Published in final edited form as:
Biochemistry. 2012 May 8; 51(18): 3723–3731. doi:10.1021/bi300368b.

The conformational change in rhomboid protease GlpG induced by inhibitor binding to its S'-subsites

Yi Xue[†], **Somenath Chowdhury**[‡], **Xuying Liu**[†,¶], **Yoshinori Akiyama**[§], **Jonathan Ellman**[†,‡], and **Ya Ha**[†,*]

[†]Department of Pharmacology, Yale School of Medicine, New Haven, Connecticut 06520, United States

[‡]Department of Chemistry, Yale University, New Haven, Connecticut 06520, United States

[§]Institute for Virus Research, Kyoto University, Kyoto 606-8507, Japan

Abstract

Rhomboid protease conducts proteolysis inside the hydrophobic environment of the membrane. The conformational flexibility of the protease is essential for the enzyme mechanism, but the nature of this flexibility is not completely understood. Here we describe the crystal structure of rhomboid protease GlpG in complex with a phosphonofluoridate inhibitor, which is covalently bonded to the catalytic serine, and extends into the S' side of the substrate binding cleft. Inhibitor binding causes subtle but extensive changes in the membrane protease. Many transmembrane helices tilt and shift positions, and the gap between S2 and S5 is slightly widened so that the inhibitor can bind between them. The side chain of Phe-245 from a loop (L5) that acts as a cap rotates and uncovers the opening of the substrate binding cleft to the lipid bilayer. A concurrent turn of the polypeptide backbone at Phe-245 moves the rest of the cap and exposes the catalytic serine to aqueous solution. This study, together with earlier crystallographic investigation of smaller inhibitors, suggests a simple model to explain substrate binding to rhomboid protease.

Rhomboid proteases have many important functions in biology (1–3). In *Drosophila* where the protease family was first discovered, rhomboid-1 controls the proteolytic release of epidermal growth factors from the membrane, which is essential for their activation (4–7). In mitochondria, rhomboid protease PARL (or its yeast homolog Pcp1/Rbd1) is involved in membrane dynamics and apoptosis by cleaving OPA1 (Mgm1 in yeast), a dynamin-like GTPase (8–12). Rhomboid protease AarA from *Providencia stuartii* removes a leader sequence from TatA, the major subunit of the twin arginine protein translocase, and activates the channel (13–15). Inactivation of AarA prevents the transport of a quorum sensing signal through the channel, resulting in the loss of intercellular communication. Recent breakthroughs in parasite genetics showed that rhomboid proteases also play an important role in host cell invasion by *Plasmodium falciparum* and *Toxoplasma gondii*, the

[*]To whom correspondence should be addressed: Department of Pharmacology, Yale School of Medicine, 333 Cedar Street, New Haven, Connecticut 06520, United States. Tel.: 203-785-7530; Fax: 203-785-7670; ya.ha@yale.edu.
[¶]Author present address: Division of Experimental Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, 02115

The atomic coordinates and structure factors (code 3ubb) have been deposited in the protein data bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

Supporting Information Available
Additional analyses of the crystal structure of GlpG:CAPF complex are provided in supplemental figures 1–4. This material is available free of charge via the Internet at http://pubs.acs.org.

causative agents of human malaria and toxoplasmosis, due to their activity against various transmembrane (TM) adhesins (16–22).

The Ser-His catalytic dyad of rhomboid protease is buried inside the TM domain of the protein (7, 23, 24). The protease's substrates also span the lipid bilayer (the cleavage site is near the N-terminus of the TM helix). At least three conformational states of the membrane protease can be predicted based on mechanistic features of the serine protease and the consideration that the protease's active site, which is filled with water, should not be widely exposed to lipid (Fig. 1). In the absence of substrate or inhibitor, the protease is closed (I): the native crystal structure of *E. coli* rhomboid protease GlpG shows that one of the entrances to the protease's active site is shallowly submerged below the membrane surface (24); this lateral opening is blocked by residues from a flexible loop we previously called the L5 cap (25; see schematic diagram in Fig. 1). When substrate binds to the protease, the structure around the lateral opening has to change so that the peptide can go through it to reach the active site, but details of this new conformation are not well understood (II). The majority of the substrate's TM domain, which is on the C-terminal side of the scissile bond, cannot fit inside the protease. Whether it engages in binding to the protease outside the active site or not is also currently unclear. After the nucleophilic attack of the catalytic serine on the substrate, the peptide fragment C-terminal to the scissile bond is released from the protease, which leaves the S' side of the substrate binding cleft unoccupied: the protease (acylenzyme) must change conformation again so that the lateral opening becomes closed to minimize the exposure of the aqueous active site to the lipid bilayer (III).

In this paper we describe the crystal structure of GlpG in complex with a phosphonofluoridate inhibitor, which fully traverses the S' side of the substrate binding cleft, a region occupied normally by the substrate segment between the scissile bond and the membrane-spanning sequence (dark green in Fig. 1). The crystal structure provides novel insights into the conformational changes that occur around the lateral opening and in other parts of the membrane protease to enable substrate binding.

Materials and Methods

Reagents

The detergents used in membrane protein purification and crystallization were purchased from Anatrace. Cbz-Ala^{P}(O-*i*Pr)F was prepared according to the eight step sequence previously reported by Bartlett and coworkers (26).

Protease activity and inhibition assays

A fusion protein containing maltose binding protein (MBP), the TM domain of gurken, thioredoxin (Trx) and a C-terminal octa-histidine tag was used as the substrate for GlpG. The construct is similar to that described by Strisovsky *et al.* (27), and was generated based on a MBP-Gurken-GlpG$_{91–276}$ construct (pGW475), which was initially designed for crystallographic study of the gurken-GlpG complex. The sequence of MBP-Gurken-GlpG$_{91-276}$ was subcloned into pET41b between the NdeI and XhoI sites, and the GlpG sequence was removed by double digestion with BamHI and XhoI. The Trx gene was amplified by PCR from *E. coli* genomic DNA. The PCR product was digested by BamHI/XhoI and ligated with the plasmid fragment. The recombinant fusion protein was overexpressed in *E. coli* BL21(DE3) cells: the bacteria were grown in LB media at 37°C in the presence of 40 μM kanamycin; IPTG was added (final concentration 0.4 mM) at OD$_{600}$ 0.6 to induce protein expression (37°C, 3 hours). Cell membranes were collected and resuspended in a buffer containing 50 mM sodium phosphate (pH 7.4) and 0.5 M NaCl. 2% n-decyl-β-D-maltoside (DM) was used to solubilize the membrane at room temperature. The

insoluble fraction was removed by centrifugation. The His-tagged protein was loaded onto a TALON metal affinity column (Clonetech) and eluted with 300 mM imidazole.

The cleavage reactions were performed in a 15 μL assay buffer containing 50 mM Tris (pH 8.0), 0.1 M NaCl and 0.5% NG, and each used 2 μg GlpG and 4 μg substrate fusion protein. The mixture was incubated at 37°C for 3 hours before SDS-PAGE loading buffer was added to stop the reaction. Various amounts of inhibitor were pre-incubated with the protease at 37°C for 1 hour before substrate was added.

