Published in final edited form as:
J Mol Biol. 2007 August 31; 371(5): 1304–1314.

Differences in the electrostatic surfaces of the type III secretion needle proteins PrgI, BsaL, and MxiH

Yu Wang, **Andrew N. Ouellette**, **Chet W. Egan**, **Thenmalarchelvi Rathinavelan**, **Wonpil Im**, and **Roberto N. De Guzman***
Department of Molecular Biosciences, The University of Kansas, 1200 Sunnyside Avenue, Lawrence, Kansas 66045

Abstract

Gram-negative bacteria use a needle-like protein assembly, the type III secretion apparatus, to inject virulence factors into target cells to initiate human disease. The needle is formed by the polymerization of ~120 copies of a small acidic protein that is conserved among diverse pathogens. We previously reported the structure of the BsaL needle monomer from *Burkholderia pseudomallei* by nuclear magnetic resonance (NMR) spectroscopy and others have determined the crystal structure of the *Shigella flexneri* MxiH needle. Here, we report the NMR structure of the PrgI needle protein of *Salmonella typhimurium*, a human pathogen associated with food poisoning. PrgI, BsaL, and MxiH form similar two helix bundles, however, the electrostatic surfaces of PrgI differ radically from those of BsaL or MxiH. In BsaL and MxiH, a large negative area is on a face formed by the helix α_1-α_2 interface. In PrgI, the major negatively charged surface is not on the "face" but instead is on the "side" of the two helix bundle, and only residues from helix α_1 contribute to this negative region. Despite being highly acidic proteins, these molecules contain large basic regions, suggesting that electrostatic contacts are important in needle assembly. Our results also suggest that needle packing interactions may be different among these bacteria and provide the structural basis for why PrgI and MxiH, despite 63% sequence identity, are not interchangeable in *S. typhimurium* and *S. flexneri*.

Keywords

Salmonella; type III secretion; PrgI; BsaL; MxiH; NMR; pathogenesis

INTRODUCTION

Many Gram-negative bacteria use the type III secretion apparatus (TTSA) to inject host altering effector proteins into eukaryotic cells to cause human disease. Among these bacteria is *Salmonella typhimurium*, which is one of the most common pathogens of food poisoning in the U.S. Salmonellosis contributes significantly to hospitalization and morbidity among susceptible individuals such as the elderly, children, and people with compromised immune systems caused by AIDS and chemotherapy.[1] Other pathogens that use a TTSA are *Shigella flexneri* (shigellosis),[2] *Yersinia pestis* (plague),[3] *Burkholderia pseudomallei* (melioidosis)[4] and *Pseudomonas aeruginosa* (pulmonary infections among cystic fibrosis patients).[5] No vaccines are currently approved for general use against any of these pathogens, and together

*Corresponding author: rdguzman@ku.edu, Phone: (785) 864 4923.

with the appearance of antibiotic resistant strains[6, 7] and the bioterrorism potential of *B. pseudomallei,*[8] these pathogens continue to pose major public health and safety concerns.

The TTSA is a protein assembly consisting of a base that spans both bacterial membranes and an external needle and tip complex (Fig. S1).[9] Electron microscopy of the *S. typhimurium* needle revealed a tube with dimensions 130 × 800 Å (width *vs*. length).[10] A similar structure exists on the surface of *S. flexneri* ,[11] *Y. enterocolitica* [12] and *P. aeruginosa*.[13] In *S. typhimurium*, the needle is a helical assembly of ~120 copies of PrgI,[14] which is homologous to BsaL in *B. pseudomallei* and MxiH in *S. flexneri*. Needle proteins are critical in bacterial virulence – *mxiH*[15] and *prgI*[16] null mutants are noninvasive and do not secrete any effectors. Recent mutational analysis has implicated the needle proteins in controlling the induction of type III secretion in *Shigella*[17] and *Yersinia.*[18] PrgI, BsaL, and MxiH are small acidic proteins that show sequence conservation around a central PxxP motif (where x = S, D, or N, Fig. 1). These proteins are predicted to contain α-helical secondary structure[17] and circular dichroism (CD) spectroscopy confirmed the α-helical nature of MxiH, PrgI, and BsaL.[19] Despite sequence conservation, these proteins have widely different thermal stabilities – with melting temperatures of 37°C for PrgI, 42°C for MxiH, and 54°C for BsaL.[19]

Deletion of the last five residues of PrgI,[19] MxiH,[19] and BsaL[20] prevented self-association and yielded monomeric proteins that allowed biophysical studies. The first reported atomic resolution structure of a needle monomer is the NMR structure of BsaL.[20] BsaL contains a central core domain flanked by dynamic regions that have residual α-helical conformation. Only about half of BsaL is ordered – the central region which includes the PxxP motif forms a two-helix bundle, and the rest of the molecule is partially folded or in random coil conformation.[20] Others have since completed the crystal structure of the *S. flexneri* MxiH needle protein which showed a disordered N-terminal region and a longer C-terminal helix.[21] Nevertheless, the structured region of BsaL can be superimposed to MxiH with a 3.2 Å C^{α} root-mean-square (RMS) deviation. These high resolution structures have provided a wealth of information, however, detailed understanding of how the needle proteins interact with each other and with the tip complex are currently unknown. Thus, major gaps exist regarding structures, functional dynamics, and protein-protein interactions of the needle proteins. We report here the NMR structure of the *S. typhimurium* needle protein PrgI.

RESULTS

NMR structure determination of PrgI$^{C\Delta5}$

The last five residues of PrgI were deleted to obtain monomeric PrgI$^{C\Delta5}$ similar to the approach used in the structure determinations[20, 21] of BsaL$^{C\Delta5}$ and MxiH$^{C\Delta5}$. Analysis of NMR datasets yielded near complete backbone and side chain assignments for PrgI$^{C\Delta5}$. We assigned 69 (out of 79) backbone amides (Fig. 2A), however, the amides of A2, T3, A35, A36, L43 and L44 are in rapid exchange with the solvent and could not be assigned. The C^{α}, C^{β}, and H^{α} secondary chemical shifts indicated the presence of three helices: V20-A35, P41-S62, and D70-E77 (Fig. S2). Assignment of 3D NOESY datasets yielded 1275 interproton distance restraints, or an average of 15 NOEs per residue (Table 1). Most of the NOEs are in the helical regions and all the long range NOEs are from the conserved hydrophobic residues in the central region of PrgI (Fig. S3A). Structure calculations using NMR-derived restraints (Table 1) and CYANA[22] followed by refinement in AMBER[23] yielded a well converged NMR structure of PrgI$^{C\Delta5}$ (Fig. 2B). The NMR structures yielded good structural statistics (Table 1), and over 85% of the residues are in the most favorable region of the Ramachandran plot.

