THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 30, Issue of July 23, pp. 31599–31605, 2004

Printed in U.S.A.

NMR Structure of a Type IVb Pilin from *Salmonella typhi* and Its Assembly into Pilus*

Received for publication, April 28, 2004, and in revised form, May 24, 2004
Published, JBC Papers in Press, May 24, 2004, DOI 10.1074/jbc.M404727200

Xing-Fu Xu‡, Yih-Wan Tan‡, Lam Lam§, Jim Hackett§, Mingjie Zhang§, and Yu-Keung Mok‡¶

From the ‡Department of Biological Sciences, National University of Singapore, Singapore 117543 and the §Department of Biochemistry, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

The structure of the N-terminal-truncated Type IVb structural pilin (t-PilS) from *Salmonella typhi* was determined by NMR. Although topologically similar to the recently determined x-ray structure of pilin from *Vibrio cholerae* toxin-coregulated pilus, the only Type IVb pilin with known structure, t-PilS contains many distinct structural features. The protein contains an extra pair of β-strands in the N-terminal αβ loop that align with the major β-strands to form a continuous 7-stranded antiparallel β-sheet. The C-terminal disulfide-bonded region of t-PilS is only half the length of that of toxin-coregulated pilus pilin. A model of *S. typhi* pilus has been proposed and mutagenesis studies suggested that residues on both the αβ loop and the C-terminal disulfide-bonded region of PilS might be involved in binding specificity of the pilus. This model structure reveals an exposed surface between adjacent subunits of PilS that could be a potential binding site for the cystic fibrosis transmembrane conductance regulator.

Type IV pili have long been recognized as major bacterial virulence-associated adhesins that promote bacterial attachment to host cells (1). The genome of the typhoid fever bacterium, *Salmonella typhi*, contains a large insertion or "pathogenicity island" that is not found in the chromosome of the noninvasive *Salmonella typhimurium* (2). DNA sequencing within the novel pathogenicity island (*Salmonella* pathogenicity island 7) identified a pil operon (11 genes, from PilL to PilV) that contains genes required for biosynthesis of Type IV pili in *S. typhi* (3). *S. typhi* uses the cystic fibrosis transmembrane conductance regulator (CFTR)[1] as an epithelial cell receptor for the Type IV pili prior to entry of the gastrointestinal submucosa following oral intake (4). The structural pilin protein (PilS) of the Type IV pili interacts with the first extracellular domain (residues 103–117) of CFTR (3–5). Type IV pilins are further classified into Type IVa and IVb pilins according to the length of their signal peptides and the identities of the first residues of the mature pilins. Type IVa pilin has a much shorter signal peptide than Type Ivb, and the N-terminal residue of the mature pilin is always an *N*-methylated phenylalanine, whereas the N-terminal residue of Type IVb pilin may be either methionine or leucine (6). Type IVa pili are found in *Pseudomonas aeruginosa*, *Neisseria gonorrhoeaem,* and *Mycobacterium bovis*. The pilus of *S. typhi*, the Tcp of *V. cholerae* (7), the bundle-forming pilus of enteropathogenic *Escherichia coli* (8), the longus pilus of enterotoxigenic *E. coli* (9), and the R64 thin pilus (10) all belong to Type IVb. Both mature Type IVa and IVb pilins contain an N-terminal hydrophobic region of ~24 residues (Fig. 1) that is believed to be the oligomerization domain required for Type IV pilus assembly through the "general assembly pathway" (11).

The structures of Type IVa pilin from strain MS11 of *N. gonorrhoeae* (12) and strain PAK of *P. aeruginosa* (13) have been solved by x-ray crystallography. The structure of the Type IVa pilin of strain K122–4 of *P. aeruginosa* (14) has been determined by NMR. The ladle-shaped Type IVa pilin adopts an αβ roll fold with a very long and bent N-terminal hydrophobic α-helix. Assembly of the helical pilus is proposed to form a core of coiled α1 helices banded by β-sheets, leaving hypervariable disulfide-bonded loop regions exposed for receptor binding, although the tip-located adhesin is also involved in receptor binding (15). This is in contrast to Type I pili, P pili, and Cs1 pili, which use exclusively their tip adhesins, FimH, PapG, and CooD respectively, for receptor binding (16–18). Very recently, the structure of a Type IVb pilin (Tcp of *V. cholerae*) was solved by x-ray crystallography (19). The structure of the Tcp pilin revealed a novel fold of pilin and suggested a different manner of pilus assembly compared with that of Type IVa pilins. Here, we present the NMR solution structure of the N-terminal-truncated Type IVb PilS from *S. typhi*. The *S. typhi* PilS shares a common topology with the *V. cholerae* Tcp pilin but with many distinct structural differences in both the αβ loop and D region. A structural model of *S. typhi* pilus is proposed based on the model of *V. cholerae* Tcp. The structural differences between these two Type IVb pilins allow a very different surface to be exposed on *S. typhi* pilus, which may represent a potential binding site for CFTR.

EXPERIMENTAL PROCEDURES

Expression and Purification of PilS—The DNA fragment for *S. typhi* t-PilS (residues 26–181) was subcloned between the BamHI and EcoRI sites of a modified pET-32a (Novagen) vector with S-tag and thioredoxin gene removed. The correct sequence of the construct was confirmed by Big-dye sequencing. The His-tagged protein was expressed in

* This work was supported by the A*Star BMRC Young Investigator Award (to Y.-K. M.) and by the Academic Research Fund, National University of Singapore. Some of the NMR spectra used in this work were acquired at the 750-MHz NMR spectrometer in Hong Kong University of Science and Technology. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1Q5F) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

The NMR chemical shifts (accession code 5879) have been deposited in the BioMagResBank (BMRB), Department of Biochemistry, University of Wisconsin, Madison (www.bmrb.wisc.edu.).

¶ To whom correspondence should be addressed: Dept. of Biological Sciences, 14 Science Dr. 4, National University of Singapore, Singapore 117543. Tel.: 65-68742967; Fax: 65-67792486; E-mail: dbsmokh@nus.edu.sg.

[1] The abbreviations and trivial terms used are: CFTR, cystic fibrosis transmembrane conductance regulator; t-PilS, N-terminal-truncated structural pilin; D region, C-terminal disulfide-bonded region; NOE, nuclear Overhauser effect; Tcp, toxin-coregulated pilus; NOESY, nuclear Overhauser effect spectroscopy.

