The crystal structure of the rhomboid peptidase from *Haemophilus influenzae* provides insight into intramembrane proteolysis

M. Joanne Lemieux*, Sarah J. Fischer, Maia M. Cherney, Katherine S. Bateman, and Michael N. G. James*

Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, AB, Canada T6G 2H7

Communicated by Robert M. Stroud, University of California, San Francisco, CA, November 10, 2006 (received for review October 24, 2006)

Rhomboid peptidases are members of a family of regulated intramembrane peptidases that cleave the transmembrane segments of integral membrane proteins. Rhomboid peptidases have been shown to play a major role in developmental processes in *Drosophila* and in mitochondrial maintenance in yeast. Most recently, the function of rhomboid peptidases has been directly linked to apoptosis. We have solved the structure of the rhomboid peptidase from *Haemophilus influenzae* (hiGlpG) to 2.2-Å resolution. The phasing for the crystals of hiGlpG was provided mainly by molecular replacement, by using the coordinates of the *Escherichia coli* rhomboid (ecGlpG). The structural results on these rhomboid peptidases have allowed us to speculate on the catalytic mechanism of substrate cleavage in a membranous environment. We have identified the relative disposition of the nucleophilic serine to the general base/acid function of the conserved histidine. Modeling a tetrapeptide substrate in the context of the rhomboid structure reveals an oxyanion hole comprising the side chain of a second conserved histidine and the main-chain NH of the nucleophilic serine residue. In both hiGlpG and ecGlpG structures, a water molecule occupies this oxyanion hole.

intramembrane peptidase | membrane protein | rhomboid protease | x-ray crystallography

Rhomboid peptidases belong to family S54 of intramembrane serine peptidases (1); these enzymes carry out proteolysis of single transmembrane substrates within the environment of the lipid bilayer. Rhomboid peptidases cleave their substrates in the outer leaflet of the lipid bilayer, thereby releasing an exocellular peptide signal that can, in turn, play a role in cell signaling (for review, see ref. 2). Although the catalytic mechanism for soluble serine peptidases has been well characterized, both biochemically and structurally, we are just beginning to understand the mechanism of serine peptidases in the membrane environment (3).

Although the rhomboids are a newly discovered family of peptidases, they have been identified in all kingdoms (4, 5). Their diverse functions are being revealed through genetic screens and developmental and cell biology studies (6). Rhomboid peptidases have been shown to play a role in releasing EGF from the cell surface at its C terminus. Both molecules pack between two hiGlpG molecules (Fig. 1 A and B). The first C12Es interacts with helices H3 and H6 on one of the hiGlpG molecules and with H6 on the other. The second C12Es interacts with helices H1 and H2 on one hiGlpG.
molecule and with H2 on the other. There are three lipid molecules visible in the electron density map that we have presently interpreted as phosphatidic acid (PA); the phosphoryl group has significantly higher electron density than the other atoms in the glycerol backbone and acyl chains. (Figs. 1A and B and 2). PA has been shown to provide a source of the signaling lipid diacylglycerol (18). It is possible that these lipids may be cardiolipin (CL), phosphatidylserine (PS) or phosphatidylethanolamine (PE), the main lipids found in the E. coli lipid bilayer. Two PA molecules can be seen flanking the L1 loop, a loop that is proposed to be flexible in order for substrate to bind (see below and ref. 3). Another PA is located near H6. It is unclear at present whether these lipids merely play a structural role, acting as chaperones, or whether they play a role in rhomboid function.

Comparison of hiGlpG and ecGlpG Structures. Superimposition of the Ca atoms of the hiGlpG and ecGlpG structures reveals that hiGlpG and ecGlpG share a common fold as predicted, with an rmsd of 1.09 Å for 137 equivalent Ca atom pairs. An overlay of the structures demonstrates this agreement (Fig. 3). The only areas that appear to have variability are the L1 loop, located between H1 and H2, and the H5 helix. In fact, we see relatively weak electron density for H5, suggesting some conformational variability for this helix. On the other hand, we see density for the C terminus of the molecule that has allowed us to build H6 further into the cytoplasm than that for ecGlpG. The C terminus of hiGlpG actually extends beyond the lipid bilayer boundary, similar to that seen with the GlpT structure (19).

