STRUCTURAL BIOLOGY

Cryo-EM structure of the bacterial intramembrane metalloprotease RseP in the substrate-bound state

Kikuko Asahi[1]†, Mika Hirose[2]†, Rie Aruga[1]†, Yosuke Shimizu[3]†, Michiko Tajiri[1]†, Tsubasa Tanaka[1], Yuriko Adachi[1], Yukari Tanaka[1], Mika K. Kaneko[4], Yukinari Kato[4], Satoko Akashi[1], Yoshinori Akiyama[3], Yohei Hizukuri[3]*, Takayuki Kato[2]*, Terukazu Nogi[1]*

Site-2 proteases (S2Ps), conserved intramembrane metalloproteases that maintain cellular homeostasis, are associated with chronic infection and persistence leading to multidrug resistance in bacterial pathogens. A structural model of how S2Ps discriminate and accommodate substrates could help us develop selective antimicrobial agents. We previously proposed that the *Escherichia coli* S2P RseP unwinds helical substrate segments before cleavage, but the mechanism for accommodating a full-length membrane-spanning substrate remained unclear. Our present cryo-EM analysis of *Aquifex aeolicus* RseP (*Aa*RseP) revealed that a substrate-like membrane protein fragment from the expression host occupied the active site while spanning a transmembrane cavity that is inaccessible via lateral diffusion. Furthermore, in vivo photocrosslinking supported that this substrate accommodation mode is recapitulated on the cell membrane. Our results suggest that the substrate accommodation by threading through a conserved membrane-associated region stabilizes the substrate-complex and contributes to substrate discrimination on the membrane.

INTRODUCTION

Intramembrane proteases (IMPs) that catalyze peptide bond hydrolysis within the membrane have been identified in all three domains of life. Of them, IMPs that cleave transmembrane (TM) segments of their substrates have been classified into three distinct families based on the catalytic type: site-2 protease (S2P), presenilin/signal peptide peptidase, and rhomboid (*1*, *2*). IMPs are implicated in signal transduction by releasing signaling molecules and in membrane proteostasis by removing unnecessary polypeptides from the lipid bilayer (*3*–*5*). As IMPs regulate diverse cellular processes, their dysregulation is also known to cause several types of human diseases such as cancer, Alzheimer's disease (*6*, *7*), and Parkinson's disease (*8*). Among the three IMP families, S2P is classified as the metalloprotease-type IMP (*9*). In eukaryotes, S2Ps have been shown to activate membrane-spanning precursors of transcription factors to regulate sterol homeostasis and endoplasmic reticulum stress responses (*10*–*12*). S2P homologs identified in bacteria are involved in similar cellular processes such as mediating signal transduction in envelope synthesis and stress responses (*13*, *14*). Furthermore, S2P function is related to chronic infection and persistence that leads to multidrug resistance (*15*). In *Mycobacterium tuberculosis*, deletion of an S2P homolog gene largely attenuates both acute and chronic infection (*16*). It has been reported that the *Vibrio cholerae* S2P homolog is involved in pathogenicity and entering dormancy (*17*). Finding inhibitors that target S2P homologs in pathogens is expected to lead to the development of antimicrobial agents, but inhibitors that directly target active site residues will likely cause severe adverse effects by disturbing the activities of human S2P and many soluble metalloproteases. Thus, exploring an allosteric site that is unique to each specific S2P homolog could inform rational design of selective inhibitors.

Regarding physiological function and the regulation of cleavage, *Escherichia coli* RseP (*Ec*RseP) is one of the best characterized members in the S2P family and classified as the group I S2P, which includes human S2P and most of the S2P homologs related to virulence in pathogens. *Ec*RseP was first identified in the signaling process for extracytoplasmic stress response, where it cleaves a membrane-spanning anti-sigma factor RseA to release and activate σ^E (*18*–*20*). *Ec*RseP also cleaves the sigma factor regulator FecR to induce gene expression in the ferric citrate uptake system (*21*). It has also been shown that *Ec*RseP contributes to quality control in the cell membrane by degrading remnant signal peptides derived from secreted proteins (*22*). Furthermore, *Ec*RseP cleaves an endogenous toxin HokB, which induces persister formation in *E. coli*. *Ec*RseP suppresses HokB cytotoxicity via intramembrane proteolysis (*23*), raising a possibility that *Ec*RseP is implicated in the awakening of the HokB-induced persister. In addition, systematic cleavage analysis has shown that 13 other small membrane proteins are potential substrates for *Ec*RseP (*23*). All *Ec*RseP substrates identified so far are single-pass TM proteins with an N-in/C-out topology but show no obvious homologies on the primary structure. Nevertheless, *Ec*RseP is suggested to perform the intramembrane proteolysis not in promiscuous manners, but rather in specific manners as the cleavage analysis also showed that some of small membrane proteins with the N-in/C-out topology were not efficiently cleaved by *Ec*RseP (*23*). Thus, structural data on *Ec*RseP and its substrate-bound complex are essential to elucidate the substrate discrimination mechanism.

We therefore performed crystallographic analysis and determined a three-dimensional (3D) structure of *Ec*RseP in complex with the competitive inhibitor batimastat (fig. S1) (*24*). *Ec*RseP is a four-TM protein with an N-out/C-out topology, in which three TM helices (TM1 to TM3) constitute a conserved catalytic core that coordinates the active site zinc ion. *Ec*RseP also has a conserved membrane-reentrant β sheet (MREβ sheet) close to the active site. Many soluble

[1]Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. [2]Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan. [3]Institute for Life and Medical Sciences, Kyoto University, 53 Shogoin kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan. [4]Department of Antibody Drug Development, Tohoku University Graduate School of Medicine; 2-1 Seiryo-machi, Sendai, Miyagi 980-8575, Japan.
*Corresponding author. Email: nogi@yokohama-cu.ac.jp (T.N.); tkato@protein.osaka-u.ac.jp (T.K.); yhizukur@infront.kyoto-u.ac.jp (Y.H.)
†These authors contributed equally to this work.

metalloproteases unwind their substrates before cleavage and bind it as a strand addition to a β sheet near the active site (*25*). A similar mechanism has been proposed for *Ec*RseP based on mutational and cross-linking analyses (*26*). The peptide-mimetic inhibitor batimastat bound to the MREβ sheet in an extended conformation comparable to the strand addition in the crystal structure (*24*). In addition, the binding of batimastat to MREβ appeared to be stabilized by main chain to side chain interactions with a highly conserved asparagine residue on TM3. As the mutations of this Asn residue substantially reduced the proteolytic activities, it is likely that the substrate cleavage segment is also unwound by MREβ and clamped by the conserved Asn residue for efficient cleavage within the membrane.

The batimastat-bound *Ec*RseP structure has deepened our understanding of the substrate cleavage mechanism, but it remained unclear how a full-length substrate is accommodated in *Ec*RseP. Batimastat was bound in a hydrophilic cavity surrounded by the TM helices (TM-cavity), which is likely the substrate-binding pocket (fig. S2). However, the cavity appears inaccessible to membrane-spanning substrates if substrates can only enter the cavity by lateral diffusion. *Ec*RseP possesses tandemly arranged PDZ-N and PDZ-C domains (PDZ tandem) and PDZ C-terminal (PCT) region on the periplasmic side. The periplasmic PDZ domains and the PCT region are structural elements unique to the bacterial group I S2Ps and suggested to be implicated in the substrate discrimination and accommodation (*27*, *28*). The crystallographic analysis of *Ec*RseP revealed that the PCT region consists of two helices (PCT-H1 and PCT-H2) and an intervening loop with a short 3_{10} helix (PCT-loop) (*24*). The PCT region appears to restrict the entry of the substrates because PCT-H1, PCT-H2, and PCT-loop form a hole (PCT-hole) at the membrane surface and surround the periplasmic end of the TM-cavity (figs. S1 and S2). In contrast to the small-molecule batimastat, membrane-spanning substrates will have more limited degrees of freedom when entering the TM-cavity. The periplasmic region of the substrates is generally abundant in hydrophilic residues and would have difficulty in entering the PCT-hole through the membrane. Alternatively, structure-based biochemical analyses raised the possibility that the PCT-region, together with TM4, undergoes gate-like structural changes to permit the substrate entry (*24*). Chemical modification and cross-linking analyses suggested that PCT-H2 movement is important for substrate cleavage. Consistent with this, substrate cleavage was substantially reduced when an electrostatic interaction that stabilizes the PCT-H2 conformation was disrupted by mutating an aspartate on TM4. These observations suggest that TM4 and PCT undergo cooperative conformational changes during substrate cleavage. A crystallographic analysis of an RseP ortholog from the marine γ-proteobacterium *Kangiella koreensis* (*Kk*RseP) is also consistent with a gate-like movement of the PCT region (fig. S1) (*24*). In the *Kk*RseP crystal structure, PCT-H2 and PCT-loop adopted a different conformation from that in the *Ec*RseP crystal. This conformational difference is probably due to the crystal packing interactions. In the *Kk*RseP crystal, TM4 on the crystal packing neighbor docked to the cleft created by the displacement of PCT-H2 and PCT-loop. The conformational flexibility required to adopt this conformation suggested that PCT-H2 and PCT-loop have a propensity to undergo structural changes through interactions with substrates. Nevertheless, the lack of structural data for substrate-bound complexes prevented us from determining the mechanism of structural changes and substrate binding.

To sample different conformational states, we further attempted to perform cryo–electron microscopy single-particle analysis (cryo-EM SPA) on *Ec*RseP. However, we could not obtain high-resolution structural data due to its small size and high structural flexibility. Alternatively, we performed cryo-EM SPA on an RseP ortholog from a hyperthermophilic bacterium *Aquifex aeolicus* (*Aa*RseP), which shares a common domain organization with *Ec*RseP and *Kk*RseP (Fig. 1A), and we obtained a well-resolved electron microscopy (EM) density map by using the Fab of an anti-*Aa*RseP monoclonal antibody as a fiducial marker. Although no physiological substrates have been identified for *Aa*RseP, a cell-based cleavage assay in *E. coli* confirmed that *Aa*RseP can proteolyze the model substrate used for *Ec*RseP (*29*). In this study, we successfully obtained a 3.61-Å resolution 3D reconstruction model for the active site mutant of *Aa*RseP in complex with the Fab. Unexpectedly, we observed that a ligand, presumably a polypeptide, was accommodated in the TM-cavity where the periplasmic C-terminal tail of the polypeptide was located inside the PCT-hole. We identified a candidate for the polypeptide ligand and confirmed that *Aa*RseP could cleave the proposed polypeptide on the membrane in our cell-based assay. Using in vivo crosslinking, we also demonstrated that the model substrate occupies the expected binding site in the TM-cavity when *Aa*RseP is expressed on the *E. coli* cell membrane. Our results suggest that the substrate accommodation by threading the C-terminal tail through the PCT-hole contributes to stabilization of the Michaelis, or enzyme-substrate (ES), complex and substrate discrimination on the cell membrane, which will help us develop strategies for selective regulation of intramembrane proteolysis by specific IMPs.

RESULTS

Cryo-EM SPA of *Aa*RseP(WT)

We prepared a complex of *Aa*RseP (47.8 kDa, not including the tag and the zinc ion) with an antibody Fab fragment recognizing the PDZ-N domain to produce a particle that is sufficiently large for structure determination by cryo-EM SPA. In addition, we obtained structural data at higher resolution by solubilizing the *Aa*RseP[wild type (WT)]:Fab complex with lauryl maltose neopentyl glycol (LMNG) (fig. S3). During data processing, we found that the *Aa*RseP(WT):Fab particles showed variations in the orientation of the PDZ tandem with respect to the TM region. We therefore performed discrete heterogeneity analysis by 3D classification and obtained a 3D reconstruction model at 3.95-Å resolution for one of the classes (Fig. 1B and figs. S4 to S6). The resolution was sufficient for us to assign atomic models to the density maps using crystal structures for the PDZ tandem and Fab and an AlphaFold2 prediction model for the TM domain (Fig. 1, C and D). Despite the low resolution, the secondary structures in the TM domain (TM1 to TM4, PCT-H1, PCT-H2, and MREβ sheet) agreed with the density maps well. In the final model, the Fab appears to fix the orientation of the PDZ tandem by docking to the PDZ-N domain in an orientation parallel to the membrane plane such that the V_L domain of the Fab makes contacts with the detergent micelle that covers the TM domain (Fig. 1B). The PDZ-N and PDZ-C domains adopt the clam shell-like arrangement observed in other RseP orthologs and form a pocket between them (PDZ-pocket; fig. S6) (*24*). In the refined structure, the PDZ-pocket was located above the PCT-hole that leads the TM-cavity where the active site for intramembrane cleavage is located. The overall structure of *Aa*RseP(WT) is consistent with the crystal structure of *Ec*RseP, except that the orientation of the two PDZ domains relative to the TM domain positions the PDZ-N above the PCT-hole in the *Ec*RseP structure (fig. S7, A and B).

