

# Extensive alanine scanning reveals protein–protein and protein–DNA interaction surfaces in the global regulator FlhD from *Escherichia coli*

Andrés Campos\* and Philip Matsumura

Department of Microbiology and Immunology (M/C 790),  
College of Medicine, University of Illinois at Chicago,  
835 S. Wolcott Ave., MSB E-603, Chicago,  
IL 60612-7344, USA.

## Summary

FlhD and FlhC are the transcriptional activators of the flagellar regulon. The heterotetrameric complex formed by these two proteins activates the transcription of the class II flagellar genes. The flagellar regulon consists not only of flagellar genes, but also of the chemotactic genes and some receptor proteins. Recently, a connection between the flagellar regulon and some virulence genes has been found in some species. Furthermore, FlhD, but not FlhC, regulates another non-flagellar target. As a first attempt to understand the mechanism of the flagellar transcriptional activation by FlhD and FlhC, the structure of FlhD has been solved. In order to understand the mechanism of the action of FlhD when it regulates the flagellar genes, we conducted site-directed mutagenesis based on its three-dimensional structure. Six interaction surfaces in the FlhD dimer were mapped by alanine scanning mutagenesis. Two of them are surface clusters formed by residues His-2, Asp-28, Arg-35, Phe-34 and Asn-61 located at each side of the dimer core. The other four are located in the flexible arms of the dimer. The residues Ser-82, Arg-83, Val-84, His-91, Thr-92, Ile-94 and Leu-96 are located at this region. All these residues are involved in the FlhD/FlhC interaction with the exception of Ser-82, Arg-83 and Val-84. These three residues affect the DNA-binding ability of the complex. The three-dimensional topology of FlhD and the site-directed mutagenesis results support the hypothesis of FlhC as an allosteric effector that activates FlhD for the recognition of the DNA.

## Introduction

FlhD and FlhC proteins form the heterotetrameric complex

Accepted 18 October, 2000. \*For correspondence. E-mail acampos@uic.edu; Tel. (+1) 312 413 0288; Fax (+1) 312 996 6415.

(D<sub>2</sub>C<sub>2</sub>) which is the transcriptional activator of the flagellar regulon in Eubacteria (Liu and Matsumura, 1994; Macnab, 1996; Yanagihara *et al.*, 1999). Their genes (*flhD* and *flhC*) belong to the class I flagellar operon *flhD* (the master operon). The co-ordinated flagellar regulon expression is dependent on the FlhD/FlhC complex and it occurs hierarchically (Komeda, 1982, 1986; Kutsukake *et al.*, 1990; Kutsukake and Iino, 1994). The promoters of the class II flagellar genes are the targets of the FlhD/FlhC complex. *fliA*, which belongs to the class II operons and is under the direct control of the master operon, is an alternative sigma  $\sigma^{28}$  specific for the expression of the class III operons (Ohnishi *et al.*, 1990; Liu and Matsumura, 1995). In turn, FlgM (class III), which is an anti-sigma factor, negatively regulates the class III operons (Ohnishi *et al.*, 1992; Hughes *et al.*, 1993; Kutsukake *et al.*, 1994). The hierarchy of expression of the flagellar operons parallels their roles in the flagellar assembly pathway (Aizawa, 1996; Macnab, 1996).

This genetic system consists of genes that not only code for the flagellum apparatus but also for the proteins that comprise the chemotaxis system. Also, the other four genes, which belong to the flagellar regulon and are under the control of the flagellar master operon, code for receptor proteins: Tap, a dipeptide receptor (Manson *et al.*, 1986); Tar, aspartate/maltose/Co<sup>2+</sup>/Ni<sup>2+</sup> receptor (Clarke and Koshland, 1979; Hedblom and Adler, 1980); Trg, ribose receptor (Ordal and Adler, 1974a, b; Kondoh *et al.*, 1979; Harayana *et al.*, 1982); and Tsr, serine receptor (Clarke and Koshland, 1979; Hedblom and Adler, 1980; Wang and Koshland, 1980).

There are other species besides *Escherichia coli* and *Salmonella typhimurium* in which the flagellar regulon and its products have been studied. A connection between motility and virulence has been found, and there is a growing interest in the study of the control of this system in pathogenic bacteria (Ottemann and Miller, 1997). Some indications of this connection were obtained when an analogy between some flagellar genes and genes from the system for the secretion of virulence proteins of mammalian and plant pathogens (type III secretion system) was observed (Dreyfus *et al.*, 1993; Van Gijsegem *et al.*, 1995). Even the position of certain genes suggested a relationship between these two systems. For example, the *inv* gene from *Yersinia enterocolitica* is located between

flagellar operons (Fauconner *et al.*, 1997). Furthermore, a co-ordinate expression of the flagellar genes and virulence genes has been described in different bacteria species (Allison *et al.*, 1992; Gygi *et al.*, 1995; Kapatal *et al.*, 1996; Badger and Miller, 1998; Eichelberg and Galán, 2000). Later, the induction of phospholipase under the control of the *flhD* operon was demonstrated in *Serratia liquefaciens* (Givskov and Molin, 1992, 1993; Givskov *et al.*, 1995) and in *Y. enterocolitica* (*ypIA* gene) (Young *et al.*, 1999a; Schmiel *et al.*, 2000). Recently, it has been shown that certain virulence genes can be under the control of some flagellar proteins and, consequently, under FlhD/FlhC complex transcriptional regulation. FliZ, a protein whose gene belongs to the class II operon, *fliAZY* in *Salmonella enterica*, is required for the activation of the *hilA* invasion gene in this species (Lucas *et al.*, 2000).

FlhD, but not FlhC, negatively regulates the transcription of non-flagellar genes when the cells enter the stationary phase (Prüß and Matsumura, 1996; Prüß *et al.*, 1997). FlhD is also involved in several cellular processes at the transcriptional regulation level (B. M. Prüß *et al.*, submitted).

In order to improve our knowledge about this system, we have crystallized the FlhD protein, and its three-dimensional structure has been solved to 1.8 Å resolution (Campos *et al.*, 1998; 2001). The structure provided several insights about the possible function and mechanism of action of this global regulator. In the accompanying paper (Campos *et al.*, 2001), the crystal structure of this protein demonstrates that FlhD is a dimer that contains two flexible arms formed by the C-terminus of each molecule and located outside the main dimer core. A putative helix–turn–helix (HTH) DNA-binding motif was found at the beginning of each flexible domain. It was also proposed that FlhC is an allosteric protein, able to activate FlhD by forming the FlhD/FlhC complex (Campos *et al.*, 2001). Extensive alanine scanning mutagenesis directed to several surface residues has been carried out. This collection of alanine mutants develops some insights into the protein–protein interaction between FlhD and FlhC and also into the protein–DNA interaction between FlhD (when forming the FlhD/FlhC complex) and the DNA of one promoter belonging to the class II flagellar genes, *fliA*. Our results showed that the FlhD dimer possesses at least four FlhC-interacting surfaces and two DNA-interacting surfaces. Some of the residues that are important for the FlhD/FlhC interaction and all the residues involved in the FlhD–DNA interaction are located in the flexible arms of the dimer.

## Results

### *FlhD is conserved among bacteria species*

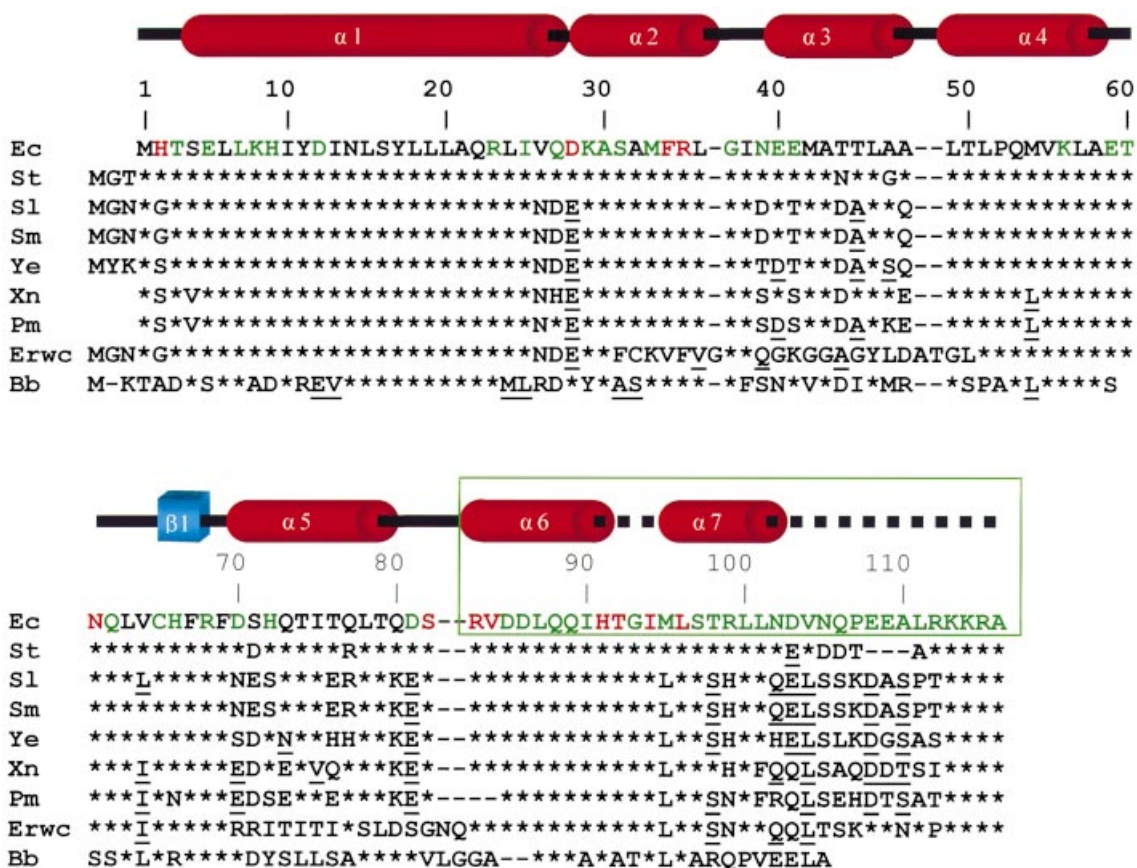
Given the importance of FlhD as a global regulator in the cell, we were interested in determining the relationship