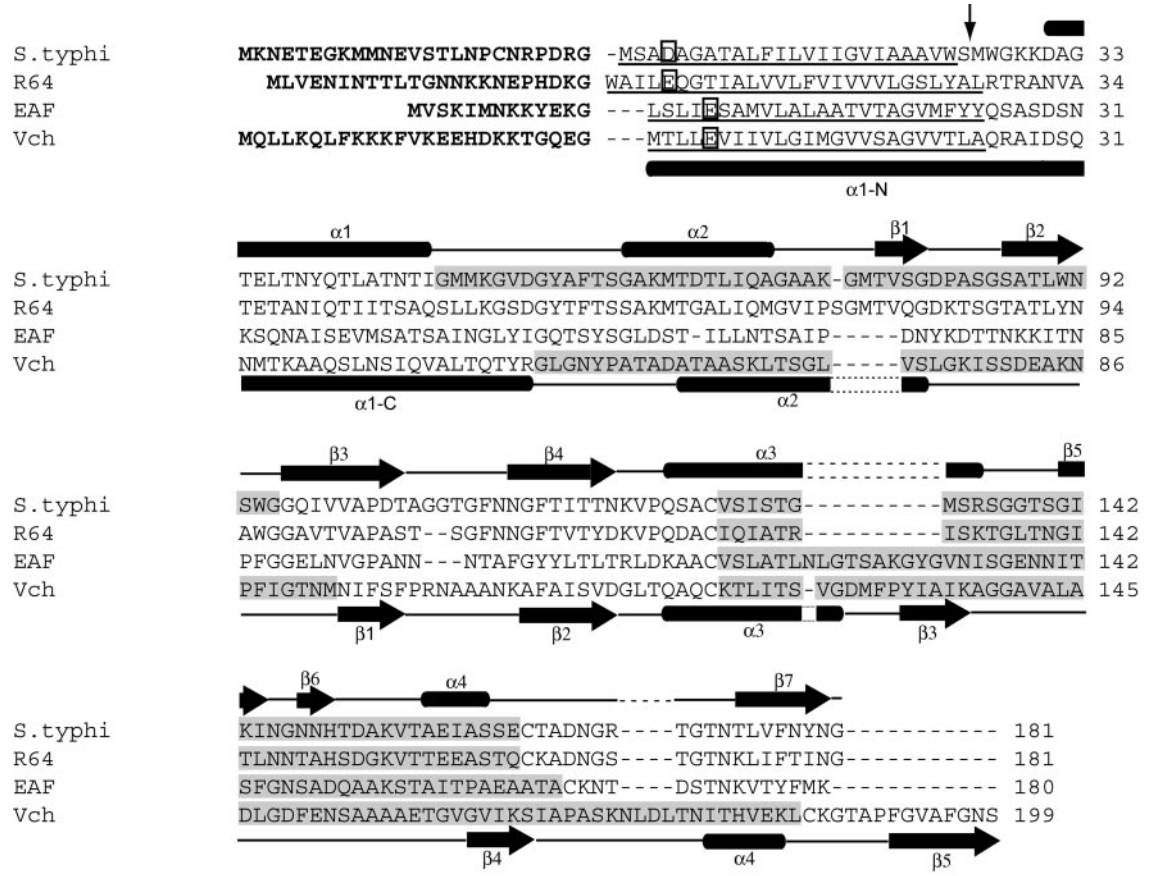

FIG. 1. **Sequence alignment and secondary structures of Type IVb pilins.** Sequence alignment of Type IVb pilins from *S. typhi* pilus (*S. typhi*), R64 thin pilus (*R64*), bundle-forming pilus of enteropathogenic *E. coli* (*EAF*), and toxin-coregulated pilus of *V. cholerae* (*Vch*) using the program CLUSTALW (36). *Arrows* and *cylinders above* the alignment indicate β-strands and α-helices of *S. typhi* t-PilS, respectively. *Arrows* and *cylinders below* the alignment indicate secondary structures of Tcp pilin as reported. The signal sequences are *bold-faced*. The N-terminal hydrophobic domains are *underlined*. The acidic residues found within the N-terminal hydrophobic domains are *boxed*. The *down arrow* indicates the starting boundary of the PilS construct used for NMR structure determination. The N-terminal αβ loop and the D region are *shaded*.

BL21(DE3) *E. coli* cells as inclusion bodies (induction with 0.3 mM final concentration of isopropyl-1-thio-β-D-galactopyranoside; cells were grown at 37 °C). Nine extra residues from the vector (GSAMADIGS) remain N-terminal to residue Met-26 of PilS after thrombin cleavage. The pellet of the expressed inclusion bodies was purified with nickel-nitrilotriacetic acid resin (Qiagen) using buffers containing 6 M GdnCl. The purified protein was refolded by rapid dilution (30×) into ice-cold buffer (50 mM Tris, pH 7.9) with stirring. Residual GdnCl was removed by dialysis, and proper folding of the soluble protein was checked by far-UV CD. The refolded protein was concentrated and cleaved with thrombin (3 units/mg protein) before further purification on a Sephacryl S-100 (Amersham Biosciences) gel filtration column (50 mM Tris, pH 7.9 + 0.5 M NaCl) for NMR studies.

The DNA fragments for PilS (residues 1–181) and its mutants were subcloned between BamHI and EcoRI site of the same vector. All mutants were confirmed by DNA sequencing. The proteins were expressed as inclusion bodies in BL21(DE3) by inducing at a cell density of A_{600} = 0.6 and at 30 °C. The inclusion body pellets were resuspended in binding buffer (20 mM Tris, pH 7.9, 0.5 M NaCl, 5 mM Imidazole) with 2% v/v Triton-X100 and 0.1 mM phenylmethylsulfonyl fluoride. The proteins were bound onto nickel-nitrilotriacetic acid resins, washed, and eluted with buffers that did not contain Triton X-100. The eluted proteins were further purified on a MonoQ ion-exchange column (Amersham Biosciences) before being used for binding assays.

NMR Spectroscopy and Structure Calculation—Purified t-PilS protein was concentrated to ~1 mM and buffer-exchanged to 50 mM phosphate, pH 6.0, with 0.5 M sodium sulfate for NMR studies. NMR experiments were performed on a 500-MHz Bruker DRX spectrometer equipped with cryoprobe or a 750-MHz Varian Inova spectrometer at 30 °C. Sequence-specific backbone assignments were done by HNCACB (20), CBCA(CO)NH (21), HNCO, and HN(CA)CO experiments using a ^{15}N-^{13}C uniformly labeled sample. Side chain assignments were obtained from H(CCO)NH, (H)CC(CO)NH (22), and HCCH-TOCSY (23) spectra. The stereo-specific assignment of methyl groups was obtained by analyzing the ^{1}H-^{13}CHSQC spectrum of a 10% ^{13}C-labeled sample (24). Hydrogen bond restraints from amide hydrogen exchange were determined by recording the ^{1}H-^{15}N HSQC spectrum of a sample left in 99% D_2O buffer for 6 h. Proton distance constraints were derived from three-dimensional ^{15}N-edited NOESY (mixing time 150 ms) and three-dimensional ^{13}C-edited NOESY. NMR data were processed using the MMRPipe/NMRDraw (25) suite and analyzed by NMRView (26).

