2-1999

Genetic and Transcriptional Analyses of the Vibrio cholerae Mannose-Sensitive Hemagglutinin Type 4 Pilus Gene Locus

Jane W. Marsh
Dartmouth College

Ronald K. Taylor
Dartmouth College

JOURNAL OF BACTERIOLOGY, Feb. 1999, p. 1110–1117 Vol. 181, No. 4
0021-9193/99/$04.00+0

Genetic and Transcriptional Analyses of the *Vibrio cholerae* Mannose-Sensitive Hemagglutinin Type 4 Pilus Gene Locus

JANE W. MARSH AND RONALD K. TAYLOR*

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

Received 7 August 1998/Accepted 24 November 1998

The mannose-sensitive hemagglutinin (MSHA) of the *Vibrio cholerae* O1 El Tor biotype is a member of the family of type 4 pili. Type 4 pili are found on the surface of a variety of gram-negative bacteria and have demonstrated importance as host colonization factors, bacteriophage receptors, and mediators of DNA transfer. The gene locus required for the assembly and secretion of the MSHA pilus has been localized to a 16.7-kb region of the *V. cholerae* chromosome. Sixteen genes required for hemagglutination, including five that encode prepilin or prepilin-like proteins, have been identified. Examination of MSHA-specific cDNAs has localized two promoters that drive expression of these genes. This evidence indicates that the MSHA gene locus is transcriptionally organized into two operons, one encoding the secretory components and the other encoding the structural subunits, an arrangement unique among previously characterized type 4 pilus loci. The genes flanking the MSHA locus encode proteins that show homology to YhdA and MreB of *Escherichia coli*. In *E. coli*, the *yhdA* and *mreB* genes are adjacent to each other on the chromosome. The finding that the MSHA locus lies between these two *E. coli* homologs and that it is flanked by a 7-bp direct repeat suggests that the MSHA locus may have been acquired as a mobile genetic element.

Type 4 pili are thin, 6- to 7-nm fibers elaborated by a wide variety of gram-negative bacterial species (50). Type 4 pili are composed of pilin subunits and are classified as either type 4a or type 4b according to amino acid sequence similarities within the amino terminus of the subunit polypeptide. The pilin monomers are synthesized as precursor proteins with a hydrophilic leader peptide of variable length that is processed at a consensus cleavage site by a type 4 prepilin peptidase during pilin secretion. The majority of type 4 pili belong to the type 4a subclass. Type 4a prepilin subunits are characterized by a short five- to six-amino acid leader sequence, which upon cleavage results in a mature pilin subunit with an N-terminal methylated phenylalanine residue. Type 4a pili tend to be distributed either peritrichously or in a polar position on the bacterial cell surface. Type 4b pili are primarily associated with enteric bacteria, namely, *Vibrio cholerae*, enteropathogenic *Escherichia coli*, and enterotoxigenic *E. coli*. A type 4b pilus is also encoded on the broad-host-range conjugal plasmid R64 (29). Type 4b prepilins tend to have leader peptides longer than those of the type 4a subclass and the N-terminal amino acid of the mature pilin subunit is variable, being either methionine, leucine, or tryptophan (13, 17, 53). Pili composed of type 4b pilin subunits have a tendency to form large bundles of laterally associated fibers. Type 4 pili of either subclass can promote bacterial aggregation, a property which may contribute to colonization mechanisms, biofilm formation, or stabilization of mating pairs (29, 37, 53).

Assembly and secretion of type 4 pili requires the expression of numerous gene products, including the structural prepilin subunit and its cognate prepilin peptidase (22). In addition, biogenesis of type 4 pili requires many ancillary proteins whose exact function in pilus assembly is unclear. These include predicted inner and outer membrane proteins and cytoplasmic nucleotide-binding proteins which may provide the energy necessary for translocation (22). The genetic organization of the loci required for type 4 pilus biogenesis tends to vary according to the pilus subclass. Genes required for the synthesis and secretion of type 4b pili tend to be organized as a single operon associated with a plasmid or pathogenicity island, whereas the genes encoding similar functions with respect to type 4a pili tend to be distributed to several locations throughout the chromosome (16, 56).

V. cholerae O1, the etiologic agent of the acute diarrheal disease cholera, expresses both subclasses of type 4 pili. The toxin coregulated pilus (TCP), a member of the family of type 4b pili, is a major determinant in the establishment of *V. cholerae* colonization of the small intestine (21, 51, 53). In addition, TCP serves as the receptor for the filamentous phage, CTXΦ, which encodes the potent cholera exotoxin that is ultimately responsible for the severe diarrhea associated with the disease (58). The genes responsible for TCP elaboration, as well as accessory colonization factor (ACF) genes, are encoded within a 12-kb operon located on the TCP-ACF pathogenicity island (28, 30). Expression of TCP and cholera toxin genes is coordinately regulated in response to specific environmental cues as part of the ToxR regulatory cascade (48).

Epidemic cholera is associated with strains of the O1 and O139 serogroups. The O1 serogroup is further divided into two biotypes, El Tor and classical, based on a variety of phenotypic characteristics such as differential antibiotic and phage sensitivities. Strains of the O139 serogroup are closely related to those of the O1 El Tor biotype (19). One feature which distinguishes the El Tor from the classical O1 biotype is the presence of the mannose-sensitive hemagglutinin (MSHA) pilus on the El Tor cell surface (14, 25). MSHA is also expressed by strains of the O139 serogroup (1). The MSHA pilus, a member of the type 4a family of pili, is not required for colonization of humans or colonization in the infant mouse cholera model (4, 51, 54). While the exact function of MSHA is un-

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

known, recent evidence demonstrates that the MSHA pilus serves as the receptor for the filamentous bacteriophage 493, which was isolated from a *V. cholerae* O139 strain and has been suggested to have a role in the horizontal evolution of *V. cholerae* (27). The MSHA pilus may also have a role in the environmental persistence or survival of El Tor *V. cholerae*. Recent studies have demonstrated that *mshA* mutants are unable to produce biofilms on abiotic surfaces (59). In this regard, biofilm formation may play an important role in mediating *V. cholerae* survival outside the living host.

Initial transposon mutagenesis identified two genetic loci required for mannose-sensitive hemagglutination of El Tor *V. cholerae*. One locus contained a set of contiguous open reading frames that showed significant homology to proteins of the general secretory pathway of gram-negative bacteria (20, 41). In a separate study, the gene encoding the major structural subunit of the pilus, *mshA*, was identified and shown to lie within a locus that included three additional genes encoding type 4 prepilin-like proteins (26). The subsequent finding that the secretory and structural gene loci are physically linked on the *V. cholerae* chromosome suggested that the genes encoding the MSHA pilus are organized into a potential pilus biogenesis operon (33). The organization of pilus biogenesis genes into a discrete cluster is atypical of type 4a pilus-encoding loci. In the present study, this genetic arrangement has been investigated by delineating the boundaries and examining the transcriptional organization of the MSHA gene locus.

MATERIALS AND METHODS

Bacterial strains and plasmids. All bacterial strains and plasmids used in this study are described in Table 1. Bacteria were grown in Luria-Bertani (LB) liquid culture or solid media. Antibiotics were used at the following concentrations unless stated otherwise: ampicillin, 100 μg/ml; polymyxin B, 50 IU/ml; streptomycin, 100 μg/ml; and kanamycin, 45 μg/ml.

