Structural basis for intramembrane proteolysis by rhomboid serine proteases

Adam Ben-Shem*†, Deborah Fass‡, and Eitan Bibi*†

Departments of *Biological Chemistry and ‡Structural Biology, The Weizmann Institute of Science, Rehovot 76100, Israel

Communicated by Arthur Karlin, Columbia University College of Physicians and Surgeons, New York, NY, November 7, 2006
(received for review October 15, 2006)

Intramembrane proteases catalyze peptide bond cleavage of integral membrane protein substrates. This activity is crucial for many biological and pathological processes. Rhomboids are evolutionarily widespread intramembrane serine proteases. Here, we present the 2.3-Å-resolution crystal structure of a rhomboid from *Escherichia coli*. The enzyme has six transmembrane helices, five of which surround a short TM4, which starts deep within the membrane at the catalytic serine residue. Thus, the catalytic serine is in an externally exposed cavity, which provides a hydrophilic environment for proteolysis. Our results reveal a mechanism to enable water-dependent catalysis at the depth of the hydrophobic milieu of the membrane and suggest how substrates gain access to the sequestered rhomboid active site.

GlpG

Intramembrane proteolysis is catalyzed by polytopic integral membrane enzymes (1). Through the controlled release of membrane-anchored dormant transcription factors or other effectors, these proteases play a pivotal role in biological and pathological processes (2–4). Several families of intramembrane proteases have been characterized, including metallo-intramembrane proteases, which regulate important cellular pathways such as sterol metabolism (5) and the unfolded protein response (6), and aspartyl intramembrane proteases such as presenilin, the catalytic subunit of γ-secretase, which generates the β-amyloid peptides involved in Alzheimer disease (3). Another family, which is the subject of our study, is the intramembrane serine proteases, termed rhomboids. The phenomenon of intramembrane proteolysis represents challenging mechanistic questions, such as how hydrolytic reactions can occur in a lipophilic environment (2).

Rhomboids are evolutionarily widespread intramembrane serine proteases (7, 8) that act as regulators of intercellular signaling (9, 10), parasite invasion (11), quorum sensing (12), mitochondrial morphology and dynamics (13, 14), and apoptosis (15). Eukaryotic and prokaryotic rhomboids, including GlpG from *Escherichia coli* (16, 17), recognize and cleave similar integral membrane protein substrates, suggesting a shared mechanism and specificity determinants (18–20). Multiple sequence alignment has identified a six-transmembrane helix (TM) core shared by all rhomboids, containing the conserved and catalytically important residues of the rhomboid family (7) (Fig. 1*A* and *B*). Other regions, including the hydrophilic N and C termini, are highly variable, sometimes absent (7, 18), and functionally dispensable (9).

Results and Discussion

To reveal the molecular organization of rhomboids and elucidate the mechanism of intramembrane proteolysis, we crystallized the rhomboid homologue from *E. coli*, GlpG. In the course of our crystallization trials, we found that trypsinization (Fig. 1*A*), which removed a cytoplasmic region at the N terminus predicted to be unstructured and left intact the six-TM core of GlpG, facilitated crystallization. Crystals of the rhomboid core, which belong to space group $P2_1$, were grown (Table 1), and phases were obtained by single isomorphous replacement with anomalous scattering (SIRAS) from an ethylmercury chloride derivative (Fig. 1*C*).

Overall Structure of the Rhomboid Core. The crystal asymmetric unit comprises two molecules of the protein (denoted A and B in the atomic coordinate file) related by nonphysiological symmetry. There are several differences between the two molecules (Cα rmsd of 0.90 Å), specifically in the orientation of TM5, the association with lipids, the local geometry of the active-site serine, and the fact that residues 244–249 are disordered in molecule A. The model of molecule B (residues 93–271) contains all of the amino acids of the rhomboid core (Fig. 1*D*). In addition to bound detergent, the structure contains a tightly bound lipid molecule (see below). The preservation of bound lipid was perhaps due to the mildness of the purification procedure. The height of GlpG normal to the membrane is ≈40 Å, in agreement with the measured average thickness of the *E. coli* membrane, including the polar surfaces (21). The N and C termini of GlpG are exposed to the cytoplasm, and the TMs are connected by short hydrophilic loops, with the exception of the region between TM1 and TM2 (loop 1) (Fig. 1 *B* and *D*).