Rhomboid Active Site. Most serine peptidases are characterized by a catalytic triad consisting of the nucleophilic O atom of the serine, a general acid/base (the imidazole ring of a histidine) that assists in the deprotonation of the nucleophile and an aspartic acid carboxylate that helps to maintain an optimal position for the imidazole ring of the histidine during catalysis. Importantly, and in addition to these groups, there is an oxyanion hole, a feature that provides electrophilic assistance to the nucleophilic attack by the serine O on the carbonyl carbon atom of the scissile bond. Rhomboid has these features as well, but the molecular details are different.

In the native unbound form of rhomboid, the residues comprising the catalytic triad and the oxyanion hole are buried and inaccessible to solvent (Fig. 4). Three loops, L1, L3 and L5, have to move substantially in order that the substrate can gain access to the active site (3). The B-factors for these loops are high, ranging from 60 Å² to 70 Å² for L1, 50 Å² for L3, and 60 Å² for L5, supporting this hypothesis. The overall B-factor for hiGlpG is 35.5 Å². A possible sequence of events could include: the destabilized area of the transmembrane helix of the substrate (20) docks to rhomboid displacing the L1 gate (Gly-29 to Ser-55 in hiGlpG). The L3 loop (Gly-109 to Gly-114) and the L5 loop (Gly-161 to Gly-165) are then displaced. The beginning and ending residues of these loops are defined principally by glycine residues that have large associated conformational flexibility, which facilitates moving these three loops to provide access for the substrate to the active site (Fig. 4).

Discussion

Rhomboid Mechanism. We have explored the binding of a short segment of a substrate polypeptide, spitz, by superimposing the side chain of His-169 and Oy of Ser-116 from hiGlpG onto the structure of chymotrypsin complexed to the turkey ovomucoid third domain, OMTKY3 [1CHO.pdb] (21). This superposition, although not very accurate, gives a reasonable position for the segment of substrate polypeptide from P2 to P1 [nomenclature of Schechter and Berger...
fitting into the rhomboid active site (Fig. 4). The substrate has been modeled as four residues surrounding the cleavage site of spitz, Ala-Ser-Gly-Ala (20). The cleavage specificities of several rhomboids indicate a preference for residues, having small side chains, such as glycine and alanine (3).

The water molecule in our structure, W40 (Fig. 4), which bridges Ser116NH and His65N\(^{-2}\)H and forms H bonds to Phe113O and Val162O, is located in the oxyanion hole; this water molecule would be displaced by the carbonyl-oxygen atom of the P1 residue of the substrate, thus placing the P1 carbonyl-carbon atom in an ideal position for nucleophilic attack by Ser116O\(^{-}\). The nucleophilic attack is facilitated by the transfer of the proton on O\(^{-}\) to the N\(^{+}\) atom of the imidazole ring of His-169. A good hydrogen bond from Ser116O\(^{-}\)...N\(^{+}\) His-169 is already established in the native enzyme, thereby defining the pathway for this proton transfer. The oxyanion hole in rhomboid hiGlpG comprises the main-chain NH of Ser-116 and the protonated N\(^{+}\) of the imidazole of His-65. This favorable tautomeric form for the imidazole ring of His-65 is ensured by the main-chain hydrogen bond from Leu61NH to N\(^{+}\)1 of His-65 (Fig. 4). Not only does this oxyanion hole stabilize the developing negative charge on the carbonyl-oxygen atom in the tetrahedral intermediate, but it also assists the nucleophilic attack by electrostatically enhancing the polarization of the carbonyl (C—O) bond of the substrate. Mutagenesis experiments have demonstrated that, in ecGlpG and YqgP, the residues equivalent to His-169 and Ser-116 from hiGlpG are essential for activity (23). In addition, mutagenesis in YqgP of the residue equivalent to His 65 in hiGlpG, which is also highly conserved, resulted in markedly reduced activity.

