Identification of Genes Encoding Components of the Swarmer Cell Flagellar Motor and Propeller and a Sigma Factor Controlling Differentiation of Vibrio parahaemolyticus

LINDA L. McCARTER* AND MIRIAM E. WRIGHT

Department of Bacteriology, The University of Wisconsin, Madison, Wisconsin 53706,1 and The Agouron Institute, La Jolla, California 920372

Received 9 December 1992/Accepted 30 March 1993

Vibrio parahaemolyticus possesses two distinct motility systems, the polar system used for swimming in liquid environments and the lateral system used for swarming over surfaces. Growth on surfaces induces swarmer cell differentiation and expression of the lateral motility system. Mutants, created by transposon mutagenesis of a clone expressing lateral flagellin and gene disruption in V. parahaemolyticus, were unable to swarm and failed to make lateral flagellin; therefore, unlike the case for the polar system, there is one gene (lafA) encoding lateral flagellin. In addition to lafA, other genes required for swarming but not for swimming were identified by gene replacement mutagenesis. The nucleotide sequence of the clone determined open reading frames (ORFs) and deduced amino acid sequences showed similarities to flagellar components of other bacteria: flagellin, hook-associated protein (HAP2), motor components, and flagellar sigma factor (σ28). Many σ28 factors have been shown to recognize cognate promoters; however, expression of lafA in Escherichia coli required LafS, and E. coli σ28 did not substitute. Also, there were no sequences preceding genes encoding flagellin or HAP2 resembling the σ28 consensus promoter. The product of the sigma-like gene seems to be a unique member of the σ28 cluster. It appears the result of requiring expression for immunodetection of flagellin clones was that the sigma locus was fortuitously cloned, since the sigma and lafA loci were not contiguous in the chromosome. This work initiates identification and placement of genes in a scheme of control for swarmer cell differentiation; three levels have been identified in the transcriptional hierarchy.

The marine bacterium and human pathogen Vibrio parahaemolyticus possesses two distinct motility systems, each appropriate for movement under different circumstances (2, 48). Translocation over surfaces, or swarming, results in radially spreading colonies and colonization of surfaces, while movement in liquid is called swimming (20, 52). The polar flagellar system (Fla) propels the bacterium in liquid medium. Growth on surfaces induces the developmental program called swarmer cell differentiation, in which septation ceases, the cells elongate, and a new flagellar system, the lateral system (Laf), is synthesized. The polar system is expressed essentially constitutively, while the lateral flagellar system is expressed only when the bacterium is in contact with surfaces or in a viscous environment, i.e., under conditions in which the polar flagellum is not functional (38). Thus, the two gene systems interact: performance of the polar system is in some way coupled to transcription of the swarmer cell gene system. To elucidate the signal transduction event by which physical information is processed and converted to a signal that leads to swarmer cell differentiation, we are defining some of the key genes and the regulatory hierarchies of the two flagellar systems.

The gene systems that encode the two motility systems are large, each composed of 40 or more genes without overlap except for shared chemotaxis components (41, 47). Flagellum-driven motility requires genes encoding components of a propeller and a motor (34). The flagellum is composed of a basal body which appears as a set of rings mounted on a rod that extends from the cytoplasm through the cell wall, a hook which couples the filament to the basal body, and the propeller, or filament, which is a hollow, semirigid helical tube that polymerizes at the tip. Proteins called HAPs (hook-associated proteins) are required for joining the filament to the hook and capping the distal tip of the filament. The mystery of flagellar assembly, in particular the nature of the flagellum-specific export pathway which is distinct from the signal sequence-dependent pathway utilized for most exported bacterial proteins, has not been solved (22, 23, 31). The motor is powered by the electrochemical membrane potential. It couples the movement of ions through a channel (which is formed by MotA in Escherichia coli) to rotation of the propeller. Another protein (called MotB in E. coli) is believed to function as a stator. Together, these two proteins are believed to act as force-generating units which transmit torque to the filament through the basal body. Additional proteins include components of the switch complex, which is important for modulation of direction of flagellar rotation and also required for both organelle assembly and energy conversion.

In this report, we begin to identify genes belonging to the lateral flagellar system and to order these genes in a hierarchical scheme of gene control. The genes encode structural components of the flagellum and the motor, as well as a sigma factor of RNA polymerase important for directing swarmer cell development. Sequence information on structural components of both motility systems should provide insight with respect to the signals for flagellar export and assembly, laf and fla promoter recognition and regulatory sites, and perhaps evolutionary clues for these two distinct systems. Defining the cascade of regulation should help us begin to understand how the lateral and polar gene systems interact.

* Corresponding author. Electronic mail (Internet) address: mccarter@psl.wisc.edu.

3361
MATERIALS AND METHODS

Bacterial strains, plages, and vectors. The wild-type *V. parahaemolyticus* strain BB2 and strain LM1017 (lafX131::lax) have been described previously (5, 38). ΔTnphoA and *E. coli* CC118 [araD139 Δ ara leu]7697 ΔlacX74 ΔphoA20 galE galK thi rpsE tpoB argE(Alm) recA1 were obtained from Colin Mangoi (17). *E. coli* MC1000 [araD139 Δ ara leu]7697 ΔlacX74 galU galK strA] and MC1000 (thiD::Tn10 Kan') were from Phil Matsumura (9), and *E. coli* YK410 (araD139 Δ lacU169 strA thi pyrC46 naA thyA his) and YK4181 (YK410 fla4) were from May Macnab (28). Cosmid pLM1457 has been described elsewhere (39), and cosmid pLM1489 contains the *V. parahaemolyticus* polar flagellin genes flaC and flaD in the vector pRK415 (25).

