2009

Structural determinants of autoproteolysis of the Haemophilus influenzae Hap autotransporter

Roma Kenjale  
*Duke University*

Guoyu Meng  
*Duke University*

Doran L. Fink  
*Washington University School of Medicine in St. Louis*

Twyla Juehne  
*Washington University School of Medicine in St. Louis*

Tomoo Ohashi  
*Duke University*

*See next page for additional authors*

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

**Recommended Citation**

https://digitalcommons.wustl.edu/open_access_pubs/2505

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Structural Determinants of Autoproteolysis of the *Haemophilus influenzae* Hap Autotransporter

Roma Kenjale, Guoyu Meng, Doran L. Fink, Twyla Juehne, Tomoo Ohashi, Harold P. Erickson, Gabriel Waksman and Joseph W. St. Geme III


Updated information and services can be found at: http://iai.asm.org/content/77/11/4704

**REFERENCES**

These include:

This article cites 32 articles, 15 of which can be accessed free at: http://iai.asm.org/content/77/11/4704#ref-list-1

**CONTENT ALERTS**

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to another ASM Journal go to: http://journals.asm.org/site/subscriptions/
Structural Determinants of Autoproteolysis of the Haemophilus influenzae Hap Autotransporter

Roma Kenjale,1,2 Guoyu Meng,4 Doran L. Fink,5† Twyla Juehne,5‡ Tomoo Ohashi,3 Harold P. Erickson,3 Gabriel Waksman,4 and Joseph W. St. Geme III1,2,*

Department of Pediatrics,1 Department of Molecular Genetics and Microbiology,2 and Department of Cell Biology,3 Duke University, Durham, North Carolina; Institute of Structural and Molecular Biology at UCL/Birkbeck, Malet Street, London, United Kingdom;4 and Department of Pediatrics, Washington University, St. Louis, Missouri5

Received 28 May 2009/Returned for modification 7 July 2009/Accepted 6 August 2009

Haemophilus influenzae is a gram-negative bacterium that initiates infection by colonizing the upper respiratory tract. The H. influenzae Hap autotransporter protein mediates adherence, invasion, and microcolony formation in assays with respiratory epithelial cells and presumably facilitates colonization. The serine protease activity of Hap is associated with autoproteolytic cleavage and extracellular release of the HapS passenger domain, leaving the HapS C-terminal domain embedded in the outer membrane. Cleavage occurs most efficiently at the LN1036-37 peptide bond and to a lesser extent at three other sites. In this study, we utilized site-directed mutagenesis, homology modeling, and assays with a peptide library to characterize the structural determinants of Hap proteolytic activity and cleavage specificity. In addition, we used homology modeling to predict the S1, S2, and S4 subsite residues of the Hap substrate groove. Our results indicate that the P1 and P2 positions at the Hap cleavage sites are critical for cleavage, with leucine preferred over larger hydrophobic residues or other amino acids in these positions. The substrate groove is formed by L263 and N274 at the S1 subsite, R264 at the S2 subsite, and E265 at the S4 subsite. This information may facilitate design of approaches to block Hap activity and interfere with H. influenzae colonization.

Haemophilus influenzae is a gram-negative coccobacillus that typically colonizes the nasopharynxes of children and adults. In addition, this organism is an important cause of localized respiratory tract and invasive disease. Nonencapsulated strains cause otitis media, sinusitis, conjunctivitis, and exacerbations of respiratory symptoms in individuals with underlying lung disease, bronchiectasis, and cystic fibrosis (21, 29). Encapsulated strains are an important cause of bacteremic diseases, such as meningitis (29). Colonization of the upper respiratory tract represents an early step in the pathogenesis of all Haemophilus disease and requires adherence to respiratory epithelium (19). Adherence is facilitated by a number of adhesins, including Hap, Hia, Hsf, HMW1/HMW2, P5, pili, and lipooligosaccharide (2, 18, 21, 26, 27).

