University of Alberta

Characterization of oligomeric state of prokaryotic rhomboid proteases

by

Padmapriya Sampath Kumar

A thesis submitted to the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Biochemistry

Fall 2012

Edmonton, Alberta

Abstract

Rhomboid proteases are a remarkable class of intramembrane enzymes that carry out cleavage of transmembrane substrates within or proximal to the lipid bilayer. These proteases have been linked to several human diseases such as cancer, diabetes and early-onset blindness. They are also involved in diverse processes including quorum sensing and cell differentiation in bacteria. To better understand the mechanisms underlying the proteolytic action and function of these proteases, we have focussed on investigating its regulation. In this thesis, the concept of oligomerization as a possible mode of regulation is examined. To assess the oligomeric state of three prokaryotic rhomboid proteases from *Haemophilus influenza* (*hi*GlpG), *Escherichia coli* (*ec*GlpG) and *Bacillus subtilis* (YqgP), sedimentation equilibrium analysis was carried out. The predominant species for the three rhomboid proteases was observed to be dimeric. To examine the effect of the membrane domain alone on dimerization, *hi*GlpG, the simplest form of rhomboid representing the core of six transmembrane domains, was studied further. Gel filtration, crosslinking and functional assay demonstrate that *hi*GlpG is dimeric and functional in dodecylmaltoside detergent solution. More importantly, co-immunoprecipitation studies establish that the dimer is present in the lipid bilayer suggesting a physiological dimer. Overall these results indicate that rhomboids form oligomers which are facilitated by the membrane domain.

This thesis also investigates the physiological role of *ec*GlpG rhomboid from *E. coli*. The potential of *E. coli* TatA as a substrate for *ec*GlpG is examined using an *in vitro* functional assay. Additionally, affinity pull-down and co-immunoprecipitation techniques are performed to identify possible substrates for this rhomboid.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to Prof. Joanne Lemieux, who has been an incredible mentor during my years as a graduate student in the lab. She has been a strong supporter and advisor, always given me the freedom to find my own answers to questions and broadened my horizons as a researcher. Her motivation, enthusiasm, and immense knowledge have guided me throughout my time of research and in writing of this thesis. I could not have imagined having a better supervisor for my graduate study.

Special thanks to my supervisory committee members, Prof. Joel Weiner and Prof. Richard Fahlman for their encouragement, guidance and insightful comments. Their guidance has helped me well and I owe them my heartfelt appreciation. I would also like to thank Prof. Emmanuelle Cordat for taking part in my thesis evaluation.

Many thanks to my colleagues who have made research stimulating and fun and have been wonderful friends: Christelle Lazareno-Saez and Elena Arutyunova for all their help, advice and creative ideas not only on technical grounds but also on a personal level; Michelle Mak who taught me many techniques when I first came into the lab and also provided me strong support in completing experiments for the paper; Cory Brooks, Melissa Morrison and Pankaj Panwar for their positive energy and lab discussions.

Working with all these people has been an incredible experience that went much further than just work.

Last but not least, I would like to thank my family for their unconditional support during these last few years. Thanks to my parents and my sister; my well-being and happiness are their only concern. The love and support of my husband, Kasturi has been my source of strength and confidence that I could accomplish almost anything. I dedicate this work to my beloved parents, sister and husband.

Table of contents

List of Tables

List of Figures

List of abbreviations

°C	Degree Celsius
AarA	*Providencia stuartii* rhomboid
Aβ	Amyloid β peptide
AMA-1	Apical membrane antigen-1
APP	Amyloid precursor protein
APS	Ammonium persulphate
Asn	Asparagine
ATP	Adenosine triphosphate
Bla	β- lactamase
BLAST	Basic Local Alignment Search Tool
BCA	Bicinchoniniic acid
Ccp1	Cytochrome c peroxidase
DDM	n-Dodecyl- β- D- maltoside
DMSO	dimethyl sulfoxide
DSP	dithiobis (succinimodylpropionate)
DTSSP	3, 3' dithiobissulfo (succinimidylpropionate)
DTT	dithiothreitol
E. coli	*Escherichia coli*
*ec*GlpG	*Escherichia coli* rhomboid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EhROM1	*Entamoeba histolytica* Rhomboid 1
ER	Endoplasmic reticulum
EDTA	Ethylene diamine tetra acetic acid
glp	Glycerol-3-phosphate
h	Hours
*hi*GlpG	*Haemophilus influenzae* rhomboid
His_6	Histidine
IPTG	Isopropyl β- D thiogalactoside

iRhom	inactive rhomboid
kDa	kilo Dalton
L5	Loop 5
LacYTM2	LacY transmembrane helix 2
LB	Luria- Bertani
M	Marker
MW	Molecular weight
min	minutes
mg	milligrams
ml	millilitre
Mgm1	Dynamin- like GTPase
MIC	Microneme protein
Ni-NTA	Nickel nitrilotriacetic acid
OD	Optical density
PARL	Presenilin-associated rhomboid like
PBS	Phosphate Buffered Saline
Pcp1	*Schizosaccharomyces pombe* mitochondrial rhomboid
PCR	Polymerase chain reaction
PfROM1	*Plasmodium falciparum* Rhomboid 1
PMSF	Phenylmethyl sulfonyl fluoride
PVDF	Polyvinyl difluoride
RHBDL2	*Homo sapiens* Rhomboid 2
S1P	Site-1 Protease
S2P	Site-2 Protease
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Ser	Serine
SREBP	Sterol Regulatory Binding Element Protein
STD	Standard
TACE	Tumor Necrosis Factor- alpha converting enzyme
TatA	Twin arginine translocase A

TEMED	Tetramethylenediamine
TgROM4	*Toxoplasma gondii* Rhomboid 4
TgROM5	*Toxoplasma gondii* Rhomboid 5
TM	Transmembrane
Tris	Tris(hydroxymethyl)aminomethane
YqgP	*Bacillus subtilis* rhomboid

Chapter 1: Introduction

1.1 Intramembrane cleaving proteases

Proteases are proteolytic enzymes that catalyze the hydrolysis of peptide bonds. They function as molecular knives cleaving long protein sequences into smaller fragments, a process that is essential for regulating the length, composition, and function of many proteins [9]. In this sense, they introduce irreversible changes to the cellular protein pool and are tightly regulated in their activity. Different proteases play pivotal roles in diverse functions including cell cycle progression, morphogenesis, wound healing, immunity and apoptosis [10]. A number of pathological conditions such as cancer, neurodegenerative and cardiovascular diseases are associated with abnormal protease functions [11-13]. Owing to the importance of regulatory roles played by these proteases, they have been identified as potential drug targets.

Previously, all proteases were thought to be water soluble enzymes occurring in an aqueous environment. However, a new class of membrane proteins have been discovered that function as proteases in the hydrophobic environment within the lipid bilayer. In a process known as Regulated Intramembrane Proteolysis (RIP), transmembrane substrates are cleaved by intramembrane proteases to release a soluble domain that is liberated into the cytosol. The released soluble domain acts as a signaling molecule at locations either proximal or distant to the parent cell. This novel proteolysis reaction is found to be crucial in many cellular processes [14-16] and

mutations of intramembrane proteases are shown to be involved in many diseases [4, 17-19].

1.2 Classification and properties of intramembrane proteases

Intramembrane cleaving proteases are polytopic membrane proteins. Based on the catalytic residues and mechanism of action, these proteases are classified as aspartyl proteases, metalloproteases and serine proteases. Aspartyl protease and metalloprotease activate water molecules to interact with scissile carbonyl carbon of substrate, generating a tetrahedral oxyanion intermediate. Protonation of the scissile amide rearranges the oxyanion intermediate to form two separate protein fragments [20, 21]. On the other hand, serine proteases possess a nucleophilic serine residue in their catalytic site which directly interacts with the carbonyl group of the scissile peptide bond [22]. Intramembrane proteases can also be classified based on their cleavage mechanism for different substrate orientations. One class of intramembrane proteases cleave type I membrane proteins (N terminus facing periplasm, C terminus facing cytoplasm) while the other set of proteases react with type II membrane proteins (C terminus facing periplasm, N terminus facing cytoplasm) [23].

1.2.1 Intramembrane metalloproteases: The first intramembrane metalloprotease was identified in the regulation of sterol and lipid

biosynthesis in mammals [24]. In response to low cholesterol, Sterol Regulatory Element Binding Protein (SREBP) is transported from endoplasmic reticulum (ER) to Golgi. SREBP is initially synthesized as a precursor protein containing four domains, a transcription activating domain, a regulatory domain and two transmembrane domains. When SREBP reaches the Golgi, it is activated by two distinct cleavage events: Initial cleavage by Site-1 Protease (S1P) separating the two membrane- bound domains of SREBP and subsequent cleavage of its transmembrane segment by intramembrane metalloprotease Site-2 Protease (S2P). Consequently, the transcription activating domain of SREBP is released and translocated from the Golgi to the nucleus where it triggers the transcription of genes controlling cholesterol biosynthesis [25] (Figure 1.1 A). Complementation cloning of factors revealed that S2P contains a conserved HEXXH motif, in which the two histidines and glutamate residues coordinate with zinc for activating the water molecule [24]. Furthermore, substrate SREBP contains aspartate-proline (NP) residues located deep within its first transmembrane segment. This NP sequence is exposed after the first cleavage event of SREBP, allowing the helix to unwind and form an extended structure, thus making it readily accessible for cleavage by S2P [26].

1.2.2 Intramembrane aspartyl protease: The second class are the aspartyl intramembrane proteases which are frequently involved in treatment of diseases as these enzymes are considered drug targets [27]. There are two

main classes of aspartyl intramembrane proteases: Presenilin and Signal peptide peptidase. Both have nine transmembrane segments and contain the catalytic aspartate residue in two conserved motifs, GxGD and YD that are located in adjacent transmembrane segments [18, 28, 29]. A third sequence, C- terminal PAL motif, is known to play a role in the formation of catalytic pore [30, 31]. Though Presenilin and SPP have identical conserved motifs, they differ in a few structural and functional aspects. While SPP functions independently, Presenilin undergoes endoproteolysis and forms the active subunit of a multiprotein γ secretase complex [32, 33]. The preference for substrate orientation is also distinctly different between these two proteases [34, 35]. While Presenilin favours cleavage of Type I membrane proteins, SPP preferentially cleaves Type II membrane proteins. As a result, the catalytic aspartates of Presenilin and SPP are in opposite orientation (Figure 1.2).

A classic example of aspartyl protease action is the proteolytic processing of amyloid precursor protein (APP) by presenilin (Figure 1.1 C). The substrate APP is first processed by β-secretase and subsequently cleaved by presenilin, the active subunit of γ- secretase complex. The γ- secretase then liberates the amyloid β peptide and an APP cytoplasmic fragment, which moves to the nucleus and regulates gene expression for the organization and function of the hippocampal synapses [36, 37]. Presenilin has been linked to Alzheimer's disease as mutations in this protein causes increase in γ-secretase activity thereby resulting in the accumulation of amyloid β- plaques in the brain [38].

1.2.3 Intramembrane serine proteases: The third and most interesting family of intramembrane proteases are the serine proteases comprising the Rhomboid family. The active site is formed by a catalytic dyad of serine and histidine residues located in two transmembrane domains [39, 40]. Unlike soluble serine proteases which utilize a Ser/His/Asn catalytic triad, rhomboids do not require asparagine for catalysis [41]. A striking difference that distinguishes rhomboids from other intramembrane proteases is that these enzymes can cleave intact membrane proteins without requiring any pre-processing of substrates. In the recent years, remarkable discoveries have shed light on the ways these proteases control and regulate many biological processes. The remainder of this chapter will focus on rhomboid protease family describing their structure, mechanism, function and regulation.

1.3 Discovery of Rhomboid proteases

Rhomboids are a recently discovered family of intramembrane proteases. The name 'rhomboid' originated from the studies of Drosophila genome as the phenotype of rhomboid mutation revealed the appearance of a rhombus-like head skeleton [42]. Genome sequences of different organisms have revealed that rhomboid homologs are found in all kingdoms of life, excluding viruses [43-46]. Rhomboid proteases are the only class of intramembrane proteases that cleave the substrates located within or close

to the membrane bilayer resulting in the release of a N- terminal domain of the substrate.

Rhomboids were first discovered by genetic experiments in *Drosophila melanogaster*. Lack-of-function mutations of Rhomboid-1 and Spitz (an epidermal growth factor signaling protein) resulted in similar phenotypic changes in the larval cuticle of Drosophila, suggesting both genes were integrated into epidermal growth factor receptor (EGFR) signaling pathway [47] (Figure 1.1 B). Subsequent work provided clues of rhomboid activity in regulation and proteolytic processing of Spitz [43, 48, 49]. Sequence analysis of Rhomboid-1 initially did not provide any information as it did not resemble any known protein or indicate its function in EGFR signaling pathway [50]. The first breakthrough came from Freeman group which presented convincing information that Rhomboid-1 was indeed a novel intramembrane serine protease that catalysed the cleavage of Spitz [39](Figure 1.1B). Spitz proteolysis was assessed in a variety of cell lines and it was found that Rhomboid-1 cleaved Spitz in all cell lines with equal efficiency. In order to identify the residues important for catalysis, mutagenesis assays were performed. Mutational analysis revealed that six mutations reduced or abolished rhomboid activity, three of which were predicted to form a putative catalytic triad (Ser/His/Asn), a characteristic feature of serine proteases. The depth of cleavage site in the transmembrane segment of Spitz also matched with the predicted depth of the active site of rhomboid. Lastly, Spitz cleavage was blocked only by the addition of serine

protease inhibitors. These results implicated that Rhomboid-1, a serine protease, was directly responsible for cleavage and activation of Spitz. Further validation was provided by three groups which used purified rhomboids in detergent solubilized state to report activity [41, 51, 52]. Taken together, these results firmly established Rhomboid-1 as an intramembrane serine protease.

Since the discovery of rhomboids, research has progressed dramatically in the direction of understanding the mechanics of the proteolytic reaction and the function of rhomboids in other homologs. Crystal structures and biochemical approaches have provided a much clearer insight into the mysterious functioning of rhomboid intramembrane proteases, which are discussed in detail in the following sections.

1.4 Architecture of rhomboids

The sequence identity between rhomboid homologs was predicted to be very low [45, 46] apart from core domain of six transmembrane segments (TMs) with its consensus active site motif. The active site is formed by serine protease motif composed of histidine followed by GxxxG and serine within a GxSx in another TM, a characteristic of serine proteases. A highly conserved tryptophan- arginine motif (WR) was also observed to be located in the first extracellular loop [52]. These signature motifs were found to be conserved in all members of the rhomboid family and were thought to be functionally

relevant [53]. However, rhomboids could not be classified based on this aspect alone, as certain transmembrane proteins that were functionally unrelated to rhomboids carried similar sequence motifs [54]. To solve this, Lemberg *et al.* combined rhomboid topologies along with conserved motifs to generate a new system of classification [46].

Rhomboids are organised into three groups: active rhomboid protease; catalytically inert iRhoms carrying two features: a long N terminal domain and a large loop between the first and second TM segments; and other inactive rhomboid homologs that cannot be clearly assigned to either of the two groups (Figure 1.3). The active proteases are further classified based on their intracellular localization: secretase rhomboids, which are involved in the secretory pathway (eg: *Drosophila* Rhomboid-1) and mitochondrial rhomboids [eg: Presenilin-associated rhomboid-like (PARL)]. Secretase rhomboids are split into two classes; secretase A carrying 6+1 TM fused to the C terminus and secretase B with the core 6TM topology. It is important to note here that bacterial rhomboids are also grouped under the secretase family as they contain the basic six TM architecture, with *Escherichia coli* rhomboid and *Haemophilus influenzae* rhomboid having 6 TM while *Bacillus subtilis* containing 6+1 TM topologies. In contrast to the secretase family, mitochondrial rhomboids have 1+ 6TM topology with an extra TM fused to the N terminus, along with a unique N –terminal mitochondrion targeting motif [46, 55-58]. As a result, the active sites of

mitochondrial rhomboids are in opposite orientation when compared to the secretase family of rhomboids.

In addition to these differences, two other variations induce substantial diversity between different classes of rhomboid proteases. The first feature is the absence of catalytic activity in iRhoms as these rhomboids do not contain the essential catalytic residues. Instead they harbour a conserved proline residue at GxSx serine motif which disrupts the active site [45, 59, 60]. Second, the lengths of the cytosolic domains are highly variable between different rhomboid homologs and fluctuate between lengthy to non-existent N terminii. Though these domains are not proteolytic, they are found to be involved in regulating the activity of rhomboids or anchoring it to the membrane [61, 62] .

1.5 Structural insights into rhomboid crystal structure

Insight into how rhomboids function in the lipid bilayer was provided by crystallography. In a span of one year, several crystal structures of two bacterial rhomboid homologs, *E. coli* and *H. influenzae* revealed complex and unexpected architecture [6, 40, 63-66] (Figure 1.4A). Though these rhomboids were purified in different detergent conditions, all exhibited structural homogeneity.

Structures revealed the presence of six TM segments clustered together to form a water-filled cavity at ~10 Å below the periplasmic face.

The water-filled cavity provided a hydrophilic environment for the cleavage of substrates within the protease. Catalytic serine and histidine were correctly positioned on TM4 and TM6 to perform nucleophilic attack on the substrate. Serine acts as the nucleophile while histidine serves as the catalytic base. Tyrosine residue (Y205 from *ec*GlPG) from TM4 stacks under the histidine base, providing support during the proteolytic reaction.

The role of asparagine was investigated for the presence of a catalytic triad as various biochemical approaches provided contradictory information. While asparagine was important for activity in Drosophila Rhomboid-1, it was not essential for human rhomboid RHBDL2 [39]. Also, the importance of asparagine varied between experiments performed *in vivo* and *in vitro* [41, 67, 68]. Contrary to other serine proteases, it has been found that asparagine located on the TM2 could only help in the stabilization of oxyanion intermediate but not assist in catalysis (Figure 1.4B). *In vitro* proteolysis experiments also suggest that rhomboids do not require co-factors or ATP molecules for proteolysis and that their activity can be regulated by different lipids in the membrane [51, 52].

Two more features were identified in the rhomboid crystal structure. The first feature is the orientation of the Loop 1 that connects TM1 and TM2. Loop 1 formed a lateral protrusion with half of its surface inserted into the membrane. A highly conserved WR motif was present near the bottom of Loop L1, where the side chains of these residues point upwards into the loop [6]. Mutational analyses revealed that substitutions of WR motif with alanine

residues either abolished or reduced the proteolytic activity of the enzyme [6, 69] (Figure 1.5). This loop was initially thought to play a role in substrate recruitment [40] but it is now thought that Loop 1, with its insertion into the membrane, helps in maintaining precise orientation of the enzyme by attaching to TM bundles of the protease during the substrate binding [63, 69, 70]. The second important feature is the Loop 5 (L5 cap) that was located right above the catalytic dyad [40, 63] (Figure 1.5). A phenylalanine residue (Phe245) in loop 5 bridges two TM segments, TM2 and TM5 blocking the accessibility of substrate to the active site. It was shown later that under different crystallization conditions, Loop 5 adopted different conformations, demonstrating its flexibility [63, 71, 72]. The role of Loop 5 has been recently addressed in the substrate gating mechanism, which is described in detail in section 1.6. For reviews please refer [2, 73, 74].

1.6 Dynamics of rhomboid protease and substrate interaction

Understanding the dynamics of substrate interaction during proteolysis remained a huge challenge because of the lack of crystallized enzyme-substrate complex. However, recent structures of rhomboid-inhibitor complexes, functional analyses and molecular dynamic simulations have provided adequate knowledge in elucidating intramembrane proteolysis. It is believed that an interplay of three complex mechanisms

leads to the cleavage of substrates. These mechanisms are explained in the following sections.

1.6.1 Gating mechanism

One issue that has been debated over a few years is the entry of substrate into the active site of rhomboid. Initial structures of rhomboid protease revealed a V- shaped orientation between TM1 and TM3, which was occupied by Loop1. Since the gap was large enough to accommodate a peptide chain, it was thought that Loop1 acted as a substrate gate [6, 40, 65], although subsequent investigations showed major conformational differences in TM5 and L5 cap, rather than Loop 1 and therefore retracted the idea [63, 64].

In addition to Loop1, two more access routes were proposed for the substrate entry. The first hypothesis arose from the observation that Loop 5 was highly flexible in the crystal structure [63]. Wang *et al.* proposed that a shift in the position of phenylalanine residue (Phe245) on L5 cap permitted the substrate to access the active site. In an alternative model, *Baker et al.* postulated that the displacement of TM5 from TM2 created an opening through which substrates could enter [53, 69]. Evidence for this model came from the tilted conformation occupied by TM5 in the crystal structure [64]. To support this idea, Urban *et al.* performed mutational experiments on TM5 and TM2 of *E. coli* rhomboid [53]. Amino acid substitutions of large aromatic residues on TM2/TM5 displaced TM5 from TM2, allowing gating activity and

increasing substrate cleavage. Conversely, disulphide bridges between TM5 and TM2 prevented entry and resulted in loss of substrate cleavage.

Brooks *et al.* provided a slight variation of substrate docking in *H. influenzae* rhomboid [72]. Similar mutagenesis studies were performed on TM2 and TM5 suggesting that their interaction was crucial for substrate access. However, the sites of interaction varied distinctly between the two rhomboid forms. A kink in the middle of the TM5 helix of *H. influenzae* rhomboid allowed the interactions to occur at the ends and not along the length of the helix, rendering a more open and flexible conformation when compared to *E. coli* rhomboid. Also, crystal structures revealed that Loop 5 was more disordered in *H. influenzae* rhomboid, supporting high mobility. These data suggested the involvement of both Loop 5 and TM5 during substrate gating. Support for this theory came from a co-crystal structure of *E. coli* rhomboid bound to isocoumarin inhibitor that showed a small movement of TM5 along with a large displacement of L5 cap [75]. However, the model was questionable since the inhibitor did not reach the gap between TM5 and TM2. Therefore it was unclear if TM5 would also move to a greater degree if a substrate was bound between the two helices.

To examine this further, Xue *et al.* performed co-crystallization studies of *E. coli* rhomboid with a synthesized phosphonofluoridate inhibitor, CAPF [76]. Crystal structure revealed that this inhibitor reached the TM2-TM5 gap to access the active site. However, the conformational changes caused by CAPF binding were similar to those observed in *E. coli rhomboid-*

isocoumarin inhibitor complex. Both L5 cap and TM5 move to accommodate the inhibitor, although L5 cap rotates and shifts its position to a large extent to expose the catalytic dyad. In comparison, the observed tilt of TM5 from TM2 is relatively small (~7.5 °), although providing sufficient gap for the inhibitor to bind between them. This difference in conformation between L5 cap and TM5 suggests that the access to the catalytic site is primarily regulated by L5 cap movement and the TM5 displacement is essential to accommodate the substrate in the substrate binding cleft (Figure 1.6).

1.6.2 Cleavage sites outside TM of the substrate

Originally it was believed that rhomboids could only cleave the transmembrane domains of substrates; however evidences have suggested that rhomboids can cleave substrates outside of their transmembrane domains. *Drosophila* growth factor Gurken contains a stretch of hydrophobic amino acids including a rhomboid cleavage site at the N terminus. Maegawa *et al.* studied this amino terminal region of Gurken by substituting it with cysteine mutations [71]. The cysteine derivatives were made to react with 4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid (AMS), a highly charged thiol-alkylating agent that was membrane impermeable. Two cysteine substitutions at positions 247 and 250, located between a scissile bond, received *E. coli* rhomboid-dependent proteolysis suggesting that cleavage recognition site resides outside the membrane. More reports confirm this theory: proteolysis of Mgm1 outside the mitochondrial inner

membrane [56], cleavage of an artificial fusion construct of *E. coli* LacY protein that has a Ser-Asp scissile bond outside the transmembrane domain [3, 71], and processing of chimeric human epidermal growth factor containing an Ala-Gly scissile bond in the N- terminal region of the transmembrane domain [77]. However, this mechanism is not yet properly understood as very few crystal structures are present to elucidate this activity. One postulated theory is that substrates, with a recognition motif above the TM helix, are structurally flexible so that they can bend and access the active site [3, 78-80].

1.6.3 Membrane thinning

Substrate entry is also assisted by membrane thinning around the proteases. Early evidence of crystal structure of *E. coli* rhomboid showed that the hydrophobic belt around rhomboid protease was narrower than the rest of the hydrophobic region of the membrane bilayer [63]. A recent structure of rhomboid in lipid environment confirms this theory that the lipid bilayer adjusts its local thickness around the rhomboid molecule [75]. Molecular dynamic simulations have revealed that L1 loop interacts with the membrane and tilts the rhomboid molecule by ~12° [6, 70, 81]. Though the implication of membrane compressions is unclear, it is believed that the tilt of rhomboid molecule along with local membrane compression submerge the gate region and maintain the position of the catalytic site beneath the membrane surface [70]. An important point to be noted here is that the molecular simulation

experiments were performed in the presence of rhomboid only without any substrate molecule. It is possible that a substrate molecule could influence lipids surrounding the rhomboid to a greater degree to facilitate the gating transition. This remains an important area that requires further investigation.

1.7 Substrate specificity

A related issue is to determine how rhomboids identify their substrates. Both prokaryotic and eukaryotic rhomboids have been shown to cleave Spitz, despite their low sequence homology [67, 82]. Expression studies have revealed that *Providencia stuartii* rhomboid AarA cleaves Spitz and activates EGFR signalling in *Drosophila* while *Drosophila* Rhomboid-1 functionally complements AarA in *Providencia stuartii* [67, 83]. The same is observed for mitochondrial rhomboids, where yeast mitochondrial rhomboid Pcp1 deficiency is restored by mammalian homolog, PARL [57]. These studies suggest that there are specific shared structural features in the substrates that enable rhomboids from different origins to recognise them [51, 52].Substrate specificity, first studied using *Drosophila* Spitz, explained the importance of helix destabilizing residues in the transmembrane region of the substrate [82]. Using chimeric substrates, the recognition motif of Spitz (ASIASGA) was mapped. In order to specifically identify the residues that were responsible for recognition by Rhomboid-1, amino acids pairs of the

motif were substituted into an uncleavable substrate. The insertion of "GA" or "GG" into the N- terminal TM region of a non-substrate molecule resulted in a cleaved product suggesting that Rhomboid-1 targeted these residues within the transmembrane helix of the uncleavable chimera. These residues (glycine and alanine) are known to be involved in destabilizing helices [51, 67]. The residues are thought to facilitate cleavage by unwinding the helix and subjecting the peptide bonds to hydrolysis [15]. The location of these residues in the transmembrane helix of the substrates is crucial as it determines the accurate recognition by rhomboids. It is predicted that these residues should occur in the initial region of the transmembrane helix which helps to position the cleavage site at the top of the transmembrane helix in the juxtamembrane region [82]. Substrates have been predicted and determined using this requirement. Urban *et al*. identified two novel targets in *Toxoplasma gondii* and *Plasmodium falciparum* using this model [82, 84, 85]. However, these helix-destabilizing residues are not required in all substrates [67, 82, 86]. For substrates whose cleavage sites are located outside the transmembrane domain, these helix-destabilizing residues are not considered essential [3, 87].

To refine the features of substrate determination, Akiyama *et al.* performed random site-directed mutagenesis of cleavage site residues. An artificial chimeric substrate was created with periplasmic β-lactamase domain, the second transmembrane domain of β-galactoside permease (LacY) fused to maltose binding protein (MBP) at the C- terminus (N'-Bla-

LacY-MBP-C'). This protein was not a physiological substrate for *E. coli* rhomboid. Surprisingly however, this fusion protein was found to be cleaved by the rhomboid both *in vivo* and *in vitro* [71, 86]. It was discovered that the amino acid residues flanking the scissile bond at P1 and P1' positions were occupied by small and negatively charged residues; a finding that was not observed in other experiments. More recently, a sequence recognition motif has been identified as the primary substrate determinant [3]. A detailed examination of TatA, a physiological substrate from *Providencia stuartii*, revealed a primary sequence recognition motif, in which three residues at positions P4, P1 and P2' require large hydrophobic, small and hydrophobic residues respectively and act as cleavage recognition determinants (Figure 1.7). This recognition motif was found to be more essential that the helix-destabilizing residues and could point to the cleavage site, not only within the transmembrane domain but also outside of it. The sequence recognition motif was found to be conserved amongst many rhomboid substrates; however it is not universal. There are rhomboids that recognize regions other than the cleavage sites of substrates. For example, mitochondrial rhomboid substrate, Ccp1 lacks P4-P1-P2' motif that is mentioned above [3, 56, 88]. Similarly, *Plasmodium falciparum* substrate EBA-175 lacks a requirement for large hydrophobic residue at P4 [89]. Additional substrate recruitment mechanisms have also been identified. Thrombodulin, a substrate for human rhomboid RHBDL2, is recognized by its cytoplasmic domain [62] and HtrA2 requires an additional factor for its recruitment to

the mitochondrial rhomboid, PARL [3, 90]. Nevertheless, the former results validate the recognition motif as a general determinant for cleavage by rhomboids. One of the objectives of my thesis is to evaluate the potential of *E. coli* TatA as a substrate using this consensus recognition motif, which is later discussed in Chapter 3. {For review, please see [74, 80, 91].

1.8 Functional insights into rhomboids

The rest of the chapter focusses on the known biological roles of rhomboids. Rhomboids represent significant targets of study as they are involved in a number of diseases. Given their ubiquitous nature, only a few functions are known [5, 92-95]. Rhomboids have been recognized to regulate a broad range of processes: initiating animal cell signaling, regulating mitochondrial fission and fusion [56, 96], disassembling adhesive junctions in eukaryotic parasites [80, 85, 97], regulating plastid translocon components in plants [98, 99], and facilitating quorum sensing in bacterial *Providencia stuartii* [83, 100]. Detailed investigations have identified at least one function of rhomboids in a dozen organisms (Table 1) [1]. The following sections will explain their functions in detail (Figure 1.8)

1.8.1 Role in signaling

Drosophila melanogaster Rhomboid-1 regulates the EGFR pathway by cleaving its activating protein, Spitz [39]. Spitz is trapped in the endoplasmic

reticulum and is transported to the Golgi by a chaperone protein, Star. In the Golgi bodies, Spitz is activated by Rhomboid-1 mediated cleavage. The released ligand acts as signal-generating component and controls multiple aspects of cell fate behaviour [101-103]. Rhomboid-1 also acts as a regulator by cleaving Star to control the level of Spitz release into the neighbouring cells [104]. A slightly modified role of rhomboids is observed in nematode, *Caenorhabditis elegans* in which rhomboid cleaves an EGF signaling protein only to amplify and propagate the ligand signal [105]. Even less understood is the function of rhomboids in mammals. Recent investigations have revealed that rhomboid expression increases in cancer cells and may have a link in the growth factor signaling pathway [62, 77, 106]. The human rhomboid RHBDL2 is found to cleave thrombodulin (a membrane receptor for thrombin) and EGF receptors in an *in vitro* assay, however the in vivo cleavage of these substrates still have to be assessed.

