Published in final edited form as:
Microbiology. 2005 June ; 151(Pt 6): 1779–1788. doi:10.1099/mic.0.27879-0.

Binding and transcriptional activation of non-flagellar genes by the *Escherichia coli* flagellar master regulator $FlhD_2C_2$

Graham P. Stafford[1], **Tomoo Ogi**[2], and **Colin Hughes**[1]

[1]University of Cambridge, Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK

[2]Genome Damage and Stability Centre, University of Sussex, Science Park Road, Falmer, Brighton BN1 9QG, UK

Abstract

The gene hierarchy directing biogenesis of peritrichous flagella on the surface of *Escherichia coli* and other enterobacteria is controlled by the heterotetrameric master transcriptional regulator $FlhD_2C_2$. To assess the extent to which $FlhD_2C_2$ directly activates promoters of a wider regulon, a computational screen of the *E. coli* genome was used to search for gene-proximal DNA sequences similar to the 42–44 bp inverted repeat $FlhD_2C_2$ binding consensus. This identified the binding sequences upstream of all eight flagella class II operons, and also putative novel $FlhD_2C_2$ binding sites in the promoter regions of 39 non-flagellar genes. Nine representative non-flagellar promoter regions were all bound *in vitro* by active reconstituted $FlhD_2C_2$ over the K_D range 38–356 nM, and of the nine corresponding chromosomal promoter–*lacZ* fusions, those of the four genes *b1904*, *b2446*, wzz_{fepE} and *gltI* showed up to 50-fold dependence on $FlhD_2C_2$ *in vivo*. In comparison, four representative flagella class II promoters bound $FlhD_2C_2$ in the K_D range 12–43 nM and were upregulated *in vivo* 30- to 990-fold. The $FlhD_2C_2$-binding sites of the four regulated non-flagellar genes overlap by 1 or 2 bp the predicted −35 motif of the $FlhD_2C_2$-activated σ^{70} promoters, as is the case with $FlhD_2C_2$-dependent class II flagellar promoters. The data indicate a wider $FlhD_2C_2$ regulon, in which non-flagellar genes are bound and activated directly, albeit less strongly, by the same mechanism as that regulating the flagella gene hierarchy.

INTRODUCTION

The motility of bacteria like *Escherichia coli*, *Salmonella typhimurium* and *Proteus mirabilis* is mediated by 15 μmm long peritrichous flagella: helical ‘propellers’ assembled on the cell surface. More than 40 genes are specifically required for flagellar biogenesis (Macnab, 1996), and these are organized with those for chemotaxis in a transcriptional hierarchy that underlies temporal and spatial control of the assembly process (Kutsukake *et al.*, 1990; Kalir *et al.*, 2001; Soutourina & Bertin, 2003). At the apex of this hierarchy is the flagellar master operon *flhDC* which assimilates environmental and physiological signals, and in *E. coli* and related bacteria is tightly regulated at the transcriptional (Dufour *et al.*, 1998; Soutourina & Bertin, 2003; Francez-Charlot *et al.*, 2003) and translational (Claret & Hughes, 2000a; Tomoyasu *et al.*, 2003) levels. This activates expression of the ‘early’ class II genes encoding membrane components of the flagellar basal body, the cytosolic and membrane proteins of the export machinery, and the sigma factor σ^{28} that switches on class III genes encoding chemotaxis proteins and the structural subunits of the flagellum (Chadsey *et al.*, 1998; Karlinsey *et al.*, 2000; Soutourina *et al.*, 1999). Activation of flagellar class II

Correspondence Graham P. Stafford, gps25@cam.ac.uk.

Europe PMC Funders Author Manuscripts

promoters is determined by the heterotetrameric $FlhD_2C_2$ complex, composed of the FlhD (13·3 kDa) and FlhC (21·5 kDa) proteins, which are closely conserved among enterobacterial species (Kutsukake *et al.*, 1990; Givskov *et al.*, 1995; Furness *et al.*, 1997; Young *et al.*, 1999; Givaudan & Lanois, 2000). While $FlhC_2$ dimers can bind to target class II promoters independently, $FlhD_2$ enhances the affinity, stability and specificity of the interaction between class II flagellar promoters and the $FlhD_2C_2$ complex (Claret & Hughes, 2002).

In addition to its role in swimming motility, the *flhDC* operon is pivotal to multicellular swarming migration (Allison & Hughes, 1991; Fraser *et al.*, 2002), during which it is strongly upregulated, and contributes to virulence-factor expression in several pathogens (Allison *et al.*, 1994; Dufour *et al.*, 1998; Young *et al.*, 1999; Kim *et al.*, 2003). In addition, microarray comparisons of mRNA levels from *E. coli* wild-type and *flhDC* mutant strains have suggested that the flagellar master operon regulates several non-flagellar genes (Prüss *et al.*, 2001, 2003). These observations have encouraged the concept of a substantial *flhDC* transcriptional 'regulon', but it is not clear whether *flhDC*-dependent regulation is determined directly by $FlhD_2C_2$ binding to non-flagellar promoters or reflects indirect pathways in which $FlhD_2C_2$ acts by influencing other regulators.

DNase footprinting and primer extension analyses of class II flagellar promoter sequences from *P. mirabilis*, *E. coli* and *S. typhimurium* (Kutsukake *et al.*, 1990; Liu & Matsumura, 1996; Claret & Hughes, 2002) have generated a consensus $FlhD_2C_2$ binding sequence of a 42–44 bp imperfect inverted repeat in which two '$FlhD_2C_2$ box' arms, each AA(C/T)G(C/G)$N_{2/3}$AAATA(A/G)CG, are separated by a non-conserved spacer of 10–12 nucleotides ($N_{10–12}$) (Claret & Hughes, 2002). To assess the nature and extent of the proposed '*flhDC* regulon', we have used this consensus recognition sequence to identify putative target genes in the *E. coli* genome. Putative target promoters identified were assayed by *in vitro* binding of reconstituted, transcriptionally active $FlhD_2C_2$, and by *in vivo* $FlhD_2C_2$-dependent transcription of the chromosomal genes.

METHODS

Computer survey of the *E. coli* genome

The *E. coli* MG1655 genome (GenBank accession no. U00096) was surveyed for putative $FlhD_2C_2$ binding sites by a heterology index (HI) based computer program, details of which can be obtained from Tomoo Ogi (t.ogi@sussex.ac.uk). Briefly, a scoring matrix was compiled from the sequences of both $FlhD_2C_2$ boxes (i.e. the arms of the inverted repeat) of 18 known class II flagellar promoters, including the 12 experimentally defined promoters originally used by Claret & Hughes (2002). Only the length of the non-conserved $N_{10–12}$ spacer was considered, not the base sequence. HI values were thus calculated for every one of the ~$4·64×10^6$ possible 42–44 base sequences of the genome by adding the HI for appropriately spaced ($N_{10–12}$) inverted $FlhD_2C_2$ box arms. This allowed identification of sites with a range of HI values from 6·1 upwards. A perfect consensus binding site would have an HI value of 0, with HI values above 20 indicating weak potential for true $FlhD_2C_2$ binding, judging from their high frequency (~500), and previous experiments on other regulons (Berg & Von Hippel, 1988; Lewis *et al.*, 1994; Fernandez de Henestrosa *et al.*, 2000).

Bacterial strains

E. coli strains were grown at 37 °C in LB broth or on LB agar. *E. coli* XL-1 Blue or XL10 Gold (Stratagene) transformants were selected by appropriate antibiotics (gentamicin, 5 μg ml^{-1}; kanamycin and ampicillin, 50 μg ml^{-1}) plus 80 mg X-Gal ml^{-1} and 20 mM IPTG.

Motile *E. coli* MC1000 wild-type and the non-motile *flhD*::*km* derivative containing a polar Tn5 transposon insertion were obtained from B. Prüss (University of Illinois, Chicago, USA). *E. coli* MC1000 strains *fur*::*km*, *fliA*::*km* and *hns*::*km* were created by P1 phage transduction from MC4100 *hns*::*km*, MG1655 *fliA*::*km* and MC4100 *fur*::*km*, obtained from Simon Andrews (University of Reading, UK) and the University of Wisconsin *E. coli* genome project (Kang *et al.*, 2004).

