Structure
Article

Rhomboid Protease Dynamics and Lipid Interactions

Ana-Nicoleta Bondar,[1,2] Coral del Val,[3] and Stephen H. White[1,2,*]
[1]Department of Physiology and Biophysics
[2]Center for Biomembrane Systems
University of California, Irvine, Irvine, CA 92697-4560, USA
[3]Department of Computer Science and Artificial Intelligence, University of Granada, Granada, E-18071, Spain
*Correspondence: stephen.white@uci.edu
DOI 10.1016/j.str.2008.12.017

SUMMARY

Intramembrane proteases, which cleave transmembrane (TM) helices, participate in numerous biological processes encompassing all branches of life. Several crystallographic structures of *Escherichia coli* GlpG rhomboid protease have been determined. In order to understand GlpG dynamics and lipid interactions in a native-like environment, we have examined the molecular dynamics of wild-type and mutant GlpG in different membrane environments. The irregular shape and small hydrophobic thickness of the protein cause significant bilayer deformations that may be important for substrate entry into the active site. Hydrogen-bond interactions with lipids are paramount in protein orientation and dynamics. Mutations in the unusual L1 loop cause changes in protein dynamics and protein orientation that are relayed to the His-Ser catalytic dyad. Similarly, mutations in TM5 change the dynamics and structure of the L1 loop. These results imply that the L1 loop has an important regulatory role in proteolysis.

INTRODUCTION

Intramembrane proteases are membrane proteins that cleave TM helices to liberate proteins that participate in important processes, such as cell signaling and gene regulation (Brown and Goldstein, 1997; McQuibban et al., 2003; Struhl and Greenwald, 1999; Urban et al., 2001; Weihofen et al., 2002; Wolfe and Kopan, 2004). The first rhomboid serine protease was identified in *Drosophila melanogaster*, where it was found to cleave the membrane-anchored Spitz growth factor that stimulates the epidermal growth factor receptor (Bier et al., 1990; Freeman et al., 1992; Urban et al., 2001). Rhomboid proteases were subsequently found to be present in almost all organisms, including *Escherichia coli* (Koonin et al., 2003; Wasserman et al., 2000). The three-dimensional structures of the *E. coli* rhomboid (GlpG) solved recently to atomic resolution provided important insights into the possible mechanism of substrate cleavage (Ben-Shem et al., 2007; Wang et al., 2006, 2007; Wu et al., 2006). Albeit invaluable, these are frozen snapshots of GlpG in a crystal environment, which provide little information on lipid interactions and dynamics.

Knowledge of how GlpG interacts with lipids is important for beginning to understand how GlpG interacts with substrates and how the composition of the lipid membrane could influence GlpG catalysis. Lipids can affect the activity of membrane proteins (Lee, 2004), and may be particularly important for membrane-embedded proteases (Kalvodova et al., 2005; Narayanan et al., 2007; Urban and Wolfe, 2005). For example, GlpG is active when reconstituted with phosphatidylethanolamine (PE) lipids, but is inactive in phosphatidylcholine (PC) (Urban and Wolfe, 2005). The highly irregular shape of GlpG revealed by the crystal structures makes the dynamics of the lipid interactions very intriguing. The role of L1 in rhomboid protease function is also poorly understood; although it extends away from the active site (Figures 1B and 1C), its deletion inactivates the protein (Wang et al., 2007). The observation that mutation of specific L1 amino acids not directly connected to the catalytic site drastically reduces the activity was interpreted to suggest that L1 has an important structural role (Baker et al., 2007). The effect of the L1 mutations could imply that perturbation of L1 is relayed as a change in the microenvironment of the catalytic groups. To begin to understand at the atomic level the dynamics and the lipid interactions of the GlpG intramembrane protease, we have performed extensive all-atom molecular dynamics (MD) simulations of wild-type and two mutant *E. coli* GlpG rhomboids embedded in two different lipid membrane environments.

GlpG consists of a catalytic TM domain and a cytosolic N-terminal domain that is not essential for catalysis (Wu et al., 2006). The recent crystal structures of the TM domain of GlpG (Ben-Shem et al., 2007; Wang and Ha, 2007; Wang et al., 2006; Wu et al., 2006) reveal six TM helices (TM1–TM6) with very different tilt angles, and the unusual periplasmic L1 loop accommodated between TM1 and TM3 (Figure 1). Alignment of these several structures shows that the greatest structural variability occurs in TM5 and loop L5 (White, 2006). The catalytic residues (S201 and H254 on TM4 and TM6, respectively) are located within the bilayer boundaries (Ben-Shem et al., 2007; Urban et al., 2001; Wang et al., 2006; Wu et al., 2006) (Figure 1D). The presence in one of the crystal structures of a phosphoglycerol lipid head group close to the catalytic dyad suggests exposure of the active site to the lipid membrane (Ben-Shem et al., 2007).

The structural events along the reaction pathway of GlpG are not known, and there is no structural information about the complex between GlpG and its substrate(s). Even though GlpG can cleave the *Drosophila* Spitz peptide (Urban et al., 2002) and the TM2 helix of LacY permease (Maegawa et al., 2005),

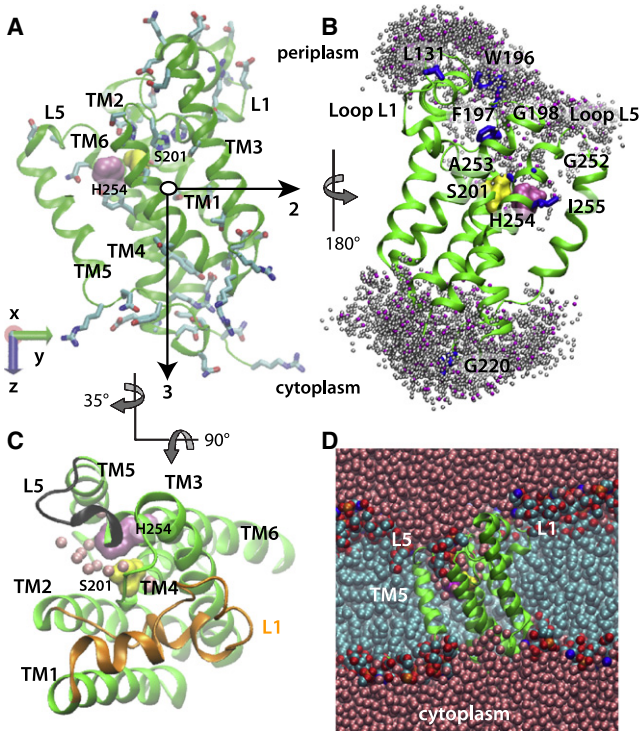

Figure 1. E. *coli* GlpG in a POPE Lipid Bilayer

(A) Starting geometry and principal axes (black arrows) of GlpG. S201 and H254 are shown as yellow and mauve surfaces, respectively; amino acid side chains with H-bonding capability are depicted as bonds.

(B) Dynamic water distribution in the active site. Locations of the water-molecule oxygen atoms located within 5 Å of the protein in a snapshot from Sim2 are shown as pink spheres. Shown in gray are water locations sampled during ten equally spaced snapshots during a subsequent 1 ns segment of Sim2. Amino acids known from experiments to be accessible to bulk water (Maegawa et al., 2007) are shown as blue bonds, or as surfaces (S201 and H254).

(C) View from the luminal side of GlpG from Sim2, depicting water molecules within 5 Å of S201 and H254 as pink spheres, L1 in orange, and L5 in black.

(D) Cut away view of GlpG in a POPE lipid bilayer, with the protein shown in green, bulk water oxygen atoms in pink, lipid alkyl carbon in cyan, lipid phosphorus in orange, lipid oxygen in red, and lipid nitrogen atoms in blue.

and may participate in quorum sensing (Clemmer et al., 2006), the physiological substrate of GlpG remains unclear. Two different scenarios have been proposed to explain how structural rearrangements of GlpG may assist the formation of the enzyme-substrate complex. In one scenario, TM5 acts as a lateral gate to control substrate access to the active site (Baker et al., 2007; Wu et al., 2006). This scenario is supported by the significant range of TM5 tilt angles (~30°) observed in the crystal structures (Ben-Shem et al., 2007; Wang et al., 2006; Wu et al., 2006) and by the effect of site-directed mutations on catalytic efficiency (Baker et al., 2007). In the second scenario, the cap loop L5 (Figure 1C) opens to allow docking of the substrate to the active site (Wang et al., 2007), consistent with the conformational variability of L5 in the crystal structures. Of course, L5 and TM5 are connected, implying that the reaction path may involve both L5 and TM5.

To derive a dynamic view of GlpG in a fluid lipid bilayer, we examined the lipid interactions of GlpG embedded in 1-palmytoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPE) and 1-palmytoyl-2-oleoyl-*sn*-glycero-3-phosphatidylethanolamine (POPC) lipid bilayers. The roles of the L1 loop and the TM5 helix were examined by modeling two different mutant GlpG phenotypes in which activity is significantly reduced (Y138S/F139S/L143S) or enhanced (L229V/F232V/W236V) (Baker et al., 2007). Our analysis of these mutants suggests that TM5 and L1 are dynamically coupled so that changes in the dynamics of one are relayed to the other.

