JOURNAL OF BACTERIOLOGY, Sept. 1996, p. 5188–5198 Vol. 178, No. 9
0021-9193/96/$04.00+0

Identification and Characterization of Additional Flagellin Genes from *Vibrio anguillarum*

KAREN MCGEE,[1] PER HÖRSTEDT,[2] AND DEBRA L. MILTON[1]*

Department of Cell and Molecular Biology[1] and the Department of Pathology,[2] Umeå University, S-901 87 Umeå, Sweden

Received 8 April 1996/Accepted 21 June 1996

Previously, the flagellar filament of *Vibrio anguillarum* was suggested to consist of flagellin A and three additional flagellin proteins, FlaB, -C, and -D. This study identifies the genes encoding FlaB, -C, and -D and a possible fifth flagellin gene that may encode FlaE. The flagellin genes map at two separate DNA loci and are most similar to the four polar flagellin genes of *Vibrio parahaemolyticus*, also located at two DNA loci. The genetic organization of these two loci is conserved between both organisms. For each gene, in-frame deletions of the entire gene, the 5′ end, and the 3′ end were made. Mutant analysis showed that each mutation, except those in *flaE*, caused a loss of flagellin from the filament. However, no obvious structural loss in the filament, as determined by electron microscopy, and only slight decreases in motility were seen. Virulence analysis indicated that all but two of the mutations gave a wild-type phenotype. The 5′-end deletions of *flaD* and *flaE* decreased virulence significantly (>10^4-fold) of infections via both the intraperitoneal and immersion routes. These results indicate that, like FlaA, FlaD and FlaE may also be involved in virulence.

For several bacteria, the flagellum has been suggested to be involved in virulence either as a motility organelle or as an organelle that carries an adhesive component. In *Campylobacter jejuni* studies, the flagellum has been suggested to aid the bacterium either in the adherence to (20) or in the internalization (7) within cultured epithelial cells. Additional studies with *C. jejuni* (32) have shown that motility and not flagellin A is required for the invasion of intestinal cells; however, flagellin A can serve as a secondary adhesin for the adherence to intestinal cultured cells. For *Pseudomonas aeruginosa*, a study utilizing isogenic motility mutants in the burned mouse model showed that motility contributes to the invasiveness of this organism (5). For *Vibrio cholerae*, earlier studies with motility mutants suggested that either motility (8) or an adhesin (2, 11), proposed to be associated with the flagellum, is needed for the colonization of intestinal tissues, whereas a later study (27) indicated that motility but not the flagellar structure is essential for the colonization of rabbits.

The polar flagellum of *Vibrio anguillarum*, the causative agent of vibriosis in marine fish, is a complex structure. The filament is believed to be composed of four flagellin proteins (23) and is covered by a sheath (25). Recently, studies have been initiated to analyze the role that the flagellum plays in the virulence of *V. anguillarum*. Chemotactic motility has been shown to be required for the invasion of rainbow trout when they are immersed in infected seawater; however, chemotactic motility is not essential for virulence after the fish integument is crossed (26). Moreover, flagellin A has been suggested to have two possible roles in virulence (23). A mutant carrying a deletion of the *flaA* gene was shown to be partially motile and to be defective in its ability to invade rainbow trout immersed in infected seawater; however, virulence was not decreased if the mutant was introduced into the fish by intraperitoneal injection. Thus, flagellin A is needed for the full motility which is essential for efficient invasion of rainbow trout. A second mutant that contained an in-frame deletion at the 3′ end of the *flaA* gene and that was also partially motile as a result of the loss of flagellin A in the filament was avirulent by intraperitoneal injection, indicating a potential second role for the flagellin protein. One proposal for this second role was that the truncated flagellin protein may remain in the cytoplasm because of the loss of export functions and thus may repress the expression of virulence genes.

To clarify the role of the flagellum and of the flagellin proteins in the virulence of *V. anguillarum*, the present study continues the characterization of the flagellar filament. Four additional flagellin genes, *flaB*, *-C*, *-D*, and *-E*, were cloned and sequenced. In-frame deletions of the full gene, the 5′ end, and the 3′ end were made in each gene similar to those of the *flaA* gene in a previous study (23). All mutations affected motility only slightly, if at all. Since motility was not strongly affected, virulence was only slightly decreased for all but two mutations, the *flaD* and *flaE* 5′-end deletions. For these two mutations, virulence was strongly decreased (10^4- to 10^5-fold) via both immersion and intraperitoneal infections.

MATERIALS AND METHODS

Bacterial strains, bacteriophage, plasmids, and media. *V. anguillarum* NB10 (serotype O1) is an isolate of our laboratory from the Gulf of Bothnia outside the Norrby Laboratory, Umeå, Sweden (24). The more competent *Escherichia coli* SY327 [Δ(*lac pro*) *argE*(Am) *rif malA recA56*] (21) was used for transformation after subcloning fragments into either pNQ705-1 or pDM4. *E. coli* S17-1 (*thi pro hsdR hsdM*$^+$ *recA* RP4-2-Tc::Mu-Km::Tn*7*) (29) was used as the donor strain for all plasmids transferred to *V. anguillarum* by a previously described method for bacterial mating (23). SY327 and S17-1 are bacteriophage lambda lysogens carrying the *pir* gene, which is needed for replication of the suicide vectors pNQ705-1 and pDM4. *E. coli* XL1-Blue (28) (*recA1 lac endA1 gyrA96 thi hsdR17 supE44 relA1* [F′ *proAB lacI*^{q} *lacZ*ΔM15 Tn*10*]) was used for bacteriophage lambda infections and for the transformation of plasmids not requiring the *pir* gene.

Construction of the *V. anguillarum* genomic library with the Lambda Zap II system from Stratagene, handling of this bacteriophage, and the excision of the pBluescript plasmid derivatives have been described previously (22). pBSFla8-1, carrying a 6.5-kb fragment with *flaA* and *flaC*, and pBSFlaD9-1, carrying a 7-kb fragment with *flaEDB*, were excised from isolated bacteriophage clones.

Plasmids used in this study are described in Table 1. The suicide vectors, pDM4 and pNQ705-1, were employed in making chromosomal deletions and insertions, respectively. pNQ705-1 is a derivative of the previously described pNQ705 (22) that was altered by exchanging the linker region of pNQ705 for the

* Corresponding author. Mailing address: Department of Cell and Molecular Biology, Umeå University, S-901 87 Umeå, Sweden. Phone: 46 90 10 25 36. Fax: 46 90 77 14 20. Electronic mail address: Tord .Hagervall@micro.umu.se.

TABLE 1. Plasmids used in this study

Plasmid[a]	Description	Reference
Parent		
pDM4	Cml^r, R6K origin requires *pir* gene, contains *sacBR* from *B. subtilis*	23
pNQ705-1	Cml^r, R6K origin requires *pir* gene	This study
flaB-containing		
pDMFlaB1	Contains a PCR fragment with the gene deletion (residues 2820 to 3047 and 4110 to 4358 [Fig. 3])	This study
pDMFlaB2	Contains a PCR fragment with a 5′-end deletion (residues 2820 to 3083 and 3264 to 3519 [Fig. 3])	This study
pDMFlaB3	Contains a PCR fragment with a 3′-end deletion (residues 3700 to 3956 and 4110 to 4358 [Fig. 3])	This study
pNQFlaB	Contains a PCR fragment with the *flaB* gene and its promoter (residues 2820 to 4358 [Fig. 3])	This study
flaC-containing		
pDMFlaC1	Contains a PCR fragment with the gene deletion (residues 1 to 255 and 1390 to 1644 [Fig. 2])	This study
pDMFlaC2	Contains a PCR fragment with a 5′-end deletion (residues 1 to 291 and 472 to 719 [Fig. 2])	This study
pDMFlaC3	Contains a PCR fragment with a 3′-end deletion (residues 960 to 1215 and 1369 to 1644 [Fig. 2])	This study
pNQFlaC4	Contains a PCR fragment downstream of *flaC* (residues 1485 to 1739 [Fig. 2])	This study
pNQFlaC	Contains a PCR fragment with the *flaC* gene and its promoter (residues 1 to 1644 [Fig. 2])	This study
flaD-containing		
pDMFlaD1	Contains a PCR fragment with a gene deletion (residues 1400 to 1628 and 2763 to 3000 [Fig. 3])	This study
pDMFlaD2	Contains a PCR fragment with a 5′-end deletion (residues 1400 to 1664 and 1845 to 2119 [Fig. 3])	This study
pDMFlaD3	Contains a PCR fragment with a 3′-end deletion (residues 2418 to 2588 and 2742 to 2899 [Fig. 3])	This study
pNQFlaD4	Contains a PCR fragment downstream of *flaD* (residues 2840 to 2930 [Fig. 3])	This study
pNQFlaD	Contains a PCR fragment with the *flaD* gene and its promoter (residues 1400 to 1628 [Fig. 3])	This study
flaE-containing		
pDMFlaE1	Contains a PCR fragment with the gene deletion (residues 21 to 251 and 1386 to 1600 [Fig. 3])	This study
pDMFlaE2	Contains a PCR fragment with a 5′-end deletion (residues 21 to 287 and 468 to 703 [Fig. 3])	This study
pDMFlaE3	Contains a PCR fragment with a 3′-end deletion (residues 981 to 1211 and 1365 to 1600 [Fig. 3])	This study
pNQFlaE4	Contains a PCR fragment downstream of *flaE* (residues 1417 to 1520 [Fig. 3])	This study
pNQFlaE	Contains a PCR fragment with the *flaE* gene and its promoter (residues 21 to 1600 [Fig. 3])	This study
ORF3-containing		
pNQORF3-1	Contains a PCR fragment from within ORF3 (residues 4310 to 4435 [Fig. 3])	This study

[a] The first two capital letters of the plasmid names indicate the parent plasmid (DM, pDM4; NQ, pNQ705-1).

linker region of pDM4 (23). For each mutation made, a PCR fragment of the chromosomal region to be altered was made and cloned into a suicide vector. For complementation of each mutation, pNQ705-1-derived plasmids containing a PCR fragment with a flagellin gene and its promoter were used. The complementary regions of the PCR fragments for each of these plasmids are given in Table 1.

The medium used routinely for *E. coli* was TYS broth (10 g of Bacto Tryptone, 5 g of Bacto Yeast Extract, and 10 g of sodium chloride [all per liter]), and for *V. anguillarum*, Trypticase soy medium (BBL) was used. The vibrio-selective medium TCBS (Difco) containing 5 μg of chloramphenicol per ml was used for the selection of *V. anguillarum* transconjugants when crosses were made with *E. coli*. Motility agar consisted of Trypticase soy broth plus 0.25% agar.

