Crystal Structures of a CTXφ pIII Domain Unbound and in Complex with a Vibrio cholerae TolA Domain Reveal Novel Interaction Interfaces

Received for publication, July 23, 2012, and in revised form, August 20, 2012. Published, JBC Papers in Press, August 31, 2012, DOI 10.1074/jbc.M112.403386

Christopher G. Ford, Subramaniapillai Kolappan, Hanh T. H. Phan, Matthew K. Waldor, Hanne C. Winther-Larsen, and Lisa Craig

From the Department of Molecular Biology and Biochemistry Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada, School of Pharmacy and Laboratory of Microbial Dynamics University of Oslo, PB 1068 Blindern 0316, Oslo Norway, and Division of Infectious Diseases, Brigham and Women’s Hospital/Harvard Medical School and Howard Hughes Medical Institute, Boston, Massachusetts 02115

Background: CTXφ infection of Vibrio cholerae confers toxigenicity.

Results: We report crystal structures of CTXφ-pIII domain N1 alone and bound to V. cholerae TolA domain C.

Conclusion: CTXφ and coliphage pIII use distinct mechanisms to bind to TolA.

Significance: These structures contribute to our understanding of a critical step in the evolution of pandemic V. cholerae with implications for emerging V. cholerae pathogens.

Vibrio cholerae colonize the small intestine where they secrete cholera toxin, an ADP-ribosylating enzyme that is responsible for the voluminous diarrhea characteristic of cholerare disease. The genes encoding cholera toxin are located on the genome of the filamentous bacteriophage, CTXφ, that integrates as a prophage into the V. cholerae chromosome. CTXφ infection of V. cholerae requires the toxin-coregulated pilus and the periplasmic protein TolA. This infection process parallels infection of V. cholerae by the Ff family of filamentous coliphage. Here we demonstrate a direct interaction between the N-terminal domain of the CTXφ minor coat protein pIII (pIII-N1) and the C-terminal domain of TolA (TolA-C) and present x-ray crystal structures of pIII-N1 alone and in complex with TolA-C. The structures of CTXφ pIII-N1 and V. cholerae TolA-C are similar to coliphage pIII-N1 and E. coli TolA-C, respectively, yet these proteins bind via a distinct interface that in E. coli TolA corresponds to a colicin binding site. Our data suggest that the TolA binding site on pIII-N1 of CTXφ is accessible in the native pIII protein. This contrast with the Ff family phage, where the TolA binding site on pIII is blocked and requires a pilus-induced unfolding event to become exposed. We propose that CTXφ pIII accesses the periplasmic TolA through retraction of toxin-coregulated pilus, which brings the phage through the outer membrane pilus secretin channel. These data help to explain the process by which CTXφ converts a harmless marine microbe into a deadly human pathogen.

Vibrio cholerae are Gram-negative, rod shaped bacteria that cause the gastrointestinal disease cholera. There are >200 known V. cholerae serogroups, yet only two, O1 and O139, cause pandemic disease. Pathogenic serogroups are distinguished from non-pathogenic strains by the acquisition of two mobile genetic elements: the Vibrio pathogenicity island (1) and the CTX element, which is a prophage of the filamentous bacteriophage CTXφ (2, 3). The Vibrio pathogenicity island contains the tcp operon encoding the toxin-coregulated pilus (TCP)4, which is necessary for V. cholerae colonization of the human intestine (4, 5), andTcpF, a soluble colonization factor protein (6, 7). The CTX element contains repetitive sequences involved in site-specific recombination of the CTXφ genome and a core region encoding the phage assembly proteins and the cholera toxin subunits A and B. Expression of tcp and the cholera toxin genes ctxA and ctxB are regulated by the same transcriptional activator, ToxT (8, 9).

CTXφ has a 7-kb single-stranded circular DNA genome that integrates into the V. cholerae chromosome as a prophage. Phage particles are produced from extrachromosomal DNA replicated as a plasmid (10). The CTXφ genome is similar in gene size and organization to Ff family filamentous phage that infect Escherichia coli (coliphage), including M13 and fd (3). The core region of the CTXφ prophage contains genes cep, orfU, ace, zot, ctxA, and ctxB. Although not homologous in

* This work was supported by the Natural Sciences and Engineering Research Council, National Institutes of Health Grant NIH R37 AI-42347, and the Howard Hughes Medical Institute.

1 Supported in part by the Kristine Bonnevie and EMBOJ fellowship, University of Oslo, and by the Research Council of Norway.

2 Supported by a New Investigator Award from the Canadian Institutes of Health Research and a Scholar Award from the Michael Smith Foundation for Health Research. To whom correspondence should be addressed. Tel: 778-782-7140; Fax: 778-782-5583; E-mail: licraig@sfu.ca.

3 The abbreviations used are: TCP, toxin-coregulated pilus; TEM, transmission electron microscopy; LB, Luria-Bertani; Sm, streptomycin; Km, kanamycin; Ap, ampicillin; Tc, tetracycline; PBS, phosphate buffered saline; BSA, bovine serum albumin; Selenomethionine; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; Ni-NTA, nickel-nitrilotriacetic acid.

4 The atomic coordinates and structure factors (codes 4G7W and 4G7X) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).
sequence, the cep, orfII, ace, and zot genes are similar in size and synteny to the M13 genes VIII, III, VI, and I, which encode the structural proteins pVIII, pIII, pVI, and pI, respectively (3, 11). In M13, the major coat protein pVIII forms the long, cylindrical phage coat that packages the phage genome. pIII, pV, and pI are minor coat proteins located at the phage tips (12–14). Although pVI and pI are very small, pIII is a large protein, present in four- to five copies at one end of the phage particle. pIII (also called g3p for gene 3 protein) mediates phage binding, uptake, and assembly (15–19).

pIII is well characterized in Ff phage, in part because of its application in phage display technology (20, 21). fd and M13 pIII amino acid sequences are almost identical, with an 18-amino acid signal peptide and a 406-amino acid mature protein organized into three distinct functional domains, N1, N2, and C (also called D1, D2, and D3 or CT), linked by glycine-rich segments of low structural complexity, LCR1 and LCR2 (19, 22, 23)). The hydrophobic C-terminal segment is required for insertion of the virion into the inner membrane and for excision after phage assembly (15, 24). CTXφ pIII has very little sequence homology to the Ff pIII proteins but is predicted to have a 14-amino acid signal peptide and 3 domains separated by serine/proline-rich linkers (25), and CTXφ pIII-C contains a very hydrophobic segment that likely represents an inner membrane anchor.

Ff coliphage bind to E. coli using their minor coat protein, pII, in a two-step process to initiate infection. First, the central N2 domain of pIII binds to the F pilus tip (26–28), which spontaneously retracts (29–31) to bring the pIII N-terminal domain, pIII-N1, into contact with the C-terminal domain of the periplasmic protein, TolA (32–35). The TolA binding site on pIII-N1 is buried by pIII-N2 (36), and the interaction with the F pilus is required not only for bringing the phage to the bacterial surface but for inducing a conformational change in pIII that exposes the TolA binding site (18, 22, 23). The mechanisms by which the F pilus retracts and the bacteriophage gains entry into the periplasm to bind to TolA are not understood. The F pilus is required for efficient infection, but Ff phage can infect E. coli lacking the F pilus, albeit much lower levels than for F+ strains (18, 27). In contrast, E. coli lacking TolA are resistant to phage infection (18, 35).

The CTXφ phage utilize a similar mechanism to infect V. cholerae, as both TCP and TolA are required for maximal infectivity (3, 11). Heilpern and Waldor (25) generated recombinant fd phage in which CTXφ pIII domains N1, N2, or N1N2 were fused to the N terminus of fd pII or to deletion mutants lacking N1N2 and assessed the ability of these hybrid phage to infect V. cholerae. fd hybrid phage displaying CTXφ pIII-N1N2 fused to domain D3 of fd pIII showed high levels of infectivity for V. cholerae. When only CTXφ pIII-N1 was present, infectivity was reduced but measurable, but when only CTXφ pIII-N2 was present, infectivity was undetectable, demonstrating that whereas N2 is important for efficient phage uptake, domain N1 is critical.

Crystal structures are available for the N-terminal domains of both M13 and fd pIII, which are 99% identical in amino acid sequence (22, 23). These structures reveal two discrete domains of similar, predominantly β-sheet fold joined by a linker and by a crossover of the C-terminal strand of N2 onto N1. N1 has a short N-terminal α-helix followed by a four-stranded β-barrel motif, and N2 is dominated by a twisted β-sheet. Not surprisingly, the glycine-rich LCR1 is disordered and is not resolved in the crystal structure. However, rather than acting as a linker between N1 and N2, this segment lies within N1 and is followed by an ordered loop and a β-strand, β5, that are part of the N1 domain. A short linker connects N1 to N2. The two domains also interact via non-covalent contacts between their β-sheet loops and by the C-terminal strand of N2, which extends across to N1 as a β-strand, β13, to form a 2-stranded β-sheet with β5 that lies across the N1 β-barrel.

