




# Bacterial rhomboid proteases mediate quality control of orphan membrane proteins

Guangyu Liu<sup>1</sup>, Stephen E Beaton<sup>2</sup>, Adam G Grieve<sup>1</sup>, Rhiannon Evans<sup>2</sup>, Miranda Rogers<sup>1</sup>, Kvido Strisovsky<sup>3</sup> , Fraser A Armstrong<sup>2</sup>, Matthew Freeman<sup>1</sup> , Rachel M Exley<sup>1,\*</sup> & Christoph M Tang<sup>1,\*\*</sup> 

## Abstract

Although multiprotein membrane complexes play crucial roles in bacterial physiology and virulence, the mechanisms governing their quality control remain incompletely understood. In particular, it is not known how unincorporated, orphan components of protein complexes are recognised and eliminated from membranes. Rhomboids, the most widespread and largest superfamily of intramembrane proteases, are known to play key roles in eukaryotes. In contrast, the function of prokaryotic rhomboids has remained enigmatic. Here, we show that the *Shigella sonnei* rhomboid proteases GlpG and the newly identified Rhom7 are involved in membrane protein quality control by specifically targeting components of respiratory complexes, with the metastable transmembrane domains (TMDs) of rhomboid substrates protected when they are incorporated into a functional complex. Initial cleavage by GlpG or Rhom7 allows subsequent degradation of the orphan substrate. Given the occurrence of this strategy in an evolutionary ancient organism and the presence of rhomboids in all domains of life, it is likely that this form of quality control also mediates critical events in eukaryotes and protects cells from the damaging effects of orphan proteins.

**Keywords** intramembrane proteolysis; membrane protein complexes; quality control; rhomboid; *Shigella*

**Subject Categories** Microbiology, Virology & Host Pathogen Interaction; Post-translational Modifications & Proteolysis; Translation & Protein Quality

**DOI** 10.15252/embj.2019102922 | Received 12 July 2019 | Revised 17 March 2020 | Accepted 24 March 2020 | Published online 27 April 2020

**The EMBO Journal (2020) 39: e102922**

See also **J Began *et al*** (May 2020) and **JD Knopf & MK Lemberg** (May 2020)

## Introduction

Multiprotein membrane complexes in bacteria mediate fundamental processes such as respiration, secretion of virulence factors and nutrient acquisition (Unden *et al*, 2014; Costa *et al*, 2015; Sheldon

*et al*, 2016). Aberrant assembly or disassembly of complexes results in orphan proteins, which usually require prompt degradation to maintain cellular proteostasis (Harper & Bennett, 2016; Juszkiwicz & Hegde, 2018). However, little is known how unincorporated, orphan components of bacterial protein complexes are sensed and eliminated from membranes.

Rhomboids are the largest family of intramembrane proteases (IMPs) and are found in all kingdoms of life. These enzymes rapidly scan membranes for their substrates (Kreutzberger *et al*, 2019) and have their active sites embedded in the lipid bilayer where they cleave their substrates in or adjacent to transmembrane domains (TMDs). A wide range of functions has been ascribed to eukaryotic rhomboids including growth factor signalling (Urban *et al*, 2001), lipid metabolism (Saita *et al*, 2018), energy production (Spinazzi *et al*, 2019), chloroplast development (Thompson *et al*, 2012), apoptosis regulation (Saita *et al*, 2017), endoplasmic reticulum (ER) protein trafficking (Fleig *et al*, 2012) and surface antigen shedding in the apicomplexan parasites (Shen *et al*, 2014).

Most insights into the catalytic mechanism of rhomboids have been gained from studies of the *Escherichia coli* rhomboid, GlpG (Wang *et al*, 2006; Wu *et al*, 2006; Ben-Shem *et al*, 2007; Xue & Ha, 2012; Zoll *et al*, 2014; Cho *et al*, 2016). The serine residue of the catalytic dyad, Ser<sup>201</sup>/His<sup>254</sup>, is embedded in rhomboids approximately 10 Å below the surface of the membrane (Wu *et al*, 2006; Ben-Shem *et al*, 2007). Current models indicate that rhomboid-mediated proteolysis is a rate-driven process, with the affinity of the enzyme for its substrate not playing an important role (Dickey *et al*, 2013; Cho *et al*, 2016). An initial “interrogation complex” is formed once the TMD of a substrate engages GlpG with an accessible catalytic site (Strisovsky *et al*, 2009; Dickey *et al*, 2013). Subsequent transition to a “scission complex” necessitates unwinding of the TMD of the substrate, driven by helix-destabilising residues such as prolines (Strisovsky *et al*, 2009; Moin & Urban, 2012; Cho *et al*, 2016). Additionally, the nature of residues at certain positions in the TMD of substrates has a major impact on cleavage (Strisovsky *et al*, 2009). In particular, the P1 residue, defined as the newly formed C-terminal residue upon

1 Sir William Dunn School of Pathology, University of Oxford, Oxford, UK

2 Inorganic Chemistry Laboratory, University of Oxford, Oxford, UK

3 Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Praha 6, Czech Republic

\*Corresponding author. Tel: +44 1865 275500; Email: rachel.exley02@path.ox.ac.uk

\*\*Corresponding author. Tel: +44 1865 275500; Email: christoph.tang@path.ox.ac.uk

[The copyright line of this article was changed on 18 May 2020 after original online publication]

cleavage, must be a small aliphatic residue, with this requirement applying to rhomboids from divergent evolutionary backgrounds (Strisovsky *et al*, 2009; Riestra *et al*, 2015; Saita *et al*, 2017). However, further understanding of the mechanism of rhomboid proteolysis is hampered by the lack of knowledge of the cognate substrate(s) of GlpG.

Despite the near-universal presence of rhomboids in bacteria (Koonin *et al*, 2003), remarkably little is known about the role of these enzymes in prokaryotes. To date, only a single substrate of a bacterial rhomboid is known; AarA is a rhomboid in *Providencia stuartii* which cleaves TatA (Stevenson *et al*, 2007), an essential component of the twin-arginine translocation (Tat) system (Palmer & Berks, 2012). TatA processing by AarA is critical for the function of the Tat system, which mediates quorum sensing in *P. stuartii* (Stevenson *et al*, 2007). Despite current knowledge of GlpG structure and activity, the only phenotypes identified for *E. coli* *glpG* mutants are an enhanced resistance to cefotaxime (Clemmer *et al*, 2006) and reduced intestinal colonisation in a murine model (Russell *et al*, 2017). However, the molecular mechanisms underlying these phenotypes are unknown.

Here, we identified and characterised two rhomboids, GlpG and Rhom7, in *Shigella sonnei*, a close relative of *E. coli* that causes bacillary dysentery (Kotloff *et al*, 2017). Similar to *E. coli*, *S. sonnei* colonises the anaerobic environment of the large intestine, where it can invade the epithelial surface by virtue of its type three secretion system (T3SS) (Marteyn *et al*, 2010). GlpG in *S. sonnei* has 99% amino acid identity with the prototypical rhomboid of *E. coli*, while Rhom7 is predicted to possess seven TMDs. A screen for their substrates identified components of three membrane respiratory complexes: HybA and FdoH for GlpG, and HybA and FdnH for Rhom7 (Abaibou *et al*, 1995; Jormakka *et al*, 2002; Pinske *et al*, 2015). We found that rhomboids exhibit exquisite selectivity by cleaving orphan substrates that are dissociated from their cognate complex, while leaving functional substrates intact. Our findings reveal that rhomboids contribute to the quality control of multiprotein membrane complexes and membrane proteostasis, with rhomboid-mediated proteolysis serving as the critical licensing step that allows downstream proteolytic degradation of orphan substrates, so preventing their aggregation.

## Results

### *Shigella sonnei* possesses two active rhomboids, GlpG and Rhom7

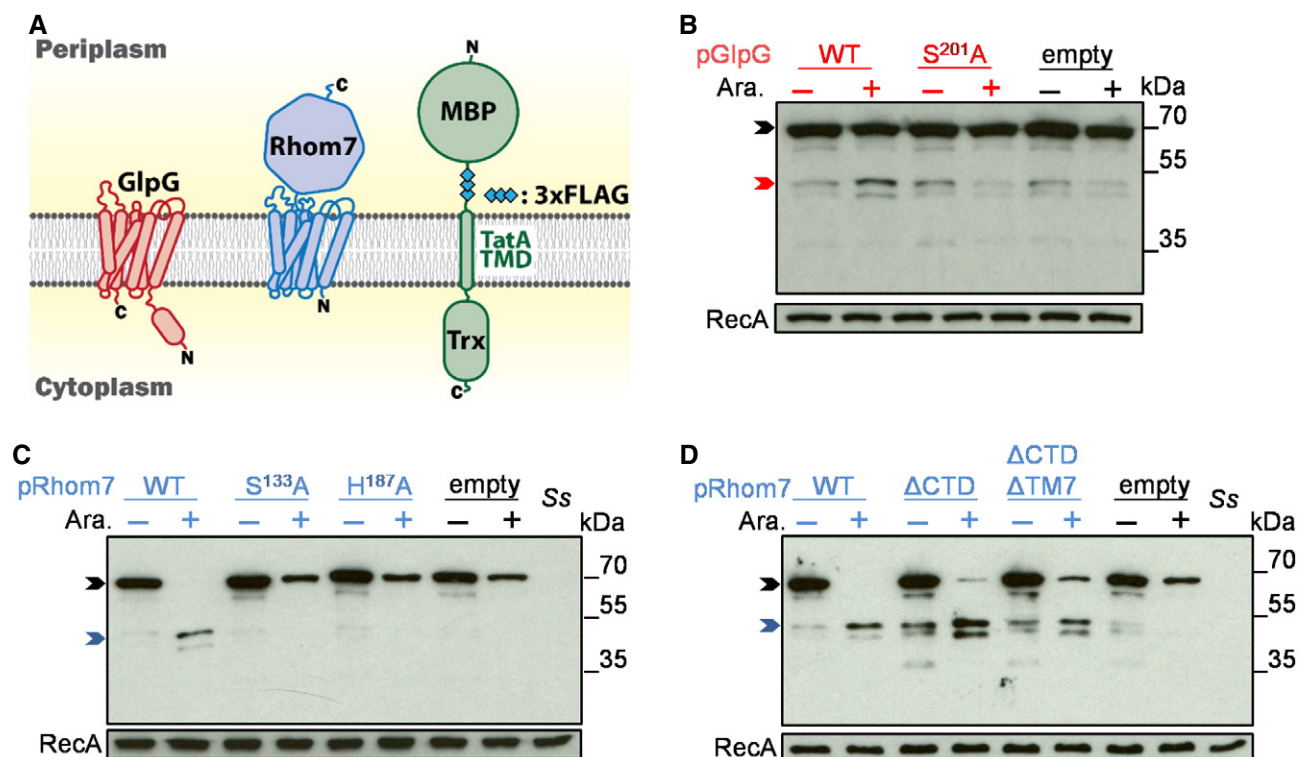
To identify rhomboid proteases in *S. sonnei*, BLASTp searches were performed using sequences of *E. coli* GlpG (accession no. YP\_026220.1) and *P. stuartii* AarA (accession no. AAA61597.1). The *S. sonnei* proteome (taxid: 300269) was used for initial analysis. BLASTp revealed two homologues in *S. sonnei*: SSON\_3661 and SSON\_0610. SSON\_3661 differs by a single amino acid (a.a.) from *E. coli* GlpG (Ala<sup>130</sup> in *S. sonnei* versus Thr<sup>130</sup> in *E. coli*) and so was designated GlpG (Fig EV1). SSON\_0610 shares 30% amino acid identity with *P. stuartii* AarA and harbours two potential catalytic residues, Ser<sup>133</sup> and His<sup>187</sup> (Fig EV1). Topology prediction by Phobius (Kall *et al*, 2007) suggests that SSON\_0610 possesses seven TMDs, as opposed to GlpG which contains six TMDs (Wang *et al*, 2006). Therefore, SSON\_0610 was designated Rhom7.

To establish whether *Shigella* GlpG and Rhom7 are active IMPs, we examined their ability to cleave an artificial substrate (AS) consisting of an N-terminal maltose-binding protein (MBP), a triple-FLAG tag (3xFLAG), the TMD of *Providencia stuartii* TatA (a.a. 1-50) and a thioredoxin domain (Trx) (Fig 1A) (Strisovsky *et al*, 2009). The AS was introduced on a plasmid into *S. sonnei* lacking chromosomal copies of *glpG* and *rhom7* (*S. sonnei*  $\Delta$ *glpG* $\Delta$ *rhom7*) with these genes expressed from plasmids instead; controls included strains expressing inactive GlpG (GlpG<sup>S201A</sup>) (Dickey *et al*, 2013) or with an empty vector. Cleavage of the AS was detected by Western blot analysis with anti-FLAG mAbs (Strisovsky *et al*, 2009) (Fig 1B and C). Both GlpG and Rhom7 cleaved the TMD of TatA within the AS, proving that both are active rhomboid proteases. GlpG<sup>S201A</sup> and Rhom7 with alanine substitution of either of the predicted catalytic residues (*i.e.* Rhom7<sup>S133A</sup> or Rhom7<sup>H187A</sup>) failed to cleave the AS (Fig 1B and C). Furthermore, we examined the effect of removing the 7<sup>th</sup> TMD and the C-terminal domain (Rhom7 $\Delta$ TMD7), or the C-terminal domain alone (Rhom7 $\Delta$ CTD) of Rhom7. Results demonstrate that neither of these features is required for Rhom7 activity against the AS (Fig 1D).

