The Structure of the *Haemophilus influenzae* HMW1 Pro-piece Reveals a Structural Domain Essential for Bacterial Two-partner Secretion*

Hye-Jeong Yeo, Takeshi Yokoyama, Katarzyna Walkiewicz, Youngchang Kim, Susan Grass, and Joseph W. St. Geme, III

Received for publication, July 13, 2007. Published, JBC Papers in Press, August 14, 2007, DOI 10.1074/jbc.M705750200

From the †Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204, the ‡Structural Biology Center, Argonne National Laboratory, Argonne, Illinois 60439, and the ††Departments of Pediatrics and Molecular Genetics & Microbiology, Duke University, Medical Center, Durham, North Carolina 27710

In pathogenic Gram-negative bacteria, many virulence factors are secreted via the two-partner secretion pathway, which consists of an exoprotein called TpsA and a cognate outer membrane translocator called TpsB. The HMW1 and HMW2 adhesins are major virulence factors in nontypeable *Haemophilus influenzae* and are prototype two-partner secretion pathway exoproteins. A key step in the delivery of HMW1 and HMW2 to the bacterial surface involves targeting to the HMW1B and HMW2B outer membrane translocators by an N-terminal region called the secretion domain. Here we present the crystal structure at 1.92 Å of the HMW1 pro-piece (HMW1-PP), a region that contains the HMW1 secretion domain and is cleaved and released during HMW1 secretion. Structural analysis of HMW1-PP revealed a right-handed β-helix fold containing 12 complete parallel coils and one large extra-helical domain. Comparison of HMW1-PP and the *Bordetella pertussis* FHA secretion domain (Fha30) reveals limited amino acid homology but shared structural features, suggesting that diverse TpsA proteins have a common structural domain required for targeting to cognate TpsB proteins. Further comparison of HMW1-PP and Fha30 structures may provide insights into the keen specificity of TpsA-TpsB interactions.

In Gram-negative bacteria, the two-partner secretion (TPS) pathway serves as a common secretion mechanism for large protein virulence factors and is essential for the virulence of many human pathogens, including *Haemophilus influenzae*, *Bordetella pertussis*, *Serratia marcescens*, and *Proteus mirabilis*, among others. The TPS pathway consists of a secreted exoprotein (referred to as a TpsA protein) and a cognate outer membrane translocator (referred to as a TpsB protein) (1–3). TpsA proteins are synthesized as preproteins 100–500 kDa in size that are processed in the course of secretion across the bacterial inner and outer membranes, yielding functional TpsA proteins (3, 4). Despite limited overall sequence conservation among the TpsA members, functional studies have established that TpsA proteins contain common features, including an atypical N-terminal signal peptide and an adjacent region of about 250 residues that forms the so-called secretion domain (1–4). Although our knowledge of TpsB proteins remains relatively limited, recent studies have established that TpsB proteins have a modular structure with a C-terminal pore-forming domain (5, 6).

The *H. influenzae* HMW1 and HMW2 proteins are high molecular weight, non-pilus adhesins that were originally identified as major targets of the human serum antibody response during acute otitis media (7). These proteins are present in ~80% of nontypeable *H. influenzae* strains and mediate adherence to a variety of epithelial cell types (4, 8). HMW1 and HMW2 are encoded by separate chromosomal loci, with each locus consisting of three genes, designated *hmwA*, *hmwB*, and *hmwC* (8–10). The *hmwA* genes encode the surface-exposed adhesins (HMW1 and HMW2), and the *hmwB* and *hmwC* genes encode accessory proteins required for proper processing and secretion of the adhesins (5, 11–14). In *H. influenzae* strain 12, the HMW1 and HMW2 proteins are synthesized as preproteins and undergo two discrete cleavage events during the process of maturation and surface localization, the first releasing the N-terminal signal peptide corresponding to residues 1–68 and the second releasing the pro-piece corresponding to residues 69–441 (see Fig. 1). In *H. influenzae* strain 12, mature HMW1 corresponds to residues 442–1536 in the HMW1 preprotepin, and mature HMW2 corresponds to residues 442–1477 in the HMW2 preprotepin. In both HMW1 and HMW2, adhesive activity resides in a ~360-amino acid region at the N terminus of the mature species, and anchoring to the bacterial surface is mediated by a 20-amino acid region at the very C terminus of the protein (13–15).