In vitro DNA binding

The FlhD and FlhC proteins were expressed separately in *E. coli* BL21 DE3 from plasmids pET11-FlhD and pET15-FlhC and recovered from insoluble pellets after cell lysis in a French pressure cell, as previously described (Claret & Hughes, 2000a). The FlhD pellet was resuspended in 6 M urea, 20 mM Tris/HCl (pH 8·0), and the protein purified by anion-exchange chromotography; the FlhC pellet was resuspended in 6 M guanidine/HCl, pH 8·0, and His-tagged FlhC protein was affinity purified by Ni-NTA chromatography. The two proteins were mixed, refolded and solubulized during dialysis against binding buffer (20 mM Tris/HCl, pH 8·0, 0·1 M NaCl, 0·1 mM EDTA, 1 mM DTT, 1 mM $MgCl_2$, 100 mg BSA ml^{-1}, 15%, w/v, glycerol) according to Claret & Hughes (2000a) to form the transcriptionally active $FlhD_2C_2$ complex. The His-tagged FlhC is comparable to the native FlhC in $FlhD_2C_2$ complex formation and DNA binding, and was able to complement the non-motile phenotype of an *flhC* mutant (not shown). The complex was pre-incubated in 20 ml binding buffer for 20 min at 25 °C, then incubated for 20 min with radiolabelled DNA probes of 175–226 bp PCR amplified from the *E. coli* chromosome with *Pfu* Turbo polymerase and oligonucleotide primers containing *Eco*RI and *Bam*HI sites. Each PCR product was digested with *Eco*RI and dephosphorylated by calf intestinal alkaline phosphatase (Roche) before 5′ phosphorylation with [γ-^{32P}]ATP (Amersham) using phage T4 polynucleotide kinase (NEB) and purification on a 5% (w/v) polyacrylamide gel containing 1× TBE (100 mM Tris/borate, pH 8·3, 2 mM EDTA). The resulting end-labelled probes were included at a concentration of 0·3 nM (i.e. providing a molar excess of protein complex) with a 1000-fold excess of non-specific competitior poly(dIdC). After electrophoresis of DNA : protein complexes through 6% polyacrylamide (0·5× TBE), the gels were dried and the relative intensities of the bound/unbound DNA were detected using a cyclone phosphorimager and quantified by OPTIQUANT software (Packard). Binding affinity was expressed as K_D, the concentration of $FlhD_2C_2$ complex required to achieve a ratio of 1 : 1 for the free : bound DNA probe. All assays were performed at least twice and the mean binding affinities (K_D) reported.

Construction of chromosomal transcriptional fusions

Plasmid pGPS123 was constructed by excision of the kanamycin resistance gene of pRS551 (Simons *et al.*, 1987) using *Xho*I and *Hind*III and replacement with the gentamicin resistance cassette of p34SGm (Dennis & Zylstra, 1998). DNA sequences used as probes in *in vitro* DNA-band shift assays were PCR amplified with *Pfu* Turbo polymerase from the *E. coli* chromosome using specific oligonucleotide primers that each generated a 5′ *Eco*RI site and a 3′ *Bam*HI site. DNA fragments were cleaned using QIAspin columns (Qiagen), digested with *Eco*RI and *Bam*HI, and ligated with *Eco*RI/ *Bam*HI digested pGPS123. After transformation into *E. coli* XL1 Blue or XL10, blue Gm^R colonies were selected, and inserts verified by PCR and sequencing (Department of Genetics, University of Cambridge) using vector-specific primers.

The *flhB*-80 promoter fragment was obtained by first amplifying the $FlhD_2C_2$ binding site plus 95 bp of 5′ DNA using primers FlhBFor (TTGAATTCATGGTGGCGTGACCACCACGTCAT) and FlhBDCRev (TAGCCGCGGTGATGCCAGAAAAAAACCCCGTCACGTTCAAGCTTAATGGTTGA

GTAAGG), then the σ^{70} promoter plus 57 bp of 3′ DNA using primers FlhBRev (CGGGATCCTTTTGTCGTCGCTCTCG) and FlhBs70 (TAGCCCGCGGTGATGCGCTTGATGGCCGAGTCTCCGATTACAAGCTTGAACGCT TTGCGC). These two primers contain SacII restriction sites (CCGCGG) at their 5′ ends, allowing ligation of the PCR products to form the 270 bp *flhB*-80 fragment construct and its cloning into pGPS123. The source of the intervening 'extra' sequence in primers FlhBs70For and FlhB-DCRev is the *flhA* gene, which contains no transcriptional features.

MC1000 strains containing each pGPS123 transcriptional-fusion-bearing plasmid were individually infected with phage λRS45, and lysates were used to produce blue Gm^R Ap^S lysogen colonies (Simons *et al.*, 1987). Fusion inserts in the chromosome were confirmed by sequencing using vector- and insert-specific primers, and correct insertion into the λ*att* site was verified according to the method of Powell *et al.* (1994).

The *flhDC* complementation plasmid p*flhDC* was constructed by amplification of the *flhDC* genes and flanking DNA containing the native promoter sequence, using the primers DC-coliF (GGAATTCTGCGCAACATCCCATTTCG) and DC-coliR (GGGAATTCCAGTTAAACAGCCTGTACTCT), restriction with *Eco*RI, and ligation into *Eco*RI-restricted pAcTrc (provided by Dr Gillian Fraser, Department of Pathology, University of Cambridge).

In vivo assay of transcription activity

Chromosomal transcriptional fusion expression was assessed as whole-cell β-galactosidase activity (Miller, 1972). Triplicate overnight cultures were diluted to OD_{600}=0·001 in LB and sampled hourly during batch culture. We performed the assays at 37 °C.

The *E. coli* MC1000 strain showed comparable motility at 37 °C and 30 °C (not shown), in agreement with observations of constant swimming speeds between 24 °C and 37 °C, and comparable FliC protein and transcript levels over a similar temperature range (Adler & Templeton, 1967; Mizushima *et al.*, 1994). Activities were recorded as means of at least three cultures. In complementation experiments, *flhDC* was provided *in trans* by expression without induction from the plasmid p*flhDC*.

RESULTS

Computational identification of putative $FlhD_2C_2$ regulon genes

To identify genes of the potential $FlhD_2C_2$ regulon, the 4639 kb DNA of the *E. coli* MG1655 genome sequence was surveyed by a statistical mechanics search program, as previously described in the characterization of the cAMP receptor protein (CRP) and LexA regulons (Berg & Von Hippel, 1988; Lewis *et al.*, 1994; Fernandez de Henestrosa *et al.*, 2000). In the program, differences between a query DNA sequence and the protein binding site consensus are assigned a penalty reflecting the frequency of each base at every position in known target sequences, and this is used to calculate a heterology index (HI). Low HI scores indicate closeness to the consensus.

The search identified 7834 genome sequences with an HI lower than 25, while 499 had an HI below 20, and 145 an HI equal to or below 18. Of the 145 sequences, 99 were within coding sequences and/or located more than 250 bp 5′ of predicted start codons. These were excluded, since the known $FlhD_2C_2$ binding sequences of flagellar class II promoters are all between 74 bp and 160 bp 5′ of their respective start codons. This left 47 putative $FlhD_2C_2$ binding sequences with an HI lower than our arbitrary cut-off of 18 or in non-coding sequence close to genes putatively transcribed alone or in operons. These are listed in Table 1 with the known or indicated function of the gene immediately 3′ of the promoter. This

number is comparable with the 69 putative LexA regulon binding sites similarly identified with an HI lower than 15 (Fernandez de Henestrosa *et al.*, 2000). The 47 promoter regions are estimated to control 86 of the 4287 predicted ORFs in the *E. coli* genome (http://genolist.pasteur.fr/Colibri/help/current.html).