Another example of signaling is best studied with AarA rhomboid of *Providencia stuartii* [100]. AarA was first identified in a screen of genes involved in quorum sensing in *Providencia stuartii* [107]. Several years later, studies revealed that AarA was a rhomboid protease and could replace the role of Rhomboid-1 in *Drosophila* [83]. An indirect genetic screening study was performed to identify genes that complemented *aarA* gene mutation in *P. stuartii* [100]. It was found that *tatA* gene from *Proteus mirabilis* and *E. coli* efficiently complemented the AarA phenotype; surprisingly *tatA* from *P. stuartii* failed to do so, implying that AarA and TatA were functionally related.

It is now known that TatA is activated by AarA- mediated cleavage of a small amino terminal extension of the protein. The released TatA multimerizes to form a large channel and is believed to be involved in the export of quorum sensing factors [83, 100]. It is interesting to note that while the presence of the amino terminal extension of TatA inhibits the formation of large complexes in *P. stuartii* [108], other bacteria do not contain this extension implying that the role of AarA in *P. stuartii* is unique [1, 109]. This is the only known example of TatA activation and regulation. The role of AarA in quorum sensing is indefinable as the identity of the quorum sensing molecule is not known yet, although it is believed that either the signaling molecule or some factor needed for the signal production is dependent on the oligomerization of TatA [5].

1.8.2 Role in protozoan parasites

The potential of rhomboids in apicomplexan parasites is one of the best elucidated functional studies of rhomboid protease. *Toxoplasma gondii* is a protozoan parasite which causes toxoplasmosis. During cell invasion, the parasite secretes microneme protein complexes (MICs) that act as cell surface adhesins and engages the host receptors for its movement into the host cell. After invasion, the MIC proteins are redistributed to the posterior side of the parasite, so that they are kept away from the host cell immune system. Once the parasite completely enters the host, the adhesins are cleaved off and the invasion is completed. Earlier studies revealed that the

microneme proteins were cleaved through an intramembrane proteolysis mechanism [110]. In an experiment performed by Urban *et al*. [82], these proteins were found to be processed by *Drosophila* Rhomboid- 1 and *P. stuartii* rhomboid AarA, implicating the role of rhomboid in host cell invasion. *T. gondii* is known to contain six rhomboid proteases, of which two rhomboids, TgROM4 and TgROM5 are known to process adhesins [111]. Lack of TgROM4 and TgROM5 result in the inability of the parasite to orient and glide in the host cell suggesting that rhomboid participation is crucial to host cell invasion [112]. TgROM5 localizes to the posterior side of the parasitic cell, while TgROM4 is evenly distributed across the cell surface [84, 111]. In addition to the MIC proteins, TgROM4 is known to cleave AMA1 adhesin and this triggers the intracellular replication of the parasite after invasion [74, 80, 97, 113]. A similar phenomenon is observed in malarial parasite *Plasmodium falciparum*, in which PfROM4 catalyses the proteolysis of adhesins. Additional roles of rhomboids have been identified in other parasites. For example, in *Entamoeba histolytica* (a parasite causing amoebic dysentery), EhROM1 is proposed to play a role in immune evasion along with parasitic invasion [114]. Host cell immune system is evaded by receptor capping; virulent surface proteins, called lectins are translocated to the posterior pole of the parasite. EhROM1 is also found to redistribute and colocalize along with lectins. These proteins are then enclosed into membrane vesicles and released into the extracellular medium to prevent recognition from the host

cell immune system [115, 116]. It is implied that EhROM1 has a role in shedding of these vesicles [117], although its exact mechanism is unclear.

1.8.3 Role in mitochondria

A sub-class of rhomboids that differ from the secretase family are the mitochondrial rhomboids. As mentioned in section 1.4, mitochondrial rhomboids have an extra transmembrane helix fused to the N terminus (1+6 TM topology). As a result, their active sites are in the opposite orientation when compared to other rhomboid classes and can cleave Type II membrane proteins.

Mitochondrial rhomboids are known to be involved in mitochondrial quality control, membrane integrity and apoptosis [94]. The dynamic morphology of mitochondria is maintained by a balance between membrane fission and fusion. Many proteins are known to be involved in governing mitochondrial membrane dynamics, of which Mgm1 GTPase facilitates membrane fusion [118]. Lack-of-function mutation of Mgm1 results in mitochondrial fragmentation, a phenotype observed to be very similar to the mitochondrial rhomboid, Pcp1 mutation. Both proteins reside in the inner mitochondrial membrane where Pcp1 cleaves a soluble GTPase domain of Mgm1 that is released into the intracellular space. Pcp1 also cleaves another protein, Cytochrome C peroxidase (Ccp1) that is involved in oxidative-stress signaling [55]. The cleavage of Mgm1 and Ccp1 by Pcp1 occurs through an

alternative topogenesis mechanism [96]. Both Mgm1 and Ccp1 are single pass transmembrane proteins. These proteins are cleaved in a second hydrophobic site other than the primary TM domain. An ATP driven motor or an ATPase pulls the protein out of the membrane, translocating the primary TM domain into the matrix. This positions the second hydrophobic region into the membrane, which is further cleaved by Pcp1. Regulation is achieved by the balancing the levels of cleaved and uncleaved isoforms of these proteins, that aid in mitochondrial fusion [96]. Similarly, OPA-1, a homolog of yeast Mgm1, is regulated by Drosophila Rhomboid-7 [119]. OPA-1 is involved in cristae membrane remodelling and dysregulation of OPA1 is believed to be involved in apoptosis. [120].

In mammals, mitochondrial rhomboid PARL is associated with cristae modelling and apoptosis instead of mitochondrial fusion [120], malfunctions of which have been linked to Parkinson's disease [121]. PINK1, a serine-threonine kinase, is found in mitochondria and cytosol. PINK1 is maintained at low levels within the healthy mitochondria and is elevated when the mitochondrial becomes dysfunctional. In healthy mitochondria, PARL catalyzes the cleavage of PINK1 and releases its mature form into the cytosol, where it is rapidly degraded in a proteasome-dependent manner [122]. Upon mitochondrial stress, PINK1 catalysis by PARL is inhibited, leading to the accumulation of PINK1 precursor on the cytosolic side. PINK1 then recruits an E3 ubiquitin ligase protein, Parkin in the cytosolic face and together they eliminate damaged organelle through mitophagy [123, 124]. Although more

evidence is required to understand the interaction between PARL, PINK1 and Parkin, evidences suggest a role of PARL in the maintenance of mitochondrial integrity [125]. For review, please refer [1, 74, 94, 99]

1.8.4 Role of iRhoms

iRhoms are pseudoproteases that are catalytically inactive. Though considered "dead enzymes", iRhoms are now known to be involved in regulating pathways that their active homologs take part in. The first evidence of iRhom function came from the genetic analysis in *D. melanogaster* [126]. Blockade of iRhom expression in the central nervous system causes the flies to exist in a sleep-like state. This resembles the effect of overexpression of *Drosophila* Rhomboid-1 and Star in EGFR pathway of the central nervous system. Further links established that iRhom acts as a suppressor for EGFR signaling pathway [127]. This is fascinating because active rhomboids initiate the activation of EGFR pathway by processing Spitz while iRhoms counteract their activity. Biochemical and genetic analysis provide an explanation for this mechanism. In *Drosophila*, iRhom binds to Spitz and Gurken ligands in the ER and direct these proteins to an ER-associated degradation (ERAD) process, instead of allowing their transport to Golgi bodies [128] (Figure 1.9). Mammalian homolog, iRhom2 is also present in ER, though its function is slightly different from its *Drosophila* counterpart. Rather than blocking the signaling pathway by triggering ERAD of ligands,

iRhom2 promotes forward trafficking of proteins from the ER to Golgi. TACE [tumor necrosis factor (TNF)-converting enzyme] is synthesized as an inactive protein in the ER. Furin, a pro-protein convertase, activates TACE in the Golgi by cleaving its inhibitory domain. The matured TACE then localizes to the cell surface to process an inflammatory cytokine, TNF. In the absence of iRhom2, TACE is not transported from ER to Golgi, does not undergo Furin-mediated maturation and does not cleave TNF (Figure 1.9). Physiological significance of iRhom2 has been studied in iRhom2 knockout mice where certain inflammatory defects have been observed [129, 130]. Given the function of iRhom2, further investigation is required to understand how iRhom2 transports TACE to the Golgi and how it interacts with the other members of the membrane trafficking machinery. Detailed information is provided in review [7].

1.9 Regulation of rhomboid activity

While a number of functions of rhomboids have been characterized, surprisingly little information is available about the regulation of their activity. Rhomboids differ from other intramembrane proteases in that they do not require any pre-processing of substrates. Hence different strategies must exist to govern rhomboid activity. The most common form of regulation is the compartmentalization where trafficking of substrates from one organelle to another is essential for contact with rhomboids. In *Drosophila*,

Star transports Spitz from ER to Golgi, where Spitz is cleaved by Rhomboid-1 and is released as an active growth factor ligand. In the absence of Star, Spitz is retained in the ER [49, 131]. Alternatively, Star can also be cleaved by Rhomboid-1; a mechanism to efficiently down-regulate the import of Spitz into Golgi [104]. This is the only case where a rhomboid is found to cleave Type II membrane protein, Star. Furthermore, additional rhomboids have been found to exist in the ER that may possibly be involved in the cleavage Star and Spitz when needed [132]. Similarly, in apicomplexan parasites, the MICs along with rhomboids are redistributed and localized to the posterior pole of the cell from where the adhesins are released and the parasitic cell is freed. Evidence has shown that the cytoplasmic domain of rhomboids play a role in reorientation of adhesins in *T. gondii* [133]. An interesting variation of this mechanism is observed in mitochondria, where the ATP levels regulate mitochondrial fusion [55]. Under low concentrations of ATP, Mgm1 is anchored to the inner mitochondrial membrane by its primary transmembrane domain and cannot be cleaved by mitochondrial rhomboid, Pcp1. This occurs in unhealthy mitochondria and inhibits its fusion to healthy mitochondria. However at high ATP levels, an ATP motor pulls the transmembrane domain of Mgm1 out of the membrane, helping to position the second hydrophobic region in the inner mitochondrial membrane. Pcp1 cleaves the hydrophobic region and the released soluble domain of Mgm1 participates in mitochondrial fusion.

It is now known that all rhomboids do not share the same mode of regulation. While compartmentalization of rhomboid and its substrates seems to be the most common form of regulation, activity can also be modulated by proteolytic release of domains in rhomboid. The N- terminal domain of PARL is cleaved by two proteolytic events: α- cleavage and β-cleavage [134]. Once PARL has reached the inner mitochondrial membrane, the mitochondrial targeting motif at the N- terminus is cleaved by an unknown protease. This process is termed α- cleavage which generates a mature form of PARL with catalytic activity. The second cleavage, called the β – cleavage cleaves the mature PARL into two fragments, a small N- terminal fragment, Pβ and a bigger fragment called PARL C- terminal fragment (PACT), with the latter retaining rhomboid protease activity. After the release, Pβ peptide reaches the nucleus and triggers mitochondrial biogenesis while PACT blocks mitochondrial fusion [125]. The status of damaged mitochondrion is maintained in two ways: the damaged mitochondrion serves as a signal for the β-cleavage of PARL. As a result, PACT stops mitochondrial fusion and prevents the damaged organelle from fusing to healthy mitochondria. The damaged mitochondrion is then isolated, providing an opportunity for the cell to repair this “compromised organelle”. In the case of extensive damage, the mitochondrion loses its electrochemical potential. At this point, PARL is further cleaved and rendered catalytically inactive. Consequently, proteolytic processing of PINK1 is inhibited, which accumulates in the cytosol and triggers mitophagy by recruiting PARKIN

[135]. Meanwhile, Pβ peptide localizes to the nucleus and activates mitochondrial biogenesis. This intricate proteolysis event that controls mitochondrial formation is an additional method of regulation of activity of rhomboid proteases.

1.10 Thesis objectives

One possible mode of regulation of rhomboids could be through oligomerization as many membrane proteins are known to regulate through this mechanism. An example of this is the sodium proton antiporter, NhaA, which dimerizes to maintain the transporter under stress conditions [136]. Our laboratory had previously found that *H. influenzae* rhomboid, *hi*GlpG formed oligomers during crystallization studies. This led us to investigate the oligomeric state of three bacterial rhomboids, *ec*GlpG (*E. coli*), *hi*GlpG and YqgP (*Bacillus subtilis*). Chapter 2 of this thesis will focus on characterizing the oligomeric state of the above-mentioned rhomboids using gel filtration, crosslinking and pull-down assays.

Given the numerous crystal structures of *ec*GlpG and that little is known about its physiological role, Chapter 3 will investigate the potential of *E. coli* TatA as a substrate for *ec*GlpG using the consensus recognition motif prediction reported by Strisvosky *et al.* [3]. Additionally, affinity pull-down and co-immunoprecipitation techniques are examined in the identification of potential substrates for this rhomboid.

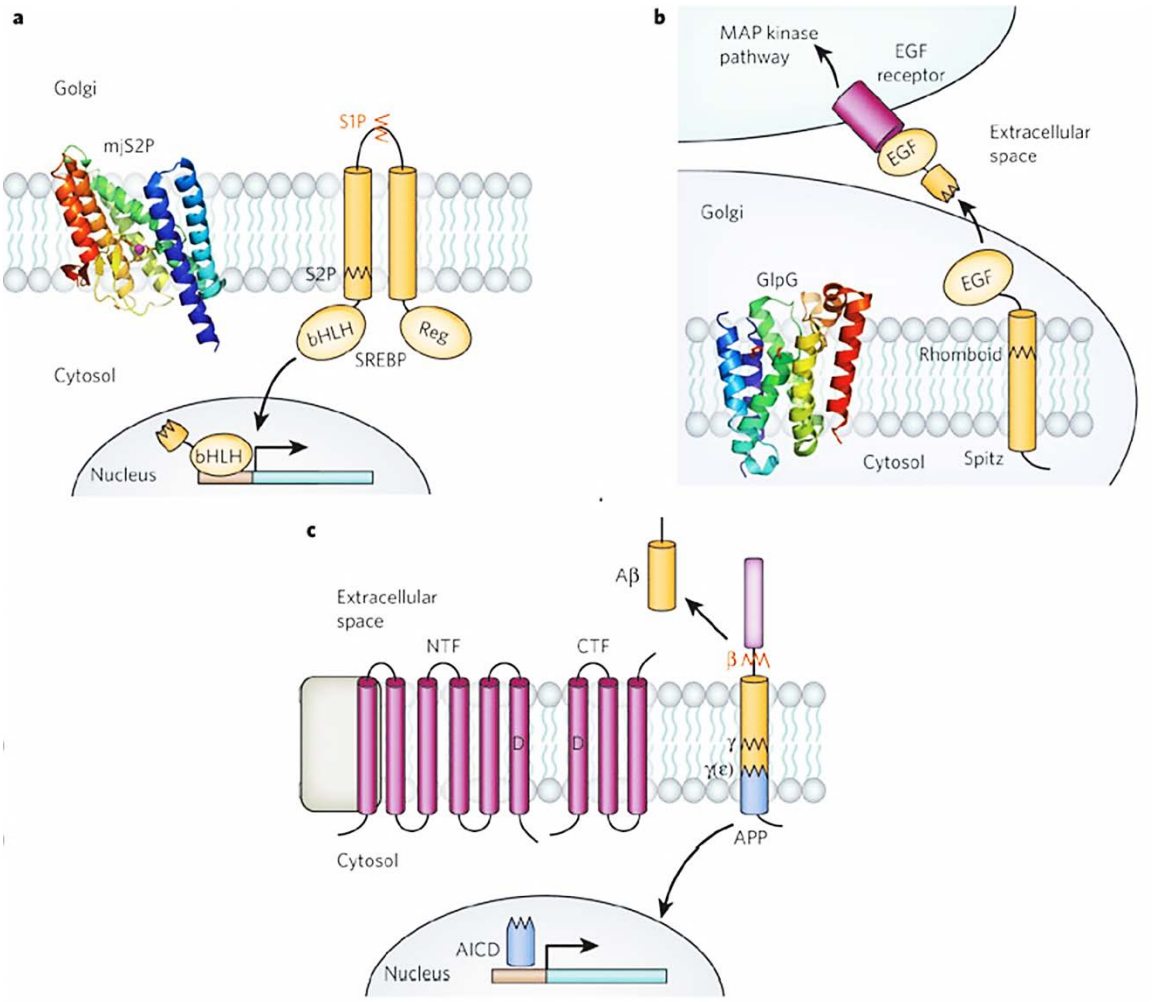


Figure 1.1: Classification of intramembrane proteases [4]: A. Activity of intramembrane metalloprotease S2P. In response to low cholesterol, Sterol Regulatory Element Binding Protein (SREBP) is transported from ER to Golgi, where it is primarily cleaved by S1P (red line) followed by S2P (black line). The released transcription activating domain of SREBP reaches the nucleus where it triggers the transcription of genes controlling cholesterol biosynthesis. **B. Activity of intramembrane serine protease, *Drosophila* Rhomboid 1.** The EGF factor Spitz is transported to the Golgi where it is activated by rhomboid-mediated cleavage. The released luminal domain activates EGF receptors on neighbouring cells. **C. Activity of intramembrane aspartyl protease, Presenilin.** β-secretase initially cleaves amyloid precursor protein (APP) (red line) at an extracellular domain followed by the proteolytic processing of the active subunit of γ-secretase, Presenilin (black line) to release Aβ, which is secreted from cells. The remaining peptide is further cleaved by γ-secretase at the ε site (second black zigzag line) to release APP intracellular domain which is thought to play a role in transcription regulation **Adapted from** Erez E, F.D., Bibi E., Nature, 2009. 459(7245): p. 371-378

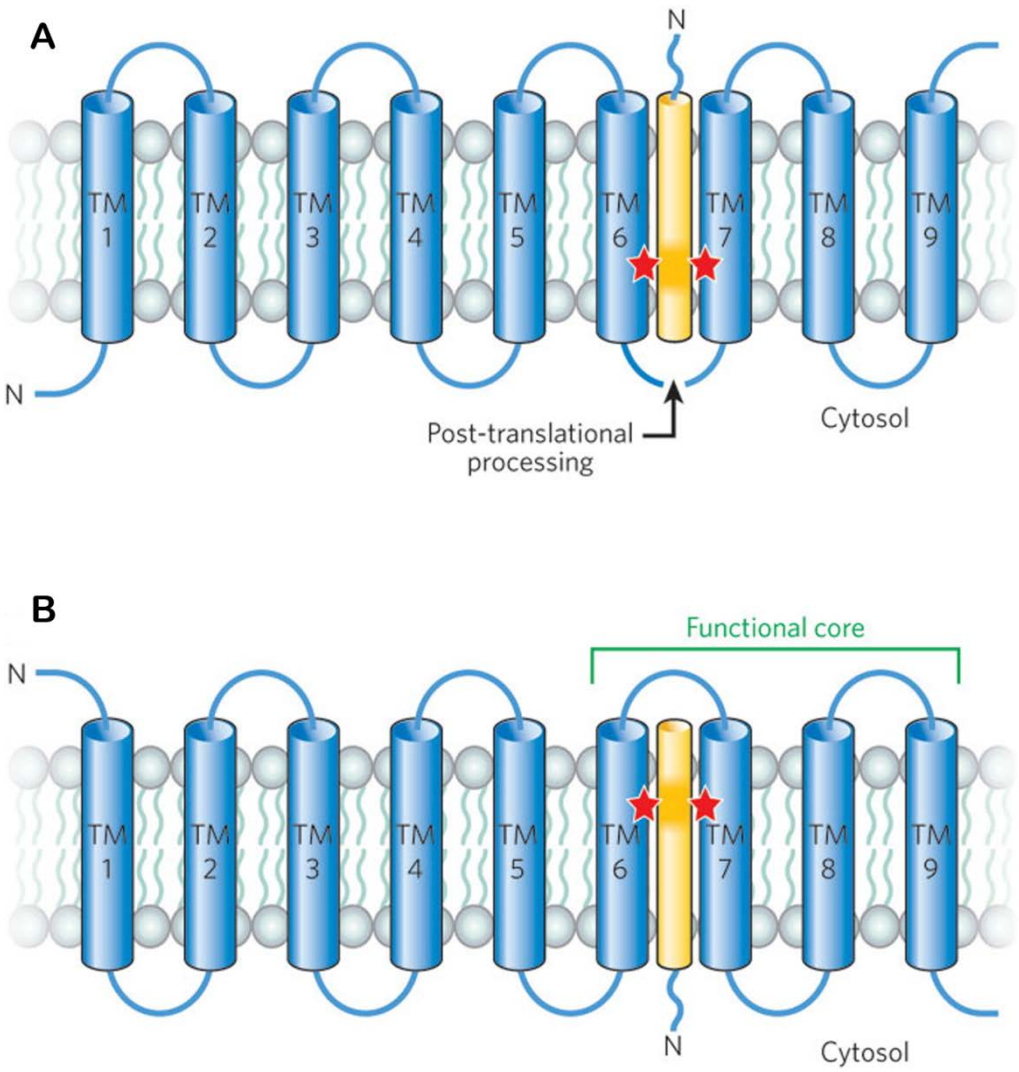


Figure 1.2: Difference in topological orientation between Presenilin and signal peptide peptidase (SPP) [4]. Presenilin favours cleavage of Type I membrane proteins, SPP preferentially cleaves Type II membrane proteins (substrates represented in yellow cylinders). As a result, the catalytic aspartates (red stars) of Presenilin and SPP are in opposite orientation. **Adapted from** Erez E, F.D., Bibi E., Nature, 2009 459(7245): p. 371-378

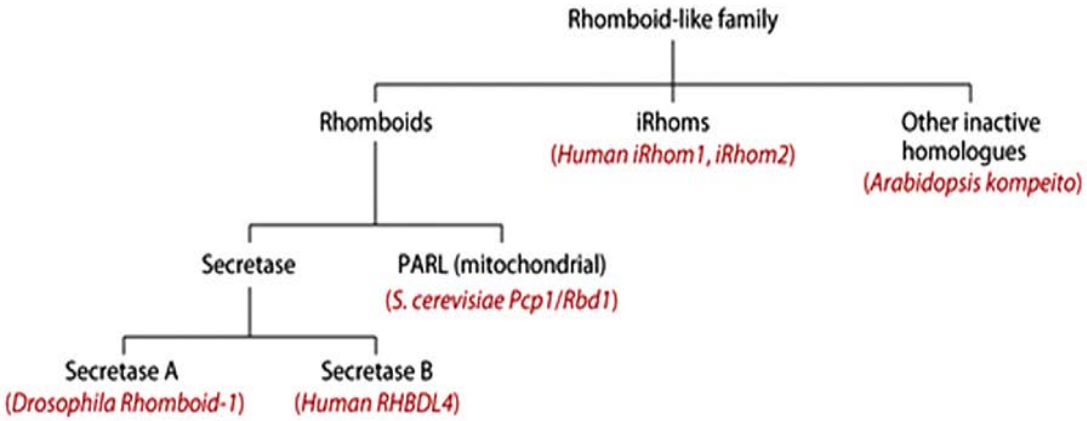


Figure 1.3: Classification of rhomboids [5]. Rhomboids are classified into three groups: active rhomboid proteases, catalytically inert iRhoms and other inactive rhomboid homologs that cannot be clearly assigned to either of the two groups. The active rhomboids are further classified based on their intracellular localization: secretase rhomboids and mitochondrial rhomboids. Bacterial rhomboids are grouped under the secretase family as they contain the basic six TM architecture. **Adapted from** Freeman, M., Annual Reviews in Genetics, 2008. **42**: p. 191-210

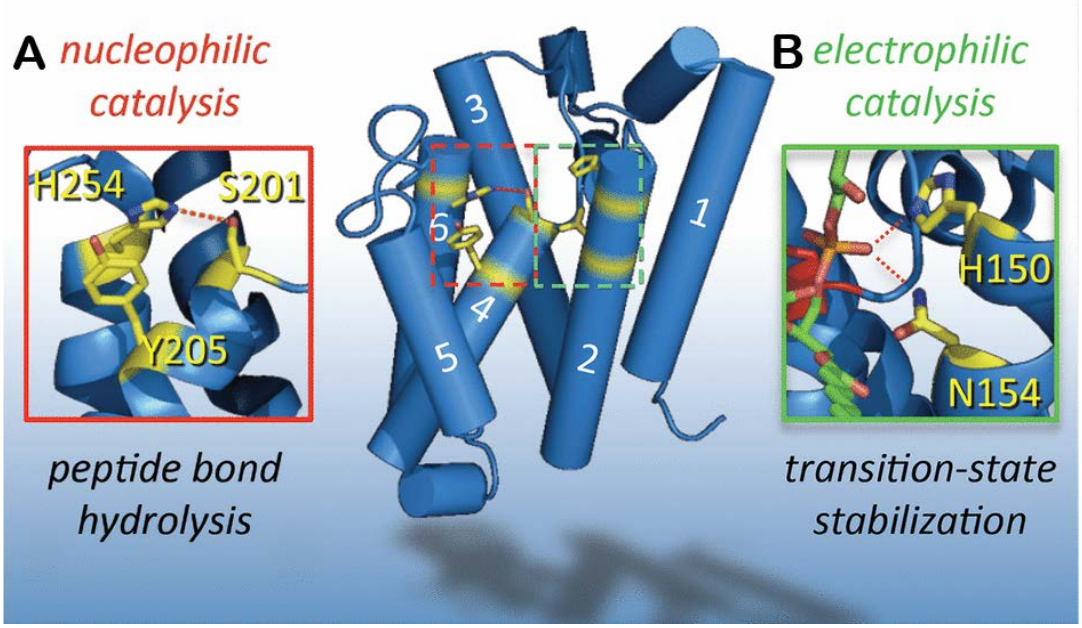


Figure 1.4: Structural features of rhomboid protease [2]: A. Nucleophilic catalysis is carried out by a catalytic dyad comprised of serine on TM4 and histidine on TM6. The hydrogen bonds between histidine and serine are depicted in red dotted lines. **B.** Electrophilic catalysis between histidine and asparagine from TM2 is shown on the right side. This is required to stabilize the oxyanion intermediate. The hydrogen bonds between the phosphate oxygen and side chains of asparagine and histidine are shown in red dotted lines. **Adapted from** Urban, S., The Biochemical journal, 2010. **425**(3): p. 501-512

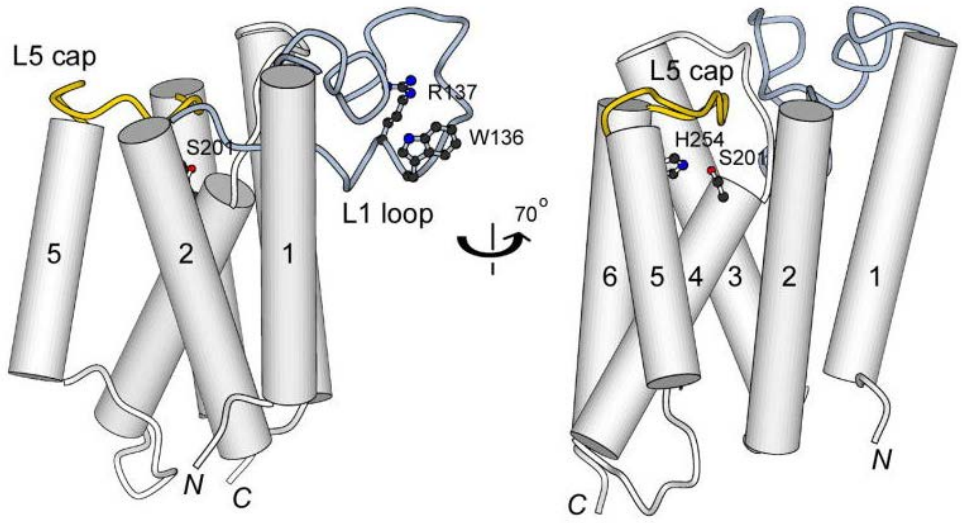
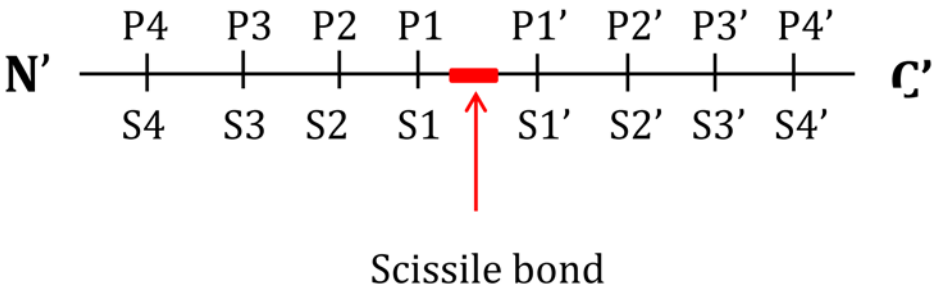


Figure 1.5: Structural features of Loop 1 and Loop 5 [6]: The highly conserved WR motif present near the bottom of Loop L1 is shown above. Also, Loop 5 (L5 cap) that is located right above the catalytic dyad is seen. The phenylalanine residue on loop 5 bridges two TM segments, TM2 and TM5, blocking the accessibility of substrate to the active site. **Adapted from** Wang, Y., et al., The Journal of Molecular Biology, 2007. **374**(4): p. 1104-1113

A.



B.