Structure determination

The recombinant GlpG core domain was purified in DM as previously described (24). For crystallization, the protein sample was concentrated to 10 mg/mL and dialyzed against a buffer containing 10 mM Tris (pH 7.4) and 0.5% NG for 5 days at 4°C. Crystals were grown in hanging drops over a well solution of 0.1 M Bis-tris propane (pH 7.0) and 3 M NaCl. Prior to soaking with the phosphonofluoridate inhibitor, the protease crystals were exchanged into a cryo-protecting solution containing 3 M NaCl, 10 mM Tris (pH 7.4), 0.5% NG and 20% glycerol. The inhibitor was dissolved in DMSO (50 mM), and diluted 10 times into the crystallization drop to reach a final concentration of 5 mM. After a 4 hour incubation at room temperature, the crystals were flash-frozen in liquid nitrogen. Diffraction data were collected from an in-house X-ray source (Rigaku Micromax X-ray generator with RAXIS-4 detector) at 100 K. The diffraction images were indexed, integrated, and scaled using *HKL2000* (28). The structure was solved by molecular replacement using the native structure of GlpG (PDB code 2ic8; ref. 24) as the probe. After several rounds of model building using *coot* (29), and refinement with *refmac5* (30), the inhibitor was incorporated into the model based on the difference Fourier map. A model of the inhibitor was initially generated by the PRODRG2 server (31), and the restraint dictionary file was created by *phenix.eLBOW*. The program-generated restraint file was modified to enforce the planar geometry of the amide and ester bonds. Incorporation of the inhibitor lowered the R factor by ~1%. Further positional and B factor refinements were carried out using *PHENIX* (32). The final model has an R_{free} value of 22.9%, and has good geometry (Table 1). The presence of the inhibitor, as well as deviations from the native structure, was confirmed by omit difference maps (Fig. 3A, Fig. S1 and Fig. S4 of the Supporting Information), and by the crystallographic refinement process (monitored by R_{free}). Many similar conformational changes are also observed in the GlpG:isocoumarin complex (37), which was solved at a higher resolution (2.09 Å).

Results

The crystal structure of the protease:inhibitor complex

Encouraged by the finding that GlpG can react with diisopropyl fluorophosphate (DFP) (33), we set out to study other organophosphate compounds, hoping that this may provide further insight into the substrate specificity and conformational change of rhomboid protease. The α-aminoalkyl phosphonate diphenyl ester class of inhibitors, which worked extremely well for soluble serine proteases (34), would be ideal for this purpose because substrate-mimicking peptide fragments could be readily incorporated into the inhibitor by organic synthesis (e.g., 35, 36). Unfortunately, the initial experiments using a number of diphenyl esters (kind gift from Professor JC Powers) showed that none of them bound to GlpG, due to either steric hindrance or poor reactivity. As a result, we switched back to inhibitors with a phosphonofluoridate functional group, which is less bulky and more reactive. We synthesized Cbz-Ala^{P}(O-*i*Pr)F (CAPF) (Fig. 2A; ref. 26), and found that it inhibited GlpG-catalyzed proteolysis of the TM fusion protein substrate with an apparent IC50 of around

50μM (27) (Fig. 2B, C). The compound also readily reacted with the crystalline protease to yield a covalent adduct whose structure could be analyzed by x-ray diffraction (Fig. 2A).

In the crystal structure, the carboxybenzyl (Cbz) end of the inhibitor occupies the S' side of the substrate binding cleft (Fig. 3A, B). Therefore, the orientation of the Ala^P moiety is opposite to that postulated for peptide substrates. The Cbz group of the inhibitor is wedged between the L5 cap (Phe-245) and a hydrophobic patch (Met-149, Phe-153) near the N-terminus of TM helix S2 (Fig. 3C). The carboxylate oxygen, which may mimic the carbonyl oxygen of substrate's P1 residue (Fig. 2A), points downward into a polar cavity surrounded by Asn-154, Ser-201, Tyr-205 and Trp-236. The Cα atom of the Ala^P moiety occupies the position of the amide nitrogen of the scissile bond, to which the catalytic histidine (His-254) is expected to donate a proton during the breakdown of the first tetrahedral intermediate. The presence of a methyl side chain at the Cα pushed the histidine up and away from the inhibitor. The phosphonyl oxygen of the inhibitor has rotated slightly out of the oxyanion hole, and is hydrogen bonded only to His-150 (Fig. S2 of the Supporting Information). The isopropyl group of the inhibitor points toward a region earlier mutagenesis study identified as the S1 pocket (37). At the base of the binding cleft, the side chains of Tyr-205 and Trp-236 show similar rotations as those observed in the complexes with DFP and isocoumarin (33, 37) (Fig. S3 of the Supporting Information). These concerted rotations may explain why mutations around Trp-236 affect the activity of the protease (40): the change of the side chain packing may facilitate the movement of Tyr-205 into an "active" position as suggested by the complexes with mechanism-based inhibitors.

The conformational flexibility of the protease

Inhibitor binding caused subtle but extensive changes in the membrane protease's structure. To help visualize these changes, we superimposed the structure of the inhibitor complex onto that of the native protease based on the Cα positions of the central helix S4 (residues 201 to 216): the catalytic serine (Ser-201) is located at the end of S4, and its position serves as a good reference point. When viewed from the outside of the cell onto the membrane plane, it appears that TM helices S2, S1, S3 and S6, as well as the peripheral L1 loop, rotated counterclockwise around S4 by a few degrees, whereas TM helix S5 rotated in the opposite direction (Fig. 4A, B). Most helices moved as rigid bodies, with a Cα r.m.s.d. of less than 0.4 Å. S6 is the only exception, and its high r.m.s.d. (1.5 Å) is due to a major rearrangement towards the helix's N-terminus (see below). The L5 cap has the largest change in structure (r.m.s.d. 2.1 Å; see below), and its center of mass shifted by 3.2 Å. This is followed by S5, L1 and S6, which moved by 1.6 Å, 1.2 Å and 1.0 Å, respectively. As a result of the helix movement, the gap between S2 and S5, to which the inhibitor is bound, has become slightly widened. A careful examination of the superposition also shows that the helices not only shifted positions, but also changed their tilting angles (Fig. 5A). S6 (10.3°) and S5 (7.5°) show the largest tilting motions. In comparison, the L1 loop rotated about 4.9°.

The conformational changes identified here are also observed in the other two GlpG:inhibitor complexes (33, 37). The movements of S2, S1, L1, S3 and S5 in the complex with the structurally unrelated isocoumarin are strikingly similar to those in the CAPF complex (Fig. 4C). The only major difference is S6, which, in the isocoumarin complex, is drawn towards the inhibitor to form a covalent bond. The movements in the DFP complex, although similar in trend, are smaller in amplitude, suggesting that the degree of conformational change may be correlated with the size of the inhibitor or substrate that occupies the active site (Fig. 4D). It is interesting that, when the native structures of *E. coli* and *H. influenzae* GlpGs are compared, similar rotations of L1, S3, S6 and S5 around the central helix S4, whose position did not change during evolution, are also observed (24, 38). This similarity lends further support to the hypothesis that the movements of various

secondary structural elements described above reflect the intrinsic conformational flexibility of the membrane protease.

