

NOTES

The Central, Surface-Exposed Region of the Flagellar Hook Protein FlgE of *Campylobacter jejuni* Shows Hypervariability among Strains

EDELTRAUD LÜNEBERG,^{1*} EDUARDO GLENN-CALVO,^{2†} MAIKE HARTMANN,¹
WERNER BÄR,^{2‡} AND MATTHIAS FROSCH¹

Institut für Hygiene und Mikrobiologie, Universität Würzburg, Würzburg,¹ and Institut für Medizinische Mikrobiologie, Medizinische Hochschule Hannover, Hannover,² Germany

Received 29 December 1997/Accepted 7 May 1998

In a previous study, we observed that monoclonal antibodies raised against the hook protein FlgE of *Campylobacter jejuni* LIO 36, isolate 5226, bound exclusively to this strain. The aim of this study was to elucidate the molecular basis for these binding specificities. The hook protein-encoding gene *flgE* of *C. jejuni* was cloned in *Escherichia coli* and sequenced. The *flgE* genes of four additional *C. jejuni* strains were amplified by PCR and also sequenced. Comparison of the deduced amino acid sequences revealed a high degree of variability in the central parts of the FlgE proteins among the strains, including variable and hypervariable domains. These findings may indicate a selective pressure of *C. jejuni* hosts, forcing the bacteria to generate variations in surface-exposed antigenic determinants.

Campylobacter jejuni is a common cause of food-borne human gastroenteritis. Motility of the bacterium is required for successful colonization of the intestinal tract and is therefore considered a major virulence determinant (15, 22). Flagella, the locomotory organelles of bacteria, are composed of three structural units: the filament, the hook, and the basal structure. The basal structure is anchored in the outer and inner membranes, whereas the hook and filament are located on the cell surface. The flagellar filament of *Campylobacter* has been extensively studied with regard to its genetics and its biochemical and immunological properties. In particular, the unsheathed flagellum is an immunodominant antigen (23, 30) and undergoes phase and antigenic variation (3, 10, 19). The flagellin protein of *C. jejuni* carries a terminal sialic acid (5, 7). This is so far the only description of a sialyl modification of bacterial flagellin. Two flagellin genes, *flaA* and *flaB*, are present on the *C. jejuni* genome and are known to be differentially expressed (8, 24).

Much less is known about the hook, which connects the filament to the basal body and functions as a joint to transmit the rotation of the rod of the basal body to the filament. In *C. jejuni*, in addition to the flagellin protein, a 92-kDa surface protein has been identified as exhibiting serological heterogeneity, and it has been speculated that this protein comprises the flagellar hook subunit (13, 20, 30). Power et al. (27) investigated purified hook proteins of *C. jejuni* and *Campylobacter coli* strains, which have molecular masses of 92.5 to 94 kDa. Immunochemical analysis with hook-specific antisera showed that serospecific epitopes were immunodominant (27). In ad-

dition, in a very recent study, Kinsella et al. (16) reported the cloning of the *flgE* gene from *C. coli*.

In a previous study in our laboratory, monoclonal antibodies (MAbs) were raised against the purified hook of *C. jejuni* LIO 36, clinical isolate 5226 (6). All of these MAbs bound exclusively to that strain, from which the hook had been isolated and used for immunization of mice. None of the MAbs reacted with any other tested *C. jejuni* LIO 36 strain. Likewise, the MAbs did not bind to any tested *C. jejuni* strain of other LIO serotypes (6). The aim of this study was to elucidate the molecular basis of this binding specificity. Therefore, we attempted to clone and express the *flgE* gene, which encodes the flagellar hook protein of *C. jejuni*, in *Escherichia coli*. To further investigate the antigenic variability of the hook protein, we sequenced the PCR-amplified *flgE* genes of different *C. jejuni* strains.