Description of the Active-Site Cavity. The rhomboid active site is located deep within the molecule in a region of unusual architecture. The catalytic serine residue (Ser-201) (8) flanks the membrane-embedded N-terminal end of TM4 (Fig. 2*A*). The catalytic histidine residue (His-254) (8) is also located inside the membrane, on TM6 (Fig. 2*A*). Highly conserved glycines in TM6 (Gly-257 and Gly-261) and TM4 (Gly-202) (18) (Fig. 1*A*) allow a particularly close packing of these two helices and approach of the catalytic residues within hydrogen bond distance of one another. Unlike the other TMs, TM4 does not span the periplasmic leaflet of the membrane. Loop 3, originating from the periplasmic end of TM3, inserts into the membrane as an extended strand (Fig. 2*A*). Abruptly, just at or adjacent to Ser-201, the polypeptide assumes a helical conformation, and the resulting TM4 traverses the cytoplasmic leaflet of the membrane. The architecture preceding TM4 is promoted by helix-breaking glycine residues (Gly-198 and Gly-199) (Fig. 1*A*). This configuration facilitates the formation above TM4 of a periplasmically exposed, water-filled cavity lined with numerous polar residues, with Ser-201 protruding into this cavity as if sitting on

Author contributions: A.B.-S. and E.B. designed research; A.B.-S. performed research; A.B.-S. and D.F. analyzed data; and A.B.-S., D.F., and E.B. wrote the paper.

The authors declare no conflict of interest.

Abbreviation: TM, transmembrane helix.

Data deposition: The coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2IRV).

See Commentary on page 401.

†To whom correspondence may be addressed. E-mail: adam.ben-shem@weizmann.ac.il or e.bibi@weizmann.ac.il.

