3-14-2009

Genetic Mapping of Secretion and Functional Determinants of the Vibrio cholerae TcpF Colonization Factor

Shelly J. Krebs
Dartmouth College

Thomas J. Kirn
Dartmouth College

Ronald K. Taylor
Dartmouth College

JOURNAL OF BACTERIOLOGY, June 2009, p. 3665–3676 Vol. 191, No. 11
0021-9193/09/$08.00+0 doi:10.1128/JB.01724-08

Genetic Mapping of Secretion and Functional Determinants of the *Vibrio cholerae* TcpF Colonization Factor▽

Shelly J. Krebs, Thomas J. Kirn,† and Ronald K. Taylor*

Department of Microbiology and Immunology, Dartmouth Medical School, Hanover, New Hampshire 03755

Received 9 December 2008/Accepted 14 March 2009

Colonization of the human small intestine by *Vibrio cholerae* requires the type IV toxin-coregulated pilus (TCP). TcpF, which is encoded within the *tcp* operon, is secreted from the bacterial cell by the TCP apparatus and is also essential for colonization. Bacteria lacking *tcpF* are deficient in colonization, and anti-TcpF antibodies are protective in the infant mouse cholera model. In order to elucidate the regions of the protein that are required for secretion through the TCP apparatus and for its function in colonization, random mutagenesis of *tcpF* was performed. Analysis of these mutants suggests that multiple regions throughout the protein influence extracellular secretion and that determinants near the C terminus are important for the function of TcpF in colonization. The TcpF proteins of certain environmental *V. cholerae* isolates with 31% to 66% identity to pathogenic *V. cholerae* TcpF showed higher similarity in regions identified as secretion determinants but diverged in regions found to be important for colonization. These environmental TcpF proteins are secreted from the pathogenic strain; however, they do not mediate colonization in the infant mouse model. Here we provide genetic evidence pointing toward regions of TcpF that influence secretion, as well as regions that play an important role in in vivo colonization.

Vibrio cholerae, a gram-negative bacterium, is the causative agent of the acute diarrheal disease cholera. More than 100,000 cases of cholera are reported every year; however, this is estimated to represent only 5 to 10% of the actual number of worldwide cases due to poor surveillance systems and frequent underreporting (http://www.who.int/topics/cholera/surveillance/en/index.html). *V. cholerae* exists in an aquatic environment and is transmitted to humans through contaminated food or water. Even though more than 200 serogroups that cause sporadic minor cases of cholera have been reported for *V. cholerae*, only two serogroups, O1 and O139, are responsible for epidemics of this disease (reviewed in reference 17). Epidemic strains of *V. cholerae* require two genetic elements to cause disease: CTXϕ, which encodes the cholera toxin, and the *Vibrio* pathogenicity island, which contains the toxin-coregulated pilus (TCP) gene cluster. Secretion of cholera toxin (CT) leads to severe diarrhea through a cascade of events that includes upregulation of adenylate cyclase activity in intestinal epithelial cells and culminates in a large efflux of electrolytes and water into the small intestinal lumen. Colonization is an essential step in the delivery of CT to epithelial cells and therefore is a prerequisite for the establishment of a productive infection. The TCP has been demonstrated to play an essential role in the colonization of the intestine (13, 46) by facilitating microcolony formation via pilus-mediated bacterial interactions (21). TCP is a type IV pilus that appears as large bundles of pilus filaments composed of polymerized TcpA pilin. Proteins involved in the biogenesis of TCP are encoded in the *tcp* operon and make up a complex machinery that is thought to span the inner membrane, periplasm, and outer membrane of the bacterium.

In addition to the well-characterized virulence factors CT and TCP, we have recently reported a third virulence factor, TcpF, which is secreted from the bacterial cell and is required for intestinal colonization (19). TcpF is encoded within the *tcp* operon and utilizes the TCP apparatus for extracellular secretion. A bacterial strain lacking *tcpF* essentially retains all of the in vitro properties associated with the elaboration of a functional TCP; however, it is as defective in colonization as a TCP-negative strain. In fact, *tcpF* mutant strains are cleared from the intestine at the same rate as those lacking TCP (22). It has been found that TcpF is expressed in vivo during human infection and generates a substantial immune response in patients infected with *V. cholerae* (11). In addition, antibodies directed against TcpF are protective against cholera infection in the infant mouse model (22), further demonstrating the importance of the role of TcpF in colonization and infection.

The mechanism by which TcpF mediates colonization is unknown, and TcpF does not share sequence identity with any other known protein. However, TcpF has many striking similarities to enterotoxigenic *Escherichia coli* (ETEC) CofJ and *Citrobacter rodentium* CfcJ (27, 45). There appear to be two other type IV pili with genetic organizations most highly related to the *tcp* operon. These are ETEC CFA/III (*cof* operon) and *C. rodentium* CFC (colonization factor *Citrobacter*; *cfc* operon) (27). The *cof* and *cfc* operons each contain a gene, *cofJ* or *cfcJ*, respectively, located in the position that corresponds to *tcpF* in the *tcp* operon. Each of these proteins has a conventional Sec-dependent signal sequence and is similar to TcpF in size. The TcpF, CofJ, and CfcJ proteins do not have homology or identity to each other, or to any other proteins, by BLAST analysis (19, 27, 45). There are no published data regarding the localization of CfcJ at this time, but CofJ is secreted extracellularly from ETEC strains (S. J. Krebs, unpublished data).

* Corresponding author. Mailing address: Dartmouth Medical School, Department of Microbiology and Immunology, 7550 Vail Building, Hanover, NH 03755. Phone: (603) 650-1632. Fax: (603) 650-1318. E-mail: ronald.k.taylor@dartmouth.edu.

† Present address: Hospital of the University of Pennsylvania, Department of Pathology and Laboratory Medicine, 6 Founders, 3400 Spruce St., Philadelphia, PA 19104.

▽ Published ahead of print on 20 March 2009.

A striking feature of type IV pilus biogenesis systems is their similarity to type II secretion systems (T2SSs) (28, 35, 36, 40). T2SSs are one of the many means by which gram-negative bacteria export proteins from the bacterial cell, and they are important for the extracellular localization of a variety of toxins and enzymes (6, 7, 39). The nascent polypeptides of these exported proteins contain a signal sequence that allows them to be transported to the periplasm via the Sec-dependent transport system (41). Transfer of the proteins across the cytoplasmic membrane is accompanied by cleavage of the signal sequence and folding of the protein within the periplasmic space. The final step of secretion across the outer membrane involves components located in both the inner and the outer membrane, where they assemble into a multiprotein apparatus powered by a cognate ATPase (reviewed by Pugsley [33] and Sandkvist [37]). Many of the proteins essential for T2SSs share homology with proteins involved in the assembly of type IV pilus systems, suggesting that the organization and function of the two machineries are similar (15, 29, 49).

Although proteins constituting T2SSs have been well defined, details concerning the molecular determinants specifying secretion of the transported proteins are only beginning to be understood. Even for T2SSs that secrete several different proteins, such as the extracellular protein secretion system of *V. cholerae*, no specific secretion signal common to all the substrates has been identified (10), although these systems demonstrate a high degree of specificity for their substrates (29). One study found at least three regions in a T2SS targeted substrate that contain information influencing secretion (10). Others have suggested that the multiple secretion determinants found throughout the substrate require proper folding of the protein to position these regions in close proximity to each other, so that they can be recognized by the secretion apparatus for transport (3, 4).

We have previously reported that the TCP apparatus is responsible for the extracellular secretion of TcpF (19), mimicking a T2SS in this regard. TcpF has a Sec signal sequence for translocation into the periplasm via the Sec apparatus and requires all of the components of the TCP apparatus, including the pilin monomer TcpA, to be secreted extracellularly (19). In this study we sought to identify regions of TcpF that are required for recognition by the TCP apparatus for extracellular secretion, as well as determinants important for its function in colonization. Identification of functional determinants could potentially facilitate the targeting of such regions for incorporation into a multivalent subunit vaccine. Since there are no homologs of TcpF and no recognized extracellular "signal sequences" in type II secreted proteins, we generated random linker mutations along with multiple internal deletions in *tcpF* and then tested the TcpF protein derivatives for extracellular secretion from *V. cholerae*. If the protein derivative was secreted, we assayed its functional ability in colonization. Using this random mutagenesis approach, regions were identified throughout TcpF that influence secretion through the TCP apparatus. We also identified determinants near the C terminus that are important for in vivo colonization. To date, *tcpF* genes from six pathogenic strains of *V. cholerae* have been sequenced, and all of these strains have nearly identical TcpF amino acid sequences. However, a study by Mukhopadhyay et al. (26) found four different alleles of TcpF in environmental strains isolated from Kolkata, India, which ranged from 31% to 70% identity to canonical TcpF at the protein level. These environmental TcpF proteins were particularly interesting because they retained a high degree of identity in the regions that we found important for secretion, but they were divergent in regions that we found important for function. These environmental TcpF proteins were expressed in a Δ*tcpF* pathogenic strain and were secreted by the pathogenic strain but could not mediate colonization in the infant mouse cholera model, consistent with the suggested roles of the regions defined by our mutagenesis study in secretion and colonization.