Cloning and expression in *E. coli* of the *flgE* gene from *C. jejuni*. A genomic expression library was constructed by ligating chromosomal DNA fragments of 5 to 6 kb from *C. jejuni* LIO 36 (isolate 5226) into the expression vector pGEX-3X, which can be used to generate glutathione *S*-transferase fusion proteins. Screening of colonies for expression of the *flgE* gene was done with MAb 02B5 by colony blotting, performed essentially as described elsewhere (21). MAb 02B5 had been raised against the purified hook filament of *C. jejuni* LIO 36, isolate 5226 (6). The 02B5-reactive clone DH5 α (pCAH32) was isolated and further analyzed. Western blot analysis revealed that a 92-kDa protein, identical to the size of the native *C. jejuni* hook protein, was synthesized in DH5 α (pCAH32) (Fig. 1). Therefore, cloning of the full-length hook gene was assumed. As can be seen in Fig. 1, proteolytic degradation of the hook subunit protein occurs in *C. jejuni* as well as in the *E. coli* host strain.

Sequence analysis of plasmid pCAH32. Sequence analysis of the 5.7-kb insert of plasmid pCAH32 revealed an open reading frame of 2,586 bp, which encoded an 862-amino-acid protein with a calculated molecular mass of 91,560 Da and an isoelectric point of 4.56. The calculated molecular mass corresponds well

* Corresponding author. Mailing address: Institut für Hygiene und Mikrobiologie, Universität Würzburg, Josef-Schneider-Str. 2, 97080 Würzburg, Germany. Phone: 49 (931) 201-3936. Fax: 49 (931) 201-3445. E-mail: elueneberg@hygiene.uni-wuerzburg.de.

† Present address: University of Costa Rica, San José, Costa Rica.

‡ Present address: Carl-Thiem-Klinikum, Cottbus, Germany.

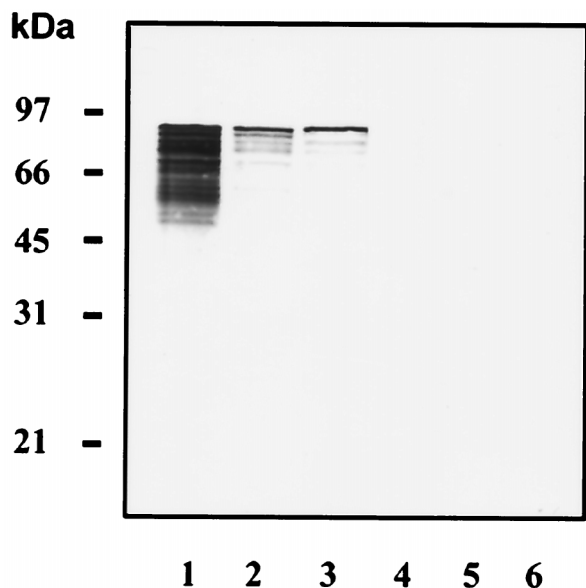


FIG. 1. Western blot analysis of native and recombinant *C. jejuni* hook protein. Bacterial cells (10^8) from fresh overnight cultures were loaded in each lane. By Ponceau S staining of nitrocellulose membranes, the presence of equal amounts of protein in all of the lanes was confirmed. Immunoblots were stained with MAb 02B5. Lanes: 1, *E. coli* DH5 α (pCAH32) (P_{tac} promoter of pGEX-3X induced by IPTG [isopropyl- β -D-thiogalactopyranoside]); 2, *E. coli* DH5 α (pCAH32) without IPTG; 3, *C. jejuni* LIO 36 (isolate 5226); 4, *C. jejuni* LIO 36 (isolate 9907); 5, *C. jejuni* LIO 36 (isolate 567); 6, *C. jejuni* LIO 36 (isolate 68).

to the molecular mass of 92 kDa of the native and recombinant hook proteins as determined by Western blot analysis (Fig. 1).

