The Hemophilus influenzae Hap Autotransporter Is a Chymotrypsin Clan Serine Protease and Undergoes Autoproteolysis via an Intermolecular Mechanism*

Doran L. Fink‖‖, Leslie D. Cope‖‖, Eric J. Hansen‖‖, and Joseph W. St. Geme III‖‖‖**

From ‖‖The Edward Mallinckrodt Department of Pediatrics and ‖‖Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110 and ‖‖Department of Microbiology, University of Texas Southwestern Medical Center, Dallas, Texas 75235-9048

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 276, No. 42, Issue of October 19, pp. 39492–39500, 2001

Received for publication, July 23, 2001, and in revised form, August 7, 2001
Published, JBC Papers in Press, August 14, 2001, DOI 10.1074/jbc.M106913200

© 2001 by The American Society for Biochemistry and Molecular Biology, Inc.

Bacterial organisms have evolved a number of pathways for presenting proteins on the cell surface and releasing proteins extracellularly. In pathogenic organisms, these pathways are essential for specific interactions that result in colonization of the host and progression to disease. The bacterial proteins that participate in these interactions are considered virulence fac-

tors and influence adherence to host cells, damage to host tissues, and interference with host defense mechanisms.

In Gram-negative bacteria, surface and extracellular pro-

teins must be transported across the inner membrane, the

periplasm, and the outer membrane. This goal is achieved by

one of five distinct secretion systems. The type V secre-
tion system is used exclusively by the autotransporter family of

proteins, typified by the IgA1 proteases of Neisseria and Hem-

ophilus species, the first such proteins to be identified (1).

Autotransporters are typically expressed as precursor polypep-
tides with at least three functional domains, including an NH₃-
terminal signal sequence, an internal passenger domain, and a

COOH-terminal translocator domain. The signal sequence di-

rects export of the polypeptide across the bacterial inner mem-

brane and is then removed by signal peptidase. Subsequently,

the translocator domain inserts into the outer membrane and

appears to fold into a β-barrel structure with a central hydro-

philic pore, allowing extrusion of the passenger domain across

the membrane. Once the passenger domain is localized on the

cell surface, one of several fates is possible. In some cases, the

passenger domain remains covalently linked to the membrane-

associated β-barrel domain. In other cases, it is cleaved but

remains cell associated. In still others, it is cleaved and then

released extracellularly. Cleavage may involve an autoproteo-

lytic event directed by protease activity in the passenger do-

main itself or may occur through the action of a separate bac-

terial protease. Autotransporter passenger domains have

been ascribed a wide variety of functions, examples of which

include adhesins, toxins, degradative enzymes, and serum re-

sistance factors (2).

Hemophilus influenzae is a Gram-negative bacterium and

represents a common cause of human disease, including both

localized respiratory tract and systemic (invasive) disease (3).

The pathogenesis of H. influenzae disease begins with coloni-

zation of the upper respiratory tract. To facilitate colonization,

H. influenzae elaborates both pilus and non-pilus proteins
called adhesins, which promote adherence to host epithelial

cells by interacting with specific host cell surface molecules.
The H. influenzae Hap protein is a non-pilus adhesin that

promotes adherence and invasion in assays with cultured hu-

man epithelial cells and also mediates bacterial aggregation

and microcolony formation (4, 5). In recent work, we dem-

onstrated that Hap is an autotransporter protein and consists of

a 110-kDa passenger domain called Hap₉ and a 45-kDa trans-

locator domain called Hapβ. The passenger domain has serine

protease activity and directs autoproteolytic cleavage via Ser²⁴³,

releasing Hap₉ into the culture supernatant and leaving Hapβ

embedded in the outer membrane (6). Of note, Ser²⁴³ is part of a

GDSGS motif that is present in a subfamily of auto-

" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Current address: Virus and Cell Biology Department, Merck & Co., Inc., West Point, PA 19486.

‖ To whom corresponding should be addressed: Dept. of Pediatrics, Washington University School of Medicine, 660 South Euclid Ave., Campus Box 6208, St. Louis, MO 63110. Tel.: 314-286-2887; Fax: 314-286-2895; E-mail: stgeme@borcim.wustl.edu.

¶ Published, JBC Papers in Press, August 14, 2001, DOI 10.1074/jbc.M106913200

‖‖ This work was supported by an Established Investigator Award from the American Heart Association (to J. W. S.), a research grant from the March of Dimes (to J. W. S.), United States Public Health Service Grant AI17621 (to E. J. H.), and funds from Wyeth-Lederle Vaccines. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†† Current address: Virus and Cell Biology Department, Merck & Co., Inc., West Point, PA 19486.

‖‖‖ This paper is available on line at http://www.jbc.org
transporter proteins, including the Neisseria and Hemophilus IgA1 proteases and a diverse group of serine protease autotransporters secreted by members of the Enterobacteriaceae (SPATEs) (7–13). Hap autoproteolytic cleavage occurs primarily at the peptide bond between Leu\(^{1036}\) and Asn\(^{1037}\), and mutation of these two residues results in nearly complete inhibition of cleavage at this site and increased cleavage at three more COOH-terminal alternate sites (6). Autoproteolysis is blocked by several serine protease-specific inhibitors, including phenylmethylsulfonyl fluoride, Pefabloc, and secretory leukocyte protease inhibitor, a component of human respiratory tract secretions.

In addition to harboring serine protease activity, the Hap passenger domain possesses the adhesive activities responsible for bacterial interaction with epithelial cells and bacterial aggregation. Thus, at first glance release of this domain from the bacterial cell surface seems counterproductive to promoting successful colonization. Indeed, mutation of the Hap active site serine to an alanine results in complete inhibition of autoproteolytic processing, full retention of Hap, on the cell surface, and increased adherence and aggregation (6). On the other hand, release of Hap, may benefit H. influenzae in ways unrelated to its adhesive activities. For example, activity against host substrates has been demonstrated for IgA1 protease (16). Finally, shedding of an adhesive domain from the cell surface may facilitate dissemination of individual organisms from microcolonies and allow spread to new sites. Such events may be controlled by the presence of host factors that modulate Hap autoproteolytic activity (e.g. secretory leukocyte protease inhibitor) or by features inherent to the mechanism of autoproteolysis itself.