The HMW1 pro-piece (HMW1-PP) shares limited homology at the N-terminal end with the secretion domain in other TpsA proteins and is critical for secretion (15–18). Recent work demonstrated that a chimeric protein containing HMW1-PP...
and a segment of the passenger domain of the *H. influenzae* Hia adhesin (Hia<sup>60–779</sup>) was secreted into the supernatant in an HMW1B-dependent manner (12). In contrast, this chimera was not secreted when co-expressed with the Hia translocator domain, demonstrating specificity between HMW1-PP and the HMW1B translocator (12). Far Western analysis established that HMW1-PP interacts directly with HMW1B, revealing in part a mechanism for how the TpsA secretion domain facilitates secretion (12). Following interaction with HMW1B, HMW1-PP is cleaved and released from the organism, leading to the mature adhesin on the bacterial surface (see Fig. 1).

The functionally characterized TpsA/TpsB pairs in the TPS family include FHA/FhaC of *B. pertussis*, ShlA/ShlB of *S. marcescens*, and HMW1/HMW1B of *H. influenzae* (5, 6, 15–19). In all of these examples, proper secretion requires that the secretion domain of TpsA and the periplasmic domain of TpsB recognize each other in the periplasm. Interestingly, previous studies either excluded HMW1 from multiple sequence alignments of TpsA secretion domains or provided a poor sequence alignment (2, 3, 20), reflecting the limited sequence homology. In contrast, this chimera suggests that there are structural elements that are common to the TPS pathway.

In an effort to advance our understanding of the structural basis of TPS, we set out to solve the crystal structure of HMW1-PP. In this report we describe the crystal structure of HMW1-PP at 1.92 Å. Despite the sequence and functional diversity among members of the TPS family, analysis of the structure of HMW1-PP suggests that TpsA proteins have a common structural domain required for targeting to the cognate TpsB protein.

**EXPERIMENTAL PROCEDURES**

**Cloning, Expression, and Purification of the HMW1 Pro-piece**—The native HMW1 pro-piece (HMW1-PP, residues 69–441) was expressed as a secreted protein by generating a DH5α derivative that contains pHMW:B:HMWC (Cam<sup>+</sup>) and pHMW1<sup>1–441</sup>:HAT (Clontech) (Amp<sup>+</sup>). The plasmid pHMW1<sup>1–441</sup>:HAT was created by ligating a DNA fragment encoding HMW1 amino acids 69–441 with an EcoRI site at the 5′ end and a SalI site at the 3′ end and then ligating this fragment into EcoRI-SalI-digested pGEX-6P-1 (GE Healthcare). The construct was transformed into *Escherichia coli* strain DL41, a Met auxotroph strain, and the SeMet-labeled protein was produced as described previously (31). After cleavage of the GST moiety from GST::HMW1-PP using PreScission protease (GE Healthcare), SeMet-labeled HMW1-PP was further purified using an anionic exchange column (Hitrap Q, 5 ml; GE Healthcare) equilibrated in buffer B. The purified fractions were finally submitted to size exclusion chromatography using HiPrep Sephacryl 16/60 S200 equilibrated with buffer C (20 mM MES, pH 5.5, 100 mM NaCl, 5% glycerol).

For the purpose of selenium multiwavelength anomalous dispersion phasing, we generated a construct expressing a GST::HMW1-PP fusion protein. The plasmid pGEX::HMW1<sup>69–441</sup> was generated by amplifying a 1.1-kb DNA fragment encoding HMW1 amino acids 69–441 with an EcoRI site at the 5′ end and a Sall site at the 3′ end and then ligating this fragment into EcoRI-SalI-digested pGEX-6P-1 (GE Healthcare). The construct was transformed into *Escherichia coli* strain DL41, a Met auxotroph strain, and the SeMet-labeled protein was produced as described previously (31). After cleavage of the GST moiety from GST::HMW1-PP using PreScission protease (GE Healthcare), SeMet-labeled HMW1-PP was further purified using an anionic exchange column (Hitrap Q) and/or a gel filtration column. For both columns, SeMet-labeled HMW1-PP showed the same elution profile as the native secreted protein. This preparation yielded about 0.8 mg of SeMet-labeled HMW1-PP/liter of culture.

