

Structure of the outer membrane translocator domain of the *Haemophilus influenzae* Hia trimeric autotransporter

Guoyu Meng^{1,6}, Neeraj K Surana^{2,3,6},
Joseph W St Geme III^{2,3,4,5,*}
and Gabriel Waksman^{1,*}

¹Institute of Structural Molecular Biology at UCL/Birkbeck, London, UK, ²The Edward Mallinckrodt Department of Pediatrics, Washington University School of Medicine, St Louis, MO, USA, ³Department of Molecular Microbiology, Washington University School of Medicine, St Louis, MO, USA, ⁴Department of Pediatrics, Duke University Medical Center, Durham, NC, USA and ⁵Department Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC, USA

Autotransporter proteins are defined by the ability to drive their own secretion across the bacterial outer membrane. The Hia autotransporter of *Haemophilus influenzae* belongs to the trimeric autotransporter subfamily and mediates bacterial adhesion to the respiratory epithelium. In this report, we present the crystal structure of the C-terminal end of Hia, corresponding to the entire Hia translocator domain and part of the passenger domain (residues 992–1098). This domain forms a β -barrel with 12 transmembrane β -strands, including four strands from each subunit. The β -barrel has a central channel of 1.8 nm in diameter that is traversed by three N-terminal α -helices, one from each subunit. Mutagenesis studies demonstrate that the transmembrane portion of the three α -helices and the loop region between the α -helices and the neighboring β -strands are essential for stability of the trimeric structure of the translocator domain, and that trimerization of the translocator domain is a prerequisite for translocator activity. Overall, this study provides important insights into the mechanism of translocation in trimeric autotransporters.

The EMBO Journal (2006) 25, 2297–2304. doi:10.1038/sj.emboj.7601132; Published online 11 May 2006

Subject Categories: membranes & transport; structural biology

Keywords: adhesion; crystal structure; microbial pathogenesis; protein secretion; trimeric autotransporter

*Corresponding authors. G Waksman, Birkbeck College, School of Crystallography, Malet Street, London WC1E 7HX, UK.

Tel.: +44 0207 631 6833; Fax: +44 0207 631 6833;

E-mail: g.waksman@mail.cryst.bbk.ac.uk and JW St Geme III, Department of Pediatrics, Duke University Medical Center, T901 Children's Health Center, Durham, NC 27710, USA.

Tel.: +1 919 681 4080; Fax: +1 919 681 2714;

E-mail: j.stgeme@duke.edu

⁶These authors contributed equally to this work

Received: 14 March 2006; accepted: 12 April 2006; published online: 11 May 2006

Introduction

The translocation of proteins across membranes is an essential task for all cell types. In pathogenic gram-negative bacteria, virulence proteins are typically extracellular and must be translocated across both the inner membrane and the outer membrane. The autotransporter family of proteins contains over 700 members and is the largest family of gram-negative bacterial virulence proteins, with functions including adherence, invasion, cytotoxicity, serum resistance, cell-to-cell spread, and proteolysis, among others (Henderson and Nataro, 2001; Pallen *et al*, 2003; Desvaux *et al*, 2004; Henderson *et al*, 2004). All autotransporter proteins are expressed as precursor proteins with three basic functional domains, including an N-terminal signal peptide, an internal passenger domain (the effector domain), and a C-terminal translocator domain (also called the β -domain) (Henderson *et al*, 2004; Jacob-Dubuisson *et al*, 2004). The process of autotransporter secretion begins with export across the inner membrane via a Sec-dependent process, initiated by the signal peptide. Subsequently, the translocator domain inserts into the outer membrane and facilitates surface localization of the passenger domain (Henderson *et al*, 2004; Jacob-Dubuisson *et al*, 2004; Cotter *et al*, 2005).

Recent studies have established that autotransporters can be separated into two distinct subfamilies, designated conventional autotransporters and trimeric autotransporters (Jacob-Dubuisson *et al*, 2004; Cotter *et al*, 2005). In conventional autotransporters, the C-terminal translocator domain is monomeric, is relatively uniform in size, and contains approximately 300 amino acids. Over the past few years, studies on the structure of the translocator domain of two prototypic conventional autotransporters have yielded conflicting results. Veiga *et al* (2002) performed biochemical and electron microscopy studies on the solubilized translocator domain of the *Neisseria gonorrhoeae* IgA1 protease and observed multimeric ring-like structures with 6–10 subunits and a central cavity ~ 2 nm in diameter. In contrast, Oomen *et al* (2004) recently solved the crystal structure of the *in vitro* folded translocator domain of the *N. meningitidis* NalP protease and found a monomeric β -barrel containing 12 transmembrane β -strands and a hydrophilic central pore, with a transmembrane N-terminal α -helix spanning the pore. The trimeric autotransporters are characterized by a short translocator domain that contains ~ 70 amino acids and forms heat-resistant, sodium dodecyl sulfate (SDS)-resistant trimers in the outer membrane (Roggenkamp *et al*, 2003; Surana *et al*, 2004). Based on phylogenetic analysis, the C-terminus of the trimeric autotransporters is uniformly distinguishable from the C-terminus of conventional autotransporters (Cotter *et al*, 2005). The prototype members of the trimeric autotransporter subfamily are the *Yersinia enterocolitica* YadA adhesin and the *Haemophilus influenzae* Hia adhesin. According to secondary structure analysis and FTIR data, the translocator domains of

YadA and Hia are predicted to contain four β -strands, leading to the proposal that the functional translocator is trimeric and contains four β -strands from each subunit in the trimer, resulting in a 12-stranded β -barrel (Hoiczky *et al*, 2000; Cotter *et al*, 2005; Wollmann *et al*, 2006). Consistent with the biochemical observations and the secondary structure predictions regarding the YadA and Hia translocator domains, crystal structures of fragments of the YadA and Hia passenger domains have revealed a trimeric architecture with three identical intertwined subunits (Nummelin *et al*, 2004; Yeo *et al*, 2004). However, no crystal structure of the translocator domain of trimeric autotransporters has yet been reported.

Based on studies of conventional autotransporters, four separate models have been proposed for translocation of autotransporter passenger domains across the outer membrane, referred to as the 'multimeric assembly model,' the 'threading model,' the 'hairpin model,' and the 'Omp85 model' (Jacob-Dubuisson *et al*, 2004). According to the multimeric assembly model, 6–10 β -barrels assemble to form a common pore and the passenger domain is translocated across the common pore (Veiga *et al*, 2002). According to the threading model, the N-terminal end of the passenger domain is targeted to the translocator domain and the passenger domain is threaded through the β -barrel from the N-terminus to the C-terminus. According to the hairpin model, translocation of the passenger domain is initiated with the C-terminal end of the passenger domain forming a hairpin structure in the β -barrel and is completed as the rest of the passenger domain sequence slides through the pore. Finally, according to the Omp85 model, the pore-forming Omp85 outer membrane protein facilitates insertion of the autotransporter translocator domain into the outer membrane and then translocates the autotransporter passenger domain through the Omp85 pore.