Mutant construction. All mutants used in this study were constructed by sequence-specific suicide plasmid integration into the *V. cholerae* chromosome. Plasmid disruption of *mshL* in C6706str2 resulted in JM127. Briefly, primers JM3 and JM4 (Table 2), carrying unique restriction enzyme sites, were used to PCR amplify an internal fragment of the *mshL* gene from C6706str2 genomic DNA. The resulting product was digested with the appropriate restriction enzymes and ligated into the multiple cloning site of the suicide vector pGP704. Following electroporation of this reaction mixture into SM10λ*pir* cells and selection on LB agar containing ampicillin, transformants were screened by colony PCR with primer pair JM3-JM4, which identified plasmid pJM8 carrying the *mshL* fragment. Plasmid pJM8 was integrated into the *V. cholerae* chromosome by conjugation of SM10λ*pir*(pJM8) with C6706str2 and selection on LB agar containing ampicillin and streptomycin. Southern blot analysis utilizing a digoxygenin-labeled probe complementary to *mshK* sequences confirmed the *mshL* plasmid integrant, JM127. Plasmid disruption of *mshQ*, *mreB*, and *yhdA* in C6706str2 was carried out in a similar fashion. Amplification of C6706str2 genomic DNA with gene-specific primer pair MSH33-MSH36, REB1-REB2, or EC1-EC2, each carrying unique restriction enzyme sites, resulted in the production of PCR fragments homologous to internal portions of *mshQ*, *mreB*, and *yhdA*, respectively. Each gene-specific fragment was digested with the appropriate restriction enzymes for ligation into the suicide vector pKAS32 (*mshQ* and *mreB*) or pGP704 (*yhdA*). S17-1λ*pir* cells were electroporated with the *mshQ* or *mreB* ligation reaction mixture followed by selection on LB agar containing ampicillin. The *yhdA* construct was electroporated into SM10λ*pir* cells followed by selection on ampicillin. The resulting transformants were screened for the correct insert by colony PCR with the appropriate sets of gene-specific primers described above. In this manner, plasmid constructs pJM226, pJM271, and pJM150, corresponding to gene fragments *mshQ*, *mreB*, and *yhdA*, respectively, were identified. Integration of each construct into the *V. cholerae* chromosome was carried out by conjugation with C6706str2. The *mshQ* and *mreB* exconjugants were selected on LB agar containing polymyxin B and ampicillin, while the *yhdA* exconjugants were selected in the presence of streptomycin and ampicillin. The *mreB* plasmid integrant strain, JM278, was identified by colony PCR using primer R6K1, which is specific to the origin of replication on the integrated pKAS32 vector, directed toward the 5′ end of *mreB* on the *V. cholerae* chromosome and primer MSH56, which is specific to sequences upstream of *mreB* but outside the region of homology on the integrating plasmid and directed toward the insertion site. The *mshQ* plasmid insertion mutant, JM227, was identified by colony PCR in a similar fashion, using primers R6K1 and MSH30. Southern blot analysis using a digoxygenin-labeled probe specific to *mshK* confirmed *yhdA* plasmid integrant strain, JM157.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Genotype or description[a]	Reference(s)
Strains		
E. coli		
SM10λ*pir*	*thi thr leu tonA lacY supE recA* [RP4-2-Tc::Mu] λ*pirR6K* Km^r	52
S17-1λ*pir*	*thi pro recA hsdR* [RP4-2Tc::Mu-Km::Tn*7*] λ*pir* Tp^r Sm^r	10
MC4100	F$^-$ *araD139* Δ(*araABC-leu*)*7679 galU galK* Δ(*lac*)X74 *rpsL thi* Sm^r	45
V. cholerae		
C6706str2	C6706, spontaneous Sm^r	54
JM127	C6706str2, *mshL*::pGP704 Ap^r	This study
JM157	C6706str2, *yhdA*::pGP704 Ap^r	This study
JM191	C6706str2, Φ(*mshA-phoA*)$_{Hyb1}$	This study
JM222	JM191, *mshF*::TT500::*mshB*	This study
JM223	C6706str2, *mshF*::TT500::*mshB*	This study
JM225	C6706str2, *yhdA*::TT500::*mshI*	This study
JM227	C6706str2, *mshQ*::pKAS32 Ap^r	This study
JM242	JM191, *ssb*::TT500::*yhdA*	This study
JM265	JM191, *yhdA*::TT500::*mshI*	This study
JM278	C6706str2, *mreB*::pKAS32 Ap^r	This study
JM282	C6706str2, *ssb*::TT500::*yhdA*	This study
Plasmids		
pGP704	*oriR6K mobRP4* Ap^r	32, 35
pKAS32	pGP704, *rpsL* Ap^r	47
pKAS46	pKAS32, Ap^r Km^r	47
pRS415	*lac* operon fusion vector, Ap^r	46
pBAD24	*araC rrnB*T$_1$T$_2$ Ap^r	18
pJM8	pGP704, ′*mshL*′ Ap^r	This study
pJM133	pGP704, JM127 *Nco*I chromosomal capture, Ap^r	This study
pJM139	pGP704, JM127 *Sma*I chromosomal capture, Ap^r	This study
pJM150	pGP704, ′*yhdA*′ Ap^r	This study
pJM218	pKAS46, ′*mshF*::TT500::*mshB*′ Ap^r	This study
pJM220	pKAS46, ′*yhdA*::TT500::*mshI*′ Ap^r	This study
pJM226	pKAS32, ′*mshQ*′ Ap^r	This study
pJM245	pKAS32, JM227 *Nsi*I chromosomal capture, Ap^r	This study
pJM246	pRS415, *mshB*::*lacZ* Ap^r	This study
pJM247	pRS415, *mshB*(rev)::*lacZ* Ap^r	This study
pJM271	pKAS32, ′*mreB*′ Ap^r	This study
pJM281	pKAS46, ′*ssb*::TT500::*yhdA*′ Ap^r	This study
pJM307	pRS415, *mshI*::*lacZ* Ap^r	This study
pJM318	pRS415, *mshI*(rev)::*lacZ* Ap^r	This study

[a] Abbreviations: Tp^r, trimethoprim resistance; Sm^r, streptomycin resistance; Km^r, kanamycin resistance; Tc^r, tetracycline resistance; Ap^r, ampicillin resistance; Gm^r, gentamicin resistance.

Transcription terminator constructs. A 500-bp fragment of the *E. coli rrnB* transcription terminator (6) (TT500) was inserted upstream of the *yhdA*, *mshI*, or *mshB* coding sequence in both C6706str2 and the isogenic *mshA-phoA* strain, JM191, by the following method. Two 500-bp fragments flanking either the *yhdA*, *mshI*, or *mshB* upstream region were generated by amplification of C6706str2 genomic DNA with a set of primer pairs, MSH42-MSH38 and MSH43-MSH44, MSP1-MSP3 and MSP4-MSP2, or MSP5-MSP6 and MSP7-MSP8, respectively. The primers carry restriction enzyme sites for ligation into the multiple cloning site of the suicide allelic exchange vector, pKAS46. A *Sac*I site was specifically incorporated into the primers of each set that flank the region to be exchanged with TT500. Ligation reaction mixtures were electroporated into S17-1λ*pir* cells, and transformants were selected on LB agar containing ampicillin. Isolated colonies were screened by PCR with the relevant flanking 5′ and 3′ primer pair described above to identify correct *yhdA*, *mshI*, and *mshB* plasmid constructs. TT500, generated by PCR amplification of pBAD24 DNA with primers TT3 and TT4, each of which carries a *Sac*I site, was introduced into the *Sac*I site of each plasmid construct. The correct orientation of the TT500 insert was determined by PCR using TT4 and the corresponding 5′ primer for each TT500 plasmid construct—MSH42, MSP1, or MSP5. S17-1λ*pir* cells carrying the *yhdA*, *mshI*, and *mshB* TT500 plasmid derivatives pJM281, pJM220, and pJM218, respec-

