

K88ab Gene of *Escherichia coli* Encodes a Fimbria-Like Protein Distinct from the K88ab Fimbrial Adhesin

FRITS R. MOOI,* MARIANNE VAN BUUREN, GERRIT KOOPMAN, BERT ROSENDAAL, AND FRITS K. DE GRAAF

Department of Microbiology, Biological Laboratory, Vrije Universiteit, 1007 MC Amsterdam, The Netherlands

Received 17 January 1984/Accepted 18 May 1984

The K88ab adhesin operon of *Escherichia coli* encodes for a fimbrial protein (the K88ab adhesin) which is involved in colonization of the porcine intestine. We characterized a structural gene (gene A) which is part of the K88ab adhesin operon and codes for an as yet unidentified polypeptide (pA). A mutation in gene A resulted in accumulation of K88ab adhesin subunits inside the cell. The nucleotide sequence of gene A was determined, and the deduced amino acid sequence suggested that pA is synthesized as a precursor containing a typical N-terminal signal peptide. The molecular weight of pA was calculated to be ca. 17,600. Gene A is preceded by a sequence showing homology with the consensus promoter. Fimbrial subunits from a number of *E. coli* strains have significant homology at their N- and C-termini. pA also contained some of these conserved sequences and showed a number of other similarities with fimbrial subunits. Therefore, it seems likely that the K88ab adhesin operon codes for a fimbrial subunit (pA) distinct from the K88ab adhesin subunit.

The specific adherence of pathogenic bacteria to mammalian cells is an essential step in the development of bacterial disease. The bacterial surface structures responsible for adherence have been named adhesins or colonization factors.

The K88ab adhesin, found on particular enterotoxigenic *Escherichia coli* strains, has been implicated in the colonization of the porcine small intestine (1, 10). It is composed of protein subunits that form extracellular filamentous structures called fimbriae or pili (20).

The genetic determinant for the K88ab adhesin has been isolated by molecular cloning (13). The resulting recombinant plasmid, pFM205, appeared to code for five polypeptides with molecular weights of 17,000 (p17), 26,000 (p26, the K88ab adhesin subunit), 27,000 (p27), 27,500 (p27.5), and 81,000 (p81) (14, 16). The subcellular location of these polypeptides has been determined previously (23), and it appeared that p17, p27, and p27.5 are located in the periplasmic space, whereas p81 is found in the outer membrane. The K88ab adhesin subunit is located on the cell surface and is probably attached to the outer membrane. Polypeptides p17, p27, and p81 have been implicated in the biosynthesis of the K88ab adhesin and are involved in modification, translocation, assembly, and anchorage of K88ab adhesin subunits (15, 16). The function of p27.5 in the biosynthesis of the K88ab adhesin has yet to be determined.

Proximal to the gene encoding p81 is located a region of ca. 1,000 base pairs which codes for an additional structural gene (designated gene A) involved in the biosynthesis of the K88ab adhesin (F. R. Mooi, Ph.D. thesis, Vrije Universiteit, Amsterdam, The Netherlands, 1982). On the basis of genetic studies, a similar conclusion has been reached for the K88ac operon (11).

This report describes the localization of gene A on the cloned K88ab DNA, its nucleotide sequence, and the effects of mutations in gene A on K88ab adhesin production. The possible function of the product of gene A (pA) is discussed on the basis of its deduced amino acid sequence.

MATERIALS AND METHODS

Bacterial strains and culture conditions. *E. coli* K-12 strain PC2028 (*recA171*) was used as the host for the plasmids studied. Cells containing derivatives of pBR322 (3) or pMC661 (4) were cultured in nutrient broth (Oxoid Ltd, Basingstoke, Hants, England) in the presence of ampicillin (500 µg/ml) or kanamycin (20 µg/ml), respectively.

Plasmids. The physical and genetic maps of the plasmids used in this study are shown in Fig. 1. The construction and characterization of pFM205, pFM77, and pFM216 have been described previously (13, 14). Plasmid pFM205 contains a 6,700-base-pair DNA fragment derived from the wild-type K88ab plasmid pRI8801, which has been inserted between the *Hind*III and *Eco*RI sites of pBR322. Plasmids pFM77 and pFM216 are deletion derivatives of pFM205.

In pFM299, a small 37-base-pair *Bst*EII fragment was deleted. This plasmid was constructed by partially digesting pFM205 with *Bst*EII and isolating the linear full-length pFM205 molecules. Subsequently, the single-stranded extensions of the linear molecules were filled in with the Klenow fragment of DNA polymerase I, and the DNA was recircularized with T4 DNA ligase. The purpose of this procedure was to introduce small insertions at the various *Bst*EII sites