Antibiotic concentrations for all *E. coli* strains were as follows: ampicillin, 100 μg/ml; chloramphenicol, 25 μg/ml. For *V. anguillarum*, chloramphenicol was used at a concentration of 5 μg/ml.

DNA techniques. Unless otherwise stated, all conditions for the various DNA techniques were as described by Sambrook et al. (28). Oligonucleotides for primers were synthesized with an Applied Biosystems DNA-RNA synthesizer (model 394). Double-strand DNA sequencing was performed by the dideoxy chain termination method with T7 DNA polymerase (Pharmacia). Both strands of the *flaC* gene and the *flaEDB* genes were sequenced by primer walking from the previously sequenced *flaA* gene for *flaC* and from the 110-bp PCR fragment used to identify the *flaD* gene locus.

Screening the genomic library. Screening of the *V. anguillarum* genomic library was done as described previously (23) with one exception. For the *flaD*-containing bacteriophage, the temperature for prehybridization and hybridization was changed to 50°C, and after hybridization, the filters were washed at 50°C for 5 min with 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and 0.5% sodium dodecyl sulfate (SDS). The probe for *flaC* was a 700-bp fragment from the downstream region of the previously sequenced *flaA* gene (23). The probe for the *flaD* gene locus was a 75-bp fragment (residues 1762 to 1837 [see Fig. 3]) derived from a 110-bp PCR fragment from the conserved 5′ end of the *V. anguillarum flaD* gene. This fragment was generated with previously described (23) primers complementary to a similar region of the *P. aeruginosa flaA* gene and chromosomal DNA from a previously described *flaA* mutant, RO8, in which most of the *flaA* gene had been deleted (23). For detection, the probes were labelled with [α-^{32}P]dCTP as described previously (23).

PCR conditions. Cycling conditions for each PCR began with a 3-min denaturation at 94°C and five cycles of the following varying annealing times: 94°C for 30 s, 50 to 55°C (1 degree change each cycle) for 30 s, 72°C for 30 s. The PCR was completed with an additional 25 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s (on the last cycle, 72°C for 5 min). *Taq* DNA polymerase (Perkin-Elmer) was used for most reactions; however, *Pfu* DNA polymerase (Stratagene), a high-fidelity polymerase, was used when creating fragments for in-frame mutagenesis to decrease the number of mutations caused by the PCR. Template DNA was obtained from either a chromosomal DNA preparation (100 ng was used) (28), bacterial cells, or a PCR. Chromosomal DNA was obtained from a single bacterial colony by touching a sterile pipette tip to a colony and then mixing the cells into the PCR mixture. When PCR-generated fragments were used, the fragments were purified from a 1.5% agarose gel with the Sephaglas kit from Pharmacia, and 10 to 100 ng was added to the PCR mixture.

Construction of the flagellin gene mutations. The construction of in-frame deletions in *V. anguillarum* by allelic exchange has been described previously in detail (23). pDM4 derivatives that were utilized for the in-frame deletion mutations and the pNQ705-1 derivatives that were used in testing for polarity effects on downstream genes are listed in Table 1.

The deletion mutations were confirmed by cloning and sequencing PCR fragments containing the mutated regions of DNA. All PCR fragments utilized were free from PCR errors except that for pDMFlaD3, which carried a base pair substitution (guanine to adenine) at position 2485 (see Fig. 3). This substitution resulted in an amino acid substitution from glycine to glutamic acid. In addition, all mutant strains were checked for growth rate variability as described previously (23).

Motility measurements. Motility was measured by movement of bacterial cells through Trypticase soy broth containing 0.25% agar. The optical density at 600 nm was determined for overnight cultures of all strains used. Equal amounts of cells in 6 μl were spotted in the center of the plates, and movement away from the center was measured after 24 h of growth at room temperature.

SDS-PAGE of flagellar proteins. The isolation of crude flagella was done as described previously (23). For separation of the flagellar proteins, SDS–11% polyacrylamide gel electrophoresis (PAGE) as described by Lugtenberg et al. (18) was used. Protein concentrations were measured with a microbicinchoninic acid protein assay kit from Pierce. Six micrograms of crude flagellar protein was applied to a gel slot from each preparation. The gels were fixed and stained with 0.1% Coomassie brilliant blue in 40% methanol–10% acetic acid and then destained in 40% methanol–10% acetic acid.

Electron microscopy. To view the structure of the flagella of each mutant, negative stains were prepared. Overnight cultures of each strain were fixed with 0.1% glutaraldehyde in 0.2 M PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] buffer (pH 7) for 5 min. The cells were washed twice in 0.2 M PIPES buffer (pH 7). Formvar-coated grids were floated on drops containing the bacterial cells. The grids were then floated on a drop of 0.2 M PIPES buffer and moved to a drop of 1% uranyl acetate (pH 4.2) and a final drop of water. Specimens were examined with a Zeiss EM 109 transmission electron microscope operated at an

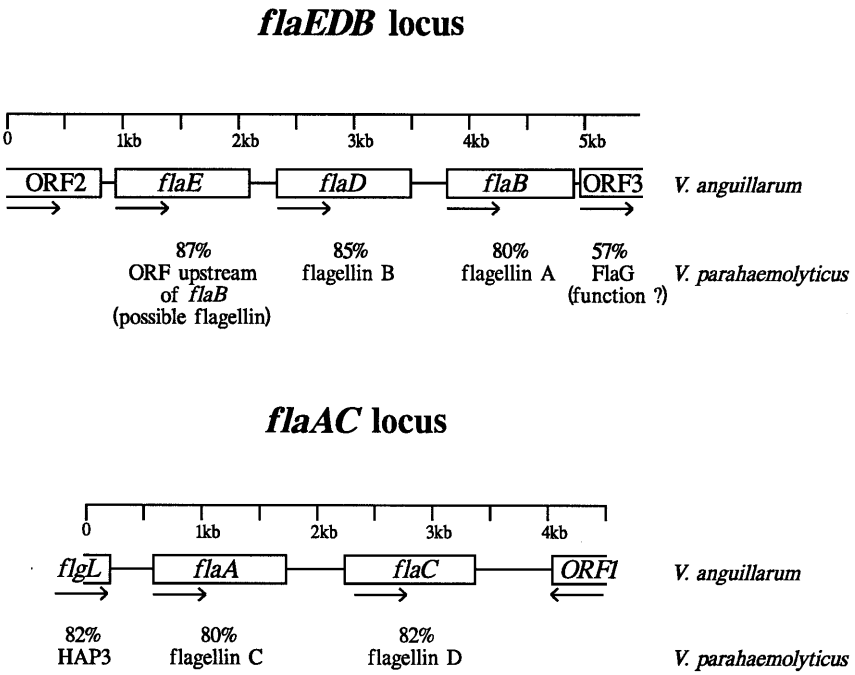


FIG. 1. Schematic diagram of the genetic organization of the flagellin gene loci of *V. anguillarum*. Flagellin genes and ORFs are indicated by open boxes for *V. anguillarum*. The homologous gene products from *V. parahaemolyticus* are listed below the respective homolog with the percent identity between the two gene products. Arrows indicate the direction of transcription.

accelerating voltage of 50 kV. All micrographs presented are a representative average of the flagellar length of at least 80% of the bacterial cells studied from two separate experiments.

Computer analysis. Data base searches were done with the Genetics Computer Group Sequence Analysis software (4) of the Genetics Computer Group, Inc. (University of Wisconsin).

Fish infections. Immersion and intraperitoneal infections were done as described previously (23).

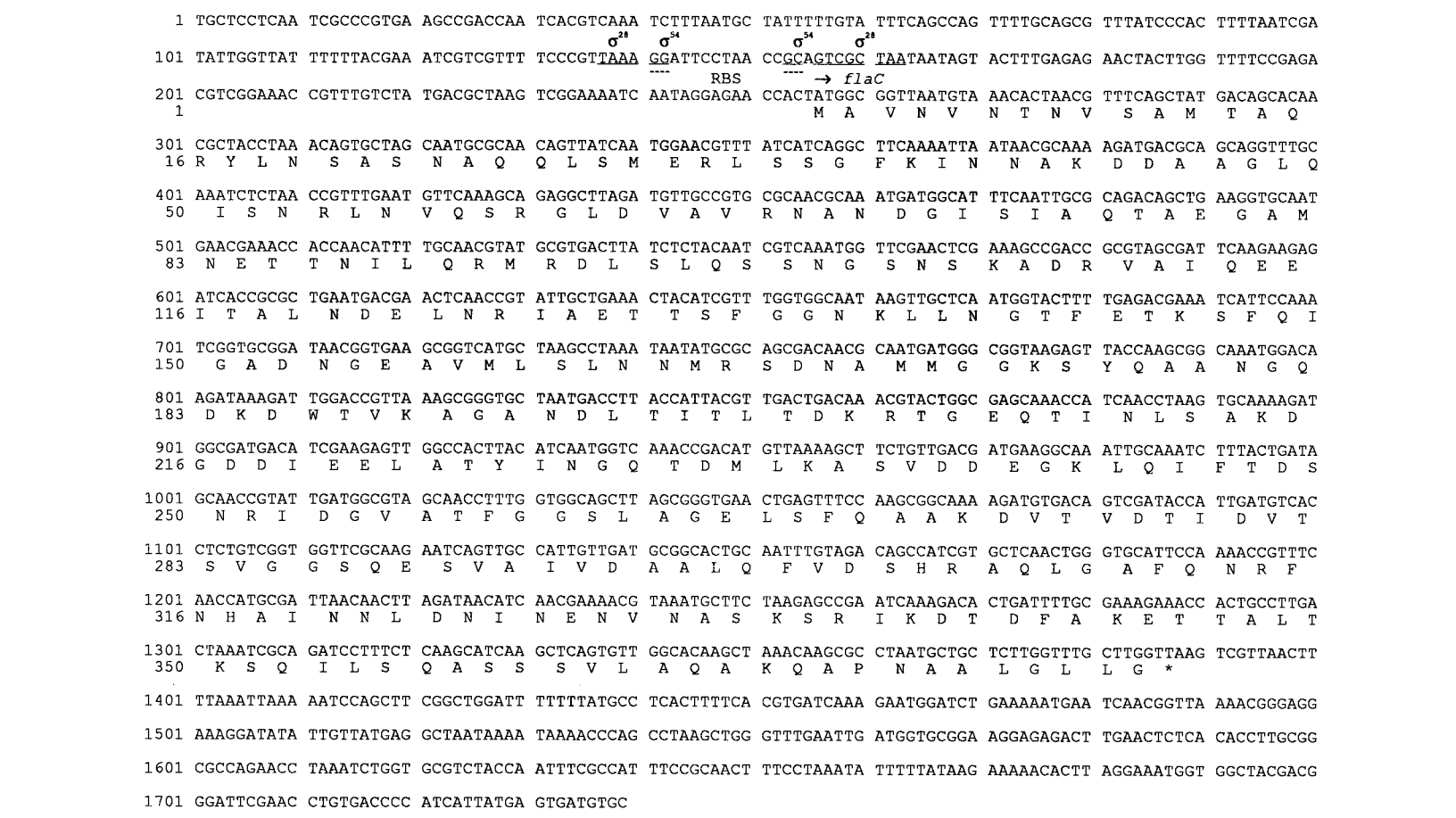
Nucleotide sequence accession numbers. The DNA sequences reported have been submitted to GenBank under accession numbers U52198 (for the *flaEDB* locus) and U52199 (for the *flaC* gene).

