REPORT DOCUMENTATION PAGE

1a.

1b. RESTRICTIVE MARKINGS

2a. **AD-A222 968**

2b.

3. DISTRIBUTION/AVAILABILITY OF REPORT
Approved for public release; distribution is unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
NMRI 90-28

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
Naval Medical Research

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
Naval Medical Command

6c. ADDRESS (City, State, and ZIP Code)
Bethesda, Maryland 20814-5055

7b. ADDRESS (City, State, and ZIP Code)
Department of the Navy
Washington, D.C. 20372-5120

8a. NAME OF FUNDING/SPONSORING ORGANIZATION
Naval Medical Research and Development Command

8b. OFFICE SYMBOL (If applicable)

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State, and ZIP Code)
Bethesda, Maryland 20814-5055

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO
61102A	3M161102BS13	AK-111	DA313955

11. TITLE (Include Security Classification)
Genomic organization and expression of Campylobacter flagellin genes

12. PERSONAL AUTHOR(S)
Guerry P, Logan SM, Thornton S, Trust TJ

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
journal article	FROM ____ TO ____	1990	8

16. SUPPLEMENTARY NOTATION
Reprinted From: Journal of Bacteriology, Apr. 1990, p. 1853-1860 Vol. 172, No. 4

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Campylobacter jejuni ; flaA gene
Campylobacter coli VC167 ; flaB gene
Campylobacter flagellin genes. antigenic variation (Over)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
Phyllis Blum, Information Services Division

22b. TELEPHONE (Include Area Code)
202-295-2188

22c. OFFICE SYMBOL
ISD/ADMIN/NMRI

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted.
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

18. Subject Terms continued:
phase variation
DNA rearrangement
flagellin genes
flagellin mutants

JOURNAL OF BACTERIOLOGY, Apr. 1990, p. 1853–1860
0021-9193/90/041853-08$02.00/0

Vol. 172, No. 4

Genomic Organization and Expression of *Campylobacter* Flagellin Genes

PATRICIA GUERRY,[1]* SUSAN M. LOGAN,[2] SCOTT THORNTON,[1] AND TREVOR J. TRUST[2]

Infectious Diseases Department, Naval Medical Research Institute, Bethesda, Maryland 20814,[1] and Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia, V8W 2Y2, Canada[2]

Received 14 August 1989/Accepted 27 December 1989

***Campylobacter coli* VC167, which undergoes an antigenic flagellar variation, contains two full-length flagellin genes, *flaA* and *flaB*, that are located adjacent to one another in a tandem orientation and are 91.5% homologous. The gene product of *flaB*, which has an M_r of 58,946, has 93% sequence homology to the gene product of *flaA*, which has an M_r of 58,916 (S. M. Logan, T. J. Trust, and P. Guerry, J. Bacteriol. 171:3031–3038, 1989). Mutational analyses and primer extension experiments indicated that the two genes are transcribed under the control of distinct promoters but that they are expressed concomitantly in the same cell, regardless of the antigenic phase of flagella being produced. The *flaA* gene, which was expressed at higher levels than the *flaB* gene in both phases, was transcribed from a typical σ^{28}-type promoter, whereas the *flaB* promoter was unusual. A mutant producing only the *flaB* gene product did not synthesize a flagellar filament and was nonmotile. Southern blot analysis indicated that flagellar antigenic variation involves a rearrangement of flagellin sequence information rather than the alternate expression of the two distinct genes.**

Campylobacter jejuni and *Campylobacter coli* are important gastrointestinal pathogens of humans, most commonly producing an acute enteritis (6, 8, 39). An essential step in the establishment of *Campylobacter* enteritis is the colonization of the viscous mucous blanket covering the epithelium of the small intestine. The motility imparted by the polar flagellum of the *Campylobacter* cell appears to play a significant role in the colonization of this mucous niche (23, 30, 31), since nonmotile variants appear unable to colonize the gastrointestinal tract of experimental animals or human volunteers (4, 7). Flagellin is also the immunodominant protein antigen recognized during infection, so flagella are clearly important virulence determinants of the thermophilic campylobacters.

Flagellar expression in *Campylobacter* spp. is subject to both phase and antigenic variation (7, 12). Phase variation refers to ability of some *Campylobacter* strains to exhibit a bidirectional transition between flagellated and nonflagellated phenotypes (7). Antigenic variation refers to the ability of other strains to reversibly express flagella of different antigenic specificities (12). In the case of *C. coli* VC167, the reversible variation corresponds to the production of flagellar filaments containing antigenically distinguishable flagellin monomers of apparent subunit molecular weight 61,500 in antigenic phase 1 (P1) cells and 59,500 in antigenic phase 2 (P2) cells. Importantly, recent studies with *C. coli* VC167 have demonstrated an in vivo preference for one flagellar phase, P2, during colonization in an animal model (19), further emphasizing the functional importance of this locomotory organelle in gastrointestinal colonization by *Campylobacter* spp.

Phase and antigenic variation of surface antigens has been described for other pathogenic bacteria, including fimbrial-phase variation in *Escherichia coli* (1), *Moraxella bovis* (25), and *Neisseria gonorrhoeae* (27, 35, 40), antigenic variation of outer membrane proteins of *Neisseria* (40, 41) and *Borrelia* (26, 33) spp., and antigenic variation of flagella in *Salmonella typhimurium* (37, 38, 42). Presumably, the ability to undergo such surface changes confers advantages to the pathogen as it encounters multifarious environments. The molecular mechanisms by which pathogens alter their surface antigens are varied but often involve programmed DNA rearrangements (5). Antigenic variation of VC167 flagella is associated with an as yet uncharacterized DNA rearrangement at a site that is not adjacent to the structural genes (11, 19). As a first step in elucidating the molecular mechanisms involved in the antigenic variation of *C. coli* flagella, a flagellin gene from VC167 was recently cloned and the nucleotide sequence was determined (22). The open reading frame encoded by the gene was found to share strong homology with the amino and carboxy termini of flagellins from other bacteria and was found to be posttranslationally modified at serine residues (22). Hybridization data indicated the presence of flagellin-related information immediately downstream of the sequenced gene (22), suggesting the presence of a second copy of flagellin. In this study, we show that there are, in fact, two full-length flagellin genes that are adjacent to one another on the VC167 chromosome in a tandem orientation. The two genes are highly related to one another but are transcribed from separate promoters concomitantly in both P1 and P2 cells. Moreover, we provide evidence that antigenic variation of flagellin in *C. coli* results from a rearrangement internal to the flagellin structural genes.