### A bioinformatic screen identifies putative substrates for GlpG and Rhom7

To define the role of these rhomboids, we subjected *S. sonnei*  $\Delta$ *glpG* $\Delta$ *rhom7* to multiple phenotypic assays including growth in complete/minimal media under aerobic or anaerobic conditions and performed Biolog MicroArrays comparing its behaviour with the wild-type strain. Deletion of the rhomboids had no effect in these assays (Fig EV2). As *glpG* is in an operon with *glpE*, which encodes a sulphur donor for the antioxidant thioredoxin 1 (Ray *et al*, 2000), we also tested whether GlpG and/or Rhom7 are involved in survival during oxidative stress. However, there was no significant difference in the recovery of wild-type *S. sonnei* and *S. sonnei*  $\Delta$ *glpG* $\Delta$ *rhom7* after exposure to hydrogen peroxide or paraquat (Fig EV3). Furthermore, loss of GlpG and Rhom7 does not alter secretion through the type three secretion system (Fig EV4) which is essential for *Shigella* virulence (Schroeder & Hilbi, 2008). Taken together, we found GlpG and Rhom7 have no detectable impact on *Shigella* under a variety of conditions.

Therefore, we searched for the substrates of GlpG and Rhom7. Rhomboids have a propensity to cleave single-pass membrane proteins with a periplasmic N-terminus and a cytosolic C-terminus, *i.e.* membrane proteins with type I or type III topology (Urban & Freeman, 2003). To identify potential substrates, we looked for candidates based on their likely location and membrane topology (Fig 2A). The predicted proteome of *S. sonnei* SS046 was interrogated, and graphical representations of the topology of all proteins ( $n = 4,911$ ) were generated using the TMHMM Server (Krogh *et al*, 2001). We manually selected type I and type III proteins, while excluding proteins with ambiguous assignments. This yielded 16 initial potential rhomboid substrates. To identify further candidates for subsequent analysis, the predicted TMDs of these 16 proteins were then used to identify further type I and III proteins using the HHpred server (Soding *et al*, 2005). Additional candidates identified by HHpred analysis were further analysed with Phobius 1.01 and TOPCONS 2.0 to detect signal peptides and to refine topology predictions (Kall *et al*, 2007; Tsirigos *et al*, 2015). Homology



**Figure 1. GlpG and Rhom7 are active rhomboids.**

- A** Topology of GlpG, Rhom7 and the artificial substrate with a maltose-binding protein (MBP) domain, triple-FLAG tag (3xFLAG), the TMD of *P. stuartii* TatA and a thioredoxin domain (Trx).
- B** Western blot analysis (probing with an anti-FLAG mAb) to detect cleavage of the artificial substrate by wild-type (WT)/inactive (S<sup>201A</sup>) GlpG encoded on pBAD33 with/without arabinose (Ara.).
- C** Western blot analysis to detect cleavage of the artificial substrate by wild-type (WT)/modified (S<sup>133A</sup> or H<sup>187A</sup>) Rhom7 encoded on pBAD33 with/without arabinose (Ara.).
- D** Activity of Rhom7 with/without its 7<sup>th</sup> TMD and/or C-terminal domain.

Data information: In (B–D), rhomboid substrates that are uncleaved, cleaved by GlpG or cleaved by Rhom7 are marked by black, red and blue arrows, respectively. Controls, empty pBAD33 (empty) and wild-type *Shigella sonnei* (Ss). RecA, loading control. Source data are available online for this figure.

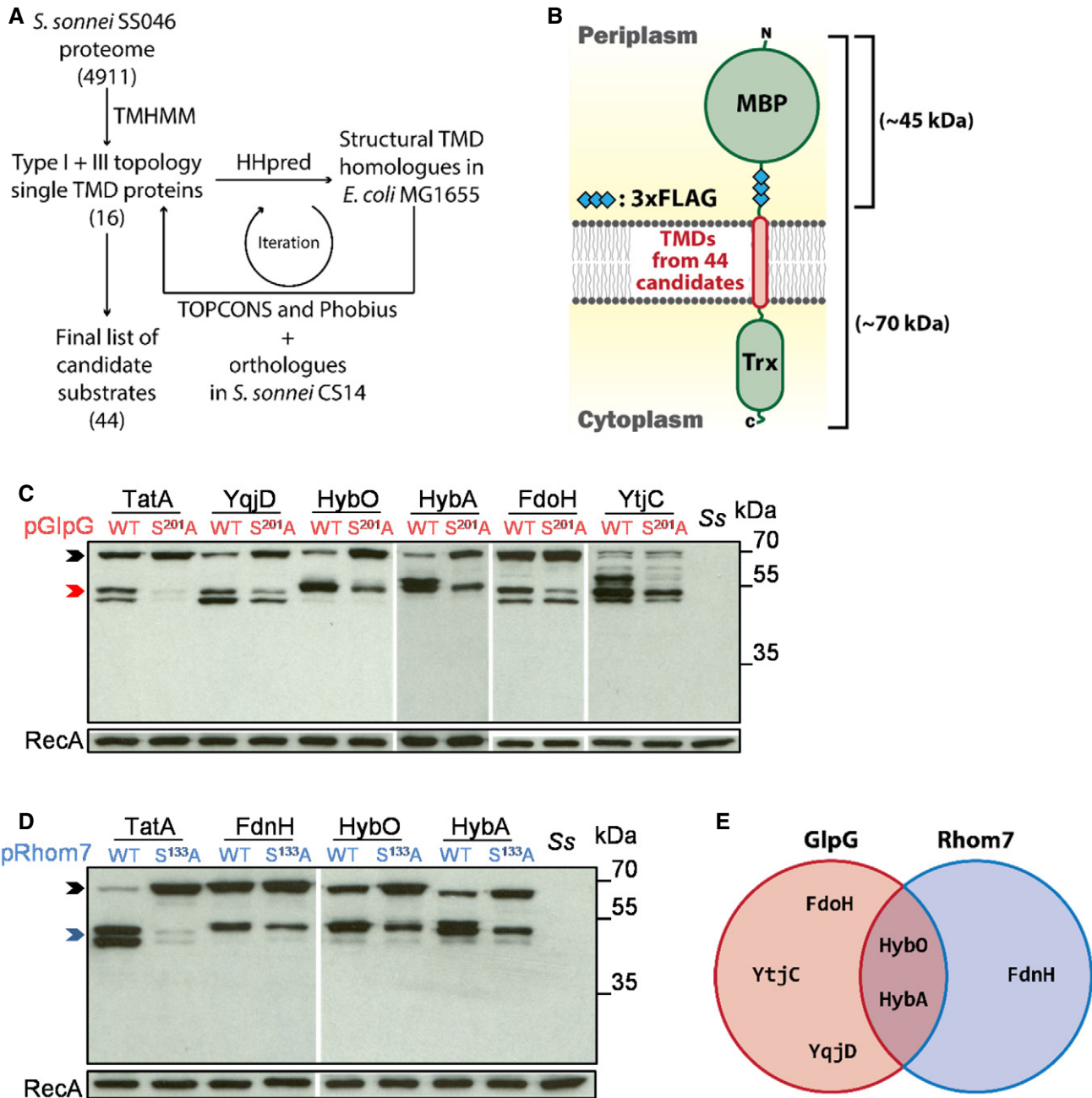
searches with HHpred were performed reiteratively until no further candidates with type I or type III topology were identified, resulting in 44 potential rhomboid substrates.

To examine whether the TMDs of these proteins are cleaved by GlpG and/or Rhom7, the TatA TMD in the AS was replaced with the predicted TMD plus 16 residues at the periplasmic aspect of each candidate (Fig 2B and Table EV1). Western blot analysis was used to assess cleavage of the candidates' TMDs in *S. sonnei* lacking *glpG* and *rhom7* with plasmid-encoded active (wild-type) or inactive versions of the proteases (GlpG<sup>S201A</sup> and Rhom7<sup>S133A</sup>) (Fig 2B and Appendix Fig S1). A positive hit was defined when more product and less full length substrate were observed in the presence of an active rhomboid compared to an inactive enzyme. TMDs from six candidates were reproducibly cleaved by GlpG and/or Rhom7 (Fig 2C–E). Plasmid-encoded GlpG cleaved the TMDs from five candidates: HybA and HybO, subunits of the hydrogenase-2 complex (Hyd-2) (Pinske *et al*, 2015); FdoH, a subunit of the formate dehydrogenase O complex (Abaibou *et al*, 1995); YqjD, a ribosome-associated protein (Yoshida *et al*, 2012);

and YtjC, a putative phosphatase (Yip & Matsumura, 2013). Plasmid-encoded Rhom7 also cleaved the TMDs from HybO and HybA, as well as FdnH, a component of formate dehydrogenase N (Jormakka *et al*, 2002).

### HybA is a physiological substrate of GlpG

Our screen for TMDs cleaved by over-expressed GlpG or Rhom7 identified two proteins belonging to the Hyd-2 complex, HybA and HybO (Pinske *et al*, 2015). Therefore, we next analysed whether full-length HybO and HybA are genuine rhomboid substrates. We constructed strains in which the chromosomal, native genes encoding HybO or HybA were tagged with a sfCherry-3xFLAG tag to allow detection of their expression and cleavage by Western blotting (Fig 3A); GlpG and Rhom7 were expressed from their native locus or from a plasmid. As Hyd-2 is only expressed in the absence of oxygen (Richard *et al*, 1999), we examined cleavage under anaerobic conditions. When analysing HybO cleavage, there was a band of unknown origin with the same molecular mass as predicted for



**Figure 2. Identification of potential GlpG and Rhom7 substrates.**

- A Workflow of the bioinformatic identification of candidate rhomboid substrates.
- B A library of artificial substrates harbouring TMDs from 44 candidate rhomboid substrates.
- C Western blot analysis of candidate substrates reproducibly cleaved by GlpG from the screen.
- D Western blot analysis of candidate substrates reproducibly cleaved by Rhom7 from the screen.
- E Venn diagram summarising putative rhomboid substrates identified by the screen.

Data information: In (C and D), rhomboid substrates that are uncleaved, cleaved by GlpG or cleaved by Rhom7 are marked by black, red and blue arrows, respectively. Controls, inactive enzymes and wild-type *Shigella sonnei* (Ss). RecA, loading control. Source data are available online for this figure.

cleaved HybO even in the absence of GlpG and Rhom7 (Fig 3B). Consequently, we were unable to assess whether either rhomboid cleaves full-length HybO. However, HybA was cleaved by both

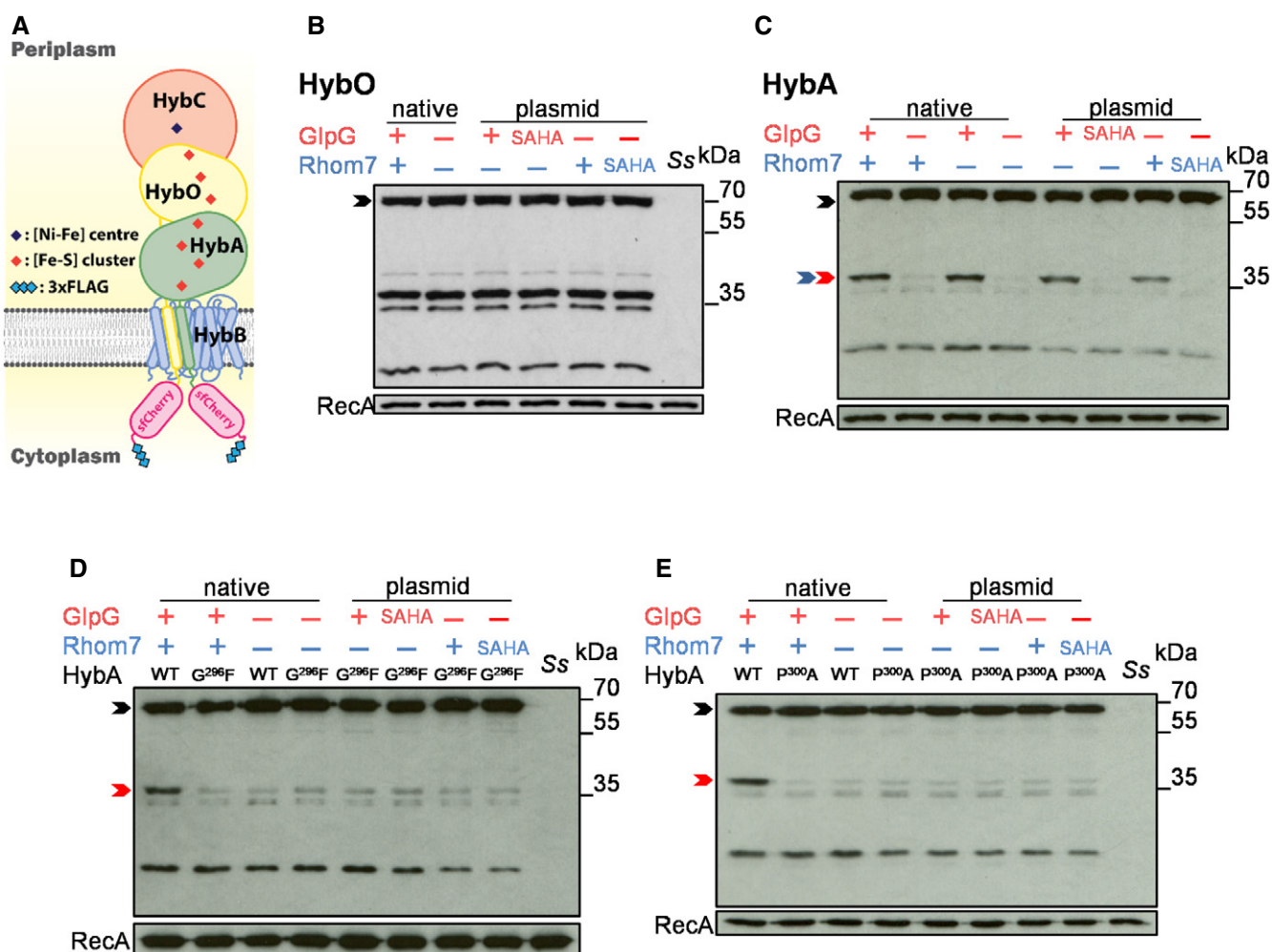
chromosomally and plasmid-encoded GlpG, as well as by plasmid-encoded Rhom7 (Fig 3C). We mapped the GlpG cleavage site in the TMD of HybA by N-terminal sequencing and identified Gly<sup>296</sup> as the