The model of conformational change proposed here differs from an earlier lateral gating model in several major aspects (39, 40). The earlier model was based on the crystal structure of an apo protein, where TM helix S5 has moved to engage in crystal packing interactions (39). The movement of S5 in the apo structure is very different from what we observed in the inhibitor complexes, both in direction and in amplitude. The apo structure also lacks changes in the other structural elements (e.g., L1, S6) that our present study suggests may occur concurrently with the S5 movement, and lacks the rotation of key side chains (Tyr-205, Trp-236) within the active site.

Conformational change in TM helix S6

In the native structure, S6 has a small kink near Gly-257 (the distance between the carbonyl oxygen of Gly-257 and the amide nitrogen of Gly-261 is 3.4 Å). Inhibitor binding causes a significant change in the helix on the N-terminal side of the kink, where the helix has deformed into two short 3_{10} helices (Fig. 5A, B). A water molecule is inserted between the two short helices, relaying hydrogen bond from the carbonyl oxygen of His-254 to the amide nitrogen of Gly-257. The insertion causes the N-terminus of the helix to bend towards S5, and pushes Gly-257 closer to Gly-261 (the distance between Gly-257's carbonyl oxygen and Gly-261's amide nitrogen has shortened to 2.9 Å). The deformed S6 is stabilized by several new interactions with a neighboring helix S3: Ser-185 from S3 forms a hydrogen bond with the water molecule bound between the two 3_{10} helices, whereas Gln-189 contributes two additional hydrogen bonds, one with the carbonyl oxygen of Ala-253, and another with the imidazole nitrogen of His-254.

Conformational change in the L5 cap

The Cbz group of the inhibitor emerges from the side of the protease, a few angstroms below the predicted membrane surface, through an opening surrounded by Met-149, Phe-153, Trp-236 and Phe-245 (Fig. 6A). The lateral opening is obstructed in the native structure by the side chain of Phe-245, which adopts a different conformation (Fig. 6B). To accommodate the inhibitor, the L5 cap and TM helix S5 moved slightly away from S2, but the amplitude of this movement is small: the distance between the Cα atoms of Phe-245 (L5) and Met-149 (S2) only increased marginally, from 8.5 Å to 9.2 Å. The side chain of Phe-245 broke off van der Waals contact with Met-149, and rotated downward (the edge of the phenyl ring now points towards the lipid; Fig. 6C). The distance between Phe-245 and Met-149 doubled from 3.8 Å to 7.6 Å. The backbone Φ/Ψ angles of Phe-245 also changed, and as a result, the region of the L5 cap downstream of Phe-245 was raised and removed from the substrate binding cleft, which exposes the catalytic dyad to aqueous solution.

In the GlpG:isocoumarin complex (37), although the inhibitor does not extend far into the S' side of the substrate binding cleft (Fig. S3 of the Supporting Information), the protein structure at the lateral opening is strikingly similar to that of the CAPF complex (Fig. 7A): the S2-S5 gap is widened to a similar degree; the N-terminal portion of the L5 cap, between Asp-243 and Met-247, can be nicely superimposed; the conformation of Phe-245's side chain is almost identical between the two structures. These similarities suggest that the observed change at the lateral opening probably reflects an intrinsic property of the membrane protease, and does not result from specific interactions with the inhibitor. In the isocoumarin complex, a detergent molecule was found at the position of the Cbz group (red in Fig. 7A). This observation raises the possibility that the hydrophobic patch (Met-149, Phe-153), with which the detergent and the Cbz group both interact, may represent the binding site for a conserved hydrophobic group on the substrate, e.g., the P2' side chain

(27). The C-terminal portion of the L5 cap, as well as the N-terminal region of S6, is very different between the two inhibitor structures. It is difficult to predict which one is a better mimic of the substrate complex, because both inhibitor structures bear features here that are artificial: the isocoumarin forms a covalent bond with His-254; the Cα methyl group in CAPF seems to have pushed the same histidine out of the active site.

Discussion

The ability of membrane proteases to change conformation is essential for their catalytic activities (41). For rhomboid protease GlpG, after its crystal structure was solved (24, 38, 39, 42), attention was immediately focused on two structural elements, the L5 cap and TM helix S5 (25, 39, 43). The fact that the closed L5 cap denies access to the Ser-His catalytic dyad suggests that the loop has to move during catalysis. Crystallographic and biochemical studies indicated that this might happen spontaneously (25, 44). Since S5 is packed loosely against the other TM helices, and its structure can be altered by the interaction with neighboring molecules in the crystal lattice (39), it has been suggested that the helix may undergo a large conformational change and function as the lateral gate for substrate entry (40). Since then, the crystal structures of GlpG in complex with an isocoumarin inhibitor (37), and with DFP (33), have been solved. In these structures, the L5 cap has lifted up from the catalytic dyad, but the movement of S5, in contrast with the prediction, is relatively small. Since neither the isocoumarin nor DFP reaches the gap between S2 and S5, it was uncertain initially whether a larger movement of S5 will occur when peptide substrate is bound between the two helices. The crystal structure reported here describes the conformation of the membrane protease in complex with an inhibitor that spans the full length of the S' side of the substrate binding cleft and reaches the S2-S5 gap. Again, the movement of S5 is found to be small: the helix tilted by 7.5°, and its center of mass shifted only by 1.6 Å. These values are comparable to those observed in the isocoumarin complex (37).

Besides the L5 cap and S5, several other structural elements also moved in the inhibitor complexes (Fig. 4). The movements are all small in amplitude, and occur in different directions. One interesting observation is that the peripheral L1 loop appears to have rotated deeper into the membrane, around an axis that is roughly parallel to the rotation axis of S5 (Fig. 5A). The possibility that their movements are coupled to each other, and to the movements of other structural elements, may explain the effects of some mutations on the L1 loop that are distal to the active site. For example, W136A and R137A, both greatly reducing GlpG's enzymatic activity, could have affected the interaction of the loop with the lipid and hindered its movement (7, 45). The conformational flexibility observed in TM helix S6, from Asn-251 to Gly-257, was not expected. The significance of this flexibility in the enzyme mechanism is also unclear. Maybe it provides L5 cap with more freedom to move, however another possibility is that a conformational change in S6 will reshape its interface with S3, which may function as the binding site for the N-terminal region of the substrate (red in Fig. 1).

The observation that the TM helices in GlpG moved only slightly in response to inhibitor binding suggests a simple model to explain substrate binding to the membrane protease, where access to the protease's active site is regulated primarily by a conformational change in the L5 cap (Fig. 7B): the rotation of Phe-245's side chain uncovers the lateral opening of the active site, which is located slightly below the membrane surface; a concurrent turn of the polypeptide backbone at Phe-245 removes the cap completely from the substrate binding cleft, exposing the catalytic dyad to aqueous solution; over the lateral opening the membrane-spanning substrate can bend 90° into the active site in an extended conformation to be cleaved. Based on similarities between the CAPF and isocoumarin structures (Fig.

