Investigations into the incorporation of GlpG rhomboid protease into nanodiscs for solution-state NMR

Brittany Semotiuk

Thesis submitted to the University of Ottawaa
in partial fulfillment of the requirements
for the Master of Science degree in Chemistry

Department of Chemistry and Biomolecular Sciences
Faculty of Science
University of Ottawa

Abstract

Rhomboids are intramembrane serine proteases that cleave transmembrane (TM) protein substrates within the phospholipid bilayer. Since the discovery of the first rhomboid protease, many homologous rhomboids have been identified in all kingdoms illustrating their biological significance. Rhomboids are key players in a variety of biological processes such as, cell signalling, protein degradation, mitochondria health, apoptosis, and pathogenicity. While the mechanism of substrate entry into the rhomboid active site is still not clear, it is thought to involve dynamics around the putative substrate gate, of which appears to be comprised of the fifth transmembrane α-helix. A powerful tool that can be used to investigate conformational dynamics around the substrate gate is solution-state nuclear magnetic resonance (NMR). However, due to the size restriction of solution-state NMR, only detergent micelles have been able to produce well-resolved ^{1}H-^{15}N HSQC spectra of rhomboids. However, the lipid membrane environment has a significant impact on rhomboid structure and function. The use of membrane-scaffolding proteins (MSPs) in the formation of nanodiscs has the potential to allow the study of rhomboid dynamics in lipid bilayers by solution-state NMR. Therefore, this thesis investigates the plausibility of incorporating rhomboid into nanodiscs that would be compatible with solution NMR with a focus on the *E. coli* rhomboid, *ec*GlpG. The formation of empty (no *ec*GlpG) and *ec*GlpG-encapsulated nanodiscs was attempted using two MSP variants. While some successful nanodisc formation was possible, MSP degradation and low yields were seen for all nanodisc samples. Further optimization or alternate nanodisc systems will be required to incorporate *ec*GlpG into more membrane-like environments in a state that is compatible with solution-state NMR.

Acknowledgements

First and foremost, I want to give my immense thanks to my supervisor, Dr. Natalie Goto, for welcoming me into her lab and imparting me with a vast array of knowledge and technical skills. I would like to thank her for the support she has always provided and the patience she has had in reviewing presentations and any written documents, such as this thesis. I also want to declare my appreciation for her understanding and caring nature, and the time she has dedicated to helping me understand more challenging topics, such as NMR. As my background is not directly in Chemistry, I am grateful for her belief in my abilities and the helping hand she lends whenever necessary. It has truly been a pleasure, and I look forward to continuing to learn from and work with her during my doctorate studies.

I am also thankful for the support and valuable feedback provided by my thesis committee, Dr. Roberto Chica and Dr. Corrie DaCosta. I would like to acknowledge Dr. Roberto Chica, Dr. Adam Damry, and Dr. John Pezacki who generously allowed use of their instruments as well as the University of Ottawa NMR facility manager, Dr. Patrick Szell, and the mass spectrometry manager, Dr. Sharon Curtis.

Additionally, I would like to thank the Goto lab, specifically to Dr. James Davey for the helpful guidance that he gave me at the beginning of my research project.

Finally, I would like to thank my family. Firstly, to my parents, Olivia, and Mark, who I thank for raising me to be who I am today and for continuing to show their support and love no matter the circumstance. I can truly say that I couldn't have done this without them. Secondly, to my partner, Troy, who has always given me love, support, and motivation to push through even the hardest times, I express my deepest thanks. Lastly, while they may not be able to read this, I

give my thanks and love to our two dogs, Einstein, and Loki, who no matter what always welcome me with love and affection.

Table of contents

List of abbreviations

ΔH5	MSP1D1ΔH5
ΔH5ΔHis	MSP1D1ΔH5 with the 6X His-tag removed
1D1	MSP1D1
1D1ΔHis	MSP1D1 with the 6X His-tag removed
4-HCCA	α-cyano-4-hydroxycinnamic acid
6X	C-terminal hexahistidine tag
AarA	*Providencia stuartii* rhomboid protease
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride
Apo-A1	Apolipoprotein A1
APP	Amyloid precursor protein
BCA	Bicinchoninic acid
bR	Bacteriorhodopsin
BSA	Bovine serum albumin
CBB	Coomassie brilliant blue
CD	Circular dichroism
cNDs	Covalently circularized nanodiscs
cND9	Covalently circularized nanodiscs with diameter of 9 nm
cND11	Covalently circularized nanodiscs with diameter of 11 nm
cytb5	Cytochrome-b5
CytD	Cytoplasmic domain
DDM	n-dodecyl β-D-maltoside

DLPC	1,2-dilauroyl-sn-glycero-3-phosphocholine
DLS	Dynamic light scattering
DMPA	1,2-dimyristoyl-sn-glycero-3-phosphate (sodium salt)
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DMPE	1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine
DMPG	1,2-dimyristoyl-sn-glycero-3-phospho-(1-rac-glycerol) (sodium salt)
DPC	n-dodecylphosphocholine
DTT	Dithiothreitol
***ec*GlpG**	*Escherichia coli* rhomboid protease
EDTA	2,2',2'',2'''-(ethane-1,2-diyldinitrilo) tetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
***Eh*ROM1**	*E. histolytica* rhomboid protease
EM	Electron microscopy
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum-Associated degradation
ExPEC	Extraintestinal Pathogenic *E. coli*
FPLC	Fast protein liquid chromatography
GlpR	Glycerol-3-phosphate regulon repressor
GPCR	G-protein coupled receptor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HSQC	Heteronuclear single quantum coherence
IMM	Inner mitochondrial membrane

IMP	Intramembrane protease
IMS	Intramembrane space
IPTG	Isopropyl B-D-thiogalactopyranoside
kDa	Kilodalton
L1	Loop 1 regions
L5	Loop 5 regions
LB	Luria-Bertani broth
LY2	Lactose permease
Lyso-Fos-12	Lysolaurylphosphocholine
Lyso-Fos-14	Lysomyristoylphosphocholine
M2M	1,2-bis(methylsulfonyl-sulfonyl) ethane
Mal-12	n-dodecyl-β-D-maltoside
MALDI-TOF	Matrix-associated laser desorption/ionization-time of flight
MIC2	Micronemal protein 2
MS	Mass spectrometry
MSP	Membrane scaffolding protein
MW	Molecular weight
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
OD	Optical density
OMM	Outer mitochondrial membrane
OmpX	Bacterial outer membrane protein X
PARL	Presenilin-associated rhomboid-like protease

PC	Phosphatidylcholine
PDB	Protein database (RCSB)
PDC	Protein-detergent complex
***pf*Rom1**	*P. falciparum* rhomboid 1
***pf*Rom4**	*P. falciparum* rhomboid 4
PINK1	(PTEN)-induced putative kinase 1
POT	Proton-coupled oligopeptide transporter
RHBDL2	Human rhomboid related protein 2
RHBDL4	Human rhomboid related protein 4
Rhom-1	*Drosophila melanogaster* Rhomboid-1 protease
Rho-7	*Drosophila melanogaster* mitochondrial rhomboid 7
RT	Room temperature
SapA	Saposin-A
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SMA	Styrene-maleic acid
SMALPs	Styrene-maleic acid lipoparticles
ssNMR	Solid-state NMR
STARD7	START domain-containing protein
Tat	Twin arginine translocase
TatA	Twin arginine translocase, component A
TatB	Twin arginine translocase, component B
TatC	Twin arginine translocase, component C

TFA	Trifluoroacetic acid
TM	Transmembrane
TMD	Transmembrane domain
TMD81	Residues 81-276 of *ec*GlpG
TRAP	Thrombospondin related anonymous protein
TRIS	2-amino-2-(hydroxymethyl-propane-1,3-diol
UIM	Ubiquitin interacting motif
WT	Wild type

List of figures

List of tables

Chapter 1: Introduction

1.1 Intramembrane proteases

The cell membrane is a hydrophobic barrier between the internal and external environment of a cell and permits a variety of crucial functions that aid in cell survival through the actions of various proteins.[1–3] Intramembrane proteases (IMPs) are a family of proteins that have multi-spanning transmembrane domains (TMDs) that catalyze the hydrolytic cleavage of peptide bonds in both solvent-exposed and buried regions of transmembrane (TM) sequences.[3] While this initially raised questions about how proteolysis could occur within a membrane environment as the presence of water molecules is limited,[4] crystal structures have revealed water molecules in the catalytic active site.[5] IMPs are involved in a vast array of biological roles which include but are not limited to the propagation of cell signalling cascades, and the processing or degradation of alternative integral membrane proteins.[3] Four different classes of IMPs have been identified since their initial discovery around 25 years ago: metallo-protease IMPs, aspartyl IMPs, glutamyl IMPs, and rhomboid proteases which will be the topic of this thesis.[6–9]

1.2 Rhomboid protease

Rhomboid proteases are serine intramembrane proteases that cleave transmembrane substrates at the N-terminal side of a transmembrane helix which contribute to a wide range of biological processes including cell signaling, protein degradation, apoptosis, and pathogenicity.[3,9–23] Eukaryotic and prokaryotic proteases have been identified in all kingdoms of life, including humans and *Escherichia coli*, the latter of which will be the focus of this thesis.[9,16,24,25] Cleavage of the substrate may produce a ligand for downstream signaling, change the functionality of a protein, or initiate degradation depending on the biological function of the rhomboid protease.

Regardless, based on sequence homology the mechanism of the catalysis is thought to remain the same across the family. Most of what is known about rhomboid protease structure and function is based on studies of the *E. coli* rhomboid protease, GlpG, in detergent micelle. [4,26–29] Although micelles are convenient for solution-phase biophysical experiments, they do not accurately represent all characteristics of the native lipid environment (discussed in 1.6.1). This thesis aims to determine whether the *E. coli* rhomboid serine protease, *ec*GlpG, function can be maintained and studied through solution-state nuclear magnetic resonance (NMR) within a more native lipid bilayer environment such as that encapsulated in nanodiscs (discussed in more detail in 1.6.4).

1.3 Biological function of rhomboids

1.3.1 Cell signalling

The first rhomboid protease to be discovered was Rhomboid-1 in *Drosophila melanogaster* which was so named due to a characteristic morphology of the head in *Drosophila* embryos with mutations in this gene.[17] Rhomboid-1 catalyzes the cleavage of the TM segment of Spitz, resulting in the secretion of Epidermal Growth Factor (EGF).[9,16] EGF acts as a signalling molecule that binds to the EGF receptor (EGFR) in *Drosophila* leading to the regulation of cell growth and development (Figure 1.1A). Subsequent rhomboid homologs were identified that also participate in the *Drosophila* EGF signalling cascade (Rhomboids 2-4),[30,31] as well as mammalian rhomboids associated with EGFR signalling (RHBDL2 and RHBDL4).[18,19]

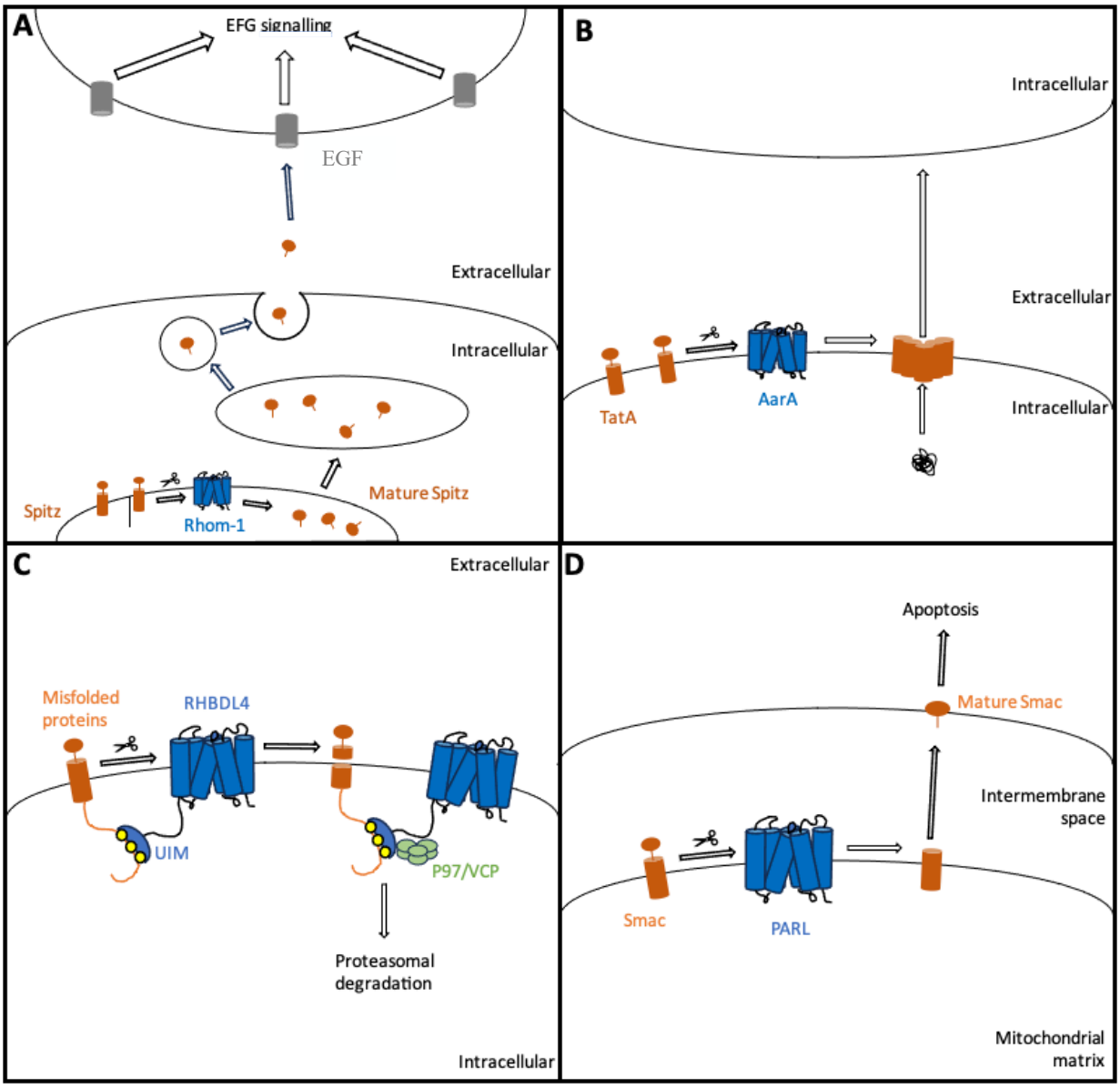

Figure 1.1: Biological role of various rhomboid proteases. (A) Cleavage of the Spitz TM helix (orange) by Rhomboid-1 (blue) releases EGF to bind EGFR and trigger downstream signalling events. (B) AarA (blue) cleavage of TatA (orange) allowing for the formation of the twin arginine translocation complex responsible for the secretion of a signalling molecule (black) critical for quorum sensing. (C) The recognition of ubiquitinated, misfolded proteins (orange) by the Ubiquitin Interacting Motifs (UIM) on RHBDL4 (blue) leading to the recruitment of p97/VCP that results in their degradation by ERAD. (D) The recognition and cleavage of Smac (orange) by PARL (blue) resulting in the translocation of the signalling molecule, mature Smac, to the outer mitochondrial membrane where it facilitates the initiation of apoptosis.

An additional rhomboid protease that is involved in cell signalling to facilitate quorum sensing and biofilm formation is the prokaryotic rhomboid, AarA, in *Providencia stuartii*;

however, the signalling mechanism differs significantly from rhomboids in eukaryotes. AarA cleaves the N-terminal inhibitory sequence on the twin-arginine translocase component A (TatA) permitting its oligomerization to form a membrane pore that is part of the twin-arginine translocation complex with TatB and TatC (Figure 1.1B).[20,21,32,33] While this cascade is known to result in the export of virulence factors that lead to quorum sensing and biofilm formation, the signalling molecule responsible for this activity has yet to be identified.

1.3.2 Protein Degradation

While the human rhomboid, RHBDL4, plays a role in cell signalling as mentioned previously, the primary role of this rhomboid is the degradation of misfolded proteins that are associated with the Endoplasmic Reticulum-Associated Degradation (ERAD) pathway.[3,22,34] RHBDL4 is located within the ER and binds ubiquitinated TM proteins resulting in cleavage of the substrate and the recruitment of the p97 AAA+ ATPase to promote the proteasomal degradation of the resulting fragments (Figure 1.1C).[22,23,34] In addition, RHBDL4 has been shown to intercept the amyloid precursor protein (APP), preventing its transport to the cell membrane where it can form the amyloid plaques that have been implicated in Alzheimer's disease.[35,36] It does this by cleaving APP into two non-amylogenic fragments, reducing the amount of APP that can form amyloid plaques.[35] While there is uncertainty on whether the human rhomboid uses the ERAD pathway to degrade APP, these results illustrate a potential role for RHBDL4 to inhibit or delay onset of APP associated diseases such as Alzheimer's.

1.3.3 Mitochondrial rhomboids and apoptosis

Presenilin-associated rhomboid-like (PARL) protein, a mitochondrial rhomboid located in the IMM, has been associated with diseases such as insulin resistance, type 2 diabetes[10] and Parkinson's.[37,38] The first substrate of PARL to be identified was the phosphatase and tensin homolog (PTEN)-induced putative kinase 1 (PINK1) which activates signalling pathways important for maintaining mitochondria quality and health.[11,37] In healthy mitochondria, PINK1 in the IMM by is cleaved by PARL, while in mitochondria suffering from oxidative stress,[39] PINK1 is not processed and builds up within the outer mitochondrial membrane (OMM).[37] The accumulation of full-length PINK1 causes phosphorylation that recruits E3 ubiquitin ligase Parkin leading to mitophagy/apoptosis of the cell.[40,41] Inhibition of this pathway through the use of PARL-specific ketoamide inhibitors leads to an increase in mitophagy, illustrating a possible therapeutic approach for treatment of Parkinson's disease.[40]

Another substrate of PARL is the START domain-containing protein, STARD7, which is important for the accumulation of phosphatidylcholine (PC) needed in the IMM for respiration and cristae morphogenesis.[12] STARD7 cleavage results in two fragments, one that moves to the cytosol while the other travels to the intramembrane space (IMS). A lower level of STARD7 cleavage products within the IMS prevents the transfer of PC across the IMS, giving rise to respiration and cristae development deficiencies.[12]

In addition to PINK1 and STARD7, PARL has also been shown to cleave mitochondrial protein Smac in apoptotic cells. Cleavage of Smac results in its release from the IMM into the cytosol where the mature Smac fragment binds to and inhibits inhibitors of apoptosis (Figure 1.1D).[42] While this pro-apoptotic mechanism seems to be a secondary method to allow the release of cytochrome c,[42] increasing understanding of the role of PARL in maintaining mitochondrial

health enhances the possibility that novel therapeutic strategies for acute asthma, and impaired mitochondrial ATP production could be developed in the future.

1.3.4 Role in pathogenicity

Rhomboid proteases have been found to play a role in pathogenicity in a vast array of species which include but are not limited to: *Plasmodium falciparum, Trichomonas vaginalis, Toxoplasma gondii, Entamoeba histolytica,* and *Aspergillus fumigatus*.[13–15,43–46]

The *T. gondii* rhomboid protease ROM4 is important for invasion of the host cell through the cleavage of its substrate, micronemal protein 2 (MIC2).[14,47] The ROM4 knock-out gives rise to accumulation of MIC2 on the surface that impairs internalisation of *T. gondii* into its host cell.[14] Similarly, the rhomboid protease from *E. histolytica*, *Eh*ROM1, plays a critical role in adhesion and phagocytosis of human red blood cells by this parasite.[44] *Eh*ROM1 knock-out parasites showed a decrease in cytotoxicity, directional and non-directional migration, and hemolytic activity which suggests multiple substrates are targeted by this protease.[44] A similarly complex role for rhomboids in parasite pathogenicity was observed for *P. falciparum,* the causative agent of malaria. There is evidence that thrombospondin related anonymous protein (TRAP) that is crucial for tissue migration may be cleaved by ROM4 allowing host cell invasion through TRAP interactions with actomyosin motors.[15] *P. falciparum* proteases *pf*ROM1 and *pf*ROM4 have also been shown to be crucial for the invasion process as they cleave adhesin proteins that enable the release of the mature parasite into the host cell.[48] Continued investigations into ROM4, *Eh*ROM1, and *pf*ROM1 and *pf*ROM4 may aid in the development of therapeutics for pathogenic diseases such as malaria.

1.3.5 *E. coli* rhomboid protease, GlpG

Unlike in apicomplexan species, the physiological role of the *E. coli* (*ec*GlpG) rhomboid protease function is not well understood. GlpG was initially thought to play a role in the regulation of glycerol metabolism; however, in studies conducted on *E. coli* GlpG knockouts, growth and replication were not affected.[49,50] More recently, a potential role for *ec*GlpG in extraintestinal pathogenic *E. coli* (ExPEC) gut colonization was identified in mouse intestinal tracts and on mucous plates[46] with a decrease in gut colonization by *ec*GlpG knock-outs both *in vitro* and *in vivo.*[46] However, it is possible that downstream effects on the GlpR gene that resides within the glpEGR operon may have been the causative agent for this phenotype, indicating that *ec*GlpG may play a role in glycerol metabolic pathways to enhance ExPEC virulence.[46] Despite little being known about the role of *ec*GlpG or its physiological substrate, GlpG is used as the model protein for the rhomboid serine protease family due to the vast amount of structural and biochemical data that has been acquired on this system, partially owing to its relative ease of sample production.

In summary, rhomboid proteases play a role in a wide variety of biological processes, all of which rely on the ability to catalyze hydrolytic cleavage reactions within the phospholipid bilayer. While these studies provide insight into the mechanism of rhomboid protease function, there is significant interest in determining the structural mechanisms by which rhomboids act so that therapeutic agents can be developed to treat the numerous disease states that rhomboids contribute to.

1.4 Mechanism of rhomboid proteolysis

Rhomboids comprise a family of serine proteases that are able to achieve proteolytic cleavage of a peptide bond within transmembrane substrates using a mechanism similar to that of

soluble serine proteases. However, unlike other serine proteases, rhomboids accomplish proteolysis through a catalytic dyad that consists of a serine and histidine. The catalytic serine acts as a nucleophile, (Figure 1.2A), which attacks the carbonyl carbon in the peptide bond to form a negatively charged tetrahedral intermediate that is stabilized by the oxyanion hole. The nucleophilic activity of the hydroxyl group on serine is enhanced by hydrogen bonding interactions with histidine to activate the imidazole side chain. Through the collapse of the tetrahedral intermediate, the N-terminal cleavage product is released, causing the formation of a covalent acyl-enzyme intermediate. Subsequent hydrolysis via formation of a second tetrahedral intermediate releases the C-terminal cleavage product which re-establishes the active site for subsequent cleavage events.[28,51–53] This slightly differs from all other know serine proteases since they utilize a catalytic triad that adds aspartic acid to the active site. In conventional serine proteases, the aspartic acid increases the basicity of histidine so it can more effectively activate the nucleophilic serine. It has been proposed that a conserved asparagine residue performs a similar function to the aspartic acid residue;[51,53] however, subsequent studies have suggested alternative hypotheses such as the stabilizing pi-stacking interactions made between the catalytic histidine and the tyrosine at position 205.[54] That being said, the proteolytic rate of *ec*GlpG is very slow (on the order of minutes) suggesting that the catalytic dyad may be sufficient to complete the proteolytic function.[28]

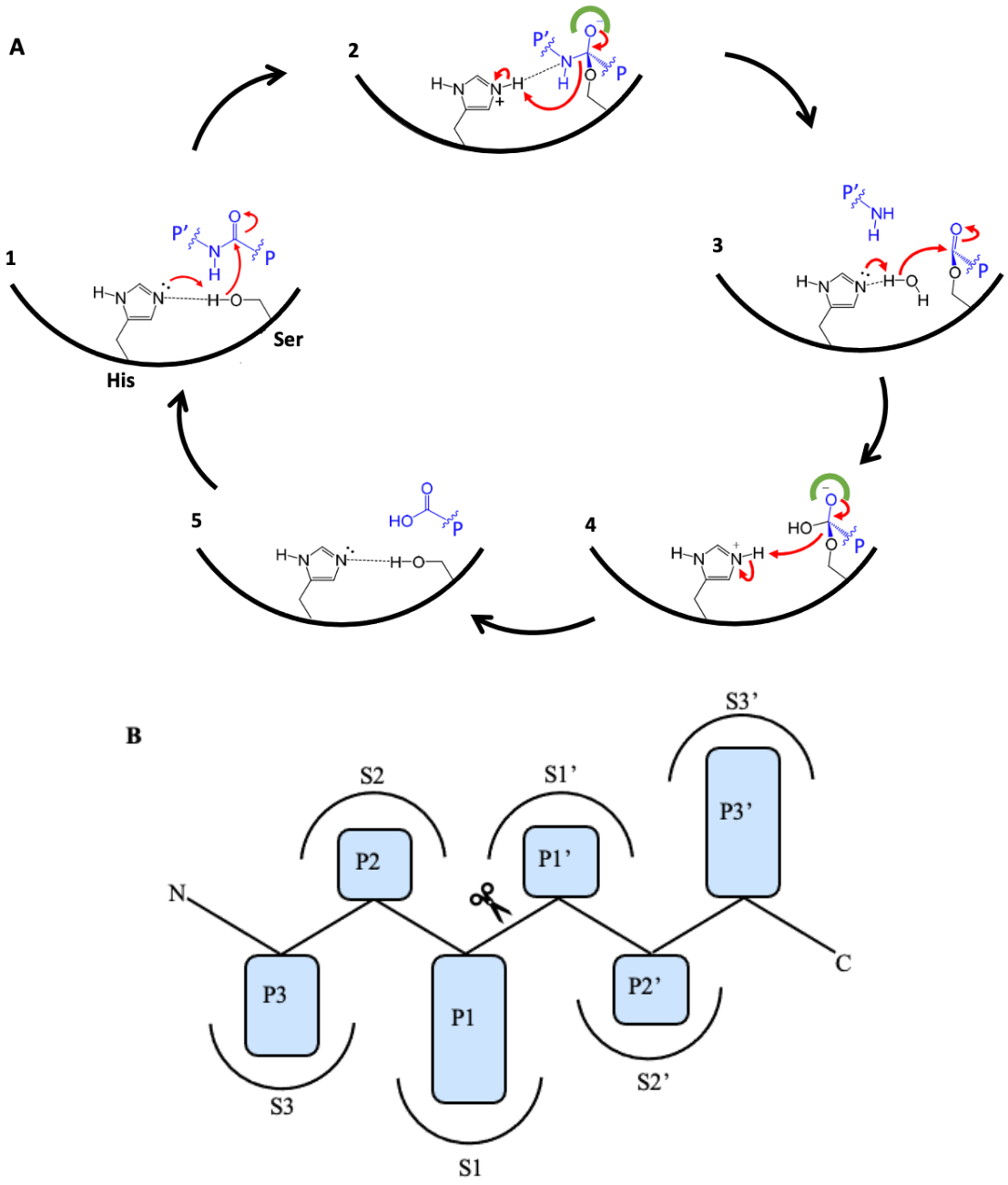

Figure 1.2: Rhomboid mechanism of peptide hydrolysis. (A) Catalytic mechanism starting with nucleophilic attack by the catalytic serine hydroxyl group activated by histidine (1), forming a tetrahedral intermediate (2), with its collapse giving rise to an acyl-enzyme intermediate (3). Activation of water forms a second tetrahedral intermediate (4) which collapses to release the cleavage products and regenerate the rhomboid active site (5). The oxyanion hole is illustrated in green, and the substrate in blue. (B) Schematic representation of nomenclature for protease active site and substrate, showing the interactions between substrate side chains (P3-P3') and substrate binding pockets (S3-S3').

As proteases cleave substrates at specific positions within their sequences, a standard nomenclature is used to define the cleavage locations and their respective protein binding sites (Figure 1.2B). This classic nomenclature defines the cleavage of the peptide bond occurring between the P1 and P1' substrate positions which correspond to the S1 and S1' protein binding sites on the rhomboid protease. With increasing distance to the N- or C-termini, the substrate positions are classified using PX and PX' respectively, where X indicates the position of the amino acid relative to the cleavage site. Similarly, the protein substrate-binding pockets are called SX and SX' for binding pockets that interact with PX and PX', respectively.

1.5 Rhomboid structure and function

1.5.1 Rhomboid substrate specificity

1.5.1.1 Consensus sequence for rhomboid proteolysis

The consensus sequence for cleavage, as identified from various substrates (Gurken, TatA, Spitz, and lactose permease (LY2)) is depicted in Figure 1.3, where a large hydrophobic residue is located in P4, a small residue is in P1, and another hydrophobic residue is in P2'.[26,55–57]

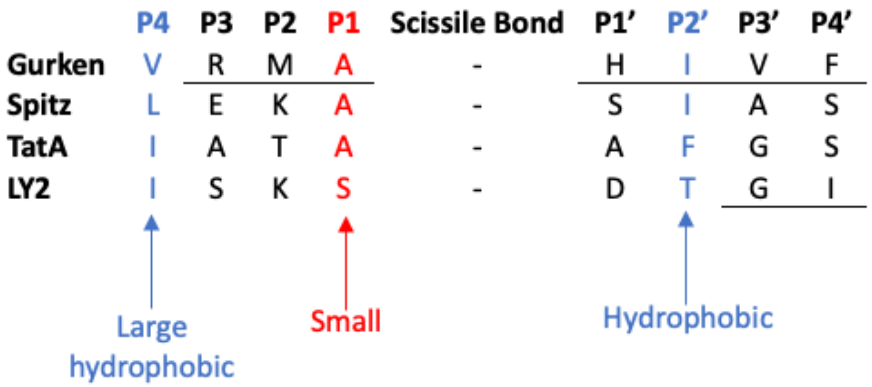

Figure 1.3: *ec*GlpG cleavage site sequences for indicated TM substrates. Regions of the sequence predicted to be within the TM segment are underlined.[26,56–58]

While this suggests that *ec*GlpG has a broad substrate sequence specificity, other factors such as the location of the consensus sequence with the TMD towards the N-terminus of the substrate on the luminal-membrane boundary[57,59–61] and the presence of helix-destabilizing residues[26,61] are also important characteristics for substrate selectivity.