between homologous sequences from other bacteria. The addition of new FlhD sequences from different species allowed us to carry out an alignment in order to determine the conservation of this protein. As expected for a global regulator, FlhD is a conserved protein among members of the Eubacteria (Fig. 1). FlhD from *E. coli* shows a 93.1% similarity (92.24% identity) to its homologues from *S. typhimurium*; 82% similarity (72.4–74.1% identity) to FlhD from *S. liquefaciens*, *Serratia marcescens*, *Y. enterocolitica*, and *Xenorhabdus nematophilus*; 78.4% similarity (69% identity) to FlhD from *Proteus mirabilis*; 63% similarity (56% identity) to FlhD from *Erwinia carotovora*; and 62% similarity (49% identity) to FlhD from *Bordetella bronchiseptica*. The first three residues of FlhD from *S. typhimurium*, *S. liquefaciens*, *S. marcescens*, *Y. enterocolitica*, *E. carotovora* and *B. bronchiseptica* were not included in the calculations, as it is believed that they are not part of the protein in *E. coli* (Soutourina *et al.*, 1999). The lowest similarity value was obtained when *E. coli* FlhD was compared with *E. carotovora* and *B. bronchiseptica* FlhDs. *B. bronchiseptica* is the only species that belongs to a different subdivision (beta subdivision; *Alcaligenaceae*) from all the other species (gamma subdivision; *Enterobacteriaceae*).

Given the high similarity value obtained when comparing the different sequences from these species, it seems that the secondary and tertiary structure of FlhD may be conserved among them. Certain important elements found in the crystal structure of FlhD from *E. coli* are conserved among the different species: Cys-65, responsible for the disulphide bond between the FlhD monomers (Campos *et al.*, 2001); Gly-93, highly conserved among the HTH DNA-binding motifs of different regulators.

### *Alanine scanning mutagenesis*

To date, nothing is known about the specific contacts between FlhD and FlhC and between the FlhD/FlhC complex and DNA. Alanine scanning mutagenesis was performed in combination with a genetic screening to identify critical residues involved in the FlhD/FlhC protein interaction and in the FlhD–DNA interaction. Site-directed mutagenesis was addressed to all charged residues on this protein located on the surface (Table 1A). Because the C-terminal structure was disordered and because of the presence of a putative HTH DNA-binding motif (Campos *et al.*, 2001), every residue in the C-terminus was also changed to alanine. We excluded from these assays residues 110 and 116, as they are already alanines. During our studies, two surface clusters that map at each side of the core dimer were found. As the initial alanine scanning was directed only to the charged residues of the FlhD N-terminal, we also carried out alanine scanning in some of the residues located at the



**Fig. 1.** Deduced amino acid sequence alignment of FlhD protein from different species. Sequence numbering is according to the *E. coli* FlhD residues (Bartlett *et al.*, 1988; Soutourina *et al.*, 1999). The secondary structure elements observed in the X-ray crystal structure of FlhD from *E. coli* are colour coded as follows: red,  $\alpha$ -helix; cyan,  $\beta$ -sheet; black, unclassified structure. Dotted lines in the secondary structure representation and elements enclosed in the green box do not display traceable electron density in the Se-Met-FlhD crystal (Campos *et al.*, 2001). Cys-65 mutation data were taken from the accompanying manuscript (Campos *et al.*, 2001). Colours in FlhD from *E. coli* residues represent the position of every residue that was changed to alanine. The phenotype of each mutation is colour coded as follows: green, swarming; red, partial swarming or non-swarming (see text). FlhD amino acid sequences are indicated as follows: Ec, *E. coli* (Bartlett *et al.*, 1988; Soutourina *et al.*, 1999); St, *S. typhimurium* (Yanagihara *et al.*, 1999; GenBank accession no. D43640); Sl, *S. liquefaciens* (Givskov *et al.*, 1995; GenBank accession no. 2126176); Sm, *S. marcescens* (GenBank accession no. AF077334); Ye, *Y. enterocolitica* (Young *et al.*, 1999b; GenBank accession no. AF081587); Xn, *X. nematophilus* (Givaudan and Lanois, 2000; GenBank accession no. AJ012828); Pm, *P. mirabilis* (Furness *et al.*, 1997; GenBank accession no. U96964); Erwc, *E. carotovora* (GenBank accession no. AF130387); and Bb, *B. bronchiseptica* (Akerley *et al.*, 1995; GenBank accession no. U17998). Identical residues to the FlhD *E. coli* sequence are represented as (\*); conserved residues are underlined, and gaps in the sequences are represented as (-). The alignment was performed manually.

periphery of these surface clusters (see below and Table 1B).

As mentioned in *Experimental procedures*, mutagenesis of *flhD* was carried out using the plasmid pXL27, which carries both *flhD* and *flhC*. After site-directed mutagenesis, a collection of *flhD* mutants was obtained (Tables 1A and B). In order to confirm that the desired changes in the DNA sequence were present in each candidate, we sequenced each *flhD* mutant gene in all the plasmids. Also, we sequenced the upstream region from the  $p_{T7}$  promoter of the expression plasmid to the start codon of *flhD* in order to confirm the integrity of this region.

The effect of each mutation on FlhD flagellar function was studied by transforming each mutant plasmid onto

strain YK4131 (*flhD*<sup>-</sup>). As FlhD, together with FlhC, is required for the transcription of the class II flagellar genes, the functionality of FlhD was evaluated by its ability to complement this *flhD*<sup>-</sup> strain. The bacteria harbouring each mutant plasmid were grown on swarming plates in order to confirm their individual phenotype (data not shown). By testing each single mutant, we identified 10 mutations (out of 52) that are somehow compromised in the function of FlhD. We found three different phenotypic displays: swarming cells in which the mutation does not affect the flagellar function of FlhD; partial swarming cells in which the FlhD flagellar activity is reduced but not abolished; and non-swarming cells in which flagellar function could not be observed (Table 1A and Fig. 1).

**Table 1A.** DNA sequence changes of *flhD'* alanine scanning mutants (first screening).

Plasmid	Base substitution(s) <sup>a</sup>	Amino acid change(s) <sup>b</sup>	Phenotype in <i>E. coli</i> YK4131
pACH2A	<sup>4</sup> CAT-GCT	H2A	Partial swarming
pACE5A	<sup>13</sup> GAG-GCG	E5A	Swarming
pACK8A	<sup>22</sup> AAA-GCA	K8A	Swarming
pACH9A	<sup>25</sup> CAC-GCC	H9A	Swarming
pACD12A	<sup>34</sup> GAC-GCC	D12A	Swarming
pACR23A	<sup>67</sup> CGT-GCT	R23A	Swarming
pACD28A	<sup>82</sup> CGT-GCC	D28A	Partial swarming
pACK29A	<sup>85</sup> AAA-GCA	K29A	Swarming
pACR35A	<sup>103</sup> CGT-GCT	R35A	Partial swarming
pACE40A	<sup>118</sup> GAA-GCA	E40A	Swarming
pACE41A	<sup>121</sup> GAA-GCA	E41A	Swarming
pACK56A	<sup>166</sup> AAG-GCG	K56A	Swarming
pACE59A	<sup>175</sup> GAA-GCA	E59A	Swarming
pACC65A	<sup>193</sup> TGT-GCT	C65A	Swarming
pACH66A	<sup>196</sup> CAC-GCC	H66A	Swarming
pACR68A	<sup>202</sup> CGT-GCT	R68A	Swarming
pACD70A	<sup>208</sup> GAC-GCC	D70A	Swarming
pACH72A	<sup>214</sup> CAC-GCC	H72A	Swarming
pACD81A	<sup>241</sup> GAT-GCT	D81A	Swarming
pACS82A	<sup>244</sup> TCC-GCC	S82A	Partial swarming
pACR83A	<sup>247</sup> CGC-GCC	R83A	Partial swarming
pACV84A	<sup>250</sup> GTT-GCT	V84A	Partial swarming
pACD85A	<sup>253</sup> GAC-GCC	D85A	Swarming
pACD86A	<sup>256</sup> GAT-GCT	D86A	Swarming
pACL87A	<sup>259</sup> CTC-GCC	L87A	Swarming
pACQ88A	<sup>262</sup> CAG-GCG	Q88A	Swarming
pACQ89A	<sup>265</sup> CAG-GCG	Q89A	Swarming
pACI90A	<sup>268</sup> ATT-GCT	I90A	Swarming
pACH 91A	<sup>271</sup> CAT-GCT	H91A	Partial swarming
pACT92A	<sup>274</sup> ACC-GCC	T92A	Non-swarming
pACG93A	<sup>276</sup> GGC-GCC	G93A	Swarming
pACI94A	<sup>280</sup> ATC-GCC	I94A	Non-swarming
pACM95A	<sup>283</sup> ATG-GCG	M95A	Swarming
pACL96A	<sup>286</sup> CTC-GCC	L96A	Partial swarming
pACS97A	<sup>289</sup> TCA-GCA	S97A	Swarming
pACT98A	<sup>292</sup> ACA-GCA	T98A	Swarming
pACR99A	<sup>295</sup> CGC-GCC	R99A	Swarming
pACL100A	<sup>298</sup> TTG-GCG	L100A	Swarming
pACL101A	<sup>301</sup> CTG-GCG	L101A	Swarming
pACN102A	<sup>304</sup> AAT-GCT	N102A	Swarming
pACD103A	<sup>307</sup> GAT-GCT	D103A	Swarming
pACV104A	<sup>310</sup> GTT-GCT	V104A	Swarming
pACN105A	<sup>313</sup> AAT-GCT	N105A	Swarming
pACQ106A	<sup>316</sup> CAG-GCG	Q106A	Swarming
pACP107A	<sup>319</sup> CCT-GCT	P107A	Swarming
pACE108A	<sup>322</sup> GAA-GCA	E108A	Swarming
pACE109A	<sup>325</sup> GAA-GCA	E109A	Swarming
pACL111A	<sup>331</sup> CTG-GCG	L111A	Swarming
pACR112A	<sup>334</sup> CGC-GCC	R112A	Swarming
pACK113A	<sup>337</sup> AAG-GCG	K113A	Swarming
pACK114A	<sup>340</sup> AAA-GCA	K114A	Swarming
pACR115A	<sup>343</sup> AGG-GCG	R115A	Swarming

