BIOCHEMISTRY
including biophysical chemistry & molecular biology

Accelerated Publication

pubs.acs.org/biochemistry

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*[,†]

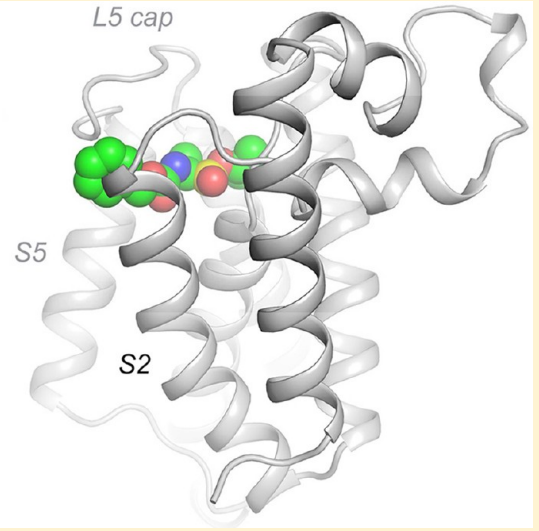
[†]Department of Pharmacology, Yale School of Medicine, New Haven, Connecticut 06520, United States

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

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

Supporting Information

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 the aqueous solution. This study, together with earlier crystallographic investigation of smaller inhibitors, suggests a simple model for explaining 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 homologue Pcp1/Rbd1) is involved in membrane dynamics and apoptosis via the cleavage of 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 the invasion of the host cell by *Plasmodium falciparum* and *Toxoplasma gondii*, the causative agents of human malaria and toxoplasmosis, respectively, because of 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 on the basis of the 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 (Figure 1). In the absence of substrate or inhibitor, the protease is closed (I): the

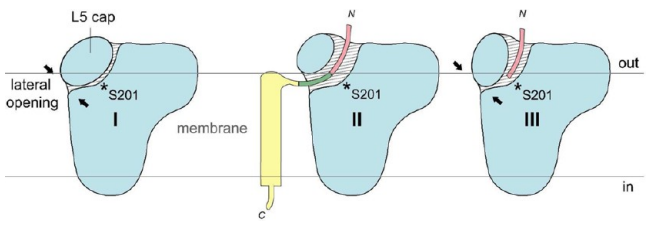


Figure 1. Schematic diagram for the three conformational states of rhomboid protease. The two horizontal lines mark the boundaries of the hydrophobic region of the membrane. The hydrophilic active site is represented by the hatched area. The catalytic serine is denoted by the asterisk. The substrate is colored red, green, and yellow. The protease cleaves between the red and green segments.

native crystal structure of *Escherichia 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

Received: January 18, 2012
Revised: April 13, 2012
Published: April 19, 2012

   dx.doi.org/10.1021/bi300368b | *Biochemistry* 2012, 51, 3723−3731

previously called the L5 cap[25] (see the schematic diagram in Figure 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 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 Figure 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 allow 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 co-workers.[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 octahistidine tag was used as the substrate for GlpG. The construct is similar to that described by Strisovsky et al.[27] and was generated on the basis of a MBP−Gurken−GlpG$_{91-276}$ construct (pGW475), which was initially designed for crystallographic study of the Gurken−GlpG complex. The sequence of the MBP−Gurken−GlpG$_{91-276}$ complex 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 polymerase chain reaction (PCR) from *E. coli* genomic DNA. The PCR product was digested with BamHI and 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 medium at 37 °C in the presence of 40 μM kanamycin; IPTG was added (final concentration of 0.4 mM) at an OD_{600} of 0.6 to induce protein expression (37 °C for 3 h). 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 (Clontech) and eluted with 300 mM imidazole.

The cleavage reactions were performed in 15 μL of assay buffer containing 50 mM Tris (pH 8.0), 0.1 M NaCl, and 0.5% NG, and each used 2 μg of GlpG and 4 μg of substrate fusion protein. The mixture was incubated at 37 °C for 3 h before sodium dodecyl sulfate−polyacrylamide gel electrophoresis loading buffer was added to stop the reaction. Various amounts of inhibitor were preincubated with the protease at 37 °C for 1 h 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 being soaked 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, yielding a final concentration of 5 mM. After a 4 h 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 a RAXIS-4 detector) at 100 K. The diffraction images were indexed, integrated, and scaled using *HKL2000*.[28] The structure was determined by molecular replacement using the native structure of GlpG [Protein Data Bank (PDB) entry 2ic8[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 on the basis of the difference Fourier map. A model of the inhibitor was initially generated by the PRODRG2 server,[31] and the restraint dictionary file was created with *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 conducted using *PHENIX*.[32] The final model has an R_{free} value of 22.9% and has good geometry (Table 1). The presence of the

Table 1. Crystallographic Statistics[a]