MATERIALS AND METHODS

Bacterial strains and growth conditions. *C. coli* VC167 serogroup LIO 8 was originally isolated from human feces and was obtained from H. Lior, National Enteric Reference Centre, Ottawa, Ontario, Canada. Derivatives of VC167 producing either P1 or P2 cells were selected as described by Harris et al. (12). *E. coli* DH5 (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) was used as the host for cloning experiments. Growth conditions were as previously described (22).

DNA sequencing. Double-stranded dideoxy sequencing was performed after alkaline denaturing by using [^{35}S]dATP (Dupont NEN Research Products, Boston, Mass.) and Se-

* Corresponding author.

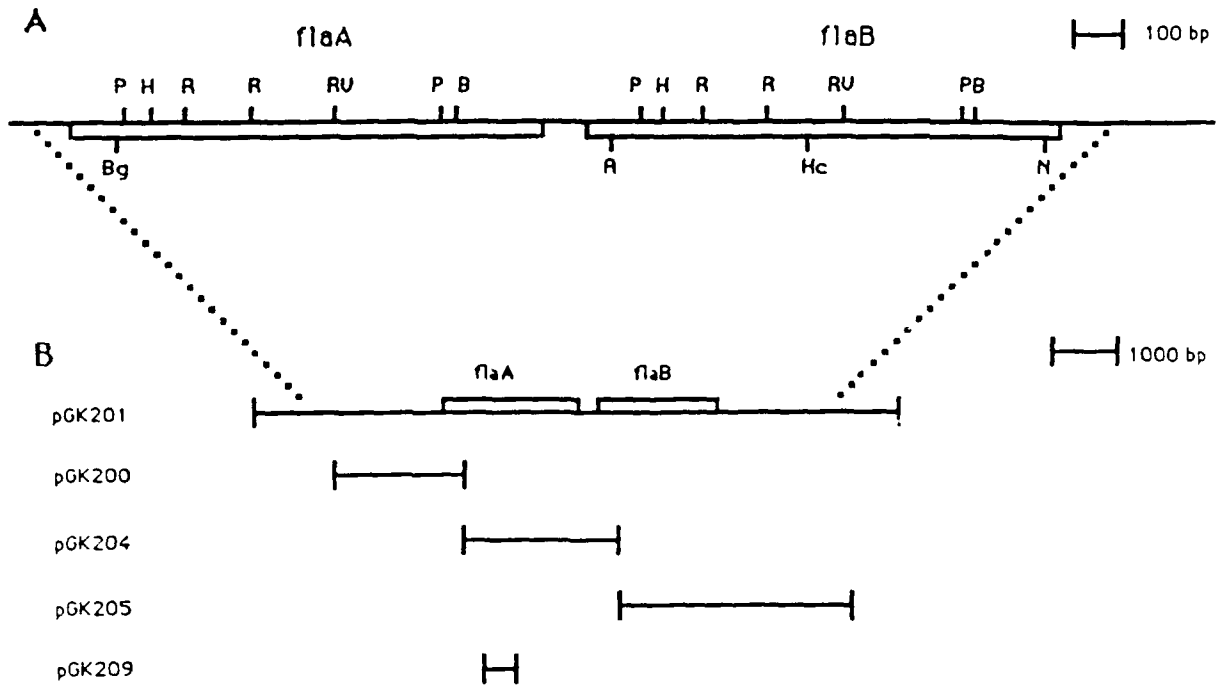

FIG. 1. Maps of VC167 flagellin gene clusters. (A) The portion of the DNA cloned on pGK201 (23) encoding two tandemly oriented flagellin genes. Transcription of the genes is from left to right, as drawn. Restriction enzymes sites marked above the line are those shared by the two genes; sites unique to one copy are indicated below the line. (B) The original clone, pGK201, containing *flaA* and *flaB*. Plasmids pGK200, pGK204, and pGK205 were used for DNA sequencing. Plasmid pGK209 contains an *Eco*RI-*Eco*RV fragment subcloned from the *flaA* gene as indicated. Restriction sites: P, *Pst*I; H, *Hin*dIII; R, *Eco*RI; RV, *Eco*RV; B, *Bcl*I; Bg, *Bgl*I; Hc, *Hin*cII; N, *Nde*I.

quenase (United States Biochemical Corp., Cleveland, Ohio) as specified by the manufacturers. The templates used were three previously described plasmids, pGK200, pGK204, and pGK205 (22; Fig. 1). Primers were commercially available pBR322 primers (New England Bio-Labs, Beverly, Mass.) and custom primers synthesized on a Biosearch 8700 synthesizer (Milligen Biosearch, Burlington, Mass.). Custom primers were synthesized at approximately 250- to 300-base-pair (bp) intervals on both strands as flagellin gene sequence became available. Codon utilization indices were calculated by the method of Sharp and Li (36).

Primer extension experiments. RNA was extracted from *Campylobacter* cells grown for 12 h at 37°C on Mueller-Hinton agar (Difco Laboratories, Detroit, Mich.) by the guanidinium-cesium chloride method described by Maniatis et al. (24). The RNA pellet was treated with RNase-free DNase (Promega Biotec, Madison, Wis.) for 10 min at 37°C, phenol extracted, and ethanol precipitated. Oligonucleotide primers were labeled with [γ-^{32}P]ATP (Dupont, NEN) as described by Maniatis et al. (24). Primer extension analysis of transcripts was performed basically as described by Fouser and Friesen (10). Briefly, 2 pmol of ^{32}P-labeled primer was mixed with 60 μg of RNA, coprecipitated, and suspended in 7.5 μl of 50 mM Tris (pH 8.3)–40 mM NaCl–0.5 mM EDTA. After heating to 65°C, the primer was annealed at 37°C for 1 h. A total of 3.5 μl of mix R (70 μM dATP, 70 μM dCTP, 70 μM dGTP, 70 μM dTTP, 30 mM $MgCl_2$, 3 mM dithiothreitol, 0.6 μg of actinomycin D per ml) and 20 U of avian myeloblastosis virus reverse transcriptase (Pharmacia, Uppsala, Sweden) was added. After incubation at 37°C for 1 h, 5 μl of sequencing dye was added, and the sample was heated for 3 min in a boiling water bath. A 5-μl amount of each extension reaction was run on an acrylamide sequencing gel in parallel with a DNA sequence ladder primed with the same oligonucleotide used in the primer extension reaction.