MATERIALS AND METHODS

Bacterial strains and growth conditions. All bacterial strains and plasmids used in this study are shown in Table 1. Strains were stored at −70°C in Luria-Bertani (LB) broth containing 30% (vol/vol) glycerol. *V. cholerae* was grown in LB broth with a starting pH of 6.5 at 30°C for 16 h (TCP-inducing conditions). When necessary, antibiotics were used in culture at the following concentrations: ampicillin (Ap), 100 μg/ml; gentamicin (Gm), 30 μg/ml; chloramphenicol (Cm), 15 μg/ml; streptomycin (Sm), 100 μg/ml generally or 1 mg/ml for selection for loss of integrated plasmids from *V. cholerae* during allelic exchange (43). 5-Bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) was used in LB agar at 40 μg/ml for enumeration in the competitive index experiments. Arabinose was used at a final concentration of 0.01%. Primers for the construction of each plasmid or strain are listed in Table 2.

Construction of *tcpF* internal deletions. *V. cholerae tcpF* internal deletions were constructed as follows. A sequence from the *V. cholerae* database identified by a BLAST search as encoding TcpF provided the basis for the design of oligonucleotide primers F25A, F25B, F25C, F25D, F80A, F80B, F80C, F80D, F140A, F140B, F140C, F140D, F221A, F221B, F221C, F221D, F260A, F260B, F260C, F260D, F270A, F270B, F270C, F270/290D, F290A, F290B, and F290C (Table 2), which were used to amplify chromosomal sequences from O395 that flank the specific internal deletion sites within the *tcpF* gene. The upstream fragment of approximately 500 bp and the downstream 500-bp fragment were inserted into the allelic exchange vector pKAS32. The *tcpF* internal deletions were then introduced into the chromosome of *V. cholerae* strain O395 by allelic exchange (43), and the constructs were verified by PCR.

Linker scanning mutagenesis. Random mutagenesis of *V. cholerae tcpF* was achieved using the GPS Linker Scanning system (New England Biolabs) and has been described previously by Nye et al. (30). Briefly, TnsABC* transposase was used to insert a Tn*7*-based transprimer into pTK10 at random locations. The transposition reaction causes a staggered cut at the transprimer insertion site that results in a 5-bp duplication. The resultant plasmids were selected on Cm and screened by PCR for insertions within the *tcpF* gene. Subsequent PmeI digestion and religation were performed to remove the bulk of the transprimer, after which plasmids were designated with the suffix "P." Religation of the mutagenized plasmids ultimately leaves a 15-bp insertion 5 bp from the duplication at the transprimer insertion site and 10 bp from the PmeI site. In four out of six frames, these 15 bp cause an insertion of 5 amino acids. In two out of six frames, an ocher mutation is created. Mutagenized plasmids were characterized by DNA sequencing of *tcpF* using primers FHISA and FHISB to determine the location of the 5-amino-acid insertion or nonsense codon. Plasmids encoding representative mutations were analyzed in *V. cholerae* O395 Δ*tcpF*. Pentapeptide insertions in the TcpF protein are named by the amino acid residue N-terminal to the insertion followed by the amino acid sequence of the insertion itself. Induction of the pBAD22-based TcpF derivatives was achieved using 0.01% arabinose with *V. cholerae* strains.

Cellular fractionation. Cells grown under TCP-expressing conditions were harvested by centrifugation, and culture supernatant fractions were removed and passed through a 0.4-μm-pore-size filter. Cellular fractionation was performed as previously described by Bose and Taylor (2). Briefly, pellets were washed, suspended in phosphate-buffered saline, and treated with polymyxin B sulfate (10 mg/ml) for 10 min on ice to release periplasmic contents. Spheroplasts were separated from the soluble periplasmic fraction by centrifugation at 10,000 × *g* for 10 min. In order to obtain total membranes, the spheroplast fraction was resuspended in 10 mM Tris-Cl (pH 8.0) and passed through a French pressure cell at 8,000 lb/in^2 to lyse the cells. After low-speed centrifugation at 10,000 × *g* at room temperature for 10 min to remove any unbroken cells, the lysate was centrifuged at 100,000 × *g* at 4°C for 30 min to pellet the membrane fraction. The inner and outer membranes were differentially solubilized as follows. The mem-

TABLE 1. Strains and plasmids used in this study

Strain or plasmid	Characteristics	Reference or source
Strains		
O395 Sm	Classical Ogawa; Sm^r	43
RT4372	O395 Δ*tcpF1*	17
RT4031	O395 Δ*tcpA11*	19
SCE4	Environmental isolate, O8 serogroup	24
SCE256	Environmental isolate, O42 serogroup	24
SCE263	Environmental isolate, O10 serogroup	24
SJK71	O395 Δ*tcpF1* (pBAD22)	This study
SJK9	O395 Δ*lacZ1* Δ*tcpF1* (pTK10)	This study
SJK7	O395 Δ*tcpF1*(pTK10)	This study
SJK252	O395 Δ*tcpC1* Δ*tcpF1*	This study
SJK94	O395 Δ*tcpF1* Δ*htrA1*	This study
SJK198	O395 *tcpF* internal deletion 1 (Δaa6–29)	This study
SJK199	O395 *tcpF* internal deletion 2 (Δaa61–80)	This study
SJK202	O395 *tcpF* internal deletion 3 (Δaa121–129)	This study
SJK205	O395 *tcpF* internal deletion 4 (Δaa201–213)	This study
SJK214	O395 *tcpF* internal deletion 5 (Δaa240–269)	This study
SJK216	O395 *tcpF* internal deletion 6 (Δaa249–311)	This study
SJK218	O395 *tcpF* internal deletion 7 (Δaa270–311)	This study
Plasmids		
pBAD22	pMB1, P*ara* promoter, *araC*; Ap^r expression plasmid	Laboratory collection
pSK60	SM107λpir/pKAS32 Δ*htrA* Ap^r/Km^r	N. A. Beck, unpublished data
pTK7	pKAS32 Δ*tcpF* Ap^r	17
pTK10	pBAD22 TcpF	This study
pSJK14	pBAD22 TcpF (S65::MFKQS)	This study
pSJK16	pBAD22 TcpF (A206::CLNGA)	This study
pSJK18	pBAD22 TcpF (Y12::CLNIY)	This study
pSJK19	pBAD22 TcpF (S288::RFKHS)	This study
pSJK20	pBAD22 TcpF (A13::CLNNA)	This study
pSJK21	pBAD22 TcpF (T177::CLNTT)	This study
pSJK26	pBAD22 TcpF (P64::CLNNP)	This study
pSJK29	pBAD22 TcpF (K36::MFKHK)	This study
pSJK30	pBAD22 TcpF (G38::CLNIG)	This study
pSJK31	pBAD22 TcpF (N123::CLNTN)	This study
pSJK37	pBAD22 TcpF (D87::CLNSD)	This study
pSJK41	pBAD22 TcpF (F59::LFKQF)	This study
pSJK42	pBAD22 TcpF (V229::FKQDV)	This study
pSJK43	pBAD22 TcpF (V153::FKQNV)	This study
pSJK44	pBAD22 TcpF (L268::FKQKL)	This study
pSJK47	pBAD22 TcpF (F52::CLNTF)	This study
pSJK52	pBAD22 TcpF (I296::CLNKI)	This study
pSJK55	pBAD22 TcpF (G205::CLNSG)	This study
pSJK61	pBAD22 TcpF (L218::FKQLV)	This study
pSJK63	pBAD22 TcpF (N291::CLNTM)	This study
pSJK244	pBAD22 TcpF (SCE4)	This study
pSJK245	pBAD22 TcpF (SCE256)	This study
pSJK247	pBAD22 TcpF (SCE263)	This study

brane pellet was resuspended in 10 mM Tris-Cl–100 mM NaCl containing 2.5% Sarkosyl, held at room temperature for 30 min, and centrifuged at 200,000 × *g* at 12°C for 1 h. The supernatant was retained as the inner membrane fraction, and the pellet containing the outer membrane was resuspended in 200 μl of 10 mM Tris-Cl (pH 8.0).

As previously described by Tripathi and Taylor (48), the purity of subcellular fractions was determined by immunoblotting and enzyme assays for known localization markers, as follows. EpsL and TcpC immunoblots were used as inner membrane and outer membrane localization markers, respectively (2, 38). As a cytoplasmic marker to test for cell lysis, glucose-6-phosphate dehydrogenase (G6PDH) activity was assayed (14). G6PDH substrate vials containing glucose-6-phosphate along with the phenazine methosulfate NADP, substrates, and 2,6-dichlorophenol indophenol (DCP-IP) (Sigma) were reconstituted with 0.3 M Tris-Cl (pH 8.0) just prior to use, as recommended by the manufacturer. One milliliter of a 1:2 dilution of the cytoplasmic fraction or 1 ml of a culture supernatant or outer membrane fraction was added to a 4.5-ml cuvette containing the reconstituted substrate, immediately followed by an overlay of 2 ml of mineral oil. The decrease in the optical density at 600 nm was measured at 1-min intervals for 5 min to monitor the reduction of blue DCP-IP to a colorless form (1). Specific activity was calculated as nanomoles of DCP-IP reduced per minute per milligram of total protein. The culture supernatant, periplasmic, and outer membrane fractions from each strain were tested for the presence or absence of these inner membrane and outer membrane control proteins and for G6PDH activity.

