Escherichia coli MT78, an avian pathogenic strain of serogroup O2, produces a variant form of type 1 fimbriae with distinct antigenic properties and apparent mol. wt of the major subunit. The fim gene cluster of strain MT78 was cloned and its sequence was determined in a region spanning upstream of fimB to the beginning of fimD. Whereas most genes were well conserved relative to fim genes previously described, comparison of the fimA gene from strain MT78 with homologous sequences from other strains of E. coli and Klebsiella pneumoniae revealed that most differences were clustered in four well defined regions. A PCR assay, based upon these variable sequences, allowed amplification of a fragment of gene fimA which is specific for most O2 strains. In addition, the sequence of the previously uncharacterised gene fimL, which is located between genes fimA and fimC, was determined.

Introduction

Type 1 fimbriae (also called F1A) are extracellular proteinaceous appendages present on the surface of most strains of Escherichia coli and of several other members of the Enterobacteriaceae. They are characterised phenotypically by mannose-sensitive haemagglutination (MSHA) of guinea-pig or chicken erythrocytes, and consist of a major protein, FimA, associated with ancillary proteins, FimF, FimG and adhesin FimH. Investigations on the fim gene cluster of E. coli have revealed an organisation of at least nine genes, seven of which are present in a single operon whose expression is controlled by an invertible element containing the promoter [1].

Several studies suggest that the expression of type 1 fimbriae is involved in the virulence of E. coli for chickens [2, 3]. E. coli strains pathogenic for chickens induce a respiratory disease that is followed frequently by a systemic infection [4]. Adhesion of bacteria to the respiratory tract of chickens, mediated by type 1 fimbriae, is most probably the first step in the infection, as demonstrated by D-mannose-inhibited bacterial adhesion to tracheal epithelial cells, and by the in-vivo expression of type 1 fimbriae in the respiratory tract of experimentally infected chickens [5, 6].

Several variants of type 1 fimbriae have already been described for avian strains of E. coli [7, 8]. A previous study [9] demonstrated that the major subunit of type 1 fimbriae isolated from O2:K1 avian strains of E. coli (one of the main serogroups associated with avian colibacillosis) differed in mol. wt and antigenic properties from that of a K12 strain [10].

To further characterise fimbriae produced by these strains, and to look for relationships with other previously described fim genes, the fim cluster from E. coli MT78, a representative serogroup O2 strain of avian origin was cloned, and its sequence was determined in a 5.2-kb region extending from upstream of fimB to the beginning of fimD. The sequence of gene fimL, the function of whose product is not known, was also determined.

Materials and methods

Bacterial strains

E. coli strain MT78 (O2:K1) was isolated in France from the trachea of a chicken suffering from colibacillosis [9]. This strain produces only type 1 fimbriae as shown by haemagglutination tests and colony hybridisation with fim, pap and sfa/foc probes (M. Dho-Moulin, unpublished data). E. coli strains HB101 [11] and AAEC185 (Δ-fim) [12] were used for constitution of the cosmid library and the subcloning steps. E. coli strain XL1-blue (Stratagene) was used for transformation with recombinant M13 bacteriophages. All the field strains tested in the PCR assays were isolated from chickens or turkeys suffering from avian colibacillosis.
Haemagglutination assays

Blood was collected from specific-pathogen-free White Leghorn chickens (strain PA12), and haemagglutination assays were carried out as described previously [2].

Purification of major fimbrial subunit FimA and determination of its N-terminal sequence

Purified type 1 fimbriae were obtained from a 3-L tryptic soy broth culture of E. coli strain MT78 as described previously [9]. The FimA protein was separated from other fimbrial components by SDS-PAGE. The 18.5-kDa protein band was electro-eluted (10 mA for 5 h) in a volatile buffer (50 mM NH₄HCO₃, SDS 0.1%) against a 3.5-kDa cut-off membrane in a BioRad apparatus (BioRad Laboratories, Richmond, CA, USA), and dialysed against distilled water. The N-terminal amino-acid sequence was determined in an automated gas-phase sequencer.