Growth conditions and media. *Vibrio* strains were propagated at 30°C in heart infusion medium (25 g of Bacto Heart Infusion [Difco] plus 20 g of NaCl per liter). Swarming was examined on solidified heart infusion media with 1.5% Bacto Agar (Difco). The minimal medium of Broach et al. (8) supplemented with 0.4% galactose, 20 mM NH4Cl, and 2% NaCl was used for conjugation and semisolid motility plates with 1.5 and 0.36% Bacto Agar, respectively. LB, 2xYT, and NZCYM media for propagation of *E. coli* strains and lambda phage were prepared as described by Maniatis et al. (35). Antibiotics (from Sigma Chemical Co.) were used at the following concentrations: 125 μg/ml (gentamicin), 100 μg/ml (kanamycin), and 10 μg/ml (tetracycline). 5-Bromo-4-chloro-3-indolyl phosphate (XPI Sigma) was used at 40 μg/ml. TnphoA and gene replacement mutagenesis. Mutagenesis with ΔTnphoA was performed by infecting a culture of *E. coli* CC118 carrying cosmid pLM1457 (tetracycline resistant). Multiple 2-ml cultures of exponentially growing cells (optical density at 600 nm of 0.4) in NZCYM medium supplemented with tetracycline were incubated with 25 μl of ΔTnphoA (at titer of >1011 PFU/ml) for 30 min to allow adsorption. Then, 2 ml of 2xYT plus tetracycline was added to each tube, and cultures were grown for 2 h at 37°C with shaking, at which time 2 ml of 2xYT plus tetracycline, kanamycin, and 0.4% glucose was added to each tube, and cells were grown overnight. Cultures were reincubated into selective media (optical density at 600 nm of 0.1) and grown to a density of 108 CFU/ml. DNA was prepared and used to transform strain CC118, selecting for tetracycline and kanamycin resistance. The procedure for gene replacement in *V. parahaemolyticus* has been outlined elsewhere (49). For mobilization of the cosmids carrying the transposon mutations into *V. parahaemolyticus*, the tra donor was pRK2013 (from Gary Ditta [15]); for selecting gene replacement, the "kick-out" plasmid was pPHJ1 (from Ron Taylor [6]). Counterselection against the *E. coli* parents was done on minimal medium with galactose as the carbon source and appropriate antibiotics for plasmid and cosmid maintenance. Strain constructions were confirmed by Southern blot analysis of restricted genomic DNA (37, 40). Transposon insertions in the *lafA* locus did not perturb the *lafA* locus and vice versa.

Immunological techniques. Western immunoblot analysis with anti-LafA antiserum 127 and anti-Fla antiserum 129 was done as previously described (39) except that proteins were transferred to 0.45-μm pore-size Magna nylon transfer membranes (MSI Micron Separations Inc.). Rabbit antibody to bacterial alkaline phosphatase was obtained from 5 Prime 3 Prime (West Chester, Pa.).

Recombinant DNA techniques. Transformations, ligations, and other general procedures were performed by the method of Maniatis et al. (35). Restriction endonucleases and T4 DNA ligase were obtained from Boehringer Mannheim Biochemicals and were used according to the manufacturer's instructions. DNA transfer for Southern blot analysis was to MS13 nylon membranes.

DNA sequencing and analysis. Recombinant DNA from pLM1457 was subcloned into M13 vector mp18 or mp19 (42). These subclones and deletions of the subclones, made with the Cyclone I Biosystem (International Biotechnologies, Inc.), were used as sequencing templates. DNA sequence was obtained for both strands, using the dideoxy-chain termination procedure of Sanger et al. (46) with the Sequenase 2.0 kit from United States Biochemical. The source of radioactivity was [α-35S]dATP (DuPont New England Nuclear). Synthetic oligonucleotides were prepared by Genosys Biotechnologies, Inc. (The Woodlands, Tex.). Sequence assembly was performed with the Lark package (EuGene and SAM version 3.2; Baylor College of Medicine, Waco, Tex.); for sequence analysis, both the EuGene program and Genetics Computer Group software package version 7.0 were used. Searches for homology were performed at the National Center for Biotechnology Information by using the Blast network service (3). Homology was examined by using a modified Dayhoff scoring matrix in the Lark and Dotplot analysis of the Genetics Computer Group packages.

Nucleotide sequence accession number. The sequence shown in Fig. 3 from nucleotides 1 to 3724 (the BglII site) has been assigned GenBank accession number L06176. The sequence shown in Fig. 4 from nucleotides 1 to 2438 (the sau3A site) has been assigned GenBank accession number L06177.

RESULTS

Cloning and transposon mutagenesis. Cloning of the locus encoding lateral flagellin has been described elsewhere (39). Briefly, a lambda bank was screened for production of lateral flagellin antigen. This bank was constructed by ligating size-selected Sau3A partial digests of *V. parahaemolyticus* chromosomal DNA in the XEMBL3 vector cleaved with BamHI. One of the antigen-producing clones was subcloned into the broad-host-range vector pLAFRII. This plasmid, pLM1457, was mutagenized with TnphoA in order to localize the structural gene, *lafA*, for sequencing and to create *lafA* mutations to be used for gene replacement in *V. parahaemolyticus*. From this mutagenesis, cosmids that failed to produce Lafa antigen of the appropriate molecular size (29 kDa) were obtained. Strains carrying some of these mutated cosmids (e.g., with transposon 1494 [Fig. 1, lane 2]) synthesized fusion proteins of a higher molecular weight that cross-reacted in Western analysis with both anti-Lafa and anti-alkaline phosphatase antisera. Intriguingly, we obtained other cosmids with transposons (e.g., transposon 1694 in [Fig. 1, lane 7]) that also failed to produce Lafa but which on restriction analysis mapped quite distal to the first class. Given the positions and orientations of the transposons (Fig. 2), it did not seem likely that mutants of the second class were simply polar upon *lafA* expression, and we postulated that there was a gene (or genes) on cosmids pLM1457 necessary for expression of *lafA*. One other interesting class of transposon insertions, represented by transposon 1692, resulted in overproduction of Lafa (Fig. 1, lane 8). Lateral flagellin appeared to be unstable in *E. coli* (long exposures of Western blots showed degradation in all strains producing Lafa); however, it was most apparent for strains carrying cosmids with transposons such as transposon 1692, which
overproduced LaF. Since there appeared to be more functions pertinent to lateral flagellar gene expression encoded by this clone than simply the flagellin structural gene, the entire region was sequenced. A restriction map, with positions of some of the transposons indicated, and the sequencing strategy are shown in Fig. 2.

