Microbiology (2000), **146**, 957–966 Printed in Great Britain

A *Clostridium difficile* gene encoding flagellin

Albert Tasteyre,[1] Marie-Claude Barc,[1] Tuomo Karjalainen,[1] Paul Dodson,[2] Susan Hyde,[2] Pierre Bourlioux[1] and Peter Borriello[2,3]

Author for correspondence: Marie-Claude Barc. Tel: +33 1 46 83 55 49. Fax: +33 1 46 83 58 83. e-mail: marie-claude.barc@cep.u-psud.fr

[1] Université de Paris-Sud, Faculté de Pharmacie, Département de Microbiologie, 5 rue JB Clément, 92296 Châtenay-Malabry cedex, France

[2] Institute of Infection and Immunity, Queen's Medical Centre, University of Nottingham, University Park, Nottingham NG7 2RD, UK

[3] PHLS Central Public Health Laboratory, 61 Colindale Ave, London NW9 5HT, UK

Six strains of *Clostridium difficile* examined by electron microscopy were found to carry flagella. The flagella of these strains were extracted and the N-terminal sequences of the flagellin proteins were determined. Four of the strains carried the N-terminal sequence MRVNTNVSAL exhibiting up to 90% identity to numerous flagellins. Using degenerate primers based on the N-terminal sequence and the conserved C-terminal sequence of several flagellins, the gene encoding the flagellum subunit (*fliC*) was isolated and sequenced from two virulent strains. The two gene sequences exhibited 91% inter-strain identity. The gene consists of 870 nt encoding a protein of 290 amino acids with an estimated molecular mass of 31 kDa, while the extracted flagellin has an apparent molecular mass of 39 kDa on SDS-PAGE. The FliC protein displays a high degree of identity in the N- and C-terminal amino acids whereas the central region is variable. A second ORF is present downstream of *fliC* displaying homology to glycosyltransferases. The *fliC* gene was expressed in fusion with glutathione *S*-transferase, purified and a polyclonal monospecific antiserum was obtained. Flagella of *C. difficile* do not play a role in adherence, since the antiserum raised against the purified protein did not inhibit adherence to cultured cells. PCR-RFLP analysis of amplified flagellin gene products and Southern analysis revealed inter-strain heterogeneity; this could be useful for epidemiological and phylogenetic studies of this organism.

Keywords: *Clostridium difficile*, flagella, flagellin, PCR, cloning

INTRODUCTION

Clostridium difficile is recognized as the major aetiological agent of nosocomial infections such as pseudomembranous colitis and antibiotic-associated colitis and diarrhoea (Bartlett *et al*., 1978; Larson *et al*., 1978; George, 1984). Several confirmed and putative virulence factors that could play a role in C. *difficile* pathogenesis have been identified. Virulence is mainly due to the production of two protein exotoxins, toxins A and B, although other toxic factors have been described (Lyerly *et al*., 1988; Perelle *et al*., 1997). There is considerable evidence that some strains are more virulent than others; a number of factors have been proposed to explain inter-strain diversity: (i) capsule, an anti-phagocytic factor (Davies & Borriello, 1990); (ii) proteolytic enzymes, which may play a role in releasing suitable substrates from available protein sources for metabolism and which could be involved in mucus penetration (Seddon & Borriello, 1992; Poilane *et al*., 1998); (iii) adhesins which are involved in mucus and cell association (Borriello *et al*., 1988b; Eveillard *et al*., 1993; Karjalainen *et al*., 1994); (iv) fimbriae, the role of which is obscure (Borriello *et al*., 1988a); and (v) flagella, the potential role of which in colonization is under study. Delmée *et al*. (1990) have developed a serogrouping method for C. *difficile* and flagella appear to be responsible for cross-reactions between strains. In other bacteria, flagella have been implicated in the adherence to mucus and cells and colonization by *Pseudomonas aeruginosa* (Ritchings *et al*., 1995; Arora *et al*., 1996), *Vibrio cholerae* (Richardson, 1991; Gardel & Mekalanos, 1996), *Vibrio anguillarum* (McGee *et al*., 1996; Milton *et al*., 1996), *Helicobacter pylori* (Eaton *et al*., 1996) and *Burkholderia pseudomallei* (Brett *et al*., 1994). Furthermore, they contribute to the invasiveness of *Campylobacter jejuni* (Morooka *et al*., 1985; Grant *et*

Abbreviation: GST, glutathione *S*-transferase.

The GenBank accession numbers for the sequences reported in this paper are AF065259 (strain 79-685) and AF077341 (strain VPI 10463).

al., 1993; Szymanski *et al*., 1995), *Salmonella typhi* (Liu *et al*., 1988) and *Proteus mirabilis* (Mobley *et al*., 1996).

One aspect of *C. difficile* virulence that has been studied by us is its interaction with target cells (Eveillard *et al*., 1993; Karjalainen *et al*., 1994). Adhesion and colonization of animal tissue by bacteria is an important step in establishing infection. It is probable that without attachment, *C. difficile* cannot colonize and will be quickly removed by non-specific host defence mechanisms.

We are interested in finding out whether flagella play a role in *C. difficile* colonization. In this study we undertook the isolation of *C. difficile* flagella from clinical strains and characterization and expression of the flagellin subunit gene. In addition, Southern analysis and PCR amplification of flagellin genes coupled with RFLP analysis were used in a preliminary attempt to differentiate between clinical isolates.

METHODS

Bacterial strains and media. Six *C. difficile* isolates were investigated and are presented in Table 1. The VPI 10463 strain was obtained from Dr Wilkins (Virginia Polytechnic Institute, Blacksburg, USA). Strains of *Clostridium perfringens*, *Bacillus subtilis* and *Clostridium sordellii* (Institut Pasteur, Paris, France) were used as negative controls. The strains were grown under anaerobic conditions on agar plates (Oxoid) supplemented with 7% horse blood (bioMérieux) or in TGY (tryptone/glucose/yeast infusion broth; Difco) for 48 h.

Escherichia coli strain XL-1 Blue MRF′ (Stratagene) was used as a host in general cloning procedures and *E. coli* BL21 (Pharmacia) was used for expression and purification of GST–FliC fusion protein.

Flagellin protein of *C. difficile*: isolation, characterization and N-terminal sequencing. Flagellin proteins were isolated by the procedure described by Delmée *et al*. (1990). The strains were grown on blood agar plates under anaerobic conditions for 48 h. Bacteria were harvested in 5 ml distilled water; the suspensions were strongly shaken for 1 min and centrifuged at 5000 ***g*** for 30 min at 4 °C. The supernatants were centrifuged at 25000 ***g*** for 1 h at 4 °C and the pellets were suspended in 100 µl PBS (pH 7·4).

SDS-PAGE was carried out as described by Laemmli (1970) using an SDS-PAGE gel (7·5%, w/v, acrylamide). Gels were stained with Coomassie blue or used for electric transfer onto nitrocellulose membrane for immunoblotting. The nitrocellulose membrane was incubated for 30 min at room temperature in blocking buffer (0·2% Tween, 3% skim milk in PBS) and then overnight in a rabbit polyclonal antiflagellin serum to serogroup A (strain W1194), kindly provided by M. Delmée (UCL, Brussels, Belgium) (1:2000 dilution). The membranes were washed in 10 mM Tris/HCl, 150 mM NaCl, 0·05% Tween 20 buffer (TNT) and bound antibodies were detected with goat anti-rabbit IgG alkaline phosphatase conjugate (1:2500 dilution; Sigma) with nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate as substrates.