The initial structures of t-PilS were generated by DYANA (27) using manually assigned unambiguous NOE restraints from ^{15}N-NOESY and ^{13}C-NOESY and dihedral angle restraints predicted by TALOS (28). Many other NOE cross-peaks in the two NOESY spectra were further assigned by CYANA (29) automatically. The unambiguous NOE restraints, automatically assigned NOE restraints, dihedral angle restraints, and hydrogen bond restraints were used for structure calculation. One hundred structures were calculated by DYANA using standard TAD protocol, and 20 conformers with the lowest target function values were selected for further energy refinement in AMBER 7.0 (30). The final ensemble of 10 structures with the lowest amber energies was checked by Procheck-NMR (31) and deposited at the Protein Data Bank.

Model of PilS Pilus—The two subunits of PilS were first docked manually using the program "O" in such a manner that α2 is in close vicinity to α4. The interface was then refined using the program MULTIDOCK with the following parameters. To define the interface, a distance cut-off of 15 Å (cut_iface) was used to include the region for multiple copy side chain rotamer representation. A relatively large distance cut-off of 200 Å (cut_jface) was used to include the whole molecule for fixed side chain representation. For mean field optimization performed at 298.0 K, a residue-residue non-bonded cut-off distance of 15 Å (cut_res_nb) was used. The criterion for convergence in terms of change in energy was set at 0.4 kcal/mol (emax). For rigid body energy minimization, a distance cut-off of 15 Å (cut_lface) across the interface was used for inclusion of residues in the calculation of non-bonded interaction. The maximum rotation step size was 1° and maxi-

mum translation step size 0.3 Å. Minimization continued until the energy of the system decreased by less than $1\times e^{-6}$ kcal/mol (ftol) for any given step. The energy cap for the number of atom-atom van der Waals clashes was set at 2.5 kcal/mol (eatmax).

Inhibition of Bacterial Entry Experiment—The procedures for inhibition of bacterial entry into human intestinal cells by recombinant PilS proteins were described elsewhere (3). Human embryonic intestinal cells INT407 were grown in tissue culture flasks and seeded in 24-well plates to obtain a monolayer in basal medium Eagle's with 15% calf serum (BME) by cultivating overnight at 37 °C in a CO_2 incubator. The medium was removed, and the cells were washed with 1× PBS. 0.35 ml of fresh BME was added, followed by 100 μl of solutions of PilS or mutant PilS to a final protein concentration of 2 μM (1× PBS was used as a control for 100% invasion) and then 50 μl of *S. typhi* bacterial cells in saline ($\sim 3.6 \times 10^7$ cells). The mixture was centrifuged for 10 min at $2{,}164 \times g$ and incubated for 2 h to allow bacterial invasion. The culture medium was removed, and the cells were washed once with 1× PBS. 0.5 ml of fresh BME with 100 μg/ml gentamycin was added, and the cells were incubated for 1 h to kill *S. typhi* that remained outside the intestinal cells. The medium was removed, and the cells were washed twice with 1× PBS to remove any trace of gentamycin. 1 ml of 0.02% Triton X-100 was added, and the cells were pipetted up and down to ensure that the intestinal cells were detached and lysed to release the *S. typhi* cells. The mixture was incubated at 37 °C for 1 h before spreading onto agar plates (with kanamycin) for cell counts.

TABLE I
Structural statistics for the S. typhi t-PilS protein

Restraints for calculation	
Total NOE restraints	2033
Intraresidue	905
Sequential	534
Medium range	225
Long range	369
Hydrogen bond distance restraints	76
Dihedral angle constraints	206
CYANA target function	1.23 ± 0.14
Maximum violation for upper limits (Å)	0.25 ± 0.08
Maximum violation for torsion angles (degree)	3.78 ± 0.76
AMBER energies (kcal/mol)	
Total energy	−3,658 ± 32
Ramachandran statistics (%)	
Most favored	68.1
Additionally allowed	27.6
Generously allowed	3.0
Disallowed	1.2
Root mean square deviation from mean structure (Å)	
Backbone atoms (secondary structure region)	0.68
All heavy atoms (secondary structure region)	1.10

RESULTS

Structure of S. typhi PilS—The construct used for the PilS structural determination is N-terminal-truncated (Met-26-Gly-181, referred as t-PilS), and the truncation of the N-terminal 25 residues is necessary to prevent the pilin from oligomerization. The three-dimensional structure of t-PilS was solved by NMR spectroscopy (Table I). Fig. 2*A* shows a stereoview of the best fit superposition of the family of 15 final structures of t-PilS. The N-terminal five residues of t-PilS (residues 26–30) are not well

