Deletion of a previously uncharacterized flagellar-hook-length control gene *fliK* modulates the σ^{54} -dependent regulon in *Campylobacter jejuni*

Nahid Kamal,^{1†} Nick Dorrell,² Aparna Jagannathan,^{1‡} Susan M. Turner,¹ Chrystala Constantinidou,¹ David J. Studholme,³ Gemma Marsden,^{4§} Jason Hinds,^{4§} Ken G. Laing,⁴ Brendan W. Wren² and Charles W. Penn¹

Correspondence

Charles W. Penn
c.w.penn@bham.ac.uk

¹School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

²Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

³Sainsbury Laboratory, John Innes Centre, Norwich NR4 7UH, UK

⁴Department of Medical Microbiology, St George's Hospital Medical School, Cranmer Terrace, Tooting, London SW17 0RE, UK

A previously unannotated, putative *fliK* gene was identified in the *Campylobacter jejuni* genome based on sequence analysis; deletion mutants in this gene had a 'polyhook' phenotype characteristic of *fliK* mutants in other genera. The mutants greatly overexpressed the σ^{54} -dependent flagellar hook protein FlgE, to form unusual filamentous structures resembling straight flagella in addition to polyhooks. The genome sequence reveals only one gene predicted to encode an orthologue of the NtrC-family activator required for σ^{54} -dependent transcription. Hence, all σ^{54} -dependent genes in the genome would be overexpressed in the *fliK* mutant together with *flgE*. Microarray analysis of genome-wide transcription in the mutant showed increased transcription of a subset of genes, often downstream of σ^{54} -dependent promoters identified by a quality-predictive algorithm applied to the whole genome. Assessment of genome-wide transcription in deletion mutants in *rpoN*, encoding σ^{54} , and in the σ^{54} -activator gene *flgR*, showed reciprocally reduced transcription of genes that were overexpressed in the *fliK* mutant. The *fliA* (σ^{28})-dependent regulon was also analysed. Together the data clearly define the roles of the alternative sigma factors RpoN and FlhA in flagellar biogenesis in *C. jejuni*, and identify additional putative members of their respective regulons.

Received 22 February 2007

Revised 23 May 2007

Accepted 27 May 2007

INTRODUCTION

Campylobacter jejuni, a spiral microaerophilic Gram-negative bacterium, is a predominant cause of acute

[†]Present address: Department of Haematological Medicine, King's College London, Rayne Institute, 123 Coldharbour Lane, London SE5 9NU, UK.

[‡]Present address: BBSRC Institute for Animal Diseases, Compton, Newbury RG20 7NN, UK.

[§]Present address: Department of Genetics, University of Leicester, Leicester LE1 7RH, UK.

Abbreviation: qRT-PCR, quantitative real-time PCR.

Supplementary tables listing the mutant genes up- or down-regulated by more than twofold compared with the wild-type and potential σ^{54} -dependent promoters in *C. jejuni* are available with the online version of this paper.

The array design is available in BμG@Sbase (accession no. A-BUGS-8; <http://bugs.sgul.ac.uk/A-BUGS-8>) and also ArrayExpress (accession no. A-BUGS-8).

bacterial enteritis worldwide (Konkel *et al.*, 2001). Putative virulence factors include the capsule (Karlyshev *et al.*, 2001), toxins (Wassenaar, 1997), surface proteins (Penn, 2001) and bipolar flagella (Wassenaar *et al.*, 1993b). The genome sequence of *C. jejuni* NCTC 11168 (Parkhill *et al.*, 2000) includes at least 42 genes predicted to be associated with flagellar expression, at more than 20 loci (Jagannathan & Penn, 2005). However, the actions of regulatory genes and proteins that control flagellar biogenesis remain incompletely defined. Reports that an intact flagellum is required for secretion of potential virulence factors (Konkel *et al.*, 2004; Song *et al.*, 2004) underscore the need to understand better the regulation of gene expression in the flagellar export pathway.

In *Salmonella enterica* serovar Typhimurium and *Escherichia coli*, a transcriptional hierarchy of gene expression is coordinated temporally with the assembly of the flagellar structure (Chilcott & Hughes, 2000; Macnab, 1996, 2003). The expression of flagellar genes is regulated in response to

environmental signals, via a 'master' operon *flhDC*. The gene expression sequence features a checkpoint, whereby expression of σ^{28} -dependent class III or 'late' filament and associated genes is contingent on release of σ^{28} from sequestration by its cognate anti-sigma factor FlgM. Release of σ^{28} follows completion of the hook structure, when a switch in export specificity, from hook-associated to filament-associated proteins, allows export of FlgM (Chilcott & Hughes, 2000). The switch is controlled by a regulatory protein, FliK, first identified by the formation of polyhooks in *fliK* mutants (Patterson-Delafield *et al.*, 1973; Silverman & Simon, 1972). Genetic studies have shown that FliK cooperates with export gateway protein FlhB (Hirano *et al.*, 1994). The exact mechanism whereby FliK senses hook length remains unclear (Ferris & Minamino, 2006; Hirano *et al.*, 2005; Keener, 2005; Minamino & Pugsley, 2005) and may depend not only on molecular interactions of proteins of the export pathway but also on temporal or kinetic aspects of hook growth (Minamino *et al.*, 2004).

Polar flagellate bacteria deviate substantially from the *E. coli*/*Salmonella* paradigm. Orthologues of the master switch genes *flhCD* are absent. Sigma factor σ^{54} is involved in flagellar expression in most polar flagellate species. This enables the inclusion of additional checkpoints in flagellar assembly, mediated by transcription activators of the NtrC class and best characterized in *Vibrio cholerae* (Prouty *et al.*, 2001), *Pseudomonas aeruginosa* (Dasgupta *et al.*, 2003; Jyot *et al.*, 2002) and *Caulobacter crescentus* (Mohr *et al.*, 1998).

In *Campylobacter* species, analysis of the promoter regions of the dual flagellin genes has identified a canonical σ^{28} -dependent promoter associated with the major flagellin gene *flaA*, and a canonical σ^{54} -dependent promoter for the minor flagellin gene *flaB*, in both *C. jejuni* (Nuijten *et al.*, 1990) and *Campylobacter coli* (Guerry *et al.*, 1991). We and others have since confirmed the involvement of σ^{54} in flagellar biogenesis, and its regulation by an NtrC family regulator FlgR (Carrillo *et al.*, 2004; Hendrixson & DiRita, 2003; Jagannathan *et al.*, 2001; Wosten *et al.*, 2004). However, several flagellar genes known from the enteric model, including the regulatory gene *fliK* that encodes a hook-length control and export specificity switch, have not been identified from the genome sequence (Parkhill *et al.*, 2000). Hendrixson & DiRita (2003) have explored the σ^{54} -dependent regulon among known flagellar genes of *C. jejuni* using an elegant reporter gene system, and have suggested that the known functional flagellar hook protein gene *Cj1729c*, annotated *flgE2*, be renamed *flgE*, and its paralogue of unclear functionality, *Cj0043*, be named *flgE2*: we adopt this suggestion. Hendrixson & DiRita (2003) also confirmed the predicted identity of *Cj0793* (named *flgS*) as the gene encoding the cognate sensor for FlgR (Jagannathan *et al.*, 2001), forming a two-component sensor-regulator system controlling σ^{54} . Wosten *et al.* (2004) have confirmed that a phospho-transfer relay mediates this two-component regulation. Using microarrays and proteomic analyses to investigate variants of

strain NCTC 11168, Carrillo *et al.* (2004) have demonstrated the importance of genes with predicted promoters dependent on σ^{54} and σ^{28} in motility-associated attenuation of a poorly motile stock strain, identified middle and late genes involved in the flagellar gene expression cascade, and demonstrated key roles for *flhA* and *fliA*.

Questions remain, however, in relation to the operation of the *C. jejuni* flagellar regulon. The functions encoded by key genes such as *fliK* whose existence remains unconfirmed in the *C. jejuni* genome are undefined, and the boundaries of the regulons controlled by the flagella-associated alternative sigma factors σ^{28} and σ^{54} are still unclear. We now present new experimental evidence for the identification and role of a novel *fliK* (*Cj0041*) gene in *C. jejuni*. Knockout mutants in this gene exhibit a 'polyhook' phenotype and greatly overexpress the hook protein FlgE, a putative σ^{54} -dependent gene product. We have used the perturbation of σ^{54} -dependent gene expression in this mutant in microarray experiments, whereby overexpression of these genes in the *fliK* knockout is confirmed by their reciprocal down-regulation in σ^{54} and *flgR* knockouts. We also present data derived from parallel studies of a σ^{28} knockout. These data firmly define the scope of the σ^{54} -dependent regulon, and we correlate σ^{54} -dependent gene expression with score-based promoter predictions.

