***Agrobacterium tumefaciens* possesses a fourth flagellin gene located in a large gene cluster concerned with flagellar structure, assembly and motility**

William J. Deakin,[†] Victoria E. Parker, Emma L. Wright,
Kevin J. Ashcroft, Gary J. Loake[‡] and Charles H. Shaw

Author for correspondence: Charles H. Shaw Tel: +44 191 374 2430. Fax: +44 191 374 2417.
e-mail: C.H.Shaw@Durham.ac.uk

Department of Biological
Sciences, University of
Durham, South Road,
Durham, DH1 3LE, UK

The authors have identified a fourth flagellin gene in a 21850 bp region of the *Agrobacterium tumefaciens* C58C1 chromosome containing at least 20 genes concerned with flagellar structure, assembly and function. Three flagellin genes, *flaA*, *flaB* and *flaC*, orientated rightward, are positioned in a tandem array at the right end, with the fourth, substantially larger gene, *flaD*, in the opposite orientation, at the left end. Between these lie four apparent operons, two transcribed in each direction (*motA*, *flhB* leftward; *flgF*, *flgB* rightward) from a divergent position approx 7.5 kb from the left end. This unifies the previously published *motA*, *flgB* and *flaABC* sequences into a single region, also containing the homologues of *flhB*, *flgF* and *flil*. Site-specific mutagenesis of *flil* results in a non-flagellate phenotype, while a Tn5-induced *flhB* mutant possesses abnormal flagella. Mutagenesis and protein profiling demonstrate that all four flagellins contribute to flagellar structure: FlaA is the major protein, FlaB and FlaC are present in lesser amounts, and FlaD is a minor component. FlaA has anomalous electrophoretic mobility, possibly due to glycosylation.

Keywords: *Agrobacterium*, flagella, flagellin, motility

INTRODUCTION

Much has been written of the process whereby *Agrobacterium tumefaciens* transfers T-DNA from the Ti-plasmid to the plant genome, leading to crown gall tumour (Sheng & Citovsky, 1996). Less is known of the events preceding the induction of Ti-plasmid virulence (*vir*) genes by plant wound phenolics. Several studies have pointed to the importance of motility and chemotaxis in rhizosphere colonization and attraction to wound sites (Ashby *et al.*, 1987, 1988; Chesnokova *et al.*, 1997; Hawes & Smith, 1989; Hawes *et al.*, 1988; Shaw *et al.*, 1991). Ti-plasmid-encoded proteins are not

only involved in events following *vir*-induction, but also in chemotaxis towards wound phenolics (Ashby *et al.*, 1987, 1988; Palmer & Shaw, 1992; Shaw *et al.*, 1988). Chemotaxis to sugars (Cangelosi *et al.*, 1990; Loake *et al.*, 1988), and to opines (Kim & Farrand, 1998), released by plants may also contribute to the rhizosphere role of *A. tumefaciens* in association with roots and established tumours.

We have embarked on a programme to characterize chemotaxis and motility genes from *A. tumefaciens* C58C1, to elucidate their role in crown gall tumour formation (Deakin *et al.*, 1997a, b; Shaw *et al.*, 1991; Wright *et al.*, 1998). We have determined the DNA sequence of the *motA* operon (Deakin *et al.*, 1997b), encoding switch and motor proteins (*motA**flhMNG*: 3978 bp), the *flgB* operon (Deakin *et al.*, 1997a), encoding flagellar basal-body and assembly proteins (*flgABCGHI**flhELP*: 7205 bp) and of three tandemly arrayed flagellin genes (Deakin, 1994) (*flaABC*: 4423 bp). Significantly, the *flaABC* sequence was con-

[†]Present address: LBMPs, Université de Genève, 1 ch. de l'Imperatrice, 1292 Chambésy-Genève, Switzerland.

[‡]Present address: Department of Cell & Molecular Biology, University of Edinburgh, Edinburgh, UK.

The GenBank/EMBL/DBJ accession number for the sequence reported in this paper is U95165.

firmed by another group from *A. tumefaciens* NT1 (Chesnokova *et al.*, 1997).

Here we report the properties and DNA sequence of 21850 bp of *Agrobacterium tumefaciens* C58C1 chromosomal DNA (Fig. 1). This sequence unifies *flaABC*, the *flgB* operon and the *motA* operon. Interestingly, the sequence includes a newly identified fourth flagellin gene, *flaD*, and two additional putative operons including the homologues of *flhB*, *flgF* and *flil*.

METHODS

Bacterial strains and plasmids. The bacterial strains and plasmids used in this work are described in Table 1. Bacteria were grown in Lab M nutrient broth no. 2 (LM broth) or on Lab M nutrient agar (LM agar) (Amersham). *Escherichia coli* strains were grown at 37 °C and *A. tumefaciens* strains at 28 °C. Antibiotics were used at the following concentrations for *E. coli* strains: ampicillin (Ap), 50 µg ml⁻¹; gentamicin (Gm), 15 µg ml⁻¹; kanamycin (Km), 50 µg ml⁻¹; neomycin (Nm), 50 µg ml⁻¹; tetracycline (Tc), 15 µg ml⁻¹. When required, 40 µl of 20 mg X-Gal ml⁻¹ in dimethyl formamide was spread over the surface of agar plates with a sterile glass loop. For *A. tumefaciens*, antibiotics were used at the following concentrations: Gm, 100 µg ml⁻¹; Km, 25 µg ml⁻¹; Nm, 100 µg ml⁻¹; rifampicin (Rif), 100 µg ml⁻¹; Tc, 10 µg ml⁻¹. Swarm plates were prepared by the addition of 0.16% agar

(Oxoid) to LM broth. The *sacB* gene was induced by the addition of 5% sucrose to LM agar plates.