A crystal structure was also determined of a fusion protein consisting of the M13 pIII-N1 and LCR1 fused to the C-terminal domain of TolA (36). The N1 fold is similar in both the pIII-N1/TolA-C fusion protein structure and the N1N2 structure but terminates just C-terminal to the disordered LCR1 in the fusion protein and, hence, lacks strand β5 and the remaining N2 domain. TolA-C interacts with N1 on a face of the β-barrel that is occupied by the β5p13 sheet and part of the N2 domain in the N1N2 structure. Thus, for TolA to bind to pIII, pIII must partially unfold, separating N1 and N2 and removing the β5p13 strand to expose the TolA binding site. The trigger that induces this conformational change is the interaction of N2 with the F pilus (18, 22, 23). This movement is proposed to involve isomerization of the Gln-212–Pro-213 bond immediately after β13 (37).

Commonalities and differences between the pIII proteins of CTXφ and the Ff phage prompted us to investigate the interaction between CTXφ and V. cholerae at a molecular level. Here we show by transmission electron microscopy (TEM) that phage binding to V. cholerae does not require TolA, but phage uptake does. We demonstrate a direct interaction between pIII-N1 and TolA-C and describe crystal structures of pIII-N1 alone and in complex with TolA-C. Our data reveal similarities in structure but surprising differences between CTXφ and the coliphage in their interactions with TolA, advancing our understanding of CTXφ infection of V. cholerae.

**EXPERIMENTAL PROCEDURES**

**Bacterial Strains**—Bacterial strains, plasmids, and primers are listed in Table 1. E. coli strains were grown in Luria-Bertani (LB) broth at 37 °C with appropriate antibiotics. V. cholerae strains were grown in LB, pH 6.5, Sm at 30 °C on a Ferris wheel (TCP-inducing conditions). Antibiotics were used at a final concentration of 200 μg/ml streptomycin (Sm), 50 μg/ml kanamycin (Km), 100 μg/ml ampicillin (Ap), and 12 μg/ml tetracycline (Tc). Anti-TcpA antibodies were a gift from Ronald Taylor (Geisel School of Medicine).

**CTXφ Transduction Assays**—V. cholerae strains O395, DH3, and TCP2 used in the infection assays were grown overnight in TCP-inducing conditions. To prepare CTXφ, V. cholerae CL101 cells were grown under the same conditions but with Km in addition to Sm. CL101 cells produce CTX-Kmφ in which the ctxA gene is replaced with the Km resistance marker (3). To produce fdΔ1-pIII<sup>CTX</sup>(15–274) phage, E. coli DH5α-λpir(fdΔ1-pII<sup>CTX</sup>(15–274)) cells were grown at 37 °C overnight with Tc, 1.2 μg/ml CL101, and E. coli DH5α-λpir cells...
were removed from the supernatant containing the phage by centrifugation and filtration on a 0.2-μm pore filter. Transduction assays were performed by mixing 75 μl of phage with 75 μl of *V. cholerae* overnight culture and shaking for 20 min at room temperature. Serial dilutions of the infection mixture were plated on LB-Sm/Km agar plates to enumerate the input bacteria. Cells were grown overnight at 37 °C, and colony-forming units (cfu) were counted. The phage transduction frequency was calculated as the ratio of the number of transductants to input *V. cholerae* cells.

For the transfection inhibition assay, the filtered cell supernatant containing CTXφ and on LB-Sm/Km plates to enumerate the input bacteria. The phage transduction frequency was calculated as the ratio of the number of transductants to input *V. cholerae* cells.

### Crystal Structures of CTXφ pIII and pIII-TolA

| Strains | Description or nucleotide sequence | Source/Reference |
|---------|------------------------------------|-----------------|
| E. coli BL21 (DE3) | F– ompT hsdS7 (r− m−) gal dcm (DE3) | Novagen |
| E. coli Rosetta-gami B (DE3) | F– ompT hsdS7 (r− m−) gal dcm lacY1 ahpC (DE3) gorS22-Tn 10 trxB pRARE (CmR, KmR, TcR) | Novagen |
| E. coli DH5α-Api | endA1 hisD17 galU44 thi-1 recA1 gyrA relA1 (lacZΔM15) Apri+ | 53 |
| V. cholerae O395 | O1 classical Ogawa, Sm2 | R. Taylor |
| V. cholerae DR3 | O395 ΔtolA-pDH149 (tolA bp 79–530 inserted into pGP704) | 11 |
| V. cholerae TCP2 | O395 ΔtcpA | 4 |
| V. cholerae CL101 | O1 El Tor, pCTX-Km | 3 |

### Plasmids

- **pET-15b**
  - T7 promoter, His-tag coding sequence, T7 terminator, lacI coding sequence, pBR322 origin, bla (ApR)
- **pET-15bpIII-N1**
  - pET vector with insertion of CTXφ pIII DNA encoding residues −5 to +96 at Ndel and BamH1 sites downstream of His-tag coding sequence
- **pET-15bpIII−ΔTM**
  - pET vector with insertion of CTXφ pIII DNA encoding residues −5 to +355 at Ndel and BamH1 sites downstream of His-tag coding sequence; contains a mutation encoding a S65T substitution in pIII−ΔTM
- **pET-15bpIII-N1+**
  - pET vector with insertion of CTXφ pIII DNA encoding residues −5 to +134 at Ndel and BamH1 sites downstream of His-tag coding sequence; contains a mutation encoding a S65T substitution in pIII−ΔTM
- **pET-15bpToIa-C**
  - pET vector with insertion of *V. cholerae* ToIa DNA encoding residues 241–356 at Ndel and BamH1 sites downstream of His-tag coding sequence
- **pCTX-Km**
  - Replicative form of CTXφ-Km DNA derived from the chromosomal CTX element of *V. cholerae* strain P27459, with cTB replaced with a KmR cassette
- **Phase. fΔ1-pIIICTX(15–274)**
  - Gene fragment encoding pIII residues −5 to 269 (corresponds to old numbering 15–274, which assumed the signal peptide spanned the first 14 residues of the unprocessed pIII) cloned into phage display vector fd-DOG in place of gene fragment encoding fd pIII domains N1 and N2

### Primers

| Primers | Description |
|---------|-------------|
| pIII-N1-for | GGGAGATTCCCTACATTGACATAGGCTACGTCCTCGG |
| pIII-N1-rev | CCGATGCTCCTAGCATCAGCTCCCTCCCTCCAGG |
| pIII−ΔTM-for | GGTAGATCGTACAGGCTACGTCCTCGG |
| pIII−ΔTM-rev | CGGAGATTCCCTACATTGACATAGGCTACGTCCTCGG |
| pIII-N1+−for | GGGAGATTCCCTACATTGACATAGGCTACGTCCTCGG |
| pIII-N1+−rev | CGGAGATTCCCTACATTGACATAGGCTACGTCCTCGG |
| ToIα-C-for | CGTGGAGTCCTATCAGTGGCAGCTACGTCCTCGG |
| ToIα-C-rev | CGTGGAGTCCTATCAGTGGCAGCTACGTCCTCGG |
| T7 promoter | TAATAGCGATCTACATAGGG |
| T7 terminator | GCTGATGTGTCACGG |

### Immunogold and Transmission Electron Microscopy

were prepared by floating glow discharged carbon-coated Formvar copper grids on transfection solutions (described above) for 10 min and then fixing with 2% paraformaldehyde in phosphate-buffered saline (PBS), pH 7.4 for 5 min. The grids were then washed twice on drops of PBS containing 0.15% glycine, blocked for 5 min with 1% bovine serum albumin (BSA) in PBS, and then incubated with monoclonal rabbit anti-β-Tac antibodies (1:100 dilution) for 30 min. After 2 washes on PBS-glycine and 1 on PBS-BSA, the grids were incubated with 10- or 15-nm gold-conjugated protein A (diluted 1:70) for 20 min. After two washes on PBS, grids were fixed with 1% glutaraldehyde in PBS for 5 min, washed twice on PBS-glycine, and then blocked for 5 min with PBS-BSA. This was followed by an incubation step on anti-M13 mouse monoclonal antibodies (anti-g8p, Progen Biotechnik GmbH, 1:100 dilution) for 30 min. After 2 washes on PBS-glycine and 1 on PBS-BSA, the grids were incubated with anti-mouse-coupled gold particles (5 nm) for 20 min. After 3 rinses on PBS and 1 on water, the grids were stained for 2 min with 2% uranyl acetate and viewed with a JEOL 1200EX TEM.