1.5.1.2 Other factors affecting rhomboid substrate selectivity

In many cases, the rhomboid cleavage site appears to be embedded in the hydrophobic phase of the membrane in a TM helix. It has been proposed that exposure of this segment to the aqueous environment of the *ec*GlpG active site can destabilize its α-helical structure to favour proteolysis.[56] This is supported by the extended β-conformation observed in a tetrapeptide that was covalently bound to the active site serine in *ec*GlpG by x-ray crystallography.[55] In addition, studies of rhomboid protease activity have shown that helix-destabilizing residues such as proline in the TM segment C-terminal to the cleavage site enhance susceptibility to cleavage.[4,26] For example, the Akiyama group found that *ec*GlpG cleavage of the TM segment from Spitz increased significantly when helix-destabilizing residues were introduced at positions +17, and +18 residues from the cleavage site.[57] Similar results were obtained for the TM helix from Gurken, where removal of the native proline residue near the cleavage site inhibited its cleavage but addition of a proline further downstream in the sequence enhanced its susceptibility to cleavage.[57] These observations suggest that the presence of helix-destabilizing residues downstream from the cleavage site may be necessary for effective proteolysis of TM helix substrates.

In addition to the presence of helix-destabilizing residues, the location of the cleavage site has an impact on proteolysis. For example, the consensus cleavage sequence for the Rhomboid-1 substrate Spitz is located just outside of the TM helix near the N-terminus.[60] If the cleavage

sequence was shifted to be closer to the TM helix, cleavage by Rhomboid-1 no longer occurred.[60] The position of cleavage sites in a number of other rhomboid substrates are typically located in the hydrophilic membrane-lumen boundary region.[57] However, sequences outside the membrane can also be cleaved, as demonstrated by the addition of a hydrophilic 7-residue sequence between positions P3' and P4' that pushed the cleavage site out of the membrane yet was still susceptible to cleavage by *ec*GlpG.[61]

Another factor that modulates substrate specificity may be the presence of soluble domains that exist outside the rhomboid TMD. Most rhomboids have at least one extramembraneous domain, the function of which depends on the type of rhomboid it occurs in. One example where the function of the soluble domain has been identified is in the human rhomboid, RHBDL2, where the soluble domain is necessary for interactions with its physiological substrate, thrombomodulin, to aid in substrate recognition and proteolysis.[59] *ec*GlpG also has a soluble domain located at the N-terminus, spanning residues 1-61,[62] the function of which remains unknown. Deletion of this domain did not change the proteolytic activity of *ec*GlpG as tested in these assays,[58,62,63] although it is possible that it may have a regulatory role in the modulation of *ec*GlpG activity for its yet-to-be-identified physiological substrate. Due to the absence of changes observed in proteolytic activity upon the removal of the *ec*GlpG cytoplasmic domain, and its interaction with micelles in a way that promotes domain-swapping dimerization, which creates difficulties in obtaining samples compatible with solution-state NMR [64,65], a truncated variant of *ec*GlpG containing residues 81-276 (TMD81), was used in this thesis.

1.5.2 Rhomboid structure

The bacterial rhomboid *ec*GlpG is widely used as a model system for the study of rhomboid function due to the abundance of structural and biochemical data that has been acquired on this variant.[5,17,27,28,54,55,59,61,62,64,66–74] Dozens of x-ray crystal structures are now available for *ec*GlpG, all of which show six TM helices connected through loops of various sizes as illustrated in Figure 1.4. A unique L1 loop, indicated in blue in Figure 1.4, connects TM helix 1 (α1) and 2 (α2) and protrudes into the membrane bilayer, whereas the L2-L6 loops consist of only 1-9 residues to connect the adjacent TM helices. Both catalytic residues, serine-201 (S201) and histidine-254 (H254) reside in the water-filled active site approximately 10 Å below the surface of the cell membrane and are located on TM helix 4 (α4) and 6 (α6) respectively (Fig. 1.4).[5] Contained within the active site, in addition to S201 and H254, are the highly conserved GxSx and AHxxGxxxG motifs that are important in ensuring α4 and α6 remain in close proximity required for hydrogen bonding between the two catalytic residues.[5,66]

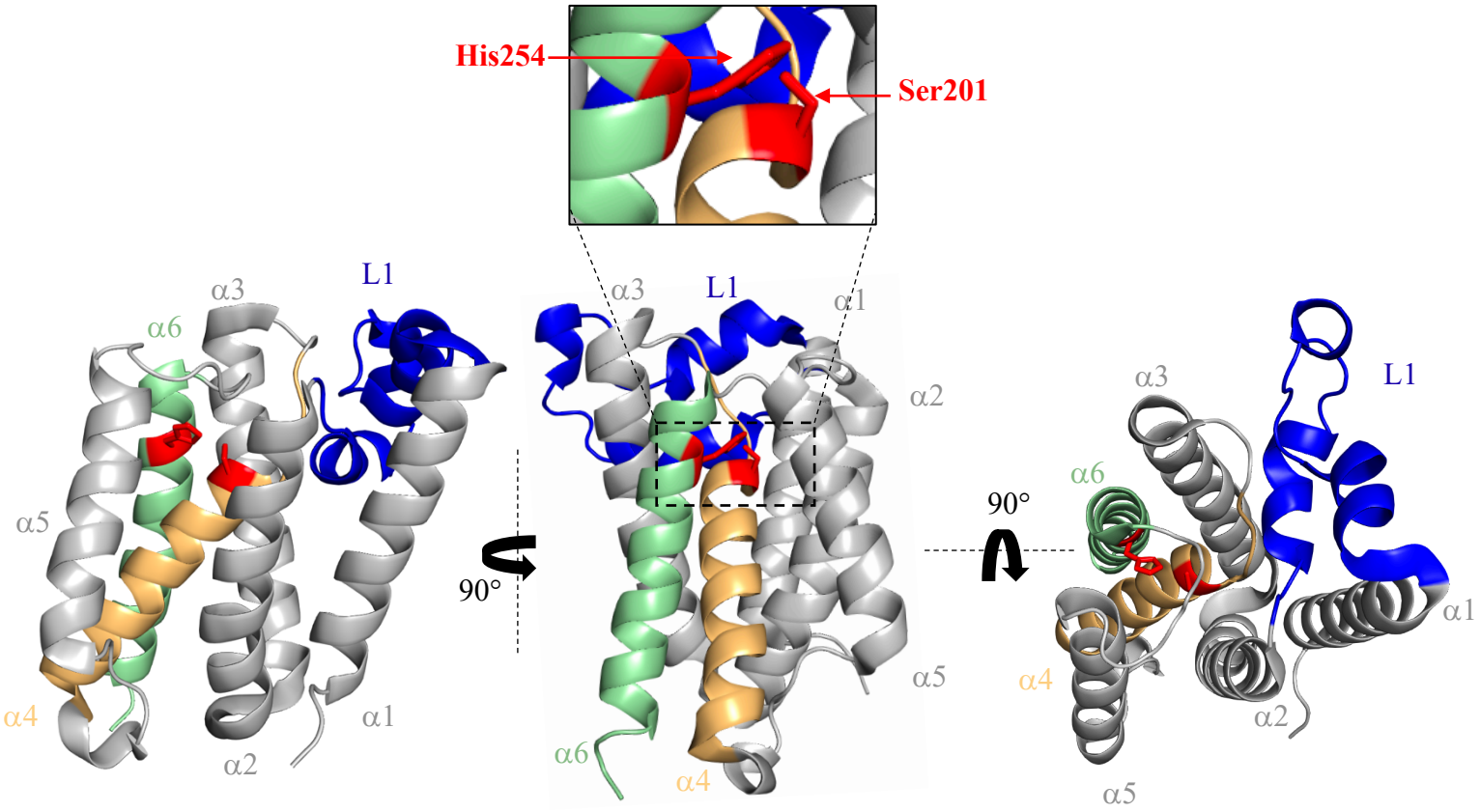

Figure 1.4: X-ray crystal structure of *ec*GlpG (PDB:2IC8).[24] The TMD includes six α-helices that adopt an up-down architecture according to CATH,[5] labelled α1-α6, connected by loops, including loop 1 (L1) that connects α1 and α2. The catalytic dyad making up the active site contains S201 on α4 and H254 on α6 (indicated in red).

Crystal structures of *ec*GlpG covalently bonded to peptide-aldehyde inhibitors (Figure 1.5A) have provided insight into interactions that are likely to stabilize formation of the tetrahedral intermediate. In these structures the backbone amide proton from S201, as well as the side chains of H150 and N154 could form interactions with the tetrahedral intermediate as part of the oxyanion hole.[55,68] In addition, these structures have revealed that the catalytic S201 carries out nucleophilic attack on the *si*-face of the scissile peptide bond (Figure 1.5B), which is unlike what has been observed in classical water-soluble serine proteases.[17,70]

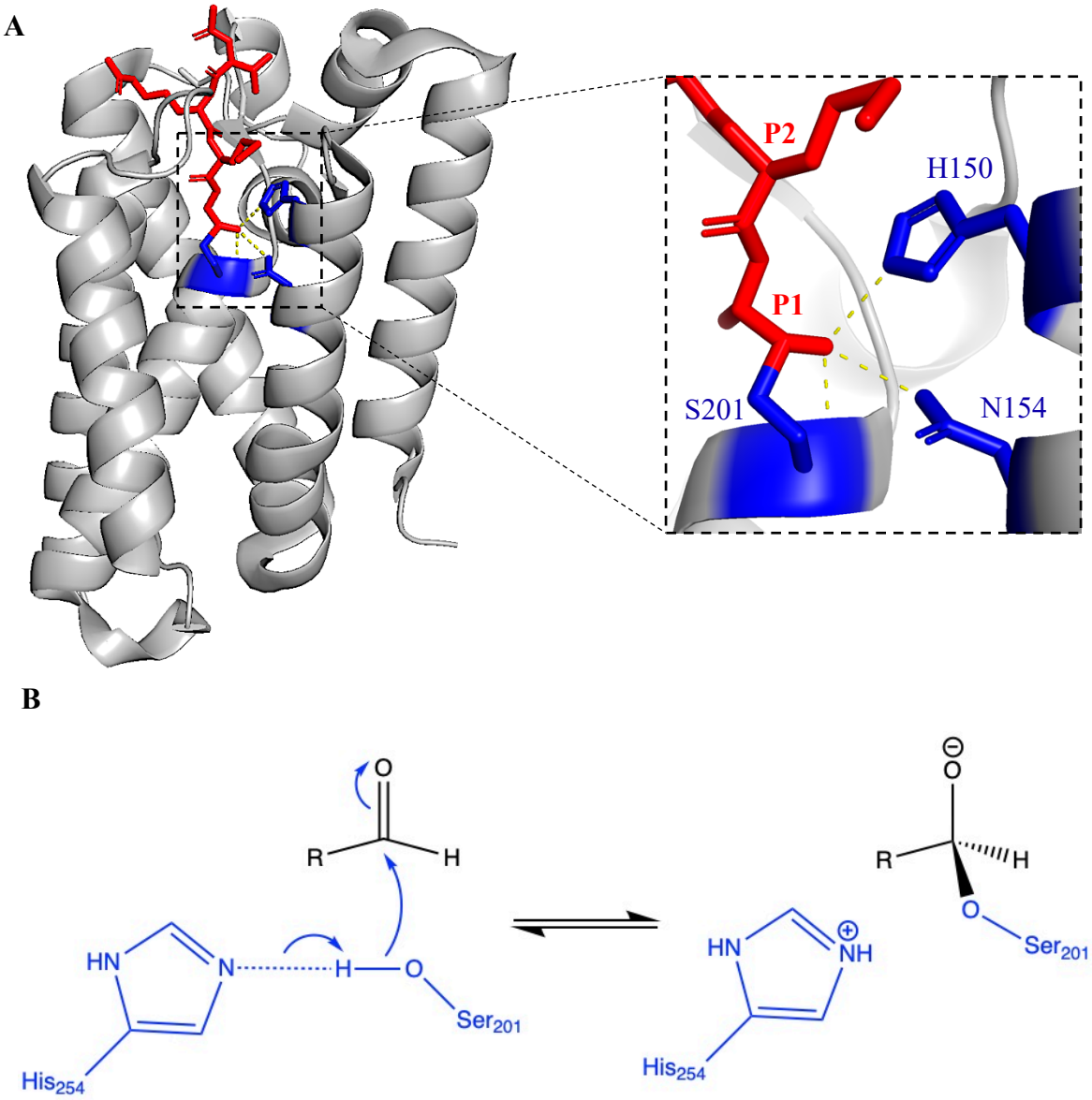

Figure 1.5: (A) *ec*GlpG in complex with a peptide-aldehyde inhibitor (illustrated in red) (PDB:5F5B). Interactions (dashed yellow lines) show the oxyanion stabilization by the side chains of H150 (2.69 Å), N154 (3.04 Å), and the backbone of S201 (3.3 Å). H150, N154, and S201 residues are illustrated in blue. (B) Illustration of the *si* face attack of the scissile peptide bond (black) by the nucleophilic serine (blue).

1.5.3 Substrate gating in rhomboid protease

As the main function of rhomboid is to perform hydrolytic cleavage of TM substrates that normally reside within the hydrophobic environment of the phospholipid bilayer, questions remain regarding the mechanism of substrate entry into the hydrophilic active site.

The α5 TM segment was initially postulated as the substrate gate following a study showing that mutations designed to reduce or enhance interactions between α2 and α5 resulted in an increase or decrease, respectively, in proteolytic activity.[27,28,54,72,75] Observation of an alternate structure for *ec*GlpG showing α5 bent ~35° away from α2 and the L5 loop shifted away from the active site gave rise to the hypothesis that this represented an open state for the enzyme (Figure 1.6A). However, it is possible that this state reflects the influence of crystal packing interactions, particularly since no other x-ray structures, out of the 42 structures that have been determined for *ec*GlpG, have shown this state.[54] Instead there has been a small subset of x-ray crystal structures showing a lateral shift in α5 by ~1 Å upon binding to a variety of inhibitors (isocoumarins, fluorophosphonates, monolactams, and chloromethyketones),[70] and ~5 Å in a 2016 study with a Gurken-derived peptide aldehyde inhibitor (Figure 1.6B).[55] However, the larger lateral shift was only observable in a ecGlpG mutant Y205F solubilised in phospholipid bicelles in complex with and without a peptide-based inhibitor. Other mutants of GlpG in bicelles or detergents failed to show this more open state.[55]

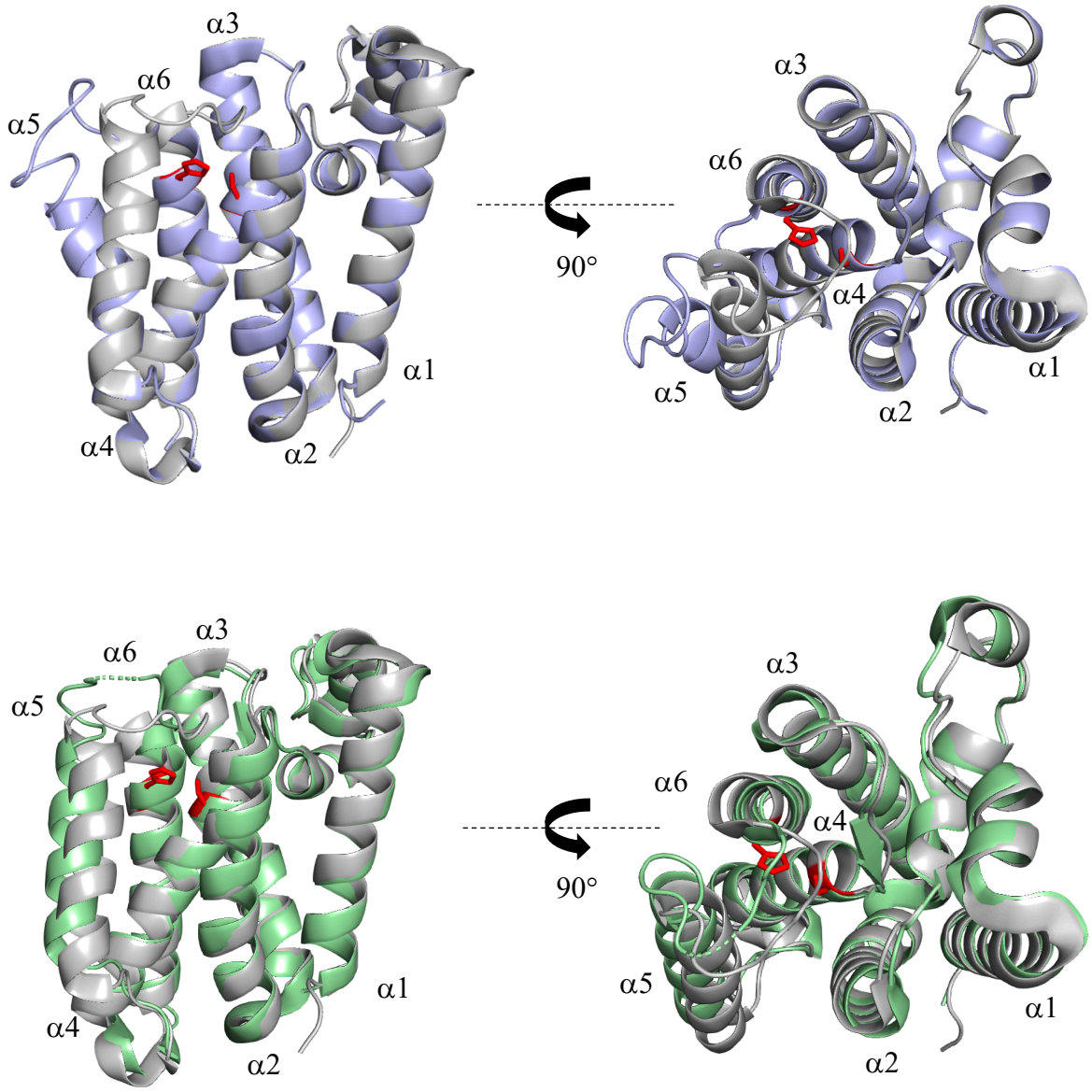

Figure 1.6: Gate-open conformation showing α5 in a 35° bend (blue) (A) or laterally displaced (green) (B) and superimposed on the closed conformation (grey). The catalytic dyad is indicated in red. (PDB: 2IC8, 2NRF, 5F5K).[5,29,55]

More recently, an series of time-resolved crystallographic structures done on *ec*GlpG in bicelles enabled the observation of discrete snapshots during proteolysis of a 13-mer substrate containing downstream helix-breaking serine and proline residues.[76] Two interesting findings arising from this work were a >90° bend in the substrate structure and a loss of crystallographic

order in the L5 loop and the entire α5 helix during the catalytic cycle.[76] The bend seen in the substrate upon formation of the tetrahedral intermediate with the catalytic serine suggests that mechanical stress induced in the peptide bond facilitates cleavage. In addition, the loss of electron density in α5 suggests that it undergoes conformational dynamics in the functional cycle.[76] In addition, the Urban group showed that the double mutant F153A/W236A designed to weaken interactions between TM helix 2 (α2) and 5 (α5) had 3-fold higher activity relative to WT *ec*GlpG in proteoliposomes with the TM substrate from TatA.[28] This observation suggests that the rate-limiting step in rhomboid proteolysis is governed by substrate entry rather than proteolysis itself.[28]

Taken together, the bulk of the evidence suggests that substrate gating involves movement of α5 which could be the rate-governing step. In spite of the importance of these measurements, up until recently there was no direct measurement of the timescale of α5 and substrate gating. Solution-state NMR has the potential to provide insight into this potential link between conformational dynamics and function. Therefore, this thesis will attempt to explore this possibility, but rather than using detergent micelles, we will investigate the potential of nanodiscs as a membrane mimicking environment as described in section 1.6.4.

1.6 Membrane mimicking environments

There are various membrane-mimetic systems that can maintain membrane proteins in a water-soluble state. Although a popular membrane mimic used in the study of rhomboid protease function and structure has been detergent micelles, other potential membrane mimics have emerged that provide an environment more similar to the native lipid environment. These include bicelles, styrene-maleic acid lipoparticles (SMALPs) and nanodiscs formed by lipid-encapsulated membrane-scaffolding proteins (MSPs), the latter which is the focus of this thesis.

1.6.1 Detergent micelles

Detergent micelles have been and continue to be one of the most popular membrane-mimetics used for the study of membrane proteins since the complex they form with integral membrane proteins (IMPs) is relatively small.[77] Micelles are characterized by a core hydrophobic phase typically formed by the aliphatic chains of amphipathic molecules, such as detergents, with the polar headgroups on the surface interacting with the surrounding aqueous environment.[78] Micelles tend to surround the TMD of integral membrane proteins to shield the hydrophobic region from the aqueous phase while maintaining the complex in a water-soluble state (Figure 1.7A).

Detergent micelles spontaneously form when the concentration exceeds the critical micelle concentration (CMC), where micelle charge, size and curvature are determined by the type of detergent used. For detergents with the same headgroup, shorter alkyl chains result in micelles that have a higher CMC value, greater water permeability in the hydrophobic phase, and more disorder.[79] However, smaller micelles are more practical for solution-phase nuclear magnetic resonance (NMR) studies which benefit from the smaller protein-detergent complexes (PDCs) that are formed.

Despite the similarities between detergent micelles and lipid bilayers, significant differences exist that could affect membrane protein structure and function. As illustrated in Figure 1.6A, detergent micelles have a greater curvature compared to lipid bilayers and use detergents which contain one acyl chain rather than two, as seen in phospholipids, resulting in a more permeable environment to lipids and/or water. This gives rise to a decrease in lateral pressure that is exerted on the protein in a detergent micelle relative to the phospholipid bilayer.[26,54,80–82] Not only does this increase the amount of water that can access the hydrophobic core, but can also

alter the conformational dynamics that underlie protein function. Despite these challenges, detergent micelles continue to be used in a large number of NMR studies of membrane proteins due to their ability to solubilize membrane proteins in small PDCs while maintaining structure and function.

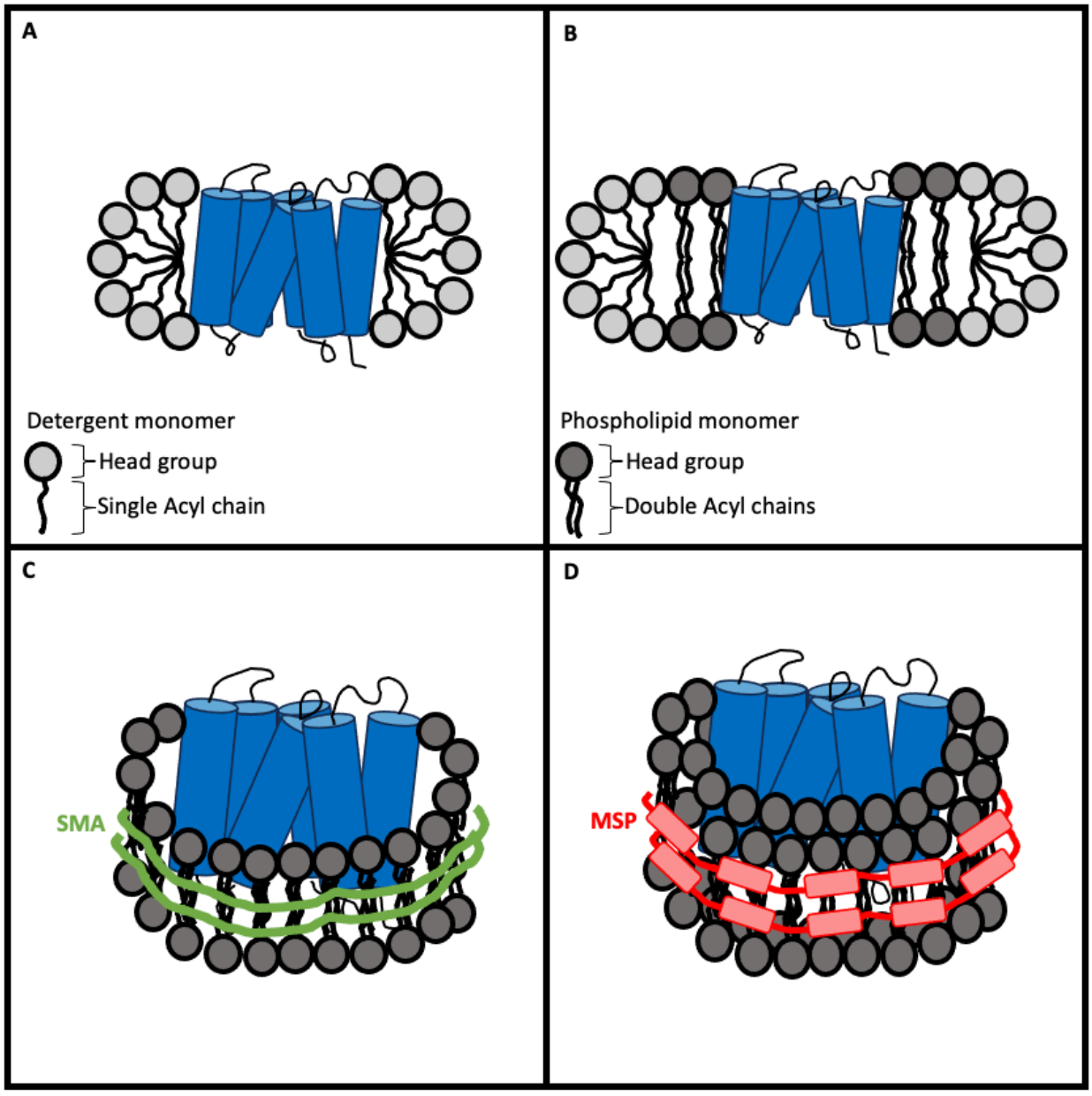

Figure 1.7: (previous page) Schematic diagram (not to scale) illustrating the solubilisation of a membrane protein in detergent micelles (A), phospholipid bicelles (B), SMALPs (C), and MSP nanodiscs (D). Detergent molecules are illustrated in light grey, and phospholipid molecules in dark grey.

1.6.2 Phospholipid bicelles

Phospholipid bicelles are formed by mixtures of long-chain lipids with short chain lipids or cholate-based detergents. These mixtures form lipid bilayers that are maintained in a soluble state by capping of hydrophobic edges with the short chain lipids or detergents (Figure 1.7B). While phospholipid bicelles have been used for structure determination of IMPs by x-ray crystallography, their larger size can make it challenging to use them for solution-state NMR.[83] However, the size of bicelles can be controlled by changing the lipid to detergent ratios, with smaller ratios giving rise to isotropic, fast-tumbling bicelles that can be used for solution-state NMR.[83] In addition, bicelles are stable over a wide range of pH values and have been able to support enzymatic activity of proteins that are non-functional in detergent micelles.[83] However, the requirement for short-chain lipids or detergents can also create problems since monomeric detergent species will be present that may bind to extra-membraneous parts of the protein and affect its stability.[84] For this reason there has been increasing interest in generating water-soluble membrane protein samples in a lipid environment that circumvent the need for detergents.

1.6.3 Styrene-maleic acid lipoparticles (SMALPs)

SMALPs are discoidal lipoparticles that are formed by a styrene-maleic anhydride (SMA) polymer that spontaneously inserts into lipid membranes and extracts membrane fragments, where the hydrophobic regions at the edges of the discs are protected by interactions with SMA (Figure 1.7C).[85] SMALPs have been successfully used to conduct solution-state NMR studies on the direct extraction of the single TM-helix-containing proteins cytochrome-b5 (cytb5) from *E. coli* membranes.[86] The polymer used in that study differed from SMA by modification of the carboxylic acid groups by amide bond-linked ethanolamine groups, which increases SMALP stability in lower

pH solutions than those formed by SMA.[87] The SMALPs formed in this study were only 5 nm in diameter, which was small enough to allow ^{1}H-^{15}N HSQC spectra of cytb5 to be obtained, showing peaks from the cytosolic domain.[86] However, no signals for the TM helix were observed, suggesting that the particle size may have still been too large for solution NMR.[86] In addition, it can be difficult to obtain SMALP samples with the high level of homogeneity required for high quality NMR spectra.[88] While optimization of the lipid to SMA ratio could potentially improve homogeneity,[88] it has been challenging to find conditions that allow solution-state NMR studies to be carried out on membrane proteins incorporated into SMALPs.

Incorporation of TMD81 into SMALPs with SMA has previously been attempted in the Goto lab, but showed a large hydrodynamic radius of 16.6 nm, with significant heterogeneity in size (± 4.1 nm). These characteristics were not compatible with solution-state NMR[85] and would therefore require further optimization of SMALP complex formation for this application.

