Urban and Dickey *Genome Biology* 2011, **12**:231
http://genomebiology.com/2011/12/10/231

Genome **Biology**

REVIEW

The rhomboid protease family: a decade of progress on function and mechanism

Sinisa Urban* and Seth W Dickey

Summary

Rhomboid proteases are the largest family of enzymes that hydrolyze peptide bonds within the cell membrane. Although discovered to be serine proteases only a decade ago, rhomboid proteases are already considered to be the best understood intramembrane proteases. The presence of rhomboid proteins in all domains of life emphasizes their importance but makes their evolutionary history difficult to chart with confidence. Phylogenetics nevertheless offers three guiding principles for interpreting rhomboid function. The near ubiquity of rhomboid proteases across evolution suggests broad, organizational roles that are not directly essential for cell survival. Functions have been deciphered in only about a dozen organisms and fall into four general categories: initiating cell signaling in animals, facilitating bacterial quorum sensing, regulating mitochondrial homeostasis, and dismantling adhesion complexes of parasitic protozoa. Although in no organism has the full complement of rhomboid function yet been elucidated, links to devastating human disease are emerging rapidly, including to Parkinson's disease, type II diabetes, cancer, and bacterial and malaria infection. Rhomboid proteases are unlike most proteolytic enzymes, because they are membrane-immersed; understanding how the membrane immersion affects their function remains a key challenge.

Gene organization and evolutionary history

Proteolysis within the membrane was discovered in seemingly rare contexts nearly 15 years ago [1-3]. It is now widely appreciated that this fascinating regulatory paradigm permeates most areas of modern cell biology [4-7]. Of the three protease families that catalyze intramembrane proteolysis, rhomboid enzymes are the only family that were not discovered from the direct study of human disease. The name 'rhomboid' has its origin deep in the rich folklore of *Drosophila* genetics. Rhomboid emerged from the historic quest to identify all genes required to organize construction of a free-living organism from a single cell [8,9]. Because genes were named after the altered appearance of the mutant larval cuticle, the mis-shaped, rhombus-like head skeleton of the mutant embryo earned *rhomboid* its name. Mutating the growth factor that rhomboid activates yielded indistinguishable head-skeleton defects, and was named *spitz* ('pointed' in German).

The *rhomboid* gene was cloned and sequenced by Bier and colleagues in 1990, revealing a seven transmembrane (7TM) protein with no homology to any sequence known at the time [10]. The *spitz* sequence was more informative, encoding a clear epidermal growth factor (EGF)-like protein [11]. The fact that *rhomboid* mirrored *spitz* phenotypically [9], and encoded a seven TM protein, led to the proposal that it might be a serpentine receptor for Spitz signaling [11]. But as sequencing of genomes from diverse organisms began to reveal rhomboid homologs in every form of cellular life [12], it became clear that rhomboid proteins may be at the core of very diverse biological regulation.

Sequence analysis, however, yielded no clues about the underlying biochemical function of rhomboid proteins, and no other homologs were as well studied as *Drosophila rhomboid*. A decade of *Drosophila* genetics had, however, set the stage for a biochemical approach: *rhomboid* was definitively implicated as an upstream activator of Spitz in the signal-sending cell, providing a framework for analyzing its molecular function [12-15]. A focused analysis of Spitz activation eventually yielded four key pieces of the puzzle [16]: (i) substochiometric levels of Rhomboid triggered Spitz proteolysis, implying that Rhomboid acts enzymatically; (ii) proteolysis depends absolutely only on four Rhomboid residues, and their identity is consistent with serine protease catalysis; (iii) Spitz proteolysis is blocked only by serine protease inhibitors; and (iv) Spitz is cleaved within its TM segment at a depth similar to that of the putative rhomboid

*Correspondence: surban@jhmi.edu
Howard Hughes Medical Institute, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

BioMed Central © 2011 BioMed Central Ltd

catalytic serine. These pieces fit together into a model in which rhomboid acts as an intramembrane serine protease for Spitz [16], which was confirmed 4 years later by reconstituting cleavage with pure proteins [17-19].

The ability to study Spitz proteolysis as a direct test of rhomboid activity was used to determine that even distant bacterial homologs are functional intramembrane serine proteases [20]. Most bacterial species are now known to encode one rhomboid protease, while some encode two, and very few encode three [21]. Rhomboid proteases are also present in many if not most Archaea, but the greatest expansion occurred in multicellular organisms and some parasitic protozoa. Although the human, mouse and *Drosophila* genomes encode at least seven *rhomboid* genes, the largest number of *rhomboid* genes are encoded by plants (13 in *Arabidopsis*), which do not have EGF signaling [21-23]. In many of these diverse organisms, at least one rhomboid has directly been demonstrated to have proteolytic activity (Table 1) [16,20,23-26].

Rhomboid proteases are found in all branches of life, yet the sequence identity across all family members is strikingly low, around only 6% [12,21,22]. We suggest that this is not despite rhomboid proteases being so widespread but because of it. This divergence is exacerbated by their sequences being predominantly transmembrane and thus experiencing a different evolutionary pressure [27]. This has made phylogenetic analyses noisy, yielding few incontrovertible conclusions and inevitably fueling debate. Of particular intrigue is their evolutionary origin: rhomboid proteins have been argued to be perhaps the most widely distributed membrane proteins in nature [21] (Figure 1). This near ubiquity is instinctively viewed as evidence of an ancient enzyme family that evolved early [12]. Although this is likely if the last universal common ancestor already encoded several different rhomboid proteins, phylogenetic analysis has also raised the possibility of a different history in which rhomboid proteins are a later invention of bacteria that rapidly spread to most other organisms [21]. This scenario requires a controversial amount of horizontal gene transfer to populate all kingdoms of life. Currently the true nature of rhomboid phylogenetic history remains a point of inconsolable debate, but three observations serve as valuable guiding principles.

Characteristic structural features

The first organizing principle emerging from sequence analysis is the observation that rhomboid proteases come in three distinct topological flavors (Figure 2) [21]. The simplest consists of the 6TM core, which itself is the smallest catalytically active unit. This form predominates in bacteria, but is also represented, albeit more rarely, in eukaryotic organisms, including animals. To this basic unit eukaryotes add a seventh TM segment following the 6TM core (6+1TM form). Five of the seven *Drosophila*, human, and mouse rhomboid proteins are of this form. Analogous 7TM forms also occur in bacteria, but are rare. Lastly, a distinct form of 7TM rhomboid proteases exists in endosymbiotic organelles, adding the seventh TM preceding the 6TM core (1+6TM form). The best studied are those imported into mitochondria [28-30], although interest in plastid-resident rhomboid proteins has recently been sparked [31]. Although the sequence analysis is clear on these three topological distinctions, their functional relevance is unclear. The expectation is that they confer different biochemical properties, although current evidence, albeit limited, indicates that many bacterial 6TM forms and eukaryotic 6+1TM forms show similar activity against surrogate substrates, including Spitz [17,32,33].

