

# Crystal structure of a rhomboid family intramembrane protease

Yongcheng Wang<sup>1</sup>, Yingjiu Zhang<sup>1†</sup> & Ya Ha<sup>1</sup>

*Escherichia coli* GlpG is an integral membrane protein that belongs to the widespread rhomboid protease family. Rhomboid proteases, like site-2 protease (S2P) and  $\gamma$ -secretase, are unique in that they cleave the transmembrane domain of other membrane proteins. Here we describe the 2.1 Å resolution crystal structure of the GlpG core domain. This structure contains six transmembrane segments. Residues previously shown to be involved in catalysis, including a Ser–His dyad, and several water molecules are found at the protein interior at a depth below the membrane surface. This putative active site is accessible by substrate through a large ‘V-shaped’ opening that faces laterally towards the lipid, but is blocked by a half-submerged loop structure. These observations indicate that, in intramembrane proteolysis, the scission of peptide bonds takes place within the hydrophobic environment of the membrane bilayer. The crystal structure also suggests a gating mechanism for GlpG that controls substrate access to its hydrophilic active site.

The concept of intramembrane proteolysis emerged from studies on sterol regulatory element-binding protein (SREBP) and amyloid precursor protein (APP)<sup>1,2</sup>. The activation of SREBP requires it to be cleaved by S2P within its first transmembrane segment<sup>3,4</sup>. To generate amyloid  $\beta$ -peptide associated with Alzheimer’s disease, the transmembrane domain of APP needs to be cleaved by  $\gamma$ -secretase<sup>5</sup>. Now we know that many proteins undergo similar scission within their transmembrane domains, and such processing is important to their biological function (for example, see refs 6, 7). In addition to S2P and  $\gamma$ -secretase, new members have also been added to the list of enzymes that catalyse this reaction (for example, see refs 8–11). The rhomboid family of proteases is among the recent additions<sup>12,13</sup>; they are not related to S2P or  $\gamma$ -secretase by amino acid sequence.

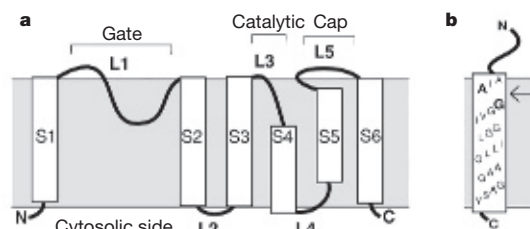
Rhomboid proteases are evolutionarily widespread<sup>14–16</sup>. *Drosophila* Rhomboid-1, the most characterized in the family<sup>17</sup>, regulates epidermal growth factor receptor signalling by cleaving the transmembrane domain of Spitz, the principal ligand for the receptor in flies, and promoting its release from signal-sending cells<sup>12</sup>. Yeast rhomboid Pcp1 regulates mitochondria membrane remodelling<sup>18</sup>. The human homologue of Pcp1, PARL, regulates cytochrome *c* release during apoptosis<sup>19</sup>. Rhomboid family member AarA of *Providencia stuartii* is probably responsible for producing an extracellular quorum-sensing signal<sup>13</sup>. In *Toxoplasma gondii*, rhomboid protease may have a role in parasite invasion<sup>20</sup>.

S2P, presenilin (the catalytic component of  $\gamma$ -secretase) and rhomboid are all integral membrane proteins<sup>5,21–23</sup>. On the basis of sequence and mutagenesis analyses, it has been proposed that S2P, presenilin and rhomboid use hydrolytic mechanisms similar to those used by soluble metallo-, aspartyl and serine proteases, respectively<sup>4,12,24</sup>. Although the location of their catalytic residues in hydrophobic regions of the protein sequence seems to suggest that their active sites are embedded within the membrane (for a review, see refs 22, 23), how these proteases catalyse peptide hydrolysis in lipid was poorly understood. Here we describe the crystal structure of a bacterial rhomboid, which is the first atomic-scale representation of any intramembrane protease, and discuss its implications for the mechanism of intramembrane proteolysis.

## Overall structure

All rhomboid proteases, including *E. coli* GlpG, have a core catalytic domain of six characteristic hydrophobic segments<sup>12,15,21,25</sup> (Fig. 1a; see also Supplementary Figs 1 and 2). They cleave type 1 transmembrane substrates a few residues inside of the membrane from the extracellular side<sup>16,21,26,27</sup> (Fig. 1b). We have solved the structure of the GlpG core domain by X-ray crystallography at 2.1 Å resolution (Table 1; see also Supplementary Table 1 and Supplementary Figs 3 and 4).