Central region forms a two-helix bundle

Only the central region (V20-V65) of PrgI$^{C\Delta5}$ converged into a single family of NMR structures and formed an ordered two helix bundle (Fig. 2B). The regions (P4-G19 and K66-I75) flanking this two helix bundle lacked long range NOEs and did not converge, however, each of the 20 NMR structures showed helical conformation for P4-G19 and K66-I75 that is consistent with the NOE patterns (Fig. S3). The well defined two helix bundle is composed of the PxxP loop (P38- P40) and parts of helix α_1 (V20-A36) and α_2 (A42-V65). The helix α_1-α_2 interface (Fig. 3) is stabilized by conserved hydrophobic contacts of residues on helix α_1 (V20, L23, V27, A30, L34) and helix α_2 (L43, L44, Y47, Y54, L56). V20 contacts the methylene groups of Q24 and R58 and the phenyl ring of Y54 (Fig. 3B). L23 packs against the methylene groups of Q26, K50, E53, Y54 and Y57 (Fig. 3); and V27 is in a hydrophobic pocket formed by Q24, L31, Y47, L51, and Y54 (Fig. 3B). The methyl group of the conserved A30 residue is in the helix α_1-α_2 interface but does not pack directly with any helix α_2 residue (Fig. 3A), suggesting that a small amino acid is required in this position. Finally, L34 packs with helix residues L31, L44, and Y47 and the PxxP loop residue P38 (Fig. 3B).

The two helix bundle is polar

The surface of the two helix bundle is polar due to residues on helix α_1 (N22, D21, Q24, T25, Q26, E29, K33, and D32), helix α_2 (Q48, S49, K50, S52, E53, N55, R58, N59, and S62), and the PxxP region (K27, S39, and D40) (Fig. 3). Most of these polar residues are pointed away from the two helix bundle and therefore are not needed in stabilizing the hydrophobic core of the two helix bundle. However, several polar residues are identical (Q48, N59, and S62) or conserved (Q24, E53, N55, and R58) among needle proteins, suggesting that they are important for function other than stabilizing the core domain. Other residues that contribute to surface polarity are the conserved Y47 and Y54. While the Y47 and Y54 phenyl rings are required in stabilizing the hydrophobic core, their hydroxyl groups are pointed towards the polar surface (Fig. 3B). Another polar moiety is the Q26-K50 hydrogen bond in the middle of the two helix bundle of PrgI (Fig. 3A), which is in a position equivalent to the MxiH E29-K53 salt bridge.[21] BsaL, however, does not have any electrostatic interactions equivalent to that of PrgI Q26-K50 or MxiH E29-K53.[24] Residues Q24, Y54, N55, and R58 are clustered together on the surface of PrgI and form a polar patch at the base of the two helix bundle (Fig. 3B). This polar patch is similar to the BsaL-BsaL contact site we identified by NMR chemical shift mapping of BsaL dimers (Wang *et al. in preparation*).

Electrostatic surfaces are different

The needle monomers are acidic proteins with theoretical pI's of 4.76 (PrgI and BsaL) and 4.47 (MxiH). Indeed, the electrostatic maps show large areas of negatively charged surfaces (Fig. 4). The location of the negatively charged surfaces, however, are radically different among PrgI, BsaL, and MxiH. In PrgI, the largest negatively charged surface is located on the "side" of the two helix bundle and spans almost the entire length of helix α_1, from the PxxP motif to the base of the two helix bundle (Fig. 4A). This is in sharp contrast to BsaL and MxiH where the negatively charged surfaces are on the "face" of the two helix bundle (Fig. 4B & C). In BsaL, the negatively charged surface is more broadly distributed on the helix α_1- α_2 interface (Fig. 4B). Likewise in MxiH, a negatively charged surface potential is on the "face" of the two helix bundle (Fig. 4C). In PrgI, residues mostly on helix α_1 (S6, D10, S13, T18, N19, D21, T25, E29, and D32) contribute to the largest negatively charged surface, whereas, in BsaL and MxiH, residues on helix α_1 and α_2 contribute to the negatively charged face. The tail of helix α_2 in PrgI and MxiH (Fig. 4A & C), which contains conserved aspartic acids (PrgI D70 and D72), also has a significant negatively charged surface.

Despite being highly acidic proteins, PrgI, BsaL, and MxiH contain contiguous regions of basic (or positive) electrostatic potentials (Fig. 4). These positive regions are located similarly among

the needle proteins, residing near the base of the two helix bundle and distal to the PxxP motifs (Fig. 4). In PrgI and BsaL, the positive regions at the base of the two helix bundle are due to residues on both helix α_1 and α_2 (Fig. 4A & B). In MxiH, however, the positively charged surface is solely due to residues on helix α_2, starting from the base of the two helix bundle and extending towards the C-terminal end of helix α_2 (Fig. 4C). In PrgI, the positively charged surface is located ~90° away from the negative "side" (Fig. 4A), whereas in BsaL and MxiH, the positively charged surfaces are 180° away, or on the flip side of the negatively charged surfaces (Fig. 4B & C).

Mutagenesis of Q26, K50, and R58

We used mutagenesis and assayed the ability of *Salmonella* to invade a cultured human epithelial cell line (Fig. S4). Wild type bacteria form functional needles, and can invade and give rise to hundreds of colonies, whereas a *prgI* null mutant strain is noninvasive because it lacks a functional PrgI protein (Figure 5). A plasmid that introduces wild type PrgI restores the invasiveness of the *prgI* null mutant. We mutagenized Q26 and K50 to probe the importance of the Q26-K50 hydrogen bond found in the middle of the two helix bundle, which is the only electrostatic contact between helix α_1 and α_2 (the rest are hydrophobic contacts). We hypothesize it contributes, albeit weakly, in stabilizing the two helix bundle, and therefore plays a role in needle assembly and virulence. Indeed, Q26A and K50D disrupted this hydrogen bond and made *S. typhimurium* noninvasive (Figure 5). Q26E has wild type invasiveness because glutamate (E) can form a salt bridge with K50. K50L cannot participate in polar interactions yet it has wild type invasiveness, suggesting it is capable of forming functional needles, and that the Q26-K50 hydrogen bond is not critical for needle assembly and virulence. Finally, changing the electrostatic surface with the R58L mutation resulted to a noninvasive phenotype (Figure 5).

DISCUSSION

TTSA needle proteins self-associate, but deletion of the last five residues was exploited to obtain monomeric $PrgI^{C\Delta5}$ [19], similar to the approach used in the structure determinations of $BsaL^{C\Delta5}$ [24] and $MxiH^{C\Delta5}$ [21]. PrgI has 52% and 63% sequence identity with BsaL and MxiH, respectively, and the central region is more conserved compared to the N-terminus (Fig. 1). $PrgI^{C\Delta5}$ (Fig. 2A) and $BsaL^{C\Delta5}$ [20] gave excellent NMR spectra compared to $MxiH^{C\Delta5}$ (data not shown), allowing NMR structure determination. The central region of $PrgI^{C\Delta5}$ (Fig. 2B), $BsaL^{C\Delta5}$ [20], and $MxiH^{C\Delta5}$ [21] form a core domain composed of two helices that are linked by a PxxP motif. Helix α_2 is longer than helix α_1, probably due to glycines and prolines at the N-termini which act as helix breakers. The first 3, 8, and 19 residues of PrgI, Bsal, and MixH, respectively, are disordered. In PrgI and BsaL, the N- and C-terminal regions flanking the two helix bundle have residual α-helical conformation and are more flexible. The first 19 residues are disordered in the $MxiH^{C\Delta5}$ crystal, however, a portion of this region might be in partial α-helical conformation when viewed by NMR as in BsaL and PrgI. The regions flanking the two helix bundle do not fold back into the core structure and give the needle proteins an elongated shape instead of a globular structure.