In this study, we report the crystal structure of the naturally folded C-terminal 107 amino acids of Hia. As predicted, the Hia translocator domain is a trimeric β -barrel with 12 transmembrane β -strands, including four strands from each subunit. The β -barrel has a central channel that is 1.8 nm in diameter and is traversed by three N-terminal α -helices, one from each subunit, providing a scaffold for folding of the passenger domains into a trimer. These results provide strong evidence that (1) conventional autotransporters and trimeric autotransporters share a common translocator domain structure and are distinguished by the number of α -helices that traverse the central channel and (2) the Hia passenger domain is translocated across the outer membrane through the Hia β -barrel via either the hairpin model or a combination of the hairpin model and a variation of the Omp85 model.

Results

Structure determination

To determine the structure of the Hia translocator domain, we began by cloning fragments of *hia* downstream of sequence encoding the signal peptide of OmpA and the streptavidin tag. We focused our efforts on the minimal translocation unit as defined by Surana *et al* (2004), corresponding to Hia_{1022–1098}, and on a larger fragment with more N-terminal sequence, corresponding to Hia_{992–1098}. Following extraction from the outer membrane, Hia_{992–1098} and Hia_{1022–1098} were purified using a streptavidin-affinity column and then ion exchange

chromatography, in each case yielding a single band on an SDS-PAGE gel corresponding to the molecular weight of a trimer (Figure 1A), analogous to native Hia. Both Hia_{992–1098} and Hia_{1022–1098} crystallized readily and diffracted to high resolution. The structure of Hia_{1022–1098} was solved using SAD phasing on the selenomethionine-derivatized crystals (Figure 1B), and the structure of Hia_{992–1098} was solved using molecular replacement and the Hia_{1022–1098} structure as a search model.

The Hia monomer

The Hia_{992–1098} structure and the Hia_{1022–1098} structure are nearly identical, containing an N-terminal helix, α 1, followed by a four-stranded β -sheet (β 1– β 4), and differing only by the length of the N-terminal helix (residues 998 (the first residue for which electron density was observed) to 1036 for Hia_{992–1098} and residues 1022–1036 for Hia_{1022–1098}) (Figures 1C and 2A). The helix packs against the inner face of the β -sheet from residue 1025 to residue 1036; this interface is primarily hydrophobic and extends to about 422 Å². The part of the helix that packs against the β -sheet (residues 1025–1036) does not coil, whereas the part that extends beyond the β -barrel (residues 998–1024) exhibits a distinct coil (see section below). Remarkably, the presumably large electrostatic dipole moment generated by the helix is neutralized by a basic residue (R1077; Figure 2F) in the β 3-strand that faces inwards towards the center of the trimeric barrel to position the guanido group just under the helix. In the trimer (see below), this residue would have a profound destabilizing effect if it did not play a very specific function in dealing with the negative charge emanating from the helix dipole. Another distinct feature of the protein surface in the helical region is the strongly basic patch at the C-terminus of the coiled part (K1019, K1022, and R1023; Figure 2E). A large loop links α 1 to β 1 (residues 1037–1044; the α 1– β 1 loop). The β 1-strand (residues 1045–1055) makes an angle of about 25° with α 1. The β 1– β 2 loop is short and is the first loop of the β -sheet facing the extracellular milieu. The β 2-strand (residues 1058–1069) follows, and the β 2– β 3 loop forms a primarily charged structure at the periplasm-facing side of the structure. The β 3-strand (residues 1074–1083) harbors R1077, while the β 3– β 4 loop, the second loop to face the extracellular milieu, contains a lysine (K1087), completing the basic patch formed by the basic residues at the base of the protruding part of helix α 1 mentioned above. Finally, the β 4-strand directs the C-terminal sequence towards the periplasmic side of the structure. In the vicinity of the C-terminus, two residues appear to form an aromatic patch, namely Trp 1098 (the very C-terminal residue) and Tyr 1096, presumably facilitating insertion into the membrane.

The Hia trimer

Both crystals of the Hia translocator domain described here contain two trimers of Hia in the asymmetric unit with two completely different packing orders. Also, both proteins are trimeric in solution (Figure 1A). Thus, the trimers seen in the asymmetric unit very likely represent the Hia trimeric translocator domain in solution.

The trimer is formed through juxtaposition of the β -sheets and stacking of the α 1-helices of three subunits along a three-fold axis running approximately through the center of the trimeric structure (Figure 2B–D). This assembly forms a