All plasmids in this table were obtained by site-directed mutagenesis carried out on pXL27 (Table 3).

**a.** The number indicates the position of the first base of the codon containing the base substitution. Nucleotide numbering is as for the *flhD*<sup>+</sup> sequence, with number 1 corresponding to the first nucleotide of the *flhD*<sup>+</sup> coding sequence, when the first ATG codon is considered as the start codon (base 274 in Bartlett *et al.*, 1988).

**b.** Amino acids in FlhD' have the same position number that they have in the intact FlhD protein. pACC65A data were taken from the accompanying manuscript (Campos *et al.*, 2001).

Some differences were observed in the mutants with partial swarming. The halo diameter was not the same in all cases, and the time of incubation required to observe partial swarming was different. Thus, the phenotypic analyses were carried out for at least 11 h.

Our data show that most of the charged surface residues in this protein can be changed to alanine with no observable effect on function. As mentioned before, only 10 out of the 52 tested residues have an effect on the FlhD function. The transcriptional activation of flagellar genes by

**Table 1B.** DNA sequence changes of *flhD'* alanine scanning mutants around the interface located at H2, D28 and R35.

Plasmid	Base substitution(s) <sup>a</sup>	Amino acid change(s) <sup>b</sup>	Phenotype in <i>E. coli</i> YK4131
pACT3A	<sup>7</sup> ACC→GCC	T3A	Swarming
pACL7A	<sup>19</sup> CTG→GCG	L7A	Swarming
pACI25A	<sup>73</sup> ATT→GCT	I25A	Swarming
pACQ27A	<sup>79</sup> CAG→GCG	Q27A	Swarming
pACA30S	<sup>88</sup> GCG→TCG	A30S	Swarming
pACS31A	<sup>91</sup> TCC→GCC	S31A	Swarming
pACM33A	<sup>97</sup> ATG→GCG	M33A	Swarming
pACF34A	<sup>100</sup> TTT→GCT	F34A	Partial swarming
pACG37A	<sup>109</sup> GGC→GCC	G37A	Swarming
pACN39A	<sup>115</sup> AAT→GCT	N39A	Swarming
pACT60A	<sup>178</sup> ACC→GCC	T60A	Swarming
pACN61A	<sup>181</sup> AAT→GCT	N61A	Partial swarming
pACQ62A	<sup>184</sup> CAA→GCA	Q62A	Swarming

All plasmids in this table were obtained by site-directed mutagenesis carried out on pXL27 (Table 3).

a. The number indicates the position of the first base of the codon containing the base substitution. Nucleotide numbering is as for the *flhD*<sup>+</sup> sequence, with number 1 corresponding to the first nucleotide of the *flhD*<sup>+</sup> coding sequence, when the first ATG codon is considered as the start codon (base 274 in Bartlett *et al.*, 1988).

b. Amino acids in FlhD' have the same position number that they have in the intact FlhD protein.

FlhD is decreased or abolished when His-2, Asp-28, Arg-35, Ser-82, Arg-83, Val-84, His-91, Thr-92, Ile-94 and Leu-96 are changed to alanine (Figs 1 and 2, and Table 1A). Hence, these are critical residues that affect the function of FlhD in two possible ways: by altering the binding of FlhD with FlhC or by altering the FlhD DNA binding with the promoter sequences of the class II flagellar genes. However, we cannot exclude the possibility of a structural change in the FlhD dimer structure.

We have found that some of these residues that affect the flagellar function of FlhD are located in the N-terminus domain. As FlhD is a dimer, when these residues are mapped in the crystal structure of FlhD, they defined two cluster surfaces in the main core dimer. These surface clusters were defined by at least three residues (His-2, Asp-28 and Arg-35) that map together at each side of the core dimer (Fig. 2). Each monomer of FlhD is twisted together in the three-dimensional structure (Campos *et al.*, 2001), and these surface clusters are not formed with residues from the same chain. For example, one of the cluster surfaces is formed by His-2<sup>A</sup>, Asp-28<sup>B</sup> and Arg-35<sup>B</sup>, where <sup>A</sup> and <sup>B</sup> indicate the monomer of the FlhD dimer to which the residue belongs (Fig. 2). The alanine scanning directed to the charged amino acids did not include all the residues around these surface cluster. As there was the possibility that other non-charged residues were part of these surface clusters, additional alanine scanning mutagenesis was directed to the residues located at the edges of the surface clusters in the dimer core (Table 1B). Of 13 residues from this second alanine screening, only two alanine changes were defective in FlhD activity when tested on swarming plates. Phe-34 and Asn-61 displayed a partial swarming phenotype in the YK4131 (*flhD*<sup>-</sup>) strain (Table 1B). These additional results showed that the surface clusters located at both sides of the FlhD dimer core are formed by the residues

His-2, Asp-28, Phe-34, Arg-35 and Asn-61. The following is the combination at one of the cluster surfaces: His-2<sup>A</sup>, Asp-28<sup>B</sup>, Phe-34<sup>B</sup>, Arg-35<sup>B</sup> and Asn-61<sup>A</sup>; and the following is the combination of the other cluster surface: His-2<sup>B</sup>, Asp-28<sup>A</sup>, Phe-34<sup>A</sup>, Arg-35<sup>A</sup> and Asn-61<sup>B</sup> (Fig. 2).

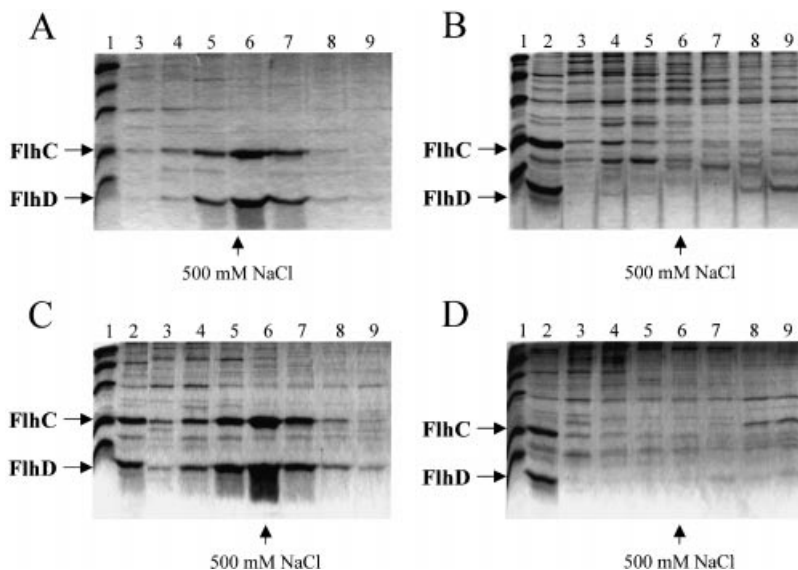
The rest of the interaction surfaces found in the first alanine scanning experiment were located at the C-terminal flexible domain of each FlhD monomer. Of 10 critical residues identified in the first alanine scanning experiments (Table 1A and Fig. 1), seven (Ser-82, Arg-83, Val-84, His-91, Thr-92, Ile-94 and Leu-96) mapped at these domains (Fig. 2). Furthermore, all these seven residues are located in the region in which the putative HTH motif was found (Fig. 1) (Campos *et al.*, 2001).

The FlhD dimer core holds two FlhD/FlhC interaction surfaces, and each C-terminus domain contains interaction surfaces for both FlhC and DNA.