The Hap adhesin is ubiquitous among isolates of H. influenzae and is a member of the autotransporter family of virulence factors that have been recognized among many gram-negative bacteria (10). Autotransporters are synthesized as precursor proteins with three functional regions, namely, an N-terminal signal sequence, an internal passenger domain, and a C-terminal β-barrel domain (11). The signal sequence targets the precursor protein to the inner membrane and is then cleaved. The C-terminal β-barrel domain inserts into the outer membrane and facilitates presentation of the passenger domain on the bacterial cell surface. Depending upon the protein, the passenger domain remains covalently attached to the β-barrel domain, is cleaved but remains loosely attached to the β-barrel domain, or is cleaved and released entirely from the cell surface (10–12). Although diverse autotransporters share a similar structural organization and a common secretion mechanism, they vary widely in function, possibly reflecting adaptations to particular bacterial pathogenic niches. Autotransporters may function as adhesins mediating tissue tropism, as proteases involved in tissue degradation, as toxins causing host tissue damage, or as mediators of serum resistance (11).

Hap is synthesized as a 155-kDa preprotein encompassing a 110-kDa passenger domain, HapS, and a 45-kDa β-barrel domain, HapB. The HapS passenger domain harbors adhesive activity that has been shown to promote interactions with human respiratory cells, as well as with extracellular matrix proteins such as fibronectin, laminin, and collagen IV (7). HapS is also responsible for bacterial aggregation via Hap-Hap interactions, contributing to microcolony formation (5). Adherence to epithelial cells and bacterial aggregation are mediated by the C-terminal 311 amino acids of HapS, whereas interaction with extracellular matrix proteins is mediated by the C-terminal 511 amino acids of HapS (7).

Beyond possessing adhesive activities, the HapS passenger domain functions as a protease that directs the autoproteolytic cleavage of HapS from HapB, resulting in release of HapS from the bacterial cell surface (6). Hap autoproteolysis has been determined to occur at least partly through intermolecular
cleavage on the surface of the bacterium and involves a catalytic triad consisting of residues His98, Asp140, and Ser243. Ser243 is part of the GDGS motif that is characteristic of many serine proteases (6, 13). In wild-type Hap, cleavage occurs most abundantly at the L1036-N1037 peptide bond, which is referred to as the primary cleavage site (13). Site-directed mutagenesis of this site and N-terminal sequencing of the resulting cleaved Hap fragments has identified three additional cleavage sites, including L1046-T1047, F1077-A1078, and F1067-S1068, termed the secondary, tertiary, and quaternary cleavage sites, respectively (see Table 2) (6). Alignment of the amino acid sequences at these cleavage sites has revealed a consensus target sequence motif that consists of (Q/R)(A/S)X(L/F) at the P4 through P1 positions (see Table 2) (6).

Hap protease activity can be blocked by selected serine protease inhibitors, including secretory leukocyte protease inhibitor, a component of human respiratory secretions (14). This inhibition results in accumulation of full-length Hap in the outer membrane, with HapS on the bacterial surface. Similar to the effect of serine protease inhibitors, mutation of the active site serine in the Hap protease domain causes retention of HapS on the bacterial surface, increased adherence to epithelial cells, increased adherence to extracellular matrix proteins, increased bacterial aggregation, and increased microcolony formation, suggesting that Hap protease activity might play a regulatory role in bacterial adherence. Indeed, autoproteolytic cleavage of Hap results in reduced adhesive activity (5).

In the present study, we sought to further elucidate the structural determinants of Hap proteolytic activity. We used site-directed mutagenesis and in vitro assays with a synthetic peptide library to characterize Hap cleavage site specificity, and we used homology modeling and site-directed mutagenesis to define the Hap substrate groove. We found that the two residues immediately N-terminal to Hap cleavage sites are critical for cleavage, with leucine preferred over larger hydrophobic residues or other amino acids in these positions. In addition, we defined the residues that likely form the Hap substrate groove.

### MATERIALS AND METHODS

#### Bacterial strains, plasmids, and growth conditions

The bacterial strains and plasmids used in the present study are listed in Table 1. Plasmid constructs encoding hap were expressed in H. influenzae strain DB117, a rec-1 derivative of
strain Rd that contains a nonfunctional hap gene due to a spontaneous nonsense mutation in hap. H. influenzae strains were stored at −80°C in brain heart infusion (BHI) broth with 20% glycerol, grown overnight on supplemented BHI agar with appropriate antibiotics, and cultivated in BHI broth as described previously (1). Escherichia coli strain DH5α was grown on Luria-Bertani (LB) agar or in LB broth and was maintained at −80°C in LB broth with 20% glycerol. Tetracycline was used at concentrations of 5 μg/ml (H. influenzae) and 12.5 μg/ml (E. coli).