Expression of CTXφ pIII Domains—The gene fragment encoding the CTXφ pIII N-terminal domain and part of the signal peptide (pIII-N1, residues −5 to +96) was PCR-amplified from *oriF1* in CTXφ-Km DNA using primers pIII-N1-for and pIII-N1-rev (Table 1). A gene fragment encoding the CTXφ pIII-soluble portion lacking the C-terminal putative transmembrane segment (pIII−ΔTM, residues 1–355) was PCR-amplified using primers pIII−ΔTM-for and pIII−ΔTM-rev (Table 1). In both PCR amplification of gene products, PCR products can be identified by their size on agarose gels. The amplified DNA was then purified and cloned into the plasmid vector pET-15b, which contains a T7 promoter, His-tag coding sequence, T7 terminator, lacI coding sequence, pBR322 origin, and bla (ApR) genes. After transformation of *E. coli* BL21 (DE3) with the recombinant plasmid, IPTG induction leads to expression of the inserted gene fragment as an insoluble protein that can be purified by metal-affinity chromatography.
products and the pET-15b expression vector (Novagen) were separately digested with Nde1 and BamHI (Fermentas), and the digestion products were ligated for 3 h at 20 °C. Insertion of genes between the Nde1 and BamHI sites results in N-terminal addition of a hexahistidine tag and linker with the sequence GSHHHHHHSGGLVPR ↓ GSHM, where the down arrow indicates a thrombin cleavage site. Ligation products were transformed into E. coli DH5α and Rosetta-gami B (DE3) cells at 42 °C using the CaCl2 heat shock method (38). Cells were incubated overnight at 37 °C on LB-Ap agar plates, and transforms were selected and grown overnight in LB-Ap broth. Plasmid was extracted from the DH5α strain and purified using the Qiagen Plasmid Purification kit. Plasmids were sequenced using the T7 promoter and terminator primers (Table 1) to verify that the insertions were correct. The pET-15b-pIII-ΔTM construct contains a mutation encoding a Ser65Thr substitution encoding pIII Asn-135 in pET-15b. The construct contains a mutation encoding a Ser65Thr substitution encoding pIII Asn-135 in pET-15b. The filtered supernatant was loaded onto a gravity column containing Ni-NTA beads (Qiagen) for affinity purification using the T7 promoter and terminator primers (Table 1) to elute the pIII-N1 and its binding partner, TolA-C or pIII-N1, was applied to the Ni-NTA beads and washed twice. The His-tagged proteins were eluted from the column with 1 ml of elution buffer (20 mM Bis-Tris, pH 6.5, 50 mM NaCl, 5 mM imidazole). Next, 1 mg of the His tag-cleaved binding partner, TolA-C or pIII-N1, was applied to the Ni-NTA beads and washed twice. The His-tagged proteins were eluted from the column with 1 ml of elution buffer (20 mM Bis-Tris, pH 6.5, 50 mM NaCl, 200 mM imidazole). Elution fractions were analyzed by SDS-PAGE. Purification of the pIII-N1-His-TolA-C Complex for Crystal Structure Determination—Purified His-TolA-C (30 mg) was loaded onto a Ni-NTA column (2 ml Ni-NTA beads), and the column was washed with 5 ml of purification buffer (20 mM Bis-Tris, pH 6.5, 50 mM NaCl). Next, ~27 mg of His tag-cleaved pIII-N1-+ was loaded onto the column and washed with 5 ml of purification buffer. Proteins were eluted from the column with elution buffer and the fractions were concentrated in the Amicon stirred cell concentrator with a 10,000 molecular mass cut off membrane and then loaded onto the Sephacryl size exclusion column. The fractions containing the complex were concentrated to 27.1 mg/ml and stored as described.

Crystal Structure of CTXφ pIII and pII-ToLA

Expression of Selenomethionine-substituted pIII-N1—Selenomethionine (SeMet)-substituted His-pIII-N1-+ was prepared by growing Rosetta-gami B (DE3) E. coli cells carrying pET-15b-pIII-N1-+ in M9 minimal media containing SeMet and select amino acids. Cells were first grown overnight in LB-Ap at 37 °C shaking. 10 ml of culture for every 1 liter of minimal media was centrifuged at 3500 × g for 20 min. The supernatant was discarded, and the cells were resuspended in 1 liter of M9 minimal medium and grown at 37 °C to an A600 of ~0.4, at which point 60 mg/liter L-SeMet, 100 mg/liter lysine, threonine, and phenylalanine, and 50 mg/liter leucine, isoleucine, and valine were added. The culture was incubated for 15 min at 37 °C with shaking before isopropyl β-D-thiogalactopyranoside was added to 1 mM to induce protein expression. Cells were grown overnight at 19 °C with shaking. Cells were pelleted by centrifugation, resuspended in lysis buffer, and lysed by sonication, and protein was purified and stored as described above.

Expression of the V. cholerae TolA C-terminal Domain—A gene fragment encoding residues 241–356 of TolA (TolA-C) was PCR-amplified from the genomic DNA of V. cholerae O395 using primers TolA-C-for and TolA-C-rev (Table 1). Additional constructs were designed by inserting stop codons at other sites in pIII-ΔTM. These include pIII-N1N2 (residues 5 to 233), pIII-N1NC (residues 5 to 330), and pIII-N1a (1 to 96). Of these constructs, only pIII-N1, pIII-N1+, and pIII-N1a produced soluble protein.

Expression and Purification of pIII and TolA Proteins—Bacterial cells transformed with the appropriate plasmid were grown overnight in 200 ml of LB broth containing 100 µg/ml Ap at 37 °C shaking. 10 ml aliquots of overnight culture were added to 1 liter of LB-Ap and grown at 37 °C to an optical density at 600 nm (A600) of ~0.6. Protein expression was induced by the addition of isopropyl β-D-thiogalactopyranoside to 1 mM and cells were grown overnight at 19 °C with shaking. Cells were pelleted by centrifugation at 4000 × g for 30 min, the supernatant was discarded, and the pellet was resuspended in lysis buffer (Bis-Tris pH 6.5, 100 mM NaCl, 10% glycerol, 0.1% Tween, 5 mM imidazole) with 10 µg/ml lysozyme and a Complete Protease Inhibitor Mixture tablet (EDTA-free, Roche Applied Science). Cells were incubated in lysis buffer at room temperature for 1 h and then lysed by sonication. Cellular debris was removed by centrifugation at 40,000 × g for 40 min, and the supernatant was filtered through a 0.45-µm membrane. The filtered supernatant was loaded onto a gravity column containing Ni-NTA beads (Qiagen) for affinity purification using the N-terminal His tag. The column was washed using 3 column volumes of wash buffer (20 mM Bis-Tris, pH 6.5, 100 mM NaCl), and protein was eluted with 200 mM imidazole. Elution fractions containing the protein of interest were concentrated using a stirred cell concentrator (Amicon) with a 3000 Da molecular mass cut off membrane and loaded onto a size exclusion column (HiPrep 26/60 Sephacryl S-100 HR, GE Health-
Crystal Structures of CTXφ pIII and pIII-TolA

**TABLE 2**

Crystallographic data collection and refinement statistics

| Data collection | SeMet-His-pIII-N1+ | SeMet-His-pIII-N1+ | pIII-D1:His-TolA-C |
|----------------|-------------------|-------------------|--------------------|
| Beamline       | SSRL 9-2          | SSRL 9-2          | SSRL 9-2           |
| Space group    | P6,22             | P6,22             | P2,1/2             |
| Cell dimensions|                  |                   |                    |
| a, b, c (Å)    | 126.33, 126.33, 127.33 | 126.25, 125.25, 128.27 | 43.38, 46.16, 101.63 |
| α, β, γ (°)    | 90, 90, 120       | 90, 90, 120       | 90, 90, 90         |
| Resolution (Å) | 3.8               | 2.9               | 1.44               |
| Wavelength (Å) | 0.918             | 0.979             | 0.98               |
| Completeness (%) | 99.9/99.7       | 99.9/99.9         | 99.6/99.6 (95.5)   |
| Observed reflections | 210,923        | 166,181           | 584,826            |
| Unique reflections | 20,639         | 16,181            | 25,290             |
| Rcryst (%)     | 13.6/28.1         | 13.1/27.6         | 11.8/66.9          |
| Rfree (%)      | 5.4/2.7           | 5.6/2.7           | 24.1/2.1           |
| Mosaicity (°)  | 0.9               | 0.9               | 0.8                |

| Refinement statistics | SeMet-His-pIII-N1+ | SeMet-His-pIII-N1+ | pIII-D1:His-TolA-C |
|-----------------------|-------------------|-------------------|--------------------|
| Resolution limits (Å) |                   |                   |                    |
| Molecules/AU          | 20.0-2.9          | 3                 |                   |
| Rcryst (%)            | 23.4              | 26.2              |                   |
| Rfree (%)             | 99.9/99.7         | 99.9/99.9         | 99.6 (95.5)        |
| Number of reflections used for refinement | 13171 | 35426 |                   |
| Number of atoms       | 2240              | 1728              |                   |
| Protein               |                   |                   |                    |
| Ligands               |                   |                   |                    |
| Water                 | 9                 | 197               |                   |
| B-factor (Å²)         |                   |                   |                    |
| Average               |                   |                   |                    |
| Protein               |                   |                   |                    |
| Chain A               | 83.5              | 12.9              |                   |
| Chain B               | 77.6              | 15.1              |                   |
| Chain C               | 74.4              |                   |                    |
| Water                 | 54.4              | 23.0              |                   |
| Root mean square deviation |               |                   |                    |
| Bond lengths (Å)      | 0.017             | 0.005             | 0.074              |
| Bond angles (°)       | 1.85              | 0.974             |                   |
| Ramachandran plot statistics (%) |         |                   |                    |
| Favored               | 90.3              | 97.4              |                   |
| Allowed               | 9.0               | 2.6               |                   |
| Generously allowed    | 0.7               |                   |                    |
| PDB accession no.     | 4G7W              | 4G7X              |                   |

---

* Overall last shell.