As the flexible domain displays characteristics of a DNA-binding motif, it was possible that the two interacting surfaces formed by His-2, Asp-28, Arg-35, Phe-34 and Asn-61 at the dimer core are needed for the interaction with FlhC. The current structure of FlhD agrees with the mutation data, which suggest that the C-terminal flexible domain (from Ser-82 to Leu-96) is critical for the interaction between FlhD and FlhC or FlhD and DNA or maybe both. Furthermore, given the fact that the C-terminal domain is highly flexible, holds a putative HTH motif and the distance between one domain and the corresponding domain in the other chain fits with the distance between two adjacent major grooves of the DNA, it was likely to be involved in binding to the promoters of the flagellar class II operons (Campos *et al.*, 2001).

In order to understand how the function of FlhD is affected in each of the alanine mutants obtained, a FlhD/FlhC complex formation assay was carried out. It is known that the FlhD/FlhC complex, but not FlhD by itself, has





**Fig. 3.** FlhD/FlhC complex formation assay.

Protein fractions eluted from a heparin column, analysed on SDS–20% polyacrylamide gels and visualized by Coomassie brilliant blue staining. Fractions were obtained by applying a NaCl gradient through the column.

A. Cell lysate of overexpressed native FlhD/FlhC complex (from pXL27) was used as a positive control.

B. Cell lysate of overexpressed FlhD alone (from pXL25) was used as a negative control.

C and D. Overexpression of two representative FlhD mutants together with native FlhC, FlhDS82A and FlhDH91A respectively. Lines: 1,  $M_r$  protein markers (BSA, 68 kDa; ovalbumin, 45 kDa; carbonic anhydrase, 29 kDa; trypsin inhibitor, 20 kDa; and lysozyme, 14 kDa); 2, clear lysate; 3–9, equal volumes of fractions 17–23 taken during elution with a gradient of NaCl. When the FlhD/FlhC complex is present, both proteins are co-eluted with 0.5 M NaCl (fraction 20, line 6). The fraction containing 0.5 M NaCl is indicated.

binding affinity to heparin (Liu and Matsumura, 1994) (Fig. 3). This interaction can be disrupted by the addition of a high salt concentration (0.5 M NaCl). In fact, this characteristic of the FlhD/FlhC complex has been exploited to purify it to homogeneity from a cell lysate. Thus, this property was used to test the ability of each FlhD mutant to interact with FlhC and form the heterotetrameric complex that is responsible for the transcriptional activation of the flagellar genes. The binding to heparin is suggestive evidence that FlhD and FlhC form the complex. As all the pAC plasmid series also carry the *flhC* gene, each *flhD* mutant gene was overexpressed in conjunction with the *flhC* gene in strain MC1000*flhD::kan* harbouring the plasmid pGp1-2<sup>cmI</sup>. In fact, the co-expression of both genes in the same cell is required to obtain the FlhD/FlhC complex. Each cell lysate was loaded directly onto a HiTrap heparin column cartridge and eluted with a linear gradient of NaCl as described in *Experimental procedures*. The fractions containing 0.5 M NaCl were checked by electrophoresis on an SDS–PAGE gel in order to ensure the presence or absence of the FlhD/FlhC complex. Figure 3 shows representative gel analyses of the FlhD/FlhC complex formation. When present, the complex proteins were eluted when 0.5 M NaCl was applied to the column (Fig. 3A and C). As mentioned before, the FlhD protein does not bind to heparin if FlhC is absent (Fig. 3B). FlhC alone was not tested, as it precipitates when FlhD is not present. Thus, the presence of both proteins in the 0.5 M NaCl fraction was interpreted as an indication that the mutation does not affect FlhD/FlhC complex formation. On the other hand, the absence of both proteins in that fraction indicated that the mutation affects the protein–protein interaction between FlhD and FlhC. Figure 3C and D show two representative complex

formation analyses. FlhDS82A was able to form the FlhDS82A/FlhC complex (Fig. 3C). FlhDH91A did not form the FlhD/FlhC complex (Fig. 3D). The same procedure was applied to each of the other mutants, and the results are summarized in Table 2.

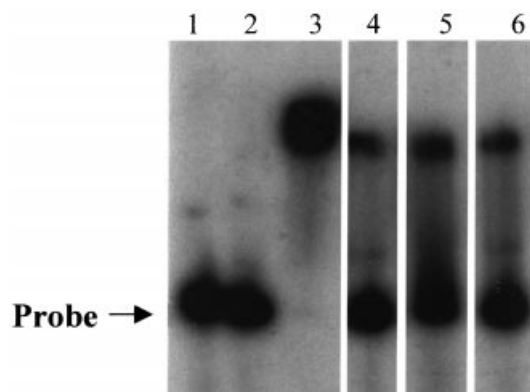
Of the 11 mutants tested (for some unknown reason, it was not possible to overexpress the FlhD/FlhC complex from the plasmid pACH2A), only three of them were able to form the FlhD/FlhC complex (Table 2), showing that the C-terminal flexible domain contains residues important for the interaction with FlhD and FlhC and also probably with FlhD and DNA. Mutations to alanine in residues Asp-28, Arg-35, Phe-34, Asn-61, His-91, Thr-92, Ile-94 and Leu-96 affected the flagellar function of FlhD at the level of complex formation with FlhC (protein–protein interaction).

**Table 2.** FlhD/FlhC complex formation assay for the mutants that affect the flagellar function of FlhD.

Mutant	Phenotype	FlhD/FlhC complex formation
FlhDH2A	Partial swarming	NA <sup>a</sup>
FlhDD28A	Partial swarming	No
FlhDF34A	Partial swarming	No
FlhDR35A	Partial swarming	No
FlhDN61A	Partial swarming	No
FlhDS82A	Partial swarming	Yes
FlhDR83A	Partial swarming	Yes
FlhDV84A	Partial swarming <sup>b</sup>	Yes
FlhDH91A	Partial swarming	No
FlhDT92A	Non-swarming	No
FlhDI94A	Non-swarming	No
FlhDL96A	Partial swarming <sup>b</sup>	No

a. NA, not applicable. Overexpression of the FlhD/FlhC complex was not obtained.

b. Non-swarming phenotype in this mutant was observed in the first hours. Longer incubations (more than 6 h) resulted in the partial swarming phenotype.



**Fig. 4.** Mobility shift analysis of purified protein (500 ng of each) performed on the *fliA* promoter (−185 to +26). Lines: 1, free probe; 2, FlhD native protein from MC1000 *fliD::kan/pGp1-2<sup>cmI</sup>/pXL25*; 3, FlhD–FlhC native protein from MC1000 *fliD::kan/pGp1-2<sup>cmI</sup>/pXL27*; 4, FlhDS82A/FlhC complex from MC1000 *fliD::kan/pGp1-2<sup>cmI</sup>/pACS82A*; 5, FlhDR83A/FlhC complex from MC1000 *fliD::kan/pGp1-2<sup>cmI</sup>/pACR83A*; and 6, FlhDV84A/FlhC complex from MC1000 *fliD::kan/pGp1-2<sup>cmI</sup>/pACV84A*.

The partial swarming that the mutants D28A, F34A, R35A, N61A, H91 and L96 exhibited in our experiments (Table 2) suggests that the interaction of FlhD and FlhC is affected (diminished) but not abolished. If this is true, the FlhD/FlhC interactions become weak when these residues are changed to alanine, and the protein–protein interaction becomes unstable. The quantity of complex formed under these conditions may be lower than the quantity that our system can detect (heparin column), but high enough to complement the *fliD*<sup>−</sup> mutant strain partially. As it was not possible to obtain the complex from the heparin column in some of the mutants that display partial complementation, their DNA-binding ability was not tested.

The FlhD/FlhC complex formation for FlhDS82A, FlhDR83A and FlhDV84A was further confirmed by passing each heparin eluate through a Superdex 200 HR column in a fast protein liquid chromatography (FPLC) system. In these three cases, the proteins (FlhD and FlhC) were co-eluted from the column at around 68 kDa, which is the approximate size of the oligomeric complex FlhD/FlhC (Liu and Matsumura, 1994). The presence of the FlhD/FlhC complex when FlhD was mutated to alanine in these residues was confirmed. As changes of Ser-82, Arg-83 and Val-84 to alanine do not affect FlhD/FlhC complex formation, the other possibility was that these changes affect the DNA binding between FlhD and the DNA. In order to confirm this possibility, DNA mobility shift analyses were carried out using the promoter region of *fliA*, a gene of a class II flagellar operon, as a template.