As predicted<sup>12,21,25</sup>, GlpG is composed of six transmembrane helices (S1–S6 in Fig. 1). Two features of the structure, however, were not anticipated. The first is an internal cavity. The amino terminus of the central helix S4 is about 10 Å below the membrane surface (Fig. 2b), which creates an aqueous cavity inside of the protein that opens to the extracellular side. The second unanticipated feature is a membrane-embedded loop, L1. All rhomboid protease family members have a large stretch of sequence between S1 and S2 that corresponds to this loop. The position of L1 in relation to the rest of the protein structure suggests that L1 is normally inserted in the outer leaflet of the lipid



**Figure 1 | The membrane topology of a rhomboid protease and its substrate.** **a**, The topology of rhomboid protease core domain. **b**, Rhomboid protease cleaves type 1 transmembrane helix near its extracellular end (arrow). For *Toxoplasma gondii* rhomboid substrate MIC2, the membrane-spanning sequence of which is shown here, cleavage occurs at an Ala–Gly (bold) bond three residues below the membrane surface<sup>45,46</sup>. *Drosophila* Spitz has a similar sequence in this general region, which is recognized and cleaved by Rhomboid-1 (ref. 35).

<sup>1</sup>Department of Pharmacology, Yale School of Medicine, 333 Cedar Street, New Haven, Connecticut 06520, USA. †Present address: College of Life Science, Jilin University, Changchun 130023, China.

bilayer (Fig. 2b). The hydrophobic carboxy-terminal region of L1 is positioned within a large V-shaped gap between S1 and S3, and extends out into the lipid phase. This region of L1, as well as the internal cavity, show the highest sequence conservation between rhomboid proteases and, as explained below, are important for the protease function.

Sequence conservation suggests that all rhomboid core structures are likely to be similar, and that the GlpG structure described here is representative of the family (Supplementary Fig. 1). S3, S4 and S6 contain a number of conserved small residues such as glycines at their packing interfaces that render close helical packing possible and also probably provide stabilizing forces<sup>28</sup>. Unlike GlpG, some rhomboid family members (for example, *Drosophila* Rhomboid-1) appear to contain an additional transmembrane helix attached to the core domain<sup>12,15</sup>. How this additional segment may affect the structure and function of the core is currently unclear.

GlpG forms a trimer in the crystal (Supplementary Fig. 5). The physiological relevance of the trimer is again unclear. Each protomer contains a complete active site, and is likely to function independently.

### Correlation with functional studies

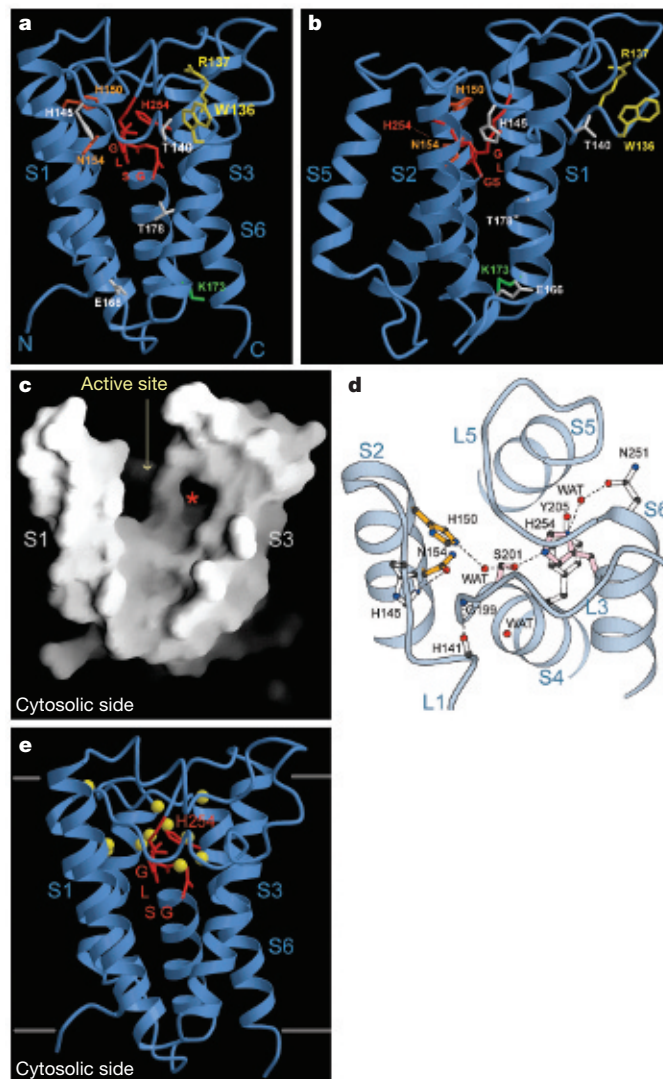
Combining results from this crystallographic study with those from mutagenesis studies on *Drosophila* Rhomboid-1, *Bacillus subtilis* YqgP, *E. coli* GlpG and others<sup>12,16,21,27</sup>, we identify two regions of GlpG structure to be important for function. The first region is the central cavity that contains all the conserved polar residues of S2, S4 and S6 (His 150, Asn 154, Ser 201, His 254). An extended loop L3, as well as the exposed N terminus of S4, also contribute polar main-chain groups to the cavity. This cavity probably represents the active site of the protease. Substituting each of the conserved residues Gly 199 (L3), Ser 201 (S4) and His 254 (S6) in the cavity invariably abolishes activity<sup>12,16,21,27</sup> (red in Fig. 3a, b). Mutations of His 150 (S2) and Asn 154 (S2) also affect the activity for some rhomboids, but not all<sup>12,21,27</sup> (orange in Fig. 3a, b; see below for further discussion).