DNA methods. Chromosomal DNA was isolated using a modification of the method of Dhaese *et al.* (1979). Promega Wizard Miniprep columns were used to isolate plasmid DNA, according to the manufacturer's instructions. Recombinant DNA techniques were carried out according to Sambrook *et al.* (1989). Restriction and modification enzymes were obtained from Northumbria Biologicals and were used according to the manufacturer's instructions. DNA fragments were isolated from agarose gel slices by melting in NaI and then by the addition of silica.

Random Tn5 mutagenesis of *A. tumefaciens* was as previously described (Shaw *et al.*, 1991). Site-specific mutagenesis was performed by the insertion of a neomycin-resistance cassette (Shaw *et al.*, 1984) into DNA fragments subcloned into an *A. tumefaciens* suicide vector, pJQ200SK (Quandt & Hynes, 1993). pJQ200SK carries the sucrose-inducible *sacB* gene, whose product is lethal to Gram-negative bacteria. Thus after conjugation into *A. tumefaciens*, by triparental mating (Ditta *et al.*, 1980), double recombination events can be selected on media containing sucrose and neomycin. Gene replacement was confirmed by two-way capillary Southern blotting (Sambrook *et al.*, 1989) onto Hybond-N membranes (Amersham). Hybridization, with probes radiolabelled using the Amersham Multiprime kit, was performed according to the manufacturer's instructions, at high stringency.

Table 1. Bacterial strains

Strain	Relevant properties	Source/reference
<i>E. coli</i>		
DH5α	<i>supE44 ΔlacU169 (φ80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i>	Woodcock <i>et al.</i> (1989)
<i>A. tumefaciens</i>		
C58C1	Wild-type chemotaxis, Ti ⁻ , Rif ^R	van Larebeke <i>et al.</i> (1974)
<i>mot-1</i>	Tn5 insertion in <i>flaA</i> , Km ^R Rif ^R	Shaw <i>et al.</i> (1991)
<i>mot-4</i>	Tn5 insertion in <i>motB</i> , Km ^R Rif ^R	Shaw <i>et al.</i> (1991)
<i>mot-9</i>	Tn5 insertion in <i>flhB</i> , Km ^R Rif ^R	Shaw <i>et al.</i> (1991)
<i>flaB</i>	Nm ^R insertion in <i>flaB</i> , Nm ^R Rif ^R	This work
<i>flaC</i>	Nm ^R insertion in <i>flaC</i> , Nm ^R Rif ^R	This work
<i>flaD</i>	Nm ^R insertion in <i>flaD</i> , Nm ^R Rif ^R	This work
<i>flaBC</i>	Deletion of <i>flaBC</i> , Rif ^R	This work
<i>flaABC</i>	Deletion of <i>flaABC</i> , Rif ^R	This work
<i>flil</i>	Nm ^R insertion in <i>flil</i> , Nm ^R Rif ^R	This work
Plasmids		
pDUB1900	pLAFR3 derivative containing ~25 kb of <i>A. tumefaciens</i> DNA, Tc ^R	Shaw <i>et al.</i> (1991)
pDUB1801	~13 kb DNA of <i>mot-1</i> (Tn5 insertion in <i>flaA</i>) in pUC18, Ap ^R Km ^R	Shaw <i>et al.</i> (1991)
pDUB1804	7 kb DNA of <i>mot-4</i> (Tn5 insertion in <i>motB</i>) in pUC18, Ap ^R Km ^R	Shaw <i>et al.</i> (1991)
pDUB1809	8.4 kb DNA of <i>mot-9</i> (Tn5 insertion in <i>flhB</i>) in pUC18, Ap ^R Km ^R	Shaw <i>et al.</i> (1991)
pJQ200SK	<i>sacB</i> suicide vector with polylinker of pBluescript SK+, Gm ^R	Quandt & Hynes (1993)
pRK2013	Helper plasmid for conjugation in <i>A. tumefaciens</i> , Km ^R	Figurski & Helinski (1979)

DNA sequencing. DNA sequencing was carried out with an Applied Biosystems 373A DNA sequencer using double-stranded DNA templates. Usually fluorescently labelled universal M13 primers from the Applied Biosystems PRISM Ready Reaction Dye Primer Cycle Sequencing Kit were used, according to the manufacturer's instructions. The oligonucleotide primer used to sequence from Tn5 (5'-GAAAC-GGGAAAGGTTCCGT-3') was synthesized on an Applied Biosystems 381A DNA synthesizer. It was used in conjunction with the Applied Biosystems *Taq* DyeDeoxy Terminator Cycle Sequencing Kit.

Searches of GenBank and the EMBL database were performed using the program FASTA. Computerized DNA sequence analyses used the UWGCG programs (Devereux *et al.*, 1984) at SEQNET, SERC Daresbury Laboratory, UK. Putative protein sequences were searched against the Owl database using the SEQNET programs Sooty/Sweep and BLASTP.

Microscopy. Light microscopy was performed with a Nikon Optiphot microscope, using phase-contrast optics. For electron microscopy, cells were stained with 1% uranyl acetate on a Formvar-coated grid, which were viewed in a Philips EM400.

Flagellar filament isolation. Flagellar filaments were isolated (Krupski *et al.*, 1985; Robinson *et al.*, 1992) from bacterial lawns washed off five agar plates using five 3 ml aliquots of 150 mM NaCl, pooled and collected in 50 ml polypropylene centrifuge tubes. The flagella were detached from the cells by vortexing for 15 s, and the cell bodies pelleted at 12000 g. The supernatants were centrifuged again for 10 min at 15000 g, and the flagella filaments pelleted by centrifuging at 100000 g, 4 °C, for 2 h. The supernatants were removed and the flagellar filaments resuspended in 200 µl HEPES buffer (10 mM HEPES, 10 µM EDTA pH 8.0, 200 µM CaCl₂). Purity of the flagellar preparations was assessed by electron microscopy and SDS-PAGE.