The conserved hydrophobic contacts at the helix α_1-α_2 interface (Fig. 3) stabilize the two helix bundle and are important in needle assembly and pathogenesis. More information is known about the effects of mutations on MxiH compared to PrgI and BsaL because Kenjale and coworkers[17] have characterized the phenotypes of over 35 MxiH point mutants (into alanine) with respect to needle assembly, secretion of effector proteins, and the ability of *S. flexneri* to invade cultured epithelial cells. In *S. flexneri*, MxiH L54A or Y57A mutation is defective in needle polymerization and secretion of effector proteins and thus is noninvasive.[17] The MxiH L54A or Y57A (equivalent to PrgI L51 and Y54) mutation is expected to decrease the

hydrophobic contacts at the helix α_1- α_2 interface and destabilize the two helix bundle. The proper folding of the two helix bundle is thus essential in needle assembly.

The polar surface of the two helix bundle is another common feature among needle proteins (Fig. 3 & 4). Conserved tyrosines contribute to surface polarity; however, the structures of PrgI, BsaL, and MxiH do not show a particular need for tyrosines in position Y47 and Y54 – phenylalanine would have maintained similar hydrophobic contacts and the overall fold of the two helix bundle. Yet, Y47 and Y54 are identical among PrgI, BsaL, and MxiH, suggesting the tyrosyl hydroxyl group is important for function. We hypothesize that the conserved tyrosine residues are important in protein-protein interactions with other needle monomers or with the tip complex.[25] Indeed, a MxiH Y50F point mutation, which is equivalent to PrgI Y47, abrogated the *in vivo* interaction between the MxiH needle and the tip complex and decreased the invasiveness of *S. flexneri* by 78% (*Zhang et al., in preparation*). Another mutation affecting a polar residue, MxiH Q51A (equivalent to PrgI Q48, Fig. 3B) showed normal needle assembly but decreased invasiveness.[17] Structurally, MxiH Q51 is pointed away from the two helix bundle and does not appear to contribute in stabilizing the core structure of the needle protein. However, MxiH Q51 is identical among the needle proteins (Fig. 1), suggesting that the decreased invasivenss of *S. flexneri* is probably due to defective protein-protein interaction. The polar surface of the needle protein is therefore important in pathogenesis.

Previous results showed that the last five residues are important in needle monomer-monomer interaction.[17, 19] In addition to the last five residues, our data suggest that electrostatic interaction is also an important component of needle assembly. This is based on the presence of distinct areas of positively and negatively charged surface potentials on PrgI, MxiH, and BsaL (Fig. 4). Furthermore, this work shows that the most significant difference among these needle proteins is in the arrangement of their electrostatic surfaces (Fig. 4), suggesting that despite primary sequence conservation, protein-protein interaction maybe different among these needle proteins. This could account for the observation that despite 63% sequence identity between PrgI and MxiH, PrgI does not complement a *S. flexneri mxiH* null mutant (*W.D. Picking, unpublished data*). The needle proteins of *S. typhimurium* and *S. flexneri* are therefore not interchangeable. This also suggests that factors affecting electrostatic interactions, such as ionic strength, metal cations, and pH will affect needle assembly. The effect of salt concentration and cations on needle assembly have not been explored in detail, however, Marlovits *et al*.[14] showed *in vitro* that increasing the pH from 8 to 10.5 dissociated the needle, but not the basal structure of the *S. typhimurium* needle apparatus. We identified a conserved tyrosine residue (equivalent to PrgI Y54) is involved in needle monomer-monomer contact by NMR chemical shift mapping of BsaL dimers (*Wang et al., in preparation*). A possible scenario is that high pH changes the protonation of this Y54 hydroxyl group, which then interferes with needle monomer-monomer contact that leads to the dissociation of the needle.

The *Salmonella* invasion assay relies on the proper assembly of functional needles, therefore, it provides an indirect readout of the PrgI-PrgI interaction involved in needle assembly. We mutagenized the polar residues (Q26, K50, R58) that contribute to the electrostatic surface of PrgI and used the *Salmonella* invasion assay to test their roles in needle assembly (Figure 5). Mutagenesis of Q26 and K50 suggest that the Q26-K50 hydrogen bond is not critical for invasiveness, but the polarity of this region, either through Q26 or K50, and the aliphatic moieties of Q26 and K50, are required in needle assembly. Mutagenesis also suggests that the hydrophobic packing interaction of the aliphatic groups of Q26 and K50 are more important in the structure of the two helix bundle than the hydrogen bonding interaction. Thus, Q26A and K50D, which have decreased aliphatic moieties, are noninvasive, whereas Q26E and K50L, which have aliphatic parts comparable to wild type, are invasive. The aliphatic part of R58, which is identical in PrgI, BsaL, and MxiH, participates in intramolecular hydrophobic contacts, however, its guanidine group is not involved in any polar interactions but contributes

prominently to the positively charged surface of the two helix bundle (Fig. 4). R58L mutation is noninvasive, suggesting that altering this surface polarity is deleterious in needle assembly.

Cordes et al.[26] showed using X-ray fiber diffraction and electron microscopy that the *S. flexneri* needle is a helical structure with 5.6 subunits/turn and a 24-Å helical pitch. Deane et al.[21] determined the crystal structure of the MxiH$^{C\Delta5}$ monomer, and used the needle helical parameters of Cordes et al.[26] to create an atomic model of the *S. flexneri* needle. To examine the electrostatic contacts between needle monomers, we modeled the *S. flexneri* needle (Fig. 6A-D) using the MxiH$^{C\Delta5}$ crystal structure and the parameters reported by Deane et al.[21] that described how the subunits are packed together. The model was built with 28 subunits of the MxiH$^{C\Delta5}$ crystal structure using the CHARMM biomolecular simulation program.[27] The first 19 N-terminal residues and the last five C-terminal residues were modeled to form regular alpha helices. The N-terminal helix was modeled to be inside the needle channel and parallel to the axis of the channel (the *Z*-axis). Each subunit was translated by 4.31 Å along the *Z*-axis and rotated by 64.3° around a 25 Å diameter of the needle channel. The model was energy minimized with a generalized Born implicit solvent model[28] in CHARMM to allow the side chains to adopt low-energy conformations and resolve the steric clash between the residues at the binding interfaces. To model the *S. typhimurium* needle, the NMR structure of PrgI was superimposed on the backbone atoms (N, C^{α}, C') of the MxiH needle and energy minimized.

Within the needle model, the two subunits with the largest surface contact were isolated and the electrostatic potentials of this dimer were calculated by solving the linearized Poisson-Boltzmann equation using the PBEQ module[29, 30] in CHARMM. Only the electrostatic potentials of residues involved in the packing interaction are shown in Fig. 6C,D. The electrostatic maps of MxiH in Fig. 4C and Fig. 6C are not identical because the former was calculated from an isolated MxiH structure and the later was that of MxiH in the assembled needle after energy minimization, which changed the side chain orientations to reduce the steric clash and optimize the MxiH-MxiH contacts in the needle. Thus, the relative side chain coordinates of MxiH in Fig. 4C and Fig. 6C are not identical, and their electrostatic surfaces will not be identical as well. Similarly, the electrostatic maps of the isolated PrgI (Fig. 4A) and that of PrgI in the assembled needle after energy minimization (Fig. 6D) will not be identical.