P1 residue (Fig EV5), *i.e.* the C-terminal residue generated upon cleavage, which is consistent with preferred P1 residues in bacterial rhomboid substrates (Strisovsky *et al*, 2009). Characteristically of rhomboid substrates, substitution of the P1 residue with a bulky amino acid (HybA<sup>G296F</sup>) rendered HybA resistant to cleavage by chromosomally or plasmid-expressed GlpG (Fig 3D). Another general feature of rhomboid substrates is the presence of a helix-destabilising residue in the TMD (Moin & Urban, 2012). The HybA TMD contains a potentially destabilising proline, and indeed, HybA<sup>P300A</sup> was not cleaved by GlpG or Rhom7 (Fig 3E). Taken together, our results confirm HybA as the first substrate of GlpG and Rhom7.

### GlpG specifically targets orphan HybA and does not affect Hyd-2 activity

Despite the visible cleavage of HybA, we noticed that > 80% of endogenous HybA remained uncleaved even when GlpG or Rhom7 was expressed from multicopy plasmids (Fig 3C). Considering that the substrate TMD must be accessible to the active site of the rhomboid in the lipid bilayer, we hypothesised that HybA might be protected when it is part of the Hyd-2 complex, but in a protease-sensitive state when it is an isolated, orphan protein. To test this hypothesis, we performed bacterial two-hybrid analysis and demonstrated that HybA interacts directly with HybB (Fig 4A) consistent



**Figure 3. HybA is a physiologic substrate of GlpG.**

A Schematic of HybA and HybO fusions.

B Western blot analysis (probing with an anti-FLAG mAb) to detect cleavage of HybO with (+)/without (-) chromosomal (native) or pBAD33-encoded rhomboids (plasmid).

C Western blot analysis to detect cleavage of HybA with (+)/without (-) chromosomal (native) or pBAD33-encoded rhomboids (plasmid).

D Western blot analysis to detect cleavage of HybA<sup>G296F</sup> by endogenous or pBAD33-encoded rhomboid.

E Western blot analysis to detect cleavage of HybA<sup>P300A</sup> by endogenous or pBAD33-encoded rhomboid.

Data information: In (B-E), rhomboid substrates that are uncleaved, cleaved by GlpG or cleaved by Rhom7 are marked by black, red and blue arrows, respectively. Wild-type *Shigella sonnei*, Ss. Wild-type (WT)/inactive (SAHA: alanine substitution of the catalytic serine and histidine residues) enzymes were pBAD33-encoded (plasmid) in *S. sonnei*  $\Delta$ glpG $\Delta$ rhom7. RecA, loading control.

Source data are available online for this figure.

with previous biochemical data (Pinske *et al*, 2015). Next, we manipulated the stoichiometry of the Hyd-2 complex by either over-expressing HybA or removing HybB. Strikingly, deletion of *hybB* led to a  $\approx$ five-fold increase in the proportion of cleaved to uncleaved HybA (Fig 4B and C). Consistent with this, over-expression of HybA in the presence of HybB resulted in a threefold increase in cleaved to

uncleaved HybA (Fig 4B and Appendix Fig S2A). Furthermore, when HybA is not replenished by *de novo* protein synthesis (by adding chloramphenicol to cells, Fig 4D and E), it is apparent that GlpG cleaves a large proportion of HybA when it is an orphan protein—*i.e.* in the absence of HybB (Fig 4D and Appendix Fig S2B). In contrast, HybA remains largely uncleaved in the presence of HybB (Fig 4E).

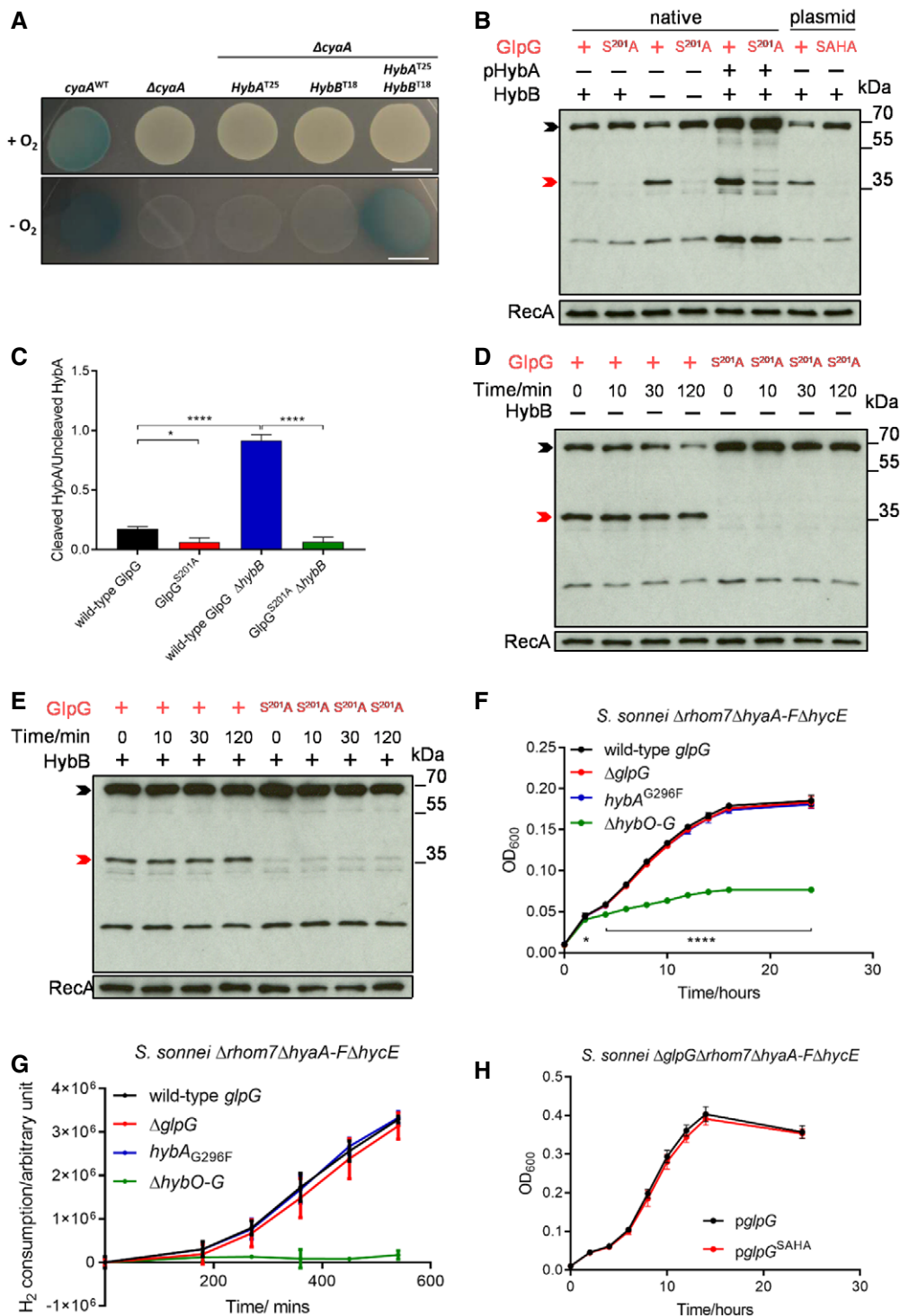


Figure 4.

**Figure 4. GlpG specifically targets orphan HybA and does not influence hydrogenase-2 activity.**

- A Bacterial two-hybrid analysis with HybA and/or HybB fused chromosomally with the T25 and T18 domains of *B. pertussis* CyaA, respectively, in the absence of endogenous CyaA. Bacteria were grown on LB agar containing 20  $\mu\text{g/ml}$  X-gal at 37°C for 14 h in the presence/absence of O<sub>2</sub>. Scale bar, 1 cm.
- B Western blot analysis to detect HybA cleavage in *S. sonnei*  $\Delta\text{rhom7}$  by wild-type (+) or inactive (S<sup>201A</sup>) GlpG expressed chromosomally (native) or from pUC19 (plasmid) with (+)/without (–) HybB. HybA was expressed from its native locus or a plasmid (pHybA).
- C Quantification of the ratio of cleaved/uncleaved HybA in strains with (+) or without (–) HybB. Mean  $\pm$  S.D. of three experiments. \**P* < 0.05; \*\*\*\**P* < 0.0001 (one-way ANOVA).
- D Western blot analysis to detect GlpG-mediated cleavage of native C-terminally sfCherry-3xFLAG-tagged HybA at indicated times after blocking protein translation by the addition of chloramphenicol at T<sub>0</sub> in the absence of HybB (–).
- E Western blot analysis to detect GlpG-mediated cleavage of tagged HybA at times after blocking protein translation by the addition of chloramphenicol at T<sub>0</sub> in the presence of HybB (+).
- F Growth of bacteria lacking hydrogenases ( $\Delta\text{hyaA-F } \Delta\text{hycE } +/– \Delta\text{hybO-G}$ ) or GlpG ( $\Delta\text{glpG}$ ), or expressing uncleavable HybA ( $\text{hybA}^{\text{G296F}}$ ) in 5% H<sub>2</sub>.
- G H<sub>2</sub> consumption by *S. sonnei*. Bacteria were grown aerobically in LB overnight and then diluted into M9 minimal media supplemented with 0.5% fumarate, 12.5  $\mu\text{g/ml}$  nicotinic acid and 0.2% casamino acids in a sealed glass chromatography vial. The headspace was purged with 10% H<sub>2</sub>/90% argon, and cultures were incubated at 37 °C with shaking at 180 rpm for 9 h. H<sub>2</sub> in the headspace was sampled and measured by gas chromatography.
- H *Shigella sonnei* growth in 10% H<sub>2</sub> with plasmid-expressed wild-type or non-functional (pGlpG<sup>S4HA</sup>) GlpG.

Data information: *P*-values were calculated by two-way ANOVA (F, H). Bars represent the mean  $\pm$  SD *n* = 3 (F–H); \**P* < 0.05; \*\*\*\**P* < 0.001. In (B, D, and E), HybA that is uncleaved or cleaved by GlpG is marked by black and red arrows, respectively. RecA, loading control. Source data are available online for this figure.

If GlpG selectively cleaves HybA when it is an orphan molecule, GlpG would be predicted to not affect Hyd-2 activity even though HybA is an essential component of the Hyd-2 complex (Pinske *et al*, 2015). To test this, we assessed whether GlpG affects anaerobic growth with H<sub>2</sub> as the main electron donor (Pinske *et al*, 2015) of strains with Hyd-2 as the only hydrogenase (in *S. sonnei*  $\Delta\text{hyaA-F } \Delta\text{hycE}$ ), and with GlpG as the only rhomboid (by deleting *rhom7*). Under these conditions, the growth of *S. sonnei* is Hyd-2-dependent as illustrated by the severe growth defect of the strain lacking Hyd-2 ( $\Delta\text{hybO-G}$ ) (Fig 4F). Of note, deletion of *glpG* or expression of uncleavable HybA<sup>G296F</sup> had no detectable impact on growth (Fig 4F). Furthermore, we measured bacterial H<sub>2</sub> consumption, a more sensitive and direct readout of Hyd-2 activity (Pinske *et al*, 2015) using the same *S. sonnei* strains during H<sub>2</sub>-dependent growth under anaerobic conditions. We observed no significant difference in the H<sub>2</sub> consumption of strains lacking GlpG or expressing non-cleavable HybA (Fig 4G); under these conditions, wild-type HybA was still cleaved by GlpG, and there was no difference in the survival of strains (Appendix Fig S3). Remarkably, even expression of GlpG from a multicopy plasmid had no impact on Hyd-2 function (Fig 4H), highlighting that GlpG has specifically evolved to target non-functional HybA.