7A), it is even tempting to speculate that Phe-245, when folded down, may lock into a rigid structure with Ala-239 and Trp-236 through van der Waals forces, and move together with S5 to fine-tune the width of the lateral opening.

The HtrA family of proteases (HtrA, DegP, DegS, Rv3671c) provides some excellent examples for the types of conformational change that occur in soluble serine proteases (other well-known examples include the activation of trypsin and thrombin; see ref. 46, 47). In *Thermotoga maritima* HtrA, the protease's active site is covered by a helical lid (48). High temperature induces a conformational change in HtrA (in which the lid is lifted), and switches on the protease (49). In *E. coli* DegP, oligomerization enables the LA loop of one molecule to insert into a neighbor, blocking and distorting its active site (50–52). In *E. coli* DegS and *Mycobacterium tuberculosis* Rv3671c, the protease's active site is not initially properly formed: peptide binding (folding stress) to a C-terminal PDZ domain in DegS (53), or disulfide bond formation (oxidative stress) in Rv3671c (54), triggers major structural rearrangements that transform the protein into a functional protease. In light of these well characterized systems, we can see that the conformational change in rhomboid protease involves some similar, but not identical, structural principles. In the absence of substrate, the protease's active site is blocked by the L5 cap (like HtrA); although the catalytic Ser-201 and His-254 are already properly aligned (unlike Rv3671c), the side chains of Tyr-205 and Trp-236 do have different conformations than those observed in the inhibitor complexes (the significance of the side chain rotations in the catalytic mechanism is not yet known). Unlike the HtrA proteases, however, the enzymatic activity of GlpG is not allosterically regulated, and the opening of the L5 cap appears to be spontaneous (25, 44). Finally, it is important to emphasize that, since rhomboid protease is embedded in the membrane, the protein structure around the lateral opening needs to be constantly adjusted so that the energetically unfavorable contact between the aqueous active site and the surrounding lipid is minimized. The conformational change in soluble serine proteases is usually not driven by such a requirement because they conduct catalysis in a homogenous and aqueous environment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Professor James C. Powers for sharing the phosphonate diphenyl esters compounds.

This work was supported by NIH grant GM082839 (Y.H.).

The abbreviations used are

CAPF	Cbz-Ala^{P}(O-*i*Pr)F
Cbz	carboxybenzyl
DFP	diisopropyl fluorophosphate
MBP	maltose binding protein
DM	n-Decyl-β-D-maltopyranoside
NG	n-Nonyl-β-D-glucopyranoside
TM	transmembrane
Trx	thioredoxin

References

1. Freeman M. Rhomboid proteases and their biological functions. Annu Rev Genet. 2008; 42:191–210. [PubMed: 18605900]
2. Hill RB, Pellegrini L. The PARL family of mitochondrial rhomboid proteases. Semin Cell Dev Biol. 2010; 21:582–592. [PubMed: 20045481]
3. Urban S. Rhomboid proteins: conserved membrane proteases with divergent biological functions. Genes Dev. 2006; 20:3054–3068. [PubMed: 17114579]
4. Mayer U, Nusslein-Volhard C. A group of genes required for pattern formation in the ventral ectoderm of the Drosophila embryo. Genes Dev. 1988; 2:1496–1511. [PubMed: 3209069]
5. Wasserman JD, Urban S, Freeman M. A family of rhomboid-like genes: Drosophila rhomboid-1 and roughoid/rhomboid-3 cooperate to activate EGF receptor signaling. Genes Dev. 2000; 14:1651–1663. [PubMed: 10887159]
6. Lee JR, Urban S, Garvey CF, Freeman M. Regulated intracellular ligand transport and proteolysis control EGF signal activation in Drosophila. Cell. 2001; 107:161–171. [PubMed: 11672524]
7. Urban S, Lee JR, Freeman M. Drosophila rhomboid-1 defines a family of putative intramembrane serine proteases. Cell. 2001; 107:173–182. [PubMed: 11672525]
8. Esser K, Tursun B, Ingenhoven M, Michaelis G, Pratje E. A novel two-step mechanism for removal of a mitochondrial signal sequence involves the mAAA complex and the putative rhomboid protease Pcp1. J Mol Biol. 2002; 323:835–843. [PubMed: 12417197]
9. McQuibban GA, Saurya S, Freeman M. Mitochondrial membrane remodelling regulated by a conserved rhomboid protease. Nature. 2003; 423:537–541. [PubMed: 12774122]
10. Herlan M, Vogel F, Bornhovd C, Neupert W, Reichert AS. Processing of Mgm1 by the rhomboid-type protease Pcp1 is required for maintenance of mitochondrial morphology and of mitochondrial DNA. J Biol Chem. 2003; 278:27781–27788. [PubMed: 12707284]
11. Cipolat S, Rudka T, Hartmann D, Costa V, Serneels L, Craessaerts K, Metzger K, Frezza C, Annaert W, D'Adamio L, Derks C, Dejaegere T, Pellegrini L, D'Hooge R, Scorrano L, De Strooper B. Mitochondrial rhomboid PARL regulates cytochrome c release during apoptosis via OPA1-dependent cristae remodeling. Cell. 2006; 126:163–175. [PubMed: 16839884]
12. Chao JR, Parganas E, Boyd K, Hong CY, Opferman JT, Ihle JN. Hax1-mediated processing of HtrA2 by Parl allows survival of lymphocytes and neurons. Nature. 2008; 452:98–102. [PubMed: 18288109]
13. Rather PN, Ding X, Baca-DeLancey RR, Siddiqui S. Providencia stuartii genes activated by cell-to-cell signaling and identification of a gene required for production or activity of an extracellular factor. J Bacteriol. 1999; 181:7185–7191. [PubMed: 10572119]
14. Gallio M, Sturgill G, Rather P, Kylsten P. A conserved mechanism for extracellular signaling in eukaryotes and prokaryotes. Proc Natl Acad Sci U S A. 2002; 99:12208–12213. [PubMed: 12221285]
15. Stevenson LG, Strisovsky K, Clemmer KM, Bhatt S, Freeman M, Rather PN. Rhomboid protease AarA mediates quorum-sensing in Providencia stuartii by activating TatA of the twin-arginine translocase. Proc Natl Acad Sci U S A. 2007; 104:1003–1008. [PubMed: 17215357]
16. Opitz C, Di Cristina M, Reiss M, Ruppert T, Crisanti A, Soldati D. Intramembrane cleavage of microneme proteins at the surface of the apicomplexan parasite Toxoplasma gondii. Embo J. 2002; 21:1577–1585. [PubMed: 11927542]
17. Zhou XW, Blackman MJ, Howell SA, Carruthers VB. Proteomic analysis of cleavage events reveals a dynamic two-step mechanism for proteolysis of a key parasite adhesive complex. Mol Cell Proteomics. 2004; 3:565–576. [PubMed: 14982962]
18. Brossier F, Jewett TJ, Sibley LD, Urban S. A spatially localized rhomboid protease cleaves cell surface adhesins essential for invasion by Toxoplasma. Proc Natl Acad Sci U S A. 2005; 102:4146–4151. [PubMed: 15753289]
19. O'Donnell RA, Hackett F, Howell SA, Treeck M, Struck N, Krnajski Z, Withers-Martinez C, Gilberger TW, Blackman MJ. Intramembrane proteolysis mediates shedding of a key adhesin during erythrocyte invasion by the malaria parasite. J Cell Biol. 2006; 174:1023–1033. [PubMed: 17000879]