RESULTS

We performed four all-atom MD simulations (Sims, Table 1) on systems comprised of one protein molecule, ~500 lipids, and 30,000 waters, for a total of ~160,000 atoms (see Experimental Procedures). Two independent simulations of wild-type GlpG were carried out in POPC (Sim1) and POPE (Sim2). In Sim3 and Sim4, we explored the dynamics of the mutant phenotypes Y138S/F139S/L143S (the triple-Ser mutant; Sim3) and L229V/F232V/W236V (the triple-Val mutant; Sim4) in POPE bilayers. Because there are no crystal structures of these mutants, starting coordinates were prepared from the wild-type crystal structure (Ben-Shem et al., 2007). Consequently, the limited timescales of our simulations may not capture the full extent of the protein conformational changes induced by the mutations. All simulations were run at least 10 nanoseconds (ns) beyond the time point at which the dimension of the unit cell reached plateau values. All histograms and average values were computed based on 10,000 sets of coordinates saved every 1 ps during the last 10 ns of each simulation run. The time-dependence plots of distances and root-mean-squared deviation (rmsd) values are for the unconstrained simulation runs.

Rhomboid Protease Orientation and Lipid Bilayer Thinning

Defining protein orientation as the angle between the bilayer normal and the third principal axis of the protein (Figure 1A), we found average orientation angles of 11.7° ± 1.8° and 15.9° ± 2.3° for POPE and POPC bilayers, respectively (Figure 2A). The difference between these angles appears to be a result of protein-bilayer interactions that optimize H bonding between protein surface groups and lipid head groups (below). Optimization is achieved mainly by the surrounding lipid molecules adjusting their geometry (Figures 2B–2D, 3A, and 3B) to accommodate the relatively stable TM part of the protein (Figures 4A and 4B).

The rearrangement of the lipid molecules led to nonuniform thinning of the membrane in the vicinity of the protein (Figure 2D). Although the average glycerol-to-glycerol thickness of the bilayer in the first 2–3 shells of lipids around the protein is only ~4 Å smaller than at remote distances (Figures 2B and 2C), the thickness varies considerably around the protein perimeter (black curves). The largest lipid rearrangements and bilayer thinning occurred on the cytoplasmic side, underneath the L1 loop (Figure 2D). The average hydrophobic thicknesses of the unperturbed, distant regions of the POPE and POPC lipid bilayers (~39 Å and ~38 Å, respectively) are compatible with earlier bilayer simulations (Jensen and Mouritsen, 2004). Bilayer thinning close to the protein was also proposed by Wang et al. (2007). However,

Table 1. Summary of MD Simulations of the GlpG Protease in Different Membrane Environments

Simulation	Protein	Lipid[a]	Active-Site Lipid	Length (ns)[b]
Sim 1	Wild-type	POPC	PGV	34.7
Sim 2	Wild-type	POPE	POPE	34.4
Sim 3	Triple-Ser mutant[c]	POPE	POPE	20.4
Sim 4	Triple-Val mutant[d]	POPE	POPE	26.9

[a] Lipids: POPC, palmitoyloleoylphosphatidylcholine; POPE, palmitoyloleoylphosphatidylethanolamine; PGV, palmitoylvaccenoylphosphatidylglycerol.
[b] Unconstrained simulation time in nanoseconds.
[c] Triple-Ser mutant: Y138S/F139S/L143S.
[d] Triple-Val mutant: L229V/F232V/W236V.

direct comparison with the thinning we observe is difficult because of the complex nature of the thinning close to GlpG (Figures 2B–2D) and the wide distribution of the protein's orientation angles in the fluid bilayer (Figure 2A), which could not have been anticipated from the X-ray structure alone.

The catalytic serine is located deep within the lipid bilayer (Figure 4C). The positions of S201 oxygen fluctuate around the mean values of ≈6 Å (POPC) and ≈5 Å (POPE) above the bilayer midplane in the periplasmic leaflet. Relative to the midplane of the 20 Å-thick detergent belt in their X-ray structure, Wang et al. (2007) estimated a distance of ≈7 Å. What distinguishes our result from that of Wang et al. (2007) is that the location of S201 is dynamic and is influenced by the composition of the lipid bilayer (Figure 4C). Even though S201 is relatively close to the bilayer midplane, it is nevertheless in a hydrophilic environment comprised of water and lipid head groups (Figures 1 and 3) due to bilayer thinning in the vicinity of the protein (Figure 2D). An aqueous environment at the active site was predicted from experiments (Maegawa et al., 2005, 2007) and is required for proteolysis.

Membrane Hydrogen Bonding of GlpG Depends on the Lipid Environment

The charged protein residues that can readily form hydrogen bonds (H bonds) with the lipid head groups are located remote from the active site, at the cytoplasmic termini of the TM segments, and on L1 (Figure 1A). Importantly, some of the lipid-protein H bonds depend on the type of lipid head group. The positively charged protein groups located at the membrane interface (Figure 1A) H bond to both POPE and POPC (Figures 3A and 3B). A detailed analysis of the protein-lipid H bonds indicates that some interactions are remarkably stable, whereas others are dynamic, breaking and reforming during the simulations. For example, the lipid H bonding of K191 (Figure 3A) is very stable (donor-acceptor distances of ∼2.7 Å for the last ∼20 ns of both Sim1 and Sim2). However, the lipid interactions of R92, R227, or R168 break and reform during the simulations. The lipid molecule close to the catalytic site in the crystal structure (Ben-Shem et al., 2007) displaces somewhat toward the bulk and is part of an H-bonded network that comprises the lipid molecule, water, the catalytic S201 and H254, and other polar amino acids (Figures 3C and 3D).

The different H-bonding capabilities of the N^+H_3 (POPE) and the bulkier $N^+(CH_3)_3$ (POPC) groups lead to significant differences in L1:membrane interactions (Figures 3E–3G). In POPE, strong H bonds between E134, R137, and a POPE amine form a structural barrier impenetrable to water molecules (Figure 3E). In contrast, during most of Sim1, E134 forms a water-mediated H bond with a POPC phosphate, which allows water molecules to penetrate deeper into the lipid bilayer (Figure 3F). Structural rearrangements of POPC toward the end of Sim1 lead to the

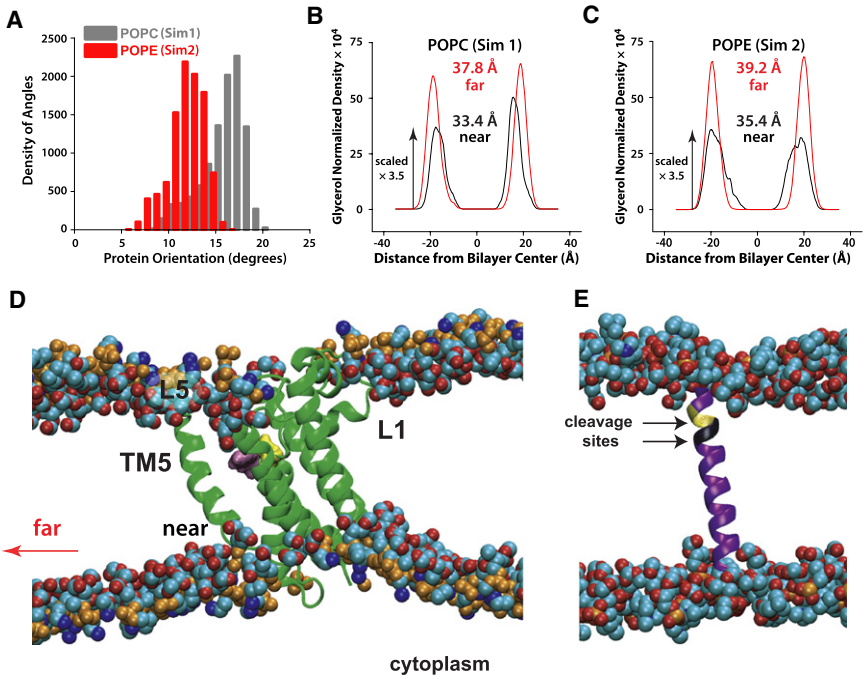

Figure 2. Local Perturbations of Lipid Bilayers Caused by GlpG Rhomboid Protease

(A) Distribution of orientational angles of GlpG relative to the membrane normal in Sim1 (gray) and Sim2 (red).