The lafA gene. The DNA sequence and the translated products of the proposed reading frames presented in Fig. 3 correspond to bp 1 to 3768 on the restriction map in Fig. 2. The open reading frame (ORF) encoding lateral flagellin, LaF, starts at position 739 with an ATG. It is preceded by sequences with features that resemble a site for translational initiation in E. coli, including a Shine-Dalgarno domain which has appropriate spacing preceding the initiation codon (16). The ORF is also preceded by two regions of inverted, dyad symmetry that show some sequence similarity to each other (indicated by the open circles in Fig. 3). There are no sequences resembling consensus flagellar promoters which have been identified preceding flagellar genes in a variety of organisms (and there are no sequences resembling ω70 or ωA consensus promoters [18]). The translated protein contains 284 amino acids with a calculated molecular size of 29,727 Da, which correlates well with the observed migration of LaF in sodium dodecyl sulfate (SDS)-polyacrylamide gels (39). LAF appears similar to other flagellins (24) in that it is quite rich in alanine (44 residues) and has only one proline, two histidines, and no cysteine or tryptophan residues. Most of the similarity occurs at the amino and carboxyl termini, with internal variabilities. Some of the flagellins with outstanding similarity scores include PSEFLAA from Pseudomonas aeruginosa, STYHIA from Salmonella paratyphi A, STYFLGHI from Salmonella typhimurium, ECOHAG from E. coli, BACFLAGA from Bacillus subtilis, SMAHAG from Serratia marcescens, TRPFLABP from Treponema phagedenis, BORFLA2 from Borrelia burgdorferi, and CAJ FLAAB from Campylobacter jejuni. For example, GCG BestFit comparison of LaF with PSEFLAA yielded the highest homology scores, with 69,4% similarity (S)/52% identity (I) (quality = 245 and quality ratio = 0,862).

Gene replacement experiments in the wild-type strain BB22 that substituted lafA::TpHoA fusions for the wild-type allele yielded mutant bacteria that were unable to swarm, failed to produce lateral flagellin subunit, and produced fusion proteins (i.e., proteins that cross-reacted with both alkaline phosphatase and LaF antisera). Polar motility (i.e., swimming in semisolid motility medium) was not affected. The apparent molecular weights on SDS-polyacrylamide gels of the fusion proteins were larger than that of either alkaline phosphatase or LaF. The fusions were regulated by growth on a surface; no fusion proteins were produced when mutant bacteria were grown in liquid me-

---

**FIG. 1.** Western blot analysis of LaF production in E. coli by clones with TpHoA insertions. Overnight cultures of E. coli CC118 carrying the following cosmids, which are derivatives of pLM1457 with transposon insertions, were concentrated twofold in sample buffer for SDS-polyacrylamide gel electrophoresis. The reference sample was wild-type V. parahaemolyticus BB22 harvested from a plate. Lanes: 0, BB22; 1, pLM1457 (parent, no insertion); 2, pLM 1494 (lafA::TpHoA [note high-molecular-weight fusion product]); 3, pLM1687 (intergenic lafA and lafB); 4, pLM1688 (lafB::TpHoA); 5, pLM1498 (lafU::TpHoA); 6, pLM1690 (lafT::TpHoA); 7, pLM1694 (lafS::TpHoA); 8, pLM1692 (lafL::TpHoA); 9, pLM1457 (parent). The blot was probed with anti-lateral flagellin antiserum and subsequently with radiolabeled protein A.