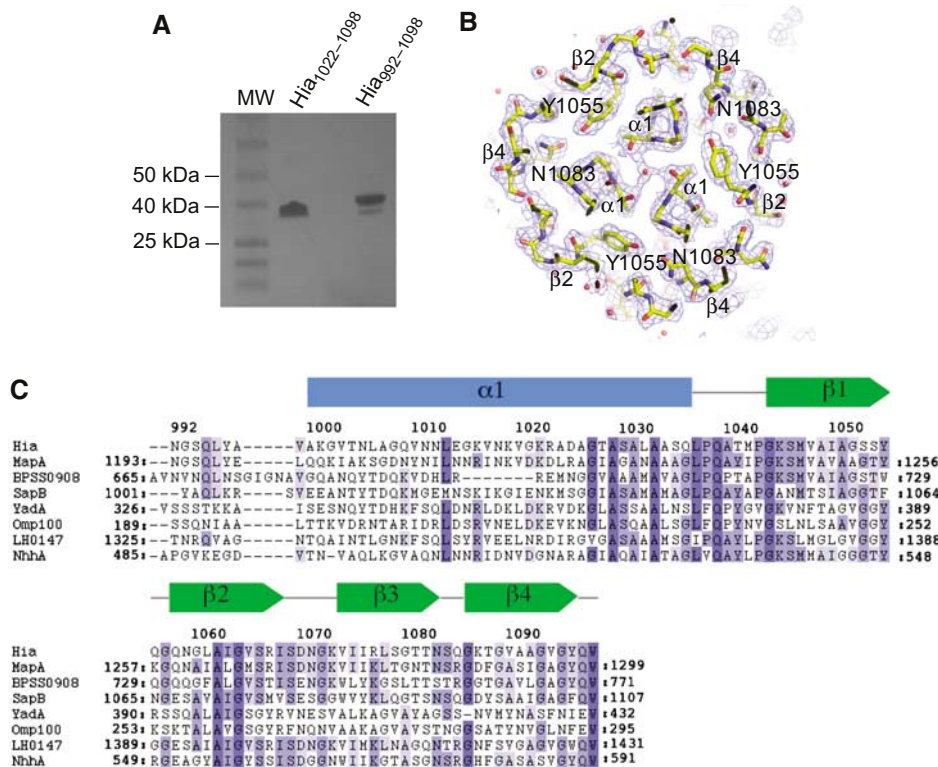


Figure 1 Purification, structure determination, and secondary structures of the Hia₁₀₂₂₋₁₀₉₈ and Hia₉₉₂₋₁₀₉₈ autotranslocating units of Hia. (A) SDS-PAGE of purified Hia₁₀₂₂₋₁₀₉₈ and Hia₉₉₂₋₁₀₉₈. Each protein (10 μg) was boiled for 5 min in a loading buffer containing 50 mM Tris pH 8.0, 5 mM 2-β-mercaptoethanol, 10% (v/v) glycerol, 2% (w/v) SDS, and 0.05% bromophenyl blue. MW: molecular weight markers. The molecular weight markers in the vicinity of the Hia bands are indicated. (B) Experimental MAD electron density after solvent flattening of the region of the Hia₁₀₂₂₋₁₀₉₈ structure around Tyr1055 contoured at 1.0σ. (C) Structure-based alignment of the translocating domain of several trimeric autotransporters. MapA, BPSS0908, SapB, YadA, Omp100, LH0147, and NhhA are the names for Hsf-like adhesin from *Pasteurella multocida*, putative surface-exposed protein from *Burkholderia pseudomallei*, putative autotransporter from *Salmonella typhi*, YadA adhesin from *Yersinia pseudotuberculosis*, outer membrane protein 100 from *Haemophilus actinomycetemcomitans*, adhesin from *Escherichia coli*, NhhA adhesin from *N. meningitidis*, respectively. Numbering on top of the sequence alignment is that of Hia. Secondary structure (green for β-strand and blue for α-helices) is shown above the Hia numbering. Dark and light blue boxes indicate strictly or moderately conserved residues, respectively. Figures 1–4 and Figure 5A were generated using PyMol (www.pymol.org).

12-stranded β-barrel surrounding three α-helices. The β-barrel alone is 29 Å in diameter and 36 Å deep (Figure 3). The pore inside the barrel is 18 Å in diameter. In the Hia₁₀₂₂₋₁₀₉₈ construct, the interface between subunits is mostly formed by the β1–β4 strand–strand and α1–α1 helix–packing interactions between juxtaposed subunits (Figures 2B and 3A). Indeed, of the 1072 Å² of buried surface area between two subunits, 621 Å² are between the β1- and β4-strands, and 191 Å² are between α1-helices. The extra sequence added to the N-terminus in the Hia₉₉₂₋₁₀₉₈ construct adds 395 Å² of buried surface area (between two subunits) towards the stability of the complex. The minimal sequence that would give rise to a trimer appears to correspond exactly to Hia₁₀₂₅₋₁₀₉₈, as residue 1025 is the first α1 residue to be buried into the β-barrel. Thus, the Hia₁₀₂₂₋₁₀₉₈ construct appears to barely exceed the minimally stable translocating domain. It is remarkable that such a minimal structure is so refractory to SDS denaturation. It is likely that the large interface, with a large polar main chain–main chain hydrogen-bonding component and a closed-circle structure, provides the resistance to SDS denaturation. Other elements of the structure appear to make significant contributions to the subunit–subunit interface, notably the α1–β1 loop. Indeed, the α1–β1 loops in all three subunits contact each other

extensively to form a primarily polar interface of 88 Å² of buried surface area (calculated between two subunits).

The electrostatic potential of the interior surface of the monomer (Figure 3B) is overall neutral, with the surfaces contributing to the α1/α1 interface being mostly hydrophobic and those contributing to the α1–β1/α1–β1 interface being mostly polar (see above). In the trimer, the outer surface of the barrel is mostly hydrophobic, consistent with its insertion in the membrane (Figure 3B). These surfaces are also bound to C8E4 molecules, the detergent that was used for crystallization. Interestingly, a ring of basic residues (mentioned above; Figure 2E) is observed where helix α1 emerges on the extracellular side of the barrel (K1019, K1022, R1023, and K1087). Whether this ‘crown’ of basic residues play a role in coordinating the lipid headgroups of the membrane is difficult to evaluate, as the positive electrostatic potential generated by this cluster of amino acids appears to be somewhat neutralized by E1025.

Strikingly, the trimeric 12-stranded β-barrel of the Hia translocator is very similar to the monomeric NalP translocator (Figure 4A and B), with the same number of strands, the same angle of the strands relative to the axis of the barrel, and very similar pore dimensions. In contrast, the 12-stranded β-barrel of the Hia translocator differs substantially