Although protease activity has been reconstituted with both the 6TM and 6+1TM rhomboid forms *in vitro*, it is only a 6TM form called GlpG from *Escherichia coli* and *Haemophilus influenzae* that has proven amenable to structural analysis [34-37]. This major breakthrough - the first atomic-resolution structure of any intramembrane protease - not only confirmed that proteolysis is intramembrane and catalyzed by a serine protease apparatus, but revealed an unanticipated and complex architecture. Although a thorough description is beyond the scope of the current discussion (see [38] for a comprehensive review), two features are characteristic (Figure 3): although most TM helices are long and run roughly perpendicular to the membrane, the fourth TM segment runs slanted relative to the others and enters the center of the protein as an extended loop, converting to an α helix at the catalytic serine. More unexpected was the orientation of the long L1 loop connecting TMs 1 and 2, which forms a lateral hairpin that lies half submerged in the membrane. This feature, which has not been encountered before or since, has major structural implications and results in a highly asymmetric protein. It is assumed that the structure of the other rhomboid forms will be analogous, and recent modeling of the mitochondrial 1+6TM form on *E. coli* GlpG hints at an unanticipated level of similarity [39].

In addition to the number of TMs, two further variations provide potential for additional rhomboid diversity. First, in all three forms, the cytosolic amino termini are highly variable, ranging from large domains to being nonexistent. The implications, however, remain unclear, at least partly because achieving well-diffracting crystals required absence of this domain, making its relationship to the catalytic core speculative. On the simplest level, these domains may house sorting signals [40].

Secondly, rhomboid proteins are often encountered that clearly lack catalytic residues. These should be

Table 1. Known rhomboid protease substrates and functions across evolution

Rhomboid	Organism	Substrate	Function	References
Rho-1	*Drosophila*	Spitz, Keren	EGFR signaling (embryo, most developing tissues, adult)	[8-16,89]
Rho-2/Stet	*Drosophila*	Gurken,?	EGFR signaling in oocyte, germline	[89,90]
Rho-3/Ru	*Drosophila*	Spitz, Keren	EGFR signaling in embryo, eye disc	[12,89,91]
Rho-4	*Drosophila*	Spitz, Keren?	EGFR signaling?	[89]
Rho-7	*Drosophila*	DmOpa1-like DmPINK1	Mitochondrial fusion, mitophagy	[30,55]
CeROM1	*C. elegans*	LIN-3L	Amplifying EGFR signaling (vulva)	[47]
AarA	*P. stuartii*	TatA	Activating twin-arginine translocon, quorum sensing	[20,17,61-62,66]
GlpG	*E. coli*	?	?	[17-20,34-36, 77-78,83-85,88]
HiGlpG	*H. influenzae*	?	?	[37,81]
YqgP	*Bacillus subtilis*	?	?	[17,18,20]
PA3086	*Pseudomonas aeruginosa*	?	?	[17,18,20]
AqRho	*Aquifex aeolicus*	?	?	[17,20]
Pcp1	*Saccharomyces cerevisiae*	Mgm1, Ccp1	Mitochondrial membrane fusion, pre-sequence removal	[28,52-54]
PARL	*Homo sapiens Mus musculus*	PINK1, others	Downregulating mitophagy, crista remodeling, anti-apoptosis	[28,29, 56-60]
AtRBL2	*Arabidopsis thaliana*	?	?	[23]
RHBDL2	*H. sapiens*	TM, EphrinB1,2,3, EGF	Blood clotting? Cell migration? Cancer?	[24,48,87]
PfROM4	*P. falciparum*	EBA175, TRAP, other adhesins	Invasion	[26,69]
PfROM1	*Plasmodium* sp.	AMA1?	Invasion? Growth	[26,69,70]
TgROM1	*Toxoplasma gondii*	?	Growth	[25,40,68,75]
TgROM2	*T. gondii*	?	?	[25,40,68]
TgROM3	*T. gondii*	?	?	[25,68]
TgROM4	*Toxoplasma gondii*	AMA1, MIC2, MIC6	Invasion	[25,68,71,72]
TgROM5	*Toxoplasma gondii*	AMA1, MIC2, MIC6	Invasion	[25,26]
EhROM1	*Entamoeba histolytica*	Lectins	Phagocytosis, immune evasion	[73,74]

Only those rhomboid proteases whose proteolytic activity has been detected are included in the table. A question mark '?' denotes unknown function. EGFR, epidermal growth factor receptor.

considered rhomboid proteins but not rhomboid proteases. Two predominant clusters are a distinct 6+1TM form in animals, called iRhom proteins [22,41], and a 6TM form that is represented widely in eukaryotes by the Derlin proteins [42-45]. Both of these have been implicated in endoplasmic reticulum-associated degradation (ERAD). Derlins have clear sequence homology near the membrane-submerged L1 loop, but also less conspicuously along their entire length, and are thus likely to adopt a GlpG-like 6TM structure. Although clearly not proteolytic, their potential similarity to other aspects of the rhomboid protease mechanism should not be discounted at this early stage (but lie beyond the scope of this review).

Localization and function

The second guiding principle stems from the tremendous diversity of organisms that encode rhomboid enzymes. Since these include organisms that do not encode any known forms of cell-to-cell communication, sequence information implies that rhomboid proteins perform an ancient and fundamental role in cell biology. This function is not essential for cell survival, however, because several lineages are missing *rhomboid* genes entirely, presumably by gene loss [21]. Although defining the cellular functions of rhomboid proteases has proven a persistent challenge, focused investigations have succeeded in documenting the function of at least one rhomboid in nearly a dozen organisms (Table 1). These functions are usually regulated by substrate trafficking, and fall into four broad categories (Figure 4).

First, rhomboid proteases initiate animal cell signaling by releasing growth factors from the membrane. This function emerged from detailed genetic study of *Drosophila* development; rhomboid proteases are localized in the Golgi apparatus and act as the signal-generating component by cleaving Spitz to initiate the pathway in neighboring cells [16,46]. Although a role in regulating