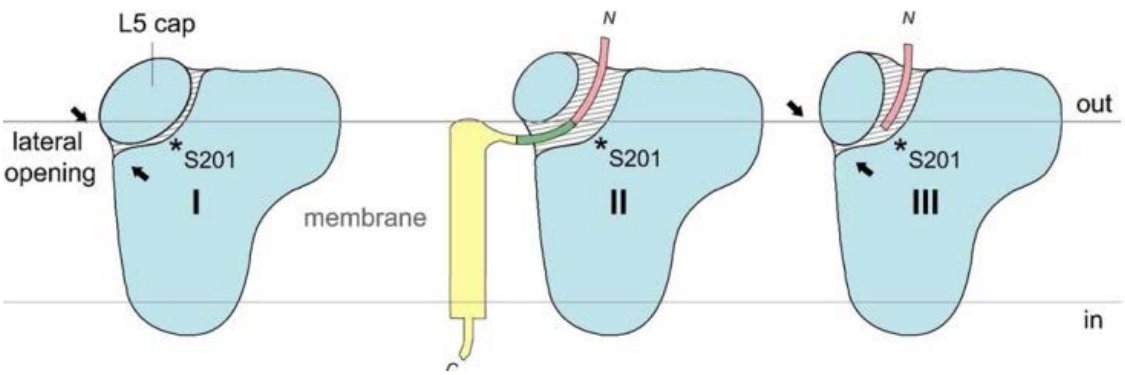


Figure 1.6: Substrate entry is regulated by L5 cap [8]: A. Nomenclature of cleavage positions of substrate and their corresponding positions in the enzyme binding site. **B.** In the presence of substrate, L5 cap opens allowing the substrate to reach the active site. Interaction with the catalytic serine (S201) releases the C- terminal peptide fragment on the S1' side of the substrate binding cleft. The enzyme changes conformation again to close the opening so that the contact of the aqueous active site to the external lipid bilayer is minimized. **Adapted from** Xue, Y. and Y. Ha. The Journal of Biological Chemistry, 2012. 287(5): p. 3099-3107

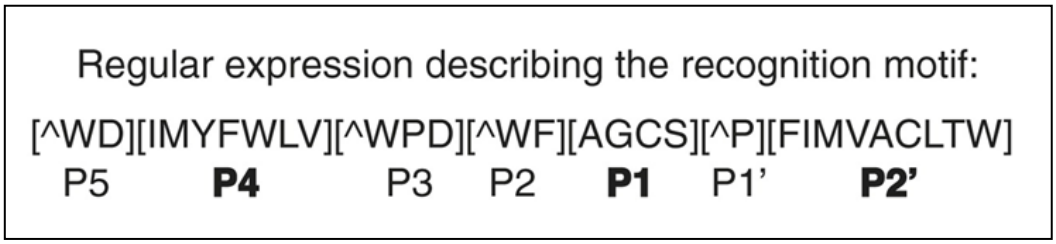


Figure 1.7: Consensus recognition motif observed in many rhomboid substrates [3]: Three residues at positions P4, P1 and P2' require large hydrophobic, small and hydrophobic residues respectively and act as cleavage recognition determinants for most rhomboid proteases. **Adapted from** Strisovsky, K., H.J. Sharpe, and M. Freeman, Molecular Cell, 2009. 36(6): p. 1048-1059

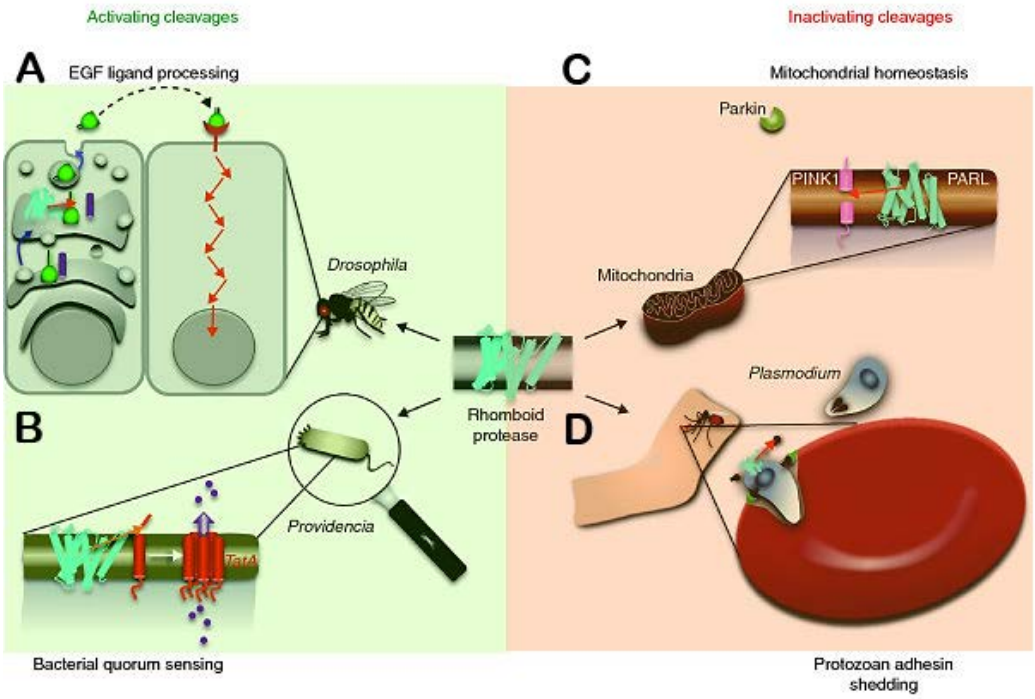


Figure 1.8: Biological roles of rhomboid proteases [1]: A. *Drosophila* Rhomboid-1 initiates EGF signaling by cleaving an EGF growth factor, Spitz in the Golgi after it is chaperoned by Star from ER. **B.** P. stuartii rhomboid AarA activates TatA by cleaving a small amino terminal extension. TatA then oligomerizes to form a pore which transports fully folded proteins across the membrane. **C.** Mitochondrial rhomboid PARL cleaves PINK-1, inhibiting Parkin recruitment and downregulates mitophagy. **D.** Malarial rhomboids dissemble the junction formed by host-parasite interactions at the end of invasion by cleaving adhesins that are located on the posterior side of the parasite. **Adapted from** Urban, S. and S.W. Dickey, Genome biology, 2011. **12**(10): p. 231

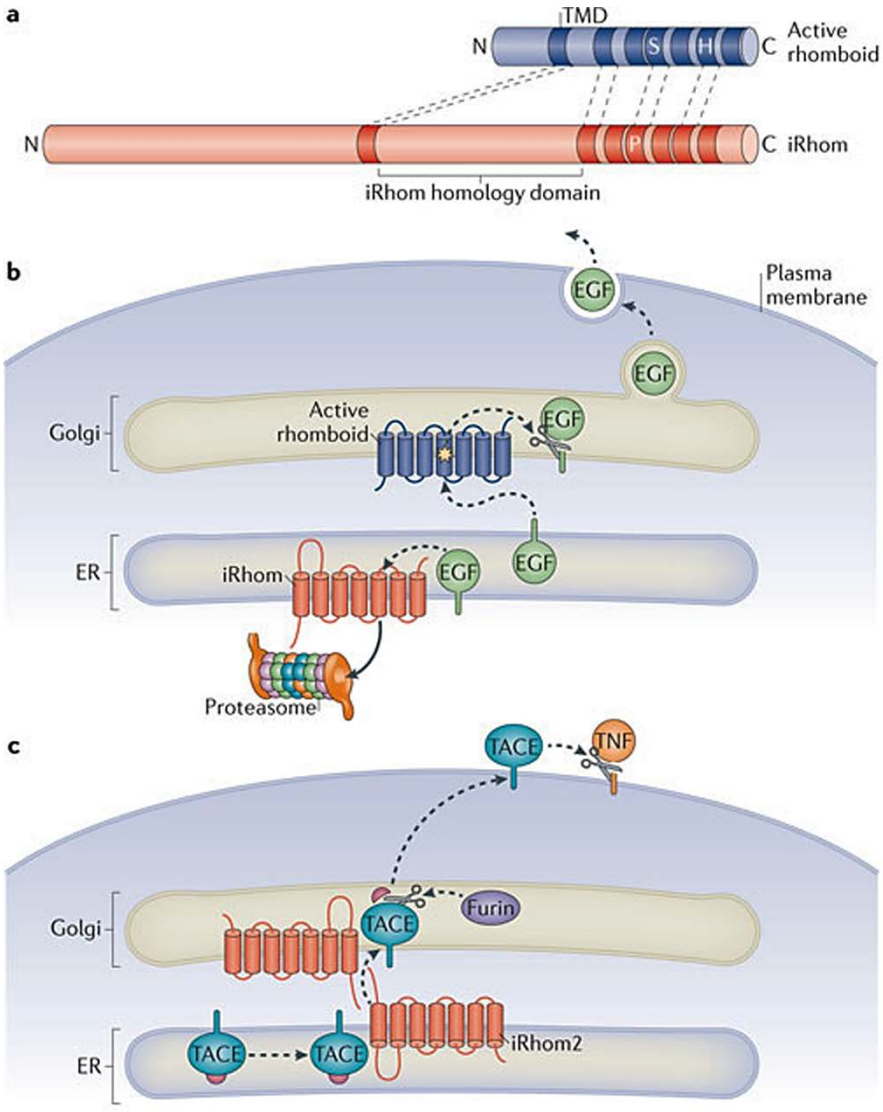


Figure 1.9: Function of iRhoms [7]. A. Differences between active rhomboids and catalytically inert iRhoms. Absence of a catalytic serine and presence of proline residue disrupts the active site, rendering iRhoms inactive. **B.** Active rhomboids and iRhoms counteract each other in EGFR signaling in *Drosophila*. While Rhomboid- 1 cleaves Spitz in the Golgi and activates EGF signaling pathway, iRhoms negatively regulate EGFR signaling by binding to Spitz in the ER and targeting it to an ER-associated degradation (ERAD). This controls the amount of Spitz available for proteolytic processing by active rhomboids in the Golgi. **C.** iRhom 2

promotes forward trafficking of proteins from the ER to Golgi in mammals. TACE is synthesized as an inactive protein in the ER. Furin, a pro-protein convertase, activates TACE in the Golgi by cleaving its inhibitory domain. The matured TACE then localizes to the cell surface to process an inflammatory cytokine, TNF. In the absence of iRhom2, TACE is not transported from ER to Golgi, does not undergo Furin-mediated maturation and does not cleave TNF. **Adapted from** Adrain, C. and M. Freeman. Nature reviews, Molecular cell biology, 2012. 13(8): p. 489-498

Chapter 1: References

1. Urban, S. and S.W. Dickey, *The rhomboid protease family: a decade of progress on function and mechanism.* Genome biology, 2011. **12**(10): p. 231.
2. Urban, S., *Taking the plunge: integrating structural, enzymatic and computational insights into a unified model for membrane-immersed rhomboid proteolysis.* The Biochemical journal, 2010. **425**(3): p. 501-12.
3. Strisovsky, K., H.J. Sharpe, and M. Freeman, *Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates.* Molecular cell, 2009. **36**(6): p. 1048-59.
4. Erez E, F.D., Bibi E., *How intramembrane proteases bury hydrolytic reactions in the membrane.* Nature, 2009. **459**(7245): p. 371-378.
5. Freeman, M., *Rhomboid Proteases and their Biological Functions.* Annual Reviews in Genetics, 2008. **42**: p. 191-210.
6. Wang, Y., Maegawa S, Akiyama Y, Ha Y. *The role of L1 loop in the mechanism of rhomboid intramembrane protease GlpG.* Journal of molecular biology, 2007. **374**(4): p. 1104-13.
7. Adrain, C. and M. Freeman, *New lives for old: evolution of pseudoenzyme function illustrated by iRhoms.* Nature reviews. Molecular cell biology, 2012. **13**(8): p. 489-98.
8. Xue, Y. and Y. Ha, *Catalytic mechanism of rhomboid protease GlpG probed by 3,4-dichloroisocoumarin and diisopropyl fluorophosphonate.* The Journal of biological chemistry, 2012. **287**(5): p. 3099-107.
9. Seife, C., *Blunting nature's Swiss army knife.* Science, 1997. **277**(5332): p. 1602-3.
10. Lopez-Otin, C. and J.S. Bond, *Proteases: multifunctional enzymes in life and disease.* The Journal of biological chemistry, 2008. **283**(45): p. 30433-7.
11. Freije JM, Balbín M, Pendás AM, Sánchez LM, Puente XS, López-Otín C., *Matrix metalloproteinases and tumor progression.* Advances in experimental medicine and biology, 2003. **532**: p. 91-107.
12. Nalivaeva NN, Fisk LR, Belyaev ND, Turner AJ, *Amyloid-degrading enzymes as therapeutic targets in Alzheimer's disease.* Current Alzheimer research, 2008. **5**(2): p. 212-24.
13. Dollery, C.M. and P. Libby, *Atherosclerosis and proteinase activation.* Cardiovascular research, 2006. **69**(3): p. 625-35.
14. Brown MS, Ye J, Rawson RB, Goldstein JL., *Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans.* Cell, 2000. **100**(4): p. 391-8.
15. Wolfe, M.S., *Intramembrane-cleaving proteases.* The Journal of biological chemistry, 2009. **284**(21): p. 13969-73.

16. Weihofen, A. and B. Martoglio, *Intramembrane-cleaving proteases: controlled liberation of proteins and bioactive peptides.* Trends in cell biology, 2003. **13**(2): p. 71-8.
17. Selkoe, D.J. and M.S. Wolfe, *Presenilin: running with scissors in the membrane.* Cell, 2007. **131**(2): p. 215-21.
18. Lichtenthaler, S.F., C. Haass, and H. Steiner, *Regulated intramembrane proteolysis--lessons from amyloid precursor protein processing.* Journal of neurochemistry, 2011. **117**(5): p. 779-96.
19. De Strooper, B. and W. Annaert, *Novel research horizons for presenilins and gamma-secretases in cell biology and disease.* Annual review of cell and developmental biology, 2010. **26**: p. 235-60.
20. Popov ME, Kashparov IV, Rumsh LD, Popov EM, *[Mechanism of action of aspartic proteases. III. Conformational characteristics of HIV-1 protease inhibitor JG-365].* Bioorganicheskaia khimiia, 1999. **25**(6): p. 418-22.
21. Kim, D.H. and S. Mobashery, *Mechanism-based inhibition of zinc proteases.* Current medicinal chemistry, 2001. **8**(8): p. 959-65.
22. Kraut, J., *Serine proteases: structure and mechanism of catalysis.* Annual review of biochemistry, 1977. **46**: p. 331-58.
23. Lemberg, M.K., *Intramembrane proteolysis in regulated protein trafficking.* Traffic, 2011. **12**(9): p. 1109-18.
24. Rawson RB, Zelenski NG, Nijhawan D, Ye J, Sakai J, Hasan MT, Chang TY., Brown MS, Goldstein JL, *Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs.* Molecular cell, 1997. **1**(1): p. 47-57.
25. Duncan EA, Davé UP, Sakai J, Goldstein JL, Brown MS., *Second-site cleavage in sterol regulatory element-binding protein occurs at transmembrane junction as determined by cysteine panning.* The Journal of biological chemistry, 1998. **273**(28): p. 17801-9.
26. Ye J, Davé UP, Grishin NV, Goldstein JL, Brown MS., *Asparagine-proline sequence within membrane-spanning segment of SREBP triggers intramembrane cleavage by site-2 protease.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(10): p. 5123-8.
27. Kopan, R. and M.X. Ilagan, *Gamma-secretase: proteasome of the membrane?* Nature reviews. Molecular cell biology, 2004. **5**(6): p. 499-504.
28. Haass, C. and H. Steiner, *Alzheimer disease gamma-secretase: a complex story of GxGD-type presenilin proteases.* Trends in cell biology, 2002. **12**(12): p. 556-62.
29. Weihofen A, Binns K, Lemberg MK, Ashman K, Martoglio B., *Identification of signal peptide peptidase, a presenilin-type aspartic protease.* Science, 2002. **296**(5576): p. 2215-8.
30. Sato C, Takagi S, Tomita T, Iwatsubo T., *The C-terminal PAL motif presenilin 1 are involved in the formation of the catalytic pore of the*

gamma-secretase. The Journal of neuroscience : the official journal of the Society for Neuroscience, 2008. **28**(24): p. 6264-71.
31. Wang J, Beher D, Nyborg AC, Shearman MS, Golde TE, Goate A., *C-terminal PAL motif of presenilin and presenilin homologues required for normal active site conformation.* Journal of neurochemistry, 2006. **96**(1): p. 218-27.
32. Edbauer D, Winkler E, Regula JT, Pesold B, Steiner H, Haass C., *Reconstitution of gamma-secretase activity.* Nature cell biology, 2003. **5**(5): p. 486-8.
33. De Strooper, B., *Aph-1, Pen-2, and Nicastrin with Presenilin generate an active gamma-Secretase complex.* Neuron, 2003. **38**(1): p. 9-12.
34. Friedmann E, Lemberg MK, Weihofen A, Dev KK, Dengler U, Rovelli G, Martoglio B., *Consensus analysis of signal peptide peptidase and homologous human aspartic proteases reveals opposite topology of catalytic domains compared with presenilins.* The Journal of biological chemistry, 2004. **279**(49): p. 50790-8.
35. Xia, W. and M.S. Wolfe, *Intramembrane proteolysis by presenilin and presenilin-like proteases.* Journal of cell science, 2003. **116**(Pt 14): p. 2839-44.
36. Wolfe MS, Xia W, Moore CL, Leatherwood DD, Ostaszewski B, Rahmati T, Donkor IO, Selkoe DJ., *Peptidomimetic probes and molecular modeling suggest that Alzheimer's gamma-secretase is an intramembrane-cleaving aspartyl protease.* Biochemistry, 1999. **38**(15): p. 4720-7.
37. De Strooper B, Saftig P, Craessaerts K, Vanderstichele H, Guhde G, Annaert W, Von Figura K, Van Leuven F., *Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein.* Nature, 1998. **391**(6665): p. 387-90.
38. Mattson, M.P., *Neurobiology: Ballads of a protein quartet.* Nature, 2003. **422**(6930): p. 385, 387.
39. Urban, S., J.R. Lee, and M. Freeman, *Drosophila rhomboid-1 defines a family of putative intramembrane serine proteases.* Cell, 2001. **107**(2): p. 173-82.
40. Wang, Y., Y. Zhang, and Y. Ha, *Crystal structure of a rhomboid family intramembrane protease.* Nature, 2006. **444**(7116): p. 179-80.
41. Maegawa, S., K. Ito, and Y. Akiyama., *Proteolytic action of GlpG, a rhomboid protease in the Escherichia coli cytoplasmic membrane.* Biochemistry, 2005. **44**(41): p. 13543-52.
42. E. Wieschaus, C.N.-V.a.G.J., *Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster.* DEVELOPMENT GENES AND EVOLUTION, 1984. **193**(5): p. 296-307.
43. Wasserman, J.D., 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 & development, 2000. **14**(13): p. 1651-63.

44. Pascall, J.C. and K.D. Brown, *Characterization of a mammalian cDNA encoding a protein with high sequence similarity to the Drosophila regulatory protein Rhomboid.* FEBS letters, 1998. **429**(3): p. 337-40.
45. Koonin EV, Makarova KS, Rogozin IB, Davidovic L, Letellier MC, Pellegrini L., *The rhomboids: a nearly ubiquitous family of intramembrane serine proteases that probably evolved by multiple ancient horizontal gene transfers.* Genome Biology, 2003. **4**(3): p. R19.
46. Lemberg, M.K. and M. Freeman, *Functional and evolutionary implications of enhanced genomic analysis of rhomboid intramembrane proteases.* Genome Research, 2007. **17**(11): p. 1634-46.
47. Mayer U, N.-V.C., *A group of genes required for pattern formation in the ventral ectoderm of the Drosophila embryo.* Genes & development, 1988. **2**(11): p. 1496-1511.
48. Guichard A, B.B., Sturtevant MA, Wickline L, Chacko J, Howard K, Bier E., *rhomboid and Star interact synergistically to promote EGFR/MAPK signaling during Drosophila wing vein development.* Development, 1999. . **126**(12): p. 2663-2776.
49. Lee JR, Urban S, Garvey CF, Freeman M., *Regulated intracellular ligand transport and proteolysis control EGF signal activation in Drosophila.* Cell, 2001. **107**(2): p. 161-71.
50. Bier, E., L.Y. Jan, and Y.N. Jan, *rhomboid, a gene required for dorsoventral axis establishment and peripheral nervous system development in Drosophila melanogaster.* Genes & Development, 1990. **4**(2): p. 190-203.
51. Urban, S. and M.S. Wolfe, *Reconstitution of intramembrane proteolysis in vitro reveals that pure rhomboid is sufficient for catalysis and specificity.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(6): p. 1883-8.
52. Lemberg MK, Menendez J, Misik A, Garcia M, Koth CM, Freeman M., *Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases.* The EMBO journal, 2005. **24**(3): p. 464-72.
53. Urban, S. and R.P. Baker, *In vivo analysis reveals substrate-gating mutants of a rhomboid intramembrane protease display increased activity in living cells.* Biological Chemistry, 2008. **389**(8): p.1107-15
54. Van den Berg B, Clemons WM Jr, Collinson I, Modis Y, Hartmann E, Harrison SC, Rapoport TA., *X-ray structure of a protein-conducting channel.* Nature, 2004. **427**(6969): p. 36-44.
55. Esser K, Tursun B, Ingenhoven M, Michaelis G, Pratje E., *A novel two-step mechanism for removal of a mitochondrial signal sequence involves the mAAA complex and the putative rhomboid protease Pcp1.* Journal of molecular biology, 2002. **323**(5): p. 835-43.
56. Herlan M, Vogel F, Bornhovd C, Neupert W, Reichert AS., *Processing of Mgm1 by the rhomboid-type protease Pcp1 is required for maintenance of mitochondrial morphology and of mitochondrial DNA.* The Journal of biological chemistry, 2003. **278**(30): p. 27781-8.

57. McQuibban, G.A., S. Saurya, and M. Freeman, *Mitochondrial membrane remodelling regulated by a conserved rhomboid protease.* Nature, 2003. **423**(6939): p. 537-41.
58. Schatz, G. and B. Dobberstein, *Common principles of protein translocation across membranes.* Science, 1996. **271**(5255): p. 1519-26.
59. Freeman, M., *Proteolysis within the membrane: rhomboids revealed.* Nature Reviews. Molecular Cell Biology, 2004. **5**(3): p. 188-97.
60. Nakagawa T, Guichard A, Castro CP, Xiao Y, Rizen M, Zhang HZ, Hu D, Bang A, Helms J, Bier E, Derynck R., *Characterization of a human rhomboid homolog, p100hRho/RHBDF1, interacts with TGF-alpha family ligands.* Developmental dynamics, 2005. **233**(4): p. 1315-31.
61. Sherratt AR, Braganza MV, Nguyen E, Ducat T, Goto NK., *Insights into the effect of detergents on the full-length rhomboid protease from Pseudomonas aeruginosa and its cytosolic domain.* Biochimica et biophysica acta, 2009. **1788**(11): p. 2444-53.
62. Lohi, O., S. Urban, and M. Freeman, *Diverse substrate recognition mechanisms for rhomboids; thrombomodulin is cleaved by Mammalian rhomboids.* Current Biology, 2004. **14**(3): p. 236-41.
63. Wang, Y. and Y. Ha, *Open-cap conformation of intramembrane protease GlpG.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(7): p. 2098-102.
64. Wu Z, Yan N, Feng L, Oberstein A, Yan H, Baker RP, Gu L, Jeffrey PD, Urban S, Shi Y., *Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry.* Nature Structure and Molecular Biology, 2006. **13**(12): p. 1084-91.
65. Ben-Shem, A., D. Fass, and E. Bibi, *Structural basis for intramembrane proteolysis by rhomboid serine proteases.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(2): p. 462-6.
66. Lemieux MJ, Fischer SJ, Cherney MM, Bateman KS, James MN., *The crystal structure of the rhomboid peptidase from Haemophilus influenzae provides insight into intramembrane proteolysis.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(3): p. 750-4.
67. Urban, S., D. Schlieper, and M. Freeman, *Conservation of intramembrane proteolytic activity and substrate specificity in prokaryotic and eukaryotic rhomboids.* Current Biology, 2002. **12**(17): p. 1507-12.
68. Clemmer KM, Sturgill GM, Veenstra A, Rather PN., *Functional characterization of Escherichia coli GlpG and additional rhomboid proteins using an aarA mutant of Providencia stuartii.* Journal of Bacteriology, 2006. **188**(9): p. 3415-9.
69. Baker RP, Young K, Feng L, Shi Y, Urban S., *Enzymatic analysis of a rhomboid intramembrane protease implicates transmembrane helix 5*

as the lateral substrate gate. Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(20): p. 8257-62.
70. Bondar, A.N., C. del Val, and S.H. White, *Rhomboid protease dynamics and lipid interactions.* Structure, 2009. **17**(3): p. 395-405.
71. Maegawa S, Koide K, Ito K, Akiyama Y., *The intramembrane active site of GlpG, an E. coli rhomboid protease, is accessible to water.* Molecular Microbiology, 2007. **64**(2): p. 435-47.
72. Brooks CL, Lazareno-Saez C, Lamoureux JS, Mak MW, Lemieux MJ. *Insights into substrate gating in H. influenzae rhomboid.* Journal of molecular biology, 2011. **407**(5): p. 687-697.
73. Lemberg, M.K. and M. Freeman, *Cutting proteins within lipid bilayers: rhomboid structure and mechanism.* Molecular Cell, 2007. **28**(6): p. 930-40.
74. Knopf, R.R. and Z. Adam, *Rhomboid proteases in plants - still in square one?* Physiologia plantarum, 2012. **145**(1): p. 41-51.
75. Vinothkumar, K.R., et al., *The structural basis for catalysis and substrate specificity of a rhomboid protease.* The EMBO Journal, 2010.
76. Xue, Y., et al., *Conformational Change in Rhomboid Protease GlpG Induced by Inhibitor Binding to Its S′ Subsites.* Biochemistry, 2012. **51**(18): p. 3723-31.
77. Adrain, C., et al., *Mammalian EGF receptor activation by the rhomboid protease RHBDL2.* EMBO reports, 2011. **12**(5): p. 421-7.
78. Reddy, T. and J.K. Rainey, *Multifaceted substrate capture scheme of a rhomboid protease.* The journal of physical chemistry. B, 2012. **116**(30): p. 8942-54.
79. Lazareno-Saez, C., C. Brooks, and M. Lemieux, *Structural Comparison of Substrate Entry Gate for Rhomboid Intramembrane Peptidases.* Biochemistry and Cell Biology, 2011: **89**(2): p. 216-23.
80. M Santos J, G.A., Soldati-Favre D., *New insights into parasite rhomboid proteases.* Molecular and biochemical parasitology, 2012. **182**(1): p. 27-36.
81. Zhou Y, Moin SM, Urban S, Zhang Y., *An Internal Water-Retention Site in the Rhomboid Intramembrane Protease GlpG Ensures Catalytic Efficiency.* Structure, 2012. **20**(7): p. 1255-63.
82. Urban, S. and M. Freeman, *Substrate specificity of rhomboid intramembrane proteases is governed by helix-breaking residues in the substrate transmembrane domain.* Molecular Cell, 2003. **11**(6): p. 1425-34.
83. Gallio M, Sturgill G, Rather P, Kylsten P., *A conserved mechanism for extracellular signaling in eukaryotes and prokaryotes.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(19): p. 12208-13.
84. Brossier F, Jewett TJ, Sibley LD, Urban S., *A spatially localized rhomboid protease cleaves cell surface adhesins essential for invasion by Toxoplasma.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(11): p. 4146-51.

85. Baker, R.P., R. Wijetilaka, and S. Urban, *Two Plasmodium rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria.* PLoS pathogens, 2006. **2**(10): p. e113.
86. Akiyama, Y. and S. Maegawa, *Sequence features of substrates required for cleavage by GlpG, an Escherichia coli rhomboid protease.* Molecular Microbiology, 2007. **64**(4): p. 1028-37.
87. Schäfer A, Zick M, Kief J, Steger M, Heide H, Duvezin-Caubet S, Neupert W, Reichert AS., *Intramembrane proteolysis of Mgm1 by the mitochondrial rhomboid protease is highly promiscuous.* Journal of molecular biology, 2010. **401**(2): p. 182-93.
88. Tatsuta T, Augustin S, Nolden M, Friedrichs B, Langer T., *m-AAA protease-driven membrane dislocation allows intramembrane cleavage by rhomboid in mitochondria.* The EMBO journal, 2007. **26**(2): p. 325-35.
89. O'Donnell RA, Hackett F, Howell SA, Treeck M, Struck N, Krnajski Z, Withers-Martinez C, Gilberger TW, Blackman MJ., *Intramembrane proteolysis mediates shedding of a key adhesin during erythrocyte invasion by the malaria parasite.* The Journal of cell biology, 2006. **174**(7): p. 1023-33.
90. Chao JR, Parganas E, Boyd K, Hong CY, Opferman JT, Ihle JN., *Hax1-mediated processing of HtrA2 by Parl allows survival of lymphocytes and neurons.* Nature, 2008. **452**(7183): p. 98-102.
91. Amarneh, B. and R.B. Rawson, *Rhomboid proteases: familiar features in unfamiliar phases.* Molecular cell, 2009. **36**(6): p. 922-3.
92. Freeman, M., *Rhomboids: 7 years of a new protease family.* Seminars in cell & developmental biology, 2008.
93. Rawson, R.B., *Intriguing parasites and intramembrane proteases.* Genes & development, 2008. **22**(12): p. 1561-6.
94. Hill, R.B. and L. Pellegrini, *The PARL family of mitochondrial rhomboid proteases.* Seminars in cell & developmental biology, 2010. **21**(6): p. 582-92.
95. Lieberman, R.L. and M.S. Wolfe, *From rhomboid function to structure and back again.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(20): p. 8199-200.
96. Herlan, M., et al., *Alternative topogenesis of Mgm1 and mitochondrial morphology depend on ATP and a functional import motor.* The Journal of Cell Biology, 2004. **165**(2): p. 167-73.
97. Santos JM, Ferguson DJ, Blackman MJ, Soldati-Favre D., *Intramembrane cleavage of AMA1 triggers Toxoplasma to switch from an invasive to a replicative mode.* Science, 2011. **331**(6016): p. 473-7.
98. Karakasis K, T.D., Ko K, *Uncovering a link between a plastid translocon component and rhomboid proteases using yeast mitochondria-based assays.* Plant and Cell physiology, 2007. **48**(4): p. 655-661.
99. Jeyaraju DV, Sood A, Laforce-Lavoie A, Pellegrini L., *Rhomboid proteases in mitochondria and plastids: Keeping organelles in shape.* Biochimica et biophysica acta, 2012.