	GlpG−CAPF
Data Collection	
cell dimensions (Å)	$a = b = 111.7$, $c = 121.8$
wavelength (Å)	1.54
resolution[b] (Å)	50.0−2.6 (2.69−2.60)
no. of observed reflections	95955
no. of unique reflections	9147
redundancy	10.5
completeness[b] (%)	99.8 (100.0)
$\langle I/\sigma \rangle$[b]	14.7 (4.8)
R_{merge}[b,c]	0.069 (0.541)
Refinement	
resolution (Å)	50.0−2.6
R_{work}/R_{free}[d]	0.211/0.229
no. of atoms	
protein	1398
CAPF	19
water	30
B factor	
protein	56.5
CAPF	56.1
water	57.2
rmsd	
bond lengths (Å)	0.008
bond angles (deg)	1.115

[a]GlpG was crystallized in space group *H*32. [b]Data for the highest-resolution shell are given in parentheses. [c]$R_{merge} = \sum|I_i - \langle I \rangle|/\sum I_i$. [d]$R_{work} = \sum|F_o - F_c|/\sum F_o$. R_{free} is the cross-validation *R* factor for the test set of reflections (5% of the total) omitted from model refinement.

dx.doi.org/10.1021/bi300368b | *Biochemistry* 2012, 51, 3723−3731

inhibitor, as well as deviations from the native structure, was confirmed by omit difference maps (Figure 3A and Figures S1 and 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 determined at a higher resolution (2.09 Å).

RESULTS

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., refs 35 and 36). Unfortunately, the initial experiments using a number of diphenyl esters (kind gift from J. C. Powers) showed that none of them bound to GlpG, because of 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) (Figure 2A)[26] and found that it inhibited GlpG-catalyzed proteolysis of the TM fusion protein substrate with an apparent IC_{50} of ~50 μM[27] (Figure 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 (Figure 2A).

In the crystal structure, the carboxybenzyl (Cbz) end of the inhibitor occupies the S′ side of the substrate binding cleft (Figure 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 and Phe-153) near the N-terminus of TM helix S2 (Figure 3C). The carboxylate oxygen, which may mimic the carbonyl oxygen of the substrate's P1 residue (Figure 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 atom 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 (Figure S2 of the Supporting Information). The isopropyl group of the inhibitor points toward a region that an 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 rotations similar to those observed in the complexes with DFP and isocoumarin[33,37] (Figure 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.

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 201–216): the catalytic serine (Ser-201) is located at the end of S4, and its position serves as a good reference point. When the structure is 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 (Figure 4A,B). Most helices moved as rigid bodies, with a Cα atom rmsd of <0.4 Å. S6 is the only exception, and its high rmsd (1.5 Å) is due to a major rearrangement toward the helix's N-terminus (see below). The L5 cap has the largest change in structure [rmsd of 2.1 Å (see below)], and its center of mass is shifted by 3.2 Å. This is followed by S5, L1, and S6, which moved by 1.6, 1.2, 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 (Figure 5A). S6 (10.3°) and S5 (7.5°) show the largest tilting motions. In comparison, the L1 loop rotated ~4.9°.

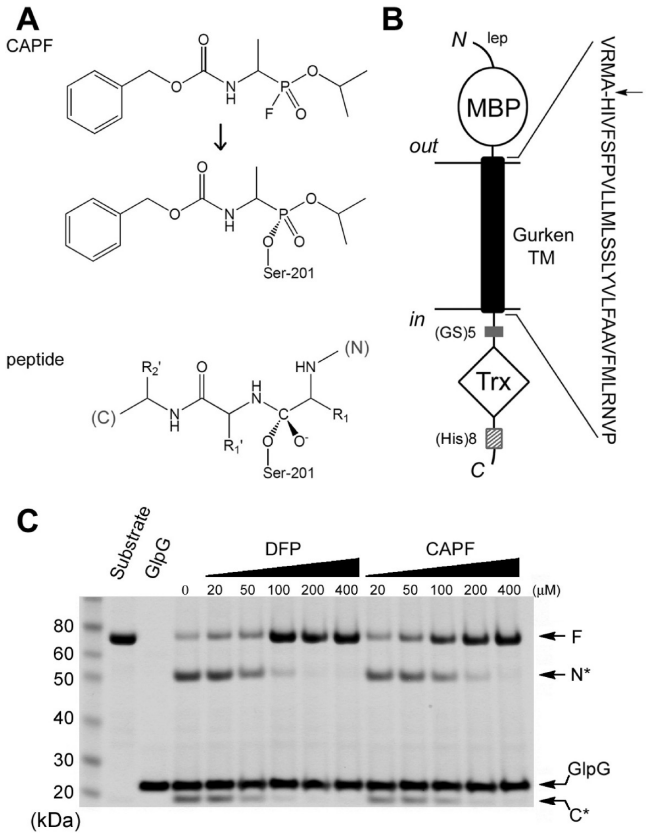


Figure 2. CAPF inhibits the proteolytic activity of GlpG. (A) Chemical structures of CAPF, its covalent adduct with the protease, and the first tetrahedral intermediate of the proteolytic reaction. (B) Diagram of the MBP–Gurken–Trx fusion protein substrate. The fusion protein consists of a leader peptide (lep), a 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.