In the present study, we sought to characterize Hap autoproteolysis more fully, including the mechanism of catalysis, the identities of all cleavage sites, and the molecular nature of the enzyme-substrate interaction. Our results demonstrate that Hap belongs to the SA (chymotrypsin) clan and contains a catalytic triad conserved among a subfamily of autotransporter serine proteases. Identification of the three alternate autoproteolytic cleavage sites in Hap revealed a consensus target sequence for Hap enzymatic activity. Interestingly, kinetic analysis and examination of a recombinant strain expressing two different derivatives of Hap established that autoproteolysis occurs at least in part via an intermolecular mechanism. These results provide insights into an expanding family of bacterial serine proteases and suggest a novel mechanism for regulation of bacterial adherence.

**EXPERIMENTAL PROCEDURES**

**Protein Sequence Alignments—**Amino acid alignments of protein sequences were performed using ClustalW v.1.8 software available on the Baylor College of Medicine Human Genome Sequencing Center website (dot.imgen bcm tmc.edu).

**Bacterial Strains and Plasmids—**The bacterial strains and plasmids used in this study are listed in Table I (4, 6, 17, 18, 20, 26). H. influenzae strains were grown as described previously (21) and were stored at –80°C in brain-heart infusion broth with 20% glycerol. Escherichia coli
strain DH5α was grown on Luria-Bertani (LB) agar or in LB broth. E. coli strains were stored at ~80 °C in LB broth with 50% glycerol. Antibiotic concentrations for H. influenzae included the following: tetracycline, 5 μg/ml; kanamycin, 50 μg/ml; chloramphenicol, 1 μg/ml; and streptomycin, 250 μg/ml. Antibiotic concentrations for E. coli included tetracycline, 12.5 μg/ml; kanamycin, 50 μg/ml; and ampicillin, 100 μg/ml.

Recombinant DNA Methods—DNA ligations, restriction endonuclease digests, and gel electrophoresis were performed according to standard techniques (22). Plasmids were introduced into E. coli strain DH5α by chemical transformation (22). H. influenzae strain DB117 was transformed using the MIV method of Herriott et al. (23).

Construction of Plasmids Encoding Mutated Hap Derivatives—Site-directed mutagenesis was performed using recombinant PCR1 techniques. To construct pHpH98α, pHpH117α, pHpD139α, and pHpD140α, pMlD100 (pUC19 containing a 6.7-kb insert with the hap gene from H. influenzae strain N187) was used as a template to amplify a 1.4-kb 5′ PCR fragment and a 0.9-kb 3′ PCR fragment that overlapped at an 18-base region, allowing amplification of a unique 6.7-kb fragment. This recombinant PCR fragment was digested with NheI and ligated into NheI-digested pMLD100. The 5′ fragment was then combined with the corresponding 3′ fragment in a 1:1 molar ratio to serve as template in generating a 2.3-kb mutated recombinant PCR fragment using external primers from the initial PCR reactions plus a 1:100 dilution of an internal primer corresponding to the 5′-end to facilitate ligation of the fragment. Each 5′ fragment was then combined with the coding sequence of the mutated site. These recombinant PCR fragments were digested with AvaI and Clal and then ligated into Avai-Clal-digested pMLD100.

For construction of the primary autoproteolytic cleavage site mutant pHapL1036S, initial PCR products (1.1 kb 5′ fragment and 1.2 kb 3′ fragment) were amplified from pMLD100 and used as templates to generate a 2.3-kb mutated recombinant PCR fragment as described above. This recombinant PCR fragment was then digested with NheI and HindIII and ligated into NheI-HindIII-digested pMLD100. The double autoproteolytic cleavage site mutants pHapL1036S::LN1036-7SG (6). pMLD100 (M1346S::LN1036-7SG) was used as the template in the initial PCR reactions.

To construct plasmid pHapL1036-7SG::G3H6, a 1.8-kb PCR fragment corresponding to the COOH-terminal 600 residues of Hap was amplified from pHDH101 (like pMLD100, but with the hap gene inserted in the opposite orientation), incorporating codons for glycine residues (6 histidine residues, and a BamHI site at the 3′ end of the hap gene (5′-end of the PCR product). This PCR fragment was digested with BamHI and NheI and ligated into BamHI-NheI-digested pHDH101. The resulting plasmid, pHDH101::G3H6, was digested with NcoI and SstI, liberating a 1.4-kb fragment that was then religated with ligation buffer and a 3′-kb insert excised from Ncol-NsiI-digested pHDH101::LN1036-7SG (6).

Construction of pHpL1036S-99 was carried out by first amplifying initial PCR fragments corresponding to residues 650–1035 (1.1-kb 5′ fragment) and 1100–1395 (1.3-kb 3′ fragment) of Hap. The primers used to create these PCR products were engineered such that the two fragments overlapped at an 18-base region, allowing amplification of a unique 2.3-kb fragment in which the codons for residues 1035 and 1100 were immediately juxtaposed, deleting the intervening sequence. This recombinant PCR fragment was digested with NheI and HindIII and ligated into NheI-HindIII-digested pMLD100.

Plasmid pJS106::trc was constructed by first amplifying a 1.6-kb PCR fragment that contained the lacZ gene, trc promoter, and polynucleotide cloning region from pTrc99A. The primers used in this amplification were engineered such that the resulting PCR product contained HindIII and PstI sites at the 5′ end and a HindIII site at the 3′ end, allowing digestion of this fragment with HindIII and ligation into HindIII-digested pHDH101. The resulting plasmid, pHDH101::trc, was then digested with Ncol and BglII, excising a fragment containing the trc promoter product in which the codons for residues 1035 and 1100 were immediately juxtaposed, deleting the intervening sequence. This recombinant PCR fragment was digested with NheI and HindIII and ligated into NheI-HindIII-digested pMLD100.