Construction of a cosmid library of MT78

Whole-cell DNA was prepared as described previously [13] from an exponentially growing culture (100 ml) of strain MT78, partially digested with Sau3A and size-fractionated by centrifugation through a sucrose 10-40% gradient. Fragments (c. 40 kb) were ligated to dephosphorylated, BamHI-linearised pHC79 [14]. Ligation products were packaged in vitro into bacteriophage λ particles (In vitro Packaging-Kit, Amersham) for subsequent infection of E. coli HB101. In a second step, large-scale preparation of cosmid from mannose-sensitive haemagglutination-positive (MSHA) clones were again partially digested and size-fractionated as before; fragments (10–15 kb) were ligated to BamHI-treated pUC18. All restriction endonucleases and DNA-modifying enzymes were purchased from Boehringer Mannheim GmbH (Germany) and used according to the manufacturer's instructions. DNA restriction endonuclease fragments were separated on agarose (0.7%) gels and recovered by phenol extraction as described previously [15].

DNA sequencing

Chosen restriction endonuclease fragments from the purified plasmids were ligated to double-stranded DNA from M13mp18 or M13mp19 bacteriophages, and XL1-blue bacteria were transformed with the recombinant molecules as described previously [16]. Sequencing reactions were performed with commercial kits (T7-Kit, Pharmacia; Sequenase-II, Amersham) on single-stranded DNA preparations of the recombinant bacteriophages with either universal M13 primers flanking the cloned fragments, or internal oligonucleotides (purchased from Eurogentec, Seraing, Belgium). Sequence analyses were performed with programmes from the GCG package [17].

PCR analysis

Bacteria from an overnight culture in 1 ml of Luria broth were collected by centrifugation, then resuspended in 200 µl of sterile water and boiled for 10 min. Bacterial debris was removed by centrifugation, and 5 µl of the supernate were tested in a 25 µl PCR assay containing 0.5 unit of Taq polymerase (Appligene, France) and 12.5 pmol of the forward and reverse primers. The primers used were: fimAO5, 5'-GTGTGATTCAAACCGTTTCG; fimAl6, 5'-AATAAGCGCGCTG-GAACG; fimA201, 5'-TCTGGCTGATACTACTACC; and fimA215, 5'-ACTTATAGATTAGTACTG. The PCR reaction conditions were as follows: 94°C for 2 min, 50°C for 45 s, 72°C for 1 min (1 cycle), 94°C for 45 s, 50°C for 45 s, 72°C for 1 min (26 cycles), then 72°C for 5 min. PCR products were analysed by electrophoresis in an agarose 1.2% gel in Tris-borate-EDTA buffer containing ethidium bromide (1 µg/ml). Bands were visualised by UV transillumination.

Results

Cloning and sequencing of the fim cluster from MT78

After constitution of the genomic library of E. coli MT78 in E. coli HB101, haemagglutination assays on 500 colonies identified two independent MSHA-positive clones. A rapid analysis of the two cosmids showed that they possessed some restriction fragments in common. DNA fragments (c. 10–15 kb) from the two cosmids were subcloned into pUC18 in the host strain AAEC185 (A-fim). One of the MSHA-positive clones thus obtained (pA080) was studied further and its restriction map was established (Fig. 1). A central HindIII fragment and a BglII-PstI fragment at its 3'-most region (a and b in Fig. 1) were subcloned into M13mp18 and their ends were sequenced. Partial sequences of fimH [18], and fimA [10] were recognised. Appropriate DNA fragments from pA080 were then cloned into bacteriophage vectors M13mp18 and M13mp19 and sequenced in both orientations.

The determined sequence (5429 nucleotides) begins 824 nucleotides upstream of fimB and ends at the beginning of fimD. It includes genes fimB (nt 825–1427), fimE (1947–2501), fimA (2982–3526), fimI (3601–4140), fimC (4177–4902), and approximately one-tenth of the fimD sequence (4968–5249). It has been deposited in the EMBL data library under the access number Z37500.

Comparisons of MT78 fimB, fimE and fimC genes with homologous sequences

The sequences of these three genes from strain MT78 were compared with those previously published for fimB, fimE [19] and fimC [20,21]. These three genes are very well conserved and only minor substitutions were observed in the MT78 sequences. Of the 13
nucleotide substitutions observed in *fimB* between strain MT78 and a K12 strain [19], two resulted in amino-acid changes: the Asp residue at position 85 and the Val residue at position 86 were changed to Glu and Ile, respectively, in the sequence. In *fimE*, one of the 11 nucleotide differences resulted in a Val to Ala substitution at the last position of the protein. The *fimC* gene differed from the sequence published by Jones *et al.* [21] by only one silent nucleotide substitution, whereas comparison with the sequence published by Klemm [20] revealed 13 nucleotide substitutions, of which four resulted in amino-acid changes at positions 29 (Met to Val in our sequence), 54 (Glu to Val), 81 (Gly to Asp) and 210 (Thr to Ala) of the precursor polypeptide.