METHODS

Strains and growth conditions. The *C. jejuni* strains used in this study were NCTC 11168, obtained from A. Karlyshev, London School of Hygiene and Tropical Medicine, to represent the genome-sequenced stock, and three knockout mutants *flgR*, *rpoN* and *fliA* (Jagannathan *et al.*, 2001). Knockouts in *Cj0041* were created in both wild-type NCTC 11168 and NCTC 11828 (also known as strain 81116). Bacteria were grown on blood agar plates, consisting of Columbia Agar Base (Oxoid) supplemented with 7% (v/v) defibrinated horse blood (TCS Microbiology), or on Mueller-Hinton (MH) agar (Oxoid), in a microaerobic atmosphere at 37 °C. For total RNA isolation, bacteria from a 24 h plate culture were resuspended in 1 ml MH broth (Oxoid) and inoculated into 10 ml MH broth to OD₆₀₀ ~0.1. Cultures were grown for 18 h with shaking at 75 r.p.m.

Bioinformatic analysis. The complete genome sequence of *C. jejuni* NCTC 11168 was obtained from the NCBI FTP server and scanned against a scoring matrix (Frech *et al.*, 1993) derived from a compilation of known σ^{54} -binding sites (Barrios *et al.*, 1999) using the promscan.pl Perl script (<http://www.promscan.uklinux.net/software.html>). Each window of 16 bases was assigned a score using the matrix. This score, known as the Kullback-Leibler distance, reflects the theoretical binding energy of the DNA-protein interaction and is calculated using the formula:

$$I_{\text{seq}}(i) = \sum_b f_{b,i} \log_2 \frac{f_{b,i}}{p_b}$$

where i is the position within the site, p_b is the frequency of that base in the genome, and $f_{b,i}$ is the observed frequency of each base at that position (from the weight matrix). Values for p_b were calculated from the mol% G+C content of the genome sequence. To facilitate comparison between predicted binding sites in different genomes,

scores were normalized. The score was divided by the maximum possible score and multiplied by 100, so that 100 was the highest possible score for a sequence window. Sites that had a score greater or equal to a threshold score were considered as matches.

In choosing a threshold score, there is a trade-off between specificity and sensitivity. All currently known σ^{54} -dependent promoters occur in intergenic regions. Therefore, if our threshold is sufficiently stringent, we would expect match sites to occur predominantly in intergenic regions. Furthermore, of those matches falling within intergenic regions, we would expect the majority to be in the same orientation as the adjacent downstream gene. Using a very stringent threshold score of 85, a disproportionately high number of sites (10 out of 37) were in the correct orientation and situated in intergenic regions. However, using this threshold would mean excluding several matches that were potentially biologically significant (e.g. those upstream of flagellar genes *flgH* and *flgG2*). Using a more relaxed threshold value of 80, the overrepresentation was not so apparent (20 out of 176 sites were in correct orientation and situated in intergenic regions). We chose the intermediate threshold value of 83, as this struck a reasonable balance between sensitivity and specificity. In the complete sequence of the *C. jejuni* chromosome, 67 sites score 83 or more, of which 16 are intergenic and in the correct orientation (see Table 1, which also includes one promoter internal to a coding region, within *Cj0063c*).

Knockout mutagenesis. Knockouts in *Cj0041* were created by inverse PCR and allelic replacement (Wren *et al.*, 1994), with kanamycin-resistance cassette insertions in both orientations designed to replace the bulk of the wild-type coding sequence, in both strain NCTC 11168 and strain NCTC 11828 (81116). A sequence extending approximately 300 bp upstream and downstream of the *Cj0041* coding region was amplified from NCTC 11168 genomic DNA with primers 5'-CCGGAATTCGTGATCTAGCTGAAAGAAAAAATG-3' (forward) and 5'-CCGGAATTCGTGCTACTAAGCTGTGAAGTTTGAG-3' (reverse), incorporating CCG clamps and *EcoRI* restriction sites at their 5' ends. The amplified fragment was cloned into pBluescript (SK+) which was then subjected to inverse PCR using primers 5'-GAGGCATGCGCGGAGCTAAATTTGACATCA-3' and 5'-GAGGCATGCGTCAATCTAGAGCTTGTTTTAGCG-3' incorporating GAG clamps and *SphI* restriction sites at their 5' ends. The amplification product was digested with *SphI*, self-ligated to circularize it, and recovered by transformation into *E. coli* DH5a. This resulted in deletion of the bulk of the coding sequence, leaving the first base of the coding sequence at the 5' end and 9 bp at the 3' end. The kanamycin-resistance cassette from pJMK30 (Wassenaar *et al.*, 1993a) was inserted into the deletion plasmid using *SphI* and transformed into *E. coli* DH5a, and transformants with insertions in forward and reverse orientations relative to the deleted coding sequence were identified by digestion of plasmid DNA with *HindIII*. The marked plasmids were transformed into both NCTC 11168 and strain 81116 (NCTC 11828), and kanamycin-resistant transformants were selected and checked by PCR for correct replacement of the wild-type allele.

RNA extraction for microarray analysis. Three independent RNA extractions were performed from wild-type NCTC 11168 and each mutant (*fliK*, *rpoN*, *fliA* and *flgR*). Briefly, 4 ml RNeasy Protect Bacteria Reagent (Qiagen) was added directly to 2 ml bacterial culture, vortexed for 5 s, incubated at room temperature for 5 min then centrifuged at 2400 g for 5 min at 4 °C. The pellet was resuspended in 200 µl TE buffer containing 1 mg lysozyme ml⁻¹ and incubated at room temperature for 10 min, vortexed every 2 min. The manufacturer's protocol was then followed, including 'on-column' DNase treatment; 50 µl RNase-free water was added to the membrane and the spin column was immediately centrifuged at 9000 g for 1 min. The concentration and purity of the total RNA was analysed using a

GeneQuant DNA/RNA calculator (GE Healthcare) and the RNA 6000 Nano LabChip system (Agilent).

Labelling of total RNA. Total RNA was used as a template for direct incorporation of fluorescent analogues (Cy3 and Cy5 dCTP) into randomly primed reverse-transcribed cDNA. Briefly, 10 µg total RNA and 3 µg random hexamers (Invitrogen) in a reaction volume of 11.5 µl were denatured at 95 °C for 5 min, snap-cooled on ice, then 5 µl First Strand Buffer (Invitrogen), 2.5 µl DTT (100 mM), 1.7 µl Cy-dye-labelled dCTP (GE Healthcare), 2.3 µl dNTP solution (5 mM each of dATP, dGTP and dTTP, plus 2 mM dCTP) and 500 U SuperScript II reverse transcriptase (Invitrogen) were added. Mutant total RNA was labelled with Cy5-dCTP and the control wild-type strain total RNA with Cy3-dCTP. The labelling reactions were incubated at 25 °C for 10 min followed by 42 °C for 90 min.

Microarray hybridizations. The original annotation of the NCTC 11168 genome contains 1654 annotated *C. jejuni* ORFs (http://www.sanger.ac.uk/Projects/C_jejuni/). Primer pairs for each ORF were designed with Primer 3 software and selected by BLAST analysis to have minimal cross-homology with all other ORFs (Hinds *et al.*, 2002a). A PCR product representing each ORF was amplified from *C. jejuni* NCTC 11168 chromosomal DNA and spotted onto CMT-GAPS II-coated glass slides (Corning) using a *MicroGrid II* microgridding robot (BioRobotics). All procedures used, including post-processing of deposited arrays, were as described previously (Hinds *et al.*, 2002b).

Microarray slides were incubated in a pre-hybridization buffer [3.5 × saline sodium citrate (SSC) buffer, 0.1 % SDS, 10 mg BSA ml⁻¹] at 65 °C for 20 min, then washed in distilled water for 1 min, followed by a 1 min wash in 2-propanol. Labelled cDNA from the control and test RNA samples was pooled and purified using the Qiagen MinElute PCR purification kit, using a two-step wash stage with 500 then 250 µl volumes of PE buffer and eluting the labelled cDNA from the MinElute column with 14 µl H₂O. The columns retained approximately 1 µl, so the final eluted volume was ~13 µl, which was made up to 30 µl with a final concentration of 4 × SSC buffer and 0.3 % SDS. The hybridization mixture was denatured at 95 °C for 2 min and cooled slowly to room temperature. A 22 × 25 mm LifterSlip coverslip (Erie Scientific) was placed over the reporter element area on the microarray, then the hybridization mixture was applied underneath the coverslip. The slide was placed in a waterproof hybridization chamber (CMT Hybridization Chambers; Corning) for hybridization in a 65 °C water bath overnight. After hybridization, slides were washed in 1 × SSC buffer with 0.06 % SDS at 65 °C for 5 min, followed by two washes in 0.06 × SSC buffer, each for 2 min at room temperature.

Microarray hybridizations were performed in triplicate for each matched pair of total RNAs, giving a total of nine hybridizations for the comparison of gene transcription in each of the *fliK*, *rpoN*, *fliA* and *flgR* mutants with that of the wild-type strain.

