

Escherichia coli F41 Adhesin: Genetic Organization, Nucleotide Sequence, and Homology with the K88 Determinant

DENISE G. ANDERSON AND STEVE L. MOSELEY*

Department of Microbiology, School of Medicine, University of Washington, Seattle, Washington 98195

Received 7 April 1988/Accepted 22 July 1988

The genetic organization of the polypeptides required for the biosynthesis of the F41 adhesin of enterotoxigenic *Escherichia coli* strains was investigated. Maxicell analysis demonstrated that a recombinant plasmid which mediated mannose-resistant hemagglutination and F41 antigen production encoded four polypeptides of 29, 30, 32, and 86 kilodaltons. The 29-kilodalton protein was identified as the F41 antigen, and the nucleotide sequence of the gene was determined. Extensive homology was observed between the region encoding the putative signal sequences of the F41 and K88 antigens and in the region immediately upstream of the antigen genes. The nucleotide sequence homology between F41 and K88 determinants was further investigated by Southern blot hybridization. A K88 probe hybridized at high stringency to all fragments shown to be essential for F41 production except for fragments internal to the F41 antigen gene.

The pathogenicity of enterotoxigenic *Escherichia coli* (ETEC) involves adherence of the pathogen to the small intestine of the host, where enterotoxins cause secretion of fluid and electrolytes. Adherence is mediated by antigenic proteins which are fimbrial in nature and referred to as adhesins (5). Several distinct types of adhesins of ETEC have been characterized, each of which confers a degree of host specificity. Many of the adhesins also confer on the bacterial cells the ability to agglutinate erythrocytes of various species in the presence of mannose, in contrast to hemagglutination mediated by type I fimbriae, which is mannose sensitive. The mannose-resistant hemagglutination (MRHA) assay is a convenient means of assessing the presence of specific adhesin types. The F41 antigen is associated with an adhesin of bovine and porcine ETEC isolates and is frequently produced by strains of serogroups O9 and O101, which also produce the K99 adhesin (6, 17, 27). The presence of the F41 antigen confers on bacterial cells the ability to agglutinate human group O erythrocytes in the presence of mannose.

In contrast to most other known adhesins of ETEC, which are plasmid encoded, F41 is chromosomally encoded (18). The genetic determinants for F41 have been cloned and shown to have nucleotide sequence homology with the plasmid-encoded K88 genes (18). In addition, several MRHA-positive strains which do not produce K88 or F41 antigens hybridized with probes from the K88 and F41-encoding DNA (18).

In this study we investigated the organization and expression of genes involved in F41 biosynthesis, and we report the nucleotide sequence of the gene encoding the F41 structural subunit. We demonstrate that homology between the F41 and K88 determinants is extensive and includes regions encoding each of the known accessory proteins but not the fimbrial subunits themselves.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture conditions. All strains used in this study were derivatives of *E. coli* K-12. Strain DH5 α (Bethesda Research Laboratories, Gaithersburg, Md.) was used as a host for the plasmids. Strain HB101

(3) was used for transposon mutagenesis, and strain JC2926 (1) was used in maxicell analysis. A derivative of the multicopy plasmid pUC18 (30), pUC18cm^r (provided by M. Yanofsky), was used as a vector for cloning. In this vector, the gene specifying β -lactamase is replaced by the chloramphenicol acetyltransferase determinant. Unless otherwise indicated, organisms were grown in Luria broth or Luria agar (12). Cells harboring pSLM204 (18) were grown in the presence of ampicillin (100 μ g/ml), and cells harboring derivatives of pUCcm^r were grown in the presence of chloramphenicol (25 μ g/ml). Kanamycin (25 μ g/ml) was used to select for transposon Tn5 insertions.

MRHA assay. Bacterial cells were tested for the ability to agglutinate human group O erythrocytes in the presence of mannose as previously described (18).

Southern blot hybridizations. Preparation, hybridization, and washing of blots under conditions of high stringency were performed as described previously (19, 20, 26).

Insertion mutagenesis. The transposon Tn5 was introduced into *E. coli* HB101(pSLM204) by infection with bacteriophage λ b221 rex::Tn5 cI857 Oam8 Pam29 (4). Transductants growing on medium containing both ampicillin and kanamycin were harvested, and total plasmid DNA was isolated. The resulting pool of plasmid DNA was used to transform strain HB101 for ampicillin and kanamycin resistance, selecting for plasmids into which the transposon had been inserted. Plasmid DNA was isolated from individual transformants, and restriction analysis identified the site of transposon insertion.

DNA manipulation and analysis. Plasmid DNA was isolated by the method described by Birnboim and Doly (2) or by cesium chloride-ethidium bromide density gradient centrifugation following alkaline lysis as described by Maniatis et al. (11). Restriction endonuclease digests were carried out according to the manufacturers' recommendations (Bethesda Research Laboratories, Gaithersburg, Md.; Boehringer Mannheim Biochemicals, Indianapolis, Ind.; International Biotechnologies, Inc., New Haven, Conn.). Digests were analyzed by electrophoresis in agarose or polyacrylamide gels.