(B and C) Hydrocarbon thickness of the (B) POPC bilayers (Sim1) and the (C) POPE bilayers (Sim2) close to the protein (black curves; computed from ∼90 lipids within the first 2–3 shells of lipids around the protein) and far from the protein (red curves) in the unperturbed region of the bilayer. The hydrocarbon thickness was estimated as the distance between the peaks of the distribution for the glycerol groups of the two lipid leaflets (Wiener and White, 1992) taken along the membrane normal and normalized by the volume of the simulation cell. In the case of the distorted density peaks for POPE lipids close to the protein (z = ∼20 Å), we used the value of z at the center of the distribution.

(D) Thinning of the lipid bilayer close to the protein; only a ∼38 Å-wide section of the bilayer containing the protein is depicted.

(E) Snapshot of the Spitz substrate in a POPC lipid bilayer. The two substrate cleavage sites (Baker et al., 2007) are shown in yellow (Ala-Ser) and black (Gly-Ala). Little perturbation of the bilayer in the vicinity of Spitz was observed.

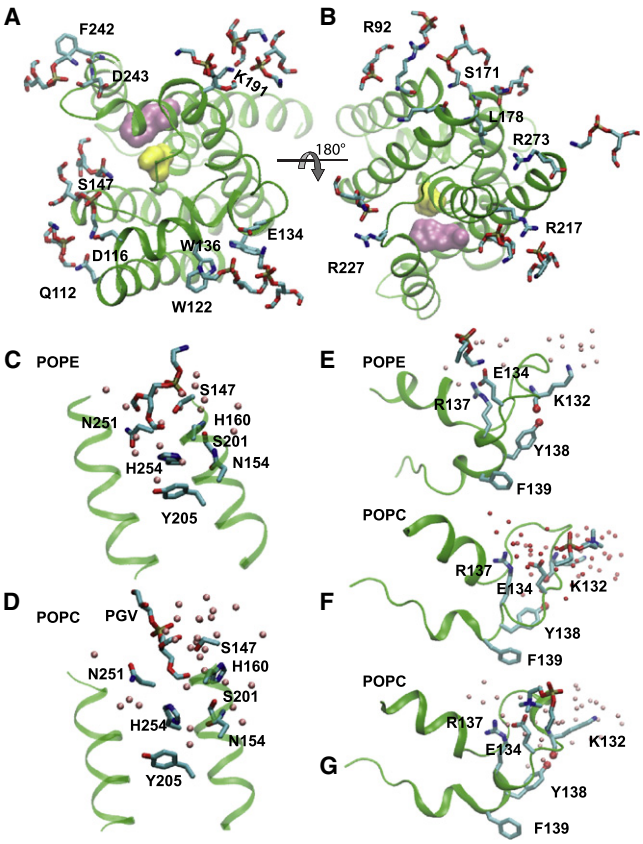

Figure 3. Lipid and Water Interactions of GlpG in POPC and POPE

(A–G) Pink spheres indicate the location of the bulk water oxygen atoms within 5 Å of the protein amino acid residues shown as bonds. Phospholipid head groups are also shown as bonds. (A) Luminal interactions of GlpG residues with POPE head groups (Sim2). (B) Cytoplasmic interactions of GlpG residues with POPE head groups. (C) Active-site interactions of GlpG in POPE membranes. (D) Active site interactions of GlpG in a POPC membrane, with a PGV lipid close to the active site (Sim1). (E) Loop L1 interactions with POPE head groups. (F and G) Loop L1 interactions with POPC head groups (F) during a segment of ~20 ns of Sim1 and (G) late in the simulation.

formation of an H bond between E134 and the POPC choline (Figure 3G). The dependence of L1:membrane interactions on lipid type may influence its local structure and intraprotein H bonding: in POPE, the H bond between Y138 and the K132 carbonyl group is stable during Sim2 at ~2.8 Å, whereas in Sim1 the presence of the Y138:K132 H bond correlates with the lipid interactions of E134 (Figures 3F, 3G, and 5D). A correlation between the L1 H bonding and hydration can also be observed in the structural analysis of the W136A mutant, in which the space initially occupied by W136 became filled with water molecules upon mutation (Wang et al., 2007).

Consistent with the stronger L1:lipid H bonding in POPE versus POPC, the root-mean-squared fluctuation (rmsf) values of the L1 atoms are somewhat smaller in POPE (Figure 4E). H bonding to lipid and intraprotein H bonding (see below) explain the structural stability of L1, indicated by small rmsd values comparable to those of the TM region (Figures 4A and 4B). Structural stability of L1 is also suggested by the structure of L1 being similar in the wild-type and in the W136A mutant in vitro (Wang et al., 2007).

The emergent pattern is that protein side chains anchor the protein to the lipid interface by H bonding. To assess further the importance of the H bonds, we performed a sequence analysis (see Supplemental Data available online) for three data sets extracted from the PFAM database version 23.0 (Finn et al., 2008): the family of Enterobacteriaceae (46 sequences), the PFAM seed (65 hand-curated sequences of rhomboids from Archea, Bacteria, Plants, and Eukaryotes), and in the complete PFAM rhomboid domain (1582 sequences).

Amino acids with H bonding capability are present in the sequences of other rhomboids at the sites at which GlpG H bonds to the lipid bilayer. For example, half of the rhomboid sequences from the PFAM database seed and 870 sequences from the full PFAM alignment have residues with H bonding capability corresponding to R217; 55 and 836 sequences of the seed and of the full PFAM alignments, respectively, have a residue with H bonding capability corresponding to E134 (Figures 3A and 3B; Tables S2 and S3). R227 is conserved in all except for two Enterobacteriacea sequences, but not in other organisms (Tables S1–S3). H bonding at the periplasmic K191 (Figure 3A) is not significantly conserved, occurring in 45% of the Enterobacteriacea sequences, 27% of the PFAM seed, and 18% of the full PFAM alignment (Tables S1–S3). Differences in the nature of residues located at the interface with lipid head groups could lead to a different response to lipids. For example, GlpG and the *Providencia stuartii* rhomboid respond differently to PC lipids (Urban and Wolfe, 2005). Although further studies are needed to understand how lipids affect rhomboid proteases, we note that differences in the L1 sequences of these two rhomboids could contribute to their different responses to lipids. For example, both GlpG and the *P. stuartii* rhomboid have E or D at position 134, R or K at position 217, and R or Y at position 227 (the *P. stuartii* rhomboid also has Arg at position 226), but they differ in the 125 position (Koonin et al., 2003; Urban et al., 2002): *P. stuartii* has Asn, whereas GlpG has Trp, which is located in a sensitive region of GlpG where its H-bonding dynamics change upon perturbation of L1 and TM5 (Figures 7E and 9D).

Dynamics of Intraprotein H-Bond Interactions

Multiple intraprotein H bonds are observed between amino acid residues with polar or charged side chains distributed asymmetrically (Figures 5A and 5B). The polar side chains are located preferentially at the cytoplasmic ends of the TM helices, on L1, and along TM3. TM5 is not connected through H bonds to other helices (Figure 5B), which could explain its rmsf values being larger than for the other TM helices (Figure 4E). In contrast, the cytoplasmic termini of the remaining five helices are interconnected by two clusters of H bonds that bridge TM1/TM2/TM3 (123-cluster, Figure 5A) and TM3/TM4/TM6 (346-cluster, Figure 5B). Both clusters involve a highly conserved amino acid residue: TM2-E166 in the 123-cluster and TM3-K173 in the 346-cluster are present as E and as R/K in 83% and 73%, respectively, of the sequences from the PFAM seed (Table S2); 69% and 58%, respectively, of the 1582 sequences from the PFAM full alignment have E166 and R/K173 (Table S3). TM3, the only TM segment involved in both H-bonding clusters, also connects to L1 via the W125:Q190 H bond (Figure 5A), and it has an array of intrahelical H bonds involving T178, S181, and

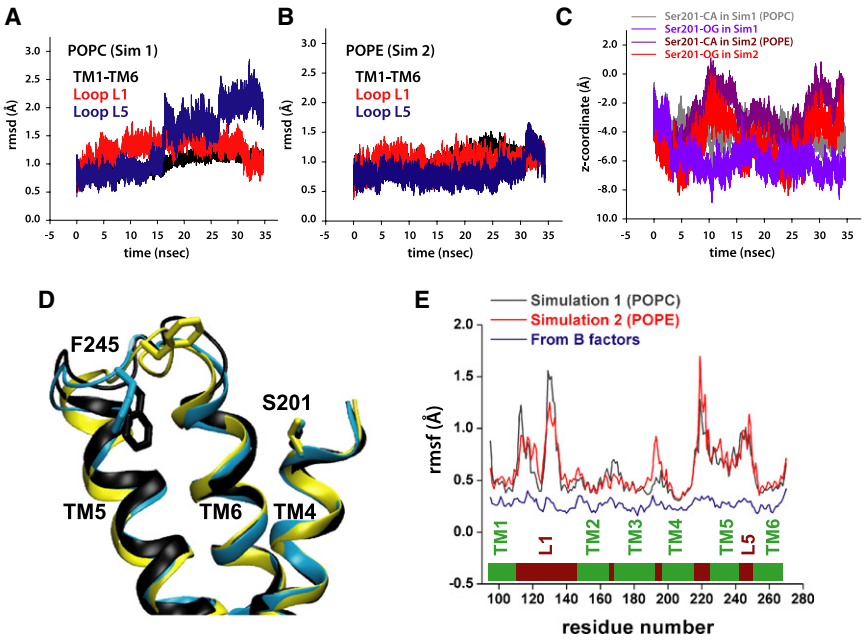

Figure 4. Dynamics of GlpG in POPC and POPE Bilayers

(A and B) Dynamics of loop L1 (red), loop L5 (blue), and of helices TM1–TM6 (black) in (A) POPC (Sim1) and (B) POPE (Sim2) bilayers measured as the rmsd (Å) relative to the starting crystal structure coordinates.