SDS-PAGE. Flagella were denatured by boiling for 5 min with an equal volume of 2× sample buffer (4% SDS, 20% glycerol, 120 mM Tris/HCl pH 6.8 and 0.005% bromophenol blue, with 10% β-mercaptoethanol added immediately before use) then separated at 150 V on a 10% running gel with a 5% stacker using an electrophoresis buffer containing 25 mM Tris/HCl pH 8.8, 190 mM glycine and 0.1% SDS. Molecular size markers were obtained from Bio-Rad and Sigma, unstained. For silver staining, after electrophoresis, the gel was successively placed in 5% formaldehyde/40% ethanol (30 min), distilled water, 50% methanol (overnight), 1 mM dithiothreitol (30 min) and finally 0.1% silver nitrate. After 30 min, the gel was washed three times with distilled water, and developed with 3% Na₂CO₃, 0.0185% formaldehyde for 2–5 min. The reaction was stopped with 20 ml 50% citric acid and the gel was washed with distilled water. Sugar modification of proteins was detected after electroblotting to nitrocellulose, using the DIG Glycan Detection Kit (Boehringer Mannheim), which coupled digoxigenin to aldehydes generated by periodic acid oxidation. Bound digoxigenin was detected by anti-digoxigenin-alkaline phosphatase conjugate.

RESULTS

Flagellin genes

Three flagellin genes, *flaA*, *flaB* and *flaC*, orientated rightward, are positioned in a tandem array at the right end of the sequenced region, with the fourth, *flaD*, in the opposite orientation, some 15 kb away, at the left end

(Fig. 1). The *flaD* ORF was identified by homology to other flagellin genes, and is predicted to be a coding region by TESTCODE (Fickett, 1982). A putative ribosome-binding site, GGGTGA, is positioned 4 bp upstream and TERMINATOR (Brendel *et al.*, 1986) identifies a putative rho-independent terminator (GC-rich stem-loop motif and 4–8 thymidine residues) 49 bp downstream of *flaD*. The upstream region shows weak similarity with the flagellar consensus promoter (TAAA...N₁₆...GCCGATAA), but good identity with the *ntxA* type promoter (GC...N₁₀...GC) found in the lateral flagellin gene of *Azospirillum brasilense* (Moens *et al.*, 1996). It is thus likely that *flaD* is transcribed as a single gene, as the preceding apparent operon, commencing with *flhB*, has a downstream sequence with a strong match to the consensus terminator (see below). *flaA*, *flaB* and *flaC* are transcribed separately (Chesnokova *et al.*, 1997), and accordingly are each flanked by flagellar (σ²⁸) consensus promoters (TAAA...N₁₆...GCCGATAA), ribosome-binding sites and terminators, although *flaC* may possess σ²⁸ and σ⁵⁴ type promoters (Chesnokova *et al.*, 1997).

The sizes of the *Agrobacterium tumefaciens* C58C1 flagellin gene ORFs are: *flaA* 921 bp, *flaB* 963 bp, *flaC* 942 bp and *flaD* 1293 bp. These encode predicted proteins of 306 amino acids (predicted molecular mass 31637 Da), 320 amino acids (predicted molecular mass 32966 Da), 313 amino acids (predicted molecular mass 32843 Da) and 430 amino acids (predicted molecular mass 45107 Da), respectively. The four *A. tumefaciens* flagellins have significant sequence identity to a number of other flagellin proteins from a wide range of bacteria, notably *Sinorhizobium meliloti*. GAP alignments between the putative *A. tumefaciens* protein sequences reveal the following close resemblances: FlaA vs FlaB, 64% identity (78% similarity); FlaA vs FlaC, 60% identity (73% similarity); FlaA vs FlaD, 56% identity (71% similarity); FlaB vs FlaC, 66% identity (78% similarity); FlaB vs FlaD, 60% identity (74% similarity); and FlaC vs FlaD, 62% identity (74% similarity). All the flagellin proteins have conserved amino- and carboxy-termini (Fig. 2), with the region between exhibiting the greatest variability in amino acid sequence and length. FlaA, FlaB and FlaC are all predicted to be of similar molecular mass. FlaD is predicted to be substantially larger, and is thus more like the flagellins of *S. meliloti* 10406 (Pleier & Schmitt, 1989) or the 414-residue lateral flagellin, LafA, of *Azospirillum brasilense*, with which it shares 35% identity (58% similarity) (Moens *et al.*, 1996).

Phenotypes of flagellin gene mutants

The Tn5 insertion site in *mot-1* (Shaw *et al.*, 1991) is near to the 5' end of *flaA* (see Fig. 1). This mutant is non-motile in swarm agar plates and electron microscopy revealed truncated flagellar filaments which were generally straighter than the (apparently) sinusoidally curved filaments of wild-type *A. tumefaciens* flagella (Shaw *et al.*, 1991). The filaments of mutant *mot-1*

Fig. 1. Map of *A. tumefaciens* flagellar cluster 1. Top line: *EcoRI* (E) and *HpaI* (H) sites within cluster 1. Boxes denote *BamHI*, *HindIII* and *EcoRI* sites within the chromosomal insert in cosmid pDUB1900. Lines below that depict the extents of the cloned fragments isolated from various mutants, the Tn5 insertion site in each denoted by ∇ . Horizontal arrows represent known extents of genes. x, y, z are ORFs of unknown function. Map positions of mutants *fla-3* and *mot-4* were erroneously assigned in a previous paper (Shaw *et al.*, 1991). *motA**fliMNG*, 3978 bp, accession no. U63290 (9 July 1996); *fliN* accession no. X65584 (16 April 1992); *flgABCGHI**fliELP*, 7205 bp, accession no. U39941 (2 November 1995); *flaABC*, 4423 bp, accession no. X80701 (29 July 1994). Complete sequenced region GenBank accession number U95165 (1 October 1997).