RESULTS

Cloning, sequencing, and characterization of the flagellin loci. Previously, protein analysis (23) indicated that the flagellum of *V. anguillarum* contains four flagellin proteins, FlaA, FlaB, FlaC, and FlaD. To date, only the *flaA* gene has been identified. To identify the additional flagellin genes, we rationalized that the predicted four flagellin genes of *V. anguillarum* may be located at two different loci on the chromosome as has recently been shown for the four *Vibrio parahaemolyticus* flagellin genes (19). If this is true, then there should be at least one flagellin gene in tandem to *flaA* and a second locus which needs to be identified.

To identify flagellin genes located at the same locus as that of the *flaA* gene, the DNA sequence upstream and downstream of *flaA* was reanalyzed (GenBank accession number L47122 [23]). No open reading frames (ORFs) or promoter sequences were found in the 350 bp downstream of *flaA*, but a partial ORF upstream was detected. This ORF encodes a 70-amino-acid polypeptide that shows 82% identity to the flagellin hook-associated protein 3 (HAP3) from *V. parahaemolyticus* (19) that is located upstream of *flaC*, the gene which has highest homology to *flaA* of *V. anguillarum* (23). This ORF will be designated *flgL*, like the gene which encodes HAP3 in *E. coli*. Therefore, a second flagellin gene may be located downstream of *flaA*, since *V. parahaemolyticus* has a flagellin gene, *flaD*, downstream of *flaC* (19). A clone was isolated, and sequencing was continued with sequences downstream of the *flaA* gene. A second flagellin gene was found, and the mutant analysis described below shows that this gene encodes *flaC*. Downstream of *flaC*, ORF1 (Fig. 1), which showed similarity to several putative GTP-binding proteins found in the GenBank, was found. The highest homology (78% identity over the entire ORF1) was to the 3′ end of a probable GTP-binding protein from *Haemophilus influenzae*. The nucleotide sequence and the deduced amino acid sequence of *flaC* are shown in Fig. 2.

Since no other flagellin gene was found in tandem to *flaA* and *flaC*, then *flaD* and *flaB* were likely to be located at another locus on the chromosome. A 75-bp PCR fragment from the N terminus of *flaD* (see Materials and Methods) was used as a probe to isolate a clone which contains a fragment with the *flaD* and *flaB* genes. The complete DNA sequence and the deduced amino acid sequence from this region are shown in Fig. 3. Surprisingly, instead of two flagellin genes, three possible flagellin genes were identified. The deduced N-terminal amino acid sequence from two of the genes was 100% identical to that previously determined for FlaD and FlaB (23). The third gene, *flaE*, possibly encodes an as-yet-unidentified flagellin protein that is highly homologous to the other flagellin genes. For the sake of comparisons in this study, we will consider *flaE* as a gene. ORF2, located upstream of *flaE*, showed no homology to other proteins in the GenBank, and ORF3, located downstream of *flaB* (Fig. 1), showed 57% identity to *flaG* from *V. parahaemolyticus* (19), which encodes a protein of unknown function. ORF3 also showed 26% identity to a potential 13-kDa protein of *Bacillus subtilis*, the gene for which is located in the intergenic region between the genes that encode for flagellin and HAP2, the flagellar cap protein. The genetic organization of both flagellin loci is shown in Fig. 1.




```
   1 TGCTCCTCAA TCGCCCGTGA AGCCGACCAA TCACGTCAAA TCTTTAATGC TATTTTTGTA TTTCAGCCAG TTTTGCAGCG TTTATCCCAC TTTTAATCGA
                                    σ28  σ54        σ54  σ28
 101 TATTGGTTAT TTTTTACGAA ATCGTCGTTT TCCCGTTAAA GGATTCCTAA CCGCAGTCGC TAATAATAGT ACTTTGAGAG AACTACTTGG TTTTCCGAGA
                                         ----  RBS   ----  → flaC
 201 CGTCGGAAAC CGTTTGTCTA TGACGCTAAG TCGGAAAATC AATAGGAGAA CCACTATGGC GGTTAATGTA AACACTAACG TTTCAGCTAT GACAGCACAA
   1                                                  M  A  V  N  V  N  T  N  V  S  A  M  T  A  Q

 301 CGCTACCTAA ACAGTGCTAG CAATGCGCAA CAGTTATCAA TGGAACGTTT ATCATCAGGC TTCAAAATTA ATAACGCAAA AGATGACGCA GCAGGTTTGC
  16 R  Y  L  N  S  A  S  N  A  Q  Q  L  S  M  E  R  L  S  S  G  F  K  I  N  N  A  K  D  D  A  A  G  L  Q

 401 AAATCTCTAA CCGTTTGAAT GTTCAAAGCA GAGGCTTAGA TGTTGCCGTG CGCAACGCAA ATGATGGCAT TTCAATTGCG CAGACAGCTG AAGGTGCAAT
  50  I  S  N  R  L  N  V  Q  S  R  G  L  D  V  A  V  R  N  A  N  D  G  I  S  I  A  Q  T  A  E  G  A  M

 501 GAACGAAACC ACCAACATTT TGCAACGTAT GCGTGACTTA TCTCTACAAT CGTCAAATGG TTCGAACTCG AAAGCCGACC GCGTAGCGAT TCAAGAAGAG
  83  N  E  T  T  N  I  L  Q  R  M  R  D  L  S  L  Q  S  S  N  G  S  N  S  K  A  D  R  V  A  I  Q  E  E

 601 ATCACCGCGC TGAATGACGA ACTCAACCGT ATTGCTGAAA CTACATCGTT TGGTGGCAAT AAGTTGCTCA ATGGTACTTT TGAGACGAAA TCATTCCAAA
 116 I  T  A  L  N  D  E  L  N  R  I  A  E  T  T  S  F  G  G  N  K  L  L  N  G  T  F  E  T  K  S  F  Q  I

 701 TCGGTGCGGA TAACGGTGAA GCGGTCATGC TAAGCCTAAA TAATATGCGC AGCGACAACG CAATGATGGG CGGTAAGAGT TACCAAGCGG CAAATGGACA
 150  G  A  D  N  G  E  A  V  M  L  S  L  N  N  M  R  S  D  N  A  M  M  G  G  K  S  Y  Q  A  A  N  G  Q

 801 AGATAAAGAT TGGACCGTTA AAGCGGGTGC TAATGACCTT ACCATTACGT TGACTGACAA ACGTACTGGC GAGCAAACCA TCAACCTAAG TGCAAAAGAT
 183  D  K  D  W  T  V  K  A  G  A  N  D  L  T  I  T  L  T  D  K  R  T  G  E  Q  T  I  N  L  S  A  K  D

 901 GGCGATGACA TCGAAGAGTT GGCCACTTAC ATCAATGGTC AAACCGACAT GTTAAAAGCT TCTGTTGACG ATGAAGGCAA ATTGCAAATC TTTACTGATA
 216 G  D  D  I  E  E  L  A  T  Y  I  N  G  Q  T  D  M  L  K  A  S  V  D  D  E  G  K  L  Q  I  F  T  D  S

1001 GCAACCGTAT TGATGGCGTA GCAACCTTTG GTGGCAGCTT AGCGGGTGAA CTGAGTTTCC AAGCGGCAAA AGATGTGACA GTCGATACCA TTGATGTCAC
 250  N  R  I  D  G  V  A  T  F  G  G  S  L  A  G  E  L  S  F  Q  A  A  K  D  V  T  V  D  T  I  D  V  T

1101 CTCTGTCGGT GGTTCGCAAG AATCAGTTGC CATTGTTGAT GCGGCACTGC AATTTGTAGA CAGCCATCGT GCTCAACTGG GTGCATTCCA AAACCGTTTC
 283  S  V  G  G  S  Q  E  S  V  A  I  V  D  A  A  L  Q  F  V  D  S  H  R  A  Q  L  G  A  F  Q  N  R  F

1201 AACCATGCGA TTAACAACTT AGATAACATC AACGAAAACG TAAATGCTTC TAAGAGCCGA ATCAAAGACA CTGATTTTGC GAAAGAAACC ACTGCCTTGA
 316 N  H  A  I  N  N  L  D  N  I  N  E  N  V  N  A  S  K  S  R  I  K  D  T  D  F  A  K  E  T  T  A  L  T

1301 CTAAATCGCA GATCCTTTCT CAAGCATCAA GCTCAGTGTT GGCACAAGCT AAACAAGCGC CTAATGCTGC TCTTGGTTTG CTTGGTTAAG TCGTTAACTT
 350  K  S  Q  I  L  S  Q  A  S  S  S  V  L  A  Q  A  K  Q  A  P  N  A  A  L  G  L  L  G  *

1401 TTAAATTAAA AATCCAGCTT CGGCTGGATT TTTTTATGCC TCACTTTTCA CGTGATCAAA GAATGGATCT GAAAAATGAA TCAACGGTTA AAACGGGAGG

1501 AAAGGATATA TTGTTATGAG GCTAATAAAA TAAAACCCAG CCTAAGCTGG GTTTGAATTG ATGGTGCGGA AGGAGAGACT TGAACTCTCA CACCTTGCGG

1601 CGCCAGAACC TAAATCTGGT GCGTCTACCA ATTTCGCCAT TTCCGCAACT TTCCTAAATA TTTTTATAAG AAAAACACTT AGGAAATGGT GGCTACGACG

1701 GGATTCGAAC CTGTGACCCC ATCATTATGA GTGATGTGC
```

FIG. 2. Nucleotide and amino acid sequence of the *flaC* gene. The consensus sequence of the potential σ^{28} promoter site is underlined with a solid line, and the possible σ^{54} promoter site is underlined with both a solid and a dashed line. Sites resembling *E. coli* ribosome binding sites (RBS) are indicated.