The search identified the binding sequences of all the seven known class II flagellar operons of *E. coli*, each with a low HI value: *flgAMN* (HI 6·1), *flgBCDEFGHIJ* (HI 6·1), *flhBAE* (HI 11·9), *fliE* (HI 12·2), *fliFGHIJK* (HI 12·2), *fliAZY* (HI 13·1) and *fliLMNOPQR* (HI 14·7). An $FlhD_2C_2$ binding site was also identified 5′ of the fliDST (HI 12·2) operon that has not been previously reported in *E. coli* but is compatible with the presence of both class II and III promoters for this operon in the closely related *S. typhimurium* (Ide *et al.*, 1999). This reflects the direct activation of 35 flagella genes by $FlhD_2C_2$. The remaining 39 putative $FlhD_2C_2$ binding sequences are located in promoter regions which control approximately 51 non-flagellar genes. Ten are 5′ of single genes or operons with products of known function, involved in global regulation (*hupB* and *hns* histone-like proteins, *groES*(*L*) protein folding, *icc* (*cpdA*) cAMP catabolite repression), the *secM* protein secretion regulator, polysaccharide synthesis (*uppS*) or modification (wzz_{fepE},O-antigen chain length control, *pmrD* LPS composition), and involvement in membrane transport (*livKHGMF*, leucine uptake, *chaC* putative calcium transporter, *gltI* glutamate transport). A further 20 genes have putative functions predicted only from the presence of conserved domains or motifs (http://www.tigr.org/), and eight only have a predicted cellular location (assessed at http://psort.nibb.ac.jp/).

In vitro binding of $FlhD_2C_2$ to target sites of the *E. coli* genome

The theoretical assumptions underlying this search approach have been validated by experimental binding affinities of the CRP and LexA proteins for target promoters of their regulons (Berg & Von Hippel, 1988; Lewis *et al.*, 1994; Fernandez De Henestrosa *et al.*, 2000). In our studies, DNA-band shift assays were used to assess recognition by the $FlhD_2C_2$ complex of 13 representative sequence targets with low HI chosen from the 47 listed in Table 1. Four were located 5′of the flagellar operons *flgB* (HI 6·1), *flhB* (HI 11·9), *fliA* (HI 13·1) and *fliL* (HI 14·7), representing a range of HI, to establish a positive control dataset for the study. Nine were upstream of the putative non-flagella targets *b1904* (HI 10·9), *b2446* (HI 11·5), *yejO* (HI 13·7), *gltI* (HI 13·7), *icc* (HI 14·1), *yegH* (HI 14·6), *hns* (HI 14·7), wzz_{fepE} (HI 14·9) and *hupB* (HI 15·5). Two sequences with HI scores outside the 18 cut-off, *rcsA* (HI 18·8) and *argR* (18·7), were assessed in parallel as negative controls. In each case, a radiolabelled probe of 150–230 bp, containing the putative $FlhD_2C_2$ binding site near its centre, was made by PCR amplification of genomic DNA from *E. coli* MC1000. Each probe was incubated with increasing concentrations of transcriptionally active heterotetrameric $FlhD_2C_2$ complex reconstituted from purified FlhD and FlhC proteins (Claret & Hughes, 2000b, 2002) in the presence of a large excess of the non-specific competitor poly(dIdC).

All 13 test probes, four flagellar and nine non-flagellar, were bound by $FlhD_2C_2$. This resulted in each case in a shift of the DNA to a slower-migrating species during native PAGE. Autoradiographs of three representative binding assays with sequences of differing HI are shown in Fig. 1A. To quantify this assay, phosphoimage analysis of the migrating band intensities allowed the ratio of bound : free DNA intensity to be plotted against protein concentration. The plots in Fig. 1B (i), (ii) show a broad range of $FlhD_2C_2$ affinities, with binding to several putative non-flagellar targets, such as *b1904*, comparable to that of the four class II flagella genes analysed.

From these plots, the apparent dissociation constant K_D was calculated for each binding site as the protein concentration at which 50% of probe was bound, in other words, the ratio of

bound : free DNA probe was 1. The class II flagellar promoter regions of *flgB*, *flhB*, *fliA* and *fliL* displayed low K_D values of 12·5 nM, 21 nM, 25 nM and 43 nM, respectively, as did those of genes *b1904*, *b2446* and wzz_{fepE}: 38 nM, 60 nM and 86 nM, respectively. The K_D values for the non-flagellar promoter regions of *yejO*, *gltI*, *icc*, *yegH*, *hns* and *hupB* were substantially higher at 220 nM, 220 nM, 167 nM, 240 nM, 165 nM and 250 nM, respectively. Although $FlhD_2C_2$ bound to the control probe of *rcsA*, the K_D of 356 nM was close to the limit of detection (450 nM), while no binding was observed to the *argR* sequence.

In vivo $FlhD_2C_2$-dependent promoter transcription

To assess whether the *in vitro* binding of $FlhD_2C_2$ correlated with $FlhD_2C_2$-dependent *in vivo* promoter activity, single-copy chromosomal *lacZ* transcriptional fusions were constructed to each of the promoter regions assessed in the DNA-band shift assay. Each of the 15 DNA probe sequences was fused to the *lacZ* reporter gene in the plasmid vector pGPS123, and the resulting construct was transferred as a single-copy fusion to the chromosome of both *E. coli* MC1000 wild-type and the isogenic *flhDC* null mutant by the phage λRS45 (Simons *et al.*, 1987). The β-galactosidase activity of each transcriptional fusion was compared in the two strains grown at 37 °C and also in the *flhDC* mutant complemented *in trans* by the p*flhDC* plasmid, which expresses $FlhD_2C_2$ from its native promoter.

During batch growth in LB medium, peak expression of the four class II flagella gene promoter regions was substantially higher in the wild-type than the *flhDC* mutant (Table 2): *flgB*, 136-fold; *flhB*, 31-fold; *fliA*, 990-fold; *fliL*, 209-fold. Expression from the strongly bound *b1904* promoter region was comparably higher, 191-fold, in the wild-type, while the *b2446*, wzz_{fepE} and *gltI* promoter fusions were activated 11-, six- and twofold, respectively. Expression of these four differentially activated non-flagellar gene fusions in the *flhDC* mutant was enhanced by at least tenfold when $FlhD_2C_2$ was provided by plasmid p*flhDC*. In addition, although expression from the *yejO* promoter fusion was not different in the wild-type and *flhDC* strains, it was marginally complemented, 1·5- to twofold, by *flhDC in trans*. The promoters of *yegH*, *hns*, *icc* and *hupB*, with HI below 18 but weak $FlhD_2C_2$ binding (K_D 160–250 nM), were not activated, as was the case for the *rcsA* and *argR* control sequences with K_D values of 356 and >450 nM.

Table 3 summarizes the findings for the 13 test promoter regions analysed experimentally: similarity to the consensus binding sequence (HI), *in vitro* $FlhD_2C_2$ binding affinity and *in vivo* promoter activation by $FlhD_2C_2$. Low binding site HI correlates with low K_D, in other words, strong binding, and in turn K_D correlates with activation. The data indicate a K_D threshold of about 100–200 nM required for *in vivo* activation by $FlhD_2C_2$, but it seems that the location of the binding site with respect to the σ^{70} promoter is also critical. All seven of the genes activated *in vivo*, including *gltI*, have putative $FlhD_2C_2$ binding sites which overlap by 1 or 2 bp the known or likely σ^{70} −35 motif. Those promoter regions that bound, but were nonetheless not activated, that is, those of *yegH*, *hns*, *icc*, *rcsA* and *argR*, do not show this binding site overlap. The only apparent exception to this is the *hupB* promoter, which was not activated in this assay.

To assess experimentally the influence of spacing between the $FlhD_2C_2$ binding site and the σ^{70} −35 sequence (TTGACA), the binding site overlapping the strongly *flhDC*-activated flagellar class II *flhB* promoter was moved 5′ 80 bp away from its −35 motif (TTGAAC), while keeping the σ^{70} promoter and $FlhD_2C_2$ binding sequence intact (Fig. 2A). The K_D for *in vitro* binding of $FlhD_2C_2$ to the resulting site *flhB*-80 was 29 nM, comparable to the 21 nM of the wild-type binding site (Fig. 2B). However, assay of the respective chromosomal *lacZ* fusions in the wild-type and *flhDC* strains showed that the σ^{70} promoter of the

uncoupled site was no longer activated by $FlhD_2C_2$ (Fig. 2C), and *in trans* overexpression of $FlhD_2C_2$ from p*flhDC* failed to restore promoter activation.