Purification of Hap protein. Hap was purified from culture supernatants of H. influenzae strain DB117/pJS106 as described previously (13).

Glycerol gradient sedimentation and electron microscopy. The sedimentation coefficient of Hap was estimated by glycerol gradient sedimentation. The purified protein was sedimented at 20°C through a 15 to 40% glycerol gradient in 0.2 M ammonium bicarbonate at 38,000 rpm for 16 h in a Beckman SW55.1 rotor (23). The glycerol gradients were calibrated with standard proteins of known S values (catalase, 11.3 S; aldolase, 7.3 S; bovine serum albumin, 4.6 S; and ovalbumin, 3.5 S). For rotary shadowing, samples from the glycerol gradient fractions were sprayed onto freshly cleaved mica, vacuum dried, and rotary shadowed with platinum (8, 23). For negative staining, purified Hap was applied to carbon-coated copper grids and stained with 2% uranyl acetate (8, 23).

Construction of mutant Hap derivatives. The plasmid pMLD100 is a pUC19 derivative that contains a 6.7-kb insert with the wild-type hap gene and was used as a template for PCR. Site-directed mutagenesis was performed using a QuikChange XL II kit from Stratagene (La Jolla, CA) according to the manufacturer’s instructions. The mutations were confirmed by sequencing. After mutagenesis, the 6.7-kb PsI fragment (Roche, Indianapolis, IN) fragment was excised from pMLD100 and ligated into PsI-digested pGJB103. The resulting plasmids were introduced into the H. influenzae strain DB117 by MII/MIV transformation (25).

Analysis of outer membrane proteins. H. influenzae strain DB117 derivatives expressing wild-type Hap or mutant Hap were grown in BHI broth with tetra-cycline to an optical density at 600 nm of 0.8. Sarkosyl insoluble outer membranes were prepared as described previously (4). Outer membrane proteins were resuspended in a small volume of 10 mM HEPES (pH 7.4), resolved on a 10% sodium dodecyl sulfate-polyacrylamide gel, and electrophoresed onto a nitrocellulose membrane. Immunoblots were probed with antiserum Rab290, which was raised against a C-terminal Hap fragment corresponding to residues 996 to 1395 and detects the membrane-bound Hap precursors, HapP, and HapP. Anti-rabbit immunoglobulin G conjugated to alkaline phosphatase (Sigma Chemical Co., St. Louis, MO) was used as a secondary antibody.

Peptide library design and purification. A 12-mer C-terminally biotinylated peptide library with the sequence NH2-MAXXXXXXALEK(biotin)-COOH was synthesized at the Tufts University Core Facility. In this peptide library, the P* residues were fixed, based on the Hap P1–P5 sequence at the primary cleavage site (Table 2). An “X” denotes random amino acids at the P1 to P5 positions that were synthesized using an inorganic mixture of 19 amino acids (except cysteine). The peptide library was resuspended in phosphate-buffered saline (PBS; pH 7.4) at a concentration of 3.4 mg/ml. In order to ensure that the peptide library was homogenously biotinylated, 0.4 mg of the resuspended library was purified over monomeric avidin agarose (Pierce, Rockford, IL) as described previously (28). The resulting purified peptide library was lyophilized and resuspended in 100 μl of assay buffer composed of 50 mM sodium phosphate (pH 7.4).