From the DNA sequence analysis, each flagellin gene appears to be a separate transcriptional unit. For each flagellin gene, a potential σ^{28} promoter sequence could be found, and for ORF3 and all of the flagellin genes, except *flaD*, a possible σ^{54} promoter could also be found. Figure 4 shows an alignment of the putative promoter sequences with the consensus sequence, and Fig. 2 and 3 indicate the locations of the suggested promoter sequences upstream of each gene.

The molecular weight of each flagellin protein was calculated from the predicted amino acid sequence, namely, 39,532 (FlaB), 40,099 (FlaC), 39,640 (FlaD), and 40,782 (FlaE). FlaA was previously predicted to have a molecular weight of 40,111 (23). Thus, all of the flagellin proteins are likely to be similar in size.

Protein sequence comparison. The predicted protein sequences from all five flagellin genes were aligned and compared for percent identity utilizing the Genetics Computer Group Bestfit analysis (Fig. 5 and Table 2). In general, all of the flagellin proteins are highly homologous at the amino and carboxy termini, while several of the flagellins have sequence homology within the central antigenic region. FlaC and FlaD are the most similar with 88% identity, and FlaB has the most similarity with FlaC and FlaD, with 72% identity. FlaA and FlaE, on the other hand, are the least similar to each other and to the other flagellin proteins, with identities of approximately 67 and 63%, respectively.

The flagellin genes were found in a GenBank search to be most similar to the flagellin genes from *V. parahaemolyticus* (19). To determine if the genetic organization as well as the protein sequence is conserved between the two species, the Genetics Computer Group Bestfit analysis was done. This analysis is schematically represented in Fig. 1 with each *V. parahaemolyticus* gene aligned below the *V. anguillarum* gene of highest identity. From this analysis, the genetic organization of the *V. anguillarum* genes *flaDB* and *flaAC* is conserved with that for the *V. parahaemolyticus* genes *flaBA* and *flaCD*, respectively. Interestingly, when the sequences flanking the flagellin genes of *V. parahaemolyticus* were compared with those of *V. anguillarum*, further conservation of genetic order was seen, and additional flagellin genes are proposed for *V. parahaemolyticus*. Upstream of *flaD* for *V. anguillarum*, there is *flaE* and ORF2. ORF2 has no homology to other proteins in the GenBank. However, the carboxy terminus of FlaE showed 87% identity to a partial ORF (57 amino acids) upstream of the *V. parahaemolyticus flaB* gene. Downstream of the *V. anguillarum flaC* gene is ORF1, which has sequence homology to putative GTP-binding proteins. A homology to ORF1 was not found downstream of the *V. parahaemolyticus flaD* gene; however, a partial ORF encoding 45 amino acids that has sequence homology to the four previously characterized (19) *V. parahaemolyticus* polar flagellin proteins was found. In conclusion, these data suggest that *V. anguillarum* contains at least four, and possibly five, flagellin genes, that *V. parahaemolyticus* contains possibly six flagellin genes, two not previously described,

```
   1 TTGGAAAGTA TGTATGGTGT GGCGGACGGT AAAACAGACC AACTGCTGCG CTTTACCAAG GCCGTAACGG GTGCGTACTT TTTTGCCCCT TCAGCTGAGA
                                                           σ54          σ54                                  σ28
 101 TGCTGCAACA TCTTGAACTA AAGTAAGTTT CAATACTGAC CTACTATCGG ACCGAGCCTT GTGCTCGGTT TTTCTTTTTT AAGCATTTTA TGTAAAGAAA
                σ28                               RBS            → flaE
 201 AGCGTCTACT GTCCGATAAA GAGATCATTC ATTCTATGGC ATGGGTCTGT TATGGCAATT ACCGTTAATA CCAATGTTTC TGCTTTGATT GCGCAGCGCC
   1                                                        M  A  I   T  V  N  T   N  V  S   A  L  I   A  Q  R  H

 301 ACTTAGGCTC AGCAAGCGAA ATGTTGAATC AATCTTTGGA GCGTCTTGCT TCCGGGAAGC GTATTAATAG TGCCAAAGAT GATGCTGCAG GTTTACAAAT
  18   L  G  S   A  S  E   M  L  N  Q   S  L  E   R  L  A   S  G  K  R   I  N  S   A  K  D   D  A  A  G   L  Q  I

 401 TTCGAATCGC CTTGAGTCGC AAATGAGGGG CTTAGATGTC GCTGTACGCA ATGCTAATGA TGGCATTTCA ATCATGCAGA CTGCTGAAGG TGCAATGCAA
  51  S  N  R   L  E  S  Q   M  R  G   L  D  V   A  V  R  N   A  N  D   G  I  S   I  M  Q  T   A  E  G   A  M  Q

 501 GAGACCACCA GCCTGCTTCA GCGTATGCGA GATCTCTCTT TGCAGTCGGC GAATGGCGCG AACAGTAAAG CGGATAGGCA AGCTTTACAG GAAGAAATGG
  84 E  T  T  S   L  L  Q   R  M  R   D  L  S  L   Q  S  A   N  G  A   N  S  K  A   D  R  Q   A  L  Q   E  E  M  G

 601 GCGCCTTGAA CGATGAATTA AACCGAATTG CTGAAACCAC CTCTTTTGGT GGTCGTAAGT TACTGAACGG TTCGTTTGGT CAATCCTCAT TTCAGATAGG
 118  A  L  N   D  E  L   N  R  I  A   E  T  T   S  F  G   G  R  K  L   L  N  G   S  F  G   Q  S  S  F   Q  I  G

 701 GGCGAGCTCT GGCGAAGCAG TGCAAGTTTC GTTAAAAAAT ATGCGTTCAG ACAGTTTGGA TATGGGGGGC TTTAGCTATG TTGCGGCAGG CATGGCGGAT
 151  A  S  S   G  E  A  V   Q  V  S   L  K  N   M  R  S  D   S  L  D   M  G  G   F  S  Y  V   A  A  G   M  A  D

 801 AGCCAGTGGC GGGTGACACA AGATAATCGG CAGCTCACCA TGAGTTATAC TGATGCCAAG GGAAAACAGC ACAACATTCA AATCCAAGCA AAAGTAGGGG
 184 S  Q  W  R   V  T  Q   D  N  R   Q  L  T  M   S  Y  T   D  A  K   G  K  Q  H   N  I  Q   I  Q  A   K  V  G  D

 901 ACGATATTGA AGAGTTGGCA ACCTACATCA ATGGGCAAAC CGATAAAGTA TCGGCGTCAG TCAACGACAA AGGACAGCTA CAGCTGTTCA TGGCGGGTAA
 218   D  I  E   E  L  A   T  Y  I  N   G  Q  T   D  K  V   S  A  S  V   N  D  K   G  Q  L   Q  L  F  M   A  G  K

1001 AGAAACCTCT GGGACTATTG ATTTTAAAGG TAGCTTAGCC AATCAGTTGC AAATGAATGT GACGGGATAT GAAGCAGTAG ACACTCTTGA TATTACCGAA
 251  E  T  S   G  T  I  D   F  K  G   S  L  A   N  Q  L  Q   M  N  V   T  G  Y   E  A  V  D   T  L  D   I  T  E

1101 GTCGGTGGTG CGCAGCGGGC GGTTGCCGTC ATTGATACCG CAATGCAGTA CGTTGATAGC CATCGCTCTG AGTTAGGCGC GTTGCAAAAC CGTTTTAACC
 284 V  G  G  A   Q  R  A   V  A  V   I  D  T  A   M  Q  Y   V  D  S   H  R  S  E   L  G  A   L  Q  N   R  F  N  H

1201 ACGCCATCAA CAACCTCGAC AACGTACATG AAAACCTCGC TTCTTCAAAC AGTCGAATTA AAGATACTGA CTATGCGAAA GAGACCACCC AAATGGTAAA
 318   A  I  N   N  L  D   N  V  H  E   N  L  A   S  S  N   S  R  I  K   D  T  D   Y  A  K   E  T  T  Q   M  V  K

1301 ACAGCAAATC CTACAACAAG TCAGTACCTC TATTCTTGCG CAAGCCAAAA AACAGCCAAA CCTAGCATTG GCCTTACTTC GCTAGCGATA AAAATCAACA
 351  Q  Q  I   L  Q  Q  V   S  T  S   I  L  A   Q  A  K  K   Q  P  N   L  A  L   A  L  L  R   *

1401 ATTATCGACC TTTATCAACA ACTCTTTAGC GAAATTATTA CCTTTTTGAT GGGTTTTTGA CCTCTTGTCC TGTTTTTTAC ACTAAACACA TTTTTTTTAA
                        σ28                    σ28
1501 GAAAATTGAA AAAAAATCCT CAAGCTTTGG AAAATACCGC CGTTATAAAA AGTAACTTTG AGAGAACTAT TTGGTTTTCC GAGACGTCGG AAACCGGAAA
                     RBS            → flaD.
1601 CATCGGAAAA TCAATTGGAG TAATCACCAT GGCAGTTAAT GTAAATACCA ACGTATCAGC GATGACAGCA CAGCGTTACC TAAACGGCGC TAGCAATGCA
   1                                 M  A  V  N   V  N  T  N   V  S  A   M  T  A   Q  R  Y  L   N  G  A   S  N  A

1701 CAACAAACAT CAATGGAACG CCTATCTTCT GGCTTTAAAA TCAACAGCGC AAAAGATGAT GCGGCAGGCC TACAAATCTC TAACCGTTTG AATGTTCAGA
  25 Q  Q  T  S   M  E  R   L  S  S   G  F  K  I   N  S  A   K  D  D   A  A  G  L   Q  I  S   N  R  L   N  V  Q  S

1801 GCCGCGGTCT TGACGTAGCA GTACGTAATG CGAATGACGG TATTTCTATT GCACAAACCG CTGAGGGTGC GATGAATGAA ACGACCAACA TTTTGCAACG
  59   R  G  L   D  V  A   V  R  N  A   N  D  G   I  S  I   A  Q  T  A   E  G  A   M  N  E   T  T  N  I   L  Q  R

1901 TATGCGTGAC TTGTCACTGC AATCAGCTAA CGGTTCAAAC TCAAAAGCAG ACCGTGTAGC GATTCAAGAA GAAGTCACCG CTCTGAATGA CGAGTTAAAC
  92  M  R  D   L  S  L  Q   S  A  N   G  S  N   S  K  A  D   R  V  A   I  Q  E   E  V  T  A   L  N  D   E  L  N

2001 CGTATCGCAG AAACCACATC ATTTGGTGGT AACAAATTAC TTAACGGTAC TTTTGAAACT AAGTCATTCC AAATCGGTGC GGATAACGGT GAAGCCGTCA
 125 R  I  A  E   T  T  S   F  G  G   N  K  L  L   N  G  T   F  E  T   K  S  F  Q   I  G  A   D  N  G   E  A  V  M

2101 TGCTTAGCCT AAACAACATG CGTAGTGACA ATGCTATGAT GGGCGGCACA AGCTACCAAG CGGCGAACGG CAAAGATAAA GATTGGTCAG TTCAAGCCGG
 159   L  S  L   N  N  M   R  S  D  N   A  M  M   G  G  T   S  Y  Q  A   A  N  G   K  D  K   D  W  S  V   Q  A  G

2201 CAGCAACGAT CTACAAATCA CACTGAAAGA TACTTCAGGC ACAGATCAGA CTATCAATAT CTCAGCTAAA GAAGGCGATG ATATTGAAGA GTTGGCGACT
 192  S  N  D   L  Q  I  T   L  K  D   T  S  G   T  D  Q  T   I  N  I   S  A  K   E  G  D  D   I  E  E   L  A  T

2301 TACATCAACG GCCAAACTGA TATGGTGAAA GCATCAGTTG ATGATGAAGG AAAACTACAA GTCTTCGCTG GCAGTAATAA AGTAGAAGGC CCTGTAACAT
 225 Y  I  N  G   Q  T  D   M  V  K   A  S  V  D   D  E  G   K  L  Q   V  F  A  G   S  N  K   V  E  G   P  V  T  F

2401 TCGCAGGCGG TCTTGCTGGT GAGCTCGGCA TGCAAGCTGG TCAAGCTGTT ACCGTAGATG TCATTGATGT GACCTCTGTG GGTGGCGCTC AAGAATCGGT
 259   A  G  G   L  A  G   E  L  G  M   Q  A  G   Q  A  V   T  V  D  V   I  D  V   T  S  V   G  G  A  Q   E  S  V

2501 TGCCATTGTT GATGCTGCAC TGCAATTTGT AGATAGCCAT CGCGCACAAT TAGGTGCATT CCAAAATCGC TTTAGCCATG CAATCAGCAA CTTAGATAAC
 292  A  I  V   D  A  A  L   Q  F  V   D  S  H   R  A  Q  L   G  A  F   Q  N  R   F  S  H  A   I  S  N   L  D  N

2601 ATTAACGAAA ACGTGAGTGC TTCTAAGAGC CGTATTAAAG ATACTGACTT TGCCAAAGAA ACAACAGCGC TTACTAAGTC GCAAATTCTT TCTCAAGCAT
 325 I  N  E  N   V  S  A   S  K  S   R  I  K  D   T  D  F   A  K  E   T  T  A  L   T  K  S   Q  I  L   S  Q  A  S

2701 CAAGCTCAGT GCTTGCTCAA GCAAAACAAG CGCCACAAGC GGCACTTAGC TTACTTGGCT AAAAGAGCCC AATAGCTCAA GAGTTACCCT TAATGAAAAT
 359   S  S  V   L  A  Q   A  K  Q  A   P  Q  A   A  L  S   L  L  G  *
```