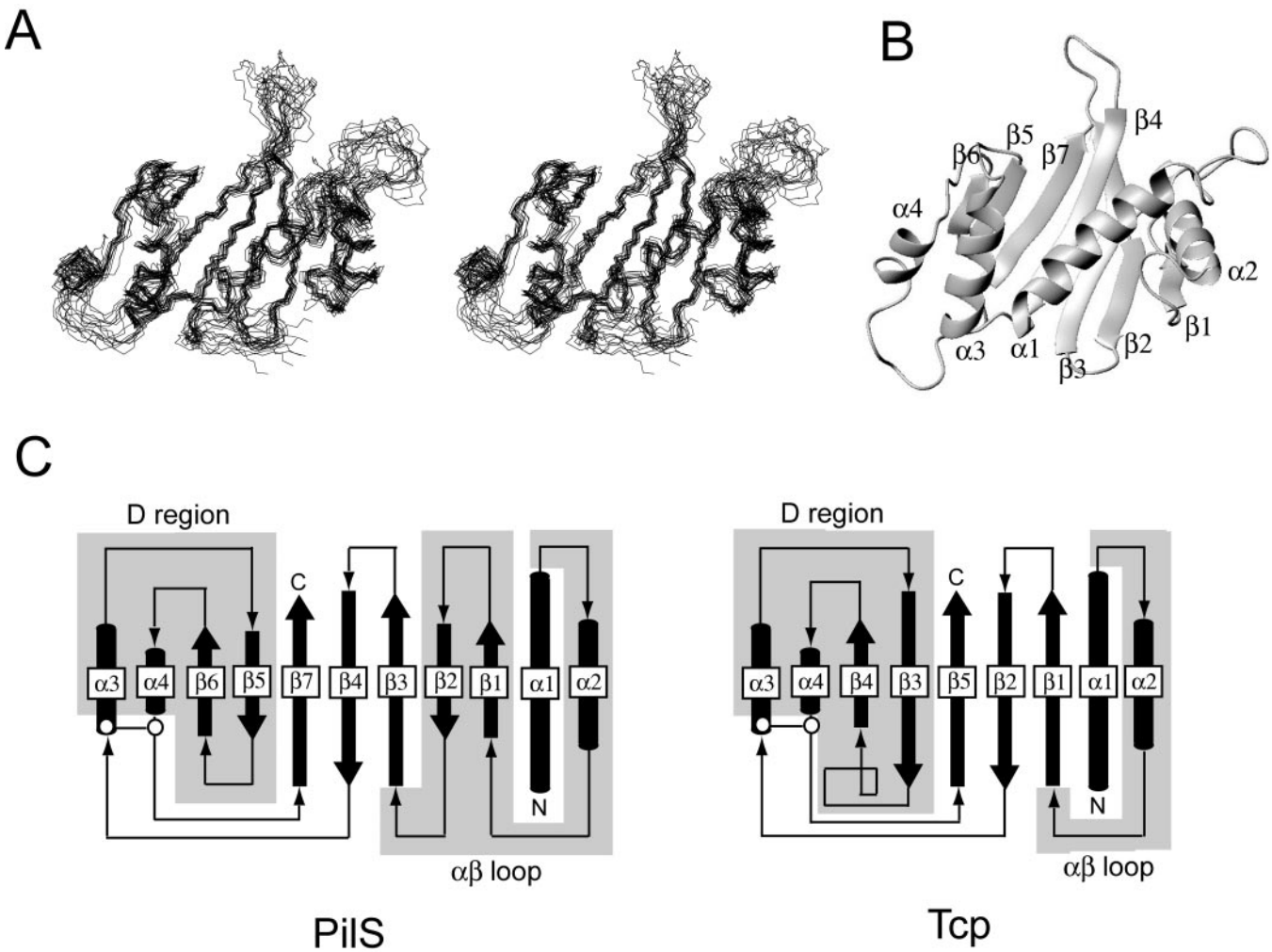

FIG. 2. **NMR structure and topology of t-PilS.** *A*, stereoview showing the best fit superposition of the backbone atoms of the final 15 structures of the *S. typhi* t-PilS protein. *B*, ribbon diagrams of the three-dimensional structure of t-PilS. *Panels A* and *B* were prepared with the program MOLMOL (37). *C*, secondary structure topologies of t-PilS and Tcp pilin. The αβ loop and D region are shown with *shaded* background. Locations of the disulfide bonds are indicated with *open circles*.

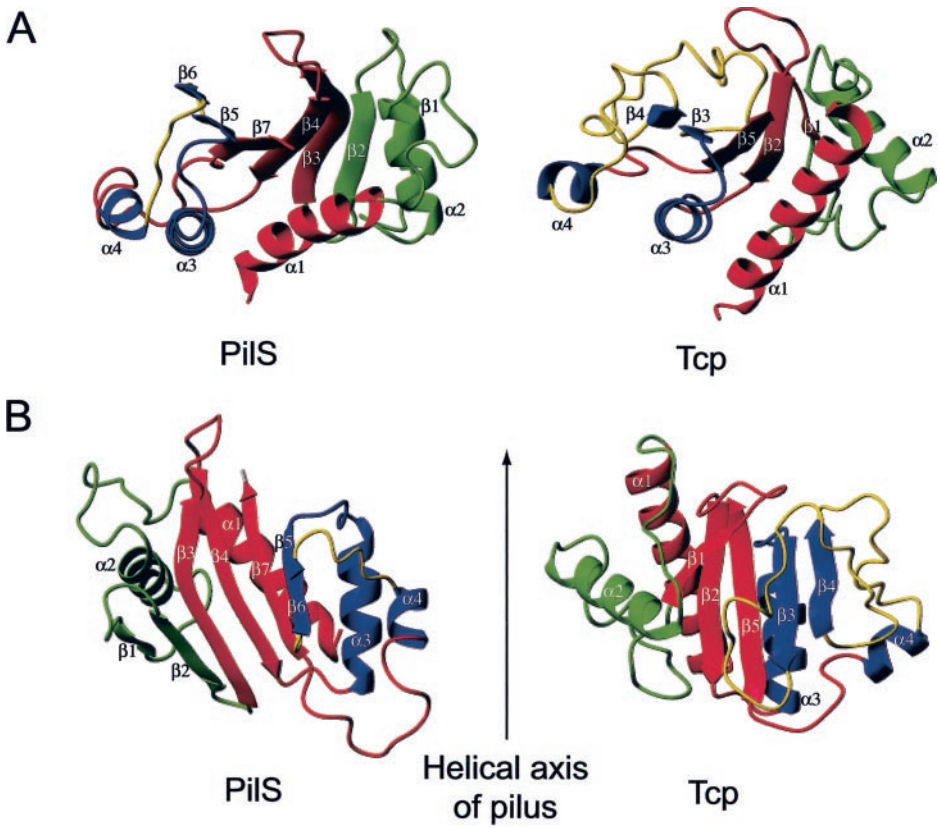

FIG. 3. **Comparison of the structures of t-PilS and Tcp pilin in two different views.** In both panels, the central hydrophobic core, the $\alpha\beta$ loop, and the D region are *red*, *green*, and *blue*, respectively. The extended β3-β4 and β4-α4 loops in Tcp pilin and their corresponding counterparts in t-PilS are *yellow*. *A*, the extended loops in Tcp pilin are highly exposed and cover most of the β-sheet. The much shortened β5-β6 and β6-α4 loops in t-PilS allow different parts of the protein to be exposed. *B*, during pilus assembly, helix α4 of PilS is aligned almost parallel to the helical axis of the pilus; helix α4 of Tcp pilin makes an angle of around 50° with the helical axis. The figure was prepared with the program MOLMOL (37).

defined because of the lack of structural restraints. The other less defined regions in the protein include the loop between helices α1 and -2 (residues 49–60), the loop between strands β1 and -2 (residues 84–86), the loop between strands β3 and -4 (residues 105–112), and the loop between helix α4 and strand β7 (residues 164–171). Residues in these regions have relatively fewer NOE restraints compared with the well defined regions and are likely to be intrinsically more flexible.