100. Stevenson LG, Strisovsky K, Clemmer KM, Bhatt S, Freeman M, Rather PN., *Rhomboid protease AarA mediates quorum-sensing in Providencia stuartii.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(3): p. 1003-8.
101. Perrimon N, P.L., *There must be 50 ways to rule the signal: the case of the Drosophila EGF receptor.* Cell, 1997 **89**(1): p. 13-16.
102. Wasserman, J.D. and M. Freeman, *An autoregulatory cascade of EGF receptor signaling patterns the Drosophila egg.* Cell, 1998. **95**(3): p. 355-64.
103. Schweitzer, R. and B.Z. Shilo, *A thousand and one roles for the Drosophila EGF receptor.* Trends in genetics : TIG, 1997. **13**(5): p. 191-6.
104. Tsruya R, Wojtalla A, Carmon S, Yogev S, Reich A, Bibi E, Merdes G, Schejter E, Shilo BZ., *Rhomboid cleaves Star to regulate the levels of secreted Spitz.* The EMBO journal, 2007. **26**(5): p. 1211-20.
105. Dutt, A., et al., *EGF signal propagation during C. elegans vulval development mediated by ROM-1 rhomboid.* PLoS Biology, 2004. **2**(11): p. e334.
106. Wang Y, Guan X, Fok KL, Li S, Zhang X, Miao S, Zong S, Koide SS, Chan HC, Wang L., *A novel member of the Rhomboid family, RHBDD1, regulates BIK-mediated apoptosis.* Cellular and molecular life sciences : CMLS, 2008. **65**(23): p. 3822-9.
107. Rather PN, Ding X, Baca-DeLancey RR, Siddiqui S., *Providencia stuartii genes activated by cell-to-cell signaling and identification of a gene required for production or activity of an extracellular factor.* Journal of bacteriology, 1999. **181**(23): p. 7185-91.
108. Fritsch MJ, Krehenbrink M, Tarry MJ, Berks BC, Palmer T., *Processing by rhomboid protease is required for Providencia stuartii TatA to interact with TatC and to form functional homo-oligomeric complexes.* Molecular microbiology, 2012. **84**(6): p. 1108-23.
109. Urban, S., *Making the cut: central roles of intramembrane proteolysis in pathogenic microorganisms.* Nature reviews. Microbiology, 2009. **7**(6): p. 411-23.
110. Opitz C, Di Cristina M, Reiss M, Ruppert T, Crisanti A, Soldati D., *Intramembrane cleavage of microneme proteins at the surface of the apicomplexan parasite Toxoplasma gondii.* The EMBO journal, 2002. **21**(7): p. 1577-85.
111. Dowse TJ, Pascall JC, Brown KD, Soldati D., *Apicomplexan rhomboids have a potential role in microneme protein cleavage during host cell invasion.* International journal for parasitology, 2005. **35**(7): p. 747-56.
112. Buguliskis JS, Brossier F, Shuman J, Sibley LD., *Rhomboid 4 (ROM4) affects the processing of surface adhesins and facilitates host cell invasion by Toxoplasma gondii.* PLoS pathogens, 2010. **6**(4): p. e1000858.

113. Tyler JS, T.M., Boothroyd JC., *Focus on the ringleader: the role of AMA1 in apicomplexan invasion and replication.* Trends in Parasitology, 2011. **27**(9): p. 410-420.
114. Baxt LA, Rastew E, Bracha R, Mirelman D, Singh U., *Downregulation of an Entamoeba histolytica rhomboid protease reveals roles in regulating parasite adhesion and phagocytosis.* Eukaryotic cell, 2010. **9**(8): p. 1283-93.
115. Espinosa-Cantellano, M. and A. Martinez-Palomo, *Entamoeba histolytica: mechanism of surface receptor capping.* Experimental parasitology, 1994. **79**(3): p. 424-35.
116. Tavares, P., P. Sansonetti, and N. Guillen, *The interplay between receptor capping and cytoskeleton remodeling in Entamoeba histolytica.* Archives of medical research, 2000. **31**(4): p. S140-2.
117. Baxt LA, Baker RP, Singh U, Urban S., *An Entamoeba histolytica rhomboid protease with atypical specificity cleaves a surface lectin involved in phagocytosis and immune evasion.* Genes & development, 2008. **22**(12): p. 1636-46.
118. Shepard, K.A. and M.P. Yaffe, *The yeast dynamin-like protein, Mgm1p, functions on the mitochondrial outer membrane to mediate mitochondrial inheritance.* The Journal of cell biology, 1999. **144**(4): p. 711-20.
119. McQuibban GA, Lee JR, Zheng L, Juusola M, Freeman M., *Normal mitochondrial dynamics requires rhomboid-7 and affects Drosophila lifespan and neuronal function.* Curr Biol, 2006. **16**(10): p. 982-9.
120. Cipolat S, Rudka T, Hartmann D, Costa V, Serneels L, Craessaerts K, Metzger K, Frezza C, Annaert W, D'Adamio L, Derks C, Dejaegere T, Pellegrini L, D'Hooge R, Scorrano L, De Strooper B., *Mitochondrial rhomboid PARL regulates cytochrome c release during apoptosis via OPA1-dependent cristae remodeling.* Cell, 2006. **126**(1): p. 163-75.
121. Whitworth AJ, Lee JR, Ho VM, Flick R, Chowdhury R, McQuibban GA., *Rhomboid-7 and HtrA2/Omi act in a common pathway with the Parkinson's disease factors Pink1 and Parkin.* Dis Model Mech, 2008. **1**(2-3): p. 168-74; discussion 173.
122. Meissner C, Lorenz H, Weihofen A, Selkoe DJ, Lemberg MK., *The mitochondrial intramembrane protease PARL cleaves human Pink1 to regulate Pink1 trafficking.* Journal of neurochemistry, 2011. **117**(5): p. 856-67.
123. Gegg ME, Cooper JM, Chau KY, Rojo M, Schapira AH, Taanman JW., *Mitofusin 1 and mitofusin 2 are ubiquitinated in a PINK1/parkin-dependent manner upon induction of mitophagy.* Human molecular genetics, 2010. **19**(24): p. 4861-70.
124. Chan NC, Salazar AM, Pham AH, Sweredoski MJ, Kolawa NJ, Graham RL, Hess S, Chan DC., *Broad activation of the ubiquitin-proteasome system by Parkin is critical for mitophagy.* Human molecular genetics, 2011. **20**(9): p. 1726-37.

125. Civitarese AE, MacLean PS, Carling S, Kerr-Bayles L, McMillan RP, Pierce A, Becker TC, Moro C, Finlayson J, Lefort N, Newgard CB, Mandarino L, Cefalu W, Walder K, Collier GR, Hulver MW, Smith SR, Ravussin E., *Regulation of skeletal muscle oxidative capacity and insulin signaling by the mitochondrial rhomboid protease PARL.* Cell metabolism, 2010. **11**(5): p. 412-26.
126. Foltenyi, K., R.J. Greenspan, and J.W. Newport, *Activation of EGFR and ERK by rhomboid signaling regulates the consolidation and maintenance of sleep in Drosophila.* Nature neuroscience, 2007. **10**(9): p. 1160-7.
127. Zettl M, Adrain C, Strisovsky K, Lastun V, Freeman M., *Rhomboid family pseudoproteases use the ER quality control machinery to regulate intercellular signaling.* Cell, 2011. **145**(1): p. 79-91.
128. Smith, M.H., H.L. Ploegh, and J.S. Weissman, *Road to ruin: targeting proteins for degradation in the endoplasmic reticulum.* Science, 2011. **334**(6059): p. 1086-90.
129. Adrain C, Zettl M, Christova Y, Taylor N, Freeman M., *Tumor necrosis factor signaling requires iRhom2 to promote trafficking and activation of TACE.* Science, 2012. **335**(6065): p. 225-8.
130. Siggs OM, X.N., Wang Y, Shi H, Tomisato W, Li X, Xia Y, Beutler B., *iRhom2 is required for the secretion of mouse TNFα.* Blood, 2012. **119**(24): p. 5769-5771.
131. Pascall, J.C., J.E. Luck, and K.D. Brown, *Expression in mammalian cell cultures reveals interdependent, but distinct, functions for Star and Rhomboid proteins in the processing of the Drosophila transforming-growth-factor-alpha homologue Spitz.* The Biochemical journal, 2002. **363**(Pt 2): p. 347-52.
132. Yogev S, S.E., Shilo BZ, *Drosophila EGFR signalling is modulated by differential compartmentalization of Rhomboid intramembrane proteases.* The EMBO journal, 2008. **27**: p. 1219-1230.
133. Sheiner, L., T.J. Dowse, and D. Soldati-Favre, *Identification of trafficking determinants for polytopic rhomboid proteases in Toxoplasma gondii.* Traffic, 2008. **9**(5): p. 665-77.
134. Sík A, Passer BJ, Koonin EV, Pellegrini L., *Self-regulated cleavage of the mitochondrial intramembrane-cleaving protease PARL yields Pbeta, a nuclear-targeted peptide.* The Journal of biological chemistry, 2004. **279**(15): p. 15323-9.
135. Jeyaraju DV, McBride HM, Hill RB, Pellegrini L., *Structural and mechanistic basis of Parl activity and regulation.* Cell death and differentiation, 2011. **18**(9): p. 1531-9.
136. Rimon, A., T. Tzubery, and E. Padan, *Monomers of the NhaA Na+/H+ antiporter of Escherichia coli are fully functional yet dimers are beneficial under extreme stress conditions at alkaline pH in the presence of Na+ or Li+.* The Journal of biological chemistry, 2007. **282**(37): p. 26810-21.

Chapter 2: The study of oligomeric state of prokaryotic rhomboid proteases

Chapter 2: The study of oligomeric state of prokaryotic rhomboid proteases

***A version of this chapter was published in:**

Padmapriya Sampathkumar, Michelle W. Mak, Sarah J. Fischer-Witholt, Emmanuel Guigard, Cyril M. Kay, M. Joanne Lemieux. 2012 *Biochim Biophys Acta* **1818:** 3090-3097

Contributions:

Dr. Joanne Lemieux and Sarah Withold (University of Alberta) purified the three prokaryotic rhomboid proteases, *hi*GlpG, *ec*GlpG and YqgP for analytical ultracentrifugation experiments. All results obtained from the analytical ultracentrifugation experiments were performed by Emmanuel Guigard from Dr. Cyril Kay's laboratory (University of Alberta). Michelle W. Mak aided in the functional assay of *hi*GlpG.

Chapter 2: Introduction

Rhomboid proteases have diverse substrate specificity; therefore different control strategies must exist to regulate the activity of these enzymes. Some common mechanisms by which rhomboids achieve regulation (as described in section 1.9) are: compartmentalization, as observed in *Drosophila* EGFR signaling where membrane trafficking is used as a mechanism to regulate the contact between enzyme and substrate [1, 2] and proteolytic activation of PARL in mammalian mitochondria [3]. In addition, regulation can also be achieved through changes in the lipid environment around the enzyme [4]. In a study performed by Urban and Wolfe, bacterial rhomboids were assayed for activity by reconstituting them in various lipid vesicles. Different membrane lipids caused different effects on rhomboid activity. While the lipids from brain stimulated *Drosophila* Rhomboid-4, the same lipid caused an inhibitory effect on *ec*GlpG. In another example, *P. stuartii* rhomboid AarA was stimulated by cardiolipin, where all the other rhomboids showed decrease in activity in the presence of this lipid. This suggests that the composition of lipids in the membrane environment around rhomboid protease influence the activity of the enzyme.

Another plausible mode of regulation can be achieved through oligomerization. Many membrane proteins are known to exist in a dynamic equilibrium between different oligomeric states to regulate their function [5, 6]. In this chapter, we access the oligomeric state of prokaryotic rhomboid

proteases from three different organisms, YqgP from *Bacillus subtilis*, *ec*GlpG from *E. coli* and *hi*GlpG from *Haemophilus influenzae*. During initial attempts of purification for crystallization studies of *H. influenzae* rhomboid, *hi*GlpG, was solved, gel filtration analysis suggested rhomboids may form oligomers [7]. In this current study, we assess the oligomeric state of the three prokaryotic rhomboid proteases by various biochemical approaches. Initial experiments using sedimentation equilibrium analysis revealed that the predominant species for all the three rhomboid proteases was a dimer. In order to determine if the membrane domain was responsible for the dimerization, we carried out further experiments with *hi*GlpG; *hi*GlpG being the simplest form of rhomboid representing the membrane domain alone with six transmembrane segments. With gel filtration and crosslinking, we confirmed that *hi*GlpG was dimeric and further examined its *in vitro* activity. In addition, a co-immunoprecipitation assay suggested that the dimeric *hi*GlpG is present within the lipid bilayer. In summary, this chapter provides evidence for the first detailed characterization of oligomeric state of rhomboid proteases. These findings may have implications in understanding the regulation of rhomboid activity and its mechanism of action.

2.1 Materials

Reagents:

Standard lab reagents were purchased from Sigma-Aldrich (ON, Canada), Fisher Scientific (ON, Canada). Other reagents are listed below.

Kits:

- QuikChange Lightining Mutagenesis Kit (Stratagene)
- QIAprep Spin Miniprep Kit (Qiagen)
- Bicinchoninic Acid (BCA) Protein Assay Kit (Thermo Scientific Pierce)

Enzymes:

- Restriction enzymes with appropriate buffers were purchased from Fermentas
- T4 DNA ligase was purchased from Invitrogen

Antibodies:

- His- Probe HRP (1:500 dilution) Santa Cruz Biotechnology, USA
- Mouse anti-flag antibody (1:10,000 dilution) Sigma-Aldrich, USA
- Secondary Rabbit anti-mouse antibody conjugated to Horseradish peroxidase (1:40000 dilution) Sigma-Aldrich, USA

Primers:

Synthesized by Integraed DNA Technolgies, USA

Culture media:

- LB (Luria Bertani) liquid media from Fisher Scientific
- LB Agar: LB liquid with 1.5% Agar

Solutions:

Phosphate buffer solution (PBS) pH 7.4	137mM NaCl, 10mM Phosphate buffer, 2.7mM KCl, pH 7.4
Tris buffered saline pH 7.5	50mM Tris, 150mM NaCl
TBST	50mM Tris, 150mM NaCl, 0.05% Tween 20
SDS-PAGE gel-loading buffer	50mM Tris-Cl pH 6.8, 100mM Dithiothreitol, 2% SDS, 0.1% Bromophenol blue, 10% glycerol
Laemmli buffer	50mM Tris, 400mM glycine, and 0.1% (w/v) SDS
Transfer buffer	2.5mM Tris, 19.2mM Glycine, 20% Methanol, 0.1%SDS
Agarose gel-loading (6X)	0.02% Bromophenol Blue, 0.02% Xylenecyanol, 30% Glycerol in H_2O
Tris/Acetate/EDTA (TAE) pH 8	0.04M Tris-Acetate, 0.001M EDTA
Tris EDTA (TE) (10X) pH 8	100mM Tris-HCl pH 7.5, 10mM EDTA

Bacterial strains:

Strains	**Genotype**
Top 10	F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG
BL21 (DE3)	F– ompT hsdSB(rB–, mB–) gal dcm (DE3)
XL 10	TetrD(mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F´ proAB lacIqZDM15 Tn10

2.2 Experimental Procedures

2.2.1 Preparation of plasmids

*Rhomboid expression plasmids: ec*GlpG was generated using PCR with an *Escherichia coli* strain DH5α genomic DNA. YqgP and *hi*GlpG were generated by PCR using genomic *Bacillus subtilis* and *Haemophilus influenzae* DNA, respectively, purchased from ATCC, USA. Using restriction digestion, the three PCR products were then ligated into pBAD.MycHisA vector (Invitrogen, Canada) to generate three prokaryotic rhomboid expression plasmids, pBAD.*ec*GlpG, pBAD.*hi*GlpG and pBAD.YqgP respectively (Figure. 2.2.1). Further experiments in this section involving *hi*GlpG were carried out with pBAD.*hi*GlpG as the template.

psTatA expression plasmid: psTatA was supplied as a gift from Dr. Matthew Freeman, MRC, Cambridge) in a pET21a vector carrying His epitope (Figure 2.2.2). Oligonucleotide primers were designed to insert a FLAG tag at the N terminus of psTatA using QuikChange Lightning Site Directed Mutagenesis Kit (Agilent Technologies, Ontario, Canada). The PCR condition was followed as per the protocol supplied by Agilent Technologies (Table 2.2.1). After the PCR, the parental dsDNA was digested with DpnI restriction enzyme. 2µl of the digested product was then added to XL10 Gold Ultracompetent cells and incubated in ice for 30 min. The cells were switched to 42°C for 30 sec for heat-shock induced transformation and immediately transferred to ice for 2 min. NZY broth was added for cells to recover at 37°C for 1hr. The cells were

then plated on 100µg/ml ampicillin, 40 µg/ml X-gal and 0.1mM IPTG containing LB agar plates and allowed to grow overnight at 37°C.

Plasmid Isolation and DNA sequencing: Single colonies were picked from the plates and grown in 5ml LB media supplemented with ampicillin at 37°C. The plasmids were then isolated using Qiagen mini prep kit (Qiagen, Toronto, Ontario). A mini-plasmid isolation protocol supplied with the kit was used for the extraction of plasmids. The plasmids were eluted in a final volume 50µl of ddH_2O. Purity of the plasmids was checked on agarose gel and concentration of DNA was determined by absorption readings (OD_{260}). The plasmid samples were sent for sequencing at The Applied Genomics Centre, Department of Medicine, University of Alberta.

2.2.2 Rhomboid expression and purification

Membrane fraction isolation: E. coli (Top10) expressing *ec*GlpG, *hi*GlpG and YqgP were grown in 5ml LB media supplemented with ampicillin (100µg/ml) overnight at 37°C. Large scale cultures (1L) were inoculated with 20ml of overnight cultures. The three cultures were grown to an OD_{600} = 1.0, 0.4 and 1.0, respectively and induced with 0.002 % arabinose at 24°C for 5 h. Cells were harvested at 7,000 rpm for 10 min using an Avanti J1.8000 rotor (Beckman, USA). Cells were resuspended in 4 volumes TBS supplemented with EDTA-free peptidase inhibitor cocktail (NEB, USA), 1 mM PMSF, 0.1 mg/ml DNase, and lysed using an EmulsiFlex-C3 (Avestin Inc, Ottawa,

Canada). Unbroken cells were pelleted in a JA17 rotor (Beckman, USA) at 15000 rpm (10,000 *g*) for 20 min. Membrane fractions were collected by ultracentrifugation in a L8-80 ultracentrifuge at 35,000 rpm (100,000*g*) in a 45Ti rotor (Beckman, USA) for 2 h.

Protein purification: Membrane fractions were homogenized in 50mM Tris, 300mM NaCl, 30mM Imidazole, 20 % glycerol, 1 % DDM pH 8.0. The solution was stirred for 30 min followed by ultracentrifugation for 30 min at 45,000 rpm (110,000 *g*) in a Ti45 rotor (Beckman, USA). The supernatant was incubated with Ni-NTA resin (Qiagen, Ontario, Canada) for 2 h. The resin was then collected and washed with 20 column volumes (CV) of 50mM Tris, 300 mM NaCl, 30mM imidazole, 20 % glycerol, 0.1 % DDM pH 8.0 followed by 20 CV of the above stated buffer with 35mM imidazole. Protein fractions were eluted in a step-wise manner with 3 times of 2 CV of the above described buffer containing 250, 500 and 1000mM Imidazole. Bicinchoninic Acid (BCA) kit (Pierce Biotechnology Inc., Thermo Scientific, Il, USA) was used to determine protein concentrations.

2.2.3 Expression and purification of *Providencia stuartii* TatA (psTatA)

Flag-psTatA-His was purified as previously described for the C100-TatA construct [4]. Briefly, the psTatA expression plasmid was transformed into BL21(DE3) cells. Cells were grown in 5ml LB media supplemented with

ampicillin overnight at 37°C. Large scale cultures (1L) were inoculated with 20ml of overnight cultures. The cultures were grown to an OD_{600} of 1.0 and induced with 1mM IPTG for 3 hours. Cells were harvested at 7000 rpm for 10 min using an Avanti J1.8000 rotor (Beckman, USA) and resuspended in 1X PBS (pH 7.4) plus 1% Triton X-100 buffer. The insoluble fraction was removed by pelleting the cells in a JA17 rotor at 15000 rpm (10,000 *g*) for 20 min. The supernatant was incubated with 1ml of ANTI-FLAG M2 Affinity Gel (Sigma-Aldrich, ON, Canada) with gentle agitation for 2 hours to facilitate protein binding. The resin was loaded onto a column and the flow-through was collected. The resin was subsequently washed with 10 CV of 1X PBS and 1% Triton X 100. Elutions of the protein were perfomed in step-wise manner by adding 3x 2CV of 0.1M glycine (pH 2.7) along with 0.1% DDM. The concentration of protein present in each eluted fraction was determined using Bicinchoninic Acid (BCA) kit (Pierce Biotechnology Inc., Thermo Scientific, Il, USA.

2.2.4 Analytical ultracentrifugation of *ec*GlpG, *hi*GlpG and YqgP

Sedimentation equilibrium experiments of the three rhomboid proteases were conducted at 20 °C in a Beckman XL-I analytical ultracentrifuge using absorbance optics, as described by Laue and Stafford [8]. Protein samples used for DDM runs were obtained after Ni-NTA. Prior to runs, samples were dialyzed for at least 48 h in 20mM Tris, 20mM NaCl pH 8.0 and 0.05 % DDM. Aliquots (110 µL, 1mg/ml) of the sample solution were

loaded into six sector CFE sample cells, allowing three concentrations to be run simultaneously. Runs were performed at a minimum of three different speeds and each speed was maintained until there was no significant difference in $r^2/2$ versus absorbance scans taken 2 h apart to ensure that equilibrium had been achieved. Sedimentation equilibrium data were evaluated using the NONLIN program, which employs a nonlinear least squares curve-fitting algorithm described by Johnson *et al*. [9]. The program allows analysis of both single and multiple data files and can be fit to models containing up to four associating species, depending upon which parameters are permitted to vary during the fitting routine. The protein's partial specific volume and the solvent density were estimated using the Sednterp program [10]. DDM amounts were quantified using thin layer chromatography.

Thin layer chromatography detection of DDM: The *hi*GlpG sample was obtained post-Ni-NTA purification. 20 µl aliquot containing 9.2 µg of protein, along with DDM standards were spotted directly onto silica glass plates (Whatman, USA). The mobile solvent phase consisted of ethylacetate/methanol 4: 1 (v/v). The plate was sprayed with 2N H_2SO_4 solution and then charred at 90°C for DDM detection [11].

2.2.5 Crosslinking assay with detergent solubilized *hi*GlpG

*hi*GlpG was expressed and purified using a similar approach described in section 2.2.2. Crosslinking was carried out using two homobifunctional crosslinking reagents, DSP and DTSSP (Pierce Protein Research Products, Thermo Fisher Scientific, Rockford, USA) (Figure 2.2.3). Post purification, *hi*GlpG was dialysed with PBS (pH 7.4) to remove Tris buffer. For each crosslinker, three aliquots of 10µg (20µl) of dialysed *hi*GlpG were taken. One aliquot was treated as control. The remaining two aliquots were treated with 1mM DSP or 1mM DTSSP respectively and incubated at room temperature for 30 min with gentle mixing. The crosslinking reactions were quenched using 1M Tris (pH 7.4) to a final concentration of 50mM. To one of the crosslinked aliquots, 1M DTT was added to a final concentration of 50mM and incubated at 37°C for 30 min. 10µl of each of the samples were then loaded onto a 4%/ 12% SDS- PAGE gel under non- reducing conditions at 100V and then transferred to a Hybond- P PVDF membrane (GE Healthcare, USA) and run at a constant voltage of 100V for 1 h at 4°C..

2.2.6 Gel filtration chromatography

Purified *hi*GlpG was digested with thrombin (9 units of thrombin/ mg of *hi*GlpG) overnight at room temperature to remove the Myc.His epitope tag. Gel filtration chromatography was carried out on a Hiload Superdex 200 10/300 analytical column (GE Healthcare, USA). Approximately 200µg of sample was injected into the column and equilibrated with 50mM Tris,

200mM NaCl, 5% glycerol (pH- 8.0) and 0.1% DDM. Samples were run at 0.3 ml/ min and the absorbance values of the fractions were determined at 280nm. To determine the elution volume of *hi*GlpG (Ve), different absorbance values were plotted against volume of the fractions. A set of standard proteins were also run on the column for calibration: thyroglobulin (MW 670k; Stokes radius 85 Å), IgG (MW 158k; Stokes radius 55 Å), ovalbumin; (MW 44k; Stokes radius 28 Å), and myoglobin (MW 17k, Stokes radius 19 Å) (Bio-rad laboratories, USA). A standard curve was plotted using the molecular weights of standard proteins versus their Ve/V_0 values from which the mass of the eluted *hi*GlpG was calculated. (V_0: column void volume).

2.2.7 Activity assay

*hi*GlpG was purified with 0.1% DDM and digested with 30U of thrombin per mg of protein for 1h at room temperature and immediately used for activity assay. Increased amount of thrombin was used in this experiment as we found that the functional assay was more efficient with freshly purified *hi*GlpG. Activity assays were performed with 15µg *hi*GlpG, 500ng of psTatA substrate, DDM to a final concentration of 0.1% and cleavage buffer (50mM Tris, pH-7.5 and 150mM NaCl). The control contained 500ng substrate, DDM and cleavage buffer without rhomboid protease. The

reaction was carried out 37° C for 1h and then stopped by adding 7µl of 4X SDS sample buffer. 20µl of the samples was resolved by 4%/16% SDS

PAGE and transferred to a Hybond- P PVDF membrane (GE Healthcare, USA) and analysed using Western blot.

2.2.8 Anti-Flag pull down assay

Plasmid constructs: His-tagged and Flag-tagged *hi*GlpG were constructed individually by Quikchange mutagenesis (Stratagene, CA) by using pBAD.*hi*GlpG.MycHisA as the template (Table 2.2.2). Mutations were verified by DNA sequencing. The PCR products were cut by NcoI/HindIII to release *hi*GlpG.His and *hi*GlpG.Flag respectively and cloned into NcoI/HindIII digested pACYCDuet1 and pET28a (Novagen, USA) to generate pACYCDuet1-*hi*GlpG-His_6 and pET28a-*hi*GlpG-Flag respectively (Figure 2.2.4). *E. coli* BL21/DE3 cells were transformed with individual clones of pACYCDuet1-*hi*GlpG, pET28a-*hi*GlpG or both (co-transformation) using the standard transformation protocol [12].

Protein expression and purification: Large scale cultures of BL21/DE3 cells harbouring pACYCDuet1-*hi*GlpG-His, pET28a-*hi*GlpG-Flag or both were grown at 37°C in LB medium supplemented with appropriate antibiotics [ampicillin (100µg/ml) for pACYCDuet1- *hi*GlpG-His, kanamycin (30µg/ml) for pET28a- *hi*GlpG -Flag or both (for co-transformed plasmids)]. Cells were grown to an OD_{600} of 0.8 and induced with 0.1mM IPTG for 6 hours. Membrane fractions were collected by ultracentrifugation using the same protocol described in section 2.2.2.

Membrane pellets were then homogenized with 50mM Tris (pH 8.0), 300mM NaCl, 10mM imidazole, 20% glycerol and protease inhibitor tablets. Membrane homogenates of *hi*GlpG-His, *hi*GlpG-Flag, and coexpressed (*hi*GlpG-His and *hi*GlpG-Flag) were aliquoted (1ml) and solubilized with 1% DDM for 30 min at 4°C followed by ultracentrifugation at 40,000 rpm (100,000 g) for 30 min at 4°C in a TLS 55 rotor (Beckman, USA). For the mixed-membrane control, 0.5ml of *hi*GlpG-His membrane fraction was mixed with 0.5 ml of *hi*GlpG-Flag membrane fraction after homogenization on ice for 30 minutes and solubilized as above. Aliquots of supernatant of each of above fractions (1ml) were incubated with 50µl of pre-equilibrated Flag affinity resin (Anti-Flag M2 Affinity Gel; Sigma, USA) for 2 h at 4°C with agitation for protein binding. Resins were washed with 20 CV of TBS supplemented with 0.1% DDM. Immunocomplexes were finally dissolved in 4X SDS sample buffer and analyzed by Western blotting. To ensure equal amounts of the protein was added to the resin, a SDS-PAGE (4%/12%) gel was prepared simultaneously and stained with Coomassie Blue R250.

2.2.9 SDS- PAGE and Western Blotting

SDS-PAGE: All protein samples were resolved using polyacrylamide gels. Different composition of the resolving gel and the stacking gel are listed in Table 2.2.3. The resolving gel and stacking gel were prepared and allowed to polymerize. The gels were run at a constant voltage of 100V at room temperature in Laemmli buffer. Following SDS-PAGE electrophoresis, the gel

was either stained with Coomassie Blue staining solution containing 50% Methanol, 10% acetic acid and 0.05% (w/v) Coomassie Brilliant Blue R-250 or transferred to a Hybond- P PVDF membrane (GE Healthcare, USA) at constant voltage of 80V.

Western Blotting: Following the transfer, the membranes were washed with 1X TBS and blocked with 3% skim milk in TBST for 1h. Primary probing was done with either His- Probe HRP (1:500 dilution; Santa Cruz Biotechnology, USA) or mouse anti-flag antibody (1:10,000 dilution; Sigma-Aldrich, USA) for 30 min. Following the incubation, the membranes were washed thrice in TBST solution for 10 min. The membranes were then incubated with secondary rabbit anti-mouse antibody conjugated to Horseradish peroxidase (1: 40000 dilution) for 30 min. After incubation, the membranes were washed four times in TBST for 10 min. The bands in the blot were detected using ECL Plus western blot detection system (GE Healthcare, USA) and visualized by ImageQuant LAS 4000 (GE Healthcare, USA).

2.3 Results

2.3.1. Overexpression

Rhomboid proteases from three prokaryotic organisms *hi*GlpG from *H. influenzae*, *ec*GlpG from *E. coli*, and YqgP from *B. subtilis*, were analyzed for protein over-expression. Final cell culture parameters for all rhomboids resulted in the use of 0.002 % arabinose with a 5 h induction time at 24°C. The proteins were solubilized in dodecylamaltoside (DDM, Anatrace, USA) using Ni-NTA column purification as previously described for *hi*GlpG [7] . Expression yields post-purification using the His-tag affinity Ni-NTA resin were as follows: 1.8 mg/L for *hi*GlpG, 2.0 mg/L *ec*GlpG, and 1.6 mg/L for YqgP. All proteins were subjected to gel filtration chromatography for further purification, resulting in a protein band on SDS-PAGE with 90-95% homogeneity (Fig. 2.3.1).