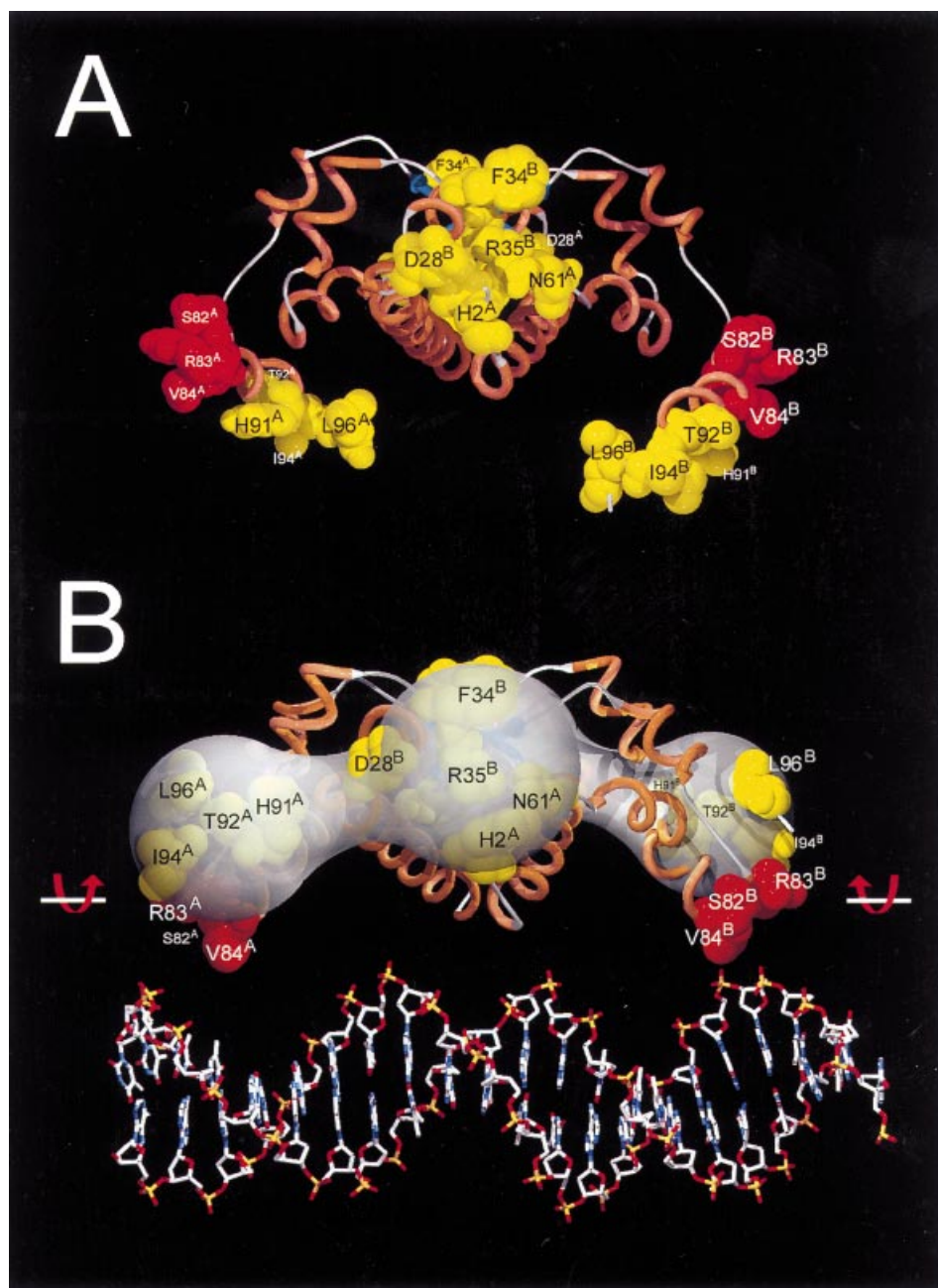
As the FlhDS82A, FlhDR83A and FlhDV84A proteins have the ability to interact with FlhC and form the FlhD/FlhC heterotetrameric complex, the DNA mobility shift

analyses were carried out with these purified complexes. *fliA* belongs to the class II flagellar genes, and its product, FlhA, is an alternative sigma factor specific for the class III flagellar genes. Thus, the template used in the experiments was the promoter region of the *fliA* operon from *E. coli* (Ohnishi *et al.*, 1990; Liu and Matsumura, 1994; 1995). FlhD alone (from pXL25) does not bind to the DNA and was used as a negative control, and the native FlhD/FlhC complex (from pXL27) was used as a positive control. The results indicated that the complex formed with these three mutants showed lower affinity than the native FlhD/FlhC complex. Figure 4 shows a representative mobility shift analysis. Complexes FlhD/FlhC, FlhDS82A/FlhC, FlhDR83A/FlhC and FlhDV84A/FlhC (500 ng of each) were tested for their ability to bind to the DNA promoter region of the *fliA* operon (−185 to +26). The native complex completely shifted the probe (1 ng) at that concentration, whereas only a partial shift was observed when any of the three FlhD mutants was tested at the same concentration (Fig. 4). In other experiments, even a high concentration of FlhD mutant/FlhC complex (1.75 μg), the DNA template (1 ng) was not completely shifted, whereas 400 ng of native complex was enough to bind completely to the same quantity of DNA template (data not shown). This result was expected, as these mutants displayed a partial phenotype when complemented with the YK4131 (*fliD*<sup>−</sup>) strain (Table 2). What is relevant here is that the C-terminal flexible domain of FlhD is not only important for the protein–protein interaction between FlhD and FlhC, but also for the DNA interaction between the FlhD/FlhC complex and the DNA *fliA* promoter sequence.

Our results indicate that the FlhD dimer holds two interface surfaces important for DNA binding located at the putative HTH motif and four interface surfaces for FlhD/FlhC interaction: two of them located at the putative HTH motif region and the other two in surface clusters at both sides of the dimer core.

## Discussion

The expanding list of DNA sequences of bacteria motility genes, including those from the flagellar master operon, reflects an increasing interest in the understanding of this system. As expected from a global regulator, FlhD is a conserved protein among Eubacteria. In our mutagenesis results indicating that changing any one of the last residues of FlhD (from 97 to 116) to alanine did not result in a loss of function, the C-terminal is the less conserved region of this protein (Fig. 1 and Table 1A). Furthermore, if this region is excluded from the calculations, the identity values when compared with *E. coli* are increased by 4% in *S. typhimurium* and by 8.1–8.4% among the other species, excluding *E. carotovora* (1.7%) and



**Fig. 5.** Hypothetical allosteric activation of FlhD by FlhC.

A. In order to visualize the partial C-terminus at both sides of the FlhD dimer, the current crystal structure of the FlhD dimer was modelled with two chain As from the crystal structure (Campos *et al.*, 2001).

B. Hypothetical representation of the FlhD/FlhC and FlhD-DNA interaction. Our results support the hypothesis that each molecule of FlhC interacts with two different regions of the FlhD dimer. As the structure of FlhC is not known, only the putative interfaces with FlhD are displayed as a blob in transparent models. Each blob represents part of each molecule of FlhC: one located at the front and the other behind the FlhD dimer model. Each molecule of FlhC may interact with the FlhD core dimer while it holds one of the C-terminal flexible domains. This allosteric interaction should generate a rotation on the flexible arms (red arrows) of the FlhD dimer, so that the residues Ser-82, Arg-83, and Val-84 can face the DNA. See text for further details. Notice that the rotation of the flexible domains can be opposite to that shown in the figure. Residues involved in the FlhD/FlhC interaction (yellow) and FlhD-DNA interaction (red) are displayed as a CPK model. The FlhD dimer and DNA models were generated with Swiss-pdb viewer (Guex and Peitsch, 1997; software available from URL: <http://www.expasy.ch/spdbv/>). Additional elements (FlhC blob representation) were added and rendered with POV-Ray 3.1 (software available from URL: <http://www.povray.org/>).

*B. bronchiseptica* (which, in fact, decreases by 1.6%). FlhD from *E. carotovora* and *B. bronchiseptica* are the most divergent sequences of the group. In fact, FlhD from *B. bronchiseptica* does not present the last residues found in the other species, again suggesting that this region is not essential. Sequence conservation reflects the functional and structural conservation of FlhD. The high degree of conservation of FlhD suggests that this protein should play similar roles in other bacteria. Furthermore, *flhDC* from different species is able to complement mutations in *E. coli*: *P. mirabilis* (Furnerss *et al.*, 1997), *S. liquefaciens* (Givskov *et al.*, 1995) and, more recently, *Y. enterocolitica* (Young *et al.*, 1999b). Thus, we expect that FlhD proteins from other species could also be involved in non-flagellar regulation as has been found in *E. coli*, i.e. the *cadBA* operon (Prüß *et al.*, 1997).

Previous results have shown that the Cys-65<sup>A</sup>:Cys-65<sup>B</sup> disulphide bond is not required for the flagellar function of FlhD (Campos *et al.*, 2001). However, Cys-65 is conserved among the different species, suggesting that the disulphide bridge could be present in all the FlhD dimer proteins. This fact reinforces the idea that the disulphide bridge may play another role in an unknown function of FlhD (Campos *et al.*, 2001). One of the most highly conserved residues in the turn of the HTH motif is a small amino acid, either glycine (most common) or alanine (Pabo and Sauer, 1984; 1992; Branden and Tooze, 1991). Gly-93 is also present in almost all the FlhD proteins from different species. However, its alanine substitution does not seem to affect the conformation of the putative HTH.

In order to study the relationship between protein structure and function on FlhD, we performed an alanine scanning based on its three-dimensional topology. Clues to the FlhD/FlhC interface surface were found in the crystal structure (Campos *et al.*, 2001) and combined with the alanine scanning and a genetic screen provided insight into the biological function of FlhD. We have identified critical residues for the function of this protein. The structural data are consistent with our genetic analysis that identifies three domains in the FlhD dimer. The domain 1 is formed by the N-terminus from each FlhD monomer. This domain consists of residues from Met-1 to Asp-81 chain A and from Thr-3 to Asp-81 chain B, which forms a compact core dimer. This domain contains the  $\alpha$ -helices 1–5 of both chains, the  $\beta$ -strand 1 of both chains, the Cys-65:Cys-65' disulphide bond and two of the four interaction surfaces for FlhC (Fig. 2). Domains 2 and 3 are formed by flexible arms that extend from domain 1. Domain 2 is composed of residues Ser-82 to Ala-116 of chain A, and domain 3 is composed of residues Ser-82 to Ala-116 of chain B. Domains 2 and 3 contain a putative HTH motif ( $\alpha$ -helices 6 and 7) that interacts with both FlhC and the DNA.