TABLE 2. Oligonucleotides used in this study

Oligonucleotide	Sequence
EC1	5′-CTAGAGATCTGGCGATTAAGTTGATTGAGCGC-3′
EC2	5′-CTAGGAATTCGCGAATCCAGCGAGCATACCG-3′
JM3	5′-CTAGTCTAGACATCGTGATCCTGTTGAAGC-3′
JM4	5′-CTAGGAATTCGCCTTGGCTTCGACATCTC-3′
KAN4	5′-TCGAGCAAGACGTTTCCCG-3′
MSP1	5′-CTAGTCTAGACACTCTCGGTAAACTTGCAC-3′
MSP2	5′-CTAGGAATTCCGACACCGCCATGCTTTGC-3′
MSP3	5′-CTAGGAGCTCTGACACCCAGCATCTGTAAGG-3′
MSP4	5′-CTAGGAGCTCGGCAAACGTAATCGATGGCG-3′
MSP5	5′-CTAGTCTAGATCGTGAAGTCGACTATG-3′
MSP6	5′-CTAGGAGCTCTCAGCGCATCAAAAAACTGGC-3′
MSP7	5′-TTTTGCCTGCTCATGGCTAG-3′
MSP8	5′-CTAGGAATTCTCTTTCATGTGAATACGCAGC-3′
MSH30	5′-GGATTCCCTACACCAGTAGCG-3′
MSH33	5′-CTAGTCTAGAGTTGGGTGTCCGCTTACATG-3′
MSH36	5′-CTAGGAATTCAGATCTTGCTCATCACCGCT-3′
MSH38	5′-CTAGGAGCTCAATACTAAAACGGGATGTCG-3′
MSH42	5′-CTAGTCTAGAGTGCGTTATATGCCAAGC-3′
MSH43	5′-CTAGGAGCTCGCACATAAGGAACGTGGC-3′
MSH44	5′-CTGAGAATTCAATACGATCAGTAGCACCGC-3′
MSH49	5′-ACAAATTCATGTTGGCTGCC-3′
MSH50	5′-GATCAAGCTTTCCGCCACTACATCCGTCAC-3′
MSH56	5′-TGTGATGGCAAGGCATTGG-3′
REB1	5′-CATGAGATCTCGATTTAGGTACCGCCAAC-3′
REB2	5′-CATGAATTCGCACTTCGATCTCTTGCAC-3′
TT3	5′-TTTAGAGCTCGCTTGGCTGTTTTGGCGGATGAGAG-3′
TT4	5′-GATGGAGCTCAAGAGTTTGTAGAAACGCAAAAAGGCCATCCG-3′
LAC1	5′-GTCATAGCTGTTTCCTGTGTG-3′
R6K1	5′-GGTTTAACGGTTGTGGACAAC-3′

tively, were mated with either C6706str2 or the isogenic *mshA-phoA* gene fusion strain, JM191, and exconjugants were selected on LB agar in the presence of polymyxin B and kanamycin. Selected exconjugants were screened for plasmid integration at the correct site on the chromosome by PCR with the KAN4 primer, specific to the *N*-acetyltransferase gene sequence on pKAS46, and a 3′ primer specific to chromosomal sequences outside of the region of homology on the integrated plasmid. Plasmid loss from all confirmed C6706str2 and JM191 integrants was carried out in the presence of streptomycin (1 mg/ml) at 30°C as previously described (47). C6706str2 streptomycin-resistant recombinants were screened by PCR for insertion of TT500 upstream of *yhdA*, *mshI*, and *mshB*, using the appropriate primer pair flanking each insertion site, and resulted in the identification of strains JM282, JM225, and JM223, respectively. Similarly, insertion of TT500 upstream of *yhdA*, *mshI*, and *mshB* in JM191 resulted in JM242, JM265, and JM222, respectively.

Operon fusion construction. PCR products carrying the promoter region upstream of *mshI* or *mshB* were generated by amplification of C6706str2 genomic DNA with primer pair EC1-EC2 or MSP5-MSP6, respectively. The resulting PCR products were blunt-end ligated into the unique *Sma*I site located upstream of the promoterless *lacZ* gene on plasmid pRS415. The *lacZ* deletion strain, MC4100, was electroporated with either ligation reaction mixture, with selection on LB agar containing ampicillin. Plasmids containing promoter inserts were identified by colony PCR using the original primer pair, EC1-EC2 or MSP5-MSP6. Orientation of the *mshI* promoter fragment was established by PCR amplification of plasmid DNA with MSP5 or MSP6 and the LAC1 primer, which is homologous to *lacZ* sequences on pRS415 and directed toward the cloned *mshI* promoter fragment. This resulted in the identification of operon fusion plasmid pJM307, carrying a 500-bp *mshI* promoter fragment in the forward orientation, and plasmid pJM318, containing the *mshI* promoter fragment in the opposite orientation [*mshI*(rev)]. A *Sal*I site on the cloned *mshB* promoter fragment and in the plasmid backbone allowed the orientation of the *mshB* promoter with respect to *lacZ*. Restriction digest of plasmid DNA identified the operon fusion pJM246 carrying a 700-bp *mshB* promoter fragment in the forward orientation and plasmid pJM247, which contains the *mshB* promoter directed away from the *lacZ* sequence [*mshB*(rev)].

Chromosomal capture. Chromosomal capture (49) was used to clone a region containing the 5′ portion of the *msh* locus and adjacent sequences. Genomic DNA from strain JM127, which carries a pGP704 plasmid insertion in the *mshL* gene, was digested with *Nco*I, self-ligated, and electroporated into strain S17-1λ*pir*, with selection on LB agar containing ampicillin. The resulting plasmid, pJM133, carries a chromosomal *Nco*I fragment containing approximately 5 kb of sequence upstream of the *yhdA* gene. The chromosomal region encompassing the 3′ portion of the *msh* locus and beyond was similarly cloned from JM127 by ligation of *Sma*I-digested chromosomal DNA and resulted in plasmid pJM139, which carries approximately 4 kb of DNA downstream of *mshC*. Chromosomal capture of sequences downstream of *mshQ* was carried out by *Nsi*I digestion of JM227 genomic DNA and resulted in the isolation of pJM245, which carries approximately 2 kb of DNA downstream of *mshQ*. Plasmid DNA from each construct was purified by anion-exchange chromatography (Qiagen, Inc.) and used in subsequent automated DNA sequence analysis.

Sequence analysis. Plasmid DNA obtained by chromosomal capture was sequenced by using a Prism dye terminator ready reaction kit (Applied Biosystems, Inc. [ABI], Foster City, Calif.). Briefly, 1 μg of purified plasmid DNA and 3.2 pmol of oligonucleotide primer were mixed with 8 μl of ABI dye terminator mix and brought to a final 20-μl volume with distilled deionized water. The mixture was subjected to 25 cycles of 96°C for 30 s, 50°C for 15 s, and 60°C for 4 min, followed by Centrisep column purification (Princeton Separations) and gel analysis on an ABI model 373 Stretch automated DNA sequencer. Oligonucleotide primers homologous to known MSHA sequences included in the chromosomal capture were designed such that sequence analysis could proceed from both ends and on both strands of the captured DNA. Each primer was used in six separate sequencing reactions to ensure the fidelity of the sequence data. Subsequent oligonucleotide primers were designed in accordance with the resulting consensus sequence. DNAstar and ABI software packages were used for sequence analysis and alignment.

5′ RACE. To identify transcription start sites, we used the 5′ RACE System for Rapid Amplification of cDNA Ends, version 2.0 (Gibco BRL). Briefly, total RNA (1 μg) isolated from log-phase C6706str2 and JM225 cells by RNeasy silica gel membrane column (Qiagen) purification was reverse transcribed into cDNA with primer MSP2 or MSP8, specific to *mshI* or *mshB*, respectively. A control tube without reverse transcriptase was included for each primer to ensure that the resulting product was due to the amplification of cDNA and not contaminating chromosomal DNA. The cDNA was GlassMax column purified (Gibco BRL), poly(dC) tailed with terminal deoxynucleotidyltransferase, and subjected to PCR using *mshI* or *mshB* gene-specific nested primer MSH50 or MSH49, respectively, and the 5′ RACE abridged anchor primer, which contains 3′ sequence complementary to the homopolymeric poly(dC) tail. The resulting PCR product was reamplified with either MSH50 (for the *mshI*-specific product) or MSH49 (for the *mshB*-specific product) and the 5′ RACE universal amplification primer, which is complementary to the abridged anchor primer. The resulting products were silica gel membrane purified on Qiaquik columns (Qiagen), and 50 ng of the purified product was used as the template for automated sequence analysis with *mshI* or *mshB* gene-specific primer MSP3 or MSP6, respectively.