Data acquisition and analysis. Slides were scanned with an Affymetrix 418 scanner (MWG Biotech) following the manufacturer's guidelines. Fluorescent spot intensities were quantified using ImaGene 5.5 (BioDiscovery) software. For each spot, background fluorescence was subtracted from the average spot fluorescence to produce a channel-specific value. The data were further analysed using GeneSpring 6.1 (Silicon Genetics) software. The geometric mean of the normalized red:green ratio was calculated for each mutant using data from nine array experiments. Low signal (<100 units in the Cy3 channel) spots were excluded. To determine which genes were over- or underexpressed, a *P* value of <0.05 was used for testing whether the normalized expression level in each mutant was significantly different from 1.0 by a two-sided one sample *t* test, using the GeneSpring 6.1 software with correction for multiple testing (Benjamini & Hochberg, 1995).

Analysis of the data allowed identification of gene expression patterns that correlated with function (Brazma *et al.*, 2000). Microarray-derived data may be affected by variables in the experimental protocol such as labelling efficiency, post-labelling purification, hybridization temperature and stringency of washes. Therefore three biological and three technical replicates were performed to produce nine datasets, and the data analysed searchingly to establish statistical significance. Fully annotated microarray data have been deposited in BμG@Sbase (accession no. E-BUGS-50; <http://bugs.sgul.ac.uk/E-BUGS-50>) and also ArrayExpress (accession no. E-BUGS-50) in accordance with MIAME requirements (Brazma *et al.*, 2001). Gene lists are available as supplementary information (see Supplementary Table S1 available with the online version of this paper), for genes showing twofold or greater changes in the mutants analysed after filtering, as described above.

RNA extraction and cDNA synthesis for real-time PCR. Cultures were grown to mid-exponential phase, according to the protocol described for the microarray experiments, and RNA was extracted similarly with an additional 1 h off-column DNase digestion step at 37 °C and an additional RNA clean-up step to minimize genomic DNA contamination. Total RNA was converted to cDNA using random hexanucleotide primers and Superscript II reverse transcriptase (Invitrogen), according to the manufacturer's protocol.

Relative quantification of gene expression using SYBR Green Real-Time PCR. Specific primers for real-time PCR (Table 2) were designed using Primer Express software (Applied Biosystems). Aliquots of cDNA (3 µl) were used as templates for real-time PCR. The positive control was the supernatant extracted after centrifugation of whole cells of NCTC 11168 boiled for 10 min. An optical 96-well microtitre plate (Applied Biosystems) was used with a 25 µl reaction volume consisting of 2 × Rainbow SYBR Green Master Mix (Bioline), 50 nM gene-specific primers, and the template. An ABI Prism 7000 Sequence Detector (Applied Biosystems) was programmed for an initial step of 10 min at 95 °C, followed by 40 cycles of 15 s at 95 °C and 1 min at 60 °C. SYBR Green technology detects double-stranded DNA so melting curves were used to monitor specificity. Gene expression levels were determined relative to a reference group using the $2^{-\Delta\Delta CT}$ method and *gyrA* as the internal control gene (Livak, 1997; Livak & Schmittgen, 2001). Signal intensity from genomic DNA contamination was monitored using RT-negative cRNA reactions as the template for real-time PCR.

Electron microscopy. Cultures were grown for 24 h on Mueller–Hinton agar and suspended in 50 µl 1.5% (w/v) potassium phosphotungstate in water (pH 7.0) or in 1% (w/v) ammonium molybdate with 70 µg ml⁻¹ bacitracin (Sigma) to just visible turbidity. Samples were immediately applied to Formvar-coated copper grids (200 mesh), and excess liquid was removed by blotting with clean filter paper before drying in air and examination in a JEOL 1200 EX transmission electron microscope at an accelerating voltage of 80 kV.

Flagellar preparation and protein analysis. Mueller–Hinton agar cultures (24 h) were suspended to OD₆₀₀ 3 and subjected to shearing by high speed blender (IKA Ultra-Turrax) for 2 min at top speed. Flagella were purified by removal of cells and gross debris by centrifugation at 3000 g for 10 min at 4 °C, treatment with 1% Triton X-100 to solubilize membrane material, and ultracentrifugation at 100 000 g for 1 h to recover flagellar filaments. Preparations were analysed by SDS-PAGE, and one gel was stained with colloidal Coomassie blue (Bio-Rad) while the other was Western-blotted with detection by a flagellin-specific mAb N2D2 raised with flagellar antigen prepared from *C. jejuni* strain 81116.

For mass spectrometric identification of the 90 kDa protein, the band was excised and cut into ~0.5 mm cubes for processing. The trypsin

digestion protocol included reduction (10 mM DTT) and alkylation (55 mM iodoacetamide) steps prior to digestion (6–10 ng trypsin µl⁻¹). Peptides were extracted with 1% (v/v) formic acid and 2% (v/v) acetonitrile and dried with matrix for MALDI-TOF analysis (Bruker Daltonics Biflex IV). Peptide masses were obtained and protein identification performed using the Mascot Peptide Mass Fingerprint program at <http://www.matrixscience.com/> with a mass cut-off of 100 p.p.m.

RESULTS

Identification and characterization of a candidate *fliK* gene

A transposon mutant in *Cj0041* was non-motile (Golden & Acheson, 2002), and a strong predicted σ^{54} -dependent canonical promoter element is associated with its closely linked upstream ORF *Cj0040* (see section on promoter prediction below). Hence, *Cj0041* was identified as a candidate σ^{54} -dependent flagellar gene. *Cj0041* showed significant sequence similarity to *Helicobacter pylori* (HP0906), which was previously shown to encode a FliK homologue of which only a small domain near the C terminus showed a marginally conserved amino acid sequence compared with known FliK sequences (Ryan *et al.*, 2005). Knockout deletion/insertion mutants in *Cj0041* were created in strains NCTC 11168 and 81116, with kanamycin-resistance cassettes inserted in both orientations. All four mutants showed a polyhook phenotype (Fig. 1a), similar in appearance to that of *fliK* mutants in *S. enterica* serovar Typhimurium (Hirano *et al.*, 1994). The identical phenotypes of mutants in different backgrounds with insertions in opposite orientations, and the presence of a separate predicted σ^{54} -dependent promoter (see Table 1 and below) for the downstream operon *flgD-flgE2*, which is likely to preclude polar effects, indicate that this phenotype is attributable to the deletion of *Cj0041*. In addition, the product of the downstream gene *flgD* (hook-capping protein) would be required for polyhook formation and must therefore be expressed in the mutants, confirming that there is no polar effect on downstream transcription in these mutants. We therefore propose that *Cj0041* encodes a FliK protein.

In addition to the polyhook structures, long, generally straight flagella-like structures were apparent, sometimes including kinks or knots of typical hook-like radius or more complex knots and twisted or braided portions (Fig. 1b–e). Others had a polyhook structure at the base and a straight distal portion, and structures alternating from straight to curved and back to straight were also seen (Fig. 1e). We did not observe transitions in width or fine structure between straight and curved portions of these structures (Fig. 1e), suggesting that they do not represent the 'polyhook-filament' structures reported in pseudo-revertants of *fliK* mutants in *S. enterica* serovar Typhimurium (Hirano *et al.*, 1994). Flagella of normal curvature were absent. In contrast, in the wild-type strain (Fig. 2) normal flagella are shown for comparison, and the

Fig. 1. (a) Mutant in *Cj0041*, strain NCTC 11168 with resistance marker in same orientation as the predicted gene, showing a polyhook flagellum; (b) straight 'flagellar' forms, comprising the majority of filamentous structures seen; (c, d) polyhook-like filaments suggestive of unwinding and entanglement of tightly coiled, elongated hook-like structures; (e) alternate straight, curved hook-like and straight filament tracts; inset shows absence of any change in fine structure or width of filament at transition from straight to curved form. Bars, 100 nm.

Table 1. High-scoring predicted σ^{54} -dependent promoters located in intergenic regions (with the exception of *Cj0062c*) and oriented towards the downstream ORF

Changes in expression in the *fliK* mutant are shown for those genes that showed a twofold or greater increase relative to wild-type.