dx.doi.org/10.1021/bi300368b | *Biochemistry* 2012, 51, 3723–3731

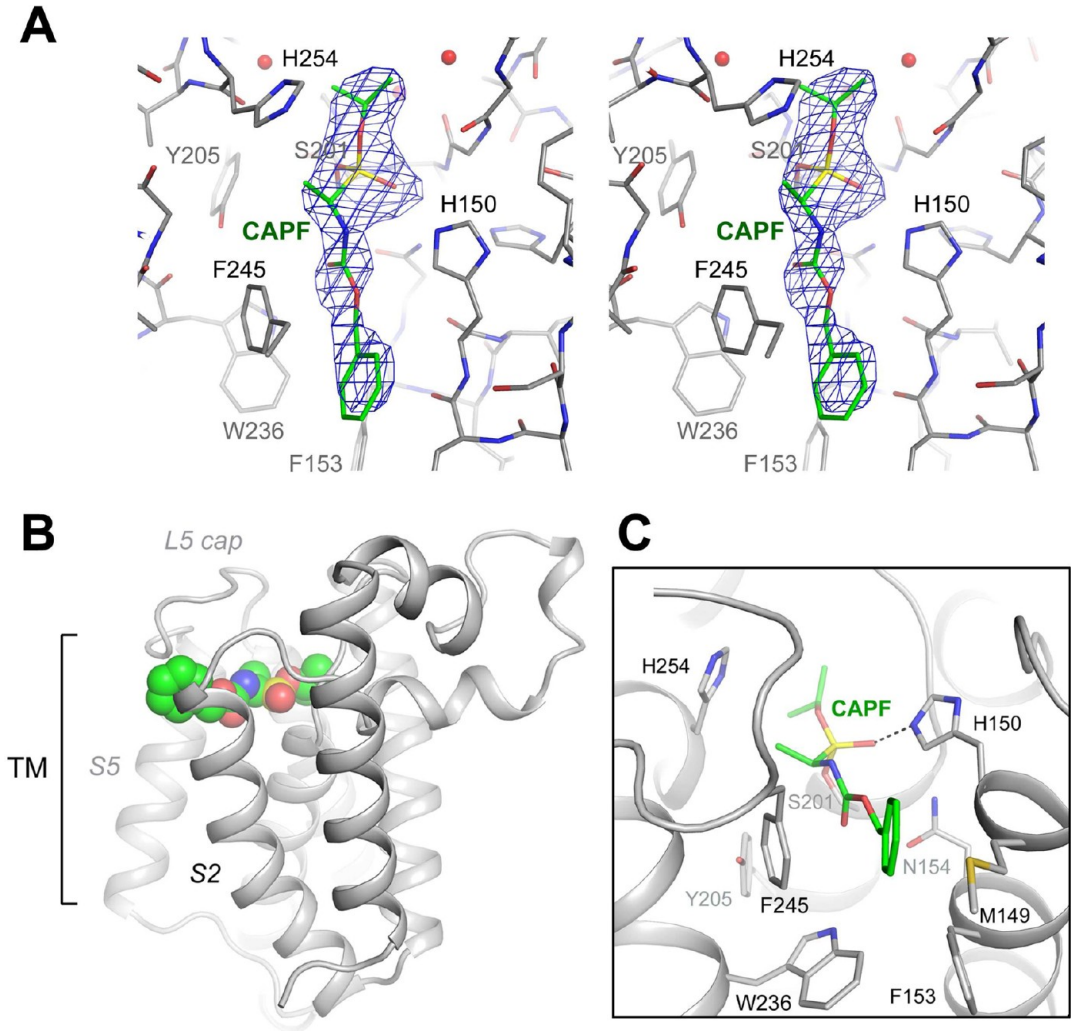


Figure 3. Structure of the CAPF complex. (A) Stereoview of the omit $F_o - F_c$ map (generated without the inhibitor) contoured at 3σ. A different view of the map is shown in Figure S1 of the Supporting Information. The inhibitor is colored green. Water molecules are shown as red spheres. (B) Overall structure. CAPF is shown as a space filling model. (C) Close-up 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.