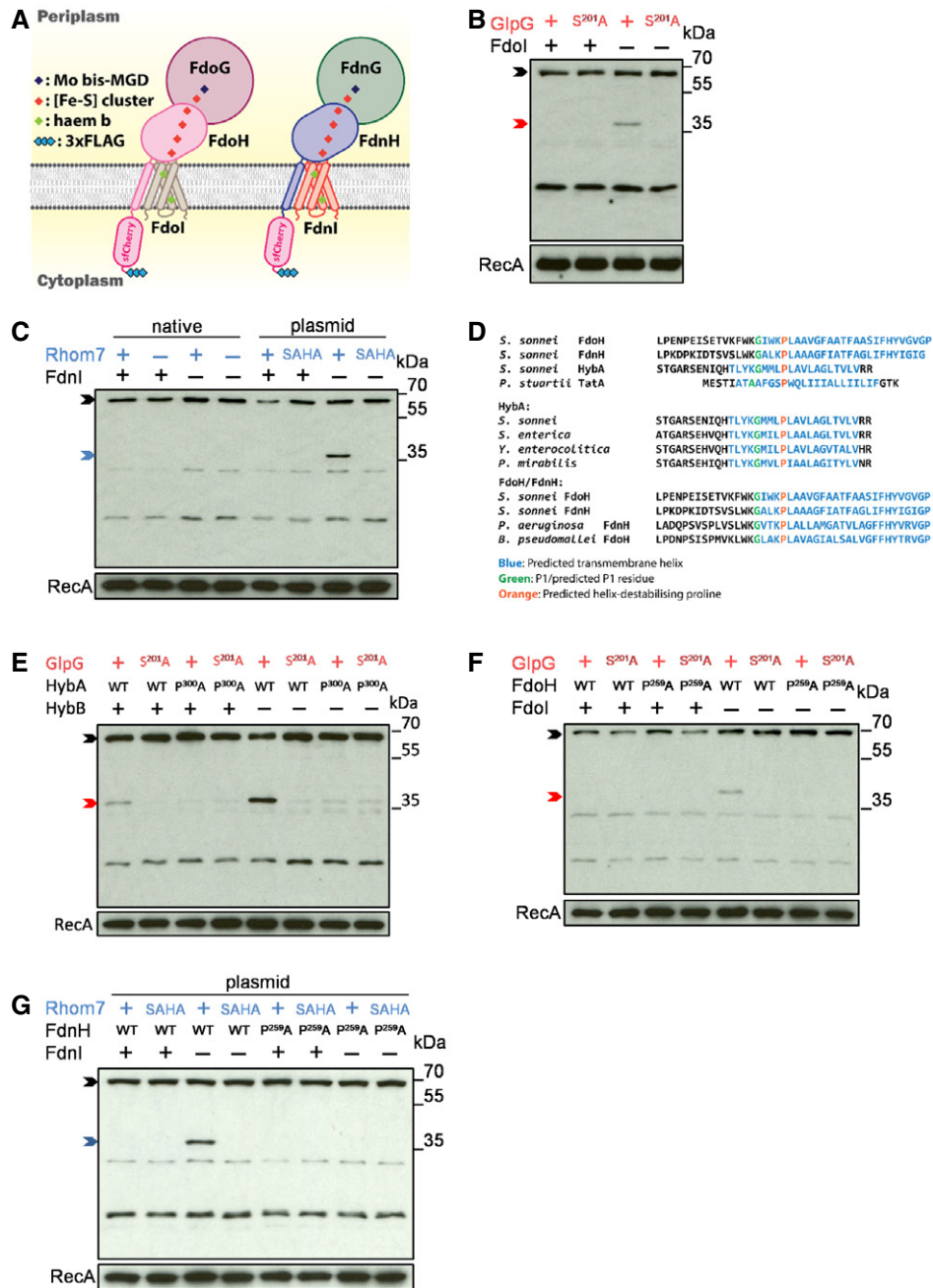
### GlpG and Rhom7 also target orphan components of formate dehydrogenases

Our data illustrate that GlpG is involved in quality control, by selectively targeting a member of a multi-subunit membrane complex when it is an orphan protein (Juszkiewicz & Hegde, 2018). Interestingly, our initial rhomboid substrate screen also identified other components of related multiprotein complexes (Fig 2C–E), *i.e.* FdoH, part of formate dehydrogenase O with FdoG and FdoI (Abaibou *et al*, 1995), and FdnH, part of formate dehydrogenase N with FdnG and FdnI (Jormakka *et al*, 2002). This led us to test whether the targeting of orphan components of multiprotein complexes by rhomboids is a general mechanism of quality control. We constructed strains in which FdoH or FdnH were tagged with the sfCherry-3xFLAG tag (Fig 5A) and assessed whether rhomboids cleave FdoH and FdnH when the stoichiometry of their complexes is perturbed. Notably,

GlpG cleavage of FdoH only became evident in the absence of FdoI (Fig 5B), while plasmid-expressed Rhom7 processes FdnH in bacteria lacking its partner FdnI (Fig 5C). Similar to HybA, the TMDs of FdnH and FdoH have helix-destabilising residues (Pro<sup>259</sup> in both), which are conserved in HybA/FdnH/FdoH orthologues in phylogenetically distant bacterial species as well as in *P. stuartii* TatA (Fig 5D). Importantly, mutating these hallmarks of rhomboid substrates (Moin & Urban, 2012) (HybA<sup>P300A</sup>, FdnH<sup>P259A</sup> and FdoH<sup>P259A</sup>) renders them resistant to cleavage by GlpG and Rhom7 even when they are orphan proteins (Fig 5E–G). In conclusion, our data demonstrate that *Shigella* rhomboids mediate quality control of orphan components from multiple multiprotein respiratory complexes and that cleavage is dependent on an evolutionarily conserved proline residue embedded in TMD of the substrate.

### Rhomboid cleavage licenses further degradation of substrates and prevents the formation of membrane aggregates of orphan HybA

To examine the fate of substrates after initial rhomboid cleavage, we inserted an N-terminal V5 epitope tag into natively expressed wild-type HybA or HybA<sup>P300A</sup> to enable detection of the periplasmic domain. Orphan HybA was cleaved by wild-type GlpG, then subjected to further degradation as evidenced by a reduction in the amount of uncleaved HybA and the appearance of bands of lower molecular mass than the GlpG-mediated cleavage product (Fig 6A). In contrast, there was no evidence of either initial cleavage or further processing of HybA<sup>P300A</sup>. Further degradation of GlpG-cleaved HybA was particularly evident when *de novo* protein translation was blocked by the addition of chloramphenicol (Fig 6B). We employed the same strategy to follow FdoH and FdnH following cleavage by GlpG and Rhom7, respectively. Interestingly, there was no evidence that cleaved, orphan FdoH or FdnH were subject to further degradation when bacteria were grown aerobically or anaerobically in LB supplemented with nitrate (Fig 6C and D, respectively). However, there was detectable processing of both these substrates when bacteria were grown anaerobically in LB supplemented with glycerol and fumarate and then subjected to brief copper stress, which is known to perturb the stability of Fe-S



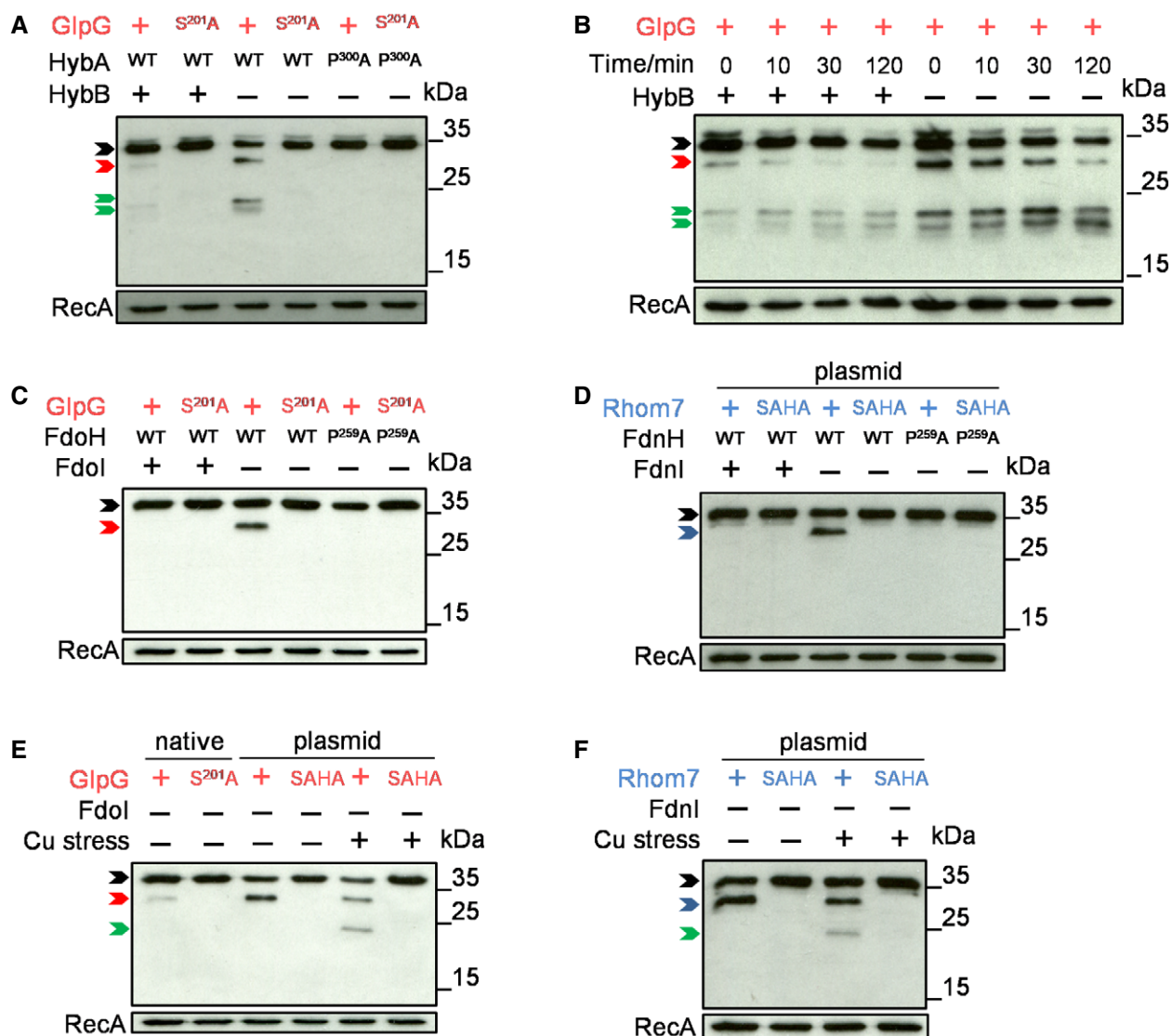
**Figure 5. GlpG and Rhom7 also target orphan components of formate dehydrogenases.**

- A Schematic of FdoH and FdnH fusions.
- B Western blot analysis (probing with an anti-FLAG mAb) to detect cleavage of FdoH cleavage in *S. sonnei*  $\Delta$ *rhom7* with (+)/without (-) Fdol.
- C Western blot analysis to detect cleavage of FdnH in *S. sonnei*  $\Delta$ *glpG* $\Delta$ *rhom7* with (+) or without (-) chromosomal Rhom7 (native) or wild-type (+)/inactive (SAHA) Rhom7 expressed from pBAD33 in *S. sonnei*  $\Delta$ *glpG* $\Delta$ *rhom7* with (+) or without (-) FdnI.
- D Alignments of HybA, FdoH and FdnH in *S. sonnei* with *Providencia stuartii* TatA and with their homologues from phylogenetically diverse organisms (*S. sonnei*, *Salmonella enterica*, *Yersinia enterocolitica*, *Proteus mirabilis*, *Pseudomonas aeruginosa* and *Burkholderia pseudomallei*) highlighting conserved glycine (green) and proline (orange) residues.
- E Western blot analysis to detect cleavage of wild-type (WT) or uncleavable (P<sup>300A</sup>) HybA by active (+) or inactive (S<sup>201A</sup>) chromosomal GlpG in the presence (+) or absence (-) of HybB.
- F Western blot analysis to detect cleavage of wild-type (WT) or modified (P<sup>259A</sup>) FdoH by active (+) or inactive (S<sup>201A</sup>) chromosomal GlpG in the presence (+) or absence (-) of Fdol.
- G Western blot analysis to detect cleavage of wild-type (WT) or modified (P<sup>259A</sup>) FdnH with active (+) or inactive (SAHA) Rhom7 expressed from pBAD33 with (+) or without (-) FdnI.

Data information: In (B, C, E-G), rhomboid substrates that are uncleaved, cleaved by GlpG or cleaved by Rhom7 are marked by black, red and blue arrows, respectively. RecA, loading control.

Source data are available online for this figure.