The hydrophilic cavity is surrounded by transmembrane helices except at its front, where there is a V-shaped gap between S1 and S3 (Fig. 2a). This opening is the likely route by which substrate enters the active site (Fig. 3c): the opening is wide enough to accommodate an  $\alpha$ -helix; its lower portion is shallow and hydrophobic, permitting favourable interactions with a hydrophobic substrate while posing no major restrictions on its sequence; its upper portion is connected to the hydrophilic cavity. In the present structure, the lateral opening is blocked by the membrane-embedded loop L1. We postulate that L1 functions as a gate, and may change conformation when substrate binds. Consistent with the possibility that L1 represents another

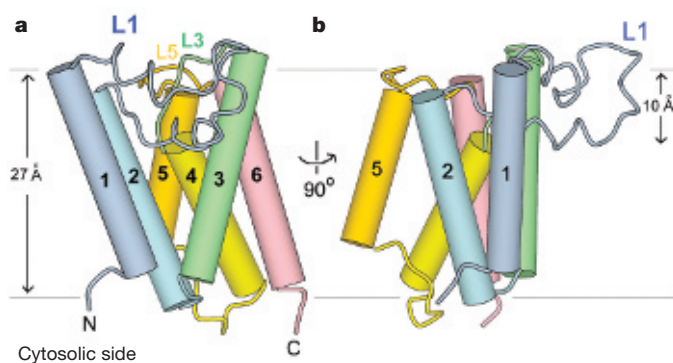
functionally important region, mutation of Trp 136 or Arg 137—two preferred residues on L1 that are 15 Å away from the active site—either abolishes or reduces protease activity<sup>12,27</sup> (yellow in Fig. 3a, b). Analysis of the structure of L1, its conservation and the role of the WR motif (tryptophan and arginine) are discussed below.

### The active site

The crystal structure supports the hypothesis that rhomboid proteases use a serine–histidine catalytic dyad<sup>27</sup>. The essential residues Ser 201 and His 254 of GlpG interact via a strong hydrogen bond<sup>29</sup>



**Figure 3 | The membrane-embedded active site of rhomboid protease.** **a, b**, Mutagenesis studies on *Drosophila* Rhomboid-1 and other rhomboid proteases mapped onto GlpG (viewing angle as in Fig. 2a, b). The GLSG (Gly 199, Ser 201) motif and His 254 are shown in red; His 150 and Asn 154 in orange; Trp 136 and Arg 137 in yellow. Alanine substitution of an arginine in Rhomboid-1 that corresponds to Lys 173 of GlpG (green) also affects activity<sup>12</sup>: this residue may have a structural function. Alanine substitutions of those residues shown in white do not affect activity<sup>12</sup>. **c**, The internal hydrophilic cavity and its front opening. For clarity, the lateral gate (L1) and the cap (L5) have been omitted. The asterisk marks the location of His 254 towards the back of the cavity. The molecular surface is generated by GRASP<sup>49</sup>. **d**, A detailed view of the active site and the hydrogen bond network surrounding the catalytic Ser 201 and His 254 (pink). WAT, water molecule (isolated red dots). Only parts of L1, S2, L3, S4, S5 and S6 are shown for clarity. **e**, Bound water molecules (yellow spheres) within the active site in relation to the conserved GLSG motif and His 254 (red). The boundaries for the hydrophobic region of the membrane are marked by horizontal lines.