For MxiH, there is complementarity between the positively and the negative charged electrostatic surfaces of two MxiH molecules (Fig. 6D-F), whereas, PrgI showed mostly positive-to-positive electrostatic contacts (Fig. 6G-I). In addition, the calculated Coulombic electrostatic interaction energy of the MxiH dimer (Fig. 6E) was 5 times lower than of the PrgI dimer (Fig. 6H). Although the packing interactions should be considered in the context of the entire needle, our modeling suggests that the interaction between PrgI molecules would have to be different from that of MxiH to maximize the complementarity between the positive-to-negative electrostatic contacts. Therefore, we conclude that the needle packing interaction in *S. typhimurium* would have to be different from that of *S. flexneri*, even though their needle monomers share 63% primary sequence identity, because the electrostatic surfaces of their needle monomers are radically different from each other.

MATERIALS AND METHODS

Expression and purification of PrgI$^{C\Delta5}$

The subcloning of the PrgI$^{C\Delta5}$ expression plasmid had been reported earlier.[19] PrgI$^{C\Delta5}$ is similar to the BsaL$^{C\Delta5}$ and MxiH$^{C\Delta5}$ constructs used in previous structure determinations where the last five residues were deleted and replaced with a C-terminal His_6-tag.[20, 21] Uniformly ^{13}C,^{15}N- and ^{15}N-labeled PrgI$^{C\Delta5}$ were obtained by growing *E. coli* BL21(DE3) with the plasmid pET22b-*prgI*$^{C\Delta5}$ in 1 liter minimal media supplemented with 2 g/L ^{13}C-glucose and 1 g/L ^{15}N ammonium chloride. Cells were grown at 37°C, induced with 1 mM

IPTG at OD_{600} 0.8, and expressed overnight at 15°C to a final OD_{600} of 2.4. Cells were centrifuged, resuspended in 40 mL buffer A (20 mM Tris pH 7.9, 500 mM NaCl) with 5 mM imidazole, and sonicated. Cellular debris was removed by centrifugation and the supernatant was loaded on a 5 mL Ni^{2+}-affinity column (Sigma) that was previously treated with 35 mL H_2O, 35 mL 50 mM $NiSO_4$, and 35 mL buffer A with 5 mM imidazole. The Ni^{2+}-column was washed with 35 mL buffer A with 60 mM imidazole and eluted with 15 mL buffer A with 1 M imidazole in 1 mL fractions. Fractions were pooled and dialyzed in NMR buffer (10 mM sodium phosphate pH 6.0, 10 mM NaCl) and protein concentration was estimated by UV absorbance at 280 nm.

NMR spectroscopy

NMR data were acquired at 25°C using a 0.8 mM protein sample on a Bruker Avance 800 MHz spectrometer, processed with NMRPipe[31], and analyzed using NMRView [32]. Resonance assignments were obtained from 2D ^{1}H-^{15}N HSQC [33] and ^{1}H-^{13}C-HMQC [34] and the following 3D datasets: HNCA [35], HNCO,[36] CBCA(CO)NH,[35] HNCACB,[37] and HBHACONH.[38] Water suppression was achieved using flip-back pulses and pulsed field gradients. The α-helices were identified from the C^{α}, C^{β}, and H^{α} secondary chemical shifts.[39] The ^{1}H, ^{13}C, ^{15}N chemical shifts were referenced indirectly using 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). Inter-proton distance restraints were obtained from 3D ^{15}N-edited NOESY-HSQC (t_{mix} = 120 ms) and ^{13}C-edited HMQC-NOESY (t_{mix} = 120 ms).

Structure calculations and refinement

Inter-proton distance restraints from the NOESY datasets were classified into upper bounds of 2.7, 3.5, 4.5, and 5.5 Å based on peak volumes, and lower bounds of 1.8 Å. The NOE patterns[40] were traced to confirm the secondary structures predicted by chemical shifts. The Phi and Psi dihedral angles of residues in α-helical regions were constrained to $\varphi = -60 \pm 20°$ and $\psi = -40 \pm 20°$.[39] The distance and dihedral angle restraints were used to calculate 200 structures using torsion angle dynamics in CYANA.[22] Several rounds of CYANA calculations followed by violation and NOE analysis were done until the global fold of the protein converged and the CYANA target function dropped below 10. The CYANA structures were further refined by restrained molecular dynamics and simulated annealing using AMBER.[23] Several rounds of *in vacuo* AMBER calculations were done until the distance violations were less than 0.5Å, then the generalized Born (GB) potential was used to account for the effect of water during molecular dynamics. The CYANA and AMBER structure calculations included cycles of iterative refinement of restraints. The surface electrostatic potential maps of PrgI, BsaL (PDB #2GOU), and MxiH (PDB #2CA5) monomers (Fig. 4) were calculated using the program APBS.[41] Structures were analyzed using PROCHECK [42] and Pymol (http://pymol.sourceforge.net).

Salmonella invasion assay

We followed published protocol to assay the ability of *S. typhimurium* to invade a cultured human epithelial cell line (Henle 407).[43] Briefly, Henle 407 was obtained from the American Type Culture Collection. *S. typhimurium* wild type (SL1344) and *prgI* null mutant (JK17)[44] strains were obtained from Dr. Bradley Jones (University of Iowa). Henle 407 cells were grown in Eagle's modified minimum essential medium with 10% newborn calf serum and 5% CO_2. Mutations were introduced in the pRK2-*prgI* plasmid[17] by PCR and transformed into *S. typhimurium* JK17 by electroporation. Bacteria were grown in media containing the antibiotics kanamycin, ampicillin, and trimethoprim. Bacterial suspension were incubated with Henle 407 cells for 60 min to allow invasion before removal by aspiration. The Henle 407 cells were washed with media containing 50 ug/ml gentamycin and lyzed with 0.9% deoxycholate to free

the entrapped bacteria. The number of bacterial colonies, which correlates with invasiveness, were estimated by serial dilution and plating.

COORDINATES AND RESONANCE ASSIGNMENTS

Structural coordinates and list of restraints were deposited at the Protein Data Bank (2JOW). Resonance assignments were deposited at the Biological Magnetic Resonance Data Bank (BMRB 15206).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We are grateful for the PrgI$^{C\Delta5}$ expression plasmid and for helpful discussion from Roma Kenjale, Lingling Zhang, William D. Picking and Wendy L. Picking and for the NMR assistance of David van der Velde. R.N.D. is supported by COBRE Protein Structure and Function Program (NCRR RR017708), MRCE for Biodefense and Emerging Infectious Diseases (NIAID 5U54 AI057160), and startup funds from the University of Kansas.

