaction mix was injected on a Superose 6 Increase 10/300 GL equilibrated in 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.03% DDM and 0.003% CHS.

2.4.6. Negative stain-EM

Holey carbon-coated grids (Quantifoil 300 mesh, Cu 1.2/1.3) were glow-discharged for 30 seconds at 15 mA. 5 µl of protein sample at varying dilutions was applied and incubated for 2 minutes. Excess buffer was blotted away, and the grids washed twice in 20 µl of ultrapure water with blotting in between. Grids were then washed with 20 µl of

uranyl acetate, blotted and incubated with 20 µl uranyl acetate for 30 seconds. Grids were blotted once more to remove excess uranyl acetate and left to air dry for a few minutes before storage.

2.4.7. TEV protease protection assay

The insertion of Hrd1 into liposomes was assessed by flotation in a Nycodenz gradient and orientation assayed by protection from TEV protease [93]. 50 ul of proteoliposomes suspension at 1 mg/ml was mixed with 50 µl of 60% (w/v) Nycodenz (MP Biomedicals Inc) prepared in phosphate buffer (pH 7.4) containing 37.7 mM Na_2HPO_4, 12.3 mM NaH_2PO_4 and 100 mM NaCl. This was placed at the bottom of a 1 ml ultracentrifuge tube and layered with 150 µl of each 30%, 20%, 10% and 0% Nycodenz solution prepared in phosphate buffer. The gradient was spun at 70,000 g inside of tube adapters in an SW55 Ti (Beckman) for 6 hours at 4°C. Six fractions were collected by pipetting from the top of the gradient and were analysed by SDS-PAGE and western blotting.

100 µl Nycodenz fractions containing the floated liposomes were incubated with 10 µg of purified TEV protease. TEV was incubated in the presence of 0%, 0.5% or 1% (w/v) of DDM or LMNG to partially solubilise liposomes. Reactions were incubated at room temperature for 1 hour before quenching with SDS-PAGE loading buffer. The cleavage extent was then analysed by SDS-PAGE.

2.4.8. Circular dichroism (CD)

CD measurements were performed on J-1500 CD Spectrophotometer (Jasco). Measurements took place in 15 mM Na_2HPO_4, 5 mM NaH_2PO_4 (pH 7.5), 50 mM NaF, 0.03% DDM and 0.003% CHS. Melting took place between 25°C and 95°C. A buffer blank was also recorded for normalisation.

2.4.9. Nano-differential scanning fluorimetry (nanoDSF)

nanoDSF measurements were performed on a Prometheus NT.48 (PR001) with Back-reflection Optics (NanoTemper Technologies GmbH). Tyrosine and tryptophan fluorescence were monitored at 330 and 350 nm, respectively. Buffer blanks were also recorded for normalisation.

2.4.10. Size exclusion with multi-angle light scattering (SEC-MALS)

SEC-MALS experiments were performed on a Superose 6 Increase 10/300 GL column connected inline to a HPLC (Shimadzu), DAWN MALS detector (Wyatt Technologies) and Optilab dRi detector (Wyatt Technologies). Protein conjugate analysis was performed with ASTRA software (Wyatt Technologies).

2.5. Cryo-EM

2.5.1. Vitrification

Grids were prepared similarly to other small membrane protein samples in Susan Lea's group [120,124]. Briefly, 4 µl purified protein, at varying concentrations, was applied to holey carbon-coated grids (Quatifoil 300 mesh, Au R1.2/1.3) glow-discharged for 30 seconds at 15-30 mA. Grids were blotted using a Vitrobot Mark IV (FEI) for 2-3 s at 100% humidity and 4°C, with a range of blot forces and a 10 second wait time. Grids were flash frozen in liquid ethane and stored in liquid nitrogen until imaging.

2.5.2. Screening

Grids were screened on a Talos Arctica operating at 200 keV. Micrographs were collected at a range of defocus values from -2.5 to -1 µm.

3. A bottom-up screen for candidate E3 ligase complexes for structural studies

3.1. Screening and characterisation

Structural biology of membrane proteins is particularly challenging because they need to be extracted from the lipid bilayer and kept in solution by a membrane mimetic throughout downstream analysis. Many of these mimetics are not compatible with traditional protein X-ray crystallography. Instead, cryo-EM has proven very successful for membrane proteins because they do not need to be crystallised, which has made some targets more amenable to structure determination [125]. This has given rise to an increasing number of membrane structures determined every year by cryo-EM, which has now begun to overtake those determined by other techniques [126]. However, for small membrane proteins cryo-EM is challenging because the bulk of the protein signal in electron micrographs lies within the detergent micelle. To determine structures by cryo-EM, a series of electron micrographs are collected of purified protein samples in vitreous ice [127]. The electron dose is kept low to reduce sample damage and thus each particle projection exhibits a low signal to noise ratio [127]. Computational methods are used to average these particle projections by identifying common signal features for alignment. This makes small membrane protein cryo-EM difficult since the high signal regions are obscured by a noisy membrane mimetic, inhibiting the ability of computational averaging [128]. Therefore, careful target selection, preparation and optimisation must be taken to generate a high-quality cryo-EM sample.

3.1.1. Identifying targets

In mammals, there are over 25 ER-associated E3 ligases, of which 11 have already been implicated in ERAD [16]. Although not all of these were implicated in ERAD, a

bioinformatic analysis was performed to identify those of which might be good candidates for cryo-EM structural studies (figures 3.1 and table 3.1). A recent interaction mapping screen by mass spectrometry revealed that many of the 25 E3 ligases do indeed form interactions with ERAD factors [40]. However, this study mapped overexpressed E3 ligases instead of using endogenously expressed protein and so this could lead to some degree of oversensitivity in the mapping results. Despite this, candidates were not ruled out at this stage until further analysis was performed. The domain organisation of these 25 E3 ligases is shown in figure 3.1. A good cryo-EM candidate for an E3 ligase alone would be a protein of sufficient MW to be above the limit for cryo-EM, of which is steadily approaching < 50 kDa with the help of nanobodies [128,129]. The domain organisation of these E3 ligases is very diverse and their lengths vary massively from 20 to 103 kDa. Correlating the overall architecture with the literature, we can see that some of these E3 ligases are related. Hrd1 and Gp78 are sequence related with each other and are homologous to yeast Hrd1. However, they both have distinct interaction networks and assemble with different partners, resulting in the degradation of different substrates [22,130]. Hrd1 and Gp78 are most different in their C-terminal domain. For example, Hrd1 possesses a HAF-H domain for interacting with its binding partner FAM8A1 [61]. Gp78 on the other hand has a G2BR domain that recruits E2 ubiquitin-conjugating enzymes and activates them through conformational changes for efficient ubiquitination [131,132]. The CUE domain of Gp78 further contributes to ubiquitination efficiency by binding to ubiquitin chains to orientate them better for ubiquitin transfer from Gp78 [132,133]. Both the VBM and VIM domains of Hrd1 and gp78 respectively, recruit p97/VCP to the ER membrane for retrotranslocation and substrate unfolding [52,53]. Surprisingly, none of the other E3 ligases have known similar domains for recruiting p97/VCP. Since p97/VCP is thought to be required for all forms of ERAD to unfold substrates before proteasome delivery, other E3 ligases likely recruit partners with p97-recruiting domains for efficient ERAD. p97/VCP also has a high affinity for

polyubiquitinated substrates, as evidenced by substrate trapping mutations, and thus likely engages directly with ERAD substrates [134]. TRC8 and RNF145 also appear related at first glance and at the sequence level they share approximately 27% identity, despite processing different classes of substrates [84]. The sequences of all E3 ligases were also locally aligned with each other using BLASTP but only the RING domains were similar between all, highlighting their diversity of sequences [135]. However, AlphaFold models of many of these E3 ligases are highly similar in TMD arrangement (some of which shown in figure 3.2), hinting at the presence of conserved structural features amongst E3 ligases in ERAD [136].

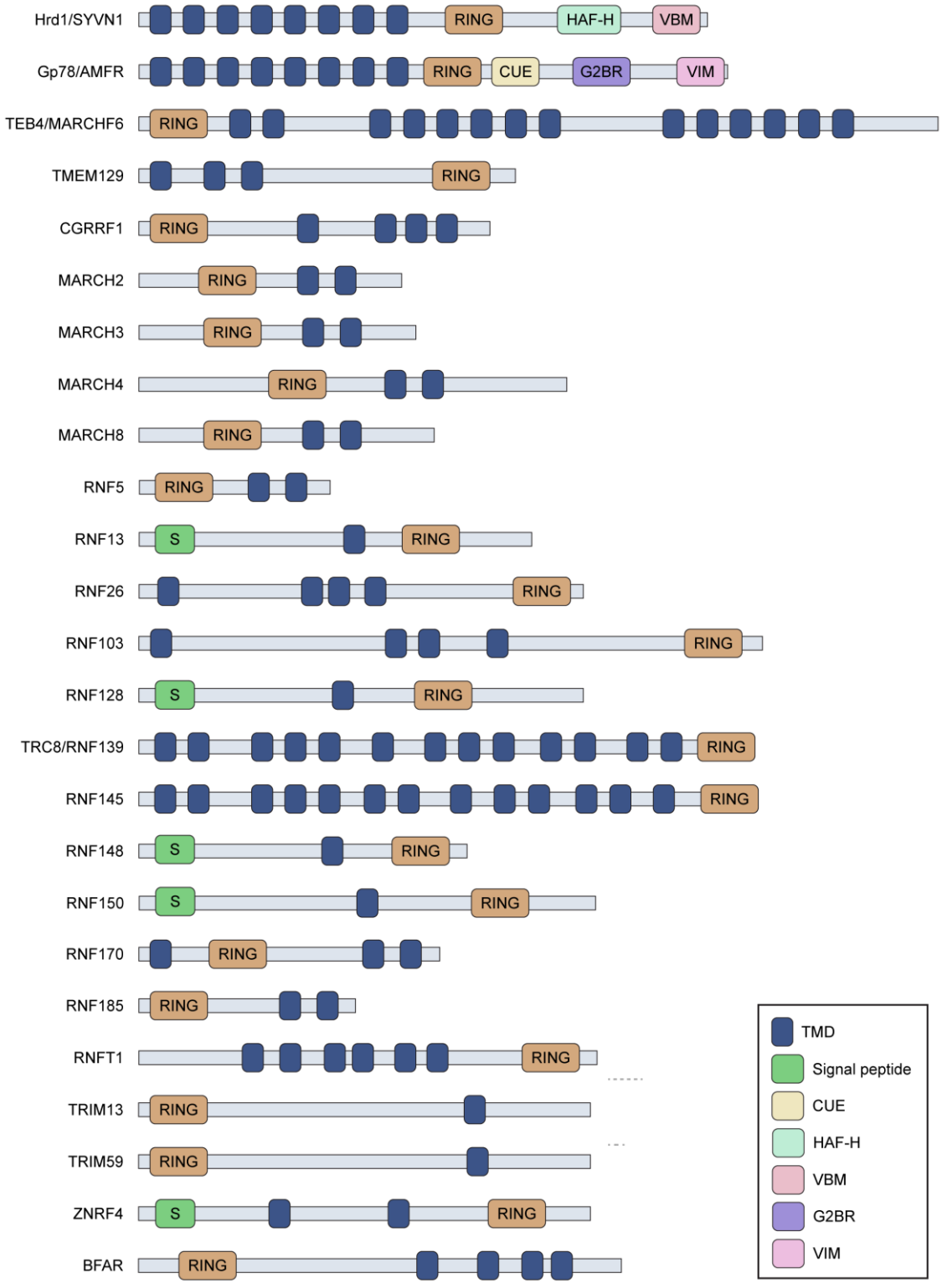


Figure 3.1. ER-associated mammalian E3 ligases and their domain organisation. Key domains and transmembrane domains (TMDs) are annotated. TMDs were annotated based on predictions from either DeepTMHMM or TMHMM [5,110]. The lengths of the

backbones represent equivalent lengths of the full-length proteins. Other domain assignments were carried out using a combination of UniProt and annotations in the literature [40,107]. The VBM and VIM domains recruit p97/VCP for retrotranslocation [52,53]. The HAF-H domain of Hrd1 is involved in recruitment of the ERAD factor FAM8A1 [61]. The G2BR and CUE domains of Gp78 cooperate to increase efficiency of ERAD substrate ubiquitination by E2 ubiquitin-conjugating enzymes [132,133].

Not all of these E3 ligases are necessarily good targets for cryo-EM since some are small and harbour high levels of intrinsic disorder, both of which would limit the achievable resolution by cryo-EM. To further assess the structure determination amenability of each E3 ligase, a brief bioinformatic analysis was performed and targets chosen based on a number of factors (table 3.1). To predict the number of TMDs, we used either DeepTMHMM or TMHMM, tools which employ a machine learning approach or hidden Markov models, respectively [5,110]. DeepTMHMM outperforms other TMD prediction algorithms and so was used for the bulk of predictions [110]. However, RNF26 had no predicted TMDs using DeepTMHMM but had 4 predicted TMD helices with TMHMM. RNF26 is a known membrane protein and so careful attention was made to use both prediction strategies for the best result [137]. At the time of performing this analysis, AlphaFold was not available but inspection of the AlphaFold model reveals at least 2 TMDs (not shown) [136]. E3s were also annotated with their percentage disorder using DISOPRED3 [138]. This is important as disordered, flexible domains will not necessarily be visible by cryo-EM since the average/ensemble structure is obtained during computational processing [127]. E3 ligases with high disorder are therefore less favourable targets. The literature was also searched to identify E3 ligases that are known to form homo-oligomers. Hrd1 and Gp78 were the only E3 ligases known to form higher-order oligomers and so these were attractive candidates for structure determination alone, since their final mass may be greater than the monomeric unit [100,139]. ERAD

interactors were qualitatively annotated based on the number of mass spectrometry hits from an interaction mapping experiment carried out on a range of E3 ligases [40]. Additionally, not all of these E3 ligases are well studied, especially in the context of ERAD. It would therefore be advantageous to focus on targets for which there was at least some biochemical studies, that could help interpret structural data. Hrd1, Gp78, MARCHF6, TRC8, RNF5, RNF145, RNF185, TMEM129, RNF170, CGRRF1 and ZNRF4 are the best studied E3 ligases and all of these have associated ERAD substrates [20]. Together, these metrics were used to annotate targets for their likely amenability for structure determination by cryo-EM (table 3.1). Good targets therefore had a high MW, higher numbers of TMDs, low disorder, known oligomers and a number of ERAD interactors that could be used to assemble higher-order complexes for structure determination. TMEM259, otherwise known as Membralin (MBRL hereafter), was also included in the bioinformatics screen. This was of particular interest because other members of the lab worked on a complex containing MBRL and identified that it assembles with the E3 ligase RNF185 in a novel ERAD pathway [105]. MBRL was shown to scaffold RNF185 along with TMUB1/2 to form a complex capable of processing an ERAD-M substrate [105]. RNF185 alone is too small a target for cryo-EM but MBRL was an attractive target since it was known to scaffold other components in the complex and contained the most TMDs. It was also shown to form homo-oligomers as endogenously tagged MBRL co-purified with untagged MBRL [105]. This analysis led to identification of Hrd1, Gp78, MARCHF6, TRC8, RNF145 and MBRL as candidates for downstream screening.

Target	MW (kDa)	TMD Number	% disorder	Known oligomers (y/n)	Number of ERAD interactors	Target amenability
Hrd1	68	8	44.2	y	+++	+++
Gp78	73	8	31.6	y	+++	+++
MARCHF6	103	14	14.4	n	+	+++
TMEM129	41	3	3.6	n	+	-
CGRRF1	38	4	21.7	n	++	+
MARCH2	27	2	34.6	n	++	-
MARCH3	29	2	36.4	n	++	-
MARCH4	46	2	41.5	n	+++	+
MARCH8	32	2	41.2	n	++	-
RNF5	20	2	28.9	n	+	--
RNF13	42	1	30.7	n	N/A	--
RNF26	48	4	15.7	n	+	+
RNF103	79	4	21.9	n	+	+
RNF128	47	1	31.8	n	+++	+
TRC8/RNF139	76	13	12.7	n	+++	+++
RNF145	75	13	15.1	n	N/A	++
RNF148	34	1	10.5	n	+	-
RNF150	48	1	34.2	n	N/A	-
RNF170	30	3	18.2	n	++	+
RNF185	20	2	35.4	n	+++	+
RNFT1	48	6	30.1	n	+	+
TRIM13	47	1	8.4	n	+	-
TRIM59	47	1	10.2	n	N/A	-
ZNRF4	47	2	33.3	n	++	-
BFAR	53	4	12.0	n	+	+
TMEM259	68	8	36.3	y	N/A	+++
SPP	41	9	14.6	y	N/A	++

Table 3.1. Bioinformatics of ER-associated E3 ligases for structural amenability. TMD number was calculated with either DeepTMHMM or TMHMM [5,110]. Percentage disorder was calculated with DISOPRED3 together with a custom Python3.7 script to calculate the percentage [138]. The literature was searched to find E3 ligases that displayed oligomeric states in biochemical data [100,105,139]. Interactions with other ERAD factors were determined from mass spectrometry experiments on different E3 ligases [40]. Target amenability was determined based on a qualitative assessment of the different biochemical and bioinformatic properties.

At the time of this bioinformatic analysis, trRosetta was used for 3D structure prediction [140]. This could accurately predict secondary structure but in our experience

of membrane proteins, the arrangement of the membrane domains often did not make sense for some targets (not shown). This has since been superseded by AlphaFold and RoseTTAFold which more accurately predict 3D protein structure, with a stronger reliance on machine-learned models of protein structure [136,141]. For reference, the corresponding AlphaFold structures of the six selected targets are shown in figure 3.2 [136]. Here we can see that these E3 ligases appear to share a broadly similar arrangement of α-helical TMDs – namely Hrd1, Gp78, TRC8 and RNF145. The Cα-backbones of two protein models can be aligned with each other and the root mean square deviation (RMSD) in distance calculated, to obtain an idea of how closely two models resemble each other [142]. Hrd1 and Gp78 are closely related homologues of each other and alignment of just the core TMDs and RING domains results in an RMSD of ~6 Å. TRC8 and RNF145 are also related to each other and align together with an RMSD of ~8 Å, with the TMDs between the two proteins aligning very well with respect to each other. Interestingly, all four of these E3 ligases feature long TMD α-helices together with shorter TMD α-helices. If all of these reside in the membrane, some level of membrane thinning towards the side of the protein with shorter TMDs would be expected, as in regions of the structure of the yeast Hrd1-Der1 complex [55]. MARCHF6 has numerous predicted TMDs, most of which are predicted with high confidence by AlphaFold (not shown) and are at an angle relative to what would be the expected membrane plane. MBRL is also unique in its arrangement of TMDs and appears to have a single long helix passing through the membrane. There is also an α-helical segment that makes a right angle at the plane of the membrane, but this is predicted with low confidence [136]. All of these E3 ligases are arranged with their RING domains on the cytosolic side and this is the strongest conserved element amongst them, adopting the typical RING domain fold as determined in crystal and NMR structures [143,144]. The main takeaway from the predicted structures is that their TMD helices bundle together in

the membrane and that the disorder is usually present in soluble cytoplasmic domains. This is important as tight, stable folds make for better structural targets.

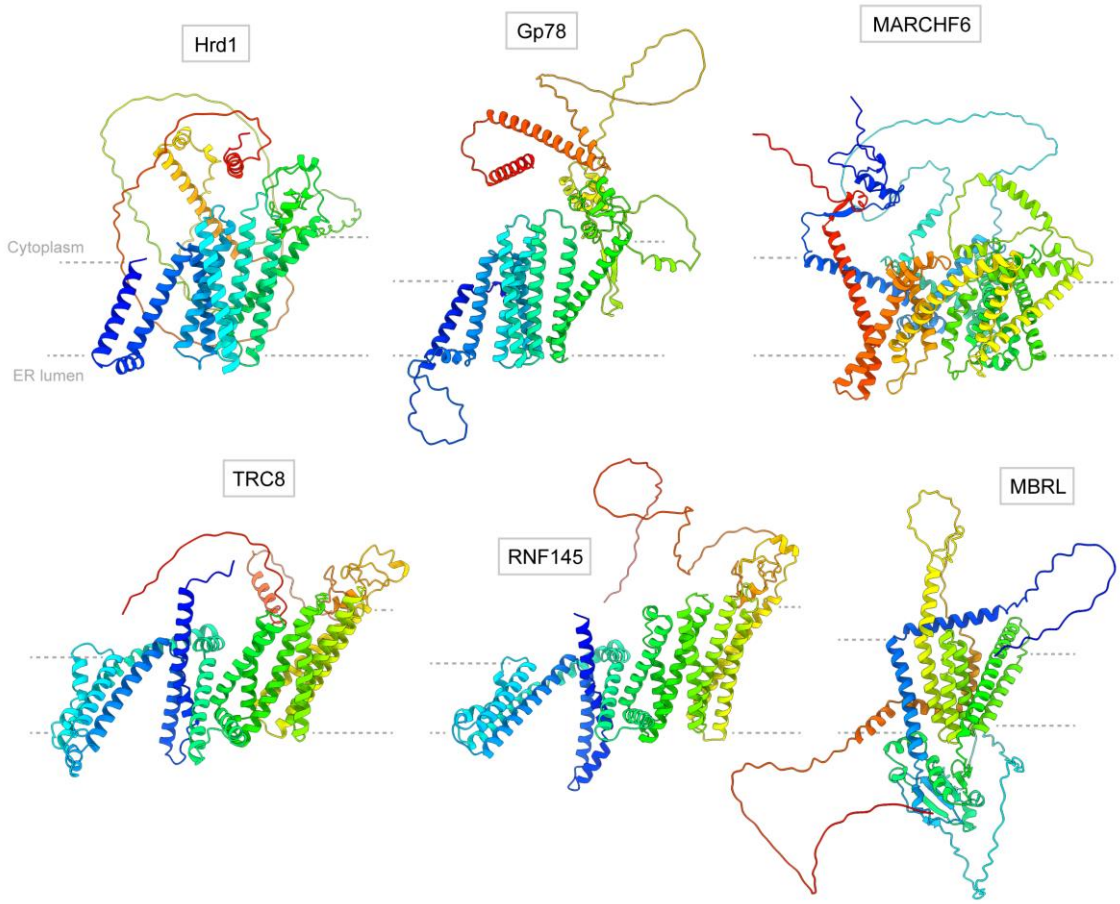


Figure 3.2. AlphaFold structure predictions for the selected targets. Models were predicted using AlphaFold2 and coloured from the N- to C-terminus in blue to red [136]. The membrane plane was approximated by inspection of protein hydrophobicity and hydrophobic regions on the exterior of the α-helical TMDs denotated as being in the plane of the membrane.

In summary, Hrd1, Gp78, MARCHF6, TRC8, RNF145 and MBRL were selected as targets for downstream screening and assessment of amenability for structure studies. Importantly, all these targets have known membrane interactors. From experience in the

lab with small membrane proteins, they often require co-purification with their native partners for good biochemical behaviour. This will be important for targets that initially behave poorly and perhaps co-expression with partners will aid in their stability and assembly. To isolate these targets and their complexes, an expression system was developed which was then used to screen for expression, amenability to purification and biochemical stability.

3.1.3. Developing a mammalian expression system

Recombinant proteins can be expressed in either bacteria, yeast, insect or mammalian cells. For expression of recombinant mammalian membrane proteins, the latter two are more heavily used [145]. This is because these expression systems better represent the environment that the endogenous protein would be exposed to. All membrane and secretory proteins produced in these systems are initially synthesised in or are targeted to the ER before maturation to a cellular compartment or the plasma membrane [4]. Recombinant proteins are best expressed in their native environment where the machinery for synthesis, folding and quality control already exists [146]. This also includes the machinery for any post-translational modifications such as O- or N-linked glycosylation, which may also be important for maturing to the correct fold [146,147].

Since we eventually wanted to pursue structural biology of multi-component complexes found within the mammalian ER, we opted for a mammalian expression system. Recent work has shown that generation of stably expressing HEK293 cells overcomes some of the drawbacks of traditional transient expression from a plasmid [111]. Transient protein expression becomes limiting when working large scale because of the large quantities of plasmid required, large biomass volume, limited transfection efficiency and high cost [111]. Instead, stable cell lines can be established using

recombinant lentiviruses to insert the transgene of interest into the genome, driven by a strong promoter [111]. Since our targets are intracellular membrane proteins, we opted for the inducible expression system whereby protein expression is driven from a human cytomegalovirus (CMV) promoter, under the control of the TetO operator. When expressed from HEK293S cells expressing the TetR repressor (HEK293S TetR cells), transgene expression can only be initiated in the presence of tetracycline or doxycycline, and thus serves as an on/off switch for expression [114]. Use of stable cell lines also allowed us to quickly screen a variety of conditions and rapidly scale up production.

To better suit our needs, this system was modified for eventual expression of multi-component complexes from a polyclonal cell line. mVenus was chosen as the fluorophore since this protein is monomeric at concentrations typically observed during purifications, and it was the brightest fluorophore of this variety at the time [118]. Some targets may not express well so having a bright fluorophore increases expression screening sensitivity. Within the pHR-CMV-TetO_2-3C-mVenus-TwinStrep plasmid, the protease site was changed to TEV, the RPTPσ signal sequence removed and the linking regions between domains increased in length. The RPTPσ signal sequence is used to designate proteins for the secretory pathway and so was not relevant for our aims. The sequence of the tags was also changed to generate backbones containing N- or C-terminal tags around the multiple cloning site (MCS). To generate cell lines for overexpression of multi-component complexes, antibiotic resistance cassettes were introduced to the backbone vectors (see methods and figure 3.3a). A stable cell line already expressing one protein would then be subject to further rounds of lentiviral transduction with recombinant lentiviruses harbouring plasmids with antibiotic resistance cassettes. Transduced cells would then be enriched by antibiotic selection and with our range of backbones, a stable cell line expressing with up to 5 different transgenes was generated.

Further changes were made to the expression system to improve protein yields and improve consistency between generation of different high expression cell lines. One of the biggest improvements in protein expression was observed when each HEK293S cell harboured more than one copy of the transgene. To achieve this, transfection efficiency for lentivirus generation was increased with higher-quality transfection reagents, in turn increasing the lentiviral titre available for transduction. Also, the effective lentiviral titre was increased by simply decreasing the number of HEK293S cells seeded during viral transductions. These changes guaranteed greater than 90% transduction efficiency, with HEK293S cells often being transduced multiple times (figure 3.3b).

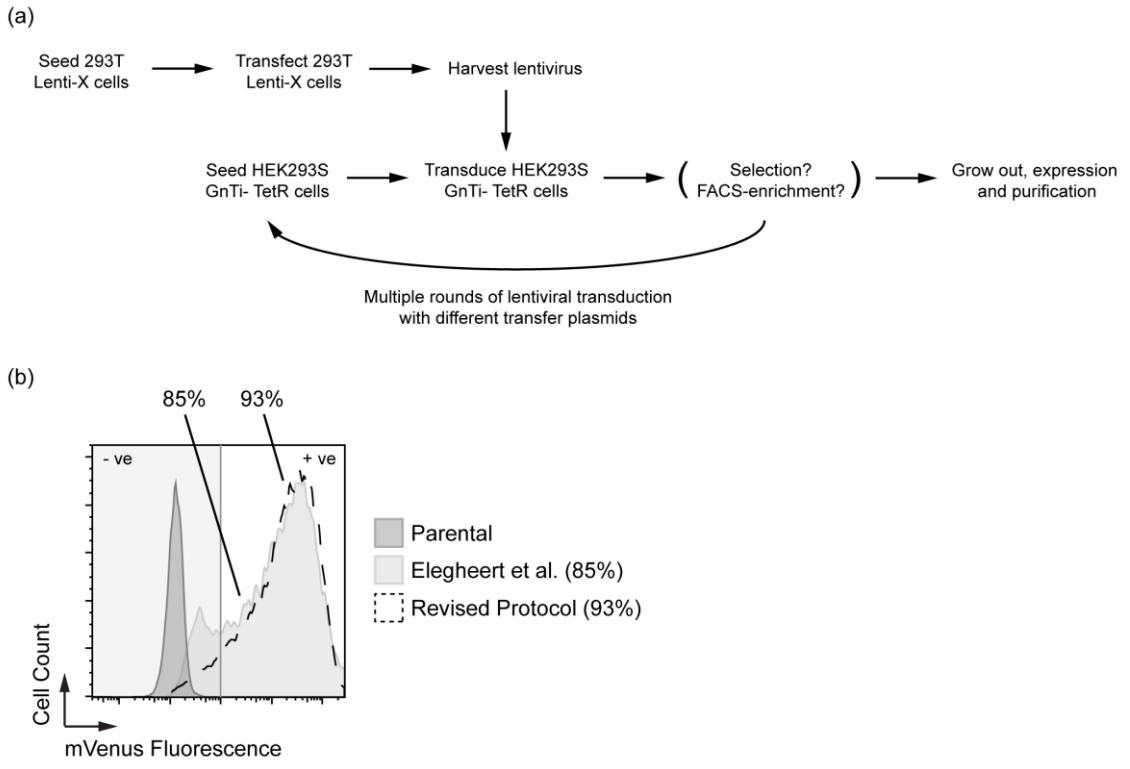


Figure 3.3. Modification of a mammalian expressions system for expression of multi-component membrane protein complexes. (a) Workflow for generation of multi-transduced, stably-expressing cell lines. (b) Comparison of expression of MBRL, as seen

by flow cytometry, from stably expressing cell lines generated with the Elgheert et al. protocol versus the revised protocol [111].

3.1.4. Expression screening

The five different mammalian E3 ligases and MBRL were cloned into pHR-CMV-$TetO_2$ for generation of stable cell lines and overexpression. All these recombinant proteins were C-terminally tagged with TEV-mVenus-TwinStrep. Stable cell lines were generated and protein expression after 48 hours induction assessed by flow cytometry and western blotting (figure 3.4). Flow cytometry was often a good way to rapidly screen expression upon addition of doxycycline. All targets showed some level of induced fluorescence signal except for RNF145, where there was only a marginal shift in fluorescence versus the parental cells (figure 3.4). The fluorescence of the MARCHF6 cell line also did not shift by a large amount upon doxycycline addition, indicating a low level of protein expression. Both RNF145 and MARCHF6 participate in sterol homeostasis, a pathway which is tightly regulated [148]. It is possible that these two proteins undergo regulated proteolysis to ensure tight regulation of sterol synthesis. RNF145 and MARCHF6 undergo auto-ubiquitination and so post-translational regulation by ERAD may be preventing high expression [84,149]. The presence of multiple peaks in the doxycycline conditions could be attributed to cells harbouring differing numbers of transgene inserts, which is especially apparent for the MBRL cell line. The viral titre also varies with plasmid size since efficiency of lentivirus assembly falls with increasing plasmid size [150]. MARCHF6 was the largest protein screened and so it was expected that its expression might be lower than the other proteins since the viral titre would be lower.

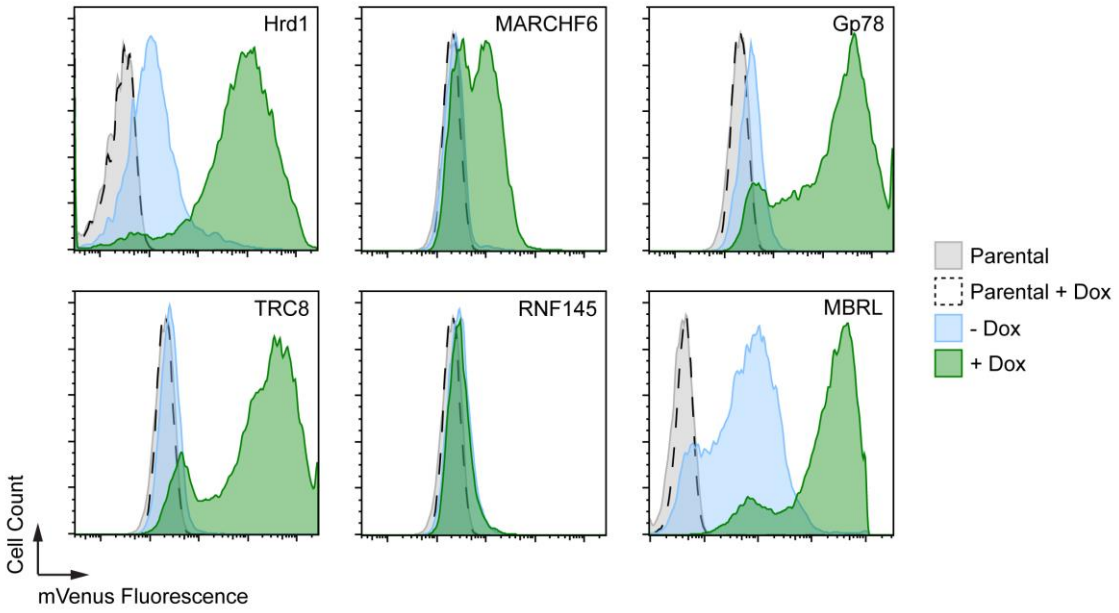


Figure 3.4. Expression screening of different protein quality control targets by flow cytometry.

Western blotting after small-scale pulldowns confirmed expression of some of these targets but for MARCHF6 and RNF145, little or no protein product could be seen (figure 3.5). In the lab's experience, tags were often cleaved from proteins that were unstable from overexpression and so pulldowns were a good way to assay the presence of an intact purification tag. Endogenous MBRL was used as a pulldown control. Some signal could be seen for MARCHF6 by western blot after a long exposure but compared to the other cell lines, this protein had very low expression. Western blotting was always an important cross-reference for flow cytometry as sometimes fluorescent signal was seen by flow cytometry, but full-length product not visualised by western blot. Aberrant fluorescent signal could be attributed to free mVenus in cells, possibly from a cleavage event after translation of recombinant protein. Since there was some fluorescent signal in the MARCHF6 cell line post-induction, this cell line was enriched for mVenus-positive cells by flow cytometry assisted cell sorting (FACS). Protein expression in the MARCHF6 cell line was induced overnight and fluorescent cells isolated. After cell sorting, this cell

line appeared more homogenous by flow cytometry, but there still was no improvement in GFP signal by western blot (figure 3.6). In future, creation of catalytically inactive versions of RNF145 and MARCHF6 may improve expression and make them more amenable to structure determination. Out of these six targets, only Hrd1, Gp78, TRC8 and MBRL were taken forward for further screening.

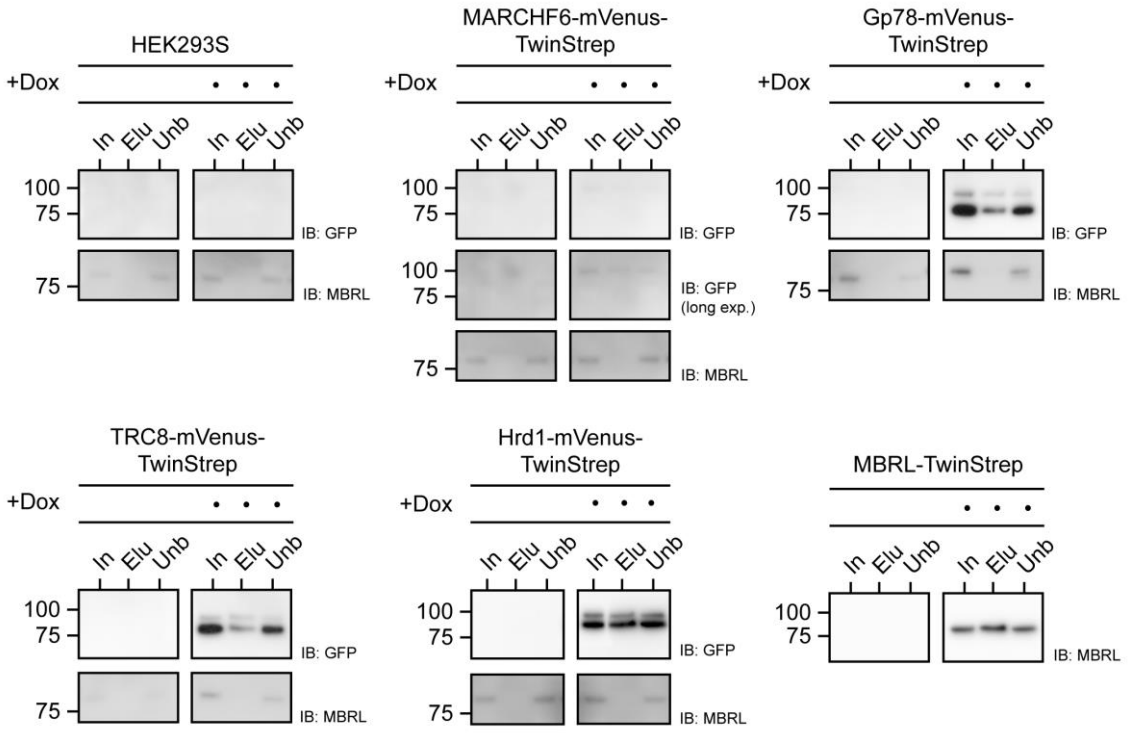


Figure 3.5. GFP pulldowns for the different stable cell lines after doxycycline-induced expression.

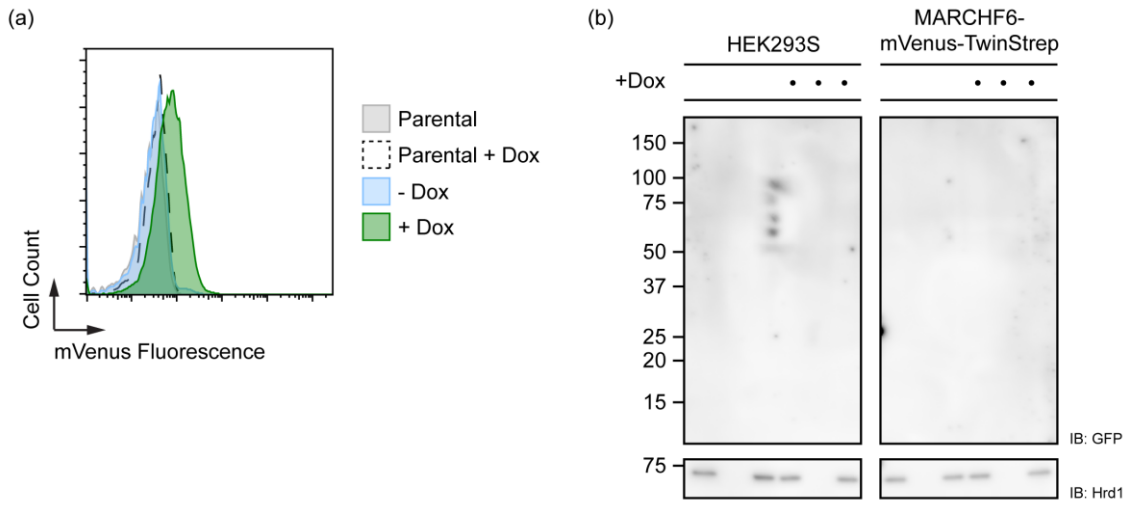


Figure 3.6. Enrichment of the MARCHF6 cell line. (a) mVenus-positive cells in figure 3.4 were isolated and re-analysed by flow cytometry after doxycycline-induced expression. (b) Expression as assessed by western blot.

3.1.5. Detergent screening

Identifying a suitable detergent for membrane protein purification is an important factor in the generation of a homogenous, purified sample. Detergents solubilise and extract membrane proteins from their lipidic environment and different proteins have distinct preferences for detergent environments [126]. The detergent environment therefore strongly influences the behaviour of the membrane protein. These differences can be attributed to interactions of protein with the detergent, the detergent head group chemistry and the detergent chain length [151]. Detergents can be classified in terms of their harshness, that is their efficiency of extraction and ability to break apart the protein-protein interactions of TMDs. Generally, the more polar the detergent head group, the harsher the detergent. Detergents also become harsher with shortening chain length and smaller head group sizes [151]. A suitable detergent condition for a membrane protein should extract a high fraction of the target protein from the membrane, although this can be overcome to some degree with a large excess of detergent during extraction. For

proteins with a tendency to aggregate, addition of more detergent may not be enough to fully extract it from the membrane. The extracted target protein should then remain stable and retain its native secondary structure. Ideally, the target should also be homogenous and not prone to aggregation. Unstable membrane proteins often tend to aggregate, and this is usually irreversible. Downstream purification techniques can remove some of this aggregation but screening for different conditions prior to purification can be more beneficial. Although detergent plays a large part in the biochemical behaviour of membrane proteins, other factors such as salt concentration and buffer pH are also important [152]. This is especially true for membrane proteins that also have soluble domains and certain additives may be needed to stabilise the overall protein fold [153].

On top of these properties, the extraction detergent must also be compatible with cryo-EM. Cryo-EM structures of membrane proteins have generally been solved in a select range of detergents or other membrane mimetics (table 3.1). A lot of our preference for detergents comes from high-resolution structural studies of membrane proteins by X-ray crystallography. Maltosides account for approximately 50% of determined X-ray crystallography structures of membrane proteins [154]. A similar preference can be seen for structures of membrane proteins determined by cryo-EM (table 3.1). There is no good information on what makes a good cryo-EM detergent, but we know that high CMC detergents like octyl glucoside (OG) can be problematic [126]. These are used less frequently, likely because of higher operating cost (from the high CMC), short acyl chain length and monosaccharide head group, the latter of which reduce protein stability [155]. High concentrations of detergent in the final preparation buffer can also make the fluid more viscous, in turn affecting blotting of the cryo-EM grids during vitrification. This can make it hard to obtain the thin ice required for imaging and a high detergent concentration could reduce contrast in EM micrographs, reducing the total signal-to-noise ratio. Detergents with large micelles, like GDN, can be problematic for small membrane proteins where the protein signal is already obscured by the noise

of the detergent micelle, thus preventing high-resolution structure determination [128]. The addition of cholesteryl hemisuccinate (CHS) has also been shown to stabilise membrane proteins [156]. Mixed sterol/detergent micelles have a more bicelle-like structure and the sterols are thought to accumulate in the less curved regions of the micelle, forming a core bicelle flanked by detergent molecules [156]. The size of the micelle is also increased with increasing sterol concentration, which may be beneficial for the stability of some proteins.

Membrane mimetic	Cryo-EM structures
Non-ionic maltosides	22
Non-ionic maltoside neopentylglycols	5
Digitonin	23
Other non-ionic (Triton, Tween-20, C12E10)	3
Zwitterionic	1
Nanodiscs	5
Amphipols	14
Other mimetics	2

Table 3.1. Choice of membrane mimetic for solved structures of membrane proteins by cryo-EM, as of August 2018 [157].

Since the solubilisation efficiency alone does not give much information on the stability of a protein in a specific detergent, we opted to screen targets solubilised in various detergents by fluorescence-based size exclusion chromatography (FSEC). This allowed us to assess the homogeneity, stability and to some extent the solubilisation efficiency of the sample. All recombinant proteins were tagged with mVenus and so solubilised membranes were loaded directly onto a SEC column without enrichment for the target protein, and protein fluorescence measured. This saved time and allowed for a wide range of detergent conditions to be screened.

During SEC, objects are separated by size as they diffuse through a column of uniformly-sized porous beads [158]. The porosity of the beads is kept constant and is critical in determining the rate at which differently sized objects pass through the column. Smaller objects diffuse more easily into the beads and so are retained for longer, thus eluting later. Objects larger than the pores in the beads are excluded from the beads and elute the earliest [158]. This volume is termed the void volume of a column and is the size-exclusion limit for a given column. All proteins exhibited some level of aggregation in all detergent conditions as seen by the peak at the void volume (figure 3.7). The final peak in all FSEC runs at ~17.5 ml is probably from free, cleaved mVenus tag. For Hrd1 in DDM-based conditions, the titration of CHS caused a shift in equilibrium of the 2 peaks around 12.5 ml, from a larger object to a smaller object. This is the opposite of what is expected, as increasing sterol concentration in detergent/sterol micelles has been shown to increase micelle size [156]. The opposite effect was seen with LMNG and DMNG, where titration of CHS caused a shift towards a larger species. GDN and DMNG-CHS exhibited the worst FSEC traces, with a large amount of material shifting towards the void, indicating a propensity for aggregation in these conditions. The 2 monodisperse peaks in DDM could be attributed to the presence of 2 distinct protein species being solubilised from the membranes. Structures of yeast Hrd1 have been solved as monomers and homodimers [55,60]. This could be an explanation, but the retention volume is dependent on the hydrodynamic radius of the particle, so these peaks could also be different protein conformations that are stabilised differently across different detergents. In the end, DDM-CHS (10:1) was chosen since this solubilised both protein species the best. Gp78 was best behaved in DDM with higher concentrations of CHS, where the peak tended to become more monodisperse and so DDM-CHS (10:1) was selected for downstream purification. Generally, Gp78 had more of an aggregating shoulder than Hrd1 and peaks tended to run wider, indicating that this protein was not as stable as Hrd1 in these conditions. TRC8 was well behaved in all detergent conditions

and so DDM-CHS (10:1) was chosen since its solubilisation efficiency was the best. MBRL also tended to aggregate and the main peak always had an associated aggregating shoulder. In DDM conditions, an additional peak appeared between 13-15 ml that was not seen in milder detergent conditions. DDM is a harsher detergent than LMNG, DMNG and GDN, and so either DDM is disassembling a MBRL complex or causing MBRL to exist in multiple conformations. Since the peaks were sharpest with DDM conditions, DDM-CHS (10:1) was selected. This condition improved the solubilisation efficiency of both protein species present.

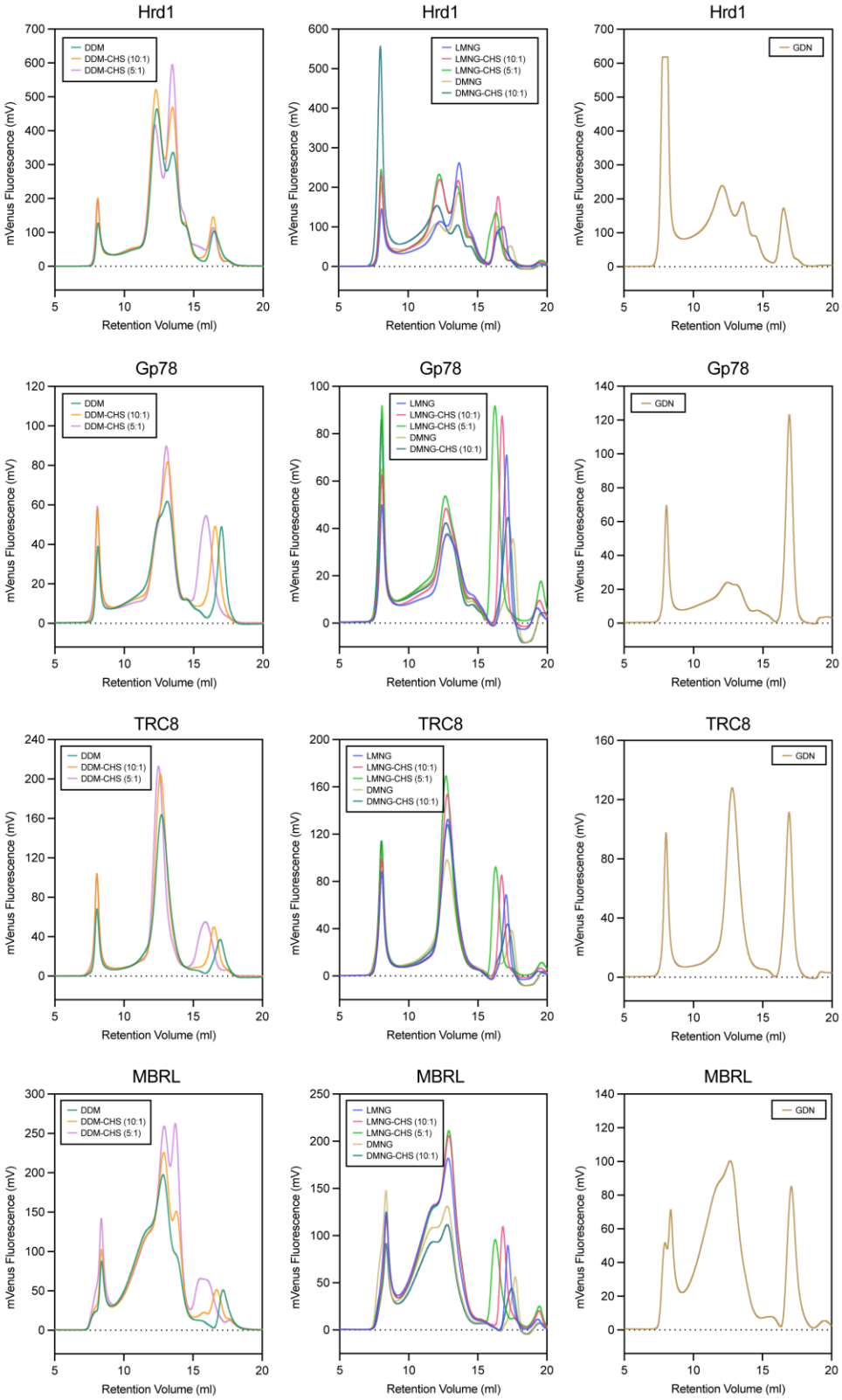


Figure 3.7. Fluorescence-based size exclusion chromatography (FSEC) detergent screening of Hrd1, Gp78, TRC8 and MBRL.

3.2. Small-scale purifications

Detergent conditions identified by FSEC were used to purify proteins at a larger scale. Membrane proteins sometimes behave quite differently at higher concentrations as they are often prone to aggregation [126]. Concentration-dependent aggregation is common with membrane proteins and so it is important to see how they behave at a larger scale before deciding on a single detergent condition [151]. 2 L cultures of suspension cells were prepared and used for small-scale purifications. Proteins were purified with the C-terminal TwinStrep tag and further polished with SEC. Final SEC traces with their representative SDS-PAGE gels are shown in figure 3.8. The mVenus-TwinStrep tag was generally cleaved off prior to SEC, except for the case of MBRL.

Hrd1 behaved the best biochemically when purified at scale. After affinity purification, the SEC trace for Hrd1 resulted in a symmetrical monodisperse peak with no sign of aggregation. The corresponding protein as assayed by SDS-PAGE was pure and contained a single band corresponding to Hrd1. Gp78 however, had a tendency to aggregate, as indicated by the aggregating shoulder left of the main peak. At ~14 ml though, there was still a somewhat monodisperse peak. TRC8 and MBRL performed poorly and were very prone to aggregation. Both proteins eluted over a wide range of volumes, indicating the existence of many states. This could be due to protein aggregation or they exist in multiple conformational states when overexpressed and isolated this way. Interestingly, during the TRC8 purification, there were additional co-purifying bands seen in the SEC input and SEC elutions. The identity of these bands was not confirmed by mass spectrometry and so they could be co-purifying endogenous partners, cleavage products or aggregated forms of TRC8. Despite this, TRC8 is known to interact with SPP, a 40 kDa membrane protein, as well as other ERAD factors [40,79]. MBRL had an aggregating shoulder as indicated by a smooth transition from the void

volume to the peak at ~14 ml. This suggests that MBRL is also poorly behaved under these conditions.

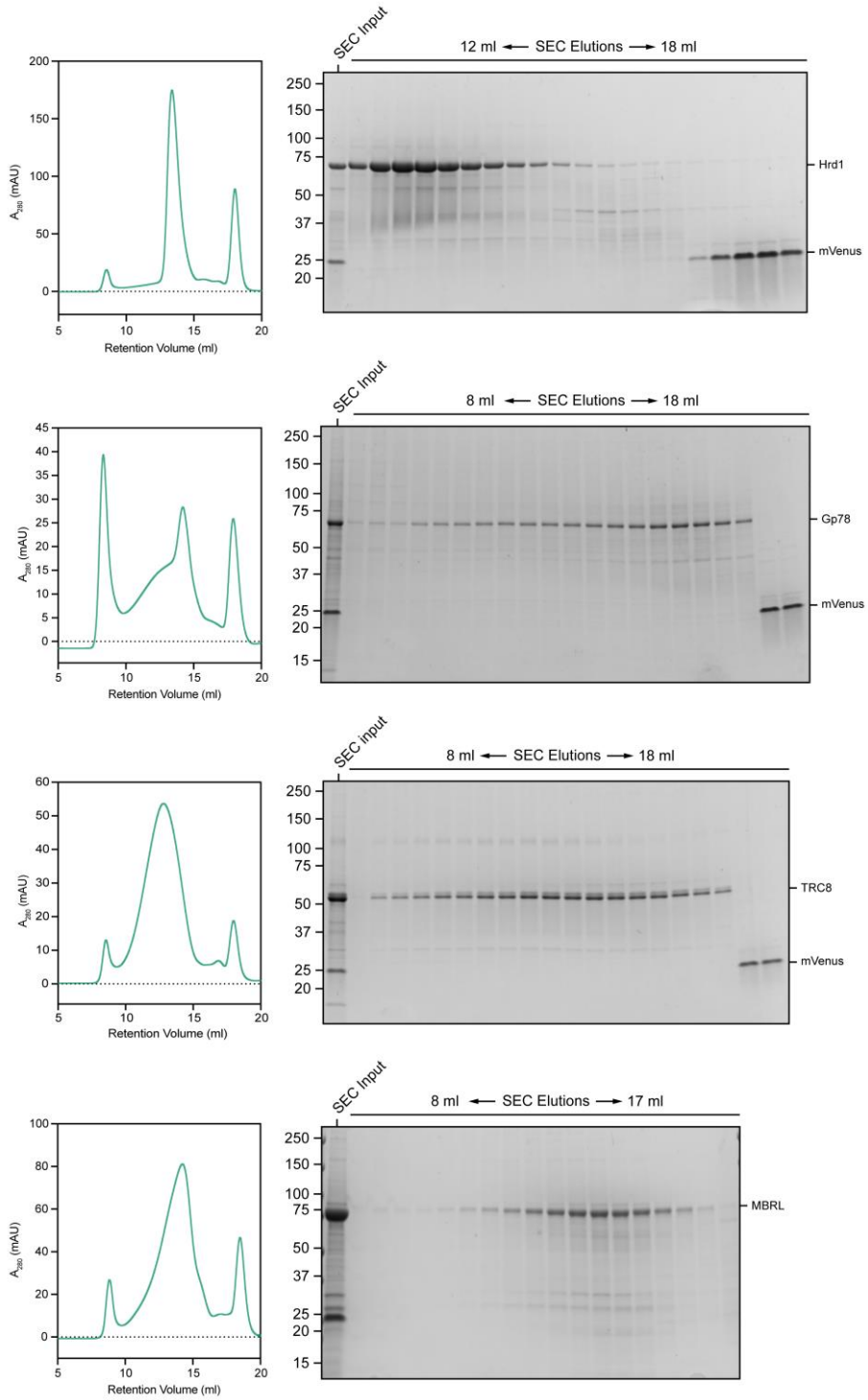


Figure 3.8. Small-scale purifications of Hrd1, Gp78, TRC8 and MBRL. All proteins were initially purified with StrepTactin and polished further by size exclusion chromatography (SEC). SDS-PAGE of the SEC elutions is also shown. Hrd1 was purified in DDM-CHS (10:1), Gp78 purified in DDM-CHS (10:2), TRC8 purified in DDM-CHS (10:1) and MBRL purified in DDM-CHS (10:1).

3.3. Electron microscopy screening

All proteins that were purified at a larger scale were further characterised by electron microscopy (EM). Samples were initially prepared by negative staining for imaging by transmission EM (figure 3.9). This is a useful tool for assessing the homogeneity of a sample but for small membrane proteins, this technique cannot be used to obtain useful high-resolution structural information [159]. There is sometimes a correlation between behaviour of proteins during negative staining and cryo-EM grid preparation. In the experience of the lab, proteins that behave poorly when imaged by negative stain, often behave poorly when imaged by cryo-EM. Characteristics of a well-behaved protein include homogenous shape and evenly dispersed particles.

To better assay behaviour of targets by negative stain EM, 2D class averages of protein particles were calculated from a series of EM micrographs using EMAN2 [160]. Briefly, 20-30 EM micrographs were collected for each target at a range of defocus values from -1.5 to -2.5 µm. This was collected on a 120 kV microscope with a spherical aberration of 2.2 mm. Images were imported into EMAN2 where CTF parameters and defocus values were estimated automatically, with X-ray pixel filter turned off and an amplitude contrast of 80. Particles were picked automatically with the inbuilt convolutional neural net after training with a series of “good”, “bad” and “background” particles. Particle boxes were extracted and the CTF auto processing tool run. 2D class

averaging was then performed with the iterative reference-free classifier, setting the class averager to mean.

By negative staining, both Hrd1 and Gp78 appeared homogenous and did not aggregate. 2D class averages of particles from these micrographs also appeared to be quite homogenous. For small membrane proteins, it is difficult to obtain any detail from 2D class averages of negative stain micrographs, hence why these objects lack any structural detail [159]. This is in part because of the granularity of the stain used for negative staining, limiting resolution to ~20 Å [161]. For these targets, there are not many large, globular folds predicted to exist outside of the membrane domains and it is expected for them to look spherical by negative staining. MBRL particles had a tendency to aggregate and were a lot more heterogenous in shape. TRC8 behaved similarly but was very aggregated and appeared to form string-like structures. This indicated that behaviour between negative stain and SEC correlated for all targets, where TRC8 and MBRL behaved the worst.

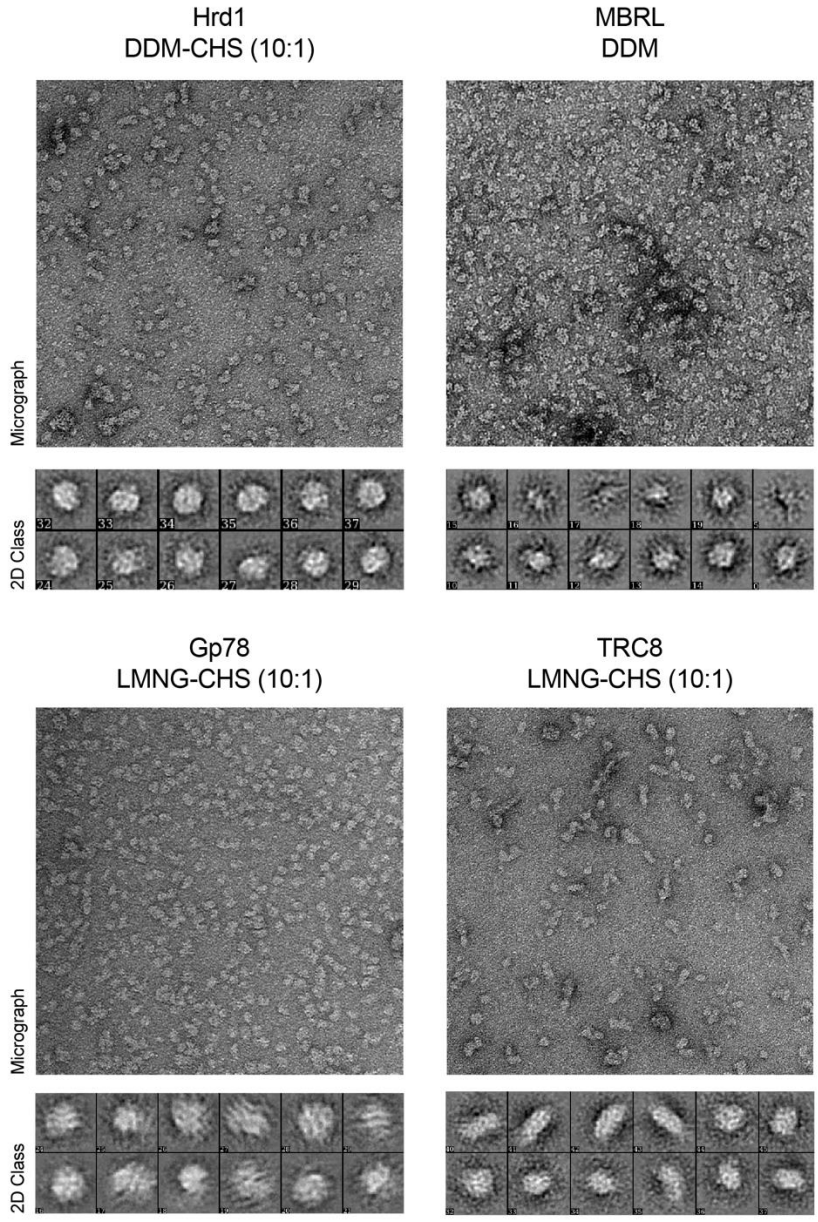


Figure 3.9. Screening of targets by negative stain electron microscopy (EM). A representative micrograph is shown for each target condition where protein particles are negatively stained with a heavy metal. 2D class averaging was performed with EMAN2 and a selection of 2D classes are shown below each micrograph [160].