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 (Figure 4C). The only major difference is S6, which, in the isocoumarin complex, is drawn toward 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 (Figure 4D). It is interesting that, when the native structures of *E. coli* and *Haemophilus influenzae* GlpGs are compared, similar rotations of L1, S3, S6, and S5 around 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 apoprotein, in which 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 and S6) that our present study suggests may occur concurrently with the S5 movement and lacks the rotation of key side chains (Tyr-205 and 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 (Figure 5A,B). A water molecule is inserted between the two short helices, relaying the 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 toward 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

dx.doi.org/10.1021/bi300368b | *Biochemistry* 2012, 51, 3723–3731

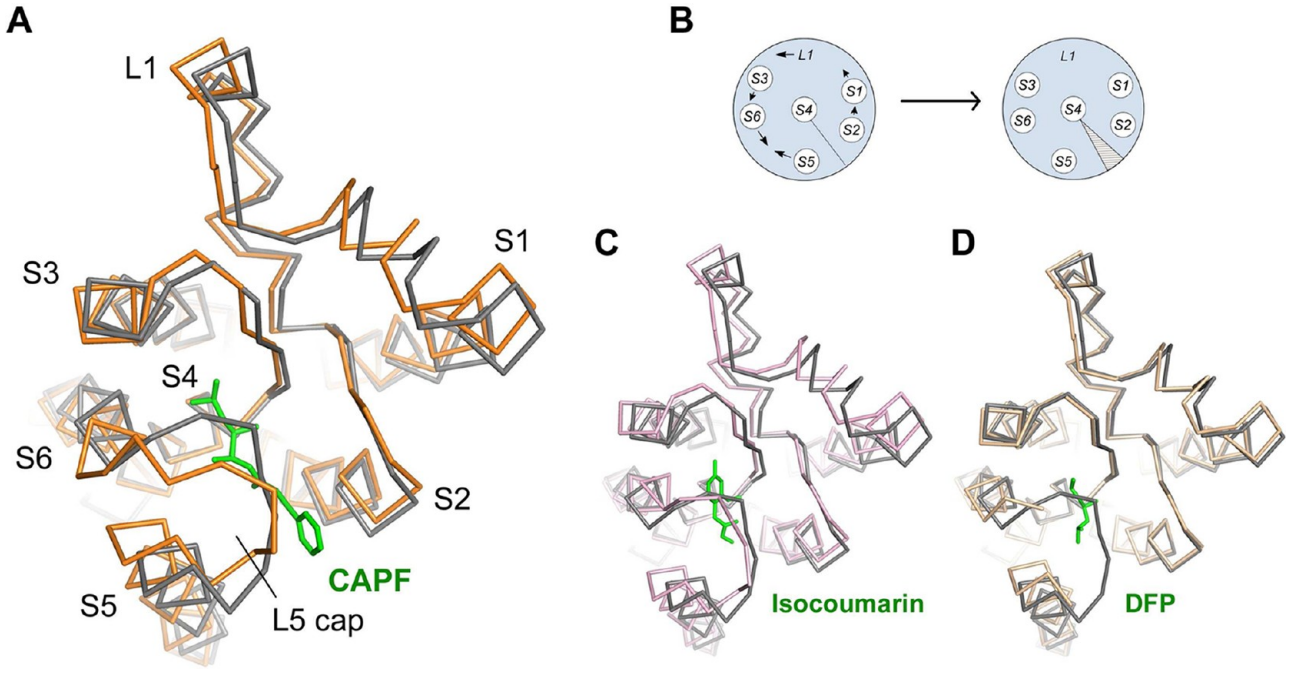


Figure 4. Movement of the TM helices. (A) Superposition of the CAPF structure (orange) onto the native structure (gray). The Cα atom trace of the protein is shown. The inhibitor (green) is represented by the stick model. The differences between the complex and native structures can be validated by the omit difference maps (Figure S4 of the Supporting Information). (B) 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) Isocoumarin structure (pink, PDB entry 2xow) and DFP structure (sand, PDB entry 3txt), respectively, similarly compared to the apo structure (gray).