2.3.2. Oligomeric state of three prokaryotic rhomboid proteases

In order to assess the oligomeric state of the three rhomboid peptidases, *hi*GlpG, *ec*GlpG and YqgP were purified in DDM and subjected to sedimentation equilibrium analysis by analytical ultracentrifugation (Table 2.3.1 and Fig. 2.3.2). All speeds for each sample, *hi*GlpG, *ec*GlpG, and YqgP are also shown (Fig. 2.3.7, 2.3.8, 2.3.9). A global mass of 135,079 Da, 141,476 Da and 203,989 Da was obtained for *hi*GlpG, *ec*GlpG and YqgP respectively. Thin layer chromatography was used to assess detergent amounts associated with

a single molecule of *hi*GlpG using densitometric analysis and it was found to be approximately 45K±5 kDa (Fig.2.3.6) [13]. This is in agreement with the amount of radiolabelled DDM associated with a membrane protein LacS that consists of 12 transmembrane segments, which was calculated to be 100K of DDM per monomer [14]. When the global mass obtained from the sedimentation equilibrium is divided by the mass of the rhomboid peptidases plus the detergent bound, we get two rhomboid molecules per species for *hi*GlpG, *ec*GlpG and YqgP (Table 2.3.1), indicating that these prokaryotic rhomboids preferentially forms a dimer in detergent solution. The best-fit analysis is in agreement with all proteins fitting to a monomer-dimer-tetramer model with a dimer being the predominant species for *hi*GlpG and *ec*GlpG. Best fit analysis with YqgP indicated that more tetramer was observed compared to *hi*GlpG and *ec*GlpG.

To further analyse the oligomeric state of rhomboids and determine whether the membrane domain is responsible for the dimerization, we focused solely on the *hi*GlpG. Sequence analysis indicates that this is the simplest form of the rhomboid family of proteins [15]. Its crystal structure indeed confirms it having only 6 transmembrane segments with no large N- or C-terminal domains [7].

2.3.3 Crosslinking studies with detergent solubilized *hi*GlpG indicate a dimeric species.

We analyzed the oligomeric state of DDM solubilized *hi*GlpG using the homobifunctional crosslinking reagents, DTSSP and DSP. These are membrane impermeant and membrane soluble reagents, both of which react covalently with amino groups, and their internal disulfide bond can be cleaved by reducing reagents such as DTT. These two crosslinkers were initially chosen to distinguish between crosslinking with loop and transmembrane segments. The molecular weight of the *hi*GlpG-MycHis is 25,061Da, however when resolved on SDS-PAGE, the protein runs at 23KDa (Fig. 2.3.1). When crosslinked samples were resolved on SDS-PAGE and blotted on a PVDF membrane for Western blot analysis, we observed that both DSP and DTSSP efficiently cross-linked *hi*GlpG and the dimeric species was running at 45KDa, which was the expected molecular weight for a dimer (Fig. 2.3.3). Followed by the addition of DTT, a reducing agent that separates the homobifunctional crosslinkers, the cross-linked products were cleaved and found to be migrating at the same molecular weight as that of the monomer. Slightly more crosslinking is observed with the membrane permeant DSP compared to the impermeant DTSSP. Two of the three lysines in *hi*GlpG are located at the base of transmembrane segments and DSP being membrane permeant may be able to penetrate the detergent micelle more readily that DTSSP. We do not observe 100% crosslinking which is typical for membrane proteins. A weak dimer band is observed in the control lane

without crosslinker suggesting a strong interaction between dimers that is not fully separated by SDS. In addition we see a faint band in the crosslinking samples at approximately 70K near the trimer range, however the absence of a trimer in the gel filtration (see below) and analytical ultracentrifugation experiments suggests this may be a minor contaminant crosslinking with *hi*GlpG.

2.3.4. Gel filtration and activity assay indicates *hi*GlpG is dimeric and functional

To determine if the dimer was indeed the predominant species for *hi*GlpG, Ni-NTA purified sample was run on an analytical gel filtration (GF) column (Fig. 2.3.4a). Examination of the profile for *hi*GlpG reveals retention time of the eluted *hi*GlpG just below the 158kDa marker. Given approximately 45K of detergent (approximately one micelle of DDM) associated with one *hi*GlpG, the gel filtration data suggests we have a dimer under these conditions. In addition, we also observe a shoulder peak at approximately half of the height of the major peak eluting at 12ml retention time. Using the standard curves, regression analysis calculation for the molecular weight for this peak indicates a mass of approximately 265kDa suggesting a tetrameric species is also present which is in agreement with the analytical centrifugation results.

In order to assess if the dimer was active, Ni-NTA purified *hi*GlpG was also subjected to a gel-based activity assay, typical for intramembrane

proteases [4, 16](Fig. 2.3.4b). *Providencia stuartii* TatA (psTatA) with an N-terminal Flag tag was used as the substrate. In the presence of *hi*GlpG, *ps*TatA was cleaved that was assessed by Western blotting using Anti-Flag antibody, indicating that the dimeric *hi*GlpG is functional.

2.3.5. Co-purification of *hi*GlpG-His and *hi*GlpG-Flag shows rhomboid form dimers in the membrane bilayer.

To test if *hi*GlpG formed dimers within the membrane, we designed a co-expression study to allow the expression of two different *hi*GlpG: one with a His tag and the other with a Flag tag. DDM- solubilized membranes were isolated from cells co-expressing *hi*GlpG-His and *hi*GlpG-Flag. Also DDM-solubilized membranes were isolated from cells expressing each individual clones of *hi*GlpG-His and *hi*GlpG-Flag to serve as controls. As a control, individual membranes of *hi*GlpG-His and *hi*GlpG-Flag were also mixed prior to the addition of detergent to ensure that dimers did not form as a result of the solubilisation step. Anti-Flag affinity gel was used to purify the proteins and the presence of dimers in the immunoprecipitated fractions was examined by Western blotting with either Anti-Flag antibody or His-Probe. Fractions expressing Flag tagged proteins immunoprecipitated with Flag affinity resin were detected by the Anti-Flag antibody (Fig. 2.3.5a). As expected, *hi*GlpG-His was not detected by Anti-Flag antibody. The blot was also subsequently probed with the His-Probe (Fig. 2.3.5b). In the co-expressed fraction purified with Flag resin, *hi*GlpG-His was detected

demonstrating that the two epitopes were found to coimmunoprecipitate. This indicated that *hi*GlpG-Flag and *hi*GlpG-His associated within the membrane. In order to rule out the possibility that the co-immunoprecipitation was due to disruption of the dimers, a control was included where the His- and Flag-tagged dimer were expressed and purified independently. The separately purified proteins were then subjected to the same detergent treatment for the co-expression study. A small but detectable signal was noted in the mixed membrane fraction when the blot was probed with the His-Probe antibody (Fig. 2.3.5b) suggesting only minor disruption of the dimers by the addition of the detergent. It is clear from this result that *hi*GlpG forms oligomers, most likely dimers, in the lipid bilayer.

2.4 Discussion

This chapter addressed the oligomeric state of different prokaryotic rhomboid proteases. Analytical ultracentrifugation results indicate that prokaryotic rhomboids *hi*GlpG, *ec*GlpG and YqgP form monomers, dimers and tetramers with dimers being the predominant species. In order to determine if the membrane domain alone was responsible for dimerization, we carried out further experiments only with *hi*GlpG as it represents the basic 6 TM rhomboid core. Confirming the analytical ultracentrifugation results, both crosslinking and gel filtration experiments show that *hi*GlpG forms a dimeric species. In addition, activity assay also reveals that the dimeric *hi*GlpG is functional.

An important question that needed to be addressed was whether the dimer formed in the membrane bilayer or if it was an effect of detergent solubilisation. In our pull-down assay, we coexpressed a His tagged *hi*GlpG and a Flag tagged *hi*GlpG and both were found to coimmunoprecipitate demonstrating that *hi*GlpG formed dimers *in vivo*. To test if the homogenization step during purification did not alter the oligomers, we conducted a control where membrane fractions harbouring only *hi*GlpG-His or *hi*GlpG-Flag were mixed prior to co-immunoprecipitation. This control was used to demonstrate that the detergent or the mechanical steps during purification did not disrupt or affect the native dimers found in the membrane. Only a faint signal was observed when Flag immunoprecipitates were probed using a His-probe, suggesting a slight disruption of the Flag homodimers to form Flag.*hi*GlpG and His.*hi*GlpG heterodimers. Coimmunoprecipitation experiments also suggest that the dimers are not easily separated. Unfortunately, we have not yet identified any conditions or reagents that could disrupt the dimer, therefore it is currently not possible to assess if the dimeric species is essential for function.

The objective of rhomboid dimerization may be because of two reasons: dimerization may assist in function and/or in the stability of the enzyme. Since rhomboids have a relaxed substrate specificity compared to other serine peptidases [17], it is necessary that rhomboid activity is regulated by other mechanisms. These experiments have shown that *hi*GlpG, the simplest rhomboid, exists as a dimer. Like *hi*GlpG, *ec*GlpG also belongs to

the secretase A family of rhomboids containing the six transmembrane core, with *ec*GlpG carrying the N-terminal cytoplasmic domain. YqgP belongs to secretase B family with seven transmembrane segments, where an extra transmembrane segment is fused to the six transmembrane core at the C terminal region [15]. Given the wide difference in topology, all these rhomboids have in common the six transmembrane core that is most likely the mechanism for dimerization. It is also tempting and not unjustified to speculate that the other secretase members along with the eukaryotic homologs form dimers.

In the crystal structure of *H. influenza* rhomboid, a physiological dimer was not observed; however a crystallographic dimer with head-to-tail packing was observed, where the two monomers have opposite topologies [7]. This is similar to that seen in various E. coli rhomboid structures forming three dimensional crystals [18-20]. Recently, a projection map from the two dimensional crystal of *ec*GlpG was published [21]. It is evident from this projection map that *ec*GlpG exists as a dimer in the asymmetric unit. Since neither a three dimensional map was provided nor the crystal structure was docked, it is difficult to predict the region of dimeric interface in the protein. Since the dimer is functional, there exists the possibility that Loop1, which is on the opposite side of the substrate entry pathway, may facilitate this dimerization. New crystals from *ec*GlpG membrane domain show the presence of a trimer, however this may be a crystallographic artifact, as both TM5 and Loop 5 point towards the centre, preventing easy access to the

active site [22, 23]. In the crosslinking assay, we do observe a faint band near the predicted trimeric molecular weight of *hi*GlpG, however it is very faint. Also, gel filtration and analytical ultracentrifugation experiments do not suggest a physiological trimer in vitro, as only dimers and tetramers are observed.

Evidence of monomer in the crystal structure along with dimer in the membrane have been observed with the sodium proton antiporter, NhaA [24]. While the three dimensional crystal depicts a monomer[25], two dimensional crystals [5] along with electron microscopy map (EM), electron spin resonance (ESR) measurements and crosslinking [26, 27]demonstrate that the antiporter exists as a dimer in the membrane. It was later revealed that dimerization was important to maintain the antiporter under conditions of extreme stress, during cell growth such as alkaline pH, salt and high temperatures [28]. The dimer interface was composed of two β- hairpins of the two monomers, forming an anti-parallel β- sheet at the periplasmic side of the membrane [27]. Examples exist of other membrane proteins that oligomerize to regulate their function. Anion exchanger, AE1 exists in equilibrium between both dimers and tetramers. While the monomers were thought as the functional units, a paper suggests that dimers are the functional units with each monomer regulating the other allosterically [6].

The fact that *ec*GlpG may homo-oligomerize was not studied in detail with *ec*GlpG [29]. YqgP is also known to oligomerize as large aggregates, however the study was carried out in the absence of detergent [30]. This

chapter describes the first study of oligomerization in rhomboid proteases. Interesting questions arise from these results. For example: does dimerization affect the function of rhomboids and influence the rate of substrate cleavage? Also, the lipids modelled as phosphatidic acid were found in *hi*GlpG structure [7] as well as a *ec*GlpG crystals were grown in lipid bicelles [21]. It would be interesting to observe if lipids play a role in oligomerization and consequently influence the catalytic activity of rhomboid proteases.

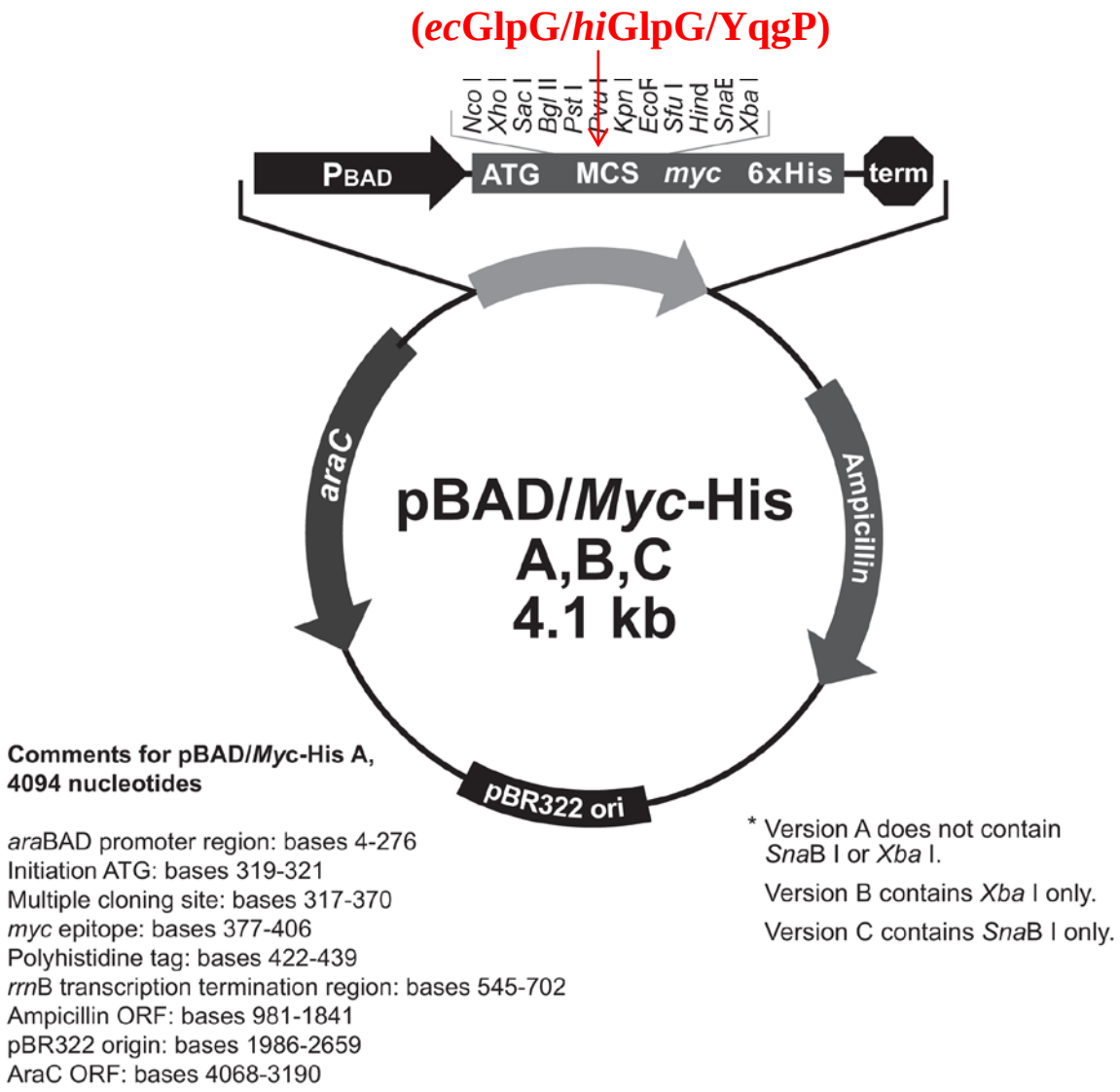


Figure 2.2.1: **Circular vector map of pBAD.MycHisA:**

Vector map of pBAD.MycHisA which was used to express pBAD.*ec*GlpG, pBAD.*hi*GlpG and pBAD.YqgP.

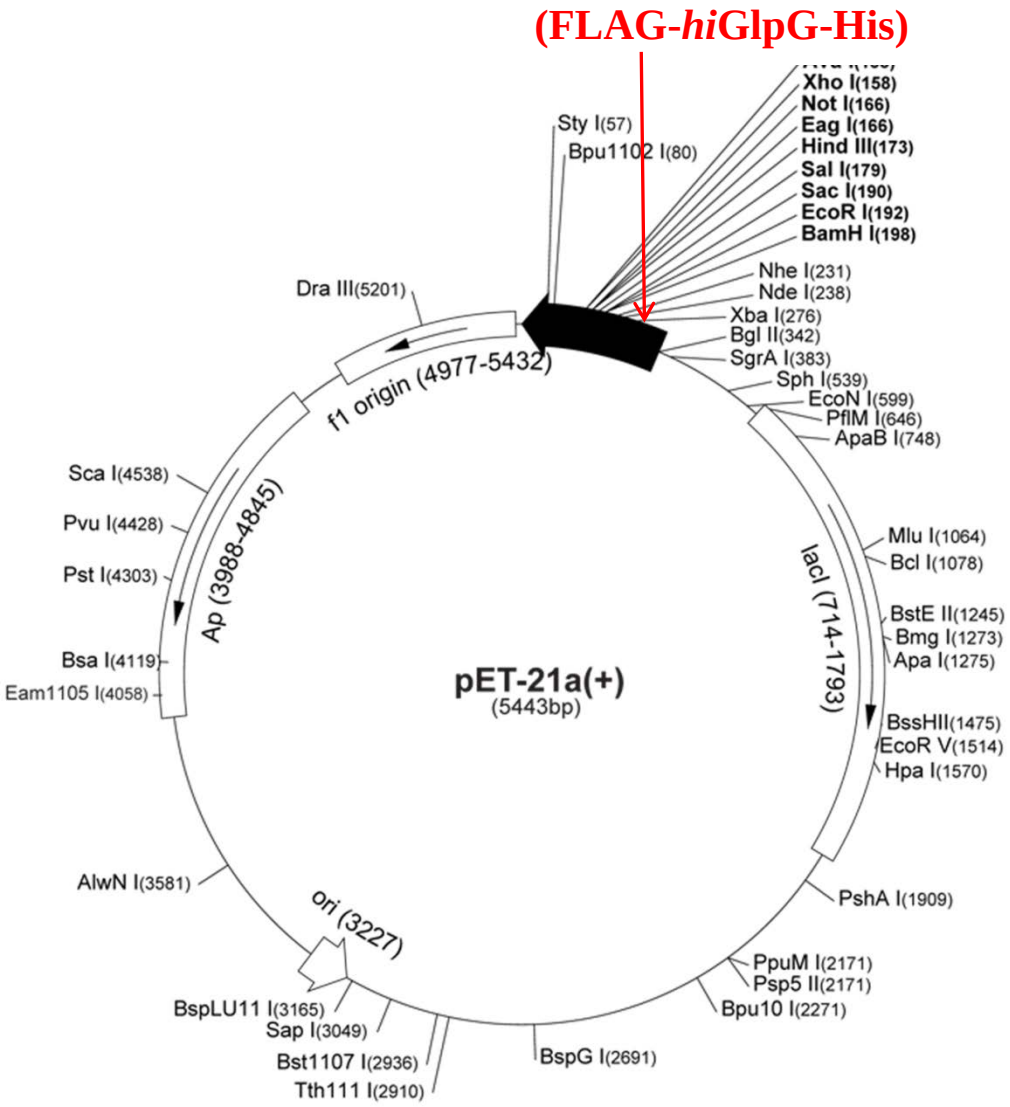
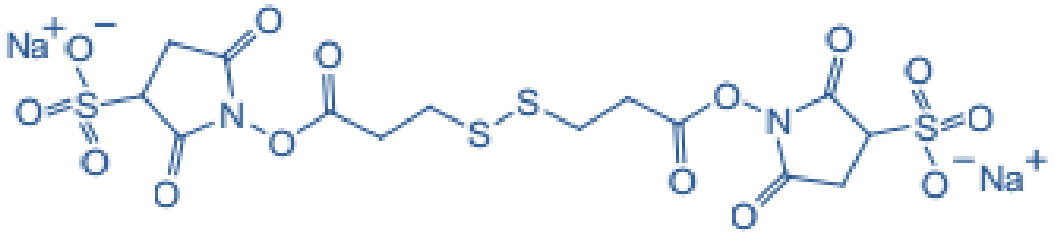
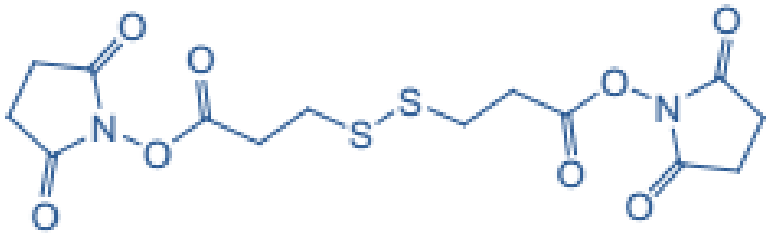


Figure 2.2.2. Circular map of pET21a:

Red arrow shows the insertion of Flag tag near the N- terminus of *hi*GlpG.His.



DTSSP - 3, 3’ dithiobissulfo (succinimidylpropionate)



DSP – dithiobis (succinimodylpropionate)

Figure 2.2.3: DTSSP and DSP Crosslinkers:

Crosslinkers containing amine-reactive NHS-ester ends that react with primary amines of the protein forming stable amide bonds.

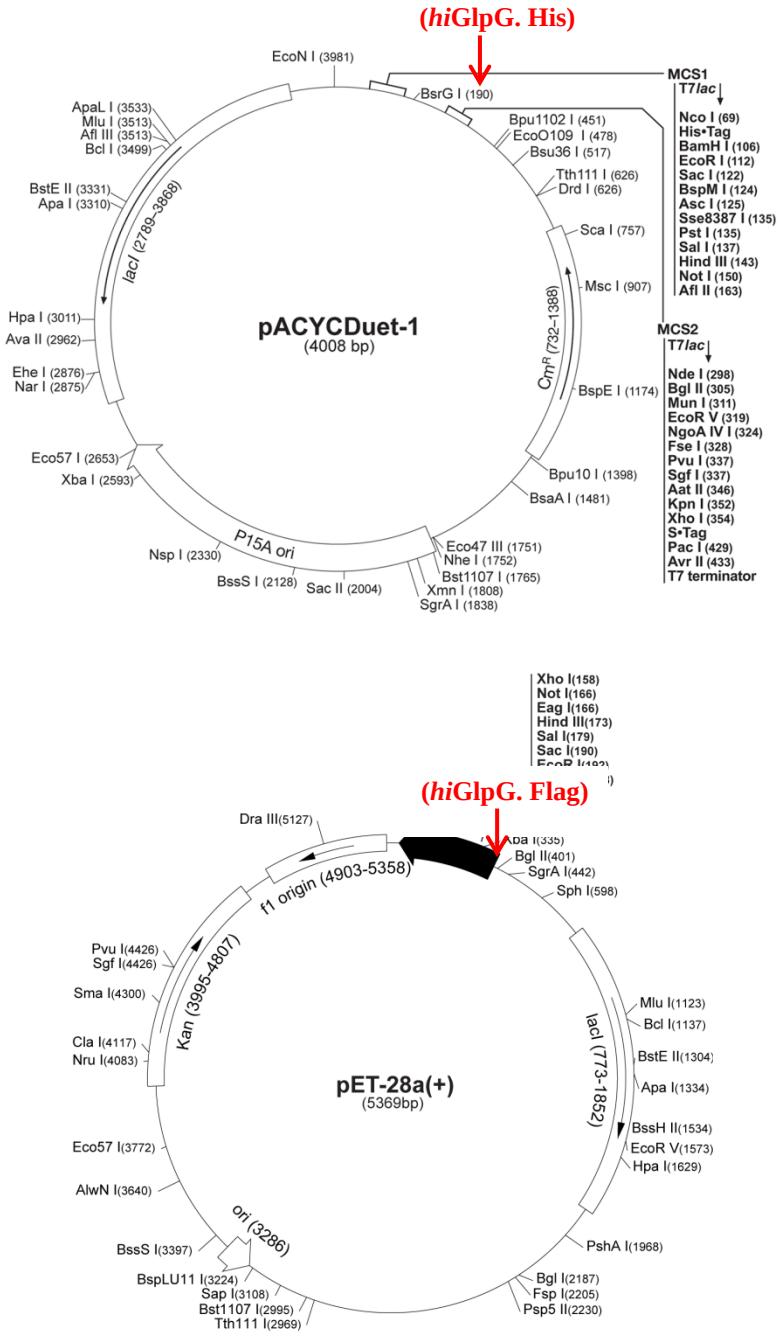


Figure 2.2.4. Vector Map of pACYCDuet1 and pET28a.

Red arrows indicate the insert sites of *hi*GlpG.His and *hi*GlpG.Flag respectively.

Primer Sequence:

Primer Name	Primer Sequence
FLAG.hiGlpG.His- Forward primer	5'-atgggtcgcggatccgattacaaggacgatgacgataagatggaatcaactattgca-3'
FLAG.hiGlpG.His- Reverse primer	5'-tgcaatagttgattccatcttatcgtcatcgtccttgtaatcggatccgcgacccat-3'

PCR conditions for Site- directed Mutagenesis:

Reagent	Volume (µl)
Plasmid template (pET21a.*hi*GlpG.His)	1
10X reaction buffer	5
Forward primer (stock 100ng/ µl)	1.25
Reverse primer (stock 100ng/ µl)	1.25
dNTP mix	1
Quik Solution reagent	1.5
ddH_2O	38
QuikChange Lightning Enzyme	1

Cycling Parameters for Site-directed Mutagenesis

Segments	Cycles	Temperature	Time
1	1	95ºC	2 min
2	18	95 ºC	20 sec
		60 ºC	10 sec
		68 ºC	2 min 45 sec
3	1	68 ºC	5 min

Table 2.2.1. Site-directed mutagenesis for pET21a.psTatA.His.

Mutagenesis PCR protocol and parameters for insertion of FLAG tag into pET21a.psTatA.His.

Primer Design:

Primer Name	Primer Sequence
Insertion of His tag- Forward primer	5'-gcgcaaaaattcgctagagcatcatcaccatcaccactaa ctggtaccatatgggaatt-3'
Insertion of His tag- Reverse primer	5'-aattcccatatggtaccagttagtggtgatggtgatgatgctcta gcgaatttttgcgc-3'
Insertion of Flag tag- Forward primer	5'-gcgcaaaaattcgctagaggattataaagatgacgacgataagt aactggtaccatatgggaatt-3'
Insertion of Flag tag- Reverse primer	5'-aattcccatatggtaccagttacttatcgtcgtcatctttataatcct ctagcgaatttttgcgc-3'

PCR conditions for Site- directed Mutagenesis:

Reagent	Volume (µl)
Plasmid template (pET21a.*hi*GlpG.His)	1
10X reaction buffer	5
Forward primer (stock 100ng/ µl)	1.25
Reverse primer (stock 100ng/ µl)	1.25
dNTP mix	1
Quik Solution reagent	1.5
ddH_2O	38
QuikChange Lightning Enzyme	1

Cycling Parameters for Site-directed Mutagenesis:

Segments	Cycles	Temperature	Time
1	1	95ºC	2 min
2	18	95 ºC	20 sec
		60 ºC	10 sec
		68 ºC	2 min 15 sec
3	1	68 ºC	5 min

Table 2.2.2. Site-directed mutagenesis PCR of pBAD.*hi*GlpG.MysHisA.

The table above lists the parameters and protocol for generating two different tagged proteins, *hi*GlpG.His and *hi*GlpG.Flag.

Resolving gel :12% and 16%

30% Acrylamide	2.4 ml for 12% gels; 3.2 ml for 16% gels
H_2O	3.4 ml for 12% gels; 2.6 ml for 16% gels
1.5M Tris (pH 8.8)	1.25 ml
10% SDS	50 µl
10% APS	25 µl
TEMED	5 µl

Stacking gel: 4%

30% Acrylamide	0.5 ml
H_2O	3.1 ml
0.5M Tris (pH 6.8)	1.25 ml
10% SDS	50 µl
10% APS	25 µl
TEMED	10 µl

Table 2.2.3. SDS-PAGE gel composition.

Compositions of the resolving gel and the stacking gel used for 4%/12% and 4%/16% gels.

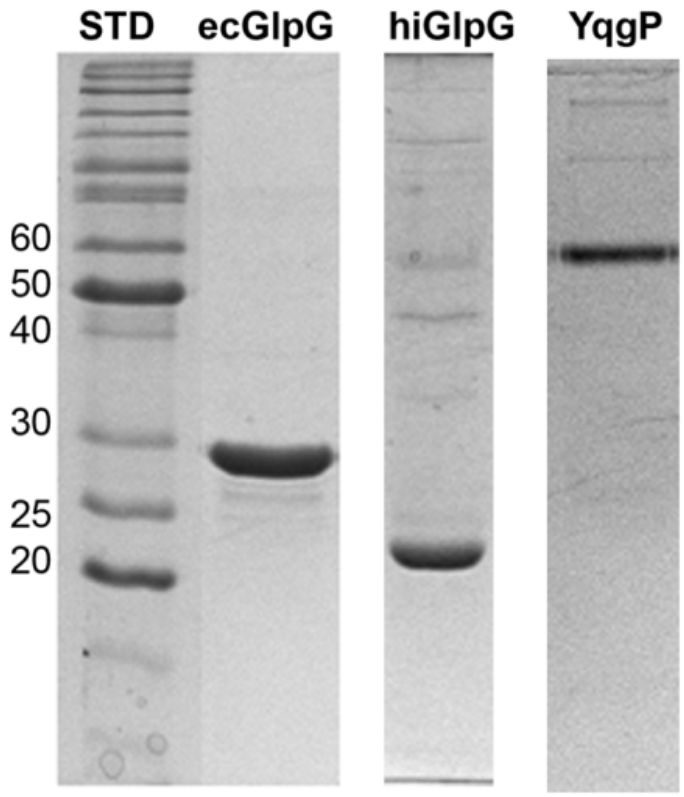


Fig. 2.3.1. Overexpression of prokaryotic rhomboid proteases.

SDS-PAGE of overexpressed prokaryotic rhomboid peptidases: *ec*GlpG from *E. coli*, *hi*GlpG from *H. influenzae*, and YqgP from *B. subtilis* (with their epitope tags removed). Molecular weights of marker (STD) (in kDa) are mentioned on the left-hand side.