Targets were also analysed in parallel by cryo-EM (figure 3.10). Protein was frozen by vitrification and imaged under a 200 kV microscope in cryogenic conditions. The

behaviour of all the targets correlated well with the SEC and negative staining data. Hrd1 particles had defined shapes and were disperse, as seen in the EM micrographs. Since the quality of the particles looked good, a small data collection was performed and 2D class averages generated with SIMPLE3.0 [162]. For processing of this dataset, EM movies were automatically processed with SIMPLE preprocess_stream where motion correction, CTF parameter estimation, particle picking and particle extraction are performed during data collection. For particle picking, a lowpass-filtered reference volume of an unrelated small membrane protein from the lab was used to automatically pick particles. 2D class averages were generated with SIMPLE 2D_stream with a mask diameter of 180 Å. For the Hrd1 sample, hints of secondary structure detail could be seen which was a promising sign. Data was also collected for the Gp78 sample since distinct micelle-like objects could be observed, despite being very concentrated. However, no secondary structure detail was seen in any of the resulting class averages. This could be attributed to a variety of reasons such as gp78 existing as a monomer too small to resolve by cryo-EM, conformational heterogeneity or partial denaturation at the air-water interface. MBRL behaved much poorer and particles were very clumpy, suggesting that they were aggregated. This correlated with previous data and so no data collection was performed. Strangely, the MBRL particles did not look particularly like a membrane protein as seen by cryo-EM as there were not many defined, clearly micelle-like particles (based on previous lab experience). One rationale for this is that MBRL was denaturing during freezing, possibly by interaction with the air-water interface. The TRC8 sample was also clumpy and there appeared to be a wide range of particle sizes present. At the time, no data was collected on this sample but in retrospect it might warrant a closer inspection since defined, micelle-like objects can be seen, despite the aggregation present. Since the Hrd1 sample looked the best at this stage, this was selected for a larger data collection. The results of this data collection are presented and discussed in the next chapter.

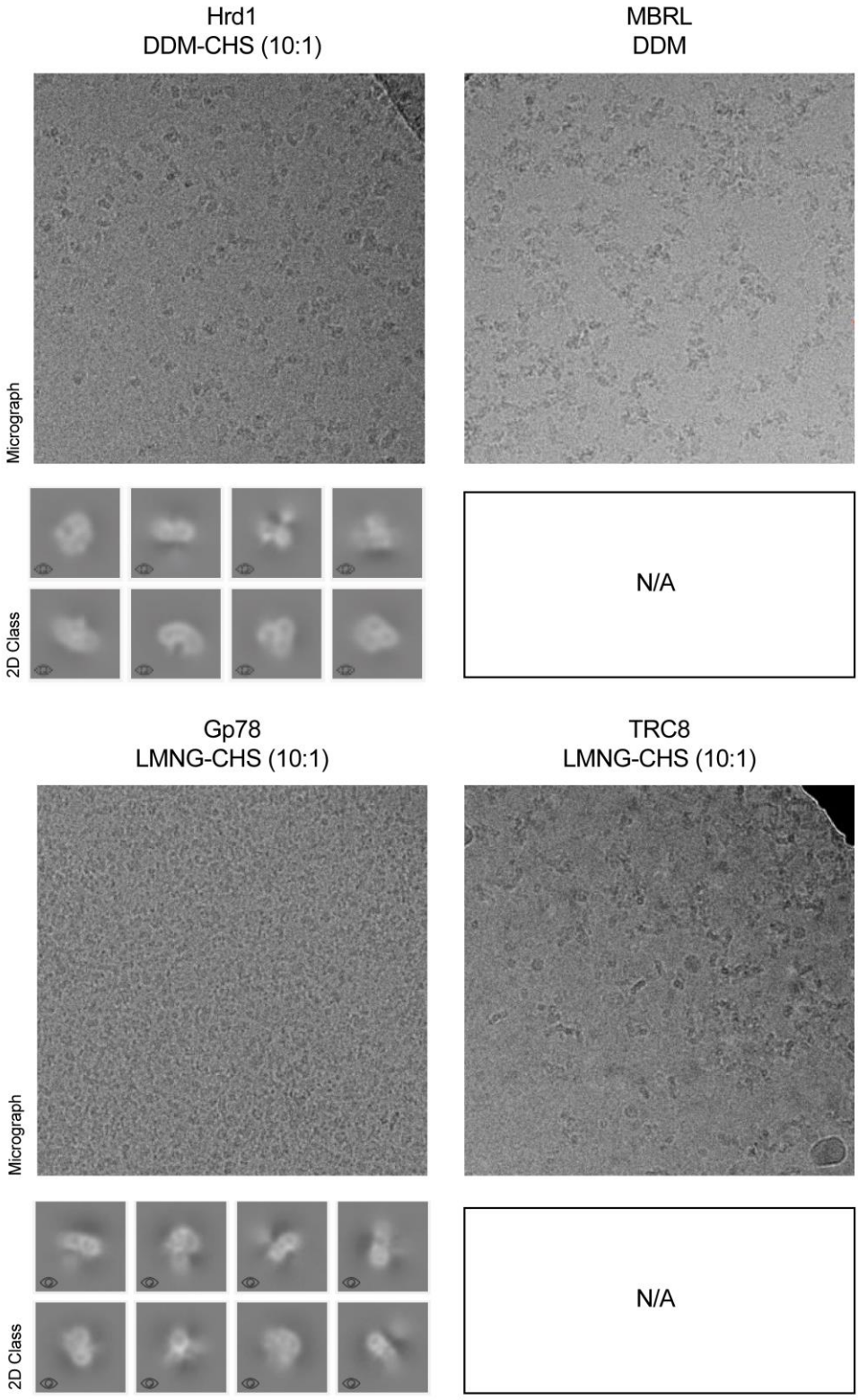


Figure 3.10. Screening of targets by cryo-EM. A representative micrograph for each target is shown where protein is frozen within vitreous ice. 2D class averages were generated with SIMPLE3.0 after a small data collection [162].

3.4. Discussion

In this chapter, an analysis of ER-associated E3 ligases was performed to assess their amenability for structure determination. These E3 ligases are diverse in their amino acid sequences and domain organisation, as well as their ERAD interaction network. Out of the E3 ligases assessed, five were chosen for downstream screening. In addition to this, the protein MBRL, that assembles an ERAD complex with the ubiquitin ligase RNF185, was included for further analysis since this was of special interest to the lab. All six targets are multi-spanning membrane proteins. Interestingly, predicted AlphaFold models suggested that all these targets would cause membrane thinning on one side of the protein. This has been suggested to be an important feature for the retrotranslocation function of the yeast Hrd1 complex [163]. Presumably, thinning of the membrane reduces the energetic barrier for moving hydrophilic polypeptides regions through the ER membrane. This may also help to partially destabilise TMDs prior to unfolding by p97/VCP.

An existing mammalian protein expression system was adapted for the recombinant expression of multi-component complexes. This was done by creation of new transfer plasmids encoding selection cassettes and modification of the existing protocol. This protocol was also optimised for transgene insertion efficiency, thereby improving protein expression and reproducibility. The capabilities of this system are explored more thoroughly in later chapters and this will undoubtedly be a useful tool for robust reconstitution of future mammalian membrane protein complexes.

Hrd1, Gp78, MARCHF6, TRC8, RNF145 and MBRL were screened for expression by flow cytometry (using their mVenus moiety) and western blot. MARCHF6 and RNF145 could not be overexpressed at detectable levels and were dropped from further characterization. Although a faint band corresponding to tagged MARCHF6 was detectable by western blot, we failed to enrich for high expressing cells using flow

cytometry. Both RNF145 and MARCHF6 are implicated in sterol metabolism where they carry out regulated proteolysis of rate-limiting enzymes in sterol synthesis [84,164]. Both these enzymes undergo auto-ubiquitination and rapid turnover and are themselves likely regulated by ERAD [84,149]. Therefore, obtaining high levels of the catalytically active forms of these E3 ligases is likely difficult. Plasmids encoding catalytically inactive mutants of RNF145 and MARCHF6 were made but were never tested for expression.

Optimal detergent conditions for Hrd1, Gp78, TRC8 and MBRL were identified, and small-scale purifications performed to understand the behaviour of these proteins. Hrd1 and Gp78 behaved the best by SEC, whilst TRC8 and MBRL were prone to aggregation in the conditions tested. These targets were further assessed by EM screening and small cryo-EM data collections were carried out for Hrd1 and Gp78. 2D class averages for Gp78 did not resolve any secondary structure detail, which is usually a prerequisite for our lab to warrant a larger data collection. This could be attributed to the target being too small to resolve by cryo-EM, conformational heterogeneity or partial denaturation during cryo-EM grid vitrification. However, minor secondary structure detail was observed after a small data collection and so Hrd1 was selected for a larger data collection. In summary, only mammalian Hrd1 was immediately amenable to structure determination and the other targets require more optimisation before cryo-EM analysis. The results of the Hrd1 data collection will be discussed in the next chapter.

4. Cryo-EM and structural analysis of mammalian Hrd1

Since the Hrd1 sample showed promising 2D class averages from a small data collection, a larger dataset was collected on a 300 kV microscope. At a later date at the NCI Center for Structural Biology, a truncated variant of Hrd1 was also analysed by cryo-EM. Both structures yielded complementary results and will be presented here.

4.1. Cryo-EM of the Hrd1 E3 ligase

4.1.1. Sample preparation

Originally, full-length Hrd1 was prepared and imaged in Oxford. This sample was purified in DDM-CHS (10:1) and tags were cleaved prior to SEC, as described in chapter 3. Material corresponding to a fraction of the main peak was concentrated to an A280 of 7.0 using a 100 kDa MWCO centrifugal concentrator, at 4°C, immediately prior to preparation of cryo-EM grids.

Later in the project, a truncated construct of Hrd1 was expressed and purified and cryo-EM data were collected. This sample was prepared exactly as the full-length sample and corresponds to a large deletion of the soluble, cytoplasmic domain of Hrd1 (figure 4.1a). The C-terminal region contains the RING domain for substrate ubiquitination as well as domains for recruitment of other ERAD factors [61]. Hrd1 is 617 residues long and residues 345-617 were deleted to yield a variant that contained only the transmembrane domains (TMDs) and the cytosolic RING domain. With a combination of AlphaFold, previous structural studies on yeast Hrd1 and our full-length Hrd1 cryo-EM map, this region was expected to be predominantly disordered and flexible and so would not be visible after cryo-EM processing, more of which is discussed later. An analysis of protein disorder also corroborates this as the majority of the C-terminus is disordered (figure 4.1c). This Hrd1 truncation eluted later by SEC and the resultant material was

homogenous in the same detergent conditions as full-length Hrd1 (figure 4.1a-b). The peak fraction from this sample was also concentrated in a 100 MWCO kDa centrifugal concentrator to an A280 of 7.0 and applied to cryo-EM grids under similar conditions as Hrd1, apart from slight changes to temperature and blot force necessary to achieve optimal ice thickness.

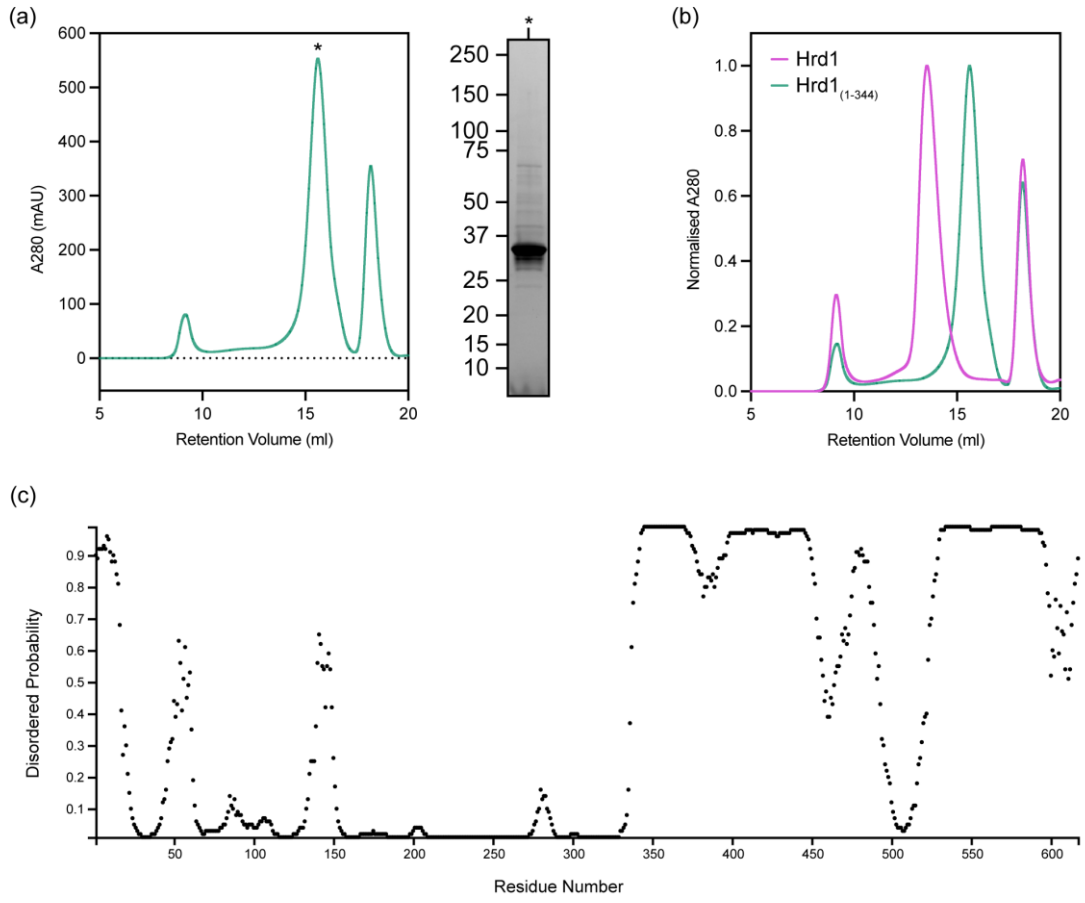


Figure 4.1. Purification of $Hrd1_{1-344}$ and a comparison with full-length Hrd1. (a) SEC profile of $Hrd1_{1-344}$ after StrepTactin purification in DDM-CHS (10:1). The annotated peak fraction was also analysed by SDS-PAGE. (b) Comparison of the SEC profiles of full-length (purple) and truncated Hrd1 (green). These two proteins run approximately 2.5 ml apart under identical chromatography conditions. (c) Disorder analysis of full-length Hrd1 with DISOPRED3, showing the disorder probability for each residue position [138].

4.1.2. Data collection

The full-length Hrd1 dataset was collected in Oxford at the Central Oxford Structural Molecular Imaging Centre (COSMIC). The $Hrd1_{1-344}$ dataset was collected at the NCI Center for Structural Biology, Cryo-EM facility. Data collection parameters and processing statistics are reported in table 4.1. Justin Deme and Susan Lea collected data for all Hrd1 samples. In Oxford, Hrd1 was imaged using a Titan Krios G3 Microscope (FEI) operating at 300 kV, equipped with a K3 direct electron detector (Gatan) and a BioQuantum energy filter (Gatan). The K3 detector was used in counting mode at a magnification of 105,000 x with a pixel size of 0.832 Å. A total dose of 54 $e^-/Å^2$ was fractioned over 40 frames. Data were collected over two separate sessions with these parameters. An initial 15,145 movies were collected followed by an additional 19,516 movies.

For the $Hrd1_{1-344}$ dataset, data were collected on a Titan Krios G4 microscope (FEI) operating at 300 kV, equipped with a cold field emission gun (Thermo Fisher Scientific), Selectris X imaging filter (Thermo Fisher Scientific) and a Falcon 4 direct electron detector (Thermo Fisher Scientific). Data were collected in counting mode in EER format at 165,000 x magnification with a pixel size of 0.693 Å. In total, 15,145 movies were collected, each with a total dose of 54.7 $e^-/Å^2$.

Sample	Hrd1 full-length	$Hrd1_{1-344}$	
Collection Location	COSMIC, UK	NCI, USA	
Data collection			
Magnification	105,000	165,000	
Voltage (kV)	300	300	
Electron exposure (e-/Å2)	54.0	54.7	
Defocus range (μm)	-0.5 to -2.5	-0.5 to -2.5	
Pixel size (Å)	0.832	0.693	
Data processing			
Map	Hrd1 full-length	$Hrd1_{1-344}$ (open)	$Hrd1_{1-344}$ (closed)
Symmetry imposed	C2	C2	C2
Initial particle images (no.)	13,635,181	3,659,052	3,659,052
Final particle images (no.)	530,152	273,396	225,767
Map resolution (Å)	3.44	3.38	3.27
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	2.4-4.0	2.8-5.2	2.8-7.0
Refinement			
Initial model used (PDB code)	None	None	None
Model resolution (Å)	3.44	3.38	3.27
FSC threshold	0.143	0.143	0.143
Model resolution range (Å)	2.4-4.0	2.8-5.2	2.8-7.0
Model composition			
Non-hydrogen atoms	3864	5446	4848
Protein residues	468	668	590
Ligands	0	0	0
B factors (Å2)			
Protein	35.9	72.9	74.96
Ligand	N/A	N/A	N/A
R.m.s. deviations			
Bond lengths (Å)	0.002	0.001	0.002
Bond angles (°)	0.372	0.334	0.448
Validation			
MolProbitiy score	1.6	1.2	1.47
Clashscore	5.37	3.19	4.4
Poor rotamers (%)	0	0	0
Ramachandran plot			
Favoured (%)	95.47	97.59	96.25
Allowed (%)	4.53	2.41	3.75
Disallowed (%)	0	0	0

Table 4.1. Cryo-EM data collection, refinement and validation statistics for full-length Hrd1 and $Hrd1_{1-344}$ in DDM-CHS.

4.1.3. Data processing

All micrograph processing was performed using the SIMPLE pipeline [162]. Motion was always corrected for in patches (20x20) with SIMPLE-unblur and patched CTF estimation performed with SIMPLE-CTFFIND. Particle picking was automatically done with SIMPLE-picker and was always used with a lowpass-filtered (20 Å) reference volume of an unrelated membrane protein of similar dimensions from other projects within our group (Lea lab). After initial 2D classification to remove "junk" particles, all later steps were performed in either cryoSPARC or RELION-3.1 [165,166]. Interconversion from cryoSPARC to RELION data was done using the UCSF pyem scripts [167]. Resolution estimates are reported from cryoSPARC and derived using gold-standard Fourier shell correlations (FSCs) with the 0.143 criterion. Directional resolution curves were calculated with the cryoSPARC wrapper for 3DFSC [168].

The data processing workflow for full-length Hrd1 is described in figure 4.2. Briefly, a smaller dataset (15,145 movies; 3,659,052 particles) was initially collected on a 300 kV Titan Krios microscope. After movie pre-processing and particle extraction with a picking template, 2D class averages were generated with SIMPLE's stream2D [162]. 2D classes that vaguely resembled protein in detergent micelles were selected and exported to cryoSPARC for further processing [165]. Additional rounds of 2D classification were performed in cryoSPARC and 260,886 "good" particles were used to generate *ab initio* models. For this sample, cryoSPARC 2D classification was always performed with a circular mask diameter of 160 Å, 40 online-EM iterations and a 200 batchsize per class. Four *ab initio* models were generated of which ~50% particles classified into a map with clearly defined transmembrane helices. Since secondary structure detail could be seen in 3D, this was a promising sign of a sample that could go to high-resolution and so more data were collected. These movies were pre-processed in SIMPLE separately but in the

same way as described previously. Cleaned particles were imported into cryoSPARC and subject to further rounds of 2D classification. Together with the particles from the initial dataset, all particles were classified against the 4 *ab initio* volumes using cryoSPARC's heterogeneous refinement classification. These volumes were lowpass-filtered to 8 Å prior to classification. Most particles classified into two prominent classes (2 and 3), representing ~40% and ~30% of the total particles, respectively. Extensive processing of particles in class 3 did not yield high resolution information. During 2D classification prior to heterogeneous refinement and in 2D classification on the particles in class 3, a distinctly different arrangement of TMs could be seen within the micelle (figure 4.2b). This led us to believe that there was a second conformation present in the dataset that could not be resolved at the time, but for which there was enough signal in both 2D and 3D classifications to separate them from the main conformation. Class 2 particles from the heterogeneous refinement were further subjected to non-uniform (NU) refinement in both C1 and C2 symmetry. For all NU-refinements of this sample, the initial volume was always lowpass-filtered to 8 Å and 3 extra final passes were performed. A lowpass-filter was used to prevent map bias during early refinement stages. The filter was reduced from the default 30 Å since refinements for Hrd1 using a very low-resolution starting map resulted in worse final resolution. Refinement against a 30 Å map likely leads to alignment on the micelle shape whereas refinement against an 8 Å map leads to alignment on the TM helices. Maps generated from C2-symmetrised refinements showed improved map details compared to C1 refinements and yielded higher nominal resolutions as assessed by gold-standard FSC curves, so C2 symmetry was used in all subsequent refinements. Per-particle motion was corrected for with Bayesian polishing in RELION and particles re-imported to cryoSPARC for an additional round of 2D classification [166]. This additional round of 2D was used to remove artefacts generated during the Bayesian polishing process. Using the polished particles, an NU-refinement was performed using the pre-polished map as reference. This was then locally refined in

cryoSPARC with a mask enveloping the protein part of the map. Per-particle defocus estimation and CTF refinements were then performed before a final round of local refinement with the same protein mask used earlier. This resulted in a final map with a resolution of 3.44 Å (figure 4.2c) and good local resolution for all TMDs (figures 4.2d and S3).

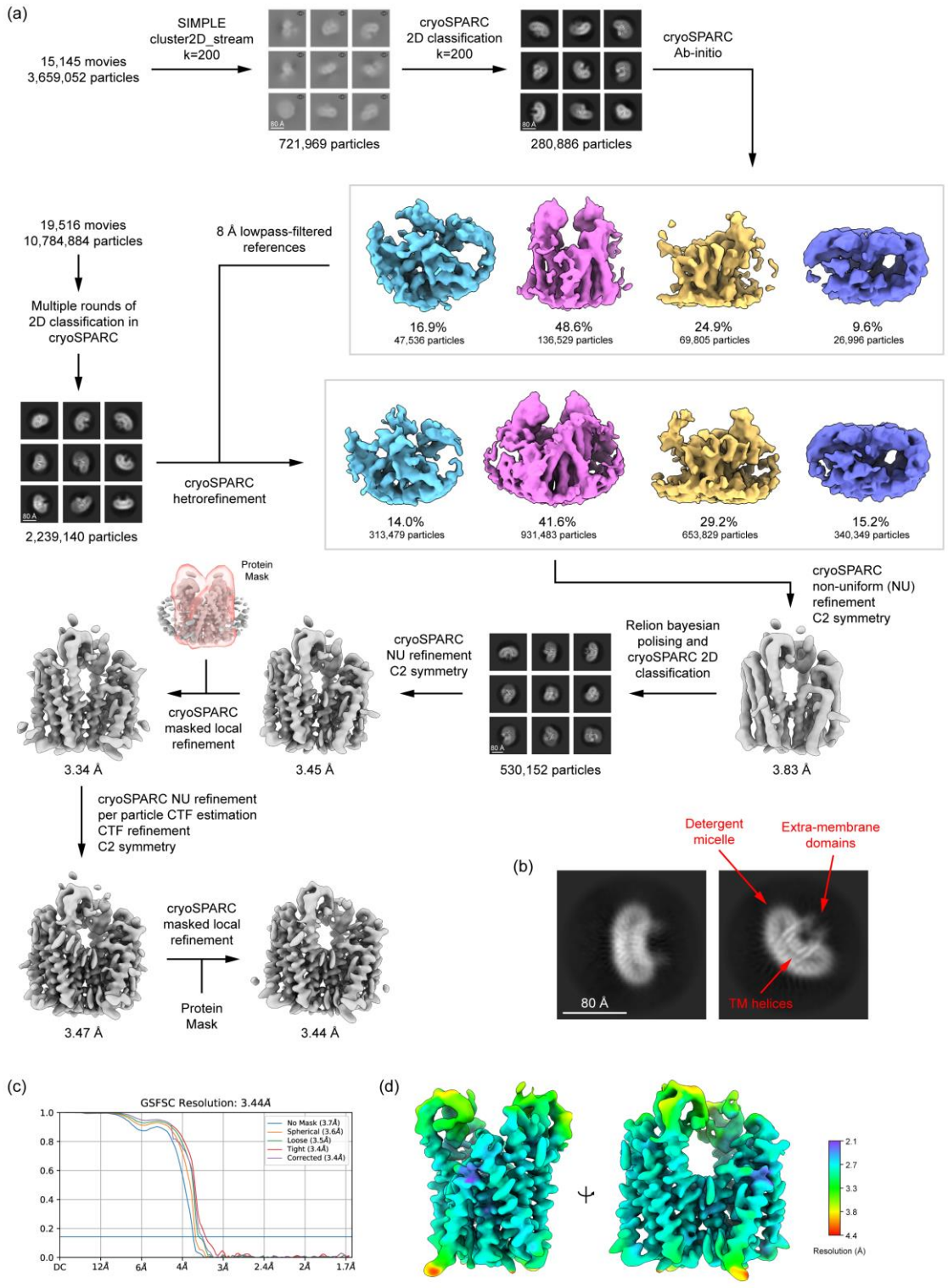


Figure 4.2. Cryo-EM processing workflow, showing local map quality, for the full-length Hrd1 dimer. (a) Overall processing workflow in SIMPLE, cryoSPARC and RELION. (b) 2D class averages of two distinct Hrd1 conformations resolved prior to heterogenous

refinement. Only the particles corresponding to the right conformation could be resolved to high-resolution. The conformations differ in the arrangement of their extra-membrane domains (RING domain) as well as the pitch of the TMDs within the micelle. (c) FSC curves for global-resolution estimation of the final map, using the 0.143 criterion for reporting resolution. (d) Local resolution estimation of the final map calculated in cryoSPARC, with the 0.143 criterion.

For the $Hrd1_{1\text{-}344}$ dataset, cryo-EM data were processed as in figure 4.3. Movies (15,145) were motion corrected, CTF estimated and particles picked in SIMPLE consistent with the full-length dataset. “Good” 2D class averages were selected and exported to cryoSPARC for further processing. 2D classification in cryoSPARC with this sample was always performed with a 180 Å mask diameter, 40 online-EM iterations and a batchsize per class of 500. The number of 2D classes assigned varied between 50 and 200, depending on the number of particles being classified. Increasing the batchsize per class to 500 improved the quality of the 2D alignments, likely due to an increase in signal-to-noise at the start of each round of 2D classification. Particles were subject to multiple rounds of 2D classification to further select for good particles before *ab initio* generation of 3 different models. This yielded two classes with secondary structure features, a key difference from the previous full-length Hrd1 dataset. Compared to the Hrd1 dataset (figure 4.2), this new class corresponded to the previously unresolvable class in the initial *ab initio* and heterogeneous refinement outputs, albeit at a now higher resolution. This resolves a new state for the Hrd1 dimer and was refined in tandem with particles corresponding to the previously resolved state. *Ab initio* class 3 eventually resolved to be the same map as that solved from the full-length dataset. This *ab initio* class was subject to an NU-refinement with C1 symmetry before an additional NU-refinement with C2 symmetry that further improved resolution and map quality. Particles were subject to Bayesian polishing in RELION to correct for per-particle motion before being re-imported

into cryoSPARC for further 2D classification. Good particles were used again in an NU-refinement with C2 symmetry yielding a higher resolution reconstruction. To further classify heterogeneity in the dataset, an additional round of multi-class *ab initio* was performed. This yielded a single high-resolution class with ~70% of the total particles that was then used in a C2 NU-refinement against a lowpass-filtered version of the *ab initio* model as a reference. This resulted in a final map with a nominal resolution of 3.38 Å, as determined by gold-standard FSC curves in cryoSPARC. To further enhance the map for modelling, deepEMhancer was used to sharpen the map with the highRes model [169]. DeepEMhancer uses a convolutional neural network trained on experimental and post-processed cryo-EM maps to enhance map detail by automatic map sharpening [169]. This enhanced the resolution of the domains outside of the detergent micelle (the RING domains) but reduced resolution for non-protein molecules such as lipids or detergent. DeepEMhancer is trained to sharpen protein and RNA features, and so non-protein signal is reduced during processing [169]. A local resolution filtered map was also produced that preserved these non-protein features but with reduced resolution in the RING domains. Processing of the other initial *ab initio* class resolved an alternative conformation of the Hrd1 dimer, which we term the “open” state. Particles belonging to this class were subjected to NU-refinement first with C1 symmetry and then with C2 symmetry, which again improved map quality and resolution. These particles were subjected to a second round of *ab initio* generation which yielded 2 similar high-resolution classes. Particles from both these classes were combined and used in a C2 NU-refinement against the map from class 1 of the initial *ab initio* to yield a slightly improved map. Particles were then corrected for per-particle motion with Bayesian polishing in RELION and re-imported into cryoSPARC followed by 2D classification to remove junk particles generated during this process. C2 NU-refinement of these particles against the previous map yielded a map with improved resolution. Particles were subject to a third round of *ab initio* generation to further discriminate heterogeneity. Particles belonging to

the most populated class were selected out and used in a C2 NU-refinement against its corresponding map, lowpass-filtered to 8 Å. The resulting map had stronger density for non-protein molecules and had a final nominal resolution of 3.27 Å. deepEMhancer was used with the highRes model to enhance resolution of the RING domains. A local resolution filtered map was also generated as before, to preserve other details within the map. A combination of all these maps were used together for model building.

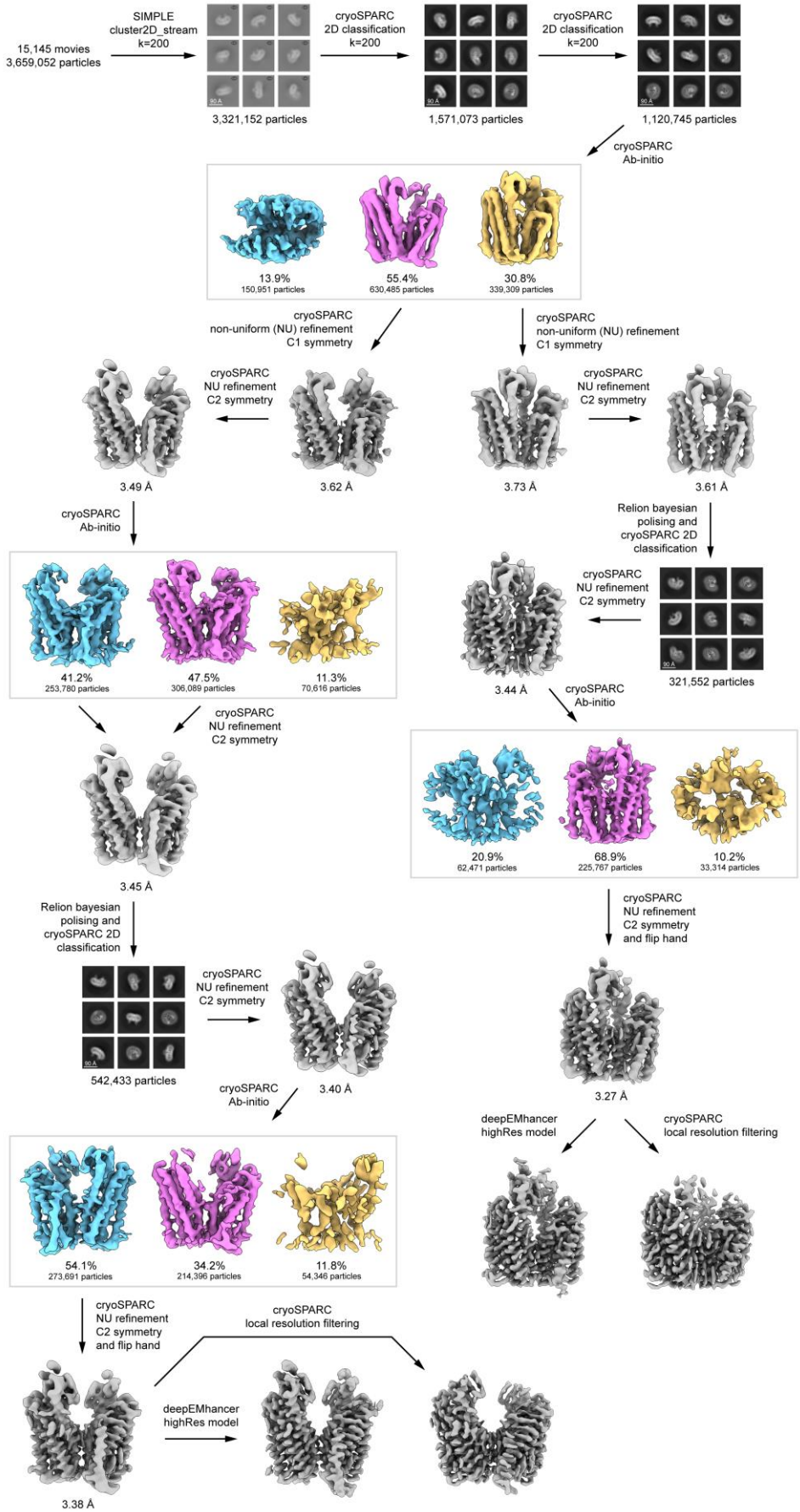

15,145 movies
3,659,052 particles
SIMPLE
cluster2D_stream
k=200
3,321,152 particles
cryoSPARC
2D classification
k=200
1,571,073 particles
cryoSPARC
2D classification
k=200
1,120,745 particles
cryoSPARC
Ab-initio
13.9%
150,951 particles
55.4%
630,485 particles
30.8%
339,309 particles
cryoSPARC
non-uniform (NU) refinement
C1 symmetry
cryoSPARC
non-uniform (NU) refinement
C1 symmetry
cryoSPARC
NU refinement
C2 symmetry
3.49 Å
3.62 Å
cryoSPARC
NU refinement
C2 symmetry
3.73 Å
3.61 Å
cryoSPARC
Ab-initio
Relion bayesian
polising and
cryoSPARC 2D
classification
41.2%
253,780 particles
47.5%
306,089 particles
11.3%
70,616 particles
cryoSPARC
NU refinement
C2 symmetry
3.44 Å
321,552 particles
cryoSPARC
Ab-initio
cryoSPARC
NU refinement
C2 symmetry
3.45 Å
20.9%
62,471 particles
68.9%
225,767 particles
10.2%
33,314 particles
Relion bayesian
polising and
cryoSPARC 2D
classification
cryoSPARC
NU refinement
C2 symmetry
542,433 particles
cryoSPARC
NU refinement
C2 symmetry
and flip hand
cryoSPARC
Ab-initio
3.40 Å
3.27 Å
deepEMhancer
highRes model
cryoSPARC
local resolution filtering
54.1%
273,691 particles
34.2%
214,396 particles
11.8%
54,346 particles
cryoSPARC
NU refinement
C2 symmetry
and flip hand
cryoSPARC
local resolution filtering
deepEMhancer
highRes model
3.38 Å

Figure 4.3. Cryo-EM processing workflow for the Hrd1$_{1\text{-}344}$ dimer. SIMPLE, cryoSPARC and RELION were used to process the data.

Through collection of two different Hrd1 samples, different dimeric conformations of Hrd1 were identified. Cryo-EM of full-length Hrd1 resulted in a 3.44 Å map of what we term the “closed” state, where the 2 RING domains are closer together and the TMD helices more perpendicular to the membrane. This map had overall better quality for non-protein density (probable lipids or detergent molecules) surrounding the TMDs compared to the Hrd1$_{1\text{-}344}$ maps. Cryo-EM of Hrd1$_{1\text{-}344}$ resolved two conformations to high resolution, a closed state exactly like the full-length map but at an improved resolution of 3.27 Å, and a new “open” state at 3.38 Å (figure 4.4). Although these maps are at a higher average resolution than the full-length map, non-protein density was not resolved as well. This could be attributed to the sample being different or that there are fewer particles in the final maps versus the full-length map (approximately half as many, table 4.1). This could also arise from heterogeneity in the dataset where some of the final particles contain these non-protein densities and others do not, resulting in weak density overall. Alternatively, these may be a result of higher resolution refinements where less ordered molecules can be filtered out as the volume incorporates higher resolution terms. However, the Hrd1$_{1\text{-}344}$ maps had overall better resolution for the cytosolic RING domains and these were resolved best in the open state, possibly because there is less direction-dependent resolution bias in the particles (figures 4.4c and 4.4f). Orientation preference (directional-bias) of vitrified protein particles is a common problem in cryo-EM and results in a loss of resolution in the axis parallel to the preferred orientation [168]. The open conformation is likely also in the full-length dataset but it was not resolved to high resolution. This could be sample related or due to differences in ice quality between the datasets. There is also a shift in the distribution of particles between the open/closed conformations in the Hrd1$_{1\text{-}344}$ and full-length datasets. In the Hrd1$_{1\text{-}344}$ dataset, the initial

ab initio round had ~55% of particles in the open state, whilst ~30% of particles were in the closed state (figure 4.3). Conversely, in the initial *ab initio* generation for the full-length dataset ~50% of particles clustered together in the closed state whilst ~25% of particles clustered together in the probable open state. This suggests that truncated Hrd1 behaves different on cryo-EM grids to the full-length sample. One rationale for this is that the view distribution of frozen particles is better with the C-terminal truncation and so the open state can be better resolved by 3D *ab initio* algorithms. Alternatively, the C-terminus of Hrd1 may be important in stabilisation of one of these states, but this is hard to rationalise since the C-terminal region was not visible in the full-length Hrd1 map. Mutational studies in yeast do however support a role for regions of the disordered C-terminal in substrate degradation [89].

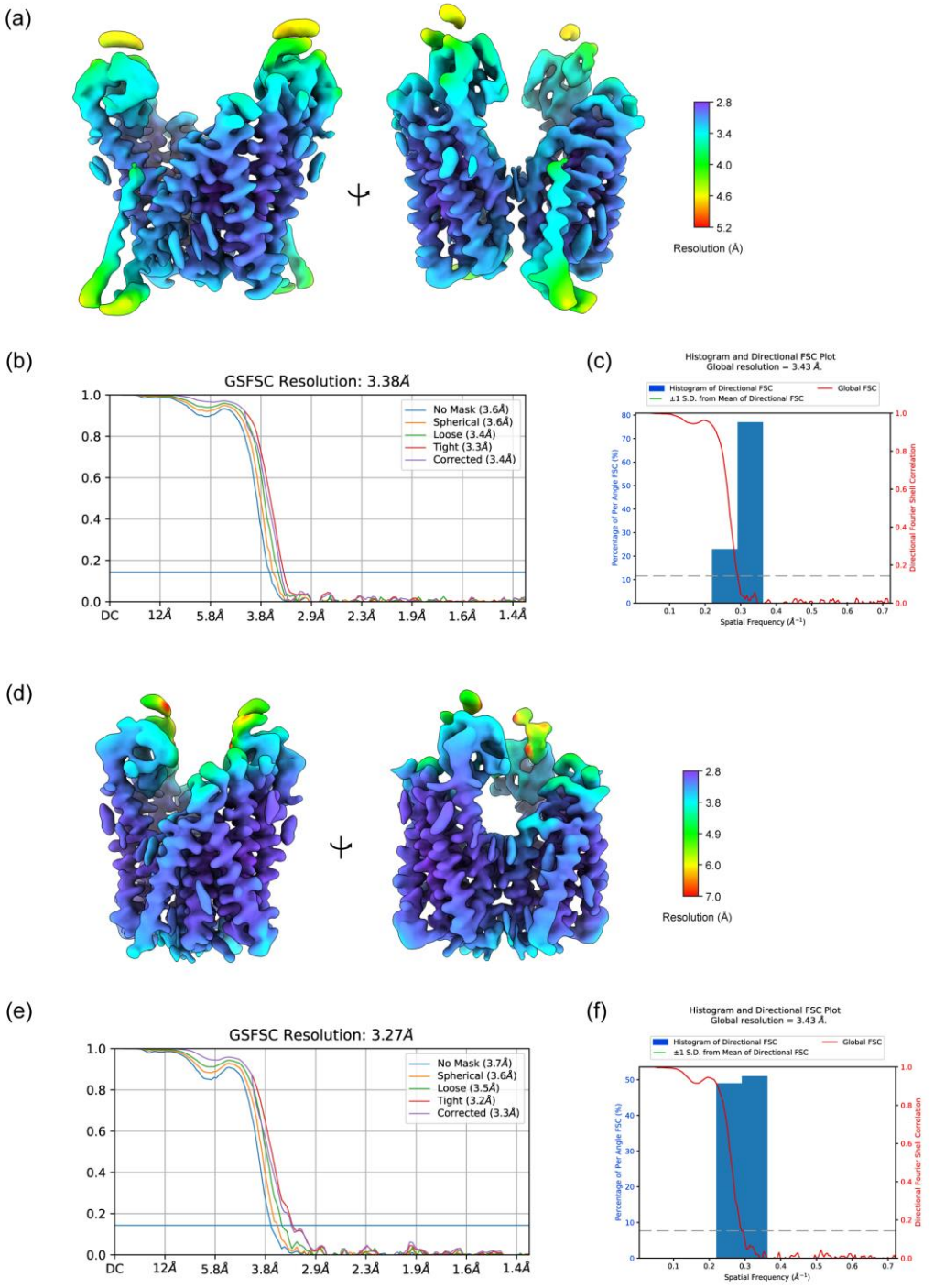


Figure 4.4. Local map quality and resolution for the Hrd1$_{1\text{-}344}$ dimer maps. (a) Local resolution estimation for the open dimer state, calculated with cryoSPARC using the 0.143 cut-off. (b) Corresponding FSC curve for the open state map where resolution is reported using the 0.143 criterion. (c) 3DFSC plot for the open map showing correlation of resolution with respect to direction. (d) Local resolution estimation of the closed dimer

state. (e) Corresponding FSC curve for the closed state map where resolution is reported as above. (f) 3DFSC plot for the closed map.

4.1.4. Model building

Together, these datasets were sufficiently detailed to build a model for Hrd1 with its TMDs and RING domain in the open and closed state. For the full-length Hrd1 dataset, a model for the closed state was built from the both the final unsharpened and globally sharpened, locally refined map at 3.44 Å. For the $Hrd1_{1\text{-}344}$ dataset, models for the open and closed state were built into the deepEMhancer and local resolution sharpened maps. Prior to model building, AlphaFold was used with the default settings to generate a model for full-length Hrd1 alone [136]. The top ranked AlphaFold model was then docked into each of the maps using the fitmap function of ChimeraX [170]. This was then subject to multiple rounds of manual building using Coot v.0.9.4 and real-space refinement in PHENIX v1.19.2-4158, using rotamer and Ramachandran restraints [171,172]. PHENIX refinements were performed with the sharpened map for modelling the TMDs and the deepEMhancer map for the RING domains. This resulted in models with statistics reported in table 4.1. The final models for the open and the closed conformations are compared with the Hrd1 AlphaFold model in figure S1. Fits for the TMDs in the closed and open maps are shown in figures S2-3.

4.2. Structural analysis of Hrd1

4.2.1. Prior structural work on yeast Hrd1

Mammalian Hrd1 shares sequence similarity (29% identity) with yeast Hrd1, for which there are numerous structures [55,60]. Initially, a Hrd1-Hrd3 complex was reconstituted into amphipols and imaged by cryo-EM, resulting in a symmetric dimer [60].

The resulting map was surprising since density for an amphipathic helix of Hrd3 significantly protruded into what would be the expected lipid bilayer, despite it not being predicted to be a TM helix [60]. The Hrd1-Hrd3 dimer structure was later refuted when a structure of yeast Hrd1-Hrd3 solubilised in DMNG/digitonin was solved by cryo-EM [55]. This structure was monomeric and was shown to be more active during *in vitro* substrate ubiquitination assays [55]. Importantly, the amphipathic helix of Hrd3 was not located deep inside the detergent micelle and thus monomeric Hrd1-Hrd3 deemed the physiologically state [55]. Placement of the amphipathic Hrd3 helix and arrangement of the Hrd1-Hrd3 dimer was thought to be an amphipol stabilisation artifact, resulting in a non-physiological structure [60]. At the same time, a structure of the yeast Hrd1-Der1 heterodimer was solved, using a fusion of Hrd1 with Usa1 (a required partner for Der1 recruitment). This complex also contained Hrd3, for which there is only one copy, and is thought to be minimally necessary for ERAD-L degradation [55,163]. However, endogenous Hrd1 in both yeast and mammals is known to form higher-order oligomers, some of which correspond to dimeric Hrd1 [62,91]. The significance of this is unknown and no structural work had been performed on mammalian Hrd1.

4.2.2. Hrd1 closed conformation

The closed state map of full-length Hrd1 and $Hrd1_{1-344}$ were of sufficient resolution to accurately build a model for Hrd1 (figure 4.5a). The RING domain of the Hrd1 map was less resolvable than in the $Hrd1_{1-344}$ map so the latter was used to model residues 42 to 336 of mammalian Hrd1. This included 7 of the 8 predicted TM helices and the RING domain. To do this, the deepEMhancer map was used to model the AlphaFold predicted RING domain in place. The AlphaFold model for the RING domain agreed with a previous NMR structure of the Hrd1 RING domain (RMSD of ~1.2 Å, figure S4) and so was a confident starting point for approximating the shape of the RING domain [143].

Importantly, this is the region of the map where resolution is worst and so side chain confidence in this area is least accurate. For the rest of the model, side chains could be modelled accurately at this resolution (figure S2). This model of Hrd1 is a dimer with C2 symmetry and is termed the closed conformation because the cytosolic RING domains are positioned closer to each other (17 Å separation, figure 4.5a). From sequence analysis and structures of yeast Hrd1, mammalian Hrd1 is thought to have an additional TM helix at the N-terminus, linked by a connecting luminal helical segment important for binding of Sel1L [55]. The AlphaFold predicted model also contains this TM helix (figure S1), further hinting at its presence. In the maps for the closed conformation, only TM helices 2-8 could be modelled and there was no additional density for TM helix 1. Even at lower map contour levels, where signal-to-noise is reduced, no density for this helix can be seen. In this conformation, this helix could be too flexible, its position too variable with respect to the micelle across different protein molecules or not folded within the micelle. All of these would result in a lack of visible density for the helix in the cryo-EM map. In this model, the 7 TM helices arrange together to support the cytosolic RING domain. For the full-length Hrd1 volume there was no apparent density for the cytosolic region after the RING domain, which is predicted to contain a high level of disorder (figure 4.1c). This was part of the rationale for designing the truncated construct. The RING domains are positioned proximal to the membrane and directly over the helical bundles within each Hrd1 monomer. Mammalian Hrd1 forms a dimer that is predominantly driven by the inter-protein interactions of TMD3 (figure 4.5a). Co-evolutionary couplings at the dimer interface between two Hrd1 molecules were identified with EVcouplings but these are hard to decipher from couplings occurring within the same TM helix or between adjacent helices of the monomer [173]. The final model for the Hrd1 monomer has a high resemblance to the AlphaFold predicted model, with an RMSD of only 1.5 Å (figure S1).

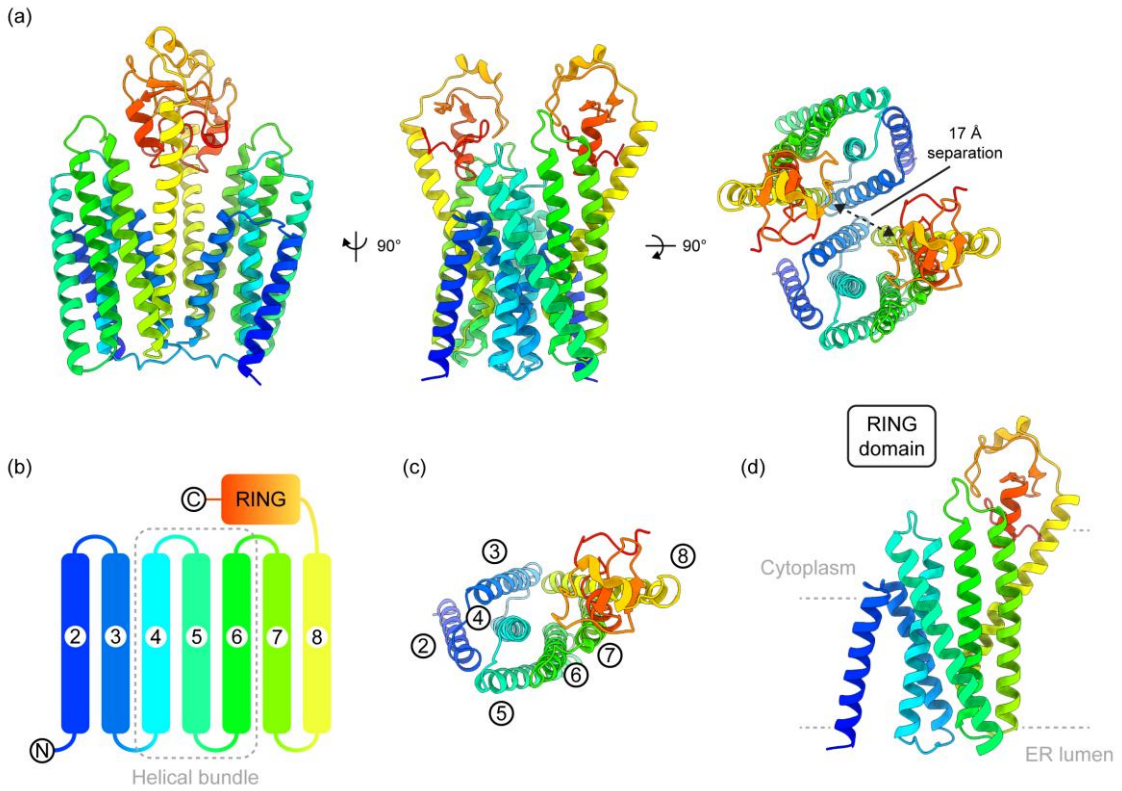


Figure 4.5. Model for the Hrd1 closed conformation. (a) Cartoon representation coloured blue to red from the N- to the C-terminus. (b) Topology schematic with a triple helical bundle indicated. TMD helices are numbered according to their known number in the full-length sequence. (c) Top-down view of a Hrd1 monomer with TMD helices annotated. (d) Monomeric Hrd1 with approximate membrane boundaries shown, as determined by TMD hydrophobicity in ChimeraX [170].

Boundaries of the lipid bilayer were determined by analysis of TMD hydrophobicity with ChimeraX and were approximately annotated in figure 4.5d [170]. Across the monomer TMDs, the height of the predicted bilayer changes significantly from TMDs 2 through 8. The length of these helices varies from 22 to 32 residues which is a significant difference. Although TMD8 is tilted relative to the bilayer, a mixture of hydrophobic residues and charged residues in the membrane and at the membrane interface, respectively, likely cause membrane thinning across the Hrd1 dimer. Residue

conservation of Hrd1 was analysed by multiple-sequence alignment with ConSurf [174]. Conservation was strongest predominantly at the base of the monomer, between TMDs 3, 6 and 8 (figure 4.6a). Residues on TMDs 3 and 8 face the opposing monomer in the Hrd1 dimer and so indicate conservation of a Hrd1 dimer. However, this interface also recruits Der1 in the yeast Hrd1-Der1 heterodimer, and so could represent a conserved interaction sites for partners [55]. Conserved residues in this region could also be important for stabilisation of the overall Hrd1 fold. Strong conservation is also seen in the RING domain which is an evolutionary conserved domain involved in substrate ubiquitination [38]. Importantly, the conserved RING domain is positioned directly above the TM helical core of Hrd1 by TMD8. Its position may be critical for substrate ubiquitination in the Hrd1 complex, where Hrd1 is thought to complex with partners to form a retrotranslocon. The opposite side of the dimer interface is weakly conserved, consistent with structural studies of yeast Hrd1 in which no interactors bound at this interface thus suggesting a similar recruitment strategy for mammalian Hrd1 partners [55,60]. There are also a series of charged, conserved residues (R246, R253) lining the cytoplasmic-facing side of TMD8, up towards the RING domain (figure 4.6b). These residues are involved in formation of a hydrophilic pocket beneath the RING domain. This hydrophilic pocket extends partially between the TM helices and into the membrane. These residues may be involved in Hrd1 functions such as positioning substrates for ubiquitination on the cytoplasmic side or for interactions with other ERAD factors, both of which should be explored in future studies on mammalian Hrd1. Regardless, this hydrophilic pocket is strongly conserved and likely plays a role in Hrd1 function.

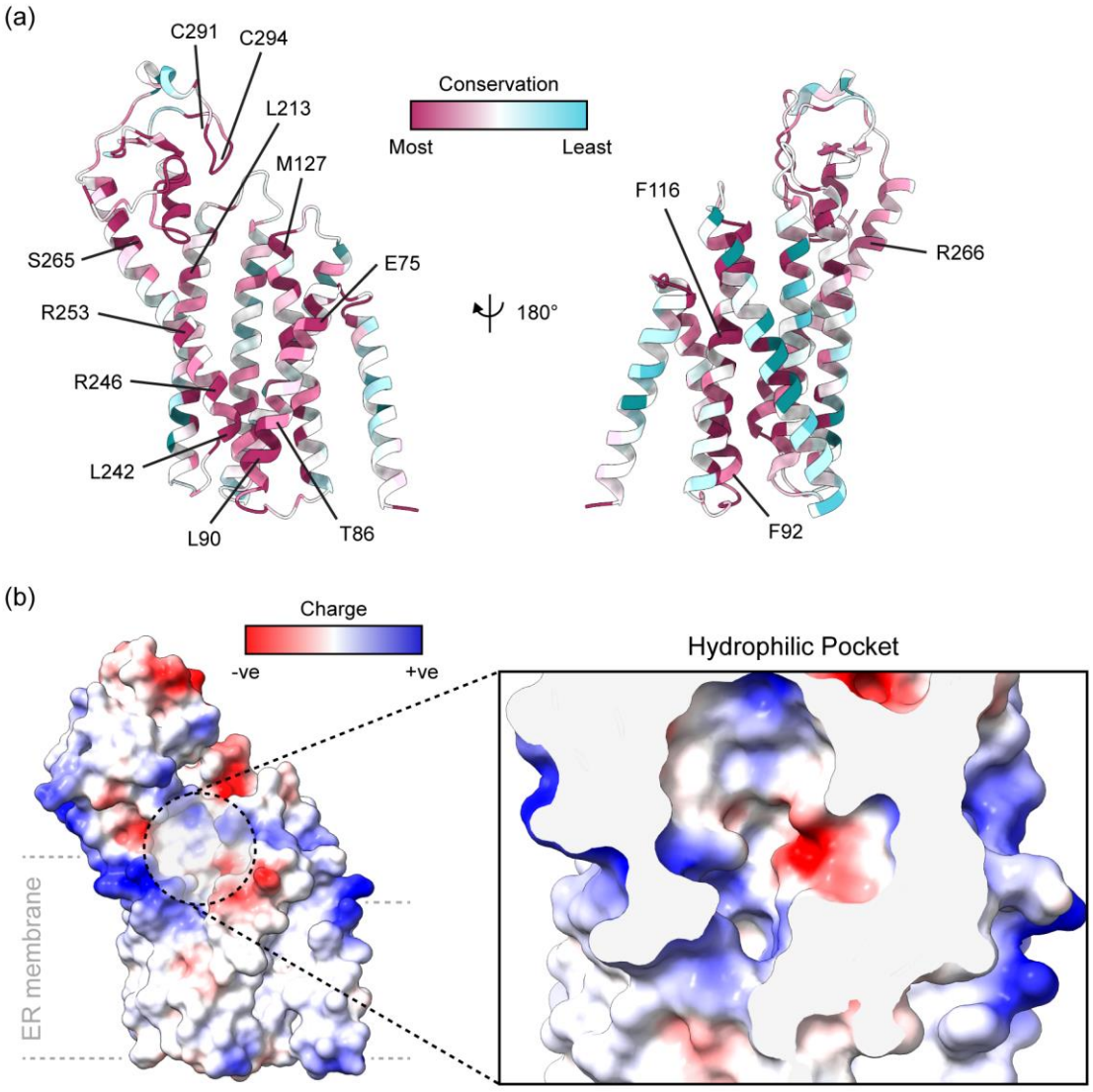


Figure 4.6. Features of the Hrd1 monomer. (a) A Hrd1 monomer coloured by conservation. Purple regions are most conserved whilst turquoise regions are least conserved. (b) Electrostatic potential surface map of a Hrd1 monomer in the closed conformation. Negatively charged surfaces are coloured red, neutral surfaces white, and positively charged surfaces blue. This monomer is orientated looking in from the dimer interface. Closer inspection reveals a hydrophilic pocket on the cytoplasmic side of the membrane, close to the RING domain.

4.2.3. Hrd1 open conformation

Also resolved in the $Hrd1_{1\text{-}344}$ dataset was an alternate conformation that could not be resolved to high resolution in the full-length dataset, although there were side views of 2D classes consistent with this conformation in the dataset. This conformation was also a C2 symmetric dimer and was of sufficient resolution to model side chains accurately throughout the TM helices (figures 4.7 and S3). The overall resolution of this map was 3.27 Å but the RING domain was the poorest resolution, with parts of it closer to 4 Å. While side chains could not be accurately modelled into the RING domain density, an AlphaFold model was loosely fit and minor adjustments made during Phenix refinements. We term this map the “open” conformation because of drastic changes at the dimer interface which lead to the RING domains being positioned distal from each other (31 Å separation, figure 4.7a). One other important difference in this model is the presence of TMD1, which was previously unobserved in both the full-length Hrd1 and $Hrd1_{1\text{-}344}$ closed conformer volumes. This allowed us to model residues 3 to 336 and build an almost complete model for TMD1 with high confidence (figure S3). In this open conformation, each Hrd1 monomer is now tilted with respect to the membrane (figure 4.7d). This has implications for the membrane environment which are discussed later.

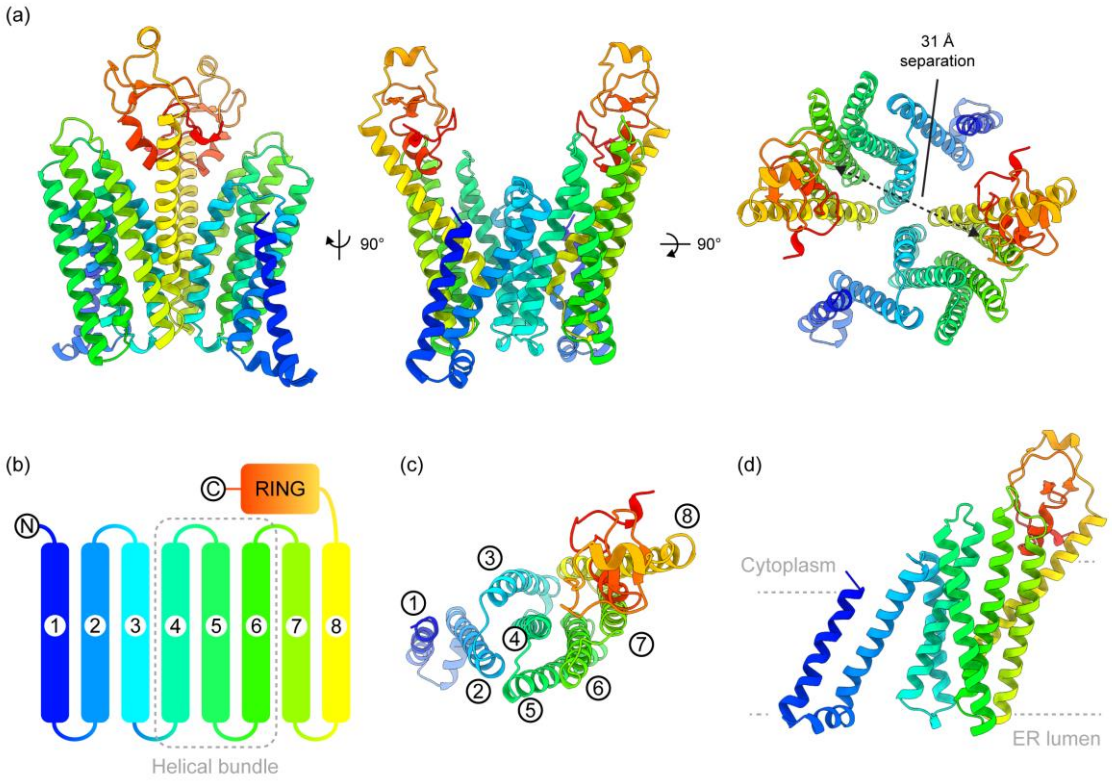


Figure 4.7. Model for the Hrd1 open conformation. (a) Cartoon representation of the open conformation coloured blue to red from the N- to the C-terminus. (b) Topology schematic depicting the TMDs and the RING domain. (c) Top-down view of the open conformation as seen in a Hrd1 monomer. TMD helices are numbered with respect to their sequence position. (d) Monomeric Hrd1 and its position in the membrane, as approximated from hydrophobicity surface measurements in ChimeraX.