Peptide library cleavage assay. In order to determine the P residue specificity of Hap, 15 μl of the purified peptide library was incubated at 37°C overnight with 8 pmol of purified HapP in a 20-μl reaction mixture. Control reactions contained either substrate only (without HapP) or HapP only (without peptide). After incubation, the samples were heated to 100°C for 5 min to stop the cleavage reaction. The sample with HapP and peptide was incubated with 0.6 μl of avidin agarose (Sigma) and rotated at room temperature for 1 h to remove the C-terminal fragments of the cleaved products, as well as uncleaved peptides. The samples were subjected to repeated lyophilization, followed by resuspension in water, finally resuspended in 10 μl of water, and then subjected to N-terminal sequencing. The peptide fragments recovered after avidin treatment were sequenced for one cycle beyond the intended P1 position to test for the removal of unreacted peptides. To determine bias present in the peptide library itself, the enzyme-free reaction without avidin treatment was sequenced in the same manner. Raw data from the sequencing reactions were obtained as molar quantities of each amino acid present in each sequencing cycle. Corrected experimental data were obtained by dividing the molar percent abundance of each residue in any given sequencing cycle by the molar percent abundance of that residue in the starting peptide library. These data were normalized to an average value of 1, resulting in a selective value greater than 1 for residues that were positively selected and a value less than 1 for residues that were negatively selected.

RESULTS

Electron microscopy of purified HapP. In order to study the architecture of HapP, we began by purifying HapP from culture supernatants of H. influenzae strain DB117 expressing wild-type Hap. As shown in Fig. 1A, the sedimentation coefficient of purified HapP, as estimated by glycerol gradient sedimentation, was found to be 5.5 S. This S value of HapP, together with its known molecular weight, gives rise to an Smax/S value of 1.53, indicating that HapP has an elongated shape (Fig. 1A) (23). To further investigate the structure of HapP, we performed rotary shadowing electron microscopy. As shown in Fig. 1B, this technique revealed structures that were ~15 nm in length and resembled very short arrowheads (Fig. 1B, middle panel). Similar arrowhead-shaped structures were also visible with negative staining transmission electron microscopy (Fig. 1B, left panels). The arrowhead shape of HapP resembled the crystal structure of E. coli hemoglobin protease (Hbp, Fig. 1B, right panel), a Hap homolog with heme-binding and serine protease activities that is expressed by clinical isolates associated with intra-abdominal infections (20, 30). Alignment of the amino acid sequences of HapP and the protease domain of Hbp revealed an identity of 41% and a similarity of 54%, as highlighted in Fig. 2A.

Homology based modeling of the HapP structure. Sequence alignment using the Pfam server (http://pfam.sanger.ac.uk) suggested that HapP consists of an N-terminal domain (residues 70 to 550) sharing 41% sequence identity with Hbp and a C-terminal domain (residues 851 to 980) based on the crystal structure of Pertactin (PDB code 1WXR) and residues 851 to 980 based on the crystal structure of Pertactin (PDB code 1DAB) using Swiss-model (http://swissmodel.expasy.org) and 3Dijigsaw (http://bmm.cancerresearchuk.org) homology modeling servers (Fig. 2B). Based on the Hbp-like shape of HapP observed by electron microscopy, we modeled the region between the protease and Pertactin-like domains on the crystal structure of Hbp, assuming that HapP51–850 shares the same β-helix fold as observed in the Hbp crystal structure and other
conventional autotransporters (31). Hap551-850 was predicted by the PredictProtein server to be a domain consisting only of β strands (http://www.predictprotein.org/).

Using the homology model of Hap551-850, we predicted the Hap substrate groove. First, we aligned Hap63-287 with the published crystal structures of alpha-chymotrypsin (PDB code 2CHA) and beta-trypsin (PDB code 5PTP). The structures of Hap63-287, alpha-chymotrypsin, and beta-trypsin are very similar, with a root mean square deviation in Cα positions of 1.0 Å between Hap63-287 and alpha-chymotrypsin and 2.2 Å between Hap63-287 and beta-trypsin (Fig. 3A). Next, we inserted the inhibitor p-sulfinotoluene in these aligned structures to highlight the catalytic triad and to predict the S subsite residues of HapS (Fig. 3A). The coordinates of inhibitor p-sulfinotoluene were extracted from the superimposed alpha-chymotrypsin structure (2CHA) and merged into the homology model of Hap63-287. By nomenclature, binding sites on a protease are called subsites (S residues), and each subsite residue interacts with a residue of the substrate (called P residues) (22). The predicted S1 subsite contains S243, N274, L263, and K240; the predicted S2 subsite contains H98, R264, and Y137; and the predicted S4 subsite contains E265 (Fig. 3B).