FIG. 3. Nucleotide and amino acid sequence of the *flaEDB* genes. The consensus sequences of the potential σ^{28} and σ^{54} promoter sites are underlined. Sites resembling *E. coli* ribosome binding sites (RBS) are indicated.

```
2801 CCAGCCTCGG CTGGATTTTT TTGCTTAGCG GCAAAGTGGA AAAAGTGTTA GTCGATGCAG CATTATCTTT CTTACTTTTT GACGCTAATC ACATTTCTCA
                                                σ28                                 σ28
2901 TGCTCGCTGC ATAAATTTTA AAAAAAGACT TATTTTTTTC TAAAGGATAT AAATCTTTCG CCGTTAAAGG ATACGAGAGA AATGAGGTGT ACCGACGTGA
     σ54          σ54                              RBS          → flaB
3001 GGTGAGAGAC ACGGAAGTAC ATCAACACAA TGCCTAAGGA GATCAATATG GCAATTAATG TAAGCACTAA CGTGTCTGCA ATGACCGCAC AGCGTTATTT
   1                                                    M   A   I   N   V     S   T   N     V   S   A     M   T   A   Q     R   Y   L

3101 AAATAACGCC GCTGATGGTA CGCAGAAATC GATGGAGCGT CTGTCATCGG GCTATAAAAT CAATAGCGCG CGTGATGATG CAGCGGGTCT ACAAATCTCA
  19  N   N   A     A   D   G   T     Q   K   S     M   E   R     L   S   S   G     Y   K   I     N   S   A     R   D   D   A     A   G   L     Q   I   S

3201 AACCGTTTAA CGTCACAAAG CCGTGGCTTA GATATGGCGG TAAGAAACGC AAATGATGGT ATTTCCATTG CACAAACCGC AGAGGGTGCA ATGAATGAAA
  52 N   R   L   T     S   Q   S     R   G   L     D   M   A   V     R   N   A     N   D   G     I   S   I   A     Q   T   A     E   G   A     M   N   E   T

3301 CCACCAACAT CTTGCAACGT ATGCGAGATC TTTCTTTGCA ATCAGCAAAC GGTTCTAACT CAAGTTCAGA ACGCCAAGCG ATCCAAGAGG AAGTCTCAGC
  86    T   N   I     L   Q   R     M   R   D   L     S   L   Q     S   A   N     G   S   N   S     S   S   E     R   Q   A     I   Q   E   E     V   S   A

3401 ACTCAATGAT GAATTAAACC GTATTGCTGA AACAACCTCG TTTGGTGGTA ATAAACTGCT AAATGGTTCG TTTGGTAATA AATCTTTCCA AATTGGTGCC
 119  L   N   D     E   L   N   R     I   A   E     T   T   S     F   G   G   N     K   L   L     N   G   S     F   G   N   K     S   F   Q     I   G   A

3501 GATTCTGGTG AAGCTGTGAT GCTTTCCATG AGCGACATGC GTTCTGACAC AAAAGCGATG GGCGGTAAGA GCTATGTCGC GACAAATGGC AAAGCGCCTG
 152  D   S   G   E     A   V   M     L   S   M     S   D   M   R     S   D   T     K   A   M     G   G   K   S     Y   V   A     T   N   G     K   A   P   D

3601 ACTGGTCAGT CACCAATGCG ACGGATCTGA CTCTGAGCTA TACTGATAAA CAAGGAGAAG CTCGCGAAGT CACCATTAAC GCTAAAGCGG GTGATGACTT
 186    W   S   V     T   N   A     T   D   L     T     L   S   Y     T   D   K     Q   G   E   A     R   E   V     T   I   N     A   K   A   G     D   D   L

3701 AGAAGAGGTT GCTACTTACA TTAATGGTCA AAATGGTGAC ATCAAAGCTT CAGTCGGTGA TGAAGGTAAA CTGCAACTGT TTGCCGCCAA TCAAAAAGTC
 219  E   E   V     A   T   Y   I     N   G   Q     N   G   D     I   K   A   S     V   G   D     E   G   K     L   Q   L   F     A   A   N     Q   K   V

3801 TCAAGCGATG TCACTATTGG TGGTGGACTG GGTACTGAAA TTGGTTTTGC TGCAGGTAAA GACGTAACCG TCAAAGACAT CAATGTCACA ACGGTTGGCG
 252  S   S   D   V     T   I   G     G   G   L     G   T   E   I     G   F   A     A   G   K     D   V   T   V     K   D   I     N   V   T     T   V   G   G

3901 GCTCTCAGGA AGCTGTAGCT TTGATTGATG GTGCACTAAA GGCCGTTGAT AGCCAACGTG CCTCGTTAGG TGCATTCCAG AACCGTTTCG GTCATGCCAT
 286    S   Q   E     A   V   A     L   I   D   G     A   L   K     A   V   D     S   Q   R   A     S   L   G     A   F   Q     N   R   F   G     H   A   I

4001 CAGTAACCTT GATAACATCA ACGAGAACGT CAACGCTTCT CGTAGTCGAA TTAAGGATAC TGATTATGCC CGTGAAACAA CCCAAATGAC GAAGTCGCAG
 319  S   N   L     D   N   I   N     E   N   V     N   A   S     R   S   R   I     K   D   T     D   Y   A     R   E   T   T     Q   M   T     K   S   Q
                                  σ54                             σ54
4101 ATTTTGCAGC AAGCGAGTAC GTCAGTTTTG GCACAGGCGA AGCCAGTCACC ATCTGCAGCT CTTAGTCTAT TGGGTTAACC GATAGGCGAA TGGAGCTTTG
 352  I   L   Q   Q     A   S   T     S   V   L     A   Q   A   K     Q   S   P     S   A   A     L   S   L   L     G   *
                                                  RBS          → ORF3
4201 TGGGGAAAGC AGTGTTTTCA CTGCCAACCT ATGAGGTGGA AGGGAGATTG TTATGGAAAT ACCATCCTAC ACATCGAACA TCCAGCCTTA CGGCTCGCAA
   1                                                          M   E   I     P   S   Y     T   S   N   I     Q   P   Y     G   S   Q

4301 AGTGGCATTA AATTTGCTTC AGAAAACGAT GGCGCGACAC GCGCTTCCTC GAAACAAAAT GAGGTGAATC GTACTGAACA GTTAAGAAAT AGGCAAAACC
  17  S   G   I   K     F   A   S     E   N   D     G   A   T   R     A   S   S     K   Q   N     E   V   N   R     T   E   Q     L   R   N     R   Q   N   Q

4401 AAAGCGTTGA AGCGGCAATC GAGCTTGCAC AGCAACGAGA GCAGATCAAC AAATCGGAAA GAGCCAAGAT GGTAGAGCAG ATGAATGAAT TTATCTCATC
  51    S   V   E     A   A   I     E   L   A   Q     Q   R   E     Q   I   N     K   S   E   R     A   K   M     V   E   Q     M   N   E   F     I   S   S

4501 TATCAACAAA GACTTAGCGT TTAGAGTCGA TGAAGAATCA GGACGGGATG TTGTGACGAT TTATGAAGCC AGTACTGGTG ACATTATTCG TCAGATCCCG
  84  I   N   K     D   L   A   F     R   V   D     E   E   S     G   R   D   V     V   T   I     Y   E   A     S   T   G   D     I   I   R     Q   I   P

4601 AACGAGGAAA TGCTAGAGGT TTTGCGACGC CTTGCCAGAC AAAAAGACCA CAGTCGAAGG GGTTCG
 117  N   E   E   M     L   E   V     L   R   R     L   A   R   Q     K   D   H     S   R   R     G   S
```

FIG. 3—*Continued.*

and that the genetic organization of these two flagellin loci are conserved between these two species (Fig. 1).