Hybridizations. DNAs were nick translated with [α-^{32}P]dCTP by using a commercial kit (Dupont, NEN). Conditions of hybridization were as described by Guerry et al. (11).

DNA manipulations. Plasmid DNA extraction procedures were as previously described (22). Total DNA extractions from *Campylobacter* cells were done by the method of Hull et al. (15). Restriction enzymes and T4 polynucleotide ligase were purchased from Boehringer Mannheim Biochemicals (Indianapolis, Ind.) or New England Bio-Labs and were used under the conditions recommended by the supplier.

Electrophoresis and Western blotting (immunoblotting). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis was performed with a minislab gel apparatus (Hoeffer Scientific Instruments, San Francisco, Calif.) by the method of Laemmli (18). Protein samples solubilized in sample buffer were stacked in 4.5% acrylamide (100 V, constant voltage), separated in either 7.5 or 12.5% acrylamide (200 V, constant voltage), and stained with Coomassie blue or transferred to nitrocellulose for immunological detection as previously described (22).

Mutational analysis. Gene disruption and replacement methods were basically those described by Labigne-Roussel et al. (17). Conjugal transfer of the disrupted flagellin gene from *E. coli* to *C. coli* VC167 was done by using as a donor a strain of DH5 harboring plasmid RK212.1 (9) under the mating conditions described by Labigne-Roussel et al. (17). Selection of *Campylobacter* transconjugants was on Muel-

ler-Hinton medium supplemented with trimethoprim (10 μg/ml) and kanamycin (100 μg/ml). Motility was determined by using the pour plate method described by Caldwell et al. (7).

RESULTS

DNA and protein sequences of the second VC167 flagellin. Plasmid pGK201, which contains a 10-kilobase-pair (kb) *Cla*I fragment from the VC167 chromosome, has been shown to contain a full-length flagellin gene (22), now termed *flaA*. Hybridization studies had suggested that a second copy of flagellin resides adjacent to the first on this plasmid (22). DNA sequence analysis of this region confirmed the presence of a second full-length flagellin gene, *flaB* (Fig. 1). The two genes are tandemly oriented and separated by 216 bp. The DNA sequence of *flaB* is compared with that of *flaA* in Fig. 2, and the deduced protein sequences of the two flagellins are compared in Fig. 3. The two DNA sequences are remarkably conserved, with an overall sequence similarity of 91.5%, suggesting that they arose as a result of gene duplication. There are regions of the two genes that are totally identical, lacking even third-base-pair changes. There is also a strong bias for codons with A or T in the third position in both genes (data not shown), as would be expected for an organism with a G+C content of 30.1 to 33% (3, 32). The *flaA* gene encodes a polypeptide of 571 amino acids (after removal of the first methionine) with a predicted posttranslationally unmodified M_r of 58,916 (22); the *flaB* gene encodes a polypeptide of 572 amino acids with a predicted posttranslationally unmodified M_r of 58,946. The two proteins share 93% homology at the amino acid level. All regions of the flagellin protein known to be posttranslationally modified (22) are identical in both molecules. Although most of the changes occur in the amino and carboxy ends of the molecules, there are two areas of major changes in the middle regions of the two proteins (Fig. 3). The first is caused by a frameshift mutation (Fig. 2) and accounts for the fact that *flaB* encodes one additional amino acid.

Analysis of the promoter regions. The sequences of the 5′ noncoding regions of the two genes are compared in Fig. 4. In the case of *flaB*, this region represents the intergenic region, which contains two nearly perfect 49-bp repeats separated by 35 nucleotides. Within these repeats are two identical stem-loop structures. These potential terminator sequences for the *flaA* gene have free-energy values of −14.8 kcal (−61.2 kJ)/mol and are separated by 60 nucleotides. The region immediately upstream of the translational start of both genes, including classical AGGA putative ribosome-binding sites, is the most highly conserved region. Both AGGA sites are 6 nucleotides from the translational initiation codon. An area upstream of *flaA* is indicated which shares strong homology to the consensus sequence for a σ^{28}-type promoter (13). The corresponding areas of *flaB* share some homology at the −35 region but little homology in the −10 region. Primer extension experiments using mRNA purified from P1 and P2 cells were performed to determine the transcriptional initiation sites and to determine which gene(s) is expressed in each flagellar phase. The mRNA from each phase was primed with oligonucleotides specific for the *flaA* (5′-TAATGCTCTCCTGTTTA-3′) and *flaB* (5′-AACGTCCCTAGCATTAA-3′) genes. These oligonucleotides bind to the coding strand of each gene at a position starting 52 bp downstream of the ATG start. Both genes appeared to be transcribed independently in both flagellar phases (Fig. 5). In addition, the *flaA*-specific oligonucleotide gave a higher relative signal in primer extension experiments than the *flaB*-specific oligonucleotide, and this relative signal strength was the same using mRNA templates from P1 and P2 cells (Fig. 5). The start point for *flaA* (Fig. 4) is that predicted from the position of the σ^{28} promoter. The start point for *flaB*, which is closer to the structural gene by 17 bp, and the putative −35 and −10 promoter regions of *flaB*, determined on the basis of this start point, are also indicated in Fig. 4.