Hemagglutination assay. Hemagglutination assays were carried out in 96-well round-bottom microtiter plates in a final volume of 200 μl. Bacteria ($\sim 10^7$) were serially diluted in Krebs-Ringer buffer (15). CD-1 mouse (Charles River Laboratories) erythrocytes were collected in the presence of heparin, washed three times, and resuspended at a final concentration of 1% in Krebs-Ringer buffer. An equal volume of serially diluted bacteria was mixed with the mouse erythrocytes and incubated at room temperature for 1 h. Hemagglutination titers are expressed as the reciprocal of the highest bacterial dilution that gave strong hemagglutination.

β-Galactosidase and alkaline phosphatase assays. β-Galactosidase assays were performed on mid-logarithmic-phase cultures of MC4100 carrying *lacZ* operon fusion plasmid pJM246, pJM247, pJM307, and pJM318. Alkaline phosphatase assays were performed on overnight cultures of *mshA-phoA* gene fusion strains JM191, JM222, JM242, and JM265. Assays were performed as previously described (32).

Nucleotide sequence accession numbers. The DNA sequence generated has been entered in the GenBank database under accession no. AF079406 (which corresponds to the partial coding sequence of SSB (single-stranded DNA-binding protein) and the complete coding sequence of YhdA) and AF079234 (which corresponds to the 3′ coding sequence of the MSHA operon including the partial coding sequence of MshD, the complete coding sequences of MshO, MshP, and MshQ, and the complete coding sequence of MreB). Accession no. AF079781 corresponds to the partial coding sequence of UvrA.

RESULTS

Identification and characterization of additional *msh* genes. Previous studies identified 13 open reading frames required for MSHA pilus biogenesis (20, 26, 33). In the present study, three additional open reading frames downstream of *mshD* were identified by chromosomal capture and sequencing of *V. cholerae* El Tor genomic DNA (Fig. 1). These genes, designated *mshO*, *mshP*, and *mshQ*, are contiguous with and oriented in the same direction as the previously described *msh* genes. Predicted protein homologies and functions for the entire MSHA gene locus and nearby genes are listed in Table 3. The predicted gene product of *mshO* resembles a type 4 prepilin subunit in that it contains a consensus prepilin peptidase cleav-

TABLE 3. Sizes, protein homologies, and predicted functions or locations of MSHA gene products

Gene product	Predicted mol wt (10^3)	Homology		Predicted function or location	Reference(s)
		Protein	%[a]		
UvrA	104	*H. influenzae* UvrA	93	Excision repair	9
SSB	16	*E. coli* SSB	98	Single-stranded DNA binding	43
YhdA	73.3	*E. coli* YhdA	60	Hypothetical protein	12
MshI	53.4	None		Inner membrane[b]	20
MshJ	24.5	*E. carotovora* OutM	48	Outer membrane	20, 42
MshK	10.9	None		Periplasmic[b]	20
MshL	60.1	*E. chrysanthemi* OutD	47	Outer membrane	31
MshM	31.7	*A. hydrophila* ExeA	56	Cytoplasmic	20, 24
MshN	42.6	None		Inner membrane[b]	20
MshE	52.5	*V. cholerae* EpsE	63	Cytoplasmic, ATP binding	20, 44
MshG	44.8	*N. gonorrhoeae* PilG	55	Inner membrane	33, 55
MshF	17.9	None		Periplasmic[b]	26
MshB	19.7	Type 4 prepilin		Minor pilin, other	26
MshA	18.4	Type 4 prepilin		Major pilin subunit	26
MshC	17.3	Type 4 prepilin		Minor pilin, other	26
MshD	19.7	Type 4 prepilin		Minor pilin, other	26
MshO	27.8	Type 4 prepilin		Minor pilin, other	This study
MshP	15.2	*E. chrysanthemi* OutG	52	Periplasm or outer membrane[b]	31
MshQ	134.8	None		Outer membrane[b]	This study
MreB	36.9	*E. coli* MreB	95	Rod cell shape	12

[a] Homologies based on gapped BLAST analysis (3).
[b] PSORT analysis (36).

age site and a C-terminal pair of cysteine residues which may form the characteristic pilin subunit disulfide hairpin (40). The predicted *mshO* gene product has a leader sequence that is only four amino acids long but contains invariant amino acids at positions −1, +1, +3, and +5 relative to the peptidase cleavage site and maintains overall N-terminal amino acid residue conservation characteristic of the type 4 prepilins. The *mshO* gene is the last of five consecutive open reading frames that encode type 4 prepilin or prepilin-like proteins. The deduced amino acid sequence of the *mshP* gene shows 52% homology to OutG of *Erwinia chrysanthemi*. This protein is a component of the *E. chrysanthemi* pectic enzyme secretion pathway and contains a type 4 prepilin leader sequence (31). While MshP does not appear to contain a consensus type 4a prepilin leader sequence, the presence of a charged amino terminus followed by a hydrophobic stretch of amino acids suggests that this protein is secreted across the bacterial inner membrane, potentially in a signal sequence-dependent manner. While the ultimate destination of the putative MshP protein to the periplasm or the outer membrane could not be predicted from its primary amino acid sequence, the OutG homology suggests that MshP may contribute to the secretion of pilin subunits and the assembly of the MSHA pilus fiber. Finally, the *mshQ* open reading frame encodes a product with a predicted molecular mass of 134.8 kDa that is likely to initiate with a valine residue, based on the location of the *mshP* stop codon. MshQ is likely to be an outer membrane protein, based on the presence of a potential signal sequence with a processing site at position 24 and the extensive predicted beta-sheet structure which occurs throughout the mature protein (36).

MSHA gene locus boundaries delineated. Chromosomal capture and sequence analysis of C6706str2 El Tor genomic DNA upstream of the secretory genes at the 5′ region of the MSHA locus identified three open reading frames (Fig. 1). Immediately upstream of the *mshI* gene lies an open reading frame predicted to encode a gene product with 60% homology to YhdA of *E. coli*, a 73.3-kDa protein of unknown function encoded within the *mreB-accB* intergenic region (Table 3) (SWISS-PROT entry P13518). The predicted molecular mass of the presumed *V. cholerae* YhdA protein is 75 kDa. Upstream of *yhdA* and oriented in the same direction lies a gene that encodes a putative *V. cholerae* SSB (Fig. 1). The predicted amino acid sequence of this protein shows 96% homology to SSB (a protein involved in DNA replication, recombination, and repair) of many bacteria (Table 3). Therefore, this *V. cholerae* gene has been named *ssb*. An additional open reading frame discovered upstream of *ssb* was designated *uvrA* due to the significant amino acid homology of its predicted product with the *Haemophilus influenzae* excision repair protein (Table 3). The organization of the DNA repair genes observed in the *V. cholerae* chromosome, with the *uvrA* gene located upstream and adjacent to *ssb* but transcribed in the opposite direction, is evident in a number of bacterial genomes (11, 43). The location of these *V. cholerae* recombination-repair genes is indicative of the upstream MSHA gene locus boundary.