appear to rotate, since *mot-1* cells were observed by light microscopy to rotate when attached to glass coverslips by their flagella (data not shown). New mutations were introduced into *flaB*, *flaC* and *flaD* (Table 1), by insertion of a neomycin-resistance cassette (see Methods). Separate mutation of *flaB*, *flaC* or *flaD* results in a motile phenotype, albeit slightly attenuated compared to wild-type, evidenced both in swarm plates and by phase-contrast microscopy (Table 2). This suggests that individually, each flagellin is not absolutely essential for flagellar function. However, a double mutation removing both *flaB* and *flaC* (Table 2) attenuates motility more severely. Complete deletion of *flaABC* results in a non-motile phenotype (Table 2), in agreement with published results (Chesnokova *et al.*, 1997).

Flagella were not detected during electron microscopic observation of the *flaABC* deletion mutant, while the flagella of the *flaABC* deletion mutant appeared morphologically normal, albeit somewhat straighter (Table 2, Fig. 3). Flagella of the *flaB*, *flaC* or *flaD* mutants were apparently similar to wild-type, although filaments from the *flaC* and *flaD* mutants showed enhanced curvature and those from the *flaD* mutant also had numerous

kinks throughout their length (Fig. 3). Taken together with the previous analysis of the *flaA* mutant, *mot-1* (Shaw *et al.*, 1991), these results suggest that FlaA is absolutely required for flagella structure in *A. tumefaciens*, with each of the other flagellins playing a minor but significant role in complete filament integrity. However, in the absence of FlaB or FlaC or FlaD a rotating flagellum capable of driving motility is formed, albeit with structural deformities. When all three of the smaller flagellins (FlaABC) are absent, no flagella are formed.

Isolation and investigation of flagellin proteins of *A. tumefaciens*

Flagella were isolated from wild-type *A. tumefaciens* C58C1 and the flagellin mutants (see Methods). Similar aliquots of the isolated filaments were analysed by SDS-PAGE; the resulting gel is shown in Fig. 4 (and summarized in Table 2). Wild-type flagella (track 7) contain two major protein bands running close together, an upper heavy band of apparent mobility 33 kDa and a slightly less prominent 32 kDa (lower) band. In addition there are two faint proteins at 47 kDa and 49 kDa. In

Mutant	Motile behaviour	Swarm size (%)†	Flagella‡	Protein(s) absent§
<i>mot-1</i> (<i>flaA</i> ⁻)	Non-motile	32	Severely truncated	33
<i>flaB</i> ⁻	Slightly attenuated	81	Normal	32
<i>flaC</i> ⁻	Slightly attenuated	81	Enhanced curvature	32
<i>flaD</i> ⁻	Slightly attenuated	94	Enhanced curvature, kinked	49
<i>flaBC</i> ⁻	Severely attenuated	71	Normal, straighter	32
<i>flaABC</i> ⁻	Non-motile	19	Not detected	32, 33, 47, 49
<i>flil</i>	Non-motile	48	Not detected	Not done

§ Size of protein band (in kDa) missing from appropriate track in Fig. 4.

Fig. 3. Electron micrographs of the *A. tumefaciens* flagellin mutants: (a) *flaABC* mutant; (b) *flaB* mutant; (c) *flaC* mutant; (d) *flaD* mutant; (e) *flaBC* mutant; (f) *flil* mutant. Bars, 0.4 μ m (a, c, d, f), 0.2 μ m (b, e).

Fig. 4. SDS-PAGE of isolated flagellar filament proteins from wild-type and mutant *A. tumefaciens* C58C1: 10% polyacrylamide gel with isolated flagellar proteins separated and silver stained. Tracks: S, size standards (Bio-Rad); 1, 10 μ l proteins from *mot-1* (*flaA*) mutant; 2, 5 μ l proteins from *flaB* mutant; 3, 5 μ l proteins from *flaC* mutant; 4, 10 μ l proteins from *flaD* mutant; 5, 5 μ l proteins from *flaABC* mutant; 6, 5 μ l proteins from *flaABC* mutant; 7, 5 μ l proteins from wild-type C58C1.

flagella from *mot-1* cells, where *flaA* is mutated, the upper heavy 33 kDa band is absent (track 1), indicating that this protein is FlaA. The lower 32 kDa band is reduced in intensity in flagella from the *flaB* (track 2) and *flaC* (track 3) mutants, and completely absent from flagella of the *flaABC* mutant (track 5). This indicates that the 32 kDa band contains both FlaB and FlaC flagellins. The faint 49 kDa band is absent from flagella of the *flaD* mutant (track 4), suggesting that this protein is FlaD. In addition to FlaABC, FlaD and the 47 kDa protein are both absent from preparations from the *flaABC* mutant (track 6).

These protein data confirm that all four flagellin genes are expressed, and that all four flagellin proteins are incorporated into flagella filaments. In common with the electron microscope results, the data indicate that without the three major flagellins (FlaABC), filament polymerization is impossible. However, FlaD may still be exported, but without an intact filament to assemble into is lost into the medium. It is thus possible that the 47 kDa protein is a distal filament protein, for example one of the hook-associated proteins (HAPs), such as HAP2 (FliD), or alternately a fifth flagellin. Southern blotting has failed to reveal a fifth gene (data not shown). In the absence of FlaA, FlaBCD are exported, and incorporated into a defective flagellum, incapable of proper assembly due to the absence of the major flagellin. Accumulation of unassembled FlaBC in a cellular compartment was ruled out by subcellular fractionation (data not shown).