Protein electrophoresis and Western blot analysis. Proteins from various cellular fractions were diluted in 2× sodium dodecyl sulfate-polyacrylamide gel electrophoresis loading buffer, and fractions containing equal amounts of total protein were separated on a sodium dodecyl sulfate–12.5% polyacrylamide gel. Protein levels were estimated using the bicinchoninic acid protein assay kit (Pierce) with bovine serum albumin as a standard. For total-protein detection, gels were stained with 0.1% Coomassie brilliant blue in 40% methanol for 1 min and destained with 50% methanol. For immunodetection, proteins were electroblotted onto a nitrocellulose membrane at 4°C in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol [pH 8.3]) with a wet transfer apparatus (Bio-Rad Laboratories) (46). The membrane was blocked with 3% bovine serum albumin in 1× Tris-buffered saline with 0.1% Tween (TBST) for 2 h. Antisera containing

TABLE 2. Oligonucleotide primers used in this study

Primer name	Sequence (5′ to 3′)	Strain or plasmid[a]
F25A	GATCGGAATTCGGCCAATATGTCCATGTTGC	SJK198
F25B	GATCGCTCTTCGATAATTATCATTAAATGC	SJK198
F25C	GATCGCTCTTCGTATCCTTATTTAGAGTGTATC	SJK198
F25D	GATCGAGATCTTAAACGACTAAGAGTCCCTTC	SJK198
F80A	GATCGGAATTCGACTACAAACAGGCTCTCGAC	SJK199
F80B	GATCGCTCTTCGATAAAACATATCTTGAGAGGT	SJK199
F80C	GATCGCTCTTCGTATGCTCGGCTTTTAGCTGGT	SJK199
F80D	GATCGAGATCTCAGGGTCAAGATAACTCCTG	SJK199
F140A	GATCGGAATTCCGTTGGCGTTGTCTGTGATTTGG	SJK202
F140B	GATCGCTCTTCGATTAACTAGTGTCTCAATTTG	SJK202
F140C	GATCGCTCTTCGAATCTTAATCAAGAGTATATTACAGG	SJK202
F140D	GATCGAGATCTGCTTTGCTATACAGTTTCACACCAC	SJK202
F221A	GATCGGAATTCACGGATTCAAGGGGGAGTGAGCAC	SJK205
F221B	GATCGCTCTTCGTAAACGACTAAGAGTCCCTTC	SJK205
F221C	GATCGCTCTTCGTTAGTAAATGAATTAAGTAAGC	SJK205
F221D	GATCGAGATCTCCCAATCATTGCGTTCTACTCTG	SJK205
F260A	GATCGGAATTCCGACAGAATCTGACGGTGCTAAG	SJK214
F260B	GATCGCTCTTCGTGCATGTGGAACTCCAACAAG	SJK214
F260C	GATCGCTCTTCGGCAGTATCTACAAATGACATGC	SJK214
F260D	GATCGAGATCTCTTAAATAGAGTATTGATTGGTTCG	SJK214
F270A	GATCGGAATTCCCCATCGACAGAATCTGACGGTGC	SJK216
F270B	GATCGCTCTTCGAACAGGGTCATAGATAACTCCTG	SJK216
F270C	GATCGCTCTTCGGTTTAAAATAGATAAGATAACAGCC	SJK216
F270/290D	GATCGAGATCTCATGAGATCACCAAGAATGTAAG	SJK216/SJK218
F290A	GATCGGAATTCGCTCAAACGACTTTCGATTGGG	SJK218
F290B	GATCGCTCTTCGATTATACAACTTCTTACCCG	SJK218
F290C	GATCGCTCTTCGAATTAAAATAGATAAGATAACAGCC	SJK218
FHISA	AGGAGGAATTCACCATGAGATATAAAAAAACCTTAATGTTATCAATC	pTK10
FHISB	AAAAAAGCTTGTGGTGGTGGTGGTGGTGTTATTTAAAGTTCTCTGAAATATGCTTTG	pTK10
BadFE4A	GATCGCAATTCACCATGAAATATAAAAAAACCTTAATG	pSJK244
BadFE4B	GATCGTCTAGATTATTTAAATTTCTCTGAATATGCC	pSJK244
BadFE256A	GATCGGAATTCACCATGAAACTTAAATTAATTTCATCATCAATC	pSJK245
BadFE256B	GATCGTCTAGATTAATCGCTTACACCCTCACAAAAC	pSJK245
BadFE263A	GATCGGAATTCCACCATGAAACTAACTATGAAGAAACTTCC	pSJK247
BadFE263B	GATCGTCTAGATTACAGTCTAAACAGTTCAC	pSJK247

[a] Strain or plasmid for which the primer was used.

polyclonal antibodies raised against TcpF–glutathione *S*-transferase were then applied at a 1:10,000 dilution in blocking buffer. A horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin G secondary antibody was added at a 1:10,000 dilution in TBST. Reactive protein bands were detected using an ECL kit (Amersham).

In vivo competition index analysis. For the in vivo competitive index determinations, strains to be tested were grown under TCP-inducing conditions and then mixed with equal cell numbers of the reference strain, SJK9. Five- to seven-day-old CD1 mice from mixed litters were orally inoculated with 50 μl of the mixture with a final concentration of 1% arabinose and were incubated at 30°C for 24 h. The bacteria were then recovered by homogenizing harvested intestines in 3 ml of 10% glycerol in water with a Tissue Terror homogenizer (Fisher Scientific, Los Angeles, CA). The homogenate was appropriately diluted and plated onto solid medium containing streptomycin and X-Gal. The competitive index was calculated by dividing the output ratio (ratio of the output of the test strain to that of the reference strain recovered from the intestine) by the input ratio (ratio of the input of the test strain to that of the reference strain in the original inoculum). The competitive index is the average for at least five mice for each strain tested, and the Student *t* test was used to determine which strains were significantly defective in colonization compared to the positive control.

RESULTS

Construction of *tcpF* mutations using linker scanning mutagenesis. Since TcpF presently has no known homologs or obvious conserved domains, targeted mutagenesis of regions of the protein likely to be important for secretion through the TCP apparatus or for in vivo colonization could not be undertaken. Instead, we employed a random mutagenesis approach to identify regions involved in the secretion and function of TcpF by using the New England Biolabs GPS LS Linker Scanning system (see Materials and Methods). Mutagenesis generated 33 unique linker insertions encoding 5 amino acids each and 18 unique ocher truncation mutations in *tcpF*. Twenty of the 33 insertions produced stable protein derivatives as determined by anti-TcpF immunoblotting (data not shown) and were studied further. None of the 18 truncated proteins were found to be stable (data not shown), suggesting that the C terminus may contribute to protein stability. The plasmids expressing derivatives of TcpF that were stable in *V. cholerae* RT4372 were sequenced, and the location of each insertion was determined. A graphical representation of the locations of these TcpF variants derived from linker scanning mutagenesis is given in Fig. 1A, and their respective 5-amino-acid insertions are listed in Fig. 1B. Internal deletions (discussed below) were also constructed throughout *tcpF* and are represented in Fig. 1.

Regions near the C terminus of TcpF are required for its function in colonization. To begin to define the determinants important for the secretion and function of TcpF, *V. cholerae* strains expressing stable TcpF derivatives (Fig. 1) were first analyzed by Western blotting for their abilities to secrete the variant TcpF proteins into the culture supernatant. Eleven of the 20 TcpF derivatives were detected in the culture supernatant (Fig. 2A). As seen previously, TcpF is represented by

A.

B.

Plasmid	LS Insertion
pSJK18	Y12::CLNIY
pSJK20	A13::CLNNA
pSJK29	K36::MFKHK
pSJK30	G38::CLNIG
pSJK47	F52::CLNTF
pSJK41	F59::LFKQF
pSJK26	P64::CLNNP
pSJK14	S65::MFKQS
pSJK37	D87::CLNSD
pSJK31	N123::CLNTN
pSJK43	V153::FKQNV
pSJK21	T177::CLNTT
pSJK55	G205::CLNSG
pSJK16	A206::CLNGA
pSJK61	L218::FKQLV
pSJK42	V229::FKQDV
pSJK44	L268::FKQKL
pSJK19	S288::RFKHS
pSJK63	N291::CLNTM
pSJK52	I296::CLNKI
Internal Deletion	
#1	Δaa6-29
#2	Δaa61-80
#3	Δaa121-129
#4	Δaa201-213
#5	Δaa240-269
#6	Δaa249-311
#7	Δaa270-311

FIG. 1. Summary of the locations of *tcpF* mutations. (A) Amino acid sequence of mature TcpF. Shaded letters represent the Sec-dependent signal sequence. The locations of the linker scanning mutations are indicated above the TcpF sequence. Internal deletions are underlined and labeled below the TcpF sequence. (B) Alteration of each TcpF derivative. Plasmids carrying each *tcpF* linker scan (LS) mutation are listed, and the insertions are given according to the nomenclature described in Materials and Methods. Chromosomal internal deletions of *tcpF* are indicated by numbers 1 through 7, and the amino acids (aa) deleted are shown.