Expression of plasmid-encoded proteins in maxicells. Maxicell analysis was performed by a modification of the method described by Sancar et al. (23). Overnight cultures of organ-

* Corresponding author.

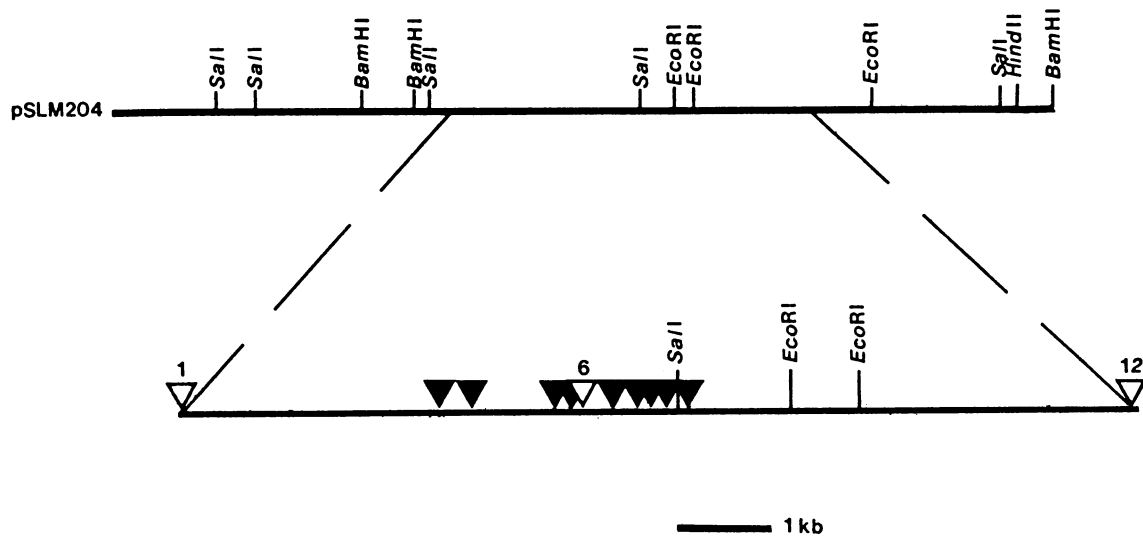


FIG. 1. Sites of transposon insertion in pSLM204. Solid triangles represent Tn5 insertions which abolished the MRHA phenotype; open triangles represent Tn5 insertions which retained the MRHA phenotype. Numbers above triangles identify specific inserts referred to in the text. All transposon insertions outside of the enlarged area are MRHA positive and are not shown. The size scale refers to the enlarged region.

isms harboring plasmids of interest were diluted and grown at 37°C in K-broth (22) to exponential phase and then exposed to UV radiation from a germicidal lamp. Following irradiation, cultures were incubated for 1 h at 37°C, and then cycloserine was added to a concentration of 200 µg/ml. Incubation was continued for 12 to 16 h at 37°C. Cells were harvested and washed three times with Hershey salts (23) and then incubated for 30 to 60 min at 37°C in Hershey medium (23). [³⁵S]Methionine (1,120 Ci/mmol) (ICN Biomedicals, Costa Mesa, Calif.) was then added to a final concentration of 5 µCi/ml, and incubation was continued for an additional hour. Cells were then washed twice with Hershey salts. The radioactive polypeptides were separated on a denaturing 12.5% (wt/vol) polyacrylamide-SDS gel as described by Laemmli (10) and visualized following autoradiography.

Immunoblot procedures. Proteins of overnight cultures of *E. coli* were separated by gel electrophoresis on a denaturing 12.5% polyacrylamide-SDS gel as described by Laemmli (10). Transfer of the proteins to nitrocellulose filters was accomplished by the methods of Towbin et al (28). Filters were preadsorbed in TN buffer (50 mM Tris hydrochloride [pH 7.5], 150 mM NaCl) with 3% bovine serum albumin (BSA) at 37°C for 45 min. Filters were then incubated for 1 h at 20°C in anti-F41 antiserum (18) which had been diluted 1:100 in TN buffer with BSA and adsorbed with *E. coli* HB101. The filters were then washed once in TN buffer with 0.1% SDS and then four times with TN buffer. The filters were then incubated for 1 h at 20°C in TN with BSA to which ¹²⁵I-protein A (10 µCi/µg) (New England Nuclear Corp., Boston, Mass.) was added to a concentration of 125 nCi/ml. The filters were then washed as described above and exposed to X-ray film with intensification screens.

DNA sequence determination. The dideoxy chain termination method described by Sanger et al. (24) was used, with the Klenow fragment of *E. coli* DNA polymerase (Boehringer Mannheim Biochemicals), avian myeloblastosis virus reverse transcriptase (New England Nuclear Corp.), and modified T7 DNA polymerase (United States Biochemicals, Cleveland, Ohio). Bacteriophage M13mp18 and M13mp19

derivatives were used as the source for single-stranded DNA templates (21).