During model building, we unexpectedly found a rod-shaped residual density inside the *Aa*RseP molecule that could not be

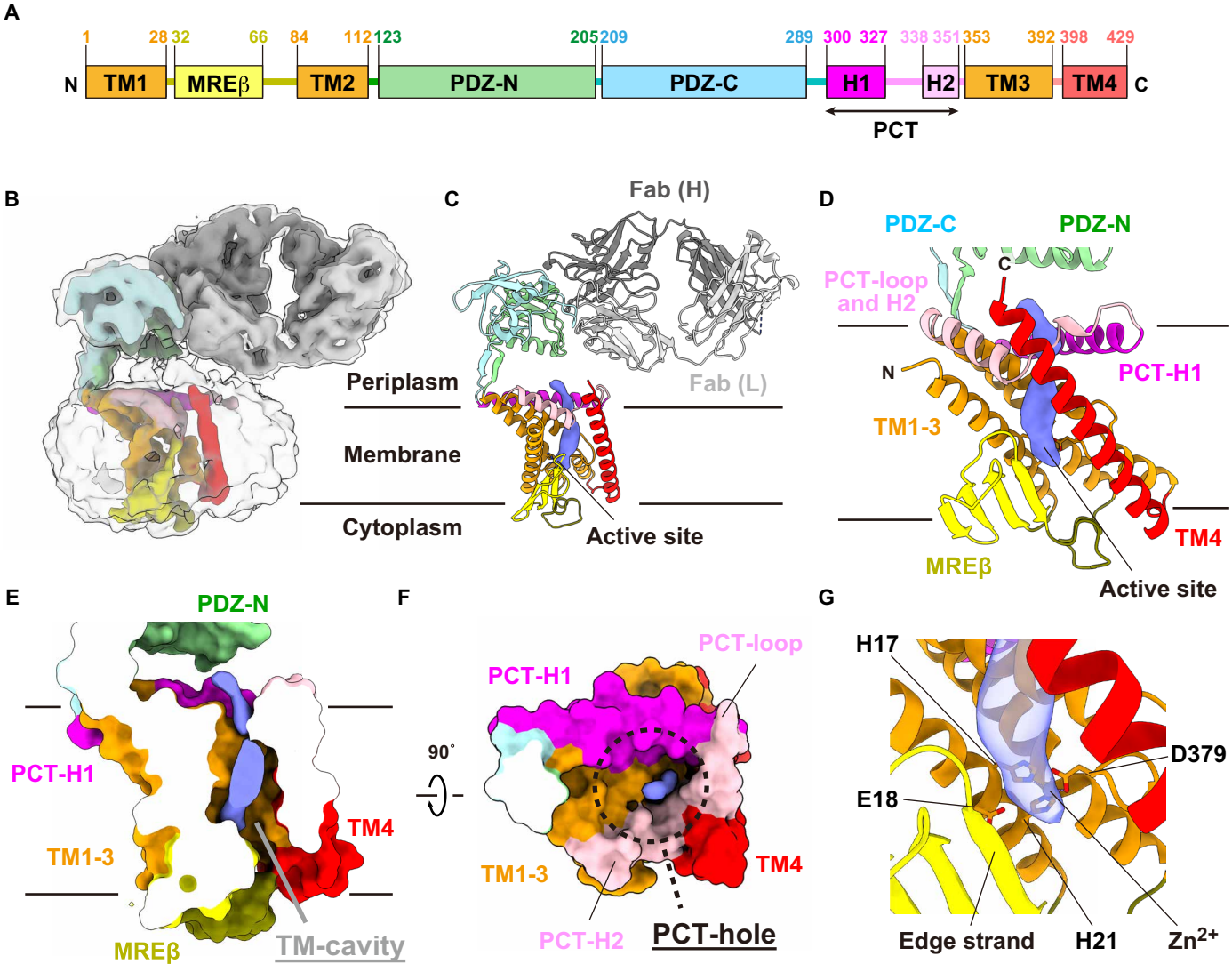

Fig. 1. Cryo-EM structure of *Aa*RseP(WT) complexed with Fab. (**A**) Domain organization of *Aa*RseP. (**B**) EM density map for *Aa*RseP(WT) complexed with Fab shown at two different contour levels. The map at the higher contour level is colored to match the assigned atomic model. (**C** and **D**) Assigned atomic models and residual densities shown in different views. The residual densities observed at 4σ level in the $F_o - F_c$ map are shown as blue surface models. (**E** and **F**) TM region of *Aa*RseP(WT) in different views. Cutaway surface models of the TM region are shown with the residual densities (blue surface) relative to the plane of the membrane in cross-section (E) and top-down (F) perspectives. (**G**) Close-up view of the active site. The side chains of zinc-coordinating residues and E18 that activates a water molecule for intramembrane proteolysis are shown as stick models. The edge of the residual densities (blue transparent surface) is located close to the zinc ion.

assigned to any portion of the expressed construct. The density entered the TM region from the PCT-hole, passed through the TM-cavity, and reached the vicinity of the zinc ion in the active site. Discontinuous but strong densities (>4σ) were also observed in the corresponding region of the $F_o - F_c$ map calculated using the EM density map and the refined atomic model (Fig. 1, D to G). To determine whether the orientation of the PDZ tandem is related to occupancy, we searched for ligand densities in all 3D classes. Residual densities were best resolved in the selected 3D class but were observed in all 3D classes. These observations suggested the presence of a cleavage product from an *E. coli* membrane protein in the TM-cavity. *Aa*RseP may not have released the substrate from the binding pocket because of the nonphysiological nature of the *E. coli* membrane protein substrate and of the expression in *E. coli* at 30°C compared to expression in *Aquifex aeolicus* at an optimal growth temperature of 95°C (*30*).

Cryo-EM SPA analysis of *Aa*RseP(E18Q) mutant

We next performed cryo-EM SPA on an active site mutant of *Aa*RseP to test if an uncleaved fragment could be observed in the TM-cavity. Specifically, we introduced the E18Q mutation, which has been reported to substantially reduce the substrate cleavage activity (*29*). We prepared the *Aa*RseP(E18Q) mutant in complex with the same Fab used with *Aa*RseP(WT) (fig. S8). We analyzed the conformational heterogeneity in the mutant through two rounds of 3D classification. We again observed rod-shaped densities within the TM region in all classes and selected a 3D class that had well-resolved densities in the TM region in each round of the 3D classification. Consequently, we

obtained a final 3D reconstruction model at 3.61-Å resolution (Fig. 2A and figs. S9 to S11), enabling us to trace the backbone for almost all residues including the PCT region (Fig. 2, B and C). In the resulting structure, the PDZ-N domain is located just above the PCT-hole, an orientation of the PDZ tandem relative to the TM region that differs slightly from that in the 3D class analyzed for *Aa*RseP(WT) (figs. S6, S7, and S11). In the density map for the E18Q mutant, the rod-shaped density was continuous from the PCT-hole to the MREβ sheet at the cytoplasmic side and comes in proximity to the active site zinc ion (Fig. 2, B to F). The cytoplasmic tail of the density was longer than that observed in the WT, and it aligned parallel to the edge strand of MREβ. The structural feature of the observed residual density is consistent with the previously proposed mechanism of strand addition for substrate docking. These observations support our hypothesis that a fragment of an *E. coli* membrane protein that is cleavable by *Aa*RseP is accommodated in the TM-cavity.

Ligand identification by mass spectrometry and in vivo cleavage experiments

We subsequently performed mass spectrometry (MS) to test the hypothesis that a fragment of an *E. coli* membrane protein is accommodated in the TM-cavity. We first performed native MS analysis on detergent-solubilized *Aa*RseP(E18Q) and observed ions with charges ranging from 10+ to 15+ under nondenaturing conditions (Fig. 3A). By deconvoluting the observed mass/charge ratio (*m*/*z*) values, the molecular mass of *Aa*RseP(E18Q) was calculated to be 49,337.6 ± 1.2, which agrees with the theoretical value calculated from the protein sequence. In the mass spectrum of the E18Q mutant, two other ion series (marked with # and † in Fig. 3A) were also observed, corresponding to molecular masses of 52,303.1 ± 2.6 and 57,625.3 ± 3.1. These mass values indicate that some populations of the E18Q mutant were associated with ligands of 2965.4 and 8287.6 Da, respectively. In contrast, no intense peaks for ligand-bound complexes were observed in the mass spectrum of *Aa*RseP(WT) whose molecular mass was calculated to be 49,338.7 ± 1.6 (Fig. 3A). These observations suggest that the uncleaved polypeptide ligand binds to the active site more stably in the E18Q mutant. We further attempted to identify the ligands by separating the purified *Aa*RseP(E18Q) sample through micro-reversed phase chromatography (Zip-Tip) and analyzing the fractions by MS. Consequently, we observed a peak of the isolated ligand at 2965.4 Da in addition to many background mass peaks at low *m*/*z* region that likely reflect co-eluted small molecules. In contrast, no ions corresponding to the 8287.6-Da ligand were observed (Fig. 3B). This absence could reflect a low initial concentration in the sample, depletion by the sample preparation, poor ionization, dilution below the detection limit from the wider range of charge states, or a combination of these factors. Subsequently, we further performed MS/MS analysis on the 2965.4-Da ligand and confirmed it to be a polypeptide by MS/MS analysis. Homology search using partial sequence showed that the isolated polypeptide matched a partial fragment of the *E. coli* endogenous enzyme heme *o* synthase (CyoE; Fig. 3B) (*31*). CyoE is composed of 296 amino acid residues and is predicted by AlphaFold2 to possess nine TM helices (https://alphafold.ebi.ac.uk/entry/A0A7T2JMK0) (*32*). The identified region

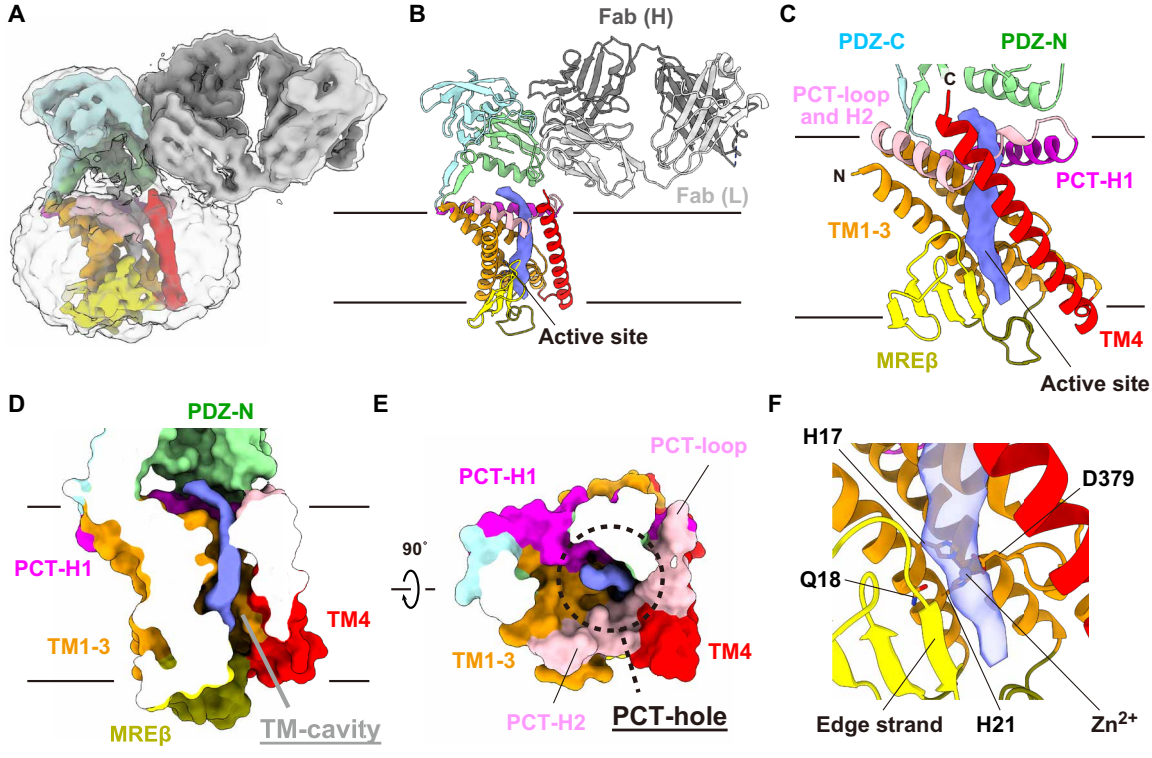

Fig. 2. Cryo-EM structure of *Aa*RseP(E18Q) complexed with Fab. Maps and models are shown for *Aa*RseP(E18Q) complexed with Fab as in Fig. 1 (B to G). (**A**) EM density map at two different contour levels. (**B** and **C**) Assigned atomic models and residual densities shown in different views. (**D** and **E**) Cutaway surface model of the TM region in different views. (**F**) Close-up view of the active site. The side chains of zinc-coordinating residues and the mutated residue (Q18) are shown as stick models. The cytoplasmic tail of the residual density (blue transparent surface) is parallel to the edge strand of the MREβ sheet.