For the N-terminal sequencing of the flagellin, proteins were electroblotted onto a Hybond-N membrane (Amersham) with CAPS buffer (Sigma) and stained with amido black. The bands of interest were excised and N-terminal analysis was performed (Polymer synthesis and Analysis unit, Department of Biochemistry, University of Nottingham). The method of Edman (1950) was used with an Applied Biosystems 473A amino acid sequencer on a PTH-RP-PHLC C18 column. Identification was performed using 610A analysis software (Applied Biosystems).

Electron and immunoelectron microscopy of flagella and strains. Flagellum preparations or 24 h cultures of *C. difficile* strains as described above were resuspended in distilled water. They were negatively stained with a phosphotungstic acid solution for 5 min and adsorbed on carbon-stabilized nitro-

Table 1. *C. difficile* strains studied

Strain	Origin*	Toxigenicity	Serogroup	N-terminal sequence	Source
51187	Diarrhoea	+	A	ND	Université Catholique de Louvain, Brussels, Belgium
B-1	PMC	+	A	MRVNTNVSAL	Queen's Medical Centre, Nottingham, UK
W1194†	PMC	+	A	MRVNTNVSAL	Université Catholique de Louvain, Brussels, Belgium
79-685	PMC	+	S3	MRVNTNVSAL	Institut de Bactériologie, Strasbourg, France
Kohn	Infant	−	A	MRVNTNVSAL	Queen's Medical Centre, Nottingham, UK
56026	Infant	−	I	ND	Université Catholique de Louvain, Brussels, Belgium

ND, Not determined.

* PMC, pseudomembranous colitis.

† Strain to which Delmée antiserum was raised.

cellulose film copper specimen grids (2 min) by placing each grid on a drop of the bacterium-staining solution. After drying, the grids were observed with an EM 301 Philips transmission electron microscope.

For immunoelectron microscopy, after adsorption of bacteria or flagellum preparation to the grids for 5 min, the grids were floated on PBS with 1% BSA for 30 min, then incubated with antiflagellin antibodies diluted 1:100 for 1 h and washed three times in PBS. They were incubated with a 1:20 dilution of protein A labelled with 15-nm-diameter colloidal gold particles (British Biocen International) for 1 h, washed three times with PBS and subsequently fixed with 3% glutaraldehyde. After three washings the grids were stained with phosphotungstic acid before observation by transmission electron microscopy. A strain of *B. subtilis* was used as a negative control.

Motility assays. Motility assays were conducted in BHI broth with 0·175% agar. The medium was placed in tubes which were inoculated from a colony by stabbing the agar with a toothpick. The tubes were incubated at 37 °C for 2 d under anaerobic conditions.

DNA techniques. Plasmid isolations were performed by the alkaline lysis procedure using a kit from Qiagen. Ligations and restriction endonuclease digestions were done according to Sambrook *et al*. (1989) and protocols provided by the vendors. The TSB (Transformation Storage Buffer) method was used for transformation of *E. coli* (Chung *et al*., 1989).

PCR amplification and nucleotide sequencing of the *C. difficile* flagellin gene. The N-terminal sequence of the *C. difficile* 79-685 flagellin and conserved motifs in the C-terminal sequences of flagellins from various bacteria (*V. cholerae*, *Vibrio parahaemolyticus*, *B. subtilis*, *Clostridium tyrobutyricum*) were used to design degenerate primers synthesized by Life Technologies. The N-terminal primer was ATGMGAGTWAATGTWTCWGCGCTY; the C-terminal primer was TTGWCGAAYTGTTGGTTWGCAGCWAGCAG. Genomic DNA was extracted from *C. difficile* strains as described by Karjalainen *et al*. (1994) or bacteria were harvested from blood agar plates, resuspended in sterile distilled water and boiled for 5 min and used directly in a standard amplification mixture. Amplifications were carried out in 100 µl volumes containing 1 µg genomic DNA, 2 µl $MgCl_2$, 1 U *Taq* polymerase (Promega), 200 pmol of each deoxynucleotide triphosphate, 1 µM of each primer and 10 µl 10 × polymerase buffer for 34 cycles consisting of denaturation at 95 °C (1 min), annealing at 55 °C (1 min) and extension at 72 °C (2 min) (Perkin Elmer Thermocycler 480). At the end of the amplification, 20 µl samples were subjected to electrophoresis on a standard 1% agarose gel alongside a PCR size marker (100 bp ladder; Pharmacia) to confirm the presence of an amplified product. Amplified products were purified by a gel extraction kit (Wizard Gel Extraction kit; Promega).

The nucleotide sequences of both strands of amplified products of the *C. difficile* flagellin of strains 79-685 and VPI 10403 were obtained by using the Taq Dye Deoxy Terminator and Dye Primer cycle sequencing protocols developed by Applied Biosystems (Perkin Elmer) using fluorescence-labelled dideoxynucleotides and primers, respectively. The labelled extension products were analysed with an ABI PRISM 310 Genetic Analyzer (Perkin Elmer). More nucleotide sequence upstream and downstream of the first amplified region of the flagellin gene was obtained by amplification by PCR from the λZapII genomic library (Karjalainen *et al*., 1994), converted into a plasmid library by excision en masse, using a gene-specific primer (ACGAACCTTCTGCTGTTTGTAC) and a primer (GGAAACAGCTATGACCATG) that hybridized with the pBluescript polylinker (M13rev). For the upstream region a PCR product of 500 bp was obtained. For the downstream region a PCR product of 1·1 kb was obtained with the gene-specific primer CTTTAGAGAATGTTACAGCAGC and a vector-specific primer GTAAAACGACGGCCAGT (M13 '−20'). Secondary structure of the flagellin protein was predicted by using the Chou and Fasman algorithm (Chou & Fasman, 1978). Additional sequencing of the flagellin gene was performed with internal primers.

Southern blotting and RFLP analysis. DNA of *C. difficile* strains was prepared as described by Karjalainen *et al*. (1994). DNA was digested with *Hin*dIII and electrically transferred to a nylon membrane (Boehringer Mannheim). The 850 bp amplified PCR product was used as a flagellin gene specific probe. The DNA probe was labelled and detected by using the ECL direct nucleic acid labelling and detection system from Amersham. Washing of membranes was performed under low stringency (0·5 × SSC at 42 °C). For RFLP amplified DNA was obtained by PCR from colonies grown on blood agar plates as described above. The DNA was digested with four restriction enzymes, *Hin*dIII, *Eco*RV, *Dra*I and *Sac*I, under the conditions recommended by the supplier (New England Biolabs). These digests were then subjected to electrophoresis on a 1% agarose gel alongside a PCR size marker (100 bp ladder; Pharmacia).

Cloning, expression, purification and identification of the fusion protein. For the cloning of the *C. difficile* 79-685 flagellin gene into an expression vector, two oligonucleotide primers, CCCCTGGGATCCATGAGAGTTAATACAAATGTAAGTGC and CCGGGAATTCCTATCCTAATAATTGTAAAACTCC, incorporating the *Bam*HI and *Eco*RI restriction site, respectively, were synthesized and used to amplify by PCR the full-length coding region of the *fliC* gene of strain 79-685 (*Taq* polymerase, Promega; 1 U per 100 µl reaction volume). The resulting 895 bp DNA product was digested with *Bam*HI and *Eco*RI and cloned in-frame into the corresponding sites of pGEX-6P-1 (Pharmacia). The nucleotide sequence of the junction between vector and insert was confirmed by sequencing analysis to be correct. The plasmid was transformed into *E. coli* BL21.