Densitometric tracing (data not shown) suggests that approximately twice as much of FlaA (the upper heavy 33 kDa band) is present in the flagellar preparations compared to FlaB and FlaC combined (the lower 32 kDa band). Thus the protein profiles and the mutant phenotypes indicate that FlaA in *A. tumefaciens* forms the major component of the flagellum, as the corresponding protein does in *S. meliloti* (Pleier & Schmitt, 1989). FlaB and FlaC are lesser components, and FlaD a minor flagellin.

Glycosylation of FlaA

The assignment of the above flagellins to the bands shown in Fig. 4 is not in accordance with the predicted protein sizes from the DNA sequence. The predicted sizes of FlaB and FlaC are 33 kDa and 32.8 kDa, respectively, and thus would be expected to co-electrophorese. However, FlaA is predicted to be smaller, at 31.6 kDa. Thus it appears that the FlaBC proteins have a higher mobility and/or FlaA a lower mobility, relative to each other, during SDS-PAGE. This may be due to modification of one or other of the proteins. Examination of the flagellin amino acid sequences suggests that FlaA may be glycosylated, due to the presence of a glycosylation signal (N-X-S/T) within the central non-conserved region. Using a glycosylation test kit, we were able to detect faint positive staining of FlaA, and negative staining of FlaBCD (Fig. 5). This indicates that alone amongst *A. tumefaciens* flagellins, FlaA may be glycosylated, which may account for its anomalous electrophoretic behaviour.

flhB and its putative operon

flhB has an ORF of 360 codons, predicted to encode a protein of 40883 Da. Downstream of *flhB*, following a short spacer of 5 bp (Fig. 1), there is an ORF (orfY) of 145 codons which TESTCODE (Fickett, 1982) predicts to be a coding region, but which has no obvious homologues in sequence databases. Preceding *flhB* is the *motA* operon terminator (Deakin *et al.*, 1997b), followed by a typical flagellar consensus promoter. TERMINATOR (Brendel *et al.*, 1986) identifies a putative rho-independent terminator (GC-rich stem-loop motif and 4–8 thymidine residues) beyond orfY, suggesting that *flhB* and orfY are expressed as a single operon. A Tn5 insertion towards the 3' end of *flhB* (mutant *mot-9*: see Fig. 1) produces a non-motile phenotype, in which elongated, straightened flagella are produced (Shaw *et al.*, 1991).

Fig. 5. Detection of sugar modification of proteins after SDS-PAGE: 10% polyacrylamide gel with isolated proteins separated, electroblotted and stained using a DIG Glycan Detection Kit. Tracks: 5, molecular mass standards (Sigma); 1, 5 μ l fetuin; 2, 5 μ l deglycosylated fetuin; 3, 5 μ l bovine serum albumin; 4, 5 μ l transferrin; 5, 5 μ l isolated flagellar filament proteins of wild-type C58C1; 6, 5 μ l creatinase. Glycosylated proteins (e.g. fetuin, transferrin) show positive staining; non-glycosylated proteins (e.g. molecular mass standards) show negative or neutral staining.

flgF and its putative operon

Between the *motA* (Deakin *et al.*, 1997b) and *flgB* (Deakin *et al.*, 1997a) operons, and transcribed in the same direction as *motA*, four ORFs appear to be flanked by a single promoter and terminator (Fig. 1). Southern blotting confirms that all four ORFs are conserved in the same orientation in *Sinorhizobium* (data not shown). The first and last ORFs, X and Z, of 312 and 134 codons respectively, are predicted by TESTCODE to be coding regions, but have no clear homologues in the databases. Mutations introduced into them (see Methods) have no effect on motility (data not shown), suggesting that they are not essential for a motile phenotype. The central ORFs in the operon, *flgF* (rod protein) and *flil*, of 224 and 473 codons, respectively, are predicted to encode proteins of 26119 and 50775 Da. Site-specific mutagenesis (see Methods) of *flil* (Table 1) results in a non-motile phenotype, with no visible flagella (Fig. 3, Table 2).

DISCUSSION

This paper demonstrates that genes concerned with motility are clustered in *Agrobacterium*, as they are in

other bacteria (Macnab, 1996). This cluster contains genes concerned with filament and basal body structure, assembly and export, and motion and switching. Such clustering may facilitate controlled expression of a set of genes required for a similar purpose. The *Agrobacterium* cluster shows considerable similarity with that from *Sinorhizobium meliloti* (Sourjik *et al.*, 1998). The gene order in the two clusters is identical from *flhB* to the third flagellin gene (*flaC* in *Agrobacterium*, *flaD* in *Sinorhizobium*). However, in the *Sinorhizobium* cluster, the *che* operon is linked downstream (i.e. to the left) of *flhB*, while the *Agrobacterium* cluster possesses an inversion which positions the fourth flagellin gene, *flaD*, and the *motBCD* operon (data not shown) downstream of *flhB*. We have no data on linkage of the *Agrobacterium che* operon (Wright *et al.*, 1998) to the cluster described here.

The data suggest that wild-type flagellar filaments possess several types of flagellin subunits, since, in the absence of FlaA subunits, a partial filament (presumably formed from FlaB or FlaC or FlaD subunits) is still made. When FlaA is present, it appears to be able to form a partially functional flagellum, when one or other of the minor flagellins is absent. Absence of both FlaB and FlaC has a more detrimental effect upon flagellar

function. When all three smaller flagellins (FlaABC) are absent, distal filament components appear not to assemble correctly. Thus, although they may be exported, they do not appear in flagellar preparations.