multiple bands (19), and it is not yet known if one or all these forms of the protein are important for its function in colonization. The lack of G6PDH activity in these culture supernatant fractions compared to the cytoplasmic fraction suggests that little or no cell lysis, which could contribute to the detection of these different forms, is occurring (data not shown). Each of the strains expressing TcpF derivatives that were detected in the culture supernatant were competed in the infant mouse colonization model against the reference strain expressing wild-type TcpF, SJK9 (Fig. 2B). The positive control was competition of SJK7 against SJK9, each expressing wild-type TcpF, which gave a competitive index of about 1, indicating that these strains colonize the infant mouse equally. The negative control was competition of SJK71 (empty vector) against SJK9 (expressing wild-type TcpF), which gave a competitive index of approximately 1 log unit less than 1. Student's *t* test showed that this difference was significant, indicating that SJK71 is defective in colonization, as predicted. As shown in Fig. 2, plasmids encoding TcpF derivatives with insertions near the N terminus (pSJK18, pSJK20, pSJK26, pSJK14) were not defective in colonization. However, plasmids encoding TcpF derivatives with insertions from the middle of the protein to near the C terminus (pSJK43, pSJK21, pSJK55, pSJK16, pSJK42, pSJK44, pSJK19) were significantly defective in colonization (Fig. 2B). The amount of protein produced by the strain containing plasmid pSJK21 was less than the amount of protein produced by the reference strain in vitro, which could account for its defect in colonization in vivo. However, the strain containing plasmid pSJK18 also produced less protein than the reference strain, and its function in colonization was not affected, suggesting that the amount of protein expressed by the strain containing plasmid pSJK21 would be sufficient for colonization if the protein were functional. Taken together, these results suggest that regions near the C terminus are required for the function of TcpF in the colonization and pathogenesis of *V. cholerae*.

Multiple regions throughout TcpF influence extracellular secretion. In order to elucidate the molecular determinants required for proper secretion of TcpF by the TCP apparatus, subcellular localization of TcpF derivatives was performed. The subcellular fractionation protocol was able to produce fairly pure fractions (data not shown), as indicated by the presence or absence of the inner membrane protein marker EpsL, the outer membrane marker TcpC, or the activity of G6PDH (see Materials and Methods). *V. cholerae* strains expressing the TcpF linker scan derivatives that could not be detected in the culture supernatant were examined to determine if these TcpF derivatives could be detected in the periplasm or the outer membrane. All of the known components of the TCP type IV pilus biogenesis apparatus are required for extracellular secretion of TcpF, including the pilin monomer, TcpA (19). In mutants missing any of these components, TcpF accumulates in the periplasm (19). Overexpression of *tcpF* in *V. cholerae* (SJK7) led to normal amounts of TcpF in the culture supernatant and also some periplasmic accumulation of the protein, likely due to higher than normal

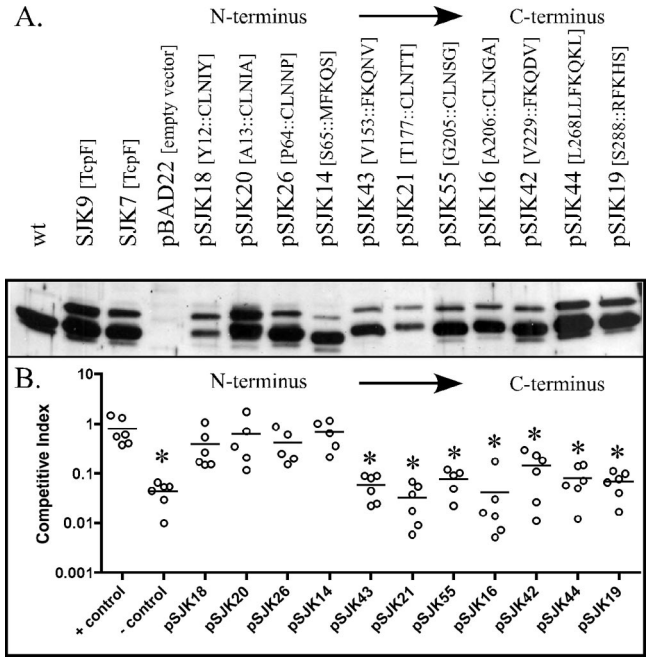

FIG. 2. Analysis of TcpF derivatives in colonization. TcpF derivatives constructed by linker scanning mutagenesis that were detected in the culture supernatant were analyzed for their functions in colonization. (A) Stable TcpF derivatives in the culture supernatant were loaded onto the protein gel according to the location of the 5-amino-acid insertion in TcpF, from the N terminus to the C terminus, and were detected by Western blotting using anti-TcpF polyclonal antibodies. Plasmids encoding the TcpF derivatives are given above the Western blot, followed by the alteration of TcpF in brackets. wt, wild type. (B) The competitive index was determined for each TcpF derivative by competing each strain in the infant mouse model against the reference strain, SJK9, expressing wild-type TcpF. The competitive index of each TcpF derivative was graphed from the N terminus to the C terminus according to the location of the 5-amino-acid insertion. Each circle represents a mouse, and each horizontal line represents the average competitive index for the five or six mice tested for one strain. The positive (+) control is the competition between SJK9 and SJK7, both expressing wild-type TcpF. The negative (−) control is the competition between SJK9, expressing wild-type TcpF, and SJK71, harboring the empty vector. Student's *t* test was used to determine which of the strains were significantly defective in colonization compared to the positive control. TcpF derivatives with 5-amino-acid insertions that mapped toward the C terminus were found to be significantly defective in colonization, as indicated by asterisks ($P < 0.01$).

levels of TcpF that occur when the gene is expressed from the plasmid (Fig. 3). Of the 20 mutants that produced stable protein derivatives from linker scanning mutagenesis, 9 mutants were totally devoid of full-length protein in the culture supernatant. However, these protein derivatives yielded full-length products that were detected in the periplasm (Fig. 3). Among these secretion-defective TcpF derivatives, six had insertions that clustered in two main regions. Plasmids pSJK29, pSJK30, pSJK47, and pSJK41 encode TcpF derivatives with 5-amino-acid insertions located near the N terminus, and plasmids pSJK63 and pSJK52 encode TcpF derivatives with insertions located near the C terminus (Fig. 1). The remaining three plasmids (pSJK37, pSJK31, and pSJK61) encode TcpF derivatives with insertions distributed at various positions throughout the protein. Strains containing plasmid pSJK63 or pSJK52, encoding a TcpF derivative that contains an insertion located near the C terminus, or pSJK47, encoding a TcpF derivative that contains an insertion located near the N terminus, produced TcpF derivatives that accumulated in the outer membrane fraction to a greater degree than wild-type TcpF (Fig. 3). To determine if these protein derivatives were held in a form associated with the outer membrane secretin, TcpC, these genes were expressed in strains lacking *tcpC* and *tcpF* and were blotted with anti-TcpF polyclonal antibodies. However, these TcpF derivatives still accumulated in the outer membrane, even in the absence of the outer membrane secretin (data not shown). To determine if the 5-amino-acid insertion might cause these protein derivatives to form β-barrels, leading to their accumulation in the outer membrane independently of interactions with other proteins, PSORT (http://www.psort.org/psortb/index.html) was utilized to analyze the amino acid sequences of these TcpF derivatives. Based on this computer algorithm, these TcpF derivatives had scores similar to that of wild-type TcpF and were not predicted to localize to the outer membrane. This finding suggests that either TcpF is associating with other components of the TCP biogenesis apparatus localized to the outer membrane or there are other components not encoded within the *tcp* operon that aid in mediating proper extracellular secretion of TcpF. Perhaps these three protein derivatives encoded on plasmids pSJK47, pSJK63, and pSJK52 retain some ability to initiate extracellular export but are defective in its completion, leading to a membrane-bound intermediate, whereas the other six protein derivatives are completely defective in interaction with the export machinery.

The six protein derivatives encoded on plasmids pSJK29, pSJK30, pSJK47, pSJK41, pSJK37, and pSJK31, which have 5-amino-acid insertions located close to the N terminus, were found to be partially degraded into two products in the periplasm: degradation product A (approximately 18 kDa) and degradation product B (approximately 13 kDa) (Fig. 3). Degradation product B was also detected at low levels in the culture supernatant. When mutants lacking the secretin, TcpC, were transformed with the plasmids encoding these protein derivatives, degradation product B was still detected in the culture supernatant (data not shown), indicating that degradation product B is released from the bacterial cell independently of the TCP apparatus. Degradation of these protein derivatives must have occurred in the periplasm and not in the supernatant, since full-length TcpF requires TcpC for secretion (19). Wild-type TcpF (SJK7) is not degraded into products A and B (Fig. 3); thus, the 5-amino-acid insertion in TcpF derivatives encoded on plasmids pSJK29, pSJK30, pSJK47, pSJK41, pSJK37, and pSJK31 may have caused changes in the protein structure, which could account for their lack of secretion. It has been well documented that accumulation of proteins in the periplasm can have harmful effects on cell viability (16, 18, 42, 44).