DISCUSSION

We have set out to identify non-flagellar *E. coli* genes directly under the control of the flagella master transcriptional regulator $FlhD_2C_2$, in other words to ascertain the nature of a putative wider regulon. We identified possible $FlhD_2C_2$ binding sites by comparing every 42–44 bp genomic sequence with the experimentally defined consensus $FlhD_2C_2$ binding site (Berg & Von Hippel, 1988; Lewis *et al.*, 1994; Claret & Hughes, 2002). The computational promoter search identified a pool of 47 sequences 5′ of putative gene promoters, with an HI below the arbitrary cut off of 18. These were located upstream of all eight class II flagellar promoters and 39 non-flagellar promoters, a figure comparable with the 69 putative *E. coli* LexA regulon binding sites identified with an HI below 15 (Lewis *et al.*, 1994; Fernandez de Henestrosa *et al.*, 2000). Nine non-flagellar sequences with HI values ranging from 10·8 to 15·5 were analysed by *in vitro* binding of reconstituted active $FlhD_2C_2$ and comparison of *in vivo* promoter activity in wild-type and *flhDC* mutant *E. coli*.

All nine non-flagellar sequences were bound by $FlhD_2C_2$, with affinities (K_D 38–250 nM) one- to 20-fold weaker than those of the four representative class II flagella promoter sequences tested (*flgB*, *flhB*, *fliA* and *fliL*; HI 6·5–14·7; K_D 12–43 nM). No false positive sites were identified, which contrasts with the LexA screen, in which the *dinJ* promoter of HI 7·1 was not bound *in vitro* (Fernandez de Henestrosa *et al.*, 2000). Like the four class II flagella promoters, four of the nine non-flagellar promoters tested were $FlhD_2C_2$ regulated: *b1904* (HI 10·5) was bound strongly (K_D 38 nM) and activated more than 30 fold; *b2446* (HI 11·5) and wzz_{fepE} (HI 14·9) had K_D values of 60 nM and 86 nM and were activated tenfold and sixfold, respectively. Activation of the weakly bound *gltI* promoter (HI 13·7, K_D 220 nM) was twofold higher in the wild-type than in the *flhDC* null mutant, but as was the case with *b1904*, *b2446* and wzz_{fepE}, it was strongly activated by *flhDC in trans*.

The four $FlhD_2C_2$ bound and regulated non-flagellar genes *b1904*, *b2446*, wzz_{fepE} and *gltI* are not in the same operon. Gene *b1904* is located at 42·81 min, adjacent to *ftn* (encoding the iron storage protein ferritin) in the 10 kb region between the *fliAZY* and flagellar *flhDC* operons (43·09 and 42·59 min, respectively). It has no known function or significant homologues, but does have a putative outer membrane lipoprotein cleavage signal (LGAC) near the N-terminus of its deduced amino acid sequence. The product of the *b2446* gene also has no known function or homology and lies at 55·18 min in an as yet anonymous region of the chromosome. It has a DNA-binding AT-hookmotif, present in many transcriptional regulators (Bustin & Reeves, 1996; Cayuela *et al.*, 2003). The wzz_{fepE} gene (13·31 min) was originally named *fepE*, putatively encoding part of the enterobactin uptake system in *E. coli*, but its role in iron transport has not been established (Ozenberger *et al.*, 1987; Murray *et al.*, 2003), and wzz_{fepE} has been shown to modulate O-antigen chain length in the polysaccharide capsule of *S. typhimurium* (Murray *et al.*, 2003). Polysaccharide is an important factor in the swarming motility of *S. typhimurium* and *P. mirabilis* (Gygi *et al.*, 1995; Toguchi *et al.*, 2000), and a recent microarray study observed upregulation of the wzz_{fepE} gene during swarming of *S. typhimurium* (Wang *et al.*, 2004). The fourth gene, *gltI*, encodes a periplasmic glutamate/aspartate binding protein (Urbanowski *et al.*, 2000) lying at 14·79 min in, or 5′ of, the *gltJKL* operon responsible for glutamate uptake. It is possible that $FlhD_2C_2$ regulation of these genes may reflect connections to motility, but this is unknown. We searched for $FlhD_2C_2$ binding sites in the genomes of the uropathogenic (UPEC) *E. coli* CFT073 (accession no. NC_004431) and the enterohaemorrhagic (EHEC) *E. coli* O157 (accession no. AE00517) to assess if there might be coregulation of flagellar and virulence

Europe PMC Funders Author Manuscripts

genes. The findings (not shown) were very similar to those for *E. coli* K-12, and no putative sites were identified in the pathogenicity islands of either organism.

In the four $FlhD_2C_2$-dependent non-flagellar promoters, the binding site overlaps the putative σ^{70} −35 motif by 1–2 bp (Fig. 3). These four σ^{70} −35 motifs have ≥50 % identity with the consensus (TTGACA), and in each case TT nucleotides form the end of the right hand (3′) $FlhD_2C_2$ box. This overlap is also seen in all flagellar class II gene promoters from *E. coli*, *S. typhimurium* and *P. mirabilis* (Fig. 3) (Kutsukake *et al.*, 1990; Liu *et al.*, 1995; Claret & Hughes, 2002). This identity strengthens the likelihood that $FlhD_2C_2$ activates flagellar and non-flagellar genes by the same mechanism, in other words, as a class I transcription factor contacting the α-C-terminal domain of the RNA polymerase holoenzyme during transcriptional initiation (Liu *et al.*, 1995; Claret & Hughes, 2002). The significance of this proximity was emphasized by uncoupling $FlhD_2C_2$ binding and promoter dependence following insertion of 80 bp of transcriptionally inert DNA between the $FlhD_2C_2$ binding site and the −35 motif. In the low-affinity non-regulated σ^{70} promoter regions of *argR*, *rcsA*, *icc/cpdA* and *hns*, the $FlhD_2C_2$ binding site does not overlap the −35 site: it is separated by 50 bp or more. The only possible exception to this is the weak $FlhD_2C_2$ binding site (HI 15·5, K_D 250 nM) of the *hupB* P3 promoter, which was not regulated in the *in vivo* assay. However, in this complex promoter region, the $FlhD_2C_2$ binding site overlaps high-affinity binding sites (K_D 5–10 nM) for the transcriptional regulators CRP and FIS (factor for inversion stimulation), which may be dominant under the conditions tested (Claret & Rouviere-Yaniv, 1996).

The range of K_D and activity values extends the possibility that $FlhD_2C_2$ can influence the expression of the hierarchy of flagellar and non-flagellar promoters via differential binding (Claret & Hughes, 2002). If the concentration of $FlhD_2C_2$ in exponentially growing *E. coli* is similar to the 35 nM estimated for vegetatively growing *P. mirabilis* (Claret & Hughes, 2000b), then the K_D values for non-regulated sites are over four times higher than those for any of the flagella genes tested here (maximum 43 nM), suggesting that they may be poorly bound by $FlhD_2C_2$ *in vivo*. The *gltI* promoter has the lowest affinity (K_D 220 nM) for $FlhD_2C_2$ binding and shows marginal activation *in vivo*, but taking all our data into account, it seems that direct and strong regulation by $FlhD_2C_2$ applies to promoters with a site that has an HI below 15 determining a K_D below ~100 nM, and which overlaps by 1–2 bp the σ^{70} −35 promoter motif.

While the differential affinity of $FlhD_2C_2$ for the binding sites of its regulon promoters seems important for the sequential activation of class II gene expression, these promoters are also subject to fine tuning by FliA (σ^{28}) (Kalir & Alon, 2004). We assessed the influence of FliA on the four novel $FlhD_2C_2$-regulated genes (wzz_{fepE}, *b2446*, *b1904* and *gltI*) by measuring the activity of the respective *lacZ* fusions in the MC1000 *fliA* :: *Km* strain and comparing to their activity in the wild-type strain: there was no difference except for a 50 % reduction in transcription of the *gltI* promoter (data not shown). The DNA sequence 5′ of the *gltI* start codon contains a putative FliA −10 motif, GACGATAA. This may explain the difference in the regulation of the *gltI* and *yejO* promoter fusions (twofold and non-regulated), which have identical binding affinities (220 nM) and σ^{70} sequences.