**FIG. 2.** Restriction map and sequencing strategy. Restriction sites shown are the ones used for cloning from cosmid pLM1457 into M13 mp18 or mp19. Abbreviations: B, BamHI; P, PstI; S, SalI; X, XbaI. The following restriction fragments were used for construction of the sequencing templates: 1, B-X (1.2 kb); 2, P-P (0.6 kb); 3, P-P (2.3 kb); and 4, S-S (2.4 kb). In addition to these templates, deletions of these clones were prepared for sequencing by using the International Biotechnologies Cyclone I Biosystem. The junctions (a) between templates 2 and 3 and templates 3 and 4 were crossed by priming the parental full-length clone (pLM1457, double stranded) with synthetic oligonucleotides homologous to sequences at one end of template 2 or 3.Overlapping sequences were obtained for both strands of DNA and are indicated by the arrows. ORFs corresponding to the genes indicated proceeded from left to right for lafA and lafB and right to left for lafl, lafS, lafT, and lafU. Only parts of the ORFs corresponding to lafl and lafu were cloned and sequenced. Representative TpHoA insertions are shown on the restriction map; filled circles indicate those insertions that form alkaline phosphatase fusion proteins, and arrows indicate the direction of transcription of phoA.
FIG. 3. Nucleotide sequences and deduced amino acid sequence for the laf4 locus. Sites that resemble E. coli ribosome binding sites (RBS) are indicated by underlined sequences which exhibit homology with the 3′ end of E. coli 16S rRNA. Arrows indicate inverted regions of dyad symmetry, and open circles indicate homologous nucleotides between these regions. Chevrons indicate a potential stem-loop structure that could serve as a transcriptional terminator. Selected restriction sites are shown to facilitate orientation to the physical map. Sequence presented corresponds to bp 1 to 3773 in Fig. 2. The coding sequences extend from nucleotides 739 to 1596 for laf4 and from 1369 to 3706 for laf5. The ORF preceding laf4 (nucleotides 1 to 431) showed no homological knowledge when GenBank was searched by using TBLASTN.
The *lafl* gene. The region between *lafl* and the next ORF extends from nucleotides 1596 to 2369 in Fig. 3. This ORF, designated *lafB*, starts with a GTG codon preceded by sequences resembling the Shine-Dalgarno sequence at a distance of 7 bp. The *lafl* and *lafB* genes do not appear to be organized in a transcriptional unit. There are sequences resembling an *E. coli* rho-independent transcriptional terminator following the *lafl* coding region (positions 1631 to 1661). Moreover, disruption of the intergenic region did not affect swarmer cell gene expression, since a strain with transposon 1697 introduced by gene replacement was able to swarm. The *lafB* gene is preceded by one region of inverted dyad symmetry that shows sequence similarity (indicated by the open circles in Fig. 3) to the two dyad regions preceding *lafl*. The deduced protein sequence is 445 amino acids in length and has a calculated molecular weight of 48,234. When the *lafB* gene was disrupted by gene replacement with *lafl*::Tphoa insertions, the mutants failed to swarm. A TBLASTN search revealed homology to STYFLID, the *S. typhimurium* fliD gene product (the distal capping protein called HAP2). Like other flagellar axial proteins which polymerize to form the flagellar basal body and filament, this protein lacks cysteine, has a 1.8% proline content, and possesses heptad repeats of hydrophobic amino acids at the carboxyl terminus (21). The best identity between the *Salmonella* and *Vibrio* proteins occurs within the carboxyl-terminal half, consistent with the putative function and location of the protein in the flagellar structure. As the distal capping protein, one end of the molecule interacts with flagellum and the other is exposed to the environment (21).

The lateral flagellar sigma gene (lafS). The DNA sequence presented in Fig. 4 corresponds to the reverse complement of the restriction map in Fig. 2, correlating with positions 6218 to 3762. Within this sequence is an ORF, designated *lafs*, that starts at nucleotide 240 with an ATG codon. It is preceded at a distance of 7 bases by sequences resembling the Shine-Dalgarno consensus. This ORF codes for a potential polypeptide of 242 amino acids with a calculated molecular size of 27,817 Da. Introduction of transposons, such as transposon 1694, into the *V. parahaemolyticus* chromosome altered swarming: *lafs*::Tphoa mutants were unable to swarm and failed to synthesize lateral flagellin. Polar motility was unaffected. When the deduced protein sequence was used to search the GenBank data base, the following similarities to proteins approximately 28 kDa in size were found: STYFLIA, the *S. typhimurium* flagellar sigma factor (58% S/33% I by BestFit analysis); PSEFLIA, the *P. aeruginosa* flagellar sigma factor (56% S/31% I); BACSIGD, the *B. subtilis* flagellar sigma factor (48% S/28% I); STMSWHIG, the *Streptomyces coelicolor* developmental sigma factor (50% S/27% I); and STMPROZ, a homolog of STMSWHIG and a minor sigma factor of *Streptomyces aureofaciens* (52% S/28% I). Three of these proteins are alternative sigma subunits of RNA polymerase that are required for flagellin synthesis (19, 45, 51). The fourth, STMSWHIG (and probably STMPROZ), is a factor required for initiation of spore development (10). In multiple sequence alignments, the other sigma factors in the $\sigma^{26}$ cluster appear more closely related to each other than to *V. parahaemolyticus*.

The sigma subunit of RNA polymerase confers the specificity of promoter selection. Inspection of the homology between the above sigma factors reveals two regions with greater than 70% similarity (Fig. 5). These two regions of similarity map in the consensus regions of alignment, designated 2.4 and 4.2, that have been identified upon comparison of many sigma factors (33). Domain 2.4, which encompasses region I in Fig. 5, is believed to be important in $-10$ recognition of promoter sequences by RNA polymerase, while domain 4.2, which encompasses region II, has been implicated as being important for $-35$ hexamer recognition. The *V. parahaemolyticus* sequence seems significantly different. In region I, the *V. parahaemolyticus* sequence varies at two positions (marked with asterisks) which were conserved among all of the other sigma factors shown. In region II, *V. parahaemolyticus* varies at six positions where the others remain identical.

Expression of the lateral flagellar gene *lafl* in *E. coli* is independent of the *E. coli* flagellar sigma factor. Many of the alternative sigma factors in the $\sigma^{26}$ cluster have been shown to recognize similar (and, it has been demonstrated in some cases, interchangeable) promoters (4.4, 18). Moreover, expression of *lafS* was examined in various *E. coli* flagellar mutant strains. Western blot analysis suggested that the *E. coli* flagellar genes, flaA, encoding the flagellar sigma, and the master regulatory genes (flhCD) that control expression of flaA, were not required for expression of *lafl* on clone pLM1457 in *E. coli*. Production of LaF was compared in the following pairs of wild-type and mutant strains: MC1000/MC1000 fliD::Tn10 (FliD $^-$ FlhC$^+$) and YK410/YK4181 (YK410 flaA). The level of expression of *lafl* from cosmid pLM1457 was similar in *E. coli* wild-type and flagellar regulatory mutant strains (Fig. 6, lanes 1 to 4). The data suggest that no *E. coli* flagellar proteins were required for expression of flaA. In contrast, expression of *V. parahaemolyticus* polar flagellin genes flaC and flaD, which are encoded by cosmid pLM1489, appeared to require *E. coli* flagellar genes for expression (lanes 5 to 8). It is unlikely that expression of *lafl* in *E. coli* is under the control of a cosmid promoter, since mutations in the lateral flagellar sigma gene (e.g., transposon 1694), which map on cosmid pLM1457 downstream of the direction of transcription of the *lafl* gene (Fig. 2), resulted in loss of expression of *lafl* (Fig. 1, lane 7). Moreover, other transposon insertions resulted in overproduction of LaF in *E. coli* (an example of which is shown in lane 8 of Fig. 1). These transposons, which all have the same orientation of insertion mapped upstream of the *lafs* gene (see transposon 1692 in Fig. 2). It seems likely that increased production of *lafl* was the result of increased transcription of *lafs* promoted by the transposon insertion. We conclude that expression of the gene encoding lateral flagellin in *E. coli* is dependent on expression of another *Vibrio* gene, *lafs*, which encodes a protein that resembles a flagellar sigma factor of RNA polymerase, and that the *E. coli* flagellar sigma cannot substitute for function of LaF.