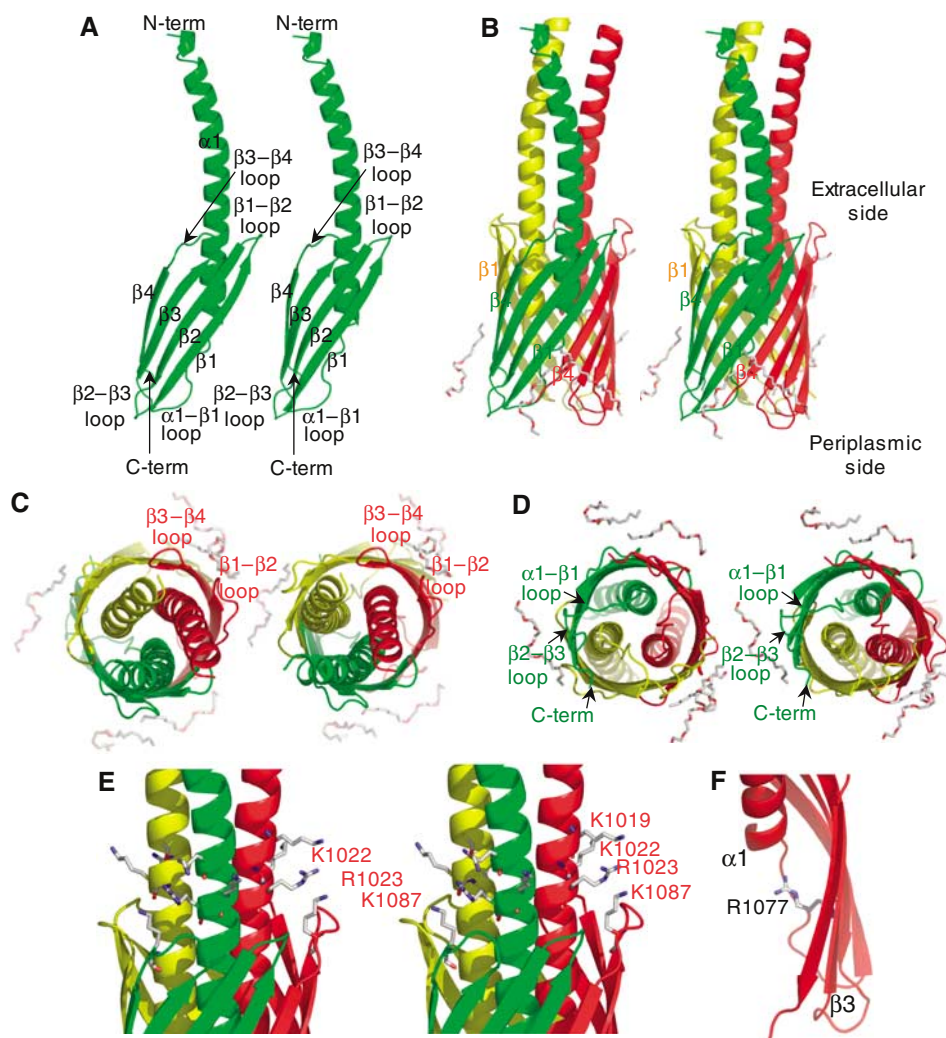


Figure 2 Structure of the Hia₉₉₂₋₁₀₉₈ and Hia₁₀₂₂₋₁₀₉₈ monomers and trimers. **(A)** Stereo ribbon diagram of the Hia₁₀₂₂₋₁₀₉₈ monomer. Secondary structures and loop regions are labeled. **(B)** Stereo ribbon diagram of the Hia₉₉₂₋₁₀₉₈ trimer. The three subunits are colored green, yellow and red. Relevant secondary structures are labeled. **(C)** Stereo ribbon diagram of the Hia₉₉₂₋₁₀₉₈ trimer viewed from the extracellular milieu side. Color coding is as in **(B)**. Relevant loops and secondary structures are labeled. **(D)** Stereo ribbon diagram of the Hia₉₉₂₋₁₀₉₈ trimer viewed from the periplasmic side. Color coding is as in **(B)**. Relevant loops and secondary structures are labeled. **(E)** Detailed molecular feature of the Hia₉₉₂₋₁₀₉₈ trimer focusing of the basic residue crown surrounding the emerging part of helix $\alpha 1$. The trimer is in ribbon representation with color coding as in **(B)**, while the highlighted residues are in stick representation color-coded in gray, red, and blue for carbon, oxygen, and nitrogen atoms, respectively. **(F)** Detailed molecular feature of the Hia₉₉₂₋₁₀₉₈ monomers focusing of Arg1077. A potential role for Arg1077 is to neutralize the dipole moment of the large N-terminal helix. During translocation, Arg1077 would need to move away and only assume the conformation observed in the crystal structure once translocation has occurred.

from the structure of TolC, a 12-stranded β -barrel that is functionally distinct from Hia (see Supplementary Figure 1). Thus, the similarity of Hia and NalP is significant and suggests a common translocation mechanism in both monomeric and trimeric autotransporters.

Functional analysis of the translocator domain of Hia

The structure of the trimeric translocator domain of Hia defines a minimal domain consisting of residues 1025–1098 and identifies the regions that appear to be important for trimer formation. However, the roles of the various regions of the structure in both translocation and trimer formation remain to be tested. Accordingly, we generated constructs consisting of the Hia signal peptide, the HAT epitope, and progressively shorter regions of the Hia C-terminus (Figure 5A), removing all but half of the membrane spanning portion of helix $\alpha 1$ (Hia₁₀₃₂₋₁₀₉₈), removing helix $\alpha 1$ entirely

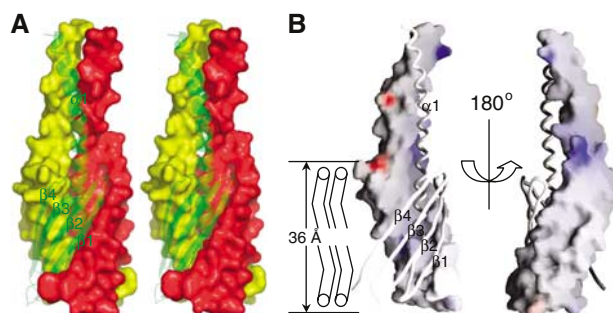


Figure 3 Surface of the Hia translocating domain. **(A)** Stereo surface diagram of two Hia₉₉₂₋₁₀₉₈ subunits of the trimer with the third subunit shown as a transparent ribbon. **(B)** Surface potential of the Hia₉₉₂₋₁₀₉₈ monomer with the adjacent subunit shown in a ribbon representation. Left and right panels differ by a 180° rotation. The membrane inserted part of the surface is schematically represented by a lipid bilayer at the left of the left panel.

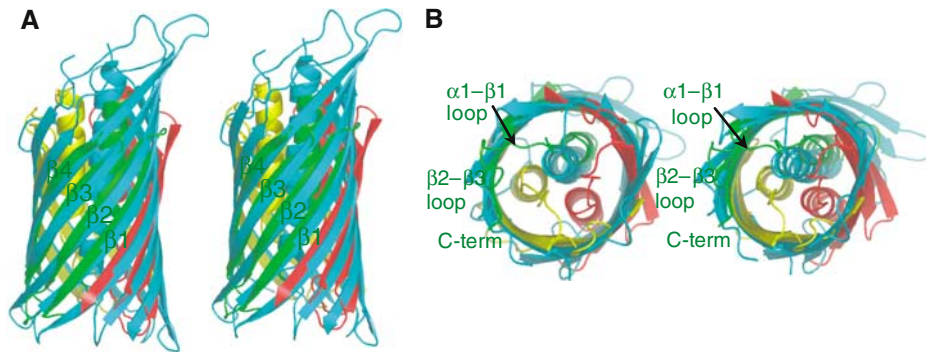


Figure 4 Comparison of the trimeric and monomeric translocator domains of Hia and NalP and detailed molecular features of the Hia translocator domain. (A) Stereo ribbon diagram of the superimposed Hia (same color coding as in Figure 2B) and NalP (color-coded in cyan) translocator domains. The orientation is the same as in Figure 2B. (B) Same as (A) but viewed from bottom (periplasm side).