Mutant construction. To analyze what possible functional and structural roles the flagellin genes have, mutagenesis was done in a manner similar to that described previously for *flaA* (23). For each gene, three different in-frame deletion mutations were made by allelic exchange. First, a full-gene deletion which took away all codons including the stop codon was made. For *flaB*, however, the last 66 bp were left since they contained a potential σ^{54} promoter for ORF3 (Fig. 3). The second mutation deleted 180 bp from the 5′ end of each gene. This N-terminal deletion was identical for all four flagellin proteins. The third mutation deleted 153 bp from the 3′ end of each gene. For three of the four flagellin proteins, this C-terminal deletion was identical. For *flaB*, this deletion was moved 66 bp upstream because of the potential overlapping promoter of ORF3. In addition, a plasmid insertion was made in the intergenic region downstream of each gene, except *flaB*, to test for potential polarity effects on downstream genes. Since the function of the ORF3 polypeptide has not been determined, a

Consensus σ^{54}	**CTGGYAYR**	**N_4**	**TTGCA**
flaA	TTGGCATA	N_4	TTGAA
flaB	GAGGTGAG	N_4	ACGGA
flaC	AAGGATTC	N_4	CCGCA
flaE	TCGGACCG	N_4	TTGTG
ORF3	TTGGCACA	N_4	AAGCA

Consensus σ^{28}	**TAAA**	**N_{15}**	**GCCGATAA**
flaA	CCAA	N_{15}	GCCGCTTC
flaB	TAAA	N_{15}	GCCGTTAA
flaC	TAAA	N_{15}	GTCGCTAA
flaD	TCAA	N_{15}	GCCGTTAT
flaE	TAAA	N_{15}	TCCGATAA

FIG. 4. DNA sequence alignment of the potential σ^{28} and σ^{54} promoter consensus sequences. The consensus sequence for each sigma factor is in bold type. The σ^{54} promoter consensus is derived from the transcription initiation sites of various NtrC- and NifA-activated promoters from several different species (9). The σ^{28} promoter consensus sequence is derived from *S. typhimurium* flagellar operons (15).

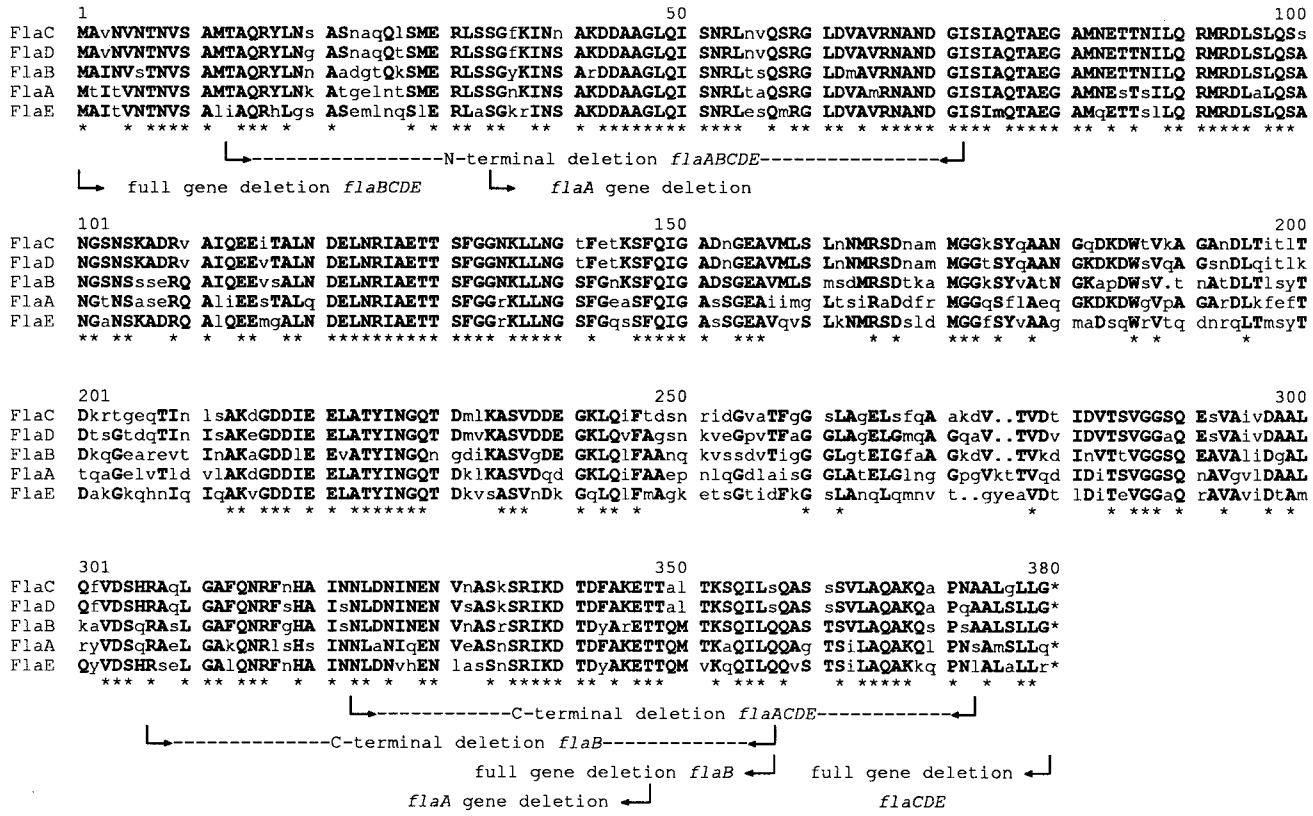


FIG. 5. Alignment of the deduced amino acid sequences of all five *V. anguillarum* flagellin proteins. Amino acids that are identical for three or more of the flagellin proteins are capitalized and in bold type, whereas those that are different are in lowercase letters. An asterisk indicates an amino acid that is identical for all five flagellin proteins. The arrows indicate regions deleted for the N-terminal deletions, the C-terminal deletions, and the full gene deletions of the *flaA*, *-B*, *-C*, *-D*, and *-E* genes.

plasmid insertion into the 5′ end of this ORF was made to assess the significance of this protein in motility. Figure 5 indicates the exact amino acids deleted for each protein, and Table 1 gives the base pairs that were deleted.

Since growth rate can affect both the motility and the virulence analyses, it was tested for each flagellin mutant and the polarity control strains. For all strains, the growth rate did not vary from that of the wild type.

Protein analysis of the flagellum. Flagellar preparations were made from each mutant to determine how each mutation affected the protein content of the flagellar filament. All mutations for each gene, except *flaE*, caused a loss of that particular flagellin protein in the flagellar filament as determined by SDS-PAGE. There are several explanations for this FlaE observation. *flaE* could be a silent gene or it could be expressed, but FlaE does not make up a part of the filament. On the other hand, FlaE could be a very minor part of the filament, and thus, the effect of the *flaE* mutations on the flagellar filament is not detectable with SDS-PAGE. Since mutations in *flaE* affect the virulence phenotype (see below), *flaE* is probably expressed; however, identification of the protein product remains to be determined. The filament protein profile for one mutant of each *fla* gene is given in Fig. 6. All intergenic plasmid insertions gave the same filament protein profile as that of the wild type (data not shown), indicating that there is no polar effect on downstream flagellar genes.

Interestingly, the molecular masses calculated from the predicted amino acid sequences of FlaB, -C, and -D (see above) differ from the sizes of these flagellin proteins determined from SDS-PAGE. FlaA, for which the predicted molecular mass agrees with the size determined from SDS-PAGE (23), is

TABLE 2. Percent identity between the flagellin protein sequences

Protein	% Identity				
	FlaA	FlaB	FlaC	FlaD	FlaE
FlaA	100	67	67	67	63
FlaB	67	100	72	73	63
FlaC	67	72	100	88	64
FlaD	67	73	88	100	64
FlaE	63	63	64	64	100

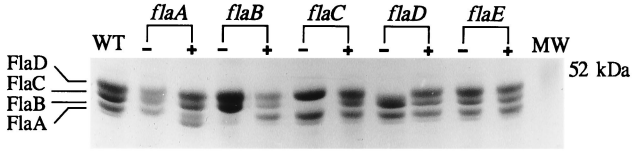


FIG. 6. Protein profiles of flagellar filament preparations. Flagella were prepared as described in Materials and Methods. Six micrograms of each flagellar preparation was electrophoresed through an SDS–11% polyacrylamide gel. A filament preparation from a *flaA* mutant was included for comparison since the FlaA and FlaB proteins migrate so close to one another. WT, a flagellar preparation from the wild type; MW, molecular mass marker; −, a full-gene deletion; +, complementation of the full-gene deletions by the insertion of a plasmid carrying the deleted flagellin gene into the chromosomal region where the gene was deleted. For the *flaA* gene, complementation of the deleted flagellin gene was done in *trans* with a plasmid carrying the *flaA* gene, pFlaA-2 (23).