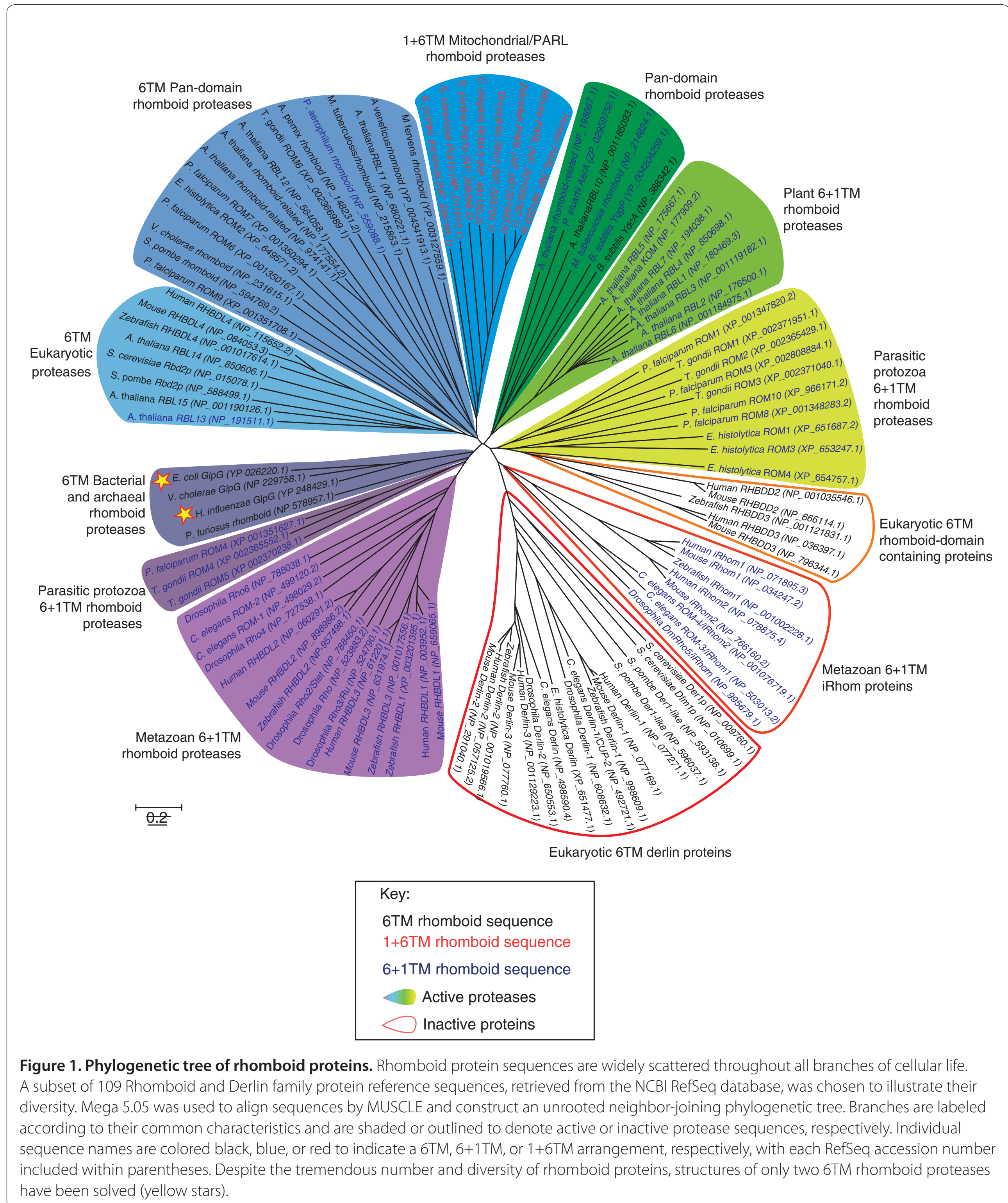

Figure 1. Phylogenetic tree of rhomboid proteins. Rhomboid protein sequences are widely scattered throughout all branches of cellular life. A subset of 109 Rhomboid and Derlin family protein reference sequences, retrieved from the NCBI RefSeq database, was chosen to illustrate their diversity. Mega 5.05 was used to align sequences by MUSCLE and construct an unrooted neighbor-joining phylogenetic tree. Branches are labeled according to their common characteristics and are shaded or outlined to denote active or inactive protease sequences, respectively. Individual sequence names are colored black, blue, or red to indicate a 6TM, 6+1TM, or 1+6TM arrangement, respectively, with each RefSeq accession number included within parentheses. Despite the tremendous number and diversity of rhomboid proteins, structures of only two 6TM rhomboid proteases have been solved (yellow stars).

EGF signaling is also seen in *Caenorhabditis elegans* vulval development, CeROM-1 has a surprisingly minor role as a target of EGF signaling that sets up a paracrine loop to amplify and spread the signal [47]. Even less is clear in mammals: recent investigations have localized rhomboid proteins to the secretory pathway and cell surface and begun to uncover increased *rhomboid* expression in cancer cells with potential links to growth