RESULTS

Localization of the cloned F41 determinants. The determinants for F41 biosynthesis are located on the 34-kilobase (kb) recombinant plasmid pSLM204, as described previously (18). Organisms harboring this plasmid produce the F41 antigen and are MRHA positive with human group O erythrocytes. In order to identify more precisely the region of pSLM204 which encodes the F41 determinants, transposon mutagenesis was used. A total of 170 independent Tn5 insertions in pSLM204 were screened for MRHA. The locations of 30 randomly selected insertions which abolished the MRHA phenotype and 29 randomly selected insertions which did not affect the phenotype were determined by restriction mapping.

The locations of selected transposon insertions in pSLM204 and their effects on the MRHA phenotype are shown in Fig. 1. The sequences required for F41 production appeared to be contained in a 10.4-kb region, the boundaries of which are defined by pSLM204::Tn5-1 and pSLM204::Tn5-12, which retained the MRHA phenotype. All Tn5 insertions outside of the 10.4-kb region retained the MRHA phenotype. Strains harboring plasmids with transposon insertions in the 7.8-kb *EcoRI*-*Bam*HI fragment generally lost the MRHA phenotype. However, one transposon insert (Fig. 1, Tn5-6) in the 7.8-kb *EcoRI*-*Bam*HI fragment retained the MRHA phenotype although surrounding inserts are MRHA negative.

Production of F41 antigen. Insertion mutagenesis of pSLM204 indicated that the sequences essential for F41 production lay within a 7.8-kb *EcoRI*-*Bam*HI fragment. We therefore cloned this fragment into pUC18cm^r. The resulting plasmid, designated pDGA9, directed synthesis of the F41 antigen and conferred MRHA ability. A restriction map of pDGA9 is shown in Fig. 2. Additional derivatives of pDGA9 were constructed in order to determine the expression and organization of the genes involved in F41 biosynthesis (Fig. 2).

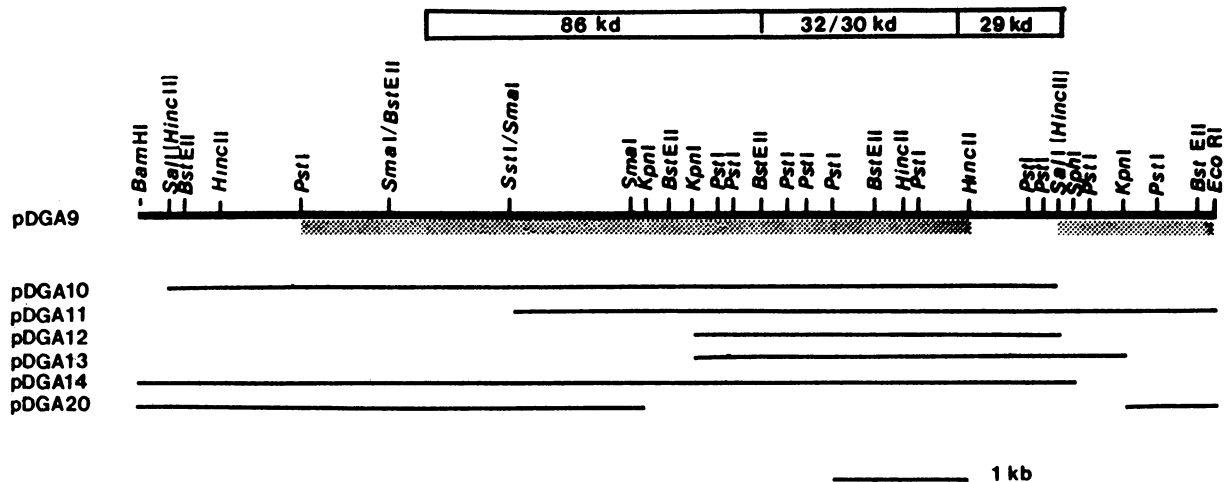


FIG. 2. Plasmid pDGA9 and derivatives. The bars represent the portions of pDGA9 incorporated into each plasmid. Vector sequences are not shown. pDGA14 and pDGA20 are deletion derivatives of pDGA9. pDGA10, pDGA12, and pDGA13 were constructed by inserting the indicated fragments into pUC18cm^r. pDGA11 was constructed by insertion of the indicated fragment into pUC19cm^r. Peptides concluded to be encoded by pDGA9 are indicated by size in boxes above the region where each peptide is encoded. Regions of DNA homology between F41 and K88, as described in the text, are indicated by the shaded bar.

In order to define the region which encodes the F41 fimbrial subunit, the production of F41 antigen by pDGA9 and its derivatives was examined by Western blot (immunoblot) analysis. Organisms harboring plasmids pDGA9, pDGA11, and pDGA13 each produced a protein with an apparent molecular mass of 29 kilodaltons (kDa) which bound F41-specific antibodies (Fig. 3). Proteins encoded by pDGA10 and pDGA12 did not bind the anti-F41 antibodies. Since pDGA13 (antigen positive) differs from pDGA12 (antigen negative) only in an additional 0.5-kb *SalI-KpnI* fragment which itself is not large enough to encode the 29-kDa F41 antigen, the gene encoding the F41 antigen was concluded to span the *SalI* site.