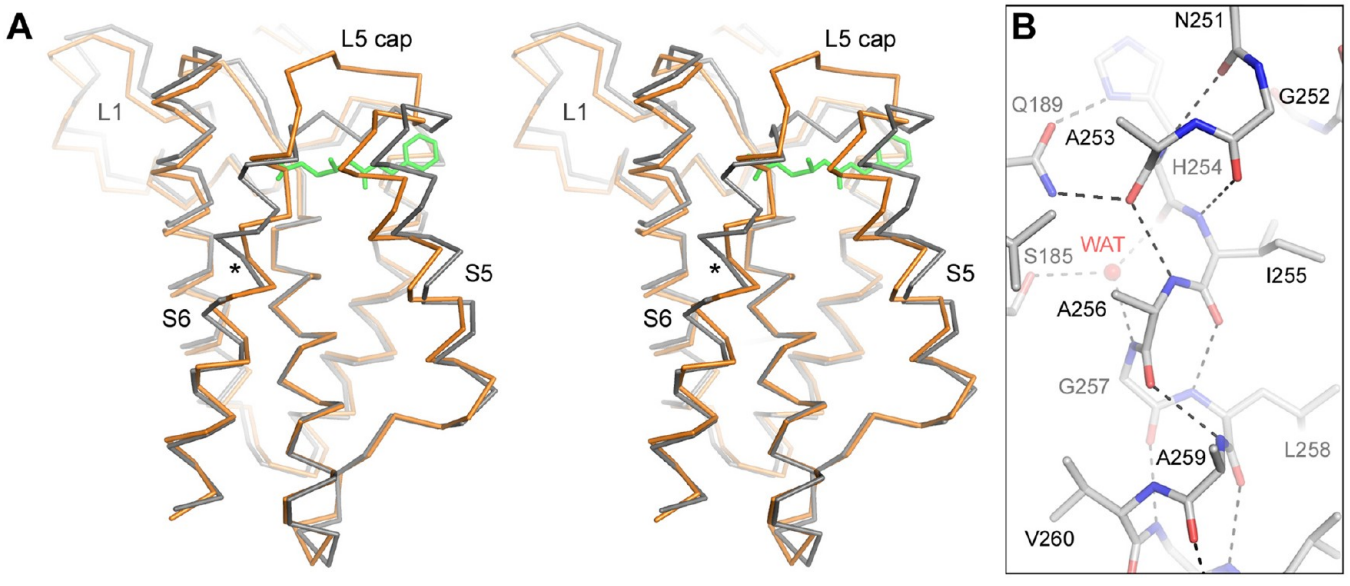


Figure 5. Conformational change in TM helix S6. (A) Stereoview of the CAPF structure (orange) superposed onto the native structure (gray). Gly-257 is denoted with an asterisk. (B) New S6 conformation stabilized by a network of hydrogen bonds (dashed lines). A water molecule, shown as a red sphere, is inserted between the two 3_{10} helices (Asn-251−Ala-253 and Ile-255−Ala-256).

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 (Figure 6A). The lateral opening is obstructed in the native structure by the side chain of Phe-245, which adopts a different conformation (Figure 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) increased only marginally, from 8.5 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 toward the lipid (Figure 6C)]. The distance between Phe-245 and Met-149 doubled from 3.8 to 7.6 Å. The backbone φ and 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

dx.doi.org/10.1021/bi300368b | *Biochemistry* 2012, 51, 3723−3731

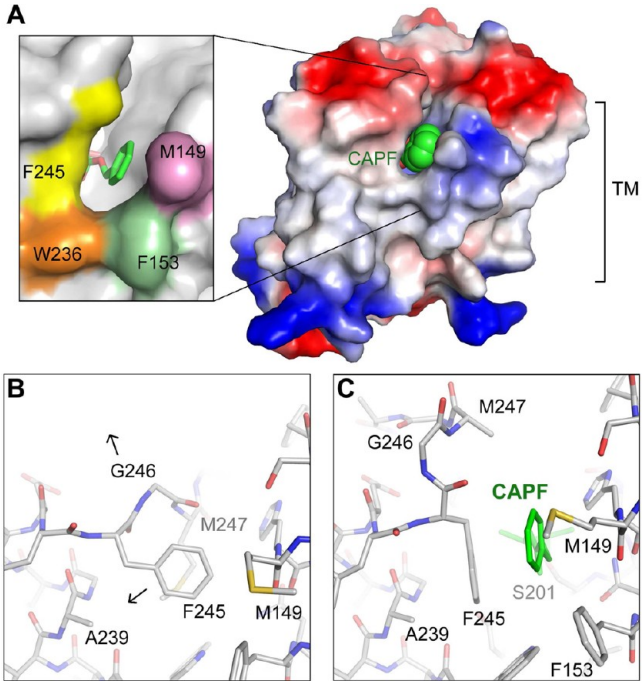


Figure 6. Conformational change of the L5 cap. (A) Molecular surface of the CAPF complex colored according to electrostatic potential (blue, positive; red, negative). The inhibitor, shown as the space filling model, is positioned a few angstroms below the predicted membrane surface. The insert highlights the residues (Phe-245, Trp-236, Phe-153, and Met-149) that define the lateral opening. (B) Conformation of the L5 cap in the native structure. (C) Conformation of the L5 cap in the CAPF structure.

substrate binding cleft, which exposes the catalytic dyad to the aqueous solution.