**Sequence of the fimI gene**

A 540-bp open reading frame, corresponding to gene *fiml* [20], was identified between genes *fimA* and *fimC*. Its analysis (Fig. 2) revealed that it encoded a 179-residue precursor containing a putative prokaryotic signal sequence that is probably cleaved after the alanine residue at position 19 [22], thus generating a 160-residue mature protein of calculated Mr 17.16 kDa. The estimated iso-electric point of mature FimI (starting from residue number 20) is 8.2, versus 4.7 for mature FimA. Protein FimI was shown to be 50–60% similar to other fimbrial subunits, with 30–35% of identical residues at equivalent positions (data not shown). It shared some constant features with all the

---

**Fig. 1.** Restriction map of recombinant plasmid pA080. Filled rectangles represent the *fim* genes, open rectangle represents vector pUC18. The invertible element is represented by a horizontal arrow between *fimE* and *fimA*. The sites for common restriction endonucleases are indicated by letters: Ba, BamHI, Bg, BglII; E, EcoRI; H, HindIII; P, PstI.

**Fig. 2.** Nucleotide and deduced amino-acid sequence of the *fimI* gene. Nucleotide sequence begins at the stop codon of *fimA* and ends in the beginning of *fimC*, whose initiation codon GTG, as proposed by Klemm [20], is underlined, 36 nucleotides downstream from the *fimI* stop codon.
fimbrial subunits, such as two cysteine residues within the amino-terminal half of the protein (positions 43 and 83) and a penultimate tyrosine residue. Its amino-acid composition revealed a high proportion of arginine and proline residues and a low content of threonine and alanine residues relative to FimA. Furthermore, like minor fimbrial subunits, it contained tryptophan residues that are not present in major subunits.

In order to search for relationships between FimI proteins, that of E. coli MT78 was compared with its counterpart from Salmonella Typhi, whose gene has recently been cloned and sequenced [23], and that of the partial sequence of FocI, from the foc operon encoding type 1C fimbriae [24, 25]. The latter was found in the sequence data library by a FastA search. The multiple alignment (Fig. 3) shows clearly that MT78 FimI is closely related to FocI, at least in the first 60 residues. The comparison of MT78 FimI with S. Typhi FimI showed a more remote relationship between the two, which nevertheless were 37.4% identical with a similarity reaching c. 60% when allowing for conservative substitutions.

Nucleotide sequence of fimA and polypeptide sequence of its product

The MT78-∗fimA gene encodes a 184-residue polypeptide which is cleaved after the alanine residue at position 23, as demonstrated by the N-terminal sequence of mature FimA (DTTPTTVNGG). Five published sequences were identified by the FastA program, that were very similar to the MT78-∗fimA sequence (Fig. 4). The strongest similarity was observed with Ecotype1f, from a serogroup O78 E. coli strain of avian origin [8], and Kpnfamb, from Klebsiella pneumoniae strain IA551 [26], with similarity levels of 93.6 and 93.7%, respectively. FimA genes from other strains showed lower levels of similarity with MT78-∗fimA: 91.2% for EcimIa01, from an E. coli K12 strain [10], and 90.9% for Ecopilaa, from a human isolate of E. coli, strain J96 [27]. These percentages were approximately the same when the amino-acid sequences of mature FimAs were compared.

A multiple alignment of FimA amino-acid sequences from strain MT78 and the other five most similar sequences (Fig. 4) showed that the most important variations between these six sequences were in four regions: aa 24–27, 65–69, 104–109 and 134–137. The first site (aa 24–27) corresponds to the characteristic N-terminus of mature FimA from strain MT78, a particular feature which has been confirmed by N-terminal sequencing of the purified protein. Secondary structure predictions [29] of FimA from strain MT78 revealed that at least two of these regions (aa 104–109 and 134–137) corresponded to hydrophilic sites which are likely to be exposed at the surface of the protein. Furthermore, these sites were associated with a significant antigenic index (Fig. 5), suggesting that they may take part in antigenic sites.