The lateral flagellar motor genes (*lafT* and *lafJ*). Located 11 bases downstream of the *lafs* gene is an ORF, designated *lafT*, that initiates with an ATG preceded at a distance of 7 bases by sequences that resemble the Shine-Dalgarno sequence. The predicted protein encoded by the gene sequence contains 285 amino acids with an estimated molecular size of 31,304 Da. Its overall hydrophobicity is consistent with it being a membrane protein (51% Ala, Leu, Ile, Val, Phe, Met, and Pro). The protein is homologous to the *E. coli* motA gene product: BestFit analysis yields 57.5% S/35.8% I (quality = 210 and quality ratio = 0.74). More-
over, upon comparison with the *E. coli* protein, six domains seem conserved (12). There are four regions, each composed of 20 to 25 mainly hydrophobic amino acids (with 1 or 2 charged amino acids per domain), that could form membrane-spanning segments according to calculations of Eisenberg et al. (14) and Kyte and Doolittle (32). Integral to the molecule are two hydrophilic, highly charged regions. The first charged region (including amino acids 52 to 73) contains four acidic, four basic, and two glutamine residues. The second charged region (amino acids 142 to 157) contains six acidic, one basic, and two glutamine residues.

The termination codon, TAA, of *lafS* overlates the initiation codon, ATG, of another ORF, designated *lafU*. This ATG is preceded at a distance of 4 bases by sequences resembling the Shine-Dalgarino sequence. The deduced protein sequence revealed homology to the *E. coli* motB gene product (BestFit analysis: 53.7% S/31.1% I; quality = 111 and quality ratio = 0.59). Three domains show conservation to domains noted for MotB (11, 50), while the central region of the molecule appears more variable. At the amino termini, there is a hydrophilic region (basic 7, acidic 3, and 2 glutamine residues in the first 23 amino acids) followed by a hydrophobic domain (including amino acids 26 to 49). This domain has characteristics of a membrane-spanning domain in terms of hydrophobicity, length, and potential α-helical structure. There is a second hydrophobic region located near the carboxyl terminus (corresponding to region C in reference 11). This region is 22 amino acids in length, long enough to span a membrane; however, it has 4 charged residues. As was observed for *motB*: TphoA fusion strains and consis-
FIG. 5. Comparison of deduced amino acid sequences for LaFS and alternative motility and developmental sigma factors. These two regions correspond to regions believed to be important for promoter recognition. Capital letters indicate amino acids of high similarity; lowercase letters indicate little or no similarity. For the consensus, specific letters indicate identity, + indicates high similarity, and * indicates cases where LaFS is different but the other sequences are identical.

**FIG. 6.** Expression of lafA in *E. coli* flagellar mutants. For Western blot analysis, overnight cultures of strains were concentrated twofold in sample buffer for SDS-polyacrylamide gel electrophoresis. Control lane 0 contains wild-type *V. parahaemolyticus* BB22 harvested from a plate. Both lateral (Laf) and polar (Fla) flagellins are produced. Three of the four polar flagellins (FlaA, FlaC, and FlaD) comigrate, and the fourth (FlaB) appears slightly smaller. Lanes: 0, strain BB22; 1, MC4100/pLM1457 (lafA+); 2, MC4100 flbD::Tn10/pLM1457; 3, YK410/pLM1457; 4, YK4181 (YK410 flaA) pLM1457; 5, MC4100/pLM1489 (flaC+ flaD+); 6, MC4100 flbD::Tn10/pLM1489; 7, YK410/pLM 1489; 8, YK4181 (YK410 flaA) pLM1489. Cross-reacting *E. coli* proteins in the background serve as controls for total amounts of cell protein loaded in each lane.

FIG. 7. When pLM1457 was used as a probe, the 0.6-kb *PsiI* fragment 1 was present in the cosmid and the chromosome (lanes 4 and 5), and this fragment was perturbed (lane 6) in a mutant strain carrying transposon 1494, which maps in the 0.6-kb *PsiI* fragment. The 1.0-kb *BglII* fragment was also present in the cosmid and chromosome (lanes 2 and 3), and this fragment was unperturbed in strains with insertions mapping in the *lafA* or *laFS* locus (lanes 9 through 14). Mutant strains with insertions in the *lafA* locus (transposons 1494 and 1687) showed perturbation of one large *BglII* fragment in the chromosome (lanes 11 and 12 versus the wild type in lane 10), and insertions in the *laFS* region (transposons 1690 and 1694) caused perturbation of a different fragment (lanes 13 and 14). Neither the 2.3-kb *PsiI* fragment 2 (lanes 4 and 5) nor the 2.4-kb *SalI* fragment (lanes 7 and 8) of the cosmid was found in the chromosome. Both of these fragments span the putative chimeric joint. Thus, either two unrelated pieces of DNA were fortuitously joined together or an original, larger clone was deleted to produce pLM1457 and the other positive clones identified. This was an unexpected result given that *λ*EMBL3 accepts 22-kb inserts and that the lambda bank was made with 16- to 22-kb fragments of DNA purified by agarose gel electrophoresis; however, it is a likely result because the antibody screen was based on selection for expression of *lafA*. It is possible that *lafA* expression requires the *laFS* locus, and thus, in effect, the screen was for the two regions to be contained on the same vector.