**Figure 6. Rhomboid cleavage licenses further degradation of substrates.**

**A** Western blot analysis (probing with an anti-V5 mAb) to detect degradation of N-terminally V5-tagged wild-type (WT) or modified (P<sup>300</sup>A) HybA in *S. sonnei*  $\Delta$ *rhom7* chromosomally expressing wild-type (+) or inactive (S<sup>201</sup>A) GlpG with (+) or without (–) HybB.

**B** Degradation of V5-tagged HybA at times after blocking protein translation at T<sub>0</sub> in the presence (+) or absence (–) of HybB.

**C** Western blot analysis of N-terminally V5-tagged wild-type (WT) or modified (P<sup>259</sup>A) FdoH in *S. sonnei*  $\Delta$ *rhom7* chromosomally expressing wild-type (+) or inactive (S<sup>201</sup>A) GlpG with (+)/without (–) FdoI.

**D** Western blot analysis of N-terminally V5-tagged wild-type (WT) or modified (P<sup>259</sup>A) FdnH in *S. sonnei*  $\Delta$ *rhom7* with wild-type (+) or inactive (SAHA) Rhom7 expressed from pBAD33 with (+)/without (–) FdnI.

**E, F** Degradation of N-terminally V5-tagged wild-type (WT) or modified (P<sup>259</sup>A) FdoH (E) or FdnH (F) in *S. sonnei*  $\Delta$ *rhom7* with wild-type (+) or inactive (S<sup>201</sup>A, SAHA) GlpG expressed chromosomally (native) or from pUC19 (plasmid) without FdoI or FdnI (–), respectively, +/- exposure to 400  $\mu$ M CuCl<sub>2</sub> for 30 min.

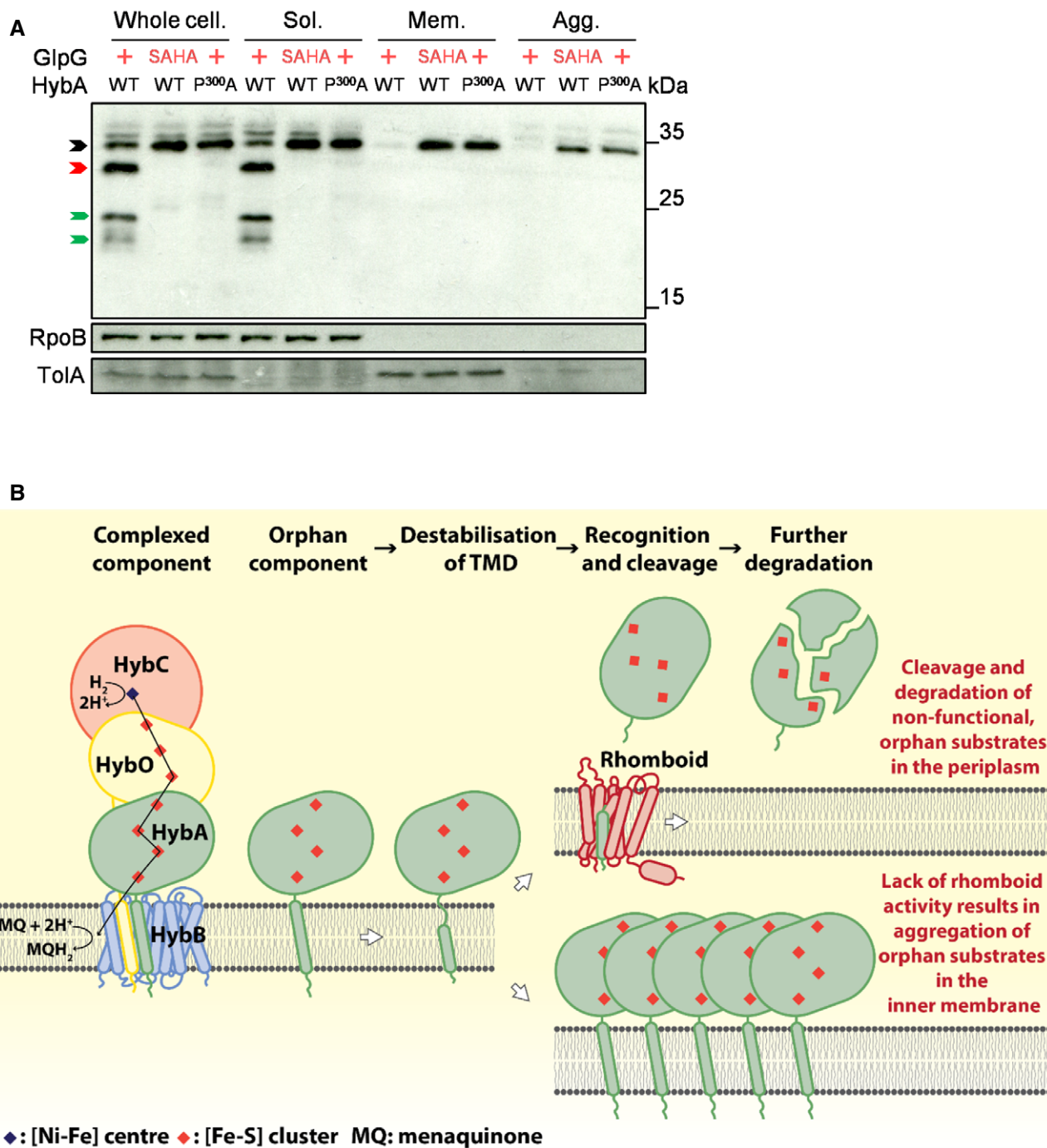
Data information: Rhomboid substrates that are uncleaved, cleaved by GlpG or cleaved by Rhom7 are marked by black, red and blue arrows, respectively. Degradation products post-rhomboid cleavage are marked by green arrows. RecA, loading control.

Source data are available online for this figure.

clusters (Macomber & Imlay, 2009). Therefore, initial rhomboid-mediated cleavage of orphan components of respiratory complexes licenses their further processing.

Given the selective activity of GlpG against orphan HybA, we reasoned that a potential role of this rhomboid is to prevent the accumulation of membrane aggregates of substrate(s) especially

when they are non-functional. Therefore, we performed cellular fractionation to characterise the location of HybA and established whether it forms aggregates following its over-expression in cells as an orphan protein (*i.e.* in the absence of HybB). Successful fractionation was determined by analysing preparations for the presence of TolA (for the whole-cell and membrane fractions) and RpoB (for the



**Figure 7. Rhomboids prevent aggregation of orphan substrates in the inner membrane.**

**A** Western blot analysis probing the localisation and status of plasmid-encoded N-terminally V5-tagged wild-type (WT) or modified (P<sup>300</sup>A) HybA in *S. sonnei*  $\Delta\rho m7\Delta h y b B$  with wild-type (+) or inactive (SAHA) GlpG expressed from pUC19. Whole-cell lysate (Whole cell), Soluble (Sol.), Membrane (detergent-solubilised, Mem.) and the Aggregate (Agg.) fractions are shown. HybA that is uncleaved, cleaved only by GlpG and further degraded post-GlpG cleavage is marked by black, red and green arrows, respectively.

**B** Model of rhomboid-mediated quality control by selectively targeting orphan components of multiprotein respiratory complexes.

Source data are available online for this figure.

whole-cell and soluble fractions) by Western blot analysis (Fig 7A) (Berardini *et al*, 1999; Le Maire *et al*, 2000; Garibyan, 2003; Rassam *et al*, 2018). Of note, while no aggregates were detected in the

presence of GlpG, we found that in the absence of GlpG, orphan HybA was retained in membranes where it formed aggregates (Fig 7A). Consistent with GlpG enabling the elimination of orphan

substrates from the membrane, significant amounts of HybA<sup>P300A</sup> (which cannot be cleaved by GlpG) also became aggregated in the membranes of cells even in the presence of active GlpG (Fig 7A), indicating that GlpG enables the elimination of orphan substrates from membranes, preventing their aggregation.

## Discussion

In this study, we characterised two rhomboids in *S. sonnei*, GlpG and Rhom7, which share sequence homology with *E. coli* GlpG and *P. stuartii* AarA, respectively. We found that both GlpG and Rhom7 are active enzymes that selectively target orphan components of respiratory complexes. The exquisite specificity of both rhomboids for orphan proteins allows them to participate in membrane quality control, by initiating the degradation and elimination of substrates only when they are not part of a functional complex. The search for rhomboid substrates in bacteria has been challenging. Our study highlights potential reasons for this. The substrates are only expressed during growth in specific conditions (*e.g.* in low oxygen environment for HybA and FdnH), cleavage can be prevented by the presence of partner proteins, and because GlpG and Rhom7 selectively target non-functional proteins, their absence might not lead to robust phenotypes.

Rhomboids typically contain a core domain consisting of six TMDs, with or without a 7<sup>th</sup> TMD, and either an additional N-terminal or C-terminal domain (Koonin *et al*, 2003; Stevenson *et al*, 2007; Saita *et al*, 2017). The role of the 7<sup>th</sup> TMD in rhomboids has not been investigated previously. We demonstrate that the 7<sup>th</sup> TMD of Rhom7 and its C-terminal domain are dispensable for cleavage of an artificial substrate. The 7<sup>th</sup> TMD could localise the C-terminal domain to the periplasm, or be involved in substrate recognition or regulate proteolytic activity (see accompanying manuscript). It has been proposed that rhomboids are among the most rapidly diffusing molecules within lipid bilayers (Kreutzberger *et al*, 2019), enabling them to patrol membranes in search of substrates. Consistent with this, GlpG is enzymatically inefficient (with a  $k_{\text{cat}} \approx 0.006/\text{s}$ ) due to its catalytic dyad, rather than the triads found at active sites of classical serine proteases (Dickey *et al*, 2013). This low efficiency would limit indiscriminate cleavage of membrane proteins (Dickey *et al*, 2013), consistent with our findings that GlpG and Rhom7 are quality control enzymes which specifically target orphan membrane proteins. A key requirement of successful quality control is that it should not perturb a functional system, supported by GlpG having no effect on Hyd-2 activity under normal growth conditions (Fig 4F and G), even though HybA is an essential component of this complex.

Our initial bioinformatic screen for identifying substrates excluded substrates predicted to have multiple TMDs, while almost half of our 44 candidates were predicted to be components of multiprotein complexes. The initial screen using TMHMM identified FdnH; however, reiterative searches with HHpred were needed to find FdoH and HybA (Fig 2A), highlighting potential limitations of bioinformatic searches relying on a single algorithm. We screened the TMDs from all candidates by over-expressing them in an AS, so they were likely to be orphan proteins. Despite this, TMDs from only six of the candidates were cleaved by GlpG and/or Rhom7, highlighting the specificity of these rhomboids. We identified HybA and FdoH as the first physiological substrates for the highly characterised enzyme GlpG

(Wang *et al*, 2006; Moin & Urban, 2012; Zoll *et al*, 2014; Cho *et al*, 2016), while HybA and FdnH are substrates for Rhom7 only when this protease is expressed from a plasmid as endogenous expression levels of Rhom7 are insufficient to result in detectable cleavage. The three rhomboid substrates we identified share several characteristics: (i) they are all components of multiprotein membrane complexes (Abaibou *et al*, 1995; Jormakka *et al*, 2002; Dubini & Sargent, 2003); (ii) they have an N-terminal periplasmic ferredoxin domain containing four [4Fe-4S] clusters; and (iii) they are secreted as folded proteins by the Tat system (Sargent *et al*, 2002). Interestingly, the only other known substrate of a bacterial rhomboid is TatA, an integral part of the Tat secretion system (Stevenson *et al*, 2007). This suggests that there might be an evolutionary link between bacterial rhomboids and the Tat system, with the possibility that GlpG and Rhom7 cleave incorrectly folded proteins which have been secreted by Tat, and thence fail to associate into a membrane complex, making them susceptible to rhomboid cleavage.

Our work reveals a form of quality control, which is based on four distinct steps (Fig 7B): (i) a substrate becomes an orphan membrane protein through either non-incorporation into or dissociation from its partner protein(s); (ii) destabilisation of the orphan substrate TMD; (iii) recognition and cleavage by a rhomboid; and finally (iv) further cleavage of the substrate in the periplasm. In this way, bacterial rhomboids offer an elegant strategy to monitor the functional status of components of membrane complexes and protect cells from any potential danger posed by orphan proteins through their accumulation, mis-incorporation into non-cognate complexes and/or localisation to inappropriate cellular sites (Juszkiewicz & Hegde, 2018). We found that in the absence of GlpG, orphan HybA accumulates in membranes as aggregates (Fig 7A). Of note, TTC19 is a substrate of the mitochondrial rhomboid PARL (Saita *et al*, 2017; Spinazzi *et al*, 2019) and is important for the function and assembly of mitochondrial complex III (cIII) by regulating the turnover of the [2Fe-2S] containing Rieske protein (Bottani *et al*, 2017), a crucial catalytic component of cIII. Together with our work, this suggests that the link between rhomboids and quality control of respiratory complexes may be conserved throughout evolution.