Gene number, name	Promoter sequence	Promoter score	Fold increase in $\Delta fliK$
<i>Cj0243c</i> , unknown	TTGGCACACCTTTTGCT	94	4.5
<i>Cj0547</i> , <i>flaG</i>	TTGGCACAGTTTTTGCT	94	
<i>Cj1026c</i> , upstream of <i>flgR</i>	TTGGAACGCTTATTGCT	88	
<i>Cj1729c</i> , <i>flgE</i>	TTGGAACAGAACTTGCT	87*	60
<i>Cj0040</i> , upstream of <i>fliK</i>	TTGGAACACTTCTTGCT	87*	11
<i>Cj1462</i> , <i>flgI</i>	TTGGAACACTTTTTGCT	86*	16
<i>Cj0887c</i> , <i>flgL</i> (<i>flaD</i>)	TTGGAACAGTTATTGCT	86*	14
<i>Cj0042</i> , <i>flgD</i>	TTGGAACACTTATTGCT	86*	17
<i>Cj0528c</i> , <i>flgB</i>	TTGGAACAGTATTGCT	85	5.9
<i>Cj1338c</i> , <i>flaB</i>	TTGGAACACTTTTTGCT	85	
<i>Cj0697</i> , <i>flgF</i> (<i>flgG2</i>)	TTGGAACACTCTTGCT	85*	13
<i>Cj0665c</i> , <i>argC</i>	TTGGAACAAAATTGCT	84	
<i>Cj0687c</i> , <i>flgH</i>	GTGGAACACTCTTGCT	84*	19
<i>Cj1029c</i> , <i>mapA</i>	TTGGCTTGAAATTGCT	84	
<i>Cj1075</i> , unknown	TTGGAATGATTATTGCT	84	
<i>Cj1397</i> , unknown	TTGGAACGGCTATAGCA	83	
<i>Cj0062c</i> †, unknown	TTGGAACGGAAAGTGCT	83	2.6
Consensus for <i>C. jejuni</i>	TTGGAACACTTTTTGCT		
Consensus‡	yTGGCACGrnnnTTGCw		

*Genes that were among the 18 genes or clusters that showed eightfold or more upregulation in *fliK* mutant microarray analysis.

†A ninefold signal increase was observed in downstream gene *Cj0059c*, but *Cj0060*–*Cj0062c* did not show such large increases: values were 6.8, 2.7 and 2.6, respectively.

‡Barrios *et al.* (1999).

Table 2. Primers used for qRT-PCR

Gene name	Primer orientation	Sequence
<i>gyrA</i>	Forward	ATGCTCTTTGCAGTAACCAAAAAA
<i>gyrA</i>	Reverse	GGCCGATTTCACGCACTTTA
<i>flgK</i>	Forward	CCATTGAAGATAAGGGCACAAAT
<i>flgK</i>	Reverse	ACTTTCATGTCACCTTGCATCAGTTC
<i>flgE2</i>	Forward	CCCTGTGGATATAGGTCCTATGTATAATC
<i>flgE2</i>	Reverse	AGTTGGTCTTGTAGTTGCATTAGCA
<i>flgD</i>	Forward	GATTATAAGCTTGGTTCATCTTCATCTACT
<i>flgD</i>	Reverse	CAGAAGAATCTCCATCGCTTGA
<i>fliA</i>	Forward	AAGATGAGCTTGTGCTTCTTACATG
<i>fliA</i>	Reverse	TTCAACGCCTATGCTAATAAGATCA
<i>flhF</i>	Forward	TGATACCTCCTGAATTTGCAAGTATT
<i>flhF</i>	Reverse	ATAGTTGCTTTCATAATCGCTTCTAGATG
<i>flgM</i>	Forward	TGGCAAATACCGCATTAAATACA
<i>flgM</i>	Reverse	CTCGCTTTATCATTTTCGCTTTTT
<i>fliI</i>	Forward	TTTGCTCAACTCATGGAAAGA
<i>fliI</i>	Reverse	CACCATCCACTAAACGGTAAAAA

discontinuity in fine structure between hook and filament is apparent at high magnification (Fig. 2b).

Characterization of polyhooks and associated flagella-like structures

Preparations were made of the filamentous structures from cultures of *fliK* (*Cj0041*) mutants by shearing, detergent treatment and differential centrifugation, and analysed by SDS-PAGE, Western blotting and protein identification by mass spectrometry. SDS-PAGE of wild-type cells revealed the expected flagellin band of about 62 kDa, as confirmed by Western blotting with an anti-flagellin mAb. However, this band was substantially decreased in the mutant 'flagellar' preparations (Fig. 3), in which a significant protein band of about 90 kDa was present, especially in strain 81116, but also visible in NCTC 11168 (lanes 6 and 7). Excision of this band and analysis by tryptic peptide

MALDI-TOF mass spectrometry confirmed its identity as FlgE, the product of *Cj1729c*, the functional hook protein in *C. coli* and *C. jejuni* (Kinsella *et al.*, 1997; Luneberg *et al.*, 1998); 27 peptide masses matched those predicted, comprising 58% of the total predicted amino acid sequence. We therefore propose that the filamentous structures seen in the mutant are composed predominantly of FlgE protein.

Microarray analysis of flagellar gene expression in *fliK* mutant cells

Identification of the flagellar biogenesis control protein FliK enabled investigation of its influence at the checkpoint where flagellar gene expression switches from σ^{54} -dependent (rod, hook) to σ^{28} -dependent (filament). We reasoned that deletion of *fliK* had prevented the switch in secretion specificity from hook to filament proteins. Furthermore, based on the results shown in Fig. 3 as well as the presence of extensive polyhooks, there was clearly overexpression of hook protein in the mutants. Thus, failure to pass the hook basal body to filament checkpoint in secretion specificity means that σ^{54} -dependent FlgE expression, normally a transient phase of the flagellar biogenesis sequence, persists at a high level. We postulated that this is likely to be at least in part dependent on a high level of transcription. Only one σ^{54} -activator protein of the NtrC family, essential for σ^{54} -dependent transcription (Buck *et al.*, 2000), is predicted from the genome sequence of *C. jejuni*. Hence, overexpression of σ^{54} -dependent genes must extend throughout the genome in this mutant, providing a tool for definition of the σ^{54} -dependent regulon by microarray analysis.

Transcription in wild-type and *fliK* mutant cells (strain NCTC 11168, cassette insertion in same orientation as *Cj0041*) was analysed. Overall, relative to the wild-type, 352 genes were significantly upregulated by at least twofold (see supplementary tables) and 110 were down-regulated in the *fliK* mutant, indicating a substantial perturbation of global

Fig. 2. Wild-type flagellar morphology of a typical bipolar flagellate cell (a) showing uniformly curved flagella; (b) a detached hook basal body and proximal filament clearly show the discontinuity in diameter and granularity at the hook–filament junction (arrow), in contrast to the continuous morphology seen in Fig. 1(e). Note that in both (a) and (b) the hook can appear straight and aligned with the filament; this may be due either to tension on the filament leading to straightening of the flexible hook structure as the specimen film dries, or to chance alignment of the curvature of the hook in the *z* dimension relative to the *x*–*y* plane of the image. Bars, 200 nm.

Fig. 3. SDS-PAGE (left) and Western blot analysis (right) of flagellar proteins isolated from wild-type and Cj0041 mutant *C. jejuni*. Lanes: 1, wild-type strain 11168; 2, Cj0041 mutant in strain 11168 with kan^r cassette in same orientation; 3, Cj0041 mutant in strain 11168 with kan^r cassette in opposite orientation; 4, protein marker, values for left-hand panel shown at left; 5, wild-type strain 81116; 6, Cj0041 mutant in strain 81116 with kan^r cassette in same orientation; 7, Cj0041 mutant in strain 81116 with kan^r cassette in opposite orientation. Dashed arrow, flagellin band identified by mAb to strain 81116 flagellin (weakly cross-reactive with strain 11168 flagellin); solid arrow, 90 kDa band overexpressed in Cj0041 mutants.

gene expression. Data relating to flagellar genes are summarized in Table 3 column 3. Known and putative flagellar genes are identified as follows: Cj0041, *fliK*; Cj0063c, *flhG* [nomenclature corresponding to that of its orthologue in *Vibrio parahaemolyticus* (Kim & McCarter, 2000)]; Cj0793, *flgS* (Hendrixson & DiRita, 2003); Cj0887c, *flgL* (Guerrey *et al.*, 2000); Cj1024, *flgR* (Jagannathan *et al.*, 2001); Cj1464, *flgM* (Pallen *et al.*, 2005); Cj1465, *flgN* (Pallen *et al.*, 2005); and Cj0697 (annotated as *flgG2*), *flgF* (Loong Chan *et al.*, 1998). Sequence analysis suggests that Cj1463, located downstream of the predicted *flgI*, resembles the N-terminal domain of known *flgJ* sequences (Pallen *et al.*, 2005). The translated sequence lacks the amidase 4 C-terminal domain (Pfam accession no. PF01832) believed to be involved in peptidoglycan hydrolysis by the FlgJ protein, but leaves intact the postulated rod-capping domain (Hirano *et al.*, 2001), and we therefore propose that Cj1463 corresponds to *flgJ*.