**The lafl gene.** The *laFS* gene is preceded by a small ORF, with the intercistronic region being 12 bases. This ORF, designated *lafl*, appears to encode 74 residues at the carboxyl terminus of a protein that shows homology to the FilL proteins of other bacteria; for example, BestFit analysis reveals approximately 52% S/S% homology with the proteins of *E. coli* and *S. typhimurium*, 49% S/S% with *Caulobacter crescentus* protein, and 45% S/S% with *B. subtilis* protein (1, 26, 29, 53). These *flIL* genes encode membrane proteins approximately 150 amino acids in length. Although the function of *flIL* is unknown, in *E. coli* and *S. typhimurium*, *flIL* is found in an operon that encodes other proteins that are members of the flagellar switch complex which is involved in energizing and switching the direction of rotation of the flagellum. Introduction of transposon mutation 1692 into the *V. parahaemolyticus* chromosome affected swimming: the *lafl::TnphoA* mutant was unable to swim and failed to synthesize lateral flagellin. Polar motility was unaffected.

**Hierarchy of control of lateral flagellar gene expression.** Genes in the swarmer cell system are expressed in response to growth on a surface. Expression of the *lafA* gene, which encodes the flagellar filament protein, requires expression of the *lafA STUV* locus encoding the sigma subunit and motor parts. Introduction of mutations in any of the identified genes in this second locus resulted in loss of swimming and loss of production of lateral flagellin in *V. parahaemolyticus*. Thus, *lafA* is placed in a scheme of regulation under the control of the sigma locus (Fig. 8). One other lateral flagellar locus has been studied (5, 38); it encodes a gene, that we now designate *lafX*, which shows surface-dependent gene expression. Strain LM1017 contains an operon fusion between the *lafX* gene and indicator luminescence genes (*lux*). It is...
unable to swarm, fails to make lateral flagellin, and produces bioluminescence only when grown on surfaces and not in liquid media (38). Introduction of sigma mutations, such as mutation 1694, into strain LM1017 did not affect expression or surface-dependent regulation of the laf::lux fusion. It appears that regulation of the lafX gene occurs at a point higher in the cascade of gene control; therefore, it is placed above the sigma locus in the hierarchy.

**DISCUSSION**

*V. parahaemolyticus* has two alternative cell types, the swimmer cell and the swarmer cell. The swimmer cell is adapted for survival in liquid environments; the rod-shaped bacterium is propelled by a single sheathed, polar flagellum. When a swimming cell contacts a surface, a series of changes occur that transforms the swimmer to a swarmer cell. The swarmer cell is adapted for colonization of surfaces; these highly elongated, superflagellated cells can rapidly move over and colonize surfaces. The two propulsive systems are distinct, having no motor or propeller components in common. Somehow the bacterium is able to assemble two distinct appendages simultaneously, and we are interested in the signals that direct export and morphogenesis of the polar and lateral organelles. Previous work has shown that the composition of the polar filament is complex; i.e., there are multiple subunits in the polymer, and the filament is sheathed by an extension of the cell outer membrane (2, 38). The polar flagellum is powered by the sodium motive force (4). The lateral flagella show more resemblance to *E. coli* flagella in that they are unsheathed, arranged peritricously, and powered by the proton motive force (1, 4). Purified

**FIG. 7.** Southern blot analysis of restricted cosmids and chromosomal DNA prepared from *V. parahaemolyticus* wild-type and transposon mutant strains probed with labeled pLM1457, showing that cosmids pLM1457 is a chimera (the lafA and lafS loci are not contiguous on the chromosome). Arrows at the left indicate approximate sizes in kilobase pairs; arrows at the right indicate fragments in restricted wild-type DNA that are perturbed in DNA prepared from transposon mutant strains. Lanes: 1, lambda DNA markers (3.5-kb band hybridizes to the probe); 2, cosmID pLM1457 restricted with *Bgl*II; 3, wild-type strain BB22 restricted with *Bgl*II; 4, cosmID pLM1457 restricted with *Pst*I; 5, BB22 restricted with *Pst*I; 6, mutant strain LM4066 (with transposon 1494) restricted with *Pst*I; 7, cosmID pLM1457 restricted with *Sal*I; 8, BB22 restricted with *Sal*I; 9, cosmID pLM1457 restricted with *Bgl*II; 10, BB22 restricted with *Bgl*II; 11, mutant strain LM4066 (with transposon 1494) restricted with *Bgl*II; 12, mutant strain LM4247 (with transposon 1687) restricted with *Bgl*II; 13, mutant strain LM4251 (with transposon 1690) restricted with *Bgl*II; 14, mutant strain LM4242 (with transposon 1694) restricted with *Bgl*II. The diagram shows the chimeric joint and selected restriction fragments from pLM1457. The asterisks indicate fragments that do not correspond with the chromosome. Locations of the transposons used to created specific mutant strains are shown. The chimeric joint is indicated by the wavy line. Abbreviations: *P*, *Pst*I; *G*, *Bgl*II; *S*, *Sal*I; *A*, Sau3A.