20. Baker RP, Wijetilaka R, Urban S. Two Plasmodium rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria. PLoS Pathog. 2006; 2:e113. [PubMed: 17040128]
21. Buguliskis JS, Brossier F, Shuman J, Sibley LD. Rhomboid 4 (ROM4) affects the processing of surface adhesins and facilitates host cell invasion by Toxoplasma gondii. PLoS Pathog. 2010; 6:e1000858. [PubMed: 20421941]
22. Santos JM, Ferguson DJ, Blackman MJ, Soldati-Favre D. Intramembrane cleavage of AMA1 triggers Toxoplasma to switch from an invasive to a replicative mode. Science. 2011; 331:473–477. [PubMed: 21205639]
23. Lemberg MK, Menendez J, Misik A, Garcia M, Koth CM, Freeman M. Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases. Embo J. 2005; 24:464–472. [PubMed: 15616571]
24. Wang Y, Zhang Y, Ha Y. Crystal structure of a rhomboid family intramembrane protease. Nature. 2006; 444:179–180. [PubMed: 17051161]
25. Wang Y, Ha Y. Open-cap conformation of intramembrane protease GlpG. Proc Natl Acad Sci U S A. 2007; 104:2098–2102. [PubMed: 17277078]
26. Bartlett PA, Lamden LA. Inhibition of chymotrypsin by phosphonate and phosphonamidate peptide analogs. Bioorg Chem. 1986; 14:357–377.
27. Strisovsky K, Sharpe HJ, Freeman M. Sequence-specific intramembrane proteolysis: identification of a recognition motif in rhomboid substrates. Mol Cell. 2009; 36:1048–1059. [PubMed: 20064469]
28. Otwinowski Z, Minor W. Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997; 276:307–326.
29. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010; 66:486–501. [PubMed: 20383002]
30. Winn MD, Murshudov GN, Papiz MZ. Macromolecular TLS refinement in REFMAC at moderate resolutions. Methods Enzymol. 2003; 374:300–321. [PubMed: 14696379]
31. Schuttelkopf AW, van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr. 2004; 60:1355–1363. [PubMed: 15272157]
32. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–221. [PubMed: 20124702]
33. Xue Y, Ha Y. The catalytic mechanism of rhomboid protease GlpG probed by 3,4-dichloroisocoumarin and diisopropyl fluorophosphonate. J Biol Chem. 2012; 287:3099–3107. [PubMed: 22130671]
34. Oleksyszyn J, Powers JC. Amino acid and peptide phosphonate derivatives as specific inhibitors of serine peptidases. Methods Enzymol. 1994; 244:423–441. [PubMed: 7845224]
35. Sampson NS, Bartlett PA. Peptidic phosphonylating agents as irreversible inhibitors of serine proteases and models of the tetrahedral intermediates. Biochemistry. 1991; 30:2255–2263. [PubMed: 1998684]
36. Bone R, Sampson NS, Bartlett PA, Agard DA. Crystal structures of alpha-lytic protease complexes with irreversibly bound phosphonate esters. Biochemistry. 1991; 30:2263–2272. [PubMed: 1998685]
37. Vinothkumar KR, Strisovsky K, Andreeva A, Christova Y, Verhelst S, Freeman M. The structural basis for catalysis and substrate specificity of a rhomboid protease. EMBO J. 2010; 29:3797–3809. [PubMed: 20890268]
38. Lemieux MJ, Fischer SJ, Cherney MM, Bateman KS, James MN. The crystal structure of the rhomboid peptidase from Haemophilus influenzae provides insight into intramembrane proteolysis. Proc Natl Acad Sci U S A. 2007; 104:750–754. [PubMed: 17210913]

39. Wu Z, Yan N, Feng L, Oberstein A, Yan H, Baker RP, Gu L, Jeffrey PD, Urban S, Shi Y. Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry. Nat Struct Mol Biol. 2006; 13:1084–1091. [PubMed: 17099694]
40. Baker RP, Young K, Feng L, Shi Y, Urban S. Enzymatic analysis of a rhomboid intramembrane protease implicates transmembrane helix 5 as the lateral substrate gate. Proc Natl Acad Sci U S A. 2007; 104:8257–8262. [PubMed: 17463085]
41. Ha Y. Structure and mechanism of intramembrane protease. Semin Cell Dev Biol. 2009; 20:240–250. [PubMed: 19059492]
42. Ben-Shem A, Fass D, Bibi E. Structural basis for intramembrane proteolysis by rhomboid serine proteases. Proc Natl Acad Sci U S A. 2007; 104:462–466. [PubMed: 17190827]
43. Brooks CL, Lazareno-Saez C, Lamoureux JS, Mak MW, Lemieux MJ. Insights into substrate gating in H. influenzae rhomboid. J Mol Biol. 2011; 407:687–697. [PubMed: 21295583]
44. Maegawa S, Koide K, Ito K, Akiyama Y. The intramembrane active site of GlpG, an E. coli rhomboid protease, is accessible to water and hydrolyses an extramembrane peptide bond of substrates. Mol Microbiol. 2007; 64:435–447. [PubMed: 17493126]
45. Wang Y, Maegawa S, Akiyama Y, Ha Y. The role of L1 loop in the mechanism of rhomboid intramembrane protease GlpG. J Mol Biol. 2007; 374:1104–1113. [PubMed: 17976648]
46. Stroud RM, Kossiakoff AA, Chambers JL. Mechanisms of zymogen activation. Annu Rev Biophys Bioeng. 1977; 6:177–193. [PubMed: 17350]
47. Bode W. The structure of thrombin: a janus-headed proteinase. Semin Thromb Hemost. 2006; 32:16–31. [PubMed: 16673263]
48. Kim DY, Kim DR, Ha SC, Lokanath NK, Lee CJ, Hwang HY, Kim KK. Crystal structure of the protease domain of a heat-shock protein HtrA from Thermotoga maritima. J Biol Chem. 2003; 278:6543–6551. [PubMed: 12458220]
49. Kim DY, Kwon E, Shin YK, Kweon DH, Kim KK. The mechanism of temperature-induced bacterial HtrA activation. J Mol Biol. 2008; 377:410–420. [PubMed: 18272173]
50. Krojer T, Garrido-Franco M, Huber R, Ehrmann M, Clausen T. Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. Nature. 2002; 416:455–459. [PubMed: 11919638]
51. Krojer T, Pangerl K, Kurt J, Sawa J, Stingl C, Mechtler K, Huber R, Ehrmann M, Clausen T. Interplay of PDZ and protease domain of DegP ensures efficient elimination of misfolded proteins. Proc Natl Acad Sci U S A. 2008; 105:7702–7707. [PubMed: 18505836]
52. Jiang J, Zhang X, Chen Y, Wu Y, Zhou ZH, Chang Z, Sui SF. Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. Proc Natl Acad Sci U S A. 2008; 105:11939–11944. [PubMed: 18697939]
53. Wilken C, Kitzing K, Kurzbauer R, Ehrmann M, Clausen T. Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease. Cell. 2004; 117:483–494. [PubMed: 15137941]
54. Biswas T, Small J, Vandal O, Odaira T, Deng H, Ehrt S, Tsodikov OV. Structural insight into serine protease Rv3671c that Protects M. tuberculosis from oxidative and acidic stress. Structure. 2010; 18:1353–1363. [PubMed: 20947023]