HtrA (DegP), characterized in *E. coli*, is a ubiquitous housekeeping protein found throughout the prokaryotes and eukaryotes. One of its primary functions in gram-negative prokaryotes is to act as a protease under conditions where proteins accumulate in the periplasm (5; reviewed in references 23 and 31). To determine whether HtrA has a role in degrading the unstable TcpF derivatives, a strain lacking both *htrA* and *tcpF* (SJK94) was transformed with plasmids encoding these proteins. In this strain, degradation products were no longer de-

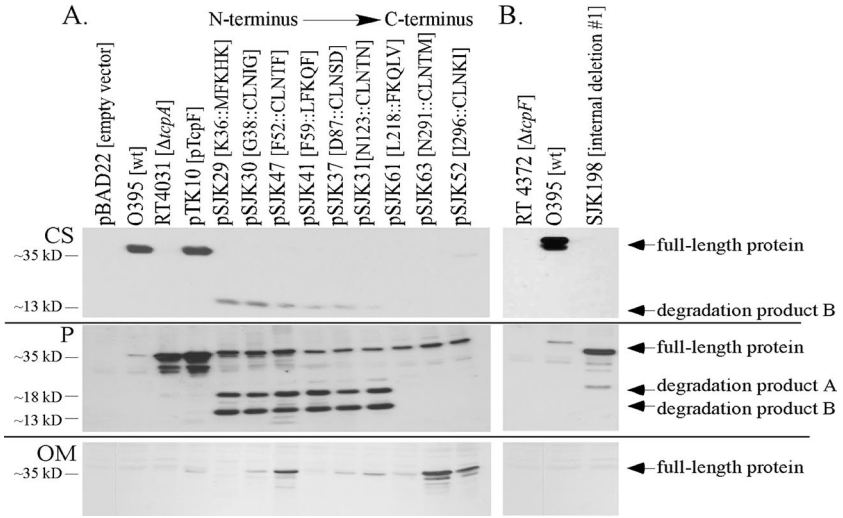

FIG. 3. Western blot analysis of TcpF linker scan derivatives and the mutant with internal deletion 1 for extracellular secretion. Western blotting with anti-TcpF polyclonal antibodies was performed on the culture supernatant (CS), periplasmic (P), and outer membrane (OM) fractions. (A) TcpF derivative fractions were loaded according to the location of the 5-amino-acid insertion, from the N terminus to the C terminus. Full-length TcpF derivatives constructed by linker scanning mutagenesis were detected in either the periplasm or the outer membrane fraction, or both, but were not detected in the culture supernatant. (B) The mutant with internal deletion 1 was detected in the periplasm but not in the outer membrane or the culture supernatant.

tected in the periplasm, and full-length protein was restored to wild-type levels, indicating that HtrA degrades these protein derivatives (Fig. 4). However, even though these TcpF derivatives were stable in the *htrA* background, they still could not be detected in the culture supernatant (Fig. 4), suggesting that the regions identified by these linker insertions are important for efficient secretion of TcpF.

Limited regions of primary sequences common among secretion-proficient TcpF proteins from environmental isolates provide a way to identify secretion determinants. To date, *tcpF* genes from six pathogenic strains of *V. cholerae* have been sequenced, and all of these strains have identical TcpF amino acid sequences. However, a study by Mukhopadhyay et al. (26) found four different alleles of *tcpF* in strains isolated from the environment in Kolkata, India, which ranged from 31% to 70% identity at the protein level to TcpF proteins from pathogenic strains. Comparison of all five sequences in a Clustal W alignment (47) reveals 15% similarity throughout mature TcpF when identical, conserved, and semiconserved residues are considered. For clarity, TcpF from pathogenic strains is re-

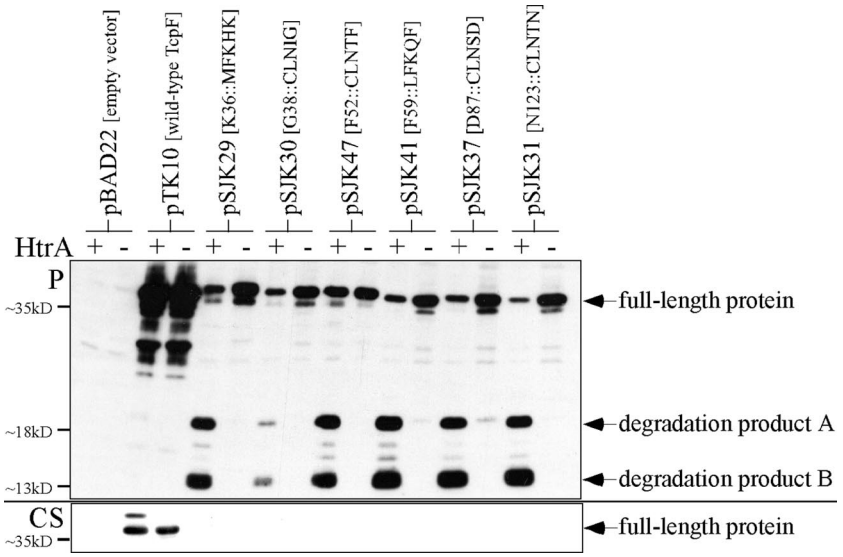

FIG. 4. Lack of HtrA restores the stability of protein derivatives that are defective in extracellular secretion. The TcpF protein derivatives that were degraded into products A and B, as shown in Fig. 3, were analyzed in the presence and absence of HtrA by Western blotting of the periplasmic (P) and culture supernatant (CS) fractions with anti-TcpF polyclonal antibodies. Although HtrA restored these protein derivatives to a stable full-length product in the periplasm, they still could not be detected in the culture supernatant.

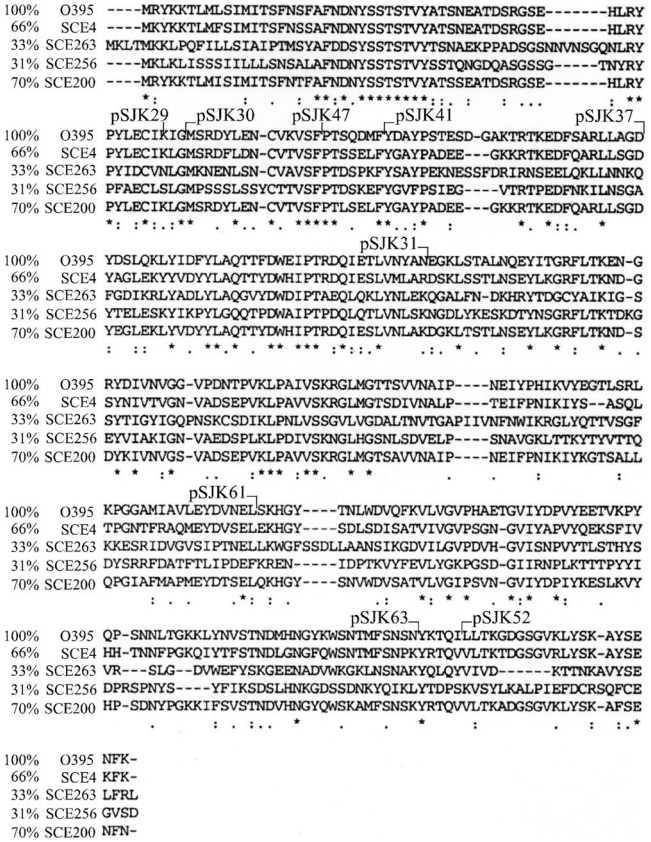

FIG. 5. Alignment of the pathogenic TcpF protein with divergent environmental TcpF proteins. Clustal W alignment reveals residues that are identical (*), conserved (:), and semiconserved (.) between the pathogenic TcpF protein and environmental TcpF proteins. Plasmids encoding linker scan protein derivatives that were found to influence secretion are annotated above the alignment with respect to the location of the 5-amino-acid insertion. The level of identity between each environmental TcpF protein and pathogenic TcpF, as reported by Mukhopadhyay et al. in 2001 (26), is given on the left. (Adapted from reference 26 with permission.)

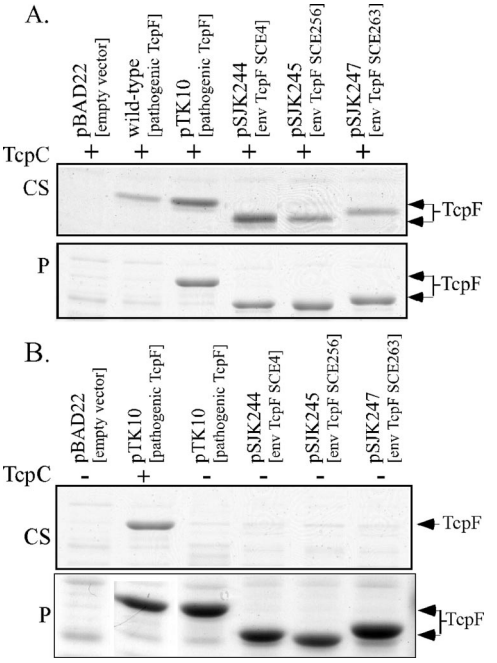

FIG. 6. Secretion of environmental TcpF proteins by the pathogenic strain. (A) Coomassie brilliant blue staining of the culture supernatant (CS) and periplasmic (P) fractions revealed that the pathogenic strain mediates the secretion of the environmental TcpF proteins. (B) In the absence of TcpC, the TCP apparatus secretin, these proteins are no longer detected in the culture supernatant, indicating that this secretion is dependent on the TCP apparatus.

ferred to below as "pathogenic TcpF," and TcpF from environmental strains is referred to below as "environmental TcpF." Alignment of these environmental TcpF proteins with the pathogenic TcpF protein revealed multiple regions throughout the protein that were conserved among the strains, and the N-terminal region was more conserved than the C-terminal region (Fig. 5). Due to the regional similarities of these environmental TcpF proteins to the pathogenic TcpF protein, we decided to utilize these natural derivatives of TcpF to further define secretion and function determinants within the highly conserved pathogenic TcpF by expressing the environmental TcpF proteins in the pathogenic strain and testing them for secretion and function.