Mutational analysis of flagellin genes. Gene replacement techniques (17) were used to generate flagellin mutants. Plasmid pGK200 (22; Fig. 1) contains a 2.3-kb *Hin*dIII fragment containing the first 311 bp of the *flaA* gene and approximately 2 kb of nonflagellin DNA sequence 5′ to the start of *flaA*. There is a unique *Bgl*II site located 162 bp downstream of the translational start of the *flaA* gene. A *Campylobacter* gene for kanamycin resistance was cloned as a *Bam*HI-ended cassette from plasmid pILL600 (17) into this *Bgl*II site. The resulting *Hin*dIII fragment, in which the *flaA* flagellin sequence had been disrupted by the kanamycin resistance gene, was transferred into the *Campylobacter* suicide vector pILL560, which is capable of being conjugally mobilized by IncP plasmids into, but is not capable of autonomous replication in, *Campylobacter* cells (17). The resulting plasmid, pGK200, was transformed into a strain of DH5 harboring the conjugative IncP plasmid RK212.2 (9). Conjugative transfer from *E. coli* to *C. coli* VC167 P1 was performed as described by Labigne-Roussel et al. (17). All kanamycin-resistant *Campylobacter* transconjugants obtained should represent events in which the kanamycin resistance gene was rescued by a crossover between the *Campylobacter* sequences flanking the kanamycin resistance gene and the corresponding regions of homology in the chromosome. Such a single crossover resulted in a duplication of sequences on the chromosome (Fig. 6A); excision of the vector DNA and one copy of the flagellin information should occur by the analogous single reciprocal crossover. In this case, in which there are two repeating regions of homology, represented by the two tandem flagellin genes, two distinct second recombinational events are possible, and both were isolated (Fig. 6A). The first class, exemplified by mutant VC167-B2, occurs when the second recombinational event occurs within the *flaA* gene. This results in a simple replacement of the kanamycin-inactivated flagellin sequences for the corresponding *flaA* chromosomal sequences. If the recombinational event occurs within the *flaB* gene, all of the chromosomal copy of *flaA* is deleted. This second class, which was the predominant class isolated (three of five transconjugants), is represented by mutant VC167-B3. Figure 6B shows a Southern blot analysis of *Hin*dIII-digested DNA from wild-type VC167 (lane 1), VC167-B2 (lane 2), and VC167-B3 (lane 3) probed with pGK201. In VC167-B2, the *Hin*dIII fragment, H3, had increased in M_r by approximately 1.4 kb, the size of the inserted kanamycin resistance gene, as compared with wild-type VC167 DNA (i.e., from 2.3 kb to 3.7 kb). In VC167-B3, the H3 fragment had also increased in size; in addition, all of the H4 fragment, which encodes the 3′ end of *flaA* and the 5′ end of *flaB*, was deleted.

Both VC167-B3, which is genotypically *flaA flaB*⁻, and VC167-B2, which is *flaA flaB*⁺, are nonmotile. No flagellin can be detected on the surface of either VC167-B2 or VC167-B3 cells after glycine extraction (20; data not shown). However, Western blot analysis of whole cell proteins using a polyclonal antibody directed against campylobacter flagellin (21) demonstrated that the *flaB* gene product was being expressed in VC167-B2 (Fig. 6C, lane 3) at low levels. No

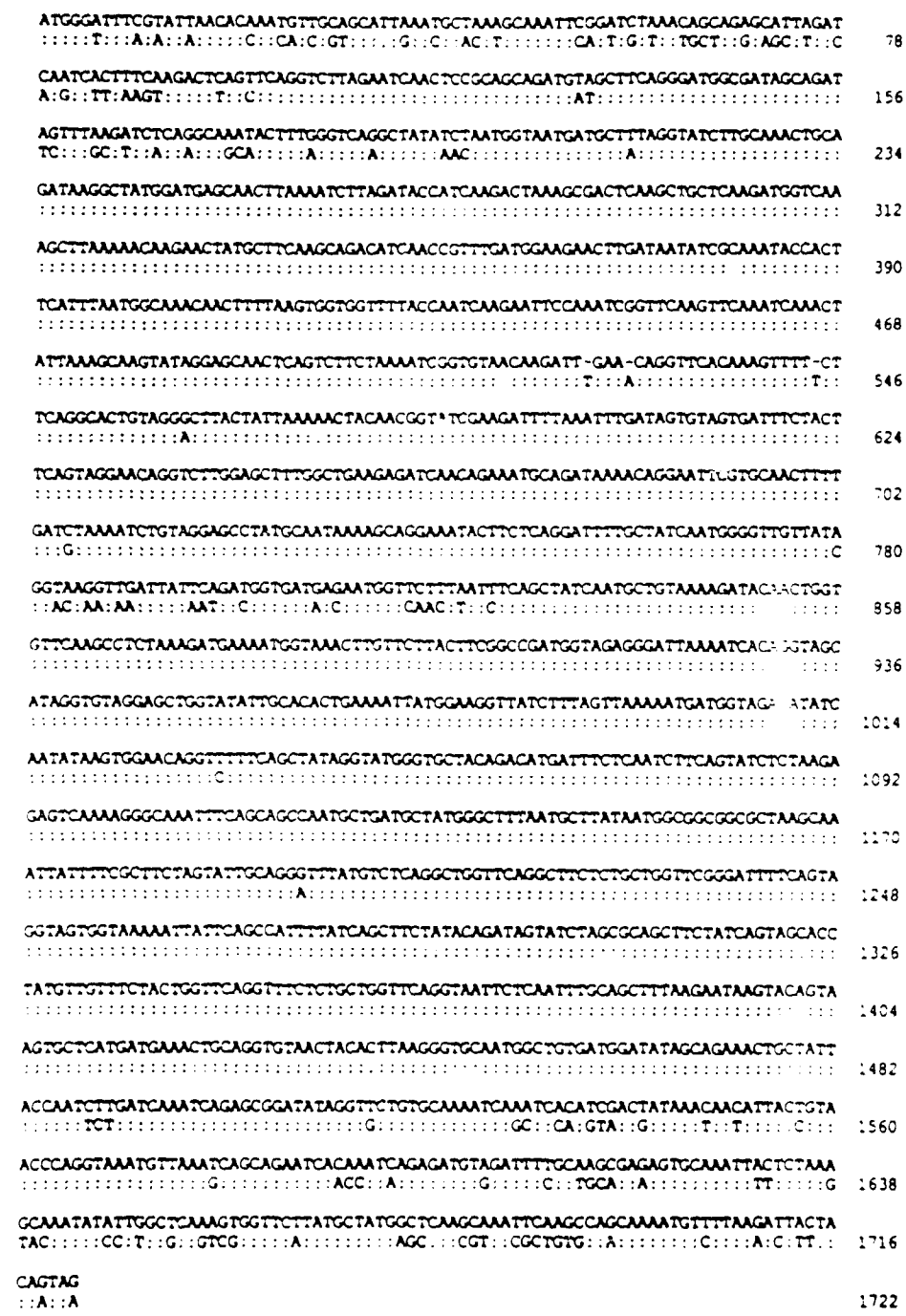

FIG. 2. Comparison of DNA sequences of *flaA* (23) (top line) and *flaB* (bottom line) genes of *C. coli* VC167 P1. Colons indicate that identical bases; changes are so indicated. Numbers at the right refer to base pairs of the *flaB* gene. The *flaB* gene contains three additional bases at positions 522, 526, and 544.

immunologically reactive material was produced in VC167-B3 (Fig. 6C, lane 4).