In the GlpG–isocoumarin complex,[37] although the inhibitor does not extend far into the S′ side of the substrate binding cleft (Figure S3 of the Supporting Information), the protein structure at the lateral opening is strikingly similar to that of the CAPF complex (Figure 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; and 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 Figure 7A). This observation raises the possibility that the hydrophobic patch (Met-149 and 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, and 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 had been determined,[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] Because 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 determined. 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. Because 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 the 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

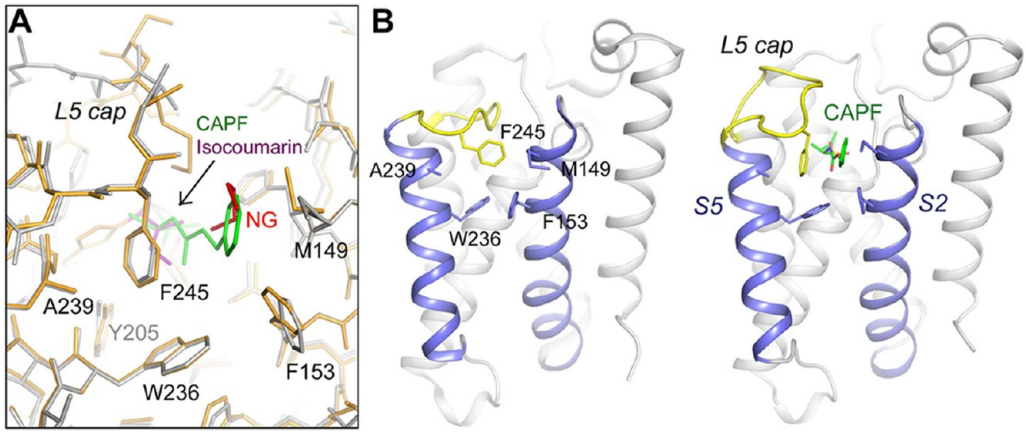


Figure 7. Substrate entry is regulated by the L5 cap. (A) Comparison between the CAPF (gray) and isocoumarin complexes (orange). CAPF and isocoumarin are colored green and magenta, respectively. The detergent molecule (NG) observed in the isocoumarin complex is colored red. (B) In the absence of substrate, 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 colored yellow. TM helices S2 and S5 are colored blue.

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 by only 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 (Figure 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 (Figure 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 the 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 Figure 1).

The observation that the TM helices in GlpG moved only slightly in response to inhibitor binding suggests a simple model for explaining 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 (Figure 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 the aqueous solution; and over the lateral opening the membrane-spanning substrate can bend 90° into the active site in an extended conformation to be cleaved. On the basis of similarities between the CAPF and isocoumarin structures (Figure 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, and 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 refs 46 and 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 allows 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 the case in Rv3671c), the side chains of Tyr-205 and Trp-236 do have conformations different from 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, because 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 homogeneous and aqueous environment.

■ ASSOCIATED CONTENT

Supporting Information

Additional analyses of the crystal structure of the GlpG−CAPF complex (Figures S1−S4). This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

The atomic coordinates and structure factors (entry 3ubb) have been deposited in the Protein Data Bank.

■ AUTHOR INFORMATION

Corresponding Author

*Department of Pharmacology, Yale School of Medicine, 333 Cedar St., New Haven, CT 06520. Telephone: (203) 785-7530. Fax: (203) 785-7670. E-mail: ya.ha@yale.edu.

Present Address

‖Division of Experimental Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02115.

Funding

This work was supported by National Institutes of Health Grant GM082839 (Y.H.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

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

■ ABBREVIATIONS

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; rmsd, root-mean-square deviation; TM, transmembrane; Trx, thioredoxin.

■ REFERENCES

(1) Freeman, M. (2008) Rhomboid proteases and their biological functions. *Annu. Rev. Genet. 42*, 191−210.

(2) Hill, R. B., and Pellegrini, L. (2010) The PARL family of mitochondrial rhomboid proteases. *Semin. Cell Dev. Biol. 21*, 582−592.

(3) Urban, S. (2006) Rhomboid proteins: Conserved membrane proteases with divergent biological functions. *Genes Dev. 20*, 3054−3068.

(4) Mayer, U., and Nusslein-Volhard, C. (1988) A group of genes required for pattern formation in the ventral ectoderm of the *Drosophila* embryo. *Genes Dev. 2*, 1496−1511.