References

1. Vugia DJ, Samuel M, Farley MM, Marcus R, Shiferaw B, Shallow S, Smith K, Angulo FJ. Invasive *Salmonella* infections in the United States, FoodNet, 1996-1999: incidence, serotype distribution, and outcome. Clin Infect Dis 2004;38(Suppl 3):S149–S156. [PubMed: 15095184]
2. Blocker A, Gounon P, Larquet E, Niebuhr K, Cabiaux V, Parsot C, Sansonetti P. The tripartite type III secreton of *Shigella flexneri* inserts IpaB and IpaC into host membranes. J Cell Biol 1999;147:683–693. [PubMed: 10545510]
3. Cornelis GR. *Yersinia* type III secretion: send in the effectors. J Cell Biol 2002;158:401–408. [PubMed: 12163464]
4. Stevens MP, Wood MW, Taylor LA, Monaghan P, Hawes P, Jones PW, Wallis TS, Galyov EE. An Inv/Mxi-Spa-like type III protein secretion system in *Burkholderia pseudomallei* modulates intracellular behaviour of the pathogen. Mol Microbiol 2002;46:649–659. [PubMed: 12410823]
5. Yahr TL, Goranson J, Frank DW. Exoenzyme S of *Pseudomonas aeruginosa* is secreted by a type III pathway. Mol Microbiol 1996;22:991–1003. [PubMed: 8971719]
6. Weinberger M, Keller N. Recent trends in the epidemiology of non-typhoid *Salmonella* and antimicrobial resistance: the Israeli experience and worldwide review. Curr Opin Infect Dis 2005;18:513–521. [PubMed: 16258325]
7. Sivapalasingam S, Nelson JM, Joyce K, Hoekstra M, Angulo FJ, Mintz ED. High prevalence of antimicrobial resistance among *Shigella* isolates in the United States tested by the National Antimicrobial Resistance Monitoring System from 1999 to 2002. Antimicrob Agents Chemother 2006;50:49–54. [PubMed: 16377666]
8. Rotz LD, Khan AS, Lillibridge SR, Ostroff SM, Hughes JM. Public health assessment of potential biological terrorism agents. Emerg Infect Dis 2002;8:225–230. [PubMed: 11897082]
9. Hueck CJ. Type III protein secretion systems in bacterial pathogens of animals and plants. Microbiol Mol Biol Rev 1998;62:379–433. [PubMed: 9618447]
10. Kubori T, Matsushima Y, Nakamura D, Uralil J, Lara-Tejero M, Sukhan A, Galan JE, Aizawa SI. Supramolecular structure of the *Salmonella typhimurium* type III protein secretion system. Science 1998;280:602–605. [PubMed: 9554854]
11. Cordes FS, Daniell S, Kenjale R, Saurya S, Picking WL, Picking WD, Booy F, Lea SM, Blocker A. Helical packing of needles from functionally altered *Shigella* type III secretion systems. J Mol Biol 2005;354:206–211. [PubMed: 16243352]
12. Journet L, Agrain C, Broz P, Cornelis GR. The needle length of bacterial injectisomes is determined by a molecular ruler. Science 2003;302:1757–1760. [PubMed: 14657497]

13. Pastor A, Chabert J, Louwagie M, Garin J, Attree I. PscF is a major component of the *Pseudomonas aeruginosa* type III secretion needle. FEMS Microbiol Lett 2005;253:95–101. [PubMed: 16239085]

14. Marlovits TC, Kubori T, Sukhan A, Thomas DR, Galan JE, Unger VM. Structural insights into the assembly of the type III secretion needle complex. Science 2004;306:1040–1042. [PubMed: 15528446]

15. Blocker A, Jouihri N, Larquet E, Gounon P, Ebel F, Parsot C, Sansonetti P, Allaoui A. Structure and composition of the *Shigella flexneri* "needle complex", a part of its type III secreton. Mol Microbiol 2001;39:652–663. [PubMed: 11169106]

16. Kubori T, Sukhan A, Aizawa SI, Galan JE. Molecular characterization and assembly of the needle complex of the *Salmonella typhimurium* type III protein secretion system. Proc Natl Acad Sci U S A 2000;97:10225–10230. [PubMed: 10944190]

17. Kenjale R, Wilson J, Zenk SF, Saurya S, Picking WL, Picking WD, Blocker A. The needle component of the type III secreton of *Shigella* regulates the activity of the secretion apparatus. J Biol Chem 2005;280:42929–42937. [PubMed: 16227202]

18. Torruellas J, Jackson MW, Pennock JW, Plano GV. The *Yersinia pestis* type III secretion needle plays a role in the regulation of Yop secretion. Mol Microbiol 2005;57:1719–1733. [PubMed: 16135236]

19. Darboe N, Kenjale R, Picking WL, Picking WD, Middaugh CR. Physical characterization of MxiH and PrgI, the needle component of the type III secretion apparatus from *Shigella* and *Salmonella*. Protein Science 2006;15:543–552. [PubMed: 16501225]

20. Zhang L, Wang Y, Picking WL, Picking WD, De Guzman RN. Solution Structure of Monomeric BsaL, the Type III Secretion Needle Protein of *Burkholderia pseudomallei*. J Mol Biol 2006;359:322–330. [PubMed: 16631790]

21. Deane JE, Roversi P, Cordes FS, Johnson S, Kenjale R, Daniell S, Booy F, Picking WD, Picking WL, Blocker AJ, Lea SM. Molecular model of a type III secretion system needle: Implications for host-cell sensing. Proc Natl Acad Sci U S A 2006;103:12529–12533. [PubMed: 16888041]

22. Guntert P. Automated NMR structure calculation with CYANA. Methods Mol Biol 2004;278:353–378. [PubMed: 15318003]

23. Case, DA.; Pearlman, DA.; Caldwell, JW.; Cheatham, TE., III; Wang, J.; Ross, WS.; Simmerling, CL.; Darden, TA.; Merz, KM.; Stanton, RV.; Cheng, AL.; Vincent, JJ.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, RJ.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, GL.; Singh, UC.; Weiner, PK.; Kollman, PA. AMBER7. University of California; San Francisco: 2002.

24. Zhang Y, Arakaki AK, Skolnick J. TASSER: an automated method for the prediction of protein tertiary structures in CASP6. Proteins 2005;61(Suppl 7):91–98. [PubMed: 16187349]

25. Espina M, Olive AJ, Kenjale R, Moore DS, Ausar SF, Kaminski RW, Oaks EV, Middaugh CR, Picking WD, Picking WL. IpaD Localizes to the Tip of the Type III Secretion System Needle of *Shigella flexneri*. Infect Immun 2006;74:4391–4400. [PubMed: 16861624]

26. Cordes FS, Komoriya K, Larquet E, Yang S, Egelman EH, Blocker A, Lea SM. Helical structure of the needle of the type III secretion system of *Shigella flexneri*. J Biol Chem 2003;278:17103–17107. [PubMed: 12571230]

27. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comp Chem 1983;4:187–217.

28. Im W, Lee MS, Brooks CL III. Generalized born model with a simple smoothing function. J Comput Chem 2003;24:1691–1702. [PubMed: 12964188]

29. Im W, Beglov D, Roux B. Continuum solvation model: Electrostatic forces from numerical solutions to the Poisson-Bolztmann equation. Comp Phys Comm 1998;111:59–75.