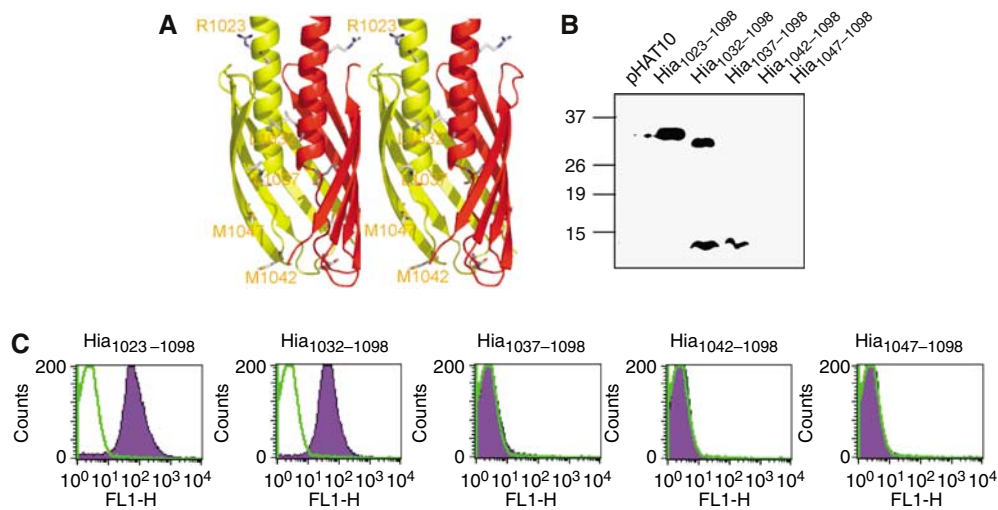


Figure 5 Analysis of Hia C-terminal regions. (A) Location of the various fragments of the Hia C-terminus made in this study. (B) Outer membrane proteins detected by Western blot with a polyclonal antiserum directed against the HAT epitope. (C) Flow cytometry analysis of the HAT epitope on the bacterial surface performed with a polyclonal antiserum directed against the HAT epitope. Strain *E. coli* BL21 (DE3)/pHAT10 (vector) was used as the negative control, and the corresponding histogram was superimposed (solid line) on results with *E. coli* expressing the indicated proteins (filled histograms).

(Hia₁₀₃₇₋₁₀₉₈), removing both helix $\alpha 1$ and the $\alpha 1$ - $\beta 1$ loop (Hia₁₀₄₂₋₁₀₉₈), or removing helix $\alpha 1$, the $\alpha 1$ - $\beta 1$ loop, and a part of strand $\beta 1$ (Hia₁₀₄₇₋₁₀₉₈; a construct designed to provide a negative control, as it is unlikely to fold properly). The resulting proteins were expressed in *E. coli* BL21(DE3), and outer membrane proteins were isolated, resolved under standard denaturing conditions, and examined by immunoblot analysis. As shown in Figure 5B, Hia₁₀₂₃₋₁₀₉₈ migrated at a molecular mass of 33 kDa, consistent with a trimer. In contrast, Hia₁₀₃₂₋₁₀₉₈ migrated at ~ 30 and ~ 10 kDa, corresponding to a trimer and monomer, respectively; Hia₁₀₃₇₋₁₀₉₈ migrated entirely at ~ 9.5 kDa, corresponding to a monomer; and Hia₁₀₄₂₋₁₀₉₈ and Hia₁₀₄₇₋₁₀₉₈ were not detectable at all, suggesting degradation. These experiments confirm that helix $\alpha 1$ is an important contributor to the stability of the trimer. Importantly, analysis by flow cytometry of bacteria expressing these proteins revealed that only Hia₁₀₂₃₋₁₀₉₈ and Hia₁₀₃₂₋₁₀₉₈ were able to translocate the HAT epitope to the bacterial surface, suggesting that Hia₁₀₃₂₋₁₀₉₈ retains translocator activity (Figure 5C). Taken together, these observations demonstrate that stable trimerization is essential for

translocator activity, as a construct partitioning between trimers and monomers on an SDS-PAGE gel (Hia₁₀₃₂₋₁₀₉₈) has reduced but significant translocator activity and a construct that migrates only as a monomer on an SDS-PAGE gel (Hia₁₀₃₇₋₁₀₉₈) has no appreciable translocator activity.

Hia residues 1023-1046 clearly form an important region of the structure, as deletions of part of this region result in reduced or impaired trimerization and translocator activity. Thus, we next performed a more detailed investigation of this region by scanning alanine mutagenesis throughout Hia residues 1023-1046. However, instead of single-site mutations, we elected to alter several residues at once. Our goal was to identify a group of residues that play a role in trimerization or translocation, then examine these residues individually. Ultimately, we generated five mutant constructs in the Hia₁₀₂₃₋₁₀₉₈ $\alpha 1$ and $\alpha 1$ - $\beta 1$ loop region: Hia_{1023-1098/1023-1031} Alanine, Hia_{1023-1098/1030-1037} Alanine, Hia_{1023-1098/1037-1042} Alanine, and Hia_{1023-1098/1042-1046} Alanine, and Hia_{1023-1098/1035-1045} Alanine. In addition, we examined the effect of deletion of residues 1037-1042 (the $\alpha 1$ - $\beta 1$ loop), a region that is highly conserved between Hia-like

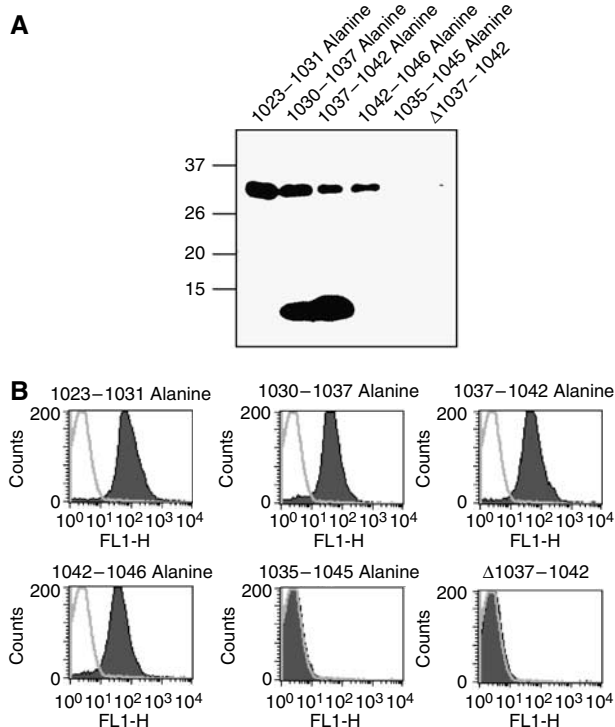


Figure 6 Analysis of group mutations within the Hia C-terminus. (A) Outer membrane proteins detected by Western blot with a polyclonal antiserum directed against the HAT epitope. (B) Flow cytometry analysis of the HAT epitope on the bacterial surface performed with a polyclonal antiserum directed against the HAT epitope. Strain *E. coli* BL21(DE3)/pHAT10 (vector) was used as the negative control, and the corresponding histogram was superimposed (solid line) on results with *E. coli* expressing the indicated Hia derivative. In panels (A) and (B), the indicated constructs are all derivatives of Hia₁₀₂₃₋₁₀₉₈.