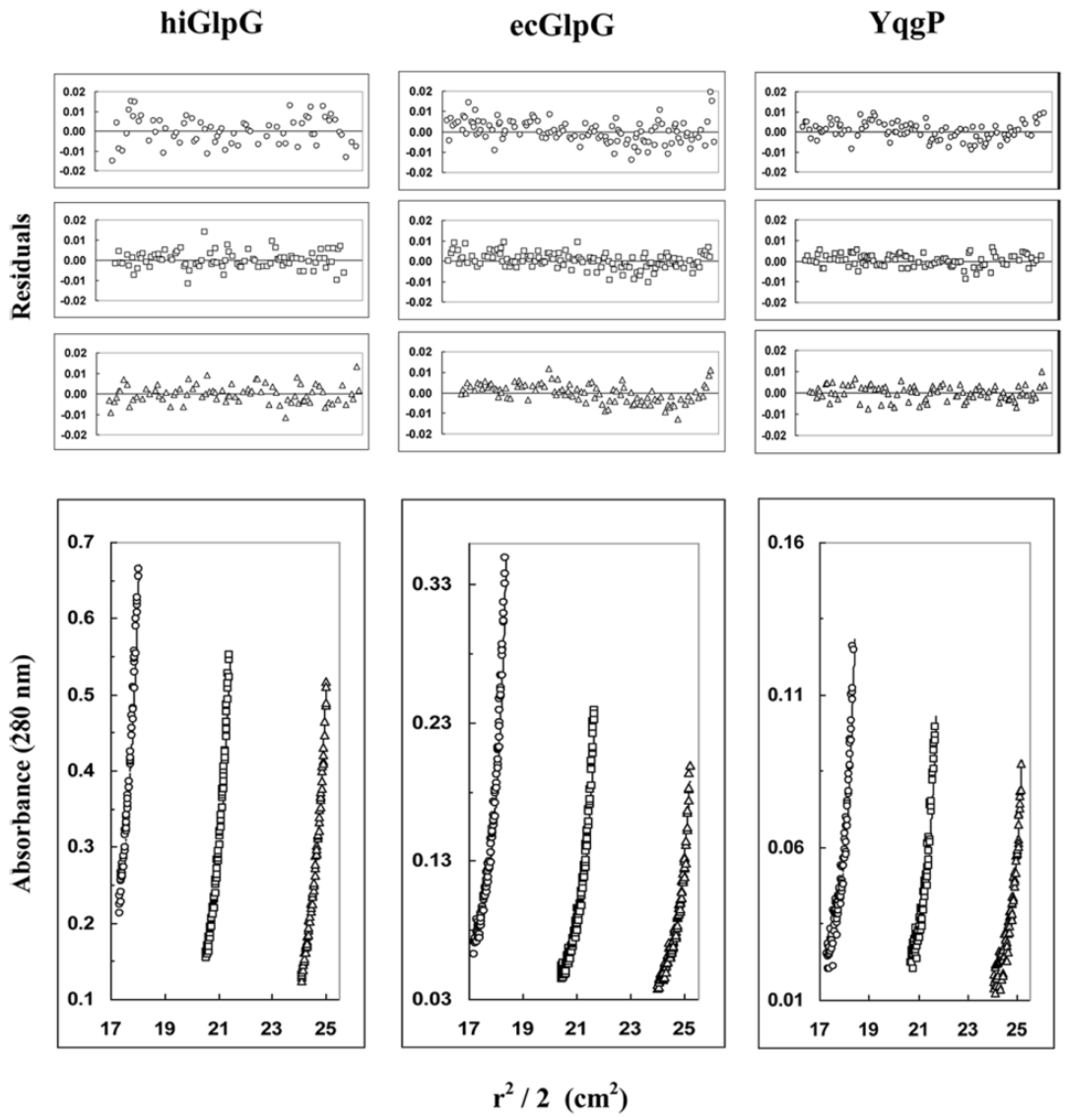


Fig. 2.3.2: Sedimentation equilibrium analysis of *hi*GlpG, *ec*GlpG and YqgP in 0.1% DDM. All three proteins were dissolved in 20mM Tris, 20mM NaCl pH 8.0 and 0.05% DDM and were each centrifuged at three different rotor speeds at 20ºC. Only the data collected at the lowest rotor speeds, which are 10,000, 9,000 and 8,000 rpm for hiGlpG, ecGlpG and YqgP respectively, are shown. The protein concentrations used were 0.53 (circles), 0.35 (squares) and 0.26 mg/mL (triangles) for hiGlpG; 0.17 (circles), 0.11 (squares) and 0.08 mg/mL (triangles) for ecGlpG; and 0.18 (circles), 0.12 (squares) and 0.09 mg/mL (triangles) for YqgP. Lower graphs illustrate $r^2/2$ versus absorbance plots; symbols represent measured data points, and solid lines represent fit lines to a monomer-dimer-tetramer model. Upper graphs illustrate the residuals from fitting the measured data points to a three-species model. The random, nonsystematic distribution of the residuals indicates a good fit of the data to the models.

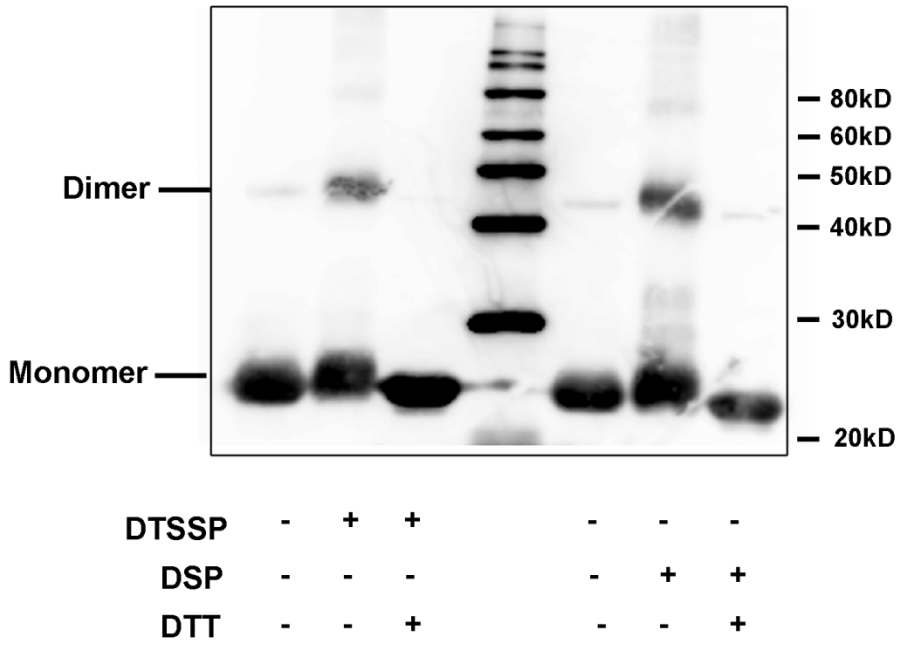


Fig. 2.3.3. Crosslinking of *hi*GlpG in detergent solution detected by Western blot using His-Probe.

His-Probe detection of detergent solubilized *hi*GlpG 10µg (20µl) aliquots incubated with 1mM DTSSP and DSP crosslinking agents, with or without the reducing agent DTT. The arrows show the position of monomer and dimer at approximately 23kDa and 45kDa respectively.

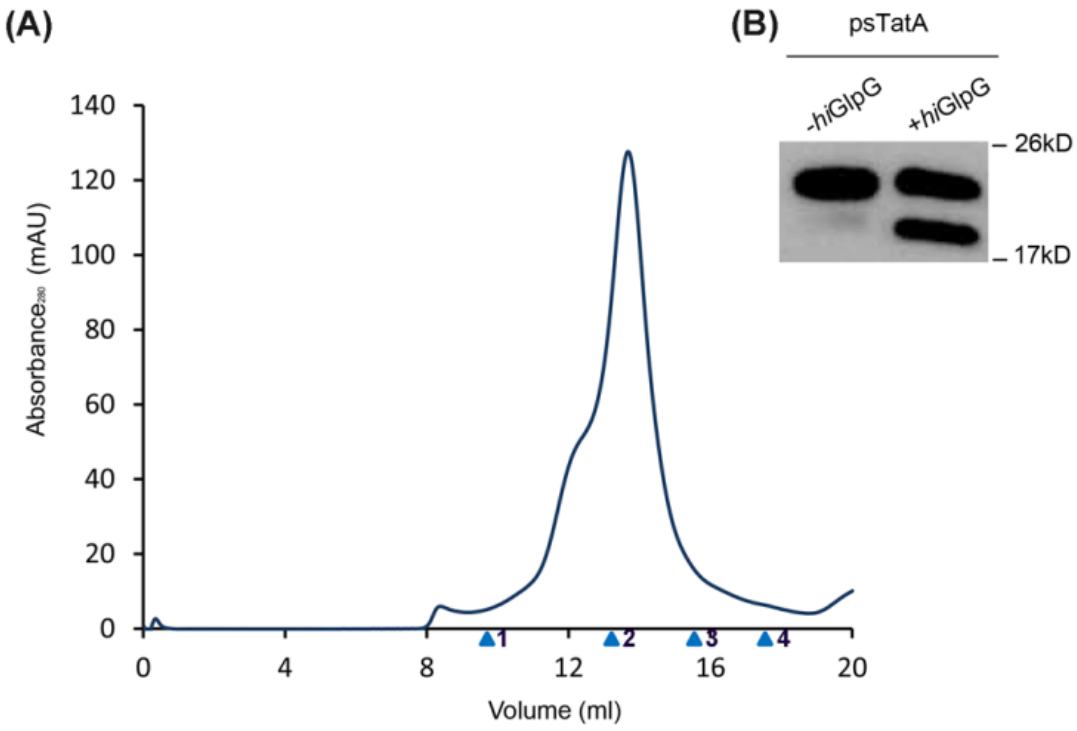


Fig. 2.3.4. Gel filtration and functional assay of *hi*GlpG.

(A) Gel Filtration of *hi*GlpG in 0.1% DDM. Approximately 200µg of Ni-NTA purified *hi*GlpG in DDM was subjected to gel filtration onto a Hiload Superdex 200 10/300 column (GE Healthcare, USA) containing 50mM Tris, 200mM NaCl and 5% glycerol, pH 8.0 supplemented with 0.1% DDM. Vo- void volume, Vt-total column volume (24ml). Arrowheads represent the standard proteins from left to right: 1. thyroglobulin (MW 670k; Stokes radius 85 Å); 2. IgG (MW 158k; Stokes radius 55 Å); 3. ovalbumin (MW 44k; Stokes radius 30.5 Å); and 4. myoglobin (MW 17k, Stokes radius 20.7 Å)

(B) Rhomboid cleavage activity on *ps*TatA. SDS-PAGE demonstrating *P. stuartii* TatA (psTatA) substrate cleavage by *hi*GlpG in 0.1% DDM. Samples run include a control on the left panel. Molecular mass markers are reported in kDa on the left-hand side.

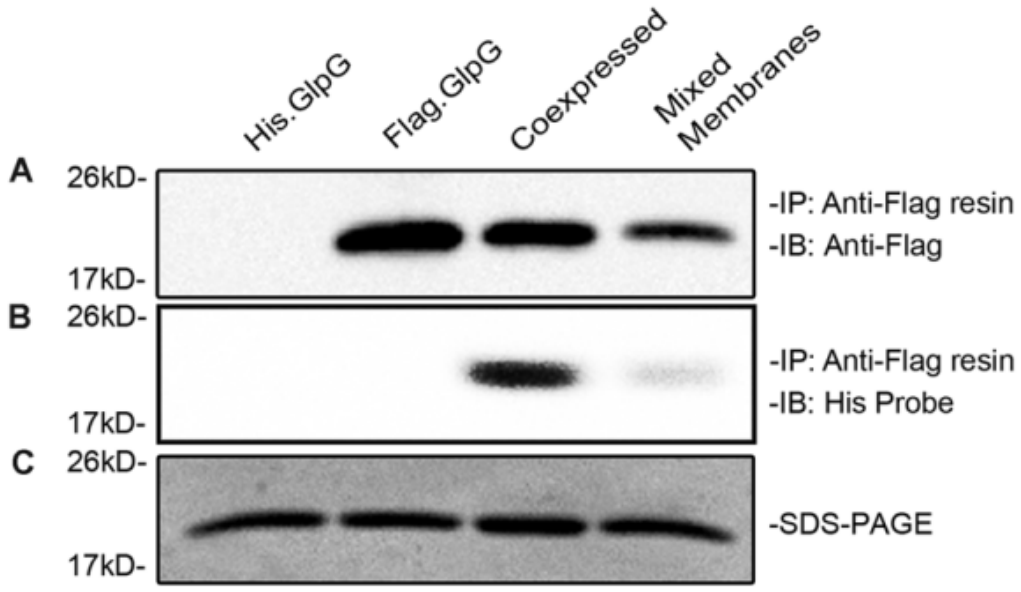


Fig. 2.3.5. Pull down assay with co-expressed His and Flag-tagged *hi*GlpG shows in vivo association.

Co-immunoprecipitation assay of *hi*GlpG molecules bearing two different immunological epitopes are shown to validate the formation of hiGlpG dimers within the membrane bilayer. *hi*GlpG-Flag and *hi*GlpG-His were expressed either independently, or coexpressed. In addition for a control, *hi*GlpG-Flag and *hi*GlpG-His membrane fractions were mixed prior to immunoprecipitation. Upon Anti-Flag immunoprecipitation, each fraction was separated on a 12% SDS-PAGE gel and analysed by immunoblotting with either (A) Anti-FLAG or (B) His-Probe. (C) A Coomassie stained gel with the four different membrane fractions is shown as a loading control to ensure equal amounts of protein was added to the resin. Molecular mass markers are reported in kDa on the left-hand side.

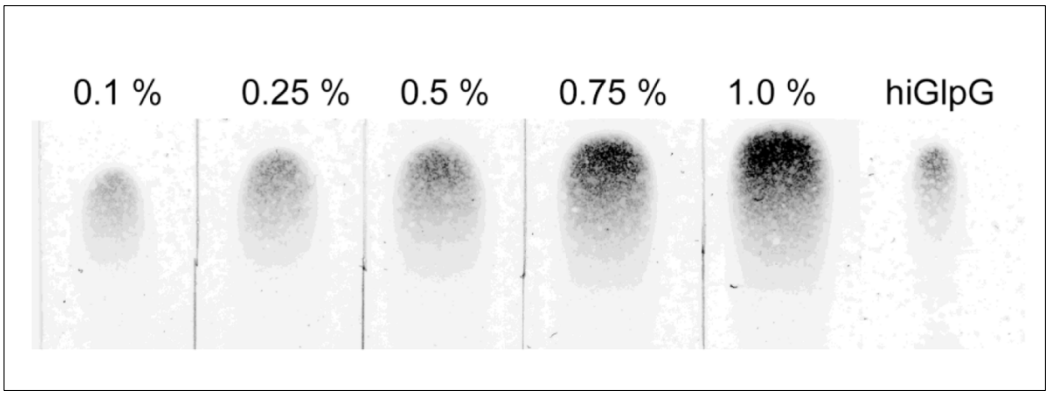


Figure 2.3.6. Thin layer chromatography of DDM associated with *hi*GlpG:

Aliquots of 10 µl of DDM standards of a given concentration were spotted on a Silica TLC plate (Whatman, USA). The hiGlpG sample obtained post-Ni-NTA purification was applied in a 20 µl aliquot containing 9.2 µg of protein.

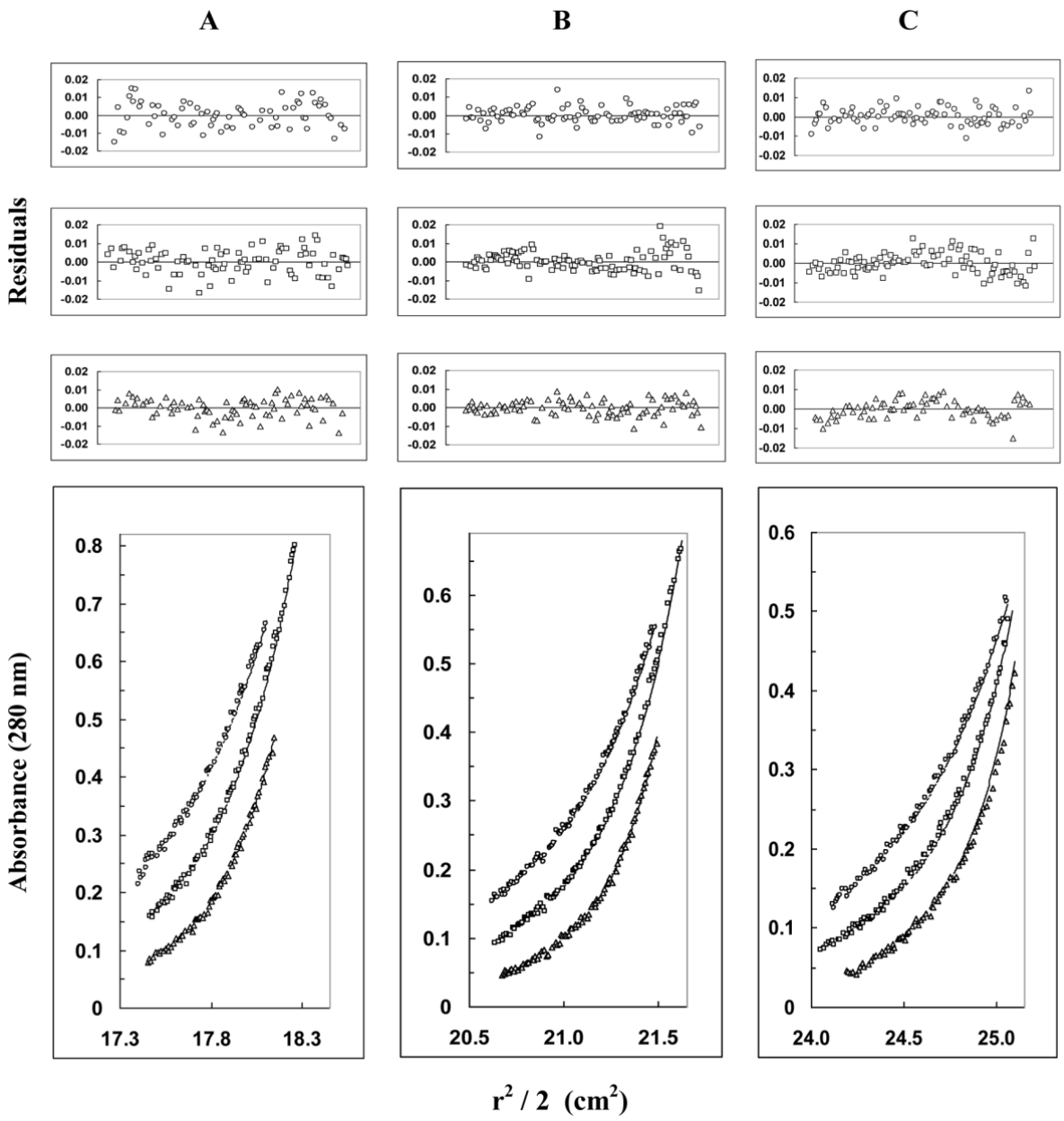


Figure 2.3.7. Sedimentation equilibrium analysis of *hi*GlpG:

Protein was dissolved in 20mM Tris, 20mM sodium chloride, 18% D_2O and 0.05 % DDM, pH 8.0 and centrifuged at 10000 rpm (circles), 12000 (squares) and 14000 rpm (triangles) at 20ºC. The protein concentrations used were 0.53, 0.35 and 0.26 mg/mL for sectors A, B and C, respectively. Lower panels: $r^2/2$ versus absorbance plots. Symbols represent measured data points, and solid lines represent the fit to a monomer-dimer-tetramer model. Upper panels: Residuals obtained from fitting the measured data points to a two-species model. The random, nonsystematic distribution of the residuals indicates a good fit of the data to the models.

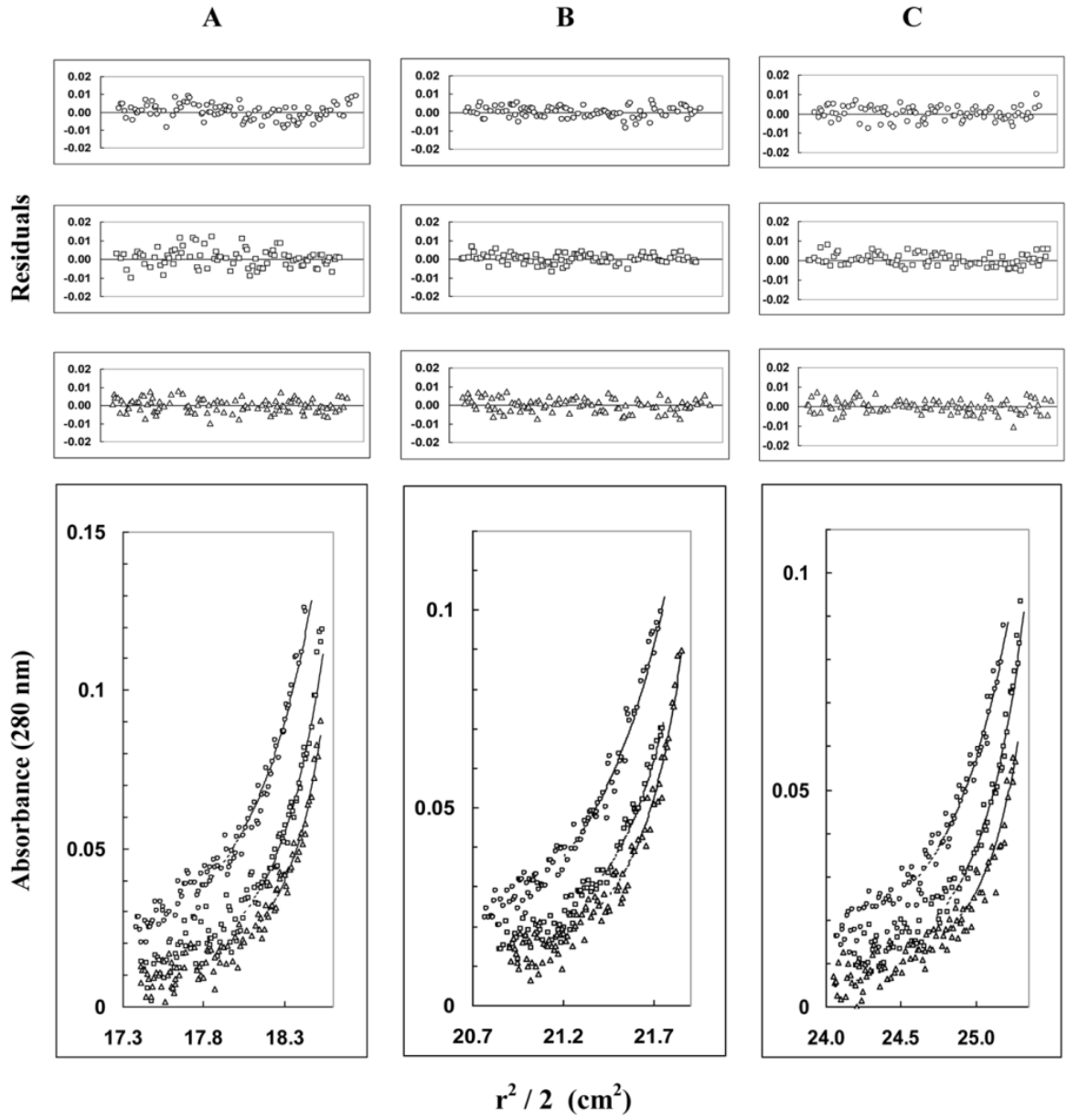


Figure 2.3.8. Sedimentation equilibrium analysis of YqgP:

Protein was dissolved in 20mM Tris, 20mM sodium chloride, 18% D_2O and 0.05 % DDM, pH 8.0 and centrifuged at 8000 rpm (circles), 10000 (squares) and 12000 rpm (triangles) at 20ºC. The protein concentrations used were 0.18, 0.12 and 0.09 mg/mL for sectors A, B and C, respectively. Lower panels: $r^2/2$ versus absorbance plots. Symbols represent measured data points, and solid lines represent the fit to a monomer-dimer-tetramer model. Upper panels: Residuals obtained from fitting the measured data points to a two-species model. The random, nonsystematic distribution of the residuals indicates a good fit of the data to the models.

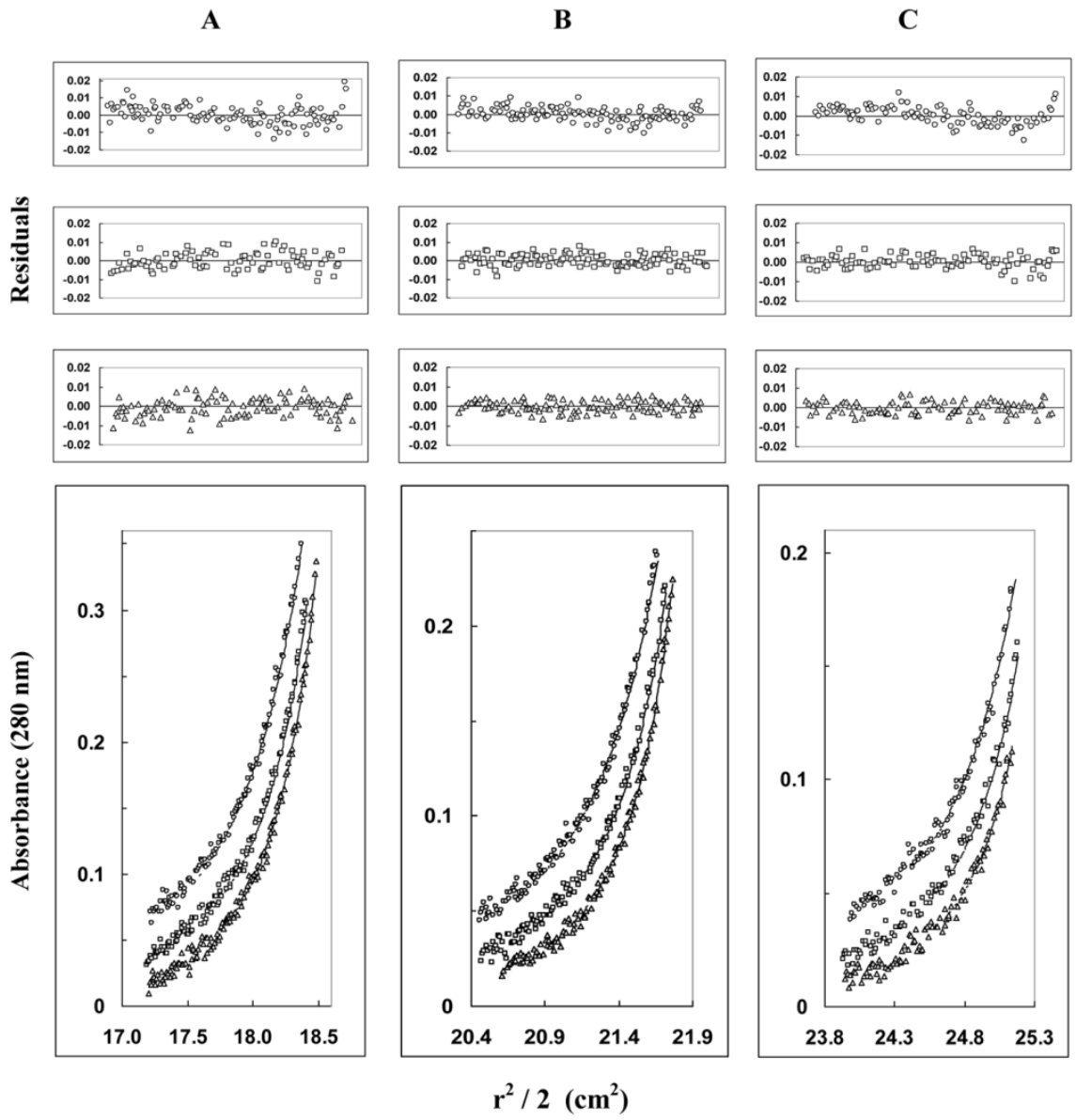


Figure 2.3.9. Sedimentation equilibrium analysis of *ec*GlpG:

Protein was dissolved in 20mM Tris, 20mM sodium chloride, 18% D_2O and 0.05 % DDM, pH 8.0 and centrifuged at 9000 rpm (circles), 11000 (squares) and 13000 rpm (triangles) at 20ºC. The protein concentrations used were 0.174, 0.115 and 0.087 mg/mL for sectors A, B and C, respectively. Lower panels: $r^2/2$ versus absorbance plots. Symbols represent measured data points, and solid lines represent the fit to a monomer-dimer-tetramer model. Upper panels: Residuals obtained from fitting the measured data points to a two-species model. The random, nonsystematic distribution of the residuals indicates a good fit of the data to the models.

Protein in DDM	Protein (Da)	Protein + Det. MW	Global Fit M.W.	Best fit	monomer/global best fit
*hi*GlpG	22124	61,967	135,079	M-**D**-T	2.3
*ec*GlpG	31772	71,615	141,476	M-**D**-T	2.0
YqgP	54700	94,543	203,989	M-T	2.2

Table 2.3.1. Summary of sedimentation equilibrium results showing are the different calculated rhomboid protease mass post thrombin cleavage, without detergent, calculated molecular weight (MW) obtained from adding the protein mass with the calculated amount detergent of detergent bound and the global fit MW. The ratio to calculate the quaternary state was obtained by dividing the global fit MW by the protein plus detergent molecular weight. Bold indicates predominant species found during the analysis.

Chapter 2: References

1. Guichard A, B.B., Sturtevant MA, Wickline L, Chacko J, Howard K, Bier E., *rhomboid and Star interact synergistically to promote EGFR/MAPK signaling.* Development, 1999 **126**(12): p. 2663-2776.
2. Bang, A.G. and C. Kintner, *Rhomboid and Star facilitate presentation and processing of the Drosophila TGF-alpha homolog Spitz.* Genes & development, 2000. **14**(2): p. 177-86.
3. Sík A, Passer BJ, Koonin EV, Pellegrini L., *Self-regulated cleavage of the mitochondrial intramembrane-cleaving protease PARL yields Pbeta, a nuclear-targeted peptide.* The Journal of biological chemistry, 2004. **279**(15): p. 15323-9.
4. Urban, S. and M.S. Wolfe, *Reconstitution of intramembrane proteolysis in vitro reveals that pure rhomboid is sufficient for catalysis and specificity.* Proceedings of National Academy of Sciences, United States of America, 2005. **102**(6): p. 1883-8.
5. Gerchman Y, Rimon A, Venturi M, Padan E., *Oligomerization of NhaA, the Na+/H+ antiporter of Escherichia coli in the membrane and its functional and structural consequences.* Biochemistry, 2001. **40**(11): p. 3403-12.
6. Salhany JM., *Stilbenedisulfonate binding kinetics to band 3 (AE 1).* Blood cells, molecules & diseases, 2001. **27**(1): p. 127-134.
7. Lemieux MJ, Fischer SJ, Cherney MM, Bateman KS, James MN., *The crystal structure of the rhomboid peptidase from Haemophilus influenzae provides insight into intramembrane proteolysis.* Proceedings of National Academy of Sciences United States of America, 2007. **104**(3): p. 750-4.
8. Laue, T.M. and W.F. Stafford, 3rd, *Modern applications of analytical ultracentrifugation.* Annual Reviews in Biophysics and Biomolecular Structure, 1999. **28**: p. 75-100.
9. Johnson ML, Correia JJ, Yphantis DA, Halvorson HR., *Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques.* Biophysical Journal, 1981. **36**(3): p. 575-88.
10. Laue, T.M., et al., *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, ed. A.J.R. S.E. Harding, J.C. Horton, eds1991, Cambridge, UK: R. Soc. Chem. 90-125.
11. Reiss-Husson, F., *Crystallization of membrane proteins.* Crystallization of nucleic acids and proteins, ed. A.D.a.R. Geige1991, Oxford: IRL Press. pp. 175–193.
12. Avery, O.T., C.M. Macleod, and M. McCarty, *Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types.* The Journal of experimental medicine, 1944. **79**(2): p. 137-58.
13. Lemieux MJ, Song J, Kim MJ, Huang Y, Villa A, Auer M, Li XD, Wang DN., *Three-dimensional crystallization of the Escherichia coli glycerol-3-phosphate transporter.* Protein Science, 2003. **12**(12): p. 2748-56.