Expression and organization of genes involved in biosynthesis of the F41 adhesin. Gene products associated with F41 biosynthesis were identified by labeling proteins encoded by

each of the recombinant plasmids in *E. coli* maxicells. Plasmid pDGA9 directed the synthesis of at least four polypeptides expressed from the cloned DNA, in addition to the vector-encoded chloramphenicol acetyltransferase. The apparent molecular masses of these polypeptides were 86, 32, 30, and 29 kDa (Fig. 4, bands a, b, c, and d, respectively). The 29- and 30-kDa proteins were within the range reported for the adhesin subunit (29.5 kDa). Plasmids pDGA9, pDGA10, and pDGA14 directed the synthesis of the 86-kDa polypeptide, whereas pDGA11, pDGA12, and pDGA13 did not. These data suggested that the gene encoding the 86-kDa polypeptide either was located within the *SalI-SstI* fragment or spanned the *SstI* site. In order to more precisely determine the location of the gene encoding the 86-kDa protein, plasmid pDGA20 was constructed by deletion of a *KpnI* fragment of pDGA9 (Fig. 2). Maxicell analysis of this plasmid demonstrated the synthesis of a 60-kDa polypeptide

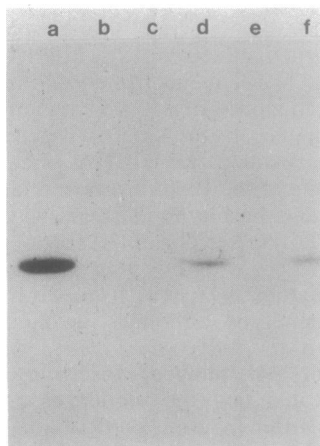


FIG. 3. F41 antigen production by pDGA9 and its derivatives. Proteins from *E. coli* harboring various plasmids were separated by SDS-polyacrylamide gel electrophoresis, transferred to nitrocellulose by electroblotting, and probed with F41 antiserum. Bound antibody was detected with ¹²⁵I-protein A and visualized by autoradiography. Lane a, pDGA13; lane b, pDGA10; lane c, pDGA12; lane d, pDGA11; lane e, pUC18cm^r; lane f, pDGA9.

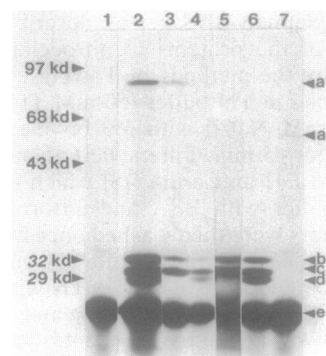


FIG. 4. Polypeptide synthesis directed by pDGA9 and its derivatives in *E. coli* maxicells. Maxicells were labeled with [³⁵S] methionine, solubilized, and analyzed on an SDS-polyacrylamide gel. An autoradiograph of the gel is shown. Lane 1, pUC18cm^r; lane 2, pDGA9; lane 3, pDGA10; lane 4, pDGA11; lane 5, pDGA12; lane 6, pDGA13; lane 7, pDGA20. Arrows a through d and a' indicate polypeptides referred to in the text. Arrow e indicates the vector-encoded chloramphenicol acetyltransferase. Positions of size standards are indicated on the left.

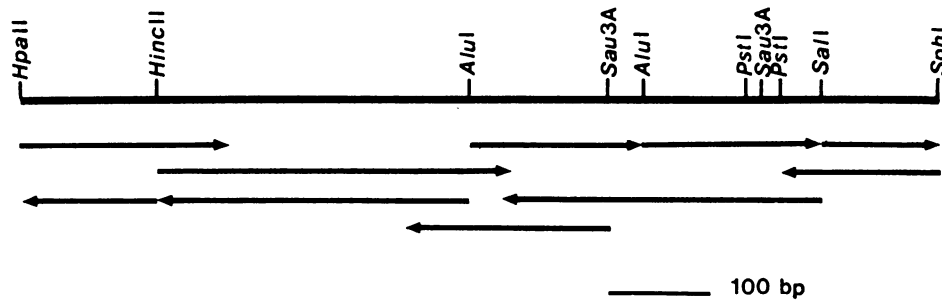


FIG. 5. Detailed physical map of the region encoding the F41 antigen. Arrows indicate fragments which were used to determine the nucleotide sequence and the direction of polymerization in the sequencing reactions.

(Fig. 4, band a'), probably a truncated form of the 86-kDa protein. Thus, the gene encoding the 86-kDa protein spanned the 1-kb *Sst-KpnI* fragment (Fig 2).

None of the plasmids constructed directed exclusive synthesis of the 29-kDa polypeptide. However, the 0.5-kb *Sall-KpnI* fragment appeared to be important for its expression. Plasmid pDGA12 directed the synthesis of only the 32- and 30-kDa polypeptides, whereas pDGA13, which contained the additional 0.5-kb *Sall-KpnI* fragment, directed the synthesis of the 32-, 30-, and 29-kDa polypeptides. Since sequences within this fragment also appear to be essential for F41 antigen production, the 29-kDa polypeptide was concluded to be the K41 antigen.