Comparison of the *C. jejuni* *flgE* sequence with the 2,553-bp *flgE* gene of *C. coli* (16) revealed 78.5% identity. On the amino acid level, the identity was found to be 79% between the two species. The 2,154-bp *flgE* gene of the closely related species *Helicobacter pylori* (25) was found to exhibit 50.6% identity to the *C. jejuni* *flgE* sequence. The amino acid sequences of the FlgE proteins of *C. jejuni* and *H. pylori* showed 49.4% identity. The G+C content of the *C. jejuni* *flgE* coding sequence amounts to 38%. Six nucleotides upstream from the start codon is a potential Shine-Dalgarno sequence. Approximately 100 bases upstream from the start codon is a putative σ^{54} promoter sequence (GGAACAGAACTTGC) (4). A possible σ^{54} promoter has also been detected upstream of the *H. pylori* *flgE* gene (25), and for *C. coli* it has been shown that *flgE*, as well as *flaB*, is under the control of the alternative sigma factor σ^{54} (1, 16). The *C. jejuni* σ^{54} *flgE* promoter sequence is completely identical to that of *C. coli* (16).

Preceding the 242-bp intergenic region upstream from *flgE* is an open reading frame (ORF A) with homology to the *ruvC* gene of *E. coli* (28). Downstream from the *flgE* gene are two further open reading frames, designated ORF B and ORF C. Following a 429-bp intergenic spacer downstream from *flgE*, which does not resemble any known transcriptional terminator sequence, is a 1,268-bp open reading frame (ORF B) oriented in the same direction of transcription. The deduced amino acid sequence of ORF B showed 43% identity to the 42.6-kDa methionine-gamma lyase *megL* of *Pseudomonas putida* (12). ORF C is incompletely present on plasmid pCAH32 and has homology to *yeeE* of *E. coli*, a hypothetical 38-kDa integral membrane protein of unknown function (2). A physical map of the 5.7-kb insert of pCAH32 is depicted in Fig. 2.

The N-terminal sequence of the *C. jejuni* FlgE protein determined in this study exhibited differences from the N-terminal sequence of *C. jejuni* VC74 serotype LIO 11, which has

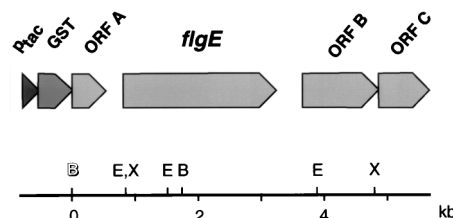


FIG. 2. Physical map of the 5.7-kb insert of plasmid pCAH32. The location of the P_{tac} promoter and the glutathione *S*-transferase (GST) gene of the cloning vector pGEX-3X and the orientation of the insert are depicted. Restriction sites for the following endonucleases are noted: E, *EcoRI*; B, *BamHI* (open letter indicates cloning site); X, *XmnI*.

been defined by protein sequencing (27). Four of 17 amino acids (positions 6, 10, 11, and 15) were not identical.

Sequence determinations of *flgE* genes from different *C. jejuni* strains. In order to elucidate the molecular basis for the binding characteristics of MAbs raised against the purified hook of *C. jejuni* LIO 36 (isolate 5226), we sequenced the *flgE* genes of different *C. jejuni* strains. All strains employed in this study are listed in Table 1. Using primers derived from the cloned *C. jejuni* *flgE* gene, we amplified the *flgE* gene of two further isolates of the LIO 36 serotype (isolates 2772 and 9907) as well as of one strain of the LIO 4 (isolate 5231) and LIO 7 (isolate 5232) serotypes, respectively. For variable regions, specific primers for PCR and sequencing were designed for each strain. The PCR fragments were purified and directly sequenced. Sequence data covering amino acid residues 102 to 736 were analyzed.