In hiGlpG there is no direct equivalent to the third member of the catalytic triad, i.e., an AspCOO\(^{-}\), or an AsnCONH\(_2\). In ecGlpG, the side-chain amide of Asn-251 forms a hydrogen bond to a water molecule, and the water, in turn, forms a hydrogen bond to N\(^{+}\) of His-254, the general base in that enzyme. hiGlpG has no equivalent water molecule, and the close contact of the imidazole ring of His-169 to Phe-155 likely precludes the binding of water or the side chain of Asn-166. A more likely candidate for a third member of the catalytic triad is the packing of the side chain of Tyr-120 against the imidazole ring of His-169, thereby ensuring some stability in the position of the general base (3).

It has not escaped our attention that the nucleophilic Ser-116 is at the beginning of helix H4 and that the NH of Ser-116 would form a good hydrogen bond to the oxyanion of the tetrahedral intermediate. Therefore, additional stabilization of the oxyanion will come from the partial positive charge on the helix dipole of helix H4. This situation is analogous to the peptide dipole stabilization provided by the helix bearing the Ser-221 in bacterial subtilisin (24). Subtilisin has the side-chain amide from Asn-155 that contributes to the formation of the oxyanion hole. Rhomboid is similar to subtilisin, but it uses a histidine side chain rather than an asparagine carboxamide.

We have solved the structure of a full-length rhomboid peptidase, namely hiGlpG. The superimposed structures of hiGlpG and ecGlpG are similar; however, their active sites have subtle differences and our structure has allowed us to identify the oxyanion hole, an essential feature for serine peptidases that is required to propose confidently a sound catalytic mechanism. Perhaps the most unexpected result of our structural analysis is to find the catalytic machinery of a soluble globular serine and the general peptidase like chymotrypsin or subtilisin below the surface of the bilayer. It is also surprising to find a catalytic site in hiGlpG resembling that of bacterial subtilisin with the nucleophilic serine and general base histidine residues near the N termini of two separate \(\alpha\)-helices. Similarly, the oxyanion hole is formed by the side chain of an asparagine in subtilisin. There is clearly an evolutionary relationship between these two families of serine peptidases.

**Materials and Methods**

**Cloning and Expression.** *H. influenzae* DNA was purchased from American Type Culture Collection (Manassas, VA). With the use of restriction digestion, PCR products were then ligated.
into pBAD-MycHisA vector (Invitrogen, Burlington, Canada). Expression was carried out in Top10 cells (Invitrogen) in Luria–Bertani medium supplemented with ampicillin. Induction of the various expression constructs were carried out as described below with arabinose.

**Membrane Fraction Isolation.** Cells were grown to an OD$_{600}$ of 0.4 and induced with 0.0002% arabinose at 24°C for 16 h. Cells were harvested at 12,227 × g for 10 min by using an AvantiJ1.8000 rotor (Beckman, Fullerton, CA). Cells were resuspended in four volumes of TBS supplemented with an EDTA-free peptidase-inhibitor mixture (NEB, Beverly, MA), 1 mM PMSF, and 0.1 mg/ml DNase and lysed by using a TEmulsiFlex-C3T, (Avestin, Ottawa, Canada). Unbroken cells were pelleted in a JA17 rotor at 10,000 × g for 20 min. Membrane fractions were collected by ultracentrifugation in a L8–80 ultracentrifuge at 100,000 × g in a 45Ti rotor (Beckman).

**Protein Purification.** Membrane fractions were homogenized in 50 mM Tris, 300 mM NaCl, 30 mM imidazole, 20% glycerol, and 1% DDM (pH 8.0). The solution was stirred for 30 min, followed by ultracentrifugation for 30 min at 110,000 × g in a 45Ti rotor.