Due to the irreversible nature of proteolysis, rhomboid activity in membranes must be carefully regulated. In eukaryotes, this can be achieved by compartmentalisation of enzymes and their substrates in different organelles; for example, rhomboid-1 is located in the endoplasmic reticulum and only cleaves Spitz that has been delivered to this site from the Golgi apparatus (Urban *et al*, 2001). Rhomboid activity can also be influenced by environmental cues such as cytosolic Ca<sup>2+</sup> concentrations (Baker & Urban, 2015). Our work illustrates another mechanism of controlling rhomboid activity through the selective cleavage of non-functional, orphan membrane proteins. This mechanism is reminiscent of regulation by compartmentalisation, but as bacteria lack organelles, compartmentalisation is achieved by sequestration of the substrate TMD in a protected molecular niche provided by its partner protein(s).

A similar mechanism is involved in the quality control of the  $\alpha\beta$  T-cell receptor ( $\alpha\beta$ -TCR) (Klausner *et al*, 1990). The TMD of the  $\alpha$ -chain of the  $\alpha\beta$ -TCR harbours two positively charged residues (Lys and Arg), making it energetically unfavourable in the membrane. However, the  $\alpha$ -chain is stabilised by the presence of Asp residues in the TMDs of other components of the  $\alpha\beta$ -TCR (Feige and Hender-shot, 2013). Although the  $\alpha\beta$ -TCR is evolutionarily and functionally

distinct from bacterial respiratory complexes, the quality control of both systems relies on partner proteins stabilising TMDs which contain atypical features for a lipophilic environment, such as charged residues in the  $\alpha\beta$ -TCR and prolines in HybA, FdoH and FdnH. Interestingly, *E. coli* GlpG can cleave truncated but not full-length MdfA, a multidrug transporter with 12 TMDs, with cleavage occurring at a type I TMD of MdfA possessing a proline (Erez & Bibi, 2009); this suggests that other rhomboids might also monitor and target metastable TMDs in multipass membrane proteins (see accompanying manuscript).

Although poorly understood, assembly-dependent quality control of bacterial protein complexes in the cytoplasmic membrane also involves regulated proteolysis. In *E. coli* and related organisms, this is mainly mediated by the AAA<sup>+</sup> protease, FtsH (Bittner *et al*, 2017). FtsH has been implicated in degradation of membrane proteins that lack periplasmic domains, such as SecY, PspC and F<sub>o</sub>a, when they are dissociated from their native complexes (Kihara *et al*, 1995; Akiyama *et al*, 1996; Singh & Darwin, 2011). However, FtsH does not degrade inner membrane proteins with tightly folded periplasmic domains (Kihara *et al*, 1999). Of note, all rhomboid substrates we identified possess a folded periplasmic ferredoxin domain (Abaibou *et al*, 1995; Jormakka *et al*, 2002; Beaton *et al*, 2018) and so are likely to be refractory to FtsH-mediated quality control. Rhomboids could therefore have evolved to cleave orphan substrates with periplasmic domains, complementing the activity of FtsH or indeed working in concert with other proteases as illustrated in the accompanying manuscript.

Hyd-2 and formate dehydrogenase N are required for the virulence of several enteric pathogens (Maier *et al*, 2013; Winter *et al*, 2013). Of note, HybA, FdnH and FdoH all possess 4Fe-4S clusters, which are prone to damage by exogenous stress. Therefore, the contribution of GlpG and Rhom7 to quality control and membrane protein homeostasis might be enhanced when substrates are subject to insults such as host innate immune response or metal stress (Slauch, 2011; Goldblatt, 2014; Dong *et al*, 2015). Indeed, for FdnH and FdoH, we only observed further degradation following exposure to copper stress (Fig 6E and F). Therefore, it may be that the *in vivo* colonisation defect exhibited by *E. coli* lacking GlpG results from its failure to deal with accumulated components of respiratory complexes in the membrane (Russell *et al*, 2017). Furthermore, efforts to determine the mechanisms of rhomboid catalysis have been hampered by the lack of rhomboid substrate pairs for enzymatic and structural studies. Our work identifying two physiological substrates of GlpG should enable studies of the scission process (Cho *et al*, 2016) and may facilitate the search for more substrates of bacterial rhomboids. Finally, given their efficiency in an evolutionarily ancient organism, the ability of bacterial rhomboids to recognise and selectively cleave orphan membrane proteins might mean that similar mechanisms act to regulate critical events within membranes in higher organisms.

## Materials and Methods

### Media, bacterial strains and plasmids

Bacterial strains, plasmids and primers used in this study are listed in Tables EV2–EV4, respectively. *Escherichia coli* DH5 $\alpha$  was used for

cloning and *E. coli* SM10  $\lambda$ pir (Simon *et al*, 1983) for maintenance of plasmids for conjugation. *Escherichia coli* was grown at 37°C with agitation at 180 r.p.m. in LB media or on LB agar plates containing 1.5% bacteriological agar (w/v). Antibiotics were used at the following concentrations: carbenicillin, 100  $\mu$ g/ml; streptomycin, 50  $\mu$ g/ml; and chloramphenicol, 20  $\mu$ g/ml.

For aerobic growth, *S. sonnei* CS14 (Holt *et al*, 2012) and derivatives were grown at 37°C with agitation at 180 r.p.m. in LB media. Anaerobic growth was performed in a Whitley A35 Anaerobic Station with 5% H<sub>2</sub>, 10% CO<sub>2</sub>, 85% N<sub>2</sub> or 10% H<sub>2</sub>, 10% CO<sub>2</sub>, 80% N<sub>2</sub> at 37°C with agitation at 180 r.p.m. in relevant media. Viable bacteria (colony-forming units/ml) were enumerated by serially diluting cultures in 96-well plates. An aliquot of 10  $\mu$ l from each well was plated onto LB agar and incubated at 37°C overnight before counting.

### Construction of mutants

To construct markerless mutants, we generated vectors for each mutation using pCONJ4s which was constructed with NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs) as follows. The backbone of pKNG101 (Kaniga *et al*, 1991) was amplified with primers pGL225/pGL224 and assembled with the *bla* gene amplified from pGEM-T Easy (Promega) using pGL222/pGL223, generating pCONJ1. pCONJ1 was linearised by *SpeI* digestion (New England Biolabs) and a *PmeI* restriction site added with primers pGL185/pGL186, yielding pCONJ2. Subsequently, the backbone of pCONJ2 was amplified with pGL154/pGL97 and assembled with *sfGFP* (Pedelacq *et al*, 2006) amplified with primers pGL98/pGL155 from pNCC1 sfGFP (Wormann *et al*, 2016), generating pCONJ3. pCONJ3 was then linearised by *EcoRV* digestion (New England Biolabs) and assembled with *sacB* amplified from pKNG101 using pGL247/pGL248, giving rise to pCONJ4. Finally, pCONJ4 was digested by *AflIII* (New England Biolabs) and assembled with a primer heterodimer formed by pGL342/pGL343, to form pCONJ4s. The upstream and downstream flanking regions (approximately 500 bp each) of each target gene were amplified by PCR with the relevant primers and assembled into pCONJ4s at the *PmeI* site; constructs were verified by sequencing (Source BioScience).

*Shigella sonnei* CS14 mutants were then generated by conjugation. First, 20  $\mu$ l of stationary-phase *E. coli* SM10  $\lambda$ pir harbouring a pCONJ4s derivative was mixed with 20  $\mu$ l of stationary-phase *S. sonnei* recipient, and the entire volume was spotted onto LB agar. Bacteria were incubated at 21°C until the spots dried and then at 37°C for 4 h, after which the bacteria were scraped from the plate and resuspended in 1 ml of LB, and 100  $\mu$ l was plated onto LB agar plates with antibiotics. Plates were incubated at 37°C for 14 h. Next, a single colony was used to inoculate liquid LB with carbenicillin and streptomycin and grown to stationary phase. Bacteria were collected by centrifugation, washed and diluted fivefold into 5 ml LB, and then grown for 4 h at 37°C. 100  $\mu$ l of 50-fold diluted culture was then plated on salt-free LB agar containing 10% (w/v) sucrose to select for loss of integrated plasmid; colonies were analysed by PCR to identify mutants and confirmed by sequencing (Source BioScience).

pBAD33-based vectors (Guzman *et al*, 1995) were constructed by assembling appropriate *glpG* and *rhom7* fragments into *XbaI*/*SphI*-linearised pBAD33 by NEBuilder (New England Biolabs).



Specific point mutations or addition of a sequence encoding a 3xHA tag was introduced by Gibson cloning (New England Biolabs) with appropriate primers. pUC19-based vectors (Yanisch-Perron *et al*, 1985) were constructed by assembling appropriate *glpG* fragments into *SphI/HindIII*-linearised pUC19 by NEBuilder, and Gibson cloning was used to introduce specific point mutations or sequences encoding a triple-HA tag. pLAC101 is a low-copy-number plasmid for expressing proteins under IPTG induction. pLAC101 contains the pSC101 origin (amplified with pGL1315/pGL1316 from pUAI39) (Zaslaver *et al*, 2006), the kanamycin resistance gene and *lacI* with a cloning site amplified from pNCC1 (Wormann *et al*, 2016) with primers pGL897/pGL1314 and pGL894/pGL1314, respectively. pLAC101V5-*hybA*/V5-*hybA*<sub>P300A</sub> were generated by assembling appropriate fragments into *PacI/KpnI*-linearised pLAC101. Constructs were verified by sequencing (Source BioScience). Vectors were introduced into *S. sonnei* by electroporation using a Pulse Controller Plus (Bio-Rad) (800 Ωs, 25 μF, 2.5 kV) in 2-mm-gap electroporation cuvettes (Molecular BioProducts) and transformants selected on LB agar plates containing appropriate antibiotics and 0.2% (w/v) L-arabinose (Sigma) where relevant.

### Rhomboid-dependent protein cleavage assays

Substrate cleavage assays were performed in wild-type *S. sonnei*, *S. sonnei* lacking *glpG* and/or *rhom7*, with strains expressing the rhomboids with or without protein tags. In plasmid-based assays, rhomboids were expressed from pBAD33 with arabinose induction, while substrates were expressed using pKS508. For assays performed under aerobic conditions, 50 μl of overnight culture was added to 5 ml LB, and bacteria were grown at 37°C, 180 r.p.m. for 75 min before addition of L-arabinose at a final concentration of 0.2% (w/v). Bacteria were then grown for a further 2 h and harvested for SDS-PAGE and Western blotting. Assays performed under anaerobic conditions were performed in the same way except that 100 μl of overnight culture was added to 5 ml of LB supplemented with 0.5% (w/v) fumarate and 0.5% (v/v) glycerol, and bacteria were grown for 3 h before harvesting for Western blot analysis. L-arabinose (0.2% w/v) was added 1 h after subculturing. Copper stress was achieved by adding CuCl<sub>2</sub> (400 μM final concentration) for 30 min prior to harvesting bacteria. Protein translation inhibition was achieved by adding chloramphenicol (final concentration, 100 ng/ml) to exponential-phase cultures.

### HybA localisation and aggregation status

Assays were performed with *S. sonnei* Δ*rhom7*Δ*hybB* containing N-terminally V5-tagged wild-type (WT) or modified (P<sup>300A</sup>) HybA expressed from pLAC101 with IPTG induction together with wild-type (+) or inactive (SAHA) GlpG constitutively expressed from pUC19. One ml of bacteria grown overnight anaerobically was added to 50 ml LB supplemented with 0.5% (w/v) fumarate and 0.5% (v/v) glycerol, and bacteria were grown anaerobically for 2 h before the addition of IPTG (1 mM). Bacteria were incubated for another hour and then pelleted. Bacterial pellets were resuspended and incubated in 2 ml lysis buffer (50 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 2 mM phenylmethylsulfonyl fluoride, pH 7.4) supplemented with lysozyme (1 mg/ml), DNase I (0.05 mg/ml) and protease inhibitor cocktail (Roche) for 30 min at room temperature.

Bacteria were sonicated on ice and spun at 1,000× g for 30 min at 4°C. A 30 μl aliquot of supernatant was collected as the “whole-cell” fraction for Western blot analysis, while the remaining 970 μl was ultra-centrifuged at 45,000 × g for 30 min. After ultra-centrifugation, 30 μl of the supernatant was collected as the “Soluble” fraction while the rest was decanted. The pellet was resolubilised by 970 μl lysis buffer supplemented with 1% Triton X-100 (v/v) at 4°C for 30 min prior to another round of ultra-centrifugation at 45,000× g for 30 min (Le Maire *et al*, 2000). After that, 30 μl of the supernatant was collected as the “Membrane” fraction. The pellet was resuspended and boiled in 940 μl SDS-PAGE sample buffer as the “Aggregates” fraction for Western blotting analysis.