For the expression and purification of the fusion protein, an overnight culture of *E. coli* BL21 containing pGEX-6P-1-*fliC* was diluted 1:100 into 4 l 2 × YT medium (Life Technologies) containing ampicillin and the culture was grown to OD_{600} 0·6 at 37 °C. The expression of the fusion protein was induced by adding IPTG at 1 mM for 2 h. Bacteria were collected by centrifugation and resuspended in 200 ml ice-cold PBS. The bacteria were lysed by sonication (intervals of 5 s for 1 h at 80% power; Bioblock Scientific 72442 Vibra Cell). Insoluble material was removed by centrifugation at 8000 ***g*** for 10 min, and the fusion protein was purified from the supernatant by a single-step affinity chromatography using glutathione–Sepharose-4B and protocols from Pharmacia. A 2 ml bed volume was used for each 200 ml sonicate; the column was washed three times with 20 ml PBS, followed by cleavage of the FliC moiety bound to glutathione–Sepharose with 80 units Prescission protease per 1 ml bed volume. Identification of the fusion protein was carried out by SDS-PAGE as described above and by immunoblot (serum dilution 1:2000).

A polyclonal antiflagellin serum was raised against the purified FliC recombinant protein. The gel band located at 39 kDa corresponding to the purified flagellin was cut out and injected into a rabbit. The polyclonal, monospecific antiserum was

obtained according to a protocol described previously (Karjalainen *et al*., 1994) and used at a 1:2000 dilution in Western blots.

Computer analyses. Nucleotide and protein sequence alignments were performed with the CLUSTALX program (Thompson *et al*., 1997). Homology searches were conducted with FASTA3 (EBI) or BLAST (National Institute for Biotechnology Information, Washington).

Adherence inhibition assays. Cell culture and adherence inhibition assays were performed as previously described (Karjalainen *et al*., 1994). For adherence inhibition with antibodies, bacteria were incubated with preimmune serum (dilution 1:2) or immune serum (dilution 1:2) for 1 h prior to contact of 1 h with cultured Vero cells. Non-adherent bacteria were eliminated by five washings in PBS (10 mM phosphate buffer, 150 mM NaCl) (pH 7·0) and the cells were fixed and stained with May–Grünwald–Giemsa (Sigma). The adhesion index is given as the mean number of adhering bacteria per cell (counted at a magnification of ×1000) from at least three different assays.

RESULTS

Flagella of *C. difficile*: isolation, N-terminal analysis and functionality

An SDS-PAGE profile of flagella isolated from the six *C. difficile* strains is shown in Fig. 1(a). Each extract carries a band of 39 kDa corresponding to the flagellin as shown previously by Delmée *et al*. (1990). The rabbit polyclonal flagellin antiserum raised by Delmée to crude flagellin preparations from strain W1194 reacted with a 39 kDa protein in all six strains on Western blots (Fig. 1b). In four strains, the N-terminal sequence of the flagellin was MRVNTNVSAL (Table 1).

Because the flagellin proteins of some other bacteria are glycosylated (Ge *et al*., 1988; Guerry, 1997), the possible glycosylation of the *C. difficile* flagellin was examined by the DIG Glycan detection kit (Roche), which detects glycoconjugates on proteins by immunoassay. The results revealed that flagella of *C. difficile* strain 79-685 were not glycosylated (data not shown). In addition, the functional activity of the flagella of the six strains of *C. difficile* studied was determined by investigating cell motility. All the strains were motile (data not shown).

Detection of flagella by electron microscopy

Electron microscopy was used to observe the flagella of *C. difficile* strains, either on intact bacteria or in isolated flagellum preparations (Fig. 2). Some strains, such as Kohn and W1194, had numerous flagella, whereas other strains carried fewer flagella. It was evident by immunoelectron microscopy that the flagella were labelled by the Delmée flagellin antiserum A. *B. subtilis* and a *C. sordellii* strain, used as negative controls, showed no labelled flagella.

DNA sequence analysis of the flagellin filament gene

PCR amplification with primers described in Methods allowed us to amplify an 850 bp product from the *C. difficile* genome. Sequences upstream and downstream were obtained by amplifying DNA from a genomic

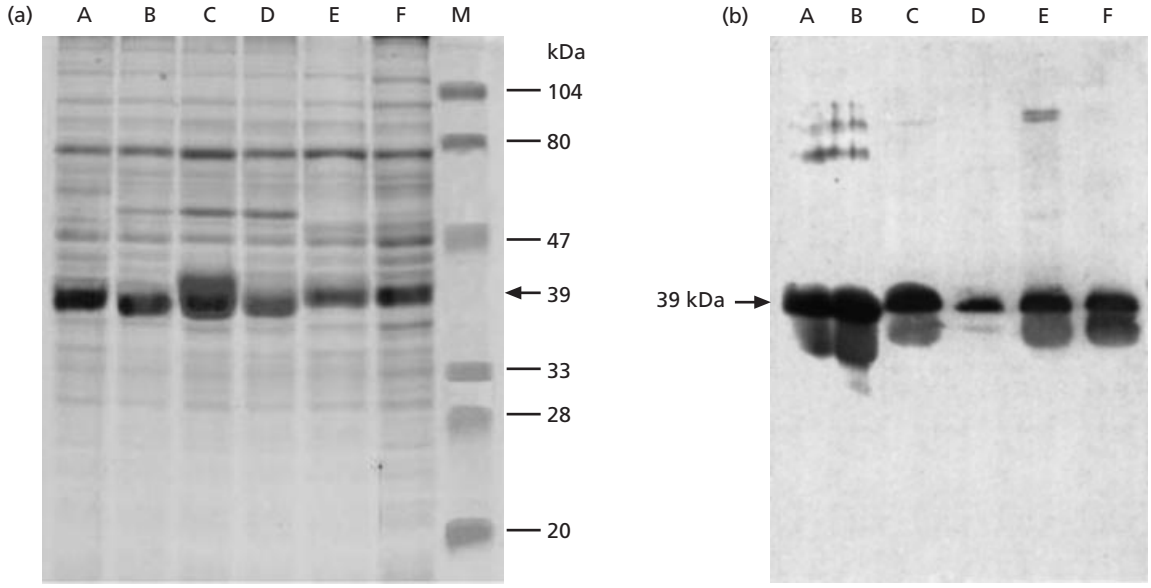

Fig. 1. (a) SDS-PAGE of surface proteins including flagellin from six strains of *C. difficile*. Isolated proteins were denatured in sample buffer and electrophoresed on a 5% stacking gel and 12·5% separating polyacrylamide gel. Proteins were isolated from strains: 79-685 (lane A); 51187 (lane B); W1194 (lane C); B-1 (lane D); 56026 (lane E); Kohn (lane F). Low-molecular-mass standards of 104, 80, 47, 33, 28 and 20 kDa (Bio-Rad) are in lane M. The arrow indicates the band corresponding to the 39 kDa flagellin. (b) Isolated surface proteins reacted with a 1:2000 dilution of polyclonal antiserum raised against flagella of strain W1194 in a Western blot. Order of strains as in Fig. 1(a).