In some bacteria, notably *S. meliloti*, 'complex' flagella have been observed (Götz *et al.*, 1982), which are composed of more than one type of flagellin (Bergman *et al.*, 1991; Pleier & Schmitt, 1989, 1991). The partial filament phenotype of the *mot-1* mutant is similar to that observed for FlaA null mutants of *S. meliloti*. Thus it is possible that *A. tumefaciens* also possesses complex flagella. The flagella of *A. tumefaciens* do appear, however, to have subtle structural differences from those of *S. meliloti* (Shaw *et al.*, 1991). Moreover, the results do not rule out the possibility that *A. tumefaciens* elaborates two distinct flagellar types, as in *Vibrio parahaemolyticus* (McCarter & Wright, 1993), or *Azospirillum brasilense* (Moens *et al.*, 1995). Intriguingly, FlaD, the larger of the four flagellins described here, shows strong sequence identity with the lateral flagellin of *Az. brasilense*. However, we have no direct evidence that supports the suggestion of two flagellar types in *A. tumefaciens*.

The purified flagellins of *A. tumefaciens* show anomalous electrophoretic behaviour. The mobility of FlaA in SDS-PAGE would be decreased if its molecular mass were increased by post-translational modification, such as the glycosylation suggested here. This has been shown to occur in a number of flagellins from different bacteria, notably the polar flagellin of *Az. brasilense* (Moens *et al.*, 1995), and the archaeon *Halobacterium halobium* (Gerl & Sumper, 1988). Other flagellin modifications include methylation in *Salmonella typhimurium* (Macnab, 1992), and phosphorylation in *Pseudomonas aeruginosa* (Kelly-Wintenberg *et al.*, 1993) and *Campylobacter coli* (Guerry *et al.*, 1991; Logan *et al.*, 1989). Glycosylation may be involved in flagellar assembly and swimming behaviour (Moens *et al.*, 1995), and may also have a role to play in attachment to the plant. Interestingly, the *mot-1* mutant has impaired root colonization ability (Shaw *et al.*, 1991), and a *flaABC* strain has reduced tumorigenicity (Chesnokova *et al.*, 1997), although in both cases it is difficult to disentangle the contribution of motility from that of attachment.

Flagellar assembly is best understood in the enteric bacteria (Macnab, 1996), but even there, substantial unsolved problems remain. In these bacteria, the filament is composed of one protein, flagellin (FliC), and flagella are built from individual flagellin monomers folded back upon themselves, like a functional cotter-pin, with the amino- and carboxy-termini defining the boundaries of the central channel (DeRosier, 1992; Namba *et al.*, 1989). These domains are thought to be important for filament formation and interaction with the neighbouring subunits. Thus they show the highest degree of sequence conservation between different bacterial types, including *A. tumefaciens*. The highly variable central domains of the flagellin monomers are on the filament surface and hence responsible for the antigenic properties of the filament.

The enteric flagellins do not possess conventional amino-terminal signal sequences to direct their export; instead they are exported via the flagellum-specific export pathway (Macnab, 1996), several members of which (FlhB, FliI, FliP) are encoded in the cluster described here. FliI and FliP are members of a superfamily of export proteins (Dreyfus *et al.*, 1993; Fenselau *et al.*, 1992; Galan *et al.*, 1992; Plano *et al.*, 1991; Venkatesan *et al.*, 1992). In *S. typhimurium* FliI resembles the F_0F_1 ATPase, and is believed to be cytoplasmic and involved in flagellum-specific export (Dreyfus *et al.*, 1993; Vogler *et al.*, 1991). FliP is probably a transmembrane protein (Malakooti *et al.*, 1994) and appears to play an early role in the specific export (and possibly assembly) of the axial flagellar proteins (Dreyfus *et al.*, 1993; Kubori *et al.*, 1992). FlhB is believed to regulate export of hook and filament proteins, by closing the gate for hook protein export, but opening it for filament proteins (Kutsukake, 1997; Williams *et al.*, 1996). Thus the phenotypes associated with mutation of *fliI*, *fliP* (Deakin *et al.*, 1997a) and *flhB* (Shaw *et al.*, 1991) in *A. tumefaciens* are consistent with their involvement in flagellum-specific export and assembly.

This paper unifies three previously published loci in *Agrobacterium* into one large cluster (Chesnokova *et al.*, 1997; Deakin *et al.*, 1997a, b). We are currently sequencing two further large regions concerned with chemotaxis and motility, at least one of which maps to the same 40 kb region of the chromosome as the cluster described here (Shaw *et al.*, 1991). It is thus likely that most *Agrobacterium* genes involved in motility and chemotaxis will map to a few large gene clusters, as is the case for other bacteria. This not only facilitates controlled expression of the genes, but also their rigorous study. The challenge will be to ascribe functional roles to the genes identified.

ACKNOWLEDGEMENTS

We thank Julia Bartley for assistance with DNA sequencing, Christine Richardson for help with electron microscopy, Kostia Bergman (Boston) for advice, Rudi Schmitt (Regensburg) for information and strains, and Jane Robinson (Columbus) for useful discussions. W. J. D. was the recipient of a postgraduate studentship from the SERC.