(5) Wasserman, J. D., Urban, S., and Freeman, M. (2000) A family of rhomboid-like genes: *Drosophila* rhomboid-1 and roughoid/rhomboid-3 cooperate to activate EGF receptor signaling. *Genes Dev. 14*, 1651–1663.
(6) Lee, J. R., Urban, S., Garvey, C. F., and Freeman, M. (2001) Regulated intracellular ligand transport and proteolysis control EGF signal activation in *Drosophila*. *Cell 107*, 161–171.
(7) Urban, S., Lee, J. R., and Freeman, M. (2001) *Drosophila* rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell 107*, 173–182.
(8) Esser, K., Tursun, B., Ingenhoven, M., Michaelis, G., and Pratje, E. (2002) 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. 323*, 835–843.
(9) McQuibban, G. A., Saurya, S., and Freeman, M. (2003) Mitochondrial membrane remodelling regulated by a conserved rhomboid protease. *Nature 423*, 537–541.
(10) Herlan, M., Vogel, F., Bornhovd, C., Neupert, W., and Reichert, A. S. (2003) Processing of Mgm1 by the rhomboid-type protease Pcp1 is required for maintenance of mitochondrial morphology and of mitochondrial DNA. *J. Biol. Chem. 278*, 27781–27788.
(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., and De Strooper, B. (2006) Mitochondrial rhomboid PARL regulates cytochrome c release during apoptosis via OPA1-dependent cristae remodeling. *Cell 126*, 163–175.
(12) Chao, J. R., Parganas, E., Boyd, K., Hong, C. Y., Opferman, J. T., and Ihle, J. N. (2008) Hax1-mediated processing of HtrA2 by Parl allows survival of lymphocytes and neurons. *Nature 452*, 98–102.
(13) Rather, P. N., Ding, X., Baca-DeLancey, R. R., and Siddiqui, S. (1999) *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. 181*, 7185–7191.
(14) Gallio, M., Sturgill, G., Rather, P., and Kylsten, P. (2002) A conserved mechanism for extracellular signaling in eukaryotes and prokaryotes. *Proc. Natl. Acad. Sci. U.S.A. 99*, 12208–12213.
(15) Stevenson, L. G., Strisovsky, K., Clemmer, K. M., Bhatt, S., Freeman, M., and Rather, P. N. (2007) Rhomboid protease AarA mediates quorum-sensing in *Providencia stuartii* by activating TatA of the twin-arginine translocase. *Proc. Natl. Acad. Sci. U.S.A. 104*, 1003–1008.
(16) Opitz, C., Di Cristina, M., Reiss, M., Ruppert, T., Crisanti, A., and Soldati, D. (2002) Intramembrane cleavage of microneme proteins at the surface of the apicomplexan parasite *Toxoplasma gondii*. *EMBO J. 21*, 1577–1585.
(17) Zhou, X. W., Blackman, M. J., Howell, S. A., and Carruthers, V. B. (2004) 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.
(18) Brossier, F., Jewett, T. J., Sibley, L. D., and Urban, S. (2005) A spatially localized rhomboid protease cleaves cell surface adhesins essential for invasion by *Toxoplasma*. *Proc. Natl. Acad. Sci. U.S.A. 102*, 4146–4151.
(19) O'Donnell, R. A., Hackett, F., Howell, S. A., Treeck, M., Struck, N., Krnajski, Z., Withers-Martinez, C., Gilberger, T. W., and Blackman, M. J. (2006) Intramembrane proteolysis mediates shedding of a key adhesin during erythrocyte invasion by the malaria parasite. *J. Cell Biol. 174*, 1023–1033.
(20) Baker, R. P., Wijetilaka, R., and Urban, S. (2006) Two *Plasmodium* rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria. *PLoS Pathog. 2*, e113.
(21) Buguliskis, J. S., Brossier, F., Shuman, J., and Sibley, L. D. (2010) Rhomboid 4 (ROM4) affects the processing of surface adhesins and facilitates host cell invasion by *Toxoplasma gondii*. *PLoS Pathog. 6*, e1000858.
(22) Santos, J. M., Ferguson, D. J., Blackman, M. J., and Soldati-Favre, D. (2011) Intramembrane cleavage of AMA1 triggers *Toxoplasma* to switch from an invasive to a replicative mode. *Science 331*, 473–477.
(23) Lemberg, M. K., Menendez, J., Misik, A., Garcia, M., Koth, C. M., and Freeman, M. (2005) Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases. *EMBO J. 24*, 464–472.
(24) Wang, Y., Zhang, Y., and Ha, Y. (2006) Crystal structure of a rhomboid family intramembrane protease. *Nature 444*, 179–180.
(25) Wang, Y., and Ha, Y. (2007) Open-cap conformation of intramembrane protease GlpG. *Proc. Natl. Acad. Sci. U.S.A. 104*, 2098–2102.
(26) Bartlett, P. A., and Lamden, L. A. (1986) Inhibition of chymotrypsin by phosphonate and phosphonamidate peptide analogs. *Bioorg. Chem. 14*, 357–377.
(27) Strisovsky, K., Sharpe, H. J., and Freeman, M. (2009) Sequence-specific intramembrane proteolysis: Identification of a recognition motif in rhomboid substrates. *Mol. Cell 36*, 1048–1059.
(28) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
(29) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D66*, 486–501.
(30) Winn, M. D., Murshudov, G. N., and Papiz, M. Z. (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol. 374*, 300–321.
(31) Schuttelkopf, A. W., and van Aalten, D. M. (2004) PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D60*, 1355–1363.
(32) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D66*, 213–221.
(33) Xue, Y., and Ha, Y. (2012) The catalytic mechanism of rhomboid protease GlpG probed by 3,4-dichloroisocoumarin and diisopropyl fluorophosphonate. *J. Biol. Chem. 287*, 3099–3107.
(34) Oleksyszyn, J., and Powers, J. C. (1994) Amino acid and peptide phosphonate derivatives as specific inhibitors of serine peptidases. *Methods Enzymol. 244*, 423–441.
(35) Sampson, N. S., and Bartlett, P. A. (1991) Peptidic phosphonylating agents as irreversible inhibitors of serine proteases and models of the tetrahedral intermediates. *Biochemistry 30*, 2255–2263.
(36) Bone, R., Sampson, N. S., Bartlett, P. A., and Agard, D. A. (1991) Crystal structures of α-lytic protease complexes with irreversibly bound phosphonate esters. *Biochemistry 30*, 2263–2272.
(37) Vinothkumar, K. R., Strisovsky, K., Andreeva, A., Christova, Y., Verhelst, S., and Freeman, M. (2010) The structural basis for catalysis and substrate specificity of a rhomboid protease. *EMBO J. 29*, 3797–3809.
(38) Lemieux, M. J., Fischer, S. J., Cherney, M. M., Bateman, K. S., and James, M. N. (2007) The crystal structure of the rhomboid peptidase from *Haemophilus influenzae* provides insight into intramembrane proteolysis. *Proc. Natl. Acad. Sci. U.S.A. 104*, 750–754.
(39) Wu, Z., Yan, N., Feng, L., Oberstein, A., Yan, H., Baker, R. P., Gu, L., Jeffrey, P. D., Urban, S., and 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.
(40) Baker, R. P., Young, K., Feng, L., Shi, Y., and Urban, S. (2007) Enzymatic analysis of a rhomboid intramembrane protease implicates transmembrane helix 5 as the lateral substrate gate. *Proc. Natl. Acad. Sci. U.S.A. 104*, 8257–8262.
(41) Ha, Y. (2009) Structure and mechanism of intramembrane protease. *Semin. Cell Dev. Biol. 20*, 240–250.
(42) Ben-Shem, A., Fass, D., and Bibi, E. (2007) Structural basis for intramembrane proteolysis by rhomboid serine proteases. *Proc. Natl. Acad. Sci. U.S.A. 104*, 462–466.