**Figure 2 | The overall structure of GlpG.** **a**, The front view of a monomer. The transmembrane segments are sequentially labelled 1–6. The two horizontal lines mark the hypothetical boundaries of the hydrocarbon region of a lipid bilayer. **b**, The side view related to **a** by a 90° rotation, as shown. These illustrations, as well as those in Figs 3a, b, d, e and 4, were generated by MOLSCRIPT<sup>47,48</sup>.

(Fig. 3d). We postulate that the lone base (His 254), like that in several other serine hydrolases<sup>30–33</sup>, is sufficient to activate the serine for direct nucleophilic attack on substrate<sup>34</sup>. Although consistent with data showing that rhomboid proteases can be inhibited by 3,4-dichloroisocoumarin<sup>12,26,27</sup>, a class-specific inhibitor for serine protease, the proposed mechanism lacks direct proof at this time, and therefore remains hypothetical. Asn 154, positioned 8 Å away on the other side of Ser 201, is not within bonding distance to His 254. This latter feature formally eliminates the catalytic triad model<sup>12</sup>, explaining why the asparagine is not essential for catalysis, and its requirement can be affected by the context of the reaction<sup>12,16,21,27</sup>.

A couple of additional interactions may assist His 254 in activating Ser 201 (Fig. 3d). A water molecule mediates hydrogen bonding between His 254 and Asn 251, a frequently observed residue at this position. The ring of His 254 stacks onto that of Tyr 205 ( $\pi$ – $\pi$  interaction). Most rhomboid proteases contain a tyrosine or phenylalanine at position 205, suggesting that this interaction may be important for the function of the dyad.

The backbone amide of Gly 199, previously thought to contribute to oxyanion binding<sup>12,34</sup>, is hydrogen bonded to a backbone carbonyl on L1, and points away from the dyad (Fig. 3d). Also, the short loop L5 tightly caps the active site from above. These features suggest that certain structural rearrangements are required to create a functional active site, an idea also raised by a previous study<sup>26</sup>.

The active site of GlpG contains a number of water molecules (Fig. 3e). The wide distribution of these water molecules, as well as the extensiveness of the protein surface that they interact with, combined with the fact that the capping L5 may open further upon substrate binding, indicate that the reactant water may enter the membrane-embedded active site by different routes from the outside aqueous solution, instead of through a single path or channel.

### The lateral gate

L1 is cradled between S1 and S3, and blocks the lateral opening of the active site (Figs 2a and 3c). It consists mainly of a regular  $\alpha$ -helix ( $\alpha$ 1) and four consecutive  $3_{10}$  helices ( $\alpha$ 2– $\alpha$ 5) (Fig. 4a). In addition to Trp 136 and Arg 137, there are clear preferences for hydrophobic residues at positions 138, 139 and 143, all facing outwards and interacting with lipid (Fig. 4b). The conserved residue His 145 forms a hydrogen bond with Asn 154 (Fig. 3d). The conservation of these interactions suggests that all rhomboid proteases may use a common structural motif for lateral gating.

L1 has an amphiphilic characteristic: its upper and inward-facing surfaces are polar whereas its lower surfaces are hydrophobic (Fig. 4b). This characteristic is consistent with its peripheral membrane location and its critical position between the active site and lipid. The membrane-embedded string of  $3_{10}$  helices is unusual. Here, the polypeptide contains several turns and a number of polar groups are exposed to lipid, which is energetically unfavourable. This feature may be indicative of a metastable and dynamic structure, consistent with its gating function.