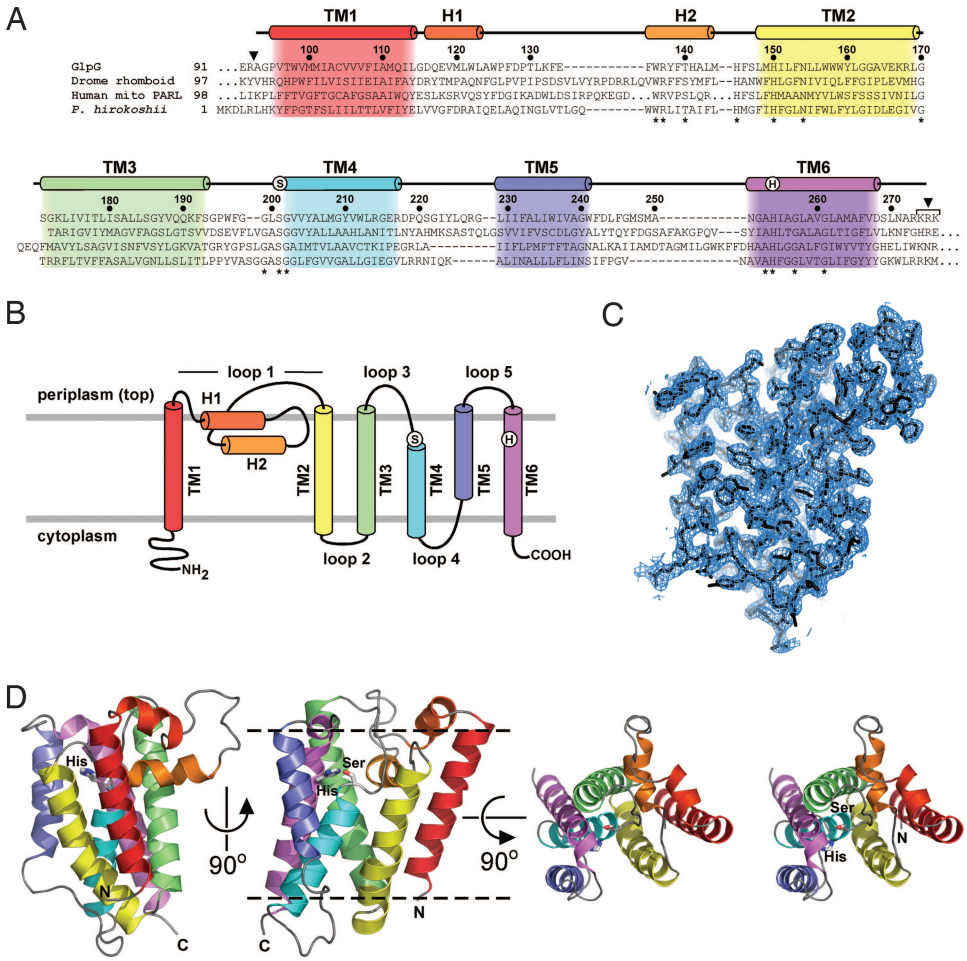

Fig. 1. Organization of GlpG functional and secondary structural elements. (*A*) Linear map of GlpG and comparison to other rhomboid proteases. Above the GlpG amino acid sequence, helices are indicated as cylinders, and the positions of the catalytic residues are marked with asterisks. Although a precise sequence alignment with rhomboid proteases from other kingdoms (Drome rhomboid = *Drosophila melanogaster* rhomboid, gi:201416977; Human mito PARL = *Homo sapiens* mitochondrial presenilins-associated rhomboid-like protein, gi:62511133; *P. hirokoshii* = rhomboid-like protein from the thermophilic archeon *Pyrococcus hirokoshii*, gi:14591283) is not possible because of low sequence similarity, predicted transmembrane regions and shared sequence motifs were used to align functional and structural regions. The starting residue numbers of the indicated sequence fragments within their full-length proteins are given. Trypsin cleavage sites in the preparation of the crystallized GlpG are indicated by arrowheads. (*B*) Schematic representation of GlpG helix orientation with respect to the membrane. (*C*) Slab of an experimental SIRAS map contoured at 1σ over a stick model of molecule B from the GlpG crystals. (*D*) Ribbon diagram of the GlpG crystal structure in three orientations with helices color-coded as in *A* and *B*; the top view is shown in stereo.

a pedestal (Fig. 2*A*). Notably, as observed in many soluble hydrolases (22), the backbone dihedral angles of the active-site serine in molecule B are in a typically disallowed region of Ramachandran space. The crystal structure demonstrates that the GlpG catalytic residues are indeed buried in the membrane, ≈11 Å from the periplasmic surface, resolving the long-standing question as to the location of the active sites of at least this class of intramembrane proteases.

Definition of the Active-Site Cavity. The hydrophilic cavity containing Ser-201 is sealed off from the bilayer on three sides (Figs. 1*D* and 2*A*). One wall of the cavity is made up of TM5, TM6, and the periplasmic loop connecting them. The opposite wall is formed from the periplasmic end of TM2. The back of the cavity is made up of the periplasmic end of TM3 and loop 3. Also contributing to the back wall of the cavity is a short non-TM helix in loop 1 (H2) that lies inside and parallel to the membrane at approximately the depth of the active site (Fig. 2*B*). H2 fills the space between TM1 and TM3, which diverge dramatically toward the periplasmic side of the membrane and therefore do not themselves close off the back side of the hydrophilic active-site cavity. The only remaining lateral entrance into the cavity containing Ser-201 is between TM2 and TM5, which interact via hydrophobic side chains within the cytoplasmic membrane leaflet but diverge toward the periplasm (Fig. 2*A*). Exposure of the GlpG active site to the lipid environment from this direction is evident in the crystal structure by a phospholipid bound between TM2 and TM5 in molecule B with its phosphate group in hydrogen bond distance of His-254, His-150, Asn-154, a water molecule, and perhaps Ser-201 (Fig. 2*A*). The position of the lipid phosphate group is analogous to that of a sulfate or acetate ion observed in crystal structures of water-soluble serine proteases (23, 24).