The data in Table 3 show that rod and hook component genes were upregulated in the *fliK* mutant, as were several regulatory genes and related gene clusters. Of 26 genes, representative of 18 contiguous clusters, that showed more than an eightfold increase in transcription in the *fliK* mutant, 21 were flagellar genes or were related to flagellar function (Table 4). The most highly expressed, with a 60-fold increase over wild-type, was Cj1729c, *flgE*. Other upregulated genes included Cj0793, encoding the two-component sensor histidine kinase FlgS (Hendrixson & DiRita, 2003); the gene cluster Cj0062c-*fliA*-*fliM*-*fliY*, which includes *fliA* encoding σ^{28} ; and the cluster *flgI*-*flgJ*-*flgM*-*flgN*-*flgL*, although in this cluster *flgM* (Cj1464) showed less than a twofold change. The remaining flagellar structural genes were unaffected, as was *rpoN* encoding σ^{54} . Cj1024 encoding FlgR was only weakly expressed, and no major change was seen in the signal. Cj0041 appeared to be down-regulated. However, this is likely to be an artefact of the deletion of the coding sequence, since upstream gene Cj0040, which appears to be co-transcribed, was upregulated.

Confirmation of microarray data by quantitative real-time PCR (qRT-PCR)

Each assay was done in triplicate on selected transcripts in the *fliK* mutant compared with the wild-type strain, and repeated on separate occasions with three different biological samples. Representative data are shown in Fig. 4, and confirm that large changes (greater than threefold) revealed by microarray analysis are matched in qRT-PCR assays. This was the case for *flgK*, *flgE*, *flgD* and *flgI*. For the early genes *fliI* and *flhF*, no statistically robust change was evident, again verifying microarray data. For *flgM*, no significant change was seen, despite changes in adjacent genes *flgI* and *flgK*, reflecting the minimal change seen in the microarray data and demonstrating that a carefully replicated and analysed microarray experiment is capable of greater discrimination than real-time PCR when values are below the twofold cut-off.

Prediction of promoters for alternative sigma factors involved in flagellar gene expression

The microarray and real-time data described above identified a clear σ^{54} -dependent regulon in *C. jejuni*. We therefore searched the genome for σ^{54} -dependent promoter sequences using the PromScan program. High-scoring potential promoter sequences (score of 83 or more) were found in 67 locations (see Supplementary Table S2), of which 15 were in non-coding regions and in the same orientation as an adjacent gene (Table 1). These sequences from *C. jejuni* had the consensus 5'-TTGGAACACT-TTTTGCT-3' sequence, similar to that of σ^{54} -dependent promoter sequences of a wide range of bacteria (Table 1) (Barrios *et al.*, 1999). Strikingly, 11 of the 15 high-scoring, intergenic potential promoters were upstream of genes or gene clusters associated with flagellar motility, and seven of these were associated with 18 genes or gene clusters that were eightfold or higher overexpressed in the *fliK* knockout mutant by microarray analysis (Table 4).

Table 3. Summary of microarray data showing flagellar genes up- (+) or down- (–) regulated in the knockout mutants shown, based on a minimum twofold change and significance at $P > 0.05$

Genes that showed no changes in any of the mutants are not listed. NC, No change observed; Invalid, genes that were knocked out in these mutants did not give a meaningful signal due to deletion of much of the coding sequence and the presence of the resistance marker. [–] or [+], statistically significant changes in gene expression of less than twofold were seen.

Gene name	Gene number	<i>fliK</i> mutant	<i>rpoN</i> mutant	<i>flgR</i> mutant	<i>fliA</i> mutant
<i>flaA</i>	Cj1339c	NC	NC	–	–
<i>flaB</i>	Cj1338c	NC	NC	–	–
<i>flaG</i>	Cj0547	NC	–	–	–
<i>flgB</i>	Cj0530c	+	–	–	NC
<i>flgC</i>	Cj0528c	+	–	–	NC
<i>flgD</i>	Cj0042	+	–	–	+
<i>flgE</i>	Cj1729c	+	–	–	+
<i>flgE2</i>	Cj0043	+	–	–	+
<i>flgF</i> (<i>flgG2</i>)	Cj0697	+	–	–	NC
<i>flgG</i>	Cj0698	+	–	–	NC
<i>flgH</i>	Cj0687c	+	–	–	+
<i>flgI</i>	Cj1462	+	–	–	[+]
<i>flgJ</i>	Cj1463	+	–	–	[+]
<i>flgK</i>	Cj1466	+	–	–	[+]
<i>flgL</i> (<i>flaD</i>)	Cj0886c	+	–	–	NC
<i>flgM</i>	Cj1464	[+]	[–]	–	–
<i>flgN</i>	Cj1465	+	[–]	–	NC
<i>flgR</i>	Cj1024	NC	NC	Invalid	NC
<i>flgS</i>	Cj0793	+	[–]	–	[+]
<i>flhB</i>	Cj0335	NC	NC	nc	+
<i>flhF</i>	Cj0064c	NC	NC	+	NC
<i>flhG</i>	Cj0063c	NC	NC	+	NC
<i>fliA</i>	Cj0061c	+	NC	NC	Invalid
<i>fliD</i>	Cj0548	NC	NC	NC	–
<i>fliE</i>	Cj0526c	+	–	–	NC
<i>fliK</i>	Cj0041	Invalid	–	–	+
<i>fliM</i>	Cj0060c	+	NC	NC	NC
<i>fliS</i>	Cj0549	NC	NC	NC	–
<i>fliY</i>	Cj0059c	+	NC	+	–
<i>rpoN</i>	Cj0670	NC	Invalid	–	NC

Sequences similar to the *E. coli* σ^{28} consensus sequence 5'-TAAAGTTT-N11-GCCGATAA-3' (Ide *et al.*, 1999; Park *et al.*, 2001) were identified upstream of *flgM* (5'-TTAAGTTT-N11-GTCGATAT-3'), the *flaG-fliD-fliS* cluster (5'-TTCATAAA-N11-GTCGATAT-3'), and the *flaA* gene (5'-TAAAATAT-N11-CACGATAT-3'), the latter resembling in position and sequence those identified previously for *C. coli* (Guerry *et al.*, 1991) and *C. jejuni* (Nuijten *et al.*, 1990). The σ^{70} –10 consensus sequence, TATAATT and –35 consensus sequence, 5'-TTTAAGTNTT-3' (Wosten *et al.*, 1998), were identified for eight other flagellar genes: *fliA*, *fliF*, *fliL*, *fliN*, *fliP*, *fliQ*, *flhA* and *flhB*.

Microarray analysis of transcription in previously described flagellar regulatory mutants

Microarray analysis of the knockout mutants described previously (Jagannathan *et al.*, 2001) in σ^{54} , its cognate activator FlgR, and σ^{28} indicated extensive changes (Table 3, columns 4–6) in flagellar gene expression in the regulatory mutants compared to the wild-type strain. Flagellar genes that did not show significant changes in levels of expression in any of the mutants are not listed. The data show that many flagellar genes, mainly rod- and hook-associated, can be categorized as σ^{54} /FlgR-dependent (expression is decreased in one or both of these mutants). Almost all of the genes that were upregulated in the *fliK* mutant were reciprocally down-regulated in *rpoN* and/or *flgR* mutants, thus providing experimentally independent confirmation of the influence of σ^{54} -dependent regulation in determining the expression changes observed in the *fliK* mutant. Genes that were down-regulated in the *fliA* mutant corresponded largely with the predicted presence of σ^{28} -dependent promoters, for *flaA*, *flaG*, *fliDS* and *flgM*.

DISCUSSION

Our observations strongly support the identification of Cj0041 as a FliK homologue, a gene product not previously recognized in *C. jejuni* although clearly homologous with HP0906, shown recently to be a FliK homologue (Ryan *et al.*, 2005). The very divergent sequence of *C. jejuni* FliK has only one small C-terminal region that is conserved among more distantly related organisms. FliK is thought to be a bifunctional protein, involved in both hook-length sensing (primarily by the N-terminal domain) and switching of specificity of the export pathway from hook-associated to filament-associated proteins (mediated by interaction of the C-terminal domain with FlhB) (Minamino *et al.*, 2004). The localization of the minimal sequence conservation seen in Cj0041 to the C-terminal domain indicates that interaction with FlhB may be more critical in determining conservation of sequence than N-terminal domain function, which appears not to be correlated with sequence. This accords with evidence that at least part of the N-terminal domain is dispensable for hook-length sensing (Hirano *et al.*, 2005).