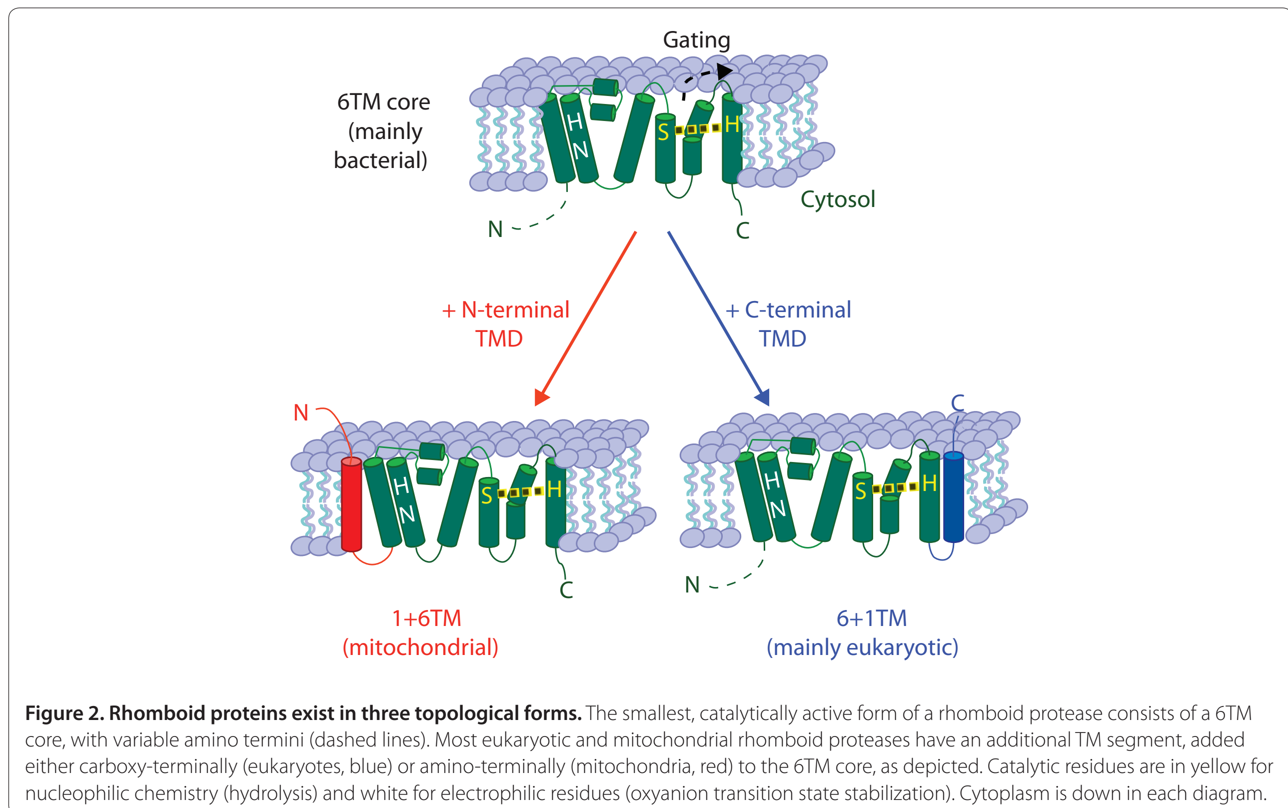

Figure 2. Rhomboid proteins exist in three topological forms. The smallest, catalytically active form of a rhomboid protease consists of a 6TM core, with variable amino termini (dashed lines). Most eukaryotic and mitochondrial rhomboid proteases have an additional TM segment, added either carboxy-terminally (eukaryotes, blue) or amino-terminally (mitochondria, red) to the 6TM core, as depicted. Catalytic residues are in yellow for nucleophilic chemistry (hydrolysis) and white for electrophilic residues (oxyanion transition state stabilization). Cytoplasm is down in each diagram.

factor signaling [24,48,49]. However, this is not limited to active rhomboid proteases; expression of the iRhom RHBDF1, which is localized in the endoplasmic reticulum in human epithelial cancer cells, increased secretion of the EGF ligand transforming growth factor-α [50]. Accordingly, RHBDF1 silencing decreased pathway activation through EGF receptor (EGFR), ERK and AKT phosphorylation, and limited tumor growth in mice [51]. The *Drosophila* homolog, however, was recently found to have the opposite effect of decreasing EGFR signaling by promoting the ERAD-mediated degradation of EGF ligands [41]. The basis of this remarkable discrepancy is currently unclear; knockout mouse studies are expected to provide clarity on the physiological roles of rhomboid proteins.

Recent studies have also placed the mitochondrial rhomboid protease at the nexus of key pathways that govern mitochondrial fusion, mitophagy and apoptosis. All mitochondrial rhomboid proteins are encoded in the nuclear genome, and imported into mitochondria. The main function of the mitochondrial rhomboid Pcp1 is to release the dynamin-like GTPase Mgm1 from the membrane [28,52,53]. Because Mgm1 is essential for mitochondrial fusion and Mgm1 cleavage occurs only in healthy mitochondria, this limits fusion to occurring between healthy organelles [54]. A similar function was described in *Drosophila* [30], but genetic interactions soon revealed further complexity in metazoans; the mitochondrial rhomboid DmRho-7 also participates in the Parkin/PINK1 pathway that malfunctions in Parkinson's disease [55]. It has recently become clear that the human mitochondrial rhomboid PARL cleaves PINK1 to suppress its ability to recruit the Parkin ubiquitin ligase onto mitochondria [56-58]. Without PARL cleavage, PINK1 accumulates in mitochondria and fails to be recruited properly to damaged mitochondria. A PARL knockout mouse suffers tremendous atrophy several months after birth resulting from malformed mitochondria and elevated apoptosis, although without mitochondrial fusion defects [29]. PARL has also been implicated in suppressing apoptosis in lymphocytes, potentially through a different substrate, High-temperature regulated A (HtrA, also called Omi) [59]. Intriguingly, mutations in PARL have recently been found in Parkinson's disease patients [58] and diabetes patients [60], although the significance of these mutations for disease remains speculative.

The third category of rhomboid function was revealed in *Providencia stuartii*, a Gram-negative bacterial pathogen. Genetic screens identified its rhomboid homolog, AarA, to be required for production of an unidentified signal for quorum sensing [61,62]. Once the similarity to rhomboid was noted [63], proteolytic activity of AarA was demonstrated against Spitz [20], and AarA was found to partially rescue tissue development of *Drosophila*

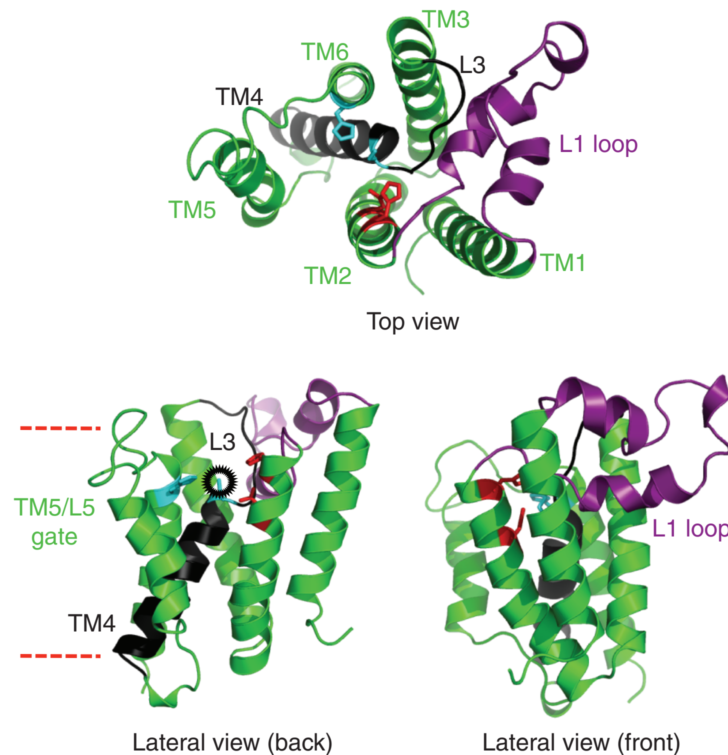

Figure 3. Structural features of the rhomboid 6TM core. The crystal structure of the 6TM core of the *E. coli* rhomboid protease GlpG (PDB 2NRF molecule A) is shown from three vantage points ('top view' is looking at the cell from the outside with the membrane in the plane of the page). The protein forms a compact helical bundle, with two characteristic features. A short and slanted TM4 (black) forms a helix below the catalytic serine (circled in the 'back' view), but an extended loop (L3) above it. This slanted trajectory and extended loop create a cavity above the serine. The L1 loop (purple) forms a hairpin structure that nestles between TMs 1 and 3 and protrudes laterally into the outer leaflet of the membrane (red dashed lines representing the membrane interface are provided only for reference). Catalytic dyad residues serine and histidine are in cyan; putative oxyanion-stabilizing electrophilic asparagine and histidine residues are in red.

mutant in *rhomboid* [64]. Historically, the intriguing similarity of activating *Drosophila* EGF signaling and producing an auto-inducer for bacterial quorum sensing, both by a rhomboid, received much attention [63,65]. But the similarity proved to be superficial when the substrate was identified to be TatA, a component of the twin-arginine translocation machinery [66]. As such, AarA removes a short amino-terminal extension, presumably to activate the machinery for signal secretion, rather than activating the signal itself. TatA from other bacteria, including *E. coli*, lacks this short extension and is immediately active, and the AarA function is therefore an exception. Nevertheless, this is the only known function for a rhomboid protease in any prokaryote, and it dramatically highlights the apparent diversity of rhomboid function even within similar bacteria.