**Crystallization and Data Collection**—Crystals of the native protein (secreted form) were grown in a 1:1 mixture of protein (6 mg/ml in buffer of 20 mM MES, pH 5.5, 100 mM NaCl, 5% glycerol) and reservoir solution containing 22–24% polyethylene glycol 6K and 0.1 M HEPES (pH 6.8–7.2) using the hanging drop vapor diffusion method at 17 °C. Native crystals (rectangular or triangular thin plate) grew to typical dimensions 200 μm × 150 μm × 40 μm within 2 weeks and diffracted to 2.4 Å when using a synchrotron beam source (Advanced Photon Source, 19BM). The SeMet-labeled protein (cytoplasmic form) often yielded needle crystals with the same conditions (polyethylene glycol 6K) used for the native crystals, and growing crystals of SeMet-labeled HMW1-PP was much more difficult. For data collection, small SeMet derivatized crystals with dimensions 70 μm × 50 μm × 40 μm were used after quick cryoprotection with reservoir solution containing 25% glycerol and direct freezing in liquid N<sub>2</sub>. From single crystals, two single-wavelength anomalous dispersion data sets were collected at 100 K (Advanced Photon Source, 19BM) at the wavelength of 0.9789 Å. One crystal diffracted to 1.92 Å with high crystal...
Crystal Structure of HMW1-PP

mosaicy, and another crystal diffracted to 2.4 Å with low crys-
tal mosaicy. Although the low resolution data set with low
mosaicy was used for single-wavelength anomalous disper-
sion phasing, the high resolution data set was used for struc-
tural refinement. The data sets were indexed and integrated
using HKL2000 and scaled with SCALEPACK (32). Both native
form and SeMet-derivatized crystals belonged to space group
I4 with unit cell dimensions a = b = 121.4 Å, c = 50.58 Å and
one molecule/asymmetric unit.

Structure Determination and Refinement—The HMW1-PP
structure was solved by the single-wavelength anomalous dis-
ersion method. Selenium atom search, initial phasing, and
density modification were performed using autoSHARP (33).
The initial model was built manually using COOT (34) and
XTALVIEW (35) and placed with 60% of the secondary struc-
tural elements as polyalanine and 10% with the side chain. After
the model building, initial refinement was carried out using
CNS1.1 (36) with rigid body, simulated annealing, and
restrained individual B factor refinement. At this step the R
factor was over 43% (R\text{free} = 48%), and then several steps of
refinement were performed using REFMAC5 (37) followed by
manual adjustment of the model. Coordinates with electron
density greater than 4σ in F\text{c} − F\text{o} maps were assigned as water
molecules if the locations were reasonable for hydrogen bond-
ing. B factor cut-off of 40.0 Å\(^2\) was applied to water molecules,
and any water molecules refining to higher values were
removed from the model. The final model has an R factor of
17.4% and a free R factor of 21.7%. Structural analysis of the final
model using the Protein Data Bank validation suite indicated
that none of the residues is in the disallowed region on the
Ramachandran plot, and almost all the residues are in the most
favored regions. A summary of the data collection and refine-
ment statistics is given in Table 1. The coordinates and struc-
ture factors for HMW1-PP have been deposited in Protein Data
Bank (code 2ODL).