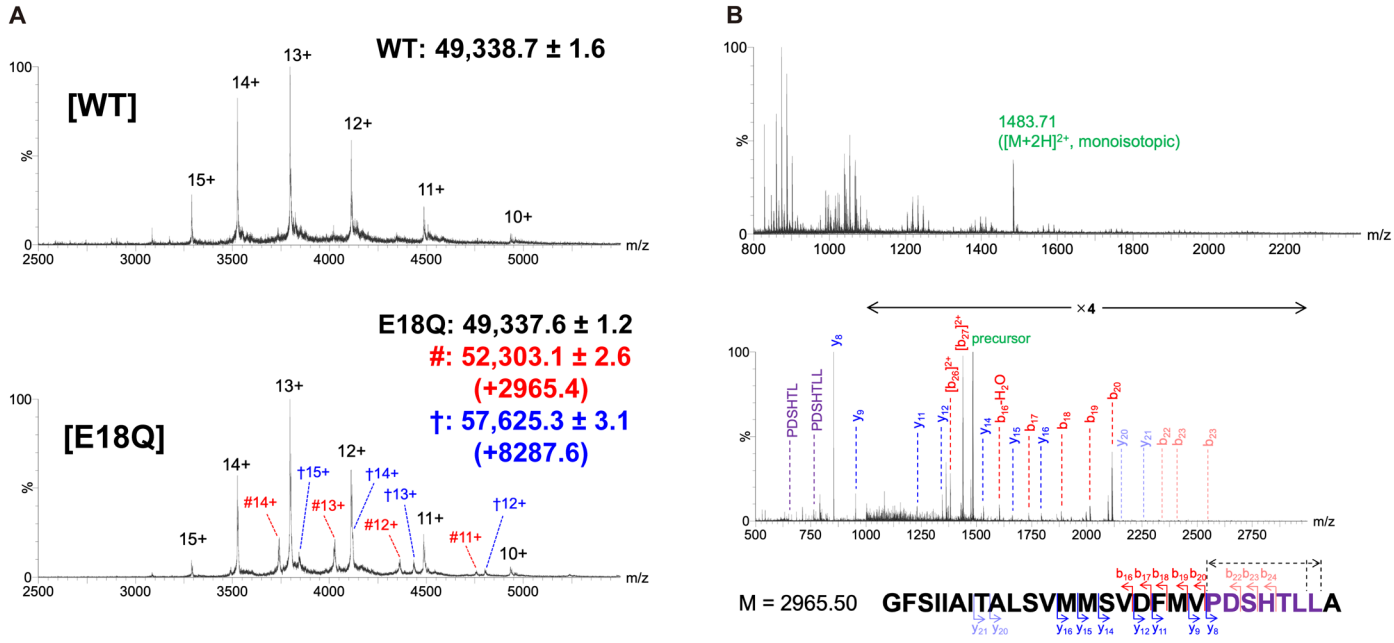

Fig. 3. Ligand identification by MS. (**A**) NanoESI mass spectra of *Aa*RseP(WT) and *Aa*RseP(E18Q) obtained under "near-native" condition [200 mM ammonium acetate containing 0.032% (w/v) DDAO (pH 7.4)]. In *Aa*RseP(E18Q), peaks of # (11 + ~14+) and † (12 + ~15+) correspond to molecular masses of 52,303.1 ± 2.6 (+2965.4) and 57,625.3 ± 3.1 (+8287.6), respectively, indicating two different ligands were bound to the protein. (**B**) Upper panel shows the mass spectrum of isolated ligands from an acetone-precipitated *Aa*RseP(E18Q) sample using micro-reverse phase chromatography with ZipTip C4. It is likely that a slight amount of chemicals co-eluted with the ligand were observed at low *m/z* region. The bottom panel shows the MS/MS spectrum with the monoisotopic ligand mass $[M+2H]^{2+}$ at *m/z* 1483.71 indicated in green. The ligand was identified from the product ions as GFSIIAITALSVMMSVDFMVPDSHTLLA (M = 2965.50, monoisotopic), corresponding to Gly-266 through Ala-293 from CyoE, an endogenous *E. coli* enzyme. Observed b ions, y ions, and internal ions are indicated in red, blue and purple, respectively.

corresponds to Gly-266 to Ala-293, 28 residues from the hydrophobic region of TM9 to the C-terminal periplasmic tail (Fig. 4A) (*33*). It is generally known that the S2P family members including RsePs cleave single-pass TM segments with an N-in/C-out orientation. One possibility is that a partial fragment of CyoE is generated during degradation in the *E. coli* cell membrane. The TM9 fragment could then be accommodated into the TM-cavity of *Aa*RseP(E18Q) in the N-in/C-out orientation that is consistent with that of typical S2P substrates.

To test if the TM9 fragment could be a substrate for *Aa*RseP, we performed an in vivo cleavage assay using model substrates of TM fragments from substrate candidates fused to an N-terminal hemagglutinin (HA) tag and a maltose-binding protein (MBP). In our previous study, we constructed a model substrate that contains the TM1 sequence (Phe-9 to Phe-29) of *E. coli* LacY, followed by the periplasmic residues of *E. coli* RseA (His-120 to Val-148), referred to as HA-MBP-RseA(LY1)148 (Fig. 4, B and C) (*34*). Our previous results showed that *Aa*RseP can efficiently cleave HA-MBP-RseA(LY1)148 (*29*). In this study, we constructed a model substrate with the CyoE fragment (Asp-257 to Val-281 on TM9 and Phe-282 to Trp-296 on the C-terminal tail), referred to as HA-MBP-CyoE(TM9+Ctail) (Fig. 4, B and C). We evaluated cleavage efficiencies based on the relative intensity for cleaved/uncleaved model substrates when coexpressed with *Aa*RseP(WT) as compared to that with *Aa*RseP(E18Q). Consistent with our previous results, the uncleaved HA-MBP-RseA(LY1)148 band almost disappeared when coexpressed with *Aa*RseP(WT). Similarly, the accumulation of uncleaved HA-MBP-CyoE(TM9+Ctail) was substantially reduced by coexpression with *Aa*RseP(WT) as compared to the condition with *Aa*RseP(E18Q), indicating that the TM9 fragment can be a substrate for *Aa*RseP (Fig. 4D). These observations are consistent with the residual density observed in the TM-cavity of *Aa*RseP(E18Q) representing a partial fragment of CyoE.

Binding mode of the substrate TM segment in the cavity

Following the results from the MS and biochemical analyses, we modeled a polypeptide chain into the residual density by locating the N terminus on the cytoplasmic side (MREβ sheet side) and the C terminus on the periplasmic side (PCT-hole side) (Fig. 5A). The resulting model consisted of 15 amino acid residues, whereas the CyoE fragment identified by the MS analysis was composed of 28 residues (Fig. 5, B and C). The main chain structure agreed well with the density map, but we could not assign the sequence due to the low quality of side chain densities. However, we could estimate the cleavage site within the polypeptide chain based on the density map and the binding mode. The density for the backbone atoms in the third residue appeared to be continuous with that of the zinc ion at the active site, suggesting the coordination by the C═O group (fig. S12). Thus, we assigned this residue as the P1 site on the substrate because the P1 C═O group generally coordinates the zinc ion in the Michaelis complex of metalloproteases. The third residue is located closest to the mutated catalytic side chain (E18Q) on TM1. In *Aa*RseP(WT), a water molecule located between the glutamate side chain and the zinc ion would perform a nucleophilic attack to cleave the peptide bond after the P1 site. From these positions, we assigned 15 residues in two groups based on their location relative to the zinc ion, P3-P1 and P1′-P12′. In the N-terminal region, the P3-P1 and P1′-P3′ residues adopted an extended conformation and aligned antiparallel to the edge strand (Leu-60 to Leu-66) of MREβ sheet (Fig. 5D). The P1′ residue and Gly-62 are positioned to form a hydrogen bond between

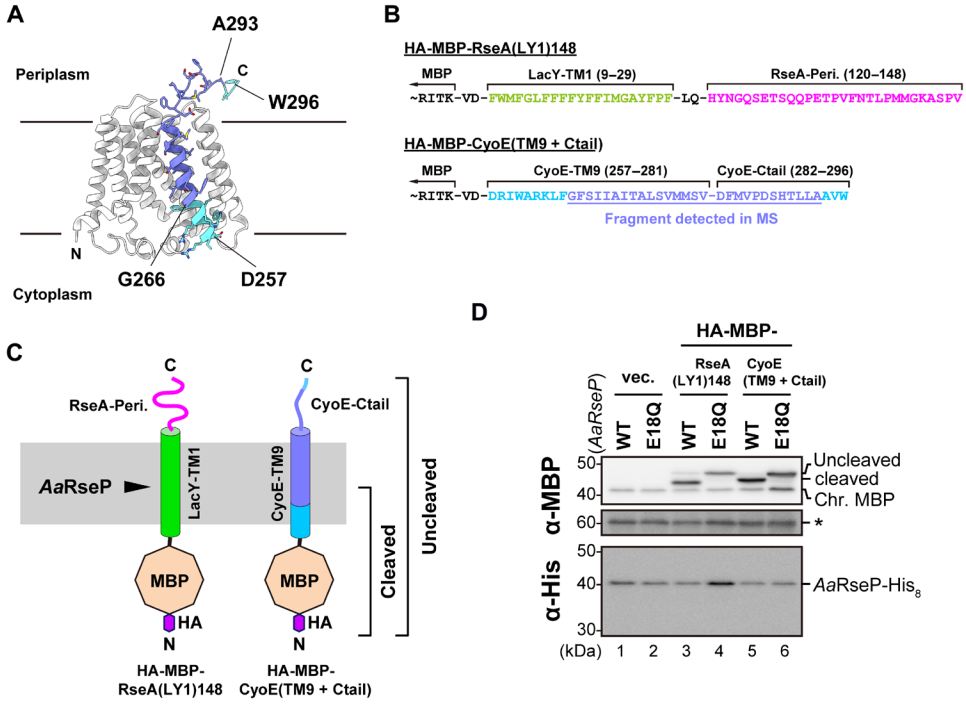

Fig. 4. In vivo cleavage of the CyoE model substrate by *Aa*RseP. (**A**) AlphaFold2 prediction model of CyoE. The residues colored medium blue (G266 to A293) corresponds to the fragment complexed with *Aa*RseP(E18Q). The region including the residues colored cyan (D257 to W296) was used in the model substrate for the cleavage assay. (**B**) Amino acid sequences of the TM and C-terminal periplasmic regions of each model substrate. (**C**) Schematic diagram of the model substrates. (**D**) In vivo cleavage assay. *E. coli* YK191 (Δ*rseA* Δ*rseP*) cells harbored plasmids for both substrate and protease: HA-MBP-RseA(LY1)148 (pYH124), HA-MBP-CyoE(TM9+Ctail) (pSY87), or vector control (pSTD689) and another one for *Aa*RseP-His_8 (pTM748) or its active site mutant (E18Q, pTM749). Cells were grown for 2 hours, and then, expression was induced for 2 hours. Total cellular proteins were TCA-precipitated and analyzed by Laemmli SDS PAGE with anti-MBP or anti-His immunoblotting. An asterisk indicates a nonspecific band serving as a loading control. Chr. MBP indicates endogenous MBP. A representative result from three biological replicates is shown.

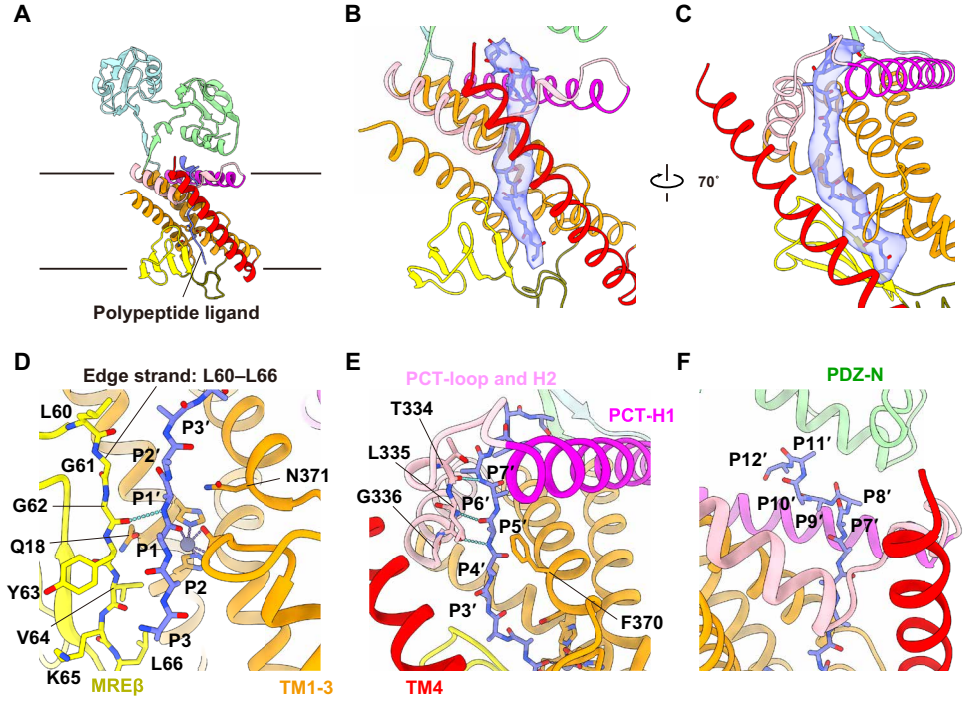

Fig. 5. Modeling the polypeptide ligand. (**A**) Overall structure of *Aa*RseP(E18Q) with the assigned polypeptide ligand. (**B** and **C**) TM region in two different views. The residual density is shown as a transparent blue surface. The assigned polypeptide ligand is shown as a stick model. (**D** to **F**) Close-up view of the regions around P3-P1 and P1′-P3′ (D), P3′-P7′ (E), and P7′-P12′ (F). Hydrogen bonds between the main chain atoms of the polypeptide ligand and *Aa*RseP are shown as cyan dotted lines.

the main chain atoms. In addition, the backbone hydrogen bond donors and acceptors on P2 and Val-64 appear oriented toward each other and may form main chain to main chain interactions in some bound conformations. This arrangement of the P3-P1 and P1′-P3′ residues is consistent with that suggested by the previous cross-linking analysis on *Ec*RseP (*26*). Cysteine residues that were introduced at Tyr-69 and Lys-71 of *Ec*RseP on the MREβ sheet (corresponding to Tyr-63 and Lys-65 on *Aa*RseP) efficiently formed disulfide bridges with the cysteine residue on the substrate, which supports the close proximity of these residues to the substrate. Crystallographic analysis of *Ec*RseP in complex with batimastat suggested that this position of the substrate backbone is clamped by the residue corresponding to N371 (*24*), and this residue is highly conserved throughout the S2P family. In our model, N371 on *Aa*RseP was located close to the substrate-like ligand (Fig. 5D) and in position to form hydrogen bonds with the main chain atoms of the P2′ residue as expected. At the P4′ site, the polypeptide chain shifts slightly to allow P4′-P7′ to pass through a narrow part of the TM-cavity created by the PCT-loop and F370 on TM3 (Fig. 5E). The P4′-P7′ residues also adopted an extended conformation and formed main chain to main chain interactions with Thr-334 to Gly-336 on the PCT-loop. In contrast to the N-terminal residues, the subsequent P8′-P12′ residues were assigned a helical conformation (Fig. 5, A and F). These five C-terminal residues protruded from the TM-cavity and make few contacts with *Aa*RseP. The C-terminal region of the polypeptide ligand was surrounded by the U-shaped arrangement of PCT-H1, PCT-loop, and PCT-H2 on the membrane surface. This arrangement appears to prevent the accommodated polypeptide chain from laterally diffusing into or out of the TM-cavity.