* Rcryst is the unweighted R value on I against symmetry mates.

* Rcryst = ∑|Fobs(hkl) − |Fcals(hkl)|/|Fobs(hkl)|.

* Rfree is the cross-validation R factor for 5% of reflections against which the model was not refined.

---

Molecular Research) by mixing 2 μl of protein with 1.6 μl of concentration-adjusted reservoir buffer and 0.4 μl of additive or detergent. SeMet-pIII-N1+ crystals grew from a 23.5 mg/ml protein solution in 20% PEG 3350, 8% Tacsimate, pH 5.0, and Foscholine-9. pIII-N1+His-TolA-C crystals grew from a 27.1 mg/ml solution in 25% PEG 6000, 100 mM MES, pH 6.0. Crystals appeared after 1–2 weeks and were flash-frozen in liquid nitrogen. Glycerol (25%) was included as a cryoprotectant for pIII-N1+.

X-ray Diffraction Data Collection, Structure Determination, and Refinement—X-ray diffraction data for SeMet-substituted pIII-N1+ and pIII-N1+His-TolA-C crystals were collected remotely on the Stanford Synchrotron Lightsource beamline 9-2 at 100 K using Blu-lce software (40). The initial test diffraction images were indexed, and x-ray fluorescence scans were performed on SeMet-substituted pIII-N1+ to select the appropriate wavelengths for multiple wavelength anomalous dispersion data collection. For pIII-N1+, data sets were collected at wavelengths corresponding to the inflection point and high energy remote from a crystal that diffracted to ~4 Å resolution, and a second data set was collected at the peak wavelength for a crystal diffracting to 2.9 Å resolution (Table 2). A native 1.44 Å resolution data set was collected for the pIII-N1+His-TolA-C complex. The dual wavelength pIII-N1+ data set was processed and scaled using iMosflm and SCALA (41, 42). The single wavelength pIII-N1+ data set and the native pIII-N1+His-TolA-C data set were processed and scaled using the XDS suite (43) to resolutions 2.9 and 1.44 Å, respectively. Matthews coefficients (44) calculated using CCP4 (41) indicated three molecules per asymmetric unit for pIII-N1+ and a single heterodimer for pIII-N1+His-TolA-C. Initial phases for pIII-N1+ were determined by multiple wavelength anomalous dispersion phasing method using SOLVE (45), and density modification was performed with RESOLVE (46) yielding an interpretable electron density map. Backbone atoms were placed and model building was carried out in COOT (47). A few cycles of rigid body refinement followed by tight NCS-restrained refinement by REFMAC (48) brought the Rcryst and Rfree values to 0.417 and 0.447, respectively. Iterative cycles of tight main chain and loose side chain restrained refinement and model building improved the model and statistics. After several cycles of TLS refinement (48), water oxygens were located using COOT and included in the refinement. The final Rcryst and Rfree are 23.4 and 26.2%, respectively. The crystal data and refinement statistics are summarized in Table 2.

The pIII-N1+His-TolA-C structure was solved by the molecular replacement method. First, structure factors were determined for pIII-N1+ using PHASER (49) with our mono-
mERIC pIII-N1 + structure as a search model. Once the orienta-
tion and position of pIII-N1 were determined, structure factors
were calculated for *V. cholerae* TolA-C within the complex
using the *E. coli* TolA-C structure in 1TOL as a model. The two
TolA-C domains have 27% amino acid sequence identity. A few
cycles of the “map improvements by atoms update and refine-
ment” option of ARP/wARP (50) made a dramatic improve-
ment to the map with clear density for main chain and side
chain atoms for both proteins, with the exception of their
N-terminal segments. The model was built manually in COOT,
and several cycles of restrained refinement were carried out.
Water oxygens were located in COOT, and after checking them
manually, WATERTIDY in CCP4 (41) was used to move the
water oxygen coordinates to the symmetry-related position
nearest to the host protein molecule. Additional cycles of
refinement and fitting brought the Rcryst to 20.6% and the Rfree
to 21.6%, with no subsequent improvement. Both pIII-N1+
and pIII-N1+His-TolA-C structures were validated using
PROCHECK and MolProbity (51, 52).

**RESULTS AND DISCUSSION**

**TolA Is Not Required for CTXϕ Binding to V. cholerae but Is
Required for Uptake**—To investigate the role of *V. cholerae*
TolA in CTXϕ infection, phage were incubated with *V. chol-
eraea* strain O395 and with an O395-derived ΔtolA deletion
mutant, DH3 (11), and visualized by immunogold labeling and
TEM. *V. cholerae* produce relatively low levels of CTXϕ that
are difficult to detect by TEM. Thus, we used the engineered fd
phage fdΔ1-pIII-CCTX(15–274), which expresses CTXϕ pIII
domains N1 and N2 in place of the native fd N1N2, causing a
switch in specificity from *E. coli* to *V. cholerae* (11). Plasmid-
expressed fdΔ1-pIII-CCTX(15–274) are produced in high levels
from *E. coli* DH5α-Apir (53). fdΔ1-pIII-CCTX(15–274) phage
were incubated with *V. cholerae* O395 and DH3 cells and applied
to TEM grids. TCP were immunolabeled with 10- or 15-nm gold
particles, and phage were labeled with 5-nm gold particles.
Although the phage and TCP are both long thin filaments of
comparable diameters (~8 nm), they are readily distinguish-
able by TEM, as the pili are several microns in length and bun-
dled, whereas the phage are typically shorter (200–400 nm)
single filaments (Fig. 1). Phage were observed in contact with
TCP in both *V. cholerae* strains. In many cases the phage con-
tact the pili via their tips (Fig. 1, A and B), but lateral interac-
tions were also observed. In some cases the phage appear to
interact directly with the *V. cholerae* surface (Fig. 1D). There
were no apparent differences in the phage interactions with
*V. cholerae* O395 (Fig. 1, A–C) and the tolA deletion strain DH3
(Fig. 1, D and E). These findings are consistent with CTXϕ
initially binding to *V. cholerae* via an interaction between pIII
and TCP in a TolA-independent manner. However, TolA is
required for the subsequent phage uptake step, as transduction
frequencies of fdΔ1-pIII-CCTX(15–274) and CTXϕ are negligible
for the TolA-deficient strain (Fig. 1F and Ref. 11).

**Interaction between CTXϕ pIII-N1 and the C-terminal
Domain of V. cholerae TolA**—Because the N-terminal domain
of coliphage pIII binds to the C-terminal domain of *E. coli* TolA
(18, 32, 33, 35, 36), we sought to demonstrate a direct interac-
tion between CTXϕ pIII-N1 and *V. cholerae* TolA-C. Recom-
binant pIII-N1 was prepared comprising residues 1–96 of the
mature CTXϕ pIII protein with an N-terminal hexahistidine
(His) tag for purification using metal affinity chromatography
(Fig. 2A). Five residues of the signal peptide are also present
in this construct, which was designed based on an earlier predic-
tion of a signal peptidase cleavage site between the 14th and
15th residue of the pre-protein. However, current bioinformat-
ics tools predict the cleavage site to lie between residues 19 and
20 (Fig. 2A).