Architecture of Loop 1. The two non-TM helices, H1 and H2, form a region of interesting architecture that protrudes from the main body of GlpG into the periplasmic leaflet of the lipid bilayer (Fig. 2*B*). Although H2 is buried with its hydrophobic residues facing the membrane interior, it also contains a number of polar residues. These residues are found to make internal hydrogen bonds to the backbone of surrounding regions of the protein. Notably, Arg-137 at the N terminus of H2 forms part of the

Table 1. Summary of crystallographic data

	Data collection	
Space group	$P2_1$ native	EtHgCl
Wavelength, Å	1.541	1.006
Unit cell parameters	49.32 × 69.53 × 67.58 Å, $\beta = 101.41°$	49.82 × 69.35 × 67.58 Å, $\beta = 101.03°$
Resolution, Å*	50–2.30 (2.38–2.30)	50–2.25 (2.33–2.25)
Completeness, %	98.4 (97.2)	99.4 (95.0)
Redundancy†	3.0 (2.8)	7.1 (5.3)
R_{sym}‡	0.081 (0.411)	0.081 (0.247)
I/σ§	9.8 (1.8)	10.8 (4.9)
Overall figure of merit¶		0.47 (0.72)
	Refinement statistics	
Total reflections/test set	17,028/1,175	
R_{work}/R_{free}¶	0.220/0.263	
rms deviations from ideality		
Bonds, Å	0.008	
Angles, °	1.301	
No. of atoms		
Protein	2,825	
Detergent/lipid	290	
Water	81	
Average B-factors		
Molecule A	32	
Molecule B	22	
Detergent/lipid	45	
Water	33	

*Values in parentheses are for the highest-resolution shell.

†Redundancy was calculated by applying Friedel's law.

‡$R_{sym} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity for i observations.

§The value for the high-resolution shell is $\langle I\rangle/\langle\sigma\rangle$ for this shell.

¶Figures of merit for data from 50 to 2.3 Å resolution obtained from SOLVE. In parentheses is the figure of merit after density modification using RESOLVE.

‖R_{work}, $R_{free} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. A set of reflections (6.9%) was excluded from refinement and used to calculate R_{free}.

tryptophan-arginine sequence found in many rhomboid proteins (7, 8) (Fig. 1*A*) and is shown here to play a major role in C-capping of H1 and in forming the hydrogen and electrostatic bonding network that stabilizes the architecture of loop 1 (Fig. 2*B*). However, the H2 helix itself, lying within and parallel to the membrane, presents a particular problem for end-capping, because this role is generally performed by polar or charged residues, or water. We find that H2 is capped at the N terminus by the backbone carbonyls of two consecutive β-hairpin turns immediately preceding the helix and at the C terminus by the backbone NHs of the glycine-rich strand above Ser-201 and by a buried water molecule. In addition to the structural role of H1/H2, it is tempting to speculate that this region is functionally involved in protein–protein interactions in view of its protrusion into the lipid bilayer.

Comparison Between Intramembrane and Soluble Serine Proteases. The rhomboid core crystal structure allows a detailed comparison between the active sites of water-soluble versus intramembrane serine proteases. One classic feature of water-soluble serine proteases is an asparagine-histidine-serine catalytic triad (25, 26). The relative positions of the GlpG active-site Ser-201 and His-254 agree well with those observed in serine proteases such as trypsin (23), with a hydrogen bond distance of ≈3.1 Å between the serine hydroxyl and the histidine imidazole ring (Fig. 3 *A* and *B*). However, as predicted (17, 20), rhomboids do not have an active-site aspartate, the third player of the triad. Instead of a hydrogen bond to an aspartate, His-254 in GlpG makes a water-mediated hydrogen bond to Asn-251 and to the backbone carbonyl of Trp-236 (Fig. 3*A*). The second characteristic feature of soluble serine proteases is an oxyanion hole, which stabilizes the tetrahedral reaction intermediate (25). GlpG contains an apparent oxyanion hole in an analogous position to soluble serine proteases. The hydrogen bond donors that contribute to the GlpG oxyanion hole are the Ser-201 backbone NH group and the side chain of the conserved Asn-154 (Fig. 3*A*). The side chain of the conserved residue His-150 and the backbone NH of Leu-200 may also contribute to stabilization of a negative charge in this region. Thus, a pronounced oxyanion hole and a hydrogen bond network for the active-site histidine in GlpG may facilitate a dyad-catalyzed hydrolytic reaction, as observed in several soluble hydrolases (27, 28).