In vivo photocrosslinking analysis to examine the substrate accommodation inside the PCT-hole

In our assigned model, the substrate-like polypeptide is threaded through the PCT-hole, which is inaccessible to a membrane-spanning polypeptide without notable structural changes in the PCT region and/or the polypeptide. However, the *Aa*RseP molecule subjected to the cryo-EM SPA was solubilized in detergent micelles, which may permit a higher degree of freedom for the interactions between *Aa*RseP and the polypeptide compared to the membrane-bound conformation. Given the possibility of structural changes, we could not exclude the possibility that the substrate binding mode in the lipid bilayer is different from that observed in the cryo-EM structure. We therefore attempted to examine whether substrates are accommodated into the same site when *Aa*RseP catalyzes proteolysis within the cell membrane. For that purpose, we performed in vivo photocrosslinking analysis in *E. coli* using the model substrate HA-MBP-RseA(LY1)148. We introduced *p*-benzoyl-L-phenylalanine (*p*BPA), a photoreactive amino acid analog, into *Aa*RseP(E18Q) at multiple residues located on the inner wall of the TM-cavity (Fig. 6A). We also introduced *p*BPA at position Y63 on the edge strand as a positive control (Fig. 6A). In *Ec*RseP, the *p*BPA residue introduced at Y69 (corresponding to Y63 on *Aa*RseP) was efficiently cross-linked with the endogenous RseA (*24*). As expected, a higher molecular weight band (~80 kDa) for the cross-linked product appeared when the *E. coli* expressing the Y63*p*BPA mutant was exposed to ultraviolet (UV) irradiation (Fig. 6B). Furthermore, cross-links were formed with high efficiency when *p*BPA was substituted with T321 on PCT-H1, I344 on PCT-H2, and F370 on TM3 (Fig. 6B). Cross-linked bands were also observed when *p*BPA was introduced at T318 on PCT-H1, L335 on PCT-loop, I341 on PCT-H2, and I363 and L367 on TM3 though the cross-link efficiencies were relatively low (Fig. 6B). These observations suggested that the model substrate containing the LacY TM1 sequence was accommodated in *Aa*RseP in a manner similar to that observed in the model. To ensure that these mutations did not alter *Aa*RseP activity, we also introduced the same set of *p*BPA mutations into *Aa*RseP(WT) and performed the in vivo cleavage assay. Consequently, we confirmed that all the *p*BPA mutants maintained cleavage activity for the model substrate (Fig. 6C), indicating that the detected cross-links were formed with the substrate in a cleavage-competent conformation. Last, we selected *p*BPA mutants that showed high cross-linking efficiency with HA-MBP-RseA(LY1)148 (T321*p*BPA on PCT-H1, I344*p*BPA on PCT-H2, and F370*p*BPA on TM3), and we repeated the cross-linking analysis with the CyoE-based substrate HA-MBP-CyoE(TM9+Ctail). Similar to the above cross-linking analysis with the LY1-based substrate, all three *p*BPA mutants indeed showed distinct cross-linking bands (Fig. 6D). These results support the notion that *Aa*RseP expressed on the cell membrane also accommodate substrates into the TM-cavity isolated by the PCT region, as observed in the cryo-EM structure of *Aa*RseP(E18Q).

DISCUSSION

Our present cryo-EM analysis unexpectedly revealed the 3D structure of the substrate-bound state for an S2P family member, specifically *Aa*RseP. The observed binding mode of the substrate-like polypeptide ligand provides us with insights into the substrate binding mechanism that is generalizable to other S2P family members, including the homologs from pathogenic bacteria. Critically, these structural data prompt us to modify the previously proposed multistep model for S2P substrate discrimination and accommodation. Ultimately, these insights may help us design selective inhibitors for the bacterial group I S2Ps. Our cryo-EM structures for both WT and active site mutant (E18Q) of *Aa*RseP solubilized in detergent micelles revealed residual densities in the TM-cavity, and for the E18Q mutant, this density could be assigned as a polypeptide ligand with identifiable structural features and mechanistically interpretable binding mode. Native MS analysis also showed the presence of ligands that did not dissociate from *Aa*RseP(E18Q) under collision energies that were sufficient to strip detergent from the complex. Furthermore, the MS/MS analysis revealed that one of the ligands bound to *Aa*RseP(E18Q) was a partial fragment of CyoE, a multipass TM protein endogenous to *E. coli*. As our biochemical analysis showed that *Aa*RseP efficiently cleaves the CyoE fragment, we assigned a 15-residue polypeptide chain to the residual density. The side-chain structures remained unresolved due to ambiguity in the density map, but the model assignment elucidated the conformation of the polypeptide backbone and its binding mode with *Aa*RseP. Most residues in the polypeptide chain adopted an extended conformation and penetrated through the TM-cavity formed inside the bundle of TM helices. In addition, the C═O group of the putative P1 site appeared to coordinate the zinc ion, indicating that the model represents the ES state. The CyoE fragment should form a TM helix before the accommodation into *Aa*RseP. Thus, our structural data suggest that the fragment is unwound by *Aa*RseP during substrate accommodation and cleavage.

Regarding the binding mode of the substrate cleavage segment, the resulting cryo-EM structure of the substrate-bound *Aa*RseP

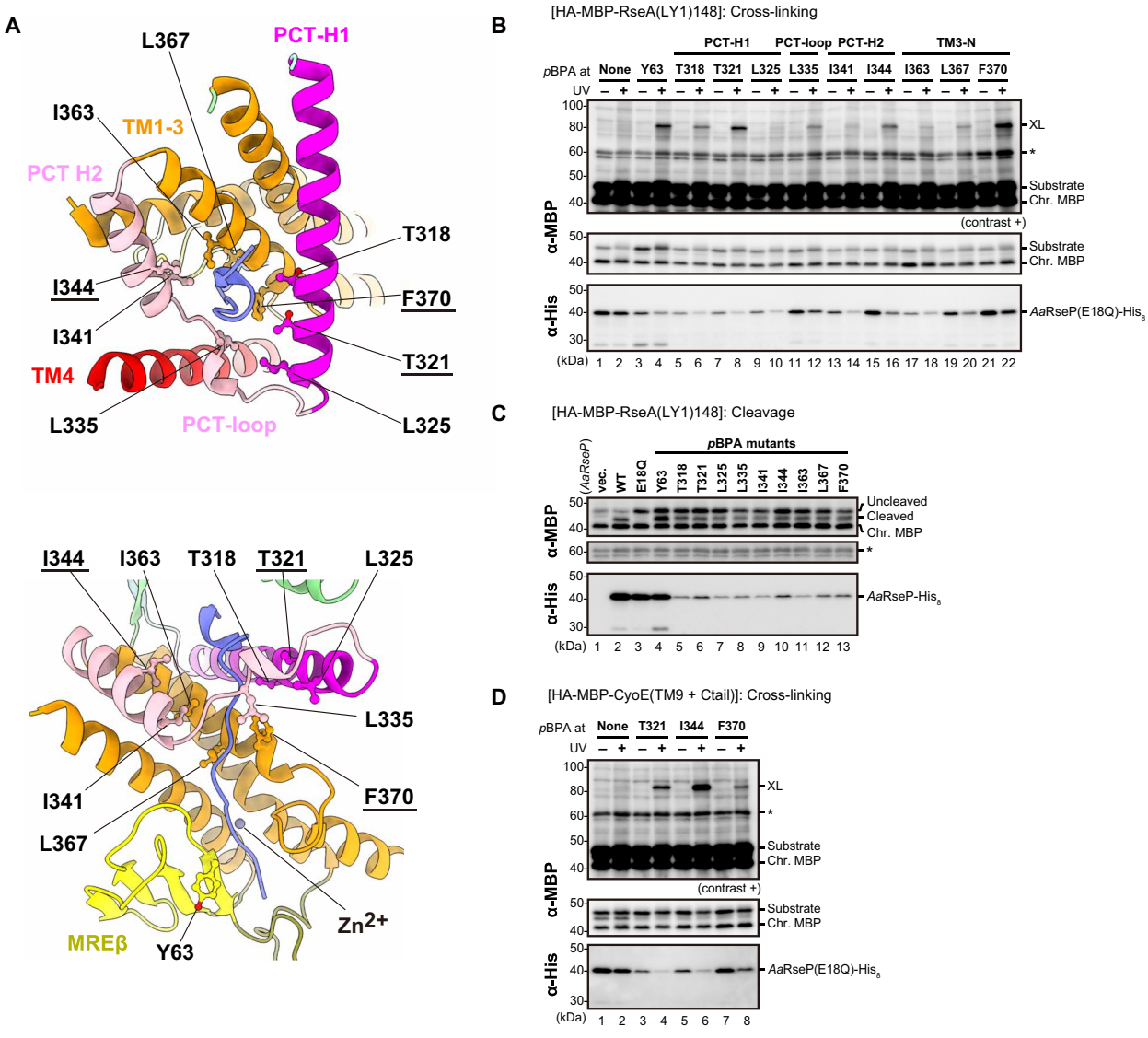

Fig. 6. In vivo photocrosslinking between *Aa*RseP and the model substrate. (**A**) Residues tested for cross-linking analysis. The residues mutated to *p*BPA are shown in ball-and-stick models. Residues where the introduced *p*BPA was efficiently cross-linked with HA-MBP-RseA(LY1)148 are underlined. TM region is shown relative to the plane of the membrane in top-down (top) and cross-section (bottom) perspectives. TM4 is omitted to visualize the assigned polypeptide ligand. (**B**) In vivo photo-cross-linking analysis. *E. coli* YK191 (Δ*rseA* Δ*rseP*) cells harboring both pEVOL-pBpF and pSY41 [HA-MBP-RseA(LY1)148] were transformed with pNO1461 [*Aa*RseP(E18Q)-His_8 (none)] or its derivatives having an amber mutation at the indicated positions. Cells were cultured in M9 medium with *p*BPA and then irradiated with ultraviolet (UV) light. TCA-precipitated proteins were analyzed as in Fig. 4D. XL indicates the predicted cross-linked products between *Aa*RseP-His_8 and the model substrate. An asterisk indicates the nonspecific band that serves as a loading control. Residue Tyr-63 on *Aa*RseP corresponds with Tyr-69 from *Ec*RseP on the edge strand of the MREβ sheet and was used as a positive control for cross-linking. A representative result from three biological replicates is shown. (**C**) Proteolytic activity of the *Aa*RseP *p*BPA mutants. *E. coli* YK191 (Δ*rseA* Δ*rseP*) cells harboring both pEVOL-pBpF and pSY41 [HA-MBP-RseA(LY1)148] were transformed with pUC118 (vec.), pNO1457 [*Aa*RseP-His_8 (WT)], pNO1461 (E18Q), or pNO1457 derivatives having an amber mutation at the indicated positions. Cells were cultured in M9 medium with *p*BPA. TCA-precipitated proteins were analyzed by Laemmli SDS-PAGE with anti-MBP and anti-His immunoblotting. An asterisk indicates the nonspecific band that serves as a loading control. A representative result from three biological replicates is shown. (**D**) In vivo photocrosslinking analysis for CyoE model substrate. Sampling and immunoblotting were performed as in (C) except that pSY137 [HA-MBP-CyoE(TM9+Ctail)] were used as the model substrate-expressing plasmid. A representative result from three biological replicates is shown.

mutant was consistent with the models that have been previously proposed based on the cross-linking and structural analyses (*24*, *26*). Specifically, the conservation of the intramembrane β sheet and the Asn clamp are further accumulating evidence that most S2P family members share this common substrate binding mechanism. Our cryo-EM structure supported the previous prediction from cross-linking and structural analyses that the polypeptide ligand interacts with the edge strand of the MREβ sheet by the strand addition mechanism. Because of the low resolution of density maps and the absence of secondary structure restraints in the refinement in our cryo-EM structure, we do not observe continuous hydrogen bonding pairs between the six N-terminal residues of the polypeptide ligand (putative P3-P1 and P1′-P3′ residues) and the edge strand. However, the P1′ residue is oriented to form a hydrogen bond, while

other residues also appear poised to form hydrogen bonds after a slight shift in binding conformation (Fig. 5D). Structure prediction by AlphaFold2 indicates that the intramembrane β sheet near the active site is a conserved structural module among the S2P family members, although the strand-order topology differs depending on the species (fig. S13). This conclusion is also supported by experimental structure determination. The crystal structure of a group III S2P from *Methanocaldococcus jannaschii*, *Mj*S2P, also has an intramembrane β sheet (*35*). In addition, a very recent cryo-EM analysis of a group III S2P from *Bacillus subtilis*, SpoIVFB, elucidated that the membrane reentrant β loop serves as the substrate-binding site (*36*). While all known substrates for the group I S2Ps including RseP have an N-in/C-out orientation, SpoIVFB accommodates an unfolded substrate Pro-σ^K in the opposite N-out/C-in orientation in the cryo-EM structure. Despite this difference in orientation, SpoIVFB also binds with the substrate cleavage segment through the main chain to main chain interactions via the membrane reentrant β loop, which was described as a β sheet augmentation in that study. Thus, it is highly possible that substrate accommodation via an extension of an intramembrane β sheet is a conserved mechanism throughout the S2P family. In addition, the previously proposed involvement of the conserved Asn residue (N371 on *Aa*RseP) in substrate binding is also supported by this current model assignment. Our previous structure-based mutational analysis on *Ec*RseP suggested that the side-chain amide group of the Asn residue at the corresponding site (N394 on *Ec*RseP) forms hydrogen bonds with the main chain C═O and N─H groups of the substrates. The density of the N371 side chain was largely disordered, but the orientation of the Cα-Cβ bond would point the side-chain atoms toward the main-chain atoms of the polypeptide ligand.