Mechanism of antigenic variation. The biochemical and genetic data presented above indicate that both *flaA* and *flaB* are expressed in VC167 P1 and VC167 P2 and that alternate expression of the two genes is not involved in flagellar antigenic variation. In other pathogens, antigenic variations often involve internal rearrangements of structural genes (4). To test this possibility, DNAs from VC167 P1 and VC167 P2 cells were digested with various restriction enzymes and analyzed by Southern blot hybridization, using ^{32}P-labeled pGK201 DNA as the probe. No rearrangements were observed when the DNAs were digested with either *Hin*dIII (Fig. 7, lanes 1 and 2) or *Eco*RV (lanes 5 and 6). However,

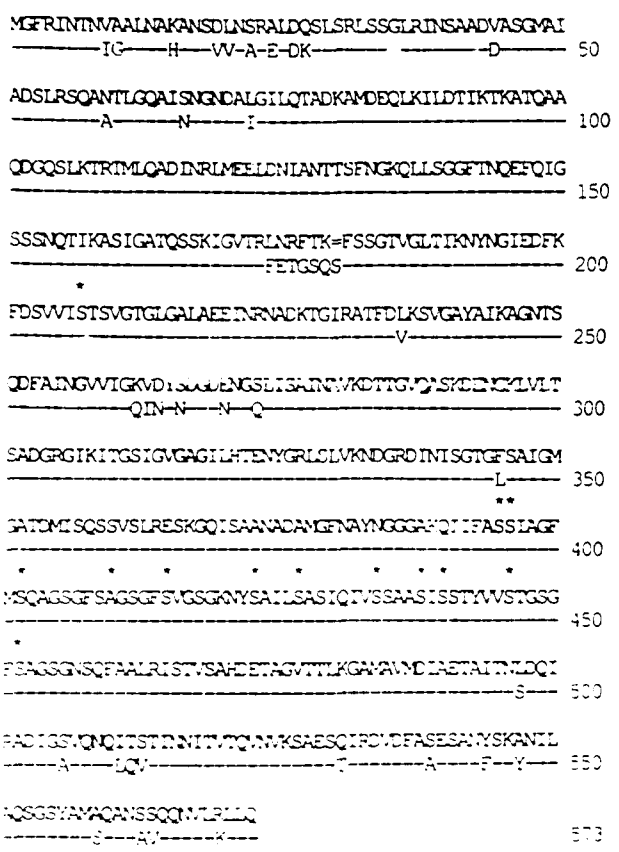

FIG. 3. Comparison of the gene products of the *flaA* and *flaB* genes of *C. coli* VC167 P1. The open reading frame encoded by *flaA* is shown on the top line (23), and that encoded by *flaB* is below it. Dashes indicate identical residues; changes are so indicated. Numbers at the right refer to the amino acid residues of the *flaA* gene. Asterisks indicate positions of serine residues known to modified (23) in both P1 and P2 flagellins. Amino acids are designated by the single-letter code.

When the DNAs were digested with *Hin*cII, a band of approximately 2 kb hybridized to the probe in VC167 P1 DNA (lane 3) but not in VC167 P2 (lane 4). To be certain that the rearrangement involved flagellin sequence information and not some of the non-flagellin-coding information present on the pGK201 probe, a similar Southern analysis was performed with plasmid pGK209 as the probe. This plasmid represents a subclone of a 323-bp *Eco*RI-*Eco*RV fragment from the middle of the *flaA* gene (Fig. 1) and corresponds to the coding region for the second of the two major internal changes between *flaA* and *flaB* (Fig. 3). When pGK209 was used to probe *Hin*cII-digested P1 and P2 DNAs, the same 2-kb fragment was observed in P1 but not P2 DNA (data not shown). This result indicates that the transition from P1 to P2 involves a rearrangement of the structural genes. The original clone of both genes, pGK201, does not contain a 2-kb *Hin*cII fragment, suggesting that partial flagellin-coding information, corresponding to an antigenically variable region, lies outside the cluster of genes present on pGK201 and that this partial copy if involved in antigenic variation.

DISCUSSION

This work confirms that *C. coli* VC167, which undergoes a flagellar antigenic variation, contains two full-length copies

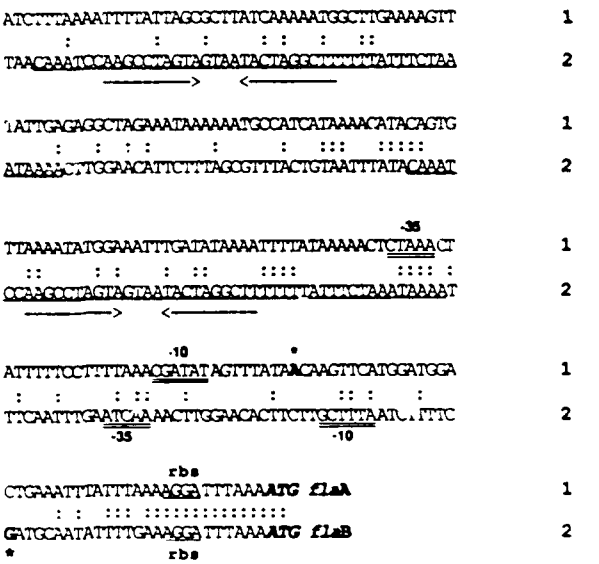

FIG. 4. Comparison of the 5′ noncoding regions of *flaA* and *flaB*. The 214-bp regions upstream of the two flagellin genes are compared. The sequence 5′ to *flaA* is on the line 1, and the sequence 5′ to *flaB* (and 3′ to *flaA*) is on line 2. Identical bases are indicated by colons. The final ATG in bold italic letters represents the translational start of each structural gene. The mRNA start points for each gene determined by primer extension experiments are indicated by asterisks. Putative ribosome-binding sites (rbs) for each gene are marked by dotted underlining. A σ^{28}-type promoter lies in the appropriate position relative to the mRNA start site on the *flaA* sequence (double underlined at the −35 and −10 positions). The consensus σ^{28} promoter contains the sequence CTAAA at the −35 position and CCGATAT at the −10 position (13). The position of a putative promoter sequence for *flaB*, determined from the position of the mRNA start point, is also indicated by double underlining. Arrows indicate two identical palindromic sequences that may function as transcriptional terminators of the *flaA* gene; these sequences are included within nearly perfect 49-bp repetitive sequences (underlined).

of flagellin. The two genes, which are tandemly oriented and 93% homologous, are separately transcribed in both antigenic phases, indicating that antigenic variation is not a result of the alternate expression of *flaA* and *flaB*. The genomic organization of VC167 flagellins resembles that described for *Rhizobium meliloti* (34), which also has two tandemly arranged flagellin genes that are 87% homologous. However, in this organism, the gene products of both genes are found in a complex flagellar filament in approximately stoichiometric amounts, as determined by amino acid sequencing data (34). The N-terminal amino acid sequence of flagellin purified from flagellar filaments of P1 and P2 VC167 cells is identical in both phases for the first 38 amino acids (20). This amino acid sequence is that encoded by the *flaA* gene; the changes in the amino terminal sequence encoded by *flaB* have not been observed in amino acid sequencing of flagellin from either VC167 P1 or P2 cells. This would suggest that the *flaB* gene product is not present as a major constituent in the flagellar filament of either phase. However, internal peptides encoded by *flaB* have been observed during protein sequencing of tryptic peptides from both P1 and P2 cells in minor amounts (data not shown), further suggesting that the *flaB* gene is expressed at low levels in both phases.