1.6.4 Membrane-scaffolding protein-based nanodiscs

Alternative soluble complexes containing lipid bilayers have also been developed using membrane-scaffolding proteins (MSPs). These proteins are derived from the human serum apolipoprotein A1 (Apo-A1) that forms the main protein component of high-density lipoproteins where Apo-A1 surrounds portions of a lipid bilayer creating a membrane disc.[89,90]

Nanodiscs can be formed *in vitro*, where two copies of the MSP protein encircle a lipid bilayer disc in an anti-parallel manner with a diameter that depends on the length of the MSP used (Figure 1.7D).[91,92] MSP forms a series of amphipathic α helices that provide a peptide belt to separate the hydrophobic phase of the lipid bilayer from the hydrophilic environment.[93,94] It has been shown that it is possible to incorporate integral membrane proteins of various sizes, including

larger membrane proteins such as G protein-couple receptors (50 kDa),[95] and rhodopsin (27 kDa),[91,96,97] into MSP-lipid nanodiscs. These complexes have allowed the study of these membrane proteins by a range of techniques including x-ray crystallography[98], electron microscopy,[92,96] and solution-state NMR.[92,96] Under most conditions studied, nanodiscs have been shown to adopt a circular disc morphology, however, at lower temperatures complexes may become more elliptical in shape.[94] This difference in shape must be taken into account when incorporating a membrane protein into the nanodisc as at least two layers of lipids should be present between the membrane protein of interest and the surrounding MSPs to form a stable complex. This usually requires some optimization of the lipid to protein ratio required to fill the nanodisc, with smaller less stable complexes[91] arising when ratios are too low, and aggregated nanodiscs when ratios are too high.[96] The choice of lipid is also vital in the formation of a nanodisc that can support the native fold and conformational dynamics of the membrane protein. This will depend on the protein being studied, since the structure and function of an IMP can be impacted by the surrounding lipids as has been seen for rhomboid proteases.[65,67,69,74,99,100]

Several MSP variants have been made to accommodate membrane proteins of different sizes. The first MSP variant developed, called MSP1, was derived from a deletion of the first 43 residues on Apo-A1 containing the N-terminal globular domain, which created nanodiscs with a diameter of ~10 nm.[101,102] It was subsequently shown that removal of an additional N-terminal 20 residues resulted in approximately the same nanodisc diameter, giving rise to the variant called 1D1.[102] 1D1 has been used to study membrane proteins of similar molecular weight to TMD81 (~25-30 kDa) by solution NMR, including the 7-TM segment protein bacteriorhodopsin (bR).[96] Thermal denaturation studies revealed significantly enhanced stability of bR in nanodiscs over detergents, allowing NMR spectra to be obtained at higher temperatures with ~90% of the expected

peaks being observed under these conditions.[96] However, these nanodisc complexes were still very large for solution NMR applications, having a molecular weight of 125-150 kDa.[93] To reduce the size of the nanodisc, truncated variants of MSPs were designed to make nanodiscs with diameters ranging from 6-8 nm.[92] A total of six variants were created through the deletions as outlined in Table 1.1, with the MSP variant identified to have the most promise for solution NMR studies being ΔH5.

Table 1.1: Properties of MSP variants used to make nanodiscs. Deleted regions in truncated mutants are in red.[92]

MSP	MSP MW (kDa)	Nanodisc Diameter (nm)	Sequence
1D1 (WT)	22	~ 9	N-α1-α2-α3-α4-α5-α6-α7-α8-α9-α10-C
ΔH4	19	~ 8	N-α1-α2-α3-α4-α5-α6-α7-α8-α9-α10-C
ΔH5	~19.5	~ 8	N-α1-α2-α3-α4-α5-α6-α7-α8-α9-α10-C
ΔH4H5	~16.5	~ 7	N-α1-α2-α3-α4-α5-α6-α7-α8-α9-α10-C
ΔH4-H6	14	~ 6	N-α1-α2-α3-α4-α5-α6-α7-α8-α9-α10-C

1.6.5 The impact of environment on *ec*GlpG

The impact of detergent micelles and bicelles on *ec*GlpG structure and function has been investigated by circular dichroism conducted on GlpG in bicelles containing DLPC or DMPC[65], showing differences that suggest hydrophobic mismatch between the hydrophobic phase of the lipid bilayer and the GlpG TMD had an impact on its secondary structure. At the same time, there have been numerous studies suggesting that *ec*GlpG has some ability to induce thinning of lipids in the local environment to reduce this hydrophobic mismatch. [67,69,74] Thinning of ~1.1 Å per leaflet

has primarily been observed in lipids that are more native to *E. coli* such as PE with PG, but is minimal (~ 0.1 Å) in other lipids such as DMPC, suggesting some specific interactions may occur between TMD81 and its native membrane lipids.[99] These membrane deformations have been supported by x-ray crystal structures showing a difference in the hydrophobic thickness of *ec*GlpG (27 Å) and that found for lipid bilayers composed of PE (~28 Å).[5,99] Furthermore, molecular dynamics simulations conducted on *ec*GlpG in PE and PC lipid bilayers illustrated the impact of lipids on the angle of the protein in the membrane relative to the bilayer as well as an average thinning of the local membrane by ~4 Å.[74] In addition to membrane thinning by *ec*GlpG, the L1 loop has been shown in molecular dynamic simulations to cause to lipid distortions in the surrounding membrane environment that increases the permeability of GlpG enabling lipids to enter the active site.[100]

On top of structural studies suggesting hydrophobic mismatch, functional studies have shown that the lipid environment has an impact on the proteolytic activity of GlpG.[62,65,67,69] Detergent micelles composed of phosphocholine or maltoside head groups with alkyl chain lengths of 10-12 carbon atoms supported the highest activity, but decreased when detergents with shorter or longer alkyl groups were used.[65] In addition, in bicelles, proteolytic activity was supported when the long-chain lipid was DLPC (Figure 1.8), but not when longer PC lipids were used (e.g. DMPC).[65] *ec*GlpG activity has also been studied in phospholipid bicelles against the peptide substrate LacYTM2, which showed greater activity in DMPC compared to DMPE or DMPG.[99] While this result was unexpected, it is possible that the native *E.coli* lipids, PE/PG, may have an inhibitory effect on *e*cGlpG potentially by modulating conformational dynamics.[99] These results

illustrate the impact that the membrane environment has on *ec*GlpG function, which highlights the importance of studying rhomboid function and structure in more native-like environments.

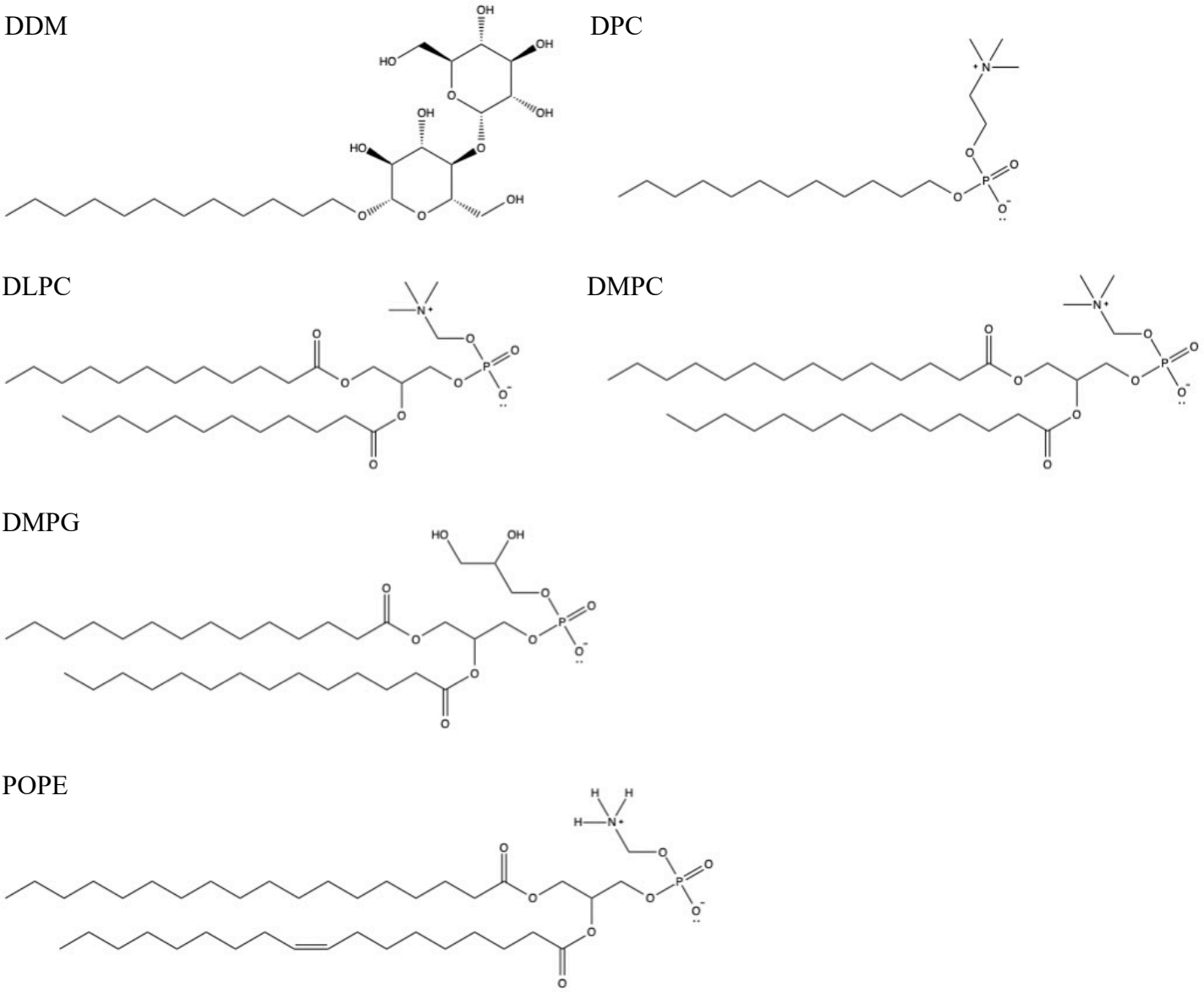

Figure 1.8: Schematic depicting chemical structure of detergents (DDM and DPC) and lipids (DLPC, DMPC, DMPG, and POPE).

1.7 Nuclear magnetic resonance (NMR) spectroscopy

We are interested in studying the structure and conformational dynamics of *ec*GlpG in an appropriate membrane-mimetic system by solution NMR, with a focus in on residues around the putative α5 gate. Some basic theory of this technique will be presented since this helps to explain some of the challenges and opportunities that are associated with it.

1.7.1 Basic principles of NMR spectroscopy

Solution NMR spectroscopy uses the magnetic properties of NMR-active nuclei and how this is affected by local chemical environment, which can be used to characterize the structure and conformational dynamics or proteins. NMR-active nuclei such as, ^{1}H, ^{15}N, and ^{13}C possess an intrinsic angular momentum with a spin number of ½.[103,104] In the NMR spectrometer these spins are exposed to an external magnetic field (B_o) which induces a dipole that interacts with the external magnetic field to adopt an antiparallel or parallel orientation relative to this field, where each of these states has a slightly different energy.[103] Upon the absorption of radiofrequency energy, transitions can be induced between the two states. This typically occurs at a characteristic frequency for each nucleus type, called the Larmor frequency (ω_o). This frequency is normalized relative to a reference frequency (v_{ref}) and reported as a chemical shift from this reference in units of parts per million (δ_{ppm}).[103,104] Therefore, the absorption frequency reflects the influence of a variety of factors such as electron density, bonding environment, and non-covalent interactions.

Since the differences between the energy states is very small, there is only a small excess population of spins in the lower energy state. This small excess of spins in the lower-energy orientation gives rise to a net magnetization vector that is oriented parallel to the external magnetic field, which is designated as the z-axis.[105] Electromagnetic radiation at ω_o tips the net

magnetization vector into the transverse plane where its precession is detected as an oscillating field in the NMR spectrometer. The intensity of the detected signal depends, in part, on modes of relaxation of the irradiated spins back to equilibrium conditions. Longitudinal relaxation (T_1) defines the rate of reestablishment of the equilibrium distribution of spins in the two states and transverse relaxation (T_2) is the loss of the coherence of the bulk magnetization vector in the transverse plane.[105] If transverse relaxation occurs quickly this will give rise to broad peaks, since the net magnetization vector in the transverse plane is what is detected by NMR. This process tends to be faster for larger molecules with slow molecular reorientation rates, which is why NMR spectra of proteins tend to have broad peaks. In cases where the complex molecular weight is very large (>50 kDa), peaks may be too broad to resolve without the use of selective labeling.

1.7.2 ^{1}H-^{15}N HSQC spectra of proteins

While 1D ^{1}H NMR is commonly used for small molecules, the ^{1}H spectrum of larger molecules such as proteins becomes difficult to interpret due to the large number of proton resonances and overlapping peaks.[106] Therefore, multidimensional NMR experiments are used to allow resolution of peaks and the correlation of nuclei through bonds or space to give structural information. The most popular type of NMR spectrum used for proteins is the 2D ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) experiment. Peaks in this spectrum occur at the chemical shift (δ_{ppm}) of ^{1}H protons along the x-axis, and the chemical shift of the directly bonded ^{15}N atom.[107] Therefore, each non-proline amino acid in a protein should give rise to a single peak in this spectrum from the backbone amide proton. In addition, the side chains of the amino acids, histidine (His), tryptophan (Trp), asparagine (Asn), glutamine (Gln), lysine (Lys), and arginine (Arg) may also be seen in this spectrum as they contain ^{1}H-^{15}N groups. The ^{1}H-^{15}N HSQC

spectrum is often considered to be a “fingerprint” spectrum, since each protein gives rise to a unique spectrum that is characteristic of that protein.[107] Typically, protein samples studied by this technique are uniformly labelled with ^{15}N during their production[106,107], usually by expression in bacterial systems that are amenable to this process.

1.7.3 NMR experiments to monitor conformational change

Even the simplest NMR experiment, that is the 1D spectrum, can also provide information on dynamic processes, such as conformational exchange. In the case where a spin undergoes exchange between 2 states, its impact on the spectrum will depend on the timescale of the process. Specifically, a spin undergoing exchange with a rate constant (k_{ex}) that is smaller than the difference in absorption frequency between the two states ($|\Delta v|$) ($k_{ex} << |\Delta v|$), will give rise to two individual peaks in the NMR spectrum with intensities that are proportional to the population of each state. As the exchange rate increases and reaches a value similar to $|\Delta v|$, the intensities of the NMR peaks decrease as the peaks broaden and shift in position towards the weighted average chemical shift of the two peaks. This type of exchange is called intermediate timescale exchange and is very difficult to detect due to the peak broadening that occurs in severe cases. Fast exchange occurs once the exchange rate exceeds the difference in absorption frequency ($k_{ex} >> |\Delta v|$) where one sharp peak representing the population-weighted average is observed. Figure 1.9 illustrates the transition of exchange rates and the expected observations for each scenario in the NMR spectrum.

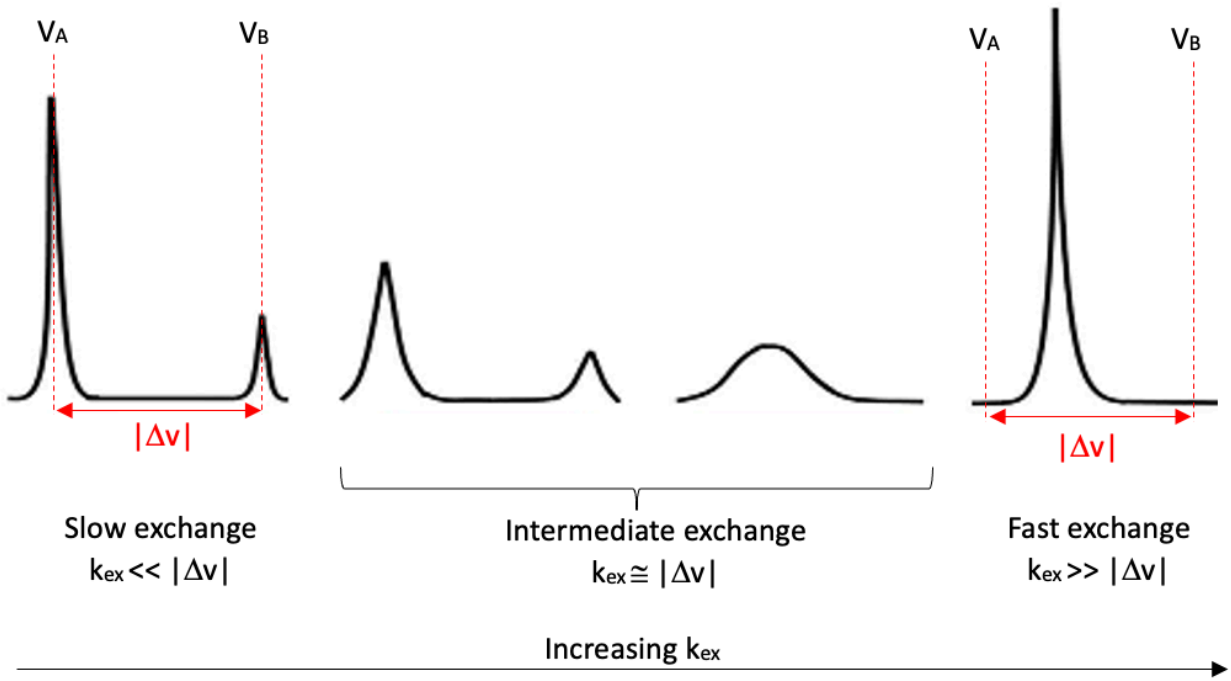

Figure 1.9: Schematic diagram representing the effect of chemical change on the NMR spectrum where v_A and v_B represent the Larmour frequencies (Hz) for the two exchanging species, A and B respectively. $|\Delta v|$ represents the differences between v_A and v_B and k_{ex} is the rate constant for exchange between the two states.

1.8 Thesis Objectives

Despite all of the structural and kinetic studies done on TMD81 in detergent micelles, bicelles, and SMALPs, the structure and timescale of conformational dynamics in the putative α5 substrate gate, and its link to rhomboid activity remains elusive. In this thesis, we investigate the feasibility of reconstituting TMD81 into MSP-based nanodiscs, and whether these complexes can produce well-resolved ^{1}H-^{15}N HSQC spectra. The results of this thesis will determine for the first time whether nanodiscs can provide a useful membrane-mimicking environment for NMR study of TMD81.

Chapter 2: Materials and methods

2.1 DNA and constructs

A truncated version of the rhomboid serine protease from *Escherichia coli* (*ec*GlpG) containing residues 81-276 (TMD81) was expressed from a pET25b vector[108–110] with a C-terminal hexahistidine tag (Figure 2.1A). The membrane-scaffolding protein, MSP1D1 (1D1) (Figure 2.1B), cloned in a pET28a vector was graciously provided by Dr. Jyh-Yeuan Lee at the University of Ottawa. A mutant of 1D1 with a deletion of residues 120-143, (MSP1D1ΔH5 (ΔH5)) (Figure 2.1C), was obtained from the Sligar lab.[92] Sequences were all verified through sequencing performed at the Ontario Genomics Innovation Center at The Ottawa Hospital Research Institute.

A ***ec*GlpG:**
MRRYPFFAALRERAGPVTWVMMIACVVVFIAMQILGDQEVMLWLAWPFDPTLKFEF
WRYFTHALMHFSLMHILFNLLWWWYLGGAVEKRLGSGKLIVITLISALLSGYVQQKF
SGPWFGGLSGVVYALMGYVWLRGERDPQSGIYLQRGLIIFALIWIVAGWFDLFGMSM
ANGAHIAGLAVGLAMAFVDSLNARKRKASLERENLYFQGSSHHHHHH

B **MSP1D1:**
GHHHHHHHDYDIPTTENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKD
LEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQEKLSPLG
EEMRDRARAHVDALRTHLAPYSDELRQRLAARLEALKENGGARLAEYHAKATEHLS
TLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLNTQ

C **MSP1D1ΔH5:**
GHHHHHHHDYDIPTTENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKD
LEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLGEEMRDRARAHVDALRTHLAPY
SDELRQRLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQGLLPVLE
SFKVSFLSALEEYTKKLNTQ

Figure 2.1: Amino acid sequences for *ec*GlpG (A), MSP1D1 (B), and MSP1D1ΔH5 (C). Amino acid residues that are removed with the addition of TEV protease to MSP sequences are highlighted in red.

2.2 Chemically competent cells

50 μL of the desired cells, *E. coli* C43 (DE3) (Lucigen), BL21 (DE3) (Novagen), or DH5α (ThermoFisher), were used to create 10X, 100X, and 1000X serial dilutions on non-selective agar-Luria-Bertani (LB) broth (Lennox, Bioshop) plates. Inverted plates were incubated for 12-16 hours at 37°C. A single colony of the desired cells was used to inoculate 30 mL of LB media which was shaken at 220 rpm and 37°C for 16-18 hours. 6 mL of preculture was used to inoculate 400 mL of LB media and allowed to shake at 37°C until an optical density measured at 600 nm (OD_{600}) using an Ultrospec 2100 Pro Spectrophotometer (GE Healthcare) of ~0.2-0.4 was reached. All solutions, and equipment used (Eppendorf tubes, falcon tubes, and the centrifuge) were kept at 4°C for the remainder of the procedure. Cultures were chilled on ice for ~20 minutes and spun using a Spinchron 15 Centrifuge (Beckman Coulter) with a C0650 fixed angle rotor at 4,700xg for 5 minutes before following previously established protocols to create chemically competent cells using the $MgCl_2$-$CaCl_2$ method.[111] In brief, pellets were resuspended in $MgCl_2$ and combined prior to being spun using a Spinchron 15 Centrifuge (Beckman Coulter) at 4,000xg for 5 minutes and resuspended in $CaCl_2$. Subsequent centrifugation was done using a Spinchron 15 Centrifuge (Beckman Coulter) at 4,000 xg for 5 minutes, after which the pellet was resuspended in $CaCl_2$ with 15% glycerol and aliquoted (50 μL) into sterile, chilled Eppendorf tubes that were immediately placed at -80°C until further use.

2.3 Transformation

TMD81 and MSP plasmids were transformed into chemically competent *E. coli* C43 (DE3) cells (Lucigen), BL21 (DE3) (Novagen) or DH5α cells using previously published protocols.[62,65] 10 ng of *ec*GlpG or 50 ng of MSP DNA was added to a 50 μL aliquot of competent cells and

allowed to incubate on ice for 30 minutes. The cells were then heat-shocked for 45 seconds at 42°C and returned to rest on ice for 2 minutes. 900 μL of 37°C Luria-Bertani (LB) broth (Lennox, Bioshop) was added to the cells before being placed in a MaxQ 4000 Benchtop Orbital Shaker (Thermo Scientific) for 60 minutes at 37°C and 300 rpm. Cells were subsequently spun at 13,200 rpm for 1 minute, 750 μL of supernatant removed, and the cell pellet resuspended in the remaining supernatant. 100 μL of the cell suspension was spread onto an LB agar plate containing 100 μg/mL ampicillin (Bioshop) for *ec*GlpG expression, or 50 μg/mL kanamycin (Bioshop) for MSP expression. Inverted plates were incubated for 12-16 hours at 37°C.

2.4 Plasmid DNA preparation

A single colony of freshly transformed DH5α cells was used to inoculate 30 mL of LB media with 100 μg/mL of ampicillin or 35 μg/mL kanamycin. Cultures were shaken at 220 rpm and 37°C for 16-18 hours. Cells were harvested by centrifugation using a Spinchron 15 Centrifuge (Beckman Coulter) with a C0650 fixed angle rotor at 5000xg and room temperature (~25°C) for 10 minutes. Plasmids were purified using the E.Z.N.A Plasmid DNA Mini Kit II (Omega Bio-Tek) following the manufacturer's protocol and the concentrations of the master DNA stocks were measured by absorbance using a Nanodrop spectrophotometer (Thermo Scientific). *ec*GlpG and MSP DNA master stocks were diluted to prepare 10 ng/μL[85] and 50 ng/μL[92] working solutions respectively as established by previously published protocols and stored at -20°C.

2.5 Expression of *ec*GlpG

Following transformation as described in section 2.3, one colony was used to inoculate 100 mL of LB media with 100 μg/mL of ampicillin. The cultures were shaken overnight at 220 rpm at 37°C and centrifuged using a Spinchron 15 Centrifuge (Beckman Coulter) with a C0650 fixed angle rotor at 5000xg and 25°C for 10 minutes and then resuspended in 1 L of LB media (100 μg/mL ampicillin), or 1 L of M9 minimal media (47.9 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.5 mM NaCl, 1 mM $MgSO_4$, 0.1% (w/v) D-glucose, 0.1% (w/v) $^{15}NH_4Cl$, 100 uM $CaCl_2$, 0.01X LB, and 100 μg/mL ampicillin). The optical density of the culture was monitored for growth by measuring the absorbance at 600 nm (OD_{600}) using an Ultrospec 2100 Pro Spectrophotometer (GE Healthcare). When OD_{600} reached 0.5 - 0.7, expression was induced with 1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG) and cultures were shaken at 220 rpm for 3 hours at 37°C or overnight at 16°C. Cells were harvested by centrifugation using an Avanti J-E Centrifuge (Beckham Coulter) at 6000xg and 4°C for 10 minutes. Pellets from 1 L of culture were combined and weighed. Yields for *ec*GlpG expression produced 0.5-0.9 mg per 1 L of culture. Pellets were stored at -20°C until needed.

2.6 Expression of MSP

One colony of freshly transformed BL21(DE3) was used to inoculate 50 mL of LB media with 35 μg/mL of kanamycin and shaken overnight at 200 rpm at 37°C. After centrifugation at 5000xg and 25°C for 10 minutes using a Spinchron 15 Centrifuge (Beckman Coulter) with a C0650 fixed angle rotor, the pellet was resuspended in 1 L of LB media (35 mg/L kanamycin). Expression was induced with 1 mM IPTG when OD_{600} was 0.6-0.8, and the cultures shaken for a maximum of 4 hours at 180 rpm, 37°C. Cells were harvested by centrifugation using an Avanti J-E Centrifuge (Beckham Coulter) at 5000xg and 4°C for 20 minutes. Pellets from 1 L of culture were combined

and resuspended in 25 mL of MSP Buffer (20 mM Tris-HCl, (pH 7.5), 100 mM NaCl, and 0.5 mM 2,2',2'',2'''-(ethane-1,2-diyldinitrilo) tetraacetic acid (EDTA)). After centrifugation using an Avanti J-E Centrifuge (Beckham Coulter) at 3000xg and 4°C for 20 minutes, pellets from 1L of culture were combined and weighted. The yield of MSP expression produced on average ~3 mg per 1 L of culture. Pellets were stored at -20°C until needed.

2.7 Purification of *ec*GlpG in detergent micelles

Frozen cell pellets were thawed on ice and resuspended in 25 mL of cold Harvest Buffer (20 mM 4-(2-hydroxyethyl) piperazine-1-ethanesulfonic acid (HEPES) pH 7.4, 100 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, and 1X Cocktail Mixture (1 mg/mL 4-2-aminoethyl benzene sulfonyl fluoride hydrochloride (AESBF), 2 μg/mL pepstatin, 1.2 μg/mL E-64, 1 μg/mL bestatin, and 1.8 μg/L phosphoramidon)). Cells were lysed on ice by sonication using a Fisher Sonic Dismembrator Model 500 at 50% intensity using two rounds of 60 one-second ON-OFF pulses separated by a 6-minute delay time. To isolate the insoluble membrane fraction, the lysate was centrifuged using an Avanti J-E centrifuge (Beckham Coulter) at 48,500xg for 1 hour at 4°C. The resulting pellet was intermittently mechanically resuspended with a metal spatula on ice every 15 minutes over a period of 1.5 hours in 1X Solubilisation Buffer (50 mM HEPES pH 7.4, 200 mM NaCl, 5 mM imidazole, 10% (w/v) glycerol, 1X Cocktail Mixture, and 0.1% (w/v) n-dodecyl β-D-maltoside (DDM)). Centrifugation in an Avanti J-E centrifuge (Beckham Coulter) at 48,500xg for 30 minutes at 4°C was used to obtain the supernatant containing the detergent-solubilised protein which was applied to 2.5 mL of nickel-NTA agarose affinity resin (Qiagen) that was equilibrated in 2X Solubilisation Buffer (50 mM HEPES buffer (pH 7.4), 200 mM NaCl, 5 mM imidazole, and 10% (w/v) glycerol). Two sequential wash steps using Wash Buffer I (50 mM HEPES pH 7.4, 500 mM NaCl, 30 mM

imidazole, 10% (w/v) glycerol, and 0.1% (w/v) n-dodecylphosphocholine (DPC)) and Wash Buffer II (50 mM HEPES pH 7.4, 200 mM NaCl, 30 mM imidazole, 10% (w/v) glycerol, and 0.1% (w/v) DPC) removed impurities. The protein was eluted using Wash Buffer II supplemented to obtain a final concentration of 250 mM imidazole. The eluted protein was applied to a Superdex-200 Increase 10/300 GL size exclusion chromatography column equilibrated in FPLC Buffer (50 mM Tris-HCl pH 7.4, 50 mM NaCl, 100 uM EDTA, and 0.1% (w/v) DPC) and run on an AKTA FPLC (GE Healthcare) that monitored absorbance at 280 nm. Aliquots of 0.5 mL were collected from the main protein peak and stored at 4°C for further use. Protein concentration and purity were assessed through a BCA assay (Pierce) and SDS-PAGE analysis.