**FIG. 8.** Scheme of control of surface-dependent gene expression for the lateral flagellar regulon. Expression of the lafA gene, encoding lateral flagellin, requires function of all of the genes presented at a higher position in the diagram. Mutations in the lafS locus or the lafX gene result in loss of the ability to synthesize lateral flagellin and inability to swarm on solidified medium. The lafX gene is placed at the highest position as yet identified because it retains its normal, surface-dependent regulation in the presence of mutations placed lower in the scheme (i.e., lafS). The lafS gene encodes a sigma factor, and the lafT and lafU genes encode motor parts. They appear to be part of a large operon, designated by the arrow.
lateral flagellar filament appears to be composed of only a single flagellin subunit (39). Purified polar flagellar filament appears to be composed of two flagellin subunits; however, disruption of one of the genes encoding filaments in *V. parahaemolyticus* indicates that there were four genes (38). In this work, we have shown that gene disruption of *lafA* created a mutant that is unable to swarm and fails to synthesize lateral flagella. Thus, there is only a single gene encoding the subunit that polymerizes to form the lateral flagellar filament.

Flagellar morphogenesis is a complex cascade of events whereby gene expression is carefully coupled to assembly of the organelle (34). This hierarchy of gene regulation was first unraveled for *E. coli* by using gene fusions to *lacZ* (27) and has now been worked out for *S. typhimurium* (30) and *C. crescentus* (13, 43). For these bacteria, expression of the many genes involved in motility and chemotaxis is grouped into at least three transcriptional classes (early, middle, and late). All of the genes of a particular class must be expressed before expression of the next, or later, class. This leads to sequential assembly of the basal body structure, addition of the hook via the HAPs, which are important for filament addition, and finally polymerization of flagellin monomers at the growing tip. Failure at any point of this assembly completely shuts down later gene expression. As depicted in Fig. 8, we have begun to identify genes in the motility system of *V. parahaemolyticus* that is used for surface translocation and to place these genes into a scheme of control. Three levels have been identified in the transcriptional hierarchy.

A key factor in flagellar regulatory cascades is an alternative sigma subunit for RNA polymerase (18). We have functionally cloned a lateral flagellar gene that appears to encode such a sigma factor while cloning the single structural gene encoding lateral flagellin. LafS, the lateral flagellar sigma factor, is related to the alternative sigma factors (σ28) of the flagellar cluster, which includes not only the sigma subunits that direct synthesis of flagellar genes from a variety of organisms but also a sigma factor from *Streptomyces coelicolor* that controls morphological differentiation, i.e., the developmental fate of hyphae (10). However, the *V. parahaemolyticus* sigma factor appears to be different; although it is highly related to these sigma factors, it is more distantly related than any of the other members of the cluster are to each other. In particular, this difference is apparent in the amino acid sequence thought to be important in conferring promoter specificity.

This observation is consistent with other facts. Some of these sigma factors have been shown to recognize similar and functionally interchangeable promoters (i.e., those of *B. subtilis*, *E. coli*, and *Streptomyces coelicolor* [10, 18]), and σ28 consensus promoter sequences have been identified preceding flagellar genes from a variety of organisms. We have determined the nucleotide sequences of two genes encoding components important for formation of the lateral filament, the flagellin monomer (*lafA*) and a HAP-like protein probably important for polymerization of the filament (*lafB*). Sequences preceding both the *lafA* and *lafB* genes show no similarity to the σ28 promoter consensus sequence. In contrast, the genes encoding the genes encoding polar flagellins that resemble the flagellar consensus sequences (37a), and expression of these genes is dependent on *E. coli* σ28 (FliA). Genetic analysis has shown that expression of *lafA* requires the product of the sigma-like gene *lafS*. Furthermore, this requirement is manifested in *E. coli*. The *E. coli* flagellar sigma factor (FliA) does not recognize the promoter for the lateral flagellin gene, and expression of *lafA* in *E. coli* is dependent on the *lafS* gene. In fact, because of this requirement, the original clone expressing *lafA* in *E. coli* is a chimera containing two distinct regions of the chromosome, the *laF* and sigma loci. On the basis of previous work that shows the lateral sigma factor is a unique member of the flagellar cluster, and analysis of the differences in protein sequence and promoter specificity may prove insightful for determining what is important for conferring specificity of promoter selection.

In *E. coli*, the *motA* and *motB* gene products are required for torque generation. Both of these proteins are found in the cytoplasmic membrane. It is thought that MotA functions as a proton channel (7) and that MotB serves as a stator connecting MotA to a stationary component of the cell wall (11). The amino terminus of MotB is thought to span the cytoplasmic membrane, with the remainder of the molecule localized to the periplasm. How these proteins function to generate torque is not clear; however, the conservation of deduced protein sequence between the proton-driven lateral flagellar motor components (LaTt and LaFu) and MotA and MotB of *E. coli* supports the present topological picture of the torque generators. There is only one region of significant difference. LaTt has four membrane-spanning domains similar to those detected for MotA (12). It also has two regions of high charge which may be important for the handling of protons. The difference in these charged regions of *V. parahaemolyticus* is that they are not separated into an acidic and a basic domain as was observed for *E. coli*; both regions contain many acidic, basic, and polar amino acids.

The arrangement of genes around *lafS* suggests a transcriptional unit, and it will be interesting to determine whether large operons or gene clusters are involved in these lateral flagellar motor genes in the sigma locus is curious. In *E. coli*, the *mot* genes are in an operon with two chemotaxis genes and require the flagellar sigma factor for expression. Also, the appearance of a gene resembling *fliL* in the sigma operon is intriguing. Although function of *FliL* is not known, it is usually found preceding a second gene, *fliM*, which is believed to be part of the motor's switch complex involved in modulating direction of flagellar rotation. Again, the gene organization appears unusual for *V. parahaemolyticus*. Mutants in the *fliM* locus of *C. crescentus* show abnormal cell division, and it has been postulated that *fliM* or *fliL* may help couple flagellar regulation to the cell cycle (53). Flagellar morphogenesis for *E. coli* is thought to be coupled to the cell cycle (43, 44). In contrast to these bacteria, synthesis of lateral flagella seems to be released from cell division events. When *V. parahaemolyticus* differentiates to a swarmer cell type, cell division ceases, the cell elongates, and hundreds of lateral flagella are synthesized. Perhaps the unique transcriptional organization of lateral flagellar genes is significant in terms of the particular timing and coordination of gene expression required for swarmer cell differentiation.