Furthermore, our present cryo-EM structures suggest the presence of an additional substrate accommodation mechanism in the bacterial group I S2Ps. The most important finding regarding the substrate binding mode is the penetration of the C-terminal tail through the PCT-hole. The PCT-hole is surrounded by the peripheral membrane-associated PCT-H1, PCT-loop, and PCT-H2 as well as TM3 and is located at the periplasmic end of the TM-cavity. Thus, the polypeptide ligand in the cryo-EM structure is separated from the lipid bilayer by the PCT region that sits above the membrane on the periplasmic side. Our previous cross-linking analysis on *Ec*RseP demonstrated that the PCT-H1 and TM3 residues located inside the TM-cavity were efficiently cross-linked with the physiological substrate RseA (*24*, *27*). This prior result indicated that RseA could be accommodated deep inside the TM-cavity, but we expected that PCT-H2 and PCT-loop move out to create space for substrates. However, our present cross-linking analysis indicates that the substrate can be accommodated inside the PCT-hole when *Aa*RseP is present in the more physiologically relevant environment of the cell membrane. Our previous and present structural analyses demonstrated that RseP orthologs share the structural feature of a PCT-region. Structure prediction by AlphaFold2 further suggests that this structural feature is common among other bacterial group I S2Ps (fig. S13) (*32*). Thus, it is possible that a unique and conserved mechanism for substrate accommodation among the bacterial group I S2Ps is substrates threading through the hole formed by the membrane-associated region. Structural analyses suggested that the substrate accommodation by group III S2Ps is largely different as their active sites appear to be accessible to membrane-spanning substrates via lateral entry due to the absence of structural elements corresponding to the PCT region (fig. S13) (*35*, *36*). In the bacterial group I S2Ps, the substrate threading is presumed to stabilize the substrate-binding state and to promote the substrate cleavage in the TM-cavity. Furthermore, there is a possibility that the threading also contributes to substrate discrimination. It is known that RseP orthologs cleave single-pass TM segments with short soluble domains. Because of the threading mechanism, internal TM segments with an N-in/C-out orientation in multipass membrane proteins would be excluded from the TM-cavity and would not act as substrates. One exception to this may be multipass membrane proteins with a C-out topology, in which the C-terminal TM helix could be thread through the PCT-hole as its orientation is consistent with that of the type II membrane protein. Our previous proteome analysis indicated that the *Ec*RseP activity substantially reduced the accumulation level of CyoE, which is a nine-TM protein with a C-out topology (*21*). It remains to be addressed whether some full-length multipass membrane proteins with a C-out topology could be substrates for bacterial group I S2Ps.

On the basis of the results from this study, we modified our model for intramembrane proteolysis by *Ec*RseP and its orthologs. We previously proposed that substrate discrimination and accommodation are regulated by three processes: (i) size exclusion, (ii) gating, and (iii) unwinding (*24*). In RsePs, the size exclusion mechanism has been proposed, wherein the PDZ tandem only permits the entry of type II TM proteins with short periplasmic domains to approach the TM region (*28*). Subsequently, PCT-H2 and PCT-loop, together with TM4, possibly serve as the substrate entry gate and undergo structural changes from closed to open conformations to accommodate the substrate into the active site. In the previous model, we had anticipated that PCT-H2 and PCT-loop would adopt a relatively closed, or partially open, conformation so as to permit the lateral entry and accommodation of substrate. However, the present study strongly suggests that the PCT-H2 and PCT-loop return to a completely closed conformation to isolate the C-terminal tail of the substrate from the lipid bilayer. After accommodation, the substrate would be locally unwound by the MREβ sheet and the conserved Asn residue in the TM-cavity (Fig. 7).

The revised model for the substrate accommodation raises another question as to how the substrates enter the PCT-hole on the membrane. If PCT-H2 maintains its helical structure on the membrane surface during the substrate entry, then substrates would need to undergo notable structural changes to pass through the PCT-H2 where the periplasmic tail remaining after the site-1 cleavage would be transiently embedded in the lipid bilayer. This process should be energetically unfavorable if it occurs, as the periplasmic soluble domains are generally abundant in hydrophilic residues and membrane deformation has a high energy penalty (Fig. 4B). Another scenario would be that PCT-H2 transiently unfolds to permit substrate entry. Evidence exists to support this model. We previously performed a cell-based Cys-scanning accessibility assay on PCT-H2 using a probe to react with Cys residues exposed to the aqueous milieu and found that all residues on PCT-H2 could be modified by the probe to comparable levels (*24*). This result suggested that PCT-H2 is not fixed. Rather, it is highly mobile on the *E. coli* membrane surface. Further observations indicated that PCT-H2 is mobile. As mentioned above, PCT-H2 and the PCT-loop are unfolded and dissociated from the membrane peripheral region in the crystal structure of *Kk*RseP (*24*). These structural changes in the PCT region were affected by the crystal packing where TM4 on the crystal packing neighbor docked to the cleft in the TM-cavity, suggesting that

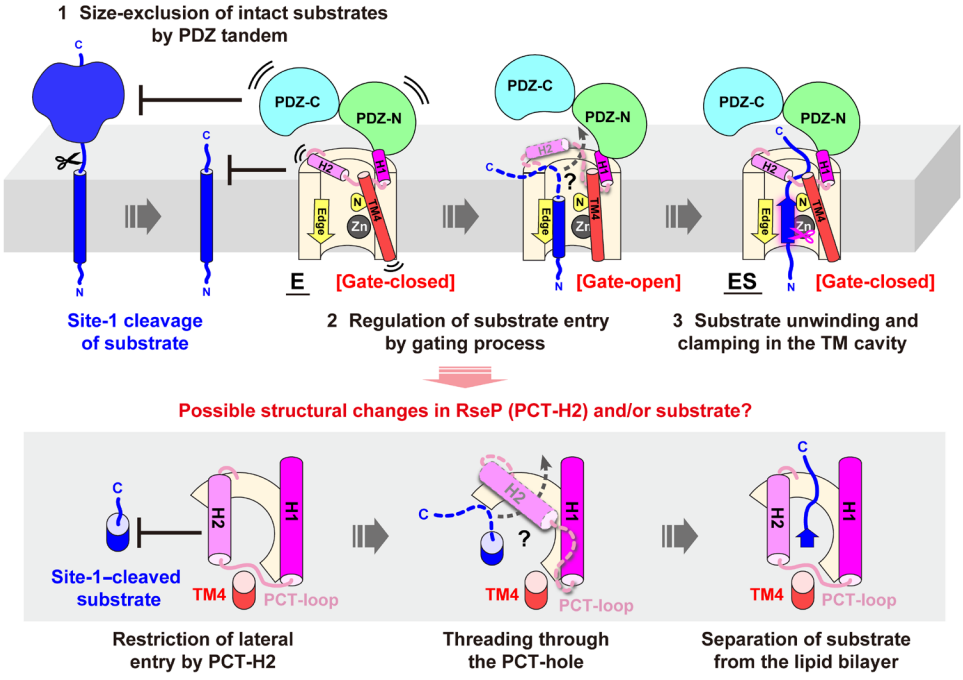

Fig. 7. Model for structure-based substrate discrimination and accommodation. RseP is presumed to regulate the substrate discrimination and accommodation by size-exclusion, gating, and unwinding processes. (i) Size exclusion: The PDZ tandem excludes bulky intact substrates by steric hindrance. After the site-1 cleavage, the size-reduced substrates become accessible to RseP. (ii) Gating: Despite conformational fluctuations in the regions surrounding the active site (PCT-H2, TM4, and PDZ tandem), substrate entry into the TM-cavity by lateral diffusion is restricted by a gate formed by the helical segment of PCT-H2 that interacts with TM4 at the periplasmic membrane surface. However, the present structural analysis demonstrated that the substrate is accommodated inside the PCT-hole. Therefore, RseP and/or substrates could undergo drastic structural changes to permit the substrate entry. Transient unfolding of PCT-H2 along with PCT-loop would allow substrates to be thread through the PCT-hole. After the entry into the TM-cavity, substrates would be separated from the lipid bilayer as the PCT-H2 region refolds and TM4 packs closely. (iii) Unwinding: In the separated TM-cavity, the substrate adopts an extended conformation and the local geometry of the intramembrane cleavage site is pre-organized by the conserved Asn residue clamp (labeled N with a yellow round triangle). This conformation, observed in this study, would restrict dissociation of unprocessed substrate out of the active site by lateral diffusion. The results from this study suggest that the accommodation inside the PCT-hole is a potential mechanism by which RseP discriminates single-pass membrane proteins from multipass membrane proteins because termini of substrates are required for threading. In this model, it is possible that the discrimination mechanism is permissive to C-out oriented C-terminal TM helices on full-length multipass membrane proteins.

these regions can undergo structural changes through interactions with substrates. The importance of this helix-swapped structure for the binding mechanism remains unclear. If PCT-H2 and PCT-loop transiently unfold, then the substrate could enter the PCT-hole without embedding the hydrophilic periplasmic domain into the hydrophobic lipid bilayer (Fig. 7).

We can now propose that substrate gating process could be a possible target for the selective inhibition of specific bacterial group I S2Ps. Our previous biochemical analysis on *Ec*RseP showed that disrupting the gate structure substantially reduced the substrate cleavage efficiency (*24*), suggesting that blocking the gating process inhibits substrate cleavage. Because the presence of a peripheral membrane-associated region corresponding to the PCT region is a conserved feature, substrate threading in the gating process is likely a conserved mechanism among the bacterial group I S2Ps. However, the sequence and structure of these putative PCT regions differ between evolutionally distant species. For instance, AlphaFold 2 predicts that the PCT region of *M. tuberculosis* contains a loop region with longer helical elements as compared to those of the RseP orthologs (fig. S13). Thus, the differences in this region alone may be sufficient to design inhibitors for specific S2Ps. In addition, the structure prediction indicates that the peripheral membrane-associated region of human S2P diverges largely from those of bacterial group I S2Ps (fig. S13), which is also preferable for the development of selective inhibitors without adverse effect.

Looking across all IMPs that cleave single-pass TM segments, lateral entry of substrates by threading appears to be common. The aspartate protease-type γ-secretase is also known to cleave single-pass TM proteins with short soluble domains (*37*), but it has the opposite preference for the substrate TM helix orientation. All known substrates of γ-secretase show an N-out/C-in topology. For γ-secretase, the standard assumption is that the substrate laterally diffuses into the active site between TM helices, but cryo-EM structures of γ-secretase show substrates such as Notch bound in a conformation where lateral diffusion in and out of the binding site is sterically blocked by a peripheral loop that connects the TM helices. Therefore, the substrates have been proposed to be threaded through this peripheral loop when they are accommodated into the internal binding site (*38*). It is possible that the substrate accommodation by threading is unique to the IMPs, which all have limited degrees of freedom on the membrane relative to soluble proteases. Design of inhibitors that target the region involved in the substrate threading may be a promising strategy for the development of the selective inhibitors for γ-secretase as well.

In summary, we have elucidated the cryo-EM structure of an *Aa*RseP mutant, providing the first 3D structure of the bacterial

group I S2P with a substrate-like polypeptide ligand accommodated into the TM cavity that houses the active site. Our in vivo photocrosslinking experiment confirmed that the membrane-spanning polypeptide is surrounded by the continuous membrane-associated helices of the PCT region and therefore separated from the lipid bilayer in the bound state. Our structural data allow us to propose a model for substrate discrimination and accommodation, including substrate threading, and suggest structural criteria for the development of selective inhibitors, including targeting the gate region that varies depending on the species. Nevertheless, the precise mechanism of substrate entry remains unknown as the low resolution of the density map and possible heterogeneity in the polypeptide ligand hampered assignment of the side-chain structure in the polypeptide model and identification of specific enzyme-substrate interactions. Further understanding of the substrate discrimination and accommodation mechanism, which should inform rational inhibitor design, awaits structure determination of *Aa*RseP or other S2P homologs in complex with their physiological substrates at higher resolutions.

MATERIALS AND METHODS

Generation of monoclonal antibody for *Aa*RseP and production of Fab

Anti-*Aa*RseP antibody 4A9 was established by immunizing GANP transgenic mice (*39*) with the PDZ tandem fragment of *Aa*RseP in the previous study (*28*). In this study, the primary structures of the heavy and light chains of 4A9 were determined by cloning their cDNAs according to the same procedure as reported previously (*24*, *40*, *41*).