The three environmental *tcpF* genes with the least identity at the protein level to pathogenic TcpF (SCE4, SCE256, and SCE263, with 66%, 33%, and 31% identity, respectively), were cloned into pBAD22 and expressed in the pathogenic strain lacking endogenous *tcpF* (RT4372). The presence of a restriction enzyme site in SCE200 (with 70% identity to the pathogenic strain) prevented further analysis of *tcpF* from this strain.

Each of the environmental TcpF proteins examined was secreted from the pathogenic strain, as revealed by Coomassie brilliant blue staining of the culture supernatant and periplasmic fractions (Fig. 6A). None of the environmental TcpF proteins accumulated in the periplasm to any greater degree than pathogenic TcpF encoded on plasmid pTK10, indicating that the environmental TcpF proteins are all secreted with the same efficiency as pathogenic TcpF (Fig. 6A). Secretion of these environmental TcpF proteins is also specific to the TCP apparatus, since a strain lacking the secretin, TcpC, could not mediate the secretion of these proteins (Fig. 6B). Consistent with this, several of the TcpF linker scan derivatives encoded on pSJK29, pSJK30, pSJK47, pSJK41, pSJK61 and pSJK63, which were found to have an influence on secretion, disrupted some of these conserved regions or conserved residues (Fig. 5).

Internal deletions were also constructed throughout *tcpF* based on the locations of the conserved regions and the linker scan data (Fig. 1). These deletions were recombined into the chromosome by allelic exchange. As seen previously with the linker scanning mutagenesis, deletions in the C terminus that were constructed in SJK216 and SJK218 (internal deletions 6 and 7) resulted in unstable proteins, suggesting that the C terminus is required for the structural stability of the protein (data not shown). Internal deletions 2 through 5, constructed throughout the gene, also resulted in unstable proteins (data not shown); however, internal deletion 1 generated a stable protein that was detected in the periplasm (Fig. 3). Internal deletion 1 removed a region on the N terminus rich in the amino acids tyrosine, serine, and threonine, all of which con-

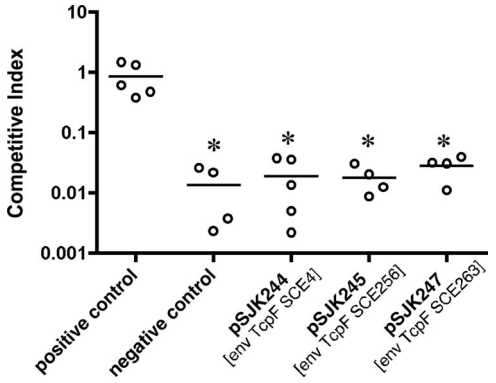

FIG. 7. Environmental TcpF proteins cannot mediate colonization in the infant mouse cholera model. The competitive index was determined for each environmental TcpF variant by competing each strain against the reference strain, SJK9, expressing wild-type TcpF. Each circle represents a mouse, and each horizontal line marks the average competitive index of the five or six mice tested for one strain. The positive control is the competition between SJK9 and SJK7, both expressing wild-type TcpF. The negative control is the competition between SJK9, expressing wild type TcpF, and SJK71, harboring the empty vector. The competitive index was determined for each mouse, and Student's *t* test was used to determine which strains were significantly defective in colonization compared to the positive control. Environmental TcpF proteins with various levels of divergence from pathogenic TcpF were found to be significantly defective in colonization compared to the positive control, as indicated by asterisks ($P < 0.016$).

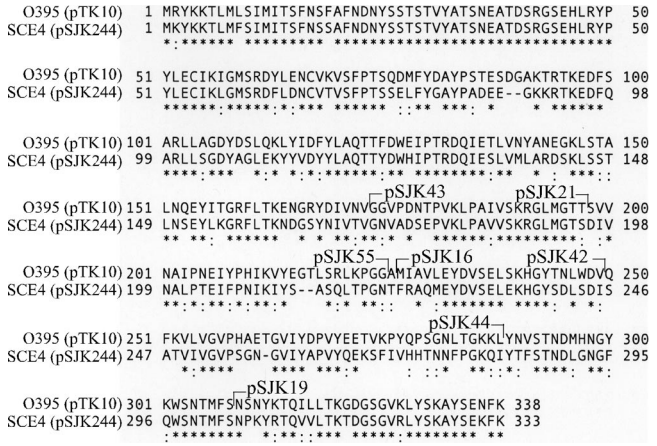

FIG. 8. Clustal W alignment of SCE4 environmental TcpF, which is 66% identical to pathogenic TcpF, reveals regions that are divergent between these proteins. TcpF derivatives that were found to be defective in colonization (Fig. 2) are represented by the designations of the plasmids encoding them, given above the alignment at the location of the 5-amino-acid insertion. Residues in the SCE environmental TcpF protein that are identical (*), conserved (:), or semiconserved (.) with respect to the pathogenic TcpF protein are indicated below the sequence.

tain hydroxyl side chains. Alignment of the environmental TcpF proteins with the pathogenic TcpF protein revealed that this region rich in hydroxyl residues is the most highly conserved among all these strains (Fig. 5). Western blotting revealed that this protein derivative accumulated in the periplasm to a greater extent than the wild type and was not detected in the culture supernatant or the outer membrane (Fig. 3), indicating that a region within the N terminus is important for the secretion of the protein. Taken together, these data suggest that multiple regions throughout TcpF are important for recognition by the TCP apparatus and for the ultimate secretion of this protein.

Lack of colonization mediated by environmental TcpF proteins suggests that regions divergent between the pathogenic and environmental proteins are important for TcpF function. Since these environmental proteins are secreted from the pathogenic strain, we wanted to determine if they could mediate the colonization of the infant mouse. Positive and negative controls were performed as described in the legend to Fig. 2. Each of the strains expressing environmental TcpF proteins was competed against SJK9, expressing pathogenic TcpF, and tested for its ability to colonize five infant mice. As seen in Fig. 7, each strain producing an environmental TcpF protein was significantly defective in colonization compared to the positive control (SJK7) ($P < 0.016$). The amount of environmental protein in the culture supernatants of the strains carrying pSJK245 and pSJK247 was slightly smaller than the amount of protein in the culture supernatant of the reference strain (Fig. 6A), which could account for their defect in colonization. However, the strain carrying pSJK244, which encodes a TcpF derivative with the highest identity to the pathogenic strain (66%), secretes quantities of TcpF similar to those secreted by the reference strain (Fig. 6A) and is still defective in colonization. Although the environmental TcpF proteins contain regions important for secretion, they must be divergent in regions important for function. Alignment of the TcpF protein of environmental strain SCE4 (which shares the highest identity with pathogenic TcpF), encoded on pSJK244, with pathogenic TcpF, encoded on pTK10, revealed regions divergent from one another that are localized mainly from the middle of the protein to the C-terminal region (Fig. 8). The strains carrying plasmids pSJK21, pSJK16, pSJK42, pSJK55, and pSJK19, which are defective in colonization (Fig. 2), express TcpF derivatives with 5-amino-acid insertions that disrupt, or are located near, regions divergent between the environmental and pathogenic TcpF proteins (Fig. 8). The correlation of these data sets provides further evidence suggesting that at least some of these regions define portions of the TcpF functional determinants with respect to mediating the colonization and pathogenesis of *V. cholerae*.

DISCUSSION

In this study, we sought to identify regions of the *V. cholerae* secreted virulence factor, TcpF, that are important for extracellular secretion and/or function in pathogenesis. We have provided genetic evidence that multiple regions dispersed throughout TcpF influence secretion by the TCP apparatus and that regions near the C terminus are important for mediating colonization. A summary of the locations of the linker scan derivatives highlighting these regions is given in Fig. 9. The findings presented here provide the foundation for more-refined mutagenesis and for future structural and immunological approaches to defining the mechanisms of secretion and function of TcpF.

To our knowledge, TcpF is the only protein secreted by the

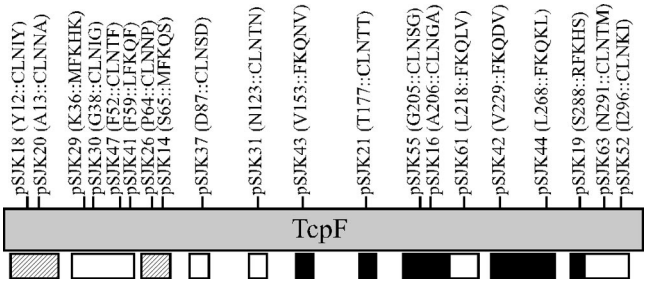

FIG. 9. Summary of TcpF secretion and colonization determinants. The shaded box represents TcpF after translocation through the Sec apparatus (mature form of the protein). TcpF derivatives generated by linker scanning mutagenesis are represented by the designations of the plasmids encoding them, given at the approximate location of the 5-amino-acid insertion. Open boxes represent regions found to be important for extracellular secretion. Striped boxes represent TcpF derivatives that are stable in the culture supernatant and are still able to mediate colonization. Filled boxes represent regions found to be important in mediating colonization.