Plasmid pJS106::trc was constructed by first amplifying a 1.6-kb PCR fragment that contained the lacZ gene, trc promoter, and polynucleotide cloning region from pTrc99A. The primers used in this amplification were engineered such that the resulting PCR product contained HindIII and PstI sites at the 5′ end and a HindIII site at the 3′ end, allowing digestion of this fragment with HindIII and ligation into HindIII-digested pHDH101. The resulting plasmid, pHDH101::trc, was then digested with Ncol and BglII, excising a fragment containing the trc promoter product in which the codons for residues 1035 and 1100 were immediately juxtaposed, deleting the intervening sequence. This recombinant PCR fragment was digested with NheI and HindIII and ligated into NheI-HindIII-digested pMLD100. The 5′ fragment was then combined with the coding sequence of the mutated site. These recombinant PCR fragments were digested with AvaI and Clal and then ligated into AvaI-Clal-digested pMLD100.

Kinetic Analysis of Autoproteolysis during Induced Expression of Hap—DB117 expressing Hap under the control of a IPTG-inducible promoter (DB117/pJS106::trc) was grown to an A600 of 1.0, at which time IPTG was added to a final concentration of 0.1 mM. A 10-ml aliquot was removed from the culture every 10 min, beginning with the addition of IPTG, and proteins were isolated from the culture supernatant and membrane fractions as described above. In order to compensate for any differences in protein amounts caused by increases in culture density over the course of the experiment or by variations in purification efficiency, volumes loaded for resolution by SDS-PAGE were adjusted according to culture density for secreted proteins and according to final protein concentration (Bio-Rad protein assay) for outer membrane proteins. After resolution by SDS-PAGE, transfer to nitrocellulose, and detection by immunoblotting, protein band intensities were quantitated by scanning densitometry using an LKB Ultrascan XL laser densitometer (Bromo, Sweden). To further ensure that protein from an equivalent number of bacterial cells had been loaded in each lane, levels of major outer membrane protein P4 (OMP P4) in each of the outer membrane fractions were quantified by scanning densitometry after detection by an anti-P4 monoclonal antibody, EPRS-2.1 (generously provided by B. Green, Wyeth-Lederle Vaccines), and anti-mouse IgG secondary antiserum conjugated to horseradish peroxidase (Sigma).

Inactivation of ompP2 gene, ompF5, and rec1 in Strain Rd—To inactivate the ompP2 gene in H. influenzae strain Rd, the cat cartridge from plasmid pUCAGAT was inserted into a PvuII site contained within the ompP2 gene from type b strain DLA2 on plasmid pEHJ191-35-L3 (26). The inactivated ompP2 gene was excised by digestion with PstI, purified by agarose gel electrophoresis, and used to transform strain Rd. To inactivate the ompP5 gene in Rd/ompP2, a 1.5-kb PCR fragment encoding the 3′ half of the Rd ompP5 gene plus 1 kb of downstream sequence was amplified from Rd chromosomal DNA, incorporating DraIII sites at each end to facilitate ligation of the fragment into DraIII-digested pUC4K (Amersham Pharmacia Biotech). The resulting 5.6-kb plasmid was then digested with PstI, excising a 2.9-kb fragment containing the region from pUC4K and ligated into the PstI site at the 5′ end of the 1.5-kb fragment generated by digestion of the 3′ end with DraIII. This ligated portion of ompP5 and downstream sequence. This 2.8-kb fragment was then ligated into PstI-digested pUC19 (New England Biolabs), creating plasmid pUC19::kan5′-P5-3′. A 1.5-kb PCR fragment encoding the 5′-half of the Rd ompP5 gene plus 1 kb of upstream sequence was amplified from Rd chromosomal DNA, incorporating an EcoRI site at the 5′ end and an XbaI site at the 3′ end to facilitate ligation of the fragment into pUC4K. The resulting 6.7-kb plasmid, pUC19::5′-kan5′-P5-3′, was linearized by digestion with XmnI and used to transform RdompP2. Finally, the rec1 gene of Rd/ompP2ompP5 was inactivated by transformation with DNA from strain BC200/rec1, which possesses a mutant rec1 gene linked to streptomycin resistance (27). Streptomycin-resistant transformants were screened for the loss of recombination activity by assaying their ability to survive exposure to 44 ergs/cm2 of ultraviolet radiation at 254 nm.
that the Hap catalytic triad is composed of Ser$^{243}$, Asp$^{140}$, and His$^{98}$. Alternatively, mutation of these residues may have affected autoproteolysis by disrupting the tertiary structure around the active site. An amino acid alignment of 10 auto-

transporter serine proteases that share the GDSGS sequence surrounding the putative active site serine revealed absolute conservation of all three of these residues (Fig. 2), providing further evidence that these amino acids represent the catalytic triad.

The P1 Residue Is Critical for Recognition of the Hap Primary Autoproteolytic Cleavage Site—Previous studies demonstrated that site-directed mutagenesis of Hap residues Leu$^{1036}$ and Asn$^{1037}$ virtually eliminated autoproteolytic cleavage of the LN1036-7 peptide bond and resulted in increased abundance of three bands, 39–43 kDa in size, in immunoblots of outer membrane proteins from bacteria expressing this Hap derivative, representing cleavage at more COOH-terminal alternate sites. Given that serine protease enzymatic activity typically relies on recognition of residues immediately NH$_2$-terminal to the cleaved peptide bond, we sought to determine whether mutation of the P1 residue alone would inhibit Hap autoproteolysis at the LN1036-7 site. As shown in Fig. 3A, examination of outer membrane proteins from DB117/

HapL1036S revealed a marked decrease in abundance of the 45-kDa Hap$_p$ species, indicating nearly complete elimination of cleavage at the LN1036-7 site. At the same time, the 110–116 kDa secreted Hap$_p$ proteins were present in slightly decreased quantity in culture supernatants (Fig. 3B). In this and other immunoblots, the Hap$_p$ proteins resulting from autoproteolytic processing at multiple sites appeared to migrate as a single band due to their close proximity in size. A very small amount of the 45-kDa cleavage product was produced, similar to the situation with HapLN1036-7SG, suggesting that additional amino acids beyond the P1 residue may also be critical for target site recognition.