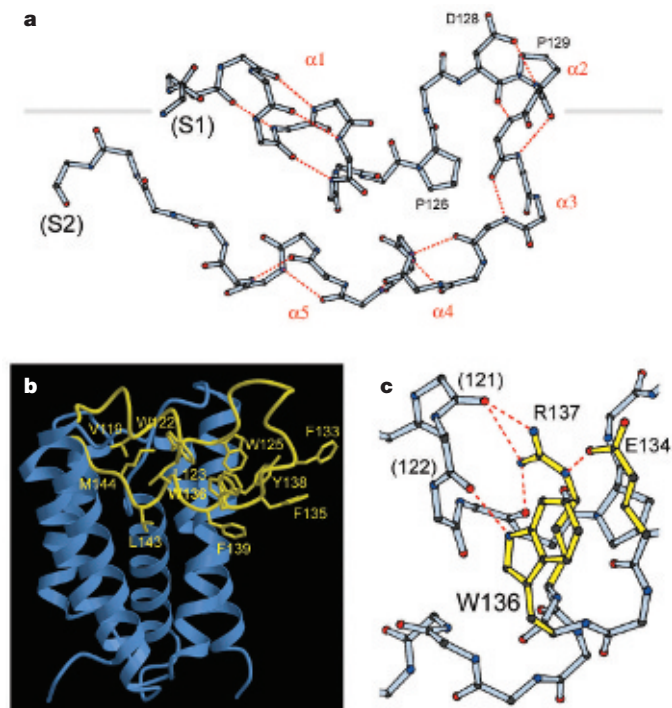
The WR motif has a structural role exclusively within L1 at the membrane–water interface (Fig. 4c). Both residues (tryptophan and arginine) emerge from membrane-embedded locations to interact with the upper structure of L1. Owing to their location towards the tip of the gate away from the main body of the protease (yellow in Fig. 3b), it seems possible that this motif is primarily selected to modulate gate dynamics at the membrane surface. This function would be consistent with previous data showing that the motif is important for activity<sup>12</sup>, but is not part of the core catalytic mechanism<sup>27</sup>, and that its requirement depends on whether the protease is in its native membrane environment<sup>27</sup>. It was noted that another unrelated membrane protein family also contains a WR motif<sup>27</sup>.

### Mechanistic implications

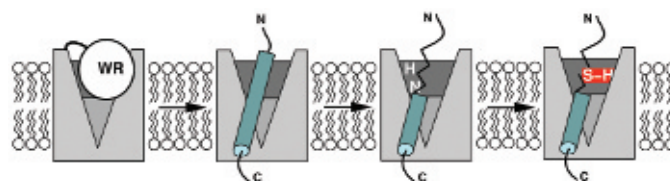
On the basis of the crystal structure described in this report, we propose a model for rhomboid-mediated intramembrane proteolysis, which is illustrated in Fig. 5. In this model, substrate laterally docks into the gap between two transmembrane helices (S1, S3) of the protease, substituting the lateral gate previously bound there. The surface features of the gap suggest that only the top portion of substrate helix unfolds in the hydrophilic cavity of the protease, and becomes subsequently cleaved (Fig. 3c). This model is consistent with a previous study showing that *Drosophila* Rhomboid-1 recognizes substrate by its top region<sup>35</sup>. The requirement for substrate to unfold fits not only the need of the nucleophilic reaction, but also the general shape of the protease active site and the preference for helix-destabilizing residues near the top of the substrate<sup>35</sup>.

### Discussion

A fundamental question concerning the mechanism of intramembrane proteolysis is how substrate moves across the phase barrier that separates membrane from water in order to reach the active site of the enzyme. The crystal structure of GlpG illustrates the physical



**Figure 4 | The structure of the lateral gate.** **a**, The main-chain conformation of L1. The approximate membrane boundary is marked by the horizontal line. **b**, Hydrophobic side chains on the membrane-facing side of L1. **c**, A detailed view of the conserved residues Trp 136 and Arg 137 (yellow), and the interactions that they participate in.



**Figure 5 | A possible mechanism of rhomboid-catalysed intramembrane proteolysis.** The white circle represents the lateral gate (L1) with the conserved WR motif. The hydrophilic internal cavity is represented by the dark grey area. The transmembrane helix of the substrate is shown in blue. Four residues are highlighted in the active site of the rhomboid protease: the serine–histidine dyad (red), and the conserved histidine and asparagine on S2 (white).