(C) Position of the catalytic S201 in POPC (Sim1) and POPE (Sim2) bilayers.

(D) Conformations sampled by L5 in POPC (Sim1) during the time periods 0–15 ns (yellow), 15–25 ns (cyan), and 25–34.5 ns (black).

(E) Rmsf (Å) of GlpG residues in Sim1 (black, POPC) and Sim2 (red, POPE) are dramatically greater than those computed from the crystal structure B factors, shown in blue (Ben-Shem et al., 2007).

S185. S181 is present as S/T in 15% of the PFAM seed sequences and in 44% of the sequences from the full alignment. T178 and S185 are less conserved than S181 (Table S3).

The cytoplasmic clusters of H bonds are present regardless of the membrane composition. The interactions between the various groups are stable during the simulations, with characteristic distances of 2.6–2.8 Å (Figures 5C and 5E). The only exception is the H bond between TM2-E166 and TM3-S171, which is stable at 2.6 ± 0.1 Å in POPE, whereas in POPC it breaks and reforms rapidly such that an H bond (donor-acceptor distance of 2.7 ± 0.1 Å) is present half of the time (Figure 5C).

The presence of the stable H-bond clusters at the cytoplasmic site likely contributes to the structural rigidity of the TM region, indicated by the small rmsd values (Figures 4A and 4B). The largest number of H bonds is observed for TM3, which H bonds to TM1, TM2, TM4, TM6, and L1; there is also an array of H bonds along TM3 (Figures 5A and 5B). Because TM3 is an H-bonding partner in both H-bond clusters and for L1, we suggest that TM3 could act as a messenger of structural and dynamical perturbations throughout the protein. Although a systematic experimental study of the effects of mutating H-bonding amino acids has not been performed yet, we note that the TM3-K173A mutant, which might perturb both H-bonding clusters, has reduced proteolytic activity (Urban et al., 2001). The seemingly normal activity of the TM2-E166A mutant in *D. melanogaster* rhomboid expressed in mammalian COS cells (Urban et al., 2001) might be due to a lipid head group substituting for the Glu side chain.

The E134:R137 and Y138:K132 H bonds (Figure 5A) may be particularly important for the local geometry of L1, and they could contribute to the rmsd values of L1 being similar to those of the TM region (Figures 4A and 4B). R137 is part of the highly conserved WR motif (Koonin et al., 2003; Urban et al., 2001), which is present in more than 80% of both the seed and full PFAM alignments (see Supplemental Data). The presence of the Y138-K132 H bond depends somewhat on the interactions of L1 with the lipid bilayer (Figures 3E–3G and 5D). Site-directed mutagenesis suggests a possible role of R137 and Y138 in GlpG function: the R137A mutant has reduced proteolytic activity in cells and is inactive in detergent (Wang et al., 2007); the Y138F mutant has a reduced activity in detergent (Baker et al., 2007). Many rhomboid sequences have amino acid residues with H-bonding capability at positions 134 and 137 (Tables S1–S3; see also [Koonin et al., 2003]). Y138 is conserved only in Enterobacteriaceae (~52%) (Table S1).

Conformational Flexibility of the Cap Loop L5

A key question about rhomboid protease function is how substrate access to the catalytic site is controlled. One scenario assigns a central role to the L5 cap loop (Figure 1C), which presumably opens upon docking of the substrate to the protease (Wang and Ha, 2007). On the timescale of our simulations, we observed two conformational transitions of L5 in Sim1, after ~15 ns and ~27 ns of unconstrained dynamics (Figure 4A). In Sim2, a conformational transition of L5 was observed after ~30 ns of unconstrained dynamics (Figure 4B). The conformational transitions of L5 occurred without any significant change of the distances between TM5 and TM6 (Figure 4D). That motions of L5 do not necessarily have to be accompanied by changes in the TM region is also suggested by the X-ray crystallographic study of Wang and Ha (2007). The differences in the dynamics of L5 in POPC and POPE lipids can be explained by L5 being located within the interfacial region (Figure 2D) of the membrane (Wang et al., 2007), where its dynamics can be affected by the lipid head groups. Indeed, the dynamics of the D243:M247 H-bonding interactions (Figure 5B) depend significantly on the lipid environment (Figure 5F).

The observation of fast structural rearrangements of L5 is consistent with its relatively large rmsf values (Figure 4E), and with the conformational variability of L5 in the crystal structures. Due to the limited length of our simulations, we cannot exclude the possibility that the conformational transitions of L5 are reversible, or that further structural rearrangements of the loop might occur on a longer timescale.

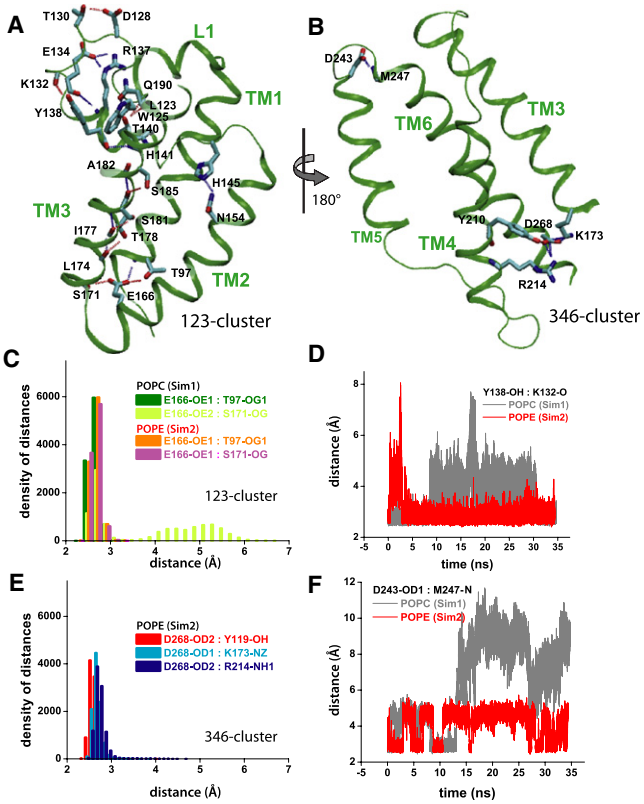

Figure 5. Intraprotein H Bonding of GlpG

(A) H bonding within L1 and TM1/TM2/TM3 (123-cluster).

(B) H bonding within TM3/TM4/TM6 (346-cluster). For simplicity, only the backbone carbonyl or amide groups are depicted for nonpolar residues.

(C) Histograms of H-bonding distances in the 123-cluster for POPC and POPE. The E166-S171 H bond breaks and reforms in POPC.

(D) Effect of lipid type on the dynamics of the Y138-K132 H bonding in the 123-cluster.

(E) Histograms of H-bonding distances within the 346-cluster in POPE. The histograms for the 346-cluster are very similar for POPE and POPC (data not shown).

(F) Dynamics of H bonds in the lipid interface of L5 depend upon lipid type.