The structure of t-PilS reinforces that Type IVb pilins adopt very different folding topologies compared with the structures of Type IVa pilin. The extended N-terminal α-helix (helix α1) is packed against three antiparallel β-strands (strands β3, -4, and -7) to form the hydrophobic core of the protein. The N-terminal $\alpha\beta$ loop on one edge of the β-sheet contains a short α-helix (helix α2) and a pair of well defined antiparallel β-strands (strands β1 and -2). This is in contrast to the sugar loop and minor β-strands found in the $\alpha\beta$ loop of MS11 and PAK pilins, respectively. The C-terminal disulfide-bonded region on the other edge of the β-sheet contains a pair of α-helices (helices α3 and -4) lying on top of two shorter antiparallel β-strands (strands β5 and -6). The seven β-strands align in an antiparallel manner to form a continuous but twisted β-sheet with all helices on one side of the β-sheet (Fig. 2, *A* and *B*). In Type IVa pilins, the hypervariable disulfide loop contains either a β-hairpin connected to a loop (12) or β-turns connected to each other (14, 32). These differences may relate to the fact that Types IVa and IVb pili have completely different receptor substrate specificities.

Although topologically similar, t-PilS contains distinct structural features in both the $\alpha\beta$ loop and D region when compared with Tcp (Fig. 2*C*). The $\alpha\beta$ loop of t-PilS contains an extra pair of antiparallel β-strands (strands β1 and -2) between helix α2 and strand β3. The corresponding region in the Tcp pilin, between helix α2 and strand β1, is replaced with an extended loop of 18 residues (Fig. 1). The extra pair of β-strands in t-PilS aligns continuously at the edge of the major β-sheet (strands β3, -4, and -7) and allows helix α2 to pack on top at an angle of ~124° to helix α1. In the Tcp pilin, helix α2 lies perpendicular to helix α1 (~90°) and across on the outside of the major β-sheet. Consequently, helix α2, strand β1, and the α1-α2 loop of t-PilS are exposed at the edge of the molecule, whereas only helix α2 is exposed at the same edge in the Tcp pilin (Fig. 3, *A* and *B*).

Even larger differences between the structures of t-PilS and Tcp pilin are found in the disulfide-bonded D region (19). The D region of Tcp pilin encompasses 65 residues and is the longest among known Type IVb pilins. The D region of bundle-forming pilus pilin from enteropathogenic *E. coli* is somewhat shorter, being composed of 49 residues. PilS and pilin of R64 thin pilus have the shortest D regions among Type IVb pilins, consisting of only 36 residues, similar in length to that of the Type IVa pilin of *N. gonorrhoeae* pilus (29 residues) (Fig. 1). The extra length in the D region of the Tcp pilin is contributed by two extended loops between strands β3 and -4 (24 residues) and between strand β4 and helix α4 (12 residues). These loops contain overlapping epitopes for protective Tcp antibodies and encompass most of the functional domain residues (19). The corresponding regions in t-PilS, between strands β5 and -6 and between strand β6 and helix α4, are only 2 and 6 residues in length, respectively. The β3-β4 loop in Tcp pilin is highly exposed on the surface of the pilus and covers most of the β-sheet on the opposite side of the N-terminal α-helix (Fig. 3*A*). In contrast to Tcp pilin, the β5-β6 loop is relatively short in t-PilS; most of the residues on the β-sheet opposite to the N-terminal helix are thus predicted to be exposed on the surface of the pilus. In addition to covering the β-sheet, the β3-β4 loop in Tcp pilin is also extended over the edge of the β-sheet at the D region. To expose functional residues found mainly in the β4-α4 loop and helix α4, the β4-α4 loop is extended and helix α4 packs loosely away from helix α3 and makes an angle of around 50° with the helical axis of the pilus (Fig. 3*B*). On the contrary, helices α3 and -4 of t-PilS align tightly with each other, and both are almost parallel along the helical axis of the pilus.

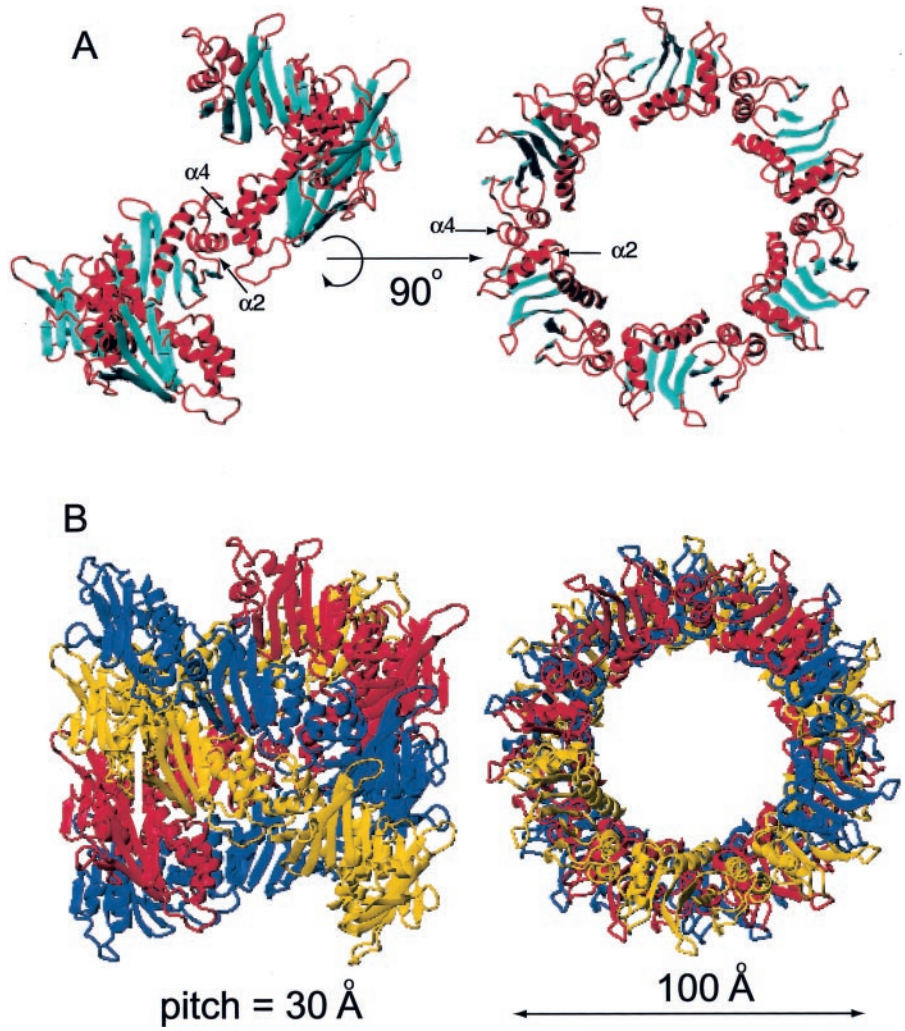

FIG. 4. **A model of PilS pilus assembly.** *A*, each left-handed helical strand of the pilus contains six subunits of PilS with their α2 helices interacting with α4 helices of an adjacent subunit. α-helices and loops are *red*; β-sheets are *cyan*. *B*, the PilS pilus can form a three-start helix with a pitch of 30 Å between each helical strand and an outer diameter of ~100 Å. The N-terminal hydrophobic helices of each subunit are packed at the central core of the pilus. The 3 helical strands of the PilS pilus are *red*, *yellow*, and *blue*. The figure was prepared with the program MOLMOL (37).