In Fig. 3, an alignment of the amino acid sequences as deduced from the nucleotide sequences of the five investigated *C. jejuni* strains, as well as of *C. coli* (16), is shown. Toward the N and C termini, there were highly conserved regions with only minor differences among all investigated *Campylobacter* strains. In contrast, in the central part of the hook protein, several regions with highly variable sequences were observed. In particular, seven distinct domains consisting of either semi-variable (>20% sequence identity), variable (10 to 20% identity), or hypervariable stretches (<10% identity) were defined. These domains are separated by conserved sequences of various lengths. The variable domains (I through VII) are depicted in Fig. 3. Domains I, III, and VII are semivariable, with 31, 40, and 46% sequence identity, respectively. Variable domain VI exhibits 17% sequence identity. Among the hypervariable domains, II, IV, and V, domain IV comprises the largest block of completely variable amino acids. The sequence pattern of hypervariable domain IV indicates a mosaiclike arrangement.

Our data show that within the species *C. jejuni*, the hook protein exhibits immense differences in amino acid composition. Moreover, these differences were also observed among three strains belonging to one serotype (LIO 36) of *C. jejuni*.

TABLE 1. *C. jejuni* strains used in this study

Serotype (isolate no.)	Source
LIO 36 (5226) ^a	W. Bär, Hannover, Germany
LIO 36 (9907)	J. Bockemühl, Hamburg, Germany
LIO 36 (567)	J. Bockemühl, Hamburg, Germany
LIO 36 (68)	J. Bockemühl, Hamburg, Germany
LIO 36 (2772)	J. P. Butzler, Brussels, Belgium
LIO 4 (5231)	W. Bär, Hannover, Germany
LIO 7 (5232)	W. Bär, Hannover, Germany

^a Purified hook protein from this strain was used for immunization of mice in order to raise MAbs. Genomic DNA from this strain was employed to establish a *C. jejuni* expression library in *E. coli*.