The C-terminal domain of *V. cholerae* TolA was delineated
based on amino acid sequence comparison with *E. coli* TolA
and secondary structure prediction using PsiPred (67). *E. coli*
TolA is a 421-residue periplasmic protein with an N-terminal
inner membrane anchor, a central α-helical segment with a
LyS1–2-Ala1–4-(Glu/Asp) repeat sequence, and a C-terminal
domain that binds to coliphage pIII-N1 as well as to colicins (18,
33). *V. cholerae* TolA is a relatively uncharacterized 356-amino
acid protein with a putative N-terminal transmembrane seg-
ment, a central Lys/Ala/Glu-rich region predicted to be an
extended α-helix, and a C-terminal domain (Fig. 2A). *V. chol-
eraea* and *E. coli* TolA share 35% sequence identity in their
N-terminal ~150 amino acids and 27% identity in their C-ter-
minal domains. The C-terminal domain of TolA (residues
241–356) was expressed and purified from *E. coli* with an
N-terminal His tag (His-TolA-C) and mixed with pIII-N1, which
had its His tag removed by thrombin digestion. The pIII-
N1-TolA-C complex was isolated by incubating the sample with
Ni-NTA beads (Qiagen), washing the beads to remove
unbound protein, and eluting bound protein using imidazole.
Samples were analyzed by SDS-PAGE (Fig. 2B). pIII-N1 bound
to the Ni-NTA-coated beads when His-TolA-C was present,
but not in its absence, suggesting that these domains interact
with each other. Similar results were obtained for reciprocal
experiments in which His-pIII-N1 was mixed with non-tagged
TolA-C (data not shown).

To further establish an interaction between pIII-N1 and
TolA-C, the proteins were loaded onto a size exclusion chro-
matography column individually and after co-incubation. His-
TolA-C, which has a calculated molecular mass of 14.8 kDa,
elutes at ~15 kDa, and pIII-N1, with a calculated mass of 11.5,
elutes at ~12 kDa (Fig. 2C). When His-TolA-C is preincu-
bated with pIII-N1, these peaks disappear, and a new large peak
elutes from the column at a mass of ~28 kDa, corresponding to
the pIII-N1:His-TolA-C complex, with a calculated mass of
26.3 kDa. The ability of pIII-N1 to remain bound to TolA-C
through size exclusion chromatography demonstrates the sta-
BILITY of this interaction and indicated that this complex was
a good candidate for x-ray structure determination.

**Crystal Structure of CTXϕ pIII-N1**—We first solved the crys-
tal structure of pIII-N1 alone. CTXϕ pIII constructs of varying
lengths were designed for expression as N-terminally His-
tagged proteins to identify crystallizable fragments or domains.
*E. coli* expression strains, growth, and induction conditions
were optimized for each construct. In addition to His-pIII-N1,
described above, two other constructs produced soluble protein
from *E. coli* Rosetta-gami B (DE3); pIII-N1 + encoding resi-
dues ~5 to 134, which includes LCR1 plus 19 residues of N2
(Figs. 2A and 3A), and pIII-N1a, encoding residues 1–96. Pro-
proteins were purified, crystallization conditions were screened, and diffraction-quality crystals were obtained for His-pIII-N1 belonging to the P6322 space group. Because pIII is not homologous in sequence to any protein of known structure, molecular replacement was not feasible for determining its structure. Thus, SeMet-substituted His-pIII-N1 was prepared for structure determination using multiple wavelength anomalous dispersion methods. SeMet pIII-N1 crystals were obtained in conditions similar to those of the native crystals, and a 2.9 Å resolution structure was solved by multiple wavelength anomalous dispersion phasing (Table 2, supplemental Fig. S1A).

There are three molecules in the asymmetric unit of the pIII-N1 crystal, all with close structural similarity (supplemental Fig. S2A). Residues 1–98 were resolved for chains A and C and 1–99 for chain B. No electron density was apparent for the N-terminal His tag or the five signal peptide residues at positions −5 to −1. Additionally, most of LCR1 (residues 100–115) and the 19 residues of N2 are not resolved in the map. Thus, we refer to this structure from here on in as pIII-N1. The core of pIII-N1 is an incomplete β-barrel with side protrusions and irregular loops at both the N- and C-terminal ends (Fig. 3, B and C). The U-shaped N-terminal loop that lies at the bottom of the β-barrel is followed by its first β-strand, β1 (residues 13–17). From β1 the polypeptide chain crosses the top of the β-barrel to form β2 (26–29) on its opposite side. A type II turn connects β2 with β3 (32–50), a long strand that is interrupted by a β-hairpin (37–44) that protrudes from the β-barrel. β3a and β3b curve around the β-barrel at almost 90° to each other. β3b is connected to another long strand, β4 (53–62), by a type I hairpin turn. The β3β4 hairpin extends away from the top of the β-barrel. Following β4, the polypeptide chain exits the β-barrel and
wraps around its flattened side formed by strands β2, β3a and β5. This loop then runs across the top of the β-barrel and folds back into it as its terminal strand, β5 (84–86). From β5 the chain forms a hairpin loop that extends away from the β-barrel on the side opposite the β3 hairpin. The β-barrel is open, as strands β1 and β5 do not contact each other. Thus, pIII-N1 is composed of an open β-barrel with several protruding loops: the N-terminal loop at the bottom of the barrel, the β3β4-hairpin protruding from its top, and the β3-hairpin and the C-terminal loop sticking out from opposite sides. The protrusions give this domain an open, splayed appearance. The long side of the β-barrel with the β3β4 hairpin forms a concave surface that resembles an open left palm, with the fingers formed by antiparallel strands β3b, β4, β1, and the C-terminal loop and the thumb, represented by the β3 hairpin (Fig. 3B). Four disulfide bonds stabilize CTX pIII-N1: between Cys-5 on the N-terminal loop and Cys-32 at the N terminus of β3a; between Cys-47 and Cys-56, joining strands β3b and β4; between Cys-75 at the top of the β-barrel on the meandering β4β5 loop and Cys-85 on the short β5 strand; between Cys-90 and Cys-96 in the C-terminal loop. The conformations of the N- and C-terminal loops have small variations among the three molecules in the asymmetric unit (supplemental Fig. S2A). Additionally, the extended β3β4 hairpin bends away from the β-barrel in chain C (Fig. 3, B and C) but is straighter in chains A and B, with a difference of ~7 Å between the Glu-51 Cαs at the hairpin tip (supplemental Fig. S2A). This difference suggests flexibility in this region, which is constrained by crystal contacts in chains A and B.

The crystal structure of CTXφ pIII-N1 was compared with its corresponding domain in pIII of fd, M13, and IF1 phage. fd and M13 pIII proteins have 99% amino acid sequence identity, and their crystal structures (domains N1 and N2 or N1N2) are essentially identical (22, 23). Molecule B of fd pIII-N1N2 (PDB code 2G3P) is shown in Fig. 3E. Although CTXφ and the coliphage pIII proteins have similar domain boundaries, size, and functions, a pairwise BLAST search (blast.ncbi.nlm.nih.gov) revealed no significant sequence identity. In comparing the CTXφ pIII-N1 structure with that of fd, structural similarities are also not immediately apparent. However, close inspection reveals similar topologies for the two domains (Fig. 3, B–G), and sequence homology is evident when the CTXφ and fd N1 amino acid sequences are aligned based on their secondary structures (Fig. 3A). Like CTXφ, fd pIII-N1 has a core β-barrel composed of...
Crystal Structures of CTXφ pIII and pIII-TolA

A

CTX MRFFLLFLLFLLSFVTA -1 (signal peptide)
fda MK-KLLAFLPLVVFYFHSH -1 (signal peptide)

CTX -SAINNUPNPMTTLLELFGPSIVQ---SVLFDCQMLDIEKDGYGFIVLSGSNEDYTVKLXPRFSQGVSPWPNSLDGASAASGKTYVCPEGRTS 96
fda AEVESQALRKLKLTNFIVWDDKTLDRYANTGVDLNATG---VYVSGPDEI GQYTVPIGLAIPE 67

LCR1

CTX -SAINNUPNPMTTLLELFGPSIVQ---SVLFDCQMLDIEKDGYGFIVLSGSNEDYTVKLXPRFSQGVSPWPNSLDGASAASGKTYVCPEGRTS 96
fda AEVESQALRKLKLTNFIVWDDKTLDRYANTGVDLNATG---VYVSGPDEI GQYTVPIGLAIPE 67

LCR2

CTX -SAINNUPNPMTTLLELFGPSIVQ---SVLFDCQMLDIEKDGYGFIVLSGSNEDYTVKLXPRFSQGVSPWPNSLDGASAASGKTYVCPEGRTS 96
fda AEVESQALRKLKLTNFIVWDDKTLDRYANTGVDLNATG---VYVSGPDEI GQYTVPIGLAIPE 67

B

\[ \beta_3\beta_4 \text{ hairpin} \]

C

\[ \beta_3 \text{ hairpin} \]

D

\[ \beta_3\beta_4 \text{ loop} \]

E

\[ \beta_3 \beta_4 \text{ hairpin} \]

F

\[ \beta_3\beta_4 \text{ hairpin} \]

G

\[ \beta_5 \text{ to N2} \]

\[ \beta_3 \text{ from N2} \]
strands β1-β4 with two similarly-placed disulfide bonds (Fig. 3, A, C, and F), but the β-barrel is more compact and lacks a fifth β-strand (for ease of comparison, we have not labeled a short β-strand between β1 and β2 of fd pIII-N1, which is absent in CTXφ N1). fd pIII-N1 has a short N-terminal α-helix not present in CTXφ pIII-N1, and its β3 and β3’4 hairpins are shorter. The fifth strand of the CTXφ pIII-N1 β-barrel is located on a 26-residue insertion between β4 and LCR1 that is not present in fd pIII (Fig. 3A). This insertion in CTXφ pIII-N1 follows β4 as a meandering loop around the outside and across the top of the β-barrel, then feeds into β5, which is followed by a disulfide-bonded C-terminal loop. The meandering loop is disulfide-bonded to β5. The LCR1 segment that follows is disordered in both the CTXφ and coliphage pIII proteins. In the coliphage pIII the LCR1 is referred to as the N1N2 linker (22, 23), yet it appears to be an integral if disordered part of N1 distinct from the linker connecting the two domains (Fig. 3). It is not clear whether CTXφ pIII-N1 possesses a two-stranded β-sheet corresponding to β5 and β13 in fd pIII, but the structural similarities between the N1 domains suggest that their N2 domains have similar orientations.