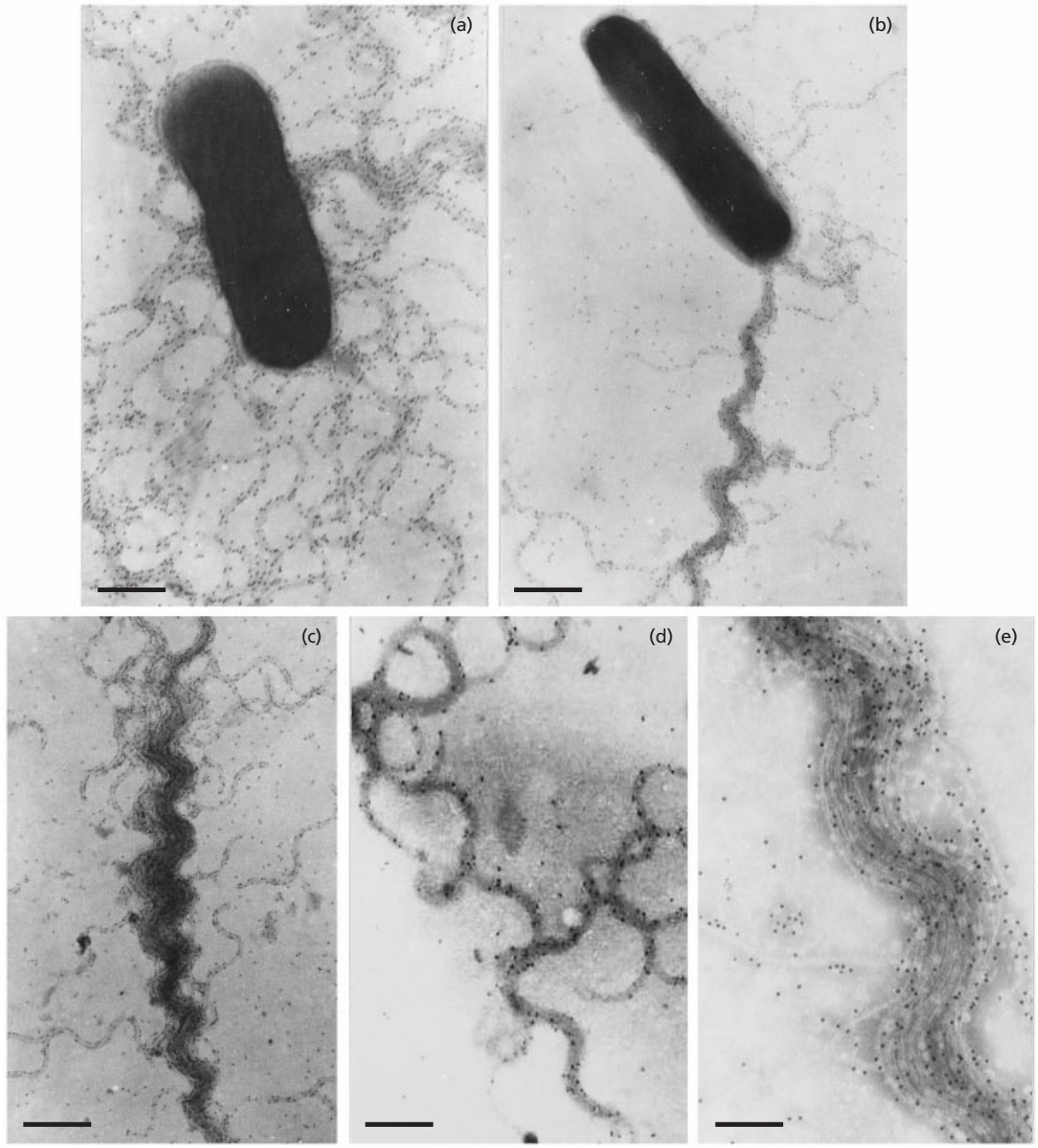

Fig. 2. Electron micrographs showing flagellated *C. difficile* strains Kohn (a) and 56026 (b) and isolated flagella from strain Kohn (c), 51187 (d) and W1194 (e). The flagella were labelled with antiflagellum serum to strain W1194 and visualized with protein A coupled to 15 nm colloidal gold particles. Bars, 1 μm.

library constructed in λZapII. The nucleotide sequence of a 1·6 kb fragment of two strains, 79-685 and VPI 10463, was determined; analysis of the DNA sequence revealed the presence of two ORFs: ORF1 composed of 870 nt corresponding to 290 amino acids, and a partial ORF 90 bases downstream. Based on the comparison of the deduced amino acid sequence with flagellin sequences from both Gram-positive and Gram-negative bacteria, ORF1 was identified as the flagellin gene of *C. difficile*, which we named *fliC*. FliC of *C. difficile* has a calculated molecular mass of 30·9 kDa: thus it differs from the estimated molecular mass of 39 kDa determined by SDS-PAGE (Fig. 1a). The FliC protein (Fig. 3) displays highest homology to the corresponding protein of *C. tyrobutyricum* (61 % identity, 76 % similarity) and exhibits features found in other flagellins: variable central part; predominantly α-helical conformation with frequent alanine and few proline residues; no signal sequence, the sequence LIAN resembling the consensus sequence N(I/L)AN that serves as an export signal for

```
79685   MRVNTNVSALIANNQMGRNVNGQSKSMEKLSSGVRIKRAADDAAGLAISEKMRAQIKGLD 60
VPI     MRVNTNVSALIANNQMGRNVNAQSKSMEKLSSGVRIKRAADDAAGLAISEKMRAQIKGLD 60
        **********************.************************************

79685   QAGRNVQDGISVVQTAEGSLEETGNILQRMRTLSLQSANEINNTEEREKIADELTQLKDE 120
VPI     QAGRNVQDGISVVQTAEGALEETGNILQRMRTLSVQSSNETNTAEERQKIADELLQLKDE 120
        ******************:***************:**:** *.:***:****** *****

79685   IERISSSTEFNGKKLLDGTSSTIRLQVGASYGTNVSGTSNNNNEIKIQLVNTASIMASAG 180
VPI     VERISSSIEFNGKKLLDGSSTEIRLQVGANFGTNVAGTSNNNNEIKVALVNTSSIMSKAG 180
        :****** **********:*: *******.:****:**********: ****:***:.**

79685   ITTASIGSMKAGGTTGTDAAKTMVSSLDAALKSLNSSRAKLGAQQNRLESTQNNLNNTLE 240
VPI     ITSSTIASLNADGTSGTNAAKQMVSSLDVALKELNTSRAKLGAQQNRLESTQNNLNNTIE 240
        **:::*.*::*.**:**:*** ******.***.**:**********************:*

79685   NVTAAESRIRDTDVASEMVNLSKMNILVQASQSMLAQANQQPQGVLQLLGZ 291
VPI     NVTAAESRIRDTDVASEMVNLSKMNILVQASQSMLAQANQQPQGVLQLLGZ 291
        **************************************************
```

Fig. 3. Sequence alignment of the deduced amino acid sequence of FliC of *C. difficile* strains 79-685 and VPI 10463. Identical residues are indicated with an asterisk; functionally identical residues are indicated with a colon. The alignment was performed with the CLUSTAL W program. The proteins show 90·6 % identity (90·3 % at the nucleotide level).