autotransporters. Western analysis of outer membrane proteins isolated from *E. coli* BL21(DE3) expressing these proteins revealed that Hia_{1023-1098/1023-1031} Alanine and Hia_{1023-1098/1042-1046} Alanine migrated at a molecular mass consistent with a trimer, while Hia_{1023-1098/1030-1037} Alanine and Hia_{1023-1098/1037-1042} Alanine were detected as two bands, corresponding to a trimer and a monomer (Figure 6A). Neither Hia_{1023-1098/1035-1045} Alanine nor Hia_{1023-1098/Δ1037-1042} was detectable in the outer membrane. As shown in Figure 6B, flow cytometry analysis of bacteria expressing these proteins revealed translocation of the N-terminal HAT epitope by all mutants except Hia_{1023-1098/1035-1045} Alanine and Hia_{1023-1098/Δ1037-1042}. Consistent with these results, analogous mutations made in the context of full-length Hia demonstrated little effect on adherence, with the exception of Hia₁₀₃₅₋₁₀₄₅ Alanine and Hia_{Δ1037-1042}, which were unable to mediate adherence (data not shown). These results confirm the importance of the α 1-helix in trimer formation and establish the importance of residues in the α 1- β 1 loop region in trimer stability and translocator activity.

Discussion

In this study, we have elucidated the structure of the C-terminal 107 amino acids of the Hia adhesin, providing the first known structure of a trimeric autotransporter translocator domain. The Hia translocator domain is a trimeric β -barrel

with 12 β -strands, including four β -strands from each sub-unit. Remarkably, the β -pore of the Hia translocator is very similar to that of the monomeric NalP translocator (Figure 4). However, the NalP structure was obtained from a protein that was purified as cytoplasmic inclusion bodies, denatured in urea, and then renatured (Oomen *et al*, 2004), raising doubts about whether the monomeric structure was native. In our study, the Hia translocator domain was extracted from the bacterial outer membrane and was not submitted to denaturation/renaturation. As a consequence, the structures of Hia₉₉₂₋₁₀₉₈ and Hia₁₀₂₂₋₁₀₉₈ are likely the native forms adopted by the Hia autotransporter. Thus, the fact that the β -barrel structures of Hia and NalP are virtually superimposable suggests that the structure of NalP is indeed that of the native form of the translocator unit of NalP, and that this type of β -barrel provides a translocator pore framework common to all autotransporters, whether conventional or trimeric. The primary distinction is that the NalP translocator is traversed by a single α -helix, while the Hia translocator is traversed by three α -helices.

In considering the mechanism of translocation of the passenger domain of Hia and other trimeric autotransporters, the observation that the three α -helices in Hia₉₉₂₋₁₀₉₈ traverse the inside rather than being outside of the β -barrel argues strongly against the multimeric assembly model. In this model, the translocating pore is a multimeric structure that is formed by multiple β -barrels. One implication of this model is that the sequences N-terminal to each β -barrel should be located outside each individual β -pore and not inside, as observed in the structures of the NalP and Hia translocator units.

The finding that the α -helices in Hia₁₀₂₃₋₁₀₉₈ and Hia₉₉₂₋₁₀₉₈ are presented on the extracellular side of the β -barrel without the need for the immediate N-terminus of Hia argues strongly against the threading model as well. In this model, the N-terminus of the passenger domain is a targeting sequence that directs the passenger domain to the β -barrel and initiates the translocation process, implying that the N-terminal end of Hia should be essential for translocation. Accordingly, in truncates of Hia that lack the N-terminus, translocation should not occur, predictions that conflict with our results.

In contrast to the situation with the multimeric assembly and the threading models, the structures of Hia₁₀₂₃₋₁₀₉₈ and Hia₉₉₂₋₁₀₉₈ appear to be consistent with the hairpin model of translocation. This model proposes that the sequence immediately N-terminal to the β -domain forms a hairpin structure in the β -barrel and allows the rest of the passenger domain sequence to slide from the periplasm to the bacterial surface. Oliver *et al* (2003a,b) have identified a conserved stretch N-terminal to the translocator domain that is present in many conventional autotransporters and may correspond to the hairpin. Based on work on the *Bordetella pertussis* BrkA conventional autotransporter, this region has chaperone activity and appears to initiate folding of the passenger domain (Oliver *et al*, 2003a). Of note, this region is absent in the trimeric autotransporter subfamily, suggesting that a nonconserved sequence may possess this function in trimeric autotransporters.

In the hairpin model, the hairpin structure is a temporary structure that ceases to exist upon completion of the translocation process, ultimately assuming the helical configuration observed in the structures of the monomeric NalP and trimeric Hia translocator units (note that this model recon-

ciles the apparent contradiction that helix $\alpha 1$ is both important in trimer stability and passenger domain translocation). A hairpin comprised of two extended polypeptides has the shape of a long block with a width of $10 \times 8 \text{ \AA}$, occupying a space similar to that of an α -helix, which has a cylindrical shape and a diameter of 10 \AA . Thus, the dimensions of the Hia translocator channel are sufficient to accommodate three hairpins, if all three passenger domains are translocated simultaneously, or one to two hairpins and one to two α -helices, if the three passenger domains are translocated sequentially. However, the tightness of the fit makes it somewhat difficult to envision polypeptide movement through the pore.

Another model invokes the assistance of the Omp85 outer membrane protein in translocation of the passenger domain to the bacterial surface. Omp85-like proteins facilitate formation of β -barrel proteins and insertion of β -barrels into the outer membrane (Voulhoux *et al*, 2003). It has been proposed that the C-terminal β -barrel domain of autotransporters serves as a targeting signal for Omp85, and that the passenger domain of autotransporters is secreted through the Omp85 pore (Oomen *et al*, 2004). However, the structure and the functional data presented here indicate that the translocator unit forms an obligate trimer through which the passenger domain passes, arguing against translocation through the Omp85 pore. Another possibility is that Omp85 serves to stabilize an open form of the autotransporter β -barrel, creating a larger pore that permits translocation of the passenger domain, in particular those regions of the passenger domain that may be partially folded in the periplasm, as observed in work on the *N. meningitidis* IgA1 protease, the *Shigella flexneri* IcsA, or the *E. coli* EspP monomeric autotransporters (Brandon and Goldberg, 2001; Veiga *et al*, 2002; Skillman *et al*, 2005). Once the passenger domain is on the bacterial surface, the pore may adopt its final 'closed' conformation as seen in the structures of the translocator units of NalP and Hia. Since Omp85 recognizes β -barrel proteins, it is likely that it interacts with some loosely associated trimer of Hia translocator units formed in the periplasm prior to membrane insertion.