Downstream of *mshQ* at the 3′ region of the MSHA locus lies a gene predicted to encode a protein with 78% homology to MreB of a number of bacteria. The MreB protein is required for the formation of the characteristic rod shape in *E. coli* (12). This *E. coli* protein is encoded within a gene cluster which is involved in the regulation of bacterial cell division (57). Due to these observations and the significant homology to the *E. coli* MreB protein, the *V. cholerae* gene has been designated *mreB*. Interestingly, the *yhdA* and *mreB* genes,

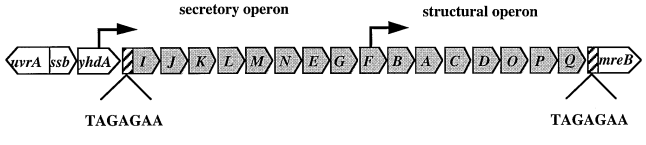

FIG. 1. Schematic representation of the MSHA gene locus and flanking open reading frames. The MSHA gene locus is 16.7 kb in length and consists of 16 contiguous open reading frames which are flanked by a 7-bp direct repeat (⎚). The three genes depicted upstream of the 5′ end (*yhdA*, *ssb*, and *uvrA*) and the one gene downstream of the 3′ end of the locus (*mreB*) are not involved in MSHA pilus biogenesis. The intragenic promoters required for the expression of MSHA are indicated by arrows.

TABLE 4. Hemagglutination titers of C6706 derivatives

Strain	Genotype	Titer
C6706str2	Wild type	128
JM157	*yhdA*::pGP704	128
JM223	*mshF*::TT500::*mshB*	<2
JM225	*yhdA*::TT500::*mshI*	<2
JM227	*mshQ*::pKAS32	32
JM278	*mreB*::pKAS32	128
JM282	*ssb*::TT500::*yhdA*	128

which are adjacent to one another on the *E. coli* chromosome, are separated by the MSHA gene locus on the *V. cholerae* chromosome. This observation is intriguing given the identification of a 7-bp direct repeat, TAGAGAA, located 5′ of *mshI* and 3′ of *mshQ* (Fig. 1). These findings suggest that the MSHA gene locus is delineated by these 7-bp direct repeats and raise the possibility that this locus was acquired by *V. cholerae* as a mobile genetic element that inserted between the *yhdA* and *mreB* genes of the ancestral *V. cholerae* chromosome.

Genetic confirmation of the MSHA gene locus boundaries. To confirm the 5′ and 3′ boundaries of the MSHA gene locus, the genes surrounding the MSHA gene locus were subjected to mutational analysis. Plasmid disruption of *yhdA* and *mreB* gene sequence in the *V. cholerae* El Tor chromosome was performed to determine if the products of these genes are involved in MSHA pilus biogenesis. Hemagglutination of mouse erythrocytes was used to quantitate MSHA on the surface of the *V. cholerae* mutants. A hemagglutination assay performed with the C6706str2 *yhdA* and *mreB* mutant strains JM157 and JM278, respectively, demonstrated that neither of these genes is required for the MSHA phenotype. Both JM157 and JM278 displayed wild-type hemagglutination titers (Table 4). The wild-type hemagglutination titer observed for the *yhdA* mutant suggests that *mshI* is the first gene in the MSHA pilus biogenesis operon, since a Tn*phoA* insertion in *mshI* abolishes hemagglutination in the El Tor strain N16961 (20).

Upstream of the *mreB* gene on the *V. cholerae* chromosome lies the *mshQ* gene. To determine if MshQ is the last protein encoded on the MSHA gene locus, plasmid disruption mutant JM227 was created and tested for its ability to hemagglutinate mouse erythrocytes. Disruption of MshQ function resulted in a 75% decrease in the hemagglutination titer compared to both the wild-type C6706str2 control strain and the *mreB* mutant (Table 4), indicating that *mshQ* is the last gene of the MSHA locus.

Transcriptional organization of the MSHA gene locus. Type 4 pilus biogenesis operons are typically organized with the pilin gene located promoter proximal, followed immediately by a sequence that decreases transcription readthrough (transcriptional down sequence) and the pilus assembly and secretory genes. This organization allows for coordinate expression of all genes while maintaining proper stoichiometry of the pilin subunits relative to the assembly and secretory components of the pilus biogenesis operon. The organization of the MSHA pilus genes illustrated in Fig. 1 is unusual in that the secretory components are encoded upstream of the pilus structural components. This arrangement can best be explained if the MSHA secretory and structural components are in fact encoded on separate operons and regulated by individual promoters. Evidence for two operons is supported by the identification of potential consensus sigma 70 promoters located upstream of both *mshI* and *mshB* coding sequences (20, 26). To ascertain whether these promoters were necessary and sufficient for *msh* gene expression, the region encompassing either the *mshI* or the *mshB* putative promoter was deleted and replaced with the *E. coli rrnB* transcription terminator (TT500) in wild-type C6706str2 and isogenic *mshA-phoA* gene fusion strains. Deletion of the promoter region upstream of *mshB* and replacement with TT500 abolished *mshA* expression, as illustrated by the low level of alkaline phosphatase activity exhibited by strain JM222 compared to the control, JM191 (Table 5). Confirmation of this expression defect was provided by the hemagglutination-negative phenotype of the isogenic C6706str2 TT500 derivative, JM223 (Table 4). These data demonstrate that a promoter is not present immediately upstream of the *mshA* structural subunit gene and indicate that regulation of *mshA* expression must originate from a promoter located upstream of *mshB*.

The possibility existed that one promoter located upstream of *mshI* was required to drive the expression of the entire MSHA gene locus. To address this possibility, the putative *mshI* promoter was replaced with TT500 in both C6706str2 and the JM191 *mshA-phoA* gene fusion strain. The TT500 *mshA-phoA* derivative, JM265, exhibited 80% of the control alkaline phosphatase activity, yet hemagglutination activity was completely abolished in the C6706str2 TT500 derivative, JM225 (Tables 4 and 5). This finding indicates that a promoter downstream of *mshI* is responsible for the majority of the transcriptional activity exhibited by the *mshA-phoA* gene fusions and that MSHA pilus biogenesis is completely dependent on expression from a promoter located upstream of *mshI*. Both JM282 and JM242, containing the TT500 terminator upstream of the *yhdA* gene, expressed wild-type levels of *mshA*, as reflected by their hemagglutination titers and alkaline phosphatase activities (Tables 4 and 5). These data indicate that a promoter upstream of *mshI* but downstream of *yhdA* is required for expression of the *msh* secretory genes, while a second promoter upstream of *mshB* but downstream of *mshI* is necessary for expression of the *mshA* structural subunit gene. These results support the two-promoter hypothesis for MSHA pilus expression, assembly, and secretion in *V. cholerae*.

Mapping the transcriptional start sites by 5′ RACE. To establish the exact locations of the promoters driving *msh* gene expression, the 5′ ends of the corresponding mRNAs were determined by 5′ RACE. Primers specific to either the *mshI* or *mshB* transcript were used for subsequent cDNA synthesis and PCR amplification. The *mshI*-specific cDNA amplified a 900-bp product that was seen only in wild-type C6706str2 cells expressing the *mshI* transcript (Fig. 2, lane 1). No PCR product was observed with cDNA prepared from JM225, the C6706str2 TT500 derivative that does not express *mshI* (Fig. 2, lane 2). A 700-bp band was specifically amplified from *mshB*-specific cDNA prepared from C6706str2 total RNA (Fig. 2, lane 3). A control cDNA reaction without the addition of reverse transcriptase gave no product, indicating that the PCR product

TABLE 5. Alkaline phosphatase activities of C6706 *mshA-phoA* derivatives

Strain	Genotype	Activity (U)[a]
C6706str2	Wild type	20 ± 3
JM191	*mshA-phoA*	1,985 ± 28
JM222	*mshF*::TT500::*mshB*	37 ± 4
JM265	*yhdA*::TT500::*mshI*	1,670 ± 158
JM242	*ssb*::TT500::*yhdA*	1,847 ± 211

[a] Average ± standard deviation of triplicate samples.