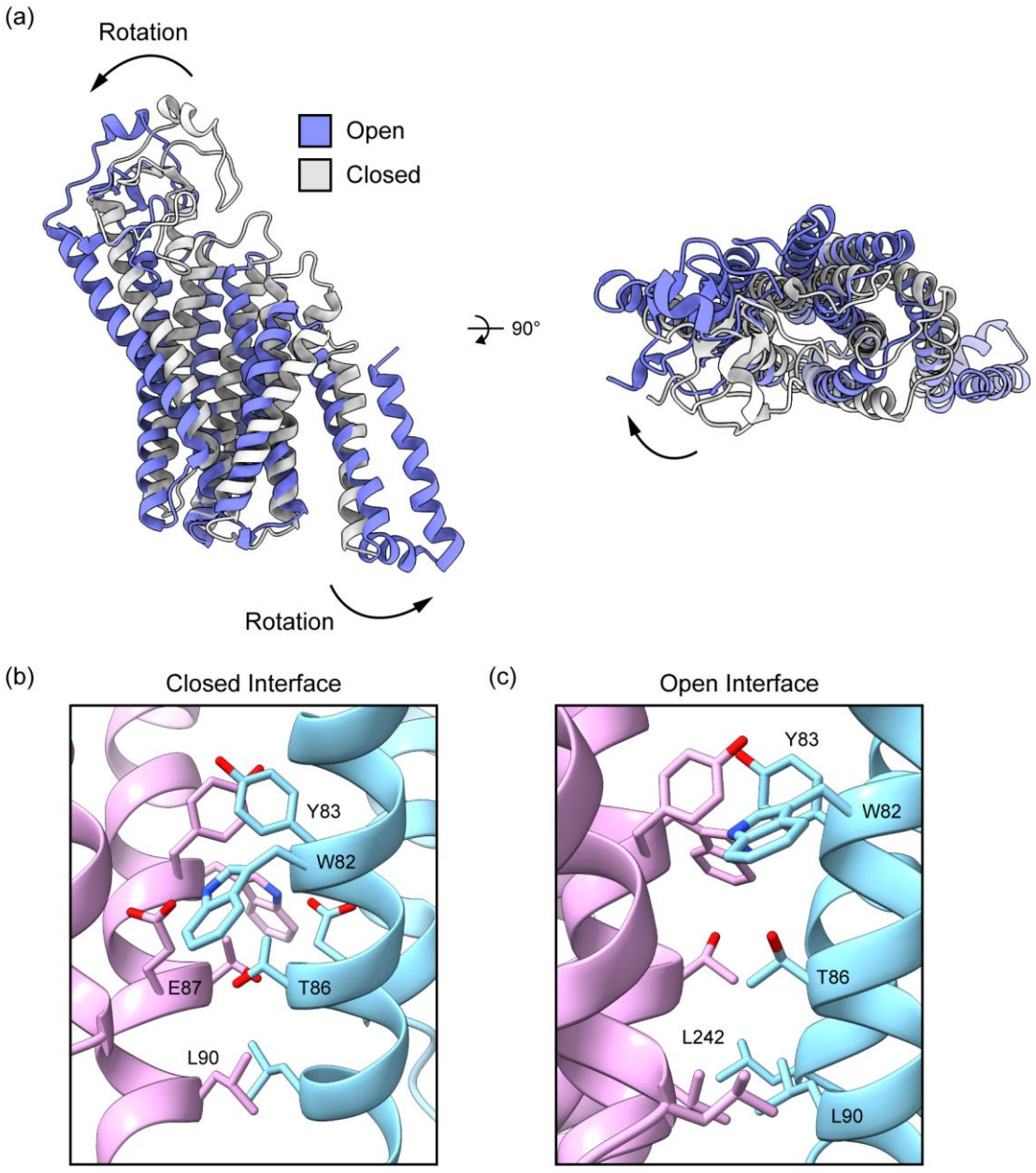


Figure 4.8. Transitioning between an open and closed Hrd1. (a) Overlayed structures of the 2 different Hrd1 conformations, generated with ChimeraX. The open conformation is coloured blue whilst the closed conformation coloured white. (b) Inter-protein residues at the interface of the Hrd1 closed conformation. (c) Inter-protein residues at the interface of the Hrd1 open conformation.

Both the open and closed maps were already orientated in the same axis after cryo-EM processing and so the models were directly compared to understand the transition between the two states (map and model overlays justifying this are shown in figure S5). To transition between the closed and open conformations, the Hrd1 monomers rotate at the dimer interface with respect to each other (figure 4.8a). This twisting is mostly in an axis that is perpendicular to the dimer interface. There is no obvious rearrangement of TMDs of the Hrd1 monomers between the closed and open conformations, and alignment RMSD of the individual monomer backbones of these conformations is 0.99 Å. Although the backbone remains largely the same, side chain residues positions have changed between these two states. To accommodate the transition to a different conformation, inter-protein contacts at the dimer interface are rearranged (figure 4.8b). The interactions at the dimer interface are mediated by TMD3 in both conformations. Here, an ensemble of π-stacking interactions, charged interactions and hydrophobic packing maintain the dimer interface. In the open conformation, there is a greater reliance on hydrophobic interactions at the base of TMD3 where L242 now contributes to dimer stabilisation. There are still π-stacking interactions between W82 and Y83 but there is now no electrostatic contribution of E87 with W82 in the opposing monomer. This rearrangement of the dimer interface essentially allows for an ordered TMD1 to fit inside of the detergent micelle (figure 4.9). The two conformations exhibit significantly different membrane environments, as visualised by the detergent micelle encapsulating the protein (figure 4.9). The open conformation has a predominantly flat detergent micelle, with a small amount of thinning at the dimer interface. The height of the detergent micelle at this interface is ~42 Å. Conversely, in the closed conformation there is significant thinning induced at the dimer interface which results in a micelle height of ~35 Å. In both these conformations, the membrane is thinnest at the dimer interface (figure 4.9). The thinning in the closed state is akin to the yeast Hrd1-Der1 interface, where Hrd1 and Der1 are proposed to each contribute a half-channel for retrotranslocation (figure S6)

[55]. These half-channels are proposed to work together to thin the membrane at the heterodimeric interface. This membrane thinning is thought to be crucial to retrotranslocation activity and so it raises the question as to whether mammalian Hrd1 alone could act as a retrotranslocon during ERAD.

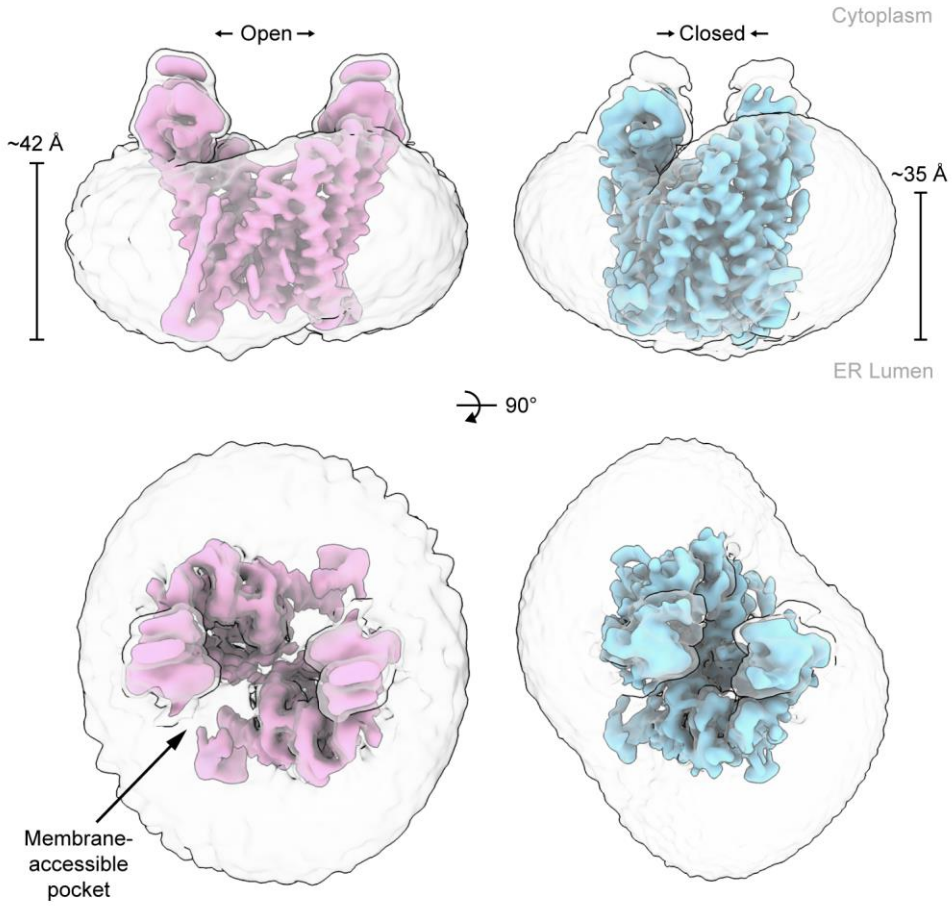


Figure 4.9. Comparison of the membrane environment between the Hrd1 conformations, as visualised from the cryo-EM maps. The detergent micelle is shown as a low contour level map with transparency overlaying a high contour level map of the protein density. Distance measurements represent membrane thickness at the dimer interface.

Yeast Hrd1 is thought to have a lateral gate for ERAD substrates composed of TM helices 3 and 8 [55,60]. This was confirmed by structural analysis as well as crosslinking studies between substrate and components of the Hrd1 complex [55,56]. In both

conformations of mammalian Hrd1, the lateral gate sits at the dimer interface where TMD3 makes contacts with the opposing monomer (figure 4.10a). This would block access to the lateral gate in mammalian Hrd1. A unique feature of the mammalian Hrd1 conformations is the presence of a membrane-accessible pocket between TM helices 2, 3 and 8 (figure 4.9 and 4.10). Although the lateral gate TMD8 is not involved heavily at the dimer interface, it forms a key part of this membrane-accessible pocket. Access to this pocket changes between the open to the closed conformations and in the open conformation, this pocket is deeper because of the presence of a stably associated TMD1. In this conformation the micelle is flatter and so this could represent a substrate accessible conformation where substrates do not have as high a barrier for entry (as the closed state) since the membrane thickness would be closer to the natural ER membrane thickness. TMD1 has been recently been implicated as having a role in the degradation of certain substrates, as identified through deep-mutational scanning [89]. Other mutants have also been identified at the entrance to the lateral gate in yeast Hrd1 that specifically affected ERAD-M substrate degradation [58]. TMD1 is not known to interact with any Hrd1 partners in yeast and if the same is true in mammals, this suggests a role of this membrane pocket in Hrd1 function. In the closed state, the membrane thinning could facilitate dislocation of substrates from the membrane by destabilisation of TMDs, as well as reducing the energetic barrier for movement of substrates across the membrane.

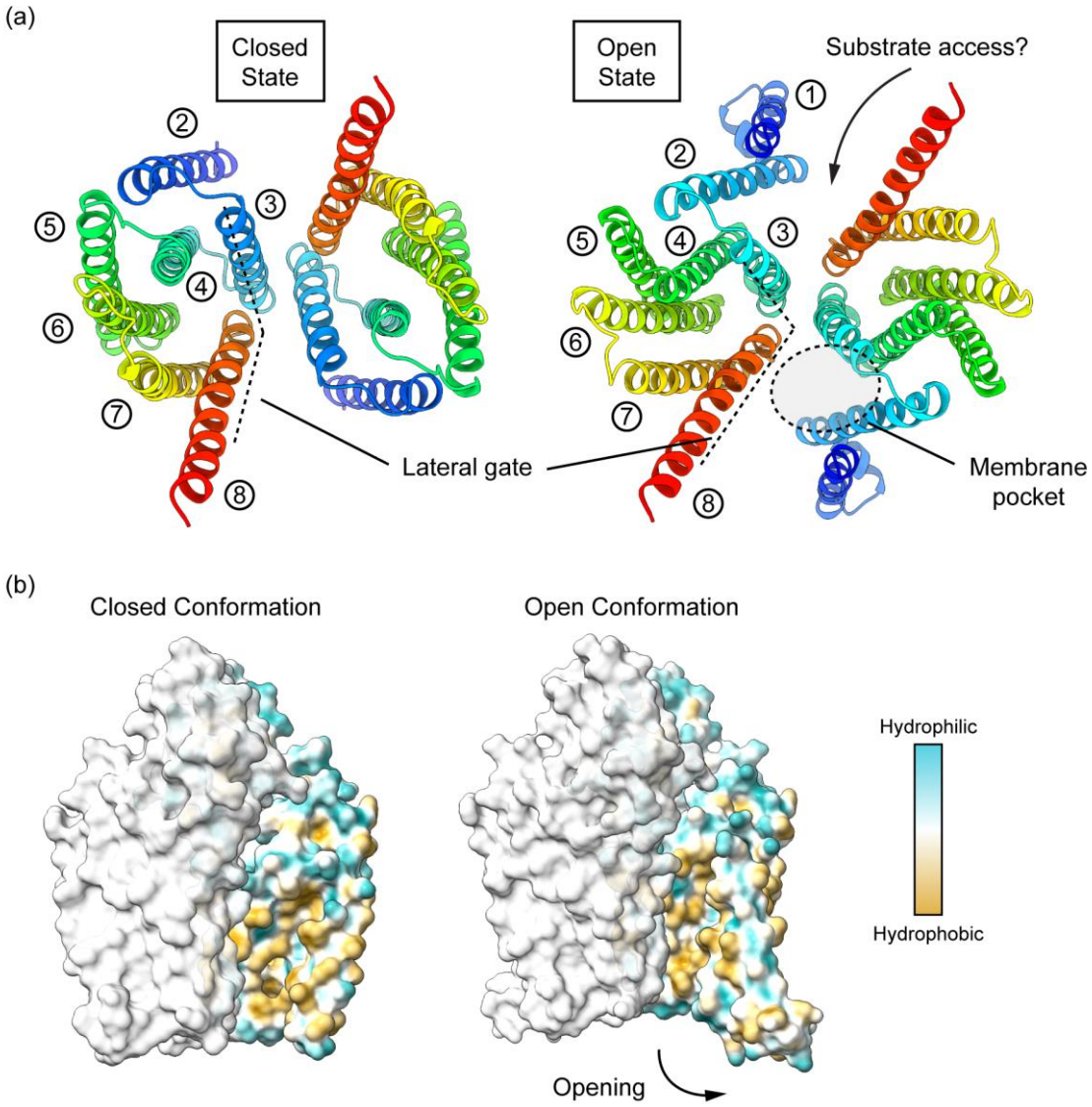


Figure 4.10. Differences in access to the lateral gate in the Hrd1 conformations. (a) Top-down view of Hrd1 in the closed and open conformations. The TMD helices are coloured blue to red from the N- to the C-terminus. (b) Hydrophobicity surface representation of the open and closed conformations of Hrd1. Hydrophobicity is coloured from orange to turquoise with decreasing hydrophobicity.

In all the Hrd1 maps, non-modelled density was also present and this was strongest in the full-length map (figure 4.11). This density is likely detergent or co-purified lipid molecules. At the dimer interface in all maps, there was strong density for a symmetric

detergent/lipid molecule (figure 4.11a). This sits at the top of the dimer interface, has a single tail and is ligated by a highly conserved arginine residue (R253) through charged interactions with the head group (figure 4.11a). The ER membrane is predominantly made of phosphatidylcholine and phosphatidylethanolamine lipids, both of which contain two fatty acid tails, and so this density is unlikely to be an endogenous lipid. Hrd1 was always purified in the presence of DDM and CHS, of which DDM has a single tail and polar head group, making it a likely candidate. There is also another large unmodelled density found between TM helices 2 and 3. No obvious side chain contacts could be identified and instead the binding surface is mostly hydrophobic. Likely, many residues contribute to this interaction by providing a hydrophobic environment for detergent/lipid binding. The significance of this molecule is unknown, however, a similar event was observed between different helices in focused cryo-EM refinements of the yeast Hrd1-Hrd3 monomer [55]. These revealed that a molecule of digitonin was stably bound at the lateral gate between TMDs 3 and 8. This raises the possibility that perhaps a sterol is bound between TMDs 2 and 3 in mammalian Hrd1. Recent mutational studies of yeast Hrd1 TM helices 1, 2 and 3 revealed that certain transmembrane residues are important for degradation of lumenal substrates [89]. This unknown cryo-EM density may therefore represent co-purified Hrd1 substrates peptides but further work is needed to confirm this. Further focused-refinements did not improve resolution in this region. Throughout the rest of the maps, there are also smaller densities along the top and bottom membrane boundaries (not shown), many of which are likely to be other detergent/lipid molecules.

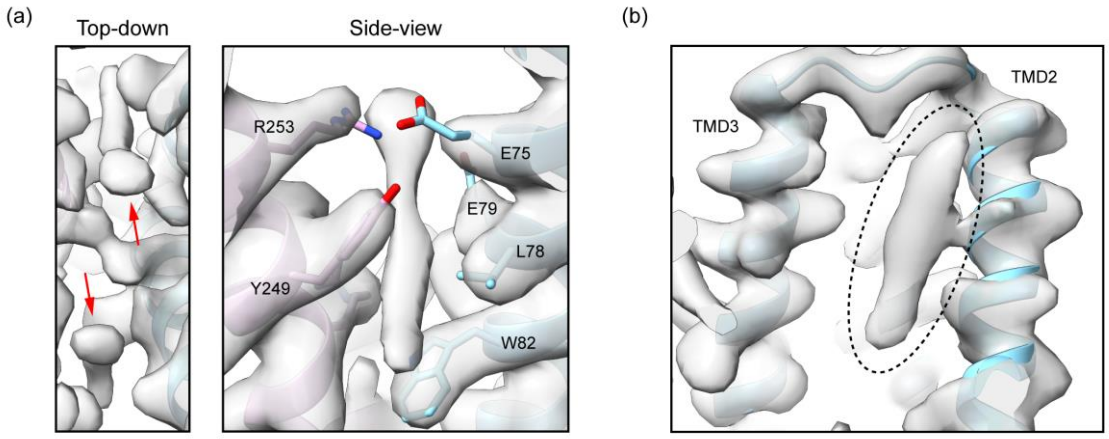


Figure 4.11. Additional non-protein density from the full-length Hrd1 map. (a) Views of non-protein density found at the dimer interface. Residues in close proximity are annotated. (b) Non-protein density found between TM helix 2 and 3.

4.2.4. Analysing in vitro assembly of purified Hrd1

Since different yeast Hrd1 monomer and dimer structures exist, some of which are less active *in vitro*, size exclusion chromatography with multi-angle light scattering (SEC-MALS) was performed to see if the mammalian Hrd1 dimer was unique to DDM detergent [55,60]. Hrd1 was purified in DDM-CHS as well as a milder LMNG-CHS detergent mix. These samples were then subject to SEC-MALS to estimate the molecular mass (figure 4.12a and 4.12b). In both DDM-CHS and LMNG-CHS, purified Hrd1 was entirely dimeric with estimated masses of 132 and 136 kDa, respectively. This shows that dimerisation is not exclusive to DDM conditions but it could still be induced by overexpression or extraction in detergents generally. This however corresponds with the mammalian Hrd1 literature where endogenous Hrd1 tagged on at least one allele, can pulldown non-tagged Hrd1, suggesting the presence of complexes containing more than one copy of Hrd1 [105].

To further probe the biophysical properties of the Hrd1 dimer, a melting temperature (T_m) assay was developed (figure 4.12c-h). Upon heating purified protein at a range of

temperatures and clearing precipitated material by centrifugation, protein fluorescence of a tagged fluorophore fusion could be monitored by FSEC to determine the proportion of protein that is no longer correctly folded. This assay assumes that the stability of the fluorescent tag is greater than that of the attached membrane protein so that changes in fluorescent signal are due to protein unfolding or aggregation of the target protein. To test the feasibility of this assay, purified mVenus was heated across a range of temperatures and fluorescence measured by FSEC (figure 4.12c). As temperature increased, protein fluorescence of the non-void peak decreased. The maximum fluorescence of this peak was then plotted against temperature to determine an approximate T_m for mVenus (figure 4.12d). The measured melting temperature was ~80°C and this corresponded to previously reported T_m values for mVenus and for the related fluorophore eYFP [175,176]. This approach was then applied to Hrd1 (figure 4.12e) and yielded a T_m of ~68°C (figure 4.12f). mVenus melting was performed in identical conditions to Hrd1-mVenus and since the Hrd1 T_m was lower than the that of mVenus, it was assumed that the majority of denaturing in this sample was driven by Hrd1. This T_m was further corroborated by circular dichroism and nano-differential scanning fluorimetry melting assays, which yielded T_ms of ~74°C and ~70°C, respectively (figure S7-8). However, in the FSEC T_m assay, two inflection points could be observed and closer inspection of the FSEC traces revealed the presence of a second, later peak upon heating of the sample at lower temperatures (figure 4.12g). This later peak is likely monomeric Hrd1 and plotting values for this transition identified at second T_m of ~31°C. If this peak corresponds to monomeric Hrd1, it implies that the dimer interface has a T_m of ~31°C. The significance of this is that in mammalian cells, an equilibrium of dimeric and monomeric Hrd1 could be established and might serve as a way to regulate dimer formation. However, this experiment is carried out in a very artificial context. This T_m is likely dependent on the detergent condition and may also be affected by the presence of other interacting partners.

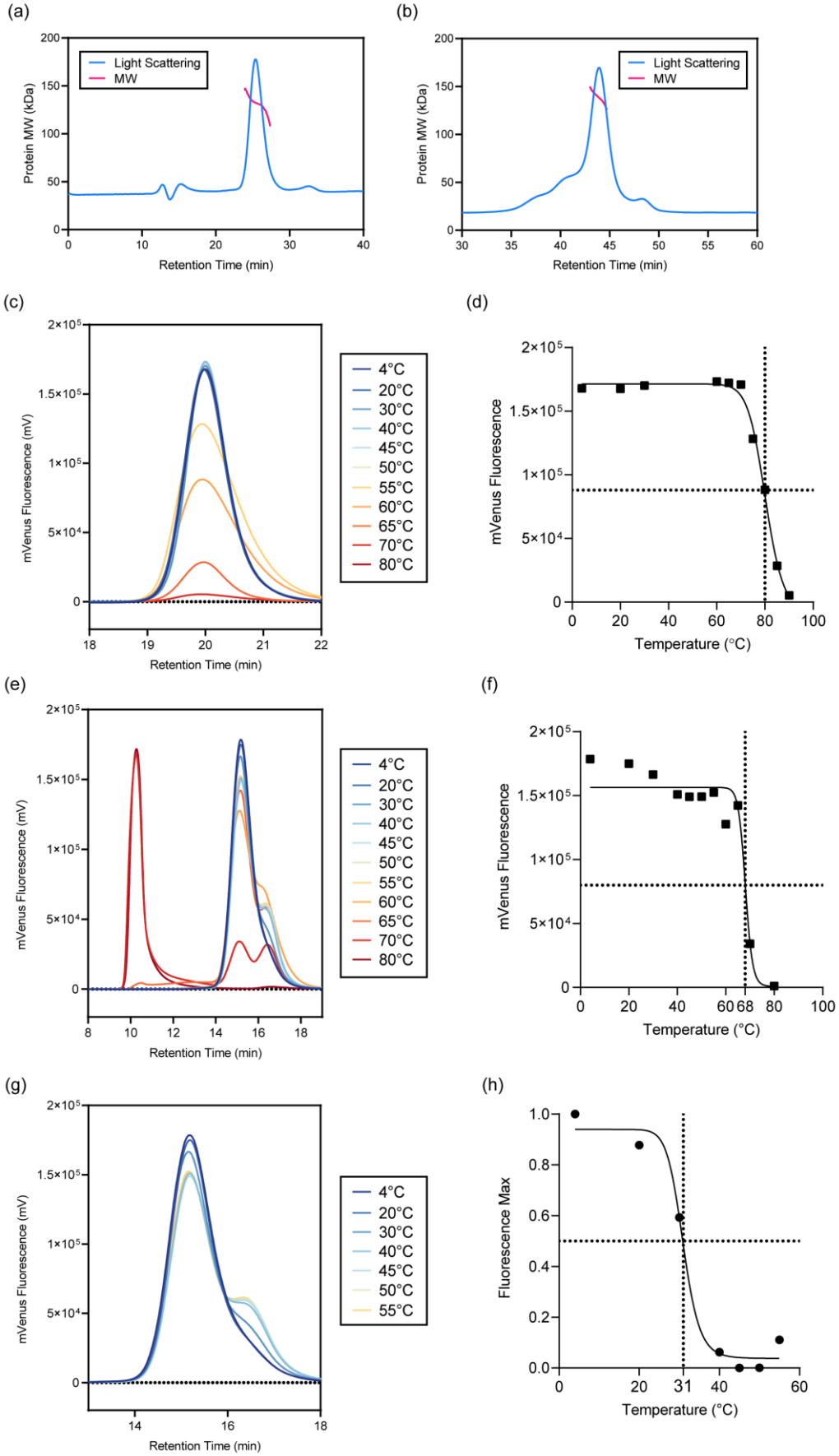

(a)
Light Scattering
MW
Protein MW (kDa)
Retention Time (min)
(b)
Light Scattering
MW
Protein MW (kDa)
Retention Time (min)
(c)
mVenus Fluorescence (mV)
Retention Time (min)
4°C
20°C
30°C
40°C
45°C
50°C
55°C
60°C
65°C
70°C
80°C
(d)
mVenus Fluorescence
Temperature (°C)
(e)
mVenus Fluorescence (mV)
Retention Time (min)
4°C
20°C
30°C
40°C
45°C
50°C
55°C
60°C
65°C
70°C
80°C
(f)
mVenus Fluorescence
Temperature (°C)
(g)
mVenus Fluorescence (mV)
Retention Time (min)
4°C
20°C
30°C
40°C
45°C
50°C
55°C
(h)
Fluorescence Max
Temperature (°C)

Figure 4.12. Analysis of purified Hrd1 by SEC-MALS and a melting temperature assay. (a) SEC-MALS of Hrd1 in DDM-CHS. The average peak mass was 132 kDa. (b) SEC-MALS of Hrd1 in LMNG-CHS. The average peak mass was 136 kDa. (c) Melting temperature (T_m) assay for an mVenus control by FSEC. Purified mVenus was incubated at a range of temperatures and fluorescence measured by FSEC. (d) T_m plot for mVenus calculated from the fluorescence maximum observed by FSEC. mVenus had a T_m of 80°C. (e) T_m assay for purified Hrd1 in DDM-CHS. (f) T_m plot for Hrd1 calculated from the fluorescence maximum in the peak between 14 and 16 minutes retention time. (g) Close-up view of the FSEC traces between 14 and 18 minutes. (h) T_m plot for the suspected dimer to monomer transition of Hrd1.

4.2.5. Comparison with yeast Hrd1 structures

Mammalian Hrd1 shares sequence similarity (29% identity) and structural similarity with the published yeast structures, as highlighted by alignments between the monomers of each structure (figure 4.14). Multiple structural studies on yeast Hrd1 have been performed and they have identified different Hrd1 states [55,60]. Initially, yeast Hrd1-Hrd3 was reconstituted in amphipols, producing a dimeric Hrd1 structure [60]. Follow up studies co-purified different Hrd1 complexes and identified a monomeric Hrd1-Hrd3, as well as a heterodimeric Hrd1-Der1 [55]. The Hrd1 monomer from these structures was aligned with the open and closed states of mammalian Hrd1, to identify differences in the backbone structure (figure 4.13). The deviation of RMSD Cα varies from 1.48 to 4.05 Å, signifying slight differences in the arrangement of the TMDs between the models. The models for the Hrd1 open and closed states closely resemble each other with a backbone RMSD of only 1.48 Å. The largest difference exists between the mammalian Hrd1 closed conformation and the yeast Hrd1-Hrd3 monomer and Hrd1-Der1 heterodimer models, with RMSDs of 4.05 and 3.60 Å, respectively. However, the structures are all broadly

similar and share the same arrangement of secondary structure. The largest differences are attributed by slight changes in the pitch of TM helices or movements of TMD3. Altogether, this suggests that mammalian Hrd1 undergoes a similar mechanism for ERAD as yeast Hrd1, since the different states align well with each other.

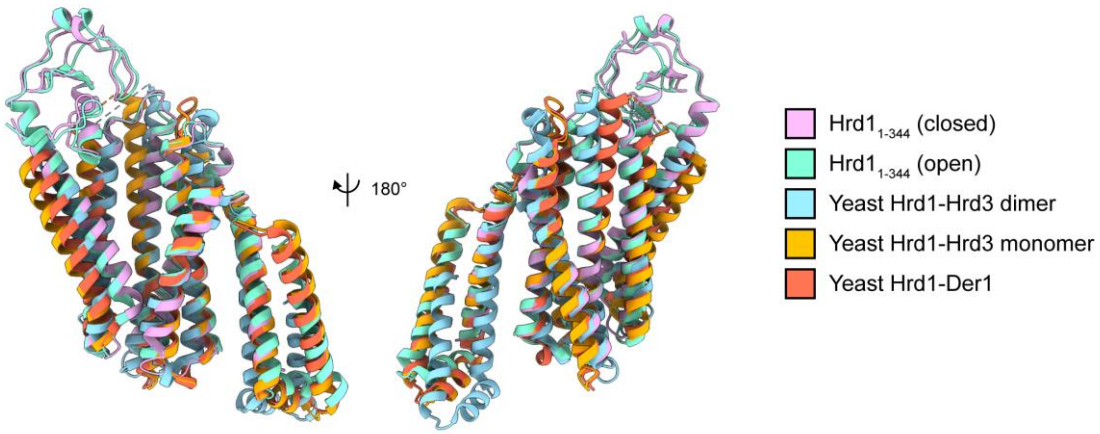


Figure 4.13. Comparison of mammalian Hrd1 with yeast Hrd1 structures. All models were aligned to the Hrd1 closed model and RMSD values reported with reference to this. Hrd1 open RMSD 1.48 Å. Hrd1 monomer from the yeast Hrd1-Hrd3 dimer in amphipol RMSD 3.39 Å [60]. Hrd1 from the yeast Hrd1-Hrd3 monomer RMSD 4.05 Å [55]. Hrd1 from the yeast Hrd1-Der1 heterodimer RMSD 3.60 Å.

The mammalian Hrd1 dimer was also aligned with the Hrd1-Der1 heterodimer, to try to understand if the open or closed state more closely resembled the yeast heterodimer (figure 4.14). Alignments were produced by aligning on a Hrd1 monomer from each structure. Interestingly, the angle between the dimeric protomers of the open state more closely resembled the angle between the Hrd1-Der1 heterodimer. Using yeast structures of Hrd1-Hrd3, the binding of Sel1L (mammalian Hrd3 homologue) was approximated in ChimeraX (figure 4.15). An AlphaFold predicted model of Sel1L was first aligned with Hrd3 to determine how similar the structures might be and if so, whether alignment in a

larger model would be a good approximation (figure 4.15a). Since a Hrd3 truncation was co-purified with Hrd1 in the yeast structures, a similar truncation was modelled with AlphaFold. The main element removed in this model is the C-terminal TMD which is responsible for ER membrane localisation, likely for increasing a chance of interaction with Hrd1 [177]. This was aligned and although there are some differences in the position of certain helices, the broad fold is quite similar, despite a Cα RMSD of 10.38 Å (figure 4.15a). The binding mode of Hrd1 to Hrd3 has been long established and is reliant on binding to the linking helix between TM helices 1 and 2 of Hrd1, a well conserved region [60]. 2 copies of the Hrd1-Hrd3 monomer were aligned with the open Hrd1 state and then the AlphaFold predicted Sel1L truncation aligned with Hrd3. Together, this produced an approximate model for the Hrd1-Sel1L interaction (figure 4.15b). Surprisingly, this model contrasts with the current belief in the literature that Hrd3/Sel1L is unable to exist as a dimer in a Hrd1 dimeric complex [55]. Previous Hrd1 structural studies showed that a Hrd1-Hrd3 dimer would result in steric clashes of the soluble Sel1L SLR motifs [55]. However, in our open model, two copies of Sel1L are compatible with the structure without steric clashes. The other issue with the two copy Hrd3/Sel1L model was that an N-terminal amphipathic helix was completely in the plane of the membrane. In the open Hrd1 model, this is not the case and only the hydrophobic portion is submerged, as would be expected. It is not possible to model Sel1L binding to the closed Hrd1 state as neither TMD1 or the linking helix between TM helices 1 and 2 were modelled, and so the orientation of Sel1L cannot be determined. If Sel1L was modelled as according to the Hrd1-Hrd3 monomer in the closed state as is, then this does indeed result in a steric clash between the Sel1L SLR domains. Furthermore, the SLR domains of Sel1L have been shown to form a homodimer by crystallography, suggesting a role for Sel1L dimers in ERAD function [178]. Given that Sel1L can only theoretically bind in the open state, this conformation could represent a transitory state for dissolution of the dimer, to allow binding with other partners for ERAD function.

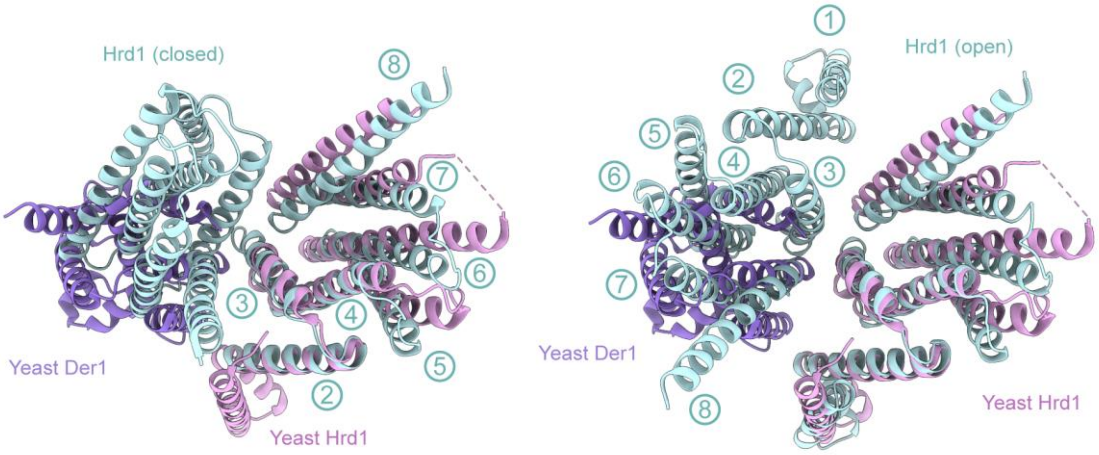


Figure 4.14. Comparison of the mammalian Hrd1 dimers with the yeast Hrd1-Der1 heterodimer. The yeast Hrd1-Der1 heterodimer was aligned with the closed and open Hrd1 states, via alignment on a Hrd1 monomer, using ChimeraX. Mammalian Hrd1 residues after 275 are omitted for clarity. Mammalian Hrd1 is coloured light blue, yeast Der1 purple and yeast Hrd1 pink. TMD numbers are labelled for the Hrd1 closed and open states.

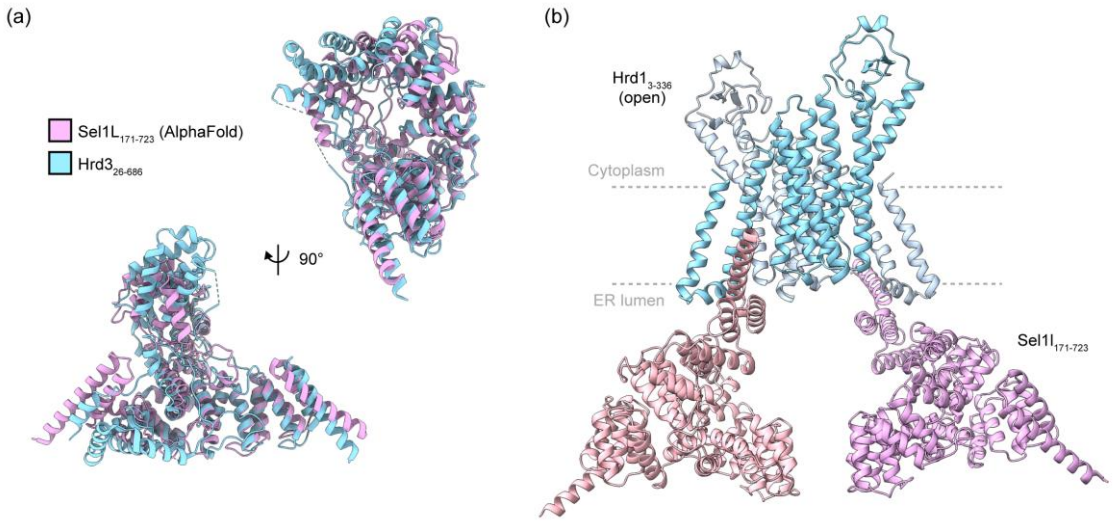


Figure 4.15. Modelling binding of Sel1L with Hrd1 using yeast Hrd1-Hrd3 as a reference. (a) AlphaFold predicted Sel1L$_{171-723}$ was aligned with the Hrd3 model from the yeast

Hrd1-Hrd3 monomer structure, using ChimeraX [55]. This alignment had an RMSD of 10.38 Å. (b) Model for interaction of Sel1L with mammalian Hrd1. The yeast Hrd1-Hrd3 monomer was first aligned to the open Hrd1 conformation. $Sel1L_{171-723}$ was then aligned to Hrd3 to produce a Hrd1-Sel1L model.

5. Discussion

In this chapter, cryo-EM of mammalian Hrd1 was performed and structures of Hrd1 in an open and closed state resolved. Cryo-EM of full-length Hrd1 only resolved a single state to high-resolution whilst imaging of a Hrd1 truncation led to determination of high-resolution maps for two distinct states. Both these conformations appear to be in the full-length Hrd1 dataset but the open state could not be resolved to high resolution. This could be due to differences in sample preparation or perhaps due to truncation of the disordered C-terminus, there is a better Hrd1 view distribution in vitreous ice which enabled better computational separation of the two conformations. Alternatively, this could be due to a combination of parameters during data collection such as differences in microscope optics and ice thickness of the cryo-EM grids. Recent mutational studies of yeast Hrd1 also support a role of the disordered C-terminus in substrate degradation and so this change in abundance could be sample dependent [89]. Despite this, the two Hrd1 conformer maps were of sufficient resolution to build models from. In both states, Hrd1 is a homodimer and the RING domains are positioned above the TMDs. In the closed Hrd1 state, the cytosolic RING domains are closer together and the micelle is drastically thinned at the dimer interface. Conversely, in the open state the monomers rotate at the dimer interface to position the RING domains further apart. This also has the added effect of making the detergent micelle flatter. The significance of these two states is difficult to dissect, but one important difference is that in the open state, the N-terminal TMD1 is stably associated with the other TMDs. Interestingly, there is a

membrane-accessible binding pocket created by TMDs 1, 2, 3 and 8, that is deeper in the open state. This hydrophobic pocket does not exist in other yeast Hrd1 structures and may be important for Hrd1 function [55,60]. The shape of this pocket changes during transition between the open and closed states. However, in both these states, the proposed lateral gate of Hrd1 (made up by TMDs 3 and 8) is obstructed and so these states may represent inactive forms of Hrd1.

Dimeric yeast Hrd1-Hrd3 was suggested to be a purification artifact as a domain swap of Hrd1 TMD1 led to obstruction of the opposing lateral gate [55]. The position of Hrd3 within the complex was also non-physiological as an amphipathic helix was submerged deep within the expected membrane [60]. In both DDM and LMNG detergents, mammalian Hrd1 was dimeric. TMD1 was not near the lateral gate and instead resides towards the outside of the micelle or is not visible, in the open and the closed states, respectively. This dimerisation is also supported by the literature, where endogenous Hrd1 forms high-order oligomers with itself [105]. The dimer interface was further probed by assessing the T_m for dimer dissolution. The proposed T_m of the dimer interface is ~31°C but more work is needed to understand if this is detergent specific and how other interacting partners affect this value.

Alignment of the open and closed states of Hrd1 with yeast structures revealed a high degree of structural conservation between mammalian and yeast Hrd1. The monomeric fold of Hrd1 is strongly conserved and so Hrd1 function is expected to be similar to yeast. The open Hrd1 state more closely resembled the Hrd1-Der1 heterodimer structure and surprisingly, in this state, two molecules of Sel1L could theoretically bind at the same time. This is in contrast to the yeast structures where a Hrd1 dimer is not compatible with two copies of Hrd3 since a steric clash would be induced. It is unknown whether yeast Hrd1 also exists in two conformations but together, this suggests that mammalian Hrd1 may act differently during ERAD. Together, the data imply that Hrd1 dimers exist both *in vivo* and recombinantly, and exist in at least two conformations.

These two conformations may represent different active states of Hrd1, binding partner component states or are a unique feature of recombinantly-purified Hrd1. Structural studies of Hrd1 with its endogenous binding partners could help delineate the role of these two states and this is explored in the next chapter.

5. Reconstituting higher-order Hrd1 E3 ligase complexes for cryo-EM

5.1. Endogenous partners of mammalian Hrd1

The core mammalian Hrd1 complex is composed of HERPUD1, FAM8A1, Sel1L and Derlins-1, -2 or -3 [20]. Aside from FAM8A1, all other subunits are essential for ERAD [61]. The interacting partners are somewhat similar to their yeast counterparts (table 5.1). Similar to yeast Hrd1, mammalian Hrd1 has been shown to assemble with a range of rhomboid pseudoproteases for ERAD function [62,100,179]. Namely, Hrd1 complexes contain Derlin-1, Derlin-2 and Derlin-3. Derlin-2 is most similar to yeast Der1 (~21% sequence identity) and is thought to be the sole driver of ERAD-L substrate degradation in mammals [100]. Derlin-1 is similar to yeast Dfm1 (~28% sequence identity), which has been implicated in Hrd1-mediated ERAD-M substrate degradation [98]. Interestingly, there is a greater expansion of Derlins in mammals which could be a result of a greater diversity of substrates processed by mammalian ERAD. Some Derlins are also present in other E3 ligase complexes and so may also represent a common mode of function within ERAD [80]. In the yeast Hrd1-Der1 heterodimer, Usa1p is required to stabilise the interaction between Hrd1 and Der1 [180]. There is no Usa1p sequence-related homologue in higher eukaryotes but HERPUD1 (Herp) is thought to be its functional equivalent. In Usa1p-deficient cells, expression of Herp partially rescued an ERAD defect [91]. The domain organisation of Herp is also similar to Usa1p, containing ubiquitin-like domains at its N-terminus, followed by C-terminal TMDs (figure 5.1) [181]. The mammalian Hrd1 complex is also known to contain FAM8A1, for which there is no functional equivalent in yeast [61]. FAM8A1 is required for stable association of Herp and therefore indirectly stabilises association of the Derlins with Hrd1 [61]. The cytoplasmic region of Hrd1 is mostly disordered but after the RING domain there is a conserved

domain termed HAF-H [61]. Binding of FAM8A1 to HAF-H was shown to promote binding of Herp to HAF-H either through increased Hrd1 affinity or exposure of an additional binding site [61]. The TMDs of Hrd1 and FAM8A1 were also shown to interact but this was not required for Herp stabilisation. This interaction mode is a departure from the yeast model where no such interaction occurs between Usa1p and Hrd1, and its significance is unknown. FAM8A1 was also proposed to drive oligomerisation of higher-order Hrd1 complexes and so special attention was made during Hrd1 complex reconstitution in this chapter [61]. One other important difference between the yeast and mammalian Hrd1 complexes is the presence of ERLEC1 (XTP3-B) in mammals. Yos9/OS-9 drive recruitment of misfolded, glycosylated substrates to the Hrd1 complex through interactions with the lumenal domains of Hrd3/Sel1L [55]. Metazoans have two orthologues of Yos9, OS-9 and XTP3-B, and only OS-9 is indispensable for glycoprotein ERAD [182]. However, more recent studies have shown that XTP3-B inhibits degradation of misfolded glycans whilst OS-9 antagonises this inhibition [183]. These two proteins likely cooperate together for finer control of ERAD of lumenal substrates and this cooperation may be important for downstream expression and reconstitution of these proteins in a Hrd1 complex.

Name	Location	Size (kDa)	Yeast Equivalent*	Sequence identity (%)
SYVN1 (Hrd1)	ER membrane	68	Hrd1	26
FAM8A1	ER membrane	44	N/A	N/A
HERPUD1 (Herp)	ER membrane	44	Likely Usa1p	Non-significant
DERL1 (Derlin-1)	ER membrane	29	Dfm1	28
DERL2 (Derlin-2)	ER membrane	28	Der1	21
DERL3 (Derlin-3)	ER membrane	19	N/A	N/A
Sel1L	ER membrane	89	Hrd3	30
OS-9	ER lumen	76	Yos-9	28
ERLEC1 (XTP3-B)	ER lumen	55	N/A	N/A
Ube2j1	Cytoplasm	35	Ubc6p	42
Ube2g2	Cytoplasm	19	Ubc7p	62
AUP1	ER membrane	46	Cue1p	Non-significant

Table 5.1. Known interactors of mammalian Hrd1 and their yeast counterparts. *Not all yeast equivalents are sequence homologues. This annotation is based on probable functional equivalency. Sequence identity calculated with BLASTP local alignments [135].

The discovery of these interactors was made through studies on ectopic expression of ERAD components by transient or stable gene expression from engineered promoters [22,184–186]. Since then, studies on endogenous Hrd1 complexes have been analysed in detail, both biochemically and by mass spectrometry [62]. These studies reveal that a range of Hrd1 complexes can be isolated and separated, each with differing levels of binding partners and MWs. Two distinct complexes containing different ratios of interacting partners were identified but it is hard to decouple these from assembly intermediates or detergent artifacts [62]. In either case, these results suggest that endogenous Hrd1 is capable of assembling with a wide range of partners and likely does have intermediate assembly states towards a full complex containing Hrd1, FAM8A1, Herp, Derlin-2 and Sel1L, a minimal complex likely required for ERAD-L [55].

The membrane topology of the core interactors of mammalian Hrd1 has been annotated in figure 5.1. These annotations were produced through TMD predictions with DeepTMHMM and UniProt annotations [107,110]. The membrane region of Hrd1 was of significant interest since there are still unanswered questions as to how substrate retrotranslocation occurs. Yeast structures provide a static view of a state likely competent for substrate engagement but offer no direct insight into substrate binding and retrotranslocation [55,60]. However, recent crosslinking studies on the yeast Hrd1-Der1 complex trace a path of substrate contact through the complex [56]. Mammalian E3 ligase complexes are especially understudied and so we sought to determine structures of mammalian Hrd1 with its membrane-associated ERAD factors. Since structures of yeast Hrd1 complexes were solved, we aimed to isolate equivalent complexes as a

starting point. A likely minimal equivalent to the Hrd1-Der1 heterodimer would contain Hrd1, FAM8A1, Herp, Derlin-2 and Sel1L [55]. Each of these proteins contains numerous TMDs, with Hrd1 and Derlin-2 contributing the largest number (figure 5.1). The TMD of Sel1L is known to be unimportant for Hrd1 binding but it may still interact with the TMDs of other partners [177]. Importantly, many of these proteins are not predicted to contain large, folded domains outside of the membrane, which may make cryo-EM studies more difficult [136]. A lack of cryo-EM signal outside of a noisy membrane mimetic could inhibit alignment of particles during 2D classification and so assembly of a large Hrd1 complex would be advantageous. Therefore, we sought complexes that behaved well biochemically.

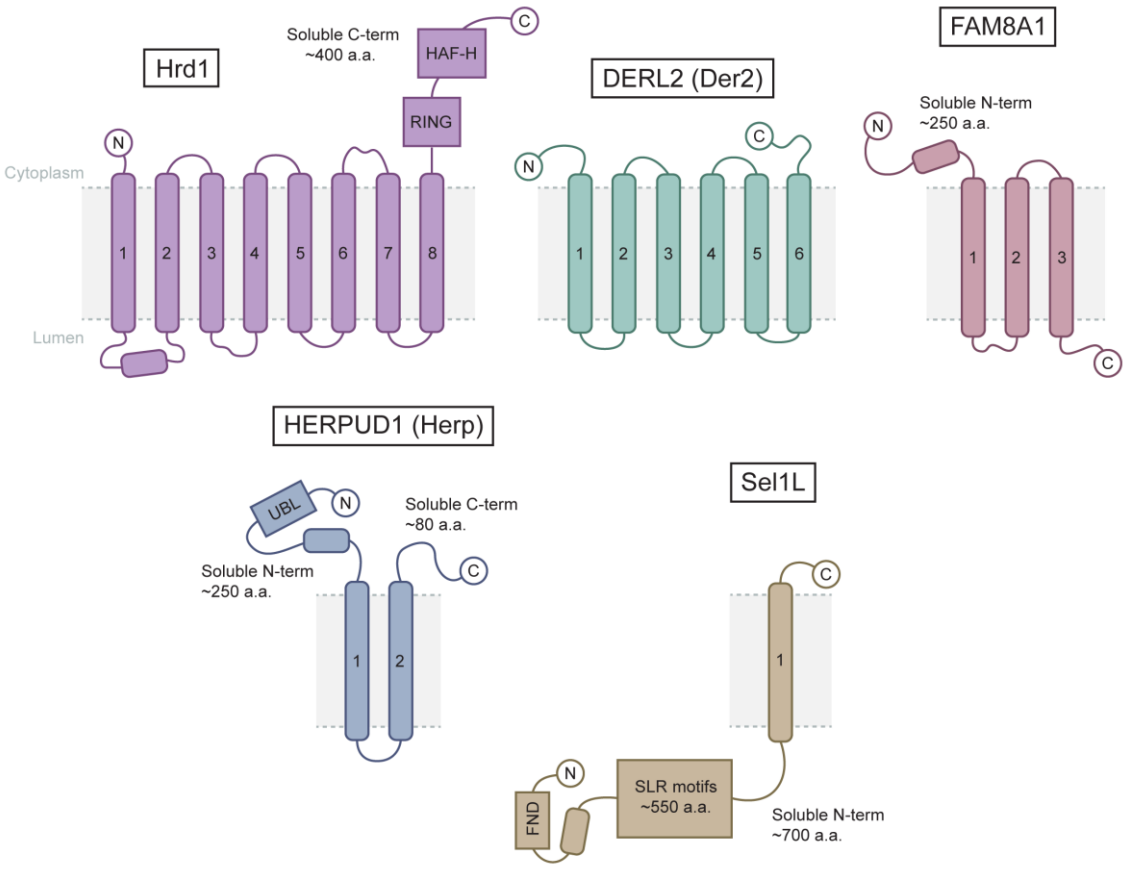


Figure 5.1. Topology of the core membrane interactors in the mammalian Hrd1 complex. TMDs are annotated based on their position in the sequence. Key domains are also

annotated. The RING domain of Hrd1 is responsible for substrate ubiquitination. The HAF-H domain is involved in recruitment of FAM8A1 to Hrd1 [61]. The ubiquitin-like domain (UBL) of Herp has unknown function in the Hrd1 complex.

5.2. Purification of higher-order Hrd1 complexes

5.2.1. Purification of Hrd1-FAM8A1

We sought to first reconstitute Hrd1 with FAM8A1 by dual affinity purification to isolate Hrd1-FAM8A1 complexes. Full-length Hrd1 and FAM8A1 constructs were used for overexpression since this would be required to guarantee an interaction between the two proteins [61]. Cells for overexpression of Hrd1-mVenus-TwinStrep were subject to a second round of lentiviral transduction with FAM8A1-6xHis vectors. Successful transductants were enriched with antibiotic selection and stable cell lines scaled up to the litre range prior to protein expression. We initially purified this complex in DDM-CHS since this condition yielded high resolution structural information for Hrd1 alone by cryo-EM. StrepTactin purification resulted in high Hrd1 enrichment, but additional bands were now visible by SDS-PAGE (figure 5.2a). This is in contrast to previous purifications of Hrd1 alone. The band running between the 37 and 50 kDa markers corresponded to FAM8A1 and upon a second affinity step, this band was further enriched. Interestingly, there was also a band close to the 25 kDa marker that remained enriched during the Ni-NTA purification, indicating it may be a part of the Hrd1-FAM8A1 complex. At the time, mass spectrometry identification of this band was not performed but based on further purifications, this band could be small amounts of an endogenous Derlin. There is also another band present above FAM8A1. Further work indicated that this is likely small amounts of Herp, which is explored later in this chapter. There was a large proportion of Hrd1 in the unbound fraction from the Ni-NTA purification, suggesting that Hrd1 expressed better than FAM8A1 or that complexes of Hrd1-FAM8A1 were unstable in

DDM-CHS. Tags were cleaved off with TEV protease and the material was subject to SEC to reduce any heterogeneity, prior to any structural studies (figure 5.2b). The cleaved mVenus tag was not usually removed by reverse-affinity chromatography since in our experience, this always separated by SEC from all membrane proteins. SEC revealed that a large proportion of material was in the void, indicative of protein aggregation. Hrd1 alone in DDM-CHS only exhibited a small void volume and so this complex had significantly different biochemical behaviour. There were 2 other main peaks between 10 and 15 ml that corresponded to differences observed by SDS-PAGE (figure 5.2c). The main difference between these peaks is the abundance of the unknown band at the 25 kDa marker. No protein of this MW was overexpressed and later studies of other recombinant complexes led us to believe this is likely an endogenous Derlin. Co-elution with this unknown band results in a larger Hrd1 complex by SEC. There also appears to be some FAM8A1 dissociation in this detergent condition, as is evidenced by free FAM8A1 eluting at 15 ml. The stoichiometry of Hrd1 to FAM8A1 does not appear entirely 1:1 but this is hard to assess by Coomassie staining since some proteins stain stronger than others. To test this, known quantities of Hrd1 were run by SDS-PAGE with known quantities of other soluble proteins of different MWs (not shown). This revealed that Hrd1 stains stronger by Coomassie than other proteins at the same concentration. The SEC profile suggests a level of instability in this detergent condition, but the sample was still homogenous enough to image on cryo-EM grids.

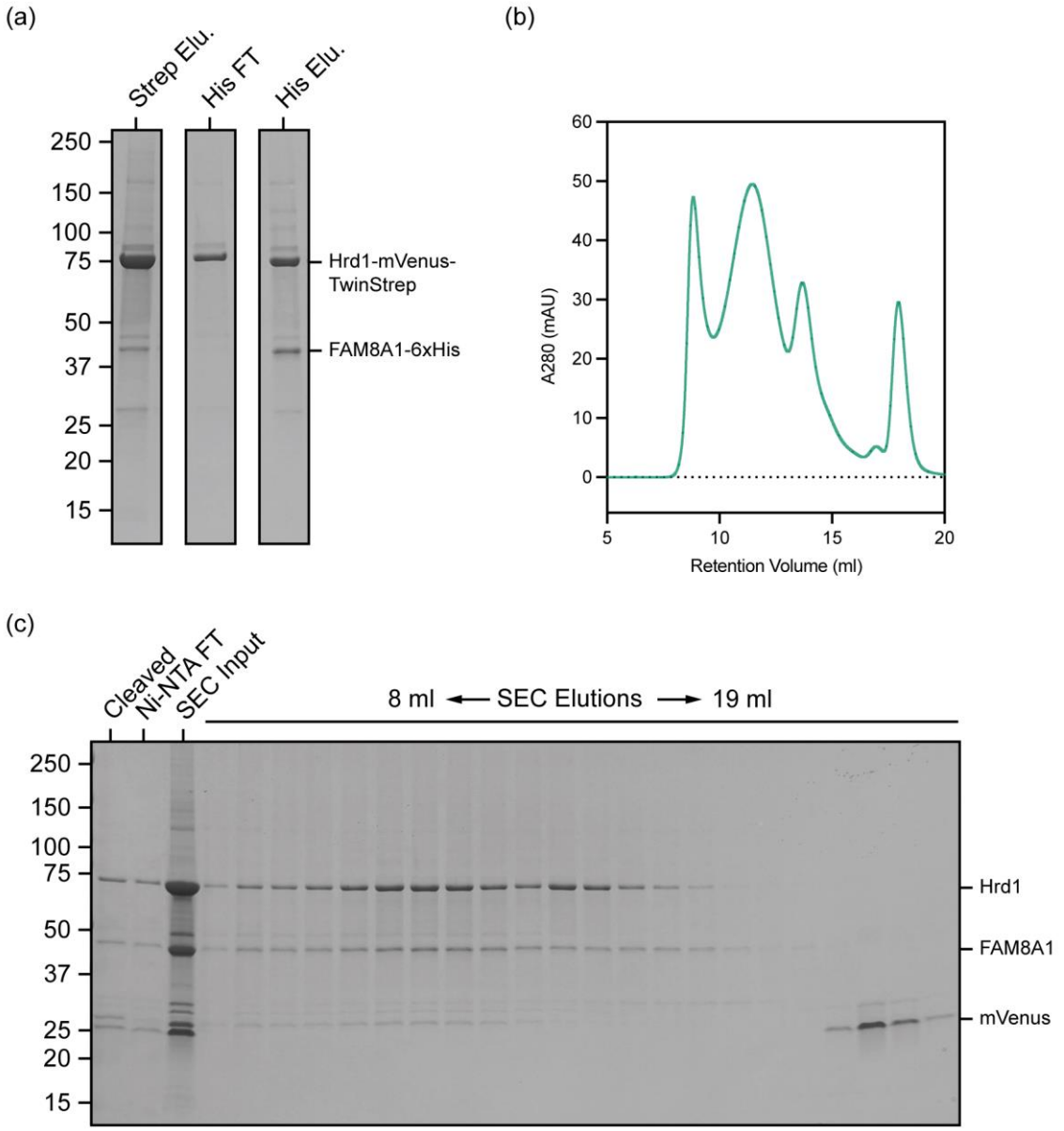


Figure 5.2. Purification of a Hrd1 complex containing FAM8A1 in DDM-CHS. (a) SDS-PAGE gels from the dual affinity purifications used to isolate Hrd1-FAM8A1. Corresponding elution fractions from the StrepTactin and Ni-NTA (His) purifications are shown. A fraction of the unbound material from the Ni-NTA purification is also shown (His FT). (b) SEC profile for Hrd1-FAM8A1. (c) SDS-PAGE gel of the TEV cleavage reaction prior to SEC, as well as the SEC elutions.

Cryo-EM micrographs for Hrd1-FAM8A1 in DDM-CHS looked promising and so a data collection was performed and data processed with SIMPLE, cryoSPARC and RELION (figure 5.3) [162,166,187]. Data were collected by Justin Deme and Susan Lea on a Titan Krios G4 operating at 300 kV, in the same way as the $Hrd1_{1-344}$ dataset. Movies were collected in counting mode in EER format at 165,000x magnification, with a pixel size of 0.693 Å. 14,796 movies were collected and processed in SIMPLE similarly to the Hrd1 datasets except that only motion correction, CTF estimation and particle picking performed. A lowpass-filtered reference map of another unrelated small membrane protein from the lab was used for automated picking. Particles (4,709,138) were extracted and multiple rounds of 2D classification performed in cryoSPARC to identify "good" particles [187]. 2D classification in cryoSPARC was performed with the same parameters as processing of the $Hrd1_{1-344}$ dataset. These were used to generate three *ab initio* classes of which ~40% of particles classified into a map with clearly defined transmembrane helices. Particles from this class (315,148) were subject to a non-uniform refinement (NU-refinement) in C1 symmetry with the *ab initio* volume as a starting reference, but the resolution of this was further improved with a C2 NU-refinement. At this point, there was a strong resemblance to the Hrd1 closed map. A second round of *ab initio* generation was performed on these particles to further reduce heterogeneity. Only one class yielded a map with defined transmembrane helices and particles from this class (155,778) were refined against the *ab initio* volume in a C2 NU-refinement. This class further improved the overall resolution of the map. Per-particle motion was corrected for with Bayesian polishing in RELION and particles re-imported into cryoSPARC for 2D classification, to remove artefacts generated during the polishing process [166]. Using the polished particles, a C2 NU-refinement was performed and this also improved the quality of the map further. A final *ab initio* generation was performed with 2 classes, to separate out any remaining heterogeneity. Most of the particles

separated out into class 1 (~70%) and a final C2 NU-refinement of these particles was performed which yielded a map with nominal resolution of 3.82 Å.