Secretion Assay—To assess the portion of HMW1-PP
required for secretion, we generated plasmids encoding HAT-
tagged HMW1\(^{1–361}\) (pHMW1\(^{1–361}\)::HAT) and HMW1\(^{1–269}\)
(pHMW1\(^{1–269}\)::HAT), using pHMW11–441::HAT as a control.
These plasmids were created by ligating a DNA fragment con-
taining 340 bp of sequence upstream of the hmw1A start codon
and the appropriate coding sequence into HindIII-digested
pHAT10 (Clontech) and were then transformed into
DH5\(\alpha\) containing 340 bp of sequence upstream of the
hmw1A start codon and the appropriate coding sequence into
HindIII-digested pHAT10 (Clontech) and were then transformed into
DH5\(\alpha\) (E. coli DH5\(\alpha\) harboring pACYC-HMW1B (12). The resulting strains
were resuspended from plates into LB broth to an absorbance
at 600 nm of 0.3 and were incubated at room temperature for 45
min and then subjected to centrifugation at 6,000 \(\times g\) at 4 °C for
10 min. The cell pellet was resuspended in 10 mm HEPES, pH
7.4, and sonicated to clarity. Culture supernatants were precipi-
tated by adding trichloroacetic acid to a final concentration of
10% (v/v), incubating for 10 min at 4 °C, and then centrifuging
at 15,600 \(\times g\) at 4 °C for 10 min. The trichloroacetic acid-pre-
cipitated proteins were resuspended in 0.2 M Tris, pH 9.0. Cell
sonicates and trichloroacetic acid precipitates were resolved on
SDS-PAGE gels and examined by Western analysis using an
antiserum against the HAT epitope.

RESULTS

Structure Determination—In earlier work, we found that the
HMW1 pro-piece (HMW1-PP, corresponding to residues
69–441) mediates interaction with the HMW1B outer mem-
brane translocator and is essential for HMW1 secretion (12, 15).
During the process of translocation of HMW1 across the
outer membrane, HMW1-PP is cleaved from the HMW1 pro-
protein, generating mature HMW1 (Fig. 1). To study the mecha-
nism of HMW1 secretion, we focused initially on obtaining
HMW1-PP as a secreted native protein. As a first step, we con-
structed a DH5\(\alpha\) derivative harboring pHMW11–441::HAT
(encoding HMW1\(^{1–441}\) fused to the HAT epitope) and
pHMW1B-HMW1C (encoding HMW1B and HMW1C). In
this strain, HMW1\(^{1–441}\) is directed to the inner membrane via
the signal sequence, processed to release amino acids 1–68,
translocated through HMW1B, and secreted extracellularly.
Using this strain, we were able to purify large amounts of
HMW1-PP from the culture supernatant. This secreted native
form of HMW1-PP yielded trigonal or tetragonal plate crystals.
Molecular replacement using Fha30 did not give a clear struc-
ture solution, and a search for heavy metal derivatives was
unsuccessful. As an alternative approach, we generated a con-
struct encoding a GST::HMW1-PP fusion protein and then
expressed this protein in E. coli DL41 in SeMet labeling
medium. Following recovery of the GST::HMW1-PP fusion
protein from the bacterial cytoplasm, we cleaved the GST moi-
ety and purified SeMet-labeled HMW1-PP. The SeMet-labeled
form of HMW1-PP showed the same chromatographic prop-
erties as the native secreted form of HMW1-PP, indicating that
the protein expressed in the cytoplasm was folded, comparable
with the native secreted form. Both cytoplasmic HMW1-PP and
native secreted HMW1-PP crystallized in the tetragonal
space group I4 with one molecule per asymmetric unit and
diffracted to high resolution. The structure of HMW1-PP was

FIGURE 1. Domain organization of HMW1 and model for TPS. The HMW1 signal sequence (residues 1–68) is shown in yellow, the HMW1 pro-piece (residues 69–441) is shown in magenta, and the mature HMW1 adhesin (442–1536) is shown in blue (a large cell surface structure) and in orange (a small C-terminal anchor). For clarity, only the secretion process across the outer membrane is presented. After Sec-dependent export of the HMW1 pro-protein across the inner membrane, the HMW1 secretion domain interacts with HMW1B, likely in the periplasm, facilitating translocation of HMW1 across the outer membrane. Sometime during this translocation process, HMW1-PP is cleaved and released to the extracellular space (S. Grass and J. W. St. Geme III, unpublished data). Ultimately, the binding domain of mature HMW1 is exposed on the tip of the adhesin, away from the bacterial cell surface, in a position that favors access to host cells.
solved using single-wavelength anomalous dispersion phasing, and the structure of the native form was identical to the SeMet-derivated form when solved by molecular replacement. The final model has an R/Rfree of 17.4/21.7% (Table 1). This model comprises all 371 amino acids of HMW1-PP except for residue 69 and includes 182 water molecules.