Finally, rhomboid proteases help to dismantle adhesive junctions in unicellular eukaryotic parasites. This is the only role that was discovered by searching for rhomboid targets using substrate specificity determinants [33]. The adhesins of *Plasmodium* and *Toxoplasma* are necessary for host-cell invasion, making them essential proteins for the survival of these obligate intracellular parasites [67]. These parasites encode six or more rhomboid proteases, two of which in each organism are known to process these adhesins at the end of the invasion program [25,26,68-70]. The precise need for this dismantling is not entirely clear, but has been thought to free the parasite from being tethered to the host plasma membrane. Recent knockdown experiments indicate that this processing is important for efficient invasion [71], although the full extent is incompletely understood and may involve later functions during parasite replication within the host cell [72]. Even the non-cell-invasive *Entamoeba histolytica* encodes a highly active rhomboid protease, which is localized to the parasite surface but which relocalizes to phagosomes during feeding and the bud neck during immune evasion, perhaps to shed surface proteins, including lectins [73,74]. The functions of other *Plasmodium* or *Toxoplasma* rhomboid proteases not involved in invasion are not yet understood [75], and many other parasites encode rhomboid enzymes whose functions have never been explored.

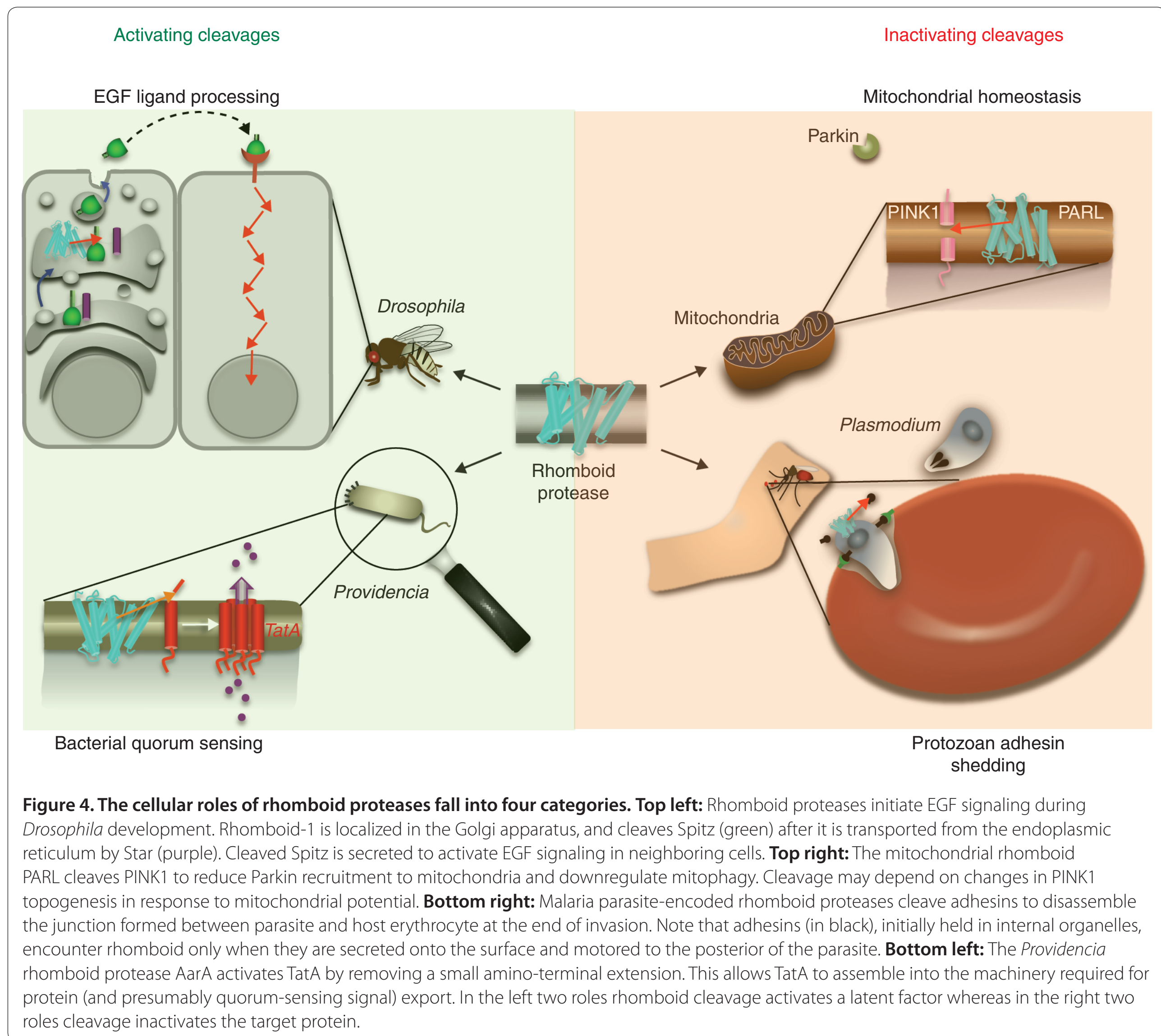

Figure 4. The cellular roles of rhomboid proteases fall into four categories. Top left: Rhomboid proteases initiate EGF signaling during *Drosophila* development. Rhomboid-1 is localized in the Golgi apparatus, and cleaves Spitz (green) after it is transported from the endoplasmic reticulum by Star (purple). Cleaved Spitz is secreted to activate EGF signaling in neighboring cells. **Top right:** The mitochondrial rhomboid PARL cleaves PINK1 to reduce Parkin recruitment to mitochondria and downregulate mitophagy. Cleavage may depend on changes in PINK1 topogenesis in response to mitochondrial potential. **Bottom right:** Malaria parasite-encoded rhomboid proteases cleave adhesins to disassemble the junction formed between parasite and host erythrocyte at the end of invasion. Note that adhesins (in black), initially held in internal organelles, encounter rhomboid only when they are secreted onto the surface and motored to the posterior of the parasite. **Bottom left:** The *Providencia* rhomboid protease AarA activates TatA by removing a small amino-terminal extension. This allows TatA to assemble into the machinery required for protein (and presumably quorum-sensing signal) export. In the left two roles rhomboid cleavage activates a latent factor whereas in the right two roles cleavage inactivates the target protein.

Mechanism

Perhaps the most powerful, yet subtle, guiding principle that can be deduced from the near ubiquity of rhomboid proteases is that they possess a biochemical property that is both very rare and highly useful: but what? Solving this riddle requires understanding the enzymatic features of rhomboid proteases, and remarkable progress has been made towards these goals (reviewed in [38]).