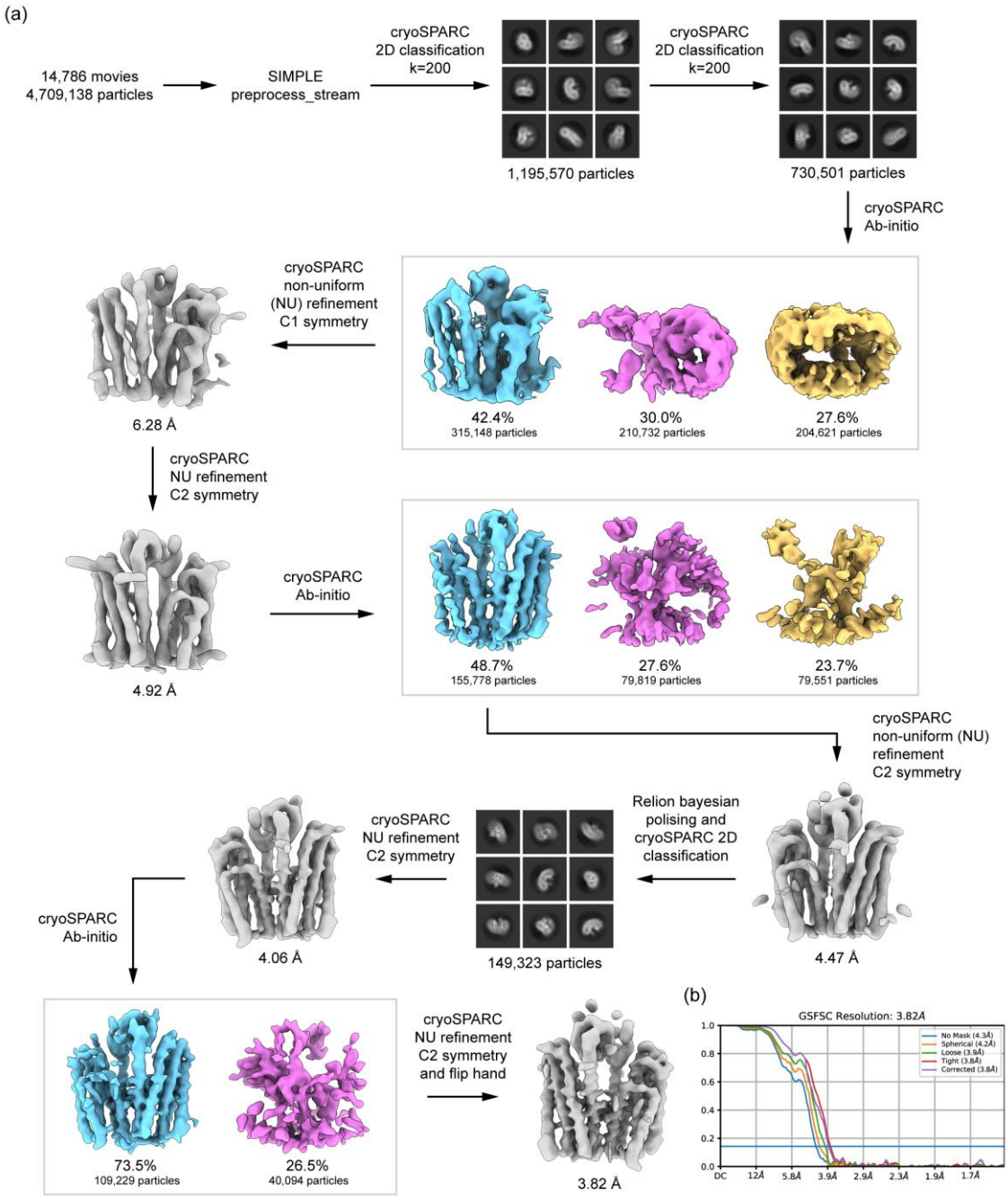


Figure 5.3. Cryo-EM processing for Hrd1-FAM8A1 in DDM-CHS. (a) Cryo-EM processing workflow. Data were processing in SIMPLE, cryoSPARC and RELION

[162,166,187]. (b) FSC curves for global-resolution estimation of the final map, using the 0.143 criterion for reporting resolution.

The high-resolution map obtained from this sample strongly resembled the closed Hrd1 state (figure 5.4). The Hrd1 closed state could be fit inside of the TMD map density and there was no obvious density for FAM8A1. Even at low map contour levels, there was no apparent density for another protein (figure 5.4). Furthermore, throughout all stages of the processing, no density was seen for FAM8A1, ruling out the possibility that this state was selected against during processing. However, it may be possible that only micelles containing Hrd1 alone yielded high resolution information and so the final particles could represent dissociated complexes. The interaction between Hrd1 and FAM8A1 is predominantly driven through FAM8A1 binding to the HAF-H domain of Hrd1 [61]. There is also an interaction at the TMD level of unknown strength [61]. HAF-H binding occurs in the soluble cytoplasmic domain of Hrd1, which was never observed in the previous datasets and this is also true for this dataset. Together, the data suggest a few possibilities. One, that the TMD domains of FAM8A1 do not interact strongly with the TMDs of Hrd1, or at least do not interact in an ordered manner. A weak interaction would result in the FAM8A1 TMDs being randomly orientated within the detergent micelle and so they are averaged out during cryo-EM processing. This effect would be even more pronounced if the stoichiometry was not 1:1 and the binding was not symmetric since symmetry-based refinements would reduce any non-symmetrical signal. Since FAM8A1 was known to be present prior to grid making, alternatively, the FAM8A1 TMDs may be in a different micelle whilst binding to the HAF-H domain still occurred. Since the HAF-H domain is flexibly linked to Hrd1 and thus unlikely to be observable without significant structural rearrangements, this would explain the lack of FAM8A1 density. To explore this possibility, particles in the final Hrd1-FAM8A1 map were re-extracted with a larger box size and 2D classification repeated to identify particles that have additional micelles

present. No obvious 2D classes were identified with more than one detergent micelle. Herp may also be required for a stable TMD interaction between Hrd1 and FAM8A1 and in its absence, there could be weak association of Hrd1 with the FAM8A1 TMDs. Finally, this could also be a detergent issue. There exist other milder detergent conditions for which this complex could be reconstituted in and perhaps in DDM-CHS, the micelle environment is not compatible with a stable Hrd1-FAM8A1 complex. This is also evidenced by the SEC profile where two main peaks were observed along with a larger void peak. There was also a lot of Hrd1 in the unbound fraction of the second affinity purification with FAM8A1 and so together, the data suggest non-optimal reconstitution conditions.

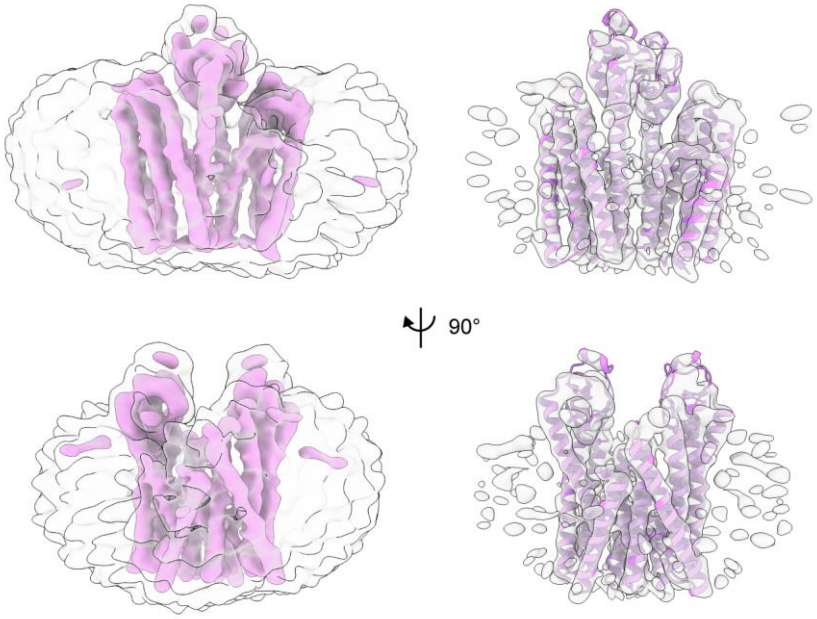


Figure 5.4. Final map for the Hrd1-FAM8A1 sample, depicted at different contour levels. The detergent micelle and lower signal regions can be seen in the transparent, low contour map. Higher signal density is shown as a high contour map in purple. The model for the closed Hrd1 state was also docked into the map using the fitmap function in ChimeraX [170]. This is shown in purple underneath a high contour Hrd1-FAM8A1 map.

We reasoned that the lack of FAM8A1 density could be detergent induced and that reconstitution of Hrd1-FAM8A1 in a milder detergent may stabilise the Hrd1-FAM8A1 interaction. Previous biochemical studies of the Hrd1-FAM8A1 complex were performed in LMNG [61]. To that end, the same complex was purified in LMNG-CHS and there was less Hrd1 in the unbound fraction during Ni-NTA purification (figure 5.5a). The abundance of co-purifying bands appeared lower but LMNG is less efficient at membrane protein solubilisation. The SEC profile revealed that the complex was more homogenous and the relative abundance of aggregated protein was decreased. There was a single main peak indicating very different behaviour to purification in DDM-CHS. SDS-PAGE analysis of the SEC elutions revealed that the complex was homogenous in composition across the peak although there was also an additional band at ~25 kDa (figure 5.5c). This is unlikely TEV protease since this was depleted by incubation with Ni-NTA resin after cleavage, and instead is possibly the same co-purifying band seen with DDM. This sample was imaged by cryo-EM but in EM micrographs, the particles appeared more clumpy and heterogenous than in the DDM condition (figure 5.6a). Some of this heterogeneity might be associated with the lower overall yield during the purification. A small data collection was performed and 2D class averages generated SIMPLE, similarly to the $Hrd1_{1\text{-}344}$ dataset [162]. Although a micelle-like object was seen in many of the 2D class averages, there was no obvious secondary structure detail (figure 5.6b). This lack of detail could be because of particle heterogeneity, thick ice resulting in low signal-to-noise, partial denaturation or other sample differences. There were no obvious Hrd1 dimer projections indicating that this sample may also not be dimeric. SEC-MALS was attempted but no reliable mass estimation could be determined for this complex.

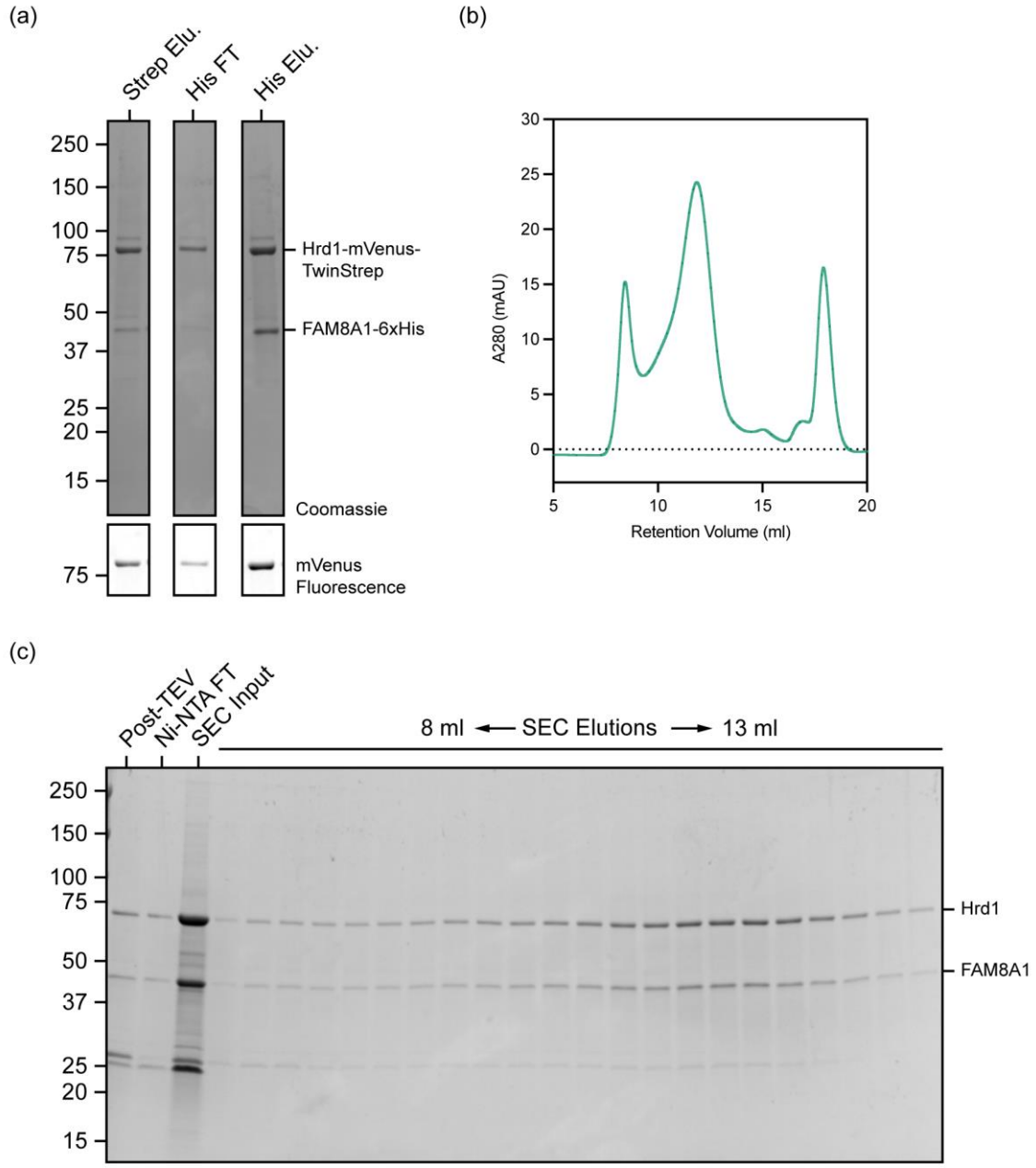


Figure 5.5. Purification of Hrd1-FAM8A1 in LMNG-CHS. (a) SDS-PAGE of the dual affinity purification elutions and the unbound fraction (FT) of the Ni-NTA purification. (b) SEC trace for the dual affinity purified material. (c) SDS-PAGE gel for the TEV cleavage reaction, reverse Ni-NTA purification flow-through (FT), SEC input and the SEC elutions.

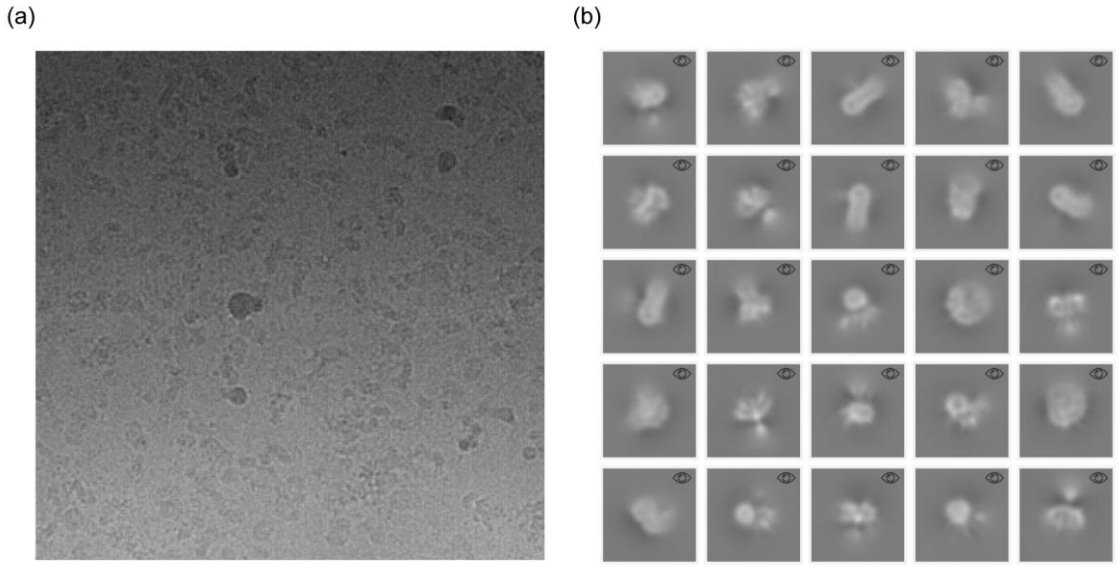


Figure 5.6. Cryo-EM screening of Hrd1-FAM8A1 in LMNG-CHS. (a) Representative micrograph of Hrd1-FAM8A1 frozen in vitreous ice. (b) Initial 2D class averages generated with SIMPLE [162].

5.2.2. Purification of Hrd1-Herp-FAM8A1

Since Herp could not be isolated with Hrd1 alone, and FAM8A1 formed a stable complex with Hrd1, it seemed reasonable that a complex containing Hrd1, FAM8A1 and Herp could be isolated. A complex containing these three proteins is a probable prerequisite to recruitment of Derlins since Herp is required to enhance the Hrd1-Derlin interaction [61]. To that end, cells expressing Hrd1-mVenus-TwinStrep were transduced with a mixed lentiviral population containing vectors for expression of untagged Herp and FAM8A1-6xHis. Antibiotic selection was used to enrich for cells expressing all three proteins.

The Hrd1-FAM8A1 complex was more stable in LMNG by SEC and therefore this detergent was used for purification including Herp. The Hrd1-Herp-FAM8A1 complex was purified as before with dual affinity purification and SEC (figure 5.7). After dual affinity purification, there was now the presence of an additional band present close to the 50

kDa marker (figure 5.7a). The SEC profile was similar to the profile of Hrd1-FAM8A1 in LMNG-CHS, and the main peak was symmetric, indicating some degree of homogeneity (figure 5.7b). SDS-PAGE of the SEC elutions showed that there was now the presence of Herp, running above the 50 kDa marker [62]. This corresponded to where full-length Herp migrated to according to the literature and was a strong indicator of a stable Hrd1-Herp-FAM8A1 complex [61]. All three proteins appeared in stochiometric amounts. This implies that FAM8A1 is required for stable association of Herp with Hrd1, as previously reported [61]. This also suggests that under these conditions one copy of FAM8A1 interacts with one copy of Herp. At the time, this material was not imaged by cryo-EM since other, more relevant complexes were pursued. However, this cell line served as the basis for introduction of further partners with the expectation of isolating larger Hrd1 complexes.

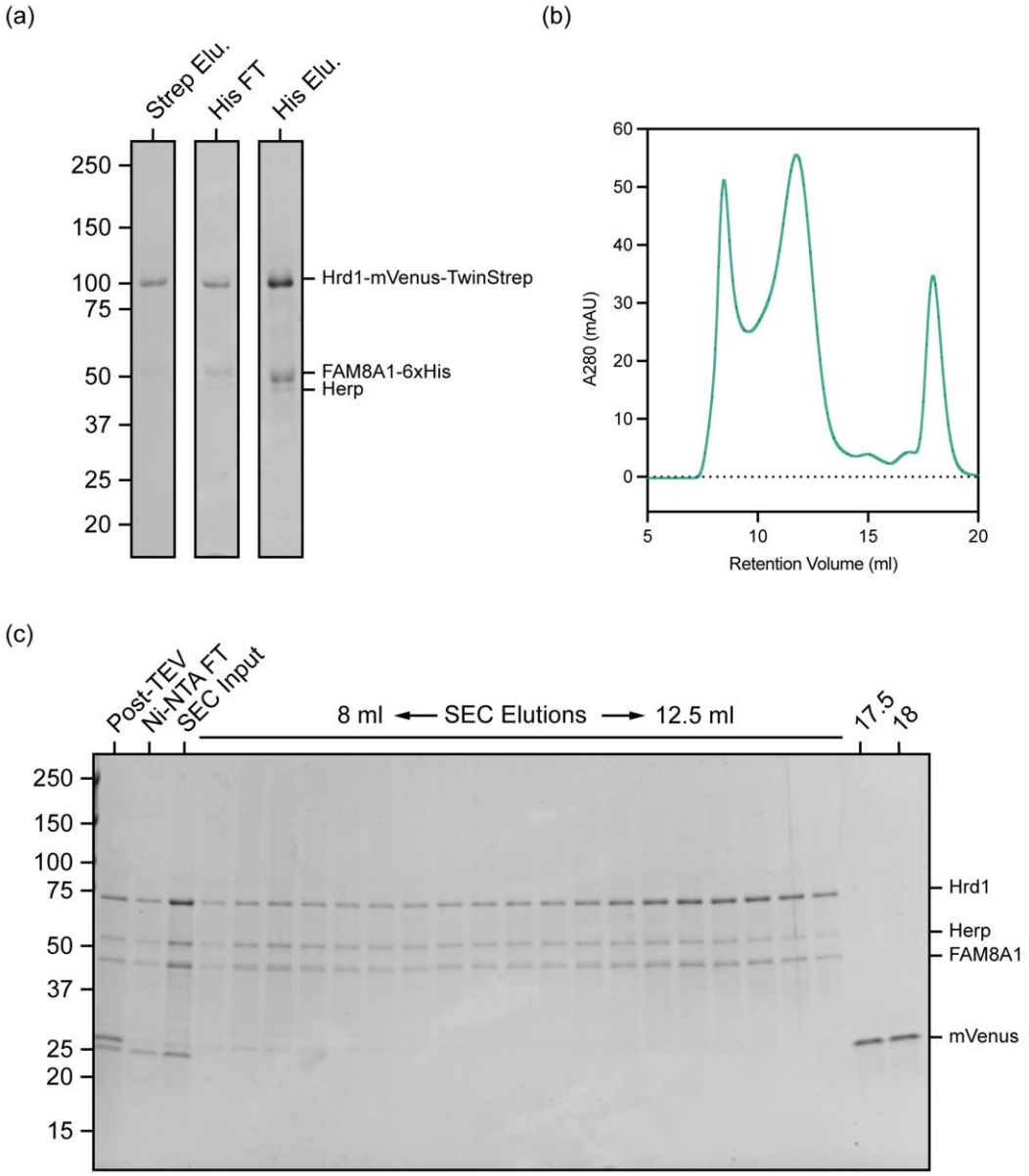


Figure 5.7. Isolation of a complex containing Hrd1, Herp and FAM8A1. (a) SDS-PAGE of elutions from the dual affinity purifications. The unbound material from the Ni-NTA (His) purification is shown as well (His FT). (b) SEC profile of the affinity purified complex. (c) Corresponding SDS-PAGE gel for the TEV cleavage reaction and the SEC elutions.

5.2.3. Purification of Hrd1-Herp-FAM8A1-DERL2

To isolate complexes containing Derlin-2 (DERL2), different approaches for overexpression were used. The suspected minimal DERL2-containing complex was Hrd1-Herp-FAM8A1-DERL2 and so initially, cells overexpressing Hrd1-mVenus-TwinStrep were transduced with a mixed lentiviral population of Herp and DERL2-6xHis. Antibiotic selection was performed to enrich for Herp and DERL2. This was followed by a further round of transduction with untagged FAM8A1. After selection for FAM8A1, a large-scale expression trial was attempted but this failed to generate high yields of complex (not shown). The overexpression burden of this combination of proteins in this expression system was likely too high. Cell lines overexpressing only three of the target proteins were then created, with the hope of capturing endogenous quantities of the fourth protein. A first attempt overexpressed Hrd1, FAM8A1 and DERL2 but this still failed to yield high levels of protein. A lot of material failed to bind during the second purification in the DERL2 pulldown. This suggested that Herp was the limiting factor and so a cell line overexpressing Hrd1, Herp and DERL2 was generated. Cells overexpressing Hrd1-mVenus-TwinStrep were transduced with lentivirus containing untagged Herp and DERL2-6xHis.

Both DDM-CHS and LMNG-CHS were trialled during an initial purification from this cell line. StrepTactin purification of Hrd1 now resulted in a range of co-purifying bands (figure 5.8a). Tagged Hrd1 was the strongest but there were now bands for what was suspected to be Herp and FAM8A1, confirmed later by mass spectrometry analysis. However, there was no obvious band for DERL2 in both detergent conditions. Western blotting revealed that there were small amounts of DERL2 co-eluting with Hrd1, but a lot of overexpressed DERL2 was in the unbound fraction (figure 5.8b). Interestingly, there was also a higher MW DERL2 species in the unbound fraction of the StrepTactin purification. This could be a non-specific cross reaction but this antibody is not reported

to cross-react in HEK cells and no evidence of this was seen in the lab [188]. Derlin-1 has been shown to form tetramers and so this band may represent oligomeric DERL2 [94]. Surprisingly, DERL2 pulldown resulted in detectable DERL2 by Coomassie staining and numerous bands were enriched. Labelled bands were sent for mass spectrometry analysis and this confirmed the presence of Hrd1, Herp, FAM8A1 and DERL2. Despite no overexpression of FAM8A1, there were near Herp-equivalent quantitates of endogenous FAM8A1 in the complex. Mass spectrometry analysis also revealed the presence of endogenous Sel1L, suggesting that there are high levels of this protein in the ER membrane. Between the two detergent conditions there were not any large differences in final complex. However, in the second purification there was more protein in the unbound fraction for DDM. This correlated with previous Hrd1-FAM8A1 purifications and indicates that the complex is more stable in milder detergents. Both conditions had similar SEC profiles and similar final yields at this scale (figure 5.9). The void peak in the LMNG condition was larger relative to the main peak but in this condition the peak at 11 ml was slightly narrower than in DDM. Overall, there was not a significant difference and a larger scale purification was repeated to obtain enough sample for cryo-EM grid preparation.

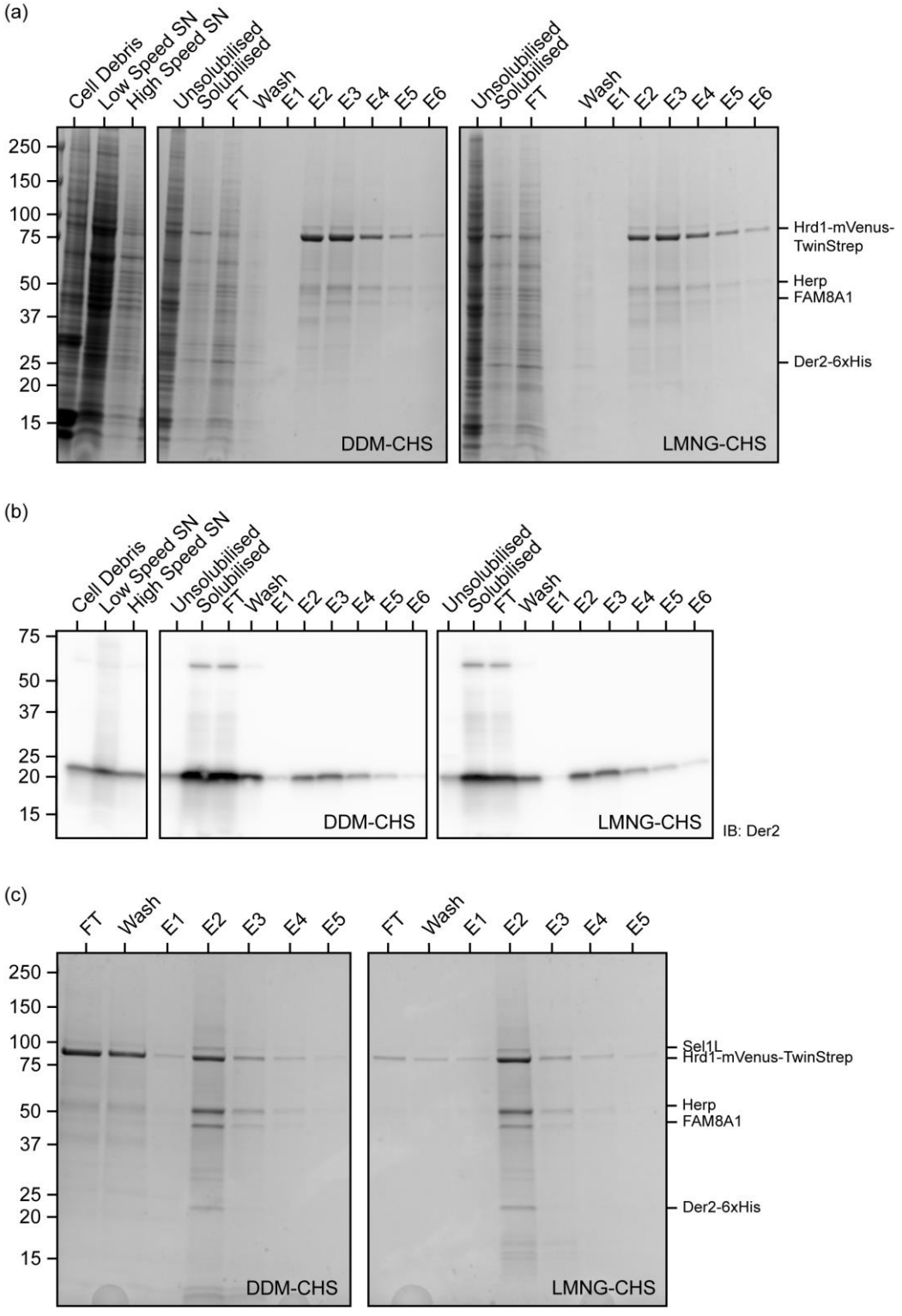


Figure 5.8. Optimising detergent conditions for purification of a Hrd1 complex containing DERL2. (a) Dual affinity purifications in DDM-CHS and LMNG-CHS for cells overexpressing Hrd1, Herp and DERL2. Cell debris is the initially pelleted material during

a low-speed spin. The low-speed supernatant (SN) and high-speed supernatant (SN) were discarded. The unsolublised membrane pellet was resuspended to the same volume as the solubilisation reaction. The unbound fraction (FT) and wash fractions are from the StrepTactin purification. Elutions are numbered E1-6. (b) DERL2 western blot for the StrepTactin purification. DERL2 is expected to run at ~25 kDa by SDS-PAGE. (c) Ni-NTA purifications of the material eluted from the StrepTactin purification. Lane names are equivalent to the StrepTactin purification.

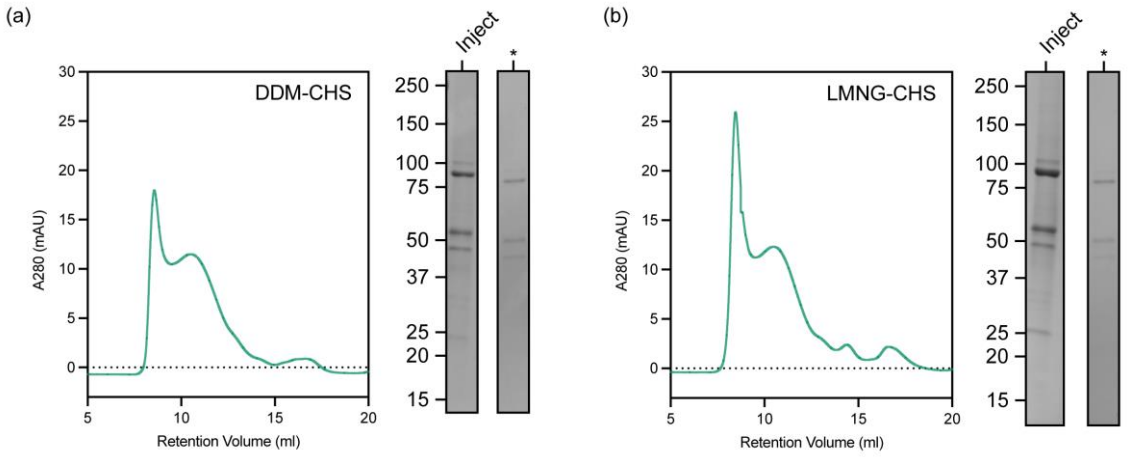


Figure 5.9. SEC profiles for DERL2-containing complexes in DDM and LMNG. (a) SEC of a DDM-CHS reconstituted DERL2-containing complex. (b) SEC of an LMNG-CHS reconstituted complex. SDS-PAGE gels for the input and peak fraction are shown for both conditions.

Larger scale purification was performed in LMNG-CHS since cryo-EM grids of small membrane proteins in this detergent are often made at lower concentrations than in DDM, thus maximising the chance of making cryo-EM grids if the yields were low. Clear bands for Hrd1, Herp, FAM8A1 and DERL2 were seen in both the StrepTactin and Ni-NTA elutions (figure 5.10a). DERL2 did not appear to be at high levels within the purified complex and it is difficult to discern if this is because DERL2 stains poorly by Coomassie

or it is has lower stoichiometry within the complex. The SEC profile was far improved from the initial screening and the main peak now eluted at ~13 ml (figure 5.10b). There was better separation from the void peak and the relative ratio between the void and main peak was improved, with less aggregation occurring. The SEC profile indicated that the material was very homogenous and SDS-PAGE of the SEC elutions identified the presence of all four proteins of the complex. Interestingly, there appeared to be a ladder of bands above DERL2, separated by equal distances (figure 5.10c). These are more obvious in the SEC input lane and their identity is unknown. No western blot for DERL2 was performed at the time so it is hard to know if these are modified forms of DERL2, other Hrd1 partners, cleavage products or contaminants. These bands appear to co-elute with the complex and so are likely part of this Hrd1 complex. Cryo-EM grids were made and a data collection performed by Justin Deme and Susan Lea. Cryo-EM data were processed as before, using a combination of SIMPLE and cryoSPARC [162,187]. Representative 2D class averages from cryoSPARC, after significant pruning of junk particles, are shown in figure 5.10d. A small proportion of particles in the dataset had some secondary structure features after 2D class averaging. These were predominantly top-down and side views, and there appeared to be at least two different types of objects present. There were distinctive S-shaped micelles as well as more flat micelles reminiscent of the Hrd1 in the open state. The pitch of the helical secondary structure relative to the micelle plane in the flat micelles is also similar to that seen in the Hrd1 open state. Unfortunately, there were not enough final particles (< 3%) to perform high resolution 3D reconstructions. A large data collection was already performed to obtain the current quality of 2D class averages and it was not feasible to collect more data. Instead, it seemed better to improve the sample quality since only a small number of particles had secondary structure. Since cryo-EM ice quality was not inhibitory for this sample, our suspicion was that this sample was limited by lack of signal outside of the micelle for alignment during cryo-EM processing. This complex is quite small and if Hrd1

is not dimeric, there would be less signal contributed by the RING domains. The other proteins in this complex are not predicted to have large, stably associated folds outside and close to the micelle [136]. Addition of further partners in this complex could help to improve alignment. One candidate for this would be Sel1L since this has large SLR motifs in the ER lumen which contain a series of helical repeats [178]. Since there is a small aggreging SEC shoulder, this may also be caused by detergent-induced heterogeneity. Only DDM resulted in a high-resolution structure for Hrd1 but DDM is known to be unsuitable for the Hrd1-FAM8A1 complex. Using an even milder detergent for reconstitution may instead improve the behaviour of this complex.

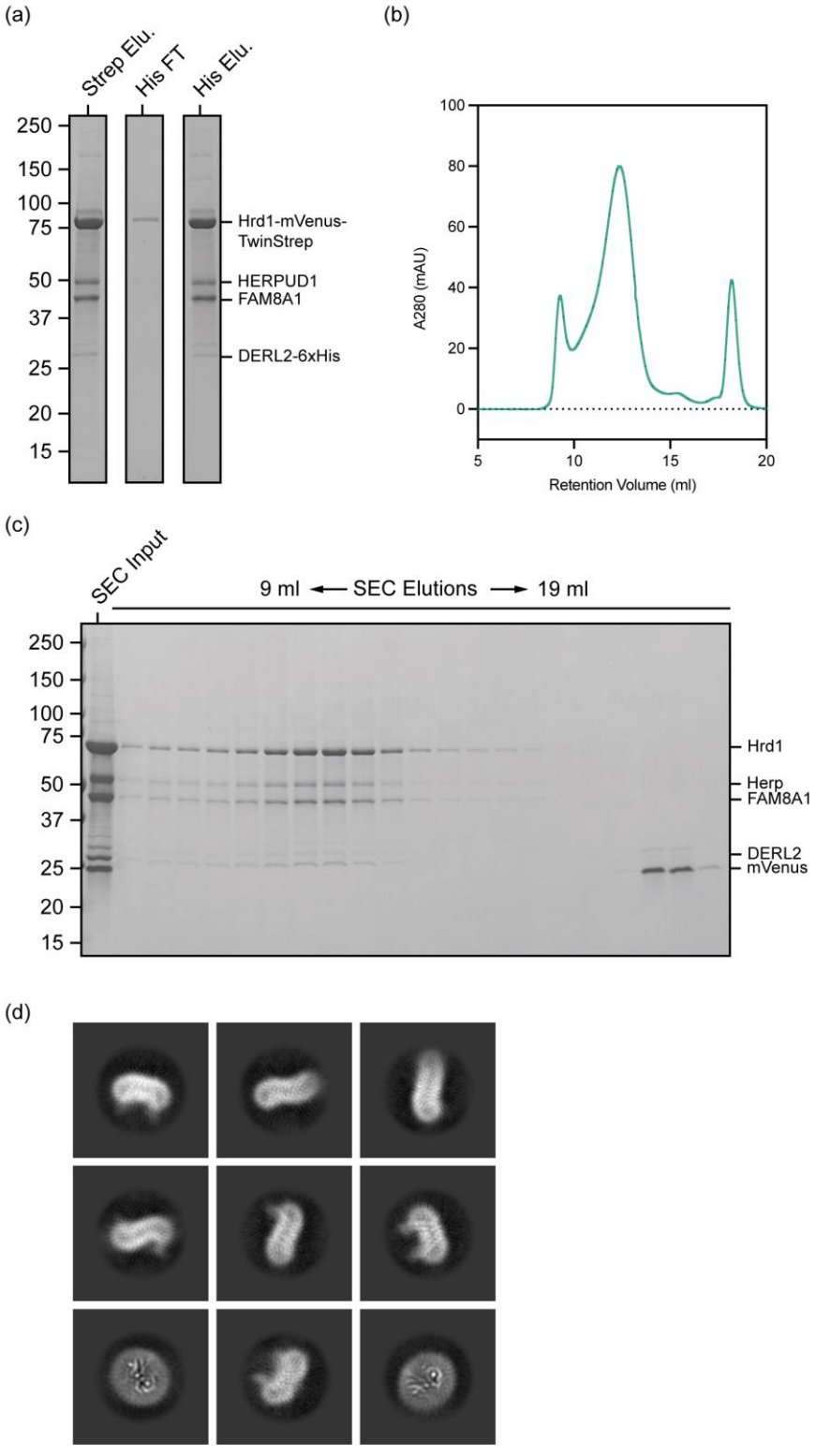


Figure 5.10. Larger scale purification and cryo-EM screening of a complex containing Hrd1, Herp, FAM8A1 and DERL2, in LMNG-CHS. (a) SDS-PAGE of the dual affinity purifications. (b) SEC profile for the dual affinity purified complex. (c) SDS-PAGE gel for

the SEC input and elutions. (d) 2D class averages from the cryo-EM screening of grids made from material in the peak SEC fraction. Cryo-EM processing was done with SIMPLE and cryoSPARC [162,187].

Purification in an even milder detergent than LMNG (GDN) was also explored as a way to improve the sample quality (figure 5.11). From a cell line expressing Hrd1, Herp and DERL2, the complex was purified as before but in the presence of GDN detergent. There were no obvious differences in purifications prior to SEC but the SEC profile of the GDN reconstitution was changed (figure 5.11a-b). The complex shifted ~0.5 ml closer to the void volume, indicating that it was now larger than in the LMNG-CHS purification. Some of this can be attributed to the fact that GDN often forms micelles with a large diameter [151]. A comparison between the concentrated material prior to SEC in both LMNG and GDN revealed that in GDN, Sel1L was more abundant within the complex (figure 5.11c). Endogenous Sel1L was observed to co-purify in previous Hrd1 complex purifications across a variety of detergent conditions, but this suggests that Sel1L binding was improved in milder detergents. SDS-PAGE of the SEC elutions revealed that the complex was very similar to the one purified in LMNG (figure 5.11d). Both the GDN and LMNG reconstitutions were done with the same cell mass input but the final DERL2 concentration was lower than in the LMNG purification. This is probably because the solubilising efficiency of GDN is lower than LMNG and so less total protein was solubilised. This sample was not prepared on cryo-EM grids since there was no significant difference in the SEC profile or of the SDS-PAGE gels, and a similar complex in GDN was prepared at the time (discussed later).

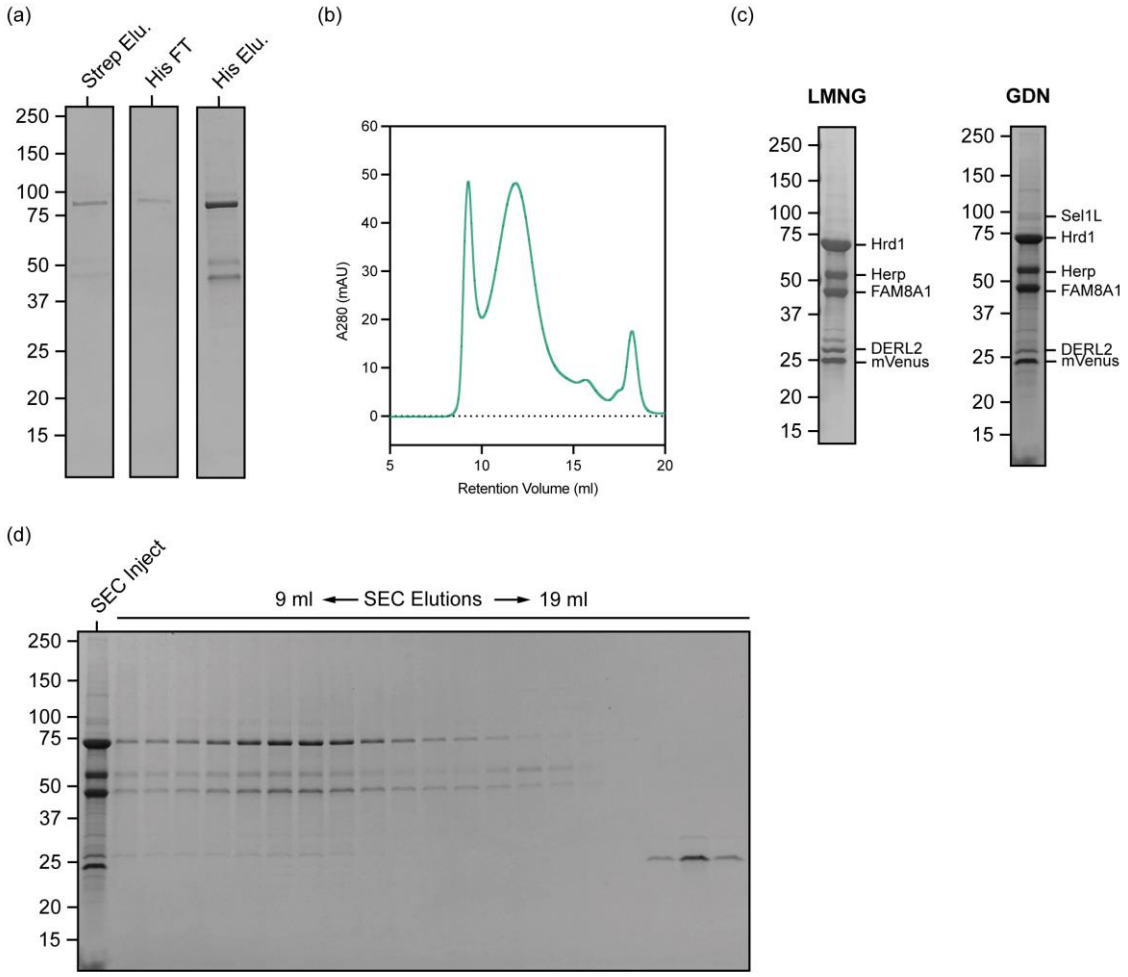


Figure 5.11. Larger scale purification of Hrd1-Herp-FAM8A1-DERL2 in GDN. (a) SDS-PAGE gels for the dual affinity purifications. (b) SEC profile for the dual affinity purified material. (c) Comparison of the concentrated material prior to SEC for the Hrd1-Herp-FAM8A1-DERL2 complex in LMNG and GDN. (d) SDS-PAGE gel of the SEC elutions.

5.2.4. Attempts to reconstitute Sel1L with Hrd1

To aid cryo-EM studies, a Hrd1 complex containing Sel1L was sought after since this would increase the overall mass of the complex significantly and theoretically provide folded mass near the micelle, to help in cryo-EM processing [55,136]. Different attempts were made to reconstitute Hrd1 with Sel1L, which proved difficult to isolate in a stable form. Initially, various truncations were made of Sel1L to recapitulate similar truncations

made for the yeast Hrd1-Hrd3 structures. In these structures, the C-terminal TMD was truncated to produce a soluble form of Hrd3 [55,60]. A stable cell line expressing Sel1L$_{1-740}$ was made to generate a soluble form of Sel1L without the C-terminal TMD, as predicted in secondary structure predictions and sequence alignments with yeast Hrd3 [136,189]. Purified Sel1L could then be added to purified Hrd1 complexes since the binding is thought to occur via a lumenal-linking region between TM helices 1 and 2 [55]. This truncation was TwinStrep tagged but it did not express well in HEK cells and the small amounts that did purify were cleaved in many places. In the literature, overexpressed Sel1L is known to be rapidly degraded whilst co-expression of Hrd1 stabilises Sel1L levels [184]. Expression of other constructs were also attempted in *E. coli* SHuffle T7 cells but with no success. To aid expression, an alternative cell line overexpressing Sel1L$_{1-740}$-6xHis and Hrd1-mVenus-TwinStrep was made, with the plan to co-purify the two proteins together (figure 5.12a). This was done in LMNG-CHS and showed an improvement in yield and a clear band for Hrd1. SEC analysis revealed that the complex was very heterogenous and the final yields still low (figure 5.12d). The Sel1L truncation may have caused poor behaviour of the complex and so full-length Sel1L was co-expressed with Hrd1 and purified in different detergents, to see if this improved yield and final behaviour by SEC (figures 4.12b-c). The DDM-CHS purified material behaved similarly to the previous purification and final yields were not much improved. In LMNG-CHS with full-length Sel1L, yields were much improved. The SEC profile was still heterogenous and exhibited some degree of aggregation (figure 5.12d). SDS-PAGE of the LMNG SEC elutions revealed that the ratio of Sel1L to Hrd1 changed across the peaks, suggesting the presence of different complexes (figure 5.12e). Cryo-EM of the complex revealed that it was very aggregated and heterogenous, so no data collection was performed (figure 5.12f). This correlated with the poor behaviour by SEC.

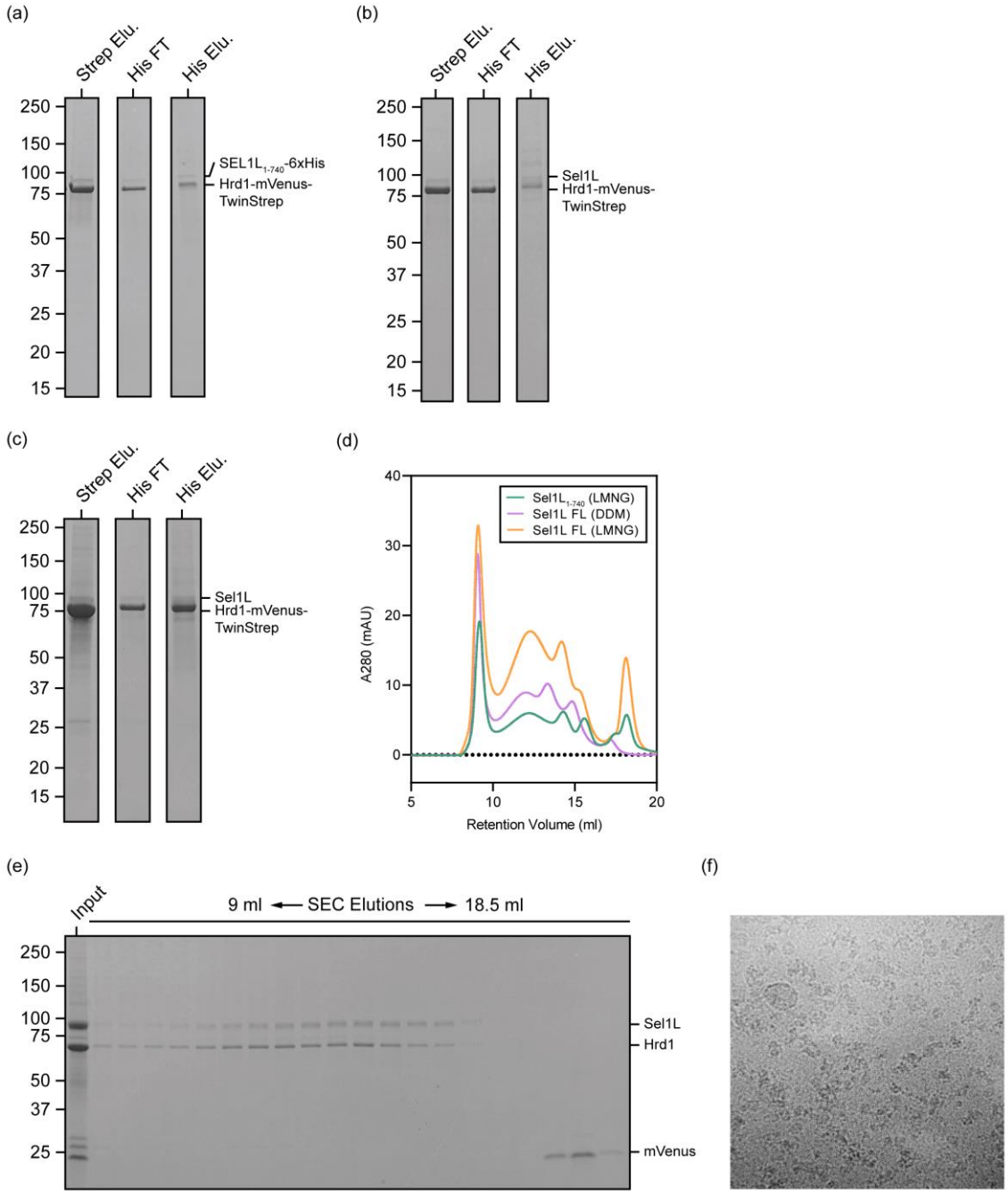


Figure 5.12. Co-purification of different Sel1L variations with Hrd1 in different detergent conditions. SDS-PAGE gels from the dual affinity purifications of Hrd1-Sel1L$_{1-740}$ in LMNG-CHS, Hrd1-Sel1L in DDM-CHS and Hrd1-Sel1L in LMNG-CHS are shown in (a), (b) and (c), respectively. (d) SEC profiles for the purifications of Sel1L$_{1-740}$ and full-length Sel1L with Hrd1. Coloured green is the Sel1L$_{1-740}$-Hrd1 purification in LMNG-CHS, purple the Sel1L-Hrd1 purification in DDM-CHS and in orange is the Sel1L-Hrd1 purification in

LMNG-CHS. (e) SDS-PAGE gel of the SEC elutions from the Hrd1-Sel1L purification in LMNG-CHS. (f) Cryo-EM micrograph of Hrd1-Sel1L in LMNG-CHS.

Sometime after generating cell lines for Sel1L overexpression, attempts were made to isolate the lumenal interacting partners of Sel1L, namely OS-9 and XTP3-B [62,183,186]. A cell line was generated that overexpressed Hrd1-mVenus-TwinStrep, Sel1L-6xHis and untagged OS-9. At this point, it was known that GDN was a better detergent for reconstitution of Sel1L in larger Hrd1 complexes. The cell line expressing Hrd1, Sel1L and OS-9 was scaled up and purified in GDN (figure 5.13). This resulted in less Hrd1 in the unbound fraction after Sel1L pulldown, as compared to previous attempts to co-purify Hrd1 with Sel1L. Sel1L and Hrd1 were more abundant and were equimolar by SDS-PAGE across the SEC peak (figure 5.13c). Although complexes containing OS-9 were not enriched by pulldown, there was still no OS-9 present at any stage during the purification. However, Sel1L now appeared as a doublet of bands near the 100 kDa marker. Overall yields were improved but there was still significant aggregation as indicated by the large shoulder from void peak to the peak at 13.5 ml. There was also the presence of a band near the 25 kDa marker throughout the SEC elutions. Cryo-EM imaging revealed that the frozen protein was quite heterogenous and 2D class averages from a small data collection did not contain any secondary structure (not shown). Reconstitution of Sel1L with Hrd1 proved to be difficult regardless of the purification conditions or modifications made to the Sel1L sequence. Complexes visualised by cryo-EM were always heterogenous and together, this indicates some level of instability of Sel1L that is not present in yeast Hrd3, since structures of yeast Hrd1-Hrd3 and Hrd3 alone could be determined [55,60]. This could be attributed to lack of correct folding during overexpression, non-favourable detergent conditions or lack of an important binding partner for stabilisation. The use of reducing agents was also trialled during purification but these failed to improve Sel1L stability. Overexpression of OS-9

and XTP3-B with Hrd1-mVenus-TwinStrep and Sel1L was trialled as well but this behaved similarly to overexpression of OS-9 alone (not shown). There was still no presence of OS-9 or XTP3-B and the same behaviour of Hrd1-Sel1L was observed as with expression of OS-9 alone.

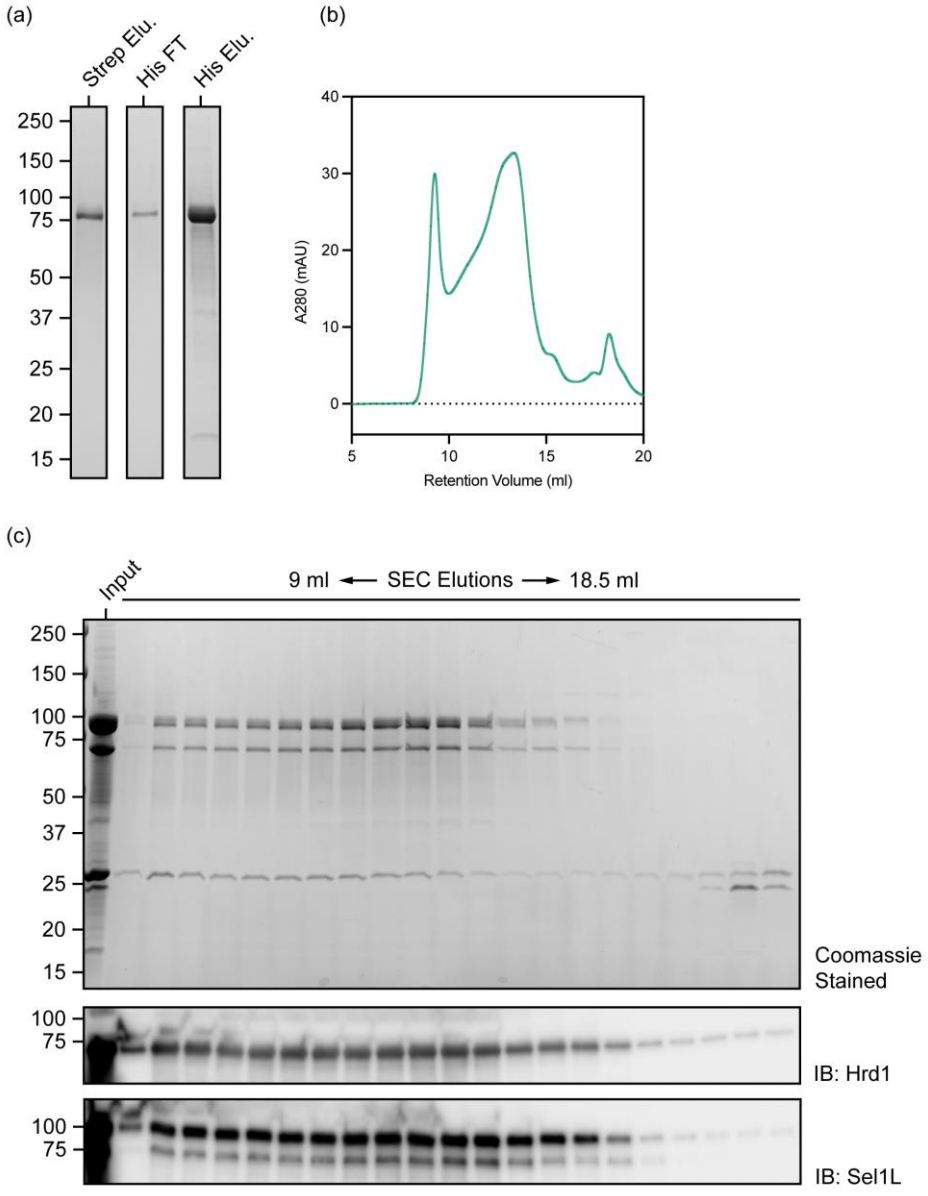


Figure 5.13. Purification of Hrd1-Sel1L during OS-9 overexpression. (a) SDS-PAGE gels from the dual affinity purifications in GDN. (b) SEC profile for Hrd1-Sel1L in GDN. (c) SDS-PAGE gel and western blots of the SEC elutions.

5.2.5. Purification of Hrd1-Herp-FAM8A1-DERL2-Sel1L

Although a 4-component complex containing Hrd1, Herp, FAM8A1 and DERL2 behaved well biochemically, it failed to yield high-resolution structural information by cryo-EM. Increasing the total mass of this complex, especially by adding mass outside of the detergent micelle, seemed like the best route for overcoming this problem. This would theoretically improve the quality of alignments during cryo-EM processing. In previous purifications, a small amount of Sel1L was always present and so boosting the expression levels of Sel1L seemed like a good approach to enrich for this. To that end, the stable cell line expressing Hrd1-mVenus-TwinStrep, Herp and DERL2-6xHis was further transduced with lentivirus containing a vector for overexpression of full-length Sel1L. This was enriched with antibiotic selection. The complex was initially reconstituted in LMNG-CHS and the StrepTactin and Ni-NTA purifications behaved very similarly to the 4-component complex in the same detergent (figure 5.14a). Upon concentration of the dual affinity purified material for SEC, it became obvious that there was now a greater abundance of Sel1L, as seen between the 75 and 100 kDa markers (figure 5.14b). Overexpression of Sel1L without enrichment by pulldown was enough to increase the presence of this protein within the complex, presumably because the Hrd1-Sel1L interaction is strong. The identity of this band was further confirmed by mass spectrometry. SEC of this complex revealed that it was very homogenous, with a single main peak that contained all 5 proteins (figure 5.14c-d). The peak eluted at a similar volume to the 4-component complex in the same conditions which was surprising, given that addition of Sel1L would cause a large increase in total mass. This implies that the majority of the hydrodynamic radius is conferred by other components in the complex or the detergent micelle. SDS-PAGE of the SEC elutions revealed that Sel1L was not as abundant as the other bands in the main peak and instead appeared to be dissociating

from the main complex or assembling a smaller complex with different stoichiometries of binding partners (figure 5.14d). The majority of Sel1L eluted closer to 15 ml in a complex containing Hrd1, Herp, FAM8A1 and Sel1L. There is no visible DERL2 in this region and so this could represent a different assembly state. The material in the 12.5 ml peak was concentrated and imaged on cryo-EM grids.

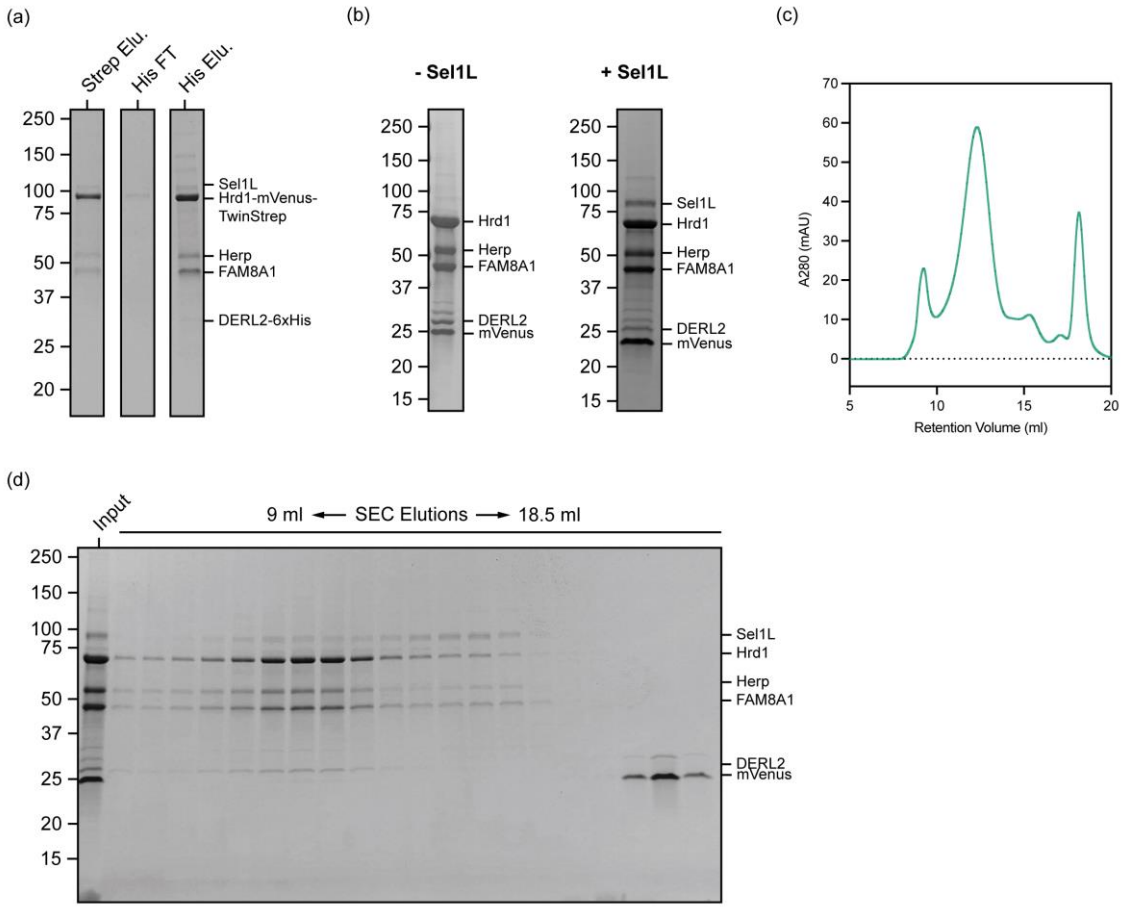


Figure 5.14. Purification of a complex containing Hrd1, Herp, FAM8A1, DERL2 and Sel1L. (a) SDS-PAGE gels from the dual affinity purifications in LMNG-CHS. (b) Comparison of the concentrated material prior to SEC, after dual affinity of Hrd1 complexes with and without Sel1L overexpression. (c) SEC profile for the dual affinity purified material after Sel1L overexpression. (d) SDS-PAGE gel for the SEC elutions.