Model of S. typhi Pilus Assembly—A model of *S. typhi* pilus was built based on our t-PilS structure and the model of Tcp (19). The postulated association interface (α2 against α4) between two adjacent t-PilS subunits was energy-refined (total energy = −34.6 kcal/mol) using the program MULTIDOCK (33). The refined dimeric subunit was used to build a single helical strand that was subsequently intertwined with another two identical helical strands to form the pilus model. The model shows that t-PilS can also be arranged into a left-handed three-start helix. Each turn of the helical strand contains six subunits of t-PilS with their α2 helices interacting with the α4 helices of an adjacent subunit in a perpendicular fashion (Fig. 4*A*). The helix α4 of t-PilS, unlike that of Tcp, is almost parallel to the helical axis of the pilus. Three such helical strands intertwine to form a cylindrical pilus with a helical pitch of around 30 Å and an outer diameter of around 100 Å (Fig. 4*B*). This arrangement of the pilus places the hydrophobic N-terminal helix (helix α1) of every subunit at the innermost core of the PilS pilus, where it can pack to form a coiled coil providing mechanical strength and stability.

In addition to the N-terminal hydrophobic helix, there are polar interfaces buried between stacked subunits of t-PilS from different helical strands. These potential "structural" charged residues at the "tip" side of the molecule are located on the α1-α2 loop (Lys-51), helix α3 (Arg-135), and helix α4 (Glu-157). Those at the "base" side of the molecule are located on the α2-β1 loop (Lys-75), the β4-α3 loop (Lys-120), and the α4-β7 loop (Asp-166 and Arg-169). These charged residues found in both the $\alpha\beta$ loop and D region are not exposed but could form salt bridges or side chains to main chain hydrogen bonds to secure the position of neighboring helical strands. In Tcp, these structural charged residues are found within the D region spanning helix α3 to strand β4. In contrast to the above mentioned residues, charged residues found among the interface formed between α2 and -4 of two adjacent subunits or on the β-sheet of each subunit are exposed on the surface of the pilus. These exposed charged residues in the $\alpha\beta$ loop are found mainly on the α1-α2 loop (Asp-54), helix α2 (Lys-63 and Asp-66), and the β1-β2 loop (Asp-82). Those found in the D region are mainly located on or around helix α4 (Asp-151, Lys-153, and Glu-162). Asp-103 and Lys-143 are the two exposed charged residues that are found on β3 and -5 of each PilS subunit, respectively (Fig. 5*A*). All these charged residues are exposed (Fig. 5*B*) and could be involved in a functional role of the PilS pilus by interacting directly with the first extracellular domain of CFTR.

Peptide Binding Surface of the S. typhi Pilus—The first extracellular domain of CFTR (residues 103–117) was identified as a binding receptor on human intestinal cells for *S. typhi* pilus (4, 5). A 10-residue peptide composed of residues 108–117 (SYDPDNKEER) of CFTR was enough to inhibit entry of *S. typhi* into intestinal cells. This inhibitory effect was not found in another 10-residue peptide composed of residues 103–112 (GRIIASYDPD) of CFTR, indicating that residues 113–117 (NKEER) of CFTR are essential for *S. typhi* pilus binding (5). Interestingly, the same 10-residue peptide was also found to be effective in inhibiting entry of *P. aeruginosa* into corneal cells (34). Because the peptide 108–117 of CFTR contains four neg-

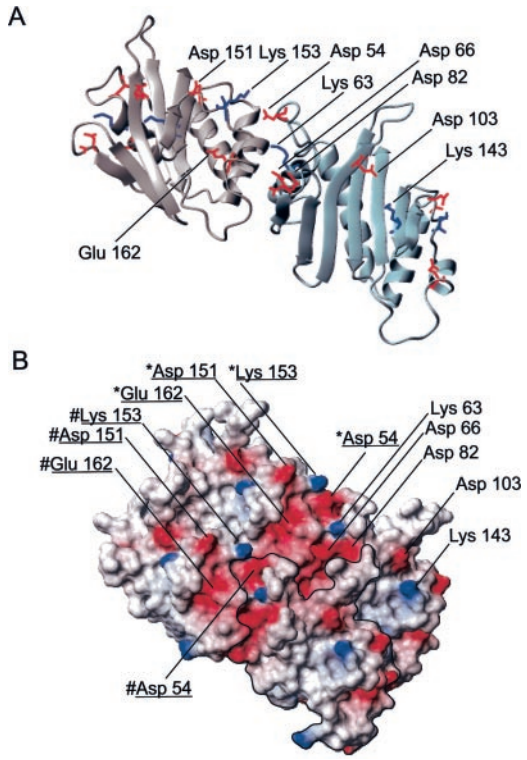

FIG. 5. **Peptide binding surface of PilS pilus.** *A*, ribbon diagram showing exposed and charged residues on t-PilS. The two subunits are *gray* and *light blue*, respectively. Exposed and charged residues at the subunit interfaces and on the β-sheet are shown in *wire-frame models*. Positively charged residues are *blue*. Negatively charged residues are *red*. Note that residues are labeled only on one of the subunits. *B*, molecular surface of four subunits of PilS on the pilus model showing the distribution of charges and locations of exposed charged residues. Positively and negatively charged surfaces are *blue* and *red*, respectively. Exposed charged residues found to be essential for bacterial cell entry into human intestinal cells are *underlined*. An *asterisk* next to the label represents functional residues from adjacent subunits on the same helical strand. A *hash sign* next to the label represents functional residues from adjacent subunits on a different helical strand. The border of one of the subunits is *highlighted* for easier identification. The figure was prepared with the program MOLMOL (37).