**TABLE 1**

| Statistics from crystallographic analysis |
|------------------------------------------|
| **Data collection**                      |
| Wavelength (Å)                           | 0.97891 |
| Unit cell (Å)                            | a = b = 121.438, c = 50.582 |
| Resolution (Å)                           | 50-1.92 (1.99-1.92) |
| Observed/unique reflections              | 182,336/28,035 |
| Completeness (%)                        | 98.8 (95.1) |
| Rmerge (%)                               | 9.9 (22.8) |
| (I/σ(I))                                 | 6.5 |
| Redundancy                               | 13.3 (10.1) |

| **Structure of the HMW1-PP** |
|-------------------------------|

**Crystal Structure of HMW1-PP**—The structure of HMW1-PP at a resolution of 1.92Å represents a complete view of residues 70–441, including the functional secretion domain. HMW1-PP is a monomer folded into a large right-handed coil termed a parallel β-helix with dimensions of ~70 Å × 40 Å × 30 Å (Fig. 2, A and B). The β-helix is formed by three parallel β-sheets, referred to as PB1, PB2, and PB3 (Fig. 2C), according to the naming of the β-sheets in the pectate lyase structure, the first β-helix fold determined (21). The entire HMW1-PP sequence contains 12 complete parallel coils in the helix (Fig. 2, A and C). At the N-terminal side, the first three β-strands are anti-parallel to the core right-handed β-helix. The first two parallel coils (β4–β5–β6 and β7–β8–β9) and the topping anti-parallel rung β1–β2–β3 are tilted from the core β-helix axis and form a slightly tapered β-helix. Interestingly, the highly conserved NPNG motif (denoted in green in Fig. 2A) forms a type I β turn and bridges this twisted N-terminal part and the rest of the helix. Following the NPNG turn, the coils become more regular and form an overall cylinder, with a triangle-shaped cross-section.

Most parallel β-helical structures appear to have a conserved amphipathic α-helix at the N terminus, which caps and shields the hydrophobic core of the β-helix at the N-terminal side (22). In HMW1-PP the N terminus is uncovered and is stabilized by the first three antiparallel strands β1, β2, and β3. Another notable feature of the HMW1-PP structure is the internal organization of the amino acids that form the core of the parallel β-helix.

**FIGURE 2. Crystal structure of HMW1-PP.** A, ribbon diagram showing the overall structure of HMW1-PP (38). The parallel β-strands forming the three parallel β-sheets PB1, PB2, and PB3 (the core β-helix fold) are colored in blue, red, and yellow, respectively. The N-terminal anti-parallel and extra-helical strands and the helices are shown in cyan and pink, respectively. For clarity, β strands are labeled only over PB3. The conserved NPNG loop connecting β11 and β12 is highlighted in green. B, ribbon diagram of HMW1-PP viewed down the β-helix axis. C, topology diagram of HMW1-PP. The parallel β-strands forming the β-helix are colored in blue, red, and yellow as in A. The noncore secondary structural elements are denoted in light blue.
The interior of the β-helix consists of mostly aliphatic hydrophobic residues but also contains two other important types of side chain stacks, namely an aromatic cluster and a hydrogen-bonded asparagine ladder. Four aromatic residues (Trp102, Phe105, Phe115, and Phe167 in B5, B6, B7, and B14, respectively) are clustered within the first four coils, thereby reinforcing the hydrophobic core of the N-terminal side. The PB1 interior of coils 4–12 is remarkable in that every β-strand in this region is strictly composed of three residues, forming a very regular and long Ile ladder, with the exception of Leu186 at β28 (Table 2). The PB2 sheet is composed of β-strands with 4–6 residues/strand and represents a very irregular face. The side chains of the PB2 interior are variable, including Leu/Val/Ile/Thr/Ala/Trp/Phe. The PB3 sheet of coils 4–12 is very regular and is composed of β-strands with four residues/strand. One side chain stack of PB3 is characterized by an Asn ladder and a contiguous Ile/Leu/ladder (Table 2). Of note, four Asn residues in PB3 (Asn199, Asn200, Asn202, and Asn202 in B16, B19, B24, and B27, respectively) are stacked in an elegant manner to produce an extensive network of hydrogen bonds. In fact, the position of the side chains is optimal for hydrogen bonding and packing. In each position of PB stacks, the nature of hydrophobic residues is very similar, thereby resulting in an overall cylinder-shaped β-helix. On the exterior of the helix, the side chains contain largely polar residues (Fig. 3), except for one stack of 13 residues at PB1 that includes seven hydrophobic residues Ile/Leu/Val (Table 2).