2.8 Purification of 1D1 and ΔH5

Frozen cell pellets were thawed on ice and resuspended in 20 mL of Resuspension A Buffer (50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 1% Triton X-100, 1 mM EDTA, and 4X Cocktail Mixture). After addition of lysozyme, cells were sonicated in the same fashion as described in section 2.7. To remove remaining DNA, 5 mM of $MgCl_2$, and 1 μg/mL of DNase I was added to the lysate and incubated on ice for 30 minutes before being centrifuged in an Avanti J-E centrifuge (Beckham Coulter) at 30,000xg for 45 minutes at 4°C. The supernatant containing the solubilized protein was applied and incubated for 5-10 minutes on 2.5 mL of nickel NTA agarose resin (Qiagen) that was equilibrated in Buffer A (50 mM Tris-HCl (pH 8.0), 500 mM NaCl) supplemented with 1% (w/v) Triton X-100. To remove impurities, four sequential wash steps were conducted using Wash Buffer I (Buffer A with 1% (w/v) Triton X-100), Wash Buffer II (Buffer A with 50 mM cholate), Wash Buffer III (Buffer A), and Wash Buffer IV (Buffer A with 20 mM imidazole). The protein was eluted using Buffer A supplemented with 500 mM of imidazole. 30

mL of purified MSP was dialyzed overnight against 4 L of Dialysis Buffer 1 (20 mM Tris-HCl (pH 7.5), 100 mM NaCl and 0.5 mM EDTA) resulting in a ~130-fold dilution factor using membrane tubing with a 6-8 kDa molecular weight cut-off (MWCO) (Spectra/Por) to remove imidazole. Using a 10 kDa MWCO centrifugal filter unit (Amicon), the protein sample was concentrated to ~500 μM and stored in the fridge at -20°C for further use. Protein concentration and purity were assessed through a BCA assay (Pierce) and SDS-PAGE analysis.

2.9 Cleavage of MSP His-Tag

Cleavage of the His-tag on either MSP construct was performed on purified samples after dialysis with Dialysis Buffer I as outlined in section 2.8 through the addition of TEV protease (Sigma Aldrich) at 1 A_{280} TEV for 100 A_{280} of MSP protein with 1 mM dithiothreitol (DTT) and left to incubate overnight at 4°C. Absorbance at 280 nm was measured using a Nanodrop spectrophotometer (Thermo Scientific) and reaction progress assessed through SDS-PAGE analysis. The protein solution (50 mL) was dialyzed against 4 L of Dialysis Buffer 2 (20 mM Tris-HCl (pH 8.0), and 500 mM NaCl) resulting in a ~80-fold dilution factor using 6-8 kDa MWCO dialysis tubing (Spectra/Por) overnight to remove EDTA. To isolate the MSP cleavage product, a reverse nickel affinity chromatography step was used, where the cleavage mixture was incubated with 2.5 mL of nickel NTA agarose affinity resin (Qiagen) for 5-10 minutes and the flow-through containing His-tag free MSP was then collected. The flow-through (50 mL) was dialyzed against 4 L of Dialysis Buffer 3 (20 mM Tris-HCl (pH 7.5), 100 mM NaCl and 0.5 mM EDTA) for a ~80-fold dilution factor using 6-8 kDa MWCO dialysis tubing (Spectra/Por) overnight. Using a 10 kDa MWCO centrifugal filter unit (Amicon), the protein sample was concentrated to ~500 μM and

stored at -20°C until further use. MSP samples were generally used within 1-2 weeks of purification.

2.10 SDS-PAGE analysis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was conducted using the Laemmli protocol[112] to assess protein purity. Equal parts of protein sample were added to 2X SDS loading buffer (125 mM Tris (pH 6.8), 20% glycerol (v/v), 4% SDS, 0.01% (v/v) bromophenol blue, and 0.02% (v/v) 2-mercaptoethanol) and boiled for one minute to denature the proteins. 4 μL of EZ-RUN pre-stained Rec protein marker (Fisher) and 10 μL of each protein sample (0.4-2.5 nmol for *ec*GlpG, 4.6 nmol for 1D1, 1.8 nmol for 1D1ΔHis, and 1.1-10.3 nmol for ΔH5) were loaded onto a 12% or 15% SDS-polyacrylamide gel (12 or 15% (v/v) acrylamide mix (acrylamide; bis-acrylamide, 29:1)(Fisher), 0.1% (w/v) SDS, 0.2% (w/v) ammonium persulfate (Fisher), 0.08% (v/v) TEMED (Fisher), and 26% (v/v) 0.39 M Tris (pH 8.8)) with a 5% stacking gel (5% (v/v) acrylamide mix (acrylamide; bis-acrylamide, 29:1)(Fisher), 0.1% (w/v) SDS, 0.2% (w/v) ammonium persulfate (Fisher), 0.2% (v/v) TEMED (Fisher), and 12 % (v/v) 0.12 M Tris (pH 6.8)). Gels were run in a Mini-PROTEAN electrophoresis cell (Bio-Rad) for 50 and 60 minutes for 12% and 15% poly-acrylamide gels respectively at 150 V in 1X Tris-Glycine Running Buffer (0.3% (w/v) Tris (pH 8.5), 1.44% (w/v) glycine and 0.1% (w/v) SDS).

Gels were stained using Coomassie Blue Stain Solution (0.3% (w/v) Coomassie brilliant blue R-250 (CBB), 50% (v/v) methanol, and 10% (v/v) acetic acid) on a rocking platform for 1 hour or overnight after being heated for 30 seconds in a microwave. Prior to destaining, the Coomassie Blue Stain was removed, and the gel rinsed with ddH_2O. Gels were destained using a

solution containing 40% (v/v) methanol and 10% (v/v) acetic acid for 1 hour after being heated for 30 seconds in a microwave. Upon completion of destaining, the gels were stored in ddH_2O.

For samples with low concentrations, gels were stained with 60 mL of SYPRO Ruby (Fisher) on a rocking platform overnight. Gels were destained using 100 mL of Wash Solution containing 10% (v/v) methanol and 7% (v/v) acetic acid for 30 minutes. Prior to imaging, the gels were washed with ddH_2O twice for a minimum of 5 minutes.

All gel images were captured on a ChemiDoc XRS+ (Bio-Rad) and integrated band intensity ratios were measured using ImageJ analysis software.

2.11 Protein Concentration Determination

TMD81, 1D1, and ΔH5 protein concentrations were determined by conducting a bicinchoninic acid (BCA) assay kit using the manufacturer's instructions (Pierce). In brief, standards of bovine serum albumin (BSA) and dilutions of TMD81, 1D1, and ΔH5 were combined with Working Solution containing 98% (v/v) Reagent A (Pierce) (1 % (w/v) BCA, 2% (w/v) sodium carbonate, 0.16% (w/v) sodium tartrate, 0.95% (w/v) sodium bicarbonate, and 0.4% (w/v) NaOH), and 2% (v/v) Reagent B (Pierce) (4% (w/v) copper (II) sulfate pentahydrate). Standards and protein samples were incubated at 37°C for 30 minutes. Absorbance was measured at 562 nm using an Ultrospec 2100 Pro Spectrophotometer (GE healthcare). Mean absorbance values for BSA standards were used to generate a standard curve which was then used to calculate sample protein concentrations.

2.12 Preparation of protein-free nanodiscs

Empty nanodiscs were prepared using a previously established protocol.[92] Cholate Buffer was made by the addition of 100 mM sodium cholate to MSP Buffer (20 mM Tris-HCl (pH 7.5), 100 mM NaCl, and 0.5 mM EDTA)) which was kept at 4°C until use. 50 mM lipid solutions were made with 1,2-dimyristoyl-sn-glycero-3-phospho-(1-rac-glycerol) (sodium salt) (DMPG), 1,2-dilauroyl-sn-glycero-3-phosphocholine (DLPC), and/or 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) in Cholate Buffer. The lipid stocks were vortexed and kept in a 50°C water bath to ensure complete dissolution after which they were kept at room temperature for further use. The ratios of lipid:MSP examined in this thesis were: 80:1, 70:1, 60:1, 50:1, 40:1, and 30:1. Mixtures were kept at room temperature on a rocking platform for 1 hour prior to the addition of 0.5-1 g of wet bio-beads (Bio-Rad) per mL of mixture and then incubated on a rocker platform overnight at room temperature. Removal of bio-beads was accomplished by poking a hole in the bottom of the Eppendorf tube containing the bio-beads with the desired sample using an 18-gauge needle (Fisher Scientific) and allowing the nanodisc solution to flow through into a new Eppendorf tube. To maximize nanodisc recovery, bio-beads were subsequently washed with 1 mL of MSP Buffer and allowed to drain into the same Eppendorf tube as the previous step. For MSP samples that had been subjected to the His-tag cleavage protocol outlined in section 2.9, the assembly mixture was then added to a Ni-NTA column equilibrated in MSP Buffer, and the flow-through collected and concentrated in a 30 kDa MWCO centrifugal filter unit (Amicon) to a final volume of ~ 1 mL. Final purification of nanodiscs was done on a Superdex-200 Increase 10/300 GL size exclusion chromatography column equilibrated in FPLC Nanodisc Buffer (50 mM Tris-HCl (pH 7.4), and 150 mM NaCl) and run on an AKTA FPLC (GE Healthcare) that monitored absorbance at 280 nm. Aliquots of 0.5 mL were collected and saved for further analysis. The approximate

molecular weight of the species in each peak on all FPLC profiles was estimated by comparison to the manufacturers graph obtained with standards ranging from 669 kDa to 13.7 kDa ran on a Superdex-200 Increase 10/300 GL SEC column.[113] The standards were not run on the FPLC machine used in this thesis. The number of trials is indicated in the caption for each figure throughout the thesis. It should be noted that most experiments were only done once, therefore the suggestions made are tentative and would need to be validated by reproducing the results over multiple trials.

2.13 Preparation of *ec*GlpG incorporated nanodiscs

Development of protocols for the incorporation of TMD81 into nanodiscs were guided by previously published work from the Wagner lab.[92] TMD81 was concentrated in a 30 kDa MWCO (Amicon) to obtain a concentration of ~200 μM which was confirmed using a BCA assay (Pierce). The ratios of lipid:MSP:TMD81 examined in this thesis were: 160:2:1, 1:2:140, 120:2:1, 100:2:1, 80:2:1, 60:2:1, 320:4:1, 280:4:1, 240:4:1, 200:4:1, 160:4:1, and 120:4:1. The assembly, collection and analysis of TMD81 nanodiscs was conducted as described in section 2.12.

2.14 ^{15}N NMR spectroscopy

For ^{15}N labelled TMD81 in detergent micelles or in nanodiscs, the protein sample was concentrated using a 30,000 Da MWCO centrifugal filter unit (Amicon) to a concentration between 50-100 μM in a 350 μL volume containing 10% (v/v) D_2O. NMR samples were inserted into a shape tube (Bruker) and analysed on a 600 MHz Bruker Avance III Spectrometer with a triple resonance cryoprobe at the University of Ottawa NMR Facility. Standard ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectra were acquired at 45°C with a spectral

width of 30 ppm spread across 128 increments in the indirect dimension.[58] NMR spectra were analysed and processed using NMRPipe[114] and NMRViewJ[115] with signal intensities normalized for the numbers of scans (typical range 208-256), receiver gain (typical range 57-203), and total protein concentration (typically between 80-250 μM).

2.15 Matrix-associated laser desorption/ionization (MALDI)

MALDI was carried out on TMD81 in detergent micelles and nanodiscs using a Microflex MALDI-TOF mass spectrometer (Bruker) set to a m/z window between 0 and 30 kDa.[62] The sandwich and the mixed method were used with saturated α-cyano-4-hydroxycinnamic (4-HCCA) acid (2 parts acetonitrile, 1 part water, and 0.1% final trifluoroacetic acid (TFA)) as the matrix for each sample for ionization of membrane proteins.[62,116] Briefly, a stainless steel MALDI sample plate was cleaned with methanol and deionized water. 1 μL of the matrix solution (1 part saturated 4-HCCA with 3 parts isopropanol) was added to the MALDI plate and allowed to dry. Protein purification and preparation was completed using the manufacturer's instructions for C18 ZipTip pipette tips (Millipore) where the tips were first equilibrated in 5% methanol in 0.1% TFA. Protein samples were then loaded onto the ZipTip and washed in 75% acetonitrile (ACN), 0.1% TFA. 1 μL of protein sample was spotted on the plate and once dried, 1 μL of matrix solution was placed on top. For the mixed method, the matrix solution was mixed with the dried protein sample through aspirating with a sterile pipette tip. All samples were allowed to dry prior to insertion into the mass spectrometer.

Chapter 3: Results

3.1 NMR spectroscopy of WT TMD81 in detergent micelles

To confirm that the purification of folded WT TMD81 could be reproduced, an ^{15}N-labeled sample was purified in DPC micelles to allow acquisition of an ^{1}H-^{15}N HSQC spectrum. WT TMD81 with a C-terminal hexahistidine tag (6X) was overexpressed in *E. coli* C43 (DE3) cells and purified by nickel affinity chromatography and size exclusion chromatography (SEC) as described in section 2.5 and 2.7. The SEC profile showed that most of the sample eluted at a volume of 12-15 mL (Figure 3.1A) in line with previous observations,[65] and as expected for a ~75 kDa PDC complex. Aliquots were taken at each stage of TMD81 purification and run on a Coomassie-stained gel (Figure 3.1B). Minimal traces of TMD81 appeared in the flow through (FT), wash 1 (W1), and wash 2 (W2) fractions whereas the elution fraction (E) shows an intense band that is expected for the 25 kDa TMD81, similar to previous results.[65] SEC fractions containing TMD81 (F8 to F11) were combined and concentrated for NMR.

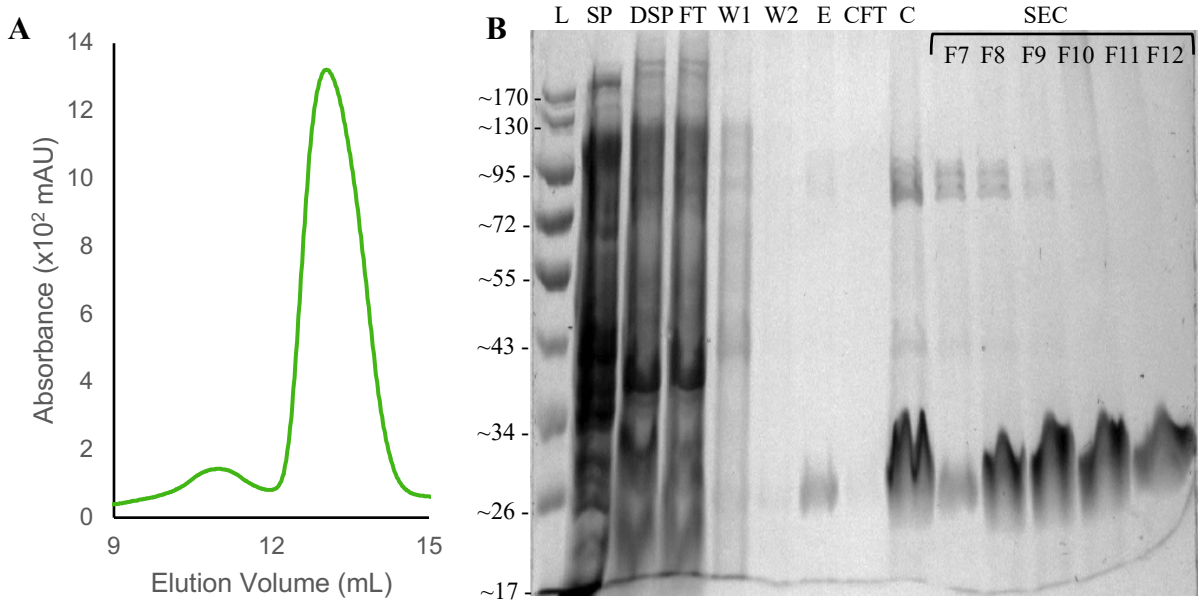

Figure 3.1: (previous page) (A) SEC run on a SP200 column and monitored at 280 nm from a purification of 4 L of ^{15}N TMD81 in GlpG buffer. (B) Coomassie-stained SDS-PAGE of fractions from the purification of ^{15}N-labeled TMD81 (100 μM). L: molecular weight ladder in kDa, SP: soluble fraction after lysis, DSP: detergent fraction after lysis, FT: flow through from the nickel affinity chromatography column, W1 and W2: fractions from the nickel affinity chromatography wash steps, E: eluted fraction, CFT: concentrator flow through, C: pooled, elution fractions after concentration, SEC (F7-F12): 0.5 mL fractions from SEC starting at an elution volume of 12 mL (F7).

A ^{1}H-^{15}N HSQC spectrum was acquired on pooled concentrated fractions from the size exclusion chromatography (Figure 3.2) and showed a pattern of peaks that was similar to previously recorded spectra.[65] This spectrum confirmed that TMD81 could be purified using established methods for NMR, forming the basis for this study of the incorporation of GlpG into nanodiscs.

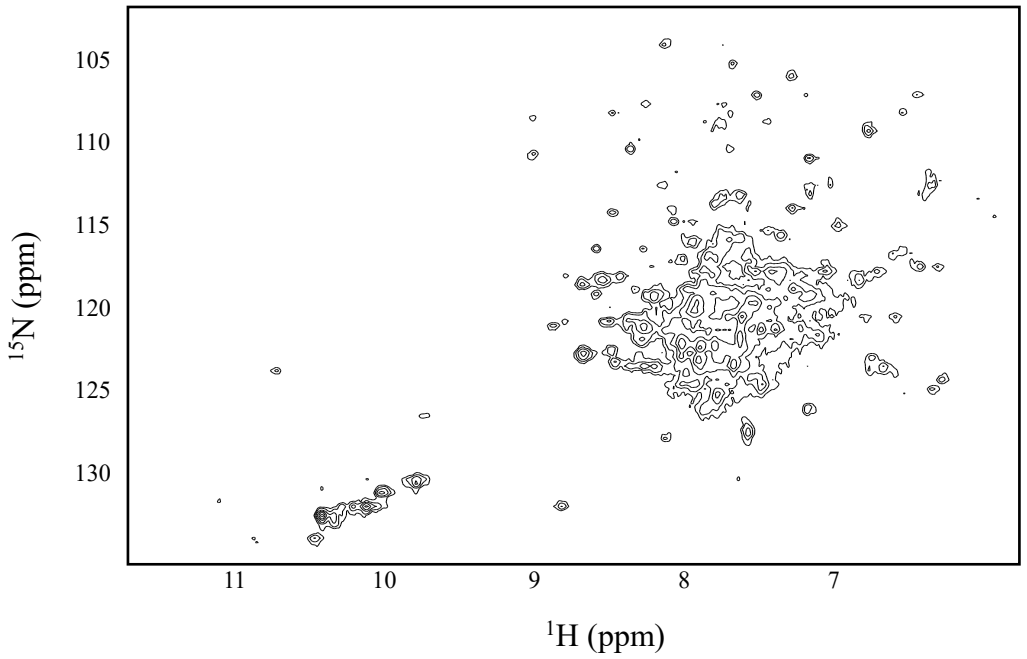

Figure 3.2: ^{1}H-^{15}N HSQC spectrum of TMD81 (100 μM) in DPC detergent micelles with 10% D_2O at 45°C. Spectrum was acquired using a 500 MHz spectrometer equipped with a triple resonance cryoprobe.

3.2 Assembly of nanodiscs with ΔH5

3.2.1 Formation trial of empty ΔH5 DMPC nanodiscs

We chose to focus our first efforts on nanodisc formation using a shorter MSP variant with a deletion of helix 5 called ΔH5 since this variant was used to successfully produce a solution-state NMR spectrum for an alpha-helical transmembrane protein with a similar molecular weight (MW) to TMD81, namely bacteriorhodopsin (bR).[93] The ΔH5 MSP was expressed in *E. coli*, purified as described in section 2.8 and analyzed by Coomassie-stained SDS-PAGE (Figure 3.3A). In purified samples, a large band was observed with a mobility consistent with the known molecular weight (MW) of ~22 kDa. The single-step purification showed there may have been some loss in the flow through (FT) although this is hard to discern due to the high concentration of protein in this sample. There is also minimal loss in the wash steps. A yield of 2 mg per L of culture was obtained with >90% purity.

To make nanodiscs, purified ΔH5 was incubated at room temperature (RT) with DMPC at a lipid:MSP molar ratio of 80:2 for one hour prior to the addition of bio-beads to remove detergent in a ~24-hour incubation. Desalting was achieved by SEC (Figure 3.3B) which showed a peak centered at ~13.5 mL which is similar to the elution volume of the 160 kDa MW standard published in the manufacturer's specifications.[113] The presence of a single peak at this elution volume suggests that successful nanodisc formation had occurred, although the FPLC profile showed that the peak was broad (~12 mL to ~15 mL) and at a slightly higher MW than expected. SDS-PAGE analysis (Figure 3.3C) of these fractions suggests the presence of ΔH5 at the expected MW of 22 kDa. Faint bands were also seen in all fractions for a ~44 kDa species that may represent an SDS-resistant dimeric state as seen in previous studies.[117] Faint bands of higher MW species (>55 kDa) were also present in all fractions which may suggest some aggregation of MSP during nanodisc

formation. Nonetheless, as the formation of empty nanodiscs with the ΔH5 MSP and DMPC and a lipid:MSP ratio of 80:2 appeared to be successful, these same conditions were used to attempt formation of TMD81-encapsulated nanodiscs.

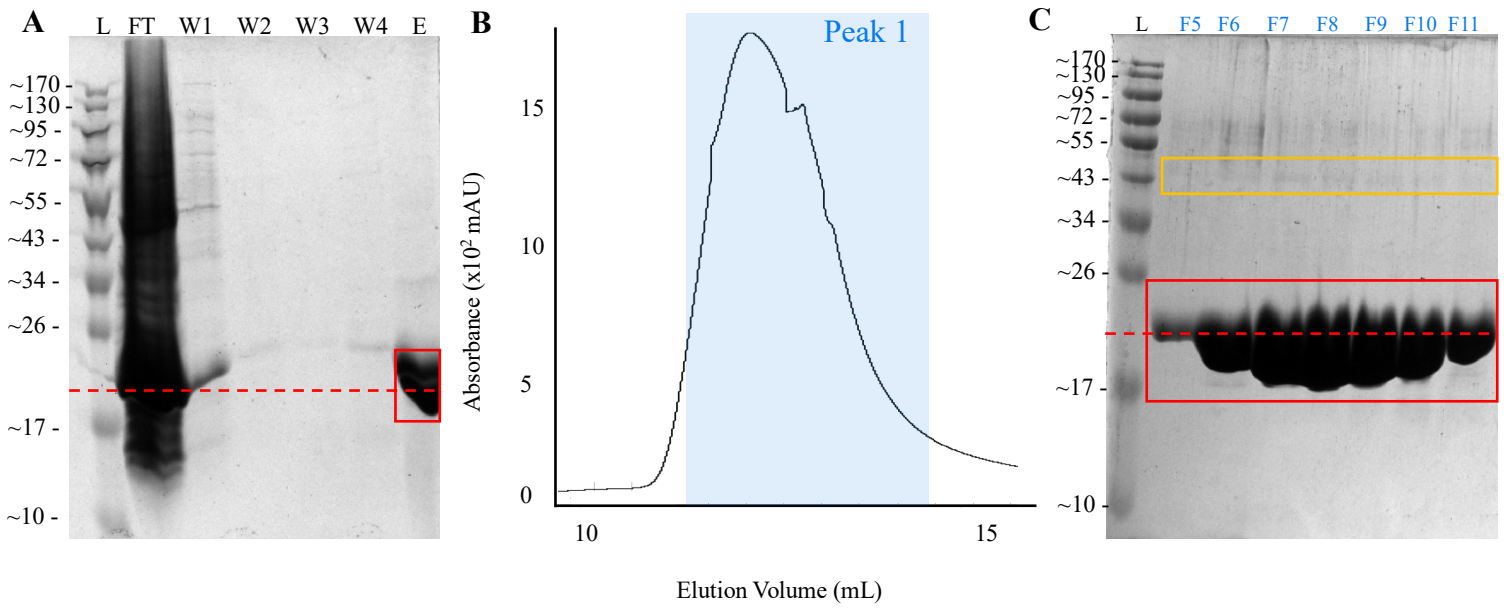

Figure 3.3: (A) Coomassie-stained SDS-PAGE of fractions taken from various steps of the nickel affinity chromatography purification of ΔH5 (100 μM). (B) SEC on a SP200 column for empty ΔH5 DMPC nanodiscs in Lipid Nanodisc buffer. Absorbance at 280 nm was monitored. Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113] (C) Coomassie-stained SDS-PAGE of fractions taken from SP200 SEC of empty ΔH5 DMPC nanodiscs in Lipid Nanodisc buffer. L: MW ladder in kDa, FT: flow through, W1, W2, W3, and W4: wash step fractions, E: elution step, SEC (F5-F11): 0.5 mL fractions from SEC starting at an elution volume of 11 mL (F5). Light red boxes: full-length ΔH5, orange box: ΔH5 dimer (44 kDa), red dashed line: expected MW for ΔH5 (22 kDa). Nanodisc reconstitution under these conditions was only attempted once.

3.2.2 Nanodisc formation trial with TMD81 in DMPC with ΔH5

ΔH5 and TMD81 were purified as described above, with similar levels of purity and yields for ΔH5 (2.5 mg/L) (Figure 3.4A), but lower yields for TMD81 (0.5 mg/L) due to some losses

during purification (Figure 3.4B). These samples were used with DMPC at a ratio of 80:2:1 (lipid:MSP:TMD81) to make nanodiscs which were analyzed by SEC and SDS PAGE.

As illustrated in Figure 3.4C, the FPLC profile showed three peaks centered at ~9 mL, ~11 mL, and ~13.5 mL. SDS-PAGE analysis (Figure 3.4D) showed 3 bands at ~15 kDa, ~22 kDa, and ~25 kDa in fractions from all peaks. While the ~22 kDa and ~25 kDa bands represent ΔH5 and TMD81 respectively, the ~15 kDa species likely corresponds to a degradation product that may have been present in the MSP sample. The main difference between each peak was the relative amount of each species, with peak 1 being mainly comprised of TMD81, and similar amounts of TMD81 and ΔH5 being present in the other two peaks. Given the large MW suggested by the elution volume of peak 1 (>669 kDa), this species may represent a larger nanodisc complex, potentially containing TMD81 in a dimeric state as has been proposed to exist in the membrane.[52,64] The other two peaks also may contain nanodisc complexes, with peak 3 at 13.5 mL occurring at an elution volume expected for that of a nanodisc-TMD81 complex that could be compatible with solution NMR. While the concentration of the sample was not sufficient to go forward with solution NMR, these results suggested the successful incorporation of some TMD81 into nanodiscs with this MSP.

Since we suspected TMD81 dimer formation was causing the formation of more than one size of nanodisc complex, incorporation was repeated using a DMPC:ΔH5:TMD81 ratio of 240:4:1 since previous studies have shown that doubling the ratio between the MSP and the membrane protein promotes incorporation of monomeric proteins.[117–119] This ratio was used with newly purified TMD81 (Figure 3.5A) and analyzed by SEC (Figure 3.5B) and SDS-PAGE (Figure 3.5C).

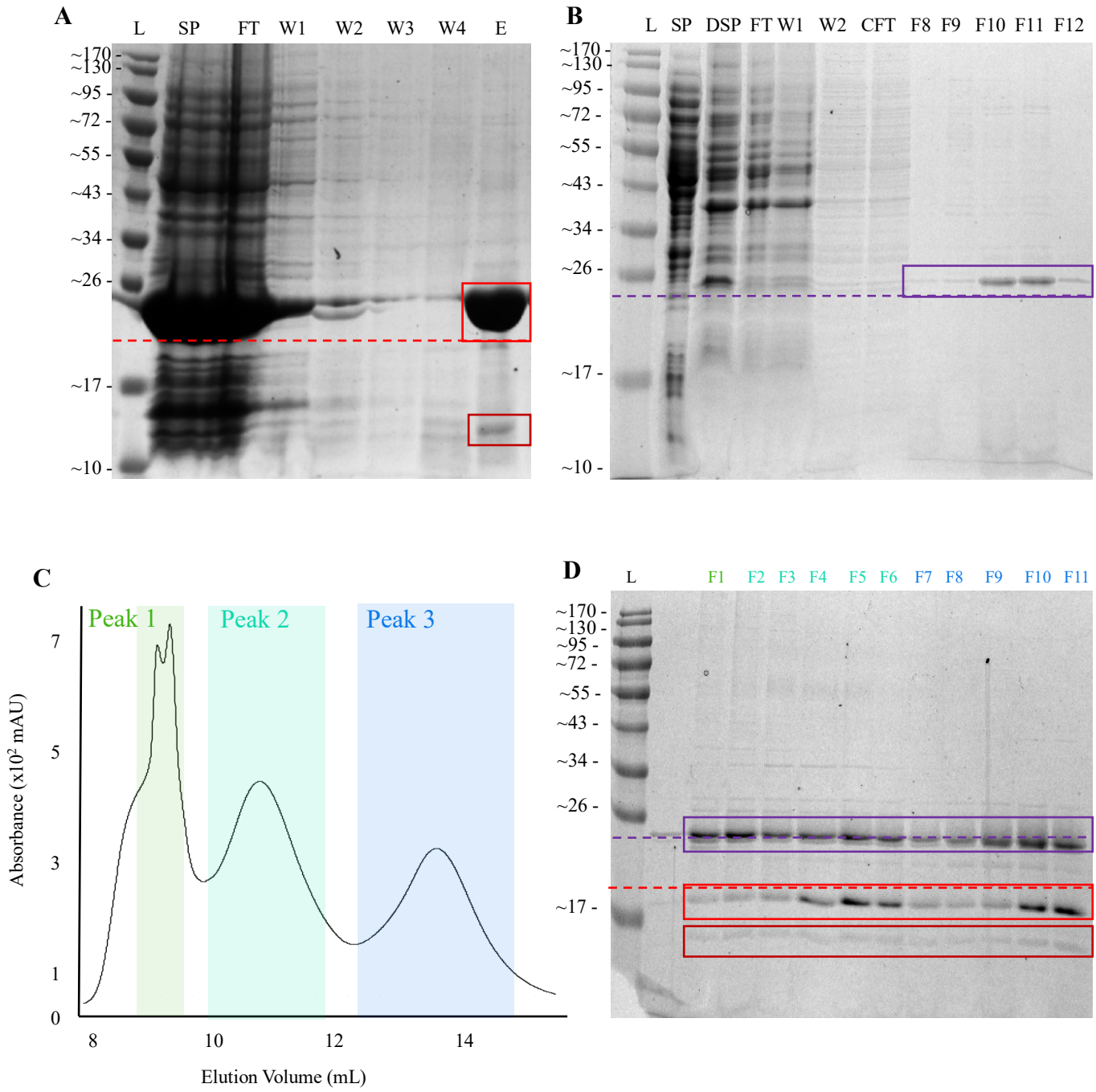

Figure 3.4: Coomassie-stained SDS-PAGE of aliquots taken of (A) ΔH5 (1 mM), and (B) TMD81 (200 µM) samples used to make nanodiscs. (C) SEC on a SP200 column monitored at 280 nm for TMD81 incorporation into ΔH5 DMPC nanodiscs in Lipid Nanodisc buffer. Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113] (D) Coomassie-stained SDS-PAGE of fractions taken from SP200 SEC of TMD81 in ΔH5 DMPC nanodiscs. L: MW ladder in kDa, SP: soluble protein, DSP: detergent soluble protein, FT: flow through, W1, W2, W3, and W4: wash step fractions, E: elution step from Ni-NTA, CFT: MW concentrator flow through, SEC (F1-F12): 0.5 mL fractions from SEC starting at an elution volume of 9 mL (F1).