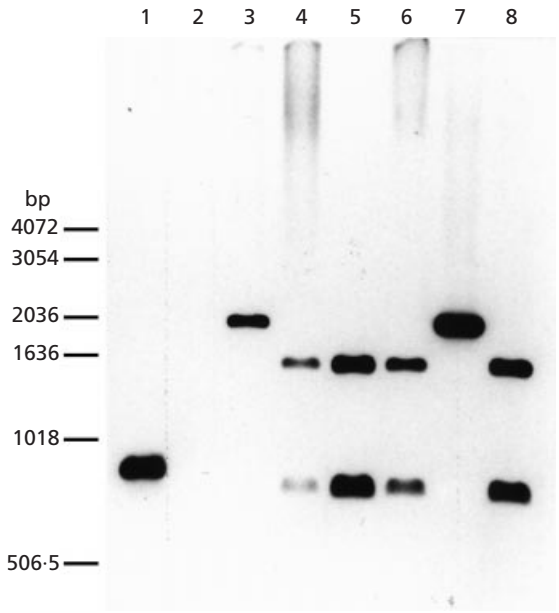

Fig. 4. Southern blot of chromosomal DNA isolated from *C. difficile* strains hybridized under low stringency to the *fliC* probe. DNA was digested with *Hin*dIII, electrophoresed and transferred to nylon membrane. Lanes: 1, 850 bp PCR product used as probe (positive control); 2, DNA from *C. sordellii* (negative control); 3, DNA from strain 79-685; 4, DNA from 51187; 5, DNA from W1194; 6, DNA from B-1; 7, DNA from 56026; 8, DNA from Kohn. The two bands in lanes 4, 5, 6 and 8 are produced by the *Hin*dIII site near base 937 of the flagellin gene. The molecular mass standards (1 kb ladder; Gibco-BRL) are shown on the left.

flagellar subunits (Heinzerling *et al.*, 1997). The gene has a G+C value of 33·5 mol %, a value higher than that of the *C. difficile* genome (28 mol %) (Sneath, 1986). Analysis of the codon usage reveals a marked preference for A or T for the third position of the triplets and a total of 29 codons are not used. A motif similar to the σ^{28} consensus binding site (TAAAGTN$_{12}$GCCGATAA) is found in the promoter: TAAAGTN$_{13}$TCCGATAA. The translational termination codon is followed by an imperfect inverted repeat that can form a stem–loop structure with a ΔG (25 °C) of $-63{\cdot}9$ kJ; it could constitute a ρ-independent transcriptional terminator.

Conservation and expression of the *fliC* gene in *C. difficile* strains

To investigate the conservation of the *fliC* gene region in different strains, we amplified by PCR the flagellin gene from the six *C. difficile* strains. All strains showed the presence of an 850 bp band corresponding to the *fliC* gene (data not shown). The amplified DNA was used as a probe in Southern hybridization of chromosomal DNA of these strains. Hybridization under low stringency conditions showed that DNAs of all isolates studied hybridized with the *fliC*-specific probe. Only one copy of the gene was present in each strain (Fig. 4). Some strains carry a *Hin*dIII site and therefore show the presence of two bands.

The amplified DNA was digested with *Hin*dIII, *Eco*RV, *Dra*I and *Sac*I restriction enzymes; the results are shown in Fig. 5. The six strains were classified into two groups according to the RFLP of the *fliC* gene.

Expression, purification and immunological detection of flagellin

Screening of the *C. difficile* genomic library with the Delmée antiserum revealed that it recognizes a second protein of 39 kDa (GenBank accession no. AF065260) which is not the flagellin (data not shown). Therefore we decided to purify the flagellin protein by recombinant technology in order to produce a polyclonal, monospecific antiserum. The coding region of *fliC* was amplified by PCR and inserted into the *E. coli* expression vector pGEX-6P-1. The GST–FliC fusion protein was purified by affinity chromatography and was cleaved with Prescission protease. As shown in Fig. 6(a), a major 39 kDa band was observed in SDS-PAGE, free of contaminating GST. This band reacted in Western blot with the Delmée antiflagellar antiserum raised against the W1194 strain (not shown). The band was excised from the gel and injected into a rabbit in order to obtain polyclonal antibodies. The monospecific antiserum thus obtained reacted with a 39 kDa protein in the flagellar preparation of all the six strains studied and with the purified flagellin protein (Fig. 6b).

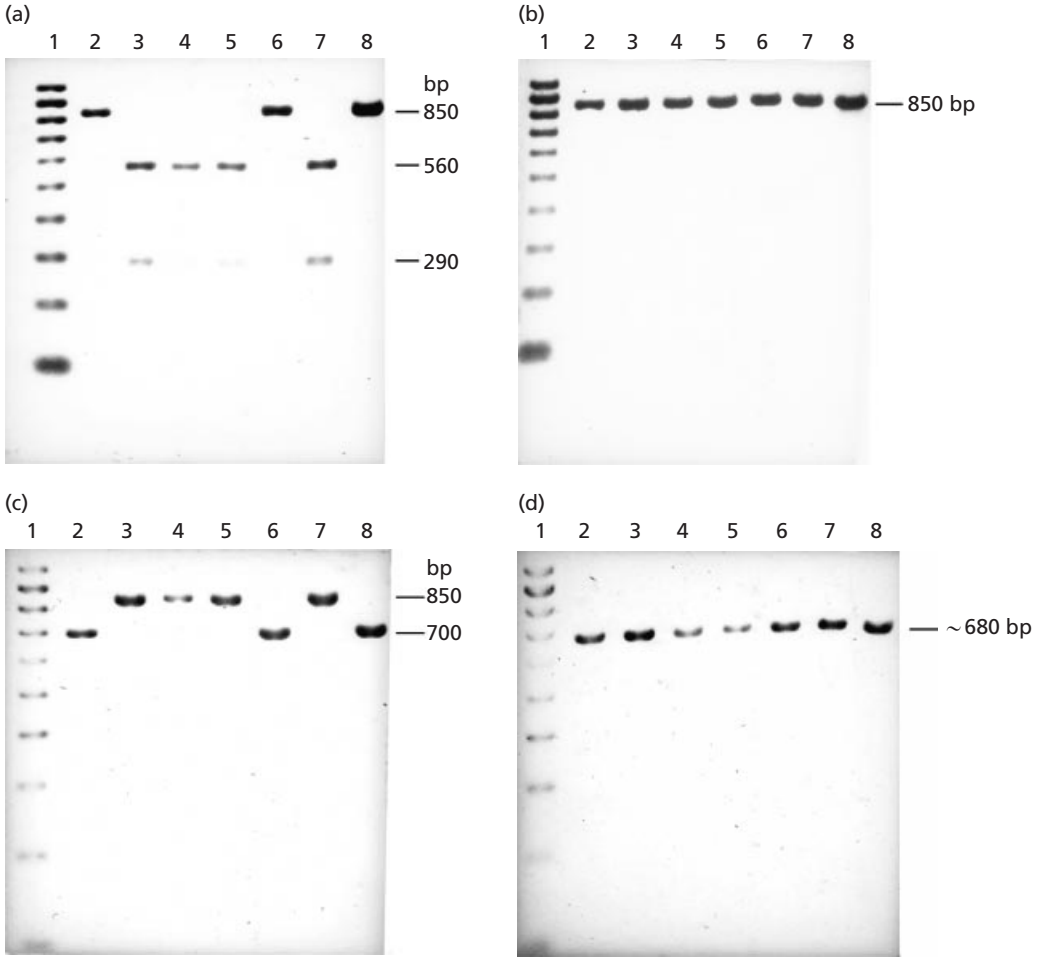

Fig. 5. RFLP patterns of PCR-amplified flagellin genes. Lanes: 1, 100 bp DNA ladder (PCR low ladder; Sigma); 2, strain 79-685; 3, strain 51187; 4, strain W1194; 5, strain B-1; 6, strain 56026; 7, strain Kohn; 8, 850 bp PCR product from the *fliC* gene of strain 79-685 (positive control). Amplified DNA was digested with *Hin*dIII (a), *Eco*RV (b), *Dra*I (c) and *Sac*I (d).