**Table 1 | X-ray refinement statistics**

Resolution (Å)	40.0–2.1
$R_{\text{work}}/R_{\text{free}}$	0.236/0.252
Number of atoms	
Protein	1,452
Detergent	191
Water	31
B-factors	
Protein	33.6
Detergent	66.1
Water	35.4
R.m.s deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.26

$R_{\text{work}} = \sum |F_o - F_c| / \sum F_o$ ,  $R_{\text{free}}$  is the cross-validation  $R$ -factor for the test set of reflections (10% of the total) omitted in model refinement. R.m.s., root mean square.

principles underlying one protein machinery that has evolved to facilitate this process. The active site of GlpG is positioned into the membrane at a distance roughly matching the cleavage site of the substrate, but remains separated from the lipid phase by protein structures that surround the active site. Substrate enters and unfolds in the active site through a protein opening that is gated by a specialized mobile structure. This lateral gate has an amphiphilic property that enables it to function smoothly at the membrane–water interface without causing denaturation or aggregation of the enzyme and its transmembrane substrate.

S2P and the presenilins may use different structural motifs to accomplish intramembrane proteolysis (for example, see refs 36, 37). Nevertheless, if lateral gating is a general feature, we note that presenilins do have a highly conserved and hydrophobic sequence—between two transmembrane segments that harbour the catalytic aspartates<sup>24,38,39</sup>—that could be well positioned to fulfil this function. Many familial Alzheimer's disease mutations are found within this sequence.

## METHODS

**Crystallization.** Crystals were grown at room temperature by the hanging-drop vapour diffusion method from a 5 mg ml<sup>-1</sup> membrane protein solution in 10 mM Tris-HCl (pH 7.6) and 20 mM nonylglucoside, over a reservoir solution of 3 M NaCl and 100 mM Bis-Tris propane (pH 7.0). Cryo-protection was achieved by stepwise transferral of the crystal to artificial mother liquor containing 25% glycerol. The crystal was flash-frozen in liquid nitrogen. Crystals of the selenomethionine-substituted protein were grown under similar conditions, except that the protein solution also contained 2 mM dithiothreitol and 0.1 mM EDTA. The selenomethionine-substituted crystals took about 1 month to grow to a full size of about 50 μm.

**Structure determination.** Numerous data sets were collected at 100 K from crystals of the native membrane protein, the selenomethionine-substituted protein, as well as those soaked with various heavy-atom salts, on beamlines X6A, X26C and X29 at Brookhaven National Laboratory–National Synchrotron Light Source (BNL–NSLS). The selenomethionine crystals were small and sensitive to radiation. Diffraction data collected from five crystals were merged to generate the final MAD data set used for solving the selenium substructure and for phasing. All diffraction images were processed by HKL2000 (ref. 40). The ten selenium sites in the asymmetric unit of the crystal were found by hkl2map<sup>41</sup>. These sites were used to calculate phases to 2.8 Å resolution. The experimental phases were improved and extended to 2.1 Å resolution by solvent flattening and histogram matching using the program dm of the CCP4 suite<sup>42</sup>, and at a later stage by incorporating phase values calculated from a refined model. The initial polypeptide model was built into the density-modified electron density map using O<sup>43</sup>, guided by the known selenium coordinates. This model was improved by iterative cycles of manual fitting using O and refinement by CNS<sup>44</sup>. On the basis of clear non-protein densities, detergent and water molecules were added to the model in the final steps of the refinement.

Received 19 July; accepted 18 September 2006.

Published online 11 October 2006.

1. Brown, M. S. & Goldstein, J. L. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* **89**, 331–340 (1997).
2. Mattson, M. P. Pathways towards and away from Alzheimer's disease. *Nature* **430**, 631–639 (2004).

