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

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 (motAfilMNG: 3978 bp), the flgB operon (Deakin et al., 1997a), encoding flagellar basal-body and assembly proteins (flgABCGHfilELP: 7205 bp) and of three tandemly arrayed flagellin genes (Deakin, 1994) (flaABC: 4423 bp). Significantly, the flaABC sequence was con-
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 CS8C1 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 fll.

**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: 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**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant properties</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5a</td>
<td>supE44 ΔlacU169 (Δ80 lacZAM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1</td>
<td>Woodcock et al. (1989)</td>
</tr>
<tr>
<td><strong>A. tumefaciens</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS8C1</td>
<td>Wild-type chemotaxis, Ti⁻, RifR</td>
<td>van Larebeke et al. (1974)</td>
</tr>
<tr>
<td>mot-1</td>
<td>Tn5 insertion in flaA, KmR RifR</td>
<td>Shaw et al. (1991)</td>
</tr>
<tr>
<td>mot-4</td>
<td>Tn5 insertion in motB, KmR RifR</td>
<td>Shaw et al. (1991)</td>
</tr>
<tr>
<td>mot-9</td>
<td>Tn5 insertion in flbB, KmR RifR</td>
<td>Shaw et al. (1991)</td>
</tr>
<tr>
<td>flaB</td>
<td>NmR insertion in flaB, NmR RifR</td>
<td>This work</td>
</tr>
<tr>
<td>flaC</td>
<td>NmR insertion in flaC, NmR RifR</td>
<td>This work</td>
</tr>
<tr>
<td>flaD</td>
<td>NmR insertion in flaD, NmR RifR</td>
<td>This work</td>
</tr>
<tr>
<td>flaBC</td>
<td>Deletion of flaBC, RifR</td>
<td>This work</td>
</tr>
<tr>
<td>flaABC</td>
<td>Deletion of flaABC, RifR</td>
<td>This work</td>
</tr>
<tr>
<td>flit</td>
<td>NmR insertion in flit, NmR RifR</td>
<td>This work</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pDUB1900</td>
<td>pLAFR3 derivative containing ~25 kb of A. tumefaciens DNA, TeR</td>
<td>Shaw et al. (1991)</td>
</tr>
<tr>
<td>pDUB1801</td>
<td>~13 kb DNA of mot-1 (Tn5 insertion in flaA) in pUC18, ApR KmR</td>
<td>Shaw et al. (1991)</td>
</tr>
<tr>
<td>pDUB1804</td>
<td>7 kb DNA of mot-4 (Tn5 insertion in motB) in pUC18, ApR KmR</td>
<td>Shaw et al. (1991)</td>
</tr>
<tr>
<td>pDUB1809</td>
<td>8-4 kb DNA of mot-9 (Tn5 insertion in flbB) in pUC18, ApR KmR</td>
<td>Shaw et al. (1991)</td>
</tr>
<tr>
<td>pJQ200SK</td>
<td>sacB suicide vector with poly linker of pBluescript SK+ , GmR</td>
<td>Quandt &amp; Hynes (1993)</td>
</tr>
<tr>
<td>pRK2013</td>
<td>Helper plasmid for conjugation in A. tumefaciens, KmR</td>
<td>Figurski &amp; Helinski (1979)</td>
</tr>
</tbody>
</table>
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 flagellar 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 2x 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 periodical acid oxidation. Bound digoxigenin was detected by antidigoxigenin–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, GGGTGGA, 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 ntrA type promoter (GC...N₂₅...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 flbB, 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 (σ²₆) consensus promoters (TAAA...N₁₆...GCCGATAA), ribosome-binding sites and terminators, although flaC may possess σ²₆ and σ₃₄ 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 32834 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 FlbB, 64% identity (78% similarity); FlaA vs FlcC, 60% identity (73% similarity); FlaA vs FlbD, 56% identity (71% similarity); FlbB vs FlcC, 66% identity (78% similarity); FlbB vs FlbD, 60% identity (74% similarity); and FlcC vs FlbD, 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, FlbB and FlcC are all predicted to be of similar molecular mass. FlbD 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
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 flaBC 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
Fig. 2. Sequence alignments of all four *A. tumefaciens* flagellins. Asterisks denote conserved residues; dots denote conservative substitutions.

**Table 2. Properties of flagellin mutants**

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

* By phase-contrast microscopy.

† Percentage of wild-type swarms, in 0.2% agar after 24 h (mean of two replicates).

‡ Appearance in electron microscope (Fig. 3 and Shaw et al., 1991).

§ 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) fill mutant. Bars, 0.4 μm (a, c, d, f), 0.2 μm (b, e).
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 flaBC 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 (FlhD), 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).
**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 *fliI*, of 224 and 473 codons, respectively, are predicted to encode proteins of 26119 and 50775 Da. Site-specific mutagenesis (see Methods) of *fliI* (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 flagellum is composed of one protein, flagellin (Fla), and flagella are built from individual flagellin monomers folded back upon themselves, like a functional cotterpin, 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, FlhF, FlhP) are encoded in the cluster described here. Flh and FlhP 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 Flh resembles the F,F ATPase, and is believed to be cytoplasmic and involved in flagellum-specific export (Dreyfus et al., 1993; Vogler et al., 1991). FlhF 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 flh, flf (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


they look like and how to find them. J Biomol Struct Dyn 3, 705–723.


Agrobacterium flagellar cluster


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