We identified at least six interaction surfaces on FlhD. Two surface clusters were located at opposite sides of the main core, and the other four mapped at the C-terminal flexible domain (Fig. 2). The putative HTH motifs at the flexible arms suggested that they are functional domains and that FlhD is part of the complex with DNA-binding properties (Campos *et al.*, 2001). This premise was further supported by the genetic analysis, which showed that Ser-82, Arg-83, Val-84, His-91, Thr-92, Ile-94 and Leu-96 are critical residues for FlhD function. Our data suggest that the surface clusters and part of the C-terminus in the dimer are important for the FlhD/FlhC interaction. On the other hand and as expected from the crystallographic studies, the FlhD C-terminus is also important for the FlhD/FlhC–DNA interaction, as shown by the mobility shift analyses (see below).

Residues His-2<sup>A</sup>, Asp-28<sup>B</sup>, Phe-34<sup>B</sup>, Arg-35<sup>B</sup>, Asn-61<sup>A</sup> and His-2<sup>B</sup>, Asp-28<sup>A</sup>, Phe-34<sup>A</sup>, Arg-35<sup>A</sup> and Asn-61<sup>B</sup> form two surface clusters in the core dimer (Fig. 2), and their topology is in agreement with the observation that the heterocomplexes present interfaces that are more planar than the homodimers (Jones and Thornton, 1996). Complex formation assay showed that these cluster surfaces form part of the FlhD–FlhC interface. In fact, when the electrostatic potential is calculated on the surface, the position of these residues involved in the FlhD/FlhC interaction correlates with the position of the neutral and positive patches located at both sides of the core dimer. Compare Fig. 2B of this work with Fig. 5E of the accompanying paper (Campos *et al.*, 2001). However, we cannot exclude the possibility that some of the mutations can affect FlhD function by changing the molecular structure, i.e. Arg-35 is not completely exposed in the surface.

As His-2 is part of these interface clusters, it is possible that Met-1 is also included in the interface. Unfortunately, only the Met-1<sup>A</sup> was observed with poor resolution in the crystal structure, and its relative position is not known (Fig. 1 of Campos *et al.*, 2001). As it is not possible to obtain a Met-1 mutant, further studies directed at solving the crystal structure of the FlhD/FlhC complex should answer this concern. The fact that the surface clusters in the main core are formed by residues from both chains supports the hypothesis that FlhD is an obligate dimer (Campos *et al.*, 2001). Even if FlhD can be separated in monomers, these surface clusters will not be present in the single polypeptide. As shown by the complex formation analyses, His-91, Thr-92, Ile-94 and Leu-96, located at the flexible domain, are also involved in the FlhD/FlhC interaction.

It was surprising that these residues are located in the putative HTH. This is an unusual function for an HTH motif. As FlhD is a dimer, all the FlhD/FlhC interface surfaces we found are represented twice (Fig. 2).

**Table 3.** *Escherichia coli* strains and plasmid genotypes.

<i>E. coli</i> strains	Genotype	Reference
YK4131	<i>flhD</i> derivative of YK410 [F <sup>-</sup> , <i>araD139</i> , $\Delta$ <i>lac(U169)</i> , <i>rpsL</i> , <i>thi</i> , <i>pyrC46</i> , <i>nalA</i> , <i>thyA</i> , <i>his</i> ]	Komeda <i>et al.</i> (1980)
MC1000 <i>flhD::kan</i>	<i>flhD flhC</i> derivative of MC1000 [F <sup>-</sup> , <i>araD139</i> , $\Delta$ ( <i>araAB</i> , <i>leu</i> )7697, $\Delta$ ( <i>lacX74</i> ), <i>galK</i> , <i>strA</i> ]	Malakooti <i>et al.</i> (1989)
Epicuran Coli XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lac<sup>R</sup>Z</i> $\Delta$ M15 Tn10 (Tet <sup>r</sup> )]	Stratagene
Plasmids		
pGp1-2 <sup>cmI</sup>	T7 RNA polymerase, Cml <sup>R</sup>	Rendone (1992)
pXL21	pAN518- <i>fliA</i> (-182 to +26)	Unpublished data
pXL25	pT7-7- <i>flhD</i>	Liu and Matsumura (1994)
pXL27	pT7-7- <i>flhD/C</i>	Liu and Matsumura (1994)
pAC series	pXL27 derivatives (see Tables 1A and B)	This study

Before the crystal structure of FlhD was obtained, nothing was known about the FlhD/FlhC–DNA interaction. The mere presence of the putative HTH motifs in the FlhD dimer suggested that FlhD is the factor in the FlhD/FlhC complex that has the ability to bind to the DNA. Our mutagenesis and the mobility shift DNA analyses are consistent with this idea. Mutants FlhDS82A, FlhDR83A and FlhDV84A are able to form the FlhD/FlhC complex, but the affinity for complex–DNA binding was reduced as shown by the mobility shift experiments. This is in agreement with our phenotypic analyses, which showed a reduced activity in the mutants (Tables 1A and 2). Ser-82, Arg-83 and Val-84, located in the first helix of each putative HTH motif, form a DNA interaction surface. The fact that only three residues in the putative HTH motif affect DNA binding was not expected. Alanine mutations in this region may be tolerated because alanine is a small amino acid with no chain beyond the  $\beta$ -carbon to interact sterically with the DNA. The FlhD component of the FlhD/FlhC complex appears to interact with the DNA, but it is not known whether FlhC also interacts with the DNA.

FlhD does not have DNA-binding activity unless it is complexed with FlhC (Liu and Matsumura, 1994). Specific DNA binding only occurs when the FlhD/FlhC complex is formed. Based on the FlhD crystal structure and on the results of genetic analyses, we believe that the DNA specificity of FlhD is given by FlhC. FlhC is an allosteric activator able to change the conformation of FlhD by holding the two C-terminal flexible domains in a suitable position, which permits interaction with the DNA (Fig. 5). In our working model, each molecule of FlhC interacts with one of the surface clusters in the main core, while another region of the same molecule holds one of the flexible arms of FlhD dimer by interacting with the residues His-91, Thr-92, Ile-94 and Leu-96 of one of the flexible arms (Fig. 5). We do not know whether FlhC is also a dimer, but two molecules of FlhC are needed to maintain both flexible arms rigid in order to confer the DNA-binding specificity. This allosteric interaction could stabilize the flexible arms of the FlhD dimer, so that the residues Ser-82, Arg-83 and Val-84 can face the DNA (Fig. 5).

In this work, we have shown that most of the charged residues in FlhD can be changed to alanine with no observable effect on flagellar function. Excluding the C-terminus (from Ser-82 to Ala-116), there are 16 charged residues that do not affect the flagellar function of FlhD when they are changed to alanine. It is possible that some of them are involved in some other function(s) of FlhD. As a global regulator, FlhD may change its DNA-binding specificity depending on the protein with which it interacts. Maybe FlhD has the ability to form different complexes inside the cell with other partners to regulate different targets. Some associations will need a strong binding of FlhD, whereas other circumstances may dictate weaker binding to another protein.

The combinatorial specificity makes FlhD a versatile and appropriate protein not only as a model for the study of protein–protein interactions and protein–DNA interactions studies, but also as a model for the study of transcriptional activation by complex formation.

## Experimental procedures

### *Bacterial strains, plasmids and media*

The genotypes of *E. coli* strains and plasmids used in this study are outlined in Table 3. *E. coli* strain Epicuran Coli XL1-Blue (Stratagene) was used to propagate the plasmids generated by *in vitro* site-directed mutagenesis. Strain YK4131 was used to examine the phenotype of the alanine scanning mutants in swarming plates. *E. coli* strain MC1000 *flhD::kan* harbouring the plasmid pGp1-2<sup>cmI</sup> was used to overexpress the FlhD/FlhC complex. Luria–Bertani broth (LB) (Miller, 1992) was used for all purposes other than swarm phenotypic analysis. Antibiotics (Sigma) were added when required at the following concentrations: penicillin G, 100  $\mu$ g ml<sup>-1</sup>; kanamycin, 30  $\mu$ g ml<sup>-1</sup>; chloramphenicol, 25  $\mu$ g ml<sup>-1</sup>. Phenotype assays were performed on tryptone soft agar plates (1% tryptone, 0.5% NaCl and 0.3% bacto-agar). Plasmid pGp1-2<sup>cmI</sup> was used for the expression of the T7 DNA polymerase, which in turn induces expression in the pT7-7 derivative plasmids of the cloned genes. Plasmids pXL25 and pXL27 (and derivatives) were used to overexpress the FlhD *E. coli* protein and the FlhD/FlhC complex respectively.