REFERENCES

- Ashby, A. M., Watson, M. D. & Shaw, C. H. (1987). A Ti-plasmid determined function is responsible for chemotaxis of *Agrobacterium tumefaciens* towards the plant wound product acetosyringone. *FEMS Microbiol Lett* **41**, 189–192.
- Ashby, A. M., Watson, M. D., Loake, G. J. & Shaw, C. H. (1988). Ti-plasmid specified chemotaxis of *Agrobacterium tumefaciens* C58C1 toward *vir*-inducing phenolic compounds and soluble factors from monocotyledonous and dicotyledonous plants. *J Bacteriol* **170**, 4181–4187.
- Bergman, K., Nulty, E. & Su, L. (1991). Mutations in the two flagellin genes of *Rhizobium meliloti*. *J Bacteriol* **173**, 3716–3723.
- Brendel, V., Hamm, G. H. & Trifonov, E. N. (1986). Terminators of transcription with RNA polymerase from *Escherichia coli*: what

- they look like and how to find them. *J Biomol Struct Dyn* 3, 705–723.
- Cangelosi, G. A., Ankenbauer, R. G. & Nester, E. W. (1990). Sugars induce the *Agrobacterium tumefaciens* virulence genes via a periplasmic binding protein and the VirA protein. *Proc Natl Acad Sci USA* 87, 6708–6712.
- Chesnokova, O., Coutinho, J. B., Khan, I. H., Mikhail, M. S. & Kado, C. I. (1997). Characterization of flagella genes of *Agrobacterium tumefaciens*, and the effect of a bald strain on virulence. *Mol Microbiol* 23, 579–590.
- Deakin, W. J. (1994). *Molecular characterisation of flagellar genes from Agrobacterium tumefaciens*. PhD thesis, University of Durham.
- Deakin, W. J., Furniss, C. S. M., Parker, V. E. & Shaw, C. H. (1997a). Isolation and characterisation of a linked cluster of genes from *Agrobacterium tumefaciens* encoding proteins involved in flagellar basal-body structure. *Gene* 189, 135–137.
- Deakin, W. J., Sanderson, J. L. C. I., Goswami, T. & Shaw, C. H. (1997b). The *Agrobacterium tumefaciens* motor gene, *motA*, is in a linked cluster of genes with the flagellar switch protein genes, *fliG*, *fliM* and *fliN*. *Gene* 189, 139–141.
- DeRosier, D. J. (1992). Whipping flagellin into shape. *Curr Opin Struct Biol* 2, 280–285.
- Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12, 387–395.
- Dhaese, P., De Greve, H., Decraemer, H., Schell, J. & Van Montagu, M. (1979). Rapid mapping of transposon insertion and deletion mutants in the large Ti-plasmids of *Agrobacterium tumefaciens*. *Nucleic Acids Res* 7, 1837–1849.
- Ditta, G., Stanfield, S., Corbin, D. & Helinski, D. R. (1980). Broad host range DNA cloning system for Gram-negative bacteria: Construction of a gene bank in *Rhizobium meliloti*. *Proc Natl Acad Sci USA* 77, 7347–7351.
- Dreyfus, G., Williams, A. W., Kawgishi, I. & Macnab, R. M. (1993). Genetic and biochemical analysis of *Salmonella typhimurium* FliI, a flagellar protein related to the catalytic subunit of the F_0F_1 ATPase and to virulence proteins of mammalian and plant pathogens. *J Bacteriol* 175, 3131–3138.
- Fenselau, S., Balbo, I. & Bonas, U. (1992). Determinants of pathogenicity in *Xanthomonas campestris* pv. *vesicatoria* are related to proteins involved in secretion in bacterial pathogens of animals. *Mol Plant-Microbe Interact* 5, 390–396.
- Fickett, J. W. (1982). Recognition of protein coding regions in DNA-sequences. *Nucleic Acids Res* 10, 5303–5318.
- Figurski, D. H. & Helinski, D. R. (1979). Replication of an origin containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. *Proc Natl Acad Sci USA* 76, 1648–1652.
- Galan, J. E., Ginocchio, C. & Costeas, P. (1992). Molecular and functional characterization of the *Salmonella* invasion gene *invA*: homology of *InvA* to members of a new protein family. *J Bacteriol* 174, 4338–4349.
- Gerl, L. & Sumper, M. (1988). Halobacterial flagellins are encoded by a multigene family. *J Biol Chem* 263, 13246–13251.
- Götz, R., Limmer, N., Ober, K. & Schmitt, R. (1982). Motility and chemotaxis in two strains of *Rhizobium* with complex flagella. *J Gen Microbiol* 128, 789–798.
- Guerry, P., Alm, R. A., Power, M. E., Logan, S. M. & Trust, T. J. (1991). Role of two flagellin genes in *Campylobacter* motility. *J Bacteriol* 173, 4757–4764.
- Hawes, M. C. & Smith, L. Y. (1989). Requirement for chemotaxis in pathogenicity of *Agrobacterium tumefaciens* on roots of soil grown pea plants. *J Bacteriol* 171, 5668–5671.
- Hawes, M. C., Smith, L. Y. & Howarth, A. J. (1988). *Agrobacterium tumefaciens* mutants deficient in chemotaxis to root exudates. *Mol Plant-Microbe Interact* 1, 182–186.
- Kelly-Wintenberg, K., South, S. L. & Montie, T. C. (1993). Tyrosine phosphate in a- and b-type flagellins of *Pseudomonas aeruginosa*. *J Bacteriol* 175, 2458–2461.
- Kim, H. & Farrand, S. K. (1998). Opine catabolic loci from *Agrobacterium* plasmids confer chemotaxis to their cognate substrates. *Mol Plant-Microbe Interact* 11, 131–143.
- Krupski, G., Götz, R., Ober, K., Pleier, E. & Schmitt, R. (1985). Structure of complex flagellar filaments in *Rhizobium meliloti*. *J Bacteriol* 162, 361–366.
- Kubori, T., Shimamoto, N., Yamaguchi, S., Namba, K. & Aizawa, S.-I. (1992). Morphological pathway of flagellar assembly in *Salmonella typhimurium*. *J Mol Biol* 226, 433–446.
- Kutsukake, K. (1997). Hook-length control of the export-switching machinery involves a double-locked gate in *Salmonella typhimurium* flagellar morphogenesis. *J Bacteriol* 179, 1268–1273.
- van Larebeke, N., Engler, G., Holsters, M., Van den Elsacker, S., Zaenen, I., Schilperoort, R. A. & Schell, J. (1974). Large plasmid in *Agrobacterium tumefaciens* essential for crown gall inducing ability. *Nature* 252, 169–170.
- Loake, G. J., Ashby, A. M. & Shaw, C. H. (1988). Attraction of *Agrobacterium tumefaciens* C58C1 towards sugars involves a highly sensitive chemotaxis system. *J Gen Microbiol* 134, 1427–1432.
- Logan, S. M., Trust, T. J. & Guerry, P. (1989). Evidence for posttranslational modification and gene duplication of *Campylobacter* flagellin. *J Bacteriol* 171, 3031–3038.
- Macnab, R. M. (1992). Genetics and biogenesis of bacterial flagella. *Annu Rev Genet* 26, 131–158.
- Macnab, R. M. (1996). Flagella and Motility. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, pp. 123–145. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.
- McCarter, L. M. & Wright, M. E. (1993). Identification of genes encoding components of the swarmer cell flagellar motor and propeller and a sigma factor controlling differentiation of *Vibrio parahaemolyticus*. *J Bacteriol* 175, 3361–3371.
- Malakooti, J., Ely, B. & Matsumura, P. (1994). Molecular characterization, nucleotide sequence and expression of the *fliO*, *fliP*, *fliQ* and *fliR* genes of *Escherichia coli*. *J Bacteriol* 176, 189–197.
- Moens, S., Michiels, K. & Vanderleyden, J. (1995). Glycosylation of the flagellin of the polar flagellum of *Azospirillum brasilense*, a Gram-negative nitrogen-fixing bacterium. *Microbiology* 141, 2651–2657.
- Moens, S., Schlöter, M. & Vanderleyden, J. (1996). Expression of the structural gene, *laf1*, encoding the flagellin of the lateral flagella in *Azospirillum brasilense* sp7. *J Bacteriol* 178, 5017–5019.
- Namba, K., Yamashita, I. & Vonderviszt, F. (1989). Structure of the core and central channel of bacterial flagella. *Nature* 342, 648–654.
- Palmer, A. C. V. & Shaw, C. H. (1992). The role of VirA and VirG phosphorylation in chemotaxis towards acetosyringone by *Agrobacterium tumefaciens*. *J Gen Microbiol* 138, 2509–2514.
- Plano, G. V., Barve, S. S. & Straley, S. C. (1991). LcrD, a membrane-