### SDS-PAGE and Western blotting

Whole-cell lysates were prepared by boiling bacteria for 10 min in SDS-PAGE sample buffer (50 mM Tris, 0.75% (w/v) SDS, 10% (v/v) glycerol, 25 mM EDTA, 1% (v/v) β-mercaptoethanol, 0.11 mM bromophenol blue) at a dilution of an OD<sub>600</sub> unit per 333 μl. A 20 μl aliquot of lysate was separated by electrophoresis (SDS-PAGE) on 12% polyacrylamide gels at 150 V for 75 min in SDS-PAGE buffer (25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS, pH 8.5). Proteins were transferred to nitrocellulose membranes using the Trans-Blot Turbo System (Bio-Rad). Tagged proteins were detected using anti-FLAG M2 antibody (Sigma, F3165, 1:1,000) and goat anti-mouse IgG-HRP (Bio-Rad, 172-1011, 1:5,000), or an anti-V5 (D3H8Q) antibody (Cell Signaling Technology, #13202, 1:1,000) and goat anti-rabbit IgG-HRP (Santa Cruz Biotechnology, sc-2004, 1:5,000). RecA was detected using anti-RecA antibody (Abcam, ab63797, 1:5,000) and goat anti-rabbit IgG-HRP (Santa Cruz Biotechnology, sc-2004, 1:5,000). Bands were detected by chemiluminescence (ECL, GE Healthcare), and band intensity was quantified using Photoshop CC (Lynda.com) as follows: a section of the blots including the relevant protein bands was generated and the total number of pixels within each box was measured, which correlates with the amount of protein.

### Bioinformatic identification of potential substrates

Type I and type III topology proteins were identified using the TMHMM Server v. 2.0 (Krogh *et al*, 2001). As TMDs of proteins are structurally similar (*e.g.* they contain non-polar residues favouring formation of alpha-helices), we used HHpred (Soding *et al*, 2005) to identify structural homologues of the 16 predicted type I and type III proteins identified by initial analysis with TMHMM. Amino acid sequences of the TMD of the 16 candidates were screened against the *E. coli* K-12 proteome (07\_Mar version, the HHpred server does not include a *Shigella* proteome), using the PDB\_mmCIF70\_3\_Jul database and default parameters (Soding *et al*, 2005). Orthologues (> 90% sequence identity) of proteins that are also present in *S. sonnei* were analysed by Phobius 1.01 (Kall *et al*, 2007) and TOPCONS 2.0 (Tsirigos *et al*, 2015) to exclude proteins with multiple TMDs.

### Construction and screen of the TMD library of potential rhomboid substrates

pKS508 was linearised by digestion with *KpnI* and *XbaI* and ligated with annealed primers encoding TMDs + 16 additional amino acids of 44 predicted type I and III membrane proteins. Plasmids encoding

the TMDs are listed in Table EV3 and were verified by sequencing (Source BioScience) before being co-transformed with pBAD33 encoding active/inactive GlpG/Rhom7 into *S. sonnei*  $\Delta$ glpG $\Delta$ rhom7. Cleavage of the 44 artificial substrates by GlpG or Rhom7 *in vivo* was assessed once. Assays were repeated with all artificial substrates that displayed more product/less substrate after expression of an active rhomboid compared to an inactive enzyme.

### Cleavage site identification

Wild-type *S. sonnei*, or *S. sonnei* in which the chromosomal copy of *glpG* was edited to encode the inactive mutant S201A (S201A), both expressing HybA fused to the sfCherry-3xFLAG tag at the C-terminus, was grown anaerobically to an OD<sub>600</sub> = 0.5, collected by centrifugation and lysed by CellLytic B Cell Lysis Reagent (Sigma) supplemented with lysozyme (Sigma) and protease inhibitor cocktail (Roche). The chimeric HybA substrate and its cleavage products were isolated by affinity chromatography using ANTI-FLAG M2 Affinity Agarose Gel (Sigma) and eluted into 50 mM Tris, 150 mM NaCl, 15% glycerol. Purified proteins were separated on SDS-PAGE, electroblotted onto a PVDF PSQ membrane (Millipore) and subjected to sequential Edman degradation on a Procise Protein Sequencing System (491 Protein Sequencer, PE Applied Biosystems).

### Bacterial two-hybrid assay

Bacterial two-hybrid assay was performed based on a previously developed system with CyaA, an adenylate cyclase from *Bordetella pertussis* with modifications (Karimova *et al*, 1998). T25 and T18 fragments of *B. pertussis* CyaA were fused to the C-terminus of HybA and HybB and expressed from their native site on the chromosome of *S. sonnei* lacking endogenous CyaA. A 20  $\mu$ l aliquot of stationary-phase bacteria was spotted on LB agar supplemented with 0.5% (w/v) fumarate, 0.5% (v/v) glycerol, 1 mM IPTG and 40  $\mu$ g/ml X-gal. Spots were dried and incubated either aerobically or anaerobically with 10% H<sub>2</sub>, 10% CO<sub>2</sub>, 80% N<sub>2</sub> at 37°C for 14 h.

### Assays of hydrogenase-2 activity

*Shigella sonnei* possesses three potential hydrogenases, Hyd-1 (encoded by *hyaA-E*), Hyd-2 and Hyd-3 (encoded by the *hyc* operon) (McNorton & Maier, 2012; Pinske & Sawers, 2016). *Shigella sonnei*  $\Delta$ hyaA- $\Delta$ hycE $\Delta$ rhom7 was generated to assess the impact of GlpG and HybA cleavage on Hyd-2 activity. Growth of strains was measured under anaerobic conditions with H<sub>2</sub> and fumarate as previously described (Dubini *et al*, 2002). One millilitre of overnight cultures of *S. sonnei* grown aerobically in LB at 37°C with agitation at 180 r.p.m. was centrifuged at 21,100 $\times$  g for 1 min, the bacterial pellets were washed twice with M9 minimal media supplemented with 0.5% fumarate, 12.5  $\mu$ g/ml nicotinic acid and 0.2% casamino acids, and resuspended in the same media, and the number of bacteria was quantified by measuring the OD<sub>600</sub>. Cultures were transferred to a Whitley A35 Anaerobic Workstation supplied with 5% H<sub>2</sub>, 10% CO<sub>2</sub> and 85% N<sub>2</sub> and diluted to a starting OD<sub>600</sub> of 0.01 in 50 ml of supplemented minimal media that had been pre-incubated in 5% H<sub>2</sub>, 10% CO<sub>2</sub> and 85% N<sub>2</sub> overnight. Bacteria were grown at 37°C with agitation at 180 r.p.m. and the OD<sub>600</sub> measured hourly for 24 h.

Hydrogen uptake was measured by gas chromatography. *Shigella sonnei* strains were prepared as described above and then diluted in 20 ml of supplemented minimal media to an OD<sub>600</sub> = 0.01 in glass gas chromatography vials in a Whitley A35 Anaerobic Workstation. Vials were sealed with #25 Suba-Seal Septa (Sigma-Aldrich) and sealed with Parafilm before removal from the anaerobic chamber. The headspace was purged with 10% H<sub>2</sub> and 90% argon (BOC gases) for 10 min (100 scc/min total flow rate) through a mass flow controller (Sierra Instruments) using a Sterican neural therapy needle (0.80  $\times$  120 mm, 21 G  $\times$  4 $\frac{1}{4}$ "', B Braun) at room temperature before transfer to a shaking incubator at 37°C, 250 r.p.m. Samples (15  $\mu$ l) of the headspace gas were taken periodically using a Hamilton syringe, and the amount and composition of the sampled gas were analysed using a Gas Chromatograph System 7890A (Agilent Technologies) with ShinCarbon ST 100/120 Column, and N<sub>2</sub> as the carrier gas, at 32°C. Each run was stopped approximately 3 min post-injection to ensure capture of the entire H<sub>2</sub> peak. Prior to measurements, a 15- $\mu$ l air injection was applied to verify the gas chromatograph. At each time point, two independent samples of headspace were analysed for each test sample and one sample was analysed for the negative control. After each injection, the Suba-Seal was reinforced with Parafilm. The amount of H<sub>2</sub> in the headspace was calculated by establishing a standard curve with known concentrations of H<sub>2</sub>, and measuring the area underneath the H<sub>2</sub> peak in experiments, quantified by Agilent EZChrom Elite. Uptake of H<sub>2</sub> at a given time point T<sub>x</sub> is calculated by the formula: H<sub>2</sub> uptake T<sub>x</sub> = amount of H<sub>2</sub> in the headspace at T<sub>0</sub> - amount of H<sub>2</sub> in the headspace at T<sub>x</sub>.

### Analysis of Type Three Secretion System activity

Overnight cultures of *S. sonnei* grown aerobically in LB were used to inoculate 10 ml TSB and grown at 37°C with agitation at 180 r.p.m until an OD<sub>600</sub> of 1.5 was reached. Bacteria were harvested by centrifugation at 21,100 $\times$  g, washed and resuspended in PBS to an OD<sub>600</sub> of 5. Congo red was added (200  $\mu$ g/ml final concentration) to the bacterial suspension, which was incubated at 37°C for 15 min to induce secretion. Bacteria were then incubated on ice for 5 min, centrifuged at 21,100 $\times$  g and 150  $\mu$ l of supernatant extracted, mixed with an equal volume of 2 $\times$  SDS-PAGE sample buffer and heated at 100°C for 10 min prior to electrophoresis (20  $\mu$ l) on 10% acrylamide gels. Proteins were stained using SilverXpress™ Silver Staining Kit (Thermo Fisher Scientific).

### Oxidative stress assays

For H<sub>2</sub>O<sub>2</sub> stress, 500  $\mu$ l of an overnight culture of *S. sonnei* grown aerobically in LB was used to inoculate 50 ml LB and grown until the OD<sub>600</sub> reached 0.5, then 5 ml was aliquoted into 50 ml falcon tubes. Bacteria were pre-incubated with 200  $\mu$ M H<sub>2</sub>O<sub>2</sub> for 20 min before the addition of H<sub>2</sub>O<sub>2</sub> (final concentration 2 mM) and incubated for a further 15 min before enumeration of viable cells by plating to LB agar.

For paraquat (PQ), 500  $\mu$ l of overnight cultures of *S. sonnei* grown aerobically in LB was used to inoculate 50 ml M9 minimal media (supplemented with 0.4% glucose, 12.5  $\mu$ g/ml nicotinic acid) and grown to an OD<sub>600</sub> of 0.5; then, 5 ml was aliquoted into 50-ml tubes. Bacteria were then incubated with 200  $\mu$ M PQ for 15 min

before the addition of PQ to a final concentration of 10 mM and incubated for another 45 min before enumeration by plating.

### Statistical analysis

Two-way ANOVA was used to calculate *P* values for OD<sub>600</sub> measurements. One-way ANOVA was used to calculate *P* values for CFU counts and Western blot band intensities. Both one-way and two-way ANOVA values were calculated by GraphPad Prism.

**Expanded View** for this article is available online.

### Acknowledgements

We thank Professor Frank Sargent for helpful discussions, Dr Daniel Mesquita Da Fonseca for helpful discussions and experimental guidance, and Dr Joanna Szczepaniak from the for TolA anti-serum. GL was funded by an EPA Scholarship; work in CMT's laboratory is supported by the Wellcome Trust and BBSRC. Rhiannon Evans is supported by BBSRC grant (BB/N006321/1) awarded to FAA. SEB is grateful to the Dakota Foundation. KS acknowledges support from the ERDF/ESF project No. CZ.02.1.01/0.0/0.0/16\_019/0000729, from Czech Science Foundation (project no. 18-09556S) and from the Czech Academy of Sciences (RVO: 61388963), and thanks Petra Rampírová and Zdeněk Voburka for technical assistance.

### Author contributions

GL, SEB, RME, CMT conceived and designed experiments. GL performed most experiments and analysed data with help from MR; both were supervised by RME. SEB and RE performed hydrogen uptake/generation experiments with GL; KS performed N-terminal sequencing on cleaved HybA. AGG, MF, RE, KS and FAA provided suggestions and contributed to experimental design. GL, RME, and CMT wrote the paper with input from all co-authors.

### Conflict of interest

The authors declare that they have no conflict of interest.