Identification of Alternative Hap Autoproteolytic Cleavage Sites Reveals a Conserved Motif—Next we set out to identify the alternate cleavage sites responsible for production of the 39–43-kDa Hap$_p$ minor cleavage products. To facilitate recovery of these proteins for NH$_2$-terminal amino acid sequencing, we first constructed a Hap derivative that contains both the mutated primary cleavage site (LN1036-7SG) and a 3xGly-6xHis tag at the C terminus. Ultimately this derivative was expressed in an H. influenzae strain Rd mutant that does not express major outer membrane proteins P2 and P5, which consistently contaminated early preparations of His-tagged Hap$_p$ minor cleavage products from strain DB117. Talon bead affinity purification of 6xHis-tagged proteins from the outer membrane of Rd/ompP2/ompP5/rec1/pHapLN1036-7SG::G3H6 lead to the recovery of three protein species migrating at 43, 39, and 41 kDa, in order of abundance (not shown). NH$_2$-terminal amino acid sequencing of the 43-kDa band revealed the presence of equivalent amounts of two proteins, one beginning with Thr$^{1047}$ (TAETQK) and the other beginning with Ala$^{1048}$ (AETQKS), indicating that autoproteolytic cleavage at the secondary site occurs either between Leu$^{1046}$ and Thr$^{1047}$ between Thr$^{1047}$ and Ala$^{1048}$, or both. NH$_2$-terminal amino acid sequencing of the 39- and 41-kDa bands resulted in unambiguous assignment of the tertiary site between Phe$^{1067}$ and Ala$^{1078}$ (ALEAAL) and the quaternary site between Phe$^{1067}$ and Ser$^{1068}$ (SDPPLL). Alignment of the five residues flanking each side of the primary, secondary, tertiary, and quaternary autoproteolytic cleavage sites allowed construction of a consensus target sequence for Hap enzymatic activity, namely (Q/R)/S/Al/X(L/F) at the P4 through P1 positions (Fig. 4). There is no apparent similarity of sequence at the P$^-$ positions.

RESULTS

The Hap Catalytic Triad Consists of His$^{98}$, Asp$^{140}$, and Ser$^{243}$—In considering the mechanism of Hap autoproteolysis, we first sought to identify the residues that participate with Ser$^{243}$ in the Hap catalytic site. Examination of the Hap predicted amino acid sequence revealed two aspartic acid residues ~100 residues amino-terminal to Ser$^{243}$ (Asp$^{139}$ and Asp$^{146}$) and two histidine residues 120–140 residues amino-terminal to Ser$^{243}$ (His$^{98}$ and His$^{117}$), reminiscent of the SA (chymotrypsin) clan of serine proteases. Alignment of Hap sequences from eight different clinical isolates of nontypable H. influenzae revealed absolute conservation of all four of these residues (data not shown). With this information in mind, we changed each of these residues individually to an alanine using site-directed mutagenesis and then expressed the resulting mutant proteins in H. influenzae strain DB117. As shown in Fig. 1, examination of outer membrane proteins from DB117/

pHapD140A and DB117/pHapH98A demonstrated accumulation of the 155-kDa full-length Hap$_p$ cleavage products, identical to earlier observations with DB117/pHapS243A. Similarly, examination of proteins in culture supernatants from these two strains demonstrated the absence of the 110-kDa secreted Hap$_p$ protein (Fig. 1). In contrast, mutation of either Asp$^{139}$ or His$^{117}$ had no effect on autoproteolysis. Taken together, these data suggest

Identification of Hap Autoproteolytic Cleavage Sites—Plasmid pHapLN1036-7SG::G3H6 was transformed into strain Rd/ompP2/ompP5/ rec1. A 1-liter culture of Rd/ompP2/ompP5/rec1/pHapLN1036-7SG::G3H6 was grown to an A$_{600}$ of 0.8, and Sarkosyl-insoluble outer membrane proteins were purified using a scaled-up version of the protocol described by Carlone et al. (24). The pellet consisting of the outer membrane fraction was resuspended in 5 ml of 20 mM Tris, 6 M guanidinium chloride, 100 mM sodium chloride, pH 8.0. 1 ml of a 50% slurry of Talon beads (CLONTECH) was added to the sample, and the mixture was incubated at 4 °C for 3 h. The Talon beads were then washed 4 times with 5 ml of 20 mM Tris, 8 M urea, 100 mM sodium chloride, 10 mM imidazole, pH 8.0, and histidine-tagged outer membrane proteins were eluted from the beads in 0.5 ml of 20 mM Tris, 8 M urea, 100 mM sodium chloride, 10 mM imidazole, 100 mM EDTA, pH 8.0. The eluted proteins were resolved by SDS-PAGE using 7.5% SDS-polyacrylamide gels and electrotransferred to a polyvinylidene difluoride membrane (28). After staining with Coomassie Brilliant Blue R-250, three protein bands migrating at 43, 41, and 39 kDa, respectively, were excised from the membrane and submitted to Midwest Analytical (St. Louis, MO) for amino-terminal sequence determination performed by automated Edman degradation using a Perkin-Elmer Applied Biosystems model 477A sequencing system.