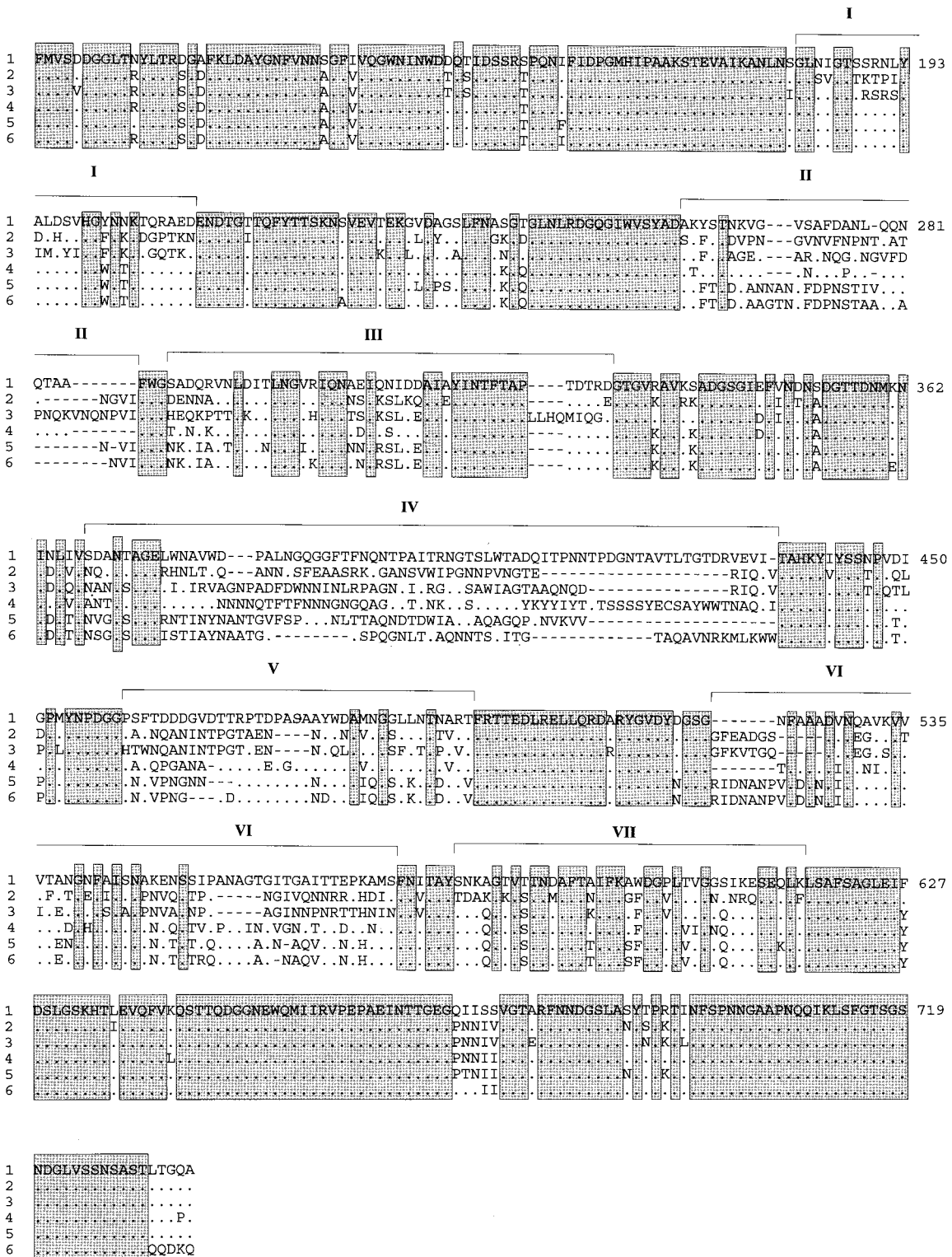


FIG. 3. Alignment of the amino acid sequences of FlgE hook proteins from different *C. jejuni* strains. 1, LIO 36 (isolate 5226 [sequence derived from the cloned *flgE* gene]); 2, LIO 36 (isolate 2772); 3, LIO 36 (isolate 9907); 4, LIO 4 (isolate 5231); 5, LIO 7 (isolate 5232); 6, *C. coli*. Identical amino acids are represented by dots. Dashes represent gaps introduced for optimal alignment. Boxed sequences indicate amino acids which are identical not only among *C. jejuni* strains but also to *C. coli* FlgE (16). Numbering of amino acids on the right refers to the cloned *C. jejuni* FlgE sequence.

The majority of the amino acids that are conserved among *C. jejuni* strains are also conserved in *C. coli* (Fig. 3). The same holds true for the variable regions: in domains with high variability among *C. jejuni* strains, the *C. coli* sequence also differs from the *C. jejuni* sequences.

Moreover, apart from hypervariable domain IV, *C. jejuni* LIO 7 (isolate 5232) reveals an FlgE sequence that is generally nearly identical to that of *C. coli*. This finding might confirm that the division of these thermophilic campylobacters into the distinct species *C. jejuni* and *C. coli* is rather artificial.

Conclusions. With a molecular mass in the range of 90 to 94 kDa, the hook protein of *Campylobacter* species is larger than that of any other bacterial hook known so far. The closely related species *H. pylori* and *Helicobacter mustelae* have FlgE proteins of 78 and 87 kDa, respectively (25), whereas the hook proteins from enterobacteria and spirochetes are considerably smaller, ranging from 42 to 50 kDa (2, 11, 14, 17, 18). The differences in size of hook subunit proteins from bacteria belonging to different genera have been discussed with regard to the requirement for motility in a viscous environment, such as the mucus in the gastrointestinal tract (26, 27). Nevertheless, the relation of hook protein size to hook function is still unclear.