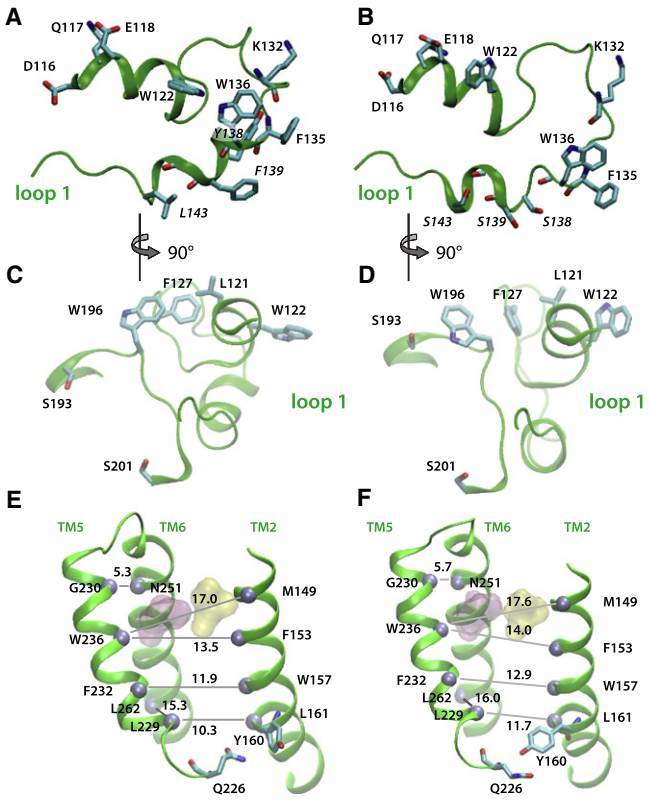

Figure 6. Simulation Snapshots Illustrating Relay of Loop L1 Perturbations to the Substrate Access Site in the Triple-Ser Mutant

(A–F) Snapshots of the wild-type and the triple-Ser Y138S/F139S/L143S mutant conformations are shown in the (A, C, and E) left-hand and (B, D, and F) right-hand panels, respectively. (A) Periplasmic view of L1 in the wild-type protein. All distances are in Å. (B) Periplasmic view of L1 in the triple-Ser mutant. (C) View of wild-type L1 viewed from the active site in the wild-type protein. (D) View of triple-Ser L1 from the active site. (E) Mean spacing between selected C_α atoms (gray spheres) in TM2/TM5/TM6 for the wild-type protein. (F) Mean C_α spacings for the triple-Ser mutant. For (E) and (F), the standard deviations of the distances are $\leq$ 0.6 Å.

Loop L1 Is a Key Determinant of the Protein Orientation within the Bilayer

The drastically perturbed proteolytic activity of some L1 mutants (Baker et al., 2007) and the lack of activity in the absence of L1 (Wang et al., 2007) indicate an important role for L1 in GlpG function. To gain insights into its mechanistic role, we performed an MD simulation (Sim3) of the L1 triple-Ser mutant Y138S/F139S/L143S (Figures 6A and 6B), which is inactive in proteolysis (Baker et al., 2007). The triple-Ser mutation affects the orientation of the protein within the lipid bilayer and the location of the catalytic Ser. On the timescale of the present simulations, the average tilt of the protein relative to the membrane normal is 14.9° ± 2.1°, as compared to 11.7° ± 1.8° in the wild-type (Figure 7A), and the catalytic Ser is buried ~2 Å deeper into the membrane. These results suggest that L1 participates in controlling the orientation of the protein within the lipid bilayer, which is important for the productive interaction with the substrate. A role for loop L1 in maintaining and adjusting the orientation of GlpG has also been suggested by Wang et al. (2007).

Perturbation of L1 Is Transmitted to the Active Site and TM2/TM5

The triple-Ser mutation of loop L1 causes structural rearrangements of L1 and, unexpectedly, the perturbation is transmitted to the active site and the TM2/TM5 helices (Figures 6 and 7). Optimization of the H bonding of the three substituted Ser residues causes small structural rearrangements of L1 (Figures 6A and 6B) that, in part, minimize the energy penalty for burying polar hydroxyl groups within the membrane hydrophobic region. The lipid-mediated interaction between the L1 residues W122 and W136 weakens (Figure 7D), and W122 flips with its $N_{\varepsilon 1}$-H group toward the bulk water (Figures 6A and 6B). Perturbation of the hydrophobic cluster formed by L121, L127, and W196 and the repositioning of F127 relative to W196 (Figure 7E) allow W196 to reorient toward the substrate access site, where it H bonds to TM3-S193 (Figures 6C, 6D, and 7E). That is, perturbation of L1 leads in our simulation to the formation of a new H bond 7 residues downstream from the catalytic S201. Moreover, the mutation induces the formation of an H bond between TM2-Y160 and TM5-Q225 (Figures 6E and 7E), perturbs the relative distance between TM2 and TM5

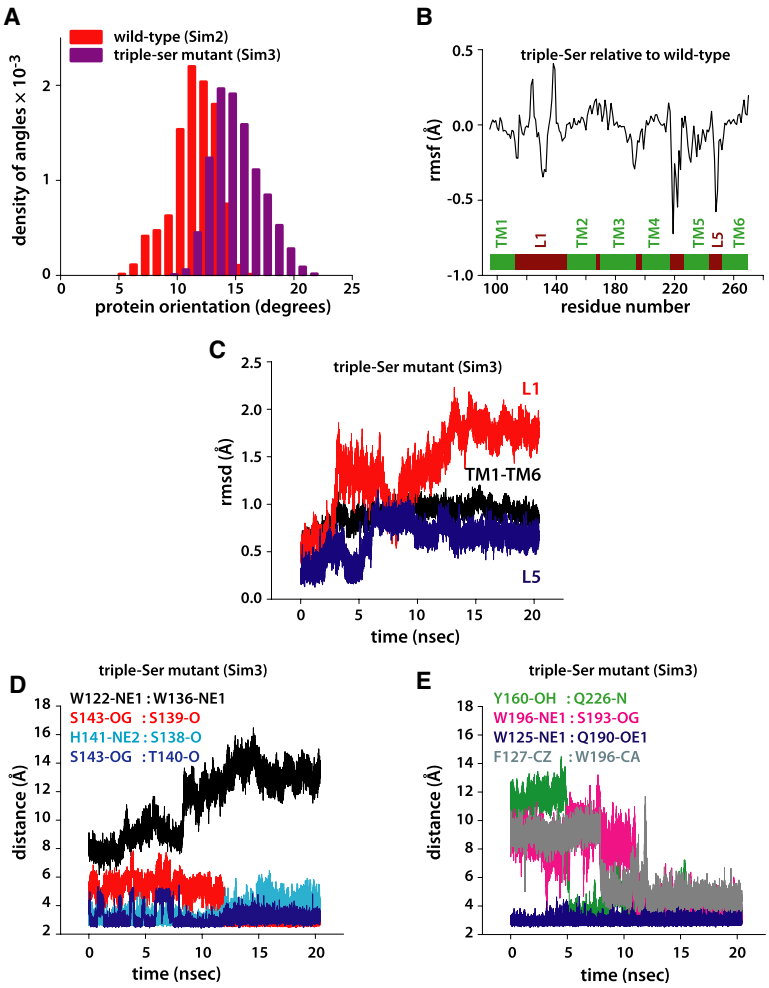

Figure 7. Dynamics of the Triple-Ser Mutant in POPE

(A) Histograms of the orientation of the protein relative to the membrane normal (measured in degrees) in the wild-type (red, Sim2) and the triple-Ser mutant (wine, Sim3).

(B) Rmsf (in Å) of the amino acid residues in the triple-Ser mutant relative to the wild-type protein.

(C) Rmsd (in Å) of the triple-Ser mutant relative to the equilibrated structure of wild-type GlpG in POPE (Sim2).

(D) Dynamics of selected distances in the triple-Ser mutant involving amino acids from L1.

(E) Dynamics of selected distances in the triple-Ser mutant involving amino acids not in the L1 loop.

(Figures 6E and 6F), and affects the fluctuations of amino acids as far away as the TM5-L5 region (Figure 7B). Thus, the changes in L1 cause a cascade of collectively important small-scale structural events that reach the substrate access site and TM2/TM5.

The changes in the structure, dynamics, and orientation of GlpG in the L1 triple-Ser mutant suggest a coupling between L1, the active site, and TM2/TM5. Because L1 interacts with TM3 (Figure 5A), which, in turn, is part of the two H-bonding clusters connecting the cytoplasmic termini of the helices (Figures 5A and 5B), structural and dynamical perturbations of L1 can be transmitted by TM3 to TM1, TM2, TM4, and TM6. The hydrophobic contacts between residues on the cytoplasmic sides of TM2 and TM5 (Figure 6E) (Ben-Shem et al., 2007; Wang et al., 2006; Wu et al., 2006) can further propagate the perturbation to TM5. Indeed, even in the early stages of Sim3, changes in the dynamics of interactions between L1 atoms (Figure 7D) were accompanied by changes in the interactions between L1 and groups from other structural elements (e.g., F127:W196, Figure 7E), and between protein groups located at distances remote from L1 (Y160:Q226, Figure 7E).

Perturbation of TM5 Influences Protein Dynamics

The conformational variability of the TM5 helix observed in the crystal structures (Ben-Shem et al., 2007; Wang et al., 2006; White, 2006; Wu et al., 2006) and experiments on mutant phenotypes (Baker et al., 2007) were interpreted to suggest that TM5 functions as a dynamic lateral gate that controls substrate access to the active site. If TM5 is indeed the lateral access gate, it is unknown how its conformation and dynamics change upon interaction with substrates and whether this interaction leads to further changes in the rest of the protein. To begin to understand the effect of TM5 perturbations, we performed simulations (Sim4) of the TM5 triple-Val mutant (W236V/F232V/L229V) (Figures 8A and 8B). This mutant is of special interest because it only involves amino acids of TM5, causes a 4-fold increase in activity relative to the wild-type (Baker et al., 2007), and has the potential to significantly increase the space between TM2 and TM5 (Figure 6E).