Adherence inhibition assays

Involvement of the flagellin filament protein in adherence of *C. difficile* to eukaryotic cells was investigated in inhibition assays using antiflagellin antibodies raised against the purified protein. Coincubation of bacteria with antibodies at a dilution of 1:2 demonstrated no inhibition of adherence as compared with control adherence of 100% (incubation with preimmune serum), indicating that the flagellin subunit is not involved in the adherence process of *C. difficile* (data not shown).

DISCUSSION

In this study, we have characterized the *C. difficile fliC* gene encoding the flagellin filament, which was sequenced from two strains. This gene is one of the few genes that has so far been isolated from this pathogen. The *fliC* gene seems to be present in a single copy on the chromosome of the strains examined, as in most species studied. The N-terminal region of FliC exhibits an absence of amino acids that are not neutral or hydrophobic, with the exception of one arginine, which is a characteristic of these structural proteins. There is evidence suggesting that amino acids in the N-terminal domain are closely linked to export of these molecules and their subsequent polymerization during biosynthesis of the flagella (Fedorov & Efimov, 1990). A number of structural studies have shown that the N- and C-terminal regions of flagellins are well conserved and play important roles in polymerization and polymorphism of bacterial flagellin filaments (Mimori-Kiyosue *et al.*, 1997). Unlike many other flagellins, the N-terminal methionine residue of *C. difficile* flagellin is not removed post-translationally. Flagellin and several other external components of the bacterial flagellum are thought to be exported by a flagellum-specific pathway involving a central channel in the flagellum itself. Thus no signal peptide is present at the N-terminus. Instead, a consensus sequence resembling the export signal of flagellins, LIAN, is evident. The FliC protein contains three imperfect repetitive motifs, a feature often seen in surface-exposed proteins.

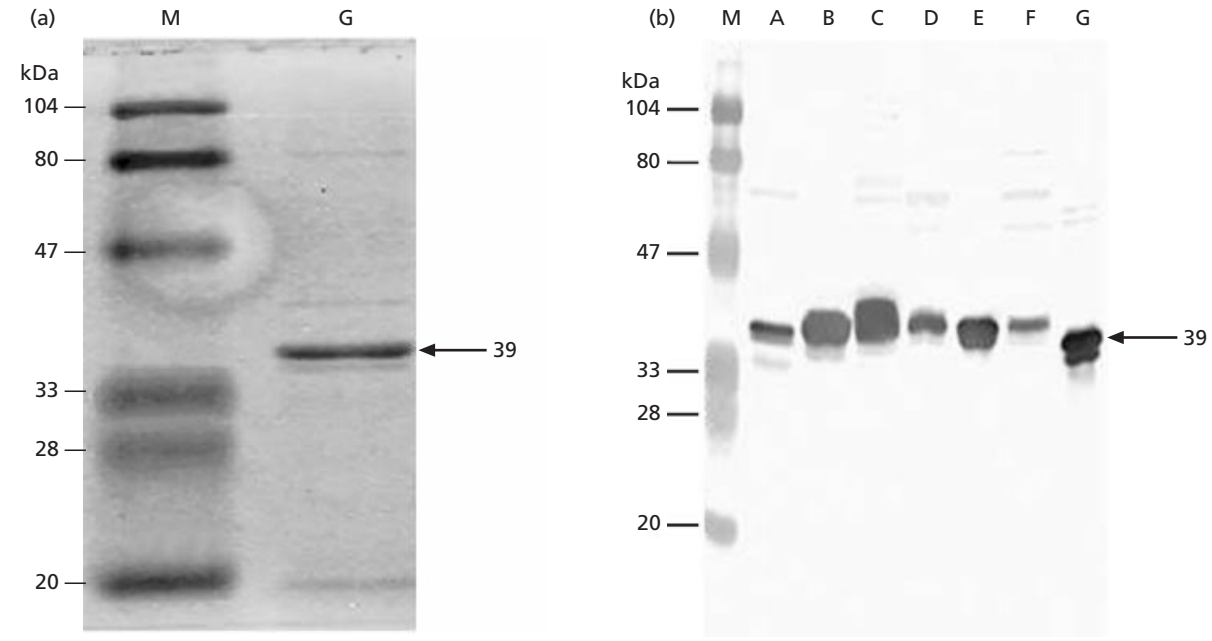

Fig. 6. (a) Purification of *C. difficile* 79-685 FliC protein. SDS-PAGE showing the protein marker (lane M) as in Fig. 1(a) and FliC eluted from glutathione–Sepharose after digestion of GST–FliC with Prescission protease (lane G). A major band at 39 kDa is observed. (b) Isolated *C. difficile* surface proteins reacted with a 1:2000 dilution of polyclonal antiserum raised against purified FliC. Proteins were isolated from strains: 79-685 (lane A); 51187 (lane B); W1194 (lane C); B-1 (lane D); 56026 (lane E); Kohn (lane F). Lane G, FliC eluted from glutathione–Sepharose after digestion of GST–FliC with Prescission protease. The arrow indicates the band corresponding to the 39 kDa flagellin.

Analysis of the promoter structure reveals the presence of a motif resembling the consensus sequence for σ^{28} regulated promoters. However, the distance between the −10 and −35 motifs is 16 nt instead of the usual 15. σ^{28}, which is involved in transcription of the flagellar and chemotaxis genes, was originally found in *B. subtilis* (Gilman & Chamberlin, 1983). Although transcription of some flagellar genes is initiated by σ^{54}, no consensus motif for this factor is present in the *C. difficile fliC* promoter.

The flagella of six strains of *C. difficile* were isolated. The molecular mass of *C. difficile* flagellin, 39 kDa, is in the middle of the range of other characterized flagellin molecules, which have been reported to have molecular masses ranging from 15 to 62 kDa (Arnold *et al.*, 1998; Joys, 1988; Sakamoto *et al.*, 1992; Wilson & Beveridge, 1993). *C. difficile* flagellins did not display heterogeneity between the different strains studied nor did they have multiple molecular masses. Immunoblotting and immunogold labelling of strains with a polyclonal antiserum raised against purified flagellin demonstrated that all *C. difficile* strains reacted with the antiserum, in contrast to strains of *C. sordellii* and *B. subtilis*. This result suggests that the flagellin of each strain contains cross-reacting epitopes due to the presence of the flagellin monomers. Although flagellins generally are structurally conserved in the N- and C-termini, the internal region is divergent and accounts for serological distinctiveness. Prediction of antigenic determinants in FliC of *C. difficile* using the Hopp and Woods algorithm (Hopp & Woods, 1981) revealed the highest probability for the presence of such motifs in the central, variable region between aa 100 and 130. The central region has been proposed to be the region that is exposed to the outer environment (Sakamoto *et al.*, 1992). Like other flagellins, the central portion has a surplus of acidic residues over basic residues and, by implication, a net negative charge. It appears that most bacterial cell surface proteins carry a net negative charge (Wilson & Beveridge, 1993). We suspected that the ORF downstream of the *fliC* gene could encode a glycosyltransferase that could be involved in the *in situ* glycosylation of flagella. If FliC of *C. difficile* were glycosylated, this could explain the difference in molecular mass observed on SDS-PAGE (39 kDa) and that estimated from the nucleotide sequence (31 kDa). In fact, using the DIG Glycan Detection kit we demonstrated that *C. difficile* strain 79-685 flagella are non-glycosylated. The fact that the cloned gene expressed in *E. coli* produces a flagellin with almost the same molecular mass as the native protein from *C. difficile* suggests that the protein undergoes post-translational modification other than glycosylation, since *E. coli* is not able to glycosylate proteins. Flagellar filaments can contain phosphorylated tyrosines or serines, *ε-N*-methyl-lysine, or they can be sulfated glycoproteins.