Mutagenesis analysis of the predicted substrate groove. To assess the accuracy of the predicted substrate groove, we performed site-directed mutagenesis, changing each of the predicted subsite residues individually and assessing the effect of each mutation on Hap autoproteolytic activity (Table 3). Wild-type residues were changed to amino acids that differ in terms of polarity, charge, or length of side chain. For the predicted S1 residues, we changed K240 from a charged amino acid to a nonpolar amino acid (K240A), L263 from a nonpolar amino acid to a charged polar amino acid (L263R), and N274 from an uncharged polar amino acid to a charged polar amino acid (N274R). For the predicted S2 subsite, we changed Y137 from an aromatic amino acid to a nonpolar amino acid (Y137A) and R264 from a charged polar amino acid to a nonpolar amino acid (R264A). For the predicted S4 subsite, we changed E265 from a charged polar amino acid to nonpolar amino acid (E265A) or an aromatic amino acid (E265W). In Western blot analysis of outer membrane preparations from strain DB117 expressing wild-type Hap, a 45-kDa band corresponding to Hap was apparent, representing cleavage at the primary cleavage site (Fig. 4A). By comparison, the N274R and L263R mutations at the predicted S1 subsite disrupted autoproteolysis, resulting in accumulation of the 155-kDa Hap precursor (Fig. 4A). Similarly, the R264A mutation at the predicted S2 subsite and the E265W mutation at the predicted S4 subsite also disrupted autoproteolysis. The E265A mutation had only a modest effect on autoproteolysis, as evidenced by partial accumulation of Hap precursor in the outer membrane. In contrast, the Y137A and K240A mutations had no effect on autoproteolysis, suggesting that Y137 and K240 may not be essential for mediating interactions between Hap and its substrate (Fig. 4A). Taken together, these mutagenesis data provide further support that residues N274, L263, R264, and E265 may indeed form the substrate binding groove in Hap predicted by our homology model (Fig. 3B).

Analysis of Hap cleavage specificity. In earlier studies we found that mutation of the P1 residue of the Hap cleavage site disrupted cleavage (13). To extend these studies, we studied Hap

![FIG. 1. (A) Glycerol gradient sedimentation of Hap. Hap sedimented at 5.5 S, with an Smax/S of 1.53. The positions of standard proteins are indicated by arrows from the left as follows: aldolase (7.3 S), bovine serum albumin (BSA) (4.6 S), and ovalbumin (3.5 S). SM, starting material. (B) Electron microscopy of purified Hap. Negative staining electron microscopy (upper and lower left panels) and rotary shadow electron microscopy images (middle panel) of purified Hap revealed an elongated arrowlike shape, which resembled the crystal structure of E. coli Hbp (right panel). Note that Hap in the rotary shadowed images is slightly larger than the negative stain images because of the platinum shell. The right panel shows the space-filling model and the overlaid ribbon model of the crystal structure of E. coli Hbp (PDB 1WXR), demonstrating an arrowlike shape. The serine protease domain is highlighted in red.](http://iai.asm.org/)

---

VOL. 77, 2009 Hap AUTOPROTEOLYSIS 4707

Downloaded from http://iai.asm.org on March 26, 2014 by Washington University in St. Louis
cleavage specificity by systematically mutating the P1 to P4 residues at cleavage sites and analyzing the effect of these mutations on autoproteolysis (Table 3). The P1 residue at the primary (L1036-N1037) and secondary (L1046-T1047) sites is a leucine. In contrast, the P1 residue at the tertiary (F1077-A1078) and quaternary (F1067-S1068) sites is a phenylalanine (Table 2). To address the possibility that leucine at the P1 position is preferred over the bulkier phenylalanine, leucine in the P1 position at the primary cleavage site was mutated to phenylalanine (L1036F). In addition, the bulky phenylalanine in the P1 position at both the tertiary and quaternary cleavage sites was mutated to leucine (F1077L and F1067L, respectively). As shown in Fig. 4B, the previously characterized L1036S mutation (6) that changes the P1 residue at the primary cleavage site to a serine resulted in loss of cleavage at the primary cleavage site and enhanced cleavage at the secondary, tertiary, and quaternary cleavage sites. The L1036F mutation in the present study resulted in slightly diminished cleavage at the primary cleavage site and enhanced cleavage at the secondary cleavage site (Fig. 4B). The F1077L mutation resulted in slightly increased cleavage at the tertiary cleavage site (Fig. 4B). When the L1036S and the F1077L mutations or the L1036F and the F1077L mutations were combined, cleavage at the tertiary site was markedly increased, whereas cleavage at the primary, secondary, and tertiary sites was markedly diminished (Fig. 4B). The F1067L mutation resulted in a marked increase in cleavage at the quaternary site (Fig. 4B). Taken together, these findings suggest that cleavage after leucine in the P1 position is more efficient than cleavage after the bulkier phenylalanine in that position.