The analysis of proteins in preparations of the filamentous structures seen in Fig. 2 showed that the predominant protein was FlgE, which, since the majority of the filamentous structures were of the straight form, must therefore be present in both the curved and straight structures seen. There was also a smaller amount of flagellin protein, its identity confirmed by Western blot analysis. The filamentous structures we report occasionally showed the alternation of straight filaments with a curved form with radius approximating that of the flagellar hook, then back to the straight structure, in the absence (Fig. 1e) of any alteration in width or fine structure normally associated with the hook–filament transition. It is known

Table 4. Genes showing more than eightfold transcription increase in the *fliK* knockout

Gene number	Name and function	Fold increase in <i>fliK</i> knockout
Cj1729c	<i>flgE</i>	60
Cj1328	<i>neuC2</i> , flagellar-modification region	34
Cj0025c	Possible glutamate symporter	31
Cj0043	<i>flgE2</i>	22
Cj0687c	<i>flgH</i>	19
Cj0042	<i>flgD</i>	17
Cj1327	<i>neuB2</i> , flagellar-modification region	17
Cj1462	<i>flgI</i>	16
Cj1463	<i>flgJ</i>	14
Cj0886c	<i>flgL</i>	14
Cj0698	<i>flgG</i>	14
Cj1466	<i>flgK</i>	14
Cj0697	<i>flgG2</i>	13
Cj0613	Possible periplasmic phosphate-binding protein	13
Cj1294	Putative aminotransferase, flagellar-modification region*	12
Cj1242	Unknown, function hypothetical	12
Cj0053c	<i>trmU</i> tRNA methylase	11
Cj0040	Unknown, function hypothetical, upstream of <i>fliK</i>	11
Cj1567c	<i>nuoM</i> , probable NADH dehydrogenase subunit	10
Cj1341c	<i>maf7</i> , flagellar-modification region, unknown function	10
Cj1295	Unknown function, flagellar-modification region*	9
Cj0059c	<i>fliY</i>	9
Cj1340c	<i>maf6</i> , flagellar-modification region, unknown function	9
Cj1467	Conserved hypothetical, downstream of <i>flgK</i>	9
Cj1566c	<i>nuoN</i> , probable NADH dehydrogenase subunit	8
Cj1645	<i>Tkt</i> , transketolase	8

*Downstream of Cj1293, known σ^{54} -dependent flagellin-modification gene (Goon *et al.*, 2003).

that the switch in export specificity from hook to filament components in *S. enterica* serovar Typhimurium is irreversible, resulting from cleavage of FlhB upon activation of the export switch (Hirano *et al.*, 2005; Minamino &

Macnab, 2000); this implies that export-specificity reversal to produce alternating hook-like and filament-like structures is not possible. We conclude that there are alternate modes of packing of the FlgE protein into either hooks/

Fig. 4. qRT-PCR assays of the relative change in gene expression of *flgK*, *flgE*, *flgD*, *fliA*, *flhF*, *flil* and *flgM* in the *fliK* mutant compared with wild-type cells. The error bars represent the range determined using the $2^{-\Delta\Delta CT}$ method with $\Delta\Delta CT + s$ and $\Delta\Delta CT - s$, where s is SD of the value for $\Delta\Delta CT$. These results represent one dataset; the results were reproducible using different RNA extracts from two additional independent cultures.

polyhooks or straight filaments. However, in *C. jejuni*, as in many other genera, there is both a second partial *fliK*-like gene (*Cj0849c*) and an upstream gene, *Cj0848c*, proposed to be named *flhX*, and predicted to encode a small protein that is homologous to the C-terminal domain of FlhB (Pallen *et al.*, 2005), and we cannot preclude some degree of dual secretion of FlgE and flagellin that might be mediated by the products of these genes. We cannot confirm at present whether the flagellin protein detected by Western blot analysis is FlaA or FlaB, either by Western blot or from transcription data, due to similarities in their antigenicity and sequence, but the σ^{54} -dependence of *flaB* expression would be compatible with this flagellin being FlaB. We have no evidence of the possible role or location of this flagellin in the structures seen.

The reciprocal datasets we have obtained, in which genes that are overexpressed in the *fliK* mutant are under-expressed in the regulatory mutants in *flgR* and *rpoN*, provide experimentally independent verification of the σ^{54} -dependent regulon, further confirmed for selected genes by qRT-PCR. Our microarray data do not show changes in 'early' genes in the expression cascade (Chilcott & Hughes, 2000) that are not associated with the alternative sigma factors. These genes include those encoding the MS ring and export apparatus. Our data show, however, that rod and hook genes associated with the 'middle' σ^{54} -dependent level of expression, are clearly identified by their substantial upregulation in the *fliK* mutant and clear down-regulation in *rpoN* and *flgR* mutants. Genes in the final 'late' stage of the expression cascade are down-regulated in *fliA* mutants; in some cases, e.g. *flaG*, these are also down-regulated in *rpoN* and *flgR* mutants. This is presumably because they require σ^{28} for transcription, and the activity of σ^{28} is predicted to be dependent on completion of the σ^{54} -dependent hook structure and consequent export of FlgM, thus being indirectly dependent on σ^{54} . These genes are distinguished not because they lack σ^{54} -dependency, but because they are σ^{28} -dependent. Confirmatory evidence for their independence of σ^{54} is provided by their unchanged level of transcription in the *fliK* knockout. The regulator *flgM* is now shown to be significantly dependent on σ^{28} , but also co-transcribed with *flgI* from its σ^{54} -dependent promoter.

The overexpression of rod/hook-associated, σ^{54} -dependent genes in the *fliK* mutant suggests a specific and powerful (but presumably transient in the wild-type) activation of σ^{54} -dependent genes triggered by an unknown signal that presumably relates to completion of the earlier MS ring/export pathway structures. Furthermore, after progression through the checkpoint that controls progression from σ^{54} -dependent to σ^{28} -dependent, there is clearly a specific down-regulation, or a termination of their prior activation, of σ^{54} -dependent genes (the absence of this process in the *fliK* mutant leads to locked-on overexpression of these genes). We postulate that both these changes may result from activation of σ^{54} , leading to transcription of σ^{54} -dependent genes at an early stage in flagellar biogenesis, signalled by an unknown mechanism, upon completion of

the MS ring and export pathway components, a structure which is encoded by genes that are not σ^{54} -dependent. We postulate that upon addition of the rod and hook structures, the signal ceases, leading to termination of σ^{54} -dependent gene activation. However, the aberrant outcome of hook construction in the *fliK* mutant argues that its gene product too must play a part in termination of σ^{54} -dependent gene expression. Thus, in cells that progress flagellar growth to completion of a normal-length hook but no further (e.g. the *fliD* and *fliS* knockout mutants that we have constructed that express hooks without filaments; unpublished data), hook overexpression (polyhook formation) is prevented, presumably by a mechanism dependent on normally functioning FlhK at the completion of the hook. Alternatively, it could be postulated that the key to down-regulation of σ^{54} -dependent genes after hook completion is expression of a σ^{28} -dependent gene product. This might be prevented in *fliK* mutants because in the absence of a switch in export-pathway specificity, export of FlgM and hence activation of σ^{28} could not occur. However, the phenotype of *fliA* deletion mutants is not the expression of a polyhook (Jagannathan *et al.*, 2001); hence, this suggestion appears incorrect. Possibly intracellular accumulation of FlhK following hook completion [there is evidence that it is secreted maximally to the exterior prior to hook completion in *S. enterica* (Minamino *et al.*, 1999)] may have a direct effect on transcription of class III genes, for instance via an anti-activator effect on FlgR or even a direct anti- σ^{54} effect, thus shutting down class III gene transcription, as also suggested for *flgE* expression by Hirano *et al.* (2005).

The observation that transcription of *fliA*, encoding σ^{28} and other genes in the *Cj0062c-fliY* gene cluster, appears from the *fliK* mutant experiment to be σ^{54} -dependent is novel; these genes were earlier predicted to be transcribed from a σ^{70} -dependent promoter (Carrillo *et al.*, 2004; Wosten *et al.*, 2004). Transcription is presumably initiated from the apparent σ^{54} -dependent canonical promoter element located upstream of *Cj0062c* and within the coding sequence of *Cj0063c* (*flhG*) (Table 3). This transcription unit would not include *Cj0063c* (*flhG*) or *Cj0064c* (*flhF*). There is, however, contradictory evidence from the microarray data derived from the *fliA* and *flgR* mutants regarding control of *flhF*, *flhG*, *fliM* and *fliY* expression, and we are currently attempting to clarify the transcription regulation of this region. The expression of transcripts of *fliA* in the *fliK* mutant suggests that σ^{28} protein would be available prior to completion of the hook structure. Hence, either expression of flagellin and other filament-related proteins potentially commences at this time, or FlgM is indeed active in inhibiting σ^{28} activity prior to hook completion, contrary to the reservations of Hendrixson & DiRita (2003).

We note that *flaG*, encoding a flagellin-like protein of unknown function, echoes almost exactly the transcription pattern of the other flagellins; this is compatible with the prediction of both σ^{28} -dependent and σ^{54} -dependent

promoters upstream (see above), but conflicts with the data for *fliDS*, which appear to be in the same transcription unit, and there is also contradiction in the lack of reciprocal upregulation of *flaG* in the *fliK* mutant. Furthermore, Carrillo *et al.* (2004) have reported that in a *fliA* deletion mutant, transcription of *flaG*, like that of *fliDS*, is abolished, indicating that only the σ^{28} -dependent promoter is active, and this is more compatible with our observations on transcription of *fliDS*. We conclude that the apparent down-regulation of *flaG* in the *flgR* and *rpoN* knockouts may be artefactual, possibly due in the case of the *flgR* mutant to some cross-reactivity of probes between *flaG* and the other flagellin genes. We note that our data (see supplementary tables) indicate strong σ^{28} -dependency for *Cj0977*, as reported by Carrillo *et al.* (2004) and recently implicated in the virulence of strain 81-176 (Goon *et al.*, 2006).