![Fig. 4. Mechanism of H. influenzae GlpG.](image)

**(A)** Top view from the periplasmic space of hiGlpG. Residues in the active site are labeled in gray. Water (WAT40) located in the active site between Ser-116 and His-65 is shown in red. A model of the D. melanogaster rhomboid substrate spitz (magenta) has been docked manually into the active site. (C) Wall-eyed stereoview of spitz docked into the active site of hiGlpG. The hydroxyl oxygen of S116 is hydrogen bonding with H169. H169 is stabilized by Y120. The carbonyl oxygen from the P1 residue is stabilized by H65 and the backbone amide of S116, forming the oxyanion hole. (D) Proposed catalytic mechanism of hiGlpG (ChemDraw).
The monoclinic data diffracted to higher resolution and space groups were obtained: monoclinic C2 and orthorhombic at the Advanced Light Source beam line 8.3.1. Two different were directly flash-cooled in liquid nitrogen. Data were collected

Tris (pH 8.0), 0.05% C12E8, 20 mM NaCl, and 10% glycerol (see hadex200 (16/60) column (Amersham, Piscataway, NJ) in 50 mM

ford, MA) and subjected to detergent exchange on a Spe-

concentrated by using a 30K centrifugal filter (Millipore, Bed-

collected and washed with 20 column volumes (CV) of 50 mM

(Beckman, USA). The supernatant was incubated with Ni-NTA resin (Qiagen, Ontario, Canada) for 2 h. The resin was then collected and washed with 20 column volumes (CV) of 50 mM Tris, 300 mM NaCl, 30 mM imidazole, 20% glycerol, and 0.1% DDM (pH 8.0), followed by 20 CV of the above stated buffer with 35 mM imidazole. Protein fractions were eluted in a step-wise manner with 3

Crystals of hiGlpg were obtained in 25% PEG 4000, 0.1 M citrate (pH 6.0), 1 M NaCl, 3% ethanol, and 15% glycerol and grew to

kDa molecular mass cutoff to a concentration of 5 mg/ml. Crystals of hiGlpg were obtained in 25% PEG 4000, 0.1 M citrate (pH 6.0), 1 M NaCl, 3% ethanol, and 15% glycerol and grew to dimensions of 100 μm × 50 μm × 20 μm in 1–2 weeks. Crystals were directly flash-cooled in liquid nitrogen. Data were collected at the Advanced Light Source beam line 8.3.1. Two different space groups were obtained: monoclinic C2 and orthorhombic P212121. The monoclinic data diffracted to higher resolution and had one molecule in the AU. Molecular replacement was carried out by using MolRep (25) with the ecGlpG coordinates. Refinement was carried out with Refmac5 (26). See SI Table 1 for statistical results.

We thank Dr. Ya Ha for generously making available the E. coli GlpG coordinates (2IC8.pdb); Drs. James Holden (ALS, BL 8.3.1), Ernst Bergmann, and Jonathan Parrish for assistance with data collection; all members of the M.N.G.J. laboratory for their support; and Wendy Kasinec for administrative assistance. X-ray diffraction data were collected at beam line 8.3.1 of the Advanced Light Source (ALS) at the Lawrence Berkeley National Laboratory, Berkeley, CA, under an agreement with the Alberta Synchrotron Institute (ASI). The ALS is operated by the Department of Energy and supported by the National Institutes of Health (Bethesda, MD). Beam line 8.3.1 was funded by the National Science Foundation, the University of California, and Henry Wheeler. The ASI synchrotron access program is supported by grants from the Alberta Science and Research Authority (ASRA) and the Alberta Heritage Foundation for Medical Research (AHFMR). This work has been supported by the Canadian Institute for Health Research (CIHR). M.J.L is supported by a fellowship scholarships from CHIR, CIHR strategic training program for membrane proteins and cardiovascular disease, and Alberta Heritage Foundation for Medical Research. M.N.G.J. acknowledges support from the Canada Research Chair’s Program.

Crystallographic Analysis. Protein for crystallization was concentrated by using Milipore Ultrafree centrifugal concentrators, 30 kDa molecular mass cutoff to a concentration of 5 mg/ml. Crystals of hiGlpg were obtained in 25% PEG 4000, 0.1 M citrate (pH 6.0), 1 M NaCl, 3% ethanol, and 15% glycerol and grew to


PNAS