The relative signal strength in both primer extension and

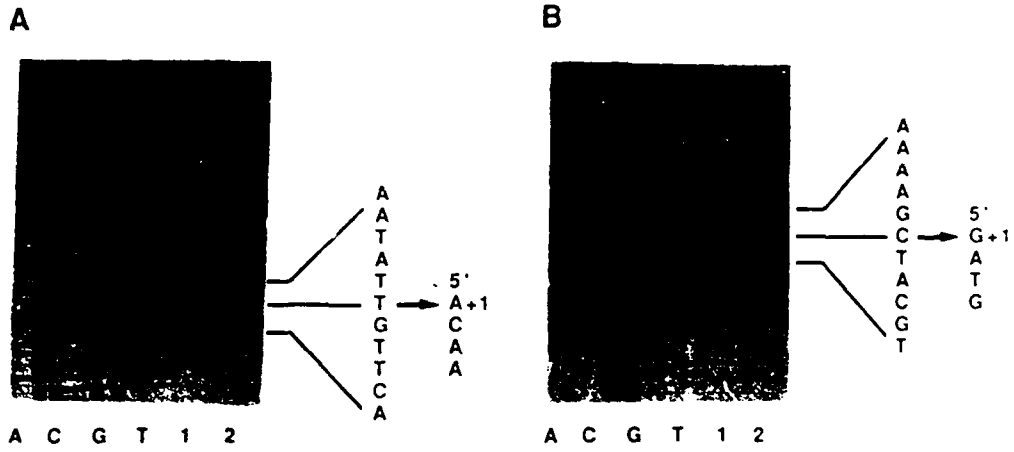

FIG. 5. Determination of *flaA* and *flaB* transcription start sites by primer extension. (A) *flaA*-specific primer extended on mRNA templates from VC167 P1 (lane 1) and VC167 P2 (lane 2) cells. (B) *flaB*-specific primer extended on mRNA templates from VC167 P1 (lane 1) and VC167 P2 (lane 2) cells. The corresponding sequencing ladders for each reaction are shown (ACGT), and the 5' end of the mRNA sequence is displayed at the right.

Western blot analyses suggests that the *flaA* gene is expressed at higher levels than *flaB* in both phases. In other systems, bacterial flagellins comprise up to 5% of the total cell protein and are transcribed under the control of strong promoters resembling that of the *B. subtilis* σ^{28} promoter (2, 29). A sequence bearing a strong homology to the σ^{28} promoter is present 5' to the *flaA* gene in the appropriate position for the promoter, as determined from the mRNA start point. The putative promoter of the *flaB* gene is unusual and presumably represents a weaker promoter. The presence of a strong σ^{28}-like promoter 5' to the *flaA* gene raises the question of why the *flaA* gene is not expressed at immunologically detectable levels in *E. coli* (22). Fusion of the *flaA* promoter to a promoterless chloramphenicol acetyltransferase gene indicates that the promoter can function in *E. coli* (data not shown). The lack of detectable *Campylo-*

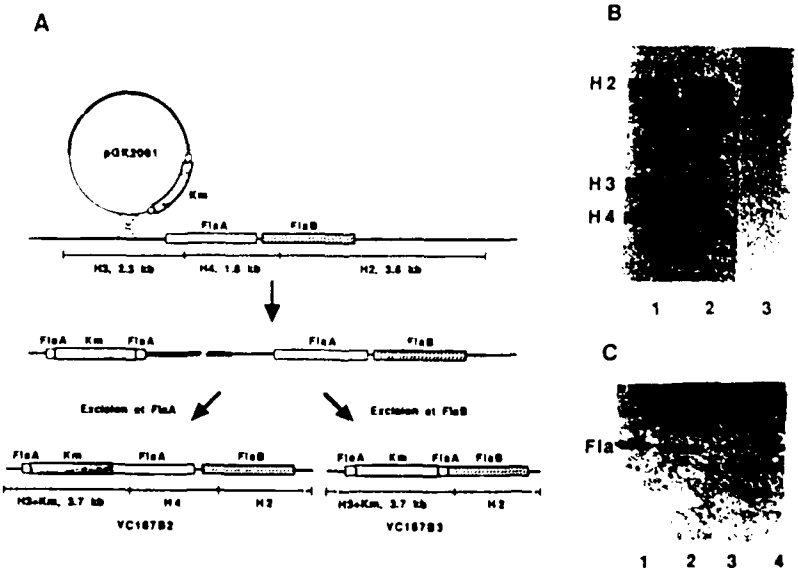

FIG. 6. Construction and characterization of flagellin mutants. (A) Construction of flagellin mutants by gene replacement techniques (not to scale). A *Campylobacter* kanamycin resistance gene was cloned into a unique *Bgl*II site on plasmid pGK200, and the resulting 3.7-kb *Hin*dIII fragment was subcloned in the suicide vector, pILL560 (17), to generate pGK2001. After conjugal transfer into VC167 P1 cells, a single crossover event allowed pGK2001 to integrate into the VC167 chromosome as indicated, resulting in a partial triplication of flagellin sequences. Excision of vector sequences could occur at regions of homology in *flaA* or *flaB*, as indicated, generating two classes of mutants, typified by VC167-B2 and VC167-B3. (B) Southern blot analysis of VC167-B2 and VC167-B3 mutants. Total DNA from each mutant and VC167 wild-type was digested with *Hin*dIII, electrophoresed, transferred to a nitrocellulose membrane, and probed with pGK201. Lanes: 1, VC167 P1; 2, VC167-B2; 3, VC167-B3. The bands in VC167 represent the following *Hin*dIII bands cloned in pGK201 (top to bottom): H2 (3.8 kb), H3 (2.3 kb), and H4 (1.8 kb). (C) Western blot analysis of whole-cell lysates, using rabbit polyclonal antiserum SML2 (22) at a 1:10,000 dilution. Lanes: 1, VC167 P1; 2, VC167 P2; 3, VC167-B2; 4, VC167-B3.