Figure 1. Effect of mutations of Hap catalytic triad residues on the processing and secretion of Hap$_p$. Shown is an analysis of outer membrane proteins (lanes 1–6) and culture supernatants (lanes 7–12) from late-exponential phase cultures of DB117/pJS106 (containing wild type Hap), DB117/pHapS243A, DB117/pHapD139A, DB117/ pHapD140A, DB117/pHapH117A, and DB117/pHapH98A. Proteins were assessed by immunoblot using antiserum Rab730, which reacts pHapD140A, DB117/pHapH117A, and DB117/pHapH98A. Proteins were assessed by immunoblot using antiserum Rab730, which reacts with full-length Hap$_p$ and Hap$_p$. The lanes of the gel were loaded as follows: lanes 1 and 7, DB117/pJS106; lanes 2 and 8, DB117/ pHapS243A; lanes 3 and 9, DB117/pHapD139A; lanes 4 and 10, DB117/ pHapD140A; lanes 5 and 11, DB117/pHapH117A; lanes 6 and 12, DB117/pHapH98A. Dots indicate the 155-kDa full-length Hap protein and the 45-kDa Hap$_p$ preferred cleavage product. The arrow indicates the 110-kDa Hap$_p$ secreted protein.

Effect of mutations of Hap catalytic triad residues on the processing and secretion of Hap$_p$. As shown in Fig. 1, examination of outer membrane proteins from DB117/

HapL1036S revealed a marked decrease in abundance of the 45-kDa Hap$_p$ species, indicating nearly complete elimination of cleavage at the LN1036-7 site. At the same time, the 110–116 kDa secreted Hap$_p$ proteins were present in slightly decreased quantity in culture supernatants (Fig. 3B). In this and other immunoblots, the Hap$_p$ proteins resulting from autoproteolytic processing at multiple sites appeared to migrate as a single band due to their close proximity in size. A very small amount of the 45-kDa cleavage product was produced, similar to the situation with HapLN1036-7SG, suggesting that additional amino acids beyond the P1 residue may also be critical for target site recognition.
To confirm the conclusions resulting from NH$_2$-terminal amino acid sequencing of the His-tagged Hap$_{H9252}$ species and to assess whether the P1 residue is critical for recognition of the alternate cleavage sites, we performed site-directed mutagenesis. As shown in Fig. 3, mutation of the P1 residues at both the primary and secondary sites (HapL1036S, L1046S) resulted in nearly complete elimination of both the 45- and 43-kDa cleavage products and a corresponding decrease in abundance of the 110–116 kDa secreted Hap$_S$ proteins. The finding that mutation of Leu 1046 disrupted cleavage at the secondary site suggests that this site probably occurs between Leu 1046 and Thr 1047, although it is also possible that Leu 1046 is important as the P2 residue for cleavage between Thr 1047 and Ala 1048. Unlike the situation with mutation of the primary site alone, mutation of both the primary and secondary sites did not result in increased abundance of Hap$_S$ proteins.
**FIG. 4. Alignment of sequences surrounding the four Hap autoprotoelytic cleavage site.** The peptide bond where cleavage occurs is indicated by the vertical line. The numbered positions in the Hap amino acid sequence of the P1 and P1’ residues on either side of the cleaved bond are indicated in parentheses.

| 1° (1036-7) | Q S L L N A L E A |
| 2° (1046-7) | Q A E L T A E T Q |
| 3° (1077-8) | Q S L F A L E A A |
| 4° (1067-8) | R A V F S D P L L |

Consensus: Q S L L

**FIG. 3. Effect of mutations of Hap autoprotoelytic cleavage site P1 residues on the processing and secretion of Hap.** Panel A, extracellular proteins from late-exponential phase cultures of DB117/pHapL1036S, DB117/pHapL1036S::L1046S, DB117/pHapL1036S::F1067S, and DB117/pHapL1036S::F1077S. Panel B, extracellular proteins from the same strains. Proteins were assessed by immunoblot using antiserum Rab730, which reacts with full-length Hap, Haps, and Hap. The arrow indicates the 45-kDa full-length Hap protein. The dot indicates the 155-kDa full-length Hap protein. The 45-kDa Hap species resulting from cleavage at the preferred site. The asterisks indicate the 39–43-kDa Hap species resulting from cleavage at alternate sites.

**FIG. 5. Examination of Hap processing and secretion by forced intermolecular autoprotoelysis.** Panel A, extracellular proteins from late-exponential phase cultures of DB117/pHaps243A, DB117/pHapS243A::L1046S, DB117/pHapS243A::F1067S, and DB117/pHapS243A::F1077S. Panel B, extracellular proteins from the same strains. Proteins were assessed by immunoblot using antiserum Rab730, which reacts with full-length Hap, Haps, and Hap. The arrow indicates the 45-kDa full-length Hap protein and the 148-kDa Hap derivative lacking the intervening region between Hap and Haps. The dot indicates the 45-kDa Hap species resulting from cleavage at the preferred site. The asterisks indicate the 39–43-kDa Hap species resulting from cleavage at alternate sites.

is proteolytically active but lacks a 64-residue region containing all four autoprotoelytic cleavage sites, thus eliminating autoprotoelysis when expressed on its own (Fig. 5, A and B). When these two derivatives were expressed together on the same bacterial cell surface, we observed four 39–43-kDa Hap species in the outer membrane and 110–116-kDa Hap species in culture supernatants, suggesting that intermolecular autoproteolysis likely occurs only between Hap molecules on the same bacterial cell and not between Hap molecules on different cells (not shown).

To extend the observation that intermolecular cleavage occurs, we examined the possibility that released Hap is capable of cleaving the full-length Hap precursor on the cell surface. Consistent with previous results, we found that incubation of purified Hap (final concentration 2 μg/ml) with DB117/Haps243A resulted in cleavage of Haps243A at the primary site (Fig. 6) (6). However, comparison with outer membrane fractions from DB117 expressing wild-type Hap revealed relatively reduced cleavage at the primary site (reduced amount of the 45-kDa Hap protein) and no cleavage at the 3 alternate sites (no smaller Hap species) (Fig. 6). Next, we incubated purified Hap with DB117 expressing a Hap derivative containing only the β-barrel domain (Hap, beginning after the primary cleavage site). Again, there was no evidence of cleavage at the 3 alternate sites (Fig. 6). Given that the amount of purified Hap used in these digestion experiments was much greater than the quantity of Hap released into culture supernatants of DB117 expressing wild-type Hap, it seems unlikely that intermolecular...
H. influenzae Hap Is a Chymotrypsin Clan Serine Protease

Identification of His$^{188}$, Asp$^{140}$, and Ser$^{243}$ as a set of three residues involved in the Hap catalytic site provides strong evidence that Hap is a classical serine protease with a catalytic triad. The order in which these residues appear in the Hap primary sequence and the distances between them suggest that Hap is a member of the SA (chymotrypsin) clan of serine proteases (29). This clan comprises a diverse spectrum of proteases including the eukaryotic chymotrypsins, trypsins, and leukocyte elastase of the S1 family, as well as several other families of microbial proteases. Members of the SA clan share primary sequence and the distances between them suggest that the autoproteolytic mechanism is the predominant means of autoproteolysis.