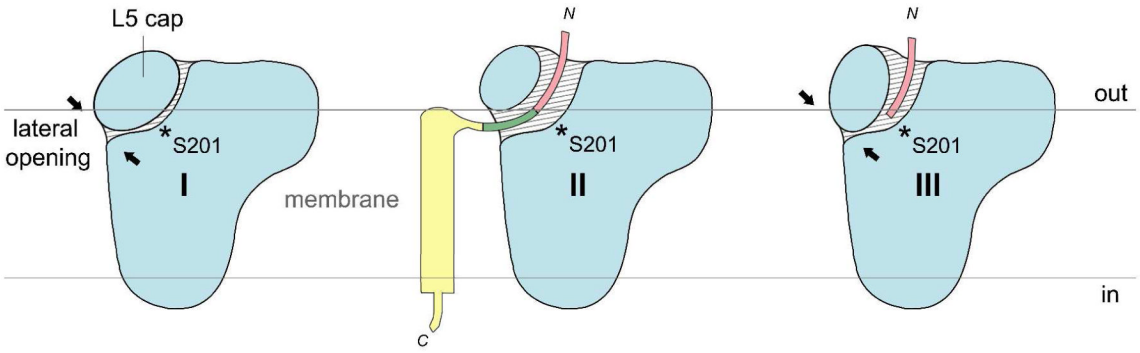

Fig. 1. A schematic diagram for the three conformational states of rhomboid protease

The two horizontal lines mark the boundaries of the hydrophobic region of the membrane. The hydrophilic active site is represented by the hatched area. The catalytic serine is denoted by the asterisk. The substrate is colored in red, green and yellow. The protease cleaves between the red and green segments.

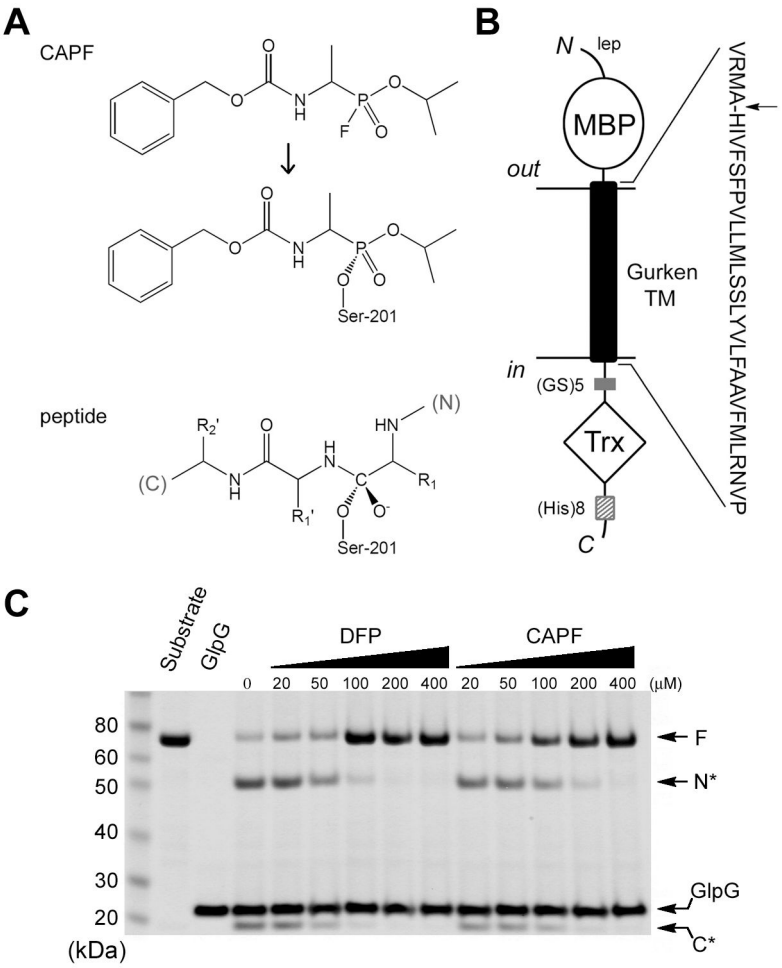

Fig. 2. CAPF inhibits the proteolytic activity of GlpG

A, The chemical structures of CAPF, its covalent adduct with the protease, and the first tetrahedral intermediate of the proteolytic reaction. **B**, A diagram of the MBP-Gurken-Trx fusion protein substrate. The fusion protein consists of a leader peptide (lep), maltose-binding protein (MBP), the transmembrane domain derived from Gurken, a short $(GS)_5$ linker, thioredoxin (Trx), and a C-terminal His-tag. The amino acid sequence of the Gurken TM domain is shown on the right with the arrow pointing to the scissile bond. **C**, *In vitro* activity and inhibition assays. GlpG cleaves the fusion substrate (F) and releases an N-terminal fragment (N*) containing MBP and a C-terminal fragment (C*) containing Trx. In the presence of CAPF or DFP, the cleavage reaction is inhibited.