Genes of the flagellar modification region, including *Cj1293–Cj1298* (this correlates with the σ^{54} -dependent promoter for this region reported by Goon *et al.*, 2003), *Cj1320*, and *Cj1324* to *Cj1344c*, were all upregulated in the *fliK* mutant, with the striking exception of the flagellin genes *flaA* and *flaB* (see Table 3 and supplementary tables). Genes in the *Cj1293–Cj1298* cluster have recently been strongly implicated in synthesis of sugars involved in flagellin glycosylation (Obhi & Creuzenet, 2005), and detailed analysis in strain 81176 has further defined the functions of many of the remaining genes (Guerrey *et al.*, 2006). Thus, a strong case can be made for co-regulation of flagellin-modification genes with the rod and hook genes described above, implying that the flagellin glycosylation mechanism may be assembled at some point after the export machinery is installed, and possibly at the same time as rod and hook construction. The flagellin gene data, however, are difficult to interpret due to cross-hybridization between *flaA* and *flaB* probes (see above), and it is possible that overexpression of *flaB* compensates for down-regulation of *flaA*. Other genes upregulated in the *fliK* mutant include many of the capsular polysaccharide genes in the region from *Cj1424c* to *Cj1444c*, suggesting co-regulation of flagella and capsule, which together form a significant fraction of the virulence-associated surface structure of the organism. Also upregulated in the *fliK* mutant are heat-shock genes associated with protein misfolding and degradation (*hrcA*, *grpE*, *dnaK*, *htrAB* and *clpA*), and genes associated with the stringent response, typified by *Cj1272c*, a *spoT* homologue (Gaynor *et al.*, 2005). These changes are consistent with a stress-related response to misdirection of the assembly of the flagellum.

Application of the promoter prediction algorithm to the entire genome including coding regions revealed large numbers of probably falsely identified promoter-like sequences. However, for putative σ^{54} -dependent promoters in intergenic locations, the method appears to have been highly successful in that many of the most strongly responsive signals in the microarray analyses correlated with high-scoring promoter predictions (Table 4). In addition to the intergenic promoters identified, high-scoring putative

σ^{54} -dependent promoter sequences were identified within coding regions, including one within *Cj0063c*, which correlates with the observed increased expression in the *fliK* mutant of the *Cj0062c–Cj0059c* (*fliA*) gene cluster (see above). The consensus sequence for σ^{54} -dependent promoters, 5'-TTGGAACACTTTTGTGCT-3', that we now correlate with gene expression data further supports the predicted consensus described by Carrillo *et al.* (2004). We have now also obtained confirmatory experimental evidence by 5' rapid amplification of cDNA ends (RACE; unpublished results) of transcript initiation at representative predicted promoters of both σ^{54} - and σ^{28} -dependent classes.

In this study we have identified and characterized the important checkpoint component FliK, highlighting the role of this protein in the switching of transcription from rod-hook genes to filament-associated genes, and delineating a single σ^{54} -dependent regulon. Further studies are required to elucidate the control of gene expression in the early stages of flagellar biogenesis in *C. jejuni*, and also to explore the significance of σ^{54} -dependent expression of other genes, including those involved in flagellin modification.

ACKNOWLEDGEMENTS

We acknowledge the Darwin Trust of Edinburgh (N.K., A.J.) for studentships, the Biotechnology and Biological Sciences Research Council (grant 6/D14520, S.M.T.; grant D15819, N.D.; grant 6/EGA16107, C.C.), and the Wellcome Trust (G.M., J.H., K.G.L., B.W.W.) for funding. We thank Shin-Ichi Aizawa for helpful discussions of the structure of the FlgE filaments described.

REFERENCES

- Barrios, H., Valderrama, B. & Morett, E. (1999). Compilation and analysis of σ^{54} -dependent promoter sequences. *Nucleic Acids Res* **27**, 4305–4313.
- Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* **57**, 289–300.
- Brazma, A., Robinson, A., Cameron, G. & Ashburner, M. (2000). One-stop shop for microarray data. *Nature* **403**, 699–700.
- Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C. A. & other authors (2001). Minimum information about a microarray experiment (MIAME) – toward standards for microarray data. *Nat Genet* **29**, 365–371.
- Buck, M., Gallegos, M. T., Studholme, D. J., Guo, Y. & Gralla, J. D. (2000). The bacterial enhancer-dependent σ^{54} (σ^N) transcription factor. *J Bacteriol* **182**, 4129–4136.
- Carrillo, C. D., Taboada, E., Nash, J. H., Lanthier, P., Kelly, J., Lau, P. C., Verhulp, R., Mykytczuk, O., Sy, J. & other authors (2004). Genome-wide expression analyses of *Campylobacter jejuni* NCTC11168 reveals coordinate regulation of motility and virulence by *fliA*. *J Biol Chem* **279**, 20327–20338.
- Chilcott, G. S. & Hughes, K. T. (2000). Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica* serovar Typhimurium and *Escherichia coli*. *Microbiol Mol Biol Rev* **64**, 694–708.
- Dasgupta, N., Wolfgang, M. C., Goodman, A. L., Arora, S. K., Jyot, J., Lory, S. & Ramphal, R. (2003). A four-tiered transcriptional

regulatory circuit controls flagellar biogenesis in *Pseudomonas aeruginosa*. *Mol Microbiol* **50**, 809–824.

Ferris, H. U. & Minamino, T. (2006). Flipping the switch: bringing order to flagellar assembly. *Trends Microbiol* **14**, 519–526.

Frech, K., Herrmann, G. & Werner, T. (1993). Computer-assisted prediction, classification, and delimitation of protein binding sites in nucleic acids. *Nucleic Acids Res* **21**, 1655–1664.

Gaynor, E. C., Wells, D. H., MacKichan, J. K. & Falkow, S. (2005). The *Campylobacter jejuni* stringent response controls specific stress survival and virulence-associated phenotypes. *Mol Microbiol* **56**, 8–27.

Golden, N. J. & Acheson, D. W. (2002). Identification of motility and autoagglutination *Campylobacter jejuni* mutants by random transposon mutagenesis. *Infect Immun* **70**, 1761–1771.

Goon, S., Kelly, J. F., Logan, S. M., Ewing, C. P. & Guerry, P. (2003). Pseudaminic acid, the major modification on *Campylobacter* flagellin, is synthesized via the Cj1293 gene. *Mol Microbiol* **50**, 659–671.

Goon, S., Ewing, C. P., Lorenzo, M., Pattarini, D., Majam, G. & Guerry, P. (2006). A σ^{28} -regulated nonflagella gene contributes to virulence of *Campylobacter jejuni* 81–176. *Infect Immun* **74**, 769–772.

Guerry, P., Alm, R. A., Power, M. E., Logan, S. M. & Trust, T. J. (1991). Role of two flagellin genes in *Campylobacter* motility. *J Bacteriol* **173**, 4757–4764.

Guerry, P., Alm, R. A., Szymanski, C. M. & Trust, T. J. (2000). Structure, function and antigenicity of *Campylobacter* flagella. In *Campylobacter*, pp. 405–421. Edited by I. Nachamkin & M. J. Blaser. Washington, DC: American Society for Microbiology.

Guerry, P., Ewing, C. P., Schirm, M., Lorenzo, M., Kelly, J., Pattarini, D., Majam, G., Thibault, P. & Logan, S. (2006). Changes in flagellin glycosylation affect *Campylobacter* autoagglutination and virulence. *Mol Microbiol* **60**, 299–311.

Hendrixson, D. R. & DiRita, V. J. (2003). Transcription of σ^{54} -dependent but not σ^{28} -dependent flagellar genes in *Campylobacter jejuni* is associated with formation of the flagellar secretory apparatus. *Mol Microbiol* **50**, 687–702.

Hinds, J., Laing, K. G., Mangan, J. A. & Wren, B. W. (2002a). Glass slide microarrays for bacterial genomes. In *Methods in Microbiology: Functional Microbial Genomics*, pp. 83–99. Edited by B. W. Wren & N. Dorrell. London: Academic Press.

Hinds, J., Witney, A. A. & Vass, J. K. (2002b). Microarray design for bacterial genomes. In *Methods in Microbiology: Functional Microbial Genomics*, pp. 67–82. Edited by B. W. Wren & N. Dorrell. London: Academic Press.

Hirano, T., Yamaguchi, S., Oosawa, K. & Aizawa, S. (1994). Roles of FlhK and FlhB in determination of flagellar hook length in *Salmonella typhimurium*. *J Bacteriol* **176**, 5439–5449.