Nucleotide sequence of the F41 gene. Since the immunological and maxicell data suggested that the region encoding the F41 subunit spanned the *Sall* site of pDGA13, we undertook a DNA sequence analysis of that region of the plasmid. A more detailed restriction map of the region upstream of *Sall* was constructed, and appropriate restriction fragments were cloned into M13mp18 and M13mp19 for nucleotide sequence determination (Fig. 5). An 897-base open reading frame was identified (Fig. 6). Of two potential initiator codons at positions 94 and 154, only the codon at position 94 was preceded by sequences indicative of the presence of a ribosome-binding site (25). Assuming that translation is initiated at this site, a polypeptide of 277 amino acids would be encoded by the sequence. The amino-terminal amino acid sequence suggested the presence of a signal sequence for transmembrane secretion of the protein. A signal peptidase cleavage site was predicted between Ala at position 22 and Ala at position 23 by the weighted matrix method of von Heijne (31). Cleavage at this site would result in a mature polypeptide of 255 amino acids, with an amino-terminal sequence in close agreement with the F41 amino acid sequence reported previously (6). The only discrepancy is Ile at position 34, which was reported as Leu.

The sequence analysis also revealed homology with the previously reported nucleotide sequence upstream of and extending into the coding region of the gene encoding the K88 antigen (7) (Fig. 6).

Genetic homology of F41 and K88 accessory proteins. The genetic determinants of the K88 and F41 adhesins show extensive DNA homology (18). Since no homology was observed in the nucleotide sequences of the mature K88 and F41 fimbrial subunits, we concluded that the genetic homology of the adhesin determinants is confined to intergenic regions or sequences encoding accessory proteins involved in adhesin biosynthesis. The genetics and functions of the accessory proteins in K88 biosynthesis have been studied extensively (15, 16, 29). Since accessory proteins of F41 which share DNA homology to K88 accessory proteins

probably have similar functions, we sought to more precisely localize the regions of DNA homology. The F41-encoding plasmid pDGA9 was digested with *PstI*, *BstEII*, and *HincII* and then probed in Southern blot experiments under conditions of high stringency with pMK005, a recombinant plasmid which mediates K88 expression (8). With the exception of fragments in the 1.2-kb *PstI* fragment (at the left of the fragment shown in Fig. 2) and the fragments internal to the F41 antigen gene, every restriction fragment hybridized to the K88 probe (data not shown). Regions of homology are shown in Fig. 2. This indicated that extensive nucleotide sequence homology exists throughout the accessory protein-encoding genes of the K88 and F41 adhesins.

Since previous probes for K88 and F41 genetic determinants were entirely cross-reactive (18), we constructed F41- and K88-specific probes by using fragments internal to the adhesin subunits. The 617-bp *HincII-PstI* fragment of pDGA9, which is internal to the F41 subunit gene, did not hybridize at high stringency with pMK005, which encodes K88. Conversely, the 382-bp *EcoRI-HincII* fragment of pMK005, which is internal to the K88 subunit gene, did not hybridize at high stringency with pDGA9.

DISCUSSION

We have determined the genetic organization of at least four polypeptides which are involved in biosynthesis of F41 and have identified and sequenced the gene encoding the F41 antigen. We have shown that the homology between the K88 and F41 determinants spans the entire region which mediates adhesin biosynthesis, exclusive of the antigen genes themselves.

The genetic organization and apparent molecular masses of the proteins required for biosynthesis of F41 are similar to those reported for the accessory proteins of other fimbrial adhesins, including type 1, P fimbriae, K99, and K88 (5). The 86-kDa protein associated with F41 production is presumably functionally analogous to the 76- to 86-kDa proteins encoded by other adhesin determinants (5, 16). These large proteins have been suggested to serve as an anchor or pore in the outer membrane through which the adhesin subunits are attached or extruded (5, 16). Maxicell analysis of *E. coli* harboring plasmids which contain the region immediately downstream of the gene encoding the F41-associated 86-kDa protein demonstrated that this region directs the synthesis of two proteins of 32 and 30 kDa. While these may represent precursor and mature forms of the same protein, we have been unable to demonstrate that inhibition of protein processing results in the accumulation of the 32-kDa protein (data not shown). Biosynthesis of other adhesins has been shown to involve secreted accessory proteins of similar sizes (5, 9,