Putative Modes of Substrate Accessibility. The crystal structure suggests possible mechanisms by which rhomboid substrates gain access to the protease active site. Known rhomboid substrates are single-span transmembrane proteins that are cleaved near the TM periplasmic edge (29). It is clear from our rhomboid core crystal structure that a substrate transmembrane helix approaching laterally cannot gain direct access to the enzyme active site. A potential substrate accessibility crevice lies between TM2 and TM5 in the periplasmic leaflet of the membrane (Figs. 4 and 2*A*). Alternatively, speculative large conformational changes in loop 1 may provide another crevice between TM1 and TM3. However, in both cases the active site is located deep inside the protein, far from the edges of these crevices. Two plausible solutions to this problem are rearrangement of the enzyme TM helices and/or disruption of the substrate helical conformation.

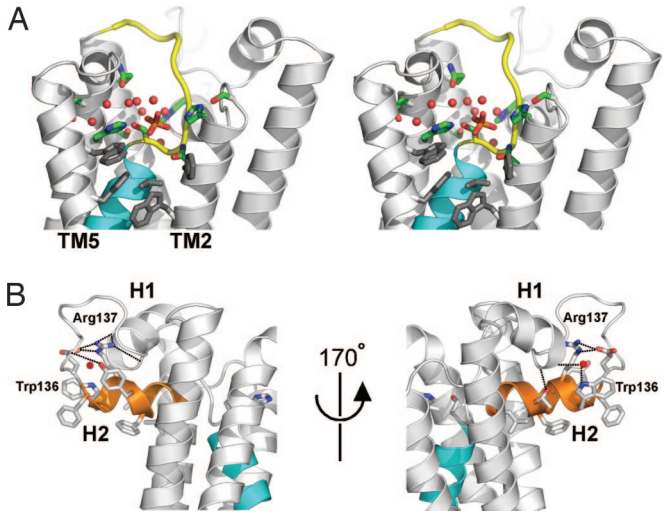

Fig. 2. Distinctive structural features of GlpG. (*A*) Stereoview into the active-site cavity of GlpG. The side chains of polar residues lining the cavity are shown in color. The oxygen atoms of water molecules are shown as red balls. For clarity, only the phosphate group (red/orange tetrahedron) of the bound phospholipid is shown. A cluster of hydrophobic residues lining the lip of the cavity and closing the gap between TM2 and TM5 is shown in dark gray. The loop between TM3 and TM4 that leads down into the core of the enzyme and the active-site Ser-201 is in yellow, and TM4 is in cyan. (*B*) The membrane-embedded helix H2 lies perpendicular to the TM helices within the membrane. Some hydrogen bonding interactions stabilizing the local geometry of helices H1 and H2 are indicated as dotted lines. Hydrophobic residues from H2 facing the lipid interior are also shown.

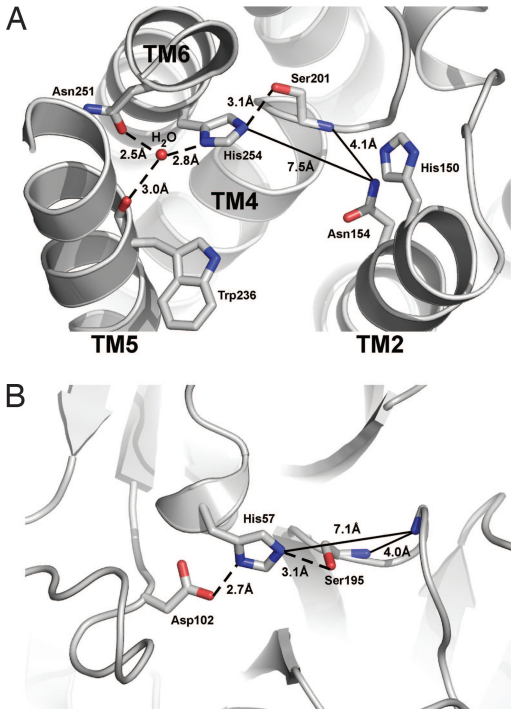

Fig. 3. Comparison of the active sites of GlpG and trypsin. The GlpG active site viewed from the membrane (*A*) compared with that of the soluble serine protease trypsin (*B*) (PDB entry 1TLD) (23). Dashed lines indicate hydrogen bond distances. Solid lines indicate distances that demonstrate the resemblance between the positions of the oxyanion holes in the two structures.