The P2 residue at the primary and tertiary cleavage sites is leucine, while the P2 residue at the secondary cleavage site is glutamic acid and at the quaternary cleavage site is valine. To
determine whether leucine is preferred in the P2 position, we initially changed the leucine at the primary cleavage site to polar residues serine (L1035S) and threonine (L1035T). The L1035S mutation resulted in diminished cleavage at the primary site and enhanced cleavage at the secondary, tertiary, and quaternary sites. The L1035T mutation virtually eliminated all cleavage, resulting in accumulation of the Hap precursor (Fig. 4B). These data indicate that the P2 position is important for substrate recognition. In order to further test the P2 position, we changed the P2 leucine at the primary cleavage site to a glutamic acid (L1035E), which is also present in the P2 position at the secondary cleavage site. The L1035E mutation showed wild-type cleavage (Fig. 4B). Next, we changed the glutamic acid in the P2 position at the secondary cleavage site to leucine (E1045L). The E1045L mutation resulted in prominent cleavage at the secondary cleavage site and almost no cleavage at the primary, tertiary, and quaternary cleavage sites (Fig. 4C). To further assess cleavage preference at the P2 position, we mutated the valine in the P2 position at the quaternary cleavage site to leucine (V1066L). As shown in Fig. 4C, the V1066L mutation resulted in enhanced
catalytic groove likely includes the L263, R264, E265, and N274 amino acids. 

**DISCUSSION**

In this study, we have investigated the Hap cleavage specificity and substrate groove in detail. Our results demonstrate that the P1 and P2 residues are important determinants of Hap cleavage specificity, that small hydrophobic amino acids are favored in the P1 and P2 positions, and that the substrate groove likely includes the L263, R264, E265, and N274 amino acids.

Our earlier work demonstrated that Hap is a member of the SA (chymotrypsin) clan of serine proteases (6). Interactions between the P1 residue and the S1 subsite pocket encompassing the residues adjacent to the active site serine are important determinants of cleavage specificity of chymotrypsinlike proteases (9). For chymotrypsin, cleavage specificity correlates to the hydrophobicity of the P1 residue and is characterized by preferential cleavage after tryptophan, tyrosine, or phenylalanine (S1034F), a much bulkier residue. 