14. Friesen, R.H., J. Knol, and B. Poolman., *Quaternary structure of the lactose transport protein in the detergent-solubilized and membrane-reconstituted state.* The Journal of biological chemistry, 2000. **275**(51): p. 40658.
15. Lemberg, M.K. and M. Freeman, *Functional and evolutionary implications of enhanced genomic analysis of rhomboid intramembrane proteases.* Genome Research, 2007. **17**(11): p. 1634-46.
16. Brooks CL, Lazareno-Saez C, Lamoureux JS, Mak MW, Lemieux MJ., *Insights into substrate gating in H. influenzae rhomboid..* Journal of molecular biology, 2011. **407**(5): p. 687-697.
17. Strisovsky, K., H.J. Sharpe, and M. Freeman, *Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates.* Molecular Cell, 2009. **36**(6): p. 1048-59.
18. Wang, Y., Y. Zhang, and Y. Ha, *Crystal structure of a rhomboid family intramembrane protease.* Nature, 2006. **444**(7116): p. 179-80.
19. Wu Z, Yan N, Feng L, Oberstein A, Yan H, Baker RP, Gu L, Jeffrey PD, Urban S, Shi Y., *Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry.* Nature Structure and Molecular Biology, 2006. **13**(12): p. 1084-91.
20. Ben-Shem, A., D. Fass, and E. Bibi, *Structural basis for intramembrane proteolysis by rhomboid serine proteases.* Proceedings of National Academy of Sciences United States of America, 2007. **104**(2): p. 462-6.
21. Vinothkumar, K.R., *Structure of rhomboid protease in a lipid environment.* Journal of Molecular Biology, 2011**407**(2): p. 232-47.
22. Vinothkumar KR, Strisovsky K, Andreeva A, Christova Y, Verhelst S, Freeman M., *The structural basis for catalysis and substrate specificity of a rhomboid protease.* The EMBO Journal, 2011. **29**(22): p. 3797-809.
23. Xue Y, Chowdhury S, Liu X, Akiyama Y, Ellman J, Ha Y., *Conformational Change in Rhomboid Protease GlpG Induced by Inhibitor Binding to Its S′ Subsites.* Biochemistry, 2012. **51**(18): p. 3723-31.
24. Padan E, Kozachkov L, Herz K, Rimon A., *NhaA crystal structure: functional-structural insights.* The Journal of experimental biology, 2009. **212**(Pt 11): p. 1593-603.
25. Hunte C, Screpanti E, Venturi M, Rimon A, Padan E, Michel H., *Structure of a Na+/H+ antiporter and insights into mechanism of action and regulation by pH.* Nature, 2005. **435**(7046): p. 1197-202.
26. Hilger D, Jung H, Padan E, Wegener C, Vogel KP, Steinhoff HJ, Jeschke G., *Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na+/H+ antiporter of E. coli.* Biophysical journal, 2005. **89**(2): p. 1328-38.
27. Hilger D, Polyhach Y, Padan E, Jung H, Jeschke G., *High-resolution structure of a Na+/H+ antiporter dimer obtained by pulsed electron paramagnetic resonance distance measurements.* Biophysical Journal, 2007. **93**(10): p. 3675-83.

28. Rimon, A., T. Tzubery, and E. Padan, *Monomers of the NhaA Na+/H+ antiporter of Escherichia coli are fully functional yet dimers are beneficial under extreme stress conditions at alkaline pH in the presence of Na+ or Li+.* The Journal of biological chemistry, 2007. **282**(37): p. 26810-21.
29. Maegawa, S., K. Ito, and Y. Akiyama, *Proteolytic action of GlpG, a rhomboid protease in the Escherichia coli cytoplasmic membrane.* Biochemistry, 2005. **44**(41): p. 13543-52.
30. Lei, X., et al., *Soluble oligomers of the intramembrane serine protease YqgP are catalytically active in the absence of detergents.* Biochemistry, 2008. **47**(46): p. 11920-9.

Appendix: Supplementary Figures

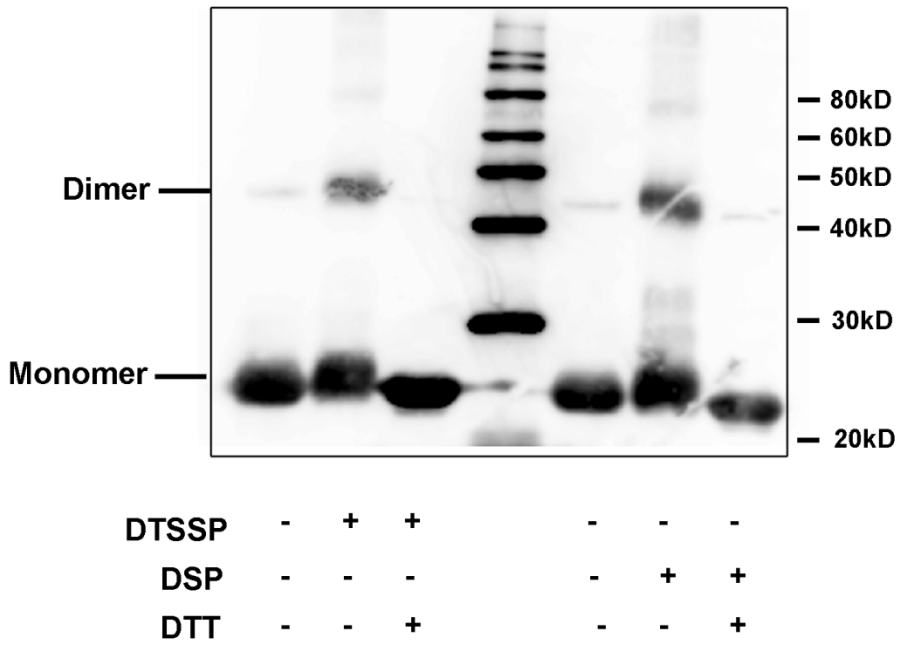


Fig. 2.3.3. Crosslinking of *hi*GlpG in detergent solution detected by Western blot using His-Probe.

His-Probe detection of detergent solubilized *hi*GlpG 10µg (20µl) aliquots incubated with 1mM DTSSP and DSP crosslinking agents, with and without the reducing agent DTT. The arrows show the position of monomer and dimer at approximately 23kDa and 45kDa respectively.

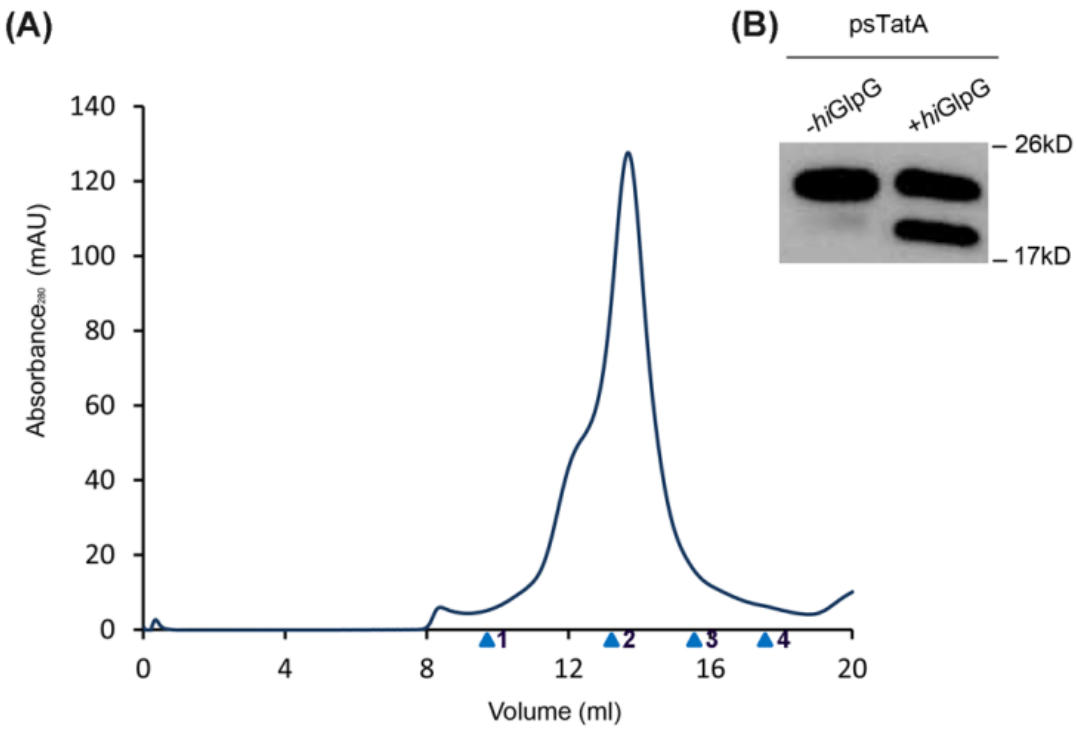


Fig. 2.3.4. Gel filtration and functional assay of *hi*GlpG.

A. Gel Filtration of *hi*GlpG in 0.1% DDM. Approximately 200µg of Ni-NTA purified *hi*GlpG in DDM was subjected to gel filtration onto a Hiload Superdex 200 10/300 column (GE Healthcare, USA) containing 50mM Tris, 200mM NaCl and 5% glycerol, pH 8.0 supplemented with 0.1% DDM. Vo- void volume, Vt-total column volume (24ml). Arrowheads represent the standard proteins from left to right: 1. thyroglobulin (MW 670k; Stokes radius 85 Å); 2. IgG (MW 158k; Stokes radius 55 Å); 3. ovalbumin; (MW 44k; Stokes radius 30.5 Å); and 4. myoglobin (MW 17k, Stokes radius 20.7 Å)

B. Rhomboid cleavage activity on *ps*TatA. SDS-PAGE demonstrating *P. stuartii* TatA (psTatA) substrate cleavage by *hi*GlpG in 0.1% DDM. Samples run include a control on the left panel. Molecular mass markers are reported in kDa on the left-hand side.

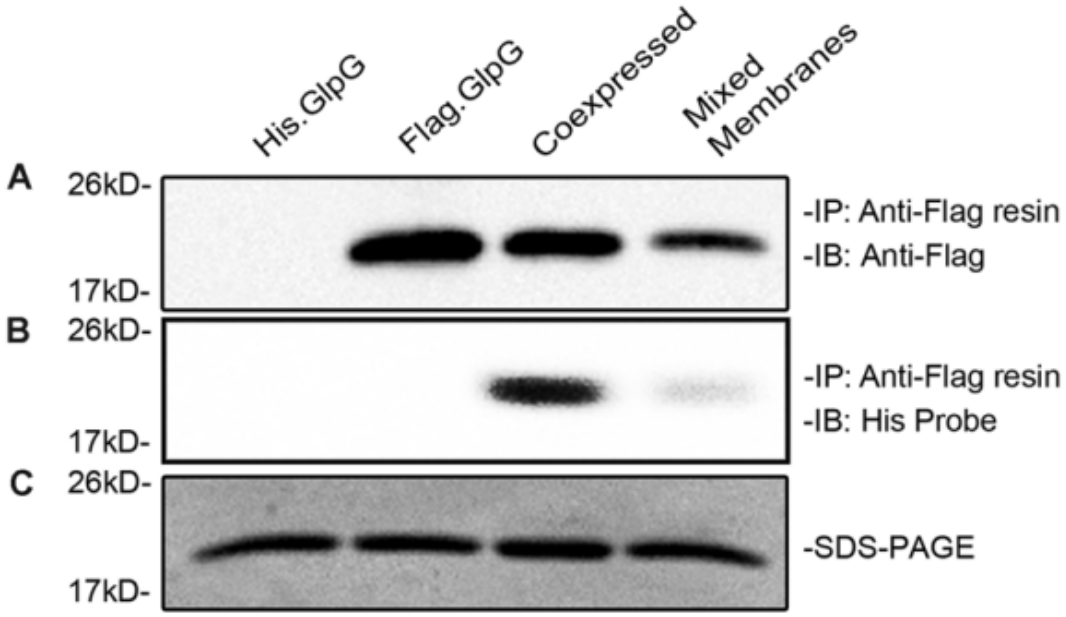


Fig. 2.3.5. Pull down assay with co-expressed His and Flag-tagged *hi*GlpG shows in vivo association.

Co-immunoprecipitation assay of *hi*GlpG molecules bearing two different immunological epitopes are shown to validate the formation of hiGlpG dimers within the membrane bilayer. *hi*GlpG-Flag and *hi*GlpG-His were expressed either independently, or coexpressed. In addition for a control, *hi*GlpG-Flag and *hi*GlpG-His membrane fractions were mixed prior to immunoprecipitation. Upon Anti-Flag immunoprecipitation, each fraction was separated on a 12% SDS-PAGE gel and analysed by immunoblotting with either (A) Anti-FLAG or (B) His-Probe. (C) A Coomassie stained gel with the four different membrane fractions is shown as a loading control to ensure equal amounts of protein was added to the resin. Molecular mass markers are reported in kDa on the left-hand side.

Chapter 3: Functional studies of E. coli rhomboid protease ecGlpG

Chapter 3: Functional studies of *ec*GlpG

Contributions:

Mass Spectrometry services by provided by Jack Moore, Institute for Biomolecular Design, University of Alberta.

Chapter 3: Introduction

As described in Chapter 1, rhomboid proteases are implicated in many functions including growth factor signalling, parasitic invasion and the activation of twin arginine translocase A (TatA) to release a quorum sensing factor [3]. However, the function of *E. coli* rhomboid *ec*GlpG remains unknown. It has been shown previously that *ec*GlpG could cleave the heterologous Spitz and Gurken proteins [4]. *ec*GlpG is also known to cleave LacY and MdfA proteins suggesting its role in degrading misfolded proteins [2, 5, 6]. A recognition motif has been identified for all rhomboid protease substrates, in which the residues flanking the scissile bond at positions P4, P1 and P2' were found to be important for substrate binding and cleavage [1]. Literature also describes the preference of "a hydrophilic region encompassing the cleavage site and helix-destabilizing residues in the downstream hydrophobic region" for *ec*GlpG substrates [8].

In this chapter, the potential of TatA in *E. coli* (ecTatA) as a substrate for *ec*GlpG is evaluated. Combinations of residues that followed the consensus sequence in TatA, described by Strisovsky et al. [1], were analyzed. Using this recognition motif, a potential cleavage site consisting of seven amino acids in ecTatA was introduced into pET21a-C100-Flag background replacing the seven amino acids in its TMD and assessed for *in vitro* cleavage by *ec*GlpG. Simultaneously, full length ecTatA was also assessed for cleavage. In a separate set of experiments, a substrate trapping strategy using affinity

pull-down was developed whereby solubilized membranes of *E. coli* were applied to a Ni-NTA resin column bound with *ec*GlpG to capture substrates of *E .coli* rhomboid. A modified approach to the affinity pull-down assay was also developed by using Anti-c-Myc agarose gel (Sigma-Aldrich, USA) instead of Ni-NTA resin in small scale volumes. Overall, this chapter describes the cleavage assays performed with ecTatA and designing strategies to trap substrates bound to *ec*GlpG.

3.1 Materials and Methods

3.1.1 Materials

All materials used for these experiments are listed in Chapter 2.1

3.2 Experimental procedures

3.2.1 Preparation of plasmids

Plasmids: pET21a-C100-*ps*Tat-Flag designed in our lab [9] was used as a template to generate pET21a-C100-*ec*Tat-Flag using the QuikChange Lightning mutagenesis kit (Stratagene, USA) to replace seven amino acids of psTatA to *ec*TatA (Figure 3.2.1). These specific seven amino acids observed the consensus sequence predicted by Strisovsky *et al*. [1] with the exception of one amino acid at P4 position (Figure 3.1). Primers for mutation and parameters for mutagenesis are listed in Table 3.2.1.

Full length ecTatA was kindly provided by Dr. Joel Weiner's group (University of Alberta) in a pMS119EH vector carrying a His epitope at the C terminus. Primers were designed to replace the His epitope with a Flag tag (Table 3.2.1). pBAD.*ec*GlpG was used as *E. coli* rhomboid expression plasmid described in section 2.2.1.

3.2.2 Protein Expression and purification

E. coli rhomboid *ec*GlpG was purified as described in section 2.2.2. In addition, detergent solubilized membrane fractions from *E. coli* Top10 cells (containing empty pBAD vector) were used for substrate trapping experiments. These solubilized membrane fractions were collected after ultracentrifugation step using the same protocol mentioned above.

For expression of TatA, vectors (pET21a-C100*ps*Tat-Flag, pET21a-C100*ec*Tat-Flag and pMS119EH-ecTatA-Flag) were transformed into competent *E.coli* BL21(DE3) cells and grown overnight at 37°C in LB media supplemented with ampicillin (100µg/ml). Large scale cultures were grown with overnight cultures as inoculum. When the OD_{600} reached 1.0, the cultures were induced with 1M IPTG and grown for 3 h. Purification of the proteins was performed as described in section 2.2.3. The eluted proteins were checked on SDS-PAGE for molecular weight and the protein concentrations were measured using Bicinchonic Acid (BCA) Protein Assay kit.

3.2.3 *In vitro* cleavage assay

C100ecTat-Flag and full length ecTatA-Flag were assayed for substrate cleavage using *ec*GlpG. This assay was carried based on the protocols previously described by [9, 10]. Briefly, 500ng of C100.ecTatA. Flag substrate was titrated with different concentrations of rhomboid (5µg, 10µg and 15µg), while full-length ecTatA was used in increasing concentration

(5µg, 10µg and 15µg) with the enzyme concentration maintained at 500ng. DDM detergent was brought to a final concentration of 0.1% and cleavage buffer (50 mM Tris, pH-7.5 and 150mM NaCl) was added in a total reaction volume of 20µl. The control contained 500ng of psTatA substrate, 15µg of rhomboid, DDM (final 0.1%) and cleavage buffer. The reaction was carried out 37° C for 2h and then stopped by adding 7µl of 4X SDS sample buffer. 20µl of the samples was resolved by 4%/16% SDS PAGE and transferred to a Hybond- P PVDF membrane (GE Healthcare, USA) for Western blot analysis. Blots were probed with mouse anti-flag primary antibody (1:10,000 dilution; Sigma-Aldrich, USA) for 30 min followed by secondary rabbit anti-mouse antibody conjugated to horseradish peroxidase (1: 40000 dilution) for 30 min. The bands in the blot were detected using ECL Plus western blot detection system (GE Healthcare, USA) and visualized by ImageQuant LAS 4000 (GE Healthcare, USA).

3.2.4 Substrate trapping using affinity pull-down

Two solubilized membrane fractions expressing *ec*GlpG were incubated with 0.5 ml Ni-NTA resin (Qiagen, Ontario, Canada) each for 2 h at 4° C. After incubation, the resins were transferred into two columns and the flowthrough from individual columns were collected. Detergent solubilized membrane fraction from Top10 was added to one column and incubated for 15 minutes at 4°C while the other served as control. The columns were washed with 20 column volumes (CV) of 50mM Tris, 300mM NaCl, 30mM

imidazole, 20 % glycerol, 0.1 % DDM pH 8.0 followed by 20 CV of the above stated buffer with 35 mM imidazole. Protein fractions were eluted in a step-wise manner with 3 times of 2 CV of the above described buffer containing 250, 500 and 1000mM Imidazole. 10 µl of the eluted protein samples were run on 4%/12% SDS-PAGE and protein bands in the test sample were compared to the control sample. Distinctly different bands that were found in the test sample but not in the control sample were cut and sent for Mass - spectrophotometry identification.

3.2.5 Co-immunoprecipitation

An improved approach to the substrate trapping experiment was developed to identify substrates for *ec*GlpG. As Ni-NTA resin allowed high nonspecific binding of other proteins, the affinity pull-down assay was modified with Anti-c-Myc agarose gel (Sigma-Aldrich, USA). Preliminary co-immunoprecipitation studies were carried using the protocol supplied by Sigma-Aldrich. Briefly, two solubilized membrane fractions expressing *ec*GlpG were added to 25 µl resin each in a 1.5ml tube. The resins were allowed to settle and then centrifuged at 9300 rpm (8,000 *g*) in a microcentrifuge to discard the liquid. The resins were then washed four times with 500 µl of 1X PBS plus 0.1% DDM. Further steps were performed with tubes incubated at 4° C. 250 µl of homogenized membrane fractions (from Top10 cells containing empty pBAD vector) were added to the tubes. The samples were incubated on a shaker at 4° C for 1h, washed four times

with 1X PBS with 0.1% DDM and centrifuged. The supernatants were discarded. 12.5 µl of 4X SDS sample buffer was added to each tube and vortexed. The tubes were then centrifuged and the supernatants were aspirated carefully avoiding the agarose. The supernatants were further divided into two tubes where one tube was incubated at 95° C for 15 minutes to denature the soluble proteins and the other tube was incubated at room temperature. 10 µl of the samples were loaded onto a 4%/12% SDS-PAGE and analyzed.

3.2.6 SDS-PAGE and Western Blotting

SDS- PAGE and Western blotting were followed as described in section 2.2.9.

3.3 Results

3.3.1. *In vitro* cleavage assay suggests that *ec*TatA is not a physiological substrate for *ec*GlpG

Many rhomboid proteases have been predicted to rely on conserved recognition motifs within their substrates: they require small residues at P1 and large, hydrophobic residues at P4 and P2' (Figure 3.2.1) [1]. This experiment verified the ability of ecTatA to serve as a potential substrate for *ec*GlpG using the predicted sequence motif. ecTatA contained 7 amino acids "MGGISIW" at the N terminus of its transmembrane domain, that when compared with the sequence specific recognition motif, differed only in the P4 position (Figure 3.2.1). We cloned the seven amino acids of ecTatA into a chimeric substrate, C100psTatAFlag that has been previously used for the study of rhomboid cleavage *in vitro* [9, 10], in which the seven amino acids of psTatA were replaced with "MGGISIW" sequence of ecTatA.

Two protein sequences of ecTatA were found in the protein database. While the shorter sequence was about 89 amino acids (starting with "MGGISIW"), the longer sequence had 103 amino acids, with a small N terminal extension (Figure 3.3.1). Alternately, we also tested the complete protein sequence of ecTatA that contained the N- terminal extension to test if *ec*GlpG cleaved the extension or if there was a different recognition motif preference for *ec*GlpG, other than the mentioned sequence motif.

In the cleavage assay, purified C100ecTatAFlag were titrated with increasing concentrations of *ec*GlpG, as the C100 protein was found to

aggregate. On the other hand, ecTatA was used in increasing concentrations for the full-length ecTatA with *ec*GlpG maintained at a constant concentration. Western blot with Anti-Flag antibody was used to detect the cleavage products (Figure 3.3.2 and Figure 3.3.3). As a positive control, C100psTatAFlag was used that has been previously shown to be cleaved by *ec*GlpG [10]. Western blot analysis showed the presence of a cleaved product in the positive control at around 15kDa. However, no cleavage was seen in C100.ecTatA.Flag (Figure 3.3.2) and full length ecTatA.Flag (Figure 3.3.3) (even when the blot was exposed to a longer time period). Taken together, these observations suggest that *E. coli* TatA is not cleaved by *ec*GlpG and may not be a physiological substrate for the rhomboid.

3.3.2 Substrate trapping using affinity pull-down

The substrate trapping method was tested as a preliminary strategy to develop a method for identifying novel proteins that interacted with *ec*GlpG. The *in vitro* pull down experiments were designed in such a way that *ec*GlpG bound Ni-NTA resin served as the bait while the interacting proteins (prey) were present in the solubilized membranes from Top10 (containing empty vector). A control containing only the rhomboid-Ni-NTA resin was included. Proteins were eluted and analysed on 12% gels. One distinct protein band was observed in the co-elution of *ec*GlpG plus solubilized membranes that were not seen in the control (Figure 3.3.4). Trypsin digestion of in-gel proteins revealed the possibility of three different soluble proteins

interacting with *ec*GlpG: 50S ribosomal protein L11 (16kDa), 50S ribosomal protein L6 (19kDa) and universal stress protein UP12 (18kDa). While this technique is not specific and may encourage non-interacting partners to bind to the resin, it would also be interesting to investigate further if these protein-pair interactions have a physiological significance.

3.3.3 Co-immunoprecipitaion using Anti-c-Myc agarose gel

An improved method of substrate trapping was developed using Anti-c-Myc agarose gel, as the Ni-NTA resin allowed high nonspecific binding of other proteins. The experiment was carried performed in small scale volume and used only 25µl of the Myc resin. A similar procedure was performed where the solubilized membranes expressing *ec*GlpG were bound to the Myc resin and homogenized membrane fractions were added to the resin bound *ec*GlpG. The samples were eluted and split into half. One aliquot was subjected to boiling at 95°C for 5 minutes to assess if contaminating soluble proteins were removed while the other aliquot was incubated at room temperature. On the SDS-PAGE, *ec*GlpG was detected near 34kDa (Figure 3.3.5). In contrast to the Ni-NTA resin, fewer contaminants were observed on the gel. However, *ec*GlpG appeared faint on the gel, suggesting that either the resin or the solubilized membrane expressing *ec*GlpG were used in low volumes for the assay. This was designed as a preliminary experiment to standardize the volume of resin used. While further experiments have not

been performed with this technique, this improvised technique can serve as a useful method to study *ec*GlpG- substrate interactions.

3.4 Discussion

In this chapter, we describe the functional characterization of *ec*GlpG. *ec*GlpG has been shown to cleave substrates from other organisms [10] and a few chimeric substrates [8], but neither its physiological substrate nor its biological function in *E. coli* have been defined. Work described in this chapter attempted in identifying substrates through two scenarios: 1. *In vitro* cleavage of TatA and 2. Designing strategies to identify the network of rhomboid- substrate interactions.

Currently, the only identified prokaryotic rhomboid substrate is TatA (Twin arginine translocase A) from *P. stuartii*. TatA contains an additional 7 amino acid extension at the N-terminus, which is cleaved by *P. stuartii* rhomboid, AarA to render it functional [3]. Two sequences of *E. coli* TatA were found in the protein database, one with a protein sequence of 89 amino acids and another with a slightly longer sequence (Full length ecTatA -103 amino acids) (figure 3.3.1). The predication software for secondary structures of the full length ecTatA revealed the presence of a similar N-terminal extension as observed in *P. stuartii* TatA [7]. Led by this, we decided to investigate the proteolytic processing of full length ecTatA by *ec*GlpG.

Also when we compared the sequence of ecTatA to the sequence specific recognition motif predicted by Strisvosky et al., we identified only

one combination of residues that closely resembled the recognition motif; "MGGISIW" at the N terminus of its transmembrane domain of ecTatA that when compared with the predicted recognition motif differed only in the P4 position. Increased concentrations of enzymes were used for the cleavage assay with the chimeric substrate, as we found that this protein formed large aggregates. On the contrary, different titrations of full length ecTatA were performed maintaining the enzyme concentration constant. We did not observe any cleaved product with either of the two substrates used, even when the blot was over-exposed. Therefore, we conclude that TatA in *E. coli* is not a physiological substrate for *ec*GlpG.

The *E. coli* rhomboid protease, glpG is encoded in Glycerol-3 phosphate regulon (glpEGR) which contains glpE coding for the thiosulfate cyanide sulfurtransferase and glpR acting as a receptor under glycerol deprivation conditions [11]. However the chromosomal context of glpG in this regulon does not direct to a physiological function or indicate any substrate in *E. coli* [12, 13]. As an initial attempt to identify rhomboid-protein interactions, we performed affinity pull-down assays using purified *ec*GlpG. Affinity pull-down assays have been performed for many membrane proteins. A recent study reports the detection of γ- secretase complex components using affinity purification [14, 15]. Microsomal membranes prepared from rat brain were incubated with a γ-secretase inhibitor coupled to biotin. By pulldown using streptavidin, the bound proteins were eluted and detected using LC-MS/MS analysis. Two proteins associated to γ-

secretase in rat brain were identified, suggesting that this method can be used for identification of protein-protein interactions.

Purified *ec*GlpG and added solubilized membranes onto the rhomboid bound Ni-NTA resin. One distinct band observed on the SDS-PAGE was detected by mass-spectrometry after affinity pull-down. Three proteins were identified: 50S ribosomal protein L11 (16kDa), 50S ribosomal protein L6 (19kDa) and universal stress protein UP12 (18kDa). Ribosomal proteins L11 and L6 are soluble proteins that bind to 23 S ribosomal RNA governing translational fidelity. UP12, universal stress protein 12 are expressed in under heat shock and stress conditions [16]. While some of these identifications may be contaminating proteins, as for example, the ribosomal proteins L6 and L11 (as they are localized in the ribosomes), UP12 may be an interesting candidate that may function in the regulation of *ec*GlpG during stress response. UP12 is characterized as a dimer in which the C- terminus is found to dimerize that contributes to structural stability [17]. Truncation of 6 amino acids in the C- terminus allows the access to an unknown protease resulting in the further cleavage of 18 amino acid residues at the C- terminus that prevents dimer formation. The dimeric species of UP12 has been proven to be important for its activity in osmotic stress conditions. It may be possible that the regulation of activity of this protein may be controlled by the proteolytic action of ecGlpG by binding to the cytoplasmic domain of ecGlpG and making it accessible to the active site. However, this requires further investigation.

The Ni-NTA resin gave a high background of contaminating proteins binding to the resin. This problem was reduced by using Anti-c-Myc agarose gel. However, the bands on the SDS-PAGE appeared faint and were hard to detect. One possible explanation is that low volumes of resin and solubilized membranes were used for this assay. This was designed as a preliminary experiment to standardize the volume of resin used. While further experiments have not been performed with this technique, it would be interesting to observe difference in substrate binding to *ec*GlpG, using the catalytic mutant S201A of *ec*GlpG along with intact *ec*GlpG expressed in ecGlpG knockout Top10 cells. This improvised technique can serve as a useful method to study *ec*GlpG- substrate interactions.