TABLE 3. Virulence and motility analyses of flagellin mutants from *V. anguillarum*[a]

Strain	Mutation	Motility (% of wild type)		LD_{50} of mutant strain	
		Mutant strains	Mutant strains with the complementing plasmid	Intraperitoneal injection (no. of bacteria)	Immersion (no. of bacteria/ml)
NB10	None (wild type)	100		22	2×10^2
RO2	*flaA* 5′-end deletion	56	87	6	1×10^5
DM16	*flaA* 3′-end deletion	58	84	8×10^5	$>1 \times 10^7$
RO8	*flaA* gene deletion	53	87	41	6×10^5
KD19	*flaB* 5′-end deletion	94	103	28	2×10^2
KD21	*flaB* 3′-end deletion	94	97	65	2×10^2
KD17	*flaB* gene deletion	87	99	13	2×10^4
KD4	*flaC* 5′-end deletion	91	96	21	4×10^2
KD6	*flaC* 3′-end deletion	80	93	10	6×10^2
KD2	*flaC* full-gene deletion	96	97	45	4×10^3
KD7	*flaC* intergenic insertion	96	ND[b]	6	6×10^2
KD12	*flaD* 5′-end deletion	83	97	8×10^6	$>1 \times 10^7$
KD14	*flaD* 3′-end deletion	72	96	3×10^2	3×10^4
KD10	*flaD* full-gene deletion	87	100	30	1×10^4
KD15	*flaD* intergenic insertion	115	ND	26	3×10^3
KD27	*flaE* 5′-end deletion	92	93	8×10^5	$>1 \times 10^7$
KD29	*flaE* 3′-end deletion	95	101	5	1×10^4
KD25	*flaE* full-gene deletion	94	94	8	1×10^4
KD30	*flaE* intergenic insertion	97	ND	3	5×10^3
KD23	ORF3 plasmid insertion	89	ND	36	2×10^4

[a] Flagellin A results have been reported previously (23).
[b] ND, not determined.

calculated to be larger than FlaB, -C, and -D. However, all three of these proteins migrate slower in SDS-PAGE. One possible explanation for this difference is posttranslational modification of FlaB, -C, and -D.

Functional analysis. The functional analysis of the mutant flagella was done to ascertain the importance of each flagellin protein in maintaining wild-type motility. Light microscopy was utilized, and the motilities of all strains were similar to that of the wild type. In addition, each mutant was tested for its ability to swim in 0.25% Trypticase soy agar to attain some measure of the speed of movement. The results are given in Table 3. The loss of FlaB, FlaC, or FlaE had almost no effect on motility, whereas, the loss of FlaD had a slightly greater effect, resulting in 10 to 25% less motility than that of the wild type. The ORF3 mutant, KD23, showed an 11% decrease in the wild-type motility, indicating a potential role late in the regulatory cascade of flagellar biogenesis. Interestingly, a plasmid insertion in the intergenic region between *flaD* and *flaB* caused a slightly hypermotile effect.

Structural analysis. To observe any structural differences in the mutant flagellar filaments, electron microscopy was done. All mutants except the *flaB* full-gene deletion mutant KD17 and the ORF3 insertion mutant, KD23, appeared to have flagella of wild-type length. Surprisingly, KD17 and KD23 had very elongated flagella (Fig. 7). Since only the full-gene deletion and not the 5′- and 3′-end deletions of *flaB* caused flagellar elongation, it is likely that in the KD17 mutant, a region was deleted from the possible σ^{54} promoter of ORF3 which is essential for its function, such as an upstream activating sequence (30). A second explanation could be a polar effect on an unidentified flagellar gene downstream of ORF3.

Virulence analysis. Previously, FlaA was shown to be essential for the virulence of *V. anguillarum* (23) (virulence results are given in Table 3 for comparison). To determine if any of the other flagellin proteins are also essential for virulence, the 50% lethal dose (LD_{50}) for infection via the immersion and intraperitoneal routes in a rainbow trout model was determined for each mutant and for the polarity control strains containing plasmid insertions in the intergenic regions. The results are presented in Table 3. As expected, since motility was altered only slightly, all mutations, except the 5′-end deletions of *flaD* and *flaE*, showed only slight, if any, increases in the LD_{50} for infection via the immersion route. Because 10-fold fluctuations are often seen with the wild type via immersion infections, these small differences are probably not significant. However, 5′-end deletions of *flaD* and *flaE* showed significant increases in the LD_{50} (10^4- to 10^5-fold) via both routes of infection. Hence, FlaD and FlaE, like FlaA (23), may have a role in virulence unrelated to motility. Strains that contained plasmids inserted into the intergenic regions showed no effect on virulence, suggesting that the decreases seen with the virulence assay are not due to polarity effects on downstream virulence genes.

Since the flagellin genes are rather homologous to one another, there is the potential for these genes to recombine with one another to create a recombinant, active gene. Therefore, each flagellin mutant strain and the polarity control strains were isolated from the kidney of a fish which had died from vibriosis, usually 3 to 4 days after infection, and PCR was done to confirm that the original mutation was still present. In addition, motility was also tested on all strains isolated from the fish to ensure that suppression of the flagellin mutations did not occur during the infection. All strains had motility rates similar to those before infection (data not shown).

Complementation studies. To confirm that the effects observed with each mutation were due to that particular alteration, complementation studies were done. Since we have not found a plasmid that stably replicates in *V. anguillarum* during the infection process, complementation was achieved by subcloning each flagellin gene into a suicide vector and then inserting these plasmids into the chromosome at the locus of the mutated flagellin gene. Protein analysis of the flagellum showed that all mutants regained a wild-type protein profile of the flagellum when complemented with a wild-type copy of the

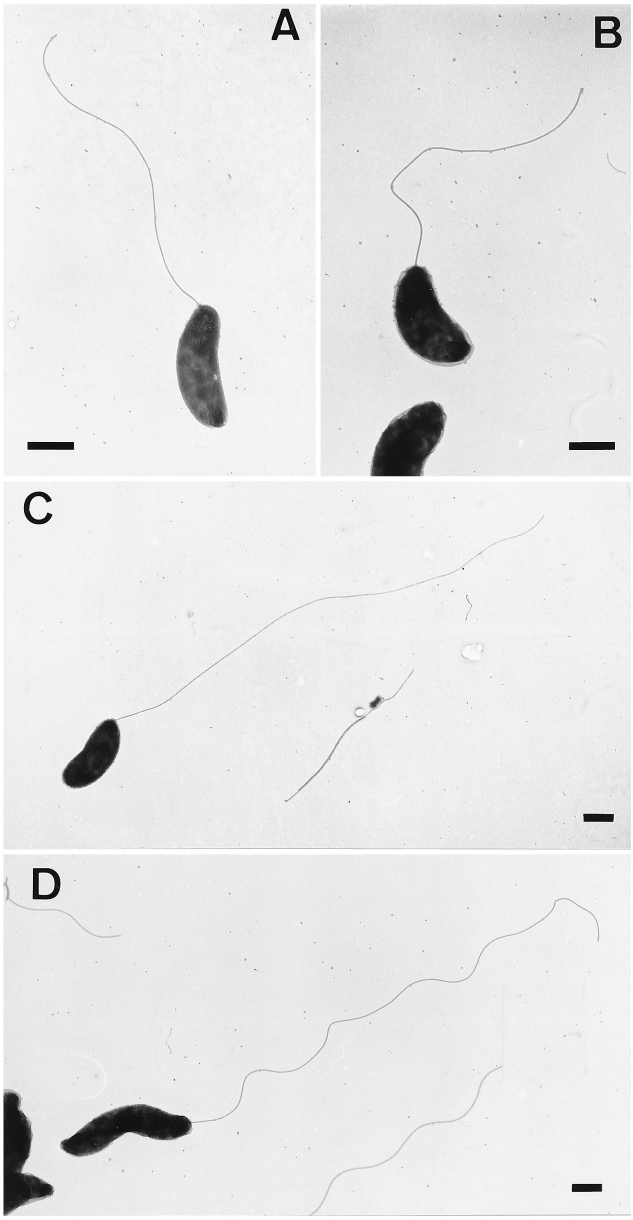


FIG. 7. Negative staining of the flagella from mutant strains with an elongated phenotype. Overnight cultures were fixed with 0.1% glutaraldehyde, spread onto Formvar-coated grids, and then negatively stained with 1% uranyl acetate (pH 4.2). (A) NB10, wild type; (B) KD17-B, the complemented *flaB* full-gene deletion mutant; (C) KD17, the *flaB* full-gene deletion mutant; (D) KD23, the ORF3 insertion mutant. Bars, 0.5 μm. All micrographs represent an approximate average flagellar length for at least 80% of the bacteria examined from two separate specimen preparations.

altered gene. One example for each gene set is shown in Fig. 6. For those mutants that showed a decrease in motility, the wild-type motility was regained as measured by the soft agar test (Table 3). Structural analysis (Fig. 7) showed that the *flaB* full-gene deletion mutant regained a normal-length flagellum, indicating that any effects on ORF3 or another unidentified gene due to the possible loss of promoter function were corrected by the gene insertion. Complementation of the 5′-end deletions of *flaD* and *flaE* did not give back the wild-type virulence phenotype. Since the stability of the plasmid could have an effect on the complementation, all complemented strains were reisolated from fish that had died of vibriosis (usually 3 to 4 days after infection) and then tested for chloramphenicol resistance, which was present on the inserted plasmid carrying complementing *fla* genes. One hundred colonies from each strain were tested, and 97% or more were still positive for the presence of the plasmid. One possible explanation for these negative complementation results for the virulence phenotypes could be that these mutations have a phenotype dominant to that of the wild type.

DISCUSSION

The animal model for *V. anguillarum* is ideal for studying the invasive mechanism of a bacterial pathogen since the assay for host invasion uses a natural route of infection. In other words, the fish are allowed to swim in seawater containing bacteria. The virulence of bacterial strains via this route of infection can be compared with the virulence via intraperitoneal infection to determine if there is a defect in the invasive capabilities of the bacterium. By the use of this model, the flagellum of *V. anguillarum* has been shown previously to be essential for crossing the integument of rainbow trout. In a previous study, we showed that chemotactic motility was needed for efficient invasion into the fish host, but once the fish integument was crossed, chemotactic motility was not required for the progression of vibriosis (26). Similarly, mutations in the *flaA* gene also showed a requirement for FlaA in the invasion of the fish host (23). Presumably, this is due to the decrease in motility that occurs with the loss of FlaA in the flagellar filament. Comparatively, the FlaB, -C, -D, and -E deletions had no significant effect on invasion of the fish host. However, motility was decreased only slightly, if at all, for the FlaBCDE mutants, thereby confirming the requirement for motility in invasion of the fish host.