### Alanine scanning by site-directed mutagenesis

*In vitro* site-directed mutagenesis of *flhD* was performed in accordance with the QuikChange site-directed mutagenesis kit (Stratagene) protocol. Plasmid pXL27, which carries the complete *flhD* and *flhC* genes, was used for each *flhD* mutagenesis experiment according to the instructions of the manufacturer. Changes were introduced through individual and internal mutagenic oligonucleotide primers, complementary to *flhD* except at the position of the desired mutation (Table 1A and B). The changes were located in the middle of the primer with  $\approx$  8–16 bases of correct sequences on both sides. *In vitro* site-directed mutagenesis reactions were performed using 50 ng of pXL27 as the template for each reaction. Twenty cycles were completed with the following parameters: 95°C for 30 s, 55°C for 1 min and 70°C for 8 min. Plasmids carrying each *flhD* mutant were transformed into *Escherichia coli* XL1-Blue competent cells after 1 h of digestion with *DpnI*. Each *flhD* mutant plasmid was then purified using the QIAprep spin miniprep kit (Qiagen), sequenced (Sanger *et al.*, 1977) using the CircumVent thermal cycle dideoxy DNA sequencing kit (New England Biolabs) and transformed into the *flhD*<sup>-</sup> strain YK4131. For each plasmid generated (Table 1A and B), the whole sequence of the *flhD* mutant gene was confirmed. The upstream region from the PT7 to the start codon of *flhD* was also sequenced to exclude any random change that might possibly affect the expression level of the *flhD* gene. Phenotypic analyses were carried out in swarming plates at 30°C in a humid box for 5–6 h.

### Genetic screening of *FlhD* alanine mutants in swarming plates

Phenotypic analyses were carried out by growing the cells (*E. coli* YK4131) harbouring the desired plasmids on LB at 37°C overnight. Two microlitres of an overnight culture were used to inoculate each swarming plate supplied with the appropriate antibiotic when required. Incubation of the swarming plates was carried out at 30°C into a humid camera to avoid evaporation of the media. The swarm diameter of each strain was monitored every 30 min for 11 h.

### *FlhD* and *FlhD/FlhC* complex purification and complex formation assay

Overexpression and purification of *FlhD* native protein was carried out using the two pT7 systems (Tabor, 1990) in *E. coli* MC1000 *flhD::kan* harbouring plasmids pGpP1-2<sup>cmI</sup> and pXL25, as described previously (Campos *et al.*, 1998; 2001).

Overexpression of *E. coli* *FlhD/FlhC* complex was carried out using the two-plasmid pT7 system (Tabor, 1990) in *E. coli* MC1000 *flhD::kan* harbouring plasmids pGpP1-2<sup>cmI</sup> and pXL27 (or pXL27 *flhD* mutant derivatives; Table 3) in 1 l of LB broth as described previously (Liu and Matsumura, 1994). pXL27 and derivatives contain the complete coding sequence of *flhD* and *flhC*. Overinduced culture cells were disrupted by sonication in 20 mM Tris (pH 7.9) and centrifuged at 31 000 *g* for 30 min. The *FlhD/FlhC* complex was then purified to homogeneity by loading the supernatant onto a HiTrap heparin column cartridge (Amersham Pharmacia

Biotech) in a Bio-Rad Econo system. The heterotetramer complex was eluted with a linear gradient of NaCl from 0 to 1.4 M in 20 mM Tris (pH 7.9). *FlhD/FlhC* complex unbinds the heparin at a 0.5 M NaCl concentration. Protein fractions from the heparin column were analysed on SDS–20% polyacrylamide gels and visualized by Coomassie brilliant blue staining. Soluble *FlhD/FlhC* protein was stored at –20°C until use for DNA mobility shift assay when applicable.

*FlhD/FlhC* complex size was confirmed by loading each heparin-purified complex onto a Superdex 200 HR 10/30 filtration column (Pharmacia Biotech) in an ÄKTA<sub>FFPLC</sub> system (Amersham Pharmacia Biotech). The column was equilibrated using BSA (65 kDa), trypsin inhibitor (20 kDa) and lysosyme (14 kDa). Chromatography of each *FlhD/FlhC* complex was carried out using 20 mM Tris, pH 7.9, and 0.5 M NaCl as a buffer at a flow rate of 0.5 ml min<sup>-1</sup>.

### Labelling of *fliA* promoter region

The *fliA* promoter sequence template used in this study for mobility shift assays was prepared as follows. Amplification of the *fliA* promoter region from *E. coli* was carried out by polymerase chain reaction (PCR) (Saiki *et al.*, 1988) using primers *fliA*<sub>5'</sub>B, 5'-GCGCATCCGGCAACATAAAG-3' and *fliA*<sub>3'</sub>B, 5'-CCTTCAGCGGTATAGAGTG-3'. The primers were used to amplify a 211 bp fragment whose 5' and 3' termini extended to –185 and +26, respectively, relative to the transcriptional +1 position of the *fliA* gene (Liu and Matsumura, 1995). The plasmid pXL21 was used as the template (Table 3).

The PCR reaction was performed in 100 µl volumes of Tris buffer (75 mM, pH 8.8) containing 200 mM ammonium sulphate, 0.01% Tween 20, 2 mM magnesium chloride, 0.4 mM each dNTP, 50 pmol of each primer, 50 ng of template DNA and *Taq* DNA polymerase (5 U; MBI Fermentas). Thirty cycles were completed with the following parameters: 95°C for 30 s, 50°C for 1 min and 72°C for 1 min. The PCR product was then 5' labelled with T4 polynucleotide kinase (Promega) with [ $\gamma$ -<sup>32</sup>P]-ATP (Amersham Pharmacia Biotech). One microlitre of DNA was incubated with 10 U of T4 polynucleotide kinase in 10 µl of 1× buffer (distributed by the supplier) and incubated for 30 min at 37°C. Further purification of the labelled fragment was carried out using the QIAquick nucleotide removal kit (Qiagen) according to the manufacturer's instructions and by elution of the probe with 100 µl of 10 mM Tris-Cl, pH 8.5, supplied with the kit. The probe was stored at –20°C until use in the DNA mobility shift assays.

### DNA mobility shift assay

To test the ability of each *FlhD/FlhC* complex for DNA binding, the heparin-eluted complexes were used. Mobility shift assay was carried out according to the method of Fried and Crothers (1981). The binding reactions were carried out in a 20 µl volume of 5 mM Tris (pH 7.9), 50 mM KCl, 5 mM EDTA, 1 mM dithiothreitol, 5% glycerol, 10 ng of poly-(dI–dC)-(dI–dC) (Amersham Pharmacia Biotech), 1 ng of <sup>32</sup>P-labelled DNA and 50–1750 ng of purified protein. After incubation at 30°C for 30 min, the samples were loaded

onto 5% polyacrylamide gels, and the bands were visualized by autoradiography.

## Acknowledgements

The authors thank Azucena Rosas for her help in the DNA characterization of the alanine scanning mutants, and Bryan Shimkos and Rhonda T. Fleming for their help in the second alanine scanning. We also acknowledge Peggy O'Neill for critical reading of the manuscript. This work was supported by National Institutes of Health grant GM59484 to P.M.

## References

- Aizawa, S.-I. (1996) Flagellar assembly in *Salmonella typhimurium*. *Mol Microbiol* **19**: 1–5.
- Akerley, B.J., Cotter, P.A., and Miller, J.F. (1995) Ectopic expression of the flagellar regulon alters development of the *Bordetella*–host interaction. *Cell* **80**: 611–620.
- Allison, C., Lai, H.-C., and Huges, C. (1992) Co-ordinate expression of virulence genes during swarm-cell differentiation and population migration of *Proteus mirabilis*. *Mol Microbiol* **6**: 1583–1591.
- Badger, J., and Miller, V.L. (1998) Expression of invasins and motility are coordinately in *Yersinia enterocolitica*. *J Bacteriol* **180**: 793–800.
- Bartlett, D.H., Frantz, B.B., and Matsumura, P. (1988) Flagellar transcriptional activators FlbB and Flal: Gene sequences and 5' consensus sequences of operons under FlbB and Flal control. *J Bacteriol* **170**: 1575–1581.
- Branden, C., and Tooze, J. (1991) DNA recognition by proteins with the helix-turn-helix motif. In *Introduction to Protein Structure*. Branden, C., and Tooze, J. (eds). New York: Garland Publishing, pp. 85–110.
- Campos, A., Matsumura, P., and Volz, K. (1998) Crystallization and preliminary X-ray analysis of FlhD from *Escherichia coli*. *J Struct Biol* **12**: 269–271.
- Campos, A., Zhang, R.-G., Alkire, A., Matsumura, P., and Westbrook, E. (2001) Crystal structure of the global regulator FlhD from *Escherichia coli* at 1.8 Å resolution. *Mol Microbiol* **39**: 000–000(in press).
- Clarke, S., and Koshland, D.E. (1979) Membrane receptors for aspartate and serine bacterial chemotaxis. *J Biol Chem* **254**: 9695–9702.
- Dreyfus, G., Williams, A.W., Kawagishi, I., and Macnab, R.M. (1993) Genetic and biochemical analysis of *Salmonella typhimurium* Flil, a flagellar protein related to the catalytic subunit of the F<sub>0</sub>F<sub>1</sub> ATPase and to virulence proteins of mammalian and plant pathogens. *J Bacteriol* **175**: 3131–3138.
- Eichelberg, K., and Galán, J.E. (2000) The flagellar sigma factor FliA ( $\sigma^{28}$ ) regulates the expression of *Salmonella* genes associated with the centisome 63 type III secretion system. *Infect Immun* **68**: 2735–2743.
- Fauconnier, A., Allaoui, A., Campos, A., Elsen, A.V., Cornelis, G.R., and Bollen, A. (1997) Flagellar *flhA*, *flhB* and *flhE* genes, organized in an operon, cluster upstream from the *inv* locus in *Yersinia enterocolitica*. *Microbiology* **143**: 3461–3471.
- Fried, M., and Crothers, D.M. (1981) Equilibria and kinetics of lac repressor–operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res* **9**: 6505–6525.
- Furness, R.B., Fraser, G.M., Hay, N.A., and Hughes, C. (1997) Negative feedback from a *Proteus* class II flagella export defect to the *flhDC* master operon controlling cell division and flagellum assembly. *J Bacteriol* **179**: 5585–5588.
- Givaudan, A., and Lanois, A. (2000) *flhDC*, the flagellar master operon of *Xenorhabdus nematophilus*. requirement for motility, lipolysis, extracellular hemolysis, and full virulence in insects. *J Bacteriol* **182**: 107–115.
- Givskov, M., and Molin, S. (1992) Expression of extracellular phospholipase from *Serratia liquefaciens* is growth-phase dependent, catabolite repressed and regulated by anaerobiosis. *Mol Microbiol* **6**: 1363–1374.
- Givskov, M., and Molin, S. (1993) Secretion of *Serratia liquefaciens* phospholipase from *Escherichia coli*. *Mol Microbiol* **8**: 229–242.
- Givskov, M., Ebert, L., Christiansen, G., Benedik, M.J., and Molin, S. (1995) Induction of phospholipase- and flagellar synthesis in *Serratia liquefaciens* is controlled by expression of the flagellar master operon *flhD*. *Mol Microbiol* **15**: 445–454.
- Guex, N., and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer. An environment for comparative protein modeling. *Electrophoresis* **18**: 2714–2723.
- Gygi, D., Bailey, M.J., Allison, C., and Hughes, C. (1995) Requirement for FlhA in flagella assembly and swarm-cell differentiation by *Proteus mirabilis*. *Mol Microbiol* **4**: 761–769.
- Harayana, S., Engstrom, P., Wolf-Watz, H., Iino, T., and Hazelbauer, G.L. (1982) Cloning of *trg*, a gene for a sensory transducer in *Escherichia coli*. *J Bacteriol* **122**: 206–214.
- Hedblom, M.L., and Adler, J. (1980) Genetic and biochemical properties of *Escherichia coli* mutants with defects in serine chemotaxis. *J Bacteriol* **144**: 1048–1060.
- Hughes, K.T., Gillen, K.L., Semon, M.J., and Karlinsey, J.E. (1993) Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator. *Science* **19**: 1277–1280.
- Jones, S., and Thornton, J.M. (1996) Principles of protein–protein interactions. *Proc Natl Acad Sci USA* **93**: 13–20.
- Kapatral, V., Olson, J.W., Pepe, J.C., Miller, V.L. and S.A., (1996) Minnich temperature-dependent regulation of *Yersinia enterocolitica* class III flagellar genes. *Mol Microbiol* **19**: 1061–1071.
- Komeda, Y. (1982) Fusions of flagellar operons to lactose genes on a Mu *lac* bacteriophage. *J Bacteriol* **150**: 16–26.
- Komeda, Y. (1986) Transcriptional control of flagellar genes in *Escherichia coli* K-12. *J Bacteriol* **168**: 1315–1318.
- Komeda, Y., Kutsukake, K., and Iino, T. (1980) Definition of additional flagellar genes in *Escherichia coli* K12. *Genetics* **94**: 277–290.
- Kondoh, H., Ball, C.B., and Adler, J. (1979) Identification of a methyl-accepting chemotaxis protein for the ribose and galactose chemoreceptors of *Escherichia coli*. *Proc Natl Acad Sci USA* **76**: 260–264.
- Kutsukake, K., and Iino, T. (1994) Role of the FliA–FlgM regulatory system on the transcriptional control of the

- flagellar regulon and flagellar formation in *Salmonella typhimurium*. *J Bacteriol* **176**: 3598–3605.
- Kutsukake, K., Ohaya, Y., and Iino, T. (1990) Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. *J Bacteriol* **172**: 741–747.
- Kutsukake, K., Iyoda, S., Ohnishi, K., and Iino, T. (1994) Genetic and molecular analyses of the interaction between the flagellum-specific sigma and anti-sigma factors in *Salmonella typhimurium*. *EMBO J* **13**: 4568–4576.
- Liu, X., and Matsumura, P. (1994) The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. *J Bacteriol* **176**: 7345–7351.
- Liu, X., and Matsumura, P. (1995) An alternative sigma factor controls transcription of flagellar class-III operons in *Escherichia coli*. gene sequence, overproduction, purification and characterization. *Gene* **164**: 81–84.
- Lucas, R.L., Lostroh, C.P., diRusso, C.C., Spector, M.P., Wanner, B.L., and Lee, C.A. (2000) Multiple factors independently regulate *hilA* and invasion gene expression in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* **182**: 1872–1882.
- Macnab, R.M. (1996) Flagella and motility. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, Vol. 1, 2nd edn. Neidhardt, F.C., Curtiss, R. III, Ingraham, J.L., Lin, E.C.C., Low, J.K.B., Magasanik, B., et al. (eds). Washington, DC: American Society for Microbiology Press, pp. 123–145.
- Malakooti, J., Komeda, Y., and Matsumura, P. (1989) DNA sequence analysis, gene product identification, and localization of flagellar motor components of *Escherichia coli*. *J Bacteriol* **171**: 2728–2734.
- Manson, M.D., Blank, V., Grade, G., and Higgins, C.F. (1986) Peptide chemotaxis in *Escherichia coli* involves the Tap signal transducer and the dipeptide permease. *Nature* **321**: 253–256.
- Miller, J.H. (1992) *A Short Course in Bacterial Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Ohnishi, K., Kutsukake, K., Suzuki, H., and Iino, T. (1990) Gene *fliA* encodes an alternative sigma factor specific for flagellar operons in *Salmonella typhimurium*. *Mol Gen Genet* **221**: 139–147.
- Ohnishi, K., Kutsukake, K., Suzuki, H., and Iino, T. (1992) A novel transcriptional regulation mechanism in the flagellar regulon of *Salmonella typhimurium*: an anti-sigma factor inhibits the activity of the flagellum-specific sigma factor, sigma<sup>F</sup>. *Mol Microbiol* **6**: 3149–3157.
- Ordal, G.W., and Adler, J. (1974a) Isolation and complementation of mutants in galactose taxis and transport. *J Bacteriol* **117**: 517–526.
- Ordal, G.W., and Adler, J. (1974b) Properties of mutants in galactose taxis and transport. *J Bacteriol* **117**: 517–526.
- Ottemann, K.M., and Miller, J.F. (1997) Roles for motility in bacterial–host interactions. *Mol Microbiol* **24**: 1109–1117.
- Pabo, C.O., and Sauer, R.T. (1984) Protein-DNA recognition. *Annu Rev Biochem* **53**: 293–321.
- Pabo, C.O., and Sauer, R.T. (1992) Transcriptional factors: Structural families and principles of DNA recognition. *Annu Rev Biochem* **61**: 1053–1095.
- Prüß, B.M., and Matsumura, P. (1996) A regulator of the flagellar regulon of *Escherichia coli*, *fliD*, also affects cell division. *J Bacteriol* **178**: 668–674.
- Prüß, B.M., Markovic, D., and Matsumura, P. (1997) The *Escherichia coli* flagellar transcriptional activator *fliD* regulates cell division through induction of the acid response gene *cadA*. *J Bacteriol* **179**: 3818–3821.
- Rendone, M.B. (1992) Molecular analysis of FlhD and FlhC function in *Escherichia coli*. Master degree thesis, University of Illinois, Chicago.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., et al. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487–491.
- Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**: 5463–5467.
- Schmiel, D.H., Young, G.M., and Miller, V.L. (2000) The *Yersinia enterocolitica* phospholipase gene *ypIA* is part of the flagellar regulon. *J Bacteriol* **182**: 2314–2320.
- Soutourina, O., Kolb, A., Krin, E., Laurent-Winter, C., Rimsky, S., Danchin, A., and Bertin, P. (1999) 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 *fliDC* master operon. *J Bacteriol* **181**: 7500–7508.
- Tabor, S. (1990) Expression using the T7 RNA polymerase/promoter system. In *Current Protocols in Molecular Biology*, Vol. 2. Ausubel, F.A., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds). New York: Greene Publishing and Wiley-Interscience, pp. 16.2.1–16.2.11.
- Van Gijsegem, F.V., Gough, C., Zischek, C., Niqueux, E., Arlat, M., Genin, S., et al. (1995) The *hrp* gene locus of *Pseudomonas solanacearum*, which controls the production of a type III secretion system encodes eight proteins related to components of the bacterial flagellar biogenesis complex. *Mol Microbiol* **15**: 1095–1114.
- Wang, E.A., and Koshland, D.E. (1980) Receptor structure in bacterial sensing system. *Proc Natl Acad Sci USA* **77**: 7157–7161.
- Yanagihara, S., Iyoda, S., Ohnishi, K., Iino, T., and Kutsukake, K. (1999) Structure and transcriptional control of the flagellar master operon of *Salmonella typhimurium*. *Genes Genet Syst* **74**: 105–111.
- Young, G.M., Schmiel, D.H., and Miller, V.L. (1999a) A new pathway for the secretion of virulence factors by bacteria: the flagellar export apparatus functions as a protein-secretion system. *Proc Natl Acad Sci USA* **96**: 6456–6461.
- Young, G.M., Smith, M.J., Minnich, S.A., and Miller, V.L. (1999b) The *Yersinia enterocolitica* motility master regulatory operon, *fliDC*, is required for flagellin production, swimming motility, and swarming motility. *J Bacteriol* **181**: 2823–2833.