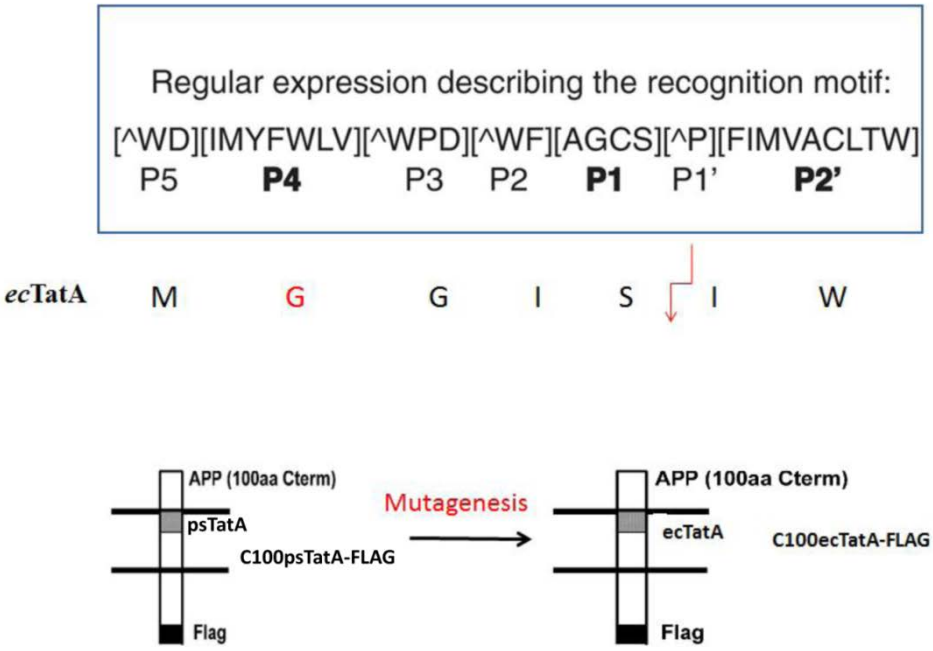


Figure 3.2.1: **Design of pET21.C100*ec*TatA.FLAG**: Combination of amino acid residues in *ec*TatA that followed the recognition motif described by Strisvosky *et. al.* [1, 2] were identified, with the exception of one amino acid at position P4. Site-directed mutagenesis of C100psTatA.FLAG to C100ecTatA.FLAG was done using the QuikChange Lightining mutagenesis kit (Stratagene) to replace seven amino acids of *ps*TatA to *ec*TatA. Red arrow indicates the possible site of cleavage.

Primer Design:

Primer Name	Primer Sequence
Insertion of 7aminoacids-Forward primer	5'- caggatatgaagttcatcatcaaaaattggtgttctttgcagaagatgtgggttcaaa Caaaatgggaggcattagtatcatagtgggcggtgttgtcatagcgacagtgatcgt 3
Insertion of 7aminoacids-Reverse primer	5'-acgatcactgtcgctatgacaacaccgcccactatgatactaatgcctcccatttt gtttgaacccacatcttctgcaaagaacaccaatttttgatgatgaacttcatatcctg-3'
Insertion of Flag tag in full length *ec*TatA- Forward primer	5'-aaagagcaggtcgaggattataaagatgacgacgataagtgacaccacc accaccac-3'
Insertion of Flag tag in full length *ec*TatA- Reverse primer	5'-gtggtggtggtggtgtcacttatcgtcgtcatctttataatcctcgacctgctcttt-3'

PCR conditions for Site- directed Mutagenesis:

Reagent	Volume (µl)
Plasmid template	1
10X reaction buffer	5
Forward primer (stock 100ng/ µl)	1.25
Reverse primer (stock 100ng/ µl)	1.25
dNTP mix	1
Quik Solution reagent	1.5
ddH_2O	38
QuikChange Lightning Enzyme	1

Cycling Parameters for Site-directed Mutagenesis:

Segments	Cycles	Temperature	Time
1	1	95ºC	2 min
2	18	95 ºC 60 ºC 68 ºC	20 sec 10 sec 2 min 15 sec
3	1	68 ºC	5 min

Table 3.2.1: Site-directed mutagenesis PCR for C100*ec*TatA.FLAG and Full Length *ec*TatA.FLAG:

The table above lists the parameters and protocol for generating two different tagged proteins, C100*ec*TatA.FLAG and Full Length *ec*TatA.FLAG.

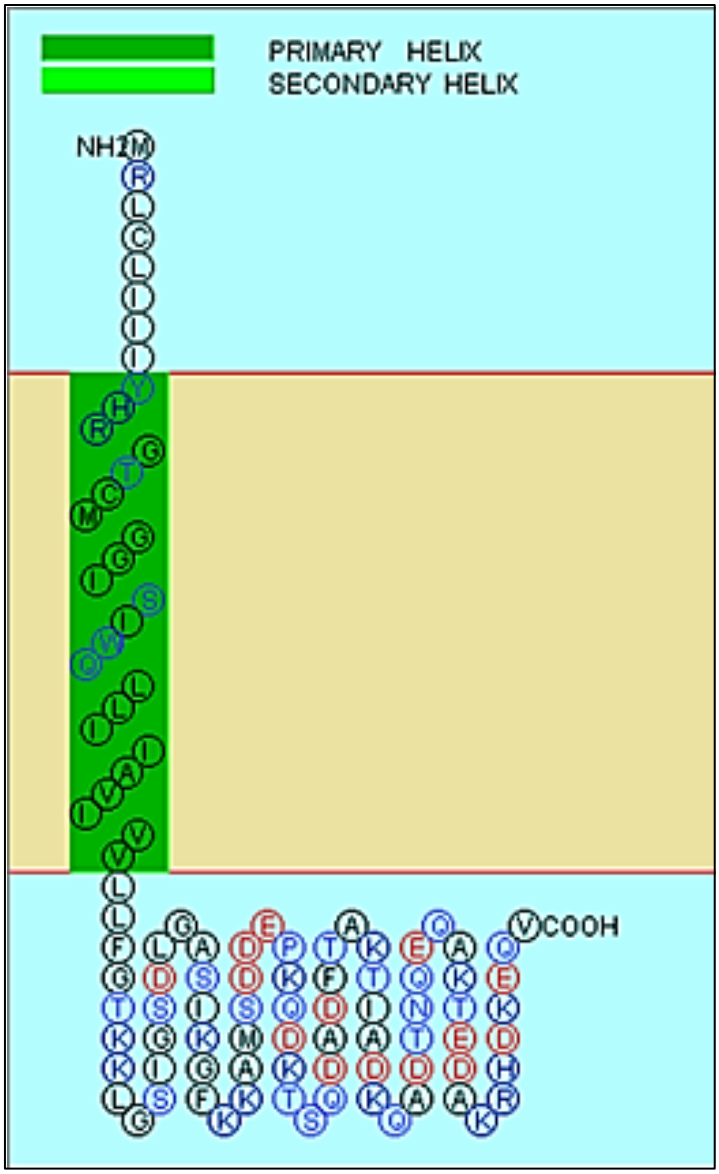


Figure 3.3.1. Secondary structure prediction of full length ecTatA using SOSUI [7]:

This protein sequence of ecTatA that contained the N- terminal extension was tested for *ec*GlpG cleavage.

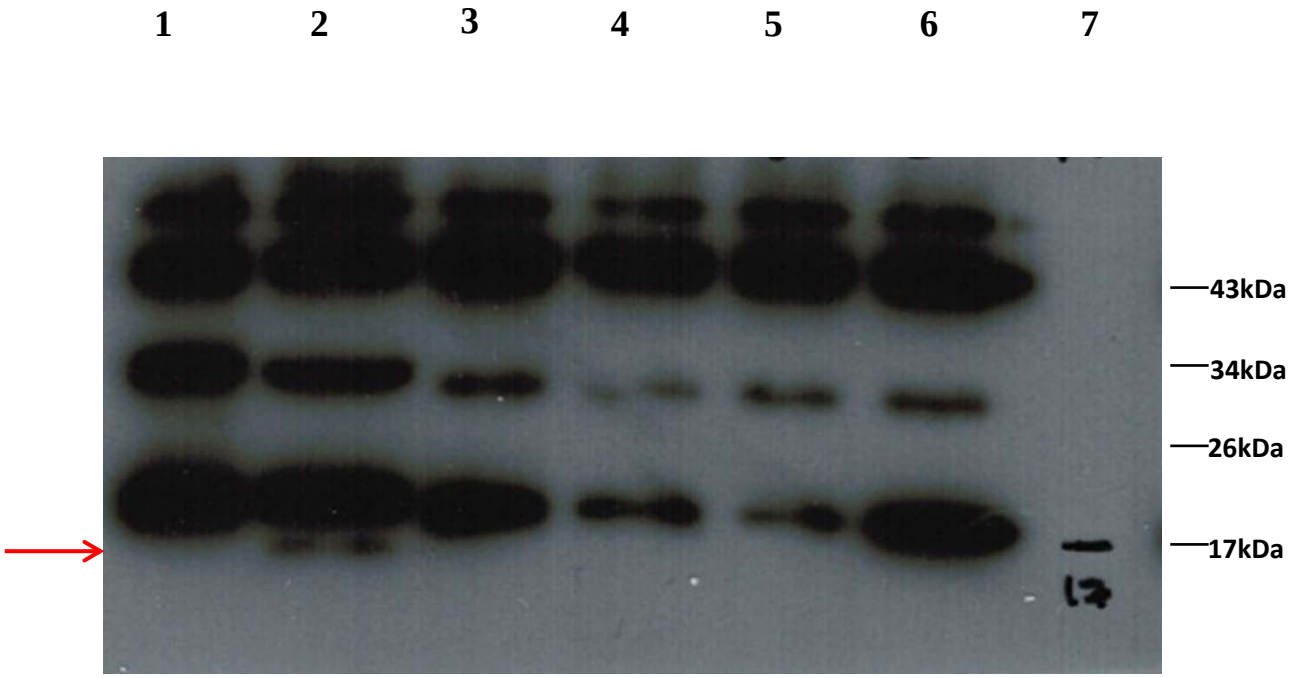


Figure 3.3.2. *In vitro* cleavage assay of C100ecTatA.Flag by *ec*GlpG detected by Western blot using Anti-Flag antibody:

Purified C100.ecTatAFlag was tested for *ec*GlpG cleavage under various titration volumes of the enzyme. No cleaved product was observed for ecTatA. On the other hand, positive control C100.psTatAFlag showed a cleaved product at around 15kDa (red arrow). Note the appearance of aggregated species, which is typical for the C100 protein. Molecular weights of the marker are shown on the right hand side.

Lane 1: C100 psTatA
Lane 2: C100 psTatA + *ec*GlpG 1:30
Lane 3: C100 ecTatA
Lane 4: C100 ecTatA + *ec*GlpG 10:1
Lane 5: C100 ecTatA + *ec*GlpG 20:1
Lane 6: C100 ecTatA + *ec*GlpG 30:1
Lane 7: Marker

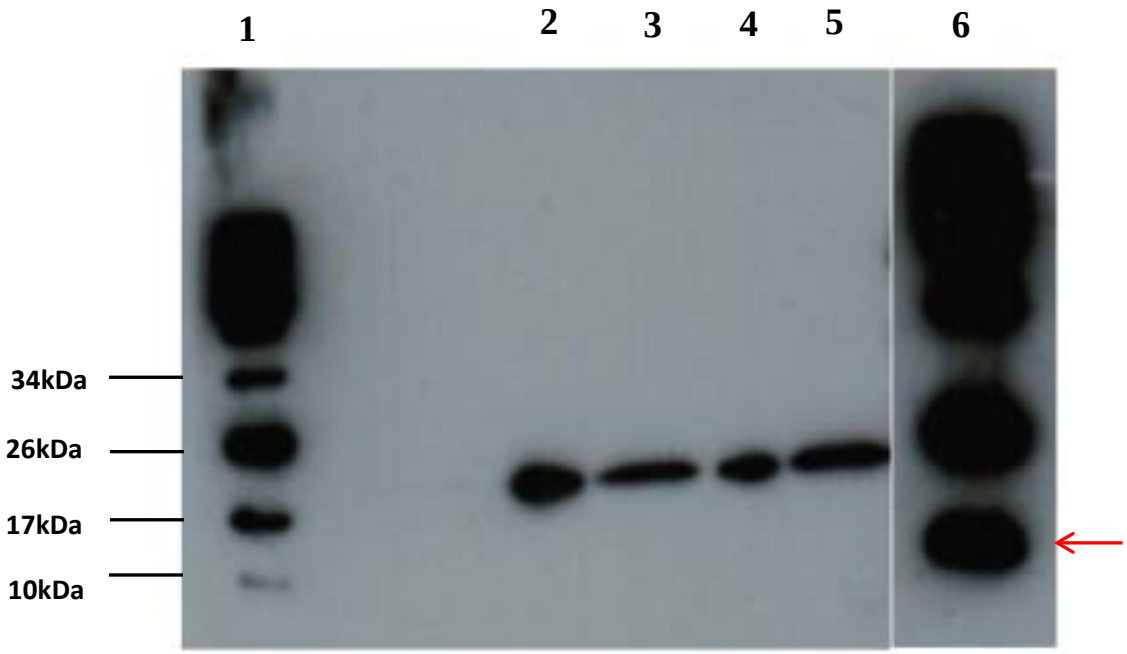


Figure 3.3.3. *In vitro* cleavage assay of Full length ecTatA.Flag by *ec*GlpG detected by Western blot using Anti-Flag antibody:

Purified ecTatAFlag was tested for *ec*GlpG cleavage under various titration volumes of the substrate. No cleaved product was observed for ecTatA. On the other hand, positive control C100.psTatAFlag showed a cleaved product at around 15kDa (red arrow).

Lane 1: Marker
Lane 2: Full ecTatA
Lane 3: Full ecTatA + *ec*GlpG 10:1
Lane 4: Full ecTatA + *ec*GlpG 20:1
Lane 5: Full ecTatA + *ec*GlpG 30:1
Lane 6: C100psTatA + *ec*GlpG 1:30

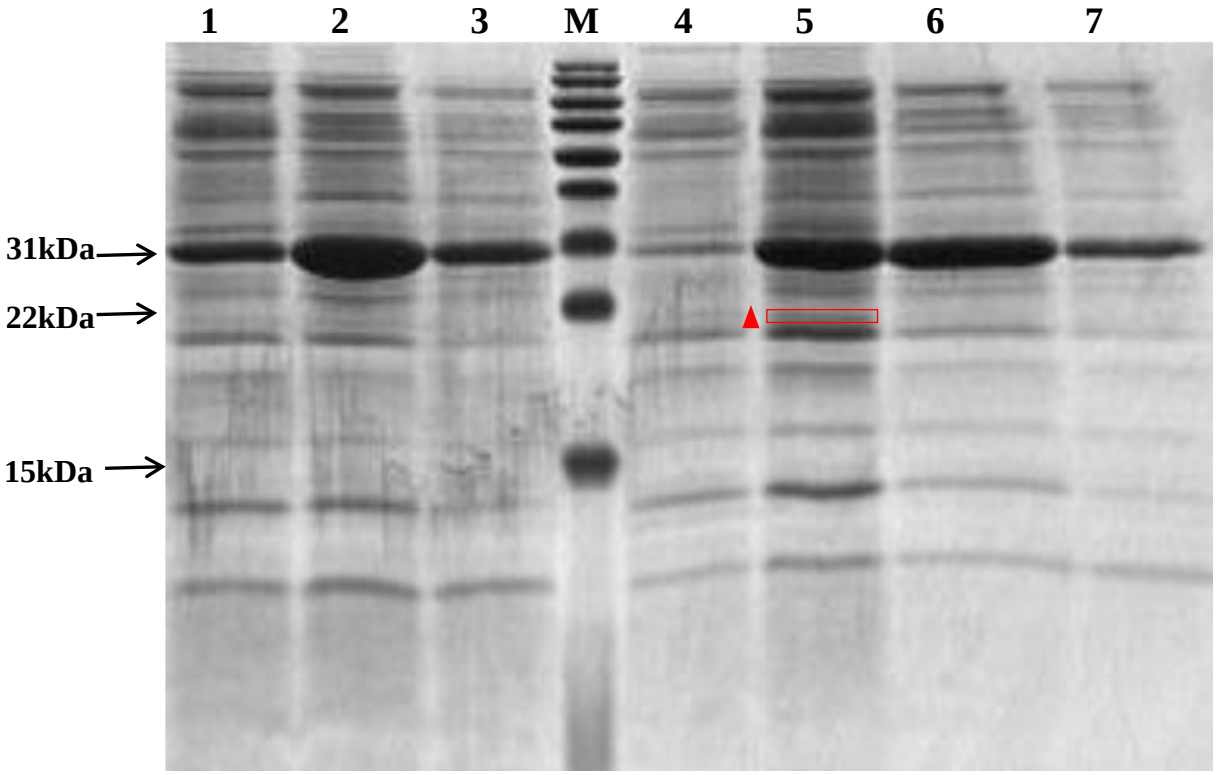


Figure 3.3.4. Preliminary assessment of substrate trapping using affinity pull-down:

Solubilized membranes from Top10 cells (containing empty pBAD vector) were added onto *ec*GlpG bound to Ni-NTA resin (Samples 4, 5, 6 and 7 are co-elutions with increasing imidazole). Negative control contained only the expressed *ec*GlpG bound to the N-NTA resin (samples 1, 2 and 3 are elutions with increasing imidazole). *ec*GlpG is observed at around 31kDa. The proteins band (marked in red box) is distinctly different and not observed in the control lane. This band was cut from the gel and analysed using Mass-spectrometry. Molecular weight markers (in kDa) are shown on the left hand side. M- Marker

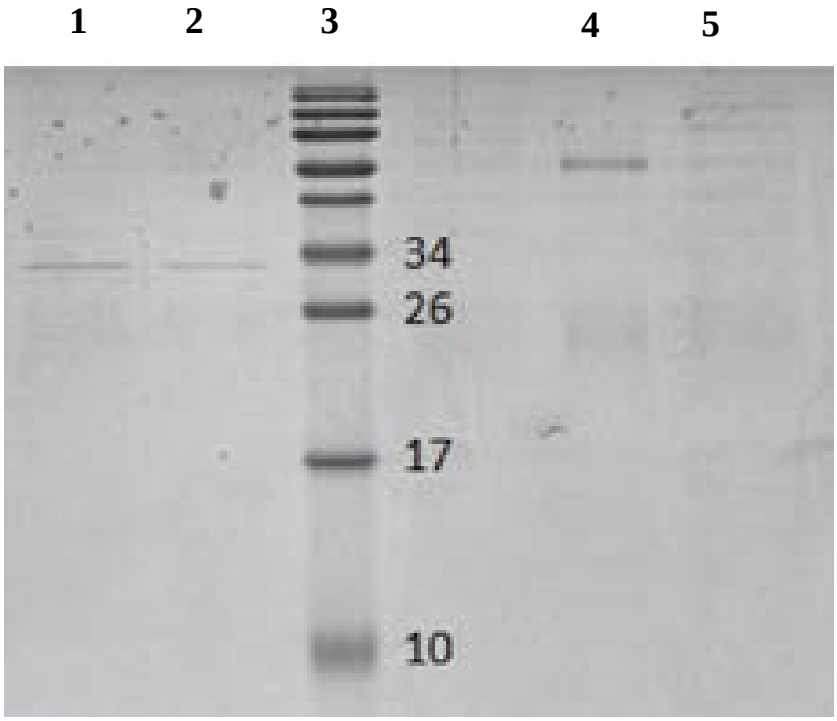


Figure 3.3.5. A modified approach to affinity pull-down: Co-immunoprecipitation using Anti-c-Myc agarose gel

Solubilized membranes from Top10 cells (containing empty pBAD vector) were added onto 250 µl of *ec*GlpG bound to 25 µl of Anti-c Myc agarose gel resin (samples 1 and 2). Negative control contained only the expressed *ec*GlpG bound to the Ni-NTA resin (samples 3 and 4). Samples 2 and 4 are boiled at 95 °C for 5 minutes prior to loading on SDS-PAGE. This was done to denature soluble proteins.

*ec*GlpG is observed at around 34kDa. The bands observed were very faint, suggesting that either the resin or the solubilized membranes were added in low volumes. Molecular weights of markers (in kDa) are indicated in the middle.

Chapter 3: References

1. Strisovsky, K., H.J. Sharpe, and M. Freeman, *Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates.* Mol Cell, 2009. **36**(6): p. 1048-59.
2. Erez, E. and E. Bibi, *Cleavage of a multispanning membrane protein by an intramembrane serine protease.* Biochemistry, 2009. **48**(51): p. 12314-22.
3. Stevenson, L.G., et al., *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**(3): p. 1003-8.
4. Urban, S., D. Schlieper, and M. Freeman, *Conservation of intramembrane proteolytic activity and substrate specificity in prokaryotic and eukaryotic rhomboids.* Curr Biol, 2002. **12**(17): p. 1507-12.
5. Maegawa, S., K. Ito, and Y. Akiyama, *Proteolytic action of GlpG, a rhomboid protease in the Escherichia coli cytoplasmic membrane.* Biochemistry, 2005. **44**(41): p. 13543-52.
6. Baker RP, Young K, Feng L, Shi Y, Urban S., *Enzymatic analysis of a rhomboid intramembrane protease implicates transmembrane helix 5 as the lateral substrate gate.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(20): p. 8257-62.
7. Hirokawa, T., S. Boon-Chieng, and S. Mitaku, *SOSUI: classification and secondary structure prediction system for membrane proteins.* Bioinformatics, 1998. **14**(4): p. 378-9.
8. Akiyama, Y. and S. Maegawa, *Sequence features of substrates required for cleavage by GlpG, an Escherichia coli rhomboid protease.* Mol Microbiol, 2007. **64**(4): p. 1028-37.
9. Brooks, C., et al., *Structural evidence for substrate gating indicates a new mechanism for a rhomboid intramembrane peptidase.* Journal of molecular biology, 2010. **407**(5): p. 687-697.
10. Urban, S. and M.S. Wolfe, *Reconstitution of intramembrane proteolysis in vitro reveals that pure rhomboid is sufficient for catalysis and specificity.* Proc Natl Acad Sci U S A, 2005. **102**(6): p. 1883-8.
11. Yang, B. and T.J. Larson, *Multiple promoters are responsible for transcription of the glpEGR operon of Escherichia coli K-12.* Biochim Biophys Acta, 1998. **1396**(1): p. 114-26.
12. Dalbey, R.E., P. Wang, and J.M. van Dijl, *Membrane proteases in the bacterial protein secretion and quality control pathway.* Microbiology and molecular biology reviews : MMBR, 2012. **76**(2): p. 311-30.
13. Butland G, Peregrín-Alvarez JM, Li J, Yang W, Yang X, Canadien V, Starostine A, Richards D, Beattie B, Krogan N, Davey M, Parkinson J, Greenblatt J, Emili A., *Interaction network containing conserved and essential protein complexes in Escherichia coli.* Nature, 2005. **433**(7025): p. 531-7.

14. Teranishi Y, Hur JY, Welander H, Frånberg J, Aoki M, Winblad B, Frykman S, Tjernberg LO., *Affinity pulldown of gamma-secretase and associated proteins from human and rat brain.* Journal of cellular and molecular medicine, 2010. **14**(11): p. 2675-86.
15. Hur JY, Teranishi Y, Kihara T, Yamamoto NG, Inoue M, Hosia W, Hashimoto M, Winblad B, Frykman S, Tjernberg LO., *Identification of novel gamma-secretase-associated proteins in detergent-resistant membranes from brain.* The Journal of biological chemistry, 2012. **287**(15): p. 11991-2005.
16. Bochkareva, E.S., A.S. Girshovich, and E. Bibi, *Identification and characterization of the Escherichia coli stress protein UP12, a putative in vivo substrate of GroEL.* European journal of biochemistry / FEBS, 2002. **269**(12): p. 3032-40.
17. Weber, A. and K. Jung, *Biochemical properties of UspG, a universal stress protein of Escherichia coli.* Biochemistry, 2006. **45**(6): p. 1620-8.

Chapter 4: Conclusions and future directions

4.1 Conclusions

The main objective of this thesis was to investigate the oligomeric state of rhomboid proteases. Previous gel filtration studies revealed that rhomboid proteases formed oligomers; however this was not extensively studied. Chapter 2 describes the first detailed oligomeric state characterization of three prokaryotic rhomboid proteases, *hi*GlpG, *ec*GlpG and YqgP. Analytical ultracentrifugation results indicated that prokaryotic rhomboids *hi*GlpG, *ec*GlpG and YqgP form monomers, dimers and tetramers with dimers being the predominant species. The membrane domain was investigated further to evaluate if it was responsible alone for dimerization. Further experiments were carried out only with *hi*GlpG as it represents the simplest rhomboid structure with the basic 6TM rhomboid core. Confirming the analytical ultracentrifugation results, gel filtration profile of the eluted *hi*GlpG revealed that the rhomboid protease was dimeric. Crosslinking studies with homobifunctional crosslinking agents also suggested that detergent solubilized *hi*GlpG existed as a dimeric species, apparent from the protein bands seen near 45 kDa on the SDS-PAGE. Additionally, *in vitro* functional assay with the chimeric substrate, C100TatAFlag demonstrated that the dimeric species was active. In order to assess if the dimers formed with in the membrane bilayer, a co-expression study of two different *hi*GlpG bearing two distinct tags was designed. Coimmunoprecipitation experiments indicated that the dimers were not easily separated and were formed in the membrane bilayer. Unfortunately, the function of the dimer *in vivo* is less

understood as we have not yet identified any conditions that could possibly disrupt the dimer and assess if the dimeric species is essential for function. Experiments looking at identifying the dimeric interface and disrupting this interface to form monomers will complete the story, thereby providing information on the physiological role of dimers of rhomboid proteases.

In Chapter 3, the mechanisms for identification of function of ecGlpG is described. We investigated the potential of *E. coli* TatA as a substrate for *ec*GlpG using the consensus recognition motif prediction reported by Strisvosky *et al.* [1]. *In vitro* cleavage assays of ecTatA performed with *ec*GlpG did not show any cleavage, however the positive control C100psTatAFlag showed a cleaved product band at around 15kDa with *ec*GlpG, suggesting that ecTatA may not be a physiological substrate for *E. coli*. We also designed strategies for identifying potential substrate of ecGlpG using affinity pull-down using Ni-NTA resin. We identified three proteins using mass spectrometry: 50S ribosomal protein L11, 50S ribosomal protein L6 and universal stress protein UP12. Though some of these interactions may be non-specific, for example, the ribosomal proteins L6 and L11 (as they are localized in the ribosomes), UP12 may be an interesting candidate that may function in the regulation of *ec*GlpG during stress response. This, however, requires further investigation. We also carried out affinity pull-down using Anti-c-Myc agarose gel as the Ni-NTA resin gave a high background of contaminating proteins. This was done as a preliminary assessment to measure the volume of resin and solubilized membrane to be used for further

experiments. Though further experiments have not been performed, this improvised method can serve as a useful technique to study *ec*GlpG-substrate interactions.

4.2 Future directions

Various biochemical approaches can be performed to shed more light into the role of dimeric rhomboid proteases within the cell (described in I, II and III). Also techniques are also described to improve characterizing the function of *ec*GlpG (IV and V).

I. The existence of dimeric species of rhomboid proteases has been established *in vitro*. However, it is yet unclear if the dimeric interface is essential for function of stability. As mentioned above, we have not been able to disrupt the dimer and examine if the functionality of the monomer. Mutating the crucial residues at this interface will help disrupt the dimer interface and the role of monomeric species can be assessed. Furthermore, NMR spectroscopy can help unravel this problem. NMR spectroscopy can be performed to experimentally identify the actual dimerization interface between two *hi*GlpG molecules in solution.

II. We hypothesize that Loop 1 which is on the opposite face of the substrate entry, may facilitate dimerization. Loop 1 comprises the conserved WR motif which may serve as a "hot spot" for dimerization. Many membrane proteins have been known to contain these residues at the dimeric

interface [2]. A suggested model would be the Loops 1 from two adjacent *hi*GlpG molecules align antiparallel to each other so that the tryptophan residues form hydrogen bonds with the arginine residues in the adjacent protein molecule. The importance of this motif in dimerization can be assessed by mutating these residues to alanine and observing the presence of dimeric species in gel filtration profile or crosslinking experiments.

III. It is also of great interest to understand if the dimer affects the rate of substrate cleavage. Currently, our lab is developing a kinetic assay to address this question. Also, it would be of importance to know if the lipids surrounding the rhomboid affect dimerization. Lipids modeled as phosphatidic acid were found in the *hi*GlpG crystal structure [3] and a recent structure presented the rhomboid protease in a bicelle environment [4]. It is possible that lipids may influence oligomerization and consequently affect the catalytic activity of rhomboid proteases. Molecular dynamic simulation experiments can help achieve this by presenting a picture of dynamics of conformational states of rhomboid proteases in lipid environments performed at physiological temperatures.

IV. Trapping of *ec*GlpG substrates using affinity pull-down by Anti-c Myc agarose gel proved to be a better technique than using Ni-NTA resin. However, further experiments have not been performed with this method. It would be interesting to observe difference in substrate binding to *ec*GlpG, using the catalytic mutant S201A of *ec*GlpG along with intact

*ec*GlpG expressed in ecGlpG knockout Top10 cells. This improvised technique can serve as a useful method to study *ec*GlpG- substrate interactions.

V. An alternate approach to the above mentioned technique is to express *ec*GlpG at natural levels so that the assembly of non-physiological complexes is avoided [5]. Tandem affinity purification (TAP) affinity tag allows the identification of physiological complexes in native concentrations. Mass spectrometry can then be used to identify all the possible interacting substrates. Our lab is currently investigating this by collaborating with Dr. Emili's lab at University of Toronto.

Chapter 4: References

1. Strisovsky, K., H.J. Sharpe, and M. Freeman, *Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates.* Mol Cell, 2009. **36**(6): p. 1048-59.
2. Bogan AA, T.K., *Anatomy of hot spots in protein interfaces.* Journal of molecular biology, 1998. **280**(1): p. 1-9.
3. Lemieux MJ, Fischer SJ, Cherney MM, Bateman KS, James MN., *The crystal structure of the rhomboid peptidase from Haemophilus influenzae provides insight into intramembrane proteolysis.* Proc Natl Acad Sci U S A, 2007. **104**(3): p. 750-4.
4. Vinothkumar, K.R., *Structure of rhomboid protease in a lipid environment.* Journal of molecular biology, 2011. **407**(2): p. 232-47.
5. Link, A.J., C. Weaver, and A. Farley, *Affinity capture of TAP-tagged protein complexes: affinity purification step 2.* Cold Spring Harbor protocols, 2011. **2011**(4): p. pdb prot5607.