**DISCUSSION**

To extend our conclusion that Hap is a member of the SA clan, we used the 3D-PSSM structural modeling program (Imperial Cancer Research Fund Fold Recognition Server, www.bmm.icnet.uk/servers/3dpssm) to predict the secondary structural elements within the 300-residue Hap NH$_2$-terminal domain and to compare predicted folding patterns with a library of known protein structures. The results of this analysis

**Fig. 6. Analysis of Hap-mediated proteolysis.** Shown is an analysis of outer membrane proteins from DB117/pJS106 (containing wild-type Hap), DB117/pHapS243A, and DB117/pHap$_p$ (containing the Hap signal sequence fused in-frame to the Hap$_p$ β-barrel domain) after incubation with purified Hap$_p$ or phosphate-buffered saline control. Proteins were assessed by immunoblot using antiserum Rab730, which reacts with full-length Hap and Hap$_p$. The lanes in the gel were loaded as follows: lane 1, DB117/pJS106 treated with phosphate-buffered saline; lane 2, DB117/pHapS243A treated with phosphate-buffered saline; lane 3, DB117/pHapS243A treated with 2 μg of Hap$_p$; lane 4, DB117/pHap$_p$ treated with phosphate-buffered saline; lane 5, DB117/pHap$_p$ treated with 2 μg of Hap$_p$. The arrow indicates the 155-kDa full-length Hap protein. The dot indicates the 45-kDa Hap$_p$ species resulting from cleavage at the preferred site. The asterisks indicate the 39–43-kDa Hap$_p$ species resulting from cleavage at alternate sites.

**Fig. 7. Examination of the effect of Hap expression levels on the rate of Hap autoproteolytic processing.** Panel A, analysis of outer membrane proteins (Outer Membranes, time points 0 to 70 min) and extracellular proteins (Supernatants, time points 0 to 70 min) from DB117/pJS106:Trc99 (containing wild-type Hap expressed under the control of an IPTG-inducible promoter). The culture was grown to early stationary phase (A$_{500} = 1.0$) and then induced with 0.1 mM IPTG. The lanes in the gel were loaded with outer membrane proteins or extracellular proteins prepared from a 10-ml sample collected at the indicated time after induction. Proteins were assessed by immunoblot using antiserum Rab730, which reacts with full-length Hap, Haps, and Hap$_p$. The arrows indicate the 155-kDa full-length Hap protein and the 45-kDa Hap$_p$ species resulting from cleavage at the preferred site. The dot indicates the 110-kDa secreted Hap, protein. Panel B, kinetic analysis of Hap expression and autoproteolytic processing. The intensities of the protein bands shown in Panel A were quantified by scanning densitometry to approximate the amounts of full-length Hap and Hap$_p$ present in outer membrane fractions and the amounts of Hap$_p$ present in culture supernatants at each time point. Shown are the detected levels of full-length Hap plus Hap$_p$ (diamonds, solid line), Hap$_p$ alone (squares, dashed line), and Hap$_p$ (circles, dotted line).
strongly suggest that the Hap NH2-terminal domain could be modeled to fit the three-dimensional structure of bovine chymotrypsinogen C, an S1 family protease with similarities in sequence and structure to both chymotrypsin and elastase (35). Chymotrypsinogen C and the Hap NH2-terminal domain exhibit roughly 19% identity at the amino acid level. Importantly, with few exceptions, hydrophobic residues predicted to form the structural core of bovine chymotrypsinogen C are conserved in the Hap sequence. Our analysis also predicts that Hap could be modeled using the structure of trypsin from the fungal organism Streptomyces griseus, although conservation of predicted core residues is not as extensive between Hap and S. griseus trypsin as it is between Hap and bovine chymotrypsinogen C (36). Ongoing efforts to crystallize the secreted Hap protein may eventually confirm these structural predictions.

Further analysis of the Hap amino acid and nucleotide sequences highlights several features that may help to evaluate the evolutionary relationship between Hap and other SA clan members. For example, the AGT codon present at the Hap active site serine is reminiscent of AGY codons utilized by physiologically complex proteases of higher metazoans, such as blood clotting factors, and contrasts to the TCN codons conserved among more "primordial" clan members, such as trypsin, chymotrypsin, and elastase (34). Furthermore, whereas most SA clan proteases have either a proline or tyrosine at position 225 as a determinant of Na+-activated allostery, Hap has an aspartic acid at the corresponding position. Other notable differences between functionally important residues in S1 family proteases and their counterparts in the Hap sequence include substitutions for conserved serines at positions 189 and 214 by alanine and arginine, respectively, and the absence of four cysteine residues at positions 1, 122, 191, and 220 that form disulfide bonds critical to the stability of the catalytic pocket. These dissimilarities may reflect how selective pressures contributing to the evolution of Hap differed from those affecting eukaryotic S1 proteases and further suggest that the Hap catalytic pocket may represent a distinct and possibly novel structure.