For overproduction, 4A9 Fab was produced as a fusion protein with human growth hormone (hGH) and octahistidine tag. Specifically, the cDNA encoding the VH-CH1 region was amplified with primers that added an additional Ser residue at the N terminus and then was subcloned into the pSGHV0 vector (*42*, *43*). 4A9 Fab used for the cryo-EM analysis contained a dodecapeptide PA tag, GVAMPGAEDDVV, (*44*) downstream of the CH1 region, while that used for the crystallization contained no additional sequence. The cDNA encoding the light chain with the original signal sequences were subcloned into the pCAG-Neo vector (FUJIFILM Wako Pure Chemical Corporation, Osaka, Japan). Expression plasmids for the heavy and light chains were mixed and transfected into Expi293F cells (Thermo Fisher Scientific) using the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific). The culture supernatant was harvested by centrifugation and neutralized with tris-HCl (pH 8.0) at a final concentration of 20 mM. The culture supernatant was incubated with Ni-NTA agarose resin (QIAGEN) and then washed with 20 mM tris-HCl (pH 8.0), 300 mM NaCl, and 10 mM imidazole. Subsequently, the captured Fab was eluted with 20 mM tris-HCl (pH 8.0), 300 mM NaCl, and 250 mM imidazole. N-terminally hexahistidine-tagged TEV protease (His_6-TEV) was added to the elution fraction to cleave off the N-terminal hGH and His_8 tags and dialyzed against 10 mM tris-HCl (pH 8.0), and 150 mM NaCl overnight at 4°C. After removing the cleaved hGH-His_8 portion and His_6-TEV by passing the fraction through Ni-NTA agarose resin, the flow-through and wash fractions containing Fab were further purified by gel filtration on a Superdex 200 Increase 10/300 GL (Cytiva) column equilibrated in 10 mM tris-HCl (pH 7.4) and 150 mM NaCl. The peak fractions containing 4A9 Fab were concentrated by ultrafiltration using an Amicon Ultra Centrifugal Filter with a molecular weight cutoff of 30 kDa (Merck Millipore).

Crystallographic analysis of 4A9 Fab

Initial crystallization conditions were searched using the Index (Hampton Research) screening kit. Each of protein solution and crystallization buffer (0.2 μl) was dispensed into 96-well plates using a Crystal Gryphon (Art Robbins Instruments) and equilibrated against 60 μl of crystallization buffer in the reservoir by the sitting-drop vapor diffusion method. 4A9 Fab was crystallized in a solution containing 25% (w/v) polyethylene glycol 3350 and 50 mM zinc acetate. Crystals were quickly soaked in the cryoprotectant, which was prepared by mixing the crystallization buffer and ethylene glycol in a ratio of 4:1 (v/v) and frozen in liquid nitrogen. As the crystals were microcrystals, x-ray diffraction data were collected at SPring-8 beamline BL32XU (*45*) using an EIGER X9M detector (Dectris). Microfocused x-rays with a beam size of 15 μm by 10 μm at wavelengths of 1.0000 Å were used for both raster scan and data collection. A dataset with a total oscillation range of 10° and 0.1° oscillations per frame was collected from each crystal under an absorbed dose of 10 megagraqy. The partial datasets collected with the automated data collection system ZOO (*46*) were merged, integrated, and scaled using the KAMO system (*47*), which integrates BLEND (*48*), XDS, and XSCALE (*49*, *50*). Diffraction intensities were converted to structure factors using the CCP4 suite where 5% of the unique reflections were randomly selected as a test set for the calculation of free *R* factor (*51*).

For structure determination, molecular replacement was performed with Molrep in CCP4 using the crystal structure of the mouse IgG Fab raised against the PDZ tandem of an RseP homologue from *Aquifex aeolicus* [Protein Data Bank (PDB) code: 3WKM] (*28*). As a result, two Fabs were assigned in the asymmetric unit. After rebuilding and refining the initial model by ARP/wARP (*52*), models were manually modified and fit into the electron density map using the program COOT (*53*). The updated models were refined with phenix.refine (*54*) iteratively while monitoring the stereochemistry with MolProbity (*55*). Statistics for data collection and refinement are summarized in table S1, respectively.

Purification of *Aa*RseP and complex formation with Fab

For the cryo-EM analysis, the *Aa*RseP-His_8 WT and the E18Q mutant were overproduced in *E. coli* KK374 (Δ*rseA*, Δ*rseP*, and Δ*degS*) (*56*). *E. coli* KK374 cells transformed with the expression plasmids, pNO1457 or pNO1461 (*29*), were grown at 30°C to an optical density at 600 of 0.7 in LB medium [tryptone (10 g/liter), yeast extract (5 g/liter), and NaCl (10 g/liter); without pH adjustment] supplemented with ampicillin (50 μg/ml), followed by induction of overexpression with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and culture at 30°C for additional 3 hours. After harvest by centrifugation, cells were lysed by sonication in 10 mM tris-HCl (pH 7.4). Subsequently, the lysate was centrifuged at 40,000*g* for 45 min. at 277 K to separate the cell debris. The supernatant was further separated by ultracentrifuge at 200,000*g* for 60 min at 277 K. The precipitant was suspended in 10 mM tris-HCl (pH 7.4) and was ultracentrifuged again under the same conditions. Last, the precipitant was suspended in 10 mM tris-HCl (pH 7.4) and collected as membrane fraction. The total protein was quantified by the bicinchoninic acid assay using bovine serum albumin as a standard, and the protein concentration was adjusted to 10 mg/ml.

Membrane proteins were solubilized by adding the same volume of a solubilization buffer containing 40 mM tris-HCl (pH 7.4), 150 mM NaCl, and 2% *n*-dodecyl-*N*, *N*-dimethylamine-*N*-oxide (DDAO) to

the suspension of the membrane fraction. After incubation at 277 K for 1 hour, the mixture was ultracentrifuged at 210,000*g* for 60 min. at 277 K. The supernatant was mixed with Ni-NTA agarose resin, and the unbound fraction was washed out with a buffer containing 20 mM tris-HCl (pH 8.0), 300 mM NaCl, 20 mM imidazole, and 0.05% DDAO. The resin was further washed with a buffer containing 20 mM tris-HCl (pH 8.0), 300 mM NaCl, 20 mM imidazole, and 0.02% LMNG for detergent exchange. *Aa*RseP-His_8 was eluted from the resin with a buffer containing 20 mM tris-HCl (pH 8.0), 300 mM NaCl, 250 mM imidazole, and 0.02% LMNG. The elution fraction was then applied to a Superdex 200 10/300 GL size exclusion chromatography column [Cytiva, Tokyo, Japan (GE Healthcare)] to isolate the monodisperse fraction of *Aa*RseP-His_8 with a running buffer containing 10 mM tris-HCl (pH 7.4), 150 mM NaCl, 0.02% LMNG. Last, the purified *Aa*RseP-His_8 was mixed with the 4A9 Fab and was again applied to the size exclusion chromatography to separate the complex fraction with a running buffer containing 10 mM tris-HCl (pH 7.4), 150 mM NaCl, and 0.003% LMNG.

Cryo-EM grid preparation and data collection

For cryo-EM sample preparation, 2.5-μl aliquots of the purified *Aa*RseP(WT):Fab or *Aa*RseP(E18Q):Fab was applied to the glow-discharged holey carbon grids [Quantifoil Cu (for E18Q) or Au (for WT) R0.6/1.0, 200 mesh]. The grid was blotted using Vitrobot Mark IV (Themo Fisher Scientific) operated at 4°C and 100% humidity with a blotting time of 3 s and waiting time of 14 s. The frozen grids were inserted into a Titan Krios (Thermo Fisher Scientific) at 300-kV accelerating voltage equipped with a Gatan Bioquantum K3 direct electron detector (Gatan) and Cs corrector (CEOS, GmbH), which were installed at the Institute for Protein Research, Osaka University. Micrographs were acquired using SerialEM software (https://bio3d.colorado.edu/SerialEM/) at a nominal magnification of 105,000 (calibrated pixel size of 0.675 Å/pixel) with a total dose of 60 $e^{-}/Å^2$ and a defocus range of −0.8 to −1.8 μm. The detailed imaging conditions are described in Table 1.

Cryo-EM data processing

All image processing was performed using cryoSPARC (*57*). (Image processing steps are summarized in Table 1 and figs. S3 and S6.) For the *Aa*RseP(WT):Fab and *Aa*RseP(E18Q):Fab, 11,400 and 15,777 movies were processed, respectively. After Patch Motion Correction (patch-based motion correction) and Patch CTF Estimation (contrast transfer function estimation), the particles were selected using Blob Picker and noise particles were removed by Inspect Particle Picks. The images of picked particles were extracted from micrographs with a box size of 360 pixels, which were reduced to 60 pixels. A total of 1,925,238 and 4,429,254 particles for *Aa*RseP(WT) and *Aa*RseP(E18Q) were subjected to 2D Classification with the following parameters: $K = 200$, number of the final full iterations = 10, and batch size per class = 500. The resulting 2D classes were grouped into four categories: "good" classes where the domain arrangement is well-resolved, "fair" classes that are not clear but presumed to be derived from the target particles, "poor" classes with weak signal, and "noise" classes. Then, Ab-initio Reconstruction was performed using the good 2D classes, which produced one good model and two noise models. Subsequently, three steps of Heterogeneous Refinement were performed against the three ab initio models to select target particles gradually. In the first step, only the good 2D classes were subjected to Heterogeneous Refinement. In the second step, the fair 2D classes were further included in the refinement. In the last step, all 2D classes except for the noise classes were used in the refinement. After three steps of Heterogeneous Refinement, the selected particles were merged and extracted with a box size of 360 pixels, which were reduced to 180 pixels. The extracted particle stacks were subjected to Ab-initio Reconstruction with three classes and Heterogeneous Refinement. At this stage, 454,239 and 1,157,141 particles were selected for *Aa*RseP(WT) and *Aa*RseP(E18Q), respectively, and further subjected to Non-uniform (NU) Refinement (*58*). We also repeated particle picking using Template Picker in cryoSPARC after obtaining the refined 3D models. However, no substantial improvements were observed in the quality of density maps after subsequent re-refinement.

Subsequently, 3D Classification using solvent masks was performed to analyze discrete heterogeneity. For *Aa*RseP(WT), selected particles were subjected to 3D Classification into six classes, in which the PDZ tandem adopts various orientations with respect to the TM region. After three rounds of Heterogeneous Refinement, NU Refinement was performed for each 3D class. The global resolutions were higher than 4 Å for four of six classes (classes #2, #4, #5, and #6). Class #2 with 88,124 particles was refined up to 3.95 Å, in which the TM region was relatively well-resolved as compared to other three high-resolution classes The density map of class #2 was used for the model building of the *Aa*RseP(WT):Fab complex.

For *Aa*RseP(E18Q), the first round of 3D Classification produced four classes (classes #1, #2, #3, and #4) with different orientations of the PDZ tandem with respect to the TM region. After three rounds of Heterogeneous Refinement and NU Refinement, the second round of 3D Classification was performed on class #1 with the best-resolved TM region to further separate it into three 3D classes (classes #1-1, #1-2, and #1-3), which improved the quality of the densities for the TM domain and polypeptide ligand. After three rounds of Heterogeneous Refinement and NU Refinement, class #1-1 with 120,267 particles was refined up to 3.61 Å and its density map was used for the model building of the *Aa*RseP(E18Q):Fab complex.

Model assignment into cryo-EM map and refinement

The crystal structures of 4A9 Fab (PDB code: 8ZAY, this study), *Aa*RseP PDZ-N (PDB code: 6AKQ) (*59*), and *Aa*RseP PDZ-C (PDB code: 3WKM) (*28*) were used for model assignment. For the TM domain, the partial model was extracted from the full-length *Aa*RseP structure predicted by AlphaFold2 (*60*). These partial models were superposed on the density map by rigid-body fitting on ChimeraX (*61*). Subsequently, the models were manually fit into the density map using Coot. The updated models were refined using phenix.refine (*54*) while monitoring the stereochemistry with MolProbity (*55*). The Fourier shell correlation (FSC) curves between the refined model and EM density map are shown in fig. S14. For the analysis of residual densities, $F_o - F_c$ maps were calculated using servalcat (*62*, *63*). Figures for the models and maps were prepared using ChimeraX. Statistics for refinement and validation are summarized in Table 1.

Mass spectrometry

The *Aa*RseP(WT) and *Aa*RseP(E18Q) samples for the MS analysis were prepared by essentially the same procedure as those for the cryo-EM analysis except that the detergent was not exchanged from DDAO to LMNG on the Ni-NTA agarose column. Purified *Aa*RseP(WT) and *Aa*RseP(E18Q) in TBS buffer (pH 8.5) containing 0.032% (w/v) DDAO were buffer exchanged into MS buffer [200 mM

Table 1. Cryo-EM data collection, processing, refinement, and validation. RMSD, root mean square deviation; FSC, Fourier shell correlation; CC, correlation coefficient; PDB, Protein Data Bank.