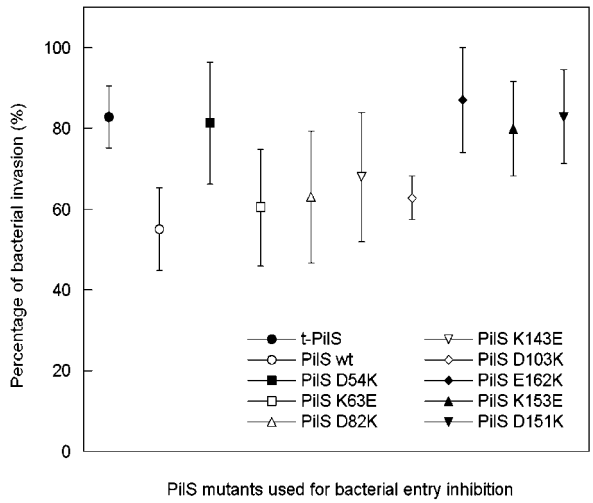

FIG. 6. **Inhibition of bacterial invasion by PilS mutants.** The percentage of bacterial invasion is an average of four separate experiments. The *error bar* indicates S.D. For each experiment, the number of bacterial colonies was compared with the control, which used only 1× PBS, set as 100% bacterial invasion. All the mutants were confirmed to have a similar CD spectrum and elution profile on a gel filtration column as did the wild-type PilS protein. t-PilS eluted at a volume corresponding to the monomeric molecular weight of PilS (data not shown).

atively charged residues and two positively charged residues, the binding site for this peptide on PilS pilus is expected to contain polar or charged residues.

Exposed and charged residues in t-PilS are mainly found on or around the helices α2 (Asp-54, Lys-63, Asp-66, and Asp-82) and α4 (Asp-151, Lys-153, and Glu-162) at the interface between two adjacent subunits and also on the exposed β-sheet (Asp-103 and Lys-143) (Fig. 5*A*). Fig. 5*B* shows the locations of all the exposed and charged residues found on the surface of PilS pilus. In an attempt to determine the exact binding site of CFTR on PilS, we have synthesized a 10-residue peptide with a sequence derived from residues 108–117 (SYDPDNKEER) of CFTR. The peptide was titrated into a ^{15}N-labeled t-PilS NMR sample, but to our surprise no binding could be observed (data not shown). The inability of t-PilS to bind CFTR is further confirmed by the loss of its ability to inhibit entry of *S. typhi* into intestinal cells (Fig. 6). Only full-length PilS protein that contains the N-terminal hydrophobic oligomerization domain retained the ability to bind CFTR and inhibit entry of *S. typhi* into intestinal cells. To identify the exposed charged residues that could involve direct interaction with CFTR, we have generated eight mutants of the full-length PilS protein and tested their abilities to inhibit entry of *S. typhi* into intestinal cells (Fig. 6). Asp-66 was not tested because the protein tends to precipitate out of solution. Results of the inhibition of bacterial entry experiment show that most of the "functional" charged residues involved in CFTR binding are located on or around helix α4 in the D region (residues Asp-151, Lys-153, and Glu-162). This agrees with the findings on Tcp pilin in which all of its functional residues are found exclusively on or around α4 (19). Unlike Tcp pilin, the residue on the α1-α2 loop of PilS (Asp-54), which is in close proximity to Lys-153 on helix α4 of an adjacent subunit, is also essential for CFTR binding. Other charged residues that are exposed on the αβ loop (Lys-63 and Asp-82) and the β-sheet (Asp-103 and Lys-143) of PilS are not essential for CFTR binding.

DISCUSSION

Type IV pili plays a crucial role in pathogenesis of many bacterial species by providing adherence to host cells. All Type IV pilins identified to date adopt an αβ roll fold with the N-terminal hydrophobic helix packed against three antiparallel β-strands to form a ladle-shaped molecule. Different classes of Type IV pilins, however, differ substantially in their N-terminal αβ loops and C-terminal disulfide-bonded D regions. Type IVa pilin has either a sugar loop (MS11) or minor β-strands (PAK) in the αβ loop, whereas Type IVb pilin has a well defined α-helix in the αβ loop that is involved in subunit interaction. The disulfide-bonded receptor binding loop in Type IVa pilin is composed of either two type I β-turns joined together (MS11 and K122–4) (14) or one type I β-turn joined to a type II β-turn (PAK) (35). The corresponding region in Type IVb pilin (Tcp) contains a much more elaborate structure of a pair of α-helices packed on top of a pair of antiparallel β-strands.

Here, we have described the structure of the Type IVb pilin from *S. typhi*, which has an extra pair of β-strands in the αβ loop and much shortened loops in the D region compared with Tcp pilin, the only Type IVb pilin with known structure. Sequence alignment shows that there is only 16% sequence identity between *S. typhi* PilS and Tcp, whereas there is close to 50% identity between PilS and the Type IVb pilin from R64 thin pilus, which is required for bacterial mating (Fig. 1) (10). We propose to classify the Type IVb pilin from *S. typhi* and R64 plasmid as a distinct subclass of Type IVb pilin having significantly shortened loops in the D region from those in Tcp and

bundle-forming pilus. Differences between the $\alpha\beta$ loop and D region of the two subclasses of Type IVb pilins affected not just the assembly of the pili but also the availability of exposed residues in the vicinity of the D region for receptor binding.

Although the structures of both Type IVb pilins belong to the same $\alpha\beta$ roll class as Type IVa pilins, the addition of extra α-helices at both ends of the β-sheet in Type IVb pilin generates a unique pattern of pilus assembly. Type IVa pilin forms either a right-handed (MS11 and PAK) or left-handed (K122–4) 1-start helix with five subunits/helical turn. The N-terminal α-helices interacting with each other at the middle of the pilus form a coiled coil structure, whereas the interface between adjacent subunits is formed by continuous β-strand hydrogen bonding. The Type IVa pilus is relatively thin and fragile with a helical pitch of around 40 Å and a diameter of around 50 Å (K122–4). Type IVb pilin, such as Tcp, adopts a left-handed three-start helix of six subunits/turn with α2 from the $\alpha\beta$ loop of one subunit interacting with α4 from the D region of an adjacent subunit. The Type IVb pilus is relatively thick and sturdy with a helical pitch of 45 Å and a diameter of 80 Å (Tcp).

The proposed model shows that *S. typhi* PilS can also be assembled into a left-handed three-start helix with six subunits/helical strand. Despite the difference in the orientation of α4 relative to the helical axis, the helical pitch (30 Å) and diameter (100 Å) of the PilS pilus model are comparable with those, 37.5 and 100 Å, respectively, obtained for the crystallographic fibers of N-terminal-truncated Tcp. The helical pitch and diameter of the PilS model, however, differ from the intact Tcp filaments, which have a 45-Å pitch and a diameter of ~80 Å (19). Addition of extra pulling forces from the N-terminal helices is expected to increase the helical pitch and reduce the diameter of the intact *S. typhi* pilus. This speculation also agrees with the measured diameter of intact *S. typhi* pilus by electron microscopy, which is around 6–8 nm (data not shown).