**Crystal Structure of CTXφ pIII-N1 in Complex with V. cholerae TolA-C**

CTXφ TolA-C. A crystal belonging to the p212121 space group diffracted to 1.44 Å resolution, with one molecule in the asymmetric unit and a solvent content of 27.5% (Table 2). This insertion in CTXφ pIII-N1, and its C-terminal loop “fingers” (Fig. 3B) clamp onto TolA-C, burying a total of 1489 Å² in surface area (Fig. 4). Two acidic side chains, Glu-37 and Asp-39, on the pIII β3 hairpin, form key salt bridges with adjacent lysines (Lys-324 and Arg-325, respectively) on α2 of TolA-C (Fig. 4B). The Arg-325 side chain on α2 of TolA inserts into the β3 hairpin and makes polar contacts with several residues, and the hydrocarbon portions of both Lys-324 and Arg-325 side chains form van der Waals interactions with Val-44 and Leu-59 of pIII-N1. Tyr-41 on the β3 hairpin makes hydrophobic contacts with Leu-282 on the 3₁₀ helix loop of TolA. Across the pIII-N1 β-barrel the β3β4 hairpin hooks around TolA (Fig. 4A), forming mainly polar interactions with the extended loop on TolA that follows α2. Thus, the curvature of the β3β4 hairpin, observed in this complex and in chain C of the unbound pIII-N1 structure (supplemental Fig. S1), provides stereochemical complementarity for TolA-C binding. Main chain hydrogen bonds between β1 of pIII-N1 and β2 of TolA-C result in a continuous anti-parallel β-sheet comprising strands β1, β4, and β3β3 from pIII and β2, β1, and β3 of TolA-C (Fig. 4, A and C). The C-terminal loop of pIII-N1 also contacts TolA via a salt bridge between Glu-92 and Lys-347 on the short loop connecting α3 and β3 of TolA-C (Fig. 4C).

![Crystal structure of CTXφ pIII-N1 at 2.9 Å resolution and comparison with fd phage pIII-N1N2.](image-url) A, shown is structure-based amino acid sequence alignment between CTXφ (NCBI locus AF220606_5) and fd pIII-N1 (G3P_BPFD). Similar residues are highlighted in yellow, and identical residues are highlighted in orange. Secondary structures are shown above the sequences. The background of the segments resolved in the crystal structures is colored; pale green for CTXφ pIII-N1, pale cyan for fd pIII-N1, and pale yellow for fd pIII-N2. The low complexity regions are delineated by dashed boxes. Cysteines are boxed in blue, and connections indicate disulfide bonds. The arrowhead points to Pro-213 in fd pIII-N1. The crystal structure of the N-terminal domain of CTXφ pIII (pIII-N1, 2.9 Å resolution) shown from the top of the β-barrel (B) and rotated ~160° about a horizontal axis to show the concave side of the β-barrel (C), which resembles an open hand. The β3β4 hairpin fingers protrude toward the reader. Cysteines and secondary structural elements are labeled, and cysteines are shown in ball-and-stick representation with sulfur atoms in yellow. D, secondary structure representation of CTXφ pIII-N1 is shown. Disulfide bonds are shown as black bars. E, the crystal structure of fd phage pIII-N1N2 (2GP3) with N1 in pale cyan and N2 in yellow is shown in the same orientation as CTXφ-pIII-N1 in C. The gray dashed line represents LCR1, which is unresolved in the crystal structure. F, fd pIII-N1 residues 2–67 from 2GP3 are shown. G, shown is a secondary structure representation of fd pIII-N1. β-Strands in fd pIII are numbered according to the corresponding strands in CTXφ-pIII-N1. Thus, a short strand between β1 and β2 in fd pIII-N1 is not included.

**V. cholerae TolA-C.** which is well resolved over residues 254–355, is an elongated domain with mixed α-helix/β-sheet structure (Fig. 4A). The N-terminal segment of TolA-C is a straight α-helix stalk (α1, residues 256–280) that associates with a three-stranded anti-parallel β-sheet flanked by two α-helices. From α1 the polypeptide chain forms a broad loop with a central 3₁₀ helix (285–287) that folds back into the protein core as an extended β-hairpin (β1, 292–301; β2, 304–315), with β1 being the central strand in the β-sheet. Between β2 and β3 lies a 4-turn α-helix (α2, 317–329), an extended loop, and a 2-turn α-helix (α3, 340–346). α2 lies along the top of the domain, parallel to the β-sheet strands and sandwiched between β2 and the 3₁₀-helix loop, interacting with these regions via hydrophobic side chains. There is a disulfide bond between Cys-292 at the N terminus of β1 and Cys-320 near the beginning of α2. α3, on the other side of the β-sheet, leads into the C-terminal strand, β3 (348–353). The β-sheet and shorter helices (α2, α3, and the 3₁₀-helix) together wrap around the C-terminal half of the α1 stalk. Overall, TolA-C is an elongated, curved structure with the crook between the β-sheet and the α1 stalk forming a concave cave face lined by α1, α3, and β3 and the convex face formed by α2, β1, and β2.

It is the convex face of TolA, mainly α2 and β2, that interacts with pIII-N1. pIII-N1 uses its open “hand” to grasp TolA. β-barrels strains β3β4β1, which form the “palm,” together with the β3 hairpin “thumb,” the protruding β3β4 hairpin, and the C-terminal loop “fingers” (Fig. 3B) clamp onto TolA-C, burying a total of 1489 Å² in surface area (Fig. 4). Two acidic side chains, Glu-37 and Asp-39, on the pIII β3 hairpin, form key salt bridges with adjacent lysine (Lys-324 and Arg-325, respectively) on α2 of TolA-C (Fig. 4B). The Arg-325 side chain on α2 of TolA inserts into the β3 hairpin and makes polar contacts with several residues, and the hydrocarbon portions of both Lys-324 and Arg-325 side chains form van der Waals interactions with Val-44 and Leu-59 of pIII-N1. Tyr-41 on the β3 hairpin makes hydrophobic contacts with Leu-282 on the 3₁₀ helix loop of TolA. Across the pIII-N1 β-barrel the β3β4 hairpin hooks around TolA (Fig. 4A), forming mainly polar interactions with the extended loop on TolA that follows α2. Thus, the curvature of the β3β4 hairpin, observed in this complex and in chain C of the unbound pIII-N1 structure (supplemental Fig. S1), provides stereochemical complementarity for TolA-C binding. Main chain hydrogen bonds between β1 of pIII-N1 and β2 of TolA-C result in a continuous anti-parallel β-sheet comprising strands β1, β4, and β3β3 from pIII and β2, β1, and β3 of TolA-C (Fig. 4, A and C). The C-terminal loop of pIII-N1 also contacts TolA via a salt bridge between Glu-92 and Lys-347 on the short loop connecting α3 and β3 of TolA-C (Fig. 4C).
Comparison of the CTXφ pIII-N1-V. cholerae TolA-C Complex with that of M13 pIII-N1-E. coli TolA-C—The crystal structure of the complex between CTXφ pIII-N1 and V. cholerae TolA-C was compared with that of a complex between M13 pIII-N1 and E. coli TolA-C (1TOL). The C-terminal domains of V. cholerae and E. coli TolA have 27% sequence identity (Fig. 4E), and their structures are very similar (1.81 Å root mean square deviation between Cα carbons for V. cholerae TolA-C residues 263–356 and E. coli TolA-C residues 334–421) (Fig. 4, A and D). The two TolA-C domains differ primarily in the loop regions; the conformations of the 3_10-helix loops differ somewhat, the β1β2 hairpin of V. cholerae TolA-C is longer and more ordered than that of E. coli TolA-C, and the α2-α3 loop of V. cholerae TolA-C is longer by two residues. As described above, N1 of the CTXφ and the Ff phage pIII proteins differ significantly in their β-sheet loops and in their N and C termini but have comparable topologies. Given the similarities between CTXφ and M13 pIII-N1 and between V. cholerae TolA-C, it might be expected that the pIII-N1-TolA-C interactions are similar in both complexes, but this is not the case. Whereas β-strands of both pIII-N1 proteins form edge-on contacts with β-strands in TolA-C to form a continuous β-sheet, the pIII β-strands and the face of the pIII-N1 β-barrel involved in this interaction differ in the two complexes. Further
thermore, CTXφ and M13 pIII-N1 bind to opposite sides of TolA-C (Fig. 4, A and D). Whereas CTXφ pIII-N1 binds to the convex face of TolA-C via its concave face, interacting with TolA-C all along the length of its β-barrel, M13 pIII-N1 binds to the concave face of TolA-C via the convex surface of its β-barrel composed of the β3β4 hairpin and the C-terminal half of β1. The N-terminal end of M13 pIII-N1 β3 hydrogen-bonds with \(E. coli\) TolA-C β3 to form a continuous 6-stranded β-sheet between the two proteins, comprising strands β3, β4, and β1 of M13 pIII-N1 and β3, β1, and β2 of TolA-C. The total buried surface area for the M13 pIII-N1 complex is 1768 Å², which is larger than that of the CTXφ pIII-N1-TolA-C complex (1489 Å²), yet both interfaces are highly complementary.