The N-terminal region of the hook protein is required for its secretion, whereas the C-terminal region is necessary for assembly of the hook filament (16). The functional relevance of the N and C termini is reflected in their conserved amino acid composition. In contrast, the functional significance of the central part of the hook protein, whose sequence varies greatly among different species, is not known. In *C. jejuni*, the large central region of the hook protein exhibits hypervariability among strains of one species and even of one serotype. The variable and hypervariable domains are likely to be exposed to the surface, since immunoelectron microscopy reveals that the MAbs employed in our studies bind to the intact hook (6). The surface exposure of variable domains may reflect the need of *C. jejuni* to generate antigenic diversity. It is conceivable that *C. jejuni* responds to the selective pressure of its hosts by altering surface-exposed antigenic determinants. Variable structures may be of selective advantage especially in areas where *C. jejuni* is endemic and reinfections of hosts occur frequently. Host immunity acquired from prior infections is useless when novel antigens are exposed during reinfections. It remains unclear if there is any functional relevance of the variable region apart from the selective advantages it confers for antigenic diversity and immune escape. Conserved and variable domains are also features of the *C. jejuni* *flaA* and *flaB* genes, which encode flagellin. These two highly homologous genes are present in the *Campylobacter* genome, and variations at the loci encoding flagellin have been shown to occur via intragenomic and intergenomic recombination (9, 29). Horizontal gene transfer may be another mechanism of generating highly variable domains in the central part of the *C. jejuni* *flgE* gene.

Nucleotide sequence accession numbers. The nucleotide sequence of the *C. jejuni* LIO 36, isolate 5226, *flgE* gene has been deposited in the EMBL database under accession no. AJ002074. The sequences of the PCR-amplified *flgE* genes from other *C. jejuni* strains have been deposited in the EMBL database under accession no. AJ224790 (LIO 36, isolate 2772), AJ224791 (LIO 36, isolate 9907), AJ224792 (LIO 4, isolate 5232), and AJ224793 (LIO 7, isolate 5231).

E.G.-C. was supported by a grant from the Deutscher Akademischer Austauschdienst.

REFERENCES

- Alm, R. A., P. Guerry, and T. J. Trust. 1993. The *Campylobacter* σ^{54} *flaB* flagellin promoter is subject to environmental regulation. *J. Bacteriol.* **175**: 4448–4455.