TCP apparatus, aside from the polymerized pilin, and it represents the first nonpilus protein identified that is specifically secreted outside the bacterial cell by a type IV pilus biogenesis apparatus. Components of the type IV pilus apparatus and the T2SS have high homology to one another (29), and secretion of TcpF further demonstrates the similarity between these two protein secretion systems. The best-characterized T2SSs are the *pul* system of *Klebsiella oxytoca*, which secretes pullulanase (34); the *xcp* system of *Pseudomonas aeruginosa*, which secretes exotoxin A (8, 24); and the *out* system of the *Erwinia* species, which secretes cellulases, pectate lyases, and endoglucanases (3, 32). *Vibrio cholerae* also has a well-characterized T2SS, the *eps* system, which secretes not only the cholera toxin but also chitinases, proteases, and attachment factors (9, 20, 39). Although these systems secrete several proteins using the same T2SS, none of the proteins share sequence homology or protein similarity, and no specific secretion signal common to all of them has been identified (32). In addition, secretion of these substrates is highly specific, as exemplified in the plant-pathogenic *Erwinia* species. Both *Erwinia chrysanthemi* and *Erwinia carotovora* export pectate lyases and cellulases; surprisingly, however, the two species are capable of exporting only their cognate proteins (12). *V. cholerae* has numerous proteins in the periplasm, but only a small fraction are known to be transported across the outer membrane into the extracellular space (9). The secretion systems, such as the *eps* or TCP system, must differentiate between the proteins and be able to recognize their cognate substrates for export. Taking this into consideration, there must be specific secretion signals that direct proteins for extracellular secretion to their cognate machineries, although the recognition mechanisms of those specific sequences may be similar for the different systems.

As an initial approach to identifying the secretion and functional determinants of TcpF, we performed linker scanning mutagenesis of *tcpF*. Analysis of the resulting TcpF derivatives revealed multiple regions of the protein that influence secretion. Multiple regions involved in substrate recognition have also been identified in the T2SS substrates *Pseudomonas aeruginosa* exotoxin A (24, 25), *K. oxytoca* PulA (10), and *Erwinia carotovora* PehA (32). More specifically, the linker scanning analysis revealed a cluster of TcpF derivatives located in the N terminus (encoded on pSJK29, pSJK30, pSJK47, and pSJK41), a cluster in the C terminus (encoded on SJK63 and SJK52), and three other derivatives scattered throughout the protein (encoded on SJK37, SJK31, and SJK61) that influence secretion through the apparatus. Some of the TcpF linker derivatives accumulated in the outer membrane even in the absence of the outer membrane secretin, TcpC. These derivatives could be interacting with other components of the TCP apparatus, or alternative factors besides the TCP biogenesis components could play a role in the secretion of TcpF. The TCP apparatus is necessary for the secretion of TcpF (19); however, the possibility remains that it may not be sufficient. Another possibility is that since there are multiple regions influencing secretion, there could also be multiple steps in export across the outer membrane, and these multiple regions might be recognized by various protein "checkpoints" during the secretion process. Clearly, there may be additional conformational secretion domains that would not be recognized by this analysis. Some of the insertions within these secretion-defective derivatives disrupt regions of residues that are conserved between the pathogenic and environmental TcpF proteins. Chromosomal internal deletion 1 also demonstrated that the N terminus is important for secretion through the apparatus, since the resultant protein accumulates in the periplasm. This deletion removes a region rich in serine, tyrosine, and threonine residues with hydroxyl side chains that is also conserved among the environmental TcpF proteins. Since these environmental proteins can be secreted by the pathogenic strain, this finding suggests the importance of these regions in extracellular secretion.

TcpF is absolutely required for colonization and *V. cholerae* pathogenesis, and *tcpF* mutants clear the infant mouse intestine with kinetics essentially identical to those of a pilin mutant (19). Antibodies directed against TcpF are protective in the infant mouse model (19), demonstrating the potential of this protein to be a target for a multivalent vaccine. The mechanisms by which TcpF mediates colonization are currently unknown; however, we hypothesize that they are similar to ETEC CofJ and *C. rodentium* CfcJ. Each of these proteins has a conventional Sec-dependent signal sequence and is similar in size to TcpF (19, 27, 45). There are no published data regarding the localization of CfcJ at this time; however, there are many striking similarities between CofJ and TcpF. For instance, both contain Sec signal sequences and are secreted extracellularly by their corresponding type IV pilus biogenesis machineries (19; T. Taniguchi, personal communication). Although the exact functions of TcpF and CofJ are currently unknown, colonization of ETEC CFA/III requires CofJ (T. Taniguchi, personal communication), just as *V. cholerae* requires TcpF for colonization (19). Due to the similar locations of these genes within their operons, the secretion of these proteins by their cognate type IV apparatuses, and their requirement for colonization, it is likely that our findings with TcpF will be applicable to the functions of CofJ as well. The linker scan studies, along with the use of the variant environmental TcpF proteins, revealed that at least some of the functional determinants of TcpF are located in regions near the C-terminal end of the protein. We propose that the function of

TcpF requires these determinants near the C terminus to be in a folded, 3-dimensional structure. Three-dimensional similarities could exist between domains of TcpF, CofJ, and CfcJ that are important for virulence, just as 3-dimensional similarities must exist between proteins secreted by the same T2SS, although they share no sequence homology. Therefore, investigation of the role of TcpF in colonization not only may help us understand the pathogenesis of *V. cholerae* but also could aid in understanding the pathogenesis of other organisms, such as ETEC.

The molecular basis of protein secretion across the bacterial outer membrane is poorly understood, although it is critical for the pathogenesis of gram-negative bacteria. It is becoming increasingly obvious that secretion determinants of T2SS are recognized by a folded, 3-dimensional conformation. We investigated the regions important for the secretion of TcpF through the TCP apparatus, mimicking a T2SS, as well as regions important for TcpF-mediated colonization, through genetic analysis. The structure of TcpF, in addition to the genetic studies presented here, will allow us to dissect the molecular interactions involved in recognition by the TCP apparatus and will also provide further information about the regions involved in colonization. We hope that these studies will provide insight into the molecular mechanisms of secretion by type IV pilus apparatuses and factors that influence colonization so that we can better understand these virulence requirements of pathogenic bacteria.

ACKNOWLEDGMENTS

We are grateful to Douglas Berg for the generous gift of the environmental strains. We also thank Lisa Craig for insightful discussions and Jennifer Day for help in constructing the internal deletions.

This work was supported by NIH grant AI25096 to R.K.T. S.J.K. was also supported by NIH training grant T32GM08704.