An important distinction between the two complex structures is that pIII-N1 and TolA-C in the M13 \(E. coli\) system were synthesized as a single fusion protein with an 18-residue linker. This unnatural covalent linkage could potentially force an interaction between these proteins that is not biologically relevant. However, the crystal structure of a complex between phage IF1 pIII-N1 and \(E. coli\) TolA-C, obtained from proteins that were expressed separately (54), is highly similar to that of the M13 pIII-N1-TolA-C complex. Thus, it appears that CTXφ pIII and the coliphage indeed bind to distinct sites on TolA.

The function of TolA in \(V. cholerae\) and \(E. coli\) is not known with certainty. TolA forms an inner membrane complex with TolQ, TolR, and TolB, encoded by the tolQRAB gene cluster conserved in Gram-negative bacteria (55) and is involved in outer membrane integrity, energy transduction, and cell division (56–58). \(E. coli\) and \(V. cholerae\) tolQRAB mutants are viable but are sensitive to detergents and antibiotics, suggesting a role for these proteins in maintaining outer membrane integrity (11, 59, 60). \(E. coli\) tolQRAB mutants are resistant to filamentous phage infection as well as to colicins, bacteriocidal molecules that are produced by some \(E. coli\) strains and act upon other \(E. coli\) strains (61, 62). Like the coliphage, colicins utilize a two-step mechanism for uptake, first binding to outer membrane receptors on \(E. coli\), then interacting with TolA upon transfer into the periplasm. Very recently a crystal structure of \(E. coli\) TolA-C was solved in complex with a polypeptide from the N-terminal domain of colicin A, TA53–107 (PDB code 3QDR (63)). Interestingly, TA53–107 binds to the side of TolA opposite to that used by coliphage pIII-N1, which corresponds to the CTXφ pIII-N1 binding site on \(V. cholerae\) TolA (Fig. 5). The main interacting segment in TA53–107 is β-strand β5, which associates with TolA-C primarily via edge-on backbone hydrogen bonds. Antiparallel strands β4 and β5 of TA53–107 form a continuous β-sheet with the three-stranded TolA-C β-sheet (Fig. 5A) and correspond to β1 and β4 in CTXφ pIII-N1 (Fig. 5B). TA53–107 β5 also makes van der Waals contacts with TolA α2. The total buried surface area for this interaction (1303 Å²) is similar to that of the pIII-TolA interaction (1489 Å²). Several of the TolA residues involved in these interactions are in equivalent positions, but they are not conserved between \(E. coli\) and \(V. cholerae\), and thus, colicin A is unlikely to bind to \(V. cholerae\) TolA.

**Accessibility of the TolA Binding Site on Native CTXφ pIII—** Because the TolA binding site on CTXφ pIII-N1 lies outside of the region that interacts with domain N2 in the fd and M13 coliphage, we wondered if it might be accessible to TolA binding in its native form. In Fig. 6A the CTXφ pIII-N1-TolA-C complex and the fd pIII-N1N2 structure are presented with both N1 domains in approximately the same orientation. This shows that if N2 is located in the same position relative to N1 in CTXφ pIII as it is in fd pIII, it would not block TolA binding. To test the accessibility of the TolA binding site on CTXφ pIII-N1, phage were incubated with an excess of TolA-C before infection of \(V. cholerae\) O395 cells to see if TolA-C binding reduced infection levels. The transduction efficiency was reduced 3-fold for CTXφ and 5-fold for fdΔ1-pIII\(\text{CTX} (15–274)\) when pre-treated with TolA-C (Fig. 6B), suggesting that TolA-C can bind to pIII-N1 in its native conformation. Thus, although TCP binding is important for bringing CTXφ to the \(V. cholerae\) surface, it may not be required to unfold pIII and expose the TolA binding site. This relationship is similar to that seen for the IF1 and IKe coliphage, where the N1 and N2 domains are not tightly apposed but instead are arranged like beads-on-a-string, and thus their TolA binding sites are accessible in the native pIII (54, 64). The TolA binding domains of CTXφ, M13, fd, IF1, and IKe phage are all structurally similar regardless of whether their respective pilus binding domains are tightly associated or not. It is possible that CTXφ pIII also adopts a beads-on-a-string arrangement for N1 and N2 rather than the close packing seen in fd and M13 phage. We were unable to express CTXφ N1N2 in a soluble form to examine this possibility. However, the unique location of the TolA binding site on CTXφ pIII-N1 and the ability of free

![Crystal Structures of CTXφ pIII and pIII-ToLA](image-url)
TolA-C to reduce phage infectivity together suggest that the TolA binding site is accessible in the native protein.

Despite differences between CTX/H9278 and the coliphage with respect to their mode of binding to TolA, strong parallels between the two systems support a common two-step infection process whereby phage bind first to the bacterial pili via pIII-N2 and pilus retraction draws the phage tip into the periplasm and presents N1 to TolA (Fig. 7). Unlike the E. coli F pilus, TCP has not been shown to retract and lacks a “retraction” ATPase present in retractile Type IV pili. However, the TCP role in TcpF secretion presumably entails a piston-like motion involving cycles of pilus extension and retraction to extrude TcpF across the outer membrane. Such a mechanism has been proposed for the closely related type II secretion system (65), which also lacks a retraction ATPase. TCP retraction may occur randomly and spontaneously in the absence of a retraction ATPase (Fig. 7A), with CTXφ being drawn in serendipitously to the benefit of both the phage and V. cholerae but to the detriment of the human host.

Conclusions—We demonstrate here direct interactions between CTXφ and V. cholerae TCP and between pIII-N1 and TolA and present x-ray crystal structures of pIII-N1 alone and in complex with TolA-C. Despite the structural similarities between the CTXφ and coliphage pIII-N1 domains and between V. cholerae and E. coli TolA-C, the molecular interfaces are entirely different between the two pIII-N1-TolA-C complexes. The unique interaction between CTXφ pIII-N1 and V. cholerae TolA-C may abrogate the need for a pilus-induced conformational change in pIII to prime it for binding to TolA. CTXφ infection may involve spontaneous retraction of TCP to transport the phage tip through the pilus secretin channel. An emerging understanding of the Type IV pilus machinery and that of the related type II secretion system provide a structural framework to investigate the CTXφ infection process, with implications for understanding coliphage uptake in E. coli.