CFTR has long been recognized as the cellular receptor for *S. typhi* entry. Human cells expressing wild-type CFTR ingested significantly more *S. typi* than cells expressing ΔF508 CFTR (5). This entry can be inhibited either with a peptide derived from the first extracellular domain of CFTR (5) or with soluble recombinant protein of PilS (3). The CFTR peptide, but not a scrambled one, can effectively neutralize the cell entry inhibition afforded by the PilS protein and vice versa (4). Because the proposed receptor of PilS on CFTR carry mostly charged residues, we have generated mutants for most of the exposed and charged residues on PilS according to the model to pinpoint its receptor binding site. Based on data from inhibition of bacterial entry by full-length PilS protein, the binding site for CFTR on PilS pilus might be at the interface between adjacent subunits with charged residues not only from the D region but also from the $\alpha\beta$ loop. Depending on the conformation of the bound CFTR peptide, the binding site could be between adjacent subunits on the same helical strand or even between subunits on different helical strands (Fig. 5*B*). In Tcp, functional charged residues that extend outward to the surface of the Tcp filament for interaction with other pili, bacteriophage, or host cells are found exclusively on or adjacent to helix α4 in the D region (19).

This finding also explains why the monomeric t-PilS is unable to bind CFTR, because the complete binding site could be comprised of charged residues from different subunits. CFTR binding is strictly dependent on the assembly of PilS pilin into pilus and might be for improved specificity and cooperativity of binding. This is the first report on structural study of PilS from *S. typhi*. Identification of the structure of t-PilS and its pattern of pilus assembly, as well as candidate residues that are essential for receptor binding, not only provides the basis for other structural studies but also opens the possibility of rational drug or vaccine design against this adhesin that is commonly used by *S. typhi* for host invasion.

Acknowledgments—We thank Daiwen Yang for providing technical support on NMR experiments setup and K. Swaminathan for advice in using the software "O."

REFERENCES

1. Hahn, H. P. (1997) *Gene* **192,** 99–108
2. Zhang, X.-L., Morris, C., and Hackett, J. (1997) *Gene* **202,** 139–146
3. Zhang, X.-L., Tsui, S.-M. I., Yip, M.-C. C., Fung, W.-Y. A., Wong, K.-H. D., Dai, X.-Y., Yang, Y.-H., Hackett, J., and Morris, C. (2000) *Infect. Immun.* **68,** 3067–3073
4. Tsui, S.-M. I., Yip, M.-C. C., Hackett, J., and Morris, C. (2003) *Infect. Immun.* **71,** 6049–6050
5. Pier, G. B., Grout, M., Zaidi, T., Meluleni, G., Mueschenborn, S. S., Banting, G., Ratcliff, R., Evans, M. J., and Colledge, W. H. (1998) *Nature* **393,** 79–82
6. Sakai, D., and Komano, T. (2002) *J. Bacteriol.* **184,** 444–451
7. Manning, P. A. (1997) *Gene* **192,** 63–70
8. Bieber, D., Ramer, S. W., Wu, C.-Y., Murray, W. J., Tobe, T., Fernandez, R., and Schoolnik, G. K. (1998) *Science* **280,** 2114–2117
9. Giron, J. A., Gomez-Duarte, O. G., Jarvis, K. G., and Kaper, J. B. (1997) *Gene* **192,** 39–43
10. Kim, S.-R., and Komano, T. (1997) *J. Bacteriol.* **179,** 3594–3603
11. Soto, G. E., and Hultgren, S. J. (1999) *J. Bacteriol.* **181,** 1059–1071
12. Parge, H. E., Forest, K. T., Hickey, M. J., Christensen, D. A., Getzoff, E. D., and Tainer, J. A. (1995) *Nature* **378,** 32–38
13. Hazes, B., Sastry, P. A., Hayakawa, K., Read, R. J., and Irvin, R. T. (2000) *J. Mol. Biol.* **299,** 1005–1017
14. Keizer, D. W., Slupsky, C. M., Kalisiak, M., Campbell, A. P., Crump, M. P., Sastry, P. A., Hazes, B., Irvin, R. T., and Sykes, B. D. (2001) *J. Biol. Chem.* **276,** 24186–24193
15. Rudel, T., Scheuerpflug, I., and Meyer, T. F. (1995) *Nature* **373,** 357–359
16. Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J., and Knight, S. D. (1999) *Science* **285,** 1061–1066
17. Kuehn, M. J., Heuser, J., Normark, S., and Hultgren, S. J. (1992) *Nature* **356,** 252–255
18. Sakellaris, H., Munson, G. P., and Scott, J. R. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 12828–12832
19. Craig, L., Taylor, R. K., Pique, M. E., Adair, B. D., Arvai, A. S., Singh, M., Lloyd, S. J., Shin, D. S., Getzoff, E. D., Yeager, M., Forrest, K. T., and Tainer, J. A. (2003) *Mol. Cell* **11,** 1139–1150
20. Wittekind, M., and Mueller, L. (1993) *J. Magn. Reson. B* **101,** 201–205
21. Grzesiek, S., and Bax, A. (1993) *J. Biomol. NMR* **3,** 185–204
22. Grzesiek, S., Anglister, J., and Bax, A. (1993) *J. Magn. Reson. B* **101,** 114–119
23. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., and Forman-Kay, J. D. (1993) *J. Magn. Reson.* **B 101,** 333–337
24. Szyperski, T., Neri, D., Leiting, B., Otting, G., and Wüthrich, K. (1992) *J. Biomol. NMR* **2,** 323–334
25. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293
26. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR* **4,** 603–614
27. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J. Mol. Biol.* **273,** 283–298
28. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
29. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) *J. Mol. Biol.* **319,** 209–227
30. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. R., Cheatham, T.E., III, DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) *Comp. Phys. Commun.* **91**, 1–41
31. Laskowski, R. A., Rullman, J. A., Mac-Arthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
32. Campbell, A. P., Bautista, D. L., Tripet, B., Wong, W. Y., Irvin, R. T., Hodges, R. S., and Sykes, B. D. (1997) *Biochemistry* **36,** 12791–12801
33. Jackson, R. M., Gabb, H. A., and Sternberg, J. E. (1998) *J. Mol. Biol.* **276,** 265–285
34. Zaidi, T. S., Lyczak, J., Preston, M., and Pier, G. B. (1999) *Infect. Immun.* **67,** 1481–1492
35. Campbell, A. P., McInnes, C., Hodges, R. S., and Sykes, B. D. (1995) *Biochemistry* **34,** 16255–16268
36. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
37. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graph. Model.* **14,** 51–55