## References

- Abaibou H, Pommier J, Benoit S, Giordano G, Mandrand-Berthelot M-A (1995) Expression and characterization of the *Escherichia coli* *fdo* locus and a possible physiological role for aerobic formate dehydrogenase. *J Bacteriol* 177: 7141–7149
- Akiyama Y, Kihara A, Ito K (1996) Subunit *a* of proton ATPase *F*<sub>0</sub> sector is a substrate of the FtsH protease in *Escherichia coli*. *FEBS Lett* 399: 26–28
- Baker RP, Urban S (2015) Cytosolic extensions directly regulate a rhomboid protease by modulating substrate gating. *Nature* 523: 101–105
- Beaton SE, Evans RM, Finney AJ, Lamont CM, Armstrong FA, Sargent F, Carr SB (2018) The structure of hydrogenase-2 from *Escherichia coli*: implications for H<sub>2</sub>-driven proton pumping. *Biochem J* 475: 1353–1370
- Ben-Shem A, Fass D, Bibi E (2007) Structural basis for intramembrane proteolysis by rhomboid serine proteases. *Proc Natl Acad Sci USA* 104: 462–466
- Berardini M, Foster PL, Loechler EL (1999) DNA polymerase II (*polB*) is involved in a new DNA repair pathway for DNA interstrand cross-links in *Escherichia coli*. *J Bacteriol* 181: 2878–2882
- Bittner L-M, Arends J, Narberhaus F (2017) When, how and why? Regulated proteolysis by the essential FtsH protease in *Escherichia coli*. *Biol Chem* 398: 625–635
- Bottani E, Cerutti R, Harbour ME, Ravaglia S, Dogan SA, Giordano C, Fearnley IM, D'Amati G, Viscomi C, Fernandez-Vizarra E et al (2017) TTC19 plays a husbandry role on UQCRC1 turnover in the biogenesis of mitochondrial respiratory complex III. *Mol Cell* 67: 96–105 e4
- Cho S, Dickey SW, Urban S (2016) Crystal structures and inhibition kinetics reveal a two-stage catalytic mechanism with drug design implications for rhomboid proteolysis. *Mol Cell* 61: 329–340
- Clemmer KM, Sturgill GM, Veenstra A, Rather PN (2006) Functional characterization of *Escherichia coli* GlpG and additional rhomboid proteins using an *aarA* mutant of *Providencia stuartii*. *J Bacteriol* 188: 3415–3419
- Costa TR, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G (2015) Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol* 13: 343–359
- Dickey SW, Baker RP, Cho S, Urban S (2013) Proteolysis inside the membrane is a rate-governed reaction not driven by substrate affinity. *Cell* 155: 1270–1281
- Dong TG, Dong S, Catalano C, Moore R, Liang X, Mekalanos JJ (2015) Generation of reactive oxygen species by lethal attacks from competing microbes. *Proc Natl Acad Sci USA* 112: 2181–2186
- Dubini A, Pye RL, Jack RL, Palmer T, Sargent F (2002) How bacteria get energy from hydrogen: a genetic analysis of periplasmic hydrogen oxidation in *Escherichia coli*. *Int J Hydrogen Energy* 27: 1413–1420
- Dubini A, Sargent F (2003) Assembly of Tat-dependent [NiFe] hydrogenases: identification of precursor-binding accessory proteins. *FEBS Lett* 549: 141–146
- Erez E, Bibi E (2009) Cleavage of a multispinning membrane protein by an intramembrane serine protease. *Biochemistry* 48: 12314–12322
- Fleig L, Bergbold N, Sahasrabudhe P, Geiger B, Kaltak L, Lemberg MK (2012) Ubiquitin-dependent intramembrane rhomboid protease promotes ERAD of membrane proteins. *Mol Cell* 47: 558–569
- Feige MJ, Hendershot LM (2013) Quality control of integral membrane proteins by assembly-dependent membrane integration. *Mol Cell* 51: 297–309
- Garibyan L (2003) Use of the *rpoB* gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair* 2: 593–608
- Goldblatt D (2014) Recent advances in chronic granulomatous disease. *J Infect* 69(Suppl 1): S32–S35
- Guzman L-M, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177: 4121–4130
- Harper JW, Bennett EJ (2016) Proteome complexity and the forces that drive proteome imbalance. *Nature* 537: 328–338
- Holt KE, Baker S, Weill FX, Holmes EC, Kitchen A, Yu J, Sangal V, Brown DJ, Coia JE, Kim DW et al (2012) *Shigella sonnei* genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe. *Nat Genet* 44: 1056–1059
- Jormakka M, Törnroth S, Byrne B, Iwata S (2002) Molecular basis of proton motive force generation structure of formate dehydrogenase-N. *Science* 295: 1863–1868
- Juzskiewicz S, Hegde RS (2018) Quality control of orphaned proteins. *Mol Cell* 71: 443–457
- Kall L, Krogh A, Sonnhammer EL (2007) Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res* 35: W429–W432
- Kaniga K, Delor I, Cornelis GR (1991) A wide-host-range suicide vector for improving reverse genetics in Gram-negative bacteria inactivation of the *blaA* gene of *Yersinia enterocolitica*. *Gene* 109: 137–141

- Karimova G, Pidoux J, Ullmann A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci USA* 95: 5752–5756
- Kihara A, Akiyama Y, Ito K (1995) FtsH is required for proteolytic elimination of uncomplexed forms of SecY, an essential protein translocase subunit. *Proc Natl Acad Sci USA* 92: 4532–4536
- Kihara A, Akiyama Y, Koreaki I (1999) Dislocation of membrane proteins in FtsH-mediated proteolysis. *EMBO J* 18: 2970–2981
- Klausner RD, Lippincott-Schwartz J, Bonifacino JS (1990) The T cell antigen receptor insights into organelle biology. *Annu Rev Cell Biol* 6: 403–431
- Koonin EV, Makarova KS, Rogozin IB, Davidovic L, Letellier MC, Pellegrini L (2003) The rhomboids a nearly ubiquitous family of intramembrane serine proteases that probably evolved by multiple ancient horizontal gene transfers. *Genome Biol* 4: R19
- Kotloff KL, Riddle MS, Platts-Mills JA, Pavlinac P, Zaidi AKM (2017) Shigellosis. *Lancet* 391: 801–812
- Kreutzberger AJB, Ji M, Aaron J, Mihaljevic L, Urban S (2019) Rhomboid distorts lipids to break the viscosity-imposed speed limit of membrane diffusion. *Science* 363: eaao0076
- Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305: 567–580
- Le Maire M, Champeil P, Møller JV (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim Biophys Acta Biomembr* 1508: 86–111
- Macomber L, Imlay JA (2009) The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. *Proc Natl Acad Sci USA* 106: 8344–8349
- Maier L, Vyas R, Cordova CD, Lindsay H, Schmidt TS, Brugiroux S, Periaswamy B, Bauer R, Sturm A, Schreiber F et al (2013) Microbiota-derived hydrogen fuels *Salmonella* typhimurium invasion of the gut ecosystem. *Cell Host Microbe* 14: 641–651
- Marteyn B, West NP, Browning DF, Cole JA, Shaw JG, Palm F, Mounier J, Prevost MC, Sansonetti P, Tang CM (2010) Modulation of *Shigella* virulence in response to available oxygen *in vivo*. *Nature* 465: 355–358
- McNorton MM, Maier RJ (2012) Roles of H<sub>2</sub> uptake hydrogenases in *Shigella flexneri* acid tolerance. *Microbiology* 158: 2204–2212
- Moin SM, Urban S (2012) Membrane immersion allows rhomboid proteases to achieve specificity by reading transmembrane segment dynamics. *Elife* 1: e00173
- Palmer T, Berks BC (2012) The twin-arginine translocation (Tat) protein export pathway. *Nat Rev Microbiol* 10: 483–496
- Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* 24: 79–88
- Pinske C, Jaroschinsky M, Linek S, Kelly CL, Sargent F, Sawers RG (2015) Physiology and bioenergetics of [NiFe]-hydrogenase 2-catalyzed H<sub>2</sub>-consuming and H<sub>2</sub>-producing reactions in *Escherichia coli*. *J Bacteriol* 197: 296–306
- Pinske C, Sawers RG (2016) Anaerobic formate and hydrogen metabolism. *EcoSal Plus* 7: ESPe0011e2016
- Rassam P, Long KR, Kaminska R, Williams DJ, Papadakos G, Baumann CG, Kleanthous C (2018) Intermembrane crosstalk drives inner-membrane protein organization in *Escherichia coli*. *Nat Commun* 9: 1082
- Ray WK, Zeng G, Potters MB, Mansuri AM, Larson TJ (2000) Characterization of a 12-kilodalton rhodanese encoded by *glpE* of *Escherichia coli* and its interaction with thioredoxin. *J Bacteriol* 182: 2277–2284
- Richard DJ, Sawers G, Sargent F, McWalter L, Boxer DH (1999) Transcriptional regulation in response to oxygen and nitrate of the operons encoding the [NiFe] hydrogenases 1 and 2 of *Escherichia coli*. *Microbiology* 145: 2903–2912
- Riestra AM, Gandhi S, Sweredoski MJ, Moradian A, Hess S, Urban S, Johnson PJ (2015) A *Trichomonas vaginalis* rhomboid protease and its substrate modulate parasite attachment and cytolysis of host cells. *PLoS Pathog* 11: e1005294
- Russell CW, Richards AC, Chang AS, Mulvey MA (2017) The rhomboid protease GlpG promotes the persistence of extraintestinal pathogenic *Escherichia coli* within the gut. *Infect Immun* 85: e00866-16
- Saita S, Nolte H, Fiedler KU, Kashkar H, Venne AS, Zahedi RP, Kruger M, Langer T (2017) PARL mediates Smac proteolytic maturation in mitochondria to promote apoptosis. *Nat Cell Biol* 19: 318–328
- Saita S, Tatsuta T, Lampe PA, König T, Ohba Y, Langer T (2018) PARL partitions the lipid transfer protein STARD7 between the cytosol and mitochondria. *EMBO J* 37: e97909
- Sargent F, Berks BC, Palmer T (2002) Assembly of membrane-bound respiratory complexes by the Tat protein-transport system. *Arch Microbiol* 178: 77–84
- Schroeder GN, Hilbi H (2008) Molecular pathogenesis of *Shigella* spp.: controlling host cell signaling, invasion, and death by type III secretion. *Clin Microbiol Rev* 21: 134–156
- Sheldon JR, Laakso HA, Heinrichs DE (2016) Iron acquisition strategies of bacterial pathogens. *Microbiol Spectr* 4: VMBF-0010-2015.
- Shen B, Buguliskis JS, Lee TD, Sibley LD (2014) Functional analysis of rhomboid proteases during *Toxoplasma* invasion. *MBio* 5: e01795–14
- Simon R, Priefer U, Pühler A (1983) A broad host range mobilization system for *in vivo* genetic engineering transposon mutagenesis in gram negative bacteria. *Bio Technol* 1: 784–791
- Singh S, Darwin AJ (2011) FtsH-dependent degradation of phage shock protein C in *Yersinia enterocolitica* and *Escherichia coli*. *J Bacteriol* 193: 6436–6442
- Slauch JM (2011) How does the oxidative burst of macrophages kill bacteria? Still an open question. *Mol Microbiol* 80: 580–583
- Soding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33: W244–W248
- Spinazzi M, Radaelli E, Horre K, Arranz AM, Gounko NV, Agostinis P, Maia TM, Impens F, Morais VA, Lopez-Lluch G et al (2019) PARL deficiency in mouse causes complex III defects, coenzyme Q depletion, and Leigh-like syndrome. *Proc Natl Acad Sci USA* 116: 277–286
- Stevenson LG, Strisovsky K, Clemmer KM, Bhatt S, Freeman M, Rather PN (2007) Rhomboid protease AarA mediates quorum sensing in *Providencia stuartii* by activating TatA of the twin-arginine translocase. *Proc Natl Acad Sci USA* 104: 1003–1008
- Strisovsky K, Sharpe HJ, Freeman M (2009) Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates. *Mol Cell* 36: 1048–1059
- Thompson EP, Llewellyn Smith SG, Glover BJ (2012) An Arabidopsis rhomboid protease has roles in the chloroplast and in flower development. *J Exp Bot* 63: 3559–3570
- Tsirigos KD, Peters C, Shu N, Kall L, Elofsson A (2015) The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res* 43: W401–W407
- Uden G, Steinmetz PA, Degreif-Dunnwald P (2014) The aerobic and anaerobic respiratory chain of *Escherichia coli* and *Salmonella enterica*: enzymes and energetics. *EcoSal Plus* 6 <https://doi.org/10.1128/ecosalplus.ESP-0005-2013>
- Urban S, Lee JR, Freeman M (2001) *Drosophila* rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell* 107: 173–182



- Urban S, Freeman M (2003) Substrate specificity of rhomboid intramembrane proteases is governed by helix-breaking residues in the substrate transmembrane domain. *Mol Cell* 11: 1425–1434
- Wang Y, Zhang Y, Ha Y (2006) Crystal structure of a rhomboid family intramembrane protease. *Nature* 444: 179–180
- Winter SE, Winter MG, Xavier MN, Thiennimitr P, Poon V, Keestra AM, Laughlin RC, Gomez G, Wu J, Lawhon SD et al (2013) Host-derived nitrate boosts growth of *E. coli* in the inflamed gut. *Science* 339: 708–711
- Wormann ME, Horien CL, Johnson E, Liu G, Aho E, Tang CM, Exley RM (2016) *Neisseria cinerea* isolates can adhere to human epithelial cells by type IV pilus-independent mechanisms. *Microbiology* 162: 487–502
- Wu Z, Yan N, Feng L, Oberstein A, Yan H, Baker RP, Gu L, Jeffrey PD, Urban S, Shi Y (2006) Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry. *Nat Struct Mol Biol* 13: 1084–1091
- Xue Y, Ha Y (2012) Catalytic mechanism of rhomboid protease GlpG probed by 3,4-dichloroisocoumarin and diisopropyl fluorophosphonate. *J Biol Chem* 287: 3099–3107
- Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene* 33: 103–119
- Yip SH, Matsumura I (2013) Substrate ambiguous enzymes within the *Escherichia coli* proteome offer different evolutionary solutions to the same problem. *Mol Biol Evol* 30: 2001–2012
- Yoshida H, Maki Y, Furuike S, Sakai A, Ueta M, Wada A (2012) YqjD is an inner membrane protein associated with stationary-phase ribosomes in *Escherichia coli*. *J Bacteriol* 194: 4178–4183
- Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, Liebermeister W, Surette MG, Alon U (2006) A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat Methods* 3: 623–628
- Zoll S, Stanchev S, Began J, Skerle J, Lepšik M, Peclínovská L, Majer P, Strisovský K (2014) Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate-peptide complex structures. *EMBO J* 33: 2408–2421



**License:** This is an open access article under the terms of the Creative Commons Attribution 4.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.