There is now proof beyond doubt that rhomboid enzymes are serine proteases. This includes reconstitution of proteolysis with pure proteins [17,19], protease inhibitor profiling [16,17,76], extensive analysis of residues essential for activity [16,18,19,77], and structural visualization of catalytic residues and with a covalently bound inhibitor [34-37,78]. Moreover, the initial paradox of how water is delivered to the membrane-immersed active site for hydrolysis was largely addressed by structural analyses [34-37]: the active site lies submerged about 10 Å below the presumed membrane surface, but with an open cavity above the active site for water access (Figure 3).

Structure-function analyses of rhomboid proteases have also revealed several unusual proteolytic properties that make them unlike most serine proteases. These differences are clear evidence of convergent evolution to a serine protease mechanism down an independent path. First, structural analysis indicates that nucleophilic catalysis is achieved by a histidine-serine catalytic pair, rather than the more common aspartate-histidine-serine catalytic triad [34-37]. Catalytic dyads have been noted in

a minority of exceptional serine proteases [79]. The identity of the residues that stabilize the oxyanion transition state is uncertain, but this stabilization is most likely mediated by asparagine and/or histidine side-chains [36,78] (Figure 3). Use of an asparagine for oxy-anion stabilization is uncommon but strikingly analogous to the mechanism of the conventional serine protease subtilisin [80].

The third unusual catalytic property of rhomboid proteases relates to the direction in which substrates lie across the active site cleft relative to the catalytic resi-dues. Although initially thought to be similar to nearly all other serine proteases [34,37], identification of the sub-strate gate on the opposite side of GlpG relative to expec-tation mandated that substrates approach the catalytic residues from the so-called 'si' face [35,77,81]. This stereochemical arrangement is very uncommon and had only been encountered in α/β-hydrolyses [82]. Consistent with this stereochemistry are rhomboid's resistance to most canonical serine protease inhibitors and a weak but specific sensitivity to monocyclic β-lactams [16,17,76]. It should be stressed that the definitive evidence for substrate orientation, identity of the oxyanion hole, and the nature of substrate stabilization await a co-structure with a peptide substrate.

Rhomboid proteases have been studied largely within the framework of an established serine protease prece-dent as a way to interpret rhomboid mechanism, which is instructive but does not help to understand how they are different. Although deciphering the specifics of the cata-lytic chemistry is essential for designing effective inhibi-tors, the key functional properties of rhomboid enzymes that are relevant to the cell are unlikely to be determined by its catalytic mechanism. These defining features most likely result from membrane-immersion of the enzyme, and more recent investigations have started to study rhomboid proteases as integral membrane proteins directly.

The greatest impact of membrane immersion is on how substrates and rhomboid proteases behave (as reviewed in [38]). The closed ring of TM segments observed in the first crystal structure suggested that something must move to clear a path for lateral substrate entry [34-37]. Only mutations that weaken TM5 packing with TM2 were found to enhance protease activity by up to ten-fold, thereby identifying the gate functionally [77,81,83]. This dramatic enhancement also revealed that gate opening is the rate-limiting step for intramembrane proteolysis. Molecular dynamics simulations and structural analysis in a bicelle also suggest membrane thinning surrounding GlpG, but its mechanistic implications remain unclear [84,85]. Investigating the role of the membrane in greater detail promises to reveal the defining features of the rhomboid proteolysis system.

Frontiers

The *rhomboid* gene was identified in the *Drosophila* screens of the late 1970s and early 1980s [8], and it was cloned and sequenced about a decade later [10]. It took another decade, until 2001, for its biochemical function as an intramembrane serine protease to be revealed [16]. It has now been a decade since that turning point, and advances in the intervening period have culminated in rhomboid proteases becoming widely regarded as the best understood of all intramembrane proteases [38]. Biochemical insights and defined roles in parasitic protozoa (reviewed in [6]) place rhomboid study on the cusp of becoming applicable in a therapeutic setting. A major lingering obstacle is a rudimentary understanding of its unusual enzymatic mechanism, but these questions are being pursued intensively, and momentum towards a sophisticated understanding is building [38].

By contrast, defining the cellular roles of rhomboid proteases has been a slow process [86]. Although even early biochemical insights have led to the identification of substrates that can be cleaved, whether these candidates are indeed physiological targets, and if so, whether they truly represent a major rhomboid function, remain unknown. For example, although the study of human RHBDL2 over the past 7 years has uncovered at least three well-cleaved substrates (thrombomodulin [24], B-type ephrins [87], and EGF [48]), it is still unclear which, if any, are actual physiological targets, and whether cleavage represents a *bona fide* contribution to cellular function. Perhaps the most humbling example is *E. coli* GlpG, whose atomic details have been revealed in over a dozen structures and countless mutants, yet its cellular function remains a complete mystery [88]. In reality, it is not the ability to find substrate candidates but rather their validation that has proven to be the bottleneck in these studies. Refining search algorithms is unlikely to contribute much towards solving this problem. The urgent need is for approaches with which to study enzymes under physiological settings on a higher throughput scale. This, in turn, will focus biochemical investigations by providing physiological targets and new functional contexts.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We are grateful to Rosanna Baker for help with illustrating Figure 4. Work in our laboratory is supported by the Howard Hughes Medical Institute, and the David and Lucile Packard Foundation.

Published: 27 October 2011

References

1. Sakai J, Duncan EA, Rawson RB, Hua X, Brown MS, Goldstein JL: **Sterol-regulated release of SREBP-2 from cell membranes requires two sequential cleavages, one within a transmembrane segment.** *Cell* 1996, **85:**1037-1046.