Datasets	*Aa*RseP(WT):Fab	*Aa*RseP(E18Q):Fab
Data collection and processing		
Magnification	105,000×	105,000×
Voltage (kV)	300	300
Total dose (e^{-}/Å^2)	60	60
Cs (mm)	0.019	0.024
Defocus range (μm)	−0.8 to −1.8	−0.8 to −1.8
Pixel size (Å/pixel)	0.675	0.675
Program	cryoSPARC v4.4.0	cryoSPARC v4.4.0
Symmetry imposed	*C*1	*C*1
No. of initial particle images	1,925,238	4,429,254
No. of final particle images	88,124	120,267
Map resolution (Å)	3.95	3.61
FSC threshold	0.143	0.143
Map resolution range (Å)	3.26–35.5	2.98–52.5
Model refinement		
Program	PHENIX-1.21-5207	PHENIX-1.21-5207
Initial models		
*Aa*RseP TM	AlphaFold2	AlphaFold2
*Aa*RseP PDZ-N	6AKQ	6AKQ
*Aa*RseP PDZ-C	3WKM	3WKM
Fab	8ZAY (this study)	8ZAY (this study)
Real or reciprocal refinement	Real space	Real space
Model-Map CC (mask)	0.82	0.83
Model resolution (Å)	4.2	4.0
FSC threshold	0.5	0.5
Model composition		
Non-hydrogen atoms	6725	6800
Protein residues	862	862
Ligand	1	1
Polypeptide ligand residues		15
Average *B*-factors (Å^2)		
Protein	185.74	147.06
Ligand	224.64	190.72
Polypeptide ligand residues		156.27
RMSD from ideality		
Bond length (Å)	0.003	0.003
Bond angle (°)	0.728	0.677
Validation		
MolProbity score	2.21	2.14
Clashscore	16.84	13.64
Poor rotamers (%)	2.30	2.71
Ramachandran plot		
Favored/allowed/disallowed (%)	96.71/3.29/0	97.06/2.94/0
PDB code	9J82	9J83

ammonium acetate (pH 7.4) containing 0.032% (w/v) DDAO] using a centrifugal buffer exchange device (Micro Bio-Spin 6, Bio-Rad, Hercules, CA, USA).

Mass spectral data were acquired using a SYNAPT G2 HDMS instrument (Waters, Milford) equipped with a nano-ESI source. The samples were deposited in platinum-coated borosilicate capillaries with a 5-μm inner diameter purchased from HUMANIX (Hiroshima, Japan).

The ligand peptides were isolated from acetone-precipitated *Aa*RseP(E18Q) samples using micro-reverse phase chromatography with ZipTip C4 (Millipore, Billerica, MA, USA). The solvent for equilibration and binding was 40% acetonitrile and 0.1% formic acid, and that for elution was 80% acetonitrile and 0.1% formic acid. The eluted peptides were then diluted where the solvent composition was adjusted to 50% acetonitrile and 0.1% formic acid, followed by analysis using nanoESI MS and MS/MS. The ligand peptide was identified from the product ions in combination with database search using ProteinProspector (https://prospector.ucsf.edu/prospector/mshome.htm).

Media, strains, plasmids, and oligonucleotides for biochemical analysis

L medium [Bacto Tryptone (10 g/liter), Bacto Yeast Extract (5 g/liter), and NaCl (5 g/liter); pH adjusted to 7.2 by using NaOH] and M9 medium (without $CaCl_2$) (*64*) supplemented with thiamine (2 μg/ml) and 0.4% glucose were used for the cultivation of *E. coli* cells. Ampicillin (50 μg/ml), chloramphenicol (20 μg/ml), and/or spectinomycin (50 μg/ml) were added for selecting transformants and for growing plasmid-harboring cells.

E. coli K-12 strains, plasmids, and oligonucleotides used in this work are listed in tables S2 to S4, respectively. Construction of the individual strain and plasmids are described in Supplementary Materials.

Construction of the plasmids used in biochemical analyses are described as follows. Plasmids encoding *Aa*RseP with a point mutation were constructed by standard site-directed mutagenesis using appropriate combinations of mutagenic primers and a template plasmid. pSY87 and pSY137 [HA-MBP-CyoE(TM9+Ctail)] was constructed as follows. A 0.07-kb SalI/HindIII–digested polymerase chain reaction (PCR) product (having a predicted TM9 and C-tail periplasmic region of the *cyoE* gene) amplified from an ASKA library plasmid encoding His_6-CyoE (*65*) with primers P1/P2 was inserted into the same site of pYH20 to obtain pSY85. Then, the 1.5-kb EcoRI/HindIII–digested fragment of pSY85 was inserted into pSTD689 and pHM1548, yielding pSY87 and pSY137, respectively. To construct pSY41 (plasmid encoding a substrate for *p*BPA suppression), a 1.5-kb PCR product [HA-MBP-RseA(LY1)148-encoding gene with TAA stop codon] amplified from pYH20 using primers P3/P4 was connected with the EcoRI/HindIII–digested pHM1548 vector fragment using In-Fusion HD cloning kit (Takara Bio).

In vivo cleavage assay of *Aa*RseP and the derivatives

In vivo cleavage of the CyoE model substrate by *Aa*RseP was analyzed essentially as described previously (*26*, *66*). Precultured cells were inoculated into M9 medium supplemented with each of the 20 amino acids (20 μg/ml), thiamine (2 μg/ml), and 0.4% glucose and grown at 30°C for 2 hours with shaking. After addition of 1 mM IPTG and 1 mM adenosine 3′,5′-cyclic monophosphate (cAMP), the cells were further cultured for 2 hours, and then total proteins were precipitated by trichloroacetic acid (TCA) treatment, washed with acetone, and suspended in SDS sample buffer with 2-mercaptoethanol (2ME). The proteins were analyzed by Laemmli SDS–polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting using an Immobilon membrane filter (MilliporeSigma) and appropriate antibodies: the rabbit polyclonal anti-MBP (*56*) or the anti-His antibodies from the Penta-His HRP Conjugate Kit (QIAGEN). Proteins were visualized by Lumino image analyzer LAS-4000 mini (Cytiva) using ECL or ECL Prime Western Blotting Detection Reagents (Cytiva).

Site-directed in vivo photo crosslinking and evaluation of the proteolytic activity of the *p*BPA mutants

Site-directed in vivo photocrosslinking was carried out using essentially the same method as previously described (*27*). *E. coli* cells harboring three plasmids, pEVOL-pBpF, a model substrate-expressing plasmid and a plasmid encoding an *Aa*RseP(E18Q)-His_8 derivative were grown at 30°C for 2 hours in an M9 medium supplemented with thiamine (2 μg/ml), 0.4% glucose, and 0.5 mM *p*BPA (Bachem AG) with shaking. After addition of 1 mM IPTG and 1 mM cAMP, the cells were further cultured for 2 hours. After adding spectinomycin (200 μg/ml) to stop further protein synthesis, a portion of the culture was withdrawn and UV-irradiated for 10 min at 4°C using a B-100 AP UV lamp (365 nm; UVP LLC.). Proteins were TCA-precipitated and dissolved in SDS sample buffer with 2ME. To evaluate the proteolytic function of the *Aa*RseP *p*BPA mutants, the cells were cultured in the same manner except that the derivatives of *Aa*RseP(Glu-18, proteolytically active form) were used. Cultured cells were TCA-sample prepared and analyzed by immunoblotting.

Supplementary Materials

This PDF file includes:

Figs. S1 to S14
Tables S1 to S4
References

REFERENCES AND NOTES

1. L. Sun, X. Li, Y. Shi, Structural biology of intramembrane proteases: Mechanistic insights from rhomboid and S2P to γ-secretase. *Curr. Opin. Struct. Biol.* **37**, 97–107 (2016).
2. M. S. Wolfe, Intramembrane proteolysis. *Chem. Rev.* **109**, 1599–1612 (2009).
3. M. S. Brown, J. Ye, R. B. Rawson, J. L. Goldstein, Regulated intramembrane proteolysis: A control mechanism conserved from bacteria to humans. *Cell* **100**, 391–398 (2000).
4. H. A. Beard, M. Barniol-Xicota, J. Yang, S. H. L. Verhelst, Discovery of cellular roles of intramembrane proteases. *ACS Chem. Biol.* **14**, 2372–2388 (2019).
5. N. Kühnle, V. Dederer, M. K. Lemberg, Intramembrane proteolysis at a glance: From signalling to protein degradation. *J. Cell Sci.* **132**, (2019).
6. D. J. Selkoe, Alzheimer's disease: Genes, proteins, and therapy. *Physiol. Rev.* **81**, 741–766 (2001).
7. L. Chavez-Gutierrez, L. Bammens, I. Benilova, A. Vandersteen, M. Benurwar, M. Borgers, S. Lismont, L. Zhou, S. Van Cleynenbreugel, H. Esselmann, J. Wiltfang, L. Serneels, E. Karran, H. Gijsen, J. Schymkowitz, F. Rousseau, K. Broersen, B. De Strooper, The mechanism of γ-Secretase dysfunction in familial Alzheimer disease. *EMBO J.* **31**, 2261–2274 (2012).
8. M. Spinazzi, B. De Strooper, PARL: The mitochondrial rhomboid protease. *Semin. Cell Dev. Biol.* **60**, 19–28 (2016).
9. L. N. Kinch, K. Ginalski, N. V. Grishin, Site-2 protease regulated intramembrane proteolysis: Sequence homologs suggest an ancient signaling cascade. *Protein Sci.* **15**, 84–93 (2006).
10. R. B. Rawson, N. G. Zelenski, D. Nijhawan, J. Ye, J. Sakai, M. T. Hasan, T. Y. Chang, M. S. Brown, J. L. Goldstein, Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs. *Mol. Cell* **1**, 47–57 (1997).
11. R. B. Rawson, The site-2 protease. *Biochim. Biophys. Acta* **1828**, 2801–2807 (2013).
12. N. Caengprasath, T. Theerapanon, T. Porntaveetus, V. Shotelersuk, MBTPS2, a membrane bound protease, underlying several distinct skin and bone disorders. *J. Transl. Med.* **19**, 114 (2021).
13. J. S. Schneider, M. S. Glickman, Function of site-2 proteases in bacteria and bacterial pathogens. *Biochim. Biophys. Acta* **1828**, 2808–2814 (2013).