ACKNOWLEDGMENTS

We thank M. Silverman, B. Bassler, R. Showalter, and C. Turnbough for much productive interaction. We also thank R. Cox, J. Eisen, A. Kwon, A. Leslie, L. Rovin, and J. Trzeciak for helping with the transposon mutagenesis and gene replacements.

This work was supported by Public Health Service grant GM43196 to L.L.M. from the National Institutes of Health.

REFERENCES

1. Albertini, A. M., T. Caramori, W. D. Crabb, F. Scofione, and A. Galizzi. 1991. The *fadA* locus of *Bacillus subtilis* is part of a large operon coding for flagellar structures, motility functions, and an ATPase-like polypeptide. J. Bacteriol. 173:3573–3579.
25. Keen, N. T., S. Tamaki, D. Kobayashi, and D. Trollering. 1988. Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. Gene 70:191–197.

26. Kihara, M., M. Homma, K. Kutsukake, and R. M. Macnab. 1989. Flagellar switch of Salmonella typhimurium: gene sequences and deduced protein sequences. J. Bacteriol. 171:3247–3257.

27. Komeda, Y. 1986. Transcriptional control of flagellar genes in Escherichia coli K-12. J. Bacteriol. 168:1315–1318.

28. Komeda, Y., and T. Iino. 1979. Regulation of expression of the flagellin gene (hag) in Escherichia coli K-12: analysis of hag-lac gene fusions. J. Bacteriol. 139:721–729.

29. Kono, S. C., and D. E. Koshland. 1986. Sequence of the flaA (cheC) locus of Escherichia coli and discovery of a new gene. J. Bacteriol. 166:1007–1012.

30. Kutsukake, K., Y. Ohya, and T. Iino. 1990. Transcriptional analysis of the flagellar regulon of Salmonella typhimurium. J. Bacteriol. 172:741–747.

31. Kuwajima, G., I. Kawagishi, M. Homma, J. Asaka, E. Kondo, and R. Macnab. 1989. Export of an N-terminal fragment of Escherichia coli flagella by a flagellum-specific pathway. Proc. Natl. Acad. Sci. USA 86:4953–4957.

32. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. J. Mol. Biol. 157:105–132.

33. Lonetto, M., M. Grisbok, and C. A. Gross. 1992. The σ70 family: sequence conservation and evolutionary relationships. J. Bacteriol. 174:3843–3849.

34. Macnab, R. M., and J. S. Parkinson. 1991. Genetic analysis of the bacterial flagellum. Trends Genet. 7:196–200.

35. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

36. Manoil, C., and J. Beckwith. 1985. TphoA: a transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA 82:8129–8133.

37. Martin, M., R. Showalter, and M. Silverman. 1989. Identification of a locus controlling expression of luminescence genes in Vibrio harveyi. J. Bacteriol. 171:2406–2414.

38. McCarther, L. Unpublished data.

39. McCarther, L., M. Hilmen, and M. Silverman. 1988. Flagellar dynamometer controls swarmer cell differentiation of V. parahaemolyticus. Cell 54:345–351.

40. McCarther, L., and M. Silverman. 1989. Iron regulation of swarmer cell differentiation of V. parahaemolyticus. J. Bacteriol. 171:731–736.

41. McCarther, L. L., and M. Silverman. 1987. Phosphate regulation of gene expression in Vibrio parahaemolyticus. J. Bacteriol. 165:3441–3449.

42. McCarther, L., and M. Silverman. 1990. Surface-induced swarmer cell differentiation of Vibrio parahaemolyticus. Mol. Microbiol. 4:1057–1062.

43. Messing, J. 1983. New M13 vectors for cloning. Methods Enzymol. 101:20–78.

44. Newton, A., and N. Ohta. 1990. Regulation of the cell division cycle and differentiation in bacteria. Annu. Rev. Microbiol. 44:689–719.

45. Nishimura, A., and Y. Hirota. 1989. A cell division regulatory mechanism controls the flagellar region in Escherichia coli. Mol. Gen. Genet. 216:340–3465.

46. Ohnishi, K., K. Kutsukake, H. Suzuki, and T. Iino. 1990. Gene flaA encodes an alternative sigma factor specific for flagellar operons in Salmonella typhimurium. Mol. Gen. Genet. 221:139–147.

47. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.

48. Sar, N., L. McCarther, M. Simon, and M. Silverman. 1989. Chemotactic control of the two flagellar systems of Vibrio parahaemolyticus. J. Bacteriol. 172:334–241.

49. Shinoda, S., and K. Okamoto. 1977. Formation and function of Vibrio parahaemolyticus lateral flagella. J. Bacteriol. 129:1266–1271.
49. Silverman, M., R. Showalter, and L. McCarter. 1991. Genetic analysis in Vibrio. Methods Enzymol. 204:515-536.
50. Stader, J., P. Matsumura, D. Vacante, G. E. Dean, and R. M. Macnab. 1986. Nucleotide sequence of the Escherichia coli motB gene and site-limited incorporation of its product into the cytoplasmic membrane. J. Bacteriol. 166:244–252.
51. Starnbach, M. N., and S. Lory. 1992. The $\text{fliA (rpoF)}$ gene of Pseudomonas aeruginosa encodes an alternative sigma factor required for flagellin synthesis. Mol. Microbiol. 6:459–469.
52. Uliätzur, S. 1974. Induction of swarming in Vibrio paraheomolyticus. Arch. Microbiol. 101:357–363.
53. Yu, J., and L. Shapiro. 1992. Early Caulobacter crescentus genes $\text{fliL}$ and $\text{fliM}$ are required for flagellar gene expression and normal cell division. J. Bacteriol. 174:3327–3338.