Frozen protein on cryo-EM grids was deemed homogenous enough for data collection. Justin Deme and Susan Lea collected data on a Titan Krios G4 operating at 300 kV, in the same way as other samples in this chapter. Data were collected in counting mode in EER format at 165,000X magnification, with a pixel size of 0.693 Å. 19,259 movies (4,404,470 particles) were collected in total and these were initially processed in SIMPLE, as previously done (figure 5.15a) [162]. Good particles (3,607,352) were exported to cryoSPARC and after multiple rounds of 2D classification, cleaner particles were subject to *ab initio* generation of three models [187]. Classes 1 and 3 yielded maps that had features resembling protein secondary structure, but further NU-refinements did not increase the resolution of these maps any further.

2D classification of the particles in the *ab initio* classes revealed two distinct protein conformations/complexes (figure 5.15a). Class 1 particles resolved into a flatter micelle-like object, whilst particles in class 3 were arranged in an S-shaped micelle object (figure 5.15b-c). This correlated with previous results from 2D classification of the Hrd1-Herp-FAM8A1-DERL2 complex without Sel1L, but the resolution of these 2D classes and of the *ab initio* models was higher. In the 2D class averages, the TMDs inside of the micelle were more clearly resolved and some connectivity could be seen through the micelle in the maps (figure 5.16). The presence of Sel1L was not obvious in the cryo-EM maps or 2D classes, despite Sel1L containing a large soluble domain. This soluble domain is made up of SLR repeats and if the Hrd1-Sel1L arrangement is similar to the yeast Hrd1-Hrd3 arrangement, these would be expected to be close to the micelle surface [55,178]. No region of Sel1L was visible and this suggests that either particles were picked during the processing that did not contain Sel1L or that the soluble domains of Sel1L underwent denaturation during grid preparation. Denaturation at the air-water interface during cryo-EM grid preparation is a common problem and can be prevented by sample modification prior to freezing [190,191]. These approaches were not explored at the time and are a possible avenue for future structural studies on these complexes. Despite the high

numbers of starting particles, the final map was low resolution. This is likely limited by the particle number in the first instance and most particles were excluded during the processing. Many of the particles did not yield high-resolution alignments during 2D classification and this could be attributed to lack of signal in the particles and/or particle heterogeneity. Denaturation of Sel1L may be contributing to this but the addition of Sel1L alone appeared enough to enhance the quality of the cryo-EM data, supposedly by addition of an extra TMD within the micelle.

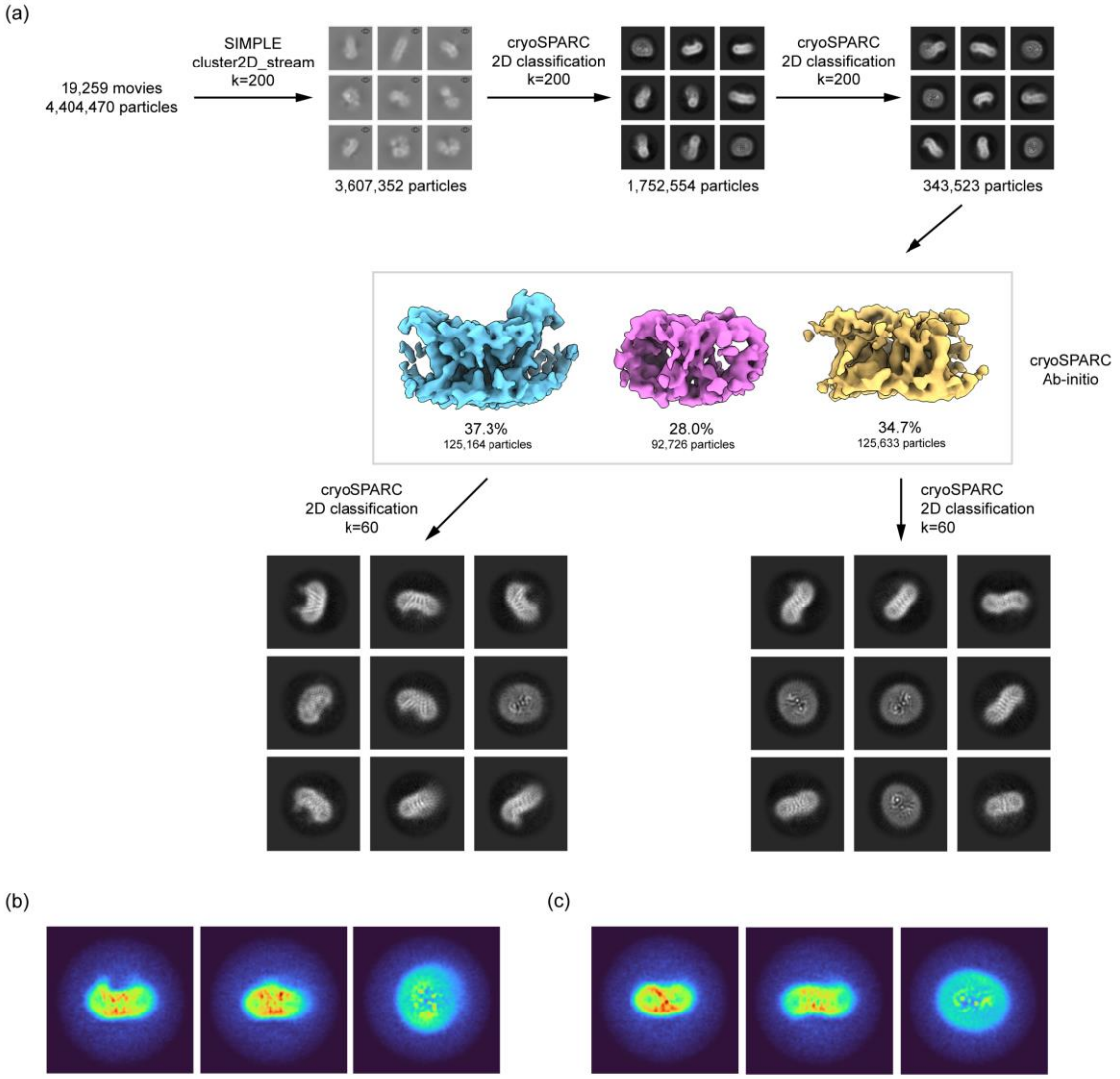


Figure 5.15. Cryo-EM processing workflow for the Hrd1-Herp-FAM8A1-DERL2-Sel1L sample. (a) SIMPLE and cryoSPARC processing workflow [162,187]. (b) and (c)

represent real space slices of the class 1 and 2 *ab initio* volumes, respectively. These are cross-sectional views of signal in the *ab intio* maps, viewed from the x, y and z planes.

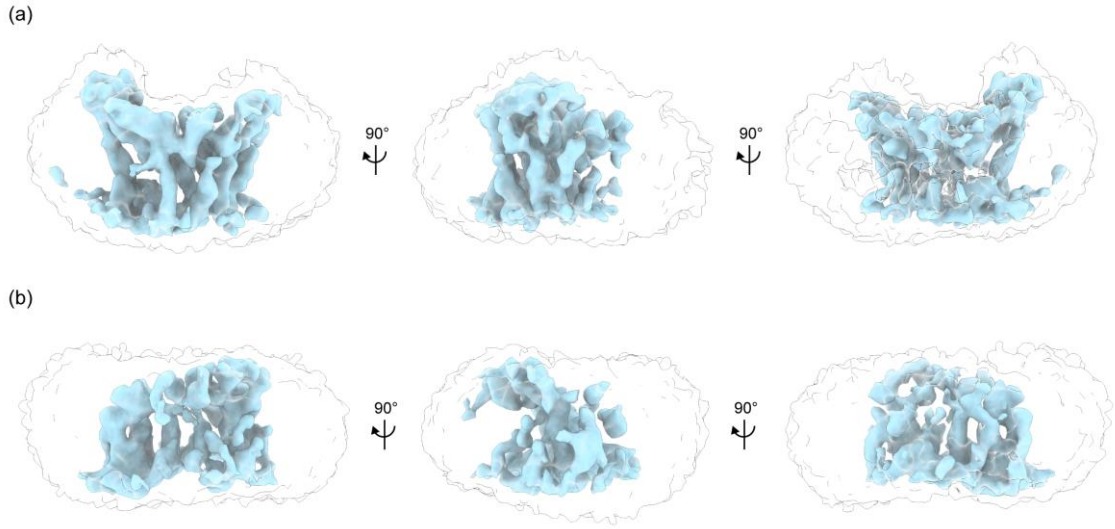


Figure 5.16. Cryo-EM maps of *ab initio* classes 1 and 3. (a) and (b) are overlays at different contour levels for *ab initio* class 1 and 2 maps, respectively. The high contour map representing probable protein density is shown in light blue and a low contour map representing probable detergent density is overlayed with transparency.

Previous reconstitutions of Sel1L with Hrd1 revealed that this interaction was best stabilised when reconstituted in GDN detergent and so cells expressing Hrd1, Herp, DERL2 and Sel1L was scaled up for a GDN reconstitution. Using this detergent could improve overall sample quality and thus aid structural studies. This may also stabilise Sel1L within the complex and limit any denaturation that may be occurring during cryo-EM grid preparation. The complex was purified at a similar scale to the LMNG condition and the affinity purifications behaved similarly (figure 5.16a). SEC of the dual affinity purified material revealed a larger complex, running closer to the void peak (figure 5.16b). This change in elution volume was seen with previous Hrd1 complexes when GDN was used and is likely partially driven by the increase in the detergent micelle size.

This shift correlated to a ~1 ml change in retention volume and also coincides with protein changes within the main peak (figure 5.16c). Sel1L is now more abundant in the main peak and is no longer dissociating into a second complex with a smaller hydrodynamic radius (figure 5.16c). GDN reconstitution therefore improved Sel1L stability within the complex. The behaviour of Herp within this complex was also slightly altered and there was now a small amount of Herp eluting closer to 15 ml, suggesting that Herp is not as stably associated in GDN. Comparison of the concentrated material prior to SEC for the LMNG-CHS and GDN conditions reveals a higher abundance of Sel1L in GDN. This was a good sign and so this sample was also imaged by cryo-EM. A multi-million particle dataset was collected but despite numerous rounds of 2D classification in cryoSPARC, no strong secondary structure detail was resolved during 2D class averaging (figure 5.16e). Some connectivity that resembled TMDs could be seen within the detergent micelle but this was very low resolution and only present in a few 2D class averages. The presence of Sel1L was not obvious, despite its higher abundance within the complex. The larger GDN micelle could be obscuring cryo-EM signal within the detergent micelle, thus preventing high resolution particle alignment during processing. This sample appears the most homogenous of all complexes imaged but the LMNG-CHS sample yielded the best cryo-EM data for a large, multi-component Hrd1 complex.

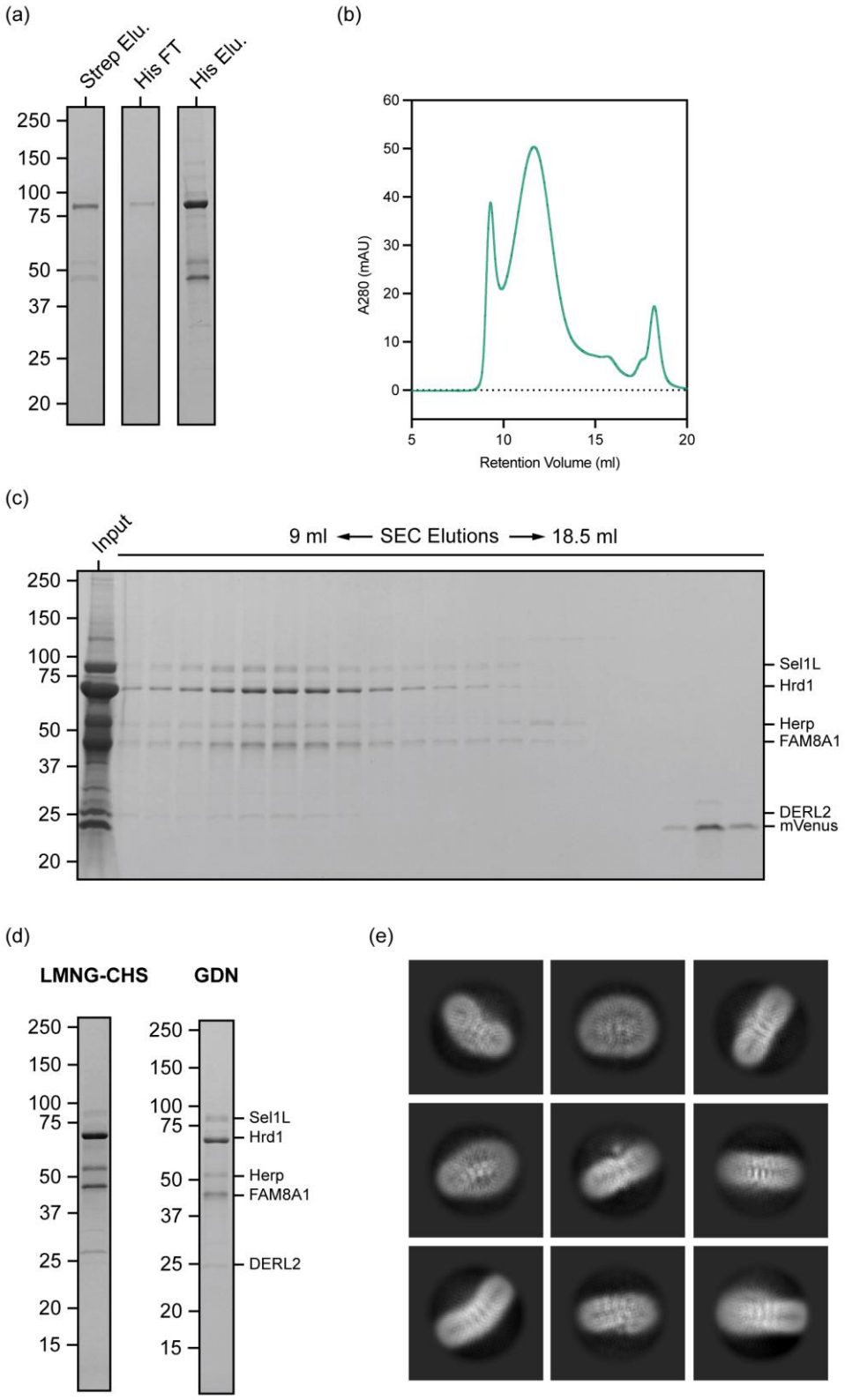


Figure 5.17. Purification of the Hrd1-Herp-FAM8A1-DERL2-Sel1L complex in GDN. (a) SDS-PAGE gels for the dual affinity purifications. (b) SEC profile for the dual affinity

purified complex. (c) SDS-PAGE gel for the SEC elutions. (d) Comparison of the concentrated material prior to SEC for the LMNG-CHS and GDN purified Hrd1-Herp-FAM8A1-DERL2-Sel1L complex. (e) 2D class averages of the GDN-reconstituted complex. These were produced with a combination of processing in SIMPLE and cryoSPARC [162,187].

5.2.6. Purification of DERL1 with a Hrd1 complex

At the time of making cell lines for overexpression of DERL2 with Hrd1 complexes, a cell line was also made for expression of DERL1 in a similar context. DERL1 is known to associate with mammalian Hrd1 and thought to assemble a similar complex with the same interacting partners as DERL2 [62]. AlphaFold predictions also support this [136]. DERL1 shares a sequence identity of 36% with DERL1, indicating some level of structural conservation [135]. A cell line was generated for overexpression of Hrd1-mVenus-TwinStrep, Herp and DERL1-6xHis, assuming that Herp and FAM8A1 would be required for DERL1 assembly. FAM8A1 was not expressed because for cell lines producing DERL2-containing Hrd1 complexes, endogenous levels of FAM8A1 were high enough for assembly of Herp with Hrd1.

This cell line was purified at a similar scale to the DERL2-containing complex purifications and the complex extracted in LMNG-CHS (figure 5.18). Hrd1 was still abundantly expressed in this cell line and large amounts were enriched during the StrepTactin purification (figure 5.18a). Strikingly, a band for DERL1 between the 25 and 20 kDa marker was seen during the StrepTactin purification, something that was never observed for DERL2. Either DERL1 stains better by Coomassie than DERL2 or its expression is much higher. There were no obvious bands for Herp or FAM8A1 in the initial purification but pulldown on the His tag of DERL1 enriched for these proteins, albeit at low levels. There was also significant abundance of Sel1L within this complex after

the dual affinity purification. Between the 50 and 75 kDa markers, there was also the abundance of an additional band that was never been seen before with other Hrd1 complex purifications. The SEC profile of this complex was quite heterogenous, some of which could be attributed to low final yields and background signal from contaminating proteins (figure 5.18b). The 12.5 ml peak eluted at a similar location to an equivalent DERL2-containing complex in the same detergent, implying that the complex is of similar hydrodynamic radius (figure 5.18c). The elution fractions contained low abundance of the proteins of interest and Hrd1 was the most abundant (figure 5.18d). DERL1, Herp and FAM8A1 could be seen co-eluting with Hrd1, as well as the unknown band seen from the purifications, running just below Hrd1 near the 75 kDa marker. There was also a large amount of what appeared to be contaminating TEV protease in the SEC elutions. The significance of this interaction is unknown but it was observed in GDN purifications of other Hrd1 complexes. Although this complex was not pursued any further, it suggests that DERL1 assembles with Hrd1 through interaction with Herp and FAM8A1, in a similar manner to DERL2. If true, this mode of Hrd1 interaction appears conserved across DERL1 and DERL2, and a similar mode of interaction with DERL3 may also exist.

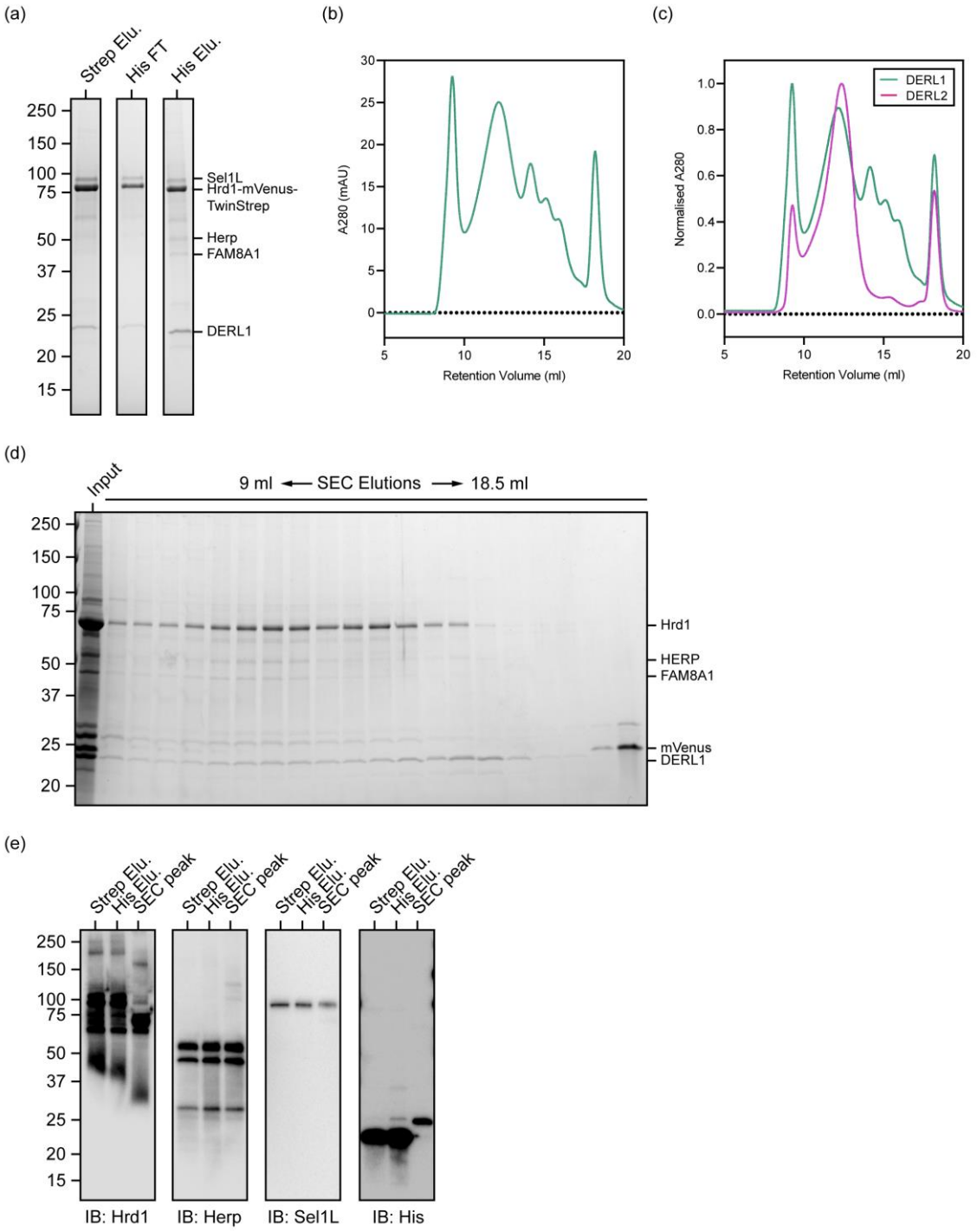


Figure 5.18. Isolation of a Hrd1 complex containing DERL1. (a) SDS-PAGE gels for dual affinity purifications. (b) SEC profile of the dual affinity purified material. (c) Comparison of the DERL1- and DERL2-containing Hrd1 complexes by SEC in LMNG-CHS. The DERL1-containing complex is shown in green whilst the DERL2-containing complex is

shown in purple. (d) SDS-PAGE of the SEC elutions from (b). (e) Western blots for different partners of the Hrd1 complex at different purification stages.

5.2.7. Attempts to reconstitute substrate with Hrd1

To gain insight into substrate binding to the Hrd1 complex, attempts were made to reconstitute an artificial substrate with Hrd1. The α-chain of the T-cell antigen receptor (TCRα) is a well characterised ERAD substrate that undergoes rapid degradation when unassembled [192]. This is dependent on Hrd1-mediated ERAD and so this substrate was used in reconstitution trials [34]. Michael van de Weijer in Pedro Carvalho's group had previously reproduced a substrate degradation assay for Hrd1 where TCRα was made as a fusion with the Tac antigen [193,194]. This fusion is based off of previous work where replacement of the TMD of the stable Tac antigen with that of the TMD of TCRα (TCRαTM) resulted in it being rapidly degraded by ERAD [195]. This fusion was cloned with a downstream His tag and initially co-expressed with Hrd1-mVenus-TwinStrep alone. This resulted in poor yields and the substrate did not express well, likely because of Hrd1-mediated degradation. Instead, the RING domain of Hrd1 was disrupted by mutation of cysteine 291 to alanine. This was co-expressed with the artificial TAC-TCRαTM substrate and purified in LMNG-CHS (figure 5.19). Small amounts of Hrd1-substrate complex were isolated after dual affinity purification and western blotting confirmed that a small amount of substrate was associated with the complex. SEC of the purified complex revealed that it assembled a large complex close to the void volume (figure 5.19b). SDS-PAGE of the SEC elutions revealed fractional amounts of substrate associated with the complex, but this eluted close to the void volume (figure 5.19c). There was contaminating TEV protease near the 25 kDa marker again, the significance of which is still unknown and probably a non-physiological interaction. More interestingly, there were numerous bands present above Hrd1, around the 100 kDa marker (figure

5.19c). From previous western blots for Sel1L, the band closest to 100 kDa likely is Sel1L whilst the band above the 100 kDa marker has never been seen before in Hrd1 reconstitutions. In a previous experiment, a doublet of Sel1L was observed but this unknown band does not correlate with the MW of the doublet band. Sel1L is known to be glycosylated in cells and so this could represent an alternative glycosylation state, since there are no known other Hrd1 partners of this MW [177]. An attempt was made to also reconstitute this substrate with a complex containing Hrd1, Herp, FAM8A1 and DERL2 but to no success as the yield was too low. Optimisation of these reconstitutions may yield a stable Hrd1-substrate complex that could be studied by cryo-EM. This would provide valuable insight into substrate recognition by ER-associated E3 ligase complexes.

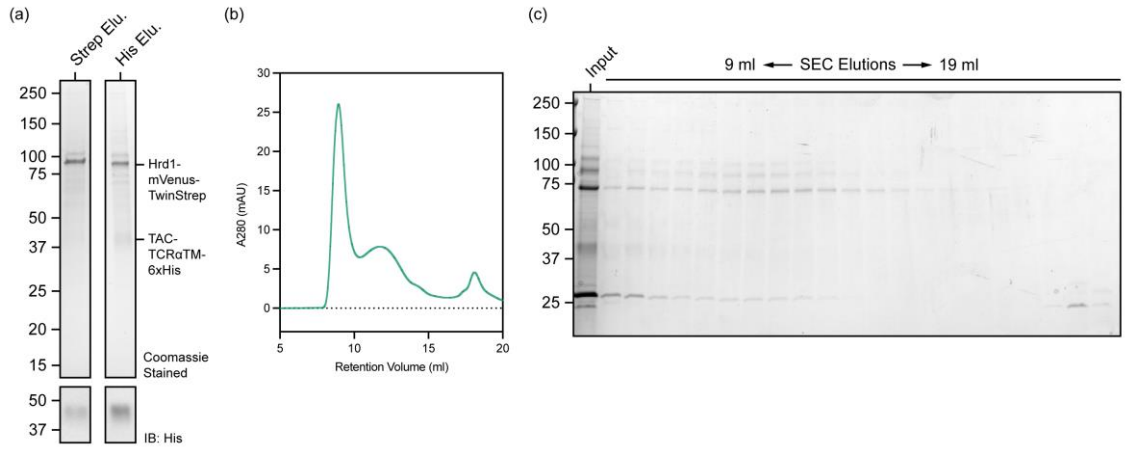


Figure 5.19. Purification of Hrd1 with the artificial substrate TAC-TCRαTM. (a) SDS-PAGE of the dual affinity purifications in LMNG-CHS. (b) SEC profile of the dual affinity purified complex. (c) SDS-PAGE gel of the SEC elutions.

5.3. Discussion

In this chapter, various higher-order Hrd1 complexes were produced via recombinant expression in a mammalian cell expression system. A series of increasingly larger

complexes were produced by sequential expression of different Hrd1 partners, to identify the components required for minimal assembly of a “complete” Hrd1 complex. Hrd1 assembled with FAM8A1 stably in what appeared to be a stochiometric complex, in both DDM and LMNG detergents. Both these samples were imaged by cryo-EM but only the DDM sample yielded high-quality 3D reconstructions. No FAM8A1 was seen in the cryo-EM maps but Hrd1 was arranged in a dimeric state reminiscent of the closed Hrd1 dimer. The bulk of the interaction between Hrd1 and FAM8A1 occurs in the disordered cytoplasmic tail of Hrd1 where FAM8A1 engages a conserved HAF-H domain [61]. Although an interaction between the TMDs of Hrd1 and FAM8A1 is also reported, this could not be seen in the cryo-EM maps [61]. The cryo-EM data suggest that the Hrd1-FAM8A1 interaction is not particularly stable in LMNG and thus is maybe averaged out during cryo-EM processing. The bulk of the cytoplasmic region of Hrd1 was not observed in previous cryo-EM maps and the same was true for this map. However, a complex of Hrd1-FAM8A1 served as a perquisite for building other complexes. Herp was stably isolated with this complex to produce Hrd1-Herp-FAM8A1, and the components of this complex appeared stochiometric with each other. The literature suggests that this complex could exist as a dimer but this is unknown from our data [61]. DERL2 was then assembled with this complex through further overexpression. This required that only Hrd1, Herp and DERL2 be overexpressed since the additional overexpression burden of FAM8A1 expression reduced total expression. Despite this, endogenous FAM8A1 was isolated in the complex to produce Hrd1-Herp-FAM8A1-DERL2, a complex proposed to exist from other studies of Hrd1 [61,62]. DERL2 appeared less abundant in this complex than the other proteins but this could be attributed to poor Coomassie staining. This complex was reconstituted in both LMNG and GDN but only the LMNG condition was subject to cryo-EM analysis at the time. This failed to yield high resolution structural information but some 2D class averages had secondary structure detail within the detergent micelle. There appeared to be two different types of particles present, some

with an S-shaped micelle and others with a flatter micelle. The lack of resolution was attributed to the size of the complex since this complex is relatively small for cryo-EM and the bulk of the cryo-EM signal was suspected to be within the detergent micelle. To further increase the size of the complex, Sel1L was reconstituted with it. Attempts to reconstitute Sel1L alone with Hrd1 failed to yield good cryo-EM data but Sel1L was shown to express better and more stably associate with Hrd1 when expressed as a full-length protein and reconstituted using milder detergents. Sel1L was then overexpressed with Hrd1, Herp and DERL2 and this led to reconstitution of a 5-component complex also containing FAM8A1. The LMNG reconstituted complex was subject to cryo-EM and this had improved resolution in the 2D class averages. This sample still failed to achieve high resolution but low-resolution maps were generated. This sample also contained two different types of particles with some micelles resembling an S-shape whilst others were flatter. These two different states are akin to the yeast Hrd1-Der1 complex where in the cryo-EM processing, about ~50% of the complex had Der1 in a flipped orientation [55]. This is likely a reconstitution artifact, perhaps because the Hrd1-Der1 interface is not particularly stable. It is difficult to assess if this is also true for the mammalian Hrd1-DERL2 complex given the low-resolution of the maps. Interestingly, the angle of the TMDs relative to the detergent micelle in the flatter micelle particles was reminiscent of that in the Hrd1 open dimer. The significance of this can only be understood with higher resolution structural information. Cryo-EM density for Sel1L was not immediately obvious in any of the maps or class averages produced from this sample, despite it containing a series of large SLR repeats, suspected to be in close proximity to the micelle [55,178]. Either these domains are flexibly linked to the detergent micelle and so are averaged out during processing, or they undergo denaturation during cryo-EM grid preparation. The addition of a single TMD to the micelle could have been enough to improve the signal in that region for alignment during cryo-EM processing, resulting in the observed improvement in resolution with this sample. Reconstitution of this complex in the milder

detergent GDN resulted in better overall biochemical behaviour, and Sel1L was more abundant in the complex. However, this reduced the quality of the cryo-EM data and so was not an ideal modification alone for cryo-EM structural studies. The data suggests that it was worth optimising the 4-component and 5-component complexes for further structural studies since these could robustly be reconstituted and yielded the best cryo-EM data of all the complexes studied. Lack of signal for alignment during cryo-EM processing may still be inhibitory and so development of small binders was chosen as a tool to overcome this, more of which is covered in the next chapter. Additionally, a complex containing Hrd1, Herp, FAM8A1 and DERL1 was also isolated, in a similar manner to the DERL2-containing complexes. This implied that DERL1 assembles with Hrd1 similarly to DERL2 and may represent a conserved interaction mode of Derlins. Attempts were also made to reconstitute an artificial substrate with Hrd1 but these require more optimisation before cryo-EM analysis. Structural analysis of Hrd1-substrate conjugates would provide valuable insight into recognition and retrotranslocation of ERAD substrates by E3 ligase complexes.

6. Generation of high-affinity binders to Hrd1 complexes

6.1. Nanobodies as tools in cryo-EM

Antigen binding fragments (Fabs) have been widely used in membrane protein structure determination by X-ray crystallography [196]. Fabs promote formation of stable crystal contacts, thereby increasing the chance of obtaining higher quality diffraction data from a protein crystal [197]. These have significant advantages over full immunoglobulins, such as immunoglobulin G (IgG), since full immunoglobulins have flexible hinge regions between the Fab and crystallisable fragment (Fc) domains [198]. Immunoglobulins are also glycosylated, further adding to structural heterogeneity [199]. Furthermore, full immunoglobulins are difficult to produce recombinantly and so are not an attractive starting point for use in structural biology [198]. Fabs represent the smallest unit for antigen binding, are more rigid and generally have less glycosylation than the Fc domain, making them better binders for use in structure determination techniques [198,200]. Over the years, this technology has been transferred to cryo-EM where Fabs have been used to compensate for the size of protein particles, allowing cryo-EM structure determination of proteins originally too small to visualise [128]. Small proteins visualised by cryo-EM often lack recognizable shape features that facilitate initial image alignment during computational processing [201]. By attachment of a high-affinity binder, this increases the size and mass of the particle, thereby aiding in particle alignment [202]. Since then, many structures have been solved by cryo-EM, where proteins are in complex with Fabs or other single chain variable fragments (scFv) [128].

Nanobodies are an emerging tool that has become increasingly popular in X-ray crystallography and cryo-EM [129]. These are derived from a type of immunoglobulin from camelids that has only one heavy chain made up of three immunoglobulin domains (figure 6.1) [203]. Nanobodies have properties which make them more advantageous as structure determination aids. Their variable region is made up of a single domain (termed

VHH), that confers all the binding activity [204]. This makes it easier to clone and express VHHs in bacteria, as well as in vectors for phage display [204]. These nanobodies are unique in that their complementarity determining region 3 (CDR3) is longer than the classical CDR3 and thus is able to bind a greater diversity of surfaces, sometimes inserting deeply into the active site of enzymes [205]. Given their small size (~15 kDa), they can bind small regions not necessarily accessible by Fab fragments, making them highly versatile tools [205]. After immunisation of camelids, large *in vitro* libraries are created for display by phage, yeast cells or on ribosomes [205,206]. However, nanobodies are not without some drawbacks. The classical approach for generation of nanobodies involves first immunising a camelid with the target antigen [207]. Blood is then collected, lymphocytes isolated, mRNA extracted and then a VHH library generated for use in display methods to enrich for antigen binders [207]. The antigen must therefore be stable and remain folded after mixing with an adjuvant prior to injection, and remain stable inside of the camelid, whose body temperatures are slightly higher (> 38°C) than other mammals [208]. This presents a challenge for membrane proteins reconstituted in detergent, which are likely to fall apart during this process due to the lack of detergent present [205]. Instead, membrane proteins are often reconstituted into detergent-free mimetics such as liposomes or amphipols, which may not always be possible [120,209,210]. Targeting specific conformations can also be difficult as ligands or other stabilising reagents are likely to dissociate unless they bind very tightly [205]. Furthermore, access to animals is not always guaranteed.

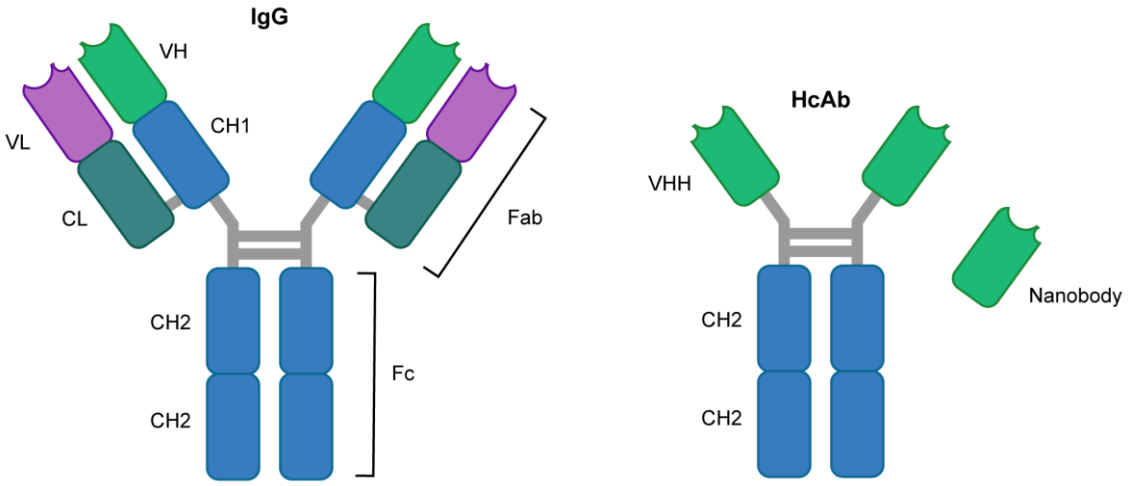


Figure 6.1. Comparison of IgG with the heavy chain antibody (HcAb) from camelids. The antigen binding region of IgG is comprised of a heavy chain (VH) and a light chain (VL) [211]. HcAbs consist of just two heavy chains with a single variable domain (termed VHH) [212]. Nanobodies are comprised of only the VHH domain and retain the full antigen-binding potential of a HcAb [212].

6.1.1. Synthetic nanobodies

Alternatively, nanobodies can be generated by an entirely *in vitro* selection process [205]. This process overcomes some of the difficulties of immunisation whilst enabling more control throughout. Zimmerman et. al developed three synthetic nanobody (sybody) libraries based on the observed binding modes of nanobodies (figure 6.2) [205]. Each library features a distinct antigen-binding surface that is predominantly driven by the length and arrangement of CDR3 [205]. Amino acid residues within the CDRs of these three sybody libraries were randomised with an optimal balance of charged, polar, aromatic and non-polar residues to maintain a moderate hydrophobicity of the binding surface [205]. Reducing the hydrophobicity of the binding surface was previously shown to reduce the number of “sticky”, false-positive binders for membrane proteins [213]. These libraries were used with a variety of display methods to maximise library size and

diversity during antigen presentation [205]. Large diversity is required in the initial display stages since there is no equivalent affinity maturation mechanism equivalent to the camelid immune system [205,214]. The process of sybody generation involves one round of ribosome display, followed by two rounds of phage display. During the second round of phage display, weaker binders with fast off-rates are competed off by addition of free (non-biotinylated) target protein [123]. Finally, an enzyme-linked immunosorbent assay (ELISA) is used to identify binders. Binders are then assayed further by pulldown and co-elution by SEC [123]. This process can theoretically be performed within a single month, does not require camelids, and allows ligands and buffers to be controlled to some degree at all stages. This is advantageous for membrane proteins since they can be reconstituted in detergent.

Since complexes containing mammalian Hrd1 failed to yield high-resolution cryo-EM maps, we sought to increase the mass, stability and homogeneity of the complexes by addition of a high-affinity binder. This would improve the quality of particle alignments during image processing and facilitate high-resolution structure determination. We opted to generate nanobodies by camelid immunisation and synthetically.

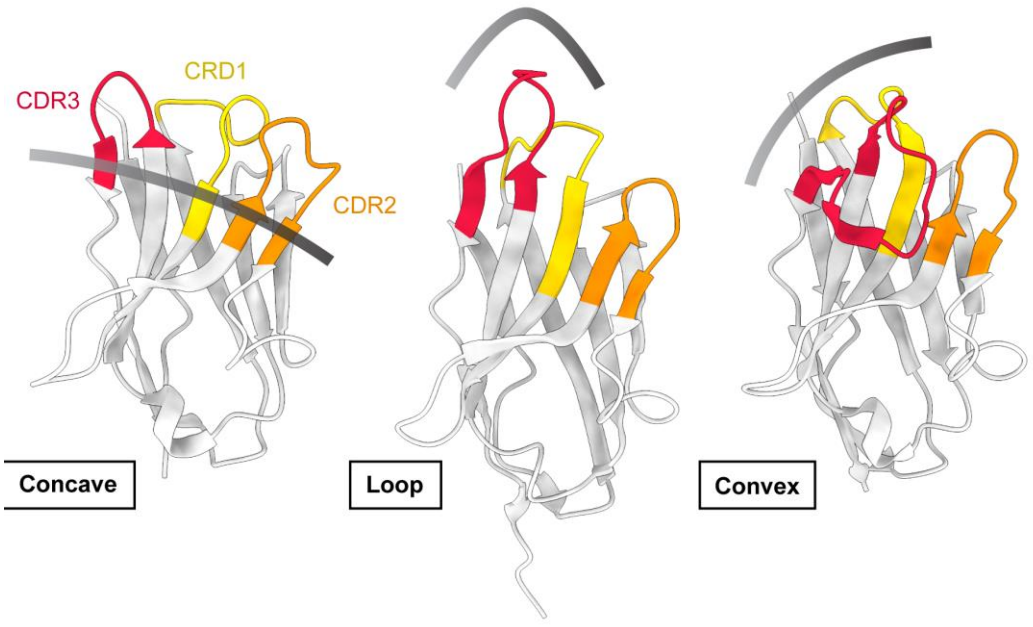


Figure 6.2. Models for binders from the three synthetic nanobody libraries. The concave, loop and convex sybody models were taken from previously solved structures of protein-nanobody complexes (concave PDB: 3K1K; loop PDB: 3P0G; convex PDB: 1ZVH) [210,215,216]. CDR domains 1, 2 and 3 are coloured yellow, orange and red respectively. The shape of the binding surface is denotated by the black line. Figure adapted from Zimmerman et al. [205].

6.2. Generation of nanobodies from llamas

To immunise llamas, a Hrd1 complex had to be reconstituted into liposomes to improve stability in the host organism [205]. This would increase the chance of successfully generating nanobodies that bind to an assembled Hrd1 complex. Ray Owens and his team at The Oxford Protein Production Facility, UK, carried out immunisation and initial panning to identify nanobody binders. A biotinylated complex was also produced as panning took place on NeutrAvidin-coated plates [217]. The panning took place in detergent and did not require liposome reconstitution. Therefore, two reagents needed to be established: a Hrd1-containing complex reconstituted into liposomes and the equivalent biotinylated complex in detergent.

6.2.1. Liposome reconstitution of Hrd1 complexes

Liposome reconstitutions were first trialled with Hrd1 alone. During reconstitution, membrane proteins are typically extracted in detergent to solubilise them before exchange into liposomes [218]. Detergent is then removed by dialysis, detergent spin columns or by addition of nonpolar polystyrene beads to adsorb the detergent [218]. In the literature, liposome reconstitutions are usually done from detergents with high off-rates (such as DDM), presumably because exchange with lipids is faster and therefore efficiency is higher [219,220]. At the time, there was only one published protocol where

a membrane protein extracted in LMNG was exchanged into a liposome [221]. Since Hrd1 complexes containing additional binding partners were more stable in milder detergents, liposome reconstitution would need to be eventually performed after LMNG extraction. Therefore, trialling reconstitution with Hrd1 alone would serve as a good starting point.

To that end, Hrd1-mVenus-TwinStrep was purified in DDM-CHS and LMNG-CHS (figure 6.3). During this purification, the tags were not cleaved in order to preserve protein fluorescence. mVenus fluorescence could then be used to follow Hrd1 during liposome reconstitution. The SEC profiles for the final purified proteins are shown in figure 6.3 and both behaved as expected from previous purifications and the Hrd1 detergent screening. The material in the SEC peak fraction for each of these conditions was concentrated and frozen prior to liposome exchange. Material was reconstituted into liposomes as previously described for camelid immunisation [120]. To assay liposome reconstitution, liposomes were floated away from non-reconstituted protein on a Nycodenz gradient (figure 6.4) [93]. Liposomes floated to the 10% Nycodenz density region and these were fluorescent upon exposure to blue light (figure 6.4a). Minimal fluorescent precipitated protein was seen at the bottom of the centrifuge tube, indicating good reconstitution. Fractions from the floatation gradient were assessed for Hrd1 content by SDS-PAGE and western blotting (figure 6.4b). Hrd1-mVenus was clearly seen in the 10% fraction and the same band was fluorescent. Western blotting also confirmed the presence of Hrd1-mVenus, as well some higher MW species. One of these species corresponded to dimeric Hrd1 and suggests that recombinant Hrd1 may exist as a dimer in a lipid environment. The reconstitution behaved the same for the DDM and LMNG conditions and so this indicated that LMNG-extracted Hrd1 complexes could be reconstituted into liposomes.

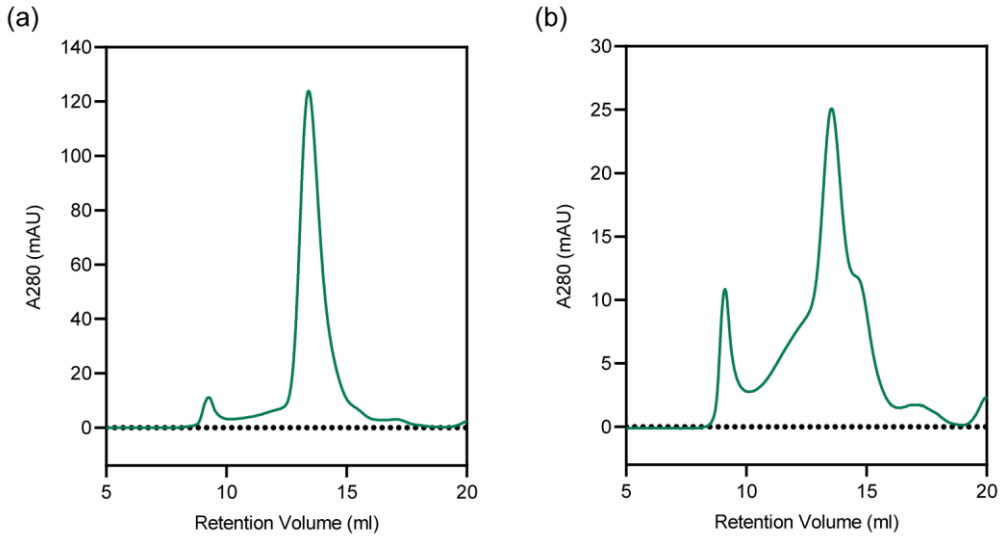


Figure 6.3. Purification of Hrd1-mVenus in DDM-CHS and LMNG-CHS. (a) and (b) are SEC profiles of Hrd1 after StrepTactin purification in DDM-CHS and LMNG-CHS, respectively.

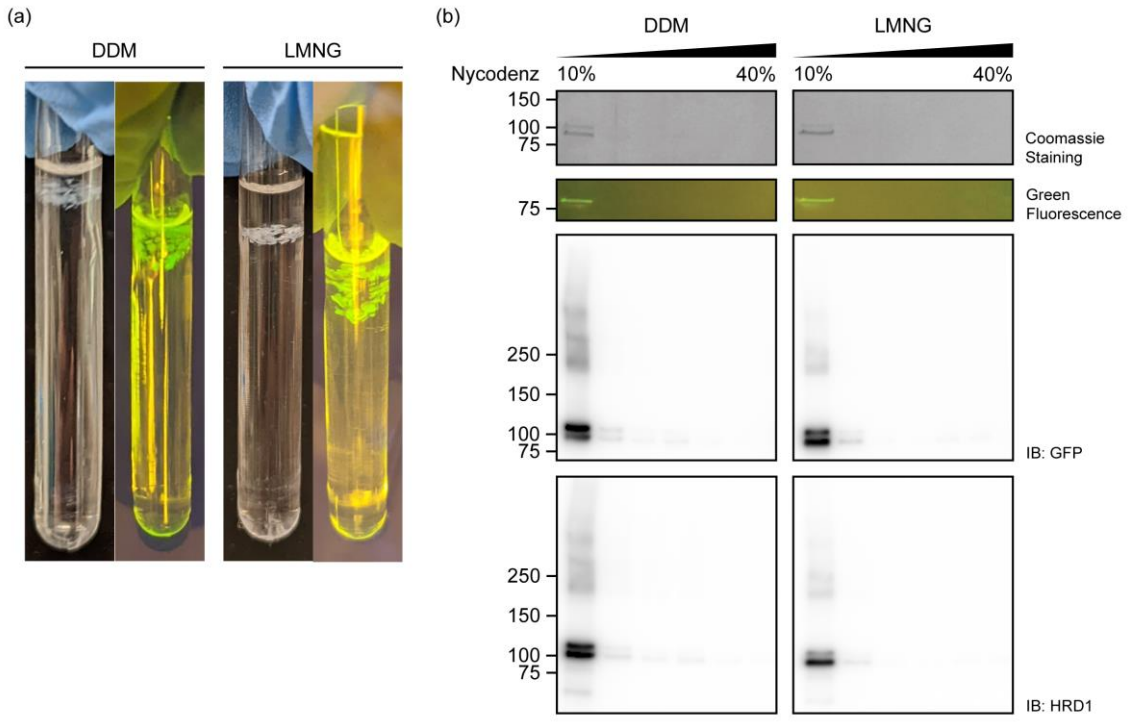


Figure 6.4. Floatation assay for Hrd1 proteoliposomes. (a) Visual inspection of liposomes after floatation in a Nycodenz gradient. Liposomes (white) floated to the top of the gradient. Exposure to blue light shows fluorescent protein in the liposome fraction.

(b) SDS-PAGE analysis of the Nycodenz gradient fractions. Proteins were imaged by Coomassie staining, in-gel fluorescence and western blot.

The liposome-containing fraction from the floatation assay was also subject to a protease protection assay to determine the insertion orientation of Hrd1 (figure 6.5). Since the TEV-mVenus-TwinStrep tag was still present, TEV protease could be added to remove the tag. Cleavage would only occur if the C-terminus was outside of the liposomes and so this was used to determine the abundance of different orientations of Hrd1. Detergent addition during the TEV cleavage reactions acts as a solubilisation control and enables accessibility of TEV to both sides of the liposomes, thereby enabling full cleavage. Addition of TEV without detergent resulted in near complete cleavage of Hrd1 (figure 6.5). This indicated that Hrd1 was almost entirely incorporated with its C-terminus on the outside of the liposome, as was also observed for yeast Hrd1 in a similar assay [93]. Some of this orientation bias is likely driven by the large, soluble C-terminus of which requires more energy to incorporate across partially solubilised liposomes membrane during reconstitution [222]. Additional freeze/thaw cycles have been shown to improve orientation bias with ABC transporters but in the case of Hrd1-mVenus, this did not help [223]. A mixed orientation population would be advantageous during camelid immunisation since both sides would be presented to the host immune system. Interestingly, there were also higher MW Hrd1 species observed during this experiment, some of which correspond to dimeric Hrd1. This provides more evidence that Hrd1 exists as a dimer in lipid environments. In both DDM and LMNG conditions, the TEV cleavage patterns were the same and so the proteoliposomes were likely of similar quality.

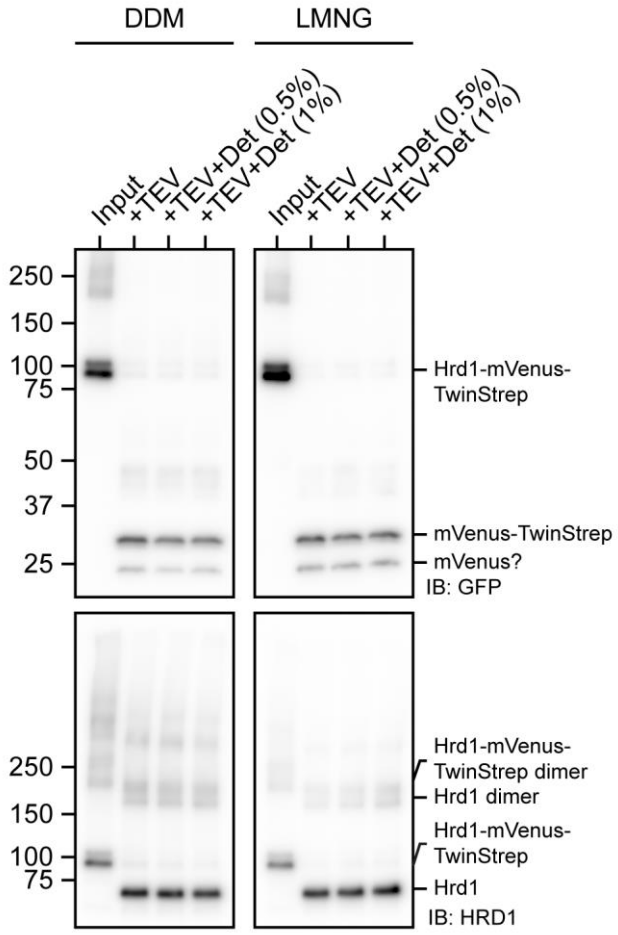


Figure 6.5. TEV protease protection assay for proteoliposomes containing Hrd1.

Since the Hrd1 liposome reconstitution was successful after LMNG extraction, a complex containing Hrd1, Herp, FAM8A1 and DERL2 was reconstituted into liposomes (figure 6.6). For immunisation of llamas, ~600 µg of protein was required and so large quantities of cells expressing Hrd1-mVenus-TwinStrep, Herp and DERL2-6xHis were grown. These were purified as before in the previous chapter and the final SEC profile indicated that the complex was as homogenous as seen previously (figure 6.6a). This was reconstituted into the same liposome mix as Hrd1 alone and a small amount floated on a Nycodenz gradient. The 10% fraction contained all 4 proteins but only small amounts of DERL2 (figure 6.6b). This behaviour of DERL2 was also seen in previous Coomassie stained gels and thought to be a result of weak Coomassie staining by

DERL2. Llamas were immunised with these proteoliposomes at the Oxford Protein Production Facility, UK. The material was mixed with an adjuvant prior to two separate immunisation rounds, separated by 30 days. Approximately four other antigens were immunised alongside this sample and this is a common practice to best utilise the immunisation animal [207].

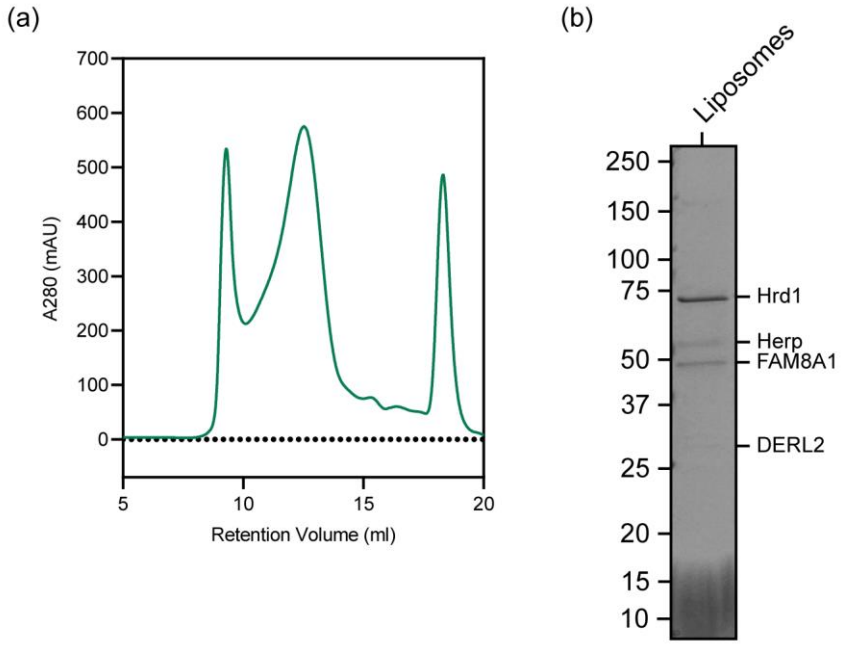


Figure 6.6. Large-scale purification of the Hrd1-Herp-FAM8A1-DERL2 complex and its reconstitution into liposomes. (a) Final SEC profile after dual affinity purification. (b) Proteoliposome fraction containing Hrd1-Herp-FAM8A1-DERL2 after floatation on a Nycodenz gradient. Tags were cleaved prior to reconstitution.

6.2.2. Generating reagents for panning

To identify binders after immunisation, display methods are used with the cloned VHH library [207]. This took place on NeutrAvidin-coated plates and so the antigen had to be biotinylated for binding to the plate. An Avi-tagged version of the Hrd1 complex used for immunisation had to be generated. The AviTag is a 15 amino acid peptide that is

biotinylated by the *E. coli* biotin ligase (BirA) and is used to covalently link biotin to proteins [224]. Since generating an Avi-tagged Hrd1 complex similar to the immunisation complex would take more time, Avi-tagged Hrd1 and DERL2 were initially created. Nanobody hits would theoretically be generated to many surfaces of the Hrd1-Herp-FAM8A1-DERL2 complex and so it seemed reasonable that some binders could be pulled out with the individual components alone. Hrd1 is the largest component in this complex and contributes the largest soluble domain, therefore some bias in hits could be expected towards Hrd1 [205]. Avi-tagging of Hrd1 and DERL2 would also identify which proteins tolerate Avi-tagging and would aid in decision making for tagging of the Hrd1-Herp-FAM8A1-DERL2 complex. Hrd1-Avi and DERL2-Avi constructs were cloned and cell lines generated for their expression and purification, using a C-terminal TwinStrep tag. Both Hrd1 and DERL2 were purified in DDM-CHS using a TwinStrep purification and SEC (figure 6.7a-b). Hrd1 behaved as previously whilst a significant portion of DERL2 was in the void volume. Material from the peak fractions of Hrd1 and DERL2 at 13.5 and 15 ml, respectively, were taken and concentrated for biotinylation overnight by BirA. Biotinylated protein was re-run by SEC to separate excess biotin, ATP and BirA (figure 6.7a and 6.7b). There was a significant yield decrease after biotinylation, likely due to protein adsorption to the membranes of the protein concentrators. Despite this, there were good yields for Hrd1 and this remained homogenous. The behaviour of DERL2 was somewhat improved and the yields were high enough for panning against the VHH library. To assay the extent of biotinylation and completeness of the reaction, both Hrd1 and DERL2 were bound to avidin and analysed by SDS-PAGE (figure 6.7c) [224]. The avidin-biotin interaction is strong and if the sample is not heated prior to SDS-PAGE, size shifts can be observed due to avidin-target interactions [224]. The commercial avidin was significantly aggregated and did not migrate correctly by SDS-PAGE but shifts in both Hrd1 and DERL2 were observed in the presence of avidin. The majority of both proteins increased in MW, suggesting that they were close to 100% biotinylated.