Prüss and colleagues (Prüss *et al.*, 2001) have postulated an extended $FlhD_2C_2$ regulon in *E. coli*, chiefly based on microarray studies in which the expression of several non-flagellar genes was influenced by the *flhDC* operon. However, their assay of corresponding multicopy plasmid transcriptional fusions did not establish direct activation of these genes by $FlhD_2C_2$. Similarly, we found no *in vitro* binding of $FlhD_2C_2$ to the promoter of *mreB* (data not shown), a gene identified in these microarray comparisons. Further microarrays have indicated regulation of genes encoding enzymes of the Entner–Doudoroff pathway by

$FlhD_2C_2$ via the *aer* gene. This has a characteristic σ^{28} (FliA)-dependent promoter sequence (TAAA–N_{15}–GCCGACAT) 5′ of its translational start site (Park *et al.*, 2001), but no putative $FlhD_2C_2$ binding sequence, suggesting that regulation proceeds indirectly from $FlhD_2C_2$ to *fliA* then *aer*.

Our data indicate that regulation of many other of these microarray-highlighted genes is probably indirect, since we did not find any putative promoter $FlhD_2C_2$ binding sequences with an HI below 18, or indeed 22 (the lowest HI of these genes is that of *napF*: 22·9). The microarray comparisons did not detect regulation of the $FlhD_2C_2$ regulon genes established in our study, *b1904*, *b2446*, wzz_{fepE} and *gltI*. It is possible that their transcript signals are below the detection level of the microarray technology, as was the case in microarray comparisons in *Yersinia enterocolitica* (Kapatral *et al.*, 2004), which failed to detect *flhDC* regulation of the class II and III genes *flhA*, *fleC* and *cheY*, which RT-PCR analysis confirmed were subject to two- to 12-fold regulation.

The main role for $FlhD_2C_2$ is the tight control of gene expression underlying flagella biosynthesis. Our study indicates a wider $FlhD_2C_2$ regulon, in which direct activation of non-flagellar promoters follows binding of the flagellar master regulator at the RNA polymerase binding site, though less avidly than to class II flagellar promoters.

Acknowledgments

We thank Laurent Claret, Gillian Fraser and Richard Hayward for useful discussions and Keith Hayward for computer assistance. This work was supported by a Wellcome Trust Program grant (C. H.) and a Uehara Memorial Foundation overseas research fellowship (T. O.).

Abbreviations

CRP	cAMP receptor protein
HI	heterology index

REFERENCES

Adler J, Templeton B. The effect of environmental conditions on the motility of *Escherichia coli*. J Gen Microbiol. 1967; 46:175–184. [PubMed: 4961758]

Allison C, Hughes C. Closely linked genetic loci required for swarm cell differentiation and multicellular migration by *Proteus mirabilis*. Mol Microbiol. 1991; 5:1975–1982. [PubMed: 1766373]

Allison C, Emody L, Coleman N, Hughes C. The role of swarm cell differentiation and multicellular migration in the uropathogenicity of *Proteus mirabilis*. J Infect Dis. 1994; 169:1155–1158. [PubMed: 8169413]

Berg OG, Von Hippel PH. Selection of DNA binding sites by regulatory proteins. II. The binding specificiity of the cyclic AMP receptor protein to recognition sites. J Mol Biol. 1988; 200:709–723. [PubMed: 3045325]

Bustin M, Reeves R. High-mobility-group chromosomal proteins, architectural components that facilitate chromatin function. Prog Nucleic Acid Res Mol Biol. 1996; 54:35–100. [PubMed: 8768072]

Cayuela ML, Elias-Arnanz M, Penalver-Mellado M, Padmanabhan S, Murillo FJ. The *Stigmatella aurantiaca* homolog of *Myxococcus xanthus* high-mobility-group A-type transcription factor CarD: insights into the functional modules of CarD and their distribution in bacteria. J Bacteriol. 2003; 185:3527–3537. [PubMed: 12775690]

Chadsey MS, Karlinsey JE, Hughes KT. The flagellar anti-sigma factor FlgM actively dissociates *Salmonella typhimurium* σ^{28} RNA polymerase holoenzyme. Genes Dev. 1998; 12:3123–3136. [PubMed: 9765212]

Claret L, Hughes C. Rapid turnover of FlhD and FlhC, the flagellar regulon transcriptional activator proteins, during *Proteus* swarming. J Bacteriol. 2000a; 182:833–836. [PubMed: 10633123]

Claret L, Hughes C. Functions of the subunits in the $FlhD_2C_2$ transcriptional master regulator of bacterial flagellum biogenesis and swarming. J Mol Biol. 2000b; 303:467–478. [PubMed: 11054284]

Claret L, Hughes C. Interaction of the atypical prokaryotic transcription activator $FlhD_2C_2$ with early promoters of the flagellar gene hierarchy. J Mol Biol. 2002; 321:185–199. [PubMed: 12144778]

Claret L, Rouviere-Yaniv J. Regulation of HU alpha and HU beta by CRP and FIS in *Escherichia coli*. J Mol Biol. 1996; 263:126–139. [PubMed: 8913296]

Dennis JJ, Zylstra GJ. Plasposons: modular self-cloning minitransposon derivatives for rapid genetic analysis of Gram-negative bacterial genomes. Appl Environ Microbiol. 1998; 64:2710–2715. [PubMed: 9647854]

Dufour A, Furness RB, Hughes C. Novel genes that upregulate the *Proteus mirabilis flhDC* master operon controlling flagellar biogenesis and swarming. Mol Microbiol. 1998; 29:741–751. [PubMed: 9723914]

Fernandez de Henestrosa AR, Ogi T, Aoyagi S, Chafin D, Hayes JJ, Ohmori H, Woodgate R. Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. Mol Microbiol. 2000; 35:1560–1572. [PubMed: 10760155]

Francez-Charlot A, Laugel B, Van Gemert A, Dubarry N, Wiorowski F, Castanie-Cornet MP, Gutierrez C, Cam K. RcsCDB His-Asp phosphorelay system negatively regulates the *flhDC* operon in *Escherichia coli*. Mol Microbiol. 2003; 49:823–832. [PubMed: 12864862]

Fraser GM, Claret L, Furness R, Gupta S, Hughes C. Swarming-coupled expression of the *Proteus mirabilis hpmBA* haemolysin operon. Microbiology. 2002; 148:2191–2201. [PubMed: 12101306]

Furness RB, Fraser GM, Hay NA, Hughes C. Negative feedback from a *Proteus* class II flagellum export defect to the *flhDC* master operon controlling cell division and flagellum assembly. J Bacteriol. 1997; 179:5585–5588. [PubMed: 9287017]

Givaudan A, Lanois A. *flhDC*, the flagellar master operon of *Xenorhabdus nematophilus*, requirement for motility, lipolysis, extracellular hemolysis, and full virulence in insects. J Bacteriol. 2000; 182:107–115. [PubMed: 10613869]

Givskov M, Eberl L, Christiansen G, Benedik MJ, Molin S. Induction of phospholipase and flagellar synthesis in *Serratia liquefaciens* is controlled by expression of the flagellar master operon *flhD*. Mol Microbiol. 1995; 15:445–454. [PubMed: 7783615]

Gygi D, Rahman MM, Lai HC, Carlson R, Guard-Petter J, Hughes C. A cell-surface polysaccharide that facilitates rapid population migration by differentiated swarm cells of *Proteus mirabilis*. Mol Microbiol. 1995; 17:1167–1175. [PubMed: 8594335]

Ide N, Ikebe T, Kutsukake K. Reevaluation of the promoter structure of the class 3 flagellar operons of *Escherichia coli* and *Salmonella*. Genes Genet Syst. 1999; 74:113–116. [PubMed: 10586520]

Kalir S, Alon U. Using a quantitative blueprint to reprogram the dynamics of the flagella gene network. Cell. 2004; 117:713–720. [PubMed: 15186773]

Kalir S, McClure J, Pabbaraju K, Southward C, Ronen M, Leibler S, Surette MG, Alon U. Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. Science. 2001; 292:2080–2083. [PubMed: 11408658]

Kang Y, Durfee T, Glasner JD, Qiu Y, Frisch D, Winterberg KM, Blattner FR. Systematic mutagenesis of the *Escherichia coli* genome. J Bacteriol. 2004; 186:4921–4930. [PubMed: 15262929]

Kapatral V, Campbell JW, Minnich SA, Thomson NR, Matsumura P, Pruss BM. Gene array analysis of *Yersinia enterocolitica* FlhD and FlhC: regulation of enzymes affecting synthesis and degradation of carbamoylphosphate. Microbiology. 2004; 150:2289–2300. [PubMed: 15256571]