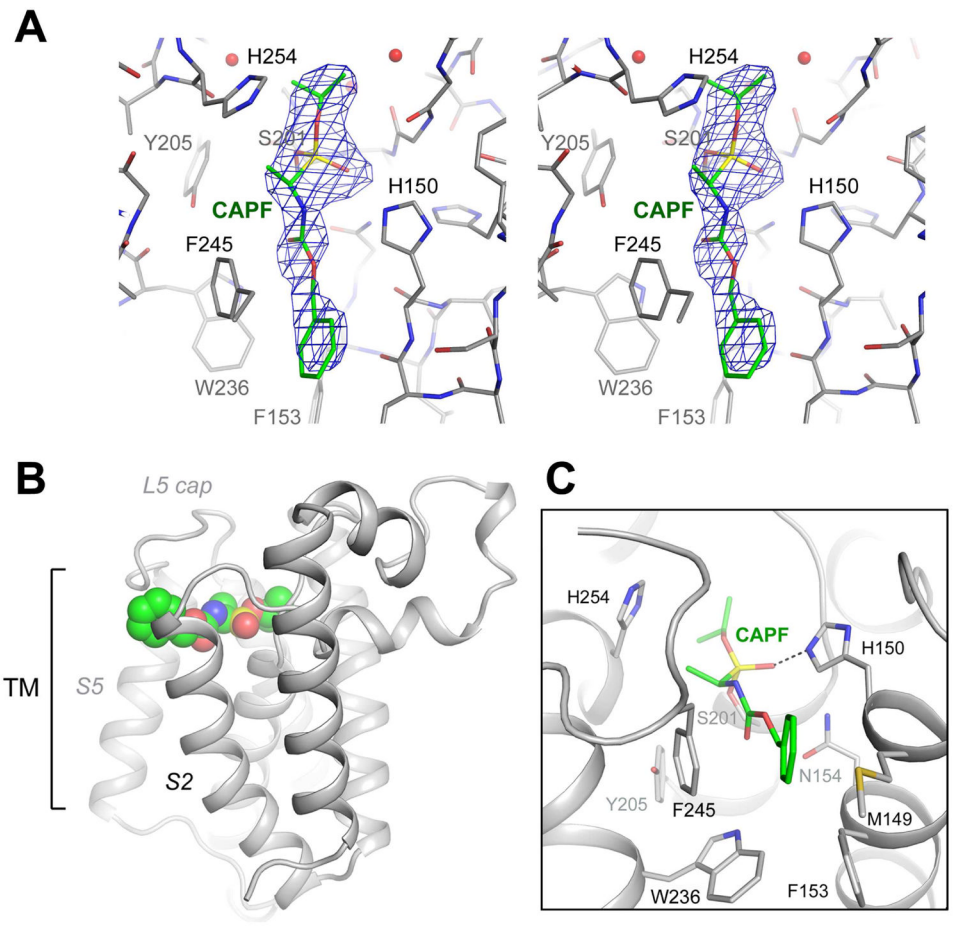

Fig. 3. The structure of the CAPF complex

A, A stereo view of the omit Fo-Fc map (generated without the inhibitor) contoured at 3σ. A different view of the map is shown in Fig. S1 of the Supporting Information. The inhibitor is shown in green. Water molecules are shown as red spheres. **B**, The overall structure. CAPF is shown as the space filling model. **C**, A close-up view of the inhibitor. CAPF is covalently bound to Ser-201. The hydrogen bond between the phosphoryl oxygen and His-150 is represented by the dashed line.

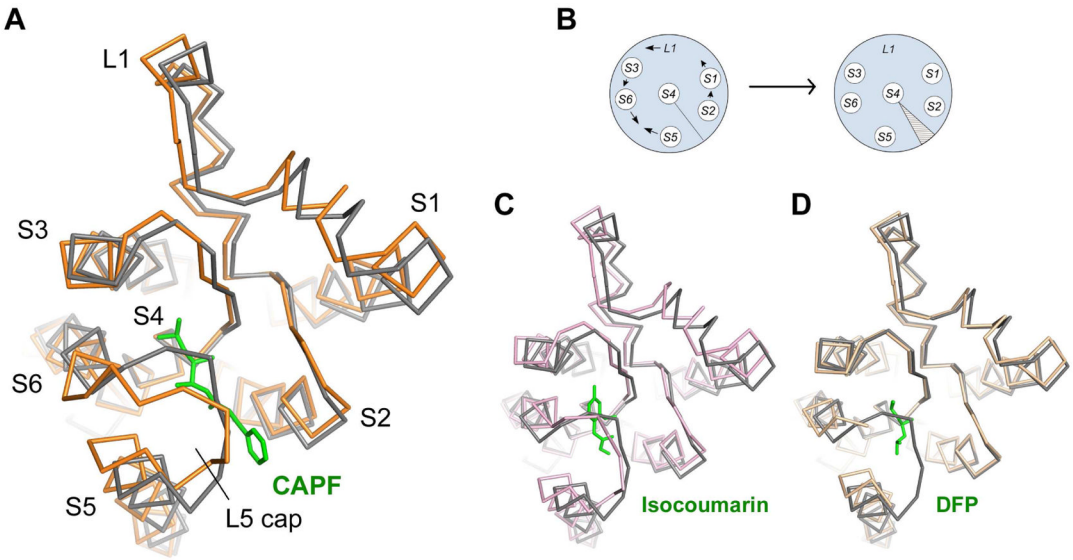

Fig. 4. The movement of the TM helices

A, The superposition of the CAPF structure (orange) onto the native structure (grey). The C_α trace of the protein is shown. The inhibitor (green) is represented by the stick model. The differences between the complex and native structures can be validated by the omit difference maps (Fig. S4 of the Supporting Information). **B**, A schematic diagram illustrating the movement of the TM helices and L1 loop. The hatched area represents the widened gap between S2 and S5 upon inhibitor binding. **C** and **D**, The isocoumarin structure (pink, PDB 2xow) and the DFP structure (sand, PDB 3txt), are similarly compared to the apo structure (grey).

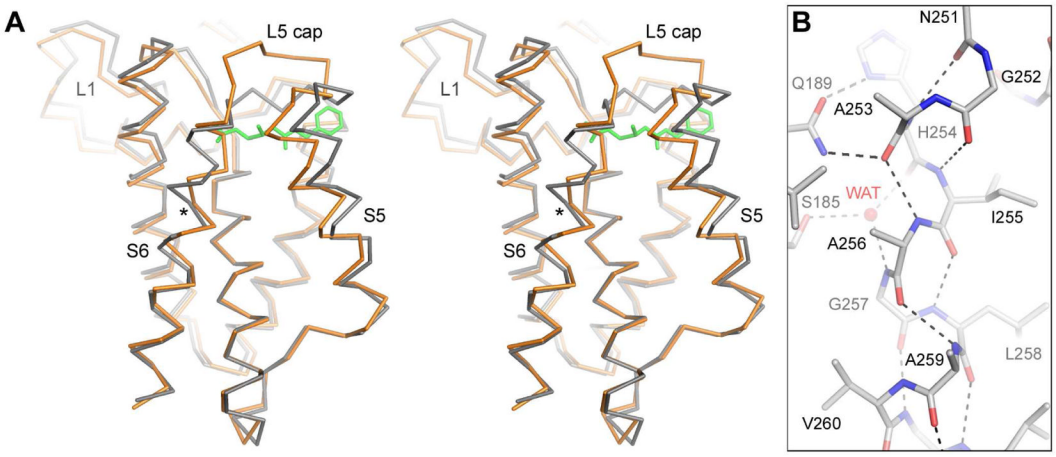

Fig. 5. The conformational change in TM helix S6

A, A stereo view of the CAPF structure (orange) superposed onto the native structure (grey). Gly-257 is indicated by the asterisk. **B**, The new S6 conformation is stabilized by a network of hydrogen bonds (dashed lines). A water molecule, shown as red sphere, is inserted between the two 3_{10} helices (Asn-251 to Ala-253, Ile-255 to Ala-256).