30. Nina M, Beglov D, Roux B. Atomic Radii for Continuum Electrostatics Calculations Based on Molecular Dynamics Free Energy Simulations. J Phys Chem B 1997;101:5239–5248.

31. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–293. [PubMed: 8520220]

32. Johnson BA. Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol 2004;278:313–352. [PubMed: 15318002]

33. Grzesiek S, Bax A. The importance of not saturating H2O in protein NMR. Application to sensitivity enhancement and NOE measurements. J Am Chem Soc 1993;115:12593–12594.

34. Tolman JR, Chung J, Prestegard JH. Pure-phase heteronuclear multiple-quantum spectroscopy using field gradient selection. J Magn Reson 1992;98:462–467.

35. Grzesiek S, Dobeli H, Gentz R, Garotta G, Labhardt AM, Bax A. ^{1}H, ^{13}C, and ^{15}N NMR backbone assignments and secondary structure of human interferon-gamma. Biochemistry 1992;31:8180–8190. [PubMed: 1525157]

36. Muhandiram DR, Kay LE. Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. J Magn Reson Ser B 1994;103:203–216.

37. Wittekind M, Mueller L. HNCACB, a high sensitivity 3D NMR experiment to correlate amide proton and nitrogen resonances with the α-carbon and β-carbon resonances in proteins. J Magn Reson 1993;101B:201–205.

38. Grzesiek S, Bax A. Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins. J Biomol NMR 1993;3:185–204. [PubMed: 8477186]

39. Wishart DS, Nip AM. Protein chemical shift analysis: a practical guide. Biochem Cell Biol 1998;76:153–163. [PubMed: 9923684]

40. Wuthrich, K. NMR of Proteins and Nucleic Acids. Wiley-Interscience; 1986.

41. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A 2001;98:10037–10041. [PubMed: 11517324]

42. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 1996;8:477–486. [PubMed: 9008363]

43. Osiecki JC, Barker J, Picking WL, Serfis AB, Berring E, Shah S, Harrington A, Picking WD. IpaC from *Shigella* and SipC from *Salmonella* possess similar biochemical properties but are functionally distinct. Mol Microbiol 2001;42:469–481. [PubMed: 11703668]

44. Klein JR, Fahlen TF, Jones BD. Transcriptional organization and function of invasion genes within Salmonella enterica serovar Typhimurium pathogenicity island 1, including the prgH, prgI, prgJ, prgK, orgA, orgB, and orgC genes. Infect Immun 2000;68:3368–3376. [PubMed: 10816487]

ABBREVIATIONS USED

IPTG
isopropyl-beta-D-thiogalactopyranoside

NMR
nuclear magnetic resonance

NOE
nuclear Overhauser effect

NOESY
NOE spectroscopy

HSQC
heteronuclear single quantum coherence

HMQC
heteronuclear multiple quantum coherence

RMSD
root mean square deviation

TTSA
type III secretion apparatus

TTSS
type III secretion system

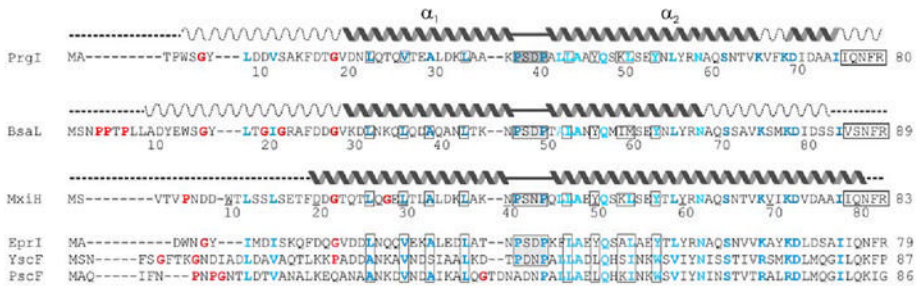

Figure 1.

The middle regions of the needle proteins are more conserved compared to their N-termini. The structural features of PrgI, BsaL, and MxiH are indicated as well-structured helices (α_1 and α_2, solid wavy lines), regions with partial α-helical character (dashed wavy lines), and random coil regions (dashed lines). Conserved hydrophobic residues (blue) at the helix α_1-α_2 interface are boxed and prolines and glycines (red) that act as helix breakers are indicated. The PxxP motif and the last five residues of PrgI, BsaL, and BsaL that were deleted are boxed. The needle proteins are from *Salmonella typhimurium* (PrgI), *Burkholderia pseudomallei* (BsaL), *Shigella flexneri* (MxiH), *Escherichia coli O157:H7* (EprI), *Yersinia pestis* (YscF), and *Pseudomonas aeruginosa* (PscF).

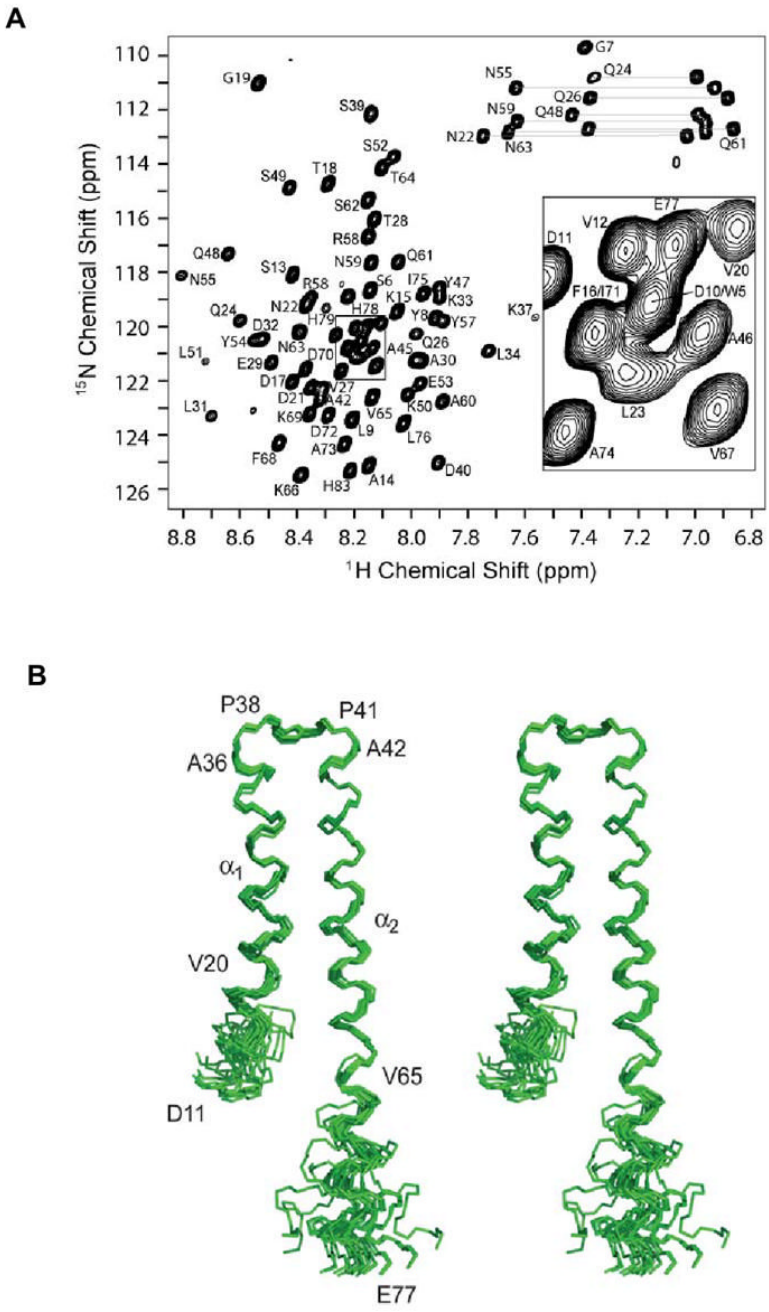

Figure 2.