Interestingly, the importance of helix breaking residues in the periplasmic quarter of rhomboid substrates has been recognized (30). These residues could facilitate entry of the substrate into the active-site cavity. Furthermore, helix breaking is essential for exposing the substrate scissile bond. The active-site region of GlpG is filled with polar residues and exposed backbone hydrogen bond donors and acceptors (Fig. 2*A*), which could interact with the rhomboid substrate and stabilize a nonhelical conformation in the vicinity of the cleavage site (Fig. 4).

We have crystallized the rhomboid core of GlpG and solved its structure at high resolution. On the basis of this structure, we propose several key principles that underlie the mechanism of integral membrane protein proteolysis by rhomboids. The crystal structure clearly demonstrates that a sequestered active-site serine-histidine dyad is located well inside the membrane, and that a cavity exposed to the periplasm allows hydrolytic water accessibility to the active site. In addition, the cavity may provide the hydrophilic environment necessary for stabilizing a nonhelical conformation in the rhomboid substrate with a scissile bond exposed to the active-site serine, as well as stabilizing the charged species along the reaction coordinate. The mode of entry of a TM substrate to the cavity and the active site can also be predicted from the crystal structure. The specialized structural

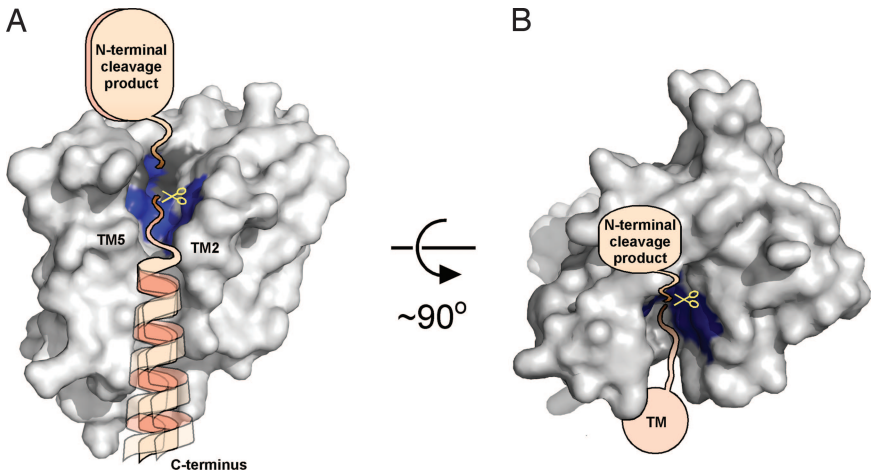

Fig. 4. A model for interaction of substrates with GlpG. A surface representation of GlpG is shown with the polar residues in the active-site cavity in blue. A schematic representation of a substrate TM (straight or tilted) and its external domain is shown. The top quarter of the substrate TM is portrayed in a nonhelical conformation that penetrates into the active-site cavity in between TM2 and TM5. (*A*) View from the membrane. (*B*) View from the periplasm.

elements observed in GlpG appear to be found, according to sequence inspection, in rhomboid-like proteases from all three kingdoms of life (Fig. 1*A*). The architectural principles that allow water-mediated proteolysis within the lipid bilayer interior may apply to other intramembrane proteases as well, in view of the likelihood that such enzymes must stabilize charged intermediates within the membrane and enable scissile bond exposure within their helical substrates.

Materials and Methods

Cloning of *glpG*. The *glpG* gene was amplified from wild-type *E. coli* K12 by colony PCR using primers sense 5′-TATT*CCATG-G*TGATGATTACCTCTTTTGC and antisense 5′-TATT*CTC-GAG*TTATTTTCGTTTTCGCGCATTG (NcoI and XhoI restriction sites are italicized). *GlpG* was cloned into plasmid PET-28a (Novagen) with a His_6-tag-encoding sequence at the 3′ end of the gene. The final construct was sequenced across the entire GlpG-coding region in both directions.