The results presented above led us to consider the possibility that these short variable regions might be specific to strains of serogroup O2. Such relationships between particular serogroups and variants of P

Fig. 3. Multiple alignment of FimI products. The deduced amino-acid sequence of MT78 FimI was aligned with that of S. Typhi [23] (ST FimI) and with the translated ORF identified in the foc sequence [25], downstream of focA gene. The latter is an incomplete ORF and its product is here designated EC ‘FocI’. Consensus represents identical residues at the same position in two or three sequences. Gaps, represented by dots, were introduced to optimise the alignment. Alignment was performed with the ‘FileUp’ programme in the GCG package [17], then edited manually.
fimba have been shown previously [30]. To address this question, a PCR assay was designed, in which two sets of primers were used. The first PCR was carried out with primers fimA05 and fimA16 to detect, with a low stringency, the presence of gene fimA. The primers fimA201 and fimA215 in the second PCR were chosen to amplify gene fimA with a high degree of specificity corresponding to the MT78 sequence. The sequence of the forward primer fimA201 corresponded to the characteristic N-terminal region of mature FimA, whereas the reverse primer fimA215 corresponded to the third variable region (aa 104–109 of the FimA precursor). A set of 62 field isolates of E. coli of avian origin, belonging to serogroups O78 (22 strains), O2 (11 strains, including MT78), O1 (five strains) or other serogroups (24 strains), was tested in a double PCR assay. The results for 19 representative strains are shown in Fig. 6. Gene fimA was amplified from all of the strains tested, giving a 331-bp product, whereas only strains of the O2 serogroup gave positive results in the second PCR, yielding a characteristic 267-bp fragment (Fig. 6, lanes 7–16). None of the 22 O78 and five O1 strains tested gave positive results in this specific PCR (data not shown). Furthermore, all of the non-typable avian strains of E. coli (not O1, O2 or O78) gave negative results in this PCR assay. One of the O2 strains failed to give the 267-bp specific product (data not shown).

Discussion

The fim cluster of MT78, an avian pathogenic strain of E. coli, was cloned and sequenced. Among the five

![Fig. 4. Multiple alignment of FimA amino-acid sequences. fimA genes from MT78 and various strains of E. coli and K. pneumoniae were translated, then aligned with the 'PileUp' programme in the GCG package. Consensus represents identical residues at the same position in five or six sequences. Only differences relative to the consensus sequence are shown. Stars represent gaps, which were added in order to optimise the alignment. The sequences are designated by their mnemonics in the sequence database: Ecopilaa is from a human clinical isolate of E. coli, J96 [27]; Ecfima01, from an E. coli K12 strain [10]; Ecotypelf from an avian strain of E. coli, (PDI-386), serogroup O78 [6]; Kpnfamb from a urinary isolate of K. pneumoniae, strain IA551 [26]; and Kpnmfimb, from another urinary isolated of K. pneumoniae, strain IAS565 [28].](image)
Fig. 5. Structure predictions for FimA polypeptide. The hydrophilicity, surface probability and antigenic index were calculated for MT78-FimA polypeptide with the Peptidestructure and Plotstructure programmes of the GCG package [17, 29]. The hydrophilicity plot of the Kyte-Doolittle algorithm was used with a window size of 7. Peaks represent hydrophilic regions in the first plot and surface-exposed residue in the second plot. Numbering of the amino-acid residues begins at the N-terminal end of the signal peptide.
genes sequenced in this work, *fimB* and *fimE*, whose products are involved in the regulation of phase variation, were the most conserved relative to the sequences already present in the data library. The sequence of gene *fimC*, which encodes a chaperone-like protein, was almost identical to the homologous sequence provided by Jones *et al.* [21], but differed from that published by Klemm [20] by 13 nucleotide substitutions. As it has been shown that FimC exhibits a relative specificity towards its substrates [31], it can be speculated that the four amino-acid substitutions observed between MT78 FimC and the homologous protein from *E. coli* K12 might be related to the sequence differences of the major and minor subunits between the two strains (see below).