Light red boxes: full-length ΔH5, dark red boxes: degraded ΔH5, purple boxes: TMD81, red dashed line: expected MW of ΔH5 (22 kDa), purple dashed line: expected MW of TMD81 (25 kDa). Nanodisc reconstitution under these conditions was only attempted once.

As shown in Figure 3.5B, the SEC profile showed a single broad peak eluting between the expected elution volume, ~11 mL and ~14 mL. However, SDS-PAGE analysis of SEC aliquots (Figure 3.5C) had only one band with a mobility that corresponds to that of ΔH5 (Figure 3.4A) and no band seen for TMD81. This suggests formation of TMD81-encapsulalted nanodiscs was not successful and perhaps empty ΔH5 nanodiscs was seen. However, this experiment was only attempted once and should be repeated to make more conclusive statements.

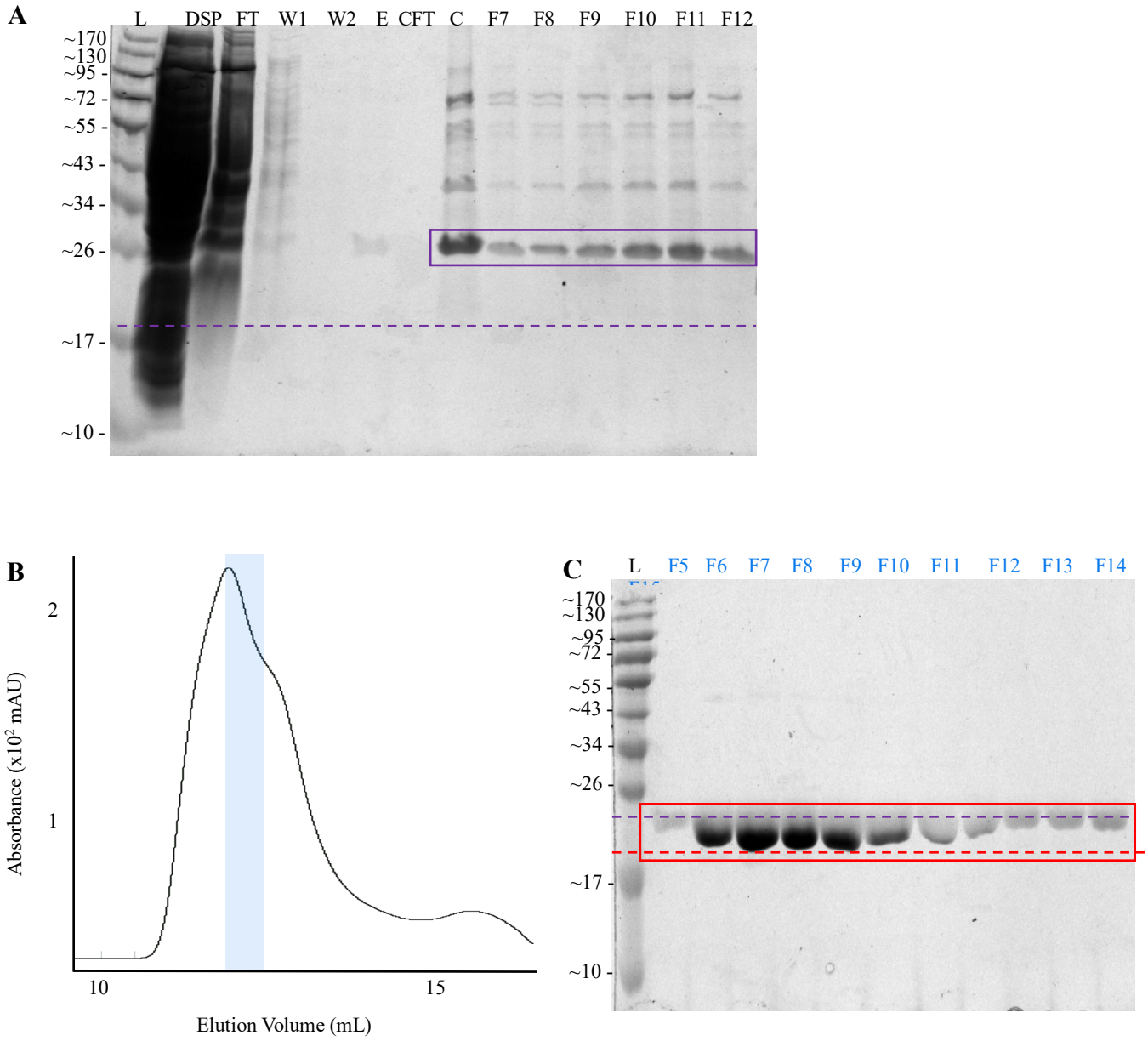

Figure 3.5: (A) Coomassie-stained SDS-PAGE of TMD81 (40 μM) samples used to make nanodiscs. (B) SEC on a SP200 column monitored at 280 nm for TMD81 incorporation in ΔH5 DMPC nanodiscs in Lipid Nanodisc buffer. Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113](C) Coomassie-stained SDS-PAGE of fractions taken from SP200 SEC of TMD81 in ΔH5 DMPC nanodiscs. L: molecular weight ladder in kDa, DSP: detergent fraction after lysis, FT: flow through from the nickel affinity chromatography column, W1 and W2: fractions from the nickel affinity chromatography wash steps, E: eluted fraction, CFT: concentrator flow through,C: pooled, elution fractions after concentration, F9(A), SEC (F5-F15): 0.5 mL fractions from SEC starting at an elution volume of 11 mL (F5). Red box: full-length ΔH5, purple box: TMD81, red dashed line: expected MW of ΔH5 (22 kDa), purple dashed line: expected

MW of TMD81 (25 kDa). Nanodisc reconstitution under these conditions was only attempted once.

3.2.3 Investigation of the effect of removing the ΔH5 His-tag in nanodisc formation

While these results suggest that it was possible to incorporate TMD81 into nanodiscs using the ΔH5 MSP, FPLC profiles showed heterogeneity in complex formation. In an attempt to reduce this heterogeneity, we tested the effect of removing the hexahistidine tag from ΔH5 (ΔH5ΔHis) on nanodisc formation. ΔH5 was purified and then incubated with DTT and TEV protease (TEV+) for 24 hours. As shown in the Coomassie-stained gel (Fig. 3.6A), a new band appeared in the TEV sample that was not present in the control with a MW of ~19 kDa which corresponds to the expected size of ΔH5ΔHis. Unfortunately, it is apparent that cleavage by TEV was not complete, although it was possible to remove some of the uncleaved MSP by passing the mixture through a nickel-NTA resin and collecting the flow-through fraction (FT lane in Figure 3.6A). While it is challenging to determine the exact proportion of uncleaved ΔH5 that remains in the purified sample due to the large amount of sample that was loaded on the gel (Figure 3.6A, lane C), the final ΔH5ΔHis sample (500 μM) and purified TMD81 (Figure 3.6B) was used in the incorporation of TMD81 into DMPC nanodiscs.

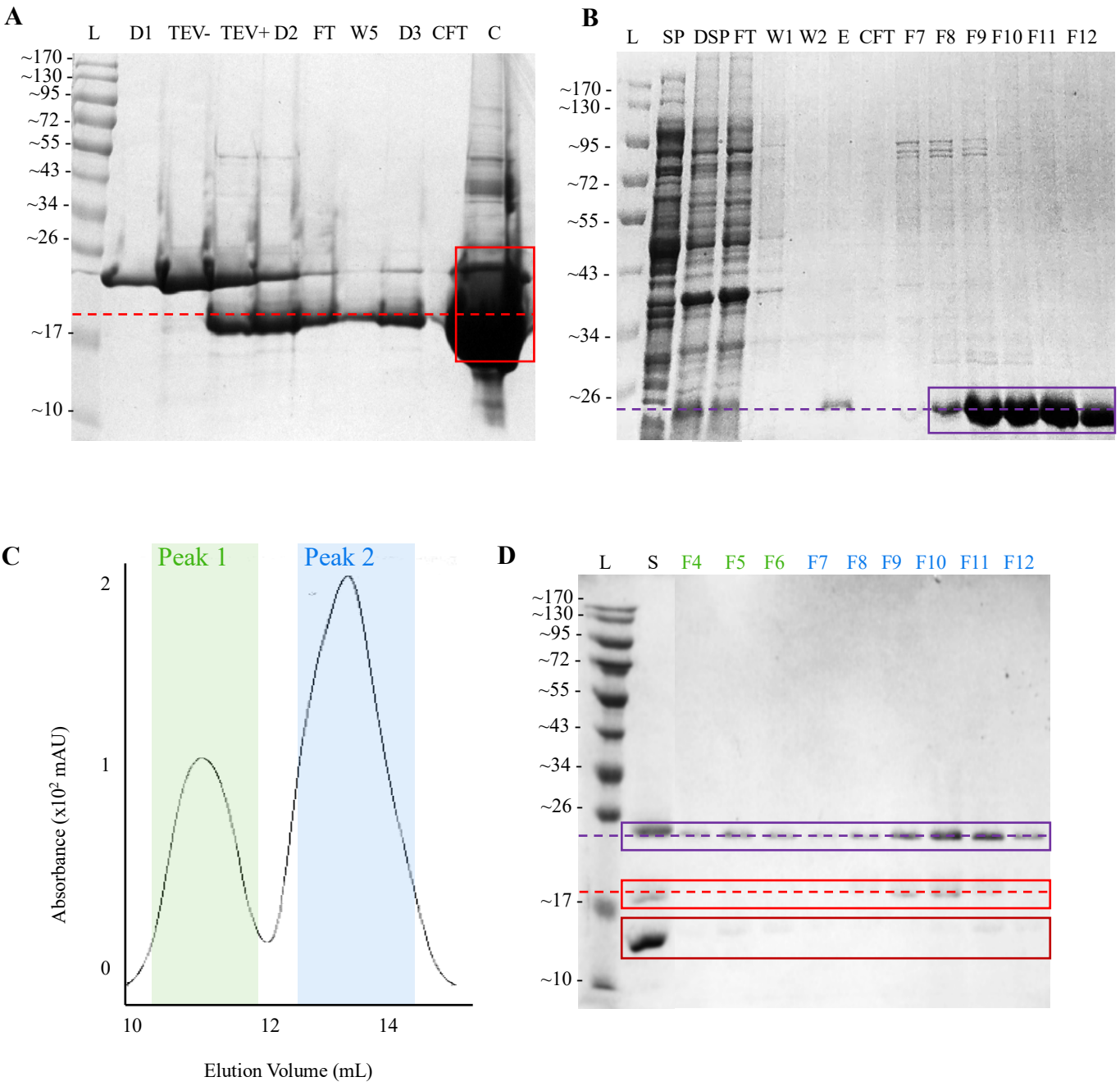
A
L D1 TEV- TEV+ D2 FT W5 D3 CFT C
~170
~130
~95
~72
~55
~43
~34
~26
~17
~10
B
L SP DSP FT W1 W2 E CFT F7 F8 F9 F10 F11 F12
~170
~130
~95
~72
~55
~43
~34
~26
C
Peak 1
Peak 2
Absorbance (x10² mAU)
2
1
0
10
12
14
Elution Volume (mL)
D
L S F4 F5 F6 F7 F8 F9 F10 F11 F12
~170
~130
~95
~72
~55
~43
~34
~26
~17
~10

Figure 3.6: (previous page) Coomassie-stained SDS-PAGE of aliquots taken during (A) TEV protease cleavage of ΔH5 (500 μM) and (B) TMD81 (300 μM) purifications. (C): SEC on a SP200 column monitored at 280 nm for TMD81 incorporation in ΔH5ΔHis DMPC nanodiscs in Lipid Nanodisc buffer. Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113] (D) Coomassie-stained SDS-PAGE of fractions taken from SP200 SEC of TMD81 in ΔH5ΔHis DMPC nanodiscs. L: MW ladder in kDa, D1: sample post-first dialysis, TEV-: D1 sample with DTT but no TEV after 24 hours, TEV+: post-TEV incubation with D1 sample and DTT for 24 hours, D2: sample post-second dialysis, FT: flow through, W5: contaminants removed during Ni-NTA wash step, D3: post-third dialysis sample, CFT: concentrator flow through, C: pooled, elution fractions after concentration, S: sample taken after nanodisc reconstitution, DSP: detergent fraction after lysis, SEC (F4-F12): 0.5 mL fractions from SEC starting at an elution volume of 11 mL (F4). Light red boxes: full-length ΔH5, dark red box: degraded ΔH5, purple boxes: TMD81, red dashed line: expected MW for ΔH5ΔHis (19 kDa), purple dashed line: expected MW for TMD81 (25 kDa). Nanodisc reconstitution under these conditions was only attempted once.

When analyzed by SEC, peaks centered at ~10.5 mL (peak 1) and ~13 mL (peak 2) were seen in the profile (Figure 3.6C). The 13 mL elution volume is similar to the elution volume of the 160 kDa MW standard as published in the manufacturer's specifications.[113] This is consistent with expected values for TMD81 in nanodiscs as some lipids will be replaced during reconstitution.[92,93] The Coomassie-stained gel shows the presence of two species (Figure 3.6D); a ~25 kDa species as expected for TMD81, and a 19 kDa band which corresponds to the expected MW for ΔH5ΔHis, although further analysis would be required to confirm these assignments. The gel also reveals the presence of a species with a MW of approximately 15 kDa that had a greater band intensity than the full-length MSP in the nanodisc mixture prior to the SEC step (Figure 3.6C, lane S). This suggests that a significant amount of ΔH5ΔHis was lost to proteolysis either before or during incorporation into nanodiscs. In spite of this loss of full-length ΔH5ΔHis, these results suggest that some TMD81 could still be incorporated into nanodiscs, albeit with a relatively low yield.

To assess the stability of the nanodiscs, fractions from peak 2 were combined, concentrated, and then subjected to a second round of SEC. As illustrated in figure 3.7A, the sample eluted as

one peak at an elution volume of 10.5 mL which is slightly smaller than that of the 65 kDa MW standard published by the manufacturer. [113] The low absorbance of this peak suggests that a large proportion of sample may have been lost during the concentration and SEC steps (Fig. 3.6C). In addition, this sample showed the presence of a 15 kDa species after concentration (S lane in Figure 3.7B), which may account for some of the sample loss through proteolysis.

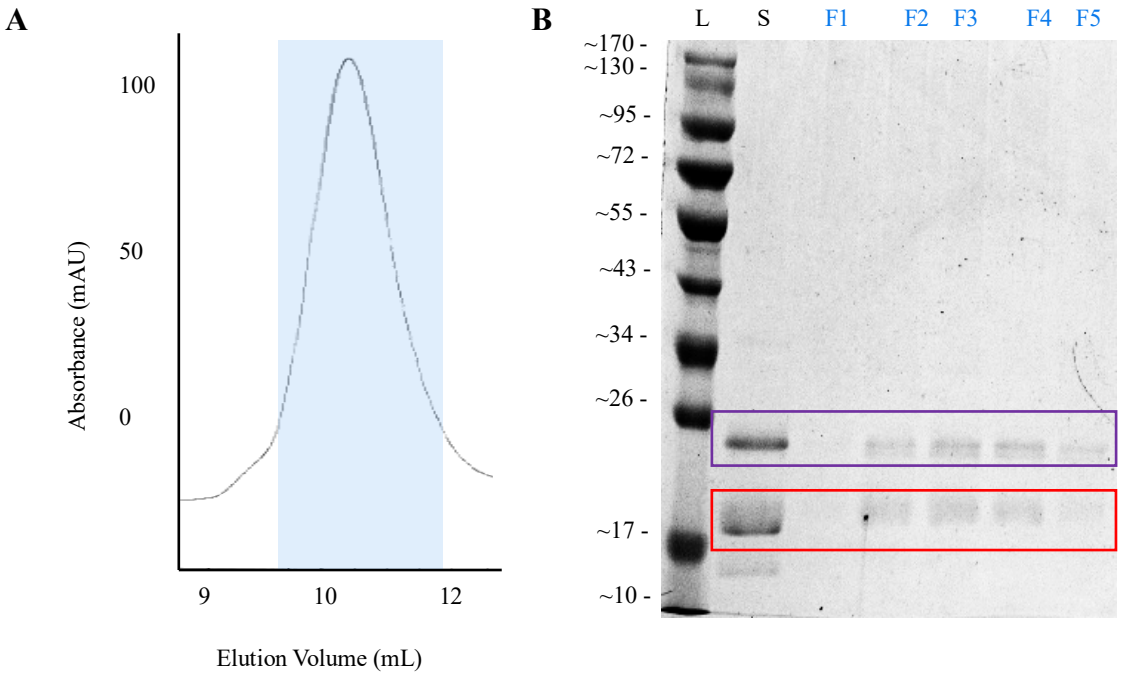

Figure 3.7: (A) SEC profile of nanodiscs (30 μM) run on SP75 column monitored at 280 nm. Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113] (B) Coomassie-stained SDS-PAGE of fractions from (A). L: molecular weight ladder in kDa, S: concentrated sample applied to column, SEC (F1-F5): 0.5 mL fractions from SEC starting at an elution volume of 9 mL (F5). Light red box: ΔH5, purple box: TMD81, red dashed line: expected MW for ΔH5ΔHis (19 kDa), purple dashed line: expected MW for TMD81 (25 kDa). Nanodisc reconstitution under these conditions was only attempted once.

3.3 Assembly of nanodiscs with 1D1

As the results for nanodisc formation with ΔH5 MSP showed some success but ultimately produced heterogenous mixtures that were not appropriate for solution NMR, we decided to evaluate the longer MSP variant, 1D1. The larger nanodiscs formed by 1D1 allows more lipids to separate the encapsulated protein from the surrounding MSP, reducing their interactions which may facilitate more uniform nanodisc formation.[92] As the TMD81 L1 loop (Figure 1.4) which protrudes into the membrane bilayer[67,69] could potentially interact with ΔH5 in a smaller nanodisc, this may have interfered with nanodisc formation. We began these trials by variation of the lipid:1D1 ratio to identify conditions that would produce nanodiscs of a size that would be compatible with solution NMR.

Since it has been established that TMD81 function is modulated by the type of lipid it is reconstituted into,[65] we chose to investigate the utility of nanodiscs for three lipids (DMPC, DMPG, and DLPC) over six ratios (160:2, 140:2, 120:2, 100:2, 80:2, and 60:2).

3.3.1 Investigation of lipid:protein ratios for empty nanodiscs with 1D1

To confirm that protocols for 1D1 nanodiscs could be successfully reproduced in our lab, we first attempted to make empty nanodiscs using DMPC. The purification of 1D1 was completed as described in section 2.8 with aliquots on the Coomassie-stained gel (Figure 3.8) showing the presence of strong bands above the 26 kDa MW marker. The migration of 1D1 is slower than expected for a 24.6 kDa protein possibly due to some interactions that remain despite the addition of a reducing agent, such as SDS and β-mercaptoethanol. However, this observation would need to be confirmed by running another gel with a ladder placed on either side to see if the slower migration was reproducible and not due to discrepancies in gel making protocols. Therefore, all

suggestions made for the nanodisc samples using this purified sample are tentative and would need to be validated. Although some 1D1 was lost in the FT and wash steps (Fig. 3.8, lanes FT, W1 – W4), the single-step purification still gave sufficient yield (2 mg/L) and purity (>80%) to use in nanodisc formation trials.

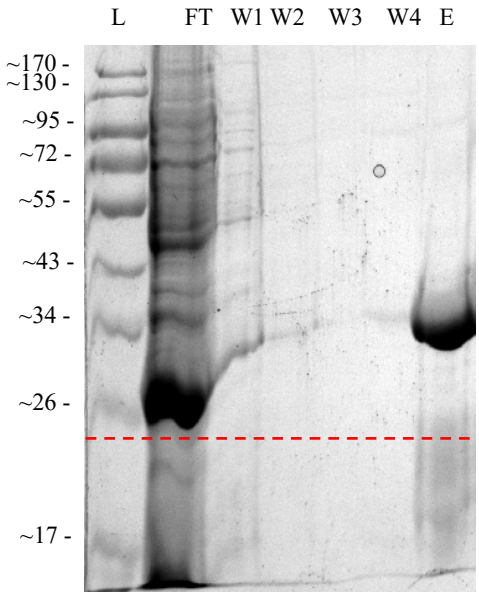

Figure 3.8: Coomassie-stained SDS-PAGE of fractions taken through various steps of the purification of 1D1 (500 μM) in Lipid Nanodisc buffer. L: MW ladder in kDa, FT: flow through, W1, W2, W3, and W4: wash steps, E: elution step. Red dashed line: expected MW for 1D1 (24.1 kDa).

Purified and concentrated 1D1 (500 μM) was used with DMPC to form nanodiscs, which was then subjected to SEC (Figure 3.9). For all ratios the SEC profile showed a peak centered at ~12-12.5 mL (peak 1) and another peak at ~15.5-16 mL (peak 2). Gel analysis shows 1D1 is present in both peaks with higher levels in peak 1 for lipid:1D1 ratios of 120:2-160:2 (Figure 3.10). The elution volume of peak 1 corresponds to a ~160 kDa complex that is within the expected range for successful nanodisc formation. However, the elution volume for peak 2 suggests a MW that is significantly smaller (~30 kDa), that could be due to free 1D1. As illustrated in Figure 3.9, the intensity of peak 1 increases relative to the intensity of peak 2 as the lipid to protein ratio increases,

similar to results from previous studies.[92,120] For lower lipid:protein ratios, specifically 60:2 and 80:2, a ~50 kDa species is seen in the SDS-PAGE gel for peak 1 that may correspond to an SDS-resistant MSP dimer, similar to those formed by ΔH5 (Figure 3.3C). However, in almost all samples there appeared to be evidence of MSP degradation to various degrees, which was particularly pronounced for the 80:2 and 160:2 sample (Figure 3.10). In addition, the bands within the gels show inconsistent migration that may be caused by the interactions with phospholipids in the sample or indicate a problem with the way the experiment was run (eg. inconsistent gel casting, errors in buffers or gel recipes, physical issues with gel apparatus, etc). Therefore, these results would need to be validated by additional trials.

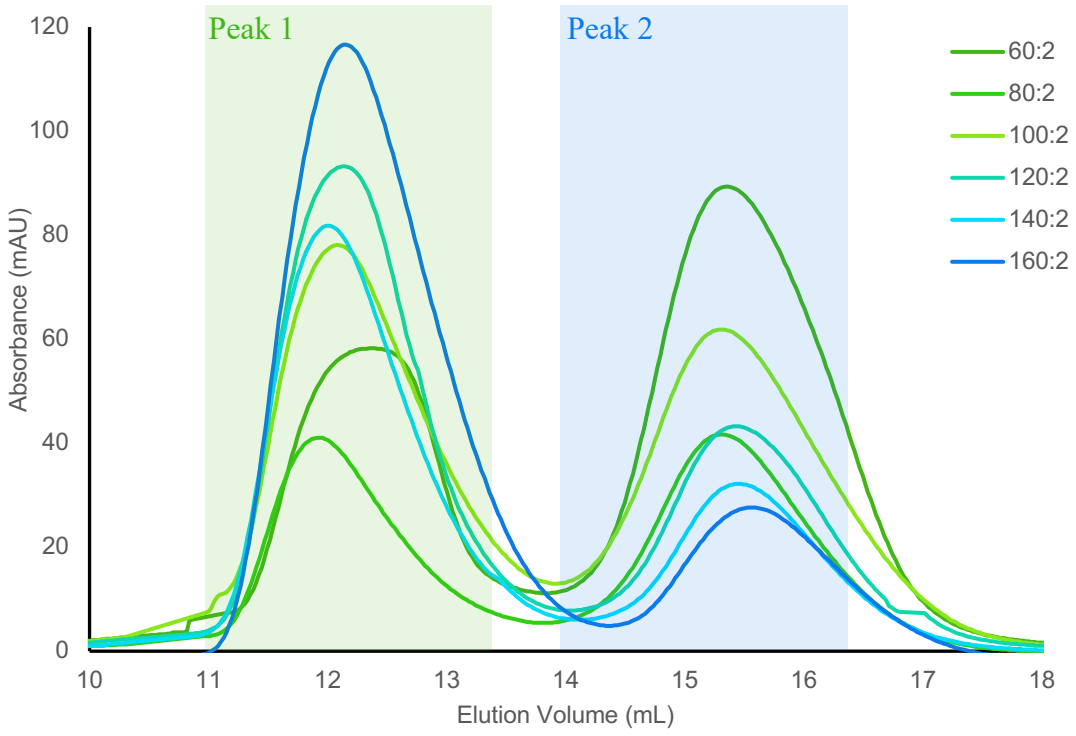

Figure 3.9: SEC profile on a SP200 column for empty 1D1 DMPC nanodiscs monitored by absorbance at 280 nm for DMPC:1D1 ratios: 60:2 (dark green), 80:2 (green), 100:2 (light green), 120:2 (teal), 140:2 (light blue), 160:2 (blue). Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113]

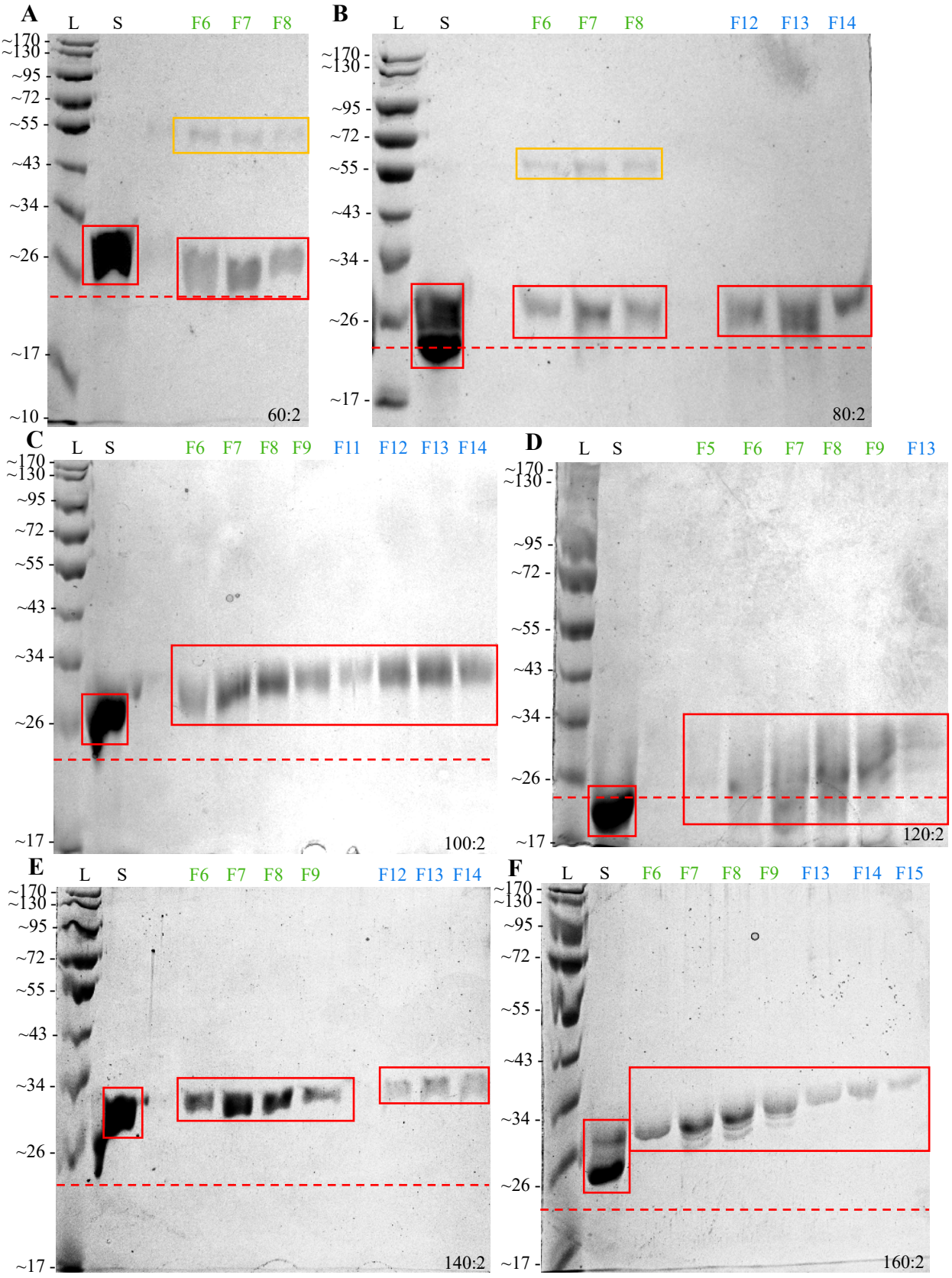
A
L S F6 F7 F8
~170
~130
~95
~72
~55
~43
~34
~26
~17
~10
60:2
B
L S F6 F7 F8 F12 F13 F14
~170
~130
~95
~72
~55
~43
~34
~26
~17
80:2
C
L S F6 F7 F8 F9 F11 F12 F13 F14
~170
~130
~95
~72
~55
~43
~34
~26
~17
100:2
D
L S F5 F6 F7 F8 F9 F13
~170
~130
~95
~72
~55
~43
~34
~26
~17
120:2
E
L S F6 F7 F8 F9 F12 F13 F14
~170
~130
~95
~72
~55
~43
~34
~26
~17
140:2
F
L S F6 F7 F8 F9 F13 F14 F15
~170
~130
~95
~72
~55
~43
~34
~26
~17
160:2

Figure 3.10: (previous page) Coomassie-stained SDS-PAGE gels for samples indicated peaks taken from SEC of nanodiscs made from DMPC:1D1 ratios of (A) 60:2, (B) 80:2, (C) 100:2, (D) 120:2, (E) 140:2, and (F) 160:2. L: MW ladder in kDa, S: sample taken after nanodisc reconstitution, SEC (F5-F15): 0.5 mL fractions from SEC starting at an elution volume of 11 mL (F5). Light red boxes: full-length 1D1, orange boxes: 1D1 dimers (~ 50 kDa), red dashed line: expected MW of 1D1 (24.6 kDa). While nanodisc reconstitution under these conditions was attempted twice for ratios 60:2-120:2, only one trial is represented for all ratios in this figure.