REFERENCES

1. **Aiuchi, T., S. Nakajo, and K. Nakaya.** 2004. Reducing activity of colloidal platinum nanoparticles for hydrogen peroxide, 2,2-diphenyl-1-picrylhydrazyl radical and 2,6-dichlorophenol indophenol. Biol. Pharm. Bull. **27:**736–738.
2. **Bose, N., and R. K. Taylor.** 2005. Identification of a TcpC-TcpQ outer membrane complex involved in the biogenesis of the toxin-coregulated pilus of *Vibrio cholerae*. J. Bacteriol. **187:**2225–2232.
3. **Chapon, V., M. Czjzek, M. El Hassouni, B. Py, M. Juy, and F. Barras.** 2001. Type II protein secretion in gram-negative pathogenic bacteria: the study of the structure/secretion relationships of the cellulase Cel5 (formerly EGZ) from *Erwinia chrysanthemi*. J. Mol. Biol. **310:**1055–1066.
4. **Chapon, V., H. D. Simpson, X. Morelli, E. Brun, and F. Barras.** 2000. Alteration of a single tryptophan residue of the cellulose-binding domain blocks secretion of the *Erwinia chrysanthemi* Cel5 cellulase (ex-EGZ) via the type II system. J. Mol. Biol. **303:**117–123.
5. **Clausen, T., C. Southan, and M. Ehrmann.** 2002. The HtrA family of proteases: implications for protein composition and cell fate. Mol. Cell **10:**443–455.
6. **Dow, J. M., M. J. Daniels, F. Dums, P. C. Turner, and C. Gough.** 1989. Genetic and biochemical analysis of protein export from *Xanthomonas campestris*. J. Cell Sci. Suppl. **11:**59–72.
7. **Filloux, A., M. Bally, G. Ball, M. Akrim, J. Tommassen, and A. Lazdunski.** 1990. Protein secretion in gram-negative bacteria: transport across the outer membrane involves common mechanisms in different bacteria. EMBO J. **9:**4323–4329.
8. **Filloux, A., M. Bally, M. Murgier, B. Wretlind, and A. Lazdunski.** 1989. Cloning of *xcp* genes located at the 55 min region of the chromosome and involved in protein secretion in *Pseudomonas aeruginosa*. Mol. Microbiol. **3:**261–265.
9. **Folster, J. P., and T. D. Connell.** 2002. The extracellular transport signal of the *Vibrio cholerae* endochitinase (ChiA) is a structural motif located between amino acids 75 and 555. J. Bacteriol. **184:**2225–2234.
10. **Francetić, O., and A. P. Pugsley.** 2005. Towards the identification of type II secretion signals in a nonacylated variant of pullulanase from *Klebsiella oxytoca*. J. Bacteriol. **187:**7045–7055.
11. **Hang, L., M. John, M. Asaduzzaman, E. A. Bridges, C. Vanderspurt, T. J. Kirn, R. K. Taylor, J. D. Hillman, A. Progulske-Fox, M. Handfield, E. T. Ryan, and S. B. Calderwood.** 2003. Use of in vivo-induced antigen technology (IVIAT) to identify genes uniquely expressed during human infection with *Vibrio cholerae*. Proc. Natl. Acad. Sci. USA **100:**8508–8513.
12. **He, S. Y., M. Lindeberg, A. K. Chatterjee, and A. Collmer.** 1991. Cloned *Erwinia chrysanthemi out* genes enable *Escherichia coli* to selectively secrete a diverse family of heterologous proteins to its milieu. Proc. Natl. Acad. Sci. USA **88:**1079–1083.
13. **Herrington, D. A., R. H. Hall, G. Losonsky, J. J. Mekalanos, R. K. Taylor, and M. M. Levine.** 1988. Toxin, toxin-coregulated pili, and the *toxR* regulon are essential for *Vibrio cholerae* pathogenesis in humans. J. Exp. Med. **168:**1487–1492.
14. **Hillar, A., L. Van Caeseele, and P. C. Loewen.** 1999. Intracellular location of catalase-peroxidase hydroperoxidase I of *Escherichia coli*. FEMS Microbiol. Lett. **170:**307–312.
15. **Hobbs, M., and J. S. Mattick.** 1993. Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. Mol. Microbiol. **10:**233–243.
16. **Isaac, D. D., J. S. Pinkner, S. J. Hultgren, and T. J. Silhavy.** 2005. The extracytoplasmic adaptor protein CpxP is degraded with substrate by DegP. Proc. Natl. Acad. Sci. USA **102:**17775–17779.
17. **Kaper, J. B., J. G. J. Morris, and M. M. Levine.** 1995. Cholera. Clin. Microbiol. Rev. **8:**48–86.
18. **Kim, D. Y., and K. K. Kim.** 2005. Structure and function of HtrA family proteins, the key players in protein quality control. J. Biochem. Mol. Biol. **38:**266–274.
19. **Kirn, T. J., N. Bose, and R. K. Taylor.** 2003. Secretion of a soluble colonization factor by the TCP type 4 pilus biogenesis pathway in *Vibrio cholerae*. Mol. Microbiol. **49:**81–92.
20. **Kirn, T. J., B. A. Jude, and R. K. Taylor.** 2005. A colonization factor links *Vibrio cholerae* environmental survival and human infection. Nature **438:**863–866.
21. **Kirn, T. J., M. J. Lafferty, C. M. Sandoe, and R. K. Taylor.** 2000. Delineation of pilin domains required for bacterial association into microcolonies and intestinal colonization by *Vibrio cholerae*. Mol. Microbiol. **35:**896–910.
22. **Kirn, T. J., and R. K. Taylor.** 2005. TcpF is a soluble colonization factor and protective antigen secreted by El Tor and classical O1 and O139 *Vibrio cholerae* serogroups. Infect. Immun. **73:**4461–4470.
23. **Lipinska, B., M. Zylicz, and C. Georgopoulos.** 1990. The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase. J. Bacteriol. **172:**1791–1797.
24. **Lu, H. M., and S. Lory.** 1996. A specific targeting domain in mature exotoxin A is required for its extracellular secretion from *Pseudomonas aeruginosa*. EMBO J. **15:**429–436.
25. **McVay, C. S., and A. N. Hamood.** 1995. Toxin A secretion in *Pseudomonas aeruginosa:* the role of the first 30 amino acids of the mature toxin. Mol. Gen. Genet. **249:**515–525.
26. **Mukhopadhyay, A. K., S. Chakraborty, Y. Takeda, G. B. Nair, and D. E. Berg.** 2001. Characterization of VPI pathogenicity island and CTXϕ prophage in environmental strains of *Vibrio cholerae*. J. Bacteriol. **183:**4737–4746.
27. **Mundy, R., D. Pickard, R. K. Wilson, C. P. Simmons, G. Dougan, and G. Frankel.** 2003. Identification of a novel type IV pilus gene cluster required for gastrointestinal colonization of *Citrobacter rodentium*. Mol. Microbiol. **48:**795–809.
28. **Nouwen, N., N. Ranson, H. Saibil, B. Wolpensinger, A. Engel, A. Ghazi, and A. P. Pugsley.** 1999. Secretin PulD: association with pilot PulS, structure, and ion-conducting channel formation. Proc. Natl. Acad. Sci. USA **96:**8173–8177.
29. **Nunn, D.** 1999. Bacterial type II protein export and pilus biogenesis: more than just homologies? Trends Cell Biol. **9:**402–408.
30. **Nye, M. B., J. D. Pfau, K. Skorupski, and R. K. Taylor.** 2000. *Vibrio cholerae* H-NS silences virulence gene expression at multiple steps in the ToxR regulatory cascade. J. Bacteriol. **182:**4295–4303.
31. **Pallen, M. J., and B. W. Wren.** 1997. The HtrA family of serine proteases. Mol. Microbiol. **26:**209–221.
32. **Palomäki, T., R. Pickersgill, R. Riekki, M. Romantschuk, and H. T. Saarilahti.** 2002. A putative three-dimensional targeting motif of polygalacturonase (PehA), a protein secreted through the type II (GSP) pathway in *Erwinia carotovora*. Mol. Microbiol. **43:**585–596.
33. **Pugsley, A. P.** 1993. The complete general secretory pathway in gram-negative bacteria. Microbiol. Rev. **57:**50–108.
34. **Pugsley, A. P., O. Francetić, K. Hardie, O. M. Possot, N. Sauvonnet, and A. Seydel.** 1997. Pullulanase: model protein substrate for the general secretory pathway of gram-negative bacteria. Folia Microbiol. (Praha) **42:**184–192.
35. **Pugsley, A. P., and O. Possot.** 1993. The general secretory pathway of *Klebsiella oxytoca*: no evidence for relocalization or assembly of pilin-like PulG protein into a multiprotein complex. Mol. Microbiol. **10:**665–674.
36. **Russel, M.** 1998. Macromolecular assembly and secretion across the bacterial cell envelope: type II protein secretion systems. J. Mol. Biol. **279:**485–499.

37. **Sandkvist, M.** 2001. Type II secretion and pathogenesis. Infect. Immun. **69:**3523–3535.
38. **Sandkvist, M., L. P. Hough, M. M. Bagdasarian, and M. Bagdasarian.** 1999. Direct interaction of the EpsL and EpsM proteins of the general secretion apparatus in *Vibrio cholerae*. J. Bacteriol. **181:**3129–3135.
39. **Sandkvist, M., L. O. Michel, L. P. Hough, V. M. Morales, M. Bagdasarian, M. Koomey, V. J. DiRita, and M. Bagdasarian.** 1997. General secretion pathway (*eps*) genes required for toxin secretion and outer membrane biogenesis in *Vibrio cholerae*. J. Bacteriol. **179:**6994–7003.
40. **Sauvonnet, N., G. Vignon, A. P. Pugsley, and P. Gounon.** 2000. Pilus formation and protein secretion by the same machinery in *Escherichia coli*. EMBO J. **19:**2221–2228.
41. **Schreiber, W., K. D. Stone, M. A. Strong, L. J. J. DeTolla, M. Hoppert, and M. S. Donnenberg.** 2002. BfpU, a soluble protein essential for type IV pilus biogenesis in enteropathogenic *Escherichia coli*. Microbiology **148:**2507–2518.
42. **Skórko-Glonek, J., A. Wawrzynow, K. Krzewski, K. Kurpierz, and B. Lipinska.** 1995. Site-directed mutagenesis of the HtrA (DegP) serine protease, whose proteolytic activity is indispensable for *Escherichia coli* survival at elevated temperatures. Gene **163:**47–52.
43. **Skorupski, K., and R. K. Taylor.** 1996. Positive selection vectors for allelic exchange. Gene **169:**47–52.
44. **Spiess, C., A. Beil, and M. Ehrmann.** 1999. A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. Cell **97:**339–347.
45. **Taniguchi, T., Y. Akeda, A. Haba, Y. Yasuda, K. Yamamoto, T. Honda, and K. Tochikubo.** 2001. Gene cluster for assembly of pilus colonization factor antigen III of enterotoxigenic *Escherichia coli*. Infect. Immun. **69:**5864–5873.
46. **Taylor, R. K., V. L. Miller, D. B. Furlong, and J. J. Mekalanos.** 1987. Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. Proc. Natl. Acad. Sci. USA **84:**2833–2837.
47. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
48. **Tripathi, S. A., and R. K. Taylor.** 2007. Membrane association and multimerization of TcpT, the cognate ATPase ortholog of the *Vibrio cholerae* toxin-coregulated-pilus biogenesis apparatus. J. Bacteriol. **189:**4401–4409.
49. **Vignon, G., R. Kohler, E. Larquet, S. Giroux, M. C. Prevost, P. Roux, and A. P. Pugsley.** 2003. Type IV-like pili formed by the type II secreton: specificity, composition, bundling, polar localization, and surface presentation of peptides. J. Bacteriol. **185:**3416–3428.