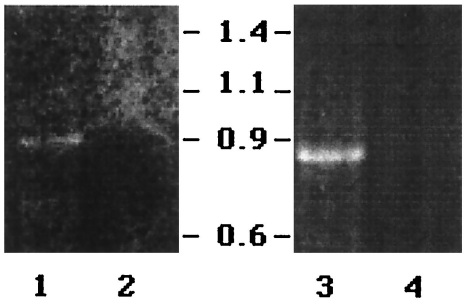

FIG. 2. PCR amplification of gene-specific cDNAs. RNA isolated from C6706str2 or JM225 was reverse transcribed into cDNA followed by 5′ RACE with primers specific to *mshI* (lane 1 and 2) and *mshB* (lanes 3 and 4). Lane 1, C6706str2; lane 2, JM225; lane 3, C6706str2; lane 4, C6706str2, negative control cDNA reaction in the absence of reverse transcriptase. DNA standards are indicated in kilobase pairs.

observed for *mshB* is specific to the *mshB* transcript and not contaminating DNA (Fig. 2, lane 4).

The PCR products obtained from the 5′ RACE were subjected to sequence analysis to localize the start of transcription of each potential operon. Interestingly, both transcription start sites were found to be intragenic. The *mshI* transcription start site maps to the middle of the *yhdA* gene, while the start of *mshB* transcription lies within the *mshF* gene (Fig. 1). A sigma 70 consensus promoter was identified at an appropriate position upstream of each transcription start site (Fig. 3).

Confirmation of promoter activity. To demonstrate transcriptional activity, the putative promoters identified by 5′ RACE were cloned into the promoterless *lacZ* reporter plasmid pRS415. Activities of the resulting *mshI* and *mshB* operon fusions were determined by β-galactosidase assay of *E. coli* cells carrying pJM307 and pJM246, respectively. In addition, the activity of each promoter fragment in the reverse orientation was tested. As shown in Table 6, each promoter supported expression of the *lacZ* gene at a level consistently higher than that for the pRS415 control. The *mshI* promoter expressed on plasmid pJM307 demonstrated a greater level of transcription than the cloned *mshB* promoter on plasmid pJM246 (5,700 versus 1,500 U). The *mshI* promoter fragment cloned in the reverse orientation with respect to the *lacZ* gene on plasmid pJM318 expressed low-level β-galactosidase activity (Table 6). This result is in contrast to the relatively high level of β-galactosidase expression observed when an inverted *mshB* promoter fragment was placed upstream of the *lacZ* gene on plasmid pJM247 (Table 6). This level of *lacZ* gene expression indicates that there is promoter activity in the opposite direction and suggests a potential mechanism for negative regulation of pilin subunit gene expression. These data indicate that the two promoters required for MSHA type 4 pilus biogenesis are functional in vivo.

```
mshI:       TTGTTCCATTTAACGATAAACGGTATGCTCGCTGGA
mshB:       AGGAGTTGATTGTGAAAAAGTGCTATTTTTGCTCTA
Consensus:  TTGACA ---- 17 bp ---- TATAAT --- +1
             -35                   -10
```

FIG. 3. Comparison of MSHA and sigma 70 consensus promoters. Nucleotide sequences upstream of the 5′ ends of the *mshI* and *mshB* transcripts are depicted. The putative start of transcription is designated +1. The −35 and −10 promoter regions are indicated in boldface. The sigma 70 consensus sequence is indicated below the *mshI* and *mshB* promoter sequences.

TABLE 6. β-Galactosidase activities of *msh* promoter constructs

Plasmid	Construct	Activity (U)[a]
pRS415	Vector	59 ± 5
pJM307	*mshI*::*lacZ*	5,682 ± 87
pJM318	*mshI*(rev)::*lacZ*	139 ± 8
pJM246	*mshB*::*lacZ*	1,576 ± 233
pJM247	*mshB*(rev)::*lacZ*	637 ± 89

[a] Average ± standard deviation of triplicate samples.

DISCUSSION

In this study, the entire MSHA type 4 pilus biogenesis gene locus has been defined. Three additional open reading frames located downstream of the structural gene locus originally described by Jonson et al. (26) have been identified. The newly identified *msh* genes have been designated *mshO*, *mshP*, and *mshQ*. The amino acid sequence of the predicted *mshO* gene product contains a short, basic leader sequence and a conserved processing motif characteristic of the type 4 prepilins. This finding increases the total number of type 4 prepilins encoded in the MSHA gene locus to five. A similar cluster of genes encoding five type 4 prepilins has been described for the *exe* secretory locus of *Aeromonas hydrophila* (23). While a type 4 pilus is not associated with the *exe* gene locus, it is interesting that a similar pilin subunit gene arrangement exists at the MSHA gene locus and that several of the predicted MSHA gene products show significant homology to proteins involved in the bacterial general secretory pathway (Table 3). The *Pseudomonas aeruginosa* type 4 pilus responsible for colonization, twitching motility, and phage sensitivity requires the expression of seven type 4 prepilin and prepilin-like proteins which are encoded on three distinct loci of the bacterial chromosome (2). The MshA pilin subunit has previously been shown to be the major structural subunit of the MSHA pilus (26); however, the role of the additional pilin proteins in MSHA biogenesis remains unclear. The possibility exists that the MSHA pilus is a heteropolymer of pilin subunits or that accessory pilin subunits are required for the assembly of a pilus secretory apparatus. Recent evidence has demonstrated that the type 4 prepilin peptidase required for MshA prepilin processing is also required for processing of the Eps prepilin-like proteins associated with *V. cholerae* extracellular secretion (34). These observations demonstrate the similarities between extracellular secretion and type 4 pilus assembly and raise the possibility that the MSHA genetic element was at one time involved in extracellular secretion.

Many of the genes located upstream of the *mshA* pilin subunit gene encode homologs of general secretory pathway components (Table 3) (20, 41). These homologies predict that the region upstream of the *mshA* pilin structural gene is involved in the secretion and assembly of MSHA pilin subunits into a functional pilus on the bacterial cell surface. Previous studies have demonstrated that the genes encoding the secretory components are physically linked to the *mshA* structural locus on the *V. cholerae* chromosome (30). This observation suggested that the contiguous set of uniformly oriented genes required for MSHA pilus biogenesis may be organized into a single operon. This genetic arrangement would be unusual since type 4 pilus biogenesis operons are typically arranged such that the major pilin subunit gene is promoter proximal, providing a means to generate greater expression of pilin monomers relative to the secretory components. This report demonstrates that two promoters are required to drive expression of the MSHA gene locus. Analysis of various transcription terminator

insertions in an *mshA-phoA* gene fusion strain demonstrated that *mshA* prepilin subunit gene expression is dependent on promoter activity generated upstream of *mshB* but downstream of *mshI*. This analysis indicates that the MSHA locus is organized as two operons. One operon encodes five type 4 prepilin subunits including the major structural subunit of the pilus, MshA; the second operon encodes proteins predicted to be involved in the assembly and secretion of the pilin structural subunits and is located immediately upstream of the operon encoding the pilin subunits.

The genetic organization of the MSHA gene locus more closely resembles that found for the family of type 4b pili than type 4a pili. The type 4b pili, including the bundle-forming pilus of enteropathogenic *E. coli* and TCP of *V. cholerae*, are encoded in a single locus which is located on a plasmid (bundle-forming pilus) or a pathogenicity island (TCP) (13, 28). Within these operons, the ancillary components essential for pilus biogenesis are typically encoded downstream of the major pilin subunit gene. However, the MSHA locus differs in that the genes required for assembly and secretion of the pilus are located directly upstream of the genes encoding the pilin subunit. This unusual genetic arrangement has been described for several fimbriae of *E. coli* whose structural subunit genes are thought to be regulated by independent promoters in order to maintain proper stoichiometry of pilin subunits relative to the secretion components (5, 39). A similar mechanism for pilin subunit regulation may exist for the MSHA locus.