In addition to its role in invading the host, FlaA was shown to have a second role once the fish integument was crossed (23). How the FlaA protein is involved in virulence at this level is unknown. However, it was suggested that, possibly, when the conserved C terminus was deleted from FlaA, a potential truncated protein that could not be exported from the cytoplasm to the flagellar filament was made. This truncated protein may then regulate virulence factors such that, even after an intraperitoneal injection, the strain carrying this mutation decreased virulence 10^4-fold. This second phenotype was also seen with the 5′-end deletion mutants of *flaD* and *flaE*, and this phenotype may be dominant to the wild-type phenotype since complementation of the mutant phenotype was not possible. Interestingly, these mutants carried a deletion at the amino terminus of the proteins instead of the carboxy terminus as for FlaA. In support of this hypothesis, the flagellum-specific export pathway has been suggested to involve the recognition of a structural conformation of the flagellar proteins at the secondary level or higher (10). The amino termini of two flagellar proteins, *E. coli* flagellin (16) and *Caulobacter* flagellar hook protein (14), have been shown to be essential for the secretion of these proteins. In addition, a recent study (3) identified a potential antisense RNA that is induced in vivo and that is transcribed from the 3′ end of a flagellin gene in *V. cholerae*. Although the biological significance of this antisense transcript is unknown, Camilli and Mekalanos (3) speculated that it may down-regulate flagellin synthesis. If this is true, then with further speculation the involvement of flagellin proteins in the regulation of virulence determinants could be explained in the following way. In vitro or outside the fish model, the flagellin antisense RNA is uninduced. This allows flagellar biosynthesis,

which is required for the motility needed to enter the fish host. Once *V. anguillarum* enters the fish host, motility is not required for the progression of vibriosis, and thus the transcription of antisense RNA molecules which repress the expression of FlaA, FlaD, and FlaE is induced. Once the flagellin genes are repressed, then virulence genes required for the progression of disease after the bacterium enters the fish are derepressed.

The analysis of the flagellar proteins of *V. anguillarum* by SDS-PAGE indicated that FlaA, -B, -C, and -D differed in molecular mass (i.e., 40, 41, 42, and 45 kDa, respectively), even though the DNA sequence predicted proteins that are almost identical in size (i.e., 40.1, 39.5, 40, and 39.6 kDa). This observation suggests posttranslational modification. Other organisms have been shown to have flagellar proteins that are posttranslationally modified. For example, the flagellar filaments of *P. aeruginosa* contain phosphorylated tyrosines (13), and those of *Campylobacter* spp. have been suggested to have phosphorylated serines (17). *Salmonella typhimurium* contains ε-*N*-methyllysine residues in its flagellar filaments (12), and halobacterial flagellins are sulfated glycoproteins (31). The function, if any, of the modification of the flagellin proteins remains to be determined.

Although the polar flagella of *V. anguillarum* and *V. parahaemolyticus* appear to be structurally very similar with this initial examination, there is one noticeable difference when comparing the results of the *V. parahaemolyticus* study (19) with those of this study and the previous *V. anguillarum* study (23). McCarter (19) reported that the loss of any single flagellin protein from the flagellar filament only slightly decreased the motility of *V. parahaemolyticus*. While the same is true if either FlaB, -C, -D, or -E is removed from the filament of *V. anguillarum*, loss of FlaA from the filament diminished motility to 50% of that for the wild type (23). Several possible explanations for this difference may be made. First, the flagellin localization within the filaments of the two species may be different. Thus, the structural role of the flagellin homologs may be different. Second, the two ORFs upstream of the *flaB* gene and downstream of the *flaD* gene in *V. parahaemolyticus* may code for two additional flagellin proteins that may structurally differentiate the flagellar filament of these two organisms. Clearly, localization of the flagellin proteins within the filament of *V. anguillarum* and *V. parahaemolyticus* may help elucidate this difference in phenotype.

Recently, flagellar length was suggested to be regulated at the level of export since cells that contain multiple copies of the flagellin gene have abnormally long filaments (1). In this study, we show that a plasmid insertion in ORF3 causes an elongated flagellar filament. This suggests that either ORF3, for which no function is known, or a downstream gene has a possible role either in regulating export of flagellin proteins or in the regulation of flagellin gene expression. If ORF3 has a role in export, then it could aid in the identification of the unidentified export machinery for the flagellar proteins.

In summary, this work continues the analysis of the complex structure of the flagellum of *V. anguillarum* and of the role that the flagellum plays in the virulence of this organism. Four additional flagellin genes, *flaBCDE*, were identified, and the protein products of three of these genes have been shown to be a part of the flagellar filament. The potential for this many flagellin subunits present in a flagellar filament is unusual. There is only one bacterium to have been reported with five flagellin genes, *Halobacterium halobium* (6). Virulence analysis of strains with defects in these genes suggests that FlaD and FlaE, like FlaA, may play a role in the regulation of virulence.

ACKNOWLEDGMENT

This work was supported by grant 42.0045/92 from the Swedish Council for Forestry and Agricultural Research.

REFERENCES

1. **Aizawa, S.-I.** 1996. Flagellar assembly in *Salmonella typhimurium*. Mol. Microbiol. **19:**1–5.
2. **Attridge, S. R., and D. Rowley.** 1983. The role of the flagellum in the adherence of *Vibrio cholerae*. J. Infect. Dis. **147:**864–872.
3. **Camilli, A., and J. J. Mekalanos.** 1995. Use of recombinase gene fusions to identify *Vibrio cholerae* genes induced during infection. Mol. Microbiol. **18:**671–683.
4. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
5. **Drake, D., and T. C. Montie.** 1988. Flagella, motility, and invasive virulence of *Pseudomonas aeruginosa*. J. Gen. Microbiol. **134:**43–52.
6. **Gerl, L., and M. Sumper.** 1988. Halobacterial flagellins are encoded by a multigene family. J. Biol. Chem. **263:**13246–13251.
7. **Grant, C. C. R., M. E. Konkel, W. Cieplak, Jr., and L. S. Tompkins.** 1993. Role of flagella in adherence, internalization, and translocation of *Campylobacter jejuni* in nonpolarized and polarized epithelial cell cultures. Infect. Immun. **61:**1764–1771.
8. **Guentzel, M. N., and L. J. Berry.** 1975. Motility as a virulence factor for *Vibrio cholerae*. Infect. Immun. **11:**890–897.
9. **Gussin, G. N., C. W. Ronson, and F. M. Ausubel.** 1986. Regulation of nitrogen fixation genes. Annu. Rev. Genet. **20:**567–591.
10. **Homma, M., D. DeRosier, and R. M. Macnab.** 1990. Flagellar hook and hook-associated proteins of *Salmonella typhimurium* and their relationship to other axial components of the flagellum. J. Mol. Biol. **211:**819–832.
11. **Jones, G. W., and R. Freter.** 1976. Adhesive properties of *Vibrio cholerae*: nature of the interaction with isolated rabbit brush border membranes and human erythrocytes. Infect. Immun. **14:**240–245.
12. **Joys, T. M., and J. Kim.** 1978. Identification of ε-*N*-methllysine residues in the phase-1 flagella protein of *Salmonella typhimurium*. Microbios Lett. **7:**65–68.
13. **Kelly-Wintenberg, K., T. Anderson, and T. C. Montie.** 1990. Phosphorylated tyrosines in the flagellum filament protein of *Pseudomonas aeruginosa*. J. Bacteriol. **172:**5135–5139.
14. **Kornacker, J. G., and A. Newton.** 1994. Information essential for cell-cycle-dependent secretion of the 591-residue *Caulobacter* hook protein is confined to a 21-amino-acid sequence near the N-terminus. Mol. Microbiol. **14:**73–85.
15. **Kutsukake, K., Y. Ohya, and T. Iino.** 1990. Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. J. Bacteriol. **172:**741–747.
16. **Kuwajima, G., I. Kawagishi, M. Homma, J.-I. Asaka, E. Kondo, and R. M. Macnab.** 1989. Export of an N-terminal fragment of *Escherichia coli* flagellin by a flagellum-specific pathway. Proc. Natl. Acad. Sci. USA **86:**4953–4957.
17. **Logan, S. M., T. J. Trust, and P. Guerry.** 1989. Evidence for posttranslational modification and gene duplication of *Campylobacter* flagellin. J. Bacteriol. **171:**3031–3038.
18. **Lugtenberg, B., J. Meijers, R. Peters, P. van der Hoek, and L. van Alphen.** 1975. Electrophoretic resolution of the "major outer membrane protein" of *Escherichia coli* K12 into four bands. FEBS Lett. **58:**254–258.
19. **McCarter, L. L.** 1995. Genetic and molecular characterization of the polar flagellum of *Vibrio parahaemolyticus*. J. Bacteriol. **177:**1595–1609.
20. **McSweegan, E., and R. I. Walker.** 1986. Identification and characterization of two *Campylobacter jejuni* adhesins for cellular and mucous substrates. Infect. Immun. **53:**141–148.
21. **Miller, V. L., and J. J. Mekalanos.** 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. J. Bacteriol. **170:**2575–2583.
22. **Milton, D. L., A. Norqvist, and H. Wolf-Watz.** 1992. Cloning of a metalloprotease gene involved in the virulence mechanism of *Vibrio anguillarum*. J. Bacteriol. **174:**7235–7244.
23. **Milton, D. L., R. O'Toole, P. Hörstedt, and H. Wolf-Watz.** 1996. Flagellin A is essential for the virulence of *Vibrio anguillarum*. J. Bacteriol. **178:**1310–1319.
24. **Norqvist, A., Å. Hagström, and H. Wolf-Watz.** 1989. Protection of rainbow trout against vibriosis and furunculosis by the use of attenuated strains of *Vibrio anguillarum*. Appl. Environ. Microbiol. **55:**1400–1405.
25. **Norqvist, A., and H. Wolf-Watz.** 1993. Characterization of a novel chromosomal virulence locus involved in expression of a major surface flagellar sheath antigen of the fish pathogen *Vibrio anguillarum*. Infect. Immun. **61:**2434–2444.
26. **O'Toole, R., D. L. Milton, and H. Wolf-Watz.** 1996. Chemotactic motility is required for invasion of the host by the fish pathogen *Vibrio anguillarum*. Mol. Microbiol. **19:**625–637.
27. **Richardson, K.** 1991. Roles of motility and flagellar structure in pathogenicity of *Vibrio cholerae*: analysis of motility mutants in three animal models. Infect. Immun. **59:**2727–2736.

28. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. Bio/Technology **1:**787–796.
30. **Thöny, B., and H. Hennecke.** 1989. The −24/−12 promoter comes of age. FEMS Microbiol. Rev. **63:**341–358.
31. **Wieland, F., G. Paul, and M. Sumper.** 1985. Halobacterial flagellins are sulfated glycoproteins. J. Biol. Chem. **260:**15180–15185.
32. **Yao, R., D. H. Burr, P. Doig, T. J. Trust, H. Niu, and P. Guerry.** 1994. Isolation of motile and non-motile insertional mutants of *Campylobacter jejuni*: the role of motility in adherence and invasion of eukaryotic cells. Mol. Microbiol. **14:**883–893.