	10	20	30	40	50	60	70	80	90
K88	<u>AACGGAGCCCGGGATGGTTTTACGGTAATCCGGAAAAATAAGGGTTACCGATT</u> <u>CAGTTTATTATTTGTGGA-TATCAAGGGGTTAAT</u>								
F41	TGAATCCGCAGGGATGGTTTTGCGGTAATCCGGAAAAATAAGGATTACCGATT								CAGTTTATTATTTATGAAATGTC <u>CAAGGT</u> GATTAT-
	100	110	120	130	140	150	160	170	180
K88	<u>TTTATGAAAAAGACTCTGATTGCA</u>								
F41	TTTATGAAAAAGACTCTGATTGCACTGGCTGTGGCTGCATCAGCGGCAGTATCTGGTTCAGTGATGGCTGCTGATTGGACGGAAGGTCAA								
		MetLysLysThrLeuIleAlaLeuAlaValAlaAlaSerAlaAlaValSerGlySerValMetAlaAlaAspTrpThrGluGlyGln							
		5	10	15	20	25			
	190	200	210	220	230	240	250	260	270
	CCAGGTGACATTATTATTGGTGGTGAATTACATCACCATCTGTTAAATGGCTATGGAAGACTGGAGAGGGACTTTCATCTTTTAGCAAT								
	ProGlyAspIleIleIleGlyGlyGluIleThrSerProSerValLysTrpLeuTrpLysThrGlyGluGlyLeuSerSerPheSerAsn								
	30	35	40	45	50	55			
	280	290	300	310	320	330	340	350	360
	ACTACAAATGAAATTGTTAAACGGAAGTTGAATATTTCTGTTCCAACGGATGAATTATTTTAGCAGCGAAGATGAGTGATGGGATTAA								
	ThrThrAsnArgIleValLysArgLysLeuAsnIleSerValProThrAspGluLeuPheLeuAlaAlaLysMetSerAspGlyIleLys								
	60	65	70	75	80	85			
	370	380	390	400	410	420	430	440	450
	GGTGTTCCTAGGGAATCACCTCATTCTAAGATTGAAATGGCATCTTATGATGGTAGTGTTATTACACCTAGTTTCACTTCAAATACA								
	GlyValPheValGlyAsnThrLeuIleProLysIleGluMetAlaSerTyrAspGlySerValIleThrProSerPheThrSerAsnThr								
	90	95	100	105	110	115			
	460	470	480	490	500	510	520	530	540
	GCAATGGATATTGCTGTAAAAGTAAAAACTCAGGTGATAATACTGAGCTAGGGACTCTTCTGTTCCTTTGTCATTTGGTGCAGGAGTT								
	AlaMetAspIleAlaValLysValLysAsnSerGlyAspAsnThrGluLeuGlyThrLeuSerValProLeuSerPheGlyAlaAlaVal								
	120	125	130	135	140	145			
	550	560	570	580	590	600	610	620	630
	GCAACTATTTTTGATGGCAATACTACTGATAGCGCTGTAGCGCATATTACCAGTGGTCTGCTGGTACAGTATTTGAAGGGCTTGTTAAT								
	AlaThrIlePheAspGlyPheThrThrAspSerAlaValAlaHisIleThrSerGlySerAlaGlyThrValPheGluGlyLeuValAsn								
	150	155	160	165	170	175			
	640	650	660	670	680	690	700	710	720
	CCAGGTCGATTTACTGATCAGAATATAGCCATAAATGGAATGGACTCTCAAAGCTGAAATGGCTGGTTATGTAGAAAAGTTAATGCCA								
	ProGlyArgPheThrAspGlnAsnIleAlaTyrLysTrpAsnGlyLeuSerLysAlaGluMetAlaGlyTyrValGluLysLeuMetPro								
	180	185	190	195	200	205			
	530	540	550	560	570	580	590	600	610
	GGGAAAAGTGCTTCAACCTCTTATAGTGGTTCCACAATTGGGATGACCTCAGTCACCCCAACTATACTTCTGCAGATAAGGCATCTTAT								
	GlyLysSerAlaSerThrSerTyrSerGlyPheHisAsnTrpAspAspLeuSerHisProAsnTyrThrSerAlaAspLysAlaSerTyr								
	210	215	220	225	230	235			
	620	630	640	650	660	670	680	690	700
	CTCTCTTATGGATCTGGTGTCTTCTGCAGGTAGTACTTTAGTTATGAATTTAAATAAGGATGTTGCGGGTGCAGTTGAATGGGTGGCTCCA								
	LeuSerTyrGlySerGlyValSerAlaGlySerThrLeuValMetAsnLeuAsnLysAspValAlaGlyArgLeuGluTrpValAlaPro								
	240	245	250	255	260	265			
	710	720	730	740	750	760	770	780	790
	GTGACTATACCGTTATTATAGTTAATCATCTTTCTGCCCTCCCCTGTTTTATTTTCGGGGGAAGGGCGGTAGTTGGGGGCAGTCAG								
	ValThrIleThrValIleTyrSer *								
	270	275							

FIG. 6. Nucleotide sequence of the gene encoding the F41 antigen, and the deduced amino acid sequence of the F41 antigen. The portion of the sequence of the gene encoding the K88 antigen (6) with homology to the F41 sequence is shown directly above the F41 sequence. Nucleotides which are identical in K88 and F41 are underlined in the K88 sequence. Gaps in the K88 and F41 sequences necessary to maintain alignment are indicated by hyphens. Sequences encoding a presumptive ribosome-binding site (25) are underlined in the F41 sequence. *, Terminator codon.