2. 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. *Mol Cell* 1997, **1:**47-57.
3. Wolfe MS, Xia W, Ostaszewski BL, Diehl TS, Kimberly WT, Selkoe DJ: Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity. *Nature* 1999, **398:**513-517.
4. Selkoe DJ, Wolfe MS: Presenilin: running with scissors in the membrane. *Cell* 2007, **131:**215-221.
5. Haffner C, Haass C: Cellular functions of gamma-secretase-related proteins. *Neurodegener Dis* 2006, **3:**284-289.
6. Urban S: Making the cut: central roles of intramembrane proteolysis in pathogenic microorganisms. *Nat Rev Microbiol* 2009, **7:**411-423.
7. Osborne TF, Espenshade PJ: Evolutionary conservation and adaptation in the mechanism that regulates SREBP action: what a long, strange tRIP it's been. *Genes Dev* 2009, **23:**2578-2591.
8. Jurgens G, Wieschaus E, Nusslein-Volhard C, Kluding H: Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. *Roux Arch Dev Biol* 1984, **193:**267-282.
9. Mayer U, Nusslein-Volhard C: A group of genes required for pattern formation in the ventral ectoderm of the *Drosophila* embryo. *Genes Dev* 1988, **2:**1496-1511.
10. Bier E, Jan LY, Jan YN: *rhomboid*, a gene required for dorsoventral axis establishment and peripheral nervous system development in *Drosophila melanogaster*. *Genes Dev* 1990, **4:**190-203.
11. Rutledge BJ, Zhang K, Bier E, Jan YN, Perrimon N: The *Drosophila spitz* gene encodes a putative EGF-like growth factor involved in dorsal-ventral axis formation and neurogenesis. *Genes Dev* 1992, **6:**1503-1517.
12. Wasserman JD, Urban S, Freeman M: A family of rhomboid-like genes: *Drosophila* rhomboid-1 and roughoid/rhomboid-3 cooperate to activate EGF receptor signalling. *Genes Dev* 2000, **14:**1651-1663.
13. Guichard A, Biehs 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:**2663-2676.
14. Wasserman JD, Freeman M: An autoregulatory cascade of EGF receptor signalling patterns the *Drosophila* egg. *Cell* 1998, **95:**355-364.
15. Schweitzer R, Shaharabany M, Seger R, Shilo B-Z: Secreted Spitz triggers the DER signalling pathway and is a limiting component in embryonic ventral ectoderm determination. *Genes Dev* 1995, **9:**1518-1529.
16. Urban S, Lee JR, Freeman M: *Drosophila* rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell* 2001, **107:**173-182.
17. Urban S, Wolfe MS: 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:**1883-1888.
18. Lemberg MK, Menendez J, Misik A, Garcia M, Koth CM, Freeman M: Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases. *EMBO J* 2005, **24:**464-472.
19. Maegawa S, Ito K, Akiyama Y: Proteolytic action of GlpG, a rhomboid protease in the *Escherichia coli* cytoplasmic membrane. *Biochemistry (Mosc)* 2005, **44:**13543-13552.
20. Urban S, Schlieper D, Freeman M: Conservation of intramembrane proteolytic activity and substrate specificity in eukaryotic and prokaryotic Rhomboids. *Curr Biol* 2002, **12:**1507-1512.
21. 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 Biol* 2003, **4:**R19.
22. Lemberg MK, Freeman M: Functional and evolutionary implications of enhanced genomic analysis of rhomboid intramembrane proteases. *Genome Res* 2007, **17:**1634-1646.
23. Kanaoka MM, Urban S, Freeman M, Okada K: An *Arabidopsis* Rhomboid homolog is an intramembrane protease in plants. *FEBS Lett* 2005, **579:**5723-5728.
24. Lohi O, Urban S, Freeman M: Diverse substrate recognition mechanisms for rhomboids; thrombomodulin is cleaved by mammalian rhomboids. *Curr Biol* 2004, **14:**236-241.
25. Brossier F, Jewett TJ, Sibley LD, Urban S: A spatially localized rhomboid protease cleaves cell surface adhesins essential for invasion by *Toxoplasma*. *Proc Natl Acad Sci U S A* 2005, **102:**4146-4151.
26. Baker RP, Wijetilaka R, Urban S: Two *Plasmodium* Rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria. *PLoS Pathog* 2006, **2:**e113.
27. Oberai A, Joh NH, Pettit FK, Bowie JU: Structural imperatives impose diverse evolutionary constraints on helical membrane proteins. *Proc Natl Acad Sci U S A* 2009, **106:**17747-17750.
28. McQuibban GA, Saurya S, Freeman M: Mitochondrial membrane remodelling regulated by a conserved rhomboid protease. *Nature* 2003, **423:**537-541.
29. 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:**163-175.
30. 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:**982-989.
31. Karakasis K, Taylor D, Ko K: Uncovering a link between a plastid translocon component and rhomboid proteases using yeast mitochondria-based assays. *Plant Cell Physiol* 2007, **48:**655-661.
32. Strisovsky K, Sharpe HJ, Freeman M: Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates. *Mol Cell* 2009, **36:**1048-1059.
33. Urban S, Freeman M: Substrate specificity of rhomboid intramembrane proteases is governed by helix-breaking residues in the substrate transmembrane domain. *Mol Cell* 2003, **11:**1425-1434.
34. Wang Y, Zhang Y, Ha Y: Crystal structure of a rhomboid family intramembrane protease. *Nature* 2006, **444:**179-180.
35. 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. *Nat Struct Mol Biol* 2006, **13:**1084-1091.
36. Ben-Shem A, Fass D, Bibi E: Structural basis for intramembrane proteolysis by rhomboid serine proteases. *Proc Natl Acad Sci U S A* 2007, **104:**462-466.
37. 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:**750-754.
38. Urban S: Taking the plunge: integrating structural, enzymatic and computational insights into a unified model for membrane-immersed rhomboid proteolysis. *Biochem J* 2010, **425:**501-512.
39. Jeyaraju DV, McBride HM, Hill RB, Pellegrini L: Structural and mechanistic basis of Parl activity and regulation. *Cell Death Differ* 2011, **18:**1531-1539.
40. Sheiner L, Dowse TJ, Soldati-Favre D: Identification of trafficking determinants for polytopic rhomboid proteases in *Toxoplasma gondii*. *Traffic* 2008, **9:**665-677.
41. 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:**79-91.
42. Lilley BN, Ploegh HL: A membrane protein required for dislocation of misfolded proteins from the ER. *Nature* 2004, **429:**834-840.
43. Ye Y, Shibata Y, Yun C, Ron D, Rapoport TA: A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. *Nature* 2004, **429:**841-847.
44. Sato BK, Hampton RY: Yeast Derlin Dfm1 interacts with Cdc48 and functions in ER homeostasis. *Yeast* 2006, **23:**1053-1064.
45. Knop M, Finger A, Braun T, Hellmuth K, Wolf DH: Der1, a novel protein specifically required for endoplasmic reticulum degradation in yeast. *EMBO J* 1996, **15:**753-763.
46. Lee JR, Urban S, Garvey CF, Freeman M: Regulated intracellular ligand transport and proteolysis control EGF signal activation in *Drosophila*. *Cell* 2001, **107:**161-171.
47. Dutt A, Canevascini S, Froehli-Hoier E, Hajnal A: EGF signal propagation during *C. elegans* vulval development mediated by ROM-1 rhomboid. *PLoS Biol* 2004, **2:**e334.
48. Adrain C, Strisovsky K, Zettl M, Hu L, Lemberg MK, Freeman M: Mammalian EGF receptor activation by the rhomboid protease RHBDL2. *EMBO Rep* 2011, **12:**421-427.
49. 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. *Cell Mol Life Sci* 2008, **65:**3822-3829.
50. Zou H, Thomas SM, Yan ZW, Grandis JR, Vogt A, Li LY: Human rhomboid

family-1 gene RHBDF1 participates in GPCR-mediated transactivation of EGFR growth signals in head and neck squamous cancer cells. *FASEB J* 2009, **23**:425-432.