The isolation of the *fliC* gene and isolation of a monospecific antiserum will allow further investigations as to the role of flagella in the pathogenic process. It has been suggested that the virulence of *C. difficile* strains is

not solely attributable to toxin production; other factors such as presence of flagella could contribute to virulence. The role of flagella in microbial pathogenesis factors such as virulence, adherence, invasiveness or colonization has been demonstrated for numerous bacteria (Grant *et al*., 1993; Scherer *et al*., 1993; Zhang *et al*., 1993; Grossman *et al*., 1995; Pruckler *et al*., 1995; Tamura *et al*., 1995; Milton *et al*., 1996; Mobley, 1996; Bosshardt *et al*., 1997; Kennedy *et al*., 1997; Rosalski *et al*., 1997; West *et al*., 1997; Feldman *et al*., 1998). In our experiments *in vitro*, no inhibition of adherence was shown with antibodies against the recombinant flagellum subunit. However, lack of adherence does not mean there is no role in colonization or virulence. We are planning to investigate further the role of the flagellin cap in adhesion (Arora *et al*., 1998) and that of flagella in colonization and virulence using, for example, animal models.

The gene isolated here could be a potential biomarker to assess intraspecies genetic variation as there appears to be a divergent region in the amino acid sequence of the protein, a feature of surface-located proteins in bacteria (Whittam, 1995; Winstanley *et al*., 1996; Winstanley & Morgan, 1997). An epidemiological survey using the combined PCR-RFLP method and nucleotide sequencing is in progress on a larger number of *C. difficile* isolates.

ACKNOWLEDGEMENTS

This work was supported in part by the FAIR Programme of the European Union (Contract CT95-0433), ACC-SV6 programme (Actions Concertées Coordonnées des Sciences du Vivant) of the Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche of France and a Medical Research Council, Programme Grant (no. G9122850). We thank Professor Delmée (Université Catholique de Louvain, Belgium) for kindly providing the antiflagellin antiserum and *C. difficile* strains 51187 and 56026 and for serogrouping Kohn and B1 strains. We thank Mrs Nick Powell (Institute of Infection and Immunity, Nottingham, UK) for helping us with electron microscopy.

REFERENCES

Arnold, F., Bedouet, L., Batina, G., Robreau, G., Talbot, F., Lecher, L. & Malcoste, R. (1998). Biochemical and immunological analysis of the flagellin of *Clostridium tyrobutyricum* ATCC 25755. *Microbiol Immunol* **42**, 23–31.

Arora, S. K., Ritchings, B. W., Almira, E. C., Lory, S. & Ramphal, R. (1996). Cloning and characterization of *Pseudomonas aeruginosa fliF*, necessary for flagella assembly and bacterial adherence to mucin. *Infect Immun* **64**, 2130–2136.

Arora, S. K., Ritchings, B. W., Almira, E. C., Lory, S. & Ramphal, R. (1998). The *Pseudomonas aeruginosa* flagellar cap protein, FliD, is responsible for mucin adhesion. *Infect Immun* **66**, 1000–1007.

Bartlett, J. G., Chang, T. W., Gurwiyh, M., Gorbach, S. L. & Onderdonk, A. M. (1978). Antibiotic associated pseudomembranous colitis due to toxin producing *clostridia*. *N Engl J Med* **298**, 531–534.

Borriello, S. P., Davies, H. A. & Barclay, F. E. (1988a). Detection of fimbriae amongst strains of *Clostridium difficile*. *FEMS Microbiol Lett* **49**, 65–67.

Borriello, S. P., Welch, A. R., Barclay, F. E. & Davies, M. A. (1988b). Mucosal association by *Clostridium difficile* in the hamster gastrointestinal tract. *J Med Microbiol* **25**, 191–196.

Bosshardt, S. C., Benson, R. F. & Field, B. S. (1997). Flagella are a positive predictor for virulence in *Legionella*. *Microb Pathog* **23**, 107–112.

Brett, P. J., Mah, D. C. & Wood, D. E. (1994). Isolation and characterization of *Pseudomonas pseudomallei* flagellin proteins. *Infect Immun* **62**, 1914–1918.

Chou, P. Y. & Fasman, G. D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol Relat Areas Mol Biol* **47**, 45–148.

Chung, C. T., Niemala, S. L. & Miller, H. R. (1989). One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci USA* **86**, 2172–2175.

Davies, H. A. & Borriello, S. P. (1990). Detection of capsule in strains of *Clostridium difficile*. *Microb Pathog* **9**, 141–146.

Delmée, M., Avesani, V., Delferriere, N. & Burtonboy, G. (1990). Characterization of flagella of *Clostridium difficile* and their role in serogrouping reactions. *J Clin Microbiol* **28**, 2210–2214.

Eaton, K. A., Suerbaum, S., Josenhams, C. & Krakowka, S. (1996). Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes. *Infect Immun* **64**, 2445–2448.

Edman, P. (1950). Preparation of phenylhydantoins from some natural amino acids. *Acta Chem Scand* **4**, 277–282.

Eveillard, M., Fourel, V., Barc, M. C., Kerneis, S., Coconnier, M. H., Karjalainen, T., Bourlioux, P. & Servin, A. (1993). Identification and characterization of adhesive factors of *Clostridium difficile* involved in adhesion to human colonic enterocyte-like Caco-2 and mucus-secreting HT29 cells in culture. *Mol Microbiol* **7**, 371–381.

Fedorov, O. V. & Efimov, A. V. (1990). Flagellin as an object for supramolecular engineering. *Protein Eng* **3**, 411–413.

Feldman, M., Bryan, R., Rajan, S., Scheffler, L., Tang, B. S. L. & Prince, A. (1998). Role of flagella in pathogenesis of *Pseudomonas aeruginosa* pulmonary infection. *Infect Immun* **66**, 43–51.

Gardel, C. L. & Mekalanos, J. J. (1996). Alterations in *Vibrio cholerae* motility phenotypes correlate with changes in virulence factor expression. *Infect Immun* **64**, 2246–2255.

Ge, Y., Li, C., Slaughter, C. A. & Charon, N. W. (1988). Structure and expression of the FlaA periplasmic flagellar protein of *Borrelia burgdorferi*. *J Bacteriol* **180**, 2418–2425.

George, W. L. (1984). Antimicrobial agent associated colitis and diarrhea: historical background and clinical aspects. *Rev Infect Dis* **6**, 208–213.

Gilman, M. Z. & Chamberlin, M. J. (1983). Development and genetic regulation of the *Bacillus subtilis* genes transcribed by σ28 RNA polymerase. *Cell* **35**, 285–293.

Grant, C. C., Konkel, M. E., Cieplak, W. J. & Tompkins, L. S. (1993). Role of flagella in adherence, internalization and translocation of *Campylobacter jejuni* in nonpolarized and polarized epithelial cell cultures. *Infect Immun* **61**, 1764–1771.

Grossman, D. A., Witham, N. D., Burr, D., Lesmana, M., Rubin, F. A., Schoolnik, G. K. & Parsonnet, J. (1995). Flagellar serotypes of *Salmonella typhi* in Indonesia: relationship among motility, invasiveness, and clinical illness. *J Infect Dis* **171**, 212–216.

Guerry, P. (1997). Nonlipopolysaccharide surface antigens of *Campylobacter* species. *J Infect Dis* **176** (suppl. 2), S122–S124.

Heinzerling, H. F., Olivares, M. & Burne, R. A. (1997). Genetic and transcriptional analysis of *flgB* flagellar operon constituents in the oral spirochete *Treponema denticola* and their heterologous expression in enteric bacteria. *Infect Immun* **65**, 2041–2051.