15). Mooi et al. (13, 16) have suggested that a 27-kDa protein which is encoded in a similar position in the K88 operon is involved in transport and modification of the fimbrial subunit in the periplasmic space. Unlike the accessory proteins of the other adhesins, however, the 30- and 32-kDa proteins of F41 appear to be encoded in a transcriptional unit separate from that of the 86-kDa protein. These two proteins are encoded 3' to the 86-kDa protein determinant, yet they are still expressed when the region encoding the 86-kDa protein is deleted, regardless of the orientation relative to the vector-encoded promoter. These observations suggest that the 30- and 32-kDa proteins are transcribed from a separate promoter.

Two additional accessory proteins associated with K88 production (14) were not detected by maxicell analysis of plasmids mediating production of F41. Preliminary sequencing data, however, suggest the presence of additional genes upstream of the region encoding the 86-kDa protein associated with F41 production and downstream of the F41 antigen gene (manuscript in preparation). The region upstream of the 86-kDa protein gene is homologous to a region of the K88 operon previously shown to encode a fimbrial-like protein (14). Although expression of the encoded protein has not been demonstrated, Mooi and co-workers reported that deletion of the gene in K88 prevented expression of fimbriae (14). A gene immediately downstream of the F41 antigen gene appears to be in the same transcriptional unit as the F41 antigen gene and thus is likely to be involved in fimbria production. However, organisms harboring pDGA14 (from which this region has been deleted) are MRHA positive, suggesting that it is not essential for fimbria production. Mooi and co-workers (16) have reported a 27.5-kDa protein encoded in a similar position in the K88 operon which also appears to be involved in, but not essential for, fimbria production.

We have shown that the F41 antigen is encoded in a region downstream of genes encoding at least three accessory proteins, a genetic organization similar to that reported for the K88 determinants (13, 16). Like other fimbrial subunits, the encoded F41 protein is a precursor of the mature protein, containing a signal sequence of 22 amino acids. The predicted amino acid composition of the encoded protein differs significantly from the previously reported amino acid composition of the F41 antigen purified from strain B41M (6). DNA sequence analysis of the F41 subunit gene of strain B41 (unpublished observations) does not reveal divergence extensive enough to explain this discrepancy.

Transposon mutagenesis data raise the possibility that the F41 subunit gene may be preceded by its own promoter. One transposon insert which mapped immediately upstream of the subunit gene retained MRHA ability (Fig. 1). Kehoe et al. (8) observed a similar effect of a Tn5 insert in pMK005 which mapped immediately upstream of the K88 subunit gene. The nucleotide sequence of the F41 subunit gene shows a region resembling a Pribnow box located upstream of the gene, but there is no apparent -35 region. Occasionally however, Tn5 insertions are nonpolar (4), and transcription of the F41 subunit in this strain may result from a promoter internal to Tn5 or from the fortuitous creation of promoter sequences at the Tn5-F41 gene junction.

Limited homology was observed between the predicted amino acid sequence of the F41 fimbrial subunit and reported sequences of other adhesin antigens. Characteristics of the F41 subunit which are shared by K88, K99, type 1, CFAI, and PapA include a hydrophobic carboxy-terminal end and a conserved tyrosine in the penultimate amino acid position

(5). While most fimbrial subunits are characterized by a pair of cysteines (7, 9), F41 is similar to K88 and CFAI in lacking this amino acid.

The DNA homology between K88 and F41 determinants is extensive and appears to include sequences encoding each of the accessory proteins involved in adhesin biosynthesis but not in genes encoding the fimbrial antigens themselves. Comparison of nucleotide sequences from the region immediately upstream of the F41 and K88 subunit genes reveals a high degree of homology in the 102 nucleotides that precede the subunit genes (Fig. 6). The homology extends into the first 21 nucleotides of the region encoding the signal sequence of the subunits. With the exception of the signal sequence, however, Southern blot hybridization and nucleotide sequencing data indicate that the genes for the K88 and F41 subunits are entirely nonhomologous. With extensive homology in genes that flank the fimbrial subunit genes but complete lack of homology in the subunit genes themselves (Fig. 2), it appears that in the evolution of K88 and F41, a fimbrial subunit gene has been replaced by an entirely new subunit gene.

Since previous K88 and F41 probes are entirely cross-reactive (18), it has been impossible to determine whether any strains harbor the genetic information for production of both K88 and F41. Using the 617-bp *HincII-PstI* fragment (Fig. 2), we have constructed a probe which is internal to the F41 subunit and does not hybridize to the K88 determinant. At least one K88-producing strain has been identified which hybridizes at high stringency to the F41-specific probe (R. Schneider, personal communication). The extensive homology between K88 and F41 accessory proteins and the presence of an F41 subunit gene, or a sequence homologous to it, in a K88-producing strain suggest a potential mechanism for antigenic variation in response to antibody pressure. We are currently attempting to complement mutations in F41 accessory proteins with K88 proteins in order to determine whether the accessory systems of the two adhesins are interchangeable. The ability of an adhesin synthesis and assembly system to rapidly evolve by a genetic exchange involving only the fimbrial subunit gene would have significant implications for vaccine development strategies based on antifimbrial antibody production.

ACKNOWLEDGMENTS

This work was supported by grants 85-CRCR-1-1907 and 87-CRCR-1-2426 from the U.S. Department of Agriculture.