FIG. 7. Southern blot analysis of VC167 P1 and P2 DNAs. DNAs from P1 and P2 cells were digested with various restriction enzymes, electrophoresed, transferred to a nitrocellulose membrane, and probed with pGK201. Lanes: 1, VC167 P1 DNA, restricted with *Hin*dIII; 2, VC167 P2 DNA, restricted with *Hin*dIII; 3, VC167 P1 DNA, restricted with *Hin*cII; 4, VC167 P2 DNA, restricted with *Hin*cII; 5, VC167 P1 DNA, restricted with *Eco*RV; 6, VC167 P2 DNA, restricted with *Eco*RV. Fragments H2 (3.8 kb), H3 (2.3 kb), and H4 (1.8 kb) are visible in the *Hin*dIII digests; the *Hin*cII fragment visible in lane 3 is approximately 2 kb. Vector sequences alone do not hybridize to *Campylobacter* DNA under the conditions used (11).

bacter flagellin in *E. coli* may be due either to the fact that the gene product is not posttranslationally modified (22) or to codon utilization differences. There is a positive correlation between the degree of codon bias and the level of gene expression, with more highly expressed genes having the strongest constraints (36). The codon utilization index, which is a measure of codon usage bias (36), is 0.455 for the *hag* gene of *E. coli* (16) and 0.232 for the *flaA* of *C. coli*. This lower value for *flaA* reflects the presence of suboptimum codons in the *E. coli* host. Since both the *hag* and *flaA* genes are transcribed by strong σ^{28}-type promoters, this would suggest that translational barriers contribute to the lack of expression of *flaA* in the foreign host.

Most plain bacterial flagellins, including the VC167 flagellins, share considerable homology at the amino and carboxy regions of the molecule, which function in assembly and transport (14). The central nonfunctional regions of flagellins are antigenically diverse. It is surprising, therefore, that over half the amino acid changes occurring between the *flaA* and *flaB* gene products of VC167 occur in the amino and carboxy regions. The central regions of *flaA* and *flaB* do have two major areas of change, but throughout this putative variable region, there are only 15 total amino acids changes. The data indicate that the *flaA* gene represents the major component of the flagellar filament and that the gene product of *flaB* is not assembled into a functional flagellum in the absence of the *flaA* gene product. The simplest explanation for this finding would be that not enough gene product is synthesized under the *flaB* promoter to assemble a functional flagellum. The changes in the amino and carboxy regions of the *flaB* gene may also represent mutations affecting assembly or transport (14). Alternatively, it may be that the VC167 flagellum resembles that of *Caulobacter crescentus*, in which distinct flagellins assemble into different regions of the filament (28), and that both the *flaA* and *flaB* gene products are required for functional flagella in VC167. In an effort to address these questions, we have attempted to construct *flaA*⁻ *flaB* mutants. However, all insertions of the kanamycin resistance cassette into the *flaB* gene in *E. coli* to date have resulted in deletions or rearrangements of the *Campylobacter* DNA. Efforts to construct such mutants are continuing in an effort to elucidate the role of the *flaB* gene in the structural organization of the VC167 flagellum.

The mechanism of antigenic variation of VC167 flagella involves a rearrangement of some internal flagellin sequence information with undefined endpoints. The current data suggest that the *flaA* gene may represent a high-level expression locus and that the P1-to-P2 shift involves insertion of alternate genetic information in the form of a partial cassette into *flaA* at some point downstream of the amino-terminal sequences. Some of this information must reside outside of the flagellin gene cluster that has been cloned and sequenced, since the DNA cloned in pGK201 does not contain the 2-kb *Hin*cII fragment observed during the rearrangement. However, we cannot rule out the possibility that the *flaB* gene also contributes sequence information to *flaA*. To clarify the molecular mechanism of antigenic variation, we are raising antisera to synthetic peptides based on the observed changes between the *flaA* and *flaB* proteins of P1 cells, and we are currently determining the DNA sequences of the *flaA* and *flaB* genes cloned from P2 cells. The relationship of this rearrangement within the structural genes to the previously described rearrangement at a nonadjacent locus (11, 19) is also being investigated.

ACKNOWLEDGMENTS

This work was supported in part by U.S. Navy Research and Development Command Research Work Unit no. 61102A3M161 102BS13 AK.111 and by a grant from the Medical Research Council of Canada to T.J.T.

We thank Agnes Labigne-Roussel and Donald G. Guiney, Jr., for strains and Li-Ting Liang, Gary Majam, and Steve Martin for expert technical assistance.

LITERATURE CITED

1. **Abraham, J. M., C. S. Freitag, J. R. Clements, and B. I. Eisenstein.** 1985. An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **82**:5724–5727.
2. **Arnosti, D. N., and M. J. Chamberlin.** 1989. Secondary sigma factor controls transcription of flagellar and chemotaxis genes in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **86**:830–834.
3. **Belland, R. J., and T. J. Trust.** 1982. Deoxyribonucleic acid sequence relatedness between thermophilic members of the Genus *Campylobacter*. J. Gen. Microbiol. **128**:2515–2522.
4. **Black, R. E., M. M. Levine, M. L. Clements, T. P. Hughes, and M. J. Blaser.** 1988. Experimental *Campylobacter jejuni* infections in humans. J. Infect. Dis. **157**:472–479.
5. **Borst, P., and D. R. Greaves.** 1987. Programmed gene rearrangements altering gene expression. Science **235**:658–667.
6. **Butzler, J. P., and M. B. Skirrow.** 1979. *Campylobacter* enteritis. Clin. Gastroenterol. **8**:737–765.
7. **Caldwell, M. B., P. Guerry, E. C. Lee, J. P. Burans, and R. I. Walker.** 1985. Reversible expression of flagella in *Campylobacter jejuni*. Infect. Immun. **50**:941–943.
8. **Chowdhury, M. N. H.** 1984. *Campylobacter jejuni* enteritis: a review. Trop. Geogr. Med. **36**:215–222.
9. **Figurski, D. H., and D. R. Helinski.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc. Natl. Acad. Sci. USA **76**:1648–1652.
10. **Fouser, L., and J. Friesen.** 1986. Mutations in a yeast intron demonstrate the importance of specific conserved nucleotides for the two stages of nuclear mRNA splicing. Cell **45**:81–93.
11. **Guerry, P., S. M. Logan, and T. J. Trust.** 1988. Genomic rearrangements associated with antigenic variation in *Campylobacter coli*. J. Bacteriol. **170**:316–319.
12. **Harris, L. A., S. M. Logan, P. Guerry, and T. J. Trust.** 1987.