51. Yan Z, Zou H, Tian F, Grandis JR, Mixson AJ, Lu PY, Li LY: Human rhomboid family-1 gene silencing causes apoptosis or autophagy to epithelial cancer cells and inhibits xenograft tumor growth. *Mol Cancer Ther* 2008, **7**:1355-1364.
52. 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. *J Biol Chem* 2003, **278**:27781-27788.
53. Sesaki H, Southard SM, Hobbs AE, Jensen RE: Cells lacking Pcp1p/Ugo2p, a rhomboid-like protease required for Mgm1p processing, lose mtDNA and mitochondrial structure in a Dnm1p-dependent manner, but remain competent for mitochondrial fusion. *Biochem Biophys Res Commun* 2003, **308**:276-283.
54. Herlan M, Bornhovd C, Hell K, Neupert W, Reichert AS: Alternative topogenesis of Mgm1 and mitochondrial morphology depend on ATP and a functional import motor. *J Cell Biol* 2004, **165**:167-173.
55. 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**:168-174; discussion 173.
56. Jin SM, Lazarou M, Wang C, Kane LA, Narendra DP, Youle RJ: Mitochondrial membrane potential regulates PINK1 import and proteolytic destabilization by PARL. *J Cell Biol* 2010, **191**:933-942.
57. Deas E, Plun-Favreau H, Gandhi S, Desmond H, Kjaer S, Loh SH, Renton AE, Harvey RJ, Whitworth AJ, Martins LM, Abramov AY, Wood NW: PINK1 cleavage at position A103 by the mitochondrial protease PARL. *Hum Mol Genet* 2010, **20**:867-879.
58. Shi G, Lee JR, Grimes DA, Racacho L, Ye D, Yang H, Ross OA, Farrer M, McQuibban GA, Bulman DE: Functional alteration of PARL contributes to mitochondrial dysregulation in Parkinson's disease. *Hum Mol Genet* 2011, **20**:1966-1974.
59. 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**:98-102.
60. 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 Metab* 2010, **11**:412-426.
61. Rather PN, Orosz E: Characterization of aarA, a pleiotrophic negative regulator of the 2'-N-acetyltransferase in *Providencia stuartii*. *J Bacteriol* 1994, **176**:5140-5144.
62. 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. *J Bacteriol* 1999, **181**:7185-7191.
63. Gallio M, Kylsten P: *Providencia* may help find a function for a novel, widespread protein family. *Curr Biol* 2000, **10**:R693-694.
64. Gallio M, Sturgill G, Rather P, Kylsten P: A conserved mechanism for extracellular signaling in eukaryotes and prokaryotes. *Proc Natl Acad Sci U S A* 2002, **99**:12208-12213.
65. Urban S: Rhomboid proteins: conserved membrane proteases with divergent biological functions. *Genes Dev* 2006, **20**:3054-3068.
66. Stevenson LG, Strisovsky K, Clemmer KM, Bhatt S, Freeman M, Rather PN: Rhomboid protease AarA mediates quorum-sensing in *Providencia stuartii* by activating TatA of the twin-arginine translocase. *Proc Natl Acad Sci U S A* 2007, **104**:1003-1008.
67. Sibley LD: Invasion and intracellular survival by protozoan parasites. *Immunol Rev* 2011, **240**:72-91.
68. Dowse TJ, Pascall JC, Brown KD, Soldati D: Apicomplexan rhomboids have a potential role in microneme protein cleavage during host cell invasion. *Int J Parasitol* 2005, **35**:747-756.
69. 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. *J Cell Biol* 2006, **174**:1023-1033.
70. Srinivasan P, Coppens I, Jacobs-Lorena M: Distinct roles of *Plasmodium* rhomboid 1 in parasite development and malaria pathogenesis. *PLoS Pathog* 2009, **5**:e1000262.
71. 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 Pathog* 2010, **6**:e1000858.
72. 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**:473-477.
73. 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 Dev* 2008, **22**:1636-1646.
74. 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. *Euk Cell* 2010, **9**:1283-1293.
75. Brossier F, Starnes GL, Beatty WL, Sibley LD: Microneme rhomboid protease TgROM1 is required for efficient intracellular growth of *Toxoplasma gondii*. *Eukaryotic cell* 2008, **7**:664-674.
76. Pierrat OA, Strisovsky K, Christova Y, Large J, Ansell K, Bouloc N, Smiljanic E, Freeman M: Monocyclic beta-lactams are selective, mechanism-based inhibitors of rhomboid intramembrane proteases. *ACS Chem Biol* 2011, **6**:325-335.
77. 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. *Proc Natl Acad Sci U S A* 2007, **104**:8257-8262.
78. Vinothkumar KR, Strisovsky K, Andreeva A, Christova Y, Verhelst S, Freeman M: The structural basis for catalysis and substrate specificity of a rhomboid protease. *EMBO J* 2010, **29**:3797-3809.
79. Ekici OD, Paetzel M, Dalbey RE: Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci* 2008, **17**:2023-2037.
80. Polgar L: The catalytic triad of serine peptidases. *Cell Mol Life Sci* 2005, **62**:2161-2172.
81. Brooks CL, Lazareno-Saez C, Lamoureux JS, Mak MW, Lemieux MJ: Insights into substrate gating in *H. influenzae* rhomboid. *J Mol Biol* 2011, **407**:687-697.
82. Paetzel M, Dalbey RE, Strynadka NC: Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. *Nature* 1998, **396**:186-190.
83. Urban S, Baker RP: *In vivo* analysis reveals substrate-gating mutants of a rhomboid intramembrane protease display increased activity in living cells. *Biol Chem* 2008, **389**:1107-1115.
84. Bondar AN, del Val C, White SH: Rhomboid protease dynamics and lipid interactions. *Structure* 2009, **17**:395-405.
85. Vinothkumar KR: Structure of rhomboid protease in a lipid environment. *J Mol Biol* 2011, **407**:232-247.
86. Freeman M: Rhomboids: 7 years of a new protease family. *Semin Cell Dev Biol* 2009, **20**:231-239.
87. Pascall JC, Brown KD: Intramembrane cleavage of ephrinB3 by the human rhomboid family protease, RHBDL2. *Biochem Biophys Res Commun* 2004, **317**:244-252.
88. 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*. *J Bacteriol* 2006, **188**:3415-3419.
89. Urban S, Lee JR, Freeman M: A family of Rhomboid intramembrane proteases activates all membrane-tether EGF ligands in *Drosophila*. *EMBO J* 2002, **21**:4277-4286.
90. Schulz C, Wood CG, Jones DL, Tazuke SI, Fuller MT: Signaling from germ cells mediated by the rhomboid homolog stet organizes encapsulation by somatic support cells. *Development* 2002, **129**:4523-4534.
91. Urban S, Brown G, Freeman M: EGF receptor signaling protects smooth-cuticle cells from apoptosis during *Drosophila* ventral epidermis development. *Development* 2004, **131**:1835-1845.

doi:10.1186/gb-2011-12-10-231
Cite this article as: Urban S, Dickey SW: The rhomboid protease family: a decade of progress on function and mechanism. *Genome Biology* 2011, **12**:231.