The structures of the TM segments did not change significantly in the triple-Val mutant (Figure 9A), but, due to the smaller size of the Val relative to the wild-type residues (Trp, Phe, and Leu), the accessibility of the catalytic Ser from the lateral side increased (Figures 8A and 8B). This change alone helps explain the enhanced activity of the triple-Val mutant. However, the dynamical changes in the protein go beyond just increasing access to the catalytic dyad. Relative to wild-type (Figure 8A), the mutation shifted the distributions of the protein orientation angles toward larger values (Figure 9B), perturbed the fluctuations of residues throughout the protein (Figure 9C), and led to structural rearrangements of L1 and L5 (Figures 9A and 9D). Importantly, the dynamics of intraprotein interactions that are affected by the triple-Ser mutation of L1 are also perturbed in the triple-Val mutant (e.g., F127-W196, W122-W136) (Figures 7D, 7E, and 9D). This observation supports our suggestion of a communication between remote regions of GlpG (Figure 9E).

The long-distance effects of mutating L1 or TM5 demonstrated here suggest long-range coupling in GlpG and help explain why GlpG's enzymatic activity is drastically perturbed when residues remote from the active are mutated. Long-range coupling has also been proposed, for example, for *Vibrio cholerae* RTX cysteine protease (Lupardus et al., 2008) and for the *E. coli* dihydrofolate reductase (Rod et al., 2003). Further computations and experiments are necessary to understand fully why catalytic activity is significantly reduced in triple-Ser

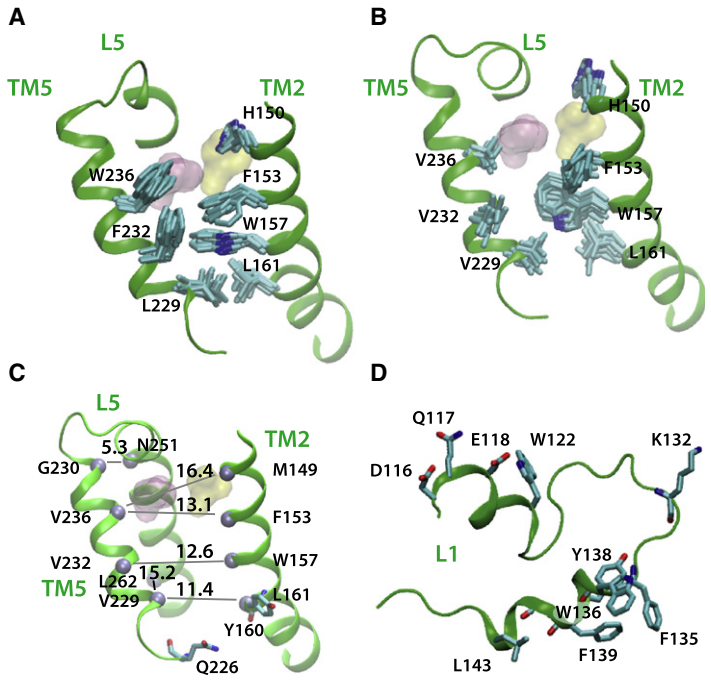

Figure 8. Simulation Snapshots Showing Structural Changes in TM5 of the Triple-Val Mutant in POPE

(A–D) Snapshots of wild-type (Sim2) and mutant (Sim4) conformations are shown in the (A and C) left-hand and (B and D) right-hand panels, respectively. (A) Loop L5, TM2, and TM5 in the wild-type protein. (B) The same features as for the triple-Val mutant. (C) Effect of the triple-Val mutant on the average distances between TM5, TM2, and TM6 C_α atoms (gray spheres) (compare with the distances in Figure 6E). All distances are average values (given in Å); standard deviations $\leq$ 0.6 Å. (D) Snapshot of loop L1 at the end of Sim4 (compare with Figure 6A).

mutants, but increased in triple-Val mutants. One potentially important difference between the two mutants is that, although in both the interactions between W193, S193, and F127 are perturbed (Figures 7E and 9D), only the (inactive) triple-Ser mutant appears to be locked in a conformation that strengthens the interactions between these groups.

Systematic computer simulation studies of additional GlpG mutants are necessary to understand why perturbing certain interactions affects the catalytic activity and which steps in GlpG's mechanism of action are affected by specific classes of mutations. For example, it is unclear why the double mutation W236A/F153A results in a 10-fold increase of activity, whereas the triple-Val mutant has a more modest 4-fold increase of activity relative to the wild-type (Baker et al., 2007).

DISCUSSION

We have carried out MD simulations of the wild-type *E. coli* GlpG rhomboid protease in POPC and POPE and two GlpG mutants in POPE. The structure of the TM helices of GlpG in bilayers is close to the crystallographic structures (Figures 4A and 4B). The TM5 helix and the L5 loop, which have the greatest structural variability among the three GlpG structures (White, 2006), are mobile regions of GlpG in lipid bilayers (Figures 4A, 4B, and 4E). The S201/H254 catalytic dyad is positioned deep in the lipid bilayer, where it nevertheless remains well hydrated at all times (Figures 1B and 4C). Differences in the H bonding of the lipid head groups with the protein affect the orientation (Figure 2A) and dynamics (Figures 4A, 4B, and 4E) of the protein. The orientation of GlpG within the lipid bilayer may be an essential determinant of enzyme efficiency, because orientation can influence the likelihood of a productive interaction between the catalytic dyad and the substrate cleavage site. Our simulations of the triple-Ser and triple-Val mutations reveal subtle, but important, changes in protein conformation and dynamics that are relayed between L1 and TM5 (Figures 6–9).

Optimization of the interactions of the protein with the lipid bilayer is mainly achieved by the structural rearrangement of lipid molecules. These rearrangements lead to changes in the hydrophobic thickness of the bilayer in the vicinity of the protein due to H bonding between surface residues and lipid head groups (Figures 2B–2D). This localized bilayer thinning close to the protein may influence the conformation (Wang et al., 2007) or orientation of the substrate, which by itself does not perturb the lipid bilayer (Figure 2E). Based on their estimate of 30% bilayer thinning, Wang et al. (2007) suggested that hydrophobic mismatch between the substrate and the bilayer around GlpG could force an unwinding of the cleavage site. When measured as the distances between the peaks in the average distributions of glycerol moieties in the two leaflets, the 4–5 Å thinning of the lipid bilayer indicated by our simulations (Figures 2B and 2C) could be interpreted to suggest that only the top 3 residues of the substrate would unwind. However, the lipid bilayer might undergo further perturbations when the substrate and GlpG are within a close distance, and the thinning of the lipid bilayer around GlpG might induce tilting of the substrate in addition to unwinding.

Due to the different H-bonding capabilities of the POPC and POPE head groups, the details of the protein-lipid interactions naturally depend on the lipid head group. This effect of the head group H-bonding capability is most prominent in the case of the L1 loop, where H-bond bridging between E134, R137, and a lipid molecule forms a barrier impenetrable to water molecules in POPE, but not in POPC (Figures 3E–3G). The dependence of the conformation of L1 on protein interactions with the lipid bilayer could allow L1 to act as a sensor of the GlpG environment. This sensitivity, together with the perturbing effect of L1 mutations on the protein orientation and dynamics, likely contributes to the effect of lipids on GlpG's proteolytic activity (Urban and Wolfe, 2005) and helps rationalize why the effect of certain L1 mutations (W136A, R137A) depends on whether the mutant is assessed in detergent or in various lipid environments (Lemberg et al., 2005; Maegawa et al., 2005; Urban et al., 2001; Wang et al., 2007). Further studies are required for a detailed understanding of how intramembrane proteases are affected by their environment. The functional role of H bonding between lipids and surface protein groups has been proved recently for the secondary multidrug transporter LmrP, whose function requires H bonding between a protein aspartate residue and the lipid membrane (Hakizimana et al., 2008).