Karlinsey JE, Tanaka S, Bettenworth V, Yamaguchi S, Boos W, Aizawa SI, Hughes KT. Completion of the hook-basal body complex of the *Salmonella typhimurium* flagellum is coupled to FlgM secretion and *fliC* transcription. Mol Microbiol. 2000; 37:1220–1231. [PubMed: 10972838]

Kim DJ, Boylan B, George N, Forst S. Inactivation of *ompR* promotes precocious swarming and *flhDC* expression in *Xenorhabdus nematophila*. J Bacteriol. 2003; 185:5290–5294. [PubMed: 12923104]

Kutsukake K, Ide N. Transcriptional analysis of the *flgK* and *fliD* operons of *Salmonella typhimurium* which encode flagellar hook-associated proteins. Mol Gen Genet. 1995; 247:275–281. [PubMed: 7770032]

Kutsukake K, Ohya Y, Iino T. Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. J Bacteriol. 1990; 172:741–747. [PubMed: 2404955]

Lewis LK, Harlow GR, Gregg-Jolly LA, Mount DW. Identification of high affinity binding sites for LexA which define new DNA damage-inducible genes in *Escherichia coli*. J Mol Biol. 1994; 241:507–523. [PubMed: 8057377]

Liu X, Matsumura P. The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. J Bacteriol. 1994; 176:7345–7351. [PubMed: 7961507]

Liu X, Matsumura P. Differential regulation of multiple overlapping promoters in flagellar class II operons in *Escherichia coli*. Mol Microbiol. 1996; 21:613–620. [PubMed: 8866483]

Liu X, Fujita N, Ishihama A, Matsumura P. The C-terminal region of the alpha subunit of *Escherichia coli* RNA polymerase is required for transcriptional activation of the flagellar level II operons by the FlhD/FlhC complex. J Bacteriol. 1995; 177:5186–5188. [PubMed: 7665504]

Macnab, RM. Flagella and Motility. In: Neidhardt, FC., editor. Escherichia coli and Salmonella typhimurium, Cellular and Molecular Biology. Washington, DC: American Society for Microbiology; 1996. p. 123-145.

Miller, JH. Experiments in Molecular Genetics. Cold Spring Harbour, NY: Cold Spring Harbor Laboratory; 1972.

Mizushima T, Tomura A, Shinpuku T, Miki T, Sekimizu K. Loss of flagellation in *dnaA* mutants of *Escherichia coli*. J Bacteriol. 1994; 176:5544–5546. [PubMed: 8071236]

Murray GL, Attridge SR, Morona R. Regulation of *Salmonella typhimurium* lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. Mol Microbiol. 2003; 47:1395–1406. [PubMed: 12603743]

Ozenberger BA, Nahlik MS, McIntosh MA. Genetic organization of multiple *fep* genes encoding ferric enterobactin transport functions in *Escherichia coli*. J Bacteriol. 1987; 169:3638–3646. [PubMed: 2956250]

Park K, Choi S, Ko M, Park C. Novel σ^F-dependent genes of *Escherichia coli* found using a specified promoter consensus. FEMS Microbiol Lett. 2001; 202:243–250. [PubMed: 11520622]

Powell BS, Rivas MP, Court DL, Nakamura Y, Turnbough CL Jr. Rapid confirmation of single copy lambda prophage integration by PCR. Nucleic Acids Res. 1994; 22:5765–5766. [PubMed: 7838735]

Prüss BM, Liu X, Hendrickson W, Matsumura P. FlhD/FlhC-regulated promoters analyzed by gene array and *lacZ* gene fusions. FEMS Microbiol Lett. 2001; 197:91–97. [PubMed: 11287152]

Prüss BM, Campbell JW, Van Dyk TK, Zhu C, Kogan Y, Matsumura P. FlhD/FlhC is a regulator of anaerobic respiration and the Entner–Doudoroff pathway through induction of the methyl-accepting chemotaxis protein Aer. J Bacteriol. 2003; 185:534–543. [PubMed: 12511500]

Simons RW, Houman F, Kleckner N. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. Gene. 1987; 53:85–96. [PubMed: 3596251]

Soutourina OA, Bertin PN. Regulation cascade of flagellar expression in Gram-negative bacteria. FEMS Microbiol Rev. 2003; 27:505–523. [PubMed: 14550943]

Soutourina O, Kolb A, Krin E, Laurent-Winter C, Rimsky S, Danchin A, Bertin P. Multiple control of flagellum biosynthesis in *Escherichia coli*, role of H-NS protein and the cyclic AMP-catabolite activator protein complex in transcription of the *flhDC* master operon. J Bacteriol. 1999; 181:7500–7508. [PubMed: 10601207]

Toguchi A, Siano M, Burkart M, Harshey RM. Genetics of swarming motility in *Salmonella enterica* serovar *typhimurium:* critical role for lipopolysaccharide. J Bacteriol. 2000; 182:6308–6321. [PubMed: 11053374]

Tomoyasu T, Takaya A, Isogai E, Yamamoto T. Turnover of FlhD and FlhC, master regulator proteins for *Salmonella* flagellum biogenesis, by the ATP-dependent ClpXP protease. Mol Microbiol. 2003; 48:443–452. [PubMed: 12675803]

Urbanowski ML, Stauffer LT, Stauffer GV. The *gcvB* gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in *Escherichia coli*. Mol Microbiol. 2000; 37:856–868. [PubMed: 10972807]

Wang Q, Frye JG, McClelland M, Harshey RM. Gene expression patterns during swarming in *Salmonella typhimurium*: genes specific to surface growth and putative new motility and pathogenicity genes. Mol Microbiol. 2004; 52:169–187. [PubMed: 15049819]

Young GM, Smith MJ, Minnich SA, Miller VL. The *Yersinia enterocolitica* motility master regulatory operon, *flhDC*, is required for flagellin production, swimming motility, and swarming motility. J Bacteriol. 1999; 181:2823–2833. [PubMed: 10217774]