(43) Brooks, C. L., Lazareno-Saez, C., Lamoureux, J. S., Mak, M. W., and Lemieux, M. J. (2011) Insights into substrate gating in *H. influenzae* rhomboid. *J. Mol. Biol. 407*, 687−697.
(44) Maegawa, S., Koide, K., Ito, K., and Akiyama, Y. (2007) 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. 64*, 435−447.
(45) Wang, Y., Maegawa, S., Akiyama, Y., and Ha, Y. (2007) The role of L1 loop in the mechanism of rhomboid intramembrane protease GlpG. *J. Mol. Biol. 374*, 1104−1113.
(46) Stroud, R. M., Kossiakoff, A. A., and Chambers, J. L. (1977) Mechanisms of zymogen activation. *Annu. Rev. Biophys. Bioeng. 6*, 177−193.
(47) Bode, W. (2006) The structure of thrombin: A janus-headed proteinase. *Semin. Thromb. Hemostasis 32*, 16−31.
(48) Kim, D. Y., Kim, D. R., Ha, S. C., Lokanath, N. K., Lee, C. J., Hwang, H. Y., and Kim, K. K. (2003) Crystal structure of the protease domain of a heat-shock protein HtrA from *Thermotoga maritima*. *J. Biol. Chem. 278*, 6543−6551.
(49) Kim, D. Y., Kwon, E., Shin, Y. K., Kweon, D. H., and Kim, K. K. (2008) The mechanism of temperature-induced bacterial HtrA activation. *J. Mol. Biol. 377*, 410−420.
(50) Krojer, T., Garrido-Franco, M., Huber, R., Ehrmann, M., and Clausen, T. (2002) Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. *Nature 416*, 455−459.
(51) Krojer, T., Pangerl, K., Kurt, J., Sawa, J., Stingl, C., Mechtler, K., Huber, R., Ehrmann, M., and Clausen, T. (2008) Interplay of PDZ and protease domain of DegP ensures efficient elimination of misfolded proteins. *Proc. Natl. Acad. Sci. U.S.A. 105*, 7702−7707.
(52) Jiang, J., Zhang, X., Chen, Y., Wu, Y., Zhou, Z. H., Chang, Z., and Sui, S. F. (2008) Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. *Proc. Natl. Acad. Sci. U.S.A. 105*, 11939−11944.
(53) Wilken, C., Kitzing, K., Kurzbauer, R., Ehrmann, M., and Clausen, T. (2004) Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease. *Cell 117*, 483−494.
(54) Biswas, T., Small, J., Vandal, O., Odaira, T., Deng, H., Ehrt, S., and Tsodikov, O. V. (2010) Structural insight into serine protease Rv3671c that protects *M. tuberculosis* from oxidative and acidic stress. *Structure 18*, 1353−1363.