3.3.2 Investigation of empty nanodisc formation with 1D1 and DMPG

We also investigated lipid:1D1 ratios for nanodiscs formed with DMPG, a lipid that is more frequently used to study membrane proteins in nanodiscs through solution-state NMR.[92,96] Nanodisc formation was attempted using previously purified 1D1 (500 μM) (Figure 3.8) and analyzed by SEC and SDS-PAGE (Figure 3.11). For all ratios, only one peak was observed in the profile, all centered around an elution volume of 12.5-13 mL which corresponds to an approximate MW of 160 kDa. In accordance with the DMPC nanodisc trials and previous studies,[92,120] there is a shift to smaller elution volumes as the DMPG:1D1 ratio increases, suggesting a small increase in nanodisc size with increasing lipid amounts. SDS-PAGE analysis (Figure 3.11B-G) suggests that proteolytic degradation of 1D1 occurred to various degrees similar to observations with DMPC nanodiscs (Figure 3.10). The relative amount of the full-length product appears to be highest at the lipid:MSP ratios 80:2 and 100:2, although the inconsistencies in gel migration possibly due to lipid-protein interactions or potential errors with gel preparation complicates the analysis (Figure 3.11B-G).

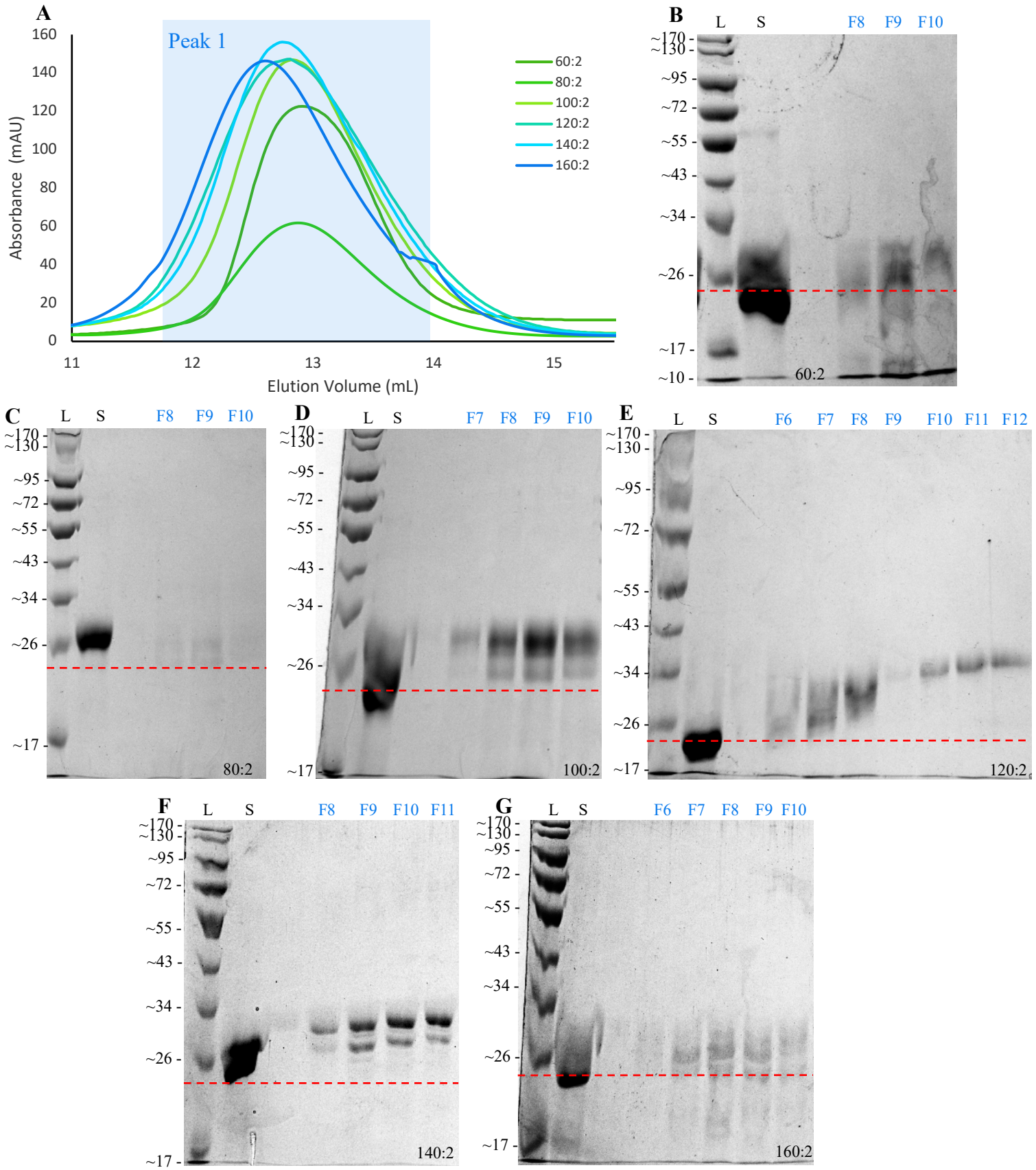
A
Peak 1
60:2
80:2
100:2
120:2
140:2
160:2
Absorbance (mAU)
Elution Volume (mL)
B
L S F8 F9 F10
60:2
C
L S F8 F9 F10
80:2
D
L S F7 F8 F9 F10
100:2
E
L S F6 F7 F8 F9 F10 F11 F12
120:2
F
L S F8 F9 F10 F11
140:2
G
L S F6 F7 F8 F9 F10
160:2
~170
~130
~95
~72
~55
~43
~34
~26
~17
~10

Figure 3.11: (previous page) (A) SEC profile on a SP200 column monitored at 280 nm for empty 1D1 DMPG nanodiscs for DMPG:1D1 ratios: 60:2 (dark green), 80:2 (green), 100:2 (light green), 120:2 (teal), 140:2 (light blue), 160:2 (blue). Absorbance at 280 nm was monitored. Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113] (B-G) Coomassie-stained SDS-PAGE gels for indicated peaks taken from SEC of nanodiscs made from DMPG:1D1 ratios of (B) 60:2, (C) 80:2, (D) 100:2, (E) 120:2, (F) 140:2, and (G) 160:2. L: MW ladder in kDa, S: sample taken after nanodisc reconstitution, SEC (F6-F12): 0.5 mL fractions from SEC starting at an elution volume of 12 mL (F6).Red dashed line: expected MW of 1D1 (24.6 kDa). While nanodisc reconstitution under these conditions was attempted twice for ratios 60:2-120:2, only one trial is represented for all ratios in this figure.

3.3.3 Investigation of nanodisc formation with 1D1 and DLPC

Since previous studies showed higher TMD81 activity in DLPC bicelles,[65] we also attempted to make nanodiscs with this lipid using 1D1. Following the same protocols used for the C-14 lipids, we observed extensive aggregation of nanodisc mixtures prior to FPLC analysis. These difficulties with empty nanodisc formation with DLPC may be a result of instability caused by ellipticity and high lateral pressure.[94] Since there are also no examples we could find in the literature of DLPC nanodiscs with an incorporated integral membrane protein, no further attempts were made to produce these nanodiscs.

3.3.4 Investigation of nanodisc formation with 1D1ΔHis and DMPC

All nanodisc attempts with 1D1 were observed to suffer from proteolysis to various degrees and may reflect loss of the hexahistidine tag which would be expected to lack structure, and therefore be more susceptible to cleavage. For this reason, we attempted to first remove the his-tag from 1D1 (1D1ΔHis) (Figure 3.11A) as this was expected to improve sample homogeneity although the 1D1ΔHis band was hard to discern due to its relatively low concentration. After purification protease inhibitor was added to each sample to inhibit any further protein degradation.[92]

1D1ΔHis (200 μM) samples (Figure 3.12A) combined with DMPC in ratios, 120:4 – 320:4 showed a similar trend to that seen previously on SEC (Figure 3.12B), with two peaks centered at 8.9-9.5 mL (peak 1) and 11.8-12 mL (peak 2) with peak 1 being the dominant species at higher lipid: 1D1ΔHis ratios. The elution volume of peak 1 corresponds to a complex that is very large (~650 kDa) suggesting aggregate formation may have occurred. Peak 2 represents a species that is ~165 kDa that could represent successful nanodisc formation. SDS-PAGE analysis (Figure 3.12C-H) was conducted using SYPRO-Ruby stain as sample concentrations were too low to detect using Coomassie stain. These gels showed a band with a slower mobility than the expected 22 kDa size (Figure 3.12C-H) as observed previously for full-length 1D1 (Figure 3.10). Other bands were also present in these samples, including a lower MW species (~17 kDa) that was present prior to SEC in some of the samples (i.e. 160:4, 280:4 and 320:4). In addition, diffuse intensity appearing at a higher MW suggests the possible presence of aggregates in all samples to varying degrees. These results were surprising as successful formation of nanodiscs should give rise to complexes that are stable in subsequent chromatography steps. The low quality of the gels due to background noise prevents accurate assignment of nanodisc compositions.

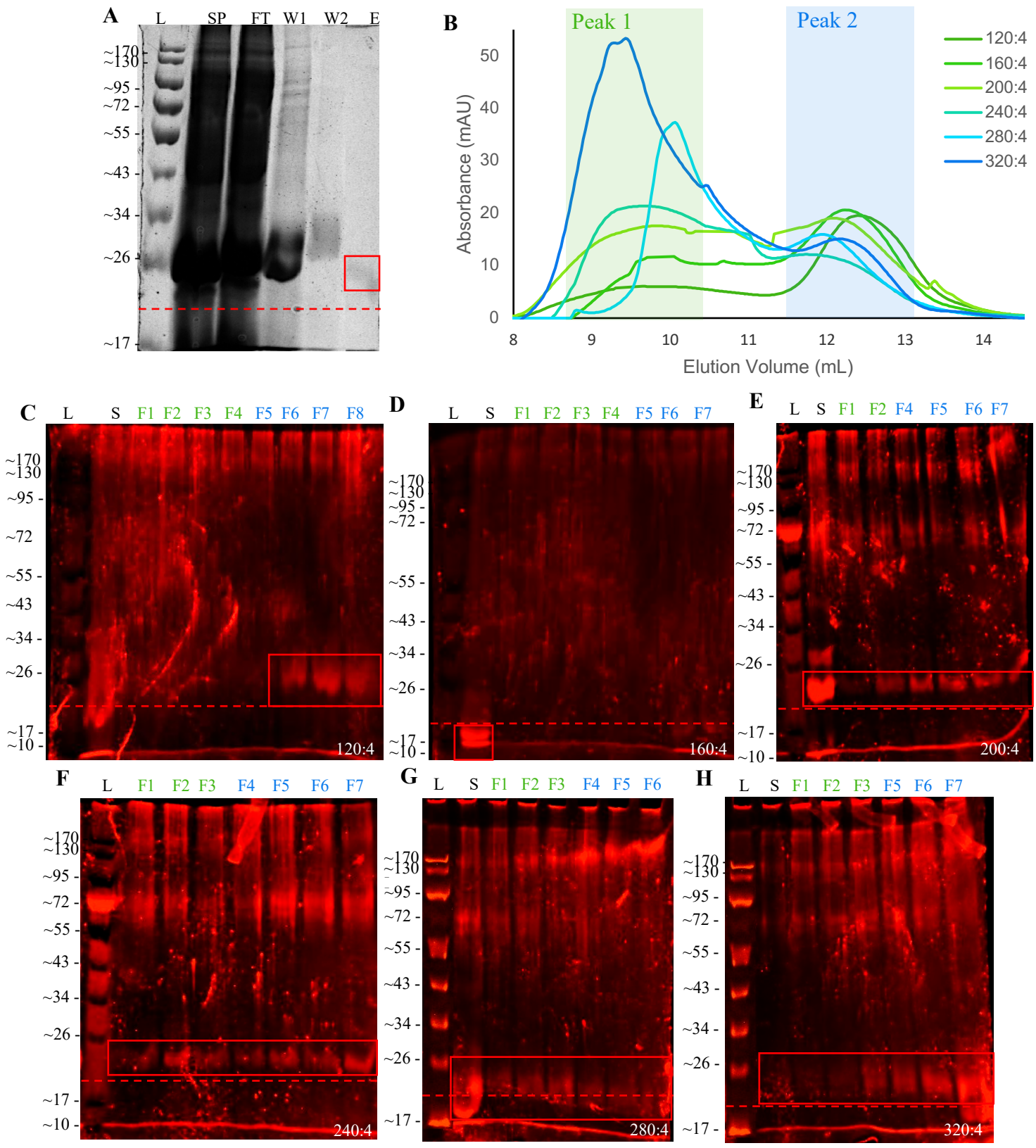
A
L
SP
FT
W1
W2
E
~170
~130
~95
~72
~55
~43
~34
~26
~17
B
Peak 1
Peak 2
120:4
160:4
200:4
240:4
280:4
320:4
Absorbance (mAU)
Elution Volume (mL)
C
L S F1 F2 F3 F4 F5 F6 F7 F8
120:4
D
L S F1 F2 F3 F4 F5 F6 F7
160:4
E
L S F1 F2 F4 F5 F6 F7
200:4
F
L F1 F2 F3 F4 F5 F6 F7
240:4
G
L S F1 F2 F3 F4 F5 F6
280:4
H
L S F1 F2 F3 F5 F6 F7
320:4

Figure 3.12: (previous page) (A) 1D1ΔHis (200 μM) purification Coomasie-stained SDS-PAGE gel (B) SEC profile on a SP200 column monitored at 280 nm for empty 1D1ΔHis DMPC nanodiscs for DMPC:1D1ΔHis ratios: 120:4 (dark green), 160:4 (green), 200:4 (light green), 240:4 (teal), 280:4 (light blue), 320:4 (blue). Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113] (C-H) SYPRO-Ruby stained SDS-PAGE gels for SEC fractions for indicated peaks taken from SEC of nanodiscs made from DMPC:1D1ΔHis ratios of (C) 120:4, (D) 160:4, (E) 200:4, (F) 240:4, (G) 280:4, (H) 320:4. L: MW ladder in kDa, M: 1D1ΔHis sample, S: sample taken after nanodisc reconstitution, SEC (F1-F7): 0.5 mL fractions from SEC starting at an elution volume of 9 mL (F1).Light red boxes: full-length 1D1, red dashed line: expected MW of 1D1ΔHis (22 kDa). While nanodisc reconstitution under these conditions was attempted twice for ratios 60:2-120:2, only one trial is represented for all ratios in this figure.

3.3.5 Investigation of TMD81-encapsulation for 1D1ΔHis and DMPC nanodiscs

Since experiments with empty nanodiscs were inconclusive, 1D1ΔHis (200 μM) nanodisc formation was attempted with purified TMD81 (200 μM) (Figure 3.13A) in DMPC in ratios 120:4:1 – 320:4:1 and analysed through SEC (Figure 3.13B) and SDS-PAGE (Figure 3.13C-H). Similar to results for empty nanodiscs, SEC showed two peaks centered at 8.6 – 9.3 mL (peak 1) and ~12 mL (peak 2) corresponding to >670 kDa aggregates and ~160 kDa nanodiscs respectively. SYPRO-stained gels all showed the presence of high MW aggregates (Figure 3.13C-H). At the lower lipid:protein ratios, specifically 120:4:1, and 160:4:1 SEC peak intensities were very low (≤20 mAU) and no species at the expected MW for 1D1ΔHis (22 kDa) or TMD81 (25 kDa) was seen in the gel. A band was visible at ~25 kDa in fractions taken from sample prepared with higher lipid:protein ratios. Due to the similarity in migration of TMD81 with the slowed migration for 1D1ΔHis it difficult to resolve the two species on the gel. In addition, the poor quality and presence of background noise makes it impossible to draw conclusions from this data.

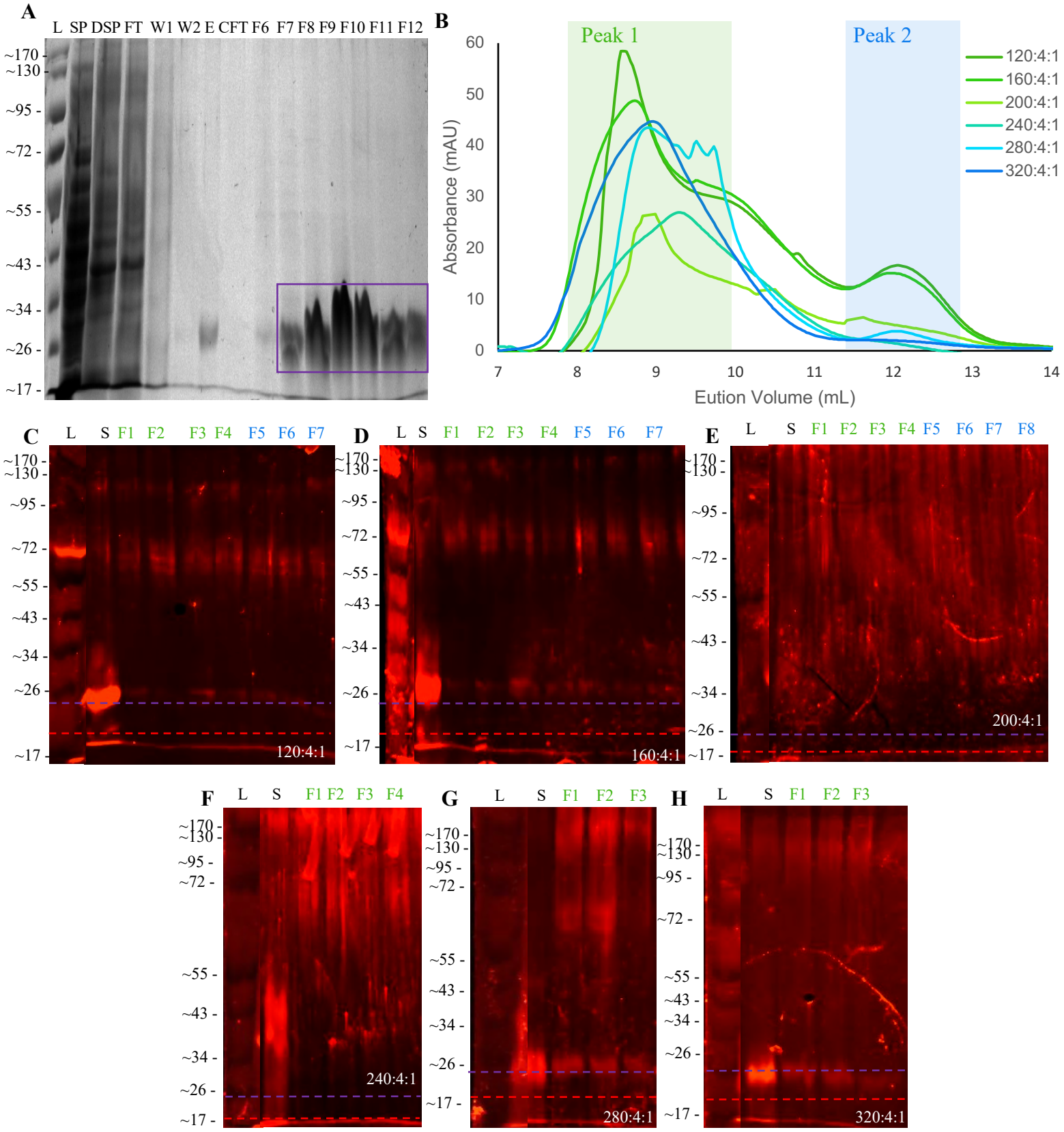
A
L SP DSP FT W1 W2 E CFT F6 F7 F8 F9 F10 F11 F12
~170
~130
~95
~72
~55
~43
~34
~26
~17
B
Peak 1
Peak 2
120:4:1
160:4:1
200:4:1
240:4:1
280:4:1
320:4:1
Absorbance (mAU)
Eution Volume (mL)
C
L S F1 F2 F3 F4 F5 F6 F7
120:4:1
D
L S F1 F2 F3 F4 F5 F6 F7
160:4:1
E
L S F1 F2 F3 F4 F5 F6 F7 F8
200:4:1
F
L S F1 F2 F3 F4
240:4:1
G
L S F1 F2 F3
280:4:1
H
L S F1 F2 F3
320:4:1

Figure 3.13: (previous page) (A) TMD81 (200 μM) purification Coomasie-stained SDS-PAGE gel (B) SEC profile on a SP200 column for TMD81 in 1D1ΔHis DMPC nanodiscs for DMPC:1D1ΔHis:TMD81 ratios: 120:4:1 (dark green), 160:4:1 (green), 200:4:1 (light green), 240:4:1 (teal), 280:4:1 (light blue), 320:4:1 (blue). Absorbance at 280 nm was monitored. Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113] (C-H) SYPRO-Ruby stained SDS-PAGE gels for indicated peaks taken from SEC of nanodiscs made from DMPC:1D1ΔHis:TMD81 ratios of (C) 120:4:1, (D) 160:4:1, (E) 200:4:1, (F) 240:4:1, (G) 280:4:1, (H) 320:4:1. L: MW ladder in kDa, M: 1D1ΔHis sample, SP: soluble fraction after lysis, DSP: detergent fraction after lysis, FT: flow through from the nickel affinity chromatography column, W1 and W2: fractions from the nickel affinity chromatography wash steps, E: eluted fraction, CFT: concentrator flow through, S: sample taken after nanodisc reconstitution, SEC (F1-F12): 0.5 mL fractions from SEC starting at an elution volume of 9 mL (F1). Red dashed line: expected MW of 1D1ΔHis (22 kDa), purple dashed line: expected MW for TMD81 (25 kDa). Nanodisc reconstitution under these conditions was only attempted once.

3.3.6 Investigation of nanodisc formation for 1D1ΔHis and DMPG

Empty nanodisc formation with 1D1ΔHis (200 μM) was attempted using DMPG in the presence of a protease inhibitor cocktail (Figure 3.14A). SEC analysis showed a peak centered at ~8.4 mL (peak 1) for ratio 320:4 that evolved into a peak centered at ~ 12 mL (peak 2) for 120:4 (Figure 3.14A). Peak 1 corresponds to a >670 kDa species whereas, peak 2 is closer to the expected MW of ~160 kDa for nanodiscs, with maximal yield at a ratio of 160:4 (Figure 3.14A). SDS-PAGE analysis shows the expected migration for a 22 kDa species, with minimal levels of uncleaved 1D1. There was evidence of higher MW aggregate in the gel for all ratios (Figure 3.14B-G). Based on peak intensities and 1D1ΔHis band intensities, 160:4 – 240:4 appear to be most promising for incorporation of TMD81. Since we were interested in analyzing a sample by solution NMR, we chose to incorporate TMD81 into 1D1ΔHis DMPG nanodiscs using the lower ratio of 120:4:1 since we expect that some lipids in the nanodisc will be replaced by TMD81.

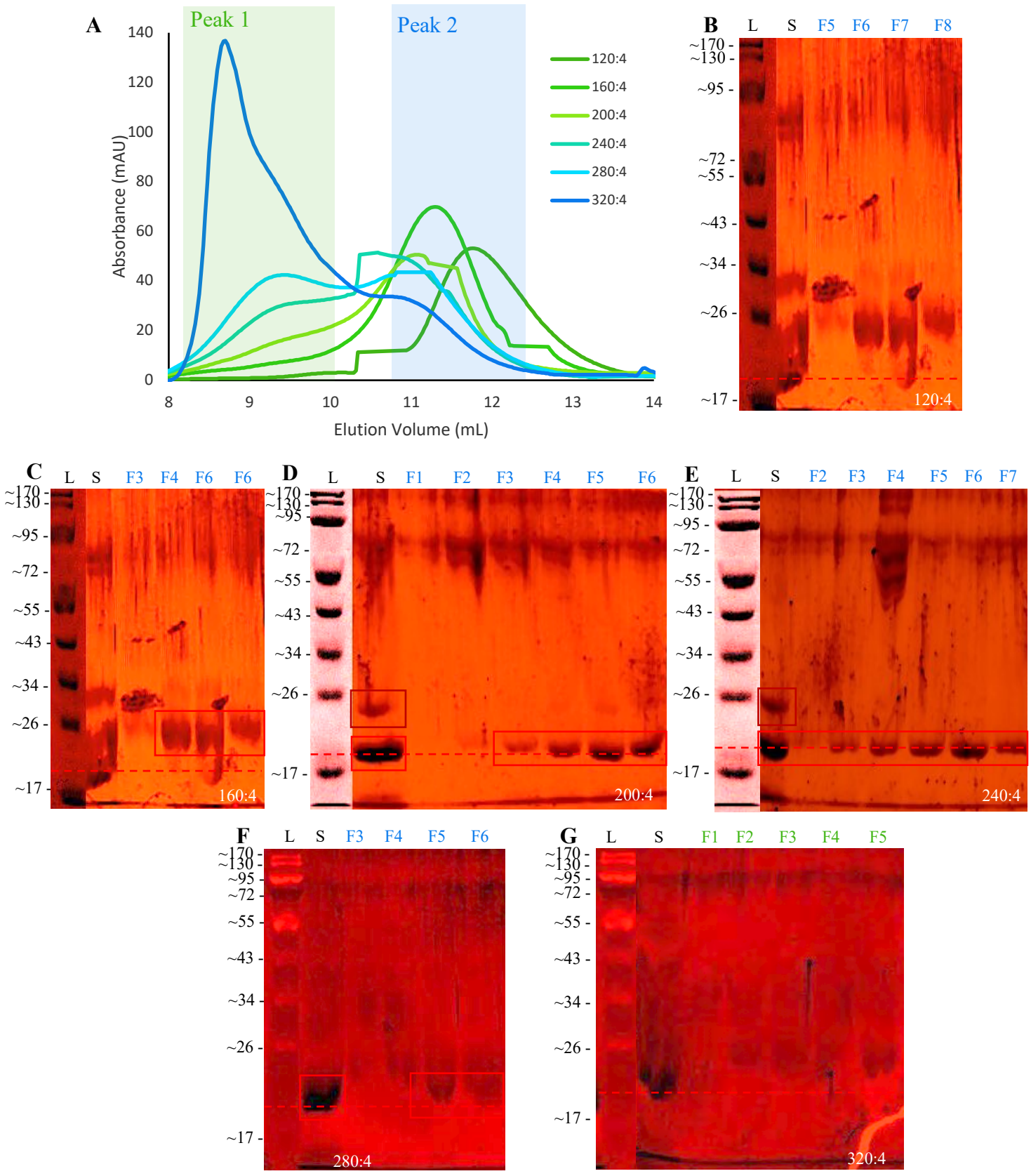
A
Peak 1
Peak 2
120:4
160:4
200:4
240:4
280:4
320:4
Absorbance (mAU)
Elution Volume (mL)
B
L S F5 F6 F7 F8
120:4
C
L S F3 F4 F6 F6
160:4
D
L S F1 F2 F3 F4 F5 F6
200:4
E
L S F2 F3 F4 F5 F6 F7
240:4
F
L S F3 F4 F5 F6
280:4
G
L S F1 F2 F3 F4 F5
320:4
~170
~130
~95
~72
~55
~43
~34
~26
~17

Figure 3.14: (previous page) (A) SEC profile on a SP200 column monitored at 280 nm for empty 1D1ΔHis DMPG nanodiscs for DMPG:1D1ΔHis ratios: 120:4 (dark green), 160:4 (green), 200:4 (light green), 240:4 (teal), 280:4 (light blue), 320:4 (blue). Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113] (B-G) SYPRO-Ruby stained SDS-PAGE gels for indicated peaks taken from SEC of nanodiscs made from DMPG:1D1ΔHis ratios of (B) 120:4, (C) 160:4, (D) 200:4, (E) 240:4, (F) 280:4, (G) 320:4. L: MW ladder in kDa, M: 1D1ΔHis sample, S: sample taken after nanodisc reconstitution, SEC (F1-F8): 0.5 mL fractions from SEC starting at an elution volume of 9 mL (F1). Dark red boxes: full-length 1D1 (25 kDa), light red boxes: 1D1ΔHis (22 kDa), red dashed line: expected MW of 1D1ΔHis (22 kDa). Nanodisc reconstitution under these conditions was only attempted once.