Europe PMC Funders Author Manuscripts

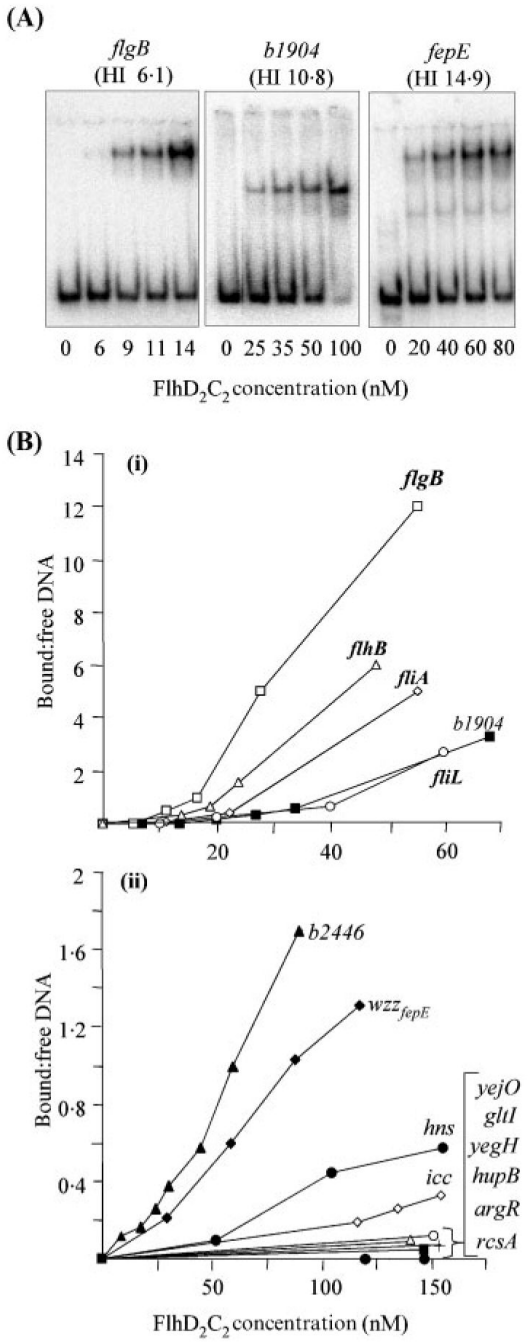
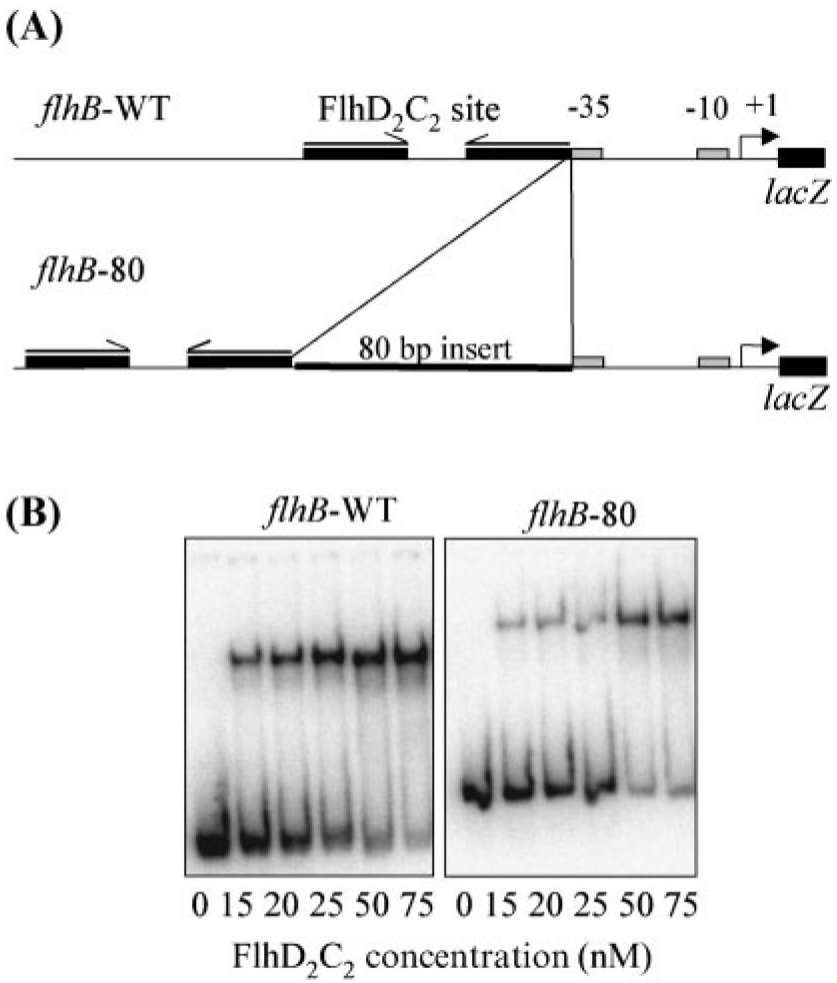


Fig. 1.

In vitro binding of $FlhD_2C_2$ to putative regulation promoters. (A) DNA-band shift. Radiolabelled nucleotide probe fragments were incubated with increasing concentrations of $FlhD_2C_2$ complex (nM) in the presence of a 1000-fold excess of non-specific competitor poly(dIdC), and analysed on 6% native polyacrylamide. (B) Ratios of bound : free $FlhD_2C_2$ probes observed in DNA-band shifts (mean of three experiments, error≤20%) plotted against $FlhD_2C_2$ protein concentration [(i) 0–70 nM; (ii) 0–175 nM].

Europe PMC Funders Author Manuscripts



(C)

	WT	*flhDC*
flhB-WT	1895	32
flhB-80	96	94

Fig. 2.

Uncoupling the $FlhD_2C_2$ binding site and promoter. (A) Representation of the *flhB*–wild-type promoter fusion region (*flhB*-WT) and the *flhB*-80 derivative in which an 80 bp transcriptionally inactive DNA sequence was inserted between the 3′ $FlhD_2C_2$ box (CCTTACTCAAACCATT) and the −35 sequence (TTGAAC). (B) DNA-band shift of *flhB*-WT and *flhB*-80 probes, performed as in Fig. 1. (C) Activity of *flhB*-WT and *flhB*-80 chromosomal promoter fusions (as in Fig. 2). The means of triplicate experiments, error≤10 %, are shown. WT, wild-type.

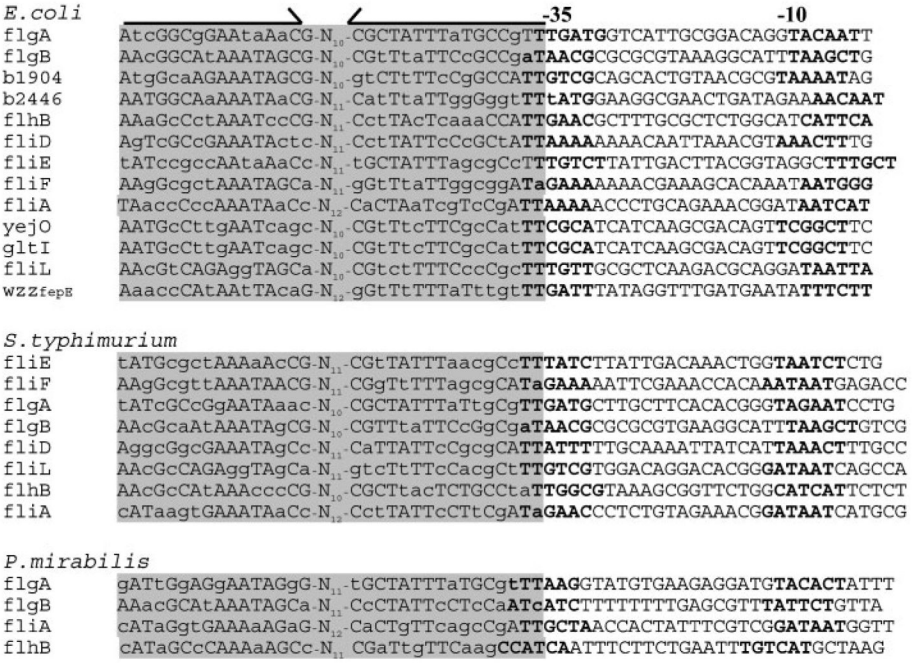


Fig. 3.

Alignment of $FlhD_2C_2$ bound promoter regions. $FlhD_2C_2$ binding sequences are highlighted in grey; known or putative −10 and −35 σ^{70} sequences promoters are shown in bold. The class II promoters from *E. coli* (including the *fliA*, *flhB* and *fliL* promoters defined by Liu & Matsumura, 1994) are aligned with the five $FlhD_2C_2$ bound and regulated non-flagellar promoters identified in this study, including the marginally regulated *yejO* promoter. The flagellar class II promoters shown from *S. typhimurium* and *P. mirabilis* have been experimentally defined by Kutsukake & Ide (1995) and Claret & Hughes (2002), respectively.

Europe PMC Funders Author Manuscripts

Table 1

Summary of putative $FlhD_2C_2$ binding sequences

Sequences (47) with a heterology index (HI) of 18 or less were identified by screening with the inverted repeat consensus (shown at the top in bold). Flagellar genes are in bold; the control *rcsA* and *argR* are included at the bottom. Nucleotides identical to or different from the consensus sequence are shown in upper- and lower-case type, respectively.