Alignment of the amino acid sequences of Hap and nine other autotransporter proteases revealed sequence similarities throughout the NH2-terminal domains and absolute conservation of all three catalytic triad residues, suggesting that these proteins comprise a subfamily of autotransporter proteases that share a common structure at the catalytic site. Included in this subgroup are Hap, the Neisseria and Hemophilus IgA1 proteases, and the SPATEs, a closely related set of putative bacterial virulence factors with distinct enzymatic functions. The existing classification system has tentatively assigned these autotransporters to the S6 family within the SA clan, based solely on conservation of a GDSGS motif and identification of the active site serine within this motif in three members of the subfamily (30, 31). Our present study provides the first conclusive biochemical evidence for assignment of an autotransporter to the SA clan and further suggests that Hap belongs to a subset of autotransporters related to but distinct from the S1 family proteases. Of note, several groups have described members of a second family of autotransporter serine proteases that possess a presumed catalytic motif seemingly distinct from that of Hap and its family members (32, 33). Based on the conservation of aspartic acid, histidine, and serine residues among its members in the same order and relative positions as the active site residues in subtilisin, this second family is referred to as the subtilases, although biochemical data to support this claim have not been reported.

Identification of the four Hap autoproteolytic cleavage sites provides evidence that Hap may be most like chymotrypsins in terms of substrate specificity. Site-directed mutagenesis confirmed that the residue at the P1 position of Hap autoproteolytic cleavage sites is critical for recognition of substrate by the Hap catalytic pocket, a feature typical of serine proteases in general. Members of the SA clan have differing substrate specificities at the P1 position, with examples including bulky hydrophobic residues such as phenylalanine and leucine in chymotrypsin-like proteases, basic residues such as arginine and lysine in trypsin-like proteases, and small hydrophobic residues such as alanine in elastase-like proteases. The presence of either leucine or phenylalanine at the P1 positions of the Hap autoproteolytic cleavage sites suggests that the Hap catalytic site substrate pocket can accommodate bulky hydrophobic residues, analogous to chymotrypsin. The presence of leucine at the P1 position of the Hap primary and secondary autoproteolytic cleavage sites suggests a preference for this residue over phenylalanine, contrasting with most chymotrypsin-like proteases but resembling bovine chymotrypsinogen C, which cleaves preferentially after leucine residues, presumably due to the influence of a threonine residue (Thr236) in the specificity pocket (35). Cleavage after large hydrophobic residues has also been demonstrated for several other IgA1 protease-like autotransporters. The Neisseria and Hemophilus IgA1 proteases cleave after proline residues within the hinge region of IgA1, and in vitro digestion of chromogenic peptide substrates by SepA from Shigella flexneri suggested that this enzyme acts on target sites with phenylalanine at the P1 position (37, 38). Further examination of Hap substrate specificity using chromogenic peptide substrates should help to address whether Hap is truly a chymotrypsin-like protease, keeping in mind that recognition of the cleavage site may require the presence of additional residues beyond those present in commonly available substrate reagents. For many SA clan proteases, substrate recognition and catalytic efficiency are dependent on residues both NH2-terminal to and COOH-terminal to the P1 position. At this point it remains unclear whether substrate positions P2 through P4 affect Hap enzymatic activity, although homologies between residues at the P2 and P4 positions among the autotransporters may suggest this possibility.

Intermolecular mechanisms of autoproteolytic processing have been demonstrated for a number of proteases. Typically, the processing event serves to remove an inhibitory peptide chain from azymogen, thus converting it into an active enzyme (19, 39–42). Autoproteolysis of Hap differs from these examples in that Hap enzymatic activity requires no modification of the protease domain subsequent to expression of the precursor on the cell surface. Rather, processing serves to alter localization of the passenger domain from the cell surface to the extracellular environment. Release of Hap via a predominantly intermolecular mechanism might therefore allow H. influenzae organisms to regulate the percentage of Hap, passenger domain associated with the cell surface by modifying the expression level of precursor molecules. In this model, one could imagine that during the initial stages of colonization, bacteria might express a relatively low level of Hap on the cell surface such that enough passenger domain is present to promote adherence to epithelial structures and formation of microcolonies but not enough to facilitate appreciable levels of autoproteolysis. Once colonization reaches a stage at which available local binding sites are saturated with dense colonies of organisms or a host immune response is elicited, bacteria might then increase Hap expression to levels that promote processing, resulting in extracellular release of Hap. Such a transition might benefit the bacteria in a variety of ways, for example, by removing an antigenic structure from the cell surface, by promoting disbursement of organisms from large colonies and
facilitating spread to new sites, and by allowing the secreted protease domain to encounter host substrates involved in subsequent stages of pathogenesis. Regulation of this transition might be further influenced by inhibitors of Hap enzymatic activity, namely secretory leukocyte protease inhibitor, which is present in normal respiratory secretions and is increased in quantity in the setting of inflammation (5).

In summary, the *H. influenzae* Hap autotransporter is a member of the SA (chymotrypsin) clan of serine proteases with a catalytic triad that consists of His\(^{98}\), Asp\(^{140}\), and Ser\(^{243}\). These residues are conserved among a subset of autotransporters that appear to constitute a distinct family within the SA clan. Furthermore, Hap cleaves after large hydrophobic residues, suggesting chymotrypsin-like substrate specificity. Autoproteolytic release of the Hap passenger domain and general insights into the diversity may provide clues as to the ultimate purpose for release of the Hap passenger domain and general insights into the diverse functions of serine proteases.

**Acknowledgments**—We thank Bruce Green for generous contribution of plasmid pUCARCAT and anti-P4 monoclonal antibody EPR5-2.1, Jerry Pinkner for valuable assistance in the purification of Hap\(_r\), and Thierry Rose, David Hendrixson, Gail Hardy, and Enrico Di Cera for helpful discussions and critical reading of the manuscript.