The first complete sequence of gene *fimI* from an *E. coli* strain has been determined. Although the presence of a *fimI* gene downstream of *fimA* [20] and of its counterpart in the *foc* operon encoding type 1C fimbriae [24] has been noticed by others, no complete sequence was available in sequence databases. However, the published sequences of *fimA* from *E. coli* [10] and *K. pneumoniae* [26] contain the beginning of a *fimI* ORF, and a truncated *foci* ORF can be recognised as well in the sequence of type 1C fimbriae [25]. Computer analyses of its deduced amino-acid sequence showed that FimI was structurally related to other fimbrial subunits. Furthermore, analysis of its amino-acid composition revealed the presence of tryptophan residues which are never found in major fimbrial subunits. Therefore, it is likely that FimI constitutes another minor fimbrial subunit. A comparison of the MT78 *fimI* sequence with homologous sequences found in databases revealed a close relationship with gene *foci*, from type 1C fimbriae of *E. coli*, and a more remote homology with gene *fimI* of *S. Typhi*. The few sequences available in databases did not reveal any obvious motif particular to FimI products that would not be present in other fimbrial subunits. Recently, Rossolini *et al.* [23] have proposed that protein FimI of *S. Typhi* could play the same role as PapH, encoded by the *pap* operon, in anchoring fimbriae to the membrane [32]. However, comparison of the sequence of MT78-FimI with either *S. Typhi* FimI or *E. coli* PapH (data not shown) provided no evidence to support this hypothesis, and more experimental data are needed to elucidate its role in the structural organisation and biogenesis of the fimbriae. Previous work has shown that strain MT78, and other *E. coli* O2 strains of avian origin, produce a variant of
type 1 fimbriae with characteristic unusual size and antigenic properties [9]. These observations were confirmed further by the nucleotide sequence of fimA and the amino-acid sequence of its product. A comparison of MT78 fimA with homologous sequences from other strains identified two closely related sequences, one from an avian pathogenic O78 strain of E. coli [8], the other from K. pneumoniae [26]. It remains to be seen whether type 1 fimbriae produced by avian strains of E. coli share particular properties conferring some specificity for avian species.

Comparison of the MT78 fimA sequence with homologous sequences from other strains of E. coli or K. pneumoniae and their subsequent multiple alignment allowed the identification of four short variable regions within this gene. A computer prediction of FimA structure showed that at least two of the corresponding sites within the polypeptide (aa 104–109 and 134–137) constitute hydrophilic clusters which are likely to be exposed at the surface of the protein. These findings suggest that these regions may be antigenic sites, and this could account for the distinct antigenic properties of MT78 FimA. It is tempting to speculate that these sites constitute hypervariable regions in the major subunit, like those recently described by Maiti et al. [33] in the major subunits of fimbriae of the P family.

The identification of short variable regions within gene fimA, and particularly the presence of an insertion at the beginning of the mature protein, characterising strain MT78, prompted us to design a PCR assay that would detect specifically strains sharing the same features as MT78. All the strains tested possessed gene fimA, but only O2 strains were identical to MT78 in the specific PCR assay in yielding a 267-bp fragment. This result showed that a major proportion of O2 strains of avian origin share with MT78 the characteristic N-terminal insertion and an identical third variable site within their major fimbrial subunit. However, the fact that one O2 strain failed to give a positive result in this second PCR assay suggests a genotype diversity within this serogroup, regarding the fimA gene at least. This is in agreement with previous findings in this laboratory, which showed that, whereas the estimated mol. wt of FimA from MT78 and the majority of O2 strains is 18.5 kDa, in one O2 strain it was 18 kDa [9]. Furthermore, such a genotype diversity has been documented by other methods among O2 strains of avian origin [34].

Several variants of type 1 fimbriae have already been described for human [10, 27, 30, 35] and avian isolates [7–9]. The present study gives more precision in showing that variations in FimA are clustered in short regions that probably correspond to antigenic sites and could be related to the serogroup of the strains. These small variations could modulate the function and specificity of type 1 fimbriae. This hypothesis is in agreement with the recent findings of Sokurenko et al. [36, 37] that minor substitutions within the adhesin FimH can result in important variations in adhesion phenotype.

Although type 1 fimbriae have long been considered as an homogeneous group, conferring some unique properties, these recent studies show that several type 1 variants exist that could be related to particular properties of these fimbriae in host-bacterial relationship. More knowledge in this field could shed new light on the controversial role of type 1 fimbriae in pathogenicity.

We are grateful to Annette Bree for her skilful assistance in the purification of fimbriate and Dr I. C. Blomfield for providing us with strain AEC185.

References
15. Lau CF, Sheu SS. Rapid and direct recovery of DNA