14. L. Kroos, Y. Akiyama, Biochemical and structural insights into intramembrane metalloprotease mechanisms. *Biochim. Biophys. Acta* **1828**, 2873–2885 (2013).
15. S. S. Kristensen, D. B. Diep, M. Kjos, G. Mathiesen, The role of site-2-proteases in bacteria: A review on physiology, virulence, and therapeutic potential. *Microlife* **4**, uqad025 (2023).
16. H. Makinoshima, M. S. Glickman, Regulation of *Mycobacterium tuberculosis* cell envelope composition and virulence by intramembrane proteolysis. *Nature* **436**, 406–409 (2005).
17. S. Almagro-Moreno, T. K. Kim, K. Skorupski, R. K. Taylor, Proteolysis of virulence regulator ToxR is associated with entry of *Vibrio cholerae* into a dormant state. *PLOS Genet.* **11**, e1005145 (2015).
18. K. Kanehara, K. Ito, Y. Akiyama, YaeL (EcfE) activates the σ^E pathway of stress response through a site-2 cleavage of anti-σ^E, RseA. *Genes Dev.* **16**, 2147–2155 (2002).
19. S. E. Ades, Regulation by destruction: Design of the σ^E envelope stress response. *Curr. Opin. Microbiol.* **11**, 535–540 (2008).
20. B. M. Alba, J. A. Leeds, C. Onufryk, C. Z. Lu, C. A. Gross, DegS and YaeL participate sequentially in the cleavage of RseA to activate the σ^E-dependent extracytoplasmic stress response. *Genes Dev.* **16**, 2156–2168 (2002).
21. T. Yokoyama, T. Niinae, K. Tsumagari, K. Imami, Y. Ishihama, Y. Hizukuri, Y. Akiyama, The *Escherichia coli* S2P intramembrane protease RseP regulates ferric citrate uptake by cleaving the sigma factor regulator FecR. *J. Biol. Chem.* **296**, 100673 (2021).
22. A. Saito, Y. Hizukuri, E. Matsuo, S. Chiba, H. Mori, O. Nishimura, K. Ito, Y. Akiyama, Post-liberation cleavage of signal peptides is catalyzed by the site-2 protease (S2P) in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13740–13745 (2011).
23. T. Yokoyama, Y. Yamagata, S. Honna, S. Mizuno, S. Katagiri, R. Oi, T. Nogi, Y. Hizukuri, Y. Akiyama, S2P intramembrane protease RseP degrades small membrane proteins and suppresses the cytotoxicity of intrinsic toxin HokB. *mBio* **14**, e0108623 (2023).
24. Y. Imaizumi, K. Takanuki, T. Miyake, M. Takemoto, K. Hirata, M. Hirose, R. Oi, T. Kobayashi, K. Miyoshi, R. Aruga, T. Yokoyama, S. Katagiri, H. Matsuura, K. Iwasaki, T. Kato, M. K. Kaneko, Y. Kato, M. Tajiri, S. Akashi, O. Nureki, Y. Hizukuri, Y. Akiyama, T. Nogi, Mechanistic insights into intramembrane proteolysis by *E. coli* site-2 protease homolog RseP. *Sci. Adv.* **8**, eabp9011 (2022).
25. F. X. Gomis-Ruth, T. O. Botelho, W. Bode, A standard orientation for metallopeptidases. *Biochim. Biophys. Acta* **1824**, 157–163 (2012).
26. K. Akiyama, S. Mizuno, Y. Hizukuri, H. Mori, T. Nogi, Y. Akiyama, Roles of the membrane-reentrant β-hairpin-like loop of RseP protease in selective substrate cleavage. *eLife* **4**, e08928 (2015).
27. T. Miyake, Y. Hizukuri, Y. Akiyama, Involvement of a membrane-bound amphiphilic helix in substrate discrimination and binding by an *Escherichia coli* S2P peptidase RseP. *Front. Microbiol.* **11**, 607381 (2020).
28. Y. Hizukuri, T. Oda, S. Tabata, K. Tamura-Kawakami, R. Oi, M. Sato, J. Takagi, Y. Akiyama, T. Nogi, A structure-based model of substrate discrimination by a noncanonical PDZ tandem in the intramembrane-cleaving protease RseP. *Structure* **22**, 326–336 (2014).
29. R. Tamura-Sakaguchi, R. Aruga, M. Hirose, T. Ekimoto, T. Miyake, Y. Hizukuri, R. Oi, M. K. Kaneko, Y. Kato, Y. Akiyama, M. Ikeguchi, K. Iwasaki, T. Nogi, Moving toward generalizable NZ-1 labeling for 3D structure determination with optimized epitope-tag insertion. *Acta Crystallogr. D Struct. Biol.* **77**, 645–662 (2021).
30. G. Deckert, P. V. Warren, T. Gaasterland, W. G. Young, A. L. Lenox, D. E. Graham, R. Overbeek, M. A. Snead, M. Keller, M. Aujay, R. Huber, R. A. Feldman, J. M. Short, G. J. Olsen, R. V. Swanson, The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**, 353–358 (1998).
31. K. Saiki, T. Mogi, Y. Anraku, Heme O biosynthesis in *Escherichia coli*: The *cyoE* gene in the cytochrome *bo* operon encodes a protoheme IX farnesyltransferase. *Biochem. Biophys. Res. Commun.* **189**, 1491–1497 (1992).
32. M. Varadi, D. Bertoni, P. Magana, U. Paramval, I. Pidruchna, M. Radhakrishnan, M. Tsenkov, S. Nair, M. Mirdita, J. Yeo, O. Kovalevskiy, K. Tunyasuvunakool, A. Laydon, A. Zidek, H. Tomlinson, D. Hariharan, J. Abrahamson, T. Green, J. Jumper, E. Birney, M. Steinegger, D. Hassabis, S. Velankar, AlphaFold protein structure database in 2024: Providing strucnture coverage for over 214 million protein sequences. *Nucleic Acids Res.* **52**, D368–D375 (2024).
33. E. D. Rivett, L. Heo, M. Feig, E. L. Hegg, Biosynthesis and trafficking of heme o and heme a: New structural insights and their implications for reaction mechanisms and prenylated heme transfer. *Crit. Rev. Biochem. Mol. Biol.* **56**, 640–668 (2021).
34. Y. Hizukuri, Y. Akiyama, PDZ domains of RseP are not essential for sequential cleavage of RseA or stress-induced σ^E activation *in vivo*. *Mol. Microbiol.* **86**, 1232–1245 (2012).
35. L. Feng, H. Yan, Z. Wu, N. Yan, Z. Wang, P. D. Jeffrey, Y. Shi, Structure of a site-2 protease family intramembrane metalloprotease. *Science* **318**, 1608–1612 (2007).
36. M. A. Orlando, H. J. T. Pouillon, S. Mandal, L. Kroos, B. J. Orlando, Substrate engagement by the intramembrane metalloprotease SpoIVFB. *Nat. Commun.* **15**, 8276 (2024).
37. D. M. Bolduc, D. R. Montagna, Y. Gu, D. J. Selkoe, M. S. Wolfe, Nicastrin functions to sterically hinder γ-secretase-substrate interactions driven by substrate transmembrane domain. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E509–E518 (2016).
38. G. Yang, R. Zhou, Q. Zhou, X. Guo, C. Yan, M. Ke, J. Lei, Y. Shi, Structural basis of Notch recognition by human γ-secretase. *Nature* **565**, 192–197 (2019).
39. N. Sakaguchi, T. Kimura, S. Matsushita, S. Fujimura, J. Shibata, M. Araki, T. Sakamoto, C. Minoda, K. Kuwahara, Generation of high-affinity antibody against T cell-dependent antigen in the *Ganp* gene-transgenic mouse. *J. Immunol.* **174**, 4485–4494 (2005).
40. M. Kaneko, Y. Kato, H. Horiuchi, M. Osawa, Molecular characterization of a human monoclonal antibody to B antigen in ABO blood type. *Immunol. Lett.* **86**, 45–51 (2003).
41. M. K. Kaneko, A. Kunita, S. Abe, Y. Tsujimoto, M. Fukayama, K. Goto, Y. Sawa, Y. Nishioka, Y. Kato, Chimeric anti-podoplanin antibody suppresses tumor metastasis through neutralization and antibody-dependent cellular cytotoxicity. *Cancer Sci.* **103**, 1913–1919 (2012).
42. D. J. Leahy, C. E. Dann III, P. Longo, B. Perman, K. X. Ramyar, A mammalian expression vector for expression and purification of secreted proteins for structural studies. *Protein Expr. Purif.* **20**, 500–506 (2000).
43. Y. Adachi, M. K. Kaneko, Y. Kato, T. Nogi, Recombinant production of antibody antigen-binding fragments with an N-terminal human growth hormone tag in mammalian cells. *Protein Expr. Purif.* **208-209**, 106289 (2023).
44. Y. Fujii, M. Kaneko, M. Neyazaki, T. Nogi, Y. Kato, J. Takagi, PA tag: A versatile protein tagging system using a super high affinity antibody against a dodecapeptide derived from human podoplanin. *Protein Expr. Purif.* **95**, 240–247 (2014).
45. K. Hirata, Y. Kawano, G. Ueno, K. Hashimoto, H. Murakami, K. Kasegawa, T. Hikima, T. Kumasaka, M. Yamamoto, Achievement of protein micro-crystallography at SPring-8 beamline BL32XU. *J. Phys. Conf. Ser.* **425**, 012002 (2013).
46. K. Hirata, K. Yamashita, G. Ueno, Y. Kawano, K. Hasegawa, T. Kumasaka, M. Yamamoto, *ZOO*: An automatic data-collection system for high-throughput structure analysis in protein microcrystallography. *Acta Crystallogr. D Struct. Biol.* **75**, 138–150 (2019).
47. K. Yamashita, K. Hirata, M. Yamamoto, *KAMO*: Towards automated data processing for microcrystals. *Acta Crystallogr. D Struct. Biol.* **74**, 441–449 (2018).
48. J. Foadi, P. Aller, Y. Alguel, A. Cameron, D. Axford, R. L. Owen, W. Armour, D. G. Waterman, S. Iwata, G. Evans, Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1617–1632 (2013).
49. W. Kabsch, Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795–800 (1993).
50. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
51. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
52. A. Perrakis, M. Harkiolaki, K. S. Wilson, V. S. Lamzin, *ARP/wARP* and molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 1445–1450 (2001).
53. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
54. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, Towards automated crystallographic structure refinement with *phenix.refine*. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
55. V. B. Chen, W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *MolProbity*: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
56. Y. Akiyama, K. Kanehara, K. Ito, RseP (YaeL), an *Escherichia coli* RIP protease, cleaves transmembrane sequences. *EMBO J.* **23**, 4434–4442 (2004).
57. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
58. A. Punjani, H. Zhang, D. J. Fleet, Non-uniform refinement: Adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
59. R. Tamura, R. Oi, S. Akashi, M. K. Kaneko, Y. Kato, T. Nogi, Application of the NZ-1 Fab as a crystallization chaperone for PA tag-inserted target proteins. *Protein Sci.* **28**, 823–836 (2019).
60. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
61. E. C. Meng, T. D. Goddard, E. F. Pettersen, G. S. Couch, Z. J. Pearson, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci.* **32**, e4792 (2023).
62. K. Yamashita, C. M. Palmer, T. Burnley, G. N. Murshudov, Cryo-EM single-particle structure refinement and map calculation using *Servalcat*. *Acta Crystallogr. D Struct. Biol.* **77** (Pt. 10), 1282–1291 (2021).
63. K. Yamashita, M. Wojdyr, F. Long, R. A. Nicholls, G. N. Murshudov, *GEMMI* and *Servalcat* restrain *REFMAC5*. *Acta Crystallogr. D Struct. Biol.* **79**, 368–373 (2023).

64. J. H. Miller, *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory Press, 1972).
65. M. Kitagawa, T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga, H. Mori, Complete set of ORF clones of *Escherichia coli* ASKA library (*A* complete *S*et of *E. coli K*-12 ORF *A*rchive): Unique resources for biological research. *DNA Res.* **12**, 291–299 (2005).
66. Y. Hizukuri, K. Akiyama, Y. Akiyama, Biochemical characterization of function and structure of RseP, an *Escherichia coli* S2P protease. *Methods Enzymol.* **584**, 1–33 (2017).
67. T. J. Silhavy, M. L. Berman, L. W. Enquist, *Experiments with Gene Fusions* (Cold Spring Harbor Laboratory Press, 1984).
68. Y. Akiyama, T. Ogura, K. Ito, Involvement of FtsH in protein assembly into and through the membrane. I. Mutations that reduce retention efficiency of a cytoplasmic reporter. *J. Biol. Chem.* **269**, 5218–5224 (1994).
69. H. Mori, K. Ito, Different modes of SecY-SecA interactions revealed by site-directed *in vivo* photo-cross-linking. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 16159–16164 (2006).
70. K. Kanehara, K. Ito, Y. Akiyama, YaeL proteolysis of RseA is controlled by the PDZ domain of YaeL and a Gln-rich region of RseA. *EMBO J.* **22**, 6389–6398 (2003).
71. R. Miyazaki, M. Ai, N. Tanaka, T. Suzuki, N. Dhomae, T. Tsukazaki, Y. Akiyama, H. Mori, Inner membrane YfgM-PpiD heterodimer acts as a functional unit that associates with the SecY/E/G translocon and promotes protein translocation. *J. Biol. Chem.* **298**, 102572 (2022).
72. T. S. Young, I. Ahmad, J. A. Yin, P. G. Schultz, An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol.* **395**, 361–374 (2010).

Acknowledgments: We are grateful to the staff of SPring-8 BL32XU (Hyogo, Japan) for the support in x-ray diffraction data collection and to the staff of the Structural Biology Research Center in the High Energy Accelerator Research Organization/KEK (Tsukuba, Japan) for the support in preliminary cryo-EM data collection. We also thank S. Thompson for the critical reading and editing of the manuscript; T. Nishizawa and Y. Lee for the valuable advice on the cryo-EM data processing; J. Takagi for his support in the early stage of this work; and S. Tabata, R. Tamura-Sakaguchi, and T. Hirose for the contribution to preliminary antibody-assisted cryo-EM analysis. **Funding:** This research is partially supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under grant numbers JP26291016, JP19H03170, and JP22H02561/23K23825 (to T.N.); JP22K06142 (to Y.H.); JP22H02571/23K23835 (to Y.Aki.); and JP21K19236 and JP24K09343 (to S.A.); by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) under JP19H05774 (to S.A.); by the Japan Agency for Medical Research and Development (AMED) under grant number JP24am0521010 (to T.N. and Y.K.); by the Research Support Project for Life Science and Drug Discovery [Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)] from AMED under grant number JP23ama121001 (to T.K.) and JP23ama121008 (to Y.K.); by the Institute for Fermentation, Osaka under grant number G-2022-2-108 (to Y.H.); and by the Daiichi Sankyo Foundation of Life Science (to S.A.). This work was also performed in part under the Cooperative Research Program (Joint Usage/Research Center program) of the Institute for Life and Medical Sciences, Kyoto University and under the Collaborative Research Program of the Institute for Protein Research, Osaka University, CR-23-02. **Author contributions:** T.N. conceived the project. T.N. and T.K. supervised the cryo-EM analysis. K.A. and R.A. prepared the cryo-EM samples. M.H., K.A., R.A., and T.T. performed the cryo-EM data collection. K.A., R.A., and T.N. processed the cryo-EM data. Y.Ada., Y.T., M.K.K., and Y.K. contributed to preparation of the Fab for structural analysis. Y.Ada. and T.N. performed the crystallographic analysis of the Fab. M.T. and S.A. performed the MS analysis. Y.Aki. and Y.H. supervised the biochemical analysis. Y.S. and Y.H. performed the cleavage assay and cross-linking analysis. Y.Aki. and Y.H. analyzed the biochemical data. T.N., Y.H., M.T., and K.A. wrote the original draft of the manuscript. M.H., T.K., Y.K., S.A., Y.Aki., and Y.H. contributed to paper preparation. TN compiled the paper. Conceptualization: T.N. Methodology: T.N., T.K., S.A., Y.K., Y.H., and Y.Aki. Investigation: T.N., K.A., M.H., R.A., T.T., Y.Ada., Y.T., M.K.K., M.T., Y.H., and Y.S. Funding acquisition: T.N., Y.H., Y.Aki., S.A., Y.K., and T.K. Supervision: T.N., T.K., S.A., Y.K., Y.H., and Y.Aki. Writing—original draft: T.N., Y.H., M.T., and K.A. Writing—review and editing: T.N., M.H., T.K., Y.K., S.A., Y.Aki., and Y.H. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The cryo-EM raw images, cryo-EM density maps, and atomic coordinates have been deposited in the Electron Microscopy Public Image Archive (EMPIAR), the Electron Microscopy Data Bank (EMDB), and the Protein Data Bank (PDB), with accession codes EMPIAR-12301, EMD-61213, and 9J82 for *Aa*RseP(WT) in complex with 4A9 Fab and EMPIAR-12302, EMD-61214, and 9J83 for *Aa*RseP(E18Q) in complex with 4A9 Fab, respectively. The atomic coordinates of 4A9 Fab have been deposited in PDB with accession code with 8ZAY. All other data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 24 October 2024
Accepted 23 January 2025
Published 26 February 2025
10.1126/sciadv.adu0925