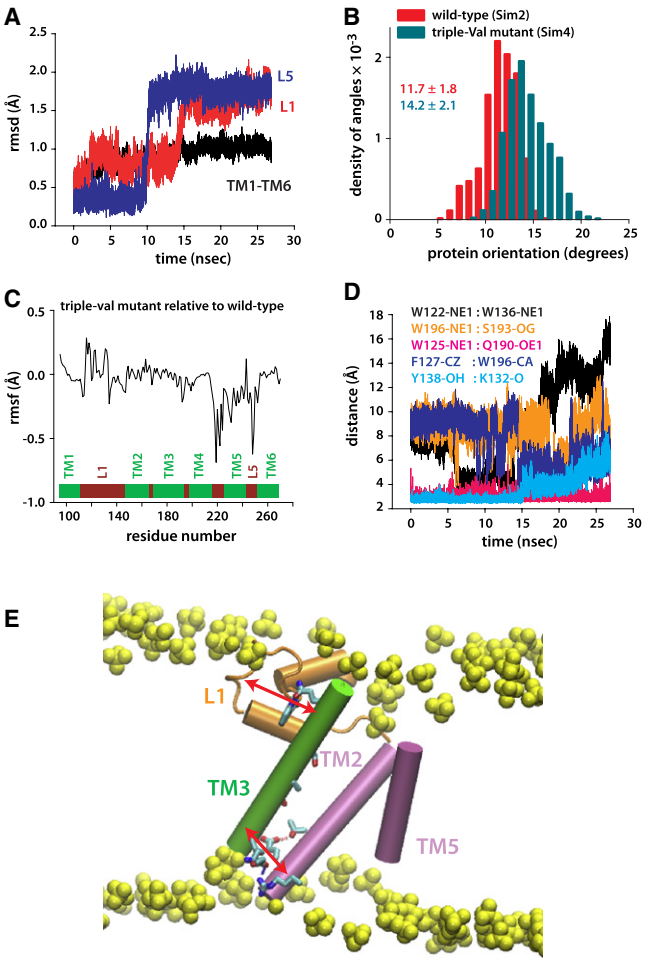

Figure 9. Protein Orientation and Dynamics of the Triple-Val Mutant

(A) Rmsd of TM1-TM6 and of loops L1 and L5 in Sim4, relative to the starting coordinates. The spontaneous and nearly simultaneous switching of L1 and L5 suggest conformational coupling between TM2/TM5 and L1.

(B) Orientation of triple-Val GlpG (compare with Figure 7A).

(C) Changes in the rmsf of triple-Val GlpG relative to the wild-type are observed not only in the region of TM5, but also in L1.

(D) Dynamics of selected intraprotein interactions in Sim4 indicating that perturbation of TM5 is transmitted to L1.

(E) Suggested relay mechanism in GlpG. Due to the array of H bonds along TM3 (green), this structural element can act as a rigid arm that relays perturbations between L1 (orange) and the TM2/TM5 helices (orange) through intrahelical and L1:TM3 H bonding.

Conformational transitions of L5 were observed both in the wild-type (Figures 4A and 4B) and in the TM5 triple-Val mutant (Figure 9A). Our observation of a conformational change of L5 in the absence of substrate (using the coordinates of GlpG in its closed conformation [Ben-Shem et al., 2007]) suggests that opening and closing of the loop is an intrinsic protein motion. Although the timescale of our simulations is too short to assess whether the cap loop closes back and then reopens, or whether further conformational changes of L5 occur on longer timescales, the significant flexibility of the cap loop indicated by experiments and our simulations supports such motions. If the optimal docking of the substrate in the active site indeed requires an opening of L5 (Wang et al., 2007), the intrinsic flexibility of L5 may allow for sampling of conformations compatible with substrate docking, even in the absence of substrate. This suggestion is inspired by the observation of the equilibrium, in the absence of the substrate, between the open and closed states of the lid domain controlling access to the adenylate kinase active (Hanson et al., 2007). However, the GlpG structure used here, which has a lipid head group bound to the active site (Ben-Shem et al., 2007), is of a closed-state enzyme. The dynamics of L5 and TM5 in the more open conformation of GlpG (Wu et al., 2006) may be different than observed here for the closed state (Ben-Shem et al., 2007).

Perturbations of L1 induced by mutation or changes in the interaction with the lipid membrane are transmitted to the catalytic site and can influence protein orientation and the location of the catalytic site (Figures 6 and 7). Furthermore, perturbation of TM5 affects the local structure and fluctuations of L1 (Figures 8C, 8D, 9A, 9C, and 9D) as well as the orientation of the protein (Figure 9B). Clusters of H bonds involving amino acids with significant conservation (TM2-E166 and TM3-K173) may be essential for the structure and dynamics of the protein and may assist the propagation of perturbations throughout the protein. This transmission of structural and dynamic changes between remote regions of the protein suggests an important role of L1 in regulating GlpG. It is possible that coupling between remote structural elements is essential for GlpG enzymatic activity (Figure 9E).

EXPERIMENTAL PROCEDURES

We used molecule B of the Ben-Shem et al. (2007) structure (PDB code 2IRV) for the starting coordinates. Amino acid residues R92, A272, and R273 were constructed with CHARMM software (Brooks et al., 1983). Missing hydrogen atoms were constructed by using the HBUILD facility of CHARMM. All titratable amino acid residues were considered in their standard protonation states. The tautomeric state of the His residues in the TM domain can be influenced by their local electrostatic environment (Baran et al., 2008). We tested the effect of the tautomeric state of H145 and H254 on GlpG dynamics and lipid interactions by comparing ~10 ns test simulations with the 2 His residues in either NE1 or in the ND1 tautomeric state. No significant effect of the tautomeric state could be observed. In the simulations discussed here, H145 and H254 are in the ND1 tautomeric state, and H141 and H150 are in NE1. The distance between N154-ND2 and H145-NE2 is 3.2 Å in the crystal structure and 3.3 ± 0.3 Å, 3.2 ± 0.3 Å, 3.1 ± 0.3 Å, and 3.3 ± 0.4 Å in Sim1, Sim2, Sim3, and Sim4, respectively.

The protein was aligned with its three principal axes along the x, y, z directions by using the VMD software (Humphrey et al., 1996) and was embedded in a solvated lipid bilayer so that its center of mass coincided with the center of the bilayer coordinates. Two independent sets of MD simulations of wild-type GlpG were performed with GlpG embedded in bilayers of 511 POPC (Sim1) and 531 POPE lipid molecules (Sim2). The coordinates of the phosphatidylglycerol head group of 1-palmitoyl-2-vaccenoyl-*sn*-glycerol-3-phosphatidylglycerol (PGV) close to the active site in the crystal structure (Ben-Shem et al., 2007) were used to construct a PGV lipid molecule in Sim1 and POPE in Sim2. We included 30,077 and 30,983 water molecules in Sim1 and Sim2, respectively. Chloride ions (one in Sim1; two in Sim2) were added for charge neutralization. The SHAKE algorithm (Ryckaert et al., 1977) was used to constrain the lengths of bonds involving hydrogen atoms. Coulomb interactions were computed by using the particle mesh Ewald summation (Darden et al., 1993). The short-range real-space interactions were cut off at 12 Å using a switching function between 8 Å and 12 Å. Temperature was kept constant by using Langevin dynamics; a Nose-Hoover Langevin piston (Feller et al., 1995; Martyna et al., 1994) was employed for pressure control (NPT, 1 bar and

300K). MD simulations were performed by using NAMD (Kalé et al., 1999; Phillips et al., 2005) with the CHARMM22 protein force field (MacKerell et al., 1998) and the CHARMM27 lipid force field (Feller and MacKerell, 2000). Waters were modeled by using TIP3P (Jorgensen et al., 1983). An integration step of 1 fs was used during the first ~2 ns of the simulation. After that, we used the reversible multiple time-step algorithm (Tuckerman and Berne, 1992) to integrate the equations of motion with time-steps of 1 fs for the bonded forces, 2 fs for the short-range nonbonded forces, and 4 fs for the long-range electrostatic forces.

During minimization, heating, and the first ~1 ns of equilibration, we used harmonic constraints of 5 kcal mol^{-1} for the protein atoms and 2 kcal mol^{-1} for ions, solvent water molecules, and lipid molecules farther than ≈15 Å from protein atoms. During the subsequent ~2 ns of equilibration, we used constraints of 2 kcal mol^{-1} for all groups, except for the lipids within ≈15 Å of the protein. All harmonic constraints were then switched off.

The triple-Ser (Y138S/F139S/L143S) and triple-Val (L229V/F232V/W236V) mutants were modeled by using a set of coordinates of wild-type GlpG equilibrated in POPE (Sim2) for ~10 ns. Sim3 and Sim4 were then performed according to the protocol described above for the wild-type. A test ~22 ns simulation on the Spitz GlpG substrate (Sim5) was performed by embedding the 23 residue TM region of Spitz in the center of a solvated lipid bilayer consisting of 280 POPC lipid molecules and 12,140 solvent water molecules. Sim1 and the Spitz-containing POPC bilayer were equilibrated at 300K; Sim2, Sim3, and Sim4 were equilibrated at 310K.

SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures, Supplemental References, and seven tables and can be found with this article online at http://www.cell.com/structure/supplemental/S0969-2126(09)00063-X.

ACKNOWLEDGMENTS

We thank Sin Urban, J. Alfredo Freites, and Douglas J. Tobias for valuable discussions; Charlie Zender, Steve Frank, and Robin Bush for their generous contribution of time on their clusters; and Joseph Farran of the University of California, Irvine High-Performance Computing Beowulf Cluster for excellent technical support. This research was supported in part by the National Institute of General Medical Sciences (GM-74637 to S.H.W.).

Received: August 14, 2008
Revised: December 21, 2008
Accepted: December 24, 2008
Published: March 10, 2009

REFERENCES

Baker, R.P., Young, K., Feng, L., Shi, Y., and Urban, S. (2007). Enzymatic analysis of a rhomboid intramembrane protease implicates transmembrane helix 5 as the lateral substrate gate. Proc. Natl. Acad. Sci. USA *104*, 8257–8262.