**TABLE 3. Autoproteolysis phenotypes of Hap mutant derivatives**

<table>
<thead>
<tr>
<th>Wild type or mutation&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Cleavage site or subsite&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Description&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hap wild type</td>
<td>NA</td>
<td>1°, 2°, and 3° cleavage</td>
</tr>
<tr>
<td>HapL1036S</td>
<td>P1 (1°)</td>
<td>No 1° cleavage; 2°, 3°, and 4° cleavage</td>
</tr>
<tr>
<td>HapL1036F</td>
<td>P1 (1°)</td>
<td>1° cleavage slightly diminished, enhanced secondary cleavage</td>
</tr>
<tr>
<td>HapF1077L</td>
<td>P1 (3°)</td>
<td>Slightly enhanced 3° cleavage</td>
</tr>
<tr>
<td>HapF1077L/L1036S</td>
<td>P1 (3°) + P1 (1°)</td>
<td>Enhanced 3° cleavage</td>
</tr>
<tr>
<td>HapF1077L/L1036F</td>
<td>P1 (3°) + P1 (1°)</td>
<td>Enhanced 3° cleavage</td>
</tr>
<tr>
<td>Hap F1067L</td>
<td>P1 (4°)</td>
<td>Enhanced 4° cleavage, some 1° and 3° cleavage</td>
</tr>
<tr>
<td>HapL1035S</td>
<td>P2 (1°)</td>
<td>Very little 1° cleavage, enhanced 2° cleavage</td>
</tr>
<tr>
<td>HapL1035T</td>
<td>P2 (1°)</td>
<td>No effect on autoproteolysis</td>
</tr>
<tr>
<td>HapL1035E</td>
<td>P2 (1°)</td>
<td>Minimal autoproteolysis, accumulation of Hap precursor</td>
</tr>
<tr>
<td>HapE1045L</td>
<td>P2 (2°)</td>
<td>Enhanced 2° cleavage</td>
</tr>
<tr>
<td>HapE1045A</td>
<td>P2 (2°)</td>
<td>Enhanced 2° cleavage</td>
</tr>
<tr>
<td>HapV1066L</td>
<td>P2 (3°)</td>
<td>Enhanced 4°, some 1° and 3° cleavage</td>
</tr>
<tr>
<td>HapS1034F</td>
<td>P3 (1°)</td>
<td>No effect on autoproteolysis</td>
</tr>
<tr>
<td>HapS1034E</td>
<td>P3 (1°)</td>
<td>No effect on autoproteolysis</td>
</tr>
<tr>
<td>HapQ1033A</td>
<td>P4 (1°)</td>
<td>No effect on autoproteolysis</td>
</tr>
<tr>
<td>HapQ1033E</td>
<td>P4 (1°)</td>
<td>No effect on autoproteolysis</td>
</tr>
<tr>
<td>HapN1037R</td>
<td>P1’ (1°)</td>
<td>No effect on autoproteolysis</td>
</tr>
<tr>
<td>HapA1038S</td>
<td>P2’ (1°)</td>
<td>No effect on autoproteolysis</td>
</tr>
<tr>
<td>HapL1039S</td>
<td>P3’ (1°)</td>
<td>No effect on autoproteolysis</td>
</tr>
<tr>
<td>HapE1040S</td>
<td>P4’ (1°)</td>
<td>No effect on autoproteolysis</td>
</tr>
<tr>
<td>HapK240A</td>
<td>S1</td>
<td>Diminished autoproteolysis, accumulation of Hap precursor</td>
</tr>
<tr>
<td>HapL263R</td>
<td>S1</td>
<td>Diminished autoproteolysis, accumulation of Hap precursor</td>
</tr>
<tr>
<td>HapN274R</td>
<td>S2</td>
<td>Diminished autoproteolysis, accumulation of Hap precursor</td>
</tr>
<tr>
<td>HapR264A</td>
<td>S2</td>
<td>Diminished autoproteolysis, accumulation of Hap precursor</td>
</tr>
<tr>
<td>HapY137A</td>
<td>S4</td>
<td>Diminished autoproteolysis, accumulation of Hap precursor</td>
</tr>
<tr>
<td>HapE265A</td>
<td>S4</td>
<td>Diminished autoproteolysis, accumulation of Hap precursor</td>
</tr>
<tr>
<td>HapE265W</td>
<td>S4</td>
<td>Diminished autoproteolysis, accumulation of Hap precursor</td>
</tr>
</tbody>
</table>

<sup>a</sup> All Hap derivatives were expressed in H. influenzae strain DB117.

<sup>b</sup> Numbers in parentheses indicate the primary (1°), secondary (2°), tertiary (3°), or quaternary (4°) cleavage sites. NA, not applicable.

<sup>c</sup> Descriptions are based on visual inspection of the blots in triplicate.
nine, with all of these residues fitting within the deep hydrophobic pocket of the protein (16, 17). However, chymotrypsin can also cleave after leucine, methionine, and histidine in the P1 position, although with lower efficiency (15, 16). In contrast, elastase prefers small hydrophobic residues such as alanine in the P1 position, reflecting a smaller S1 subsite pocket (24). In the present study, we demonstrated that Hap cleaves with high preference for leucine over phenylalanine in the P1 position, suggesting that the S1 subsite pocket in Hap is relatively shallow and accommodates small hydrophobic residues better than bulky residues. This conclusion is consistent with the relative inefficiency of cleavage at the tertiary and quaternary sites in wild-type Hap, where the P1 residue is phenylalanine.