3. Sakai, J. *et al.* Sterol-regulated release of SREBP-2 from cell membranes requires two sequential cleavages, one within a transmembrane segment. *Cell* **85**, 1037–1046 (1996).
4. Rawson, R. B. *et al.* Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs. *Mol. Cell* **1**, 47–57 (1997).
5. Haass, C. Take five—BACE and the  $\gamma$ -secretase quartet conduct Alzheimer's amyloid  $\beta$ -peptide generation. *EMBO J.* **23**, 483–488 (2004).
6. Levitan, D. & Greenwald, I. Facilitation of *lin-12*-mediated signalling by *sel-12*, a *Caenorhabditis elegans* S182 Alzheimer's disease gene. *Nature* **377**, 351–354 (1995).
7. De Strooper, B. *et al.* A presenilin-1-dependent  $\gamma$ -secretase-like protease mediates release of Notch intracellular domain. *Nature* **398**, 518–522 (1999).
8. Rudner, D. Z., Fawcett, P. & Losick, R. A family of membrane-embedded metalloproteases involved in regulated proteolysis of membrane-associated transcription factors. *Proc. Natl Acad. Sci. USA* **96**, 14765–14770 (1999).
9. Weihofen, A., Binns, K., Lemberg, M. K., Ashman, K. & Martoglio, B. Identification of signal peptide peptidase, a presenilin-type aspartic protease. *Science* **296**, 2215–2218 (2002).
10. Fluhrer, R. *et al.* A  $\gamma$ -secretase-like intramembrane cleavage of TNF $\alpha$  by the GxGD aspartyl protease SPPL2b. *Nature Cell Biol.* **8**, 894–896 (2006).
11. Friedmann, E. *et al.* SPPL2a and SPPL2b promote intramembrane proteolysis of TNF $\alpha$  in activated dendritic cells to trigger IL-12 production. *Nature Cell Biol.* **8**, 843–848 (2006).
12. Urban, S., Lee, J. R. & Freeman, M. *Drosophila* rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell* **107**, 173–182 (2001).
13. Gallio, M., Sturgill, G., Rather, P. & Kylsten, P. A conserved mechanism for extracellular signalling in eukaryotes and prokaryotes. *Proc. Natl Acad. Sci. USA* **99**, 12208–12213 (2002).
14. Wasserman, J. D., Urban, S. & Freeman, M. A family of rhomboid-like genes: *Drosophila* rhomboid-1 and rhomboid-3 cooperate to activate EGF receptor signaling. *Genes Dev.* **14**, 1651–1663 (2000).
15. Koonin, E. V. *et al.* The rhomboids: a nearly ubiquitous family of intramembrane serine proteases that probably evolved by multiple ancient horizontal gene transfers. *Genome Biol.* **4**, R19 (2003).
16. Urban, S., Schlieper, D. & Freeman, M. Conservation of intramembrane proteolytic activity and substrate specificity in prokaryotic and eukaryotic rhomboids. *Curr. Biol.* **12**, 1507–1512 (2002).
17. Sturtevant, M. A., Roark, M. & Bier, E. The *Drosophila* rhomboid gene mediates the localized formation of wing veins and interacts genetically with components of the EGF-R signaling pathway. *Genes Dev.* **7**, 961–973 (1993).
18. McQuibban, G. A., Saurya, S. & Freeman, M. Mitochondrial membrane remodelling regulated by a conserved rhomboid protease. *Nature* **423**, 537–541 (2003).
19. Cipolat, S. *et al.* Mitochondrial rhomboid PARL regulates cytochrome c release during apoptosis via OPA1-dependent cristae remodeling. *Cell* **126**, 163–175 (2006).
20. Brossier, F., Jewett, T. J., Sibley, L. D. & Urban, S. A spatially localized rhomboid protease cleaves cell surface adhesins essential for invasion by *Toxoplasma*. *Proc. Natl Acad. Sci. USA* **102**, 4146–4151 (2005).
21. Maegawa, S., Ito, K. & Akiyama, Y. Proteolytic action of GlpG, a rhomboid protease in the *Escherichia coli* cytoplasmic membrane. *Biochemistry* **44**, 13543–13552 (2005).
22. Brown, M. S., Ye, J., Rawson, R. B. & Goldstein, J. L. Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. *Cell* **100**, 391–398 (2000).
23. Wolfe, M. S. & Kopan, R. Intramembrane proteolysis: theme and variations. *Science* **305**, 1119–1123 (2004).
24. Wolfe, M. S. *et al.* Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and  $\gamma$ -secretase activity. *Nature* **398**, 513–517 (1999).
25. Daley, D. O. *et al.* Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* **308**, 1321–1323 (2005).
26. Urban, S. & Wolfe, M. S. Reconstitution of intramembrane proteolysis *in vitro* reveals that pure rhomboid is sufficient for catalysis and specificity. *Proc. Natl Acad. Sci. USA* **102**, 1883–1888 (2005).
27. Lemberg, M. K. *et al.* Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases. *EMBO J.* **24**, 464–472 (2005).
28. Russ, W. P. & Engelman, D. M. The GxxxG motif: a framework for transmembrane helix-helix association. *J. Mol. Biol.* **296**, 911–919 (2000).
29. Cleland, W. W., Frey, P. A. & Gerlt, J. A. The low barrier hydrogen bond in enzymatic catalysis. *J. Biol. Chem.* **273**, 25529–25532 (1998).
30. Wei, Y. *et al.* A novel variant of the catalytic triad in the *Streptomyces scabies* esterase. *Nature Struct. Biol.* **2**, 218–223 (1995).
31. Zhou, G. W., Guo, J., Huang, W., Fletterick, R. J. & Scanlan, T. S. Crystal structure of a catalytic antibody with a serine protease active site. *Science* **265**, 1059–1064 (1994).
32. Paetzel, M., Dalbey, R. E. & Strynadka, N. C. Crystal structure of a bacterial signal peptidase in complex with a  $\beta$ -lactam inhibitor. *Nature* **396**, 186–190 (1998).
33. Tjalsma, H. *et al.* Conserved serine and histidine residues are critical for activity of the ER-type signal peptidase SipW of *Bacillus subtilis*. *J. Biol. Chem.* **275**, 25102–25108 (2000).
34. Fersht, A. *Structure and Mechanism in Protein Science: a Guide to Enzyme Catalysis and Protein Folding* (W.H. Freeman, New York, 1999).