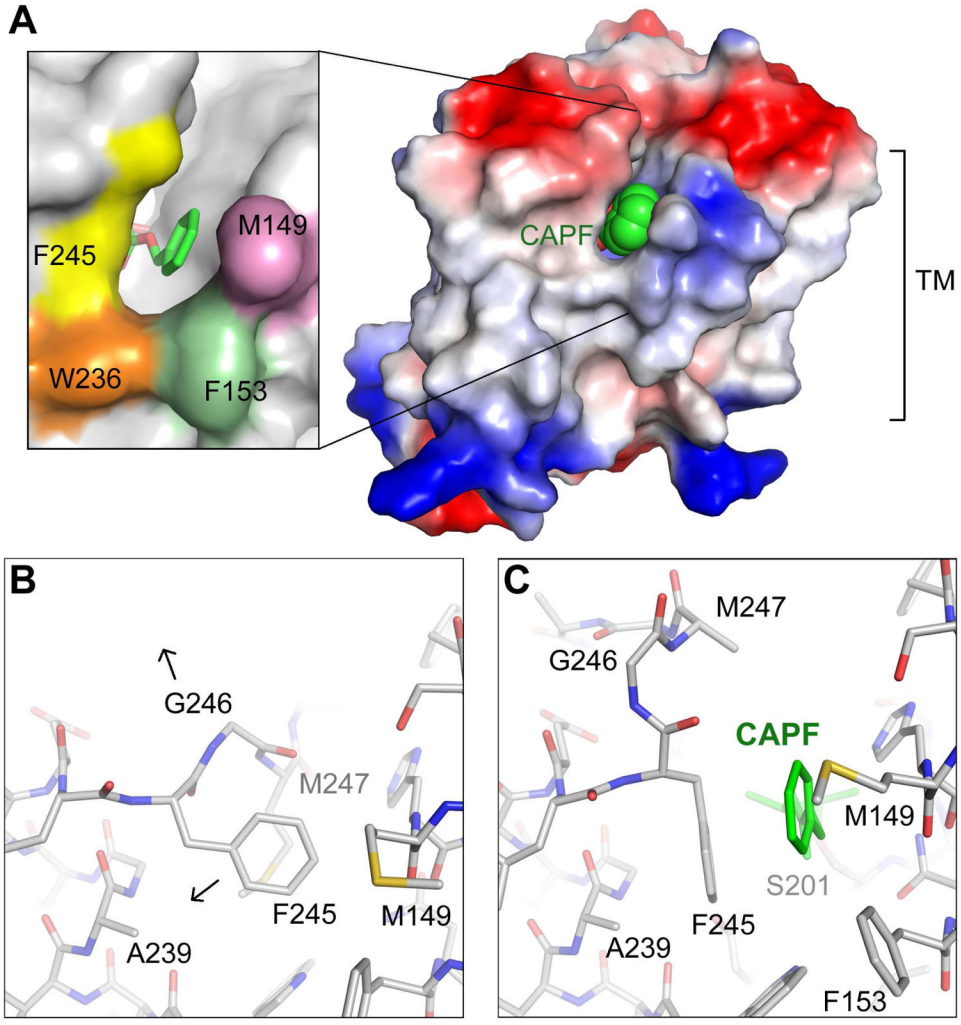

Fig. 6. The conformational change of the L5 cap

A, The molecular surface of the CAPF complex colored according to electrostatic potential (blue, positive; red, negative). The inhibitor, shown as the space filling model, is positioned a few angstroms below the predicted membrane surface. The insert highlights the residues (Phe-245, Trp-236, Phe-153, Met-149) that define the lateral opening. **B**, The conformation of the L5 cap in the native structure. **C**, The conformation of the L5 cap in the CAPF structure.

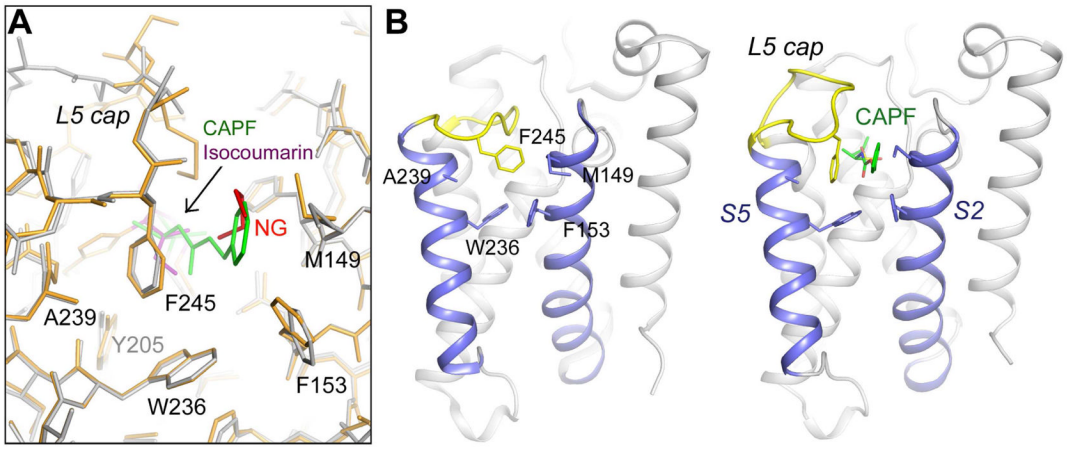

Fig. 7. Substrate entry is regulated by the L5 cap

A, A comparison between the CAPF (grey) and isocoumarin complexes (orange). CAPF and isocoumarin are colored in green and magenta, respectively. The detergent molecule (NG) observed in the isocoumarin complex is colored in red. **B**, In the absence of substrate, the L5 cap rests upon the active site, blocking its access. In the inhibitor complex, the side chain of Phe-245 has rotated to reveal the lateral opening. The L5 cap is highlighted in yellow. TM helices S2 and S5 are colored in blue.

Table 1

Crystallographic statistics. GlpG was crystallized in space group H32.

Data collection	**GlpG:CAPF**
Cell dimensions (Å)	a=b=111.7, c=121.8
Wavelength (Å)	1.54
[a] Resolution (Å)	50.0–2.6 (2.69–2.60)
Observed reflections	95,955
Unique reflections	9,147
Redundancy	10.5
[a] Completeness (%)	99.8 (100.0)
[a] $\langle I/\sigma \rangle$	14.7 (4.8)
[a,b] R_{merge}	0.069 (0.541)
Refinement	
Resolution (Å)	50.0–2.6
[c] R_{work}/R_{free}	0.211/0.229
Number of atoms	
Protein	1,398
CAPF	19
Water	30
B-factors	
Protein	56.5
CAPF	56.1
Water	57.2
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angle (°)	1.115

[a] Highest resolution shell is shown in parentheses.

[b] $R_{merge} = \Sigma|I_i - \langle I \rangle|/\Sigma I_i$

[c] $R_{work} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$. R_{free} is the cross-validation R factor for the test set of reflections (5% of the total) omitted in model refinement.