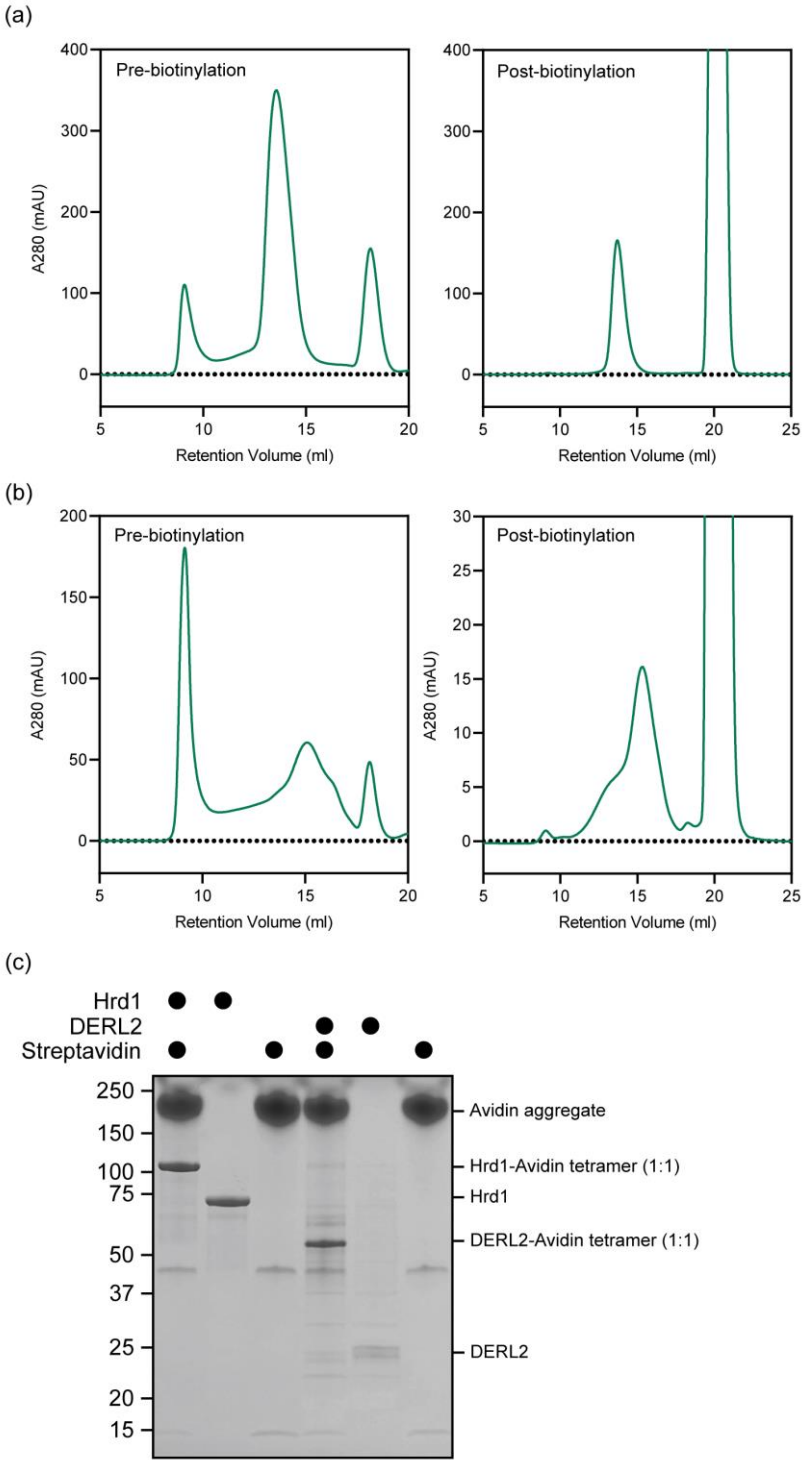


Figure 6.7. Purification and biotinylation of Hrd1 and DERL2 for nanobody panning. (a) and (b) are SEC profiles of Hrd1-Avi and DERL2-Avi in DDM-CHS before and after biotinylation, respectively. (c) Avidin-shift assay to assay the biotinylation level of Hrd1

and DERL2. The commercial avidin was very aggregated but size shifts were still observed for Hrd1 and DERL2.

With the knowledge that both Hrd1-Avi and DERL2-Avi could be purified, attempts were made to isolate an Avi-tagged complex containing Hrd1, Herp, FAM8A1 and DERL2. Cell lines for overexpression of this complex were established in a similar way as before and initially Hrd1-Avi-mVenus-TwinStrep, Herp and DERL2-6xHis were overexpressed. These cells were grown to the same scale as previous Hrd1 complexes and the complex isolated by dual affinity purification in LMNG-CHS (figure 6.8a). One minor difference with this cell line is that the quantity of endogenous FAM8A1 in the purified complex was much lower than previously observed. The final yield by SEC was lower than with the equivalent non-AviTag cell line. This became especially apparent when complex from the peak fraction was concentrated, biotinylated and subject to a second round of SEC (figure 6.8b). Given that it was not time or cost effective to operate at even larger scales of cell mass, alternative cell lines and approaches were explored. Next, a cell line was generated that overexpressed Hrd1-mVenus-TwinStrep, Herp and DERL2-Avi-6xHis. This was overexpressed and purified in the same conditions as the above cell line. This resulted in poor overexpression of all proteins and final yields were significantly reduced (figure 6.8c and 6.8d). Both these attempts were not the way forward and tagging of DERL2 with a larger C-terminal tag affected expression significantly. Although Hrd1 alone tolerated an AviTag, it seemed that the cell line from the previous chapter that expressed Hrd1, Herp and DERL2 was somewhat serendipitous in that it expressed better with improved stoichiometry for the complex components.

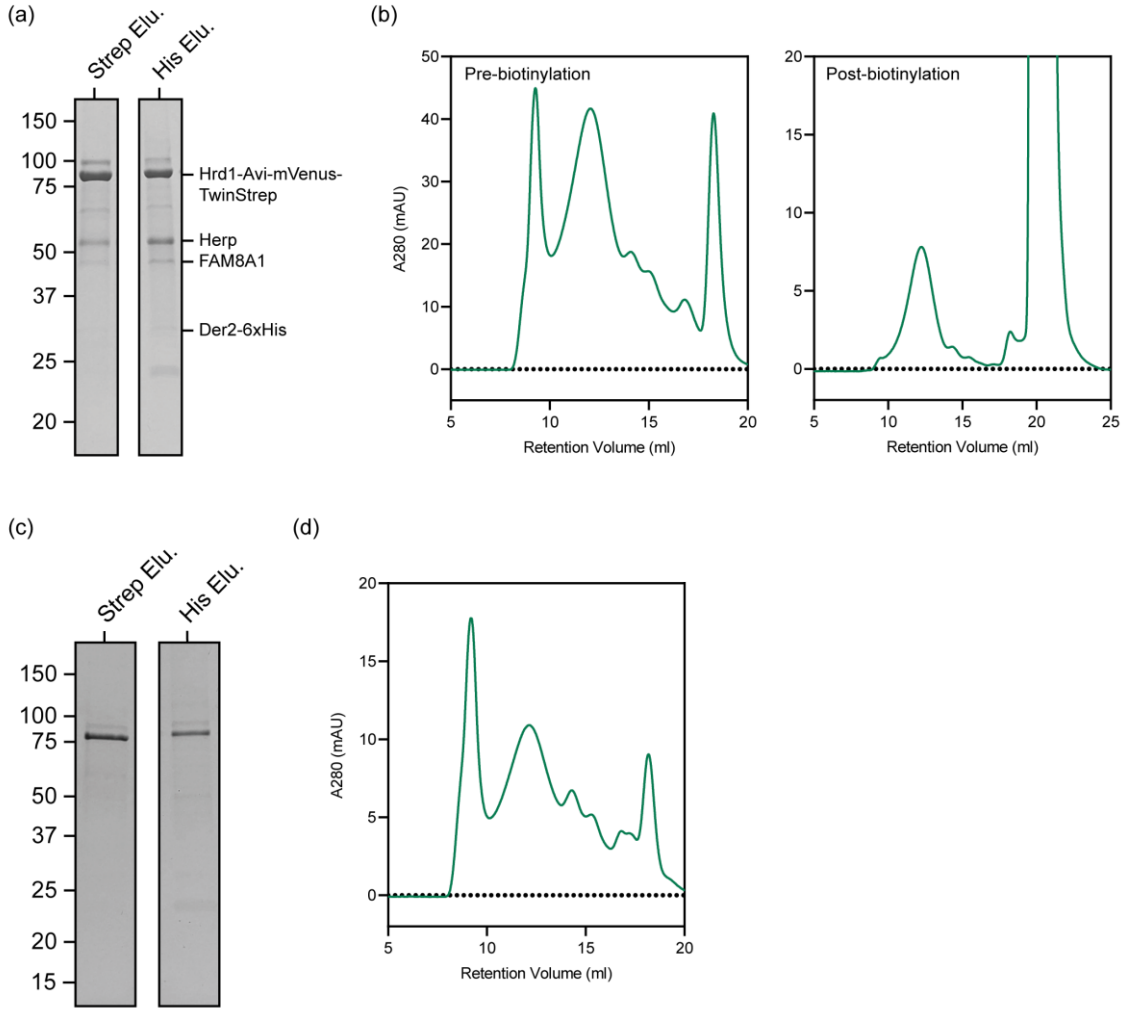


Figure 6.8. Attempts to generate a biotinylated complex containing Hrd1, Herp, FAM8A1 and DERL2. (a) StrepTactin and Ni-NTA (His) purifications from cells overexpressing Hrd1-Avi, Herp and DERL2. (b) Corresponding SEC profiles before and after biotinylation of the purified complex. (c) StrepTactin and Ni-NTA (His) purifications from cells overexpressing Hrd1, Herp and DERL2-Avi. (d) Corresponding SEC trace before biotinylation. This was not biotinylated due to low yields after the initial SEC.

We reasoned that the next best alternative would be a biotinylated complex that contained at least high levels of Hrd1, Herp and FAM8A1. A cell line was generated that overexpressed Hrd1-Avi-mVenus-TwinStrep, Herp and FAM8A1-6xHis. Dual affinity purification was performed which resulted in a complex containing Hrd1, Herp, FAM8A1

and surprisingly, high levels of endogenous DERL2 (figure 6.9a). To reduce losses during the biotinylation step, biotinylation was performed prior to SEC after dialysis and tag cleavage. This also limited exposure to protein concentrators. The resulting SEC profile had higher final yields for this 4-component complex and the main peak was more homogenous (figure 6.9b). The presence of biotin was confirmed by western blot and only Hrd1 was reactive (figure 6.9c).

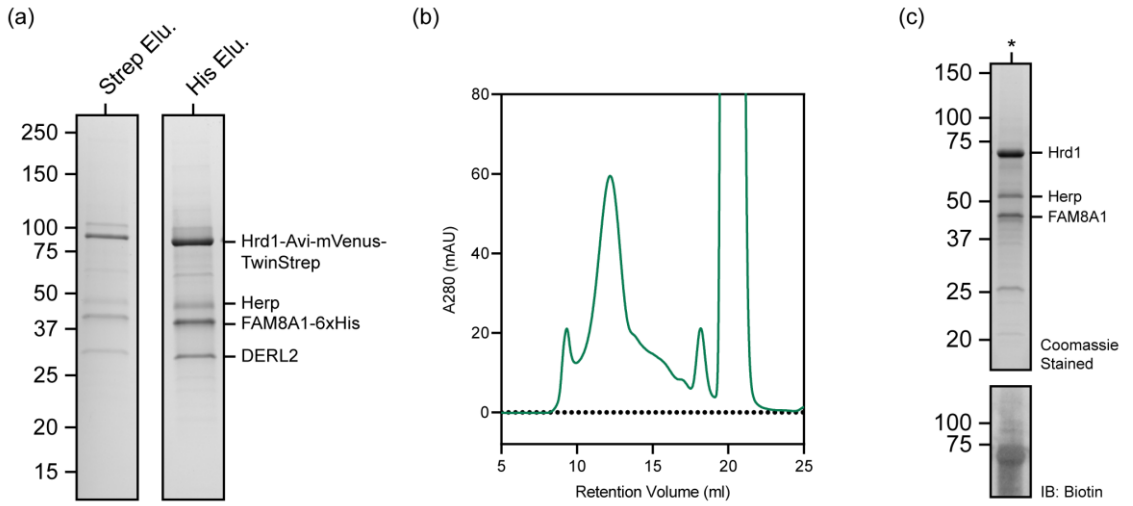


Figure 6.9. Optimised cell line and protocol for production of a biotinylated Hrd1-Herp-FAM8A1-DERL2 complex. (a) Dual affinity purifications for isolation of the complex. (b) SEC profile after biotinylation of the purified complex. (c) SDS-PAGE of the peak fraction from the SEC profile. Protein was visualised by Coomassie staining and biotinylation confirmed by western blot with Streptavidin-HRP.

6.2.3. Panning and assessment of hits after llama immunisation

Immunisation and panning for nanobody binders to the Hrd1-Herp-FAM8A1-DERL2 complex was carried out by Ray Owens and his team. Potential binders were identified after a round of nanobody display from the VHH library, followed by an ELISA screen. 10 unique CDR3 clusters were enriched in the nanobody hits, suggesting some level of

enrichment from the VHH library [123]. However, none of these hits bound to the target complex in pulldowns. The pre-post-immune serum was also tested for immunoreactivity against the complex but there was no seroconversion. Together, the data imply that none of the VHHs isolated had a high affinity for the Hrd1 complex and either they were artefacts of the ELISA or just very weak binders. It is possible that upon immunisation, the Hrd1 proteoliposomes were not as stable as expected and so no nanobodies were generated against the folded/assembled complex. Another possibility is that other antigens immunised alongside the Hrd1 complex were more immunogenic and so a reduced antibody response was generated against the Hrd1 complex [207]. Given the amount of time it would take to repeat immunisation, perhaps with a different complex, as well as the potential complications, we opted to take a different approach to generate high affinity binders.

6.3. Generating sybodies

6.3.1. Sybody strategy

Generation of sybodies is faster and more controlled than of native nanobodies from immunised camelids [205]. This would allow us to precisely control the detergent conditions and guarantee the quality of the antigen throughout the process. To generate sybodies, biotinylated and non-biotinylated target protein are required, similar to the nanobody generation process. Since the complex containing Hrd1, Herp, FAM8A1 and DERL2 was time consuming and costly to produce, we opted for screening with Hrd1 alone first. Anna Valkov and her team at the Protein Technology Core, Center for Structural Biology, NCI, had begun to establish the sybody generation protocol in-house [123]. This technology had proved successful in generating high affinity sybodies for other soluble targets within the department and so Hrd1 was to serve as the first membrane protein antigen.

Full-length Hrd1 contains significant disorder in its C-terminal region, most of which was not observed in our previous cryo-EM maps. A Hrd1 truncation was generated ($Hrd1_{1-344}$) that removed the amino acids unobserved in those maps. Sybodies that bound to visible density in the maps would be more advantageous for structure determination of larger complexes since these would likely be visible. Sybodies that bound to the flexible C-terminus would not necessarily be visible after cryo-EM processing since the signal would be averaged out [127]. Therefore, a cell line for expression of Avi-tagged $Hrd1_{1-344}$ was generated (figure 6.10). This purified as a much smaller object than full-length Hrd1, but it expressed well and large amounts of protein were biotinylated. The shift in retention volume indicates that the cytoplasmic tail contributes significantly to the hydrodynamic radius of detergent-extracted Hrd1.

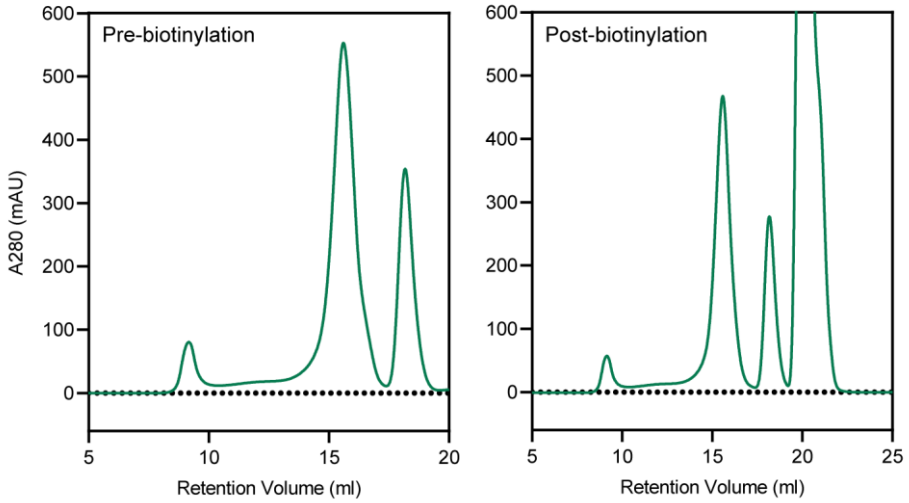


Figure 6.10. Purification and biotinylation of $Hrd1_{1-344}$-Avi. SEC profiles for StrepTactin purified $Hrd1_{1-344}$-Avi before and after biotinylation.

6.3.2. Initial screening of sybody library against $Hrd1_{1-344}$-Avi

Anna Valkov and her team performed the initial panning of the sybody libraries against $Hrd1_{1-344}$-Avi. Throughout the panning, enrichment was observed by qPCR at

each stage [123]. Binders were then identified by ELISA and these were handed over to perform binding at a larger scale. Most of the enrichment occurred in the concave (green) and loop (blue) libraries and a total of 20 sybody hits were initially screened. Small-scale *E. coli* cultures were grown and sybodies purified from each of them by Ni-NTA purification (figure 6.11). All 20 sybodies were extracted from the periplasm and this resulted in clean purifications. Some expressed better than others with C1 green expressing the worst. The sybodies exhibit slightly different MWs likely because there CDRs are of different lengths.

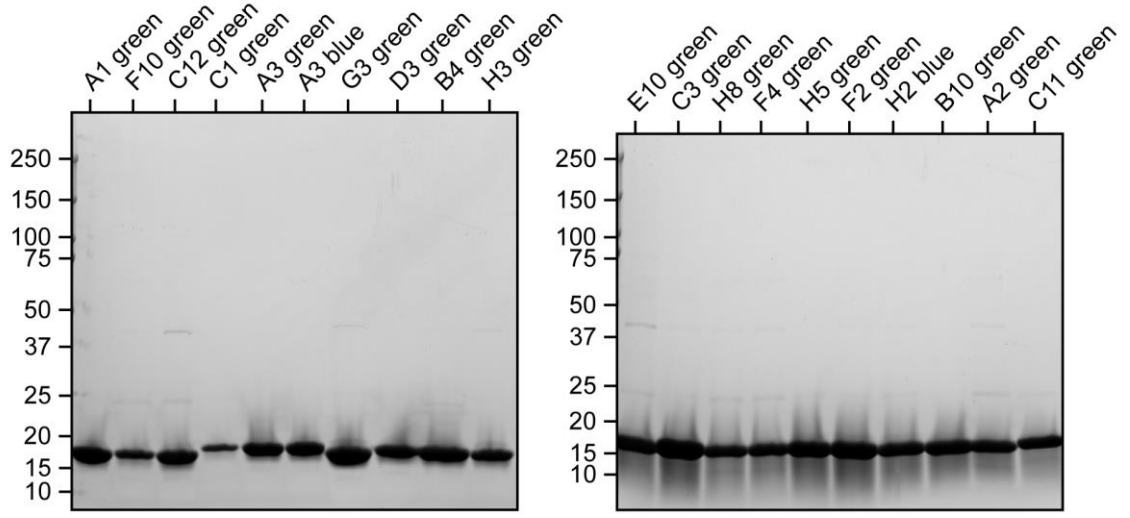


Figure 6.11. Ni-NTA purifications of 20 sybody binders identified from screening of the sybody libraries against $Hrd1_{1-344}$.

All purified sybodies were then subject to SEC, to exclude those that had poor biochemical behaviour (figure 6.12). Sybodies targeting small membrane proteins are usually enriched in hydrophobic residues within their CDRs, which can result in poor biochemical behaviour such as low expression or aggregation [123]. These are excluded from downstream analysis. A significant proportion of sybodies are also expected to interact with Superdex and Superose SEC resins, resulting in delayed elution profiles

[123]. This does not correlate with the overall quality of the binder and can be overcome somewhat by use of a Sepax column, which contains a silica-based SEC resin [123]. A diverse range of behaviour was observed by SEC for the sybody hits (figure 6.12). From these libraries, monomeric sybodies can elute anywhere after 10 ml and so a range of hits met this criterion. A1 green, A4 green, B10 green, C3 green, C12 green, F2 green, H2 green and H5 green exhibited the best biochemical behaviour as their peaks were symmetrical and they eluted over a narrow range. Some of the sybodies like A3 blue, although homogenous, eluted over a large range, indicative of interactions with the column. Those hits were excluded since binding reactions were to take place on Superose resin, given the size of the Hrd1 truncation in a detergent micelle. Other hits, such as A2 green, had poor expression and were excluded. The peak fractions of purified sybodies that behaved well were isolated and concentrated in a protein concentrator prior to binding with $Hrd1_{1-344}$.

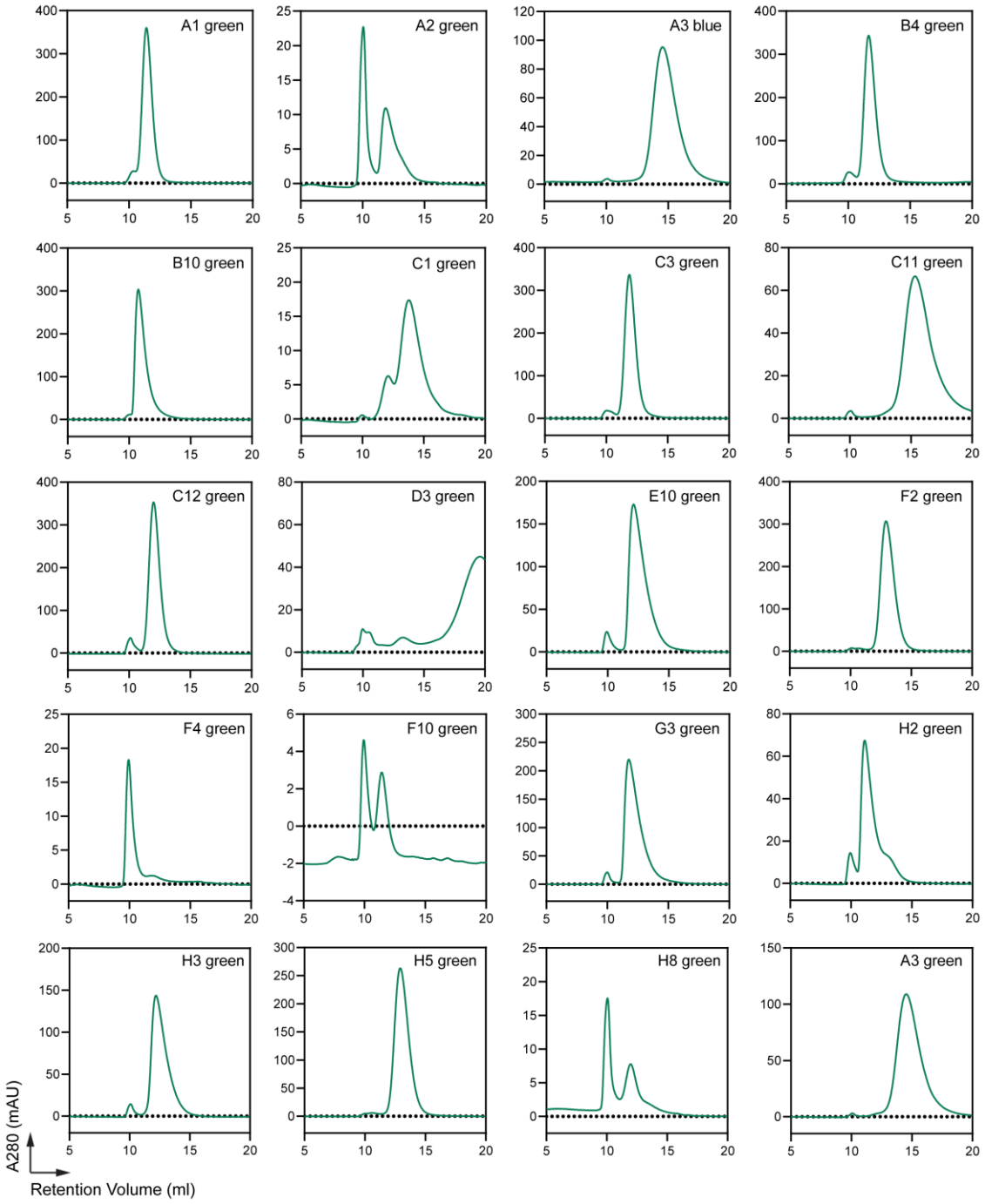


Figure 6.12. SEC profiles for the 20 purified sybody hits. All SEC was performed using a Sepax SRT-10C SEC column at the same flow rate.

A1 green, B4 green, B10 green, C3 green, C12 green, F2 green, H2 green and H5 green were selected to perform pulldown and SEC co-elution assays with Hrd1 (figure

6.13). Pulldowns were performed by incubating $Hrd1_{1-344}$ with sybodies bound to Ni-NTA resin. Small amounts of $Hrd1_{1-344}$ co-eluted with C3 green, H2 green, A1 green and C12 green (figure 6.13a). A1 green bound to the largest quantity of $Hrd1_{1-344}$ but the majority was still in the unbound fraction. Stronger binding was previously observed with other soluble targets by Anna Valkov and her group, but not with this sample. During the pulldown, high volume wash steps are performed and the final wash included a long incubation to remove weaker binders. It is possible that most of the hits were only weak binders and so were not observed to bind by pulldown. The ones that did bind were subject to SEC, but only the stronger binding A1 green and C12 green are shown (figure 6.13b). No co-elution was observed by SEC for any of the binders. This is evidenced by the lack of shift of the Hrd1 peak and the fact that all peaks in the binding reaction share the same areas as the Hrd1 and sybody control peaks (figure 6.13b). The same was true for C3 green and H2 green. Other binders have since been tested but these also suffered from the same problems. Due to time limitations, an extensive analysis of other binders has not been performed and no strong binders have been identified. Our current reasoning is that a lot of the binders identified from the ELISA are non-specific hits and so attempts are being made to optimise the protocol for sybody generation.

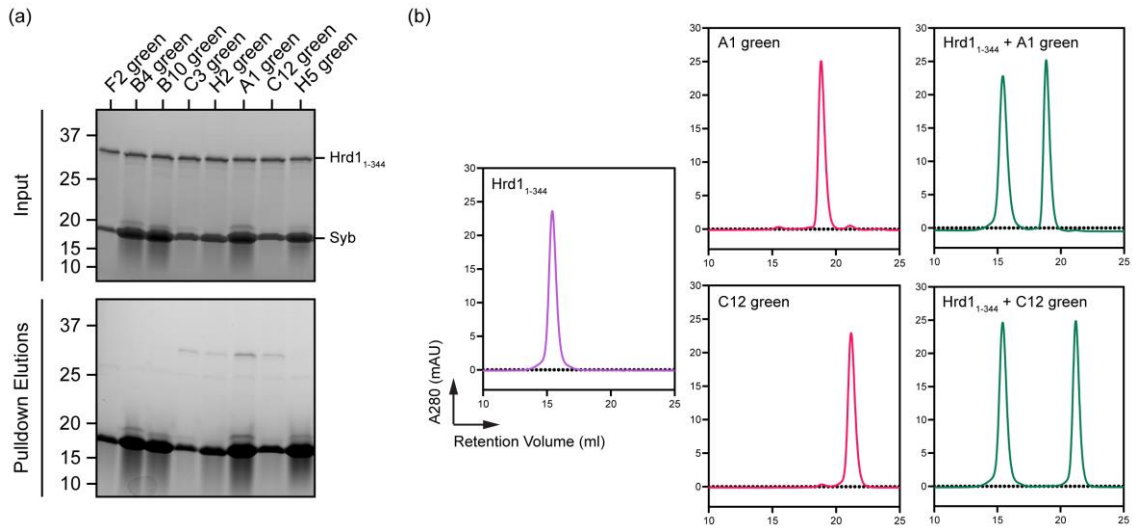


Figure 6.13. Binding analysis of sybodies with good biochemical behaviour. (a) Ni-NTA pulldowns with sybodies in the presence of $Hrd1_{1-344}$. (b) SEC profiles of the strongest sybody binders from the pulldowns. No binding was observed by SEC.

6.4. Discussion

In this chapter, attempts were made to generate high affinity binders to different Hrd1 complexes, with the eventual aim to use these to aid cryo-EM studies. Nanobodies were initially sought after from llama immunisation and this required a Hrd1 complex to be reconstituted into liposomes. Hrd1 alone was initially trialled in liposome reconstitutions and this showed that Hrd1 likely also existed as a dimer in a lipid environment. This correlated with previous structural and biochemical data in this thesis, as well as other studies of yeast and mammalian Hrd1 [55,60,84]. To screen for nanobody binders against Hrd1 and its partners, a biotinylated complex was required. This was successfully done for Hrd1 and DERL2 alone, but biotinylation of a complex containing Hrd1, Herp, FAM8A1 and DERL2 required optimisation. Purification of biotinylated Hrd1 complexes containing these four proteins required the eventual overexpression of Hrd1, Herp and FAM8A1. High levels of endogenous DERL2 were co-purified with this complex and the biotinylation protocol optimised. An equivalent complex without the AviTag was

reconstituted into liposomes and llamas immunised. After affinity maturation, the llama VHH library was isolated but this failed to demonstrate any immunoreactivity towards the Hrd1 complex. This lack of reactivity could be a result of instability of the protein complex upon immunisation, co-immunisation with other highly immunogenic antigens or that the complex only had weak immunogenicity, resulting in generation weak binders [207]. Given time constraints, a second immunisation was not possible and so generation of sybodies was pursued.

Sybody binders were identified from three different artificial nanobody libraries, each of which recognises differently-shaped protein surfaces [205]. $Hrd1_{1-344}$ was screened for binders as we reasoned that this would be an easier starting point and would allow us to optimise the procedure for enrichment and isolation of binders. Enrichment was seen by qPCR for the sybody libraries throughout multiple rounds of sybody display, and binders were identified in an ELISA. 20 sybodies were purified at scale and those with good biochemical behaviour were isolated for binding with $Hrd1_{1-344}$. Of the 20 binders purified, only 10 sybodies were deemed to have good biochemical behaviour. These 10 sybodies were used in pulldowns with $Hrd1_{1-344}$ and only 4 bound small amounts of Hrd1. One sybody exhibited the strongest interaction whilst the others were of similar binding strength. However, all sybodies failed to co-elute by SEC with $Hrd1_{1-344}$. At this point, time was limited and so further sybody screening is not reported. Since then, no other strong binders to $Hrd1_{1-344}$ have been identified and this was attributed to a high prevalence of non-specific binders in the ELISA. Steps are now being taken to increase the abundance of specific binders.

The lack of high affinity binders limits our ability to determine high-resolution structures of our purified Hrd1 complexes. Our belief is that these complexes fail to reach high resolution by cryo-EM because of a lack of signal for alignment during computational processing of the cryo-EM micrographs. Addition of small, stable, high-affinity binders is known to increase the mass, stability and homogeneity of membrane protein samples

for structure determination [120,202,209,210]. Future structural studies of mammalian Hrd1 and its complexes will require generation of either nanobodies or sybodies, or other modifications to enhance the signal for alignment.

7. Reconstitution of the RNF185 E3 ligase complex for cryo-EM

In a previous chapter, a bottom-up screen was performed to identify ERAD factors amenable to cryo-EM structure determination. One of the proteins screened was TMEM259 (MBRL), which is a component of the RNF185 E3 ligase complex [105]. This was of special interest to the Carvalho lab and so alongside Hrd1 complex purifications, further attempts were made to improve the biochemical behaviour and quality of MBRL samples.

7.1. Composition of the RNF185 E3 ligase complex

The RNF185 E3 ligase complex is composed of RNF185, MBRL, TMUB1 or TMUB2, all of which are integral membrane proteins in the ER [105]. This complex was identified from a genome-wide CRISPR-Cas9 screen for proteins that affected the stability of an artificial ERAD substrate. In yeast, the first 68 amino acids of Erg11 were shown to be sufficient for ERAD degradation [57]. These residues contain an amphipathic helix and a single TMD and are degraded in an ERAD-M-like manner at the inner nuclear membrane [57]. The mammalian homologue CYP51A1 also shares the same domain organisation, and when that same region was expressed alone (termed CYP51A1TM hereafter) in HEK293 cells, it too was an ERAD-M substrate [105,225]. This led to the identification of the components of the RNF185 E3 ligase complex, which were further validated by biochemical and mass spectrometry analysis [105]. MBRL is the largest component of this complex and contains 7-8 TMDs, as predicted by AlphaFold and DeepTMHMM [110,136]. The soluble C-terminal region is predicted to be mostly disordered with a few short helical stretches [138,226]. This protein is unique to higher eukaryotes and its knockout is perinatal lethal in mice [108,227]. Re-expression of MBRL in astrocytes rescues this phenotype but how loss of MBRL results in this drastic phenotype is unknown. MBRL scaffolds the interaction between RNF185 and the TMUBs

and is essential for ERAD degradation of CYP51A1, as well as other mislocalised membrane proteins [105,228,229]. The topologies of the components of the RNF185 E3 ligase complex are shown in figure 7.1. MBRL also possess a large lumenal domain that is predicted to contain a tight, β-sheet fold by AlphaFold [136]. This domain shares no structural similarity to other domains but it is highly conserved amongst higher eukaryotes [135,230]. RNF185 is a relatively small protein (~20 kDa) and contains 1-2 predicted TMDs, depending on the prediction method [110,136]. This protein is a RING-type E3 ligase and polyubiquitinates substrates that supposedly emerge on the cytoplasmic side of the membrane [105]. TMUBs 1 and 2 are closely related with each other and have similar overall domain architecture [105]. Both of these proteins contain 3 predicted TMDs as well as a ubiquitin-like domain (UBL). The UBL domain can recruit the p97-UFD1L-NPLOC4 complex and thus may be important for substrate retrotranslocation in RNF185-mediated ERAD [104]. The TMUBs were shown to be redundant with each other in degradation of CYP51A1TM and only knockout of both proteins resulted in total substrate stabilisation [105]. However, knockout of TMUB1 significantly reduced retrotranslocation of other substrates [104]. This redundancy highlights the possibility that TMUB binding may be mutually exclusive in the RNF185 E3 ligase complex.

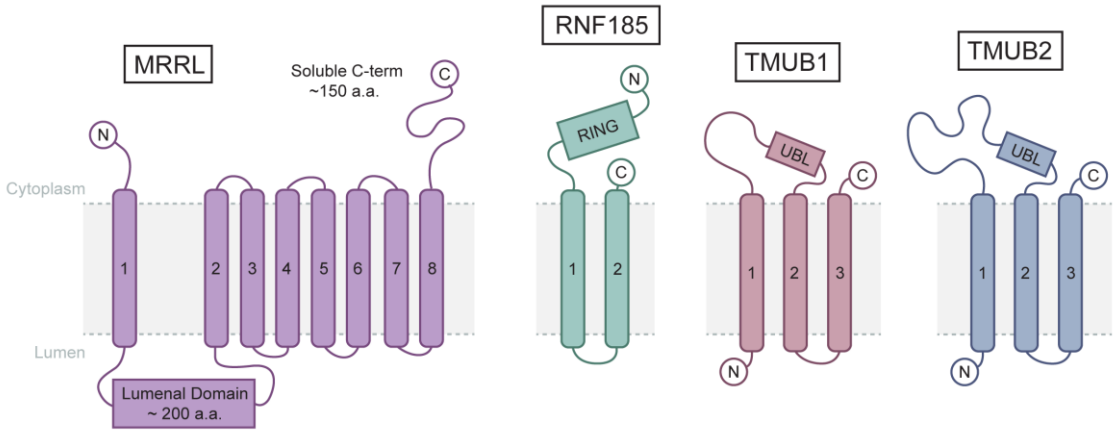


Figure 7.1. Components of the RNF185-MBRL complex. The overall topology was determined from TMD predictions with DeepTMHMM and from the literature [105,110].

7.2. Purification of H. sapiens MBRL

Previously, when ERAD factors were screened for their cryo-EM amenability, MBRL was purified in a single condition and imaged by cryo-EM. Purification of MBRL in DDM-CHS resulted in a significant aggregating shoulder by SEC (shown in figure 7.2a-b for reference). When imaged by cryo-EM, protein from the SEC peak fraction was also aggregated and cryo-EM data collection not performed as this was unlikely to result in high-resolution data (based on previous lab experience). Other detergent conditions for extraction of MBRL from membranes were also screened alongside DDM-CHS, some of which exhibited different behaviour. In DDM conditions, titration of CHS caused the appearance of an additional, later peak, corresponding to an object with a smaller hydrodynamic radius (chapter 3, figure 3.7). This implied that some level of disassembly was occurring when CHS was added. This second peak was never observed in LMNG conditions and when CHS was titrated in, the relative abundance of the aggregating shoulder decreased (chapter 3, figure 3.7).

Since the behaviour was significantly different in LMNG, MBRL was also purified at scale in LMNG-CHS, with the hope that this had improved biochemical behaviour (figure 7.2c-d). The same cell line used for expression of MBRL in chapter 3 was used to isolate MBRL by StrepTactin purification and SEC. Despite a high protein purity, there was a significant aggregating shoulder of the main peak (figure 7.2c-d). The majority of MBRL eluted over a wide range of SEC volumes which is indicative of high structural heterogeneity or interaction with the column. The aggregation could be attributed to a propensity for aggregation at higher concentrations, existence of multiple conformations or a lack of correct folding/assembly during expression. To further understand MBRL

quality in LMNG, protein from the peak SEC fraction was imaged by cryo-EM (figure 7.2f). In the previous DDM condition, detergent micelles containing protein were observed but these were very clumpy (figure 7.2e). When purified in LMNG, the protein particles had uneven shapes and sizes, and were generally very heterogenous. There was also a level of contaminating background particles present. These contaminants may be empty detergent micelles or minor contaminants from the purification process. However, no distinct micelle shapes could be seen and protein particles did not have well-defined borders as in the DDM sample. No cryo-EM data collection was performed on this sample due to its poor quality. Our belief is that when LMNG-extracted MBRL is vitrified during cryo-EM grid preparation it undergoes some level of denaturation. During the vitrification process, proteins diffuse through the liquid phase and come into contact with the air-water interface thousands of times prior to freezing [190]. This contact can cause denaturation for some proteins and is a well-documented problem with cryo-EM grid preparation [190]. A recent analysis of many different proteins frozen on cryo-EM grids found that most had a tendency to adhere to the air-water interface and were not distributed evenly in vitreous ice [231]. Partial or full denaturation would result in high particle heterogeneity and so the sample needed to be improved. If different behaviour of MBRL was observed prior to freezing, then behaviour on cryo-EM grids may be different. Throughout many of the detergent screening conditions, MBRL always exhibited some level of aggregation. To understand if the presence of the bulky fluorescence tag was affecting folding/assembly during expression, expression of MBRL without mVenus was trialled. This failed to yield any improvement in the SEC behaviour of MBRL (not shown). MBRL also contains a significant amount of disorder in its cytoplasmic C-terminal domain (~70% predicted disorder) [138]. Truncation of the disordered C-terminus failed to improve MBRL behaviour by SEC (not shown). MBRL was also purified in GDN but this too was very heterogenous by SEC and cryo-EM (not

shown). Together, the data suggest that isolation of MBRL alone in detergent does not result in homogenous protein, an important pre-requisite for cryo-EM [201].

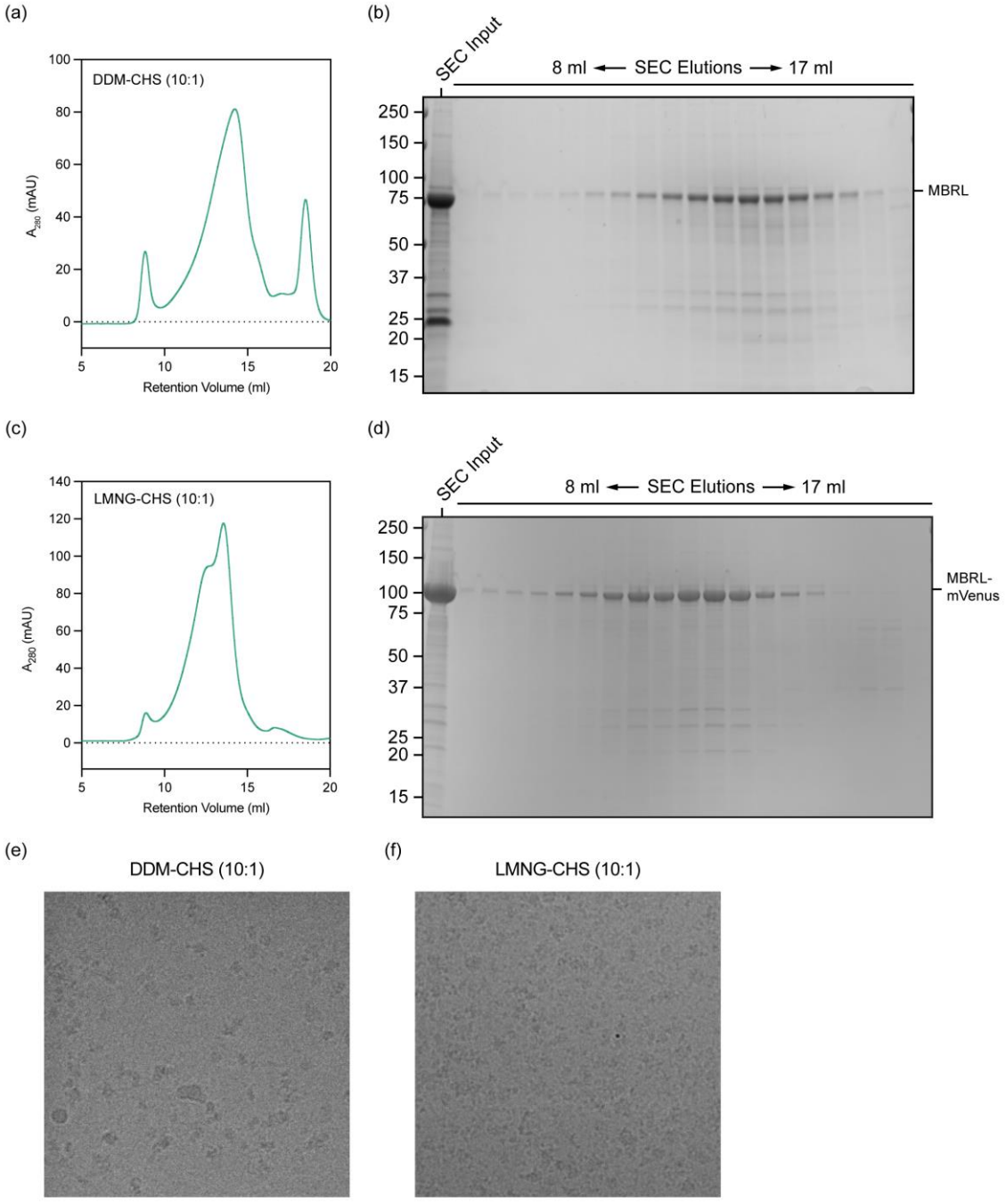


Figure 7.2. Purification of MBRL in different detergent conditions for cryo-EM. (a) SEC profile after TwinStrep purification of MBRL in DDM-CHS (10:1). (b) SDS-PAGE of the MBRL SEC elutions in DDM-CHS. (c) SEC profile after TwinStrep purification in LMNG-

CHS (10:1). (d) SDS-PAGE of the MBRL SEC elutions in LMNG-CHS. (e) and (f) are representative cryo-EM micrographs of MBRL purified in DDM and LMNG, respectively.

7.2.1. Screening other membrane mimetics

Approximately 20% of membrane protein structures solved by cryo-EM are of proteins in detergent-free membrane mimetics, such as amphipols or nanodiscs [126]. Detergents are an attractive choice because they are easy to work with, often require less optimisation towards a final sample and there is a large range of detergents with different biochemical properties [232]. Amphipols are a detergent-free membrane mimetic that can be exchanged with detergent to maintain membrane proteins in a soluble form outside of the lipid bilayer [233]. They are a class of surfactants that through their amphipathic nature, self-assemble into well-defined particles [234]. A combination of short chain length and specific number of hydrophobic segments enables a hydrophobic core to be established with only a few amphipol molecules, thus preventing large networks of amphipol forming in aqueous solution [233]. Amphipols have much higher affinities for hydrophobic TMDs than detergents and so they remain tightly associated with membrane proteins, even when there are is no free amphipol in solution [233,235]. This feature is attractive for cryo-EM since free detergent micelles are no longer present and the lack of detergent can improve ice thickness during grid preparation [236]. A8-35 and PMAL-C8 are the most popular amphipols for cryo-EM studies of membrane proteins, likely due to their high solubility in aqueous solution and high assembly homogeneity [233,236]. However, as amphipols are poorly dissociating surfactants they are not efficient at solubilising biological membranes [237,238]. Instead, membrane proteins are first extracted into detergent and then exchanged into amphipols. Amphipols are mixed with purified membrane proteins and this allows amphipol exchange. The detergent is then titrated away by dilution below the CMC or by addition

of polystyrene beads that adsorb detergent molecules [233]. The ratio of amphipol to membrane protein often needs to be optimised to ensure complete reconstitution of the protein and prevent aggregation [233]. Since MBRL behaved poorly after detergent extraction, we hypothesised that after amphipol exchange the biochemical behaviour of MBRL may improve, making it more amenable to structure cryo-EM determination.

To screen amphipol reconstitution, MBRL was initially extracted from membranes with DDM. The purified protein was then exchanged into A8-35 and PMAL C8 amphipols, following protocols for reconstitution of other small membrane proteins in amphipols [60,121]. Reconstitution efficiency was compared between the two amphipol types by SDS-PAGE (figure 7.3a). The reconstitution efficiency varied greatly between A8-35 and PMAL-C8 but remained similar across different concentrations of amphipol. MBRL was reconstituted with a greater efficiency into A8-35 amphipols than PMAL C8 amphipols (figure 7.3a). Interestingly, in all A8-35 conditions, a higher MW MBRL species corresponding to dimeric MBRL was stabilised. Homogeneity and biochemical behaviour post-amphipol exchange were assessed by FSEC (figure 7.3b-c). A8-35 reconstituted MBRL behaved very similarly to DDM-extracted MBRL. There was still an aggregating shoulder present as well as a later peak containing a smaller object, possibly monomeric MBRL. Increased concentrations of A8-35 during reconstitution resulted in a marginal increase in reconstitution efficiency but did not change the biochemical behaviour of MBRL. MBRL still behaved like its detergent counterpart by SEC. When reconstituted into PMAL C8 amphipols, MBRL was poorly behaved and there was low fluorescent signal for all conditions trialled. Low signal could be a result of protein instability and precipitation during the reconstitution or precipitation during SEC. Only at high ratios of PMAL C8 was there an observable fluorescent signal and the SEC profile for the 1:8 and 1:10 conditions did not appear homogenous. There were multiple peaks and the retention volume shifted greatly between conditions suggesting failed assembly with amphipol.

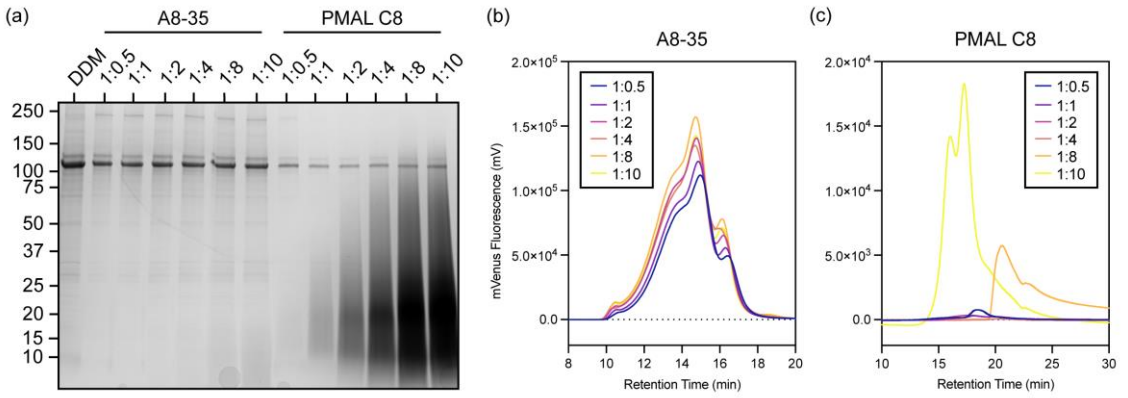


Figure 7.3. Screening amphipol reconstitution conditions for DDM-extracted MBRL. (a) SDS-PAGE gel of MBRL reconstituted into A8-35 or PMAL C8 amphipols at different protein to amphipol ratios. A DDM control of the input protein to the exchange reaction is also shown. (b) and (c) are FSEC profiles for A8-35 and PMAL C8 reconstituted MBRL, respectively.

Since exchange from DDM to amphipols did not significantly change the behaviour of MBRL, LMNG-extracted MBRL was also reconstituted into amphipols (figure 7.4). Only the A8-35 amphipol was efficient in reconstituting MBRL and so this was used to reconstitute LMNG-extracted MBRL at different reaction ratios. Once again, amphipol exchange did not alter MBRL behaviour by SEC when compared to its detergent-extracted counterpart (figure 7.4a). This suggests that MBRL is already biased with poor biochemical behaviour before amphipol exchange. Alternatively, it may be possible that there is still some detergent present in the sample, causing the amphipol FSEC profiles to appear similar to detergent conditions. However, control exchange reactions without amphipol present resulted in complete protein precipitation (not shown). A larger volume reaction condition was also run by SEC to see if UV absorbance correlated with fluorescence (figure 7.4b). Again, a higher MW MBRL species was seen after western blotting of the SEC elutions (figure 4.4c). This corresponds to a MBRL-mVenus dimer

and perhaps amphipols stabilise this state or are more resistant to SDS denaturation prior to loading on an SDS-PAGE gel. Nevertheless, the data suggest that some proportion of A8-35 MBRL could exist as a dimer after both DDM and LMNG extraction.

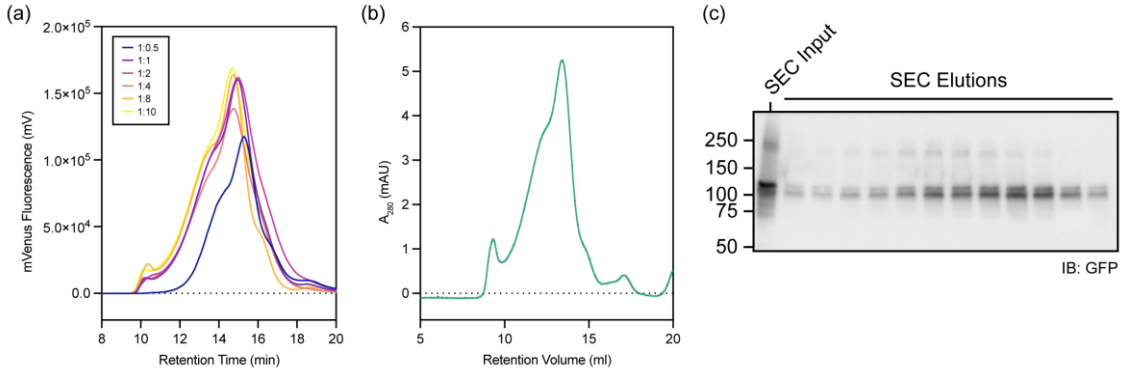


Figure 7.4. SEC screening of A8-35 amphipol reconstitution of LMNG-extracted MBRL. (a) FSEC profiles of LMNG-extracted MBRL reconstituted into A8-35 amphipols. (b) SEC profile of A8-35 reconstituted MBRL (1:1 reaction) at a larger scale. (c) Western blot of the (b) SEC elution fractions for GFP.

Although the behaviour of MBRL in A8-35 amphipols was similar by SEC to its detergent counterpart, the stabilisation of a dimeric species suggested that the protein could have increased stability. We reasoned that this material could behave differently when frozen on cryo-EM grids and so this purification was scaled up (figure 7.5). Large quantities of MBRL were purified in LMNG-CHS and exchanged to A8-35 amphipol in a 1:1 reaction. SEC was performed after reconstitution and this resulted in a profile similar to the LMNG equivalent (figure 7.5a). SDS-PAGE of the peak fraction had no obvious dimeric species present but as previously seen, the monomer band was shifted likely due to the presence of bound amphipol. Grids were prepared of this material and a small cryo-EM data collection performed (figure 7.5b). After multiple rounds of 2D classification with SIMPLE, an object resembling a membrane protein in a detergent micelle was

observed [162]. However, there was no clear secondary structure detail and 2D class averaging indicated that there was a high degree particle heterogeneity. Amphipol reconstitution therefore did not improve the sample enough for high-resolution cryo-EM studies.

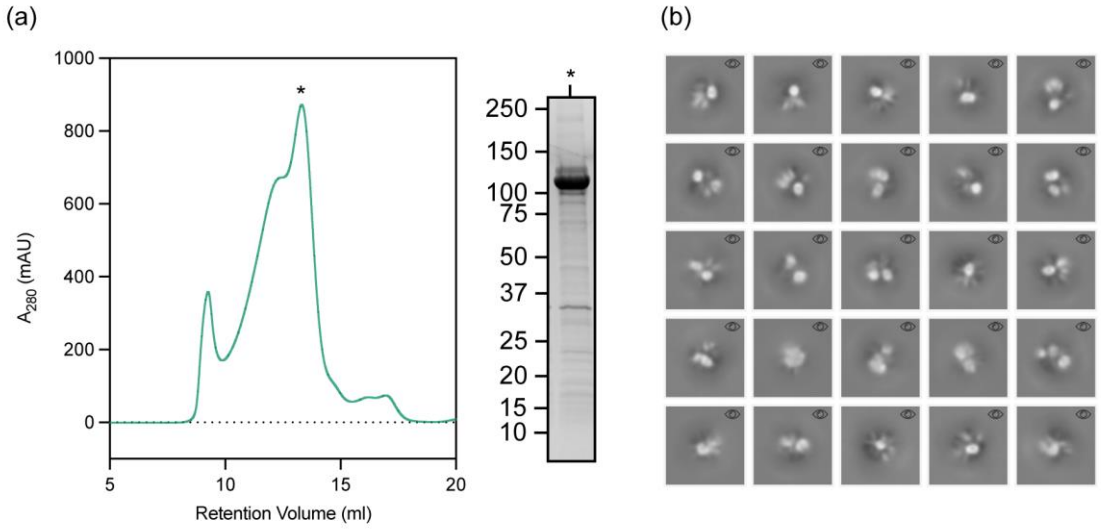


Figure 7.5. Larger-scale A8-35 amphipol reconstitution of MBRL for cryo-EM. (a) SEC profile after A8-35 amphipol reconstitution (1:1) of LMNG-extracted MBRL. SDS-PAGE of the peak elution fraction is also shown. (b) Cryo-EM 2D class averages after a small data collection of A8-35 amphipol reconstituted MBRL. Cryo-EM micrographs were processed with SIMPLE [162].

An alternative membrane mimetic that has had some success in cryo-EM of membrane proteins is the use of styrene-maleic acid (SMA) copolymers [239–241]. Reconstitution of membrane proteins in detergent can have deleterious effects on folding and function, and so extraction methods that preserve the native membrane environment may be better [236]. SMAs solubilise cellular membranes directly into nanoscale lipid particles, termed SMALPs, without the addition of detergent [242]. The extracted SMALPs are akin to protein nanodiscs in structure but they capture native lipids in a

bilayer environment [242]. The copolymer encircles captured lipids and this environment is close to the native lipid environment. However, SMA composition can be quite heterogenous and so this can influence downstream extraction efficiency and protein homogeneity [236]. Conditions for efficient reconstitution can also be difficult to establish as SMAs are pH sensitive due to their negatively charged maleic acid molecules [242]. To date, only a handful of membrane protein structures in SMALPs have been determined by cryo-EM [239–241]. Diisobutylene maleic acid (DIBMA) copolymers are an alternative to SMAs which were developed to improve solubility and in-solution behaviour, but a high-resolution cryo-EM structure is yet to be solved with DIBMAs [243]. Given the different properties of this mimetic, MBRL was reconstituted into variations of these copolymers (figure 7.6). All copolymers successfully extracted soluble MBRL from overexpressed membranes but only SMALP 200 and SMALP 140-I extracted MBRL in a form competent for StrepTactin purification (figure 7.6a). This implied that after most copolymer extractions, MBRL is not folded correctly. The StrepTactin elutions from each pulldown were also subject to FSEC, to see if small amounts of protein were eluted and to assess biochemical behaviour (figure 7.6b). A DDM-extracted control on the same column and under similar conditions eluted after ~14 minutes (not shown). The only copolymer condition to approximate this was SMALP 200 but for all other conditions, the eluted proteins were very small, possibly monomeric MBRL or other cleavage products. However, there was a significant size difference after SMALP extraction of MBRL, suggesting either MBRL disassembly, cleavage or small lipid nanoparticles. Given time constraints, this was not pursued further but in future, determining the biochemical nature of those peaks may provide be a route to high-resolution cryo-EM of MBRL.

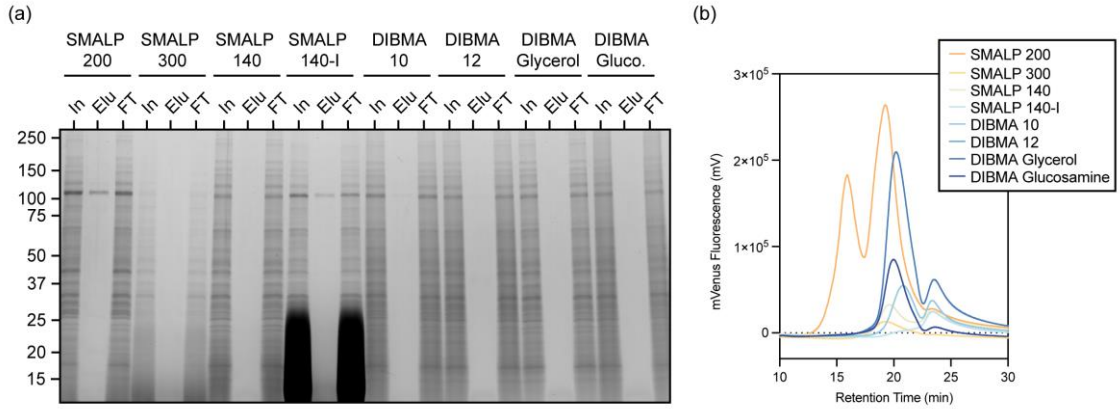


Figure 7.6. Screening for amenability of SMALP reconstitution of MBRL. (a) Efficiency of SMALP extraction of MBRL from overexpressed membranes. StrepTactin pulldowns were performed on the extracted, soluble material. In = solubilised input, elu = elution and FT = unbound material. (b) FSEC profiles of the StrepTactin elutions.

7.2.2. Screening the biochemical behaviour of MBRL homologues

Different MBRL homologues may exhibit improved biochemical behaviour and so a range of other species homologues were screened for their amenability to cryo-EM structure determination. Since MBRL is highly conserved amongst metazoans, its structure and therefore function are likely conserved. A sequence identity of at least 20% is known to preserve similar structural elements between proteins and so different homologues were identified with this cut-off [244]. Homologues from very similar organisms (*Papio Anubis*, *Pongo abelii*, *Mus musculus*), from organisms with different body temperatures (*Gallus gallus*, *Danio rerio*, *Caenorhabditis elegans*) and those with different sequence lengths (*Chinchilla lanigera*, *Sus scrofa*, *Drosophila melanogaster*), were selected for screening (figure 7.7). No homologues were identified from thermophilic organisms, which would be ideal structural candidates as their proteomes are more thermostable [245]. Ideally, AlphaFold would be used as a tool to identify homologues with distinct structural differences and therefore possible distinct

biochemical behaviour, but this was not available [136]. Instead, a range of homologues were selected with different sequence identities (figure 7.7b). The selected homologues all had similar numbers of predicted TMDs, except for that from *Canis lupus*, which is predicted to have only 6 TMDs. The MBRL homologues from *C. lanigera*, *C. lupus* and *D. melanogaster* are significantly longer than other MBRL sequences. This is attributed by the presence of large, soluble, cytoplasmic tails, the function of which is unknown in human MBRL [105]. With these selected homologues, expression, amenability to purification and biochemical stability were screened to identify homologues that were more amenable to cryo-EM than human MBRL.

(a)

Species	MW (kDa)	#TMDs
Homo sapiens	67.9	8
Papio anubis	67.7	8
Pongo abelii	65.5	8
Mus musculus	63.5	8
Gallus gallus	81.3	8
Danio rerio	77.7	8
Caenorhabditis elegans	68.6	8
Chinchilla lanigera	133.4	8
Sus scrofa	84.9	8
Canis lupus	97.3	6
Drosophila melanogaster	103	8

(b)

	H. sapiens	*P. anubis*	*P. abelii*	*M. musculus*	*G. gallus*	*D. rerio*	*C. elegans*	*C. lanigera*	*S. scrofa*	*C. lupus*	*D. melanogaster*
H. sapiens	-	96%	94%	84%	76%	70%	41%	89%	91%	88%	50%
P. anubis	96%	-	92%	84%	72%	66%	42%	88%	90%	87%	55%
P. abelii	94%	92%	-	80%	71%	66%	54%	89%	86%	83%	50%
M. musculus	84%	84%	80%	-	75%	71%	52%	86%	83%	84%	51%
G. galllus	76%	72%	71%	75%	-	75%	40%	74%	71%	75%	75%
D. rerio	70%	66%	66%	71%	75%	-	41%	68%	67%	70%	75%
C. elegans	41%	42%	54%	52%	40%	41%	-	54%	54%	41%	39%
C. lanigera	89%	88%	89%	86%	74%	68%	54%	-	87%	85%	75%
S. scrofa	91%	90%	86%	83%	71%	66%	54%	87%	-	91%	55%
C. lupus	88%	87%	83%	84%	75%	70%	41%	85%	91%	-	53%
D. melanogaster	50%	55%	50%	51%	75%	75%	39%	75%	55%	53%	-

Figure 7.7. Comparison of predicted structural and sequence differences between different species MBRLs. (a) Table for comparison of the total MW and predicted number of TMDs with DeepTMHMM [110]. TMDs were predicted with DeepTMHMM and TMHMM [5,110]. (b) Sequence similarity between different MBRL homologues was calculated by local alignment with BLASTP [135].