5′ RACE analysis of *msh* transcripts indicated that the structural operon transcript initiating upstream of *mshB* was more abundant than the secretory operon transcript which initiates upstream of *mshI*. While 5′ RACE is not quantitative, these results are consistent with the possibility that increased levels of MshA pilin subunit could be regulated at the level of transcription or through increased message stability. Contrary to this hypothesis, the operon fusion analysis in *E. coli* found the *mshI* promoter to have more transcriptional activity than the *mshB* promoter. This finding may reflect the absence of a required *V. cholerae* regulatory protein in *E. coli*. The poor consensus homology for the −35 region of the *mshB* promoter further suggests that this promoter may require an additional *V. cholerae* factor for specific activation.

There is no apparent transcriptional terminator within the secretory operon, suggesting that one long transcript which is initiated from the *mshI* promoter and contributes to the structural gene operon expression may be generated. This hypothesis is supported by studies using the *mshA-phoA* gene fusion strain, JM265, which contains a transcription terminator within the secretory operon and shows less alkaline phosphatase activity than the isogenic control strain, JM191. These results indicate that expression from the secretory operon is essential for maximum *mshA* structural subunit gene expression. Alternatively, expression of the secretory components may stabilize the MshA-PhoA hybrid protein or increase its secretion across the inner membrane in this system. In either case, evidence of a second transcript beginning at a promoter located upstream of *mshB* and in the middle of the *mshF* gene is provided by 5′ RACE analysis and suggests that expression from a long *mshI*-initiated transcript is not sufficient for MSHA structural operon expression.

The MSHA gene locus consists of 16 open reading frames flanked by two genes whose predicted gene products show significant homology to YhdA and MreB of *E. coli*. Interestingly, the genes encoding these proteins are adjacent to one another on the *E. coli* chromosome. Neither protein has any function associated with pilus formation or extracellular secretion in *E. coli*. Furthermore, inactivation of these homologs in *V. cholerae* by plasmid insertion had no effect on hemagglutination titers. These results help define the MSHA gene locus boundary given that *mshI* or *mshQ* gene disruption results in hemagglutination defects. In addition, the identification of a 7-bp direct repeat flanking these genes suggests that the MSHA gene locus may have been acquired by *V. cholerae* as a transposable or otherwise mobilizable genetic element. While no integrase function can be attributed to any of the proteins encoded on the locus, it is possible that during genetic transfer to the *V. cholerae* chromosome the integrase gene, essential for MSHA gene locus transposition, was lost.

While the MSHA type 4 pilus has no function in colonization of the mammalian intestine (51), this finding does not exclude the possibility that MSHA is an important attachment factor in the environment, where *V. cholerae* is often found associated with a variety of aquatic organisms (7). Studies indicate that many bacteria form complex communities or biofilms on solid biotic and abiotic surfaces (8, 38). Recent evidence indicates that *mshA* mutants are unable to form biofilms (59). This finding suggests that MSHA may be involved in the initial stages of *V. cholerae* biofilm production in the aquatic environment. Although further studies are necessary, the MSHA locus may represent a genetic cassette that plays a significant role in survival of *V. cholerae* in the environment. We propose that the MSHA locus, analogous to bacterial pathogenicity islands, be termed an environmental persistence island.

ACKNOWLEDGMENTS

We gratefully acknowledge the helpful comments and suggestions of Nicholas Morris, discussions with Jean-François Tomb, and the expert technical assistance of Emily Cornell and the Dartmouth Molecular Biology Core Facility.

This work was supported by National Institutes of Health grant AI-25096 (R.K.T.). J.W.M. is the recipient of a predoctoral fellowship from the National Institutes of Health (training grant AI-07519).

REFERENCES

1. **Albert, M. J.** 1994. *Vibrio cholerae* O139 Bengal. J. Clin. Microbiol. **32:**2345–2349.
2. **Alm, R. A., J. P. Hallinan, A. A. Watson, and J. S. Mattick.** 1996. Fimbrial biogenesis genes of *Pseudomonas aeruginosa*: *pilW* and *pilX* increase the similarity of type 4 fimbriae to the GSP protein-secretion systems and *pilY1* encodes a gonococcal pilC homologue. Mol. Microbiol. **22:**161–173.
3. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
4. **Attridge, S. R., P. A. Manning, J. Holmgren, and G. Jonson.** 1996. Relative significance of mannose-sensitive hemagglutinin and toxin-coregulated pili in colonization of infant mice by *Vibrio cholerae* El Tor. Infect. Immun. **64:**3369–3373.
5. **Bilge, S. S., C. R. Clausen, W. Lau, and S. L. Moseley.** 1989. Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated *Escherichia coli* to HEp-2 cells. J. Bacteriol. **171:**4281–4289.
6. **Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller.** 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. J. Mol. Biol. **148:**107–127.
7. **Colwell, R. R.** 1996. Global climate and infectious disease: the cholera paradigm. Science **274:**2025–2031.
8. **Davies, D. G., M. R. Parsek, J. P. Pearson, B. H. Iglewski, J. W. Costerton, and E. P. Greenberg.** 1998. The involvement of cell-to-cell signals in the development of a bacterial biofilm. Science **280:**295–298.
9. **de la Morena, M. L., D. R. Hendrixson, and J. W. St. Geme.** 1996. Isolation and characterization of the *Haemophilus influenzae uvrA* gene. Gene **177:**23–28.
10. **de Lorenzo, V., and K. N. Timmis.** 1994. Analysis and construction of stable phenotypes in gram-negative bacteria with Tn*5*- and Tn*10*-derived minitransposons. Methods Enzymol. **235:**386–405.
11. **de Vries, J., and W. Wackernagel.** 1993. Cloning and sequencing of the *Serratia marcescens* gene encoding a single-stranded DNA-binding protein (SSB) and its promoter region. Gene **127:**39–45.
12. **Doi, M., M. Wachi, F. Ishino, S. Tomioka, M. Ito, Y. Sakagami, A. Suzuki, and M. Matsuhashi.** 1988. Determinations of the DNA sequence of the