Expression and Purification and Crystallization of GlpG. GlpG was expressed in *E. coli* C43(DE3) (31), which was deleted of the nonessential gene *mipA*, because the MipA protein was a major contaminant in our GlpG preparations. Deletion of *mipA* was performed according to Datsenko and Wanner (32). Cells were grown at 37°C until mid-log phase, induced with 0.5 mM isopropyl-β-D-thiogalactoside, shifted to 16°C, and harvested after 14 h. All of the purification and crystallization procedures were performed at 4°C. Cells were disrupted by French press, and the membrane fraction was isolated by ultracentrifugation. GlpG was solubilized from membranes in 50 mM Hepes (pH 7.5), 0.5 M NaCl, 8% glycerol, 20 mM imidazole, and 1% *n*-dodecyl-β-D-maltopyranoside (Anatrace), and then purified by nickel affinity chromatography during which *n*-dodecyl-β-D-maltopyranoside was replaced by lauryl dimethylamine oxide (Anatrace). Trypsin (Pierce) was added at a ratio of 1:20 (wt/wt trypsin/GlpG), and the solution was left overnight at 4°C. Sequential passages through Para-aminobenzamidine (70 μl) and Sephadex G-50 (10 ml) columns were used to remove trypsin and small cleavage products. The truncated protein was then dialyzed (≈4 h) and concentrated (Vivaspin 30-kDa cutoff) to 5 mg/ml in 20 mM Hepes (pH 7.5), 90 mM NaCl, 10% glycerol, and lauryl dimethylamine oxide. Crystals of GlpG were grown in hanging drops over a reservoir solution containing 30% (wt/vol) PEG 400 (Fluka), 200 mM $CaCl_2$, and 100 mM Mes (pH 6.5).

Structure Determination. Native crystals of GlpG diffracted to 2.3 Å. The native data set used for refinement was collected at 120 K on an RU-H3R generator (Rigaku, Tokyo, Japan) equipped with an R-AXIS IV image plate system and Osmic mirrors. Data from an ethylmercury chloride derivative were collected at European Synchrotron Radiation Facility beamline ID29-1. All data sets were processed and scaled by using the programs Denzo and Scalepack. Heavy-atom positions were determined by using SOLVE (33). Using the program RESOLVE (34, 35), phases were improved by density modification exploiting noncrystallographic symmetry averaging, and a starting model that included almost all of the GlpG sequence was automatically built. The model was improved by several rounds of manual building with O (36) and refinement in CNS (37). The model was verified by using MolProbity (38) and PROCHECK (39). Aside from Ser-201 in molecule B ($\varphi = 45°$, $\psi = -110°$), no other residues were found in disallowed regions of the Ramachandran plot. The absence of a bound lipid may be responsible for Ser-201 being in a helical configuration in molecule A. Structure figures were created with PyMOL (40).

We thank B. Z. Shilo and D. S. Tawfik for critically reading the manuscript. E.B. is grateful to R. G. Brennan, M. A. Schumacher, and their colleagues for introducing him to the field of crystallography. A.B.-S. thanks N. Nelson for introducing him to the field of membrane protein biochemistry and for support in the initial stages of this project. This work was supported by grants from the Israel Science Foundation and the Y. Leon Benoziyo Institute for Molecular Medicine at the Weizmann Institute of Science (to E.B.). A.B.-S. is the recipient of a Sir Charles Clore Postdoctoral Fellowship. D.F. is the incumbent of the Lilian and George Lyttle Career Development Chair.