Sequences from the selected homologues were codon optimised for expression in HEK cells and cloned into vectors for overexpression with a C-terminal TEV-mVenus-

TwinStrep tag. The mVenus tag was used for initial expression screening by flow cytometry and in downstream detergent screening. After generation of stable cell lines with lentiviral transduction, only *G. gallus*, *P. anubis*, *P. abelii*, *S. scrofa* and *M. musculus* showed any expression by flow cytometry and western blot (not shown). Small amounts of *D. melanogaster* MBRL were seen by western blot but expression was very poor so this was excluded from further screening. In detergent screening, no fluorescent signal could be seen for *P. abelii* or *S. scrofa* in a region expected for detergent-extracted membrane proteins on a Superose 6 Increase column (figure 7.8, 8-17 ml). Therefore, the expressed protein from these homologues was likely all aggregated or cleavage products. *G. gallus* MBRL was prone to aggregation in all detergent conditions, and only a small amount of material eluted in the main peak. There was always the presence of an aggregating shoulder and a large amount of protein eluted at the void volume. This suggested that this homologue is not stable either due to overexpression, lack of native partners or detergent extraction. Of all the conditions tested, the GDN-extracted protein was the least stable whilst protein in the maltoside and neopentyl glycol conditions behaved similarly. *P. anubis* MBRL behaved similarly to human MBRL during detergent screening. In maltoside conditions, titration of CHS caused the appearance of a second peak which was possibly monomeric MBRL. In neopentyl glycol conditions this second peak is not present and titration of CHS only led to increased extraction efficiency. The GDN-extracted protein behaved the worst and had a significant tendency for aggregation. *M. musculus* also behaved similarly to human MBRL with the exception that all peaks appeared slightly sharper. This indicated that there was slightly improved homogeneity of detergent-extracted *M. musculus* MBRL. Unfortunately, none of the screened homologues exhibited significantly different behaviour to human MBRL by SEC.

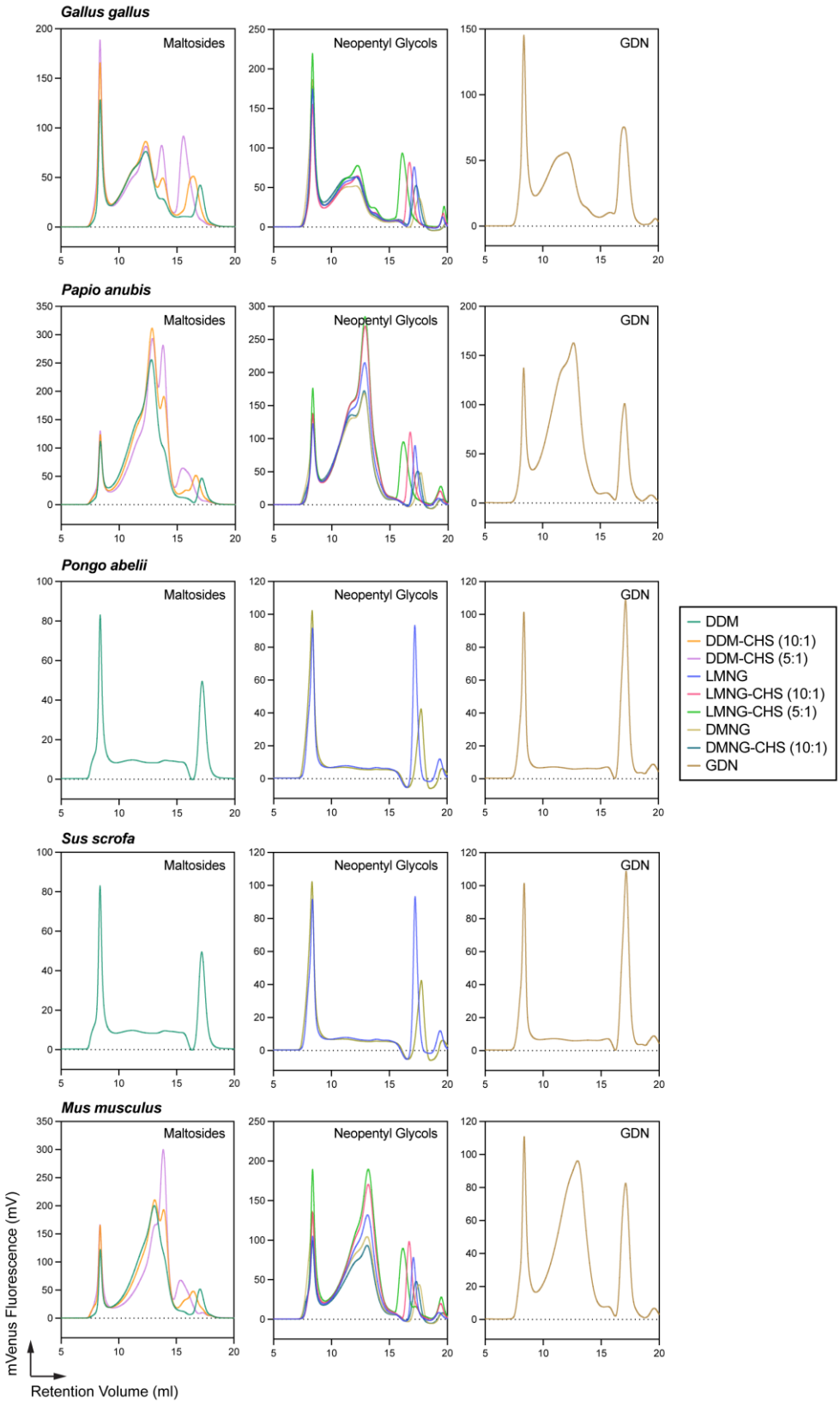


Figure 7.8. Detergent screening of MBRL homologues by FSEC. Those with expression confirmed by western blotting were subject to detergent screening. Membranes from

cells overexpressing recombinant protein were prepared and solubilised in a range of conditions. Solubilised material was run by SEC and protein fluorescence monitored.

Since only the *G. gallus*, *M. musculus* and *P. anubis* homologues expressed well and could be purified successfully in a range of detergent conditions, cell lines expressing these proteins were scaled up for larger scale purifications (figure 7.9). All proteins were extracted from membranes in LMNG-CHS (10:1) and StrepTactin purifications performed to isolate pure protein. Cryo-EM grids of LMNG-containing buffers are usually made at lower protein concentrations [246,247]. Therefore, purifying in LMNG would be advantageous if the final yields were low, maximising the chance of cryo-EM imaging. All homologues aggregated similarly to human MBRL by SEC (figure 7.9a, c and e). They all eluted over a large range of volumes (~5 ml), indicative of severe heterogeneity within the samples or column interactions. Interestingly, all homologues exhibited a main peak at ~12 ml followed by a second peak at ~14.5 ml. This behaviour was not observed with human MBRL in LMNG and suggests that these homologues may be oligomeric and prone to disassembly in LMNG. The peak fraction from each of these samples was analysed by cryo-EM (figure 7.9b, d and f). *G. gallus* and *M. musculus* MBRL exhibited distinct micelle-like particles reminiscent of protein in detergent. This was an improvement over human MBRL but the particles still remained heterogenous in shape and size, so no data collection was performed. *P. anubis* MBRL was more similar to human MBRL in its cryo-EM behaviour and it appeared to undergo denaturation during vitrification. No homologues were pursued any further given their poor behaviour by SEC or lack of expression. The differences observed in SEC between the small-scale detergent screening and larger-scale purifications may be attributed to concentration-dependent aggregation. In the future, identification of distinct structural homologues by AlphaFold prediction may prove more successful in determining a cryo-EM structure of MBRL alone.

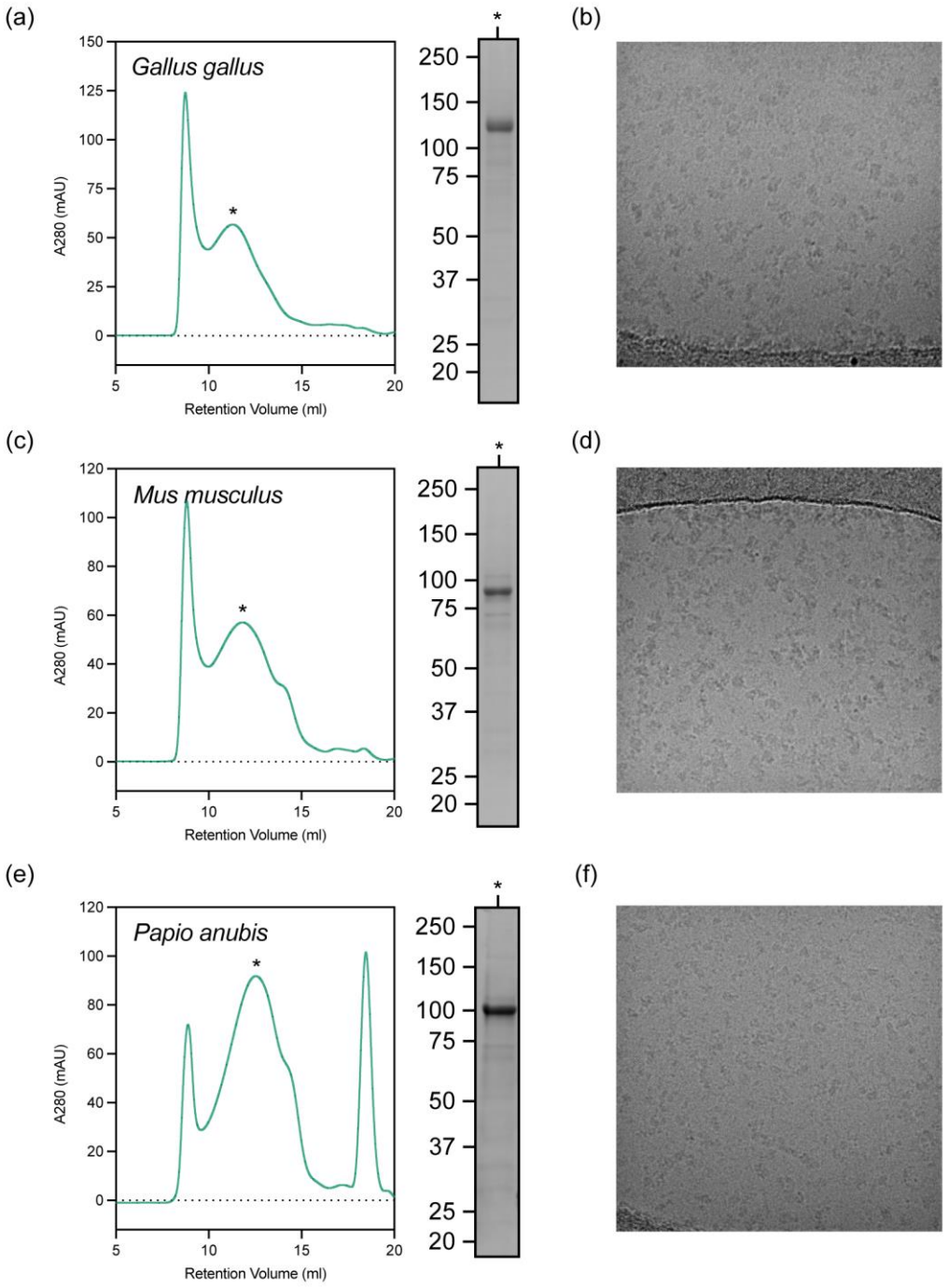


Figure 7.9. Larger-scale purifications and cryo-EM of a selection of MBRL homologues. (a) SEC profile after StrepTactin purification of *Gallus gallus* MBRL in LMNG-CHS (10:1). (b) Cryo-EM micrograph of frozen *G. gallus* MBRL. (c) SEC profile after StrepTactin purification of *Mus musculus* MBRL in LMNG-CHS (10:1). (d) Corresponding cryo-EM

micrograph of frozen *M. musculus* MBRL. (e) SEC profile after StrepTactin purification of *Papio anubis* MBRL in LMNG-CHS (10:1). (f) Cryo-EM micrograph of frozen *P. anubis* MBRL.

7.3. Co-purification of partners with MBRL

When MBRL was reconstituted alone, its biochemical behaviour was always poor despite screening a large range of different detergents and other membrane mimetics. We reasoned that reconstitution with partners of the endogenous RNF185 E3 ligase complex may improve stability, given that these proteins are known to co-purify with each other during pulldowns [105]. In the experience of the lab, recombinant membrane proteins often behave better when reconstituted with their endogenous binding partners. At present, MBRL is known to interact with RNF185, TMUB1 and TMUB2. To that end, we sought to isolate complexes containing these proteins to see if these complexes would be more amenable to cryo-EM structural studies.

7.3.1. Reconstitution of RNF185-MBRL

All the partners of the RNF185 E3 ligase complex are membrane proteins and so for reconstitution into the same detergent micelle, co-expression was required. The MBRL cell line generated previously was further transduced with lentivirus containing a vector for expression of 6xHis-mCherry-RNF185. RNF185 had previously been N-terminally tagged and was still component for assembly with MBRL and function [105]. Protein expression was confirmed by red/green fluorescence observed by flow cytometry, so larger scale purifications were performed (figure 7.10). The RNF185-MBRL complex was isolated by dual affinity purification in both DDM-CHS and LMNG-CHS (figure 7.10a-b). In both conditions, two bands were enriched on SDS-PAGE gels that were green and red fluorescent. This corresponded to MBRL and RNF185, respectively, and no major

differences were observed between the two detergent conditions. There was an additional co-purifying band at the 50 kDa marker for which no known component of the RNF185 E3 ligase complex corresponded to. Mass spectrometry identification of this band revealed that this was a DNA-binding protein and so a likely contaminant. The additional bands below the 37 kDa marker were also identified as contaminating proteins. Due to time restraints, no specificity control was included in the dual affinity purifications and so these interactions could be non-specific. SEC of this complex revealed that there was less aggregation present in both detergent conditions, compared to purifications of MBRL alone. Interestingly, the DDM condition behaved the same as the LMNG condition and there was now no presence of the additional, later peak that usual elutes close to 15 ml retention volume. This suggested that reconstitution of RNF185 with MBRL stabilised the complex. The reduction in total protein aggregation also aligns with this observation. The main SEC peak was sharper indicating improved homogeneity and appeared almost stochiometric by SDS-PAGE. The intensity of the MBRL band in the SEC peak was slightly higher than of RNF185 but this could be attributed to better Coomassie staining. One additional difference between this sample and previous MBRL samples is that the tags were not cleaved off and so the presence of soluble mVenus and mCherry tags may confer additional complex stability.

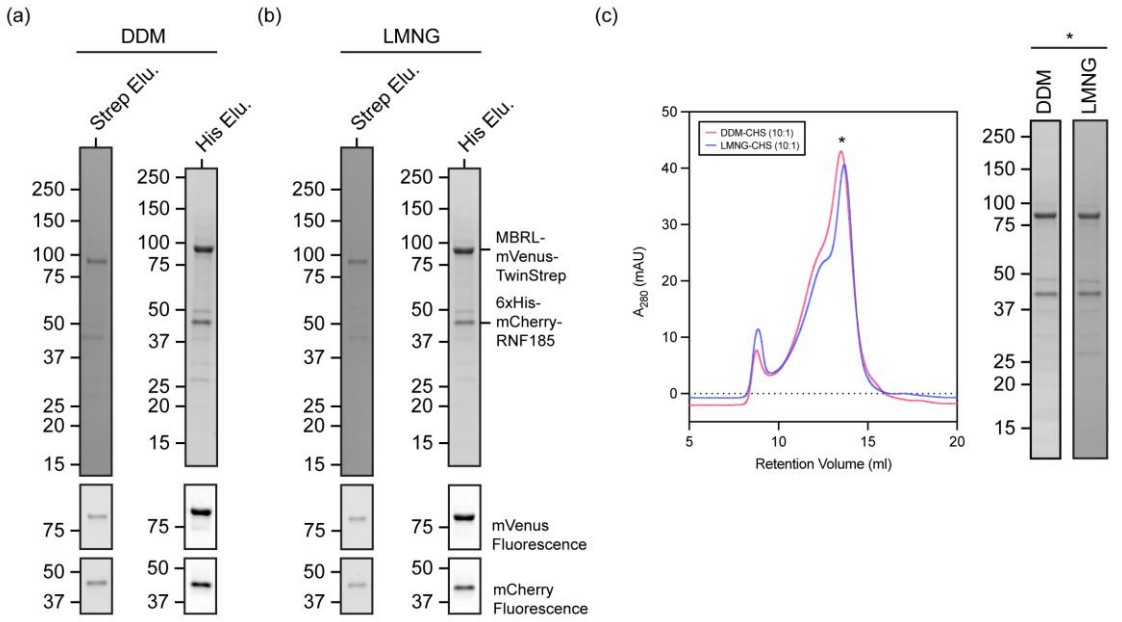


Figure 7.10. Purification of a complex containing RNF185 and MBRL. StrepTactin and Ni-NTA (His) purifications were used to enrich for the complex in either DDM-CHS or LMNG-CHS detergent conditions. (a) Dual affinity purification of the RNF185-MBRL complex in DDM-CHS. (b) Dual affinity purification of the RNF185-MBRL complex in LMNG-CHS. (c) SEC profiles for the dual affinity purified complexes. Material from the indicated peak fraction was analysed by SDS-PAGE.

The LMNG-extracted complex was imaged by cryo-EM and a small data collection performed on a Titan Krios G3 operating at 300 kV (figure 7.11). At the micrograph level, the protein particles looked similar to previous MBRL samples extracted in LMNG (figure 7.11a). Particles were aggregated and denatured. Despite this, a small data collection was performed to confirm the quality of the sample and the data processed with SIMPLE [162]. Movies were processed as previously in this thesis and particles picked with a reference volume of an unrelated small membrane protein from the lab. 2D class averages yielded no secondary structure detail and were very heterogenous (figure 7.11b). The shapes and sizes of the particles varied significantly and so the data was not processed any further. The DDM-extracted complex was also analysed by cryo-EM but

this behaved similarly to previous DDM-extracted MBRL samples and was severely aggregated. Although reconstitution of MBRL with RNF185 improved biochemical behaviour by SEC, this did not improve the behaviour of frozen protein on cryo-EM grids. This suggested that MBRL was still driving a lot of the behaviour seen on cryo-EM grids.

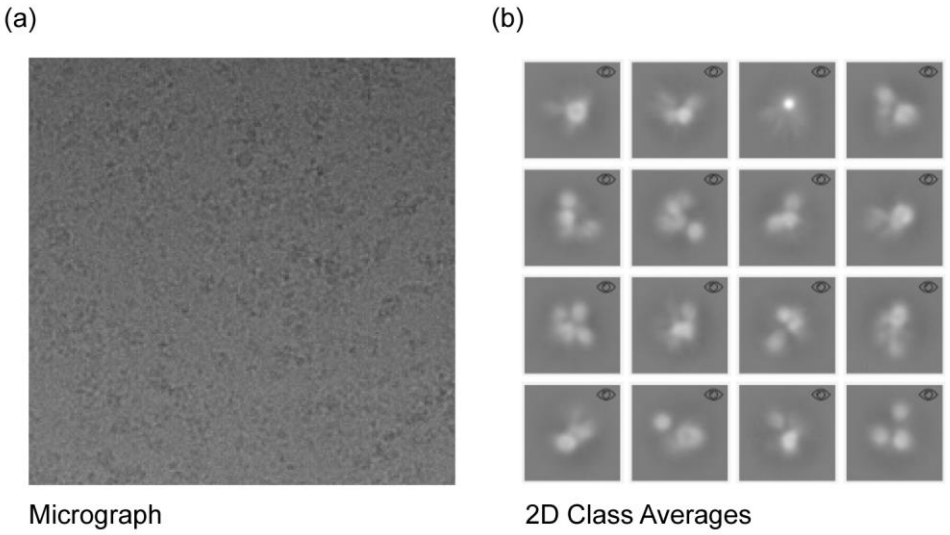


Figure 7.11. Cryo-EM analysis of the RNF185-MBRL complex in LMNG-CHS. (a) Representative cryo-EM micrograph of the complex frozen within vitreous ice. (b) 2D class averages generated with SIMPLE [162].

7.3.2. Reconstitution of RNF185-MBRL-TMUB complexes

In a further attempt to improve the homogeneity and behaviour of MBRL-containing complexes, efforts were made to reconstitute a complex containing RNF185, MBRL and the TMUBs. Since the TMUBs were shown to be somewhat redundant with each other in substrate degradation assays, cell lines overexpressing either TMUB1 or TMUB2 in addition to the other components were generated [105]. To do this, the cell line overexpressing MBRL-mVenus-TwinStrep and 6xHis-mCherry-RNF185 was further transduced with lentivirus containing untagged TMUB1 or TMUB2. After antibiotic selection to enrich cells expressing the TMUBs, large scale purifications were performed.

Membrane proteins were solubilised with LMNG-CHS since there was not a significant difference in SEC behaviour compared to the DDM-extracted RNF185-MBRL complex. This also facilitated sample preparation for cryo-EM grids due to lower final yield requirements for grid preparation. Overexpression of MBRL, RNF185 and untagged TMUBs resulted in reconstitution of complexes containing TMUBs (figure 7.12a-b). Interestingly, the ratio of MBRL reconstituted in the complex relative to RNF185 increased when TMUBs were overexpressed. The significance of this is unknown but dual affinity purification resulted in similar stochiometric levels of TMUB1/2 and RNF185 by SDS-PAGE. It is possible that the initial MBRL purification is dominating and results in high MBRL levels throughout further purifications. SEC of the purified complexes revealed that they had a large hydrodynamic radius or were very aggregated (figure 7.12c). Both complexes eluted over a large volume (~5 ml) and now there were two major peaks present. The new peak at ~12 ml corresponds to the same volume of the aggregating shoulder seen in previous purifications of MBRL and MBRL-containing complexes in LMNG. One important difference is that the final quantity of purified protein is greater in the MBRL-RNF185-TMUB1/2 purifications and so these changes may be attributed to concentration-dependent aggregation. Protein from the first major peak (~12 ml) was imaged on cryo-EM grids (figure 7.13). SDS-PAGE gels of the peak SEC fractions are shown in figure 7.13 and these reveal that the purified complex consists primarily of MBRL and only small amounts of RNF185 and TMUB1/2. RNF185 and the TMUBs appear equimolar by SDS-PAGE but the organisation of the complex appears changed from the previous RNF185-MBRL complex extracted in LMNG. Also, TMUB2 seems more abundant in the complex than TMUB1 by SDS-PAGE (figure 7.13). Since the TMUBs were untagged, it is difficult to know if this difference in abundance is due to different expression levels or is detergent-related behaviour. The lentiviral titres during the two separate TMUB1 and TMUB2 transductions are likely different and so would result in stable cell lines with different levels of transgene insertion and hence different

expression levels. The TMUBs may also express to different levels in HEK293 cells. Despite the addition of a new component to the complex, both complexes behaved similarly to other MBRL-containing complexes in LMNG on cryo-EM grids (figure 7.13). The frozen protein was aggregated and still appeared partially denatured. Reconstitution of MBRL with its known partners failed to improve the biochemical behaviour of the complex in LMNG. Other detergent conditions were trialled with different MBRL complexes but also without success (not shown).

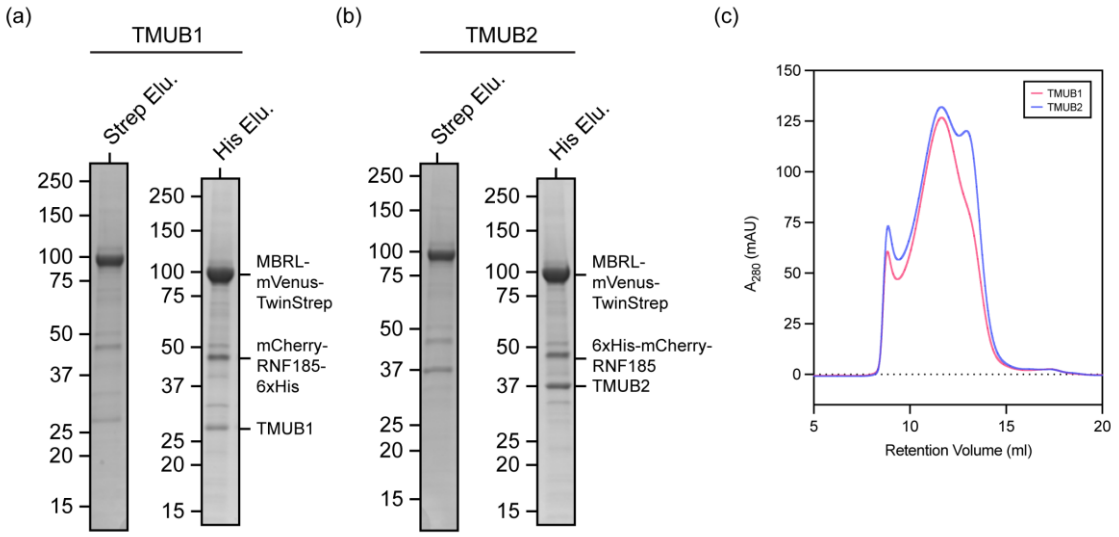


Figure 7.12. Reconstitution of complexes containing RNF185, MBRL and TMUB1/2. (a) and (b) are SDS-PAGE gels of the elution fractions after dual affinity purification by StrepTactin and Ni-NTA (His) resin, for the TMUB1-containing and TMUB2-containing complexes, respectively. (c) SEC profiles for the dual affinity purified complexes. The TMUB1-containing complex is coloured in pink, and the TMUB2-containing complex coloured in blue.

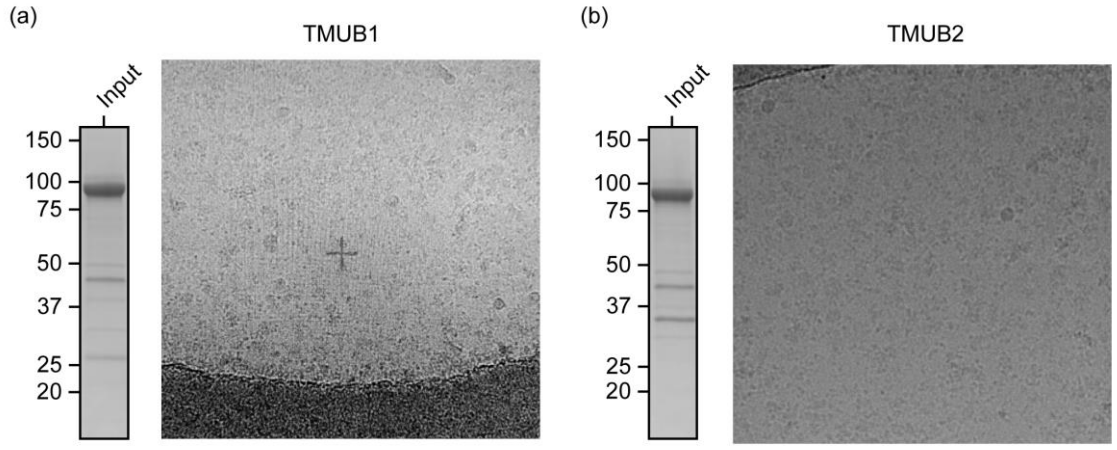


Figure 7.13. Cryo-EM analysis of the peak SEC fractions after isolation of complexes containing RNF185, MBRL and TMUB1/2. (a) SDS-PAGE gel of the peak SEC fraction concentrated for imaging on cryo-EM grids, as well as a representative cryo-EM micrograph for the TMUB1-containing complex. (b) SDS-PAGE gel of the peak SEC fraction and cryo-EM micrograph of the TMUB2-containing complex.

7.3.3. Reconstitution of MBRL with substrate

A further attempt to improve the biochemical behaviour of MBRL was made by co-expression and reconstitution with an artificial ERAD substrate. Since MBRL is part of an ERAD complex that processes membrane substrates, we reasoned that if it interacted with substrate, it may become more stable. The RNF185 E3 ligase complex is known to process an ERAD substrate made from a portion of CYP51A1 [105]. The N-terminus of CYP51A1 contains an amphipathic helix and a single TMD (CYP51A1TM) that when expressed with a C-terminal sfGFP tag, is a strong ERAD substrate [105]. For co-expression and purification with MBRL, a C-terminal 6xHis tag was added to CYP51A1TM-sfGFP and lentivirus used to introduce this into cells already expressing MBRL-TwinStrep. Dual-affinity purification was performed in LMNG and a complex containing MBRL and CYP51A1TM isolated (figure 7.14a). Small amounts of substrate were enriched after dual affinity purification and this was further subject to SEC (figure

7.14b). Since the substrate was fluorescently tagged, fluorescence was monitored and the fluorescent substrate peak co-eluted with the UV absorbance peak. Inspection by SDS-PAGE of the SEC elution fractions confirmed that substrate co-eluted with MBRL. A small amount of substrate dissociated during the purification as seen by the small peak at ~16 ml. This shows that MBRL is capable of forming a complex with substrate and this function is likely important in RNF185-mediated ERAD. It is unknown whether the other partners of the RNF185 complex also interact with substrate but given that MBRL confers the most TMDs to the complex, an ability to interact and bind to membrane substrates is sensible. By SDS-PAGE, this interaction is almost 1:1 of MBRL to substrate. Unfortunately, this did not change the biochemical behaviour of MBRL by SEC and so this sample was not analysed by cryo-EM.

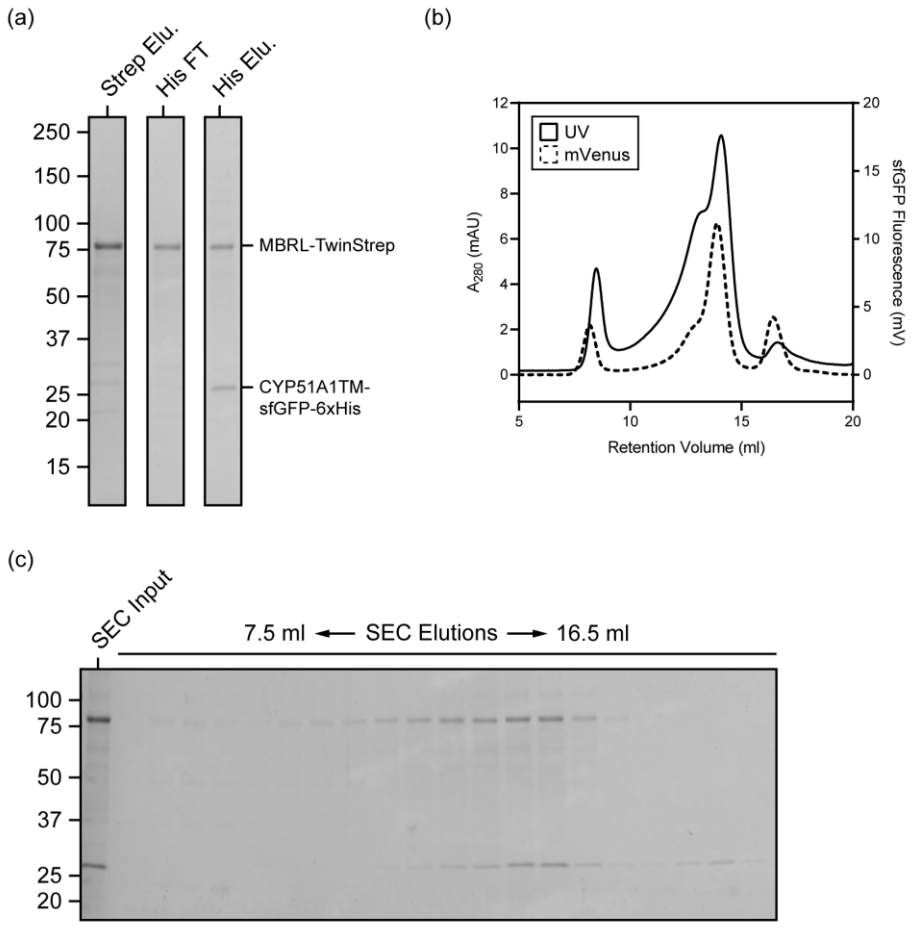


Figure 7.14. Reconstitution of an artificial substrate with MBRL. (a) SDS-PAGE gels of the elutions from the StrepTactin and Ni-NTA (His) purifications. (b) SEC profile for the dual affinity purified MBRL-substrate complex. (c) SDS-PAGE gel of the SEC input and elutions.

7.4. Isolation of the endogenous RNF185-Membralin complex

Since attempts to overexpress MBRL with and without its endogenous partners failed to yield a well-behaved, homogenous complex, we opted to purify the endogenous complex without overexpression. Overexpression of the complex or addition of numerous recombinant tags may affect assembly of the complex and/or cause protein aggregation. The endogenous complex may also associate with other stabilising factors that were not identified in previous work. Some of the work described in this section is already published in van de Weijer et al. [105]. To achieve this, endogenous MBRL was tagged with an mVenus-TwinStrep tag using a modified version of the protocol described by Fuellar et al. [117]. HEK Flp-In-293 cells were transfected with Cas12a and a PCR cassette to tag MBRL. Modified cells were enriched by antibiotic selection and single clones expressing mVenus isolated by flow cytometry assisted cell sorting. Validation of tagged MBRL clones was carried out by western blot and Cas9 knockout (KO) with MBRL guide RNAs. A clone was isolated that constitutively expressed mVenus and mVenus expression was ablated by guide RNA KO of MBRL (figure 7.15a). To probe the native partners of the MBRL complex, membranes from these CRISPR-tagged cells were harvested, solubilised in DMNG and the solubilised material analysed by SEC (figure 7.15b). MBRL generally eluted in the 12-13 ml region on an S6 increase 10/300 GL SEC column and small amounts of tagged MBRL (~100 kDa) were observed by SDS-PAGE. There was also a significant population of untagged MBRL at ~75 kDa. Therefore, at least one MBRL allele was tagged with mVenus-TwinStrep. RNF185 and the TMUBs are

relatively small membrane proteins and so would be expected to elute in the >15 ml range if they were unassembled. A large proportion of RNF185 and TMUB2 eluted in the 12-13 ml range, co-eluting with MBRL. Only a small amount of TMUB1 co-eluted with MBRL and the rest appeared to form a complex with a larger hydrodynamic radius near the 9-10 ml elution volume. DMNG-extracted RNF185 complex primarily contained TMUB2 and MBRL. This complex of proteins co-eluted across a narrow range (2-3 ml), a drastic difference from previous recombinant approaches. However, small quantities of proteins generally elute over a smaller SEC volume than in larger SEC runs [158]. The 12 ml elution fraction was taken and further analysed in a GFP-pulldown (figure 7.15c). All components of the RNF185 complex were enriched after pulldown on MBRL. Interestingly, untagged MBRL also co-eluted with the complex, a strong indication that MBRL is oligomeric within the complex.

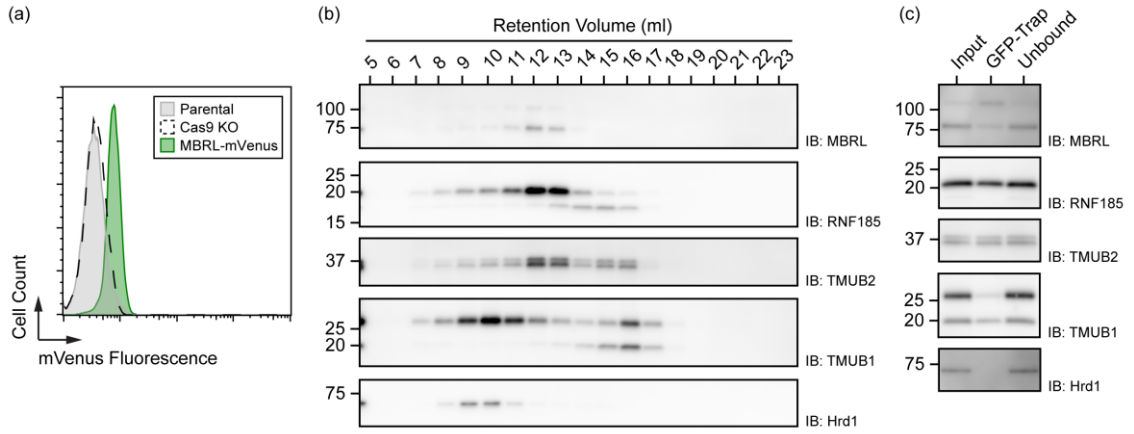


Figure 7.15. CRISPR tagging of endogenous MBRL and analysis of the native RNF185 complex assembled in membranes. (a) Flow cytometry of the CRISPR-tagged cell line and validation of tagging by CRISPR Cas9 knockout (KO). (b) Western blot of DMNG-solubilised membranes injected directly over a SEC column (S6 increase 10/300 GL). (c) GFP pulldown of the 12 ml SEC elution fraction. Hrd1 served as a pulldown control. (b) and (c) were modified from van de Weijer et al. [105].

Given that the behaviour of the endogenous complex was different by SEC, the CRISPR-tagged cell line was transferred to suspension culture and a large volume grown. Membranes were solubilised in LMNG and a StrepTactin purification performed (figure 7.16a). The yield was not high enough to visualise proteins by Coomassie staining and so western blotting was performed. All components of the RNF185 complex were enriched after MBRL pulldown and there were almost stochiometric quantities of tagged and untagged MBRL. This is likely because expression levels of untagged MBRL were greater (figure 7.15) and so the chance of associating with an untagged copy was higher than with a tagged copy. The StrepTactin eluted material was subject to FSEC and this demonstrated very different behaviour to previous SEC of recombinant MBRL complexes. Although there was some aggregated protein present, a narrow, homogenous fluorescent peak was observed at 13.5 ml. Western blot analysis of the peak indicated that it contained MBRL, RNF185, TMUB1 and TMUB2. The equimolar ratio of tagged and untagged MBRL was maintained across the SEC peak. TMUB1 eluted across a range of volumes as before and the majority of TMUB1 eluted earlier than the MBRL complex. This indicates that TMUB1 may aggregate under these conditions or forms other higher MW complexes with a larger hydrodynamic radius. Mass spectrometry analysis of RNF185 complexes did not identify any other protein partners that could stabilise the complex for overexpression [105]. Therefore, recombinant overexpression of MBRL and/or its partners is likely disrupting its biochemical behaviour and causing significant heterogeneity in the purified protein. Unfortunately, the native RNF185 E3 ligase complex is not expressed highly enough in cells and so it was not feasible to scale this cell line further for structural studies.

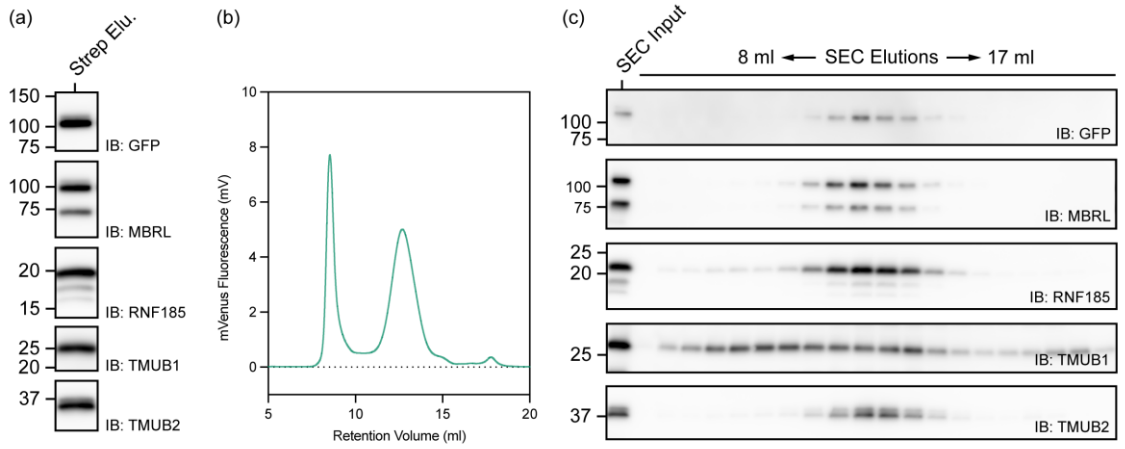


Figure 7.16. Purification of the endogenously tagged RNF185 E3 ligase complex. (a) Western blots of the StrepTactin purification of the RNF185-complex. At least one MBRL allele is endogenously tagged with mVenus-TwinStrep. (b) FSEC profile of the StrepTactin purified complex. (c) Western blots of the SEC elutions for the components of the MBRL complex. All components co-elute with MBRL except for TMUB1, that forms higher MW complexes/aggregates.

7.5. Discussion

In this chapter, attempts were made to isolate components of the RNF185 E3 ligase complex for cryo-EM structural studies. The RNF185 E3 ligase complex is composed of the E3 ligase RNF185, MBRL, TMUB1 and TMUB2, and is implicated in the degradation of membrane substrates by ERAD [105]. Of this complex, MBRL is the largest and confers the most TMDs. Amenability for MBRL solubilisation from membranes with detergent was initially screened in a previous chapter and attempts to purify this in DDM-CHS failed to yield homogenous, well-behaved protein. Other conditions for solubilisation of MBRL were trialled in this chapter. Detergent extraction with the milder detergent LMNG resulted in protein that was still significantly aggregated. Cryo-EM analysis of LMNG-extracted MBRL revealed that the purified protein was very heterogenous and frozen protein particles varied highly in size and shape. There also

appeared to be denaturation occurring during the grid preparation process, as evidenced by soluble-like protein in the cryo-EM micrographs [190]. Since other detergent conditions did not vary significantly in the detergent screening, we sought to use different membrane mimetics for extraction of MBRL from membranes.

Given the success of amphipols in reconstitution of membrane proteins for cryo-EM, MBRL was exchanged from detergent into A8-35 and PMAL C8 amphipols, with the hope of improved biochemical behaviour and stability [126]. MBRL was exchanged from DDM into amphipols and exchange efficiency was better with the A8-35 amphipol. However, amphipol reconstituted material behaved like its DDM-extracted counterpart and no improvement by SEC was observed. MBRL failed to exchange efficiently with PMAL C8. Amphipol exchange from LMNG-extracted MBRL was also trialled and this also behaved like its detergent counterpart by SEC. Western blot analysis of the SEC elutions revealed the presence of a higher MW MBRL species that corresponded to dimeric MBRL. We reasoned that overall stability may be improved and so cryo-EM behaviour of this sample may be different. Cryo-EM analysis however revealed that A8-35 reconstituted MBRL was not improved and the frozen protein looked similar to LMNG-extracted MBRL. Extraction of MBRL from membranes into SMALPs was also trialled as this was a significantly different mimetic to detergents and amphipols. Across a range of SMALPs screened, only SMALP 200 and 140-I extracted MBRL in a form competent for StrepTactin binding. SEC of the purified material revealed the presence of two symmetric peaks but due to time constraints, these were not pursued further. In future, investigation into the identity of these peaks and may lead to an improved MBRL sample for cryo-EM.

Since human MBRL was poorly behaved, other species homologues were screened for their purification amenability. A total of 10 other eukaryotic homologues were selected for expression trials and these had varied sequence diversity and a range of MWs. Only 5 of the homologues expressed in our HEK293S expression system and these were subject to detergent screening. Of the 5, only 3 of these homologues were expressed

and extracted as non-aggregated protein. Homologues that failed to express or fold correctly possibly require co-expression of other partners for stability or need species specific chaperones to guide folding and assembly. Larger scale purifications were performed with 3 homologues and material from the peak SEC fractions analysed by cryo-EM. All of these homologues were highly heterogenous and no data collection was performed. Poor MBRL behaviour was conserved across homologues indicating that expression of MBRL alone was not the approach to continue for cryo-EM analysis.

Complexes of MBRL with RNF185, TMUB1 and TMUB2 were then isolated by co-overexpression. Initially, RNF185-MBRL was isolated in different detergent conditions but these also behaved similarly to MBRL purified alone. The stoichiometric ratio of MBRL to RNF185 appeared almost equal by SDS-PAGE. Cryo-EM grids of this complex were also highly heterogenous and so complexes containing MBRL, RNF185 and TMUB1 or TMUB2 were isolated. Unfortunately, these too were also highly heterogenous and failed to yield high quality cryo-EM data. MBRL was also reconstituted with an artificial substrate to see if this stabilised and improved MBRL behaviour. This too failed to change the behaviour of MBRL by SEC but substrate was successfully reconstituted with MBRL, an interaction that is likely important in RNF185-mediated ERAD.

Given that attempts to isolate recombinant MBRL and its partners failed, endogenous MBRL complexes were isolated from a CRISPR tagged cell line. Analysis of solubilised membranes revealed that MBRL co-eluted with RNF185, TMUB2 and small amounts of TMUB1. Enrichment of this material by pulldown further showed that MBRL existed as an oligomer in this complex, since untagged MBRL co-purified with tagged MBRL. Purification from a large cell mass of this cell line was performed and all the components of the RNF185 E3 ligase complex were enriched in a StrepTactin purification of MBRL. SEC of this complex revealed that it was homogenous and separated well from aggregates in the void volume, a drastic change from previous attempts to purify this complex. Western blot analysis revealed that all 4 proteins co-eluted together. TMUB1

eluted over a wide range of volumes with a tendency to elute closer to the void volume, indicating either formation of a separate, larger complex or a tendency to aggregate under these conditions. Despite this improvement in behaviour, the expression of the endogenous RNF185 E3 ligase complex was too low to scale up. Future work on this complex may require further attempts to optimise purification of the endogenous RNF185 E3 ligase complex, since this was the best-behaved of all complexes purified. The endogenous tagging protocol could be improved to increase the number of tagged MBRL alleles, resulting in final higher yields. Expression of this complex may also be enhanced by addition of histone deacetylase inhibitors or by tagging in an alternative cell line that expresses the complex to higher levels [248].

8. Conclusions and future perspectives

During this project, we sought to understand the mechanism of ER-associated E3 ligases through structural studies. There is limited structural information available from studies on the yeast E3 ligase Hrd1 and no structures exist for any mammalian E3 ligase involved in ERAD. A dimeric yeast Hrd1-Hrd3 structure was solved that was later refuted as a purification artefact due to the unlikely position of amphipathic helix inside of the expected membrane environment [55,60]. A heterodimer of yeast Hrd1-Der1 was also solved, where Hrd1 and Der1 are each proposed to contribute a half-channel for substrate retrotranslocation [55]. The precise mechanism of retrotranslocation remains a mystery but recent crosslinking studies with substrate have begun to decipher a path through the yeast Hrd1-Der1 heterodimer [56].

8.1. The Hrd1 E3 ligase complex

To gain insight into mammalian E3 ligase mechanisms, a bottom-up screen was performed to identify E3s immediately amenable to cryo-EM. After screening of purification conditions, a variety of E3s were purified at scale and imaged by cryo-EM. Only mammalian Hrd1 was initially amenable to structure determination and resulted in determination of two distinct dimeric conformations, resolving the TMDs and the RING domains. Both these dimeric models were in a unique arrangement from the previously solved yeast Hrd1 dimer and were C2-symmetric. We termed these conformations the “closed” and “open” states given the relative position of their RING domains. The RING domain was positioned directly above the core TMDs in both conformations, a position that is likely important for ubiquitination of emerging substrates. In the closed state, the opposing RING domains are positioned close together with only 17 Å separation but in the open state, the RING domains are separated by 31 Å. In both states, the individual monomers have the same overall fold with an RMSD of < 1 Å. Instead, rearrangements

at the dimer interface facilitate rotation of the individual monomers in an axis perpendicular to the dimer interface. Importantly, in the closed state no density for TMD1 of Hrd1 was observed in the cryo-EM map but in the open state this is clearly resolved. These states also differ dramatically in their membrane environments. The detergent micelle is clearly visualised in both maps and in the closed state, the micelle is severely thinned at the dimer interface. In the open state, the micelle becomes flatter with only marginal thinning at the interface. Membrane thinning was also observed for the yeast Hrd1-Der1 heterodimer but was not as pronounced as in the mammalian closed Hrd1 state [55]. Another important difference with mammalian Hrd1, is that overexpression and extraction in DDM and LMNG both resulted in dimer formation, an event that was not observed with yeast Hrd1 [55].

The significance of the Hrd1 dimer is unknown but work on endogenous Hrd1 has revealed the existence of higher-order oligomers [56,62,91,105]. The retrotranslocation activity of yeast Hrd1 for certain substrates is known to be stimulated by auto-ubiquitination and more recently Hrd1 dimers were found to be a requirement for auto-ubiquitination [92,249,250]. Yeast Hrd1 is regulated by cycles of auto- and de-ubiquitination of conserved lysine residues near the RING domain, and for an ERAD-L substrate this was shown to enhance degradation [92]. Dimers are required for this auto-ubiquitination event and this is proposed to happen in *trans* between the individual monomers [250]. It is unknown if this event causes conformational changes but crosslinking studies suggest that auto-ubiquitination triggers dimer dissolution and perhaps drives assembly of alternative Hrd1 complexes [56]. Despite this, expression of yeast Hrd1 alone can drive retrotranslocation of some substrates indicating retrotranslocation activity [93,249]. Mammalian Hrd1 can also be ubiquitinated but whether it undergoes the same events as its yeast counterpart is unknown since there are not many lysine residues near its RING domain and auto-ubiquitination has not been observed [251].

The mammalian Hrd1 complex has different partners to the yeast system and so structural insight would help decipher mammalian ERAD mechanisms. An expression system was developed for reconstitution of multi-component membrane protein complexes from HEK cells. This was used to isolate different sub-complexes by overexpression of different combinations of partners. We successfully reconstituted a range of complexes in different detergent conditions including: Hrd1-FAM8A1, Hrd1-Herp-FAM8A1, Hrd1-Herp-FAM8A1-DERL2, Hrd1-Herp-FAM8A1-DERL2-Sel1L, Hrd1-Sel1L, Hrd1-Herp-FAM8A1-DERL1 and a Hrd1-substrate complex. Some of these complexes were analysed by cryo-EM but not all yielded high-resolution structural information. No FAM8A1 density was visible in the Hrd1-FAM8A1 map, despite the presence of FAM8A1 throughout purifications. The final map resembled the Hrd1 closed state and no density was seen for the cytoplasmic region after the RING domain. FAM8A1 predominantly interacts with Hrd1 through interactions with the HAF-H domain in the cytosolic tail of Hrd1 [61]. Interaction between the TMDs of Hrd1 and FAM8A1 has been reported but our data suggest that this interaction is weak, flexible or not symmetric, all of which would reduce visible density during cryo-EM processing [61]. Complexes containing DERL2 were also analysed by cryo-EM and these revealed the presence of two distinct complexes or conformations. Addition of Sel1L to this complex resulted in the best resolution maps but these failed to achieve the high-resolution needed for model building and map interpretation. Despite this, two states could be separated, one of which resembled S-shaped micelles and the other flatter micelle shapes. These two objects were reminiscent of the two Der1 topologies observed with the yeast Hrd1-Der1 heterodimer, where ~50% of the purified complexes had Der1 in a flipped orientation [55]. Models were not built into the maps as they were too low resolution, so the significance of these two objects is unknown. In all complexes containing Sel1L, no density for Sel1L was visible during 2D classifications or 3D reconstructions despite it containing large, soluble SLR repeats [178]. Sel1L was unstable in many conditions

trialled and we suspect that the SLR domains denature during cryo-EM grid preparation. A complex containing DERL1 was also assembled and this contained Hrd1, Herp and FAM8A1. This complex was similar to the DERL2-equivalent complex and implies a conserved interaction mode of Derlins with Hrd1. Substrate was also reconstituted with Hrd1 but due to time constraints, this was not further optimised. Together, the data identified a range of compatible Hrd1 complexes that likely all exist in cells and probably represent assembly or regulatory intermediates, or complexes that process different classes of substrates (figure 8.1).

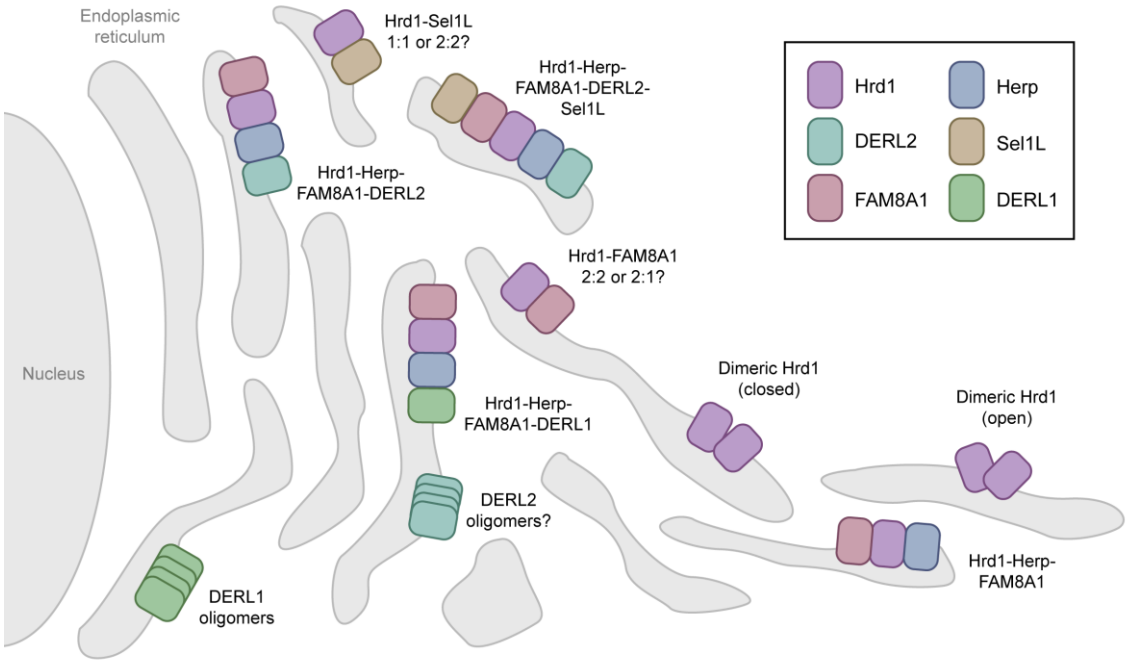


Figure 8.1. Possible endogenous Hrd1 complexes in mammalian cells. Annotated complexes are based off observations in this thesis and the literature [55,60–62,94].

8.2. The RNF185 E3 ligase complex

Attempts were also made to reconstitute components of the RNF185 E3 ligase complex for cryo-EM. MBRL was initially reconstituted in a range of membrane mimetics,

all of which failed to yield a well-behaved, homogenous sample. Cryo-EM analysis of these samples revealed high particle heterogeneity and no secondary structure was observed in any 2D class averages. Other homologues were also screened for their purification amenability but similar behaviour was observed by SEC. All of these also failed to yield homogenous samples by cryo-EM and no improvement in MBRL behaviour was observed. Given that MBRL is an integral part of the RNF185 complex, we reasoned that co-expression of components of RNF185 or the TMUBs might stabilise and therefore improve biochemical behaviour [105]. Complexes containing MBRL-RNF185 and MBRL-RNF185-TMUB1/2 were successfully assembled but these were also aggregation prone. When imaged by cryo-EM, all purified complexes were heterogenous and no secondary structure detail was observed in 2D class averages. The endogenous RNF185 complex was also isolated using CRISPR to introduce a purification tag on at least one allele of endogenous MBRL [105,117]. Separation of solubilised membrane proteins by SEC revealed that MBRL, RNF185, TMUB1 and TMUB2 co-eluted together, in a probable complex [105]. The bulk of TMUB1 however eluted earlier and so was either aggregating under the experimental conditions or it formed higher MW complexes with itself or other proteins. TMUB1 was recently shown to promote extraction of TMDs from the ER membrane and to exist as homo-oligomers, which might explain this observation [104]. Enrichment of a SEC fraction by pulldown identified that MBRL is oligomeric within the complex since untagged MBRL also enriched. Interestingly, the elution profile of the RNF185 complex was narrower than recombinant complexes and so this was purified at scale. All components of the RNF185 complex were enriched during MBRL pulldown and the SEC profile revealed that the complex was homogenous and non-aggregating. However, the final yields were too low to pursue this for structural studies.

8.3. Future directions for structural studies of mammalian E3 ligases

Without further structural insight into a mammalian Hrd1 complex with its partners or substrate, it is difficult to discern how substrates access Hrd1 and what role partners play in retrotranslocation. Instead, perhaps more questions now exist as to the significance of the mammalian Hrd1 dimer and if the open and closed state are both physiological, what role do they play in Hrd1 function. Our cryo-EM analysis of Hrd1 complexes was limited by a lack of high-quality alignments during data processing. Efforts were made to generate high-affinity binders to the Hrd1 complex to increase the size, stability and homogeneity of the samples [204,215]. Addition of small binders (nanobodies) has previously facilitated structure determination of other small membrane proteins by cryo-EM and is known to increase sample homogeneity by stabilising particular conformations [120,205,210]. Llamas were immunised with a Hrd1 complex reconstituted into liposomes but this failed to yield any specific binders. Instead, we focused on the generation of sybodies but this is still in the early stages and no high-affinity binders have been identified yet. Nanobodies or sybodies are likely critical in facilitating high-resolution cryo-EM of Hrd1 complexes and may serve as a useful tool for *in vivo* assays if binders trap particular states. Given that no density was seen for the SLR repeats of Sel1L, determining if this is a result of denaturation during cryo-EM grid preparation would be useful for future structural studies. Buffer additives can abrogate this behaviour in some cases and this would increase the effective folded mass of the complex and may help image alignments during processing [190]. The functional role of the mammalian Hrd1 dimer was also not explored. Future studies may now use structures of Hrd1 and the location of key conserved residues to probe function. The significance of the membrane pocket created between TM helices 1 and 2 could be explored by mutational analysis, to see if this disrupts the recognition and retrotranslocation of Hrd1 substrates. There were a series of strongly conserved charged residues that lined TMD8 on the cytosolic side

and these may be important for substrate positioning or interaction with other partners, both of which can now be explored. If a mammalian Hrd1 auto-ubiquitination mechanism is identified and is important for regulation like its yeast counterpart, this could now be subject to cryo-EM analysis to understand if conformational changes occur. This idea was briefly trialled *in vitro* outside of this thesis but to no success. The Hrd1-Herp-FAM8A1 complex could also be analysed by cryo-EM in DDM. If it were monomeric, it would identify Herp as a factor for dimer dissolution, since Hrd1-FAM8A1 retained dimeric Hrd1.

During our studies, we failed to determine structures of any other E3 ligases or their complexes, despite considerable effort for the RNF185 complex. Structural predictions of other E3 ligases revealed that many would theoretically thin the membrane on one side [136]. This may be an important conserved feature of ER-associated E3s and thus hints at the possibility that E3s other than Hrd1 act as translocons. Some of the E3s chosen for initial purification screening (MARCHF6 and RNF145) failed to express and these were ones involved in sterol metabolism [84,164]. Structural studies of these E3s will likely require expression of catalytically inactive versions or expression in a different host system. Although TRC8 could be purified, it was significantly heterogenous by SEC and efforts to improve sample homogeneity could make it amenable to cryo-EM analysis. TRC8 is known to form complexes with other proteins implicated in ERAD, such as signal peptide peptidase, and reconstitution with its partners may improve its biochemical behaviour [252]. Other than Hrd1, Gp78 was the most well-behaved E3 but no secondary structure detail was observed in 2D class averages. Given its similarity with Hrd1 (27% sequence identity) but different substrate preference, structural characterisation of this ligase should be pursued further with the aim of identifying unique features for substrate recruitment [20]. Cryo-EM attempts of the RNF185 complex failed due to high particle heterogeneity, despite a wide range of membrane mimetics and homologues screened. Given that the endogenous complex was very well-behaved, future attempts to solve

structures of this complex would likely involve optimisations with endogenous protein. Endogenous expression could be increased by additives such as histone deacetylase inhibitors or by using an alternative cell line where expression is higher [248]. Identification of endogenous substrates may help identify cell lines where expression of the RNF185 complex is higher. Further screening of other homologous complexes may also yield a complex amenable to cryo-EM.

The structures and work here present a step forward in our current understanding of mammalian E3 ligases and their role in ERAD. Further work is absolutely required to fully understand how mammalian Hrd1 assembles with its partners and how substrate is recognised and retrotranslocated through the lipid bilayer.