3.3.7 NMR Spectroscopy of WT TMD81 in 1D1ΔHis and DMPG nanodiscs

The incorporation of TMD81 into 1D1ΔHis DMPG nanodiscs at the ratio of 120:4:1 was conducted by combining purified TMD81 (200 μM) (Figure 3.15A) and 1D1ΔHis (200 μM) (Figure 3.12A) with DMPG and protease inhibitor. SEC analysis (Figure 3.15B) yielded a broad peak centered at ~12 mL suggesting the formation of a ~160 kDa complex. While the peak had a shoulder eluting at lower volumes, fractions that represent elution volumes 11-13 mL (fractions 5-9) were combined to prepare a sample for NMR. SDS-PAGE analysis (Figure 3.15C) of fractions from SEC and the concentrated NMR sample show two species at MW ~25 kDa and ~28 kDa. It is difficult to assign these species as the 28 kDa species is larger than predicted for either 1D1ΔHis (22 kDa) or TMD81 (25 kDa). However, since the TMD81 in the sample was ^{15}N-labeled and we observed two bands, as expected, in the gel, we proceeded to record an ^{1}H-^{15}N HSQC-TROSY experiment (Figure 3.16 labelled in red). Not unexpectedly for a complex of this size and the low concentration we anticipated of TMD81 in the sample (~15 μM), very little intensity was observed in this spectrum (Figure 3.16). Although there is a small number of very broad peaks in the central part of the spectrum, it is not possible to determine if this signal originates from TMD81.

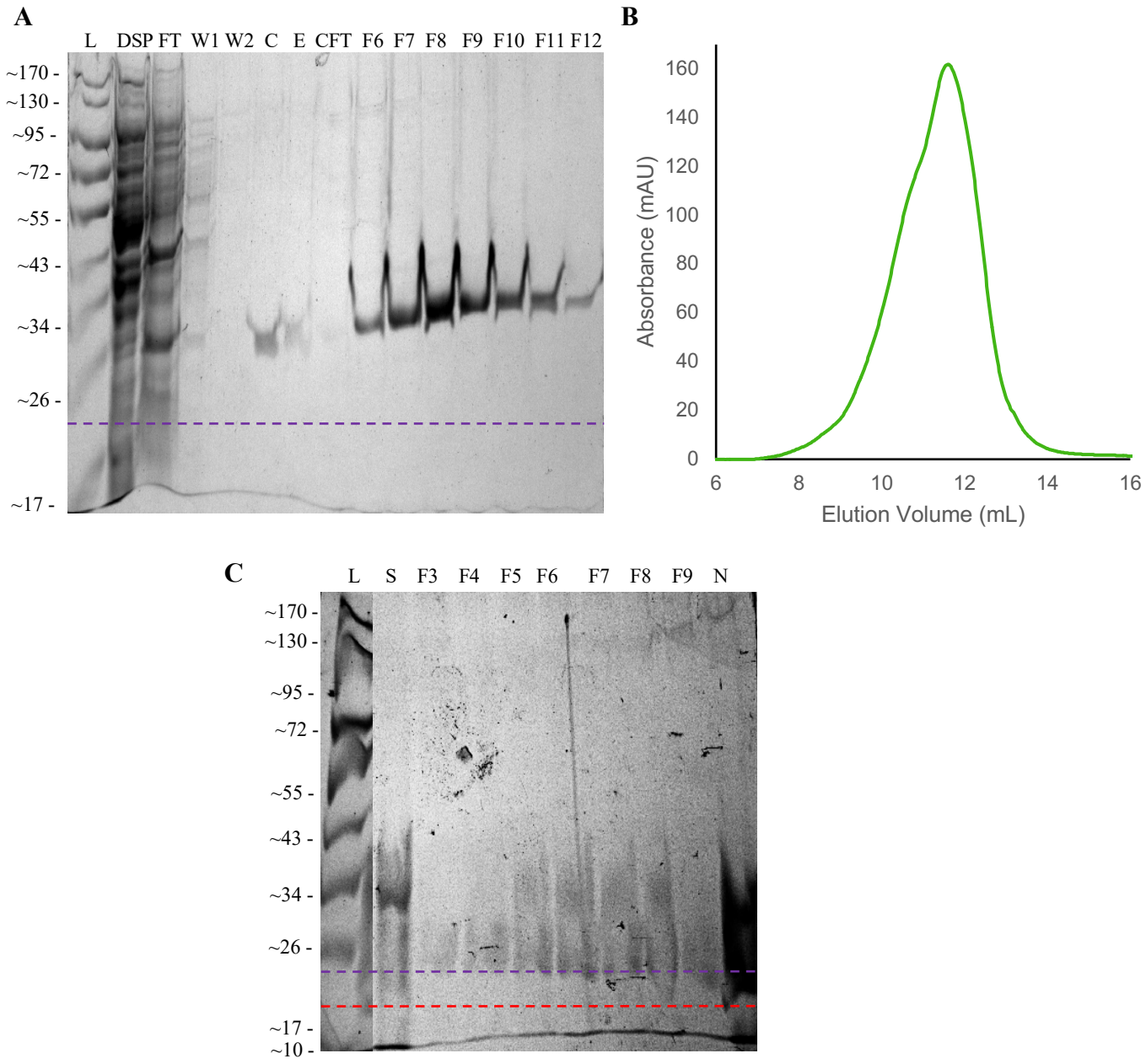

Figure 3.15: (A) Pure ^{15}N TMD81 (200 µM) sample used for nanodisc formation. (B) SEC run on a SP200 column monitored at 280 nm for ^{15}N TMD81 in 1D1ΔHis (22 kDa) DMPG nanodiscs (C) Coomassie-stained SDS-PAGE of fractions from SEC for ^{15}N TMD81 in 1D1ΔHis DMPG nanodiscs. Estimates of MW are based on SEC profiles of MW standards published by the manufacturer.[113] L: molecular weight ladder in kDa, DSP: detergent fraction after lysis, FT: flow through from the nickel affinity chromatography column, W1 and W2: fractions from the nickel affinity chromatography wash steps, C: pooled, elution fractions after concentration, E: eluted fraction, CFT: concentrator flow through, S: sample taken after nanodisc reconstitution, SEC (F3-F12): 0.5 mL fractions from SEC starting at an elution volume of 10 mL (F3), N: concentrated NMR sample (F5-9) with 10% D_2O. Red dashed line: expected MW of 1D1ΔHis (22 kDa), purple

dashed line: expected MW for TMD81 (25 kDa). Nanodisc reconstitution under these conditions was only attempted once.

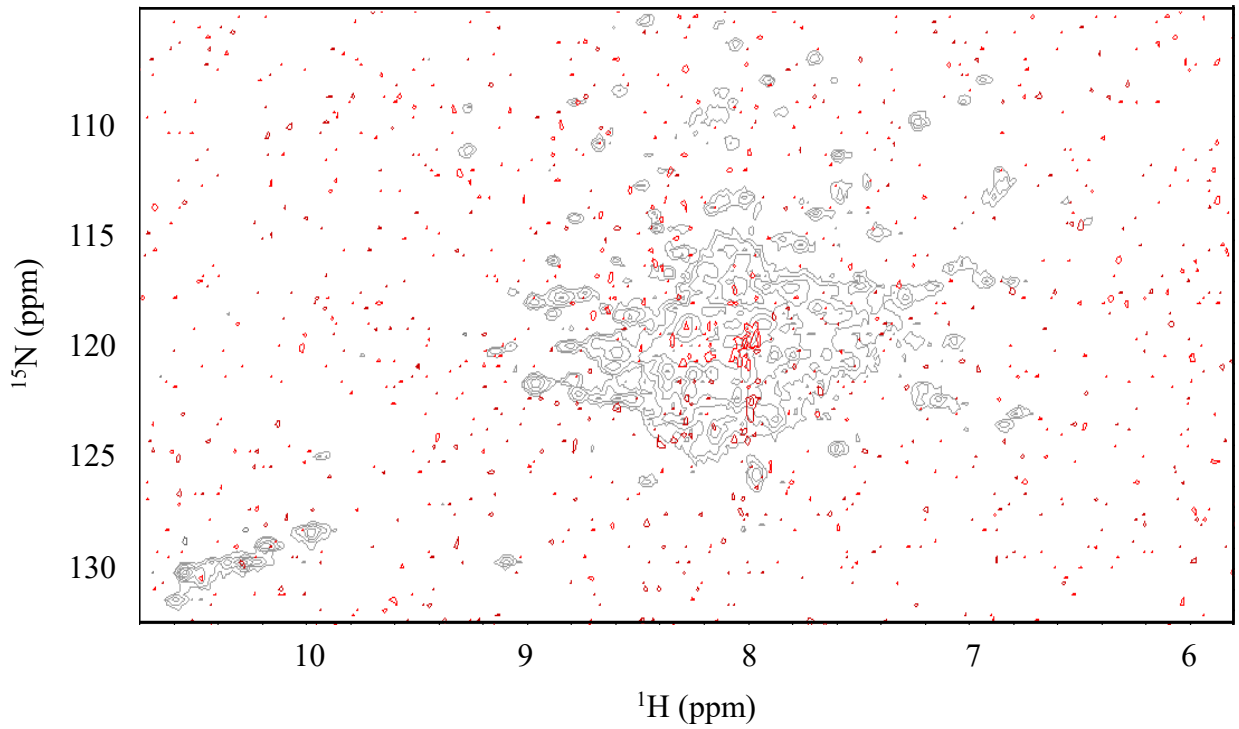

Figure 3.16: Superposition of ^{1}H-^{15}N HSQC spectrum for ^{15}N TMD81 (100 μM) in DPC detergent micelles (grey) and ^{15}N TMD81 (~15 μM) in 1D1ΔHis DMPG nanodiscs (red) each with 10% D_2O and at 45°C. Spectra were acquired using a 500 MHz spectrometer equipped with a triple resonance cryoprobe.

3.3.8 Assessment of nanodisc composition through MALDI mass spectrometry

To determine if the 1D1ΔHis samples correspond to the expected products of TEV protease cleavage we conducted matrix-assisted laser desorption/ionization (MALDI) MS for the 120:4:1 (DMPG:1D1ΔHis:TMD81) sample. As shown in Figure 3.17B in addition to the expected species at ~ 22 kDa that should correspond to 1D1ΔHis, multiple peaks of lower MW also appeared that may represent degradation products of 1D1ΔHis and/or TMD81 (~17, ~13, ~12, ~11, and ~8 kDa). Although no peak intensity corresponding to the MW of TMD81 (25 kDa) was observed, a control MALDI-MS spectrum (Figure 3.17A) acquired for a concentrated TMD81 (200 μM) sample in

DPC detergent micelles also did not show any peaks for this species. This reflects the difficulty of obtaining MALDI-MS for integral membrane proteins, a challenge that may be further exacerbated by the presence of the MSP nanodisc. Nonetheless, this experiment highlighted the problems that continue to persist with proteolysis of the MSP samples during nanodisc reconstitution, despite the inclusion of protease inhibitors in this solution.

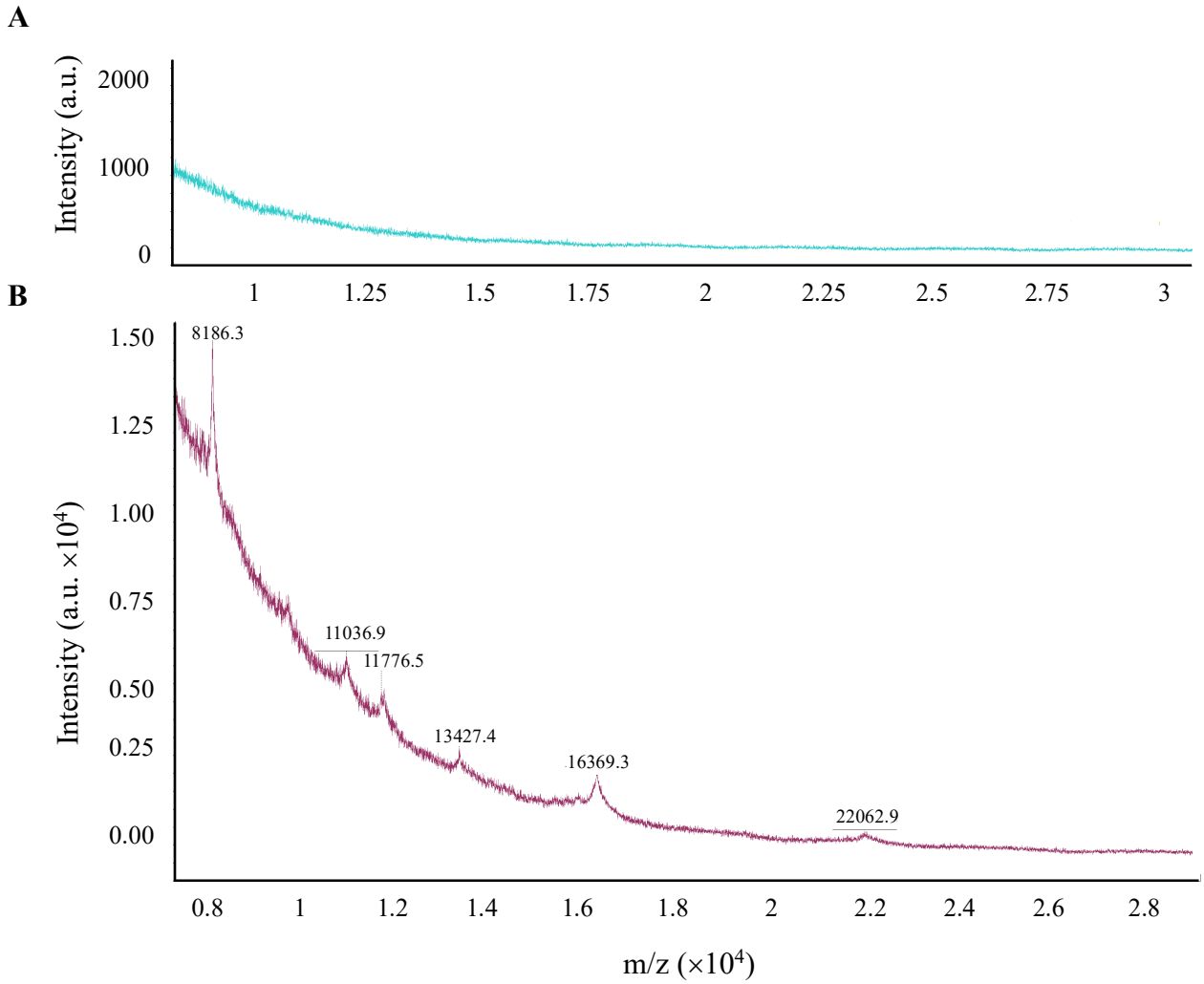

Figure 3.17: MALDI-MS for (A) TMD81 (200 μM) in DPC detergent micelles, and (B) TMD81 (~15 μM) encapsulated in 1D1ΔHis DMPG nanodiscs. Conducted on a Microflex MALDI-TOF mass spectrometer with a m/z window of 0-30 kDa. The matrix used for both samples was 4-HCCA. Spectra were acquired through the culminations of signals obtained using the sandwich and mixed method.

Chapter 4: Discussion and Conclusion

4.1 Challenges with TMD81 incorporation into MSP nanodiscs

The results for all nanodisc experiments were significantly affected by two challenges; significant proteolysis of both MSP variants (ΔH5 and 1D1) in all samples, and very low yields of nanodiscs even when nanodisc reconstitution was successful.

Proteolysis of both MSP variants, ΔH5 and 1D1, was observed in empty and TMD81-encapsulated nanodiscs regardless of the lipid and ratio used. While we attempted to reduce the proteolysis by adding protease inhibitor into the reconstitution mixture before the addition of bio-beads, proteolysis was still observed albeit to a lesser extent. This observation was puzzling as the inhibitor cocktail contained a broad range of inhibitors for cysteine (E64)[121], serine (AEBSF)[122], aminopeptidase (bestatin)[123], aspartic acid (pepstatin)[124], and metalloproteases (phosphoramidon)[125] similar to the protease cocktail used for other nanodisc preparations.[92] Since the concentrations of protease inhibitor added to the reconstitution mixture should have been sufficient for full inhibition, we propose that the protease inhibitors present in the cocktails not used in our study may have been more effective than the ones used in our work. These include D-mannitol,[126] EDTA,[127] and poly vinyl pyrrolidone K90 (phenolic inhibitor)[128], for inhibiting cysteine, aspartic acid, metalloproteases, and aminopeptidase proteases. Also, it is possible that our protease inhibitor cocktail lost potency during the nanodisc reconstitution process since the mixture is typically stored at room temperature for over 24 hours during detergent extraction with bio-beads. The temperature we used for nanodisc formation follows previously established protocols[92] to ensure that it remains above the lipid transition temperature and could be circumvented in future attempts at nanodisc reconstitution by adding protease inhibitor cocktail to the reconstitution mixture at several points in the procedure; for example, during the MSP

purification, prior to bio-bead addition, 12 hours through bio-bead incubation, and prior to SEC analysis.

While there are no studies that examine the rate of MSP proteolysis after nanodisc formation, proteolysis of 1D1 in nanodiscs with DMPC, DMPG, and 1,2-dimyristoyl-sn-glycero-3-phosphate (sodium salt) (DMPA), has been previously observed.[129] DMPC and DMPA nanodiscs showed significant instability on SEC profiles after one week of storage as seen by the presence of additional peaks that represent higher and lower MW species. While it is possible that the instability of 1D1 in DMPC and DMPA nanodiscs arose from inappropriate lipid to 1D1 ratios,[129] our investigations showed MSP degradation in all samples to varying degrees, even in lipid:MSP ratios that have been shown in previous studies to produce stabilized nanodiscs.[92,96] Since the susceptibility of MSP proteolysis has been previously observed,[102] it may be helpful in the future to subject MSP samples to an additional purification step, the most obvious choice being size-exclusion chromatography (SEC) that has been used in previous studies to get rid of MSP degradation products as well as contaminating proteases.[130]

Throughout the trials of nanodisc incorporation, yields were found to be low, even when favorable lipid to protein ratios were used. While MSP degradation was an obvious contributor to these low yields, it is also possible that oligomerization of TMD81 could have complicated the generation of homogeneous nanodisc preparations. Previous studies with full-length *ec*GlpG, as well as other rhomboid proteases (AarA[52], *hi*GlpG[52,131], YqgP[131]), have shown significant dimer formation in detergent micelles[52,58,131] through sedimentation equilibrium analysis.[131] In addition, dimerization of TMD81 was observed with the potential formation of a tetrameric state in a subset of detergent micelles types (i.e. Mal-12, Lyso-Fos-12, Lyso-Fos-14) as suggested by SEC profiles.[58] Furthermore, kinetic studies performed on AarA with its physiological substrate

revealed cooperative activity with a hill coefficient that was close to 2, suggesting the dimer is the physiologically-relevant functional state.[52] In contrast, the Urban group used Halo tagging at the C-terminus to measure photobleaching rates for *ec*GlpG in membranes of various thickness and composition, and observed the presence of monomeric protein in all cases. While these results suggest that the physiologically relevant functional state for *ec*GlpG may be monomeric, the presence of multiple oligomeric states in detergent micelles presents challenges during our reconstitution that may have increased the heterogeneity of nanodiscs in our preparations. While we attempted to reduce the oligomerization of TMD81 by doubling the ratio of MSP, this was unsuccessful and appeared to result in the formation of empty nanodiscs as illustrated in Figure 3.5.

While *ec*GlpG oligomerization may contribute to some of the heterogeneity seen in the SEC profiles, the large size of the early eluting complexes far exceeds expected MW for nanodiscs and is more likely to reflect nonspecific aggregation. Previous studies with the α-helical membrane protein, bacteriorhodopsin (~22 kDa), also gave rise to similarly large aggregates.[96] Electron microscopy suggested that this species was an amalgamation of larger nanodiscs that had formed as well as the clustering of regular nanodiscs and was dependent on the lipid: MSP ratio used.[96]

4.2 Limitations of results

Significant limitations of this thesis arise from the absence of duplicate trials to assess the reproducibility of the results and the poor quality of many of the SDS-PAGE gels. All trials shown in the results (Chapter 3), were only done once or twice, making it impossible to draw meaningful conclusions. One of the issues is that TMD81 yields are very low (0.5-0.9 mg/L of culture) and

purifications are time intensive. Therefore, there was a pressure to expand the range of conditions that could be tested with each trial at the expense of ensuring reproducibility. For future studies, the repetition of all experiments to establish reproducibility on a greater scale would allow more useful conclusions to be drawn from the results.

In addition, the inconsistent and poor quality of the SDS-PAGE gels created addition problems in the analysis of results. Many gels showed inconsistent migration across the lanes, suggesting potential problems with experimental set-up such as, inconsistent gel casting, errors in buffers and gel recipes, physical problems with the gel apparatus, etc. To avoid these problems, pre-cast gels could have been purchased and MW ladders run on both sides of all gels. Furthermore, the concentration of samples was not adjusted prior to SDS-PAGE gel analysis, which sometimes lead to over-loading problems. This sometimes gave rise to inconsistent band intensities, streaking, and seeping of samples into neighbouring lanes. While the determination of concentration was complicated by the very low yields obtained for all samples and the absorbance of the detergent, Triton-X100, at 280 nm, approximations of concentrations would have likely resulted in better resolved gels.

4.3 Future directions for TMD81 NMR studies in membrane mimicking environments

Due to the problems we experienced with MSP proteolysis and possible oligomerization, it may be useful to use circularized nanodiscs (cNDs) in future attempts to incorporate TMD81 into lipid particles.[132] Circularized MSPs have been made from linear ΔH5 and 1D1 by covalently linking C- and N- termini through the use of sortase A, a transpeptidase that is able to cleave the consensus sequence (LPGTG) and anchor the remaining peptide to a glycine residue.[133] This reaction is accomplished by attaching the consensus sequence near the C terminus and the glycine residue near the N terminus. NMR studies on ^{1}H relaxation rates indicate that lipid dynamics are

more restricted in nanodiscs made with circularized MSP compared to their linear counterparts, possibly explaining the greater thermal stability observed.[133] In addition, despite the increase in diameter of ~1 nm seen by electron microscopy for circular nanodiscs (cNDs), higher homogeneity in cNDs size was observed.[133] Two variants of cNDs, cNW9 (diameter = 9nm) and cNW11 (diameter = 11nm), have been used to study the β-barrel membrane protein, VDAC-1, by solution NMR.[133] HSQC spectra for both variants showed enhanced signal intensities and better resolution than previous ^{1}H-^{15}N HSQC spectra obtained in nanodiscs with linear MSPs, illustrating the benefits of the increased homogeneity[133] and stability mentioned above. In addition to the spectral quality improvements, it was suggested that different sized MSPs allowed the incorporation of different oligomeric states of VDAC-1 due to spectral differences observed between cND9 and cND11 nanodiscs.[133] All in all, these results suggest that cNDs may help to reduce the issues with proteolysis and oligomerization that were seen for TMD81 in nanodiscs made with linear MSPs.

In addition to MSPs, other proteins can form soluble complexes with lipids to create alternative nanodiscs that are also able to reconstitute a membrane protein into a lipid bilayer. A promising alternative is based on Saposin-A (SapA), an amphipathic protein that can exist in a closed, soluble and an open, lipid-bound state. X-ray crystal structures have shown that SapA can form dimers,[134] or larger oligomers to make various sizes of lipoparticles.[135] SapA-based nanodiscs have been used to solubilize the 7-TM helix rhodopsin II and acquire a ^{1}H-^{15}N HSQC spectrum.[135] While the spectral quality suffered from large size of the complex (>200 kDa), SapA nanodiscs were stable over a wide range of pH values, and could even tolerate being subjected to freeze-thaw cycles. There was also better homogeneity than that seen in nanodiscs as measured by SEC and negative-stain EM.[135] In addition, the α-helical proton-coupled oligopeptide transporter (POT), PepTSo2, reconstituted into SapA-discs also had greater thermostability (T_m =

72°C) than that obtained in detergent micelles (T_m = 43°C). Single-particle cryo-EM showed a that these complexes had a diameter of 12 nm, suggesting a tetrameric state for PepTSo2 that was in a nanodisc formed by four SapA molecules.[136] This observation suggests that SapA can accommodate various sizes of membrane proteins as the number of SapA subunits that form the nanodisc can vary depending on the size of the complex being formed. This would remove the need to test MSP proteins of various lengths in optimization trials, with only lipid-to-protein ratios needing investigation for optimal complex formation.[136]

4.4 The study of *ec*GlpG conformational dynamics by solid-state NMR

Solid-state NMR (ssNMR) has become a powerful tool for the study of membrane proteins in lipid bilayers,[137] often under more physiological conditions and temperatures than is possible by solution NMR.[138] During the writing of this thesis, a ssNMR was published by the Lange group that recorded 2D and 3D spectra of TMD81 in liposomes. This allowed chemical shift assignment of backbone residues and a more detailed study of the putative substrate gate.[138] In this study TMD81 mutants were made to change the strength of interactions between α-helix 2 (α2) and 5 (α5). To reduce the ability of α5 to undergo lateral motion, cysteine residues were introduced at various sites in α2 and α5. Cross-linking between cysteine residues using a 1,2-bis(methylsulfonyl-sulfonyl)ethane (M2M) linker lead to a significant decrease in proteolysis activity along with severe broadening in peaks arising from α5 as well as parts of L1, the lower part of α-helix 6, and some residues on α-helix 1, 2, and 3. This broadening suggested a more rigid overall protein conformation that was confirmed by comparison of relaxation rates in the ps-ns (R_1) and ns-ms ($R_{1\rho}$) time scales in the crosslinked versus reduced (non-crosslinked) states.[138] Despite smaller than expected differences in measured bulk relaxation rates, there was an increase

in slower timescale dynamics at the expense of faster timescale motions when the two residues were mutated to cysteine in a non-crosslinked state. Molecular dynamic simulations performed in the same study also supported a dynamic state for α5 that was similar to those seen in loop regions of TMD81. These results all strongly support the role of helix α5 in substrate gating and illustrate the promise of studying TMD81 gating dynamics through solid-state NMR.

4.5 Conclusion

In conclusion, the results observed in this thesis show two main challenges faced with TMD81 incorporation into nanodiscs; one, the high level of MSP proteolysis in all samples, and two, the low yields obtained for successful TMD81-encapsulated nanodiscs. While additional efforts could be dedicated to optimization to address these issues, alternative nanodisc systems or the use of solid-state NMR may have more promise for the study of TMD81.

Chapter 5: References

1 A. Krogh, B. Larsson, G. Von Heijne and E. L. L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes, *J. Mol. Biol.*, 2001, **305**, 567–580.

2 E. Wallin and G. Von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, *Protein Sci.*, 1998, **7**, 1029–1038.

3 M. K. Lemberg, Sampling the membrane: Function of rhomboid-family proteins, *Trends Cell Biol.*, 2013, **23**, 210–217.

4 Y. Zhou, S. M. Moin, S. Urban and Y. Zhang, An internal water-retention site in the rhomboid intramembrane protease GlpG ensures catalytic efficiency, *Structure*, 2012, **20**, 1255–1263.

5 Y. Wang, Y. Zhang and Y. Ha, Crystal structure of a rhomboid family intramembrane protease, *Nature*, 2006, **444**, 179–183.

6 R. B. Rawson, N. G. Zelenski, D. Nijhawan, J. Ye, J. Sakai, M. T. Hasan, T. Y. Chang, M. S. Brown and J. L. Goldstein, Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs, *Mol. Cell*, 1997, **1**, 47–57.

7 M. S. Wolfe, W. Xia, B. L. Ostaszewski, T. S. Diehl, W. T. Kimberly and D. J. Selkoe, Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and γ-secretase activity, *Nature*, 1999, **398**, 513–517.

8 A. Weihofen, K. Binns, K. M. Lemberg, K. Ashman and B. Martoglio, Identification of signal peptide peptidase, a Presenilin-type aspartic protease, *Science (80-.).*, 2002, **296**, 2215–2218.

9 S. Urban, J. R. Lee and M. Freeman, Drosophila Rhomboid-1 defines a family of putative intramembrane serine proteases, *Cell*, 2001, **107**, 173–182.

10 K. Walder, L. Kerr-Bayles, A. Civitarese, J. Jowett, J. Curran, K. Elliott, J. Trevaskis, N. Bishara, P. Zimmet, L. Mandarino, E. Ravussin, J. Blangero, A. Kissebah and G. R. Collier, The mitochondrial rhomboid protease PSARL is a new candidate gene for type 2 diabetes, *Diabetologia*, 2005, **48**, 459–468.

11 C. Meissner, H. Lorenz, A. Weihofen, D. J. Selkoe and M. K. Lemberg, The mitochondrial intramembrane protease PARL cleaves human PINK1 to regulate PINK1 trafficking, *J. Neurochem.*, 2011, **117**, 856–867.

12 S. Saita, T. Tatsuta, P. A. Lampe, T. König, Y. Ohba and T. Langer, PARL partitions the lipid transfer protein STARD7 between the cytosol and mitochondria, *EMBO J.*, 2018, **37**, 1–18.

13 J. wen Lin, P. Meireles, M. Prudêncio, S. Engelmann, T. Annoura, M. Sajid, S. Chevalley-Maurel, J. Ramesar, C. Nahar, C. M. C. Avramut, A. J. Koster, K. Matuschewski, A. P. Waters, C. J. Janse, G. R. Mair and S. M. Khan, Loss-of-function analyses defines vital and redundant functions of the Plasmodium rhomboid protease family, *Mol. Microbiol.*, 2013, **88**, 318–338.

14 G. Rugarabamu, J. B. Marq, A. Guérin, M. Lebrun and D. Soldati-Favre, Distinct contribution of Toxoplasma gondii rhomboid proteases 4 and 5 to micronemal protein protease 1 activity during invasion, *Mol. Microbiol.*, 2015, **97**, 244–262.