1. Brown MS, Ye J, Rawson RB, Goldstein JL (2000) *Cell* 100:391–398.
2. Wolfe MS, Kopan R (2004) *Science* 305:1119–1123.
3. Marjaux E, Hartmann D, De Strooper B (2004) *Neuron* 42:189–192.
4. Selkoe D, Kopan R (2003) *Annu Rev Neurosci* 26:565–597.
5. Rawson RB, Zelenski NG, Nijhawan D, Ye J, Sakai J, Hasan MT, Chang TY, Brown MS, Goldstein JL (1997) *Mol Cell* 1:47–57.
6. Ye J, Rawson RB, Komuro R, Chen X, Dave UP, Prywes R, Brown MS, Goldstein JL (2000) *Mol Cell* 6:1355–1364.
7. Koonin EV, Makarova KS, Rogozin IB, Davidovic L, Letellier MC, Pellegrini L (2003) *Genome Biol* 4:R19.
8. Urban S, Lee JR, Freeman M (2001) *Cell* 107:173–182.
9. Lee JR, Urban S, Garvey CF, Freeman M (2001) *Cell* 107:161–171.
10. Tsruya R, Schlesinger A, Reich A, Gabay L, Sapir A, Shilo BZ (2002) *Genes Dev* 16:222–234.
11. Brossier F, Jewett TJ, Sibley LD, Urban S (2005) *Proc Natl Acad Sci USA* 102:4146–4151.
12. Gallio M, Sturgill G, Rather P, Kylsten P (2002) *Proc Natl Acad Sci USA* 99:12208–12213.
13. Herlan M, Vogel F, Bornhovd C, Neupert W, Reichert AS (2003) *J Biol Chem* 278:27781–27788.
14. McQuibban GA, Saurya S, Freeman M (2003) *Nature* 423:537–541.
15. Cipolat S, Rudka T, Hartmann D, Costa V, Serneels L, Craessaerts K, Metzger K, Frezza C, Annaert W, D'Adamio, *et al.* (2006) *Cell* 126:163–175.
16. Clemmer KM, Sturgill GM, Veenstra A, Rather PN (2006) *J Bacteriol* 188:3415–3419.
17. Maegawa S, Ito K, Akiyama Y (2005) *Biochemistry* 44:13543–13552.
18. Urban S, Schlieper D, Freeman M (2002) *Curr Biol* 12:1507–1512.
19. Urban S, Wolfe MS (2005) *Proc Natl Acad Sci USA* 102:1883–1888.
20. Lemberg MK, Menendez J, Misik A, Garcia M, Koth CM, Freeman M (2005) *EMBO J* 24:464–472.
21. Mitra K, Ubarretxena-Belandia I, Taguchi T, Warren G, Engelman DM (2004) *Proc Natl Acad Sci USA* 101:4083–4088.
22. Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J, *et al.* (1992) *Protein Eng* 5:197–211.
23. Bartunik HD, Summers LJ, Bartsch HH (1989) *J Mol Biol* 210:813–828.
24. Johnson A, Gautham N, Pattabhi V (1999) *Biochim Biophys Acta* 1435:7–21.
25. Hedstrom L (2002) *Chem Rev* 102:4501–4524.
26. Radisky ES, Lee JM, Lu CJ, Koshland DE, Jr (2006) *Proc Natl Acad Sci USA* 103:6835–6840.
27. Paetzel M, Dalbey RE (1997) *Trends Biochem Sci* 22:28–31.
28. Wei Y, Schottel JL, Derewenda U, Swenson L, Patkar S, Derewenda ZS (1995) *Nat Struct Biol* 2:218–223.
29. Freeman M (2004) *Nat Rev Mol Cell Biol* 5:188–197.
30. Urban S, Freeman M (2003) *Mol Cell* 11:1425–1434.
31. Miroux B, Walker JE (1996) *J Mol Biol* 260:289–298.
32. Datsenko KA, Wanner BL (2000) *Proc Natl Acad Sci USA* 97:6640–6645.
33. Terwilliger TC, Berendzen J (1999) *Acta Crystallogr D* 55:849–861.
34. Terwilliger TC (2000) *Acta Crystallogr D* 56:965–972.
35. Terwilliger TC (2003) *Acta Crystallogr D* 59:38–44.
36. Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) *Acta Crystallogr A* 47:110–119.
37. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, *et al.* (1998) *Acta Crystallogr D* 54:905–921.
38. Davis IW, Murray LW, Richardson JS, Richardson DC (2004) *Nucleic Acids Res* 32:W615–W619.
39. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) *J Biomol NMR* 8:477–486.
40. DeLano WL (2002) PyMOL (DeLano Scientific, Palo Alto, CA), www.pymol.org.