As an alternative, perhaps more speculative, model for Hia surface localization, Hia may assemble and fold in the periplasm and may then be inserted into the outer membrane as a fully functional molecule by Omp85. Assembly and folding in the periplasm would avoid the need for a translocation event through the β -barrel pore and for hairpin formation.

Materials and methods

Bacterial strains, plasmids, and culture conditions

Laboratory strains employed in these studies include *E. coli* B834(DE3) (methionine auxotroph, Novagen), *E. coli* DH5 α (Life Technologies), *E. coli* XL-1 Blue (Invitrogen), *E. coli* BL21 (DE3), and *H. influenzae* strain DB117 (*rec1*⁻).

The plasmid pHMW8-7 contains *hia* from *H. influenzae* strain 11 in pT7-7. The plasmid pHAT::HiaSS contains a portion of the *hia* promoter and coding sequence for the Hia signal sequence upstream of the HAT epitope in pHAT10 (Clontech). The plasmids pHAT::Hia₁₀₂₃₋₁₀₉₈, and pHAT::Hia₁₀₄₇₋₁₀₉₈ encode the HAT epitope fused to Hia₁₀₂₃₋₁₀₉₈ and Hia₁₀₄₇₋₁₀₉₈, respectively (Surana *et al*, 2004).

E. coli strains were grown on Luria-Bertani (LB) agar or in LB broth and stored at -80°C in LB broth with 50% glycerol. *H. influenzae* strains were grown as described previously (Anderson *et al*, 1972).

Recombinant DNA methods

DNA ligations, restriction endonuclease digestions, gel electrophoresis, and PCR were performed according to standard techniques.

Plasmids were introduced into *E. coli* by electroporation. *H. influenzae* was transformed using the MIV method of Herriott *et al* (1970).

Construction of plasmids used in this study

To generate constructs for crystallography experiments, pHMW8-7 was used as a template to amplify the coding sequence for Hia residues 992–1098 and 1022–1098 with engineered *Bsa*I sites at both ends of the amplicons. The resulting PCR fragments were digested with *Bsa*I and ligated into *Bsa*I-digested pASK-IBA12 (IBA), generating pASK-Hia₉₉₂₋₁₀₉₈ and pASK-Hia₁₀₂₂₋₁₀₉₈.

To generate constructs encoding the HAT epitope fused to varying Hia C-terminal regions, pHMW8-7 was used as a template to amplify the coding sequence for Hia residues 1032–1098, 1037–1098, and 1042–1098 with an engineered *Sall* site at the 5'-end and an engineered *Bam*HI site at the 3'-end of the amplicons. These PCR fragments were digested with *Sall* and *Bam*HI and were then ligated into *Sall*-*Bam*HI-digested pHAT::HiaSS, generating pHAT::Hia₁₀₃₂₋₁₀₉₈, pHAT::Hia₁₀₃₇₋₁₀₉₈, and pHAT::Hia₁₀₄₂₋₁₀₉₈.

To generate plasmids encoding Hia with alanine substitutions or deletion of Hia residues 1037–1042, we used the QuikChange XL site-directed mutagenesis kit (Stratagene) and followed the manufacturer's recommendations. Mutant plasmids were initially transformed into *E. coli* XL-1 Blue, and nucleotide sequencing was performed. Plasmid clones with the desired nucleotide sequence were then transformed into *E. coli* BL21 (DE3).

Expression, purification, and crystallization

Expression, purification, and crystallization of Hia₉₉₂₋₁₀₉₈ and Hia₁₀₂₂₋₁₀₉₈ are described in Supplementary data.

Table 1 Data collection and structure refinement statistics

	Hia ₁₀₂₂₋₁₀₉₈	Hia ₉₉₂₋₁₀₉₈
<i>Data collection</i>		
Derivative	Se Met	Native
Source/station ^a	ID14.4	ID14.3
Wavelength (Å)	0.979	0.931
Resolution range (Å)	20–2.0	20–2.3
Observations ($I/\sigma(I) > 0$)	275 847	136 757
Unique reflections ($I/\sigma(I) > 0$)	38 510	37 888
High-resolution shell (Å)	2.11–2.00	2.42–2.30
R_{sym} (%) ^{b,c}	13.4 (48.1)	12.4 (51.8)
$\langle I/\sigma(I) \rangle^c$	12.0 (3.9)	8.3 (1.9)
Completeness ^c (%)	99.9 (100)	98.4 (93)
Redundancy ^c :	7.2 (7.3)	3.6 (2.4)
<i>Structure refinement</i>		
Resolution range (Å)	20–2.0	20–2.3
R-factor (%)	19.3	21.7
R-factor (high-resolution shell) ^d	21.1	27.3
R_{free} (%) ^e	21.5	25.3
R_{free} (high-resolution shell)	26.4	32.9
Total number of non-hydrogen atoms	3432	4606
Protein atoms	3246	4171
C8E4 molecules	0	5
Water molecules	186	331
R.m.s. deviations: ^f		
Bond length (Å)	0.005	0.011
Bond angle (deg)	0.847	1.425
Main chain B-factors (Å ²)	0.148	0.309
Side chain B-factors (Å ²)	0.392	1.280
Wilson B-factor (Å ²)	17.3	31.2
Average B-factor protein atoms (Å ²)	7.6	27.3
Average B-factor solvent atoms (Å ²)	21.0	30.6
Average B-factor (Å ²), occupancy of C8E4		27.7, 1

^aBeamline designations refer to the ESRF, Grenoble, France.

^b $R_{\text{sym}} = \sum (I - \langle I \rangle) / \sum I$.

^cOverall, high-resolution shell in parentheses.

^dHigh-resolution shell: 2.05–2.00 Å, 2.36–2.30 Å.

^e R_{free} calculated using 5% of total reflections omitted from refinement.

^fR.m.s. deviations report root mean square deviations from ideal bond lengths/angles and of B-factors between bonded atoms (Engh and Huber, 1991).

Data collection and phasing

Single anomalous diffraction (SAD) data of selenomethionine-derivatized Hia_{1022–1098} were recorded at selenium edge (0.979 Å) on ID14.4 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) using inverse beam geometry. Diffraction data for Hia_{992–1098} native crystals were recorded on ID14.3 at ESRF. Hia_{1022–1098} SAD data were integrated and scaled using the programs DENZO and SCALEPACK (Otwinowski and Minor, 1997). The programs MOSFLM and SCALA (CCP4, 1994) were used for Hia_{992–1098} diffraction data. The statistics of data collection are reported in Table I.