35. Urban, S. & Freeman, M. Substrate specificity of rhomboid intramembrane proteases is governed by helix-breaking residues in the substrate transmembrane domain. *Mol. Cell* **11**, 1425–1434 (2003).
36. Ye, J., Dave, U. P., Grishin, N. V., Goldstein, J. L. & Brown, M. S. Asparagine-proline sequence within membrane-spanning segment of SREBP triggers intramembrane cleavage by site-2 protease. *Proc. Natl Acad. Sci. USA* **97**, 5123–5128 (2000).
37. Lazarov, V. K. *et al.* Electron microscopic structure of purified, active  $\gamma$ -secretase reveals an aqueous intramembrane chamber and two pores. *Proc. Natl Acad. Sci. USA* **103**, 6889–6894 (2006).
38. Li, X. & Greenwald, I. Membrane topology of the *C. elegans* SEL-12 presenilin. *Neuron* **17**, 101510–101521 (1996).
39. Li, X. & Greenwald, I. Additional evidence for an eight-transmembrane-domain topology for *Caenorhabditis elegans* and human presenilins. *Proc. Natl Acad. Sci. USA* **95**, 7109–7114 (1998).
40. Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
41. Pape, T. & Schneider, T. R. HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Crystallogr.* **37**, 843–844 (2004).
42. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
43. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
44. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
45. Opitz, C. *et al.* Intramembrane cleavage of microneme proteins at the surface of the apicomplexan parasite *Toxoplasma gondii*. *EMBO J.* **21**, 1577–1585 (2002).
46. Zhou, X. W., Blackman, M. J., Howell, S. A. & Carruthers, V. B. Proteomic analysis of cleavage events reveals a dynamic two-step mechanism for proteolysis of a key parasite adhesive complex. *Mol. Cell. Proteomics* **3**, 565–576 (2004).
47. Kraulis, P. J. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
48. Merritt, E. A. & Murphy, M. E. Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873 (1994).
49. Nicholls, A., Sharp, K. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1991).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank V. Stojanoff, H. Robinson, A. Saxena and A. Héroux at BNL NSLS beamlines for help; T. Boggon and J. Schlessinger for sharing the crystallization robot in their laboratories; and B. Turk for sharing the fluorescence spectrometer in his laboratory. X-ray diffraction data for this study were measured at beamlines X6A, X29 and X26C of NSLS. Financial support comes principally from the US Department of Energy, and from the National Institutes of Health. This work was supported by a New Scholar Award in Aging from the Ellison Medical Foundation (to Y.H.) and a gift from the Neuroscience Education and Research Foundation (to Y.H.).

**Author Contributions** Y.W. and Y.H. purified and characterized GlpG in various detergents. Y.W. crystallized GlpG. Y.H. and Y.W. solved the structure of GlpG and wrote the paper. Y.Z. screened the expression of many constructs and conducted some initial biochemical characterizations.

**Author Information** The atomic coordinates of GlpG have been deposited in the Protein Data Bank (accession number 2IC8), and will be released upon publication of the paper. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to Y.H. ([ya.ha@yale.edu](mailto:ya.ha@yale.edu)).