**REFERENCES**

1. Henderson, I. R., Cappello, R., and Nataro, J. P. (2000) *Trends Microbiol.* 8, 529–532
2. Henderson, I. R., and Nataro, J. P. (2001) *Infect. Immun.* 69, 1231–1243
3. Turk, D. C. (1984) *J. Med. Microbiol.* 17, 23–33
4. St. Geme, J. W., de la Morena, M. L., and Falkow, S. (1994) *Mol. Microbiol.* 14, 217–233
5. Hendrixson, D. R., and St. Geme, J. W., III (1998) *Mol. Cell* 2, 841–850
6. Hendrixson, D., de la Morena, M. L., Stathopoulos, C., and St. Geme, J. W., III (1997) *Mol. Microbiol.* 26, 505–518
7. Brander, W., Schmid, H., and Karch, H. (1997) *Mol. Microbiol.* 24, 767–778
8. Eslava, C., Navarro-Garcia, F., Czezulin, J. R., Henderson, I. R., Cravito, A., and Nataro, J. P. (1998) *Infect. Immun.* 66, 3155–3163
9. Henderson, I. R., Czezulin, J., Eslava, C., Noriega, F., and Nataro, J. P. (1999) *Infect. Immun.* 67, 772–781
10. Stathopoulos, C., Provence, D. L., and Curtiss, R., III (1999) *Infect. Immun.* 67, 772–781
11. Benjelloun-Touimi, Z., Sansonetti, P. J., and Parsot, C. (1995) *Mol. Microbiol.* 17, 123–135
12. Stein, M., Kenny, B., Stein, M. A., and Finlay, B. B. (1996) *J. Bacteriol.* 178, 6546–6554
13. Al-Hasani, K., Henderson, I. R., Sakellaris, H., Rajakumar, K., Grant, T., Nataro, J. P., Robins-Browne, R., and Adler, B. (2000) *Infect. Immun.* 68, 2437–2443
14. Kilian, M., Mestecky, J., Kulchy, R., Tomana, M., and Butler, W. T. (1980) *J. Immunol.* 124, 2596–2600
15. Lin, E., Ayala, P., Larson, J., Mulk, M., Fukuda, M., Carlson, S. R., Enns, C., and So, M. (1997) *Mol. Microbiol.* 24, 1083–1094
16. Pohler, J., Langenberg, U., Walk, U., Beck, S. C., and Meyer, T. F. (1995) *Mol. Microbiol.* 17, 1073–1083
17. Setlow, J. K., Brown, D. C., Boling, E. M., Mattingly, A., and Gordon, M. P. (1968) *J. Bacteriol.* 95, 546–558
18. Tomb, J. F., Barcak, G. J., Chandler, M. S., Redfield, R. J., and Smith, H. O. (1989) *J. Bacteriol.* 171, 3796–3802
19. Wittlin, S., Rosel, J., Hofmann, F., and Stover, D. R. (1999) *Eur. J. Biochem.* 265, 384–393
20. Williams, P. J., Albritten, W. L., Slaney, L., and Setlow, J. K. (1989) *Antimicrob. Agents Chemother.* 33, 1627–1630
21. Anderson, P., Johnston, R. B., Jr., and Smith, D. H. (1972) *J. Clin. Invest.* 51, 31–38
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
23. Herriott, R. M., Meyer, E. M., and Vogt, M. (1970) *J. Bacteriol.* 101, 517–524
24. Carline, G. M., Thomas, M. L., Rumschlag, H. S., and Sottnek, F. O. (1986) *J. Clin. Microbiol.* 24, 330–332
25. Laemmli, U. K. (1970) *Nature* 227, 680–685
26. Cope, L. D., Pelzel, S. E., Latimer, J. L., and Hansen, E. J. (1990) *Infect. Immun.* 58, 3312–3318
27. Setlow, J. K., Boling, M. E., Allison, D. P., and Beattie, K. L. (1973) *J. Bacteriol.* 115, 153–161
28. Matsudaica, P. (1987) *J. Biol. Chem.* 262, 10035–10038
29. Barrett, A. J., and Rawlings, N. D. (1995) *Arch. Biochem. Biophys.* 318, 247–250
30. Rao, M. B., Tanksale, A. M., Ghatge, M. S., and Deshpande, V. V. (1998) *Biochim. Biophys. Acta* 1376, 510–515
31. Rawlings, N. D., and Barrett, A. J. (1993) *Biochem. J.* 290, 205–218
32. Oshinsky, Y., Beppo, T., and Horinouchi, S. (1997) *J. Biochem. (Tokyo)* 121, 902–913
33. Lo, R. Y., Strathdee, C. A., Shewen, P. E., and Cooney, B. J. (1991) *Infect. Immun.* 59, 3398–3406
34. Krem, M. M., and Di Cera, E. (2001) *EMBO J.* 20, 3036–3045
35. Gomis-Ruth, F. X., Gomez, M., Bode, W., Huber, R., and Aviles, F. X. (1995) *EMBO J.* 14, 4387–4394
36. Read, R. J., and James, M. N. (1988) *J. Biol. Chem.* 263, 523–522
37. Benjelloun-Touimi, Z., Yahara, M. S., Montecucco, C., Sansonetti, P. J., and Parsot, C. (1998) *Microbiology* 144, 1815–1822
38. Piault, A. G., and Bachovchin, W. W. (1994) *Methods Enzymol.* 244, 137–151
39. Kasche, V., Lummer, K., Nuhn, A., Piotraschke, E., Rieks, A., Stoeva, S., and Voelker, W. (1999) *Biochim. Biophys. Acta* 1333, 76–86
40. Michetti, M., Salamino, F., Tedesco, I., Averna, M., Minafra, R., Melleni, E., and Ponnremoli, S. (1996) *FEBS Lett.* 399, 11–15
41. Weinheimer, S. P., McCann, P. J., 3rd, O'Soye, D. R., 2nd, Stevens, J. T., Boyd, B. A., Drier, D. A., Yamanaka, G. A., Dilani, C. L., Deckman, J. C., and Cordingly, M. G. (1998) *J. Virol.* 67, 5815–5822
42. Guan, C., Cui, T., Rao, V., Liao, W., Benner, J., Lin, C. L., and Comb, D. (1996) *J. Biol. Chem.* 271, 1732–1737