Hopp, T. P. & Woods, K. R. (1981). Prediction of protein antigenic determinants from aminoacid sequences. *Proc Natl Acad Sci USA* **78**, 3824–3828.

Joys, T. M. (1988). The flagellar filament protein. *Can J Microbiol* **34**, 452–458.

Karjalainen, T., Barc, M. C., Collignon, A., Trollé, S., Boureau, H., Cotte-Laffite, J. & Bourlioux, P. (1994). Cloning of a genetic determinant from *Clostridium difficile* involved in adherence to tissue culture cells and mucus. *Infect Immun* **62**, 4347–4355.

Kennedy, M. J., Rosey, E. L. & Yancey, R. J. J. (1997). Characterization of *flaA*- and *flaB*- mutants of *Serpulina hyodysenteriae*: both flagellin subunits, FlaA and FlaB, are necessary for full motility and intestinal colonization. *FEMS Microbiol Lett* **153**, 119–128.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Larson, H. E., Honour, P., Price, A. B. & Borriello, S. P. (1978). *Clostridium difficile* and aetiology of pseudomembranous colitis. *Lancet* **1**, 1063–1066.

Liu, S. L., Ezaki, T., Miura, H., Matsui, K. & Yabuuchi, E. (1988). Intact motility as a *Salmonella typhi* invasion-related factor. *Infect Immun* **56**, 1967–1973.

Lyerly, D. M. H., Krivan, H. C. & Wilkins, T. D. (1988). *Clostridium difficile*: its disease and toxins. *Clin Microbiol Rev* **1**, 1–12.

McGee, K., Horstedt, P. & Milton, D. L. (1996). Identification and characterization of additional flagellin genes from *Vibrio anguillarum*. *J Bacteriol* **178**, 1310–1319.

Milton, D. L., Otoole, R., Horstedt, P. & Wolfwatz, P. (1996). Flagellin A is essential for the virulence of *Vibrio anguillarum*. *J Bacteriol* **178**, 1310–1319.

Mimori-Kiyosue, Y., Vonderviszt, F. & Namba, K. (1997). Locations of terminal segments of flagellin in the filament structure and their roles in polymerization and polymorphism. *J Mol Biol* **270**, 222–237.

Mobley, H. L. (1996). Defining *Helicobacter pylori* as a pathogen: strain heterogeneity and virulence. *Am J Med* **100**(5A), 2S–9S.

Mobley, H. L. T., Belas, R., Lockatell, V., Chippendale, G., Trifillis, A. L., Johnson, D. E. & Warren, J. W. (1996). Construction of a flagellum-negative mutant of *Proteus mirabilis*: effect on internalization by human renal epithelial cells and virulence in a mouse model of ascending urinary tract infection. *Infect Immun* **64**, 5332–5340.

Morooka, T., Umeda, A. & Amado, K. (1985). Motility as an intestinal colonization factor for *Campylobacter jejuni*. *J Gen Microbiol* **131**, 1973–1980.

Perelle, S., Gibert, M., Bourlioux, P., Corthier, G. & Popoff, M. (1997). Production of a complete binary toxin (actin-specific ADP-ribosyltransferase) by *Clostridium difficile* CD196. *Infect Immun* **65**, 1402–1407.

Poilane, I., Karjalainen, T., Barc, M. C., Bourlioux, P. & Collignon, A. (1998). Protease activity of *Clostridium difficile* strains. *Can J Microbiol* **44**, 157–161.

Pruckler, J. M., Benson, R. F., Moyenuddin, M., Martin, W. T. & Fields, B. S. (1995). Association of flagellum expression and intracellular growth of *Legionella pneumophila*. *Infect Immun* **63**, 4928–4932.

Richardson, K. (1991). Roles of motility and flagellar structure in pathogenicity of *Vibrio cholerae*: analysis of motility mutants in three animal models. *Infect Immun* **59**, 2727–2736.

Ritchings, B. W., Almira, E. C., Lory, S. & Ramphal, R. (1995). Cloning and phenotypic characterization of *fleS* and *fleR*, new response regulators of *Pseudomonas aeruginosa* which regulate motility and adhesion to mucin. *Infect Immun* **63**, 4868–4876.

Rosalski, A., Sidorczyk, Z. & Kotelko, K. (1997). Potential virulence factors of *Proteus* bacilli. *Microbiol Mol Biol Rev* **6**, 65–89.

Sakamoto, Y., Sutherland, K. J., Tamaoka, J., Kobayashi, T., Kudo, T. & Horikoshi, K. (1992). Analysis of the flagellin (*hag*) gene of alkalophilic *Bacillus* sp. C-125. *J Gen Microbiol* **138**, 2159–2166.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Scherer, D. C., DeBurron-Connors, C. I. & Minnick, M. F. (1993). Characterization of *Bartonella bacilliformis* flagella and effect of antiflagellin antibodies on invasion of human erythrocytes. *Infect Immun* **61**, 4962–4971.

Seddon, S. V. & Borriello, S. P. (1992). Proteolytic activity of *Clostridium difficile*. *J Med Microbiol* **36**, 307–311.

Sneath, P. H. A. (1986). Endospore-forming Gram-positive rods and cocci. In *Bergey's Manual of Systematic Bacteriology*, pp. 1104–1207. Edited by P. H. A. Sneath, N. S. Mair, M. E. Sharpe & J. G. Holt. Baltimore: Williams & Wilkins.

Szymanski, C. M., King, M., Haardt, M. & Armstrong, G. T. (1995). *Campylobacter jejuni* motility and invasion of Caco-2 cells. *Infect Immun* **63**, 4295–4300.

Tamura, Y., Kijima-Tanaka, M., Aoki, A., Ogikubo, Y. & Takahashi, T. (1995). Reversible expression of motility and flagella in *Clostridium chauvoei* and their relationship to virulence. *Microbiology* **141**, 605–610.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **24**, 4876–4882.

West, N. P., Fitter, J. T., Jakubzik, U., Rohde, M., Guzman, C. A. & Walker, M. J. (1997). Non-motile mini-transposon mutants of *Bordetella bronchiseptica* exhibit altered abilities to invade and survive in eukaryotic cells. *FEMS Microbiol Lett* **146**, 263–269.

Whittam, T. S. (1995). Genetic population structure and pathogenicity in enteric bacteria. In *Population Genetics of Bacteria* (Society for General Microbiology Symposium 52), pp. 217–245. Edited by S. Baumberg, J. P. W. Young, E. M. H. Wellington & J. R. Saunders. Cambridge: Cambridge University Press.

Wilson, D. R. & Beveridge, T. J. (1993). Bacterial flagellar filaments and their component flagellins. *Can J Microbiol* **39**, 451–472.

Winstanley, C. & Morgan, J. A. (1997). The bacterial flagellin gene as a biomarker for detection, population genetics and epidemiological analysis. *Microbiology* **143**, 3071–3084.

Winstanley, C., Coulson, M., Wepner, B., Morgan, J. A. & Hart, C. (1996). Flagellin gene and protein variation amongst clinical isolates of *Pseudomonas aeruginosa*. *Microbiology* **142**, 2145–2151.

Zhang, M. Y., Lovgren, A., Low, M. G. & Landen, R. (1993). Characterization of an avirulent pleiotropic mutant of the insect pathogen *Bacillus thuringiensis*: reduced expression of flagellin and phospholipases. *Infect Immun* **61**, 4947–4954.

Received 22 September 1999; revised 20 December 1999; accepted 7 January 2000.