(A) Assigned ^{1}H-^{15}N HSQC spectra of PrgI$^{C\Delta5}$. The insert shows an expansion of the crowded region of the spectrum. The side chain peak of W5 (not shown) is at 10.16 ^{1}H ppm and 129.89 ^{15}N ppm. **(B)** Stereoview of the superposition of 20 lowest energy structures of PrgI$^{C\Delta5}$, showing only the backbone N, C$^{\alpha}$, and C' atoms. Only the region V20-V65 can be superimposed into an ensemble of NMR structures, forming the core domain of PrgI$^{C\Delta5}$, however, the regions flanking this core domain, from W5-G19 and K66-I75, retained α-helical conformation in each of the 20 NMR structures.

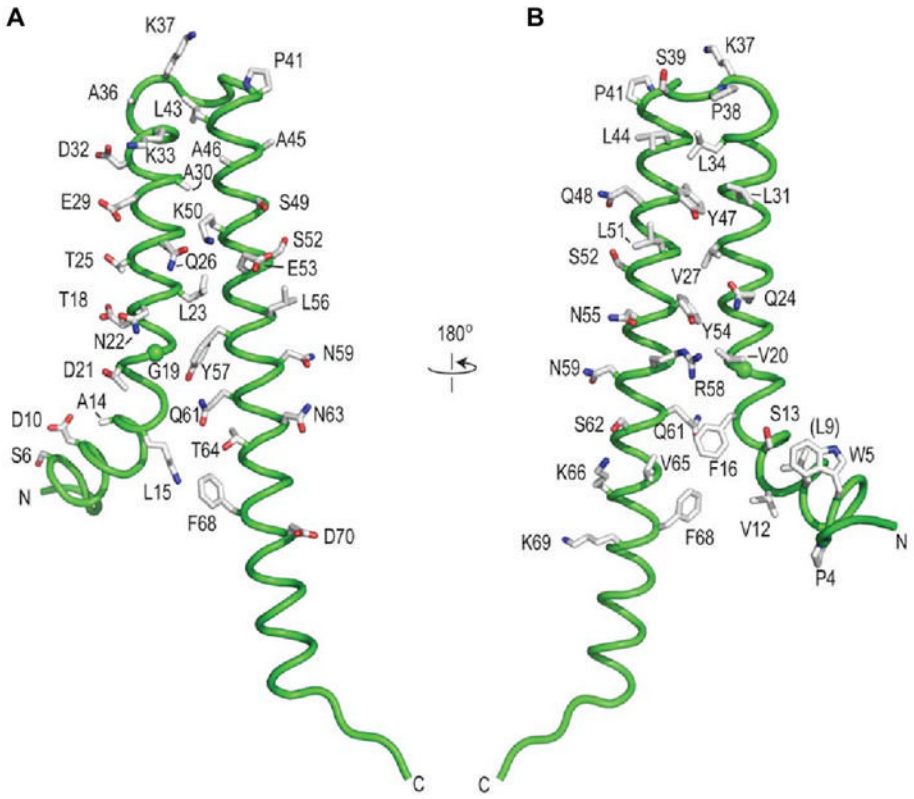

Figure 3.

$PrgI^{C\Delta 5}$ forms a two helix bundle stabilized by hydrophobic contacts at the helix α_1- α_2 interface. The side chains of the hydrophobic and polar residues are shown. (A) is oriented by a 180 rotation from (B). Atoms are colored as: red, oxygen; blue, nitrogen; gray, carbon.

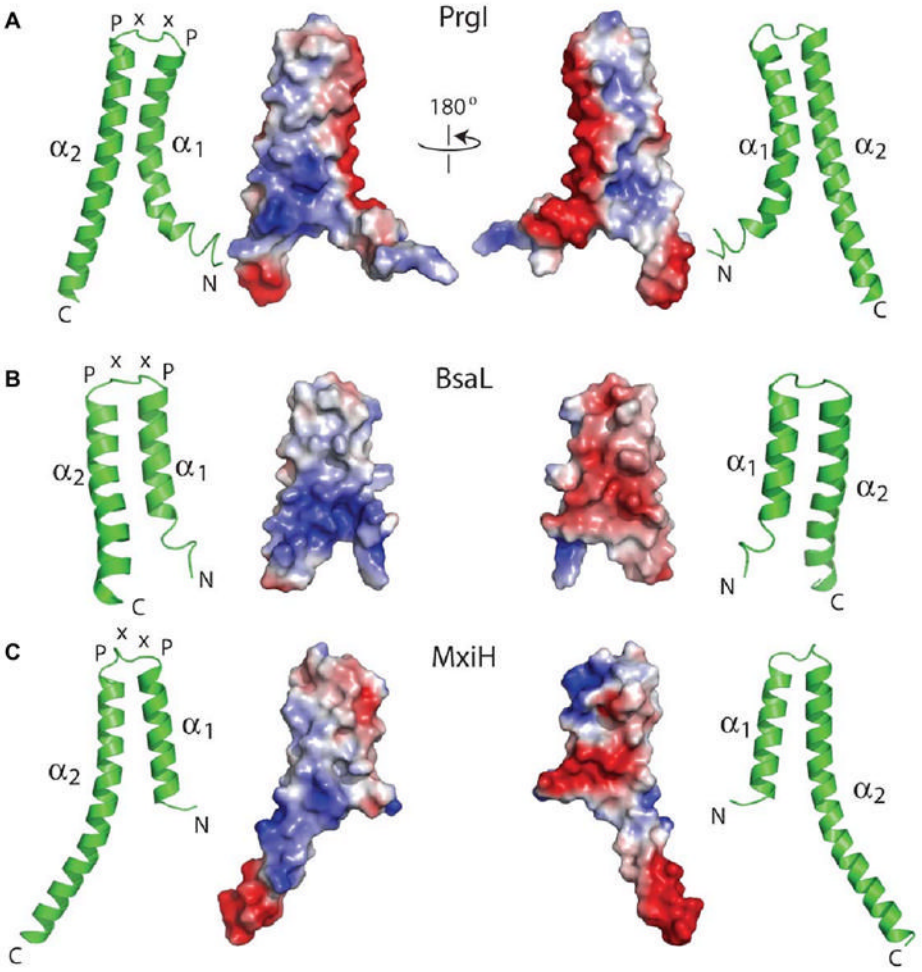

Figure 4.