Hirano, T., Minamino, T. & Macnab, R. M. (2001). The role in flagellar rod assembly of the N-terminal domain of *Salmonella* FlgJ, a flagellum-specific muramidase. *J Mol Biol* **312**, 359–369.

Hirano, T., Shibata, S., Ohnishi, K., Tani, T. & Aizawa, S. (2005). N-terminal signal region of FlhK is dispensable for length control of the flagellar hook. *Mol Microbiol* **56**, 346–360.

Ide, N., Ikebe, T. & Kutsukake, K. (1999). Reevaluation of the promoter structure of the class 3 flagellar operons of *Escherichia coli* and *Salmonella*. *Genes Genet Syst* **74**, 113–116.

Jagannathan, A. & Penn, C. W. (2005). Motility. In *Campylobacter: Molecular and Cellular Biology*, pp. 331–347. Edited by J. M. Ketley. Norwich: Horizon Bioscience.

Jagannathan, A., Constantinidou, C. & Penn, C. W. (2001). Roles of *rpoN*, *fliA*, and *flgR* in expression of flagella in *Campylobacter jejuni*. *J Bacteriol* **183**, 2937–2942.

Jyot, J., Dasgupta, N. & Ramphal, R. (2002). FleQ, the major flagellar gene regulator in *Pseudomonas aeruginosa*, binds to enhancer sites located either upstream or atypically downstream of the RpoN binding site. *J Bacteriol* **184**, 5251–5260.

Karlyshev, A. V., McCrossan, M. V. & Wren, B. W. (2001). Demonstration of polysaccharide capsule in *Campylobacter jejuni* using electron microscopy. *Infect Immun* **69**, 5921–5924.

Keener, J. P. (2005). A model for length control of flagellar hooks of *Salmonella typhimurium*. *J Theor Biol* **234**, 263–275.

Kim, Y. K. & McCarter, L. L. (2000). Analysis of the polar flagellar gene system of *Vibrio parahaemolyticus*. *J Bacteriol* **182**, 3693–3704.

Kinsella, N., Guerry, P., Cooney, J. & Trust, T. J. (1997). The *flgE* gene of *Campylobacter coli* is under the control of the alternative sigma factor σ^{54} . *J Bacteriol* **179**, 4647–4653.

Konkel, M. E., Monteville, M. R., Rivera-Amill, V. & Joens, L. A. (2001). The pathogenesis of *Campylobacter jejuni*-mediated enteritis. *Curr Issues Intest Microbiol* **2**, 55–71.

Konkel, M. E., Klena, J. D., Rivera-Amill, V., Monteville, M. R., Biswas, D., Raphael, B. & Mickelson, J. (2004). Secretion of virulence proteins from *Campylobacter jejuni* is dependent on a functional flagellar export apparatus. *J Bacteriol* **186**, 3296–3303.

Livak, K. J. (1997). ABI Prism 7700 Sequence Analysis Detection System. User bulletin no. 2, PE Applied Biosystems, AB Website, bulletin reference: 4303859B, 777802–002. <http://docs.appliedbiosystems.com/pediodocs/04303859.pdf>.

Livak, K. J. & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402–408.

Loong Chan, V., Louie, H. & Joe, A. (1998). Expression of the *flgFG* operon of *Campylobacter jejuni* in *Escherichia coli* yields an extra fusion protein. *Gene* **225**, 131–141.

Luneberg, E., Glenn-Calvo, E., Hartmann, M., Bar, W. & Frosch, M. (1998). The central, surface-exposed region of the flagellar hook protein FlgE of *Campylobacter jejuni* shows hypervariability among strains. *J Bacteriol* **180**, 3711–3714.

Macnab, R. M. (1996). Flagella and motility. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, pp. 123–145. Edited by F. C. Neidhardt. Washington, DC: American Society for Microbiology.

Macnab, R. M. (2003). How bacteria assemble flagella. *Annu Rev Microbiol* **57**, 77–100.

Minamino, T. & Macnab, R. M. (2000). Domain structure of *Salmonella* FlhB, a flagellar export component responsible for substrate specificity switching. *J Bacteriol* **182**, 4906–4914.

Minamino, T. & Pugsley, A. P. (2005). Measure for measure in the control of type III secretion hook and needle length. *Mol Microbiol* **56**, 303–308.

Minamino, T., Gonzalez-Pedrajo, B., Yamaguchi, K., Aizawa, S. I. & Macnab, R. M. (1999). FlhK, the protein responsible for flagellar hook length control in *Salmonella*, is exported during hook assembly. *Mol Microbiol* **34**, 295–304.

Minamino, T., Saijo-Hamano, Y., Furukawa, Y., Gonzalez-Pedrajo, B., Macnab, R. M. & Namba, K. (2004). Domain organization and function of *Salmonella* FlhK, a flagellar hook-length control protein. *J Mol Biol* **341**, 491–502.

Mohr, C. D., MacKichan, J. K. & Shapiro, L. (1998). A membrane-associated protein, FlhX, is required for an early step in *Caulobacter* flagellar assembly. *J Bacteriol* **180**, 2175–2185.

Nuijten, P. J., van Asten, F. J., Gaastra, W. & van der Zeijst, B. A. (1990). Structural and functional analysis of two *Campylobacter jejuni* flagellin genes. *J Biol Chem* **265**, 17798–17804.

- Obhi, R. K. & Creuzenet, C. (2005). Biochemical characterization of the *Campylobacter jejuni* Cj1294, a novel UDP-4-keto-6-deoxy-GlcNAc aminotransferase that generates UDP-4-amino-4,6-dideoxy-GalNAc. *J Biol Chem* **280**, 20902–20908.
- Pallen, M. J., Penn, C. W. & Chaudhuri, R. R. (2005). Bacterial flagellar diversity in the post-genomic era. *Trends Microbiol* **13**, 143–149.
- Park, K., Choi, S., Ko, M. & Park, C. (2001). Novel σ^F -dependent genes of *Escherichia coli* found using a specified promoter consensus. *FEMS Microbiol Lett* **202**, 243–250.
- Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T. & other authors (2000). The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403**, 665–668.
- Patterson-Delafield, J., Martinez, R. J., Stocker, B. A. & Yamaguchi, S. (1973). A new *fla* gene in *Salmonella typhimurium* – *flaR* – and its mutant phenotype-superhooks. *Arch Mikrobiol* **90**, 107–120.
- Penn, C. W. (2001). Surface components of *Campylobacter* and *Helicobacter*. *Symp Ser Soc Appl Microbiol* **30**, 25S–35S.
- Prouty, M. G., Correa, N. E. & Klose, K. E. (2001). The novel σ^{54} - and σ^{28} -dependent flagellar gene transcription hierarchy of *Vibrio cholerae*. *Mol Microbiol* **39**, 1595–1609.
- Ryan, K. A., Karim, N., Worku, M., Penn, C. W. & O'Toole, P. W. (2005). *Helicobacter pylori* flagellar hook-filament transition is controlled by a *FliK* functional homolog encoded by the gene HP0906. *J Bacteriol* **187**, 5742–5750.
- Silverman, M. R. & Simon, M. I. (1972). Flagellar assembly mutants in *Escherichia coli*. *J Bacteriol* **112**, 986–993.
- Song, Y. C., Jin, S., Louie, H., Ng, D., Lau, R., Zhang, Y., Weerasekera, R., Al Rashid, S., Ward, L. A. & other authors (2004). FlaC, a protein of *Campylobacter jejuni* TGH9011 (ATCC43431) secreted through the flagellar apparatus, binds epithelial cells and influences cell invasion. *Mol Microbiol* **53**, 541–553.
- Wassenaar, T. M. (1997). Toxin production by *Campylobacter* spp. *Clin Microbiol Rev* **10**, 466–476.
- Wassenaar, T. M., Fry, B. N. & van der Zeijst, B. A. (1993a). Genetic manipulation of *Campylobacter*: evaluation of natural transformation and electro-transformation. *Gene* **132**, 131–135.
- Wassenaar, T. M., van der Zeijst, B. A., Ayling, R. & Newell, D. G. (1993b). Colonization of chicks by motility mutants of *Campylobacter jejuni* demonstrates the importance of flagellin A expression. *J Gen Microbiol* **139**, 1171–1175.
- Wosten, M. M., Boeve, M., Koot, M. G., van Nuene, A. C. & van der Zeijst, B. A. (1998). Identification of *Campylobacter jejuni* promoter sequences. *J Bacteriol* **180**, 594–599.
- Wosten, M. M., Wagenaar, J. A. & van Putten, J. P. (2004). The FlgS/FlgR two-component signal transduction system regulates the *fla* regulon in *Campylobacter jejuni*. *J Biol Chem* **279**, 16214–16222.
- Wren, B. W., Henderson, J. & Ketley, J. M. (1994). A PCR-based strategy for the rapid construction of defined bacterial deletion mutants. *Biotechniques* **16**, 994–996.

Edited by: J. G. Shaw