Acknowledgments—We are grateful to Ronald Taylor (Geisel School of Medicine) for providing anti-TcpA antibodies. Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, a Directorate of SLAC National Accelerator Laboratory and an Office of Science User Facility operated for the United States Department of Energy Office of Science by Stanford University. We thank the staff at Beamline 9-2 for assistance with remote data collection.
REFERENCES

1. Karolis, D. K., Johnson, J. A., Bailey, C. C., Boedeker, E. C., Kaper, J. B., and Reeves, P. R. (1998) A Vibrio cholerae pathogenicity island associated with epidemic and pandemic strains. Proc. Natl. Acad. Sci. U.S.A. 95, 3139–3139

2. Pearson, G. D., Woods, A., Chiang, S. L., and Mekalanos, J. J. (1993) CTX genetic element encodes a site-specific recombination system and an integrational colonization factor. Proc. Natl. Acad. Sci. U.S.A. 90, 3750–3754

3. Riechmann, L., and Holliger, P. (1997) The C-terminal domain of TolA is essential for filamentous phage encoding TolA toxin. Science 272, 1910–1914

4. Rakonjac, J., and Model, P. (1998) Roles of pIII in filamentous phage assembly. J. Mol. Biol. 282, 25–41

5. Davis, B. M., Moyer, K. E., Boyd, E. F., and Waldor, M. K. (2000) CTX prophages in classical biotype Vibrio cholerae. J. Bacteriol. 182, 6992–6998

6. Grant, R. A., Lin, T. C., Konigsberg, W., and Webster, R. E. (1981) Structure and location of the A protein in bacteriophage f1 spheroids and I-forms. J. Mol. Biol. 539–546

7. Lopez, J., and Webster, R. E. (1982) Minor coat protein composition and the location of the A protein in bacteriophage f1 spheroids and I-forms. J. Virol. 42, 1099–1107

8. Wen, Z. Q., Overman, S. A., and Thomas, G. J. Jr. (1997) Structure and interactions of the single-stranded DNA genome of filamentous virus fd. Investigation by ultraviolet resonance raman spectroscopy. Biochemistry 36, 7810–7820

9. Crissman, J. W., and Smith, G. P. (1984) Gene-III protein of filamentous virus fd. Evidence for a carboxyl-terminal domain with a role in morphogenesis. Virology 132, 445–455

10. Deng, L. W., Malik, P., and Perham, R. N. (1999) Interaction of the globular domains of pIII protein of filamentous bacteriophage fd with the F-pili of Escherichia coli. Virology 253, 271–277

11. Rakonjac, J., and Model, P. (1998) Roles of pIII in filamentous phage assembly. J. Mol. Biol. 282, 25–41

12. Riechmann, L., and Holliger, P. (1997) The C-terminal domain of TolA is the coreceptor for filamentous phage infection of E. coli. Cell 90, 351–360

13. Stengele, I., Bross, P., Garcés, X., Giray, J., and Raschd, I. (1990) Dissection of functional domains in phage fd adsorption protein. Discrimination between attachment and penetration sites. J. Mol. Biol. 212, 143–149

14. Scott, J. K., and Craig, L. (1994) Random peptide libraries. Curr. Opin. Biotechnol. 5, 40–48

15. Rakonjac, J., Bennett, N. J., Spagnuolo, J., Gagic, D., and Russel, M. (2011) Filamentous Bacteriophage. Biology, phage display, and nanotechnology applications. Curr. Issues Mol. Biol. 13, 51–76

16. Lukbowski, J., Hennecke, F., Plückthun, A., and Wlodawer, A. (1998) The structural basis of phage display elucidated by the crystal structure of the N-terminal domains of g3p. Nat. Struct. Biol. 5, 140–147

17. Helliwell, A. J., and Waldor, M. K. (2003) pIIICTX, a predicted CTX g3p minor coat protein, can expand the host range of coliphage fd to include Vibrio cholerae. J. Bacteriol. 185, 1037–1044

18. Gray, C. W., Brown, R. S., and Marvin, D. A. (1981) Adsorption complex of filamentous fd virus. J. Mol. Biol. 146, 621–627

19. Russell, M., Whirlow, H., Sun, T. P., and Webster, R. E. (1988) Low frequency infection of F-bacteria by transducing particles of filamentous bacteriophages. J. Bacteriol. 170, 5312–5316

20. Deng, L. W., and Perham, R. N. (2002) Delineating the site of interaction on the pIII protein of filamentous bacteriophage fd with the F-pili of Escherichia coli. J. Mol. Biol. 319, 603–614

21. Marvin, D. A. (1998) Filamentous phage structure, infection, and assembly. Curr. Opin. Struct. Biol. 8, 150–158

22. Jacobson, A. (1972) Role of F pilin in the penetration of bacteriophage fl. J. Virol. 10, 835–843

23. Clarke, M., Maddera, L., Harris, R. L., and Silverman, P. M. (2008) F-pili dynamics by live-cell imaging. Proc. Natl. Acad. Sci. U.S.A. 105, 19798–197981

24. Sun, T. P., and Webster, R. E. (1987) Nucleotide sequence of a gene cluster involved in entry of E. coliins and single-stranded DNA of infecting filamentous bacteriophages into Escherichia coli. J. Bacteriol. 169, 2667–2674

25. Leevengood, S. K., Beyer, W. F., Jr., and Webster, R. E. (1991) TolA. A membrane protein involved in colinic uptake contains an extended helical region. Proc. Natl. Acad. Sci. U.S.A. 88, 5939–5943

26. Holliger, P., and Pichlmann, L. (1997) A conserved infection pathway for filamentous bacteriophages is suggested by the structure of the membrane penetration domain of the minor coat protein g3p from phage fd. Structure 5, 265–275

27. Click, E. M., and Webster, R. E. (1998) The TolQRA proteins are required for membrane insertion of the major capsid protein of the filamentous phage fd during infection. J. Bacteriol. 180, 1723–1728

28. Lukbowski, J., Hennecke, F., Plückthun, A., and Wlodawer, A. (1999) Filamentous phage infection. Crystal structure of g3p in complex with its coreceptor, the C-terminal domain of TolA. Structure 7, 711–722

29. Martin, A., and Schmid, F. X. (2003) A proline switch controls folding and domain interactions in the gene-3-protein of the filamentous phage fd. J. Mol. Biol. 331, 1131–1140

30. Huff, J. P., Grant, B. J., Penning, C. A., and Sullivan, K. F. (1990) Optimization of routine transformation of Escherichia coli with plasmid DNA. Biotechniques 9, 570–572

31. Fan, Q. R., Mosyak, L., Winter, C. C., Wagtmann, N., Long, E. O., and Wiley, D. C. (1997) Structure of the inhibitory receptor for human natural killer cells resembles haematopoietic receptors. Nature 398, 96–100

32. Gonzales, A., Moorhead, P., McPhillips, S. E., Song, J., Sharp, K., Taylor, J. R., Adams, P. D., Sauter, N. K., and Soltis, S. M. (2008) Web-Ice. Integrated data collection and analysis for macromolecular crystallography. J. Appl. Crystallogr. 41, 176–184

33. Collaborative Computational Project. Number 4 (1994) The CCP4 suite. Programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760–763

34. Leslie, A. G. W. (1992) Joint CCP4 ESF-EAMCB Newsletter on Protein Crystallography, No. 26

35. Kabsch, H. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Crystallogr. 26, 795–800

36. Matthews, B. W. (1968) Solvent content of protein crystals. J. Mol. Biol. 30, 491–497

37. Tervilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. 55, 849–861

38. Tervilliger, T. C. (2000) Maximum-likelihood density modification. Acta Crystallogr. D Biol. Crystallogr. 56, 965–972

39. Emsley, P., and Cowtan, K. (2004) Coot. Model-building tools for molec-
Crystal Structures of CTXφ pIII and pIII-TolA

The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in *E. coli*. Mol. Microbiol. 63, 1008–1025

Goemaere, E. L., Cascales, E., and Lloubès, R. (2007) Mutational analyses define helix organization and key residues of a bacterial membrane energy-transducing complex. J. Mol. Biol. 366, 1424–1436

Lazzaroni, J. C., and Portalier, R. C. (1981) Genetic and biochemical characterization of periplasmic-leaky mutants of *Escherichia coli* K-12. J. Bacteriol. 145, 351–1358

Webster, R. E. (1991) The tol gene products and the import of macromolecules into *Escherichia coli*. Mol. Microbiol. 5, 1005–1011

Nagel de Zwaig, R., and Luria, S. E. (1967) Genetics and physiology of colicin-tolerant mutants of *Escherichia coli*. J. Bacteriol. 94, 1112–1123

Sun, T. P., and Webster, R. E. (1986) fii, a bacterial locus required for filamentous phage infection and its relation to colicin-tolerant tolA and tolB. J. Bacteriol. 165, 107–115

Li, C., Zhang, Y., Vankemmelbeke, M., Hecht, O., Aleanizy, F. S., Macdonald, C., Moore, G. R., James, R., and Penfold, C. N. (2012) Structural evidence that colicin A protein binds to a novel binding site of TolA protein in *Escherichia coli* periplasm. J. Biol. Chem. 287, 19048–19057

Jakob, R. P., Geitner, A. J., Weininger, U., Balbach, J., Dobbek, H., and Schmid, F. X. (2012) Structural and energetic basis of infection by the filamentous bacteriophage Ike. Mol. Microbiol. 84, 1124–1138

Sandkvist, M. (2001) Biology of type II secretion. Mol. Microbiol. 40, 271–283

Goldberg, I., and Mekalanos, J. J. (1986) Effect of a recA mutation on cholera toxin gene amplification and deletion events. J. Bacteriol. 165, 723–731

Buchan, D. W., Ward, S. M., Lobley, A. E., Nugent, T. C., Bryson, K., and Jones, D. T. (2010) Protein annotation and modelling servers at University College London. Nucl. Acids Res. 38 Suppl, W563–W568