Baran, K.L., Chimenti, M.S., Schlessman, J.L., Fitch, C.A., Herbst, K.J., and Garcia-Moreno, B.E. (2008). Electrostatic effects in a network of polar and ionizable groups in *Staphylococcal nuclease*. J. Mol. Biol. *379*, 1045–1062.

Ben-Shem, A., Fass, D., and Bibi, E. (2007). Structural basis for intramembrane proteolysis by rhomboid serine proteases. Proc. Natl. Acad. Sci. USA *104*, 462–466.

Bier, E., Jan, L.Y., and Jan, Y.N. (1990). *rhomboid*, a gene required for dorsoventral axis establishment and peripheral nervous system developpent in *Drosophila melanogaster*. Genes Dev. *4*, 190–203.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics. J. Comput. Chem. *4*, 187–217.

Brown, M.S., and Goldstein, J.L. (1997). The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. Cell *89*, 331–340.

Clemmer, K.M., Sturgill, G.M., Veenstra, A., and Rather, P.N. (2006). Functional characterization of *Escherichia coli* GlpG and additional rhomboid proteins using an aarA mutant of *Providencia stuartii*. J. Bacteriol. *188*, 3415–3419.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: an *N*•log(*N*) method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089–10092.

Feller, S.E., and MacKerell, A.D., Jr. (2000). An improved empirical potential energy function for molecular simulations of phospholipids. J. Phys. Chem. B *104*, 7510–7515.

Feller, S.E., Zhang, Y., Pastor, R.W., and Brooks, B.R. (1995). Constant pressure molecular dynamics simulation: the Langevin piston method. J. Chem. Phys. *103*, 4613–4621.

Finn, R.D., Tate, J., Mistry, J., Coggill, P.C., Sammut, J.S., Hotz, H.-R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L.L., and Bateman, A. (2008). The Pfam protein families database. Nucleic Acids Res. *36*, D281–D288.

Freeman, M., Kimmel, B.E., and Rubin, G.M. (1992). Identifying targets of the *rough* homeobox gene of *Drosophila*: evidence that *rhomboid* functions in eye development. Development *116*, 335–346.

Hakizimana, P., Masureel, M., Gbaguidi, B., Ruysschaert, J.-M., and Govaerts, C. (2008). Interactions between phosphatidylethanolamine headgroup and LmrP, a multidrug transporter: a conserved mechanism for protein gradient sensing? J. Biol. Chem. *283*, 9369–9376.

Hanson, J.A., Duderstadt, K., Watkins, L.P., Bhattacharyya, S., Brokaw, J., Chu, J., and Yang, H. (2007). Illuminating the mechanistic roles of enzyme conformational dynamics. Proc. Natl. Acad. Sci. USA *104*, 18055–18060.

Humphrey, W., Dalke, W., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38.

Jensen, M.Ø., and Mouritsen, O.G. (2004). Lipids do influence protein function: the hydrophobic matching hypothesis revisited. Biochim. Biophys. Acta *1666*, 205–226.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. *79*, 926–935.

Kalé, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K., and Schulten, K. (1999). NAMD2: greater scalability for parallel molecular dynamics. J. Comput. Phys. *151*, 283–312.

Kalvodova, L., Kahya, N., Schwille, P., Ehehalt, R., Verkade, P., Drechsel, D., and Simons, K. (2005). Lipds as modulators of proteolytic activity of BACE: involvement of chlosterol, glyco-spingolipid, and ionic lipids in vitro. J. Biol. Chem. *280*, 36815–36823.

Koonin, E.V., Makarova, K.S., Rogozin, I.B., Davidovic, L., Letellier, M.-C., and Pellegrini, L. (2003). The rhomboids: a nearly ubiquitous family of intramembrane serine proteases that probably evolved by multiple ancient horizontal gene transfers. Genome Biol. *4*, R19.

Lee, A.G. (2004). How lipids affect the activities of integral membrane proteins. Biochim. Biophys. Acta *1666*, 62–87.

Lemberg, M.K., Menendez, J., Misik, A., Garcia, M., Koth, C.M., and Freeman, M. (2005). Mechanism of intramembrane proteolysis investigated with purified rhomboid proteases. EMBO J. *24*, 464–472.

Lupardus, P.J., Shen, A., Bogyo, M., and Garcia, K.C. (2008). Small molecule-induced allosteric activation of the *Vibrio cholerae* RTX cysteine protease domain. Science *322*, 265–268.

MacKerell, A.D., Jr., Bashford, D., Bellott, M., Dunbrack, R.L., Jr., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B *102*, 3586–3616.

Maegawa, S., Ito, K., and Akiyama, Y. (2005). Proteolytic action of G1pG, a rhomboid protease in the *escherichia coli* cytoplasmic membrane. Biochemistry *44*, 13543–13552.

Maegawa, S., Koide, K., Ito, K., and Akiyama, Y. (2007). The intramembrane active site of GlpG, an *E. coli* rhomboid protease, is accessible to water and hydrolyses an extramembrane peptide bond of substrates. Mol. Microbiol. *64*, 435–447.

Martyna, G.J., Tobias, D.J., and Klein, M.L. (1994). Constant-pressure molecular-dynamics algorithms. J. Chem. Phys. *101*, 4177–4189.

McQuibban, G.A., Saurya, S., and Freeman, M. (2003). Mitochondrial membrane remodelling regulated by a conserved rhomboid protease. Nature *423*, 537–541.

Narayanan, S., Sato, T., and Wolfe, M.S. (2007). A C-terminal region of signal peptide peptidase defines a functional domain for intramembrane aspartic protease catalysis. J. Biol. Chem. *282*, 20172–20179.

Phillips, J.C., Braun, B., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J. Comput. Chem. *26*, 1781–1802.

Rod, T.H., Radkiewicz, J.L., and Brooks, C.L., III. (2003). Correlated motion and the effect of distal mutations in dihydrofolate reductase. Proc. Natl. Acad. Sci. USA *100*, 6980–6985.

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J. Comput. Phys. *23*, 327–341.

Struhl, G., and Greenwald, I. (1999). Presenilin is required for activity and nuclear access to Notch in *Drosophila*. Nature *398*, 522–525.

Tuckerman, M., and Berne, B.J. (1992). Reversible multiple time scale molecular dynamics. J. Chem. Phys. *97*, 1990–2001.

Urban, S., and Wolfe, M.S. (2005). Reconstitution of intramembrane proteolysis in vitro reveals that pure rhomboid is sufficient for catalysis and specificity. Proc. Natl. Acad. Sci. USA *102*, 1883–1888.

Urban, S., Lee, J.R., and Freeman, M. (2001). *Drosophila* Rhomboid-1 defines a family of putative intramembrane serine proteases. Cell *107*, 173–182.

Urban, S., Schlieper, D., and Freeman, M. (2002). Conservation of intramembrane proteolytic activity and substrate specificity in prokaryotic and eukaryotic rhomboids. Curr. Biol. *12*, 1507–1512.

Wang, Y., and Ha, Y. (2007). Open-cap conformation of intramembrane protease GlpG. Proc. Natl. Acad. Sci. USA *104*, 2098–2102.

Wang, Y., Zhang, Y., and Ha, Y. (2006). Crystal structure of a rhomboid family intramembrane protease. Nature *444*, 179–183.

Wang, Y., Maegawa, S., Akiyama, Y., and Ha, Y. (2007). The role of L1 loop in the mechanism of rhomoboid intramembrane protease GlpG. J. Mol. Biol. *374*, 1104–1113.

Wasserman, J.D., Urban, S., and Freeman, M. (2000). A family of *rhomboid*-like genes: *Drosophila rhomboid-1* and *roughoid/rhomboid-3* cooperate to active EGF receptor signaling. Genes Dev. *14*, 1651–1663.

Weihofen, A., Binns, K., Lemberg, M.K., Ashman, K., and Martoglio, B. (2002). Identification of signal peptide peptidase, a presenilin-type aspartic protease. Science *296*, 2215–2218.

White, S.H. (2006). Rhomboid intramembrane protease structures galore! Nat. Struct. Mol. Biol. *13*, 1049–1051.

Wiener, M.C., and White, S.H. (1992). Structure of a fluid dioleoylphosphatidylcholine bilayer determined by joint refinement of x-ray and neutron diffraction data. III. Complete structure. Biophys. J. *61*, 434–447.

Wolfe, M.S., and Kopan, R. (2004). Intramembrane proteolysis: theme and variations. Science *305*, 1119–1123.

Wu, Z., Yan, N., Feng, L., Oberstein, A., Yan, H., Baker, R.P., Gu, L., Jeffrey, P.D., Urban, S., and Shi, Y. (2006). Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry. Nat. Struct. Mol. Biol. *13*, 1084–1091.