9. References

1. Guerriero, C.J. and Brodsky, J.L. (2012) The delicate balance between secreted protein folding and endoplasmic reticulum-associated degradation in human physiology. *Physiol Rev* 92, 537–576
2. Toyama, B.H. and Hetzer, M.W. (2013) Protein homeostasis: live long, won't prosper. *Nat Rev Mol Cell Biol* 14, 55–61
3. Jelokhani-Niaraki, M. (2022) Membrane Proteins: Structure, Function and Motion. *Int J Mol Sci* 24, 468
4. Hegde, R.S. and Keenan, R.J. (2022) The mechanisms of integral membrane protein biogenesis. *Nat Rev Mol Cell Biol* 23, 107–124
5. Krogh, A. *et al.* (2001) Predicting transmembrane protein topology with a hidden markov model: application to complete genomes11Edited by F. Cohen. *Journal of Molecular Biology* 305, 567–580
6. Zhang, X. and Shan, S. (2014) Fidelity of Cotranslational Protein Targeting by the Signal Recognition Particle. *Annual Review of Biophysics* 43, 381–408
7. Shao, S. and Hegde, R.S. (2011) Membrane protein insertion at the endoplasmic reticulum. *Annu Rev Cell Dev Biol* 27, 25–56
8. Braakman, I. and Bulleid, N.J. (2011) Protein folding and modification in the mammalian endoplasmic reticulum. *Annu Rev Biochem* 80, 71–99
9. Duttler, S. *et al.* (2013) Principles of Cotranslational Ubiquitination and Quality Control at the Ribosome. *Molecular Cell* 50, 379–393
10. Wang, F. *et al.* (2013) A Cotranslational Ubiquitination Pathway for Quality Control of Misfolded Proteins. *Molecular Cell* 50, 368–378
11. Sun, Z. and Brodsky, J.L. (2019) Protein quality control in the secretory pathway. *Journal of Cell Biology* 218, 3171–3187

12. Lukacs, G.L. and Verkman, A.S. (2012) CFTR: folding, misfolding and correcting the ΔF508 conformational defect. *Trends Mol Med* 18, 81–91
13. Berner, N. *et al.* (2018) Protein Quality Control of the Endoplasmic Reticulum and Ubiquitin–Proteasome-Triggered Degradation of Aberrant Proteins: Yeast Pioneers the Path. *Annual Review of Biochemistry* 87, 751–782
14. Ward, C.L. and Kopito, R.R. (1994) Intracellular turnover of cystic fibrosis transmembrane conductance regulator. Inefficient processing and rapid degradation of wild-type and mutant proteins. *J Biol Chem* 269, 25710–25718
15. Bhattacharya, A. and Qi, L. (2019) ER-associated degradation in health and disease – from substrate to organism. *J Cell Sci* 132, jcs232850
16. Christianson, J.C. and Carvalho, P. (2022) Order through destruction: how ER-associated protein degradation contributes to organelle homeostasis. *The EMBO Journal* 41, e109845
17. Hetz, C. *et al.* (2020) Mechanisms, regulation and functions of the unfolded protein response. *Nat Rev Mol Cell Biol* 21, 421–438
18. Gardner, B.M. *et al.* (2013) Endoplasmic reticulum stress sensing in the unfolded protein response. *Cold Spring Harb Perspect Biol* 5, a013169
19. Fribley, A. *et al.* (2009) Regulation of Apoptosis by the Unfolded Protein Response. *Methods Mol Biol* 559, 191–204
20. Krshnan, L. *et al.* (2022) Endoplasmic Reticulum–Associated Protein Degradation. *Cold Spring Harb Perspect Biol* 14, a041247
21. Chitwood, P.J. *et al.* (2018) EMC Is Required to Initiate Accurate Membrane Protein Topogenesis. *Cell* 175, 1507-1519.e16
22. Christianson, J.C. *et al.* (2011) Defining human ERAD networks through an integrative mapping strategy. *Nat Cell Biol* 14, 93–105
23. O'Donnell, J.P. *et al.* (2020) The architecture of EMC reveals a path for membrane protein insertion. *eLife* 9, e57887

24. Pleiner, T. *et al.* (2020) Structural basis for membrane insertion by the human ER membrane protein complex. *Science* 369, 433–436
25. Chen, Z. *et al.* (2023) EMC chaperone-CaV structure reveals an ion channel assembly intermediate. *Nature* DOI: 10.1038/s41586-023-06175-5
26. Miller-Vedam, L.E. *et al.* (2020) Structural and mechanistic basis of the EMC-dependent biogenesis of distinct transmembrane clients. *eLife* 9, e62611
27. Pleiner, T. *et al.* (2023) A selectivity filter in the ER membrane protein complex limits protein misinsertion at the ER. *Journal of Cell Biology* 222, e202212007
28. Hegde, R.S. and Keenan, R.J. (2011) Tail-anchored membrane protein insertion into the endoplasmic reticulum. *Nat Rev Mol Cell Biol* 12, 787–798
29. McKenna, M.J. *et al.* (2022) ATP13A1 prevents ERAD of folding-competent mislocalized and misoriented proteins. *Molecular Cell* 82, 4277-4289.e10
30. McKenna, M.J. *et al.* (2020) The endoplasmic reticulum P5A-ATPase is a transmembrane helix dislocase. *Science* 369, eabc5809
31. Hiller, M.M. *et al.* (1996) ER degradation of a misfolded luminal protein by the cytosolic ubiquitin-proteasome pathway. *Science* 273, 1725–1728
32. Shimizu, Y. *et al.* (2010) Ubiquitylation of an ERAD substrate occurs on multiple types of amino acids. *Mol Cell* 40, 917–926
33. Wang, X. *et al.* (2007) Ubiquitination of serine, threonine, or lysine residues on the cytoplasmic tail can induce ERAD of MHC-I by viral E3 ligase mK3. *J Cell Biol* 177, 613–624
34. Ishikura, S. *et al.* (2010) Serine Residues in the Cytosolic Tail of the T-cell Antigen Receptor α-Chain Mediate Ubiquitination and Endoplasmic Reticulum-associated Degradation of the Unassembled Protein. *J Biol Chem* 285, 23916–23924
35. Christianson, J.C. and Ye, Y. (2014) Cleaning up in the endoplasmic reticulum: ubiquitin in charge. *Nat Struct Mol Biol* 21, 325–335

36. Schulman, B.A. and Wade Harper, J. (2009) Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol* 10, 319–331
37. Ye, Y. and Rape, M. (2009) Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol* 10, 755–764
38. Metzger, M.B. *et al.* (2014) RING-type E3 ligases: Master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochim Biophys Acta* 1843, 47–60
39. Cruz Walma, D.A. *et al.* (2022) Ubiquitin ligases: guardians of mammalian development. *Nat Rev Mol Cell Biol* 23, 350–367
40. Fenech, E.J. *et al.* (2020) Interaction mapping of endoplasmic reticulum ubiquitin ligases identifies modulators of innate immune signalling. *eLife* 9, e57306
41. Berndsen, C.E. and Wolberger, C. (2014) New insights into ubiquitin E3 ligase mechanism. *Nat Struct Mol Biol* 21, 301–307
42. Li, W. and Ye, Y. (2008) Polyubiquitin chains: functions, structures, and mechanisms. *Cell Mol Life Sci* 65, 2397–2406
43. Schumacher, M.M. and DeBose-Boyd, R.A. (2021) Posttranslational Regulation of HMG CoA Reductase, the Rate-Limiting Enzyme in Synthesis of Cholesterol. *Annual Review of Biochemistry* 90, 659–679
44. Chua, N.K. *et al.* (2020) Squalene monooxygenase: a journey to the heart of cholesterol synthesis. *Progress in Lipid Research* 79, 101033
45. Carvalho, P. *et al.* (2006) Distinct Ubiquitin-Ligase Complexes Define Convergent Pathways for the Degradation of ER Proteins. *Cell* 126, 361–373
46. Huyer, G. *et al.* (2004) Distinct machinery is required in Saccharomyces cerevisiae for the endoplasmic reticulum-associated degradation of a multispanning membrane protein and a soluble luminal protein. *J Biol Chem* 279, 38369–38378

47. Vashist, S. and Ng, D.T.W. (2004) Misfolded proteins are sorted by a sequential checkpoint mechanism of ER quality control. *J Cell Biol* 165, 41–52
48. Chou, T.-F. *et al.* (2011) Reversible inhibitor of p97, DBeQ, impairs both ubiquitin-dependent and autophagic protein clearance pathways. *Proc Natl Acad Sci U S A* 108, 4834–4839
49. Rabinovich, E. *et al.* (2002) AAA-ATPase p97/Cdc48p, a cytosolic chaperone required for endoplasmic reticulum-associated protein degradation. *Mol Cell Biol* 22, 626–634
50. Twomey, E.C. *et al.* (2019) Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. *Science* 365, eaax1033
51. Bodnar, N.O. *et al.* (2018) Structure of the Cdc48 ATPase with its ubiquitin-binding cofactor Ufd1–Npl4. *Nat Struct Mol Biol* 25, 616–622
52. Stapf, C. *et al.* (2011) The general definition of the p97/valosin-containing protein (VCP)-interacting motif (VIM) delineates a new family of p97 cofactors. *J Biol Chem* 286, 38670–38678
53. Lim, J.J. *et al.* (2016) Structural insights into the interaction of p97 N-terminus domain and VBM in rhomboid protease, RHBDL4. *Biochemical Journal* 473, 2863–2880
54. Cooney, I. *et al.* (2019) Structure of the Cdc48 segregase in the act of unfolding an authentic substrate. *Science* 365, 502–505
55. Wu, X. *et al.* (2020) Structural basis of ER-associated protein degradation mediated by the Hrd1 ubiquitin ligase complex. *Science* 368, eaaz2449
56. Pisa, R. and Rapoport, T.A. (2022) Disulfide-crosslink analysis of the ubiquitin ligase Hrd1 complex during endoplasmic reticulum-associated protein degradation. *J Biol Chem* 298, 102373
57. Natarajan, N. *et al.* (2020) Quality Control of Protein Complex Assembly by a Transmembrane Recognition Factor. *Mol Cell* 77, 108-119.e9

58. Sato, B.K. *et al.* (2009) Misfolded membrane proteins are specifically recognized by the transmembrane domain of the Hrd1p ubiquitin ligase. *Mol Cell* 34, 212–222
59. Habeck, G. *et al.* (2015) The yeast ERAD-C ubiquitin ligase Doa10 recognizes an intramembrane degron. *J Cell Biol* 209, 261–273
60. Schoebel, S. *et al.* (2017) Cryo-EM structure of the protein-conducting ERAD channel Hrd1 in complex with Hrd3. *Nature* 548, 352–355
61. Schulz, J. *et al.* (2017) Conserved cytoplasmic domains promote Hrd1 ubiquitin ligase complex formation for ER-associated degradation (ERAD). *J Cell Sci* 130, 3322–3335
62. Hwang, J. *et al.* (2017) Characterization of protein complexes of the endoplasmic reticulum-associated degradation E3 ubiquitin ligase Hrd1. *J Biol Chem* 292, 9104–9116
63. Neuber, O. *et al.* (2005) Ubx2 links the Cdc48 complex to ER-associated protein degradation. *Nat Cell Biol* 7, 993–998
64. Ballar, P. *et al.* (2006) The role of a novel p97/valosin-containing protein-interacting motif of gp78 in endoplasmic reticulum-associated degradation. *J Biol Chem* 281, 35359–35368
65. Greenblatt, E.J. *et al.* (2011) Derlin-1 is a rhomboid pseudoprotease required for the dislocation of mutant α-1 antitrypsin from the endoplasmic reticulum. *Nat Struct Mol Biol* 18, 1147–1152
66. Albert, S. *et al.* (2020) Direct visualization of degradation microcompartments at the ER membrane. *Proceedings of the National Academy of Sciences* 117, 1069–1080
67. Quan, E.M. *et al.* (2008) Defining the Glycan Destruction Signal for Endoplasmic Reticulum-Associated Degradation. *Mol Cell* 32, 870–877
68. Avezov, E. *et al.* (2008) Endoplasmic Reticulum (ER) Mannosidase I Is Compartmentalized and Required for N-Glycan Trimming to Man5–6GlcNAc2 in Glycoprotein ER-associated Degradation. *Mol Biol Cell* 19, 216–225

69. Hebert, D.N. and Molinari, M. (2012) Flagging and docking: dual roles for N-glycans in protein quality control and cellular proteostasis. *Trends Biochem Sci* 37, 404–410
70. Mehnert, M. *et al.* (2015) The interplay of Hrd3 and the molecular chaperone system ensures efficient degradation of malfolded secretory proteins. *Mol Biol Cell* 26, 185–194
71. Ji, Y. *et al.* (2016) The Sel1L-Hrd1 Endoplasmic Reticulum-Associated Degradation Complex Manages a Key Checkpoint in B Cell Development. *Cell Rep* 16, 2630–2640
72. Leto, D.E. *et al.* (2019) Genome-wide CRISPR analysis identifies substrate-specific conjugation modules in ER-associated degradation. *Mol Cell* 73, 377-389.e11
73. Okuda-Shimizu, Y. and Hendershot, L.M. (2007) Characterization of an ERAD pathway for nonglycosylated BiP substrates, which require Herp. *Mol Cell* 28, 544–554
74. Wiertz, E.J. *et al.* (1996) The human cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the endoplasmic reticulum to the cytosol. *Cell* 84, 769–779
75. Wiertz, E.J.H.J. *et al.* (1996) Sec6l-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction. *Nature* 384, 432–438
76. Hewitt, E.W. (2003) The MHC class I antigen presentation pathway: strategies for viral immune evasion. *Immunology* 110, 163–169
77. van den Boomen, D.J.H. and Lehner, P.J. (2015) Identifying the ERAD ubiquitin E3 ligases for viral and cellular targeting of MHC class I. *Mol Immunol* 68, 106–111
78. Chevalier, M.S. *et al.* (2002) Binding of human cytomegalovirus US2 to major histocompatibility complex class I and II proteins is not sufficient for their degradation. *J Virol* 76, 8265–8275
79. Stagg, H.R. *et al.* (2009) The TRC8 E3 ligase ubiquitinates MHC class I molecules before dislocation from the ER. *J Cell Biol* 186, 685–692

80. van de Weijer, M.L. *et al.* (2014) A high-coverage shRNA screen identifies TMEM129 as an E3 ligase involved in ER-associated protein degradation. *Nat Commun* 5, 3832
81. Luo, J. *et al.* (2020) Mechanisms and regulation of cholesterol homeostasis. *Nat Rev Mol Cell Biol* 21, 225–245
82. Sever, N. *et al.* (2003) Insig-dependent ubiquitination and degradation of mammalian 3-hydroxy-3-methylglutaryl-CoA reductase stimulated by sterols and geranylgeraniol. *J Biol Chem* 278, 52479–52490
83. Sever, N. *et al.* (2003) Accelerated degradation of HMG CoA reductase mediated by binding of insig-1 to its sterol-sensing domain. *Mol Cell* 11, 25–33
84. Menzies, S.A. *et al.* (2018) The sterol-responsive RNF145 E3 ubiquitin ligase mediates the degradation of HMG-CoA reductase together with gp78 and Hrd1. *eLife* 7, e40009
85. Lee, J.N. *et al.* (2006) Sterol-regulated Degradation of Insig-1 Mediated by the Membrane-bound Ubiquitin Ligase gp78*. *Journal of Biological Chemistry* 281, 39308–39315
86. Zelcer, N. *et al.* (2014) The E3 ubiquitin ligase MARCH6 degrades squalene monooxygenase and affects 3-hydroxy-3-methyl-glutaryl coenzyme A reductase and the cholesterol synthesis pathway. *Mol Cell Biol* 34, 1262–1270
87. Gill, S. *et al.* (2011) Cholesterol-dependent degradation of squalene monooxygenase, a control point in cholesterol synthesis beyond HMG-CoA reductase. *Cell Metab* 13, 260–273
88. Sharpe, L.J. *et al.* (2019) Cholesterol increases protein levels of the E3 ligase MARCH6 and thereby stimulates protein degradation. *J Biol Chem* 294, 2436–2448
89. Peterson, B.G. *et al.* (2023) Deep mutational scanning highlights a new role for cytosolic regions in Hrd1 functionbioRxiv, 2023.04.03.535444
90. Mehnert, M. *et al.* (2014) Der1 promotes movement of misfolded proteins through the endoplasmic reticulum membrane. *Nat Cell Biol* 16, 77–86

91. Carvalho, P. *et al.* (2010) Retro-translocation of a misfolded luminal ER protein by the ubiquitin-ligase Hrd1p. *Cell* 143, 579–591
92. Peterson, B.G. *et al.* (2019) Cycles of autoubiquitination and deubiquitination regulate the ERAD ubiquitin ligase Hrd1. *eLife* 8, e50903
93. Vasic, V. *et al.* (2020) Hrd1 forms the retrotranslocation pore regulated by auto-ubiquitination and binding of misfolded proteins. *Nat Cell Biol* 22, 274–281
94. Rao, B. *et al.* (2021) The cryo-EM structure of an ERAD protein channel formed by tetrameric human Derlin-1. *Science Advances* 7, eabe8591
95. Freeman, M. (2014) The Rhomboid-Like Superfamily: Molecular Mechanisms and Biological Roles. *Annual Review of Cell and Developmental Biology* 30, 235–254
96. Adrain, C. and Cavadas, M. (2020) The complex life of rhomboid pseudoproteases. *The FEBS Journal* 287, 4261–4283
97. Goder, V. *et al.* (2008) The ER-associated degradation component Der1p and its homolog Dfm1p are contained in complexes with distinct cofactors of the ATPase Cdc48p. *FEBS Lett* 582, 1575–1580
98. Neal, S. *et al.* (2018) The Dfm1 Derlin Is Required for ERAD Retrotranslocation of Integral Membrane Proteins. *Molecular Cell* 69, 306-320.e4
99. Avci, D. *et al.* (2014) The Yeast ER-Intramembrane Protease Ypf1 Refines Nutrient Sensing by Regulating Transporter Abundance. *Molecular Cell* 56, 630–640
100. Huang, C.-H. *et al.* (2013) Derlin2 Protein Facilitates HRD1-mediated Retro-translocation of Sonic Hedgehog at the Endoplasmic Reticulum. *J Biol Chem* 288, 25330–25339
101. Nejatfard, A. *et al.* (2021) Derlin rhomboid pseudoproteases employ substrate engagement and lipid distortion to enable the retrotranslocation of ERAD membrane substrates. *Cell Rep* 37, 109840
102. Zoll, S. *et al.* (2014) Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate–peptide complex structures. *EMBO J* 33, 2408–2421

103. Wunderle, L. *et al.* (2016) Rhomboid intramembrane protease RHBDL4 triggers ER-export and non-canonical secretion of membrane-anchored TGFα. *Sci Rep* 6, 27342

104. Wang, L. *et al.* (2022) TMUB1 is an endoplasmic reticulum-resident escortase that promotes the p97-mediated extraction of membrane proteins for degradation. *Molecular Cell* 82, 3453-3467.e14

105. van de Weijer, M.L. *et al.* (2020) Quality Control of ER Membrane Proteins by the RNF185/Membralin Ubiquitin Ligase Complex. *Molecular Cell* 79, 768-781.e7

106. Jo, Y. *et al.* (2011) Membrane-associated Ubiquitin Ligase Complex Containing gp78 Mediates Sterol-accelerated Degradation of 3-Hydroxy-3-methylglutaryl-coenzyme A Reductase*. *Journal of Biological Chemistry* 286, 15022–15031

107. The UniProt Consortium (2023) UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research* 51, D523–D531

108. Jiang, L.-L. *et al.* (2019) Membralin deficiency dysregulates astrocytic glutamate homeostasis, leading to ALS-like impairment. *J Clin Invest* 129, 3103–3120

109. Stefanovic-Barrett, S. *et al.* (2018) MARCH6 and TRC8 facilitate the quality control of cytosolic and tail-anchored proteins. *EMBO Rep* 19, e45603

110. Hallgren, J. *et al.* (2022) DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networksbioRxiv, 2022.04.08.487609

111. Elegheert, J. *et al.* (2018) Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins. *Nat Protoc* 13, 2991–3017

112. Lin, Y.-C. *et al.* (2014) Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nat Commun* 5, 4767

113. Reeves, P.J. *et al.* (2002) Structure and function in rhodopsin: High-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable

mammalian cell line. *Proceedings of the National Academy of Sciences* 99, 13419–13424

114. Qin, J.Y. *et al.* (2010) Systematic Comparison of Constitutive Promoters and the Doxycycline-Inducible Promoter. *PLOS ONE* 5, e10611

115. Gibson, D.G. *et al.* (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6, 343–345

116. Porosk, L. *et al.* (2022) Predicting Transiently Expressed Protein Yields: Comparison of Transfection Methods in CHO and HEK293. *Pharmaceutics* 14, 1949

117. Fueller, J. *et al.* (2020) CRISPR-Cas12a–assisted PCR tagging of mammalian genes. *Journal of Cell Biology* 219, e201910210

118. Kremers, G.-J. *et al.* (2006) Cyan and Yellow Super Fluorescent Proteins with Improved Brightness, Protein Folding, and FRET Förster Radius,. *Biochemistry* 45, 6570–6580

119. Shaner, N.C. *et al.* (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol* 22, 1567–1572

120. Parker, J.L. *et al.* (2021) Structural basis of antifolate recognition and transport by PCFT. *Nature* 595, 130–134

121. Schmidpeter, P.A.M. *et al.* (2020) Reconstitution of membrane proteins into platforms suitable for biophysical and structural analyses. *Methods Mol Biol* 2127, 191–205

122. Cube Biotech (2022) Membrane Protein Solubilization Protocol. *Cube Biotech*. [Online]. Available: https://cube-biotech.com/products/membrane-protein-stabilization/synthetic-nanodisc-products/dibma/synthetic-nanodisc-screening-kit-mini-8x2x50-mg/18290. [Accessed: 11-Jul-2023]

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

124. Klug, Y.A. *et al.* (2021) Mechanism of lipid droplet formation by the yeast Sei1/Ldb16 Seipin complex. *Nat Commun* 12, 5892

125. Kühlbrandt, W. (2022) Forty years in cryoEM of membrane proteins. *Microscopy* 71, i30–i50

126. Choy, B.C. *et al.* (2021) A 10-year meta-analysis of membrane protein structural biology: Detergents, membrane mimetics, and structure determination techniques. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1863, 183533

127. Chua, E.Y.D. *et al.* (2022) Better, Faster, Cheaper: Recent Advances in Cryo–Electron Microscopy. *Annual Review of Biochemistry* 91, 1–32

128. Nygaard, R. *et al.* (2020) Cryo-electron microscopy analysis of small membrane proteins. *Current Opinion in Structural Biology* 64, 26–33

129. Wentinck, K. *et al.* (2022) Putting on molecular weight: Enabling cryo-EM structure determination of sub-100-kDa proteins. *Current Research in Structural Biology* 4, 332–337

130. Zhang, T. *et al.* (2015) gp78 functions downstream of Hrd1 to promote degradation of misfolded proteins of the endoplasmic reticulum. *Mol Biol Cell* 26, 4438–4450

131. Das, R. *et al.* (2009) Allosteric Activation of E2-RING Finger Mediated Ubiquitylation by a Structurally-Defined Specific E2 Binding Region of gp78. *Mol Cell* 34, 674–685

132. Liu, W. *et al.* (2014) Gp78 elongates of polyubiquitin chains from the distal end through the cooperation of its G2BR and CUE domains. *Scientific reports* 4, 7138

133. Bagola, K. *et al.* (2013) Ubiquitin binding by a CUE domain regulates ubiquitin chain formation by ERAD E3 ligases. *Mol Cell* 50, 528–539

134. Hülsmann, J. *et al.* (2018) AP-SWATH Reveals Direct Involvement of VCP/p97 in Integrated Stress Response Signaling Through Facilitating CReP/PPP1R15B Degradation. *Mol Cell Proteomics* 17, 1295–1307

135. Camacho, C. *et al.* (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421

136. Jumper, J. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589

137. Cremer, T. *et al.* (2021) The ER-embedded UBE2J1/RNF26 ubiquitylation complex exerts spatiotemporal control over the endolysosomal pathway. *Cell Reports* 34, 108659

138. Jones, D.T. and Cozzetto, D. (2015) DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* 31, 857–863

139. Li, W. *et al.* (2009) Mechanistic insights into active site-associated polyubiquitination by the ubiquitin-conjugating enzyme Ube2g2. *Proc Natl Acad Sci U S A* 106, 3722–3727

140. Du, Z. *et al.* (2021) The trRosetta server for fast and accurate protein structure prediction. *Nat Protoc* 16, 5634–5651

141. Baek, M. *et al.* (2021) Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 373, 871–876

142. Kufareva, I. and Abagyan, R. (2012) Methods of protein structure comparison. *Methods Mol Biol* 857, 231–257

143. Miyamoto, K. *et al.* (2019) Unique RING finger structure from the human HRD1 protein. *Protein Sci* 28, 448–453

144. Zhang, X. *et al.* (2013) Structural basis for role of ring finger protein RNF168 RING domain. *Cell Cycle* 12, 312–321

145. Wiseman, D.N. *et al.* (2020) Expression and purification of recombinant G protein-coupled receptors: A review. *Protein Expression and Purification* 167, 105524

146. Andréll, J. and Tate, C.G. (2013) Overexpression of membrane proteins in mammalian cells for structural studies. *Mol Membr Biol* 30, 52–63

147. Birch, J. and Quigley, A. (2021) The high-throughput production of membrane proteins. *Emerging Topics in Life Sciences* 5, 655–663

148. Johnson, B.M. and DeBose-Boyd, R.A. (2018) Underlying mechanisms for sterol-induced ubiquitination and ER-associated degradation of HMG CoA reductase. *Seminars in Cell & Developmental Biology* 81, 121–128

149. HASSINK, G. *et al.* (2005) TEB4 is a C4HC3 RING finger-containing ubiquitin ligase of the endoplasmic reticulum. *Biochemical Journal* 388, 647–655

150. Sweeney, N.P. and Vink, C.A. (2021) The impact of lentiviral vector genome size and producer cell genomic to gag-pol mRNA ratios on packaging efficiency and titre. *Molecular Therapy - Methods & Clinical Development* 21, 574–584

151. Lee, H.J. *et al.* (2022) Impact of novel detergents on membrane protein studies. *Chem* 8, 980–1013

152. Remans, K. *et al.* (2022) Protein purification strategies must consider downstream applications and individual biological characteristics. *Microbial Cell Factories* 21, 52

153. Lebendiker, M. *et al.* (2015) A Screening Methodology for Purifying Proteins with Aggregation Problems. In *Insoluble Proteins: Methods and Protocols* (García-Fruitós, E., ed), pp. 261–281, Springer

154. Parker, J.L. and Newstead, S. (2016) Membrane Protein Crystallisation: Current Trends and Future Perspectives. In *The Next Generation in Membrane Protein Structure Determination* (Moraes, I., ed), pp. 61–72, Springer International Publishing

155. Vanaken, T. *et al.* (1986) [3] Alkyl glycoside detergents: Synthesis and applications to the study of membrane proteins. In *Methods in Enzymology* 125, pp. 27–35, Academic Press

156. Thompson, A.A. *et al.* (2011) GPCR stabilization using the bicelle-like architecture of mixed sterol-detergent micelles. *Methods* 55, 310–317

157. Anatrace (2018) Cryo-EM is setting records in membrane protein structural biology. It's raising standards, too.

158. Burgess, R.R. (2018) A brief practical review of size exclusion chromatography: Rules of thumb, limitations, and troubleshooting. *Protein Expression and Purification* 150, 81–85

159. Vahedi-Faridi, A. *et al.* (2013) 3D imaging and quantitative analysis of small solubilized membrane proteins and their complexes by transmission electron microscopy. *Microscopy (Tokyo)* 62, 95–107

160. Tang, G. *et al.* (2007) EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology* 157, 38–46

161. Scarff, C.A. *et al.* (2018) Variations on Negative Stain Electron Microscopy Methods: Tools for Tackling Challenging Systems. *J Vis Exp* DOI: 10.3791/57199

162. Caesar, J. *et al.* (2020) SIMPLE 3.0. Stream single-particle cryo-EM analysis in real time. *Journal of Structural Biology: X* 4, 100040

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

164. Scott, N.A. *et al.* (2021) The E3 ubiquitin ligase MARCHF6 as a metabolic integrator in cholesterol synthesis and beyond. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 1866, 158837

165. Punjani, A. *et al.* (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14, 290–296

166. Zivanov, J. *et al.* (2018) New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* 7, e42166

167. Asarnow, D. *et al.* (2019) asarnow/pyem: UCSF pyem v0.5Zenodo

168. Tan, Y.Z. *et al.* (2017) Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat Methods* 14, 793–796

169. Sanchez-Garcia, R. *et al.* (2021) DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun Biol* 4, 1–8

170. Pettersen, E.F. *et al.* (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* 30, 70–82

171. Brown, A. *et al.* (2015) Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr D Biol Crystallogr* 71, 136–153

172. Afonine, P.V. *et al.* (2018) Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Cryst D* 74, 531–544

173. Hopf, T.A. *et al.* (2019) The EVcouplings Python framework for coevolutionary sequence analysis. *Bioinformatics* 35, 1582–1584

174. Yariv, B. *et al.* (2023) Using evolutionary data to make sense of macromolecules with a “face-lifted” ConSurf. *Protein Science* 32, e4582

175. Rana, M.S. *et al.* (2018) An improved strategy for fluorescent tagging of membrane proteins for overexpression and purification in mammalian cells. *Biochemistry* 57, 6741–6751

176. Campbell, B.C. *et al.* (2022) Chemically stable fluorescent proteins for advanced microscopy. *Nat Methods* 19, 1612–1621

177. Hattori, T. *et al.* (2021) SEL1L degradation intermediates stimulate cytosolic aggregation of polyglutamine-expanded protein. *The FEBS Journal* 288, 4637–4654

178. Jeong, H. *et al.* (2016) Crystal structure of SEL1L: Insight into the roles of SLR motifs in ERAD pathway. *Sci Rep* 6, 20261

179. Eura, Y. *et al.* (2020) Derlin-3 Is Required for Changes in ERAD Complex Formation under ER Stress. *Int J Mol Sci* 21, 6146

180. Horn, S.C. *et al.* (2009) Usa1 Functions as a Scaffold of the HRD-Ubiquitin Ligase. *Molecular Cell* 36, 782–793

181. Kokame, K. *et al.* (2000) Herp, a New Ubiquitin-like Membrane Protein Induced by Endoplasmic Reticulum Stress *. *Journal of Biological Chemistry* 275, 32846–32853

182. Tang, H.-Y. *et al.* (2014) EDEM2 and OS-9 Are Required for ER-Associated Degradation of Non-Glycosylated Sonic Hedgehog. *PLoS One* 9, e92164

183. Goot, A.T. van der *et al.* (2018) Redundant and antagonistic roles of XTP3B and OS9 in decoding glycan and non-glycan degrons in ER-associated degradation. *Molecular cell* 70, 516

184. Iida, Y. *et al.* (2011) SEL1L protein critically determines the stability of the HRD1-SEL1L endoplasmic reticulum-associated degradation (ERAD) complex to optimize the degradation kinetics of ERAD substrates. *J Biol Chem* 286, 16929–16939

185. Mueller, B. *et al.* (2008) SEL1L nucleates a protein complex required for dislocation of misfolded glycoproteins. *Proc Natl Acad Sci U S A* 105, 12325–12330

186. Hosokawa, N. *et al.* (2008) Human XTP3-B Forms an Endoplasmic Reticulum Quality Control Scaffold with the HRD1-SEL1L Ubiquitin Ligase Complex and BiP. *J Biol Chem* 283, 20914–20924

187. Punjani, A. *et al.* (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14, 290–296

188. Anti-Derlin-2 pAb (Polyclonal Antibody) (DISCONTINUED) - MBL International. *Products*. [Online]. Available: https://www.mblintl.com/products/pm019/. [Accessed: 28-Jun-2023]

189. Sievers, F. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7, 539

190. D'Imprima, E. *et al.* (2019) Protein denaturation at the air-water interface and how to prevent it. *eLife* 8, e42747

191. Zhang, Z. *et al.* (2021) Improving particle quality in cryo-EM analysis using a PEGylation method. *Structure* 29, 1192-1199.e4

192. Huppa, J.B. and Ploegh, H.L. (1997) The α Chain of the T Cell Antigen Receptor Is Degraded in the Cytosol. *Immunity* 7, 113–122

193. (1989) Pre-Golgi degradation of newly synthesized T-cell antigen receptor chains: intrinsic sensitivity and the role of subunit assembly. *J Cell Biol* 109, 73–83

194. Bonifacino, J.S. *et al.* (1990) Colocalized transmembrane determinants for ER degradation and subunit assembly explain the intracellular fate of TCR chains. *Cell* 63, 503–513

195. Soetandyo, N. *et al.* (2010) Role of intramembrane charged residues in the quality control of unassembled T-cell receptor α-chains at the endoplasmic reticulum. *Journal of Cell Science* 123, 1031

196. Wu, S. *et al.* (2012) Fabs enable single particle cryoEM studies of small proteins. *Structure* 20, 582–592

197. Carpenter, E.P. *et al.* (2008) Overcoming the challenges of membrane protein crystallography. *Current Opinion in Structural Biology* 18, 581–586

198. Ereño-Orbea, J. *et al.* (2018) Structural Basis of Enhanced Crystallizability Induced by a Molecular Chaperone for Antibody Antigen-Binding Fragments. *Journal of Molecular Biology* 430, 322–336

199. Stura, E.A. *et al.* (1993) Crystallization of Antibodies and Antibody-Antigen Complexes. *ImmunoMethods* 3, 164–179

200. van de Bovenkamp, F.S. *et al.* (2016) The Emerging Importance of IgG Fab Glycosylation in Immunity. *The Journal of Immunology* 196, 1435–1441

201. Herzik, M.A. *et al.* (2019) High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. *Nat Commun* 10, 1032

202. Kim, J. *et al.* (2019) Structure and drug resistance of the Plasmodium falciparum transporter PfCRT. *Nature* 576, 315–320

203. Hamers-Casterman, C. *et al.* (1993) Naturally occurring antibodies devoid of light chains. *Nature* 363, 446–448

204. Muyldermans, S. (2013) Nanobodies: Natural Single-Domain Antibodies. *Annual Review of Biochemistry* 82, 775–797

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

206. McMahon, C. *et al.* (2018) Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat Struct Mol Biol* 25, 289–296

207. Muyldermans, S. (2021) A guide to: generation and design of nanobodies. *The FEBS Journal* 288, 2084–2102

208. Fowler, M.E. (1998) Medicine and surgery of South American camelids: llama, alpaca, vicuna, guanaco. *Medicine and surgery of South American camelids: llama, alpaca, vicuna, guanaco.* at <https://www.cabdirect.org/cabdirect/abstract/19982213210>

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

210. Rasmussen, S.G.F. *et al.* (2011) Structure of a nanobody-stabilized active state of the β2 adrenoceptor. *Nature* 469, 175–180

211. Tiller, K.E. and Tessier, P.M. (2015) Advances in Antibody Design. *Annual Review of Biomedical Engineering* 17, 191–216

212. Yang, E.Y. and Shah, K. (2020) Nanobodies: Next Generation of Cancer Diagnostics and Therapeutics. *Frontiers in Oncology* 10

213. Seeger, M.A. *et al.* (2013) Design, construction, and characterization of a second-generation DARPin library with reduced hydrophobicity. *Protein Science* 22, 1239–1257

214. Asaadi, Y. *et al.* (2021) A comprehensive comparison between camelid nanobodies and single chain variable fragments. *Biomarker Research* 9, 87

215. Kirchhofer, A. *et al.* (2010) Modulation of protein properties in living cells using nanobodies. *Nat Struct Mol Biol* 17, 133–138

216. De Genst, E. *et al.* (2006) Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proceedings of the National Academy of Sciences* 103, 4586–4591

217. Pardon, E. *et al.* (2014) A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc* 9, 674–693

218. Rigaud, J.-L. and Lévy, D. (2003) Reconstitution of Membrane Proteins into Liposomes. In *Methods in Enzymology* 372, pp. 65–86, Academic Press

219. Seddon, A.M. *et al.* (2004) Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1666, 105–117

220. Crouch, C.H. *et al.* (2018) Optimization of Detergent-Mediated Reconstitution of Influenza A M2 Protein into Proteoliposomes. *Membranes (Basel)* 8, 103

221. Kraft, T.E. *et al.* (2015) Expression, purification, and functional characterization of the insulin-responsive facilitative glucose transporter GLUT4. *Protein Sci* 24, 2008–2019

222. Amati, A.M. *et al.* (2020) Current problems and future avenues in proteoliposome research. *Biochemical Society Transactions* 48, 1473–1492

223. Geertsma, E.R. *et al.* (2008) Membrane reconstitution of ABC transporters and assays of translocator function. *Nat Protoc* 3, 256–266

224. Fairhead, M. and Howarth, M. (2015) Site-specific biotinylation of purified proteins using BirA. *Methods Mol Biol* 1266, 171–184

225. Monk, B.C. *et al.* (2014) Architecture of a single membrane spanning cytochrome P450 suggests constraints that orient the catalytic domain relative to a bilayer. *Proc Natl Acad Sci U S A* 111, 3865–3870

226. McGuffin, L.J. *et al.* (2000) The PSIPRED protein structure prediction server. *Bioinformatics* 16, 404–405

227. Yang, B. *et al.* (2015) The critical role of membralin in postnatal motor neuron survival and disease. *eLife* 4, e06500

228. Qiu, D. *et al.* (2018) RNF185 modulates JWA ubiquitination and promotes gastric cancer metastasis. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1864, 1552–1561

229. El Khouri, E. *et al.* (2013) RNF185 Is a Novel E3 Ligase of Endoplasmic Reticulum-associated Degradation (ERAD) That Targets Cystic Fibrosis Transmembrane Conductance Regulator (CFTR). *J Biol Chem* 288, 31177–31191

230. van Kempen, M. *et al.* (2023) Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* DOI: 10.1038/s41587-023-01773-0

231. Noble, A.J. *et al.* (2018) Routine single particle CryoEM sample and grid characterization by tomography. *eLife* 7, e34257

232. Sgro, G.G. and Costa, T.R.D. (2018) Cryo-EM Grid Preparation of Membrane Protein Samples for Single Particle Analysis. *Frontiers in Molecular Biosciences* 5

233. Popot, J.-L. *et al.* (2011) Amphipols From A to Z. *Annual Review of Biophysics* 40, 379–408

234. Gohon, Y. *et al.* (2008) Bacteriorhodopsin/Amphipol Complexes: Structural and Functional Properties. *Biophysical Journal* 94, 3523–3537

235. Popot, J.-L. *et al.* (2003) Amphipols: polymeric surfactants for membrane biology research. *Cell Mol Life Sci* 60, 1559–1574

236. Autzen, H.E. *et al.* (2019) Membrane Mimetic Systems in CryoEM: Keeping Membrane Proteins in their Native Environment. *Curr Opin Struct Biol* 58, 259–268

237. Champeil, P. *et al.* (2000) Interaction of Amphipols with Sarcoplasmic Reticulum Ca2+-ATPase *. *Journal of Biological Chemistry* 275, 18623–18637

238. Picard, M. *et al.* (2006) Protective and Inhibitory Effects of Various Types of Amphipols on the Ca2+-ATPase from Sarcoplasmic Reticulum: A Comparative Study. *Biochemistry* 45, 1861–1869

239. Parmar, M. *et al.* (2018) Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1860, 378–383

240. Sun, C. *et al.* (2018) Structure of the alternative complex III in a supercomplex with cytochrome oxidase. *Nature* 557, 123–126

241. Qiu, W. *et al.* (2018) Structure and activity of lipid bilayer within a membrane-protein transporter. *Proc Natl Acad Sci U S A* 115, 12985–12990

242. Dörr, J.M. *et al.* (2016) The styrene–maleic acid copolymer: a versatile tool in membrane research. *Eur Biophys J* 45, 3–21

243. Oluwole, A.O. *et al.* (2017) Formation of Lipid-Bilayer Nanodiscs by Diisobutylene/Maleic Acid (DIBMA) Copolymer. *Langmuir* 33, 14378–14388

244. Krissinel, E. (2007) On the relationship between sequence and structure similarities in proteomics. *Bioinformatics* 23, 717–723

245. Meruelo, A.D. *et al.* (2012) Structural differences between thermophilic and mesophilic membrane proteins. *Protein Sci* 21, 1746–1753

246. Baker, M.R. *et al.* (2021) Cryo-EM structure of type 1 IP3R channel in a lipid bilayer. *Commun Biol* 4, 1–11

247. Maeda, S. *et al.* (2020) Structure, lipid scrambling activity and role in autophagosome formation of ATG9A. *Nat Struct Mol Biol* 27, 1194–1201

248. Li, R.W. and Li, C. (2006) Butyrate induces profound changes in gene expression related to multiple signal pathways in bovine kidney epithelial cells. *BMC Genomics* 7, 234

249. Baldridge, R.D. and Rapoport, T.A. (2016) Autoubiquitination of the Hrd1 Ligase Triggers Protein Retrotranslocation in ERAD. *Cell* 166, 394–407

250. Assainar, B.M. *et al.* (2023) Direct observation of autoubiquitination for an integral membrane ubiquitin ligase in ERADbioRxiv, 2023.06.20.545802

251. Harada, K. *et al.* (2016) USP19-Mediated Deubiquitination Facilitates the Stabilization of HRD1 Ubiquitin Ligase. *International Journal of Molecular Sciences* 17, 1829

252. Avci, D. *et al.* (2023) Intramembrane protease SPP defines a cholesterol-regulated abundance control of the mevalonate pathway enzyme SQSbioRxiv, 2021.07.19.452877

10. Supplementary Data

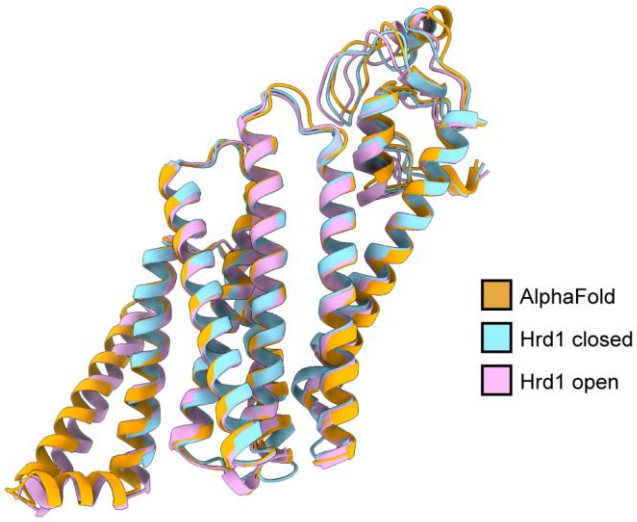


Figure S1. Alignment of the final models for the Hrd1 open and closed conformations to the AlphaFold predicted model. Models were aligned against the AlphaFold model using ChimeraX. Hrd1 closed RMSD 1.50 Å. Hrd1 open RMSD 1.90 Å.

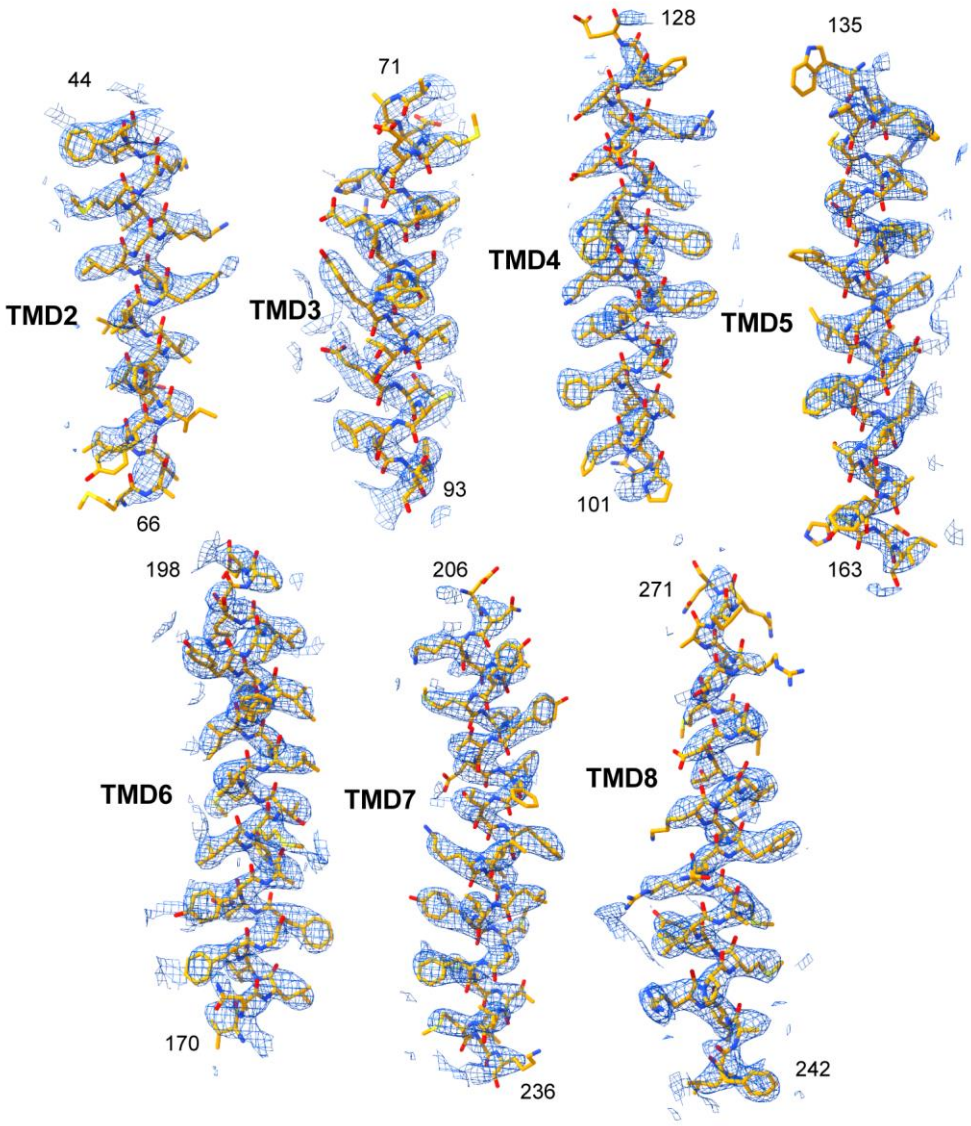


Figure S2. Close-up view of the globally-sharpened Hrd1 closed map and side-chain density for TM helices. Volume was contoured to a threshold level of 0.765.

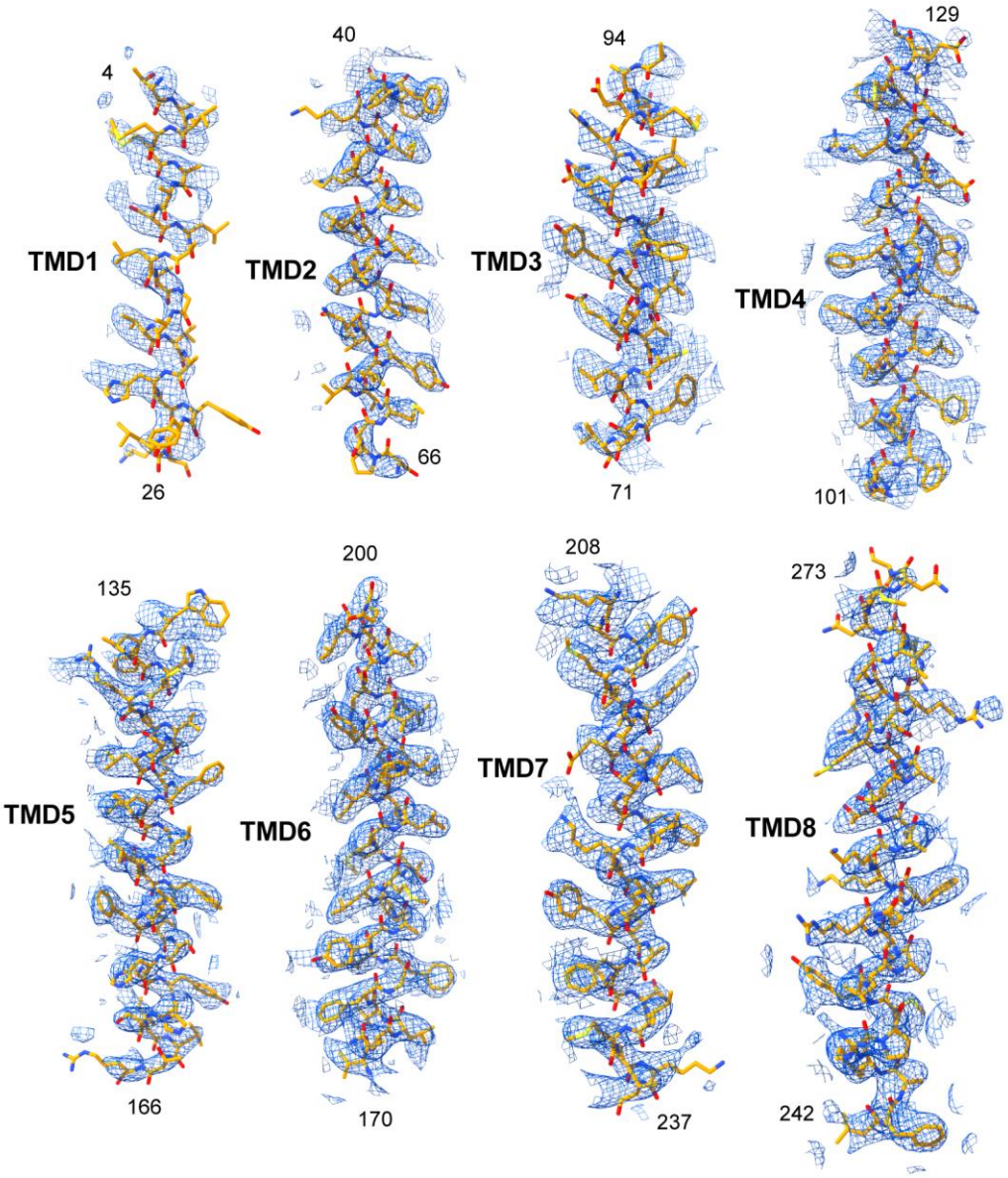


Figure S3. Close-up view of the globally-sharpened Hrd1 open map and side-chain density for TM helices. Volume was contoured at a threshold level of 0.386 for TMD1 and 0.528 for the other TMDs.

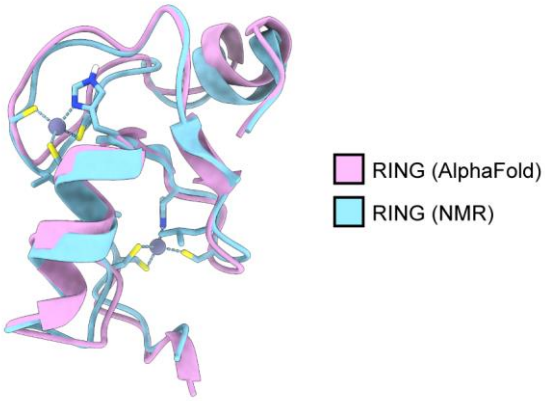


Figure S4. Alignment of the NMR structure of the Hrd1 RING domain with the AlphaFold predicted structure [136,143]. AlphaFold predicted residues 279 to 344 of Hrd1 were aligned with the NMR structure. Cα alignment was performed using ChimeraX and resulted in an RMSD of 1.12 Å [170].

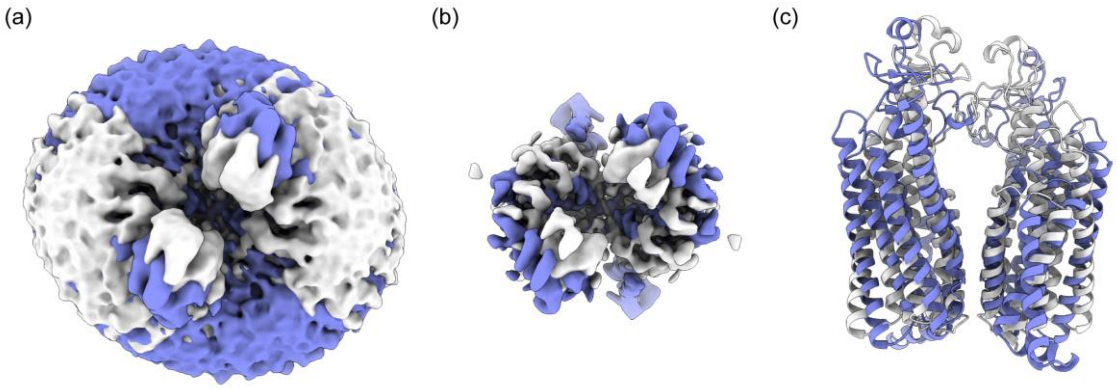


Figure S5. Overlays of the Hrd1 open and closed states. (a) Top-down view of low contour maps for the open and closed state. The open state is in blue and the closed state in white. The orientation of the open and closed maps was the same and so can be overlayed with each other without adjustment. (b) Top-down view of high contour versions of the maps in (a). (c) Side-view of models for the open and closed states overlayed without adjustment.

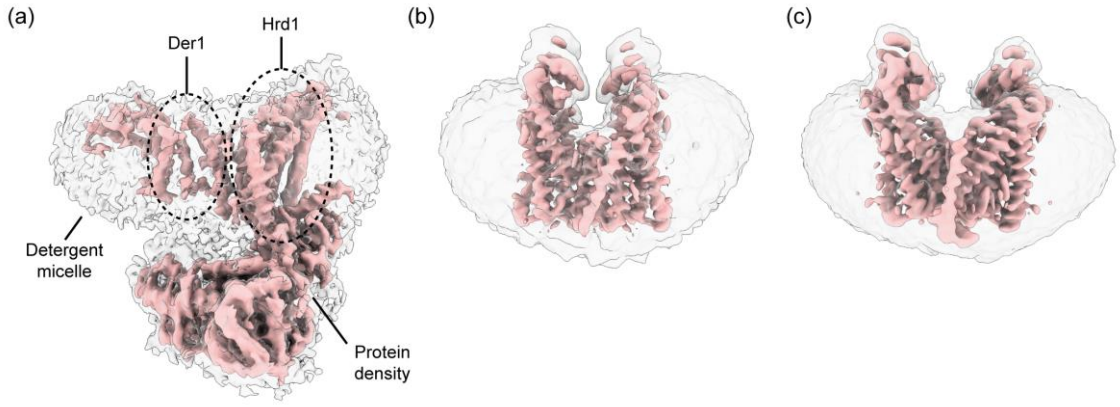


Figure S6. Comparison of the membrane thinning in the Hrd1 open and closed states with the yeast Hrd1-Der1 heterodimer. (a) Cryo-EM map of the yeast Hrd1-Der1 heterodimer in complex with Usa1 and Hrd3 [55]. Protein density is denoted by the high contour map in pink and the detergent micelle by the overlayed low contour map. High contour map at 0.0289 cut-off, low contour map at 0.0114 cut-off. (b) Cryo-EM map of the mammalian Hrd1 close state. High contour map at 0.225 cut-off, low contour map at 0.0667 cut-off. (c) Cryo-EM map of the mammalian Hrd1 open state. High contour map at 0.184 cut-off, low contour map at 0.0606 cut-off.

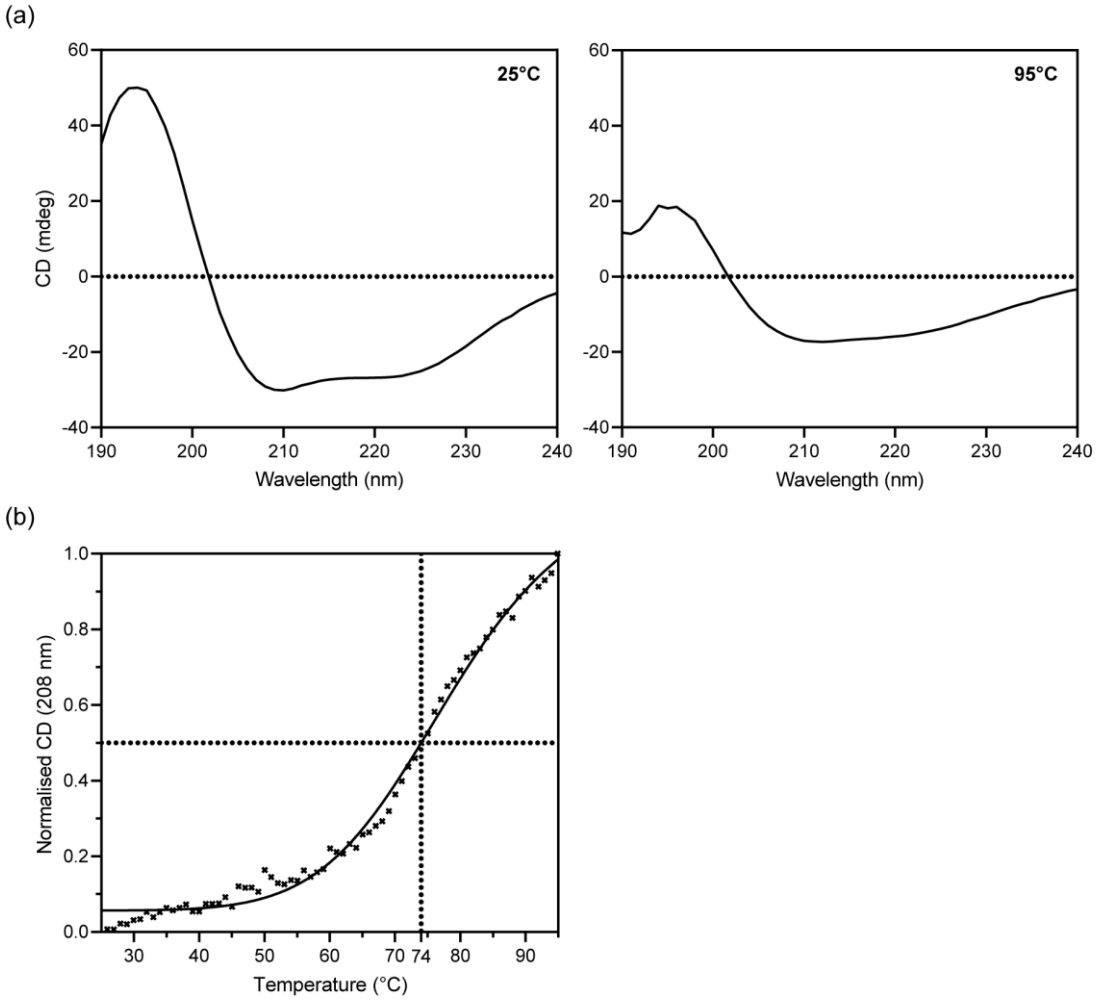


Figure S7. Circular-dichroism (CD) melting profile for Hrd1. (a) Full-spectrum CD profiles for Hrd1 at 25°C and 95°C. Full melting was not observed for Hrd1 at 95°C. (b) CD-melt profile for Hrd1 in DDM-CHS (10:1), recorded at 208 nm. The inflection point is marked at 74°C.

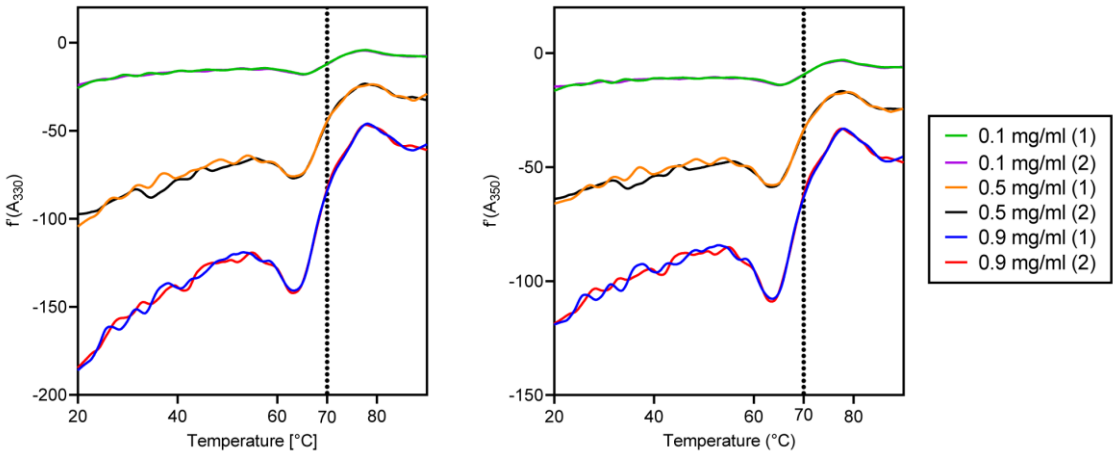


Figure S8. Nano-differential scanning fluorimetry (nanoDSF) melting profiles for Hrd1. Tyrosine and tryptophan fluorescence were monitored at 330 nm and 350 nm, respectively. (a) Change in Hrd1 fluorescence at 330 nm with temperature. (b) Change in Hrd1 fluorescence at 350 nm with temperature. The ratio of fluorescence intensities did not change but at both wavelengths, an inflection point was observed in the first derivatives of both wavelengths, indicating a change in tryptophan and tyrosine residue environments.