The electrostatic surface of **(A)** PrgI, **(B)** BsaL, and **(C)** MxiH are shown together with the ribbon structures showing the relative orientation of the two helix bundle. The structures of PrgI, BsaL, and MxiH are oriented is similar manner from top to bottom to allow comparison of the surfaces. The structures on the left are rotated by 180° on the y-axis from the right.

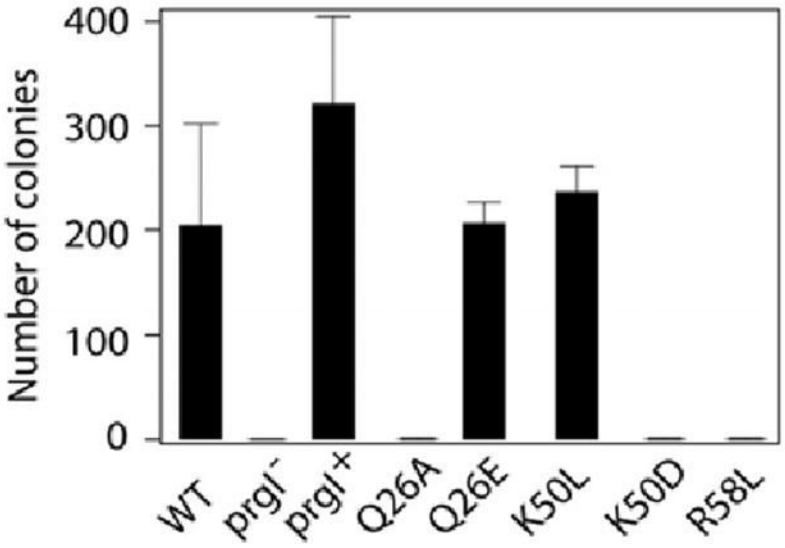

Figure 5.

Salmonella invasion assay. The number of bacterial colonies (y axis) correlates with the ability of *S. typhimurium* to invade a cultured human epithelial cell line (Henle 407). Wild type (WT) bacterium is invasive, *whereas a S. typhimurium* with a *prgI* null mutation (*prgI*-) is noninvasive. A plasmid that overexpressed a functional PrgI (prgI^{+}) restored the invasiveness of the *prgI* null mutant strain. Point mutations were introduced in the Q26, K50, and R58 residues using the PrgI plasmid. The assays were done in triplicates.

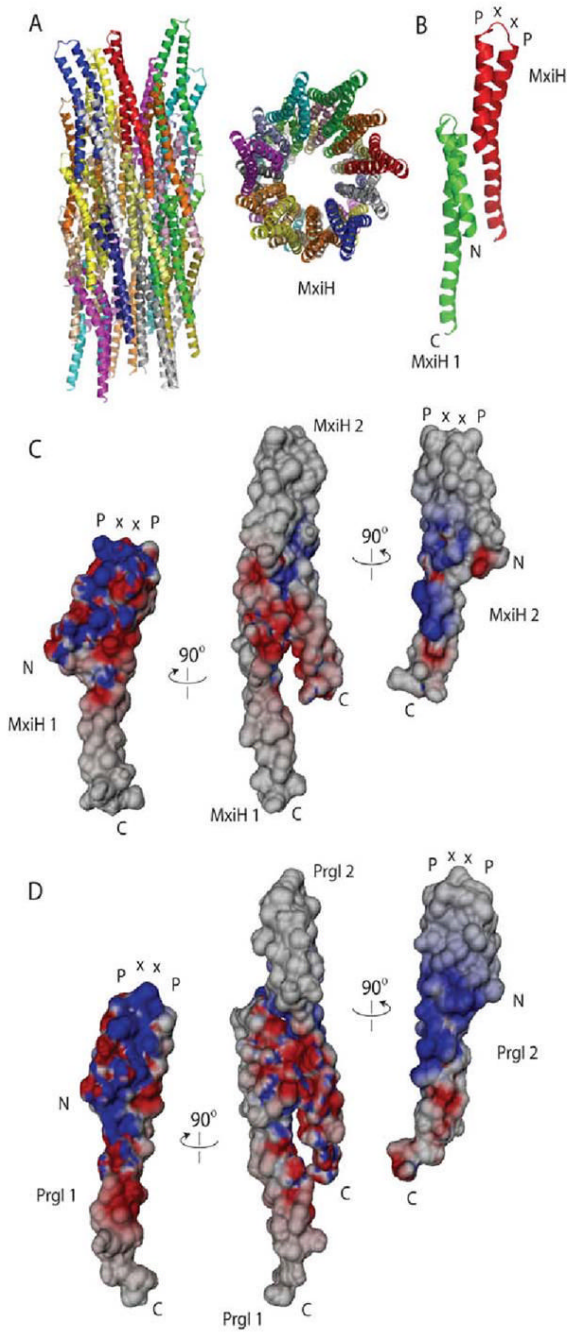

Figure 6.

(A) Model of the *S. flexneri* needle created using the crystal structure of MxiH$^{C\Delta5}$ (PDB #2CA5) and published needle packing parameters,[21,26] and visualized with the top towards the host and below towards the bacterium, and along the 25 Å needle channel looking down towards the bacterium. **(B)** Two MxiH molecules within the needle with extensive intermolecular contacts are chosen for electrostatic analysis. **(C)** The two monomers, MxiH 1 and MxiH 2, are separated from the MxiH dimer and rotated by 90° towards the viewer to reveal the electrostatic potentials of the residues involved in the needle packing interaction. The N-terminal 19 residues of MxiH were removed because they were disordered in the MxiH crystal structure. **(D)** Electrostatic analysis of two PrgI molecules (PrgI 1 and PrgI 2) in the

model of the *S. typhimurium* needle, which was created by superimposing the NMR structure of PrgI into *S. flexneri* needle. The N-terminal 19 residues of PrgI were removed in the calculation of electrostatic potentials because they did not converge into a single family of structures. The Coulombic electrostatic interaction energy is -120 kcal/mol for the MxiH dimer and -23 kcal/mol for the PrgI dimer. Molecular graphics were made using Pymol **(A and B)** and Dino (DINO: Visualizing Structural Biology (2002) http://www.dino3d.org) **(C and D)**. The amino (N) and the carboxy (C) termini and the PxxP loop are indicated.

Table 1

Structural statistics for 20 NMR structures of $PrgI^{C\Delta5}$.

Total Unambiguous Distance Restraints	1270
Intraresidue (i,i)	283
Sequential (i, i+1)	344
Medium Range (2<=|i-j|<=4)	469
Long Range (|i-j| > 4)	174
Total Dihedral Angle Restraints	101
Phi	68
Psi	33
RMS difference from mean structure (V20-V65)	
Backbone atoms of helices	0.26
All heavy atoms atoms of helices	0.60
Violation analysis	
Max distance violation (A)	0.3
Max dihedral angle violation (deg)	4.4
Energies	
mean AMBER energy (kcal mol^{-1})	-2464
mean restraint energy (kcal mol^{-1})	114
Deviation from idealized geometry	
Bond lengths (A)	0.00983 ± 0.00005
Bond angles (degree)	2.79 ± 0.01
Ramachandran plot	
Most favorable region	85.1%
Additionally allowed regions	14.9%
Generously allowed regions	0%
Disallowed regions	0%