We thank Dalia Alfi for technical assistance and Harley Moon, Gordon Dougan, Tom Casey, Robert Schneider, and Stephen Lory for helpful discussions.

LITERATURE CITED

1. Bachman, B. 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriol. Rev.* **36**:525-557.
2. Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**:1513-1523.
3. Bolivar, F., and K. Backman. 1979. Plasmids of *Escherichia coli* as cloning vectors. *Methods Enzymol.* **68**:245-267.
4. Bruijn, F. J., and J. R. Lupski. 1984. The use of transposon Tn5 mutagenesis in the rapid generation of correlated physical and genetic maps of DNA segments cloned into multicopy plasmids—a review. *Gene* **27**:131-149.
5. deGraaf, F., and F. Mooi. 1986. The fimbrial adhesins of *Escherichia coli*. *Adv. Microb. Physiol.* **28**:65-142.
6. deGraaf, F. K., and I. Rooda. 1982. Production, purification, and characterization of the fimbrial adhesive antigen F41 isolated from calf enteropathogenic *Escherichia coli* strain B41M.

- Infect. Immun. **36**:751-758.
7. Gaastra, W., F. Mooi, A. Stuitje, and F. deGraaf. 1981. The nucleotide sequence of the gene encoding the K88ab protein subunit of porcine enterotoxigenic *Escherichia coli*. FEMS Microbiol. Lett. **12**:41-46.
 8. Kehoe, M., R. Sellwood, P. Shipley, and G. Dougan. 1981. Genetic analysis of K88-mediated adhesion of enterotoxigenic *Escherichia coli*. Nature (London) **291**:122-126.
 9. Klemm, P. 1981. The complete amino acid sequence of the K88 antigen, a fimbrial protein from *Escherichia coli*. Eur. J. Biochem. **117**:617-627.
 10. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680-685.
 11. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 12. Miller, J. 1972. Experiments in molecular genetics, p. 433. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 13. Mooi, F., N. Harms, D. Bakker, and F. deGraaf. 1981. Organization and expression of genes involved in the production of the K88ab antigen. Infect. Immun. **32**:1155-1163.
 14. Mooi, F., M. vanBuuren, G. Koopman, R. Roosendaal, and F. deGraaf. 1984. K88ab gene of *Escherichia coli* encodes a fimbria-like protein distinct from the K88ab fimbrial adhesin. J. Bacteriol. **159**:482-487.
 15. Mooi, F., A. Wijffjes, and F. deGraaf. 1983. Identification and characterization of precursors in the biosynthesis of K88ab fimbria of *Escherichia coli*. J. Bacteriol. **154**:41-49.
 16. Mooi, F., C. Wouters, A. Wijffjes, and F. deGraaf. 1982. Construction and characterization of mutants impaired in the biosynthesis of the K88ab antigen. J. Bacteriol. **150**:512-521.
 17. Morris, J., C. Thorns, G. Wells, A. Scott, and W. Sojka. 1983. The production of F41 fimbriae by piglet strains of enterotoxigenic *Escherichia coli* that lack K88, K99, and 987P fimbriae. J. Gen. Microbiol. **129**:2753-2759.
 18. Moseley, S. L., G. Dougan, R. A. Schneider, and H. W. Moon. 1986. Cloning of chromosomal DNA encoding the F41 adhesin of enterotoxigenic *Escherichia coli* and genetic homology between adhesins F41 and K88. J. Bacteriol. **167**:799-804.
 19. Moseley, S. L., and S. Falkow. 1980. Nucleotide sequence homology between the heat-labile enterotoxin gene of *Escherichia coli* and *Vibrio cholerae* deoxyribonucleic acid. J. Bacteriol. **144**:441-446.
 20. Moseley, S. L., M. Samadpour-Motalebi, and S. Falkow. 1983. Plasmid association and nucleotide sequence relationships of two genes encoding heat-stable enterotoxin production in *Escherichia coli* strain H-10407. J. Bacteriol. **156**:441-443.
 21. Norrander, J., T. Kempe, and J. Messing. 1982. Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene **26**:101-106.
 22. Rupp, W., C. Wilde, and D. Reno. 1971. Exchanges between DNA strands in ultraviolet-irradiated *Escherichia coli*. J. Mol. Biol. **61**:25-44.
 23. Sancar, A., A. Hack, and W. Rupp. 1979. Simple method for identification of plasmid-encoded proteins. J. Bacteriol. **137**:692-693.
 24. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463-5467.
 25. Shine, J., and L. Dalgarno. 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342-1346.
 26. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503-515.
 27. To, S. 1984. F41 antigen among porcine enterotoxigenic *Escherichia coli* strains lacking K88, K99, and 987P pili. Infect. Immun. **43**:549-554.
 28. Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350-4354.
 29. vanDoorn, J., B. Oudega, F. Mooi, and F. deGraaf. 1982. Subcellular localization of polypeptides involved in the biosynthesis of K88ab fimbriae. FEMS Microbiol. Lett. **13**:99-104.
 30. Vieira, J., and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259-268.
 31. von Heijne, G. 1985. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **11**:4683-4690.