The T1, T2, and T3 connecting loops and turns between β-helical strands correspond to loop regions between PB1 and PB2, between PB2 and PB3, and between PB3 and PB1, respectively (Fig. 2C). The T3 turns are the most regular and short loops and contain three to five residues (predominantly three residues). The T1 and T2 turns are more variable in length and adopt several secondary structural elements, forming three protruding extra-helical domains. The large extra-helical domain in T2 is formed by α1, α15, α22, and α23 and is bent toward PB2, stabilized by forming hydrophobic cores with residues from B8, B11, B14, B18, and B21 of PB2. In contrast, the small extra-helical domain in T1 composed of β32, β36, and β40, and α2 is oriented externally and is exposed to solvent. As a result of the T1, T2, and T3 turns, the PB1 and PB3 faces are completely accessible to solvent, and the PB2 face is partially shielded by the large extra-helical domain. Of note, as shown in Fig. 3, the partial shielding of the PB2 face creates a major hydrophobic pocket decorated by basic amino acid residues, namely Lys295, Lys304, and Lys325.

**TABLE 2**
Assignment of β-strands in the core helix of HMW1-PP

| PB3  | PB2  | PB1  | Coil number |
|------|------|------|-------------|
| β1_71 → 72 | β8_75 → 78 | β9_78 → 86 |
| LQ  | DWVH | AMQVD |
| 105 → 106 β8 | 102 → 98 β8 | 94 → 89 β8 |
| PQ  | WNTIA | NRNTIK |
| 136 → 134 β8 | 128 → 123 β8 | 116 → 112 β8 |
| LQS | VREVF | LEQVM |
| 156 → 154 β12 | 149 → 145 β11 | 142 → 139 β10 |
| ITI | ILYFQ  | SDLK |
| 199 → 196 β16 | 170 → 167 β15 | 163 → 161 β13 |
| NVE | SATF  | |
| 226 → 217 β19 | 215 → 211 β18 | 204 → 202 β17 |
| NRVK | GILNV | TT |
| 262 → 259 β23 | 235 → 230 β22 | 225 → 223 β21 |
| NVAE | ALLICS | SV |
| 282 → 279 β27 | 277 → 272 β26 | 267 → 265 β25 |
| NRNT | ARWNN | PV |
| 310 → 307 β30 | 303 → 298 β29 | 287 → 285 β28 |
| NVAE | ASLVDN | SLK |
| 335 → 332 β34 | 330 → 325 β33 | 315 → 313 β31 |
| LTVK | GTRMLK | SV |
| 373 → 370 β38 | 353 → 350 β37 | 342 → 340 β36 |
| LGRTK | LYTE | DIV |
| 396 → 393 β42 | 393 → 38 β41 | 380 → 378 β39 |
| ILIA | NWLAR | SV |
| 427 → 424 β45 | 421 → 417 β44 | 404 → 402 β43 |
| IFLD | STEVF | NEK |
| 439 → 437 β47 | 434 → 432 β46 |  

**FIGURE 3.** Electrostatic potential surface of HMW1-PP using GRASP (39). Positive and negative potentials are indicated in blue and red, respectively. The major pocket formed on the PB2 face is indicated.
β1 and β2 of Fha30 are antiparallel to the core β-helix (Fig. 4A). Hence, the N terminus of HMW1-PP forms an uncovered helical structure, whereas the N terminus of Fha30 has a capped structure (Fig. 4C). The extrahelical β-hairpin formed by β7 and β8 in Fha30 does not have the counterpart in HMW1-PP. Interestingly, the major extrahelical motif in HMW1-PP and Fha30 (composed of α1/β15/β22/β23 in HMW1-PP and β16/β17/β24/β25 in Fha30) is formed at the same spatial position for each structure, despite the different nature and arrangement of the secondary structure elements. As shown in Fig. 4A, the sequence conservation is highly concentrated on β strands forming the core helix. Yet, with the low sequence identity, side chain stacks of the interior as well as the exterior of the coils are quite different, resulting in very different surface properties for the two structures and thereby conferring specificity for the interaction with the cognate TpsB protein.
Crystal Structure of HMW1-PP

![Image](60x618 to 288x734)

**FIGURE 5.** Functional analysis of the portion of HMW1-PP required for secretion. Immunoblot analysis using antiserum against the HAT epitope is shown. Lane 1, DH5\(^{\text{A}}\)/pHMW1\(^{1–441}\)-HAT; lane 2, DH5\(^{\text{A}}\)/pHMW1\(^{1–361}\)-HAT; lane 3, DH5\(^{\text{A}}\)/pHMW1\(^{1–269}\)-HAT; and lane 4, DH5\(^{\text{A}}\)/pHMW1\(^{1–269}\)-HAT, respectively, in the presence of pACYC-HMW1B. The cell pellet was resuspended and sonicated to prepare for each construct. The cell pellet was resuspended and sonicated to prepare for each construct. Culture supernatants were prepared by trichloroacetic acid precipitation. The expected molecular masses for HMW1\(^{1–441}\), HMW1\(^{1–361}\), and HMW1\(^{1–269}\) are 41, 32, and 22 kDa, respectively, as indicated by asterisks.

**STRUCTURAL REQUIREMENT FOR HMW1-PP SECRETION**—The structure of HMW1-PP is a single domain structure and contains the secretion domain, a functional region that is proposed to correspond to the N-terminal ~250 residues of TpsA proteins but is poorly defined (20, 23). In an effort to define the region of HMW1-PP required for secretion, we generated constructs encoding HMW1\(^{1–441}\), HMW1\(^{1–361}\), and HMW1\(^{1–269}\) and then expressed these proteins in the presence of HMW1B. As shown in Fig. 5, HMW1\(^{1–441}\), HMW1\(^{1–361}\), and HMW1\(^{1–269}\) were all detectable in whole-cell sonicates, but only HMW1\(^{1–441}\) and HMW1\(^{1–361}\) were associated with secretion into the culture supernatant (proteins corresponding to HMW1\(^{69–94}\) and HMW1\(^{69–361}\) were detectable in culture supernatants, reflecting cleavage of the signal peptide). (The smaller bands in lanes 2 and 3 in Fig. 5B presumably represent breakdown products of HMW1\(^{69–94}\) and HMW1\(^{69–361}\), respectively). These results demonstrate that the functional HMW1 secretion domain corresponds to a portion of HMW1-PP, larger than HMW1\(^{69–269}\) and perhaps as large as HMW1\(^{69–361}\).

**DISCUSSION**

The TPS pathway is among the most widespread mechanisms for protein secretion in Gram-negative bacteria and is characterized by specific recognition between a TpsA exoprotein and a TpsB outer membrane translocator. The nontypeable *H. influenzae* HMW1 adhesin is a prototype virulence factor secreted via the TPS pathway. In this study, we have elucidated the structure of the HMW1-PP, a region that interacts with the HMW1B outer membrane translocator and is essential for HMW1 secretion. The HMW1-PP structure adopts a monomeric β-helix with 12 complete parallel helix turns and a single large extra-helical domain. Importantly, the HMW1-PP structure shares striking similarity with Fha30, demonstrating a conserved fold critical for secretion of TpsA proteins and allowing a structure-based sequence alignment of HMW1-PP and Fha30. Of note, earlier studies failed to report a meaningful sequence alignment of HMW1 and FHA, reflecting the fact that these two prototype TpsA proteins possess low sequence similarity.

Previous studies have examined the NPNG and NPNL motifs that are located within the ~100 residue region adjacent to the signal peptide in TpsA proteins (15–18). Although all TpsA proteins contain the NPNG motif, only a subset contain the NPNL motif, including FHA but not HMW1. The structure-based sequence alignment of HMW1-PP and Fha30 demonstrates that the NPNL motif resides in a 19-residue insertion that is present in FHA and lacking in HMW1. This insertion contains the β7 and β8 strands and forms an extra-helical β-hairpin in Fha30, serving a function that may be important for FHA but is obviously not essential for all TpsA proteins. In contrast, the NPNG motif appears to be a critical structural element for all TpsA proteins. Indeed, in HMW1 the NPNL motif contributes to a type I β-turn and is largely buried in the hydrophobic core, suggesting that this β-turn is important for folding of TpsA proteins. In previous work, we found that mutation of the NPNG motif in HMW1 to IAIG resulted in a loss of processing and secretion (15). Similarly, examination of this motif in other TpsA proteins has demonstrated a crucial role in secretion (16–18). Based on the structures of HMW-PP and Fha30, we generated a multiple sequence alignment of representative TpsA proteins, revealing two distinct subsets (Fig. 6). It is noteworthy that the variable regions between HMW1-PP and Fha30 are conserved within each subset, suggesting that all members of a given subset adopt the same structure.

Since the report of *Erwinia chrysanthemi* pectate lyase C (PelC) as the first example of a β-helix structure, β-helical folds have been identified in a number of proteins associated with infection (21). Representatives include pectate lyase-related proteins, viral adhesin proteins, and bacterial virulence factors such as Bordetella spp. p69 pertactin (20, 24–27). Recently, β-helical folds have also been identified in an antibiotic resistance protein by mimicking DNA and a pollen allergen protein (28, 29). A search for structural homologs of HMW1-PP using the program DALI (30) identified a number of proteins with very high scores. The HMW1-PP fold has a high level of structural similarity (Z score of >10; root mean square deviation value of 2.4–3.6 Å) to 13 other β-helix structures that are separated into two different functional classes of proteins, namely adhesin molecules (e.g. FHA, tailspike viral adhesion protein, and p69 pertactin) and glycoside hydrolases, including PelC homologs (20, 21, 24–27). Of note, all of these molecules are secreted to the cell surface or released to the extracellular space. Interestingly, among these β-helix structures, HMW1-PP and Fha30 are unique in that the N terminus comprises antiparallel β-strands, instead of α-helical capping. As we await more TpsA structures, we speculate that the N-terminal antiparallel coil may be a signature for the TpsA β-helix fold, serving to stabilize the N-terminal side.

The functional HMW1 adhesin is anchored to the bacterial surface via a C-terminal 20-amino acid region (14). As a corollary, the N-terminal end of mature HMW1 is likely located distal from the bacterial surface. Given that the N
terminus of mature HMW1 is the site of the HMW1-binding domain, it is possible that cleavage of the HMW1-PP fragment enhances exposure of the binding domain and thereby facilitates interaction with host cells. In considering the fact that HMW1-PP adopts the same fold as some glycoside hydrolases, it is notable that HMW1 is glycosylated, raising the possibility that HMW1-PP may be capable of cleaving carbohydrates, resulting in an increase in the structural polymorphism of HMW1.

In conclusion, the structure of HMW1-PP provides the first high resolution picture of the H. influenzae HMW1 TPS pathway. Despite limited overall sequence similarity between HMW1-PP and other TpsA proteins, our structure-based alignment demonstrates remarkable sequence conservation in the secondary structural elements throughout the β-helix core structure. Thus, the HMW1-PP structure highlights a fundamental concept, namely that the essential structural and functional elements of proteins are preserved during evolution, sometimes independent of specific amino acid sequences. The structure presented here provides the foundation for rational mutagenesis studies to investigate the determinants of the highly specific interaction between HMW1-PP and HMW1B and between TpsA and TpsB proteins in general. Moreover, based on our multiple sequence alignment, homology models of the secretion domains in other HMW1-like TpsA proteins will be useful in designing new experiments to examine their functional roles.

Acknowledgments—We thank the staff of Beamline 19BM of the Structural Biology Center at Advanced Photon Source (Argonne National Laboratory) for help during data collection.

REFERENCES

1. Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C., and Alsa‘Aldeen, D. (2004) Microbiol. Mol. Biol. Rev. 68, 692–744