- Blattner, F. R., G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Colladovides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**:1453–1474.
- 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.
- Coppard, J. R., and M. J. Merrick. 1991. Cassette mutagenesis implicates a helix-turn-helix motif in promoter recognition by the novel RNA polymerase sigma factor sigma 54. *Mol. Microbiol.* **5**:1309–1317.
- Doig, P., N. Kinsella, P. Guerry, and T. J. Trust. 1996. Characterization of a post-translational modification of *Campylobacter* flagellin: identification of a sero-specific glycosyl moiety. *Mol. Microbiol.* **19**:379–387.
- Glenn-Calvo, E., W. Bär, and M. Frosch. 1994. Isolation and characterization of the flagellar hook of *Campylobacter jejuni*. *FEMS Microbiol. Lett.* **123**:299–304.
- Guerry, P., P. Doig, R. A. Alm, D. H. Burr, N. Kinsella, and T. J. Trust. 1996. Identification and characterization of genes required for post-translational modification of *Campylobacter coli* VC167 flagellin. *Mol. Microbiol.* **19**:369–378.
- Guerry, P., S. M. Logan, S. Thornton, and T. J. Trust. 1990. Genomic organization and expression of *Campylobacter* flagellin genes. *J. Bacteriol.* **172**:1853–1860.
- Harrington, C. S., F. M. Thomson-Carter, and P. E. Carter. 1997. Evidence for recombination in the flagellin locus of *Campylobacter jejuni*: implications for the flagellin gene typing scheme. *J. Clin. Microbiol.* **35**:2386–2392.
- Harris, L. A., S. M. Logan, P. Guerry, and T. J. Trust. 1987. Antigenic variation of *Campylobacter* flagella. *J. Bacteriol.* **169**:5066–5071.
- Homma, M., D. J. DeRosier, and R. M. Macnab. 1990. Flagellar hook and hook-associated proteins of *Salmonella typhimurium* and their relationship to other axial components of the flagellum. *J. Mol. Biol.* **213**:819–832.
- Inoue, H., K. Inagaki, M. Sugimoto, N. Esaki, K. Soda, and H. Tanaka. 1995. Structural analysis of the L-methionine gamma-lyase gene from *Pseudomonas putida*. *J. Biochem. (Tokyo)* **117**:1120–1125.
- Jin, T., and J. L. Penner. 1988. Role of the 92.5-kilodalton outer membrane protein of *Campylobacter jejuni* in serological reactions. *J. Clin. Microbiol.* **26**:2480–2483.
- Jwang, B., P. Dewing, E. Fikrig, and R. A. Flavell. 1995. The hook protein of *Borrelia burgdorferi*, encoded by the *flgE* gene, is serologically recognized in Lyme disease. *Clin. Diagn. Lab. Immunol.* **2**:609–615.
- Ketley, J. M. 1997. Pathogenesis of enteric infection by *Campylobacter*. *Microbiology* **143**:5–21.
- Kinsella, N., P. Guerry, J. Cooney, and T. J. Trust. 1997. The *flgE* gene of *Campylobacter coli* is under the control of the alternative sigma factor σ^{54} . *J. Bacteriol.* **179**:4647–4653.
- Li, H., J. Ruby, N. Charon, and H. Kuramitsu. 1996. Gene inactivation in the oral spirochete *Treponema denticola*: construction of a *flgE* mutant. *J. Bacteriol.* **178**:3664–3667.
- Limberger, R. J., L. L. Slivinski, and W. A. Samsonoff. 1994. Genetic and biochemical analysis of the flagellar hook of *Treponema phagedenis*. *J. Bacteriol.* **176**:3631–3637.
- 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.
- Logan, S. M., and T. J. Trust. 1983. Molecular identification of surface protein antigens of *Campylobacter jejuni*. *Infect. Immun.* **42**:675–682.
- Meyer, T. F., N. Mlawer, and M. So. 1982. Pilus expression in *Neisseria gonorrhoeae* involves chromosomal rearrangement. *Cell* **30**:45–52.
- Morooka, T., A. Umeda, and K. Amako. 1985. Motility as an intestinal colonization factor for *Campylobacter jejuni*. *J. Gen. Microbiol.* **131**:1973–1980.
- Nachamkin, I., and X. H. Yang. 1989. Human antibody response to *Campylobacter jejuni* flagellin protein and a synthetic N-terminal flagellin peptide. *J. Clin. Microbiol.* **27**:2195–2198.
- Nuijten, P. J., F. J. van Asten, W. Gastra, and B. A. van der Zeijst. 1990. Structural and functional analysis of two *Campylobacter jejuni* flagellin genes. *J. Biol. Chem.* **265**:17798–17804.
- O'Toole, P. W., M. Kostrzynska, and T. J. Trust. 1994. Non-motile mutants of *Helicobacter pylori* and *Helicobacter mustelae* defective in flagellar hook production. *Mol. Microbiol.* **14**:691–703.
- Ottmann, K. M., and J. F. Miller. 1997. Roles for motility in bacterial host interactions. *Mol. Microbiol.* **24**:1109–1117.
- Power, M. E., R. A. Alm, and T. J. Trust. 1992. Biochemical and antigenic properties of the *Campylobacter* flagellar hook protein. *J. Bacteriol.* **174**: 3874–3883.
- Takahagi, M., H. Iwasaki, A. Nakata, and H. Shinagawa. 1991. Molecular analysis of the *Escherichia coli* *ruvC* gene, which encodes a Holliday junction-specific endonuclease. *J. Bacteriol.* **173**:5747–5753.
- Wassenaar, T. M., B. N. Fry, and B. A. van der Zeijst. 1995. Variation of the flagellin gene locus of *Campylobacter jejuni* by recombination and horizontal gene transfer. *Microbiology* **141**:95–101.
- Wenman, W. M., J. Chai, T. J. Louie, C. Goudreau, H. Lior, D. G. Newell, A. D. Pearson, and D. E. Taylor. 1985. Antigenic analysis of *Campylobacter* flagellar protein and other proteins. *J. Clin. Microbiol.* **21**:108–112.