mreB gene and of the gene products of the *mre* region that function in the formation of the rod shape of *Escherichia coli* cells. J. Bacteriol. **170:**4619–4624.
13. **Donnenberg, M. S., J. A. Giron, J. P. Nataro, and J. B. Kaper.** 1992. A plasmid-encoded type IV fimbrial gene of enteropathogenic *Escherichia coli* associated with localized adherence. Mol. Microbiol. **6:**3427–3437.
14. **Finkelstein, R. A., and S. Mukerjee.** 1963. Haemagglutination. A rapid method for the differentiating *Vibrio cholerae* from El Tor vibrios. Proc. Soc. Exp. Biol. Med. **112:**355–359.
15. **Freter, R., and G. W. Jones.** 1976. Adhesive properties of *Vibrio cholerae*: nature of the interaction with intact mucosal surfaces. Infect. Immun. **14:**2246–2256.
16. **Giron, J. A., O. G. Gomez-Duarte, K. G. Jarvis, and J. B. Kaper.** 1997. Longus pilus of enterotoxigenic *Escherichia coli* and its relatedness to other type-4 pili—a minireview. Gene **192:**39–43.
17. **Giron, J. A., M. M. Levine, and J. B. Kaper.** 1994. Longus: a long pilus ultrastructure produced by human enterotoxigenic *Escherichia coli*. Mol. Microbiol. **12:**71–82.
18. **Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J. Bacteriol. **177:**4121–4130.
19. **Hall, R. H., F. M. Khambaty, M. H. Kothary, S. P. Keasler, and B. D. Tall.** 1994. *Vibrio cholerae* non-O1 serogroup associated with cholera gravis genetically and physiologically resembles O1 El Tor cholera strains. Infect. Immun. **62:**3859–3863.
20. **Hase, C. C., M. E. Bauer, and R. A. Finkelstein.** 1994. Genetic characterization of mannose-sensitive hemagglutinin (MSHA)-negative mutants of *Vibrio cholerae* derived by Tn*5* mutagenesis. Gene **150:**17–25.
21. **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.
22. **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.
23. **Howard, S. P., J. Critch, and A. Bedi.** 1993. Isolation and analysis of eight *exe* genes and their involvement in extracellular protein secretion and outer membrane assembly in *Aeromonas hydrophila*. J. Bacteriol. **175:**6695–6703.
24. **Jahagirdar, R., and S. P. Howard.** 1994. Isolation and characterization of a second *exe* operon required for extracellular protein secretion in *Aeromonas hydrophila*. J. Bacteriol. **176:**6819–6826.
25. **Jonson, G., J. Holmgren, and A. M. Svennerholm.** 1991. Identification of a mannose-binding pilus on *Vibrio cholerae* El Tor. Microb. Pathog. **11:**433–441.
26. **Jonson, G., M. Lebens, and J. Holmgren.** 1994. Cloning and sequencing of *Vibrio cholerae* mannose-sensitive haemagglutinin pilin gene: localization of *mshA* within a cluster of type 4 pilin genes. Mol. Microbiol. **13:**109–118.
27. **Jouravleva, E. A., G. A. McDonald, J. W. Marsh, R. K. Taylor, M. Boesman-Finkelstein, and R. A. Finkelstein.** 1998. The *Vibrio cholerae* mannose-sensitive hemagglutinin is the receptor for a filamentous bacteriophage from *V. cholerae* O139. Infect. Immun. **66:**2535–2539.
28. **Karaolis, D. K. R., J. A. Johnson, C. C. Bailey, E. C. Boedeker, J. B. Kaper, and P. R. Reeves.** 1998. A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. Proc. Natl. Acad. Sci. USA **95:**3134–3139.
29. **Kim, S. R., and T. Komano.** 1997. The plasmid R64 thin pilus identified as a type IV pilus. J. Bacteriol. **179:**3594–3603.
30. **Kovach, M. E., M. D. Shaffer, and K. M. Peterson.** 1996. A putative integrase gene defines the distal end of a large cluster of ToxR-regulated colonization genes in *Vibrio cholerae*. Microbiology **142:**2165–2174.
31. **Lindeberg, M., and A. Collmer.** 1992. Analysis of eight *out* genes in a cluster required for pectic enzyme secretion by *Erwinia chrysanthemi*: sequence comparison with secretion genes from other gram-negative bacteria. J. Bacteriol. **174:**7385–7397.
32. **Maloy, S. R., V. J. Stewart, and R. K. Taylor.** 1996. Genetic analysis of pathogenic bacteria. A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Marsh, J. W., D. Sun, and R. K. Taylor.** 1996. Physical linkage of the *Vibrio cholerae* mannose-sensitive hemagglutinin secretory and structural subunit gene loci: identification of the *mshG* coding sequence. Infect. Immun. **64:**460–465.
34. **Marsh, J. W., and R. K. Taylor.** 1998. Identification of the *Vibrio cholerae* type 4 prepilin peptidase required for toxin secretion and pilus formation. Mol. Microbiol. **29:**1481–1492.
35. **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.
36. **Nakai, K., and M. Kanehisa.** 1991. Expert system for predicting protein localization sites in gram-negative bacteria. Proteins **11:**95–110.
37. **O'Toole, G. A., and R. Kolter.** 1998. Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. Mol. Microbiol. **30:**295–304.
38. **O'Toole, G. A., and R. Kolter.** 1998. Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. Mol. Microbiol. **28:**449–461.
39. **Oudega, B., and F. K. De Graaf.** 1988. Genetic organization and biogenesis of adhesive fimbriae of *Escherichia coli*. Antonie Leeuwenhoek **54:**285–299.
40. **Parge, H. E., K. T. Forest, M. J. Hickey, D. A. Christensen, E. D. Getzoff, and J. A. Tainer.** 1995. Structure of the fibre-forming protein pilin at 2.6 Å resolution. Nature **378:**32–38.
41. **Pugsley, A. P.** 1993. The complete general secretory pathway in gram-negative bacteria. Microbiol. Rev. **57:**50–108.
42. **Reeves, P. J., D. Whitcombe, S. Wharam, M. Gibson, G. Allison, N. Bunce, R. Barallon, P. Douglas, V. Mulholland, S. Stevens, et al.** 1993. Molecular cloning and characterization of 13 *out* genes from *Erwinia carotovora* subspecies carotovora: genes encoding members of a general secretion pathway (GSP) widespread in gram-negative bacteria. Mol. Microbiol. **8:**443–456.
43. **Sancar, A., K. R. Williams, J. W. Chase, and W. D. Rupp.** 1981. Sequences of the *ssb* gene and protein. Proc. Natl. Acad. Sci. USA **78:**4274–4278.
44. **Sandkvist, M., V. Morales, and M. Bagdasarian.** 1993. A protein required for secretion of cholera toxin through the outer membrane of *Vibrio cholerae*. Gene **123:**81–86.
45. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
46. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. Gene **53:**85–96.
47. **Skorupski, K., and R. K. Taylor.** 1996. Positive selection vectors for allelic exchange. Gene **169:**47–52.
48. **Skorupski, K., and R. K. Taylor.** 1997. Control of the ToxR virulence regulon in *Vibrio cholerae* by environmental stimuli. Mol. Microbiol. **25:**1003–1009.
49. **Skorupski, K., and R. K. Taylor.** 1997. Cyclic AMP and its receptor protein negatively regulate the coordinate expression of cholera toxin and toxin-coregulated pilus in *Vibrio cholerae*. Proc. Natl. Acad. Sci. USA **94:**265–270.
50. **Strom, M. S., and S. Lory.** 1993. Structure-function and biogenesis of the type IV pili. Annu. Rev. Microbiol. **47:**565–596.
51. **Tacket, C. O., R. K. Taylor, G. Losonsky, Y. Lim, J. P. Nataro, J. B. Kaper, and M. M. Levine.** 1998. Investigation of the roles of toxin-coregulated pili and mannose-sensitive hemagglutinin pili in the pathogenesis of *Vibrio cholerae* O139 infection. Infect. Immun. **66:**692–695.
52. **Taylor, R. K., C. Manoil, and J. J. Mekalanos.** 1989. Broad-host-range vectors for delivery of Tn*phoA*: use in genetic analysis of secreted virulence determinants of *Vibrio cholerae*. J. Bacteriol. **171:**1870–1878.
53. **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 regulated with cholera toxin. Proc. Natl. Acad. Sci. USA **84:**2833–2837.
54. **Thelin, K. H., and R. K. Taylor.** 1996. Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by *Vibrio cholerae* O1 El Tor biotype and O139 strains. Infect. Immun. **64:**2853–2856.
55. **Tonjum, T., N. E. Freitag, E. Namork, and M. Koomey.** 1995. Identification and characterization of *pilG*, a highly conserved pilus-assembly gene in pathogenic *Neisseria*. Mol. Microbiol. **16:**451–464.
56. **Tonjum, T., and M. Koomey.** 1997. The pilus colonization factor of pathogenic neisserial species: organelle biogenesis and structure/function relationships. Gene **192:**155–163.
57. **Wachi, M., and M. Matsuhashi.** 1989. Negative control of cell division by *mreB*, a gene that functions in determining the rod shape of *Escherichia coli* cells. J. Bacteriol. **171:**3123–3127.
58. **Waldor, M. K., and J. J. Mekalanos.** 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science **272:**1910–1914.
59. **Watnick, P. I., and R. Kolter.** Personal communication.