Gene	Function	HI	Potential $FlhD_2C_2$ binding site		
	Consensus sequence	0	**AATGGCAGAAATAGCG**	**N_{10-12}**	**CGCTATTTCTGCCATT**
flgB	Flagella biogenesis	6·1	AAcGGCAtAAATAGCG	N_{10}	CGtTtaTTCcGCCgaT
flgA	Flagella biogenesis	6·1	AAcGGCAtAAATAGCG	N_{10}	CGtTtaTTCcGCCgaT
b1904	Putative lipoprotein	10·8	AtgGcaAGAAATAGCG	N_{10}	gtCTtTTcCgGCCATT
b2446	Putative transcription factor	11·5	AATGGCAaAAATAaCG	N_{11}	CatTtaTTggGggtTT
flhB	Flagella biogenesis	11·9	AAaGcCctAAATccCG	N_{11}	CctTAcTcaaaCCATT
fliD	Flagella biogenesis	12·2	AgTcGCcGAAATActc	N_{11}	CctTATTcCcGCtATT
fliE	Flagella biogenesis	12·2	AAgGcgctAAATAGCa	N_{11}	gGtTtaTTggcggATa
fliF	Flagella biogenesis	12·2	AAgGcgctAAATAGCa	N_{11}	gGtTtaTTggcggATa
fliA	Flagella biogenesis	13·1	TAaccCccAAATAaCc	N_{12}	CaCTAaTcgTcCgATT
yejO	Putative periplasmic transport	13·7	AATGcCttgAATcagc	N_{10}	CGtTtcTTCgcCatTT
gltI	Putative periplasmic glutamate transport	13·7	AATGcCttgAATcagc	N_{10}	CGtTtcTTCgcCatTT
icc	cAMP metabolism	14·1	AgTccgctAAAaActG	N_{10}	tGCTtTTcgTcgCgTT
yegH	Putative membrane transport	14·6	tATGagAtAAATAGtG	N_{11}	CatTtTTgCccCgATT
fliL	Flagella biogenesis	14·7	AAcGtCAGAggTAGCa	N_{10}	CGtctTTTCccCgcTT
hns	Histone protein, global regulation	14·7	ttaGcCAGAAAagaCG	N_{10}	CcCTATTTacaCCgaT
wzz_{fepE}	Iron transport/O antigen chain length	14·9	AAaccCAtAAtTAcaG	N_{12}	gGtTtTTTaTttgtTT
uppS	Isoprenoid synthesis	15	cAgGGaAtAAAaAaCG	N_{12}	tGCTAcTcaacCacTT
hupB	Histone protein, global regulation	15·5	AATGcactAAtaAaaa	N_{11}	gGCTAaTTCgGgCtTg
cvpA	Membrane protein, colicin V production	15·9	tggcGCtGAAAaAata	N_{12}	gaCTtTTTaTGCCtTT
b3007	Putative membrane protein	16·1	ggcGcggtAAATAGCG	N_{10}	CGCTtcTTCgctggTc
yieG	Putative membrane transport protein	16·3	AATGcgAcAAtagGCG	N_{12}	actTtTTTaTtggATT
yieH	Putative membrane phosphatase	16·3	AATGcgAcAAtagGCG	N_{12}	actTtTTTaTtggATT

Gene	Function	HI	Potential FlhD$_2$C$_2$ binding site		
livK	Periplasmic leucine transporter	16·6	AAcGGCAcAAAgggGCG	N_{11}	tatcATTcagGggATT
yhhK	Putative acetyl transferase	16·6	AAcGGCAcAAAgggGCG	N_{11}	tatcATTcagGggATT
yjeH	Putative membrane transporter	16·6	AtTGcgAtgAAatGtG	N_{10}	gGgTtTTcaccCgATT
groES	Heat-shock protein	16·6	AtTGcgAtgAAatGtG	N_{10}	gGgTtTTcaccCgATT
yhiJ	Putative membrane protein	16·8	AcTGGCAGgAtTAaaa	N_{10}	CcgTATTcaTGgagTT
ygjO	Putative cytoplasmic enzyme	16·9	ttgccCctgAAaAGgG	N_{12}	CcCTtTTTCTctttTT
ygjP	Putative cytoplasmic enzyme	16·9	ttgccCctgAAaAGgG	N_{12}	CcCTtTTTCTctttTT
lhr	ATP dependent helicase	17	cATGcaAtgAtTcagG	N_{12}	gGCTATTcCacgCATc
chaC	Ca^{2+}/H^+ antiporter accessory protein	17	AATGGCAtAAAaAatc	N_{12}	gctTAgTTaaagCtaT
pmrD	Polymyxin gene regulator	17·1	AATGGCAGAAATgaga	N_{12}	gGtTAaTTaTtgCtga
secM	General secretion system	17·4	AgaGctAtAAAaAaCG	N_{12}	gttTtcTTaaGCacTT
ggt	Periplasmic reducing agent synthesis	17·4	tAcGcgAtgAAacGac	N_{12}	CGgcAcTcCTGCCgTT
yhhA	Putative outer-membrane protein	17·4	tAcGcgAtgAAacGac	N_{11}	CGgcAcTcCTGCCgTT
yjhF	Putative gluconate membrane permease	17·4	AgTcGgAtAAtTcagG	N_{10}	tttTATTTaTGCttTc
yggF	Putative bisphosphatase	17·5	AActctttAAAcAGCG	N_{11}	gcCTtTTTCgGCaAaa
yebE	Putative membrane protein	17·5	gATGcgcaAAtaAGCG	N_{11}	CGtTATTcaaagaAaT
yajF	Putative transcriptional regulator	17·7	cATGGggtAAATgcCa	N_{11}	tGtTtTTcCatCtcTT
yraQ	Putative membrane permease	17·7	AgcGGCgcAAAaAGCG	N_{10}	CGCTgTTTgctaCgaT
yajD	Putative cytoplasmic protein	17·9	AgTGGaAaAAgaAGgG	N_{12}	CGtcAaTTCTGCCgTa
yojI	Putative membane ABC transporter	17·9	cAacGgttgAATAcCG	N_{11}	CGCgATTcaccgCccc
ycbS	Putative outer-membrane fimbrial usher	17·9	AATaGgtGgAAaAGta	N_{11}	aatgATTTCccCCtTT
b0832	Putative membrane permease protein	17·9	AaatGCgtgAtTAcCc	N_{11}	CGgaAaTTCTGCttTT
yjiQ	Putative membrane protein	17·9	tAgaGCAtgAtaAaCG	N_{11}	CGtTggTcaTcCCgaT
yfiE	Putative LysR-type gene regulator	18	AAcaGagGAgATAaCa	N_{11}	CcCTtTTaagtgCtTT
yfiK	Putative LysE-family transporter	18	AAcaGagGAgATAaCa	N_{11}	CcCTtTTaagtgCtTT
Controls					
argR	Arginine synthesis gene regulator	18·7	AATttgAtAAAatcCc	N_{12}	CatTATTTCaGCCtTc
rcsA	Capsule synthesis gene regulator	18·8	AATGGCgGcAtTAatg	N_{12}	CctTtTTgCgaaCATT

Europe PMC Funders Author Manuscripts

Table 2

In vivo transcriptional activity of $FlhD_2C_2$ bound promoters

Values shown are peak *β*-galactosidase activity (× 10^2 Miller units) of single-copy *lacZ* fusions in *E. coli* wild-type and *flhDC* null mutant strains. Mean values from three experiments, error<15%, are shown. Restoration to wild-type activity in the MC1000 *flhDC* mutant by *flhDC* expressed in trans is denoted by '+' or '−'.

Gene	Wild-type	*flhDC*
flgB	75·0	0·6+
flhB	19·0	1·5+
fliA	29·7	4·6+
b1904	72·4	0·4+
b2446	3·5	0·3+
yejO	0·1	0·1 (+)
gltI	1·9	0·9+
icc	41·0	41·8−
yegH	27·0	29·0+
fliL	125·6	9·5+
hns	5·0	6·5−
wzz_{fepE}	4·0	0·6+
hupB	10·4	10·4−
argR	1·3	1·1−
rcsA	0·3	0·4−

Table 3
Characteristics of putative regulon promoters

Overlap of the $FlhD_2C_2$ by 1–2 bp with a putative or confirmed σ^{70} −35 box is shown as '+'; another location of the site between −12 and +190 bp is indicated by '−'. ?, Not known. The K_D (nM; the concentration required to achieve 50% binding) was calculated from the plots in Fig. 1. Promoter activation is expressed as the ratio of *in vivo lacZ* fusion activity in wild-type and *flhDC* mutant strains, derived from Table 2.

Gene	HI	−35 site	K_D	Activation
flgB	6·1	+	12	136
flhB	11·9	+	21	31
fliA	13·1	+	25	990
b1904	10·8	+	38	191
b2446	11·5	+	60	11
yejO	13·7	+	220	1
gltI	13·7	+	220	2
icc	14·1	?	167	1
yegH	14·6	?	240	1
fliL	14·7	+	43	209
hns	14·7	−	165	1
wzz_{fepE}	14·9	+	86	6
hupB	15·5	+	250	1
argR	18·7	−	>450	1
rcsA	18·8	−	350	1