In chymotrypsin, proteolysis is also influenced by interaction between the P2 residue and the S2 subsite pocket (3). In this context, it is noteworthy that our work supports an important role for the P2 residue in Hap cleavage specificity, with leucine or glutamic acid favored over other amino acids. Interestingly, insertion of leucine in the P2 position at the secondary cleavage site (HapE1045L) results in a preference for this cleavage site to the exclusion of other cleavage sites. It is possible that these mutations result in local steric hindrance that causes increased binding of the mutant substrate and more efficient cleavage at the secondary cleavage site. However, without a high-resolution crystal structure, we cannot explain the apparent preference of the secondary cleavage site over the primary cleavage site in this mutant. We also found that insertion of glutamic acid at the P2 position (L1035E) of the primary cleavage site had no significant effect on autoproteolysis. Consistent with this finding, glutamic acid is the P2 residue at the native secondary cleavage site (E1045).

Hap cleavage of a random peptide library revealed high preference for leucine at the P1 and P2 positions, agreeing with our mutagenesis results. Interestingly, phenylalanine was not selected at either of these positions, despite the fact that cleavage occurs naturally at F1067 and F1077 in wild-type Hap (albeit to a lesser extent than with leucine in the P1 position). In considering this information, it is possible that the S1 and S2 subsites in the Hap proteolytic pocket have reduced affinity for phenylalanine in the context of linear peptides compared to native folded protein. Based on selectivity ratios, arginine at the P1 position and glutamic acid at the P2 position seem to be the favored amino acids after leucine. Serine and threonine have the lowest selectivity ratio at the P1 and P2 positions, indicating that they have an inhibitory effect on proteolytic cleavage, in agreement with our mutagenesis data (L1036S, L1035S, and L1035T). Further analysis of the peptide library cleavage assay revealed that glutamic acid was selected at the P3 position and that a number of amino acids were selected at the P4 position.

Considered together, our mutational analysis and our peptide library cleavage results indicate that the P1 and P2 residues are important determinants of Hap cleavage specificity. Previously, we reported a consensus cleavage site sequence based only on alignment of the primary, secondary, tertiary, and quaternary cleavage sites (6). In the present study, mutational analysis reveals a target motif that consists of XX(L/E/V)(L/F) and the peptide library cleavage assay reveals a target motif of XE(L/E)(L/R) at the P4 through P1 positions for the Hap$_\text{p}$ protease. The difference in results between mutational
analysis and peptide library cleavage may be due to the intrinsic difference between cleavage of full-length folded protein and linear peptides. Mutations introduced at the P’ residues of the primary cleavage site had no effect on autoproteolysis, indicating that these residues do not significantly affect proteolytic cleavage.

In an effort to define the Hap substrate binding groove, we developed a homology model based on the crystal structures of the E. coli Hbp and the B. pertussis Pertactin autotransporters. Guided by this model, we performed site-directed mutagenesis, changing residues at the predicted S1, S2, and S4 subsites. Our results demonstrated that L263, N274, R264, and E265 are critical for efficient autoproteolysis, supporting a role for these residues in forming the Hap substrate groove and interacting intimately with substrate residues. Mutations at Y137 and K240 that are located at the rim of S1 and S2 subsites in our model did not affect autoproteolysis.

The implication of the Hap serine protease domain in the pathogenesis of H. influenzae disease could be twofold. First, autoproteolytic release of Hap from the bacterial surface may allow the organism to modulate interactions with host tissues during various stages of colonization and infection. Second, the secreted Hap protease domain may have activity against host substrates such as extracellular matrix proteins or immune factors. We have previously determined that Hap can bind to various extracellular matrix proteins such as fibronectin, collagen IV, and laminin (7). Further investigation is needed to determine whether Hap protease activity might facilitate spread to deeper tissues or promote persistence in the face of the host immune response.

In conclusion, in this study we have provided important insights into the structural determinants of H. influenzae Hap proteolytic activity. Complete knowledge of the Hap substrate groove and cleavage specificity may facilitate the design of therapeutic agents that block Hap proteolytic activity and thereby attenuate H. influenzae disease.

ACKNOWLEDGMENTS

We thank Ben Turk for advice on the peptide library cleavage assays.

This study was supported by Public Health Service grant AI44322 to J.S.G. and Wellcome Trust grant 08087 to G.W. and G.M.

REFERENCES


Editor: A. J. Bäumler