Antigenic variation of *Campylobacter* flagella. J. Bacteriol. **169**:5066–5071.
13. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. 1988. Annu. Rev. Biochem. **57**:839–872.
14. **Homma, M., H. Fujita, S. Yamaguchi, and T. Iino.** 1987. Regions of *Salmonella typhimurium* flagellin essential for its polymerization and excretion. J. Bacteriol. **169**:291–296.
15. **Hull, R. A., R. E. Gill, P. Hsu, B. H. Minshew, and S. Falkow.** 1981. Construction and expression of recombinant plasmids encoding type 1 or D-mannose resistant pili from a urinary tract infection *Escherichia coli* isolate. Infect. Immun. **33**:933–938.
16. **Kuwajima, J., I. Asaka, T. Fujiwara, T. Fujiwar, K. Node, and E. Kondo.** 1986. Nucleotide sequence of the *hag* gene encoding flagellin of *Escherichia coli*. J. Bacteriol. **168**:1479–1483.
17. **Labigne-Roussel, A., P. Couroux, and L. S. Tompkins.** 1988. Gene disruption and replacement as a feasible approach for mutagenesis of *Campylobacter jejuni*. J. Bacteriol. **170**:1704–1708.
18. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
19. **Logan, S. M., P. Guerry, D. M. Rollins, D. H. Burr, and T. J. Trust.** 1989. In vivo antigenic variation of *Campylobacter* flagellin. Infect. Immun. **57**:2583–2585.
20. **Logan, S. M., L. A. Harris, and T. J. Trust.** 1987. Isolation and characterization of *Campylobacter* flagellins. J. Bacteriol. **169**: 5072–5077.
21. **Logan, S. M., and T. J. Trust.** 1986. Location of epitopes on *Campylobacter jejuni* flagella. J. Bacteriol. **168**:739–745.
22. **Logan, S. M., T. J. Trust, and P. Guerry.** 1989. Evidence for posttranslational modification and gene duplication of *Campylobacter* flagellin. J. Bacteriol. **171**:3031–3038.
23. **Lee, A., J. L. O'Rourke, P. J. Barrington, and T. J. Trust.** 1986. Mucus colonization by *Campylobacter jejuni* a mouse cecal model. Infect. Immun. **51**:536–546.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Marrs, C. F., W. R. Ruehl, G. K. Schoolnik, and S. Falkow.** 1988. Pilin gene phase variation of *Moraxella bovis* is caused by an inversion of the pilin genes. J. Bacteriol. **170**:3032–3039.
26. **Meier, J. T., M. I. Simon, and A. G. Barbour.** 1985. Antigenic variation is associated with DNA rearrangements in relapsing fever *Borrelia*. Cell **41**:403–409.
27. **Meyer, T. F., N. Mlawer, and M. So.** 1982. Pilus expression in *Neisseria gonorrhoeae* involves chromosomal rearrangement. Cell **30**:45–52.
28. **Minnich, S. A., N. Ohta, N. Taylor, and A. Newton.** 1988. Role of the 25-, 27-, and 29-kilodalton flagellins in *Caulobacter crescentus* cell motility: method for construction of deletion and Tn*5* insertion mutants by gene replacement. J. Bacteriol. **170**: 3953–3960.
29. **Mirel, D. B., and M. J. Chamberlin.** 1989. The *Bacillus subtilis* flagellin gene (*hag*) is transcribed by the σ^{28} form of RNA polymerase. J. Bacteriol. **171**:3095–3101.
30. **Morooka, T., A. Umeda, and K. Amako.** 1985. Motility as an intestinal colonization factor for *Campylobacter jejuni*. J. Gen. Microbiol. **131**:1973–1980.
31. **Newell, D. G.** 1986. Monoclonal antibodies directed against the flagella of *Campylobacter jejuni*: production, characterization, and lack of effect on the colonization of infant mice. J. Hyg. **96**:131–141.
32. **Owen, R. J., and S. Leaper.** 1981. Base composition, size and nucleotide sequence similarities of genome deoxyribonucleic acids from species of the genus *Campylobacter*. FEMS Microbiol. Lett. **12**:395–400.
33. **Plasterk, R. H. A., M. I. Simon, and A. G. Barbour.** 1985. Transposition of structural genes to an expression sequence on a linear plasmid causes antigenic variation in the bacterium *Borrelia hermsii*. Nature (London) **318**:257–263.
34. **Pleier, E., and R. Schmitt.** 1989. Identification and sequence analysis of two related flagellin genes in *Rhizobium meliloti*. J. Bacteriol. **171**:1467–1475.
35. **Segal, E., P. Hagblom, H. S. Seifert, and M. So.** 1986. Antigenic variation of gonococcal pilus involves assembly of separated silent gene segments. Proc. Natl. Acad. Sci. USA **83**:2177–2181.
36. **Sharp, P. M., and W.-H. Li.** 1987. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. **15**:1281–1295.
37. **Silverman, M. J., and M. Simon.** 1980. Phase variation: genetic analysis of switching mutants. Cell **19**:845–854.
38. **Silverman, M. J., J. Zeig, M. Hilman, and M. Simon.** 1979. Phase variation in *Salmonella*: genetic analysis of a recombinational switch. Proc. Natl. Acad. Sci. USA **76**:391–395.
39. **Skirrow, M. B.** 1977. *Campylobacter* enteritis: a "new" disease. Br. Med. J. **2**:9–11.
40. **Sparling, P. F., J. G. Cannon, and M. So.** 1986. Phase and antigenic variation of pili and outer membrane protein II of *Neisseria gonorrhoeae*. J. Infect. Dis. **153**:196–201.
41. **Stern, A., M. Brown, P. Nickel, and T. F. Meyer.** 1986. Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. Cell **47**:61–71.
42. **Zieg, J., M. Silverman, M. Hilmen, and M. Simon.** 1977. Recombinational switch for gene expression. Science **196**:170–172.