15 I. Ejigiri, D. R. T. Ragheb, P. Pino, A. Coppi, B. L. Bennett, D. Soldati-Favre and P. Sinnis, Shedding of TRAP by a rhomboid protease from the malaria sporozoite surface is essential

for gliding motility and sporozoite infectivity, *PLoS Pathog.*, 2012, **8**, 7.
16 A. G. Bang and C. Kintner, Rhomboid and star facilitate presentation and processing of the Drosophila TGF-α homolog Spitz, *Genes Dev.*, 2000, **14**, 177–186.
17 Y. Ha, Y. Akiyama and Y. Xue, Structure and mechanism of rhomboid protease, *J. Biol. Chem.*, 2013, **288**, 15430–15436.
18 W. Song, W. Liu, H. Zhao, S. Li, X. Guan, J. Ying, Y. Zhang, F. Miao, M. Zhang, X. Ren, X. Li, F. Wu, Y. Zhao, Y. Tian, W. Wu, J. Fu, J. Liang, W. Wu, C. Liu, J. Yu, S. Zong, S. Miao, X. Zhang and L. Wang, Rhomboid domain containing 1 promotes colorectal cancer growth through activation of the EGFR signalling pathway, *Nat. Commun.*, , DOI:10.1038/ncomms9022.
19 C. Adrain, K. Strisovsky, M. Zettl, L. Hu, M. K. Lemberg and M. Freeman, Mammalian EGF receptor activation by the rhomboid protease RHBDL2, *EMBO Rep.*, 2011, **12**, 421–427.
20 U. A. Ochsner, A. Snyder, A. I. Vasil and M. L. Vasil, Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 8312–8317.
21 L. G. Stevenson, K. Strisovsky, K. M. Clemmer, S. Bhatt, M. Freeman and P. N. Rather, Rhomboid protease AarA mediates quorum-sensing in Providencia stuartii by activating TatA of the twin-arginine translocase, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 1003–1008.
22 L. Fleig, N. Bergbold, P. Sahasrabudhe, B. Geiger, L. Kaltak and M. K. Lemberg, Ubiquitin-dependent intramembrane rhomboid protease promotes ERAD of membrane proteins, *Mol. Cell*, 2012, **47**, 558–569.
23 J. J. Lim, Y. Lee, T. T. Ly, J. Y. Kang, J. G. Lee, J. Y. An, H. S. Youn, K. R. Park, T. G. Kim, J. K. Yang, Y. Jun and S. H. Eom, Structural insights into the interaction of p97 N-Terminus domain and VBM in rhomboid protease, RHBDL4, *Biochem. J.*, 2016, **473**, 2863–2880.
24 E. Bier, L. Y. Jan and Y. N. Jan, Rhomboid, a gene required for dorsoventral axis establishment and peripheral nervous system development in Drosophila melanogaster, *Genes Dev.*, 1990, **4**, 680–681.
25 C. Nusslein-Volhard, E. Wieschaus and H. Kluding, Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster, *Roux's Arch. Dev. Biol.*, 1984, **5193**, 267–282.
26 S. M. Moin and S. Urban, Membrane immersion allows rhomboid proteases to achieve specificity by reading transmembrane segment dynamics, *Elife*, 2012, **2012**, 1–16.
27 R. P. Baker and S. Urban, Architectural and thermodynamic principles underlying intramembrane protease function, *Nat. Chem. Biol.*, 2012, **8**, 759–768.
28 S. W. Dickey, R. P. Baker, S. Cho and S. Urban, Proteolysis inside the membrane is a rate-governed reaction not driven by Substrate Affinity, *Cell*.
29 Z. Wu, N. Yan, L. Feng, A. Oberstein, H. Yan, R. P. Baker, L. Gu, P. D. Jeffrey, S. Urban and Y. Shi, Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry, *Nat. Struct. Mol. Biol.*, 2006, **13**, 1084–1091.
30 J. D. Wasserman, S. Urban and M. Freeman, A family of rhomboid-like genes: Drosophila rhomboid-1 and roughoid/rhomboid-3 cooperate to activate EGF receptor signaling, *Genes Dev.*, 2000, **14**, 1651–1663.
31 M. A. Sturtevant, M. Roark and E. Bier, The Drosophila rhomboid gene mediates the localized formation of wing veins and interacts genetically with components of the EGF-R

signaling pathway, *Genes Dev.*, 1993, **7**, 961–973.
32 U. Gohlke, L. Pullan, C. A. McDevitt, I. Porcelli, E. De Leeuw, T. Palmer, H. R. Saibil and B. C. Berks, The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 10482–10486.
33 M. Jamshad, V. Grimard, I. Idini, T. J. Knowles, M. R. Dowle, N. Schofield, P. Sridhar, Y. Lin, R. Finka, M. Wheatley, O. R. T Thomas, R. E. Palmer, M. Overduin, C. Govaerts, J.-M. Ruysschaert, K. J. Edler, T. R. Dafforn and S.-V. Berlin, Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins, *Nano Res.*, 2015, **8**, 774–789.
34 J. E. Greenblatt, A. J. Olzmann and R. R. Kopito, Making the cut: intramembrane cleavage by a rhomboid proteae promotes ERAD, *Nat. Struct. Mol. Biol.*, 2012, **19**, 979–981.
35 S. Paschkowsky, M. Hamzé, F. Oestereich and L. M. Munter, Alternative processing of the amyloid precursor protein family by rhomboid protease RHBDL4, *J. Biol. Chem.*, 2016, **291**, 21903–21912.
36 R. J. O'Brien and P. C. Wong, Amyloid precursor protein processing and alzheimer's disease, *Annu. Rev. Neurosci.*, 2011, **34**, 185–204.
37 Y. Kitagishi, N. Nakano, M. Ogino, M. Ichimura, A. Minami and S. Matsuda, PINK1 signaling in mitochondrial homeostasis and in aging (review), *Int. J. Mol. Med.*, 2017, **39**, 3–8.
38 D. V. Jeyaraju, L. Xu, M. C. Letellier, S. Bandaru, R. Zunino, E. A. Berg, H. M. McBride and L. Pellegrini, Phosphorylation and cleavage of presenilin-associated rhomboid-like protein (PARL) promotes changes in mitochondrial morphology, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 18562–18567.
39 C. Meissner, H. Lorenz, B. Hehn and M. K. Lemberg, Intramembrane protease PARL defines a negative regulator of PINK1- and PARK2/Parkin-dependent mitophagy, *Autophagy*, 2015, **11**, 1484–1498.
40 E. Poláchová, K. Bach, E. Heuten, S. Stanchev, A. Tichá, P. Lampe, P. Majer, T. Langer, M. K. Lemberg and K. Střížovský, Chemical blockage of the mitochondrial rhomboid protease PARL by novel ketoamide inhibitors reveals Its role in PINK1/Parkin-dependent mitophagy, *J. Med. Chem.*, 2023, **66**, 251–265.
41 N. Phasukkijwatana, B. Kunhapan, J. Stankovich, W. L. Chuenkongkaew, R. Thomson, T. Thornton, M. Bahlo, T. Mushiroda, Y. Nakamura, S. Mahasirimongkol, A. W. Tun, C. Srisawat, C. Limwongse, C. Peerapittayamongkol, T. Sura, W. Suthammarak and P. Lertrit, Genome-wide linkage scan and association study of PARL to the expression of LHON families in Thailand, *Hum. Genet.*, 2010, **128**, 39–49.
42 S. Saita, H. Nolte, K. U. Fiedler, H. Kashkar, A. V. Saskia, R. P. Zahedi, M. Krüger and T. Langer, PARL mediates Smac proteolytic maturation in mitochondria to promote apoptosis, *Nat. Cell Biol.*, 2017, **19**, 318–328.
43 S. Dhingra, C. H. Kowalski, A. Thammahong, S. R. Beattie, K. M. Bultman and R. A. Cramer, RbdB, a rhomboid protease critical for SREBP activation and virulence in Aspergillus fumigatus, *mSphere*, , DOI:10.1128/msphere.00678-19.
44 E. Rastew, L. Morf and U. Singh, Entamoeba histolytica rhomboid protease 1 has a role in migration and motility as validated by two independent genetic approaches, *Exp. Parasitol.*, 2015, **154**, 33–42.
45 A. M. Riestra, S. Gandhi, M. J. Sweredoski, A. Moradian, S. Hess, S. Urban and P. J.

Johnson, A Trichomonas vaginalis rhomboid protease and its substrate modulate parasite attachment and cytolysis of host cells, *PLoS Pathog.*, 2015, **11**, 1–25.

46 C. W. Russell, A. C. Richards, A. S. Chang and M. A. Mulvey, The rhomboid protease GlpG promotes the persistence of extraintestinal pathogenic Escherichia coli within the gut, *Infect. Immun.*, , DOI:10.1128/IAI.00866-16.

47 J. S. Buguliskis, F. Brossier, J. Shuman and L. D. Sibley, Rhomboid 4 (ROM4) affects the processing of surface adhesins and facilitates host cell invasion by Toxoplasma gondii, *PLoS Pathog.*, 2010, **6**, 1–14.

48 R. P. Baker, R. Wijetilaka and S. Urban, Two Plasmodium rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria, *PLoS Pathog.*, 2006, **2**, 0922–0932.

49 S. Maegawa, K. Ito and Y. Akiyama, Proteolytic action of GlpG, a rhomboid protease in the Escherichia coli cytoplasmic membrane, *Biochemistry*, 2005, **44**, 13543–13552.

50 T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner and H. Mori, Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection, *Mol. Syst. Biol.*

51 M. K. Lemberg, J. Menendez, A. Misik, M. Garcia, C. M. Koth and M. Freeman, Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases, *EMBO J.*, 2005, **24**, 464–472.

52 E. Arutyunova, P. Panwar, P. M. Skiba, N. Gale, M. W. Mak and M. J. Lemieux, Allosteric regulation of rhomboid intramembrane proteolysis, *EMBO J.*, 2014, **33**, 1869–1881.

53 K. A. Gaffney and H. Hong, The rhomboid protease GlpG has weak interaction energies in its active site hydrogen bond network, *J. Gen. Physiol.*, 2019, **151**, 282–291.

54 R. P. Baker, K. Young, L. Feng, Y. Shi and S. Urban, Enzymatic analysis of a rhomboid intramembrane protease implicates transmembrane helix 5 as the lateral substrate gate, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 8257–8262.

55 S. Cho, S. W. Dickey and S. Urban, Crystal structures and inhibition kinetics reveal a two-stage catalytic mechanism with drug design implications for rhomboid proteolysis, *Mol. Cell*, 2016, **61**, 329–340.

56 K. Strisovsky, H. J. Sharpe and M. Freeman, Sequence-specific intramembrane proteolysis: Identification of a recognition motif in rhomboid substrates, *Mol. Cell*, 2009, **36**, 1048–1059.

57 Y. Akiyama and S. Maegawa, Sequence features of substrates required for cleavage by GlpG, an Escherichia coli rhomboid protease, *Mol. Microbiol.*, 2007, **64**, 1028–1037.

58 A. C. Foo, New insights into the role of membrane interactions and conformational dynamics in intramembrane proteolysis by GlpG rhomboid, 2017, 1–188.

59 O. Lohi, S. Urban and M. Freeman, Diverse substrate recognition mechanisms for rhomboids: Thrombomodulin Is cleaved by mammalian rhomboids, *Curr. Biol.*, 2004, **14**, 236–241.

60 S. Urban and M. Freeman, Substrate specificity of rhomboid intramembrane proteases is governed by helix-breaking residues in the substrate transmembrane domain, *Mol. Cell*, 2003, **11**, 1425–1434.

61 S. Maegawa, K. Koide, K. Ito and Y. Akiyama, The intramembrane active site of GlpG, an E. coli rhomboid protease, is accessible to water and hydrolyses an extramembrane peptide bond of substrates, *Mol. Microbiol.*, 2007, **64**, 435–447.

62 A. R. Sherratt, D. R. Blais, H. Ghasriani, J. P. Pezacki and N. K. Goto, Activity-based

protein profiling of the escherichia coli GlpG rhomboid protein delineates the catalytic core, *Biochemistry*, 2012, **51**, 7794–7803.

63 C. Lazareno-Saez, E. Arutyunova, N. Coquelle and M. J. Lemieux, Domain swapping in the cytoplasmic domain of the escherichia coli rhomboid protease, *J. Mol. Biol.*, 2013, **425**, 1127–1142.

64 H. Ghasriani, J. K. C. Kwok, A. R. Sherratt, A. C. Y. Foo, T. Qureshi and N. K. Goto, Micelle-catalyzed domain swapping in the GlpG rhomboid protease cytoplasmic domain, *Biochemistry*, 2014, **53**, 5907–5915.

65 A. C. Y. Foo, B. G. R. Harvey, J. J. Metz and N. K. Goto, Influence of hydrophobic mismatch on the catalytic activity of Escherichia coli GlpG rhomboid protease, *Protein Sci.*, 2015, **24**, 464–473.

66 S. Urban, D. Schlieper and M. Freeman, Conservation of intramembrane proteolytic activity and substrate specificity in prokaryotic and eukaryotic rhomboids, *Curr. Biol.*, 2002, **12**, 1507–1512.

67 Y. Wang, S. Maegawa, Y. Akiyama and Y. Ha, The role of L1 loop in the mechanism of rhomboid intramembrane protese GlpG, *J. Mol. Biol.*, 2007, **374**, 1104–1113.

68 K. R. Vinothkumar, K. Strisovsky, A. Andreeva, Y. Christova, S. Verhelst and M. Freeman, The structural basis for catalysis and substrate specificity of a rhomboid protease, *EMBO J.*, 2010, **29**, 3797–3809.

69 K. R. Vinothkumar, Structure of rhomboid protease in a lipid environment, *J. Mol. Biol.*, 2011, **407**, 232–247.

70 Y. Xue and Y. Ha, Catalytic mechanism of rhomboid protease GlpG probed by 3,4-dichloroisocoumarin and diisopropyl fluorophosphonate, *J. Biol. Chem.*, 2012, **287**, 3099–3107.

71 Y. Xue, S. Chowdhury, X. Liu, Y. Akiyama, J. Ellman and Y. Ha, Conformational change in rhomboid protease GlpG induced by inhibitor binding to its S′ subsites, *Biochemistry*, 2012, **51**, 3723–3731.

72 Y. Xue and Y. Ha, Large lateral movement of transmembrane helix S5 is not required for substrate access to the active site of rhomboid intramembrane protease, *J. Biol. Chem.*, 2013, **288**, 16645–16654.

73 A. R. Sherratt, M. V. Braganza, E. Nguyen, T. Ducat and N. K. Goto, Insights into the effect of detergents on the full-length rhomboid protease from Pseudomonas aeruginosa and its cytosolic domain, *Biochim. Biophys. Acta - Biomembr.*, 2009, **1788**, 2444–2453.

74 A. N. Bondar, C. del Val and S. H. White, Rhomboid protease dynamics and lipid interactions, *Structure*, 2009, **17**, 395–405.

75 R. P. Baker and S. Urban, Cytosolic extensions directly regulate a rhomboid protease by modulating substrate gating, *Nature*, 2015, **523**, 101–105.

76 S. Cho, R. P. Baker, M. Ji and S. Urban, Ten catalytic snapshots of rhomboid intramembrane proteolysis from gate opening to peptide release, *Nat. Struct. Mol. Biol.*, 2019, **26**, 910–918.

77 S. F. Poget and M. E. Girvin, Solution NMR of membrane proteins in bilayer mimics: Small is beautiful, but sometimes bigger is better, *Biochim. Biophys. Acta - Biomembr.*, 2007, **1768**, 3098–3106.

78 M. Barniol-Xicota and S. H. L. Verhelst, Isolation of intramembrane proteases in membrane-like environments, *Biochim. Biophys. Acta - Biomembr.*, 2020, **1862**, 183193.

79 M. Stangl, A. Veerappan, A. Kroeger, P. Vogel and D. Schneider, Detergent properties influence the stability of the glycophorin a transmembrane helix dimer in

lysophosphatidylcholine micelles, *Biophys. J.*, 2012, **103**, 2455–2464.
80 R. M. Garavito and S. Ferguson-Miller, Detergents as tools in membrane biochemistry, *J. Biol. Chem.*, 2001, **276**, 32403–32406.
81 E. A. Morrison and K. A. Henzler-Wildman, Reconstitution of integral membrane proteins into isotropic bicelles with improved sample stability and expanded lipid composition profile, *Biochim. Biophys. Acta - Biomembr.*, 2012, **1818**, 814–820.
82 E. Van Den Brink-Van Der Laan, J. Antoinette Killian and B. De Kruijff, Nonbilayer lipids affect peripheral and integral membrane proteins via changes in the lateral pressure profile, *Biochim. Biophys. Acta - Biomembr.*, 2004, **1666**, 275–288.
83 I. Marcotte and M. Auger, Bicelles as model membranes for solid-and solution-state NMR studies of membrane peptides and proteins, *Concepts Magn. Reson. Part A Bridg. Educ. Res.*, 2005, **24**, 17–37.
84 K. S. Mineev and K. D. Nadezhdin, Membrane mimetics for solution NMR studies of membrane proteins, *Nanotechnol. Rev.*, 2017, **6**, 15–32.
85 A. I. Hassan, Development of 19 F NMR methods for the GlpG rhomboid protease in detergents and lipid nanoparticle systems, 2021, 1–102.
86 B. Krishnarjuna, T. Ravula and A. Ramamoorthy, Detergent-free extraction, reconstitution and characterization of membrane-anchored cytochrome-b5 in native lipids, *Chem. Commun.*, 2020, **56**, 6511–6514.
87 T. Ravula, S. K. Ramadugu, G. Di Mauro and A. Ramamoorthy, Bioinspired, size-tunable self-assembly of polymer–lipid bilayer nanodiscs, *Angew. Chemie - Int. Ed.*, 2017, **56**, 11466–11470.
88 K. A. Morrison, A. Akram, A. Mathews, Z. A. Khan, J. H. Patel, C. Zhou, D. J. Hardy, C. Moore-Kelly, R. Patel, V. Odiba, T. J. Knowles, M. U. H. Javed, N. P. Chmel, T. R. Dafforn and A. J. Rothnie, Membrane protein extraction and purification using styrene-maleic acid (SMA) copolymer: Effect of variations in polymer structure, *Biochem. J.*, 2016, **473**, 4349–4360.
89 M. J. Ranaghan, C. T. Schwall, N. N. Alder and R. R. Birge, Green proteorhodopsin reconstituted into nanoscale phospholipid bilayers (Nanodiscs) as photoactive monomers, *J. Am. Chem. Soc.*, 2011, **133**, 18318–18327.
90 D. J. Reid, J. E. Keener, A. P. Wheeler, D. E. Zambrano, J. M. Diesing, M. Reinhardt-Szyba, A. Makarov and M. T. Marty, Engineering nanodisc scaffold proteins for native mass spectrometry, *Anal. Chem.*, 2017, **89**, 11189–11192.
91 T. H. Bayburt, Y. V. Grinkova and S. G. Sligar, Assembly of single bacteriorhodopsin trimers in bilayer nanodiscs, *Arch. Biochem. Biophys.*, 2006, **450**, 215–222.
92 F. Hagn, M. L. Nasr and G. Wagner, Assembly of phospholipid nanodiscs of controlled size for structural studies of membrane proteins by NMR, *Nat. Protoc.*, 2018, **13**, 79–98.
93 F. Hagn, M. Etzkorn, T. Raschle and G. Wagner, Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins, *J. Am. Chem. Soc.*, 2013, **135**, 1919–1925.
94 N. Skar-Gislinge, J. B. Simonsen, K. Mortensen, R. Feidenhans'l, S. G. Sligar, B. Lindberg Møller, T. Bjørnholm and L. Arleth, Elliptical structure of phospholipid bilayer nanodiscs encapsulated by scaffold proteins: Casting the roles of the lipids and the protein, *J. Am. Chem. Soc.*, 2010, **132**, 13713–13722.
95 A. J. Leitz, T. H. Bayburt, A. N. Barnakov, B. A. Springer and S. G. Sligar, Functional reconstitution of β2-adrenergic receptors utilizing self-assembling Nanodisc technology,

Biotechniques, 2006, **40**, 601–612.
96 M. Etzkorn, T. Raschle, F. Hagn, V. Gelev, A. J. Rice, T. Walz and G. Wagner, Cell-free expressed bacteriorhodopsin in different soluble membrane mimetics: Biophysical properties and NMR accessibility, *Structure*, 2013, **21**, 394–401.
97 M. Standfuss, J. Schertler, G. F. X. Oprian and D. D. Ernst, Assembly of an activated rhodopsin−transducin complex in nanoscale lipid bilayers, *Proc. Natl. Acad. Sci. U.S.A*, 2011, **50**, 10859–10864.
98 L. Sušac, R. Horst and K. Wüthrich, Solution-NMR characterization of outer-membrane protein A from E. coli in lipid bilayer nanodiscs and detergent micelles, *ChemBioChem*, 2014, **15**, 995–1000.
99 O. Engberg, D. Ulbricht, V. Döbel, V. Siebert, C. Frie, A. Penk, M. K. Lemberg and D. Huster, Rhomboid-catalyzed intramembrane proteolysis requires hydrophobic matching with the surrounding lipid bilayer, *Sci. Adv.*, 2022, **8**, 1–10.
100 A. J. B. Kreutzberger, M. Ji, J. Aaron, L. Mihaljević and S. Urban, Rhomboid distorts lipids to break the viscosity-imposed speed limit of membrane diffusion, *Science (80-.).*, , DOI:10.1126/science.aao0076.
101 I. G. Denisov, Y. V. Grinkova, A. A. Lazarides and S. G. Sligar, Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size, *J. Am. Chem. Soc.*, 2004, **126**, 3477–3487.
102 T. K. Ritchie, Y. V. Grinkova, T. H. Bayburt, I. G. Denisov, J. K. Zolnerciks, W. M. Atkins and S. G. Sligar, Chapter 11: Reconstitution of membrane proteins in phospholipid bilayer nanodiscs, *Methods Enzymol.*, 2009, **464**, 211–231.
103 D. Marion, An introduction to biological NMR spectroscopy, *Mol. Cell. Proteomics*, 2013, **12**, 3006–3025.
104 K. Zia, T. Siddiqui, S. Ali, I. Farooq, M. S. Zafar and Z. Khurshid, Nuclear magnetic resonance spectroscopy for medical and dental applications: A comprehensive review, *Eur. J. Dent.*, 2019, **13**, 124–128.
105 T. Claridge, Introducing high-resolution NMR, 2016, 11–59.
106 I. R. Kleckner and M. P. Foster, An introduction to NMR-based approaches for measuring protein dynamics, *Biochim. Biophys. Acta - Proteins Proteomics*, 2011, **1814**, 942–968.
107 J. Vaynberg and J. Qin, Weak protein-protein interactions as probed by NMR spectroscopy, *Trends Biotechnol.*, 2006, **24**, 22–27.
108 J. T. Gerig, Fluorine NMR of proteins, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1994, **26**, 293–370.
109 J. T. Gerig, Fluorine nuclear magnetic resonance of fluorinated ligands, *Methods Enzymol.*, 1989, **177**, 3–23.
110 B. D. Sykes and W. E. Hull, Fluorine nuclear magnetic resonance studies of proteins, *Methods Enzymol.*, 1978, **49**, 270–295.
111 J. Sambrook and D. W. Russell, Molecular Cloning: a Laboratory Manual.
112 U. K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature*, 1970, **227**, 680–685.
113 Cytiva, Superdex 200 Increase 5 / 150 GL Superdex 200 Increase 10 / 300 GL Instructions for Use, 2020, 1–9.
114 F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer and A. Bax, NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR*, 1995, **6**, 277–293.

115 B. A. Johnson and R. A. Blevins, NMR View: A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR*, 1994, **4**, 603–614.

116 Y. Hioki, H. Kuyama, C. Hamana, K. Takeyama and K. Tanaka, An improved sample preparation method for the sensitive detection of peptides by MALDI-MS, *J. Mass Spectrom.*, 2013, **48**, 1217–1223.

117 M. Alami, K. Dalal, B. Lelj-Garolla, S. G. Sligar and F. Duong, Nanodiscs unravel the interaction between the SecYEG channel and its cytosolic partner SecA, *EMBO J.*, 2007, **26**, 1995–2004.

118 I. G. Denisov and S. G. Sligar, Nanodiscs for structural and functional studies of membrane proteins, *Nat. Struct. Mol. Biol. 2016 236*, 2016, **23**, 481–486.

119 T. Raschle, S. Hiller, T.-Y. Yu, A. J. Rice, T. Walz and G. Wagner, Structural and functional characterization of the integral membrane protein VDAC-1 in lipid bilayer nanodiscs, *JACS Commun.*, 2009, 17777–17779.

120 T. H. Bayburt, Y. V. Grinkova and S. G. Sligar, Assembly of single bacteriorhodopsin trimers in bilayer nanodiscs, *Arch. Biochem. Biophys.*, 2006, **450**, 215–222.

121 M. M. Grozdanović, B. J. Drakulić and M. Gavrović-Jankulović, Conformational mobility of active and E-64-inhibited actinidin, *Biochim. Biophys. Acta - Gen. Subj.*, 2013, **1830**, 4790–4799.

122 V. Bondada, J. Gal, C. Mashburn, D. W. Rodgers, K. E. Larochelle, D. E. Croall and J. W. Geddes, The C2 domain of calpain 5 contributes to enzyme activation and membrane localization, *Biochim. Biophys. Acta - Mol. Cell Res.*, 2021, **1868**, 119019.

123 Y. Poloz, A. Catalano and D. H. O'Day, Bestatin inhibits cell growth, cell division, and spore cell differentiation in Dictyostelium discoideum, *Eukaryot. Cell*, 2012, **11**, 545–557.

124 K. Matúz, J. Mótyán, M. Li, A. Wlodawer and J. Tözsér, Inhibition of XMRV and HIV-1 proteases by pepstatin A and acetyl-pepstatin, *FEBS J.*, 2012, **279**, 3276–3286.

125 B. R. Goblirsch, B. T. Arachea, D. J. Councell and M. C. Wiener, Phosphoramidon inhibits the integral membrane protein zinc metalloprotease ZMPSTE24, *Acta Crystallogr. Sect. D Struct. Biol.*, 2018, **74**, 739–747.

126 P. K. Jadhav and F. J. Woerner, Synthesis of C2-symmetric HIV-1 protease inhibitors from D-mannitol, *Bioorganic Med. Chem. Lett.*, 1992, **2**, 353–356.

127 S. Yavuz, S. Kocabay, S. Çetinkaya, B. Akkaya, R. Akkaya, A. F. Yenidunya and M. Z. Bakıcı, Production, purification, and characterization of metalloprotease from Candida kefyr 41 PSB, *Int. J. Biol. Macromol.*, 2017, **94**, 106–113.

128 A. B. Cory, J. P. Chanton, R. G. M. Spencer, O. C. Ogles, V. I. Rich, C. K. McCalley and R. M. Wilson, Quantifying the inhibitory impact of soluble phenolics on anaerobic carbon mineralization in a thawing permafrost peatland, *PLoS One*, 2022, **17**, 1–19.

129 M. Wadsäter, S. Maric, J. B. Simonsen, K. Mortensen and M. Cardenas, The effect of using binary mixtures of zwitterionic and charged lipids on nanodisc formation and stability, *Soft Matter*, 2013, **9**, 2329–2337.

130 R. Puthenveetil and O. Vinogradova, Optimization of the design and preparation of nanoscale phospholipid bilayers for its application to solution NMR, *Proteins Struct. Funct. Bioinforma.*, 2013, **81**, 1222–1231.

131 P. Sampathkumar, M. W. Mak, S. J. Fischer-Witholt, E. Guigard, C. M. Kay and M. J. Lemieux, Oligomeric state study of prokaryotic rhomboid proteases, *Biochim. Biophys. Acta - Biomembr.*, 2012, **1818**, 3090–3097.

132 M. Daniilidis, M. J. Brandl and F. Hagn, The advanced properties of circularized MSP

nanodiscs facilitate high-resolution NMR studies of membrane proteins: Advanced NMR properties of circular MSP nanodiscs, *J. Mol. Biol.*, 2022, **434**, 167861.

133 M. L. Nasr, D. Baptista, M. Strauss, Z. Y. J. Sun, S. Grigoriu, S. Huser, A. Plückthun, F. Hagn, T. Walz, J. M. Hogle and G. Wagner, Covalently circularized nanodiscs for studying membrane proteins and viral entry, *Nat. Methods 2016 141*, 2016, **14**, 49–52.

134 K. Popovic, J. Holyoake, R. Pomès and G. G. Privé, Structure of saposin A lipoprotein discs, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 2908–2912.

135 C. T. H. Chien, L. R. Helfinger, M. J. Bostock, A. Solt, Y. L. Tan and D. Nietlispach, An adaptable phospholipid membrane mimetic system for solution NMR studies of membrane proteins, *J. Am. Chem. Soc.*, 2017, **139**, 14829–14832.

136 J. Frauenfeld, R. Löving, J. P. Armache, A. F. P. Sonnen, F. Guettou, P. Moberg, L. Zhu, C. Jegerschöld, A. Flayhan, J. A. G. Briggs, H. Garoff, C. Löw, Y. Cheng and P. Nordlund, A saposin-lipoprotein nanoparticle system for membrane proteins, *Nat. Methods*, 2016, **13**, 345–351.

137 Y. Miao and T. A. Cross, Solid state NMR and protein-protein interactions in membranes, *Curr. Opin. Struct. Biol.*, 2013, **23**, 919–928.

138 C. Bohg, C. Öster, B. Türkaydin, M. Lisurek, P. Sanchez-Carranza, S. Lange, T. Utesch, H. Sun and A. Lange, The opening dynamics of the lateral gate regulates the activity of rhomboid proteases, *Sci. Adv.*, 2023, **9**, eadh3858.