SAD method was used to phase Hia_{1022–1098}. A total of 12 Se positions were determined by anomalous difference Patterson analysis using data between 20 and 2.4 Å resolution (program SOLVE) (Terwilliger and Berendzen, 1999). The parameters of the heavy atom sites were refined using the program SHARP (deLaFortelle and Bricogne, 1997). An interpretable map was obtained after solvent flattening using the program DM (CCP4, 1994) (Figure 1B). The extended phases (2.0 Å) allowed automatic tracing of the backbone using the program ARP/Warp (CCP4, 1994), producing a σ_A -weighted $2F_o - F_c$ map of excellent quality into which side chains were built manually using the program COOT (CCP4, 1994). Hia_{992–1098} was phased by molecular replacement using the program PHASER (CCP4, 1994) and the refined structure of Hia_{1022–1098} as the search model. The programs ARP/WARP and COOT were used as mentioned above to build the missing residues (998–1021).

Structure refinement

The structure models of Hia_{1022–1098} and Hia_{992–1098} were refined by conjugate gradient minimization using the program REFMAC5 (CCP4, 1994) with intermittent manual rebuilding, refining individual B-factors applying a TLS correction (six TLS group, 126

parameters) (Winn *et al*, 2001). The final model of Hia_{1022–1098} contains residues 1022–1098, one additional residue, a Ser, coming from the vector, and 186 water molecules. The final model of Hia_{992–1098} contains residues 998–1098, five C8E4 detergent molecules, and 331 water molecules. Ramachandran statistics (PROCHECK) (Laskowski *et al*, 1993) on the Hia_{992–1098} structure indicate that 97.5% of the atoms are in the most favored region, and 2.5% are in the additionally allowed regions. Similar statistics (95.7%, 4.3%) were observed for the atoms in Hia_{1022–1098} structure. The detailed structure refinement statistics are reported in Table I. Coordinates of the Hia_{1022–1098} and Hia_{992–1098} structures have been deposited to the PDB (entry codes 2GR8 and 2GR7, respectively).

Cell fractionation, protein analysis, and flow cytometry analysis of bacteria

See Supplementary data for this part of the Materials and methods section.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online.

Acknowledgements

This work was funded by grant 070001 from the Wellcome Trust (GW) and by NIH grants RO1-AI44167 (JWS) and T32-AI07172 (NKS). NKS is a member of the Medical Scientist Training Program at Washington University. We thank the personnel of beamlines ID14.4 and ID14.3 at ESRF (Grenoble) for help during data collection and Sue Grass for technical assistance.

References

- Anderson P, Johnston Jr RB, Smith DH (1972) Human serum activities against *Haemophilus influenzae*, type b. *J Clin Invest* **51**: 31–38
- Brandon LD, Goldberg MB (2001) Periplasmic transit and disulfide bond formation of the autotransported *Shigella* protein IcsA. *J Bacteriol* **183**: 951–958
- CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Cryst D* **50**: 760–763
- Cotter SE, Surana NK, St Geme III JW (2005) Trimeric autotransporters: a distinct subfamily of autotransporter proteins. *Trends Microbiol* **13**: 199–205
- deLaFortelle E, Bricogne G (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol* **276**: 472–494
- Desvaux M, Parham NJ, Henderson IR (2004) The autotransporter secretion system. *Res Microbiol* **155**: 53–60
- Engh RA, Huber R (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr A* **47**: 392–400
- Henderson IR, Nataro JP (2001) Virulence functions of autotransporter proteins. *Infect Immun* **69**: 1231–1243
- Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D (2004) Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* **68**: 692–744
- Herriott RM, Meyer EM, Vogt M (1970) Defined nongrowth media for stage II development of competence in *Haemophilus influenzae*. *J Bacteriol* **101**: 517–524
- Hoiczynk E, Roggenkamp A, Reichenbecher M, Lupas A, Heesemann J (2000) Structure and sequence analysis of *Yersinia* YadA and *Moraxella* UspAs reveal a novel class of adhesins. *EMBO J* **19**: 5989–5999
- Jacob-Dubuisson F, Fernandez R, Coutte L (2004) Protein secretion through autotransporter and two-partner pathways. *Biochim Biophys Acta* **1694**: 235–257
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* **26**: 283–291
- Nummelin H, Merckel MC, Leo JC, Lankinen H, Skurnik M, Goldman A (2004) The *Yersinia* adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. *EMBO J* **23**: 701–711
- Oliver DC, Huang G, Fernandez RC (2003a) Identification of secretion determinants of the *Bordetella pertussis* BrkA autotransporter. *J Bacteriol* **185**: 489–495
- Oliver DC, Huang G, Nodel E, Pleasance S, Fernandez RC (2003b) A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain. *Mol Microbiol* **47**: 1367–1383
- Oomen CJ, van Ulsen P, van Gelder P, Feijen M, Tommassen J, Gros P (2004) Structure of the translocator domain of a bacterial autotransporter. *EMBO J* **23**: 1257–1266
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**: 307–326
- Pallen MJ, Chaudhuri RR, Henderson IR (2003) Genomic analysis of secretion systems. *Curr Opin Microbiol* **6**: 519–527
- Roggenkamp A, Ackermann N, Jacobi CA, Truelzsch K, Hoffmann H, Heesemann J (2003) Molecular analysis of transport and oligomerization of the *Yersinia enterocolitica* adhesin YadA. *J Bacteriol* **185**: 3735–3744
- Skillman KM, Barnard TJ, Peterson JH, Ghirlando R, Bernstein HD (2005) Efficient secretion of a folded protein domain by a monomeric bacterial autotransporter. *Mol Microbiol* **58**: 945–958
- Surana NK, Cutter D, Barenkamp SJ, St Geme III JW (2004) The *Haemophilus influenzae* Hia autotransporter contains an unusually short trimeric translocator domain. *J Biol Chem* **279**: 14679–14685
- Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Cryst D* **55**: 849–861
- Veiga E, Sugawara E, Nikaido H, de Lorenzo V, Fernandez LA (2002) Export of autotransported proteins proceeds through an oligomeric ring shaped by C-terminal domains. *EMBO J* **21**: 2122–2131
- Voulhoux R, Bos MP, Geurtsen J, Mols M, Tommassen J (2003) Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**: 262–265
- Winn MD, Isupov MN, Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Cryst D* **57**: 122–133
- Wollmann P, Zeth K, Lupas AN, Linke D (2006) Purification of the YadA membrane anchor for secondary structure analysis and crystallization. *Int J Biol Macromol* (in press)
- Yeo HJ, Cotter SE, Laarmann S, Juehne T, St Geme III JW, Waksman G (2004) Structural basis for host recognition by the *Haemophilus influenzae* Hia autotransporter. *EMBO J* **23**: 1245–1256