- bound regulator of the *Yersinia pestis* low-calcium response. *J Bacteriol* **173**, 7293–7303.
- Pleier, E. & Schmitt, R. (1989).** Identification and sequence analysis of two related flagellin genes in *Rhizobium meliloti*. *J Bacteriol* **171**, 1467–1475.
- Pleier, E. & Schmitt, R. (1991).** Expression of two *Rhizobium meliloti* flagellin genes and their contribution to the complex filament structure. *J Bacteriol* **173**, 2077–2085.
- Quandt, J. & Hynes, M. F. (1993).** Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. *Gene* **127**, 15–21.
- Robinson, J. B., Tuovinen, O. H. & Bauer, W. D. (1992).** Role of divalent cations in the subunit associations of complex flagella from *Rhizobium meliloti*. *J Bacteriol* **174**, 3896–3902.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989).** *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Shaw, C. H., Watson, M. D., Carter, G. H. & Shaw, C. H. (1984).** The right hand copy of the nopaline Ti-plasmid is required for tumour formation. *Nucleic Acids Res* **12**, 6031–6041.
- Shaw, C. H., Ashby, A. M., Brown, A. P., Royal, C., Loake, G. J. & Shaw, C. H. (1988).** *virA* and *virG* are the Ti-plasmid functions required for chemotaxis of *Agrobacterium tumefaciens* towards acetosyringone. *Mol Microbiol* **2**, 413–418.
- Shaw, C. H., Loake, G. J., Brown, A. P. & 7 other authors (1991).** Isolation and characterisation of behavioural mutants and genes from *Agrobacterium tumefaciens*. *J Gen Microbiol* **137**, 1939–1953.
- Sheng, J. & Citovsky, V. (1996).** *Agrobacterium*–plant cell DNA transport: have virulence proteins, will travel. *Plant Cell* **8**, 1699–1710.
- Sourjik, V., Sterr, W., Platzter, J., Bos, I., Haslbeck, M. & Schmitt, R. (1998).** Mapping of 41 chemotaxis, flagellar and motility genes to a single region of the *Sinorhizobium meliloti* chromosome. *Gene* **223**, 283–290.
- Venkatesan, M. M., Buysse, J. M. & Oaks, E. V. (1992).** Surface presentation of *Shigella flexneri* invasion plasmid antigens requires the products of the *spa* locus. *J Bacteriol* **174**, 1990–2001.
- Vogler, A. P., Homma, M., Irikura, V. M. & Macnab, R. M. (1991).** *Salmonella typhimurium* mutants defective in flagellar filament regrowth and sequence similarity of FliI to F₀F₁ vacuolar and archaeobacterial ATPase subunits. *J Bacteriol* **173**, 3564–3572.
- Williams, A. W., Yamaguchi, S., Togashi, F., Aizawa, S. I., Kawagishi, I. & Macnab, R. M. (1996).** Mutations in *fliK* and *flhB* affecting flagellar hook filament assembly in *Salmonella typhimurium*. *J Bacteriol* **178**, 2960–2970.
- Woodcock, D. M., Crowther, P. J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S. S., Michael, M. Z. & Graham, M. W. (1989).** Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res* **17**, 3469–3478.
- Wright, E. L., Deakin, W. J. & Shaw, C. H. (1998).** A chemotaxis cluster from *Agrobacterium tumefaciens*. *Gene* **220**, 83–89.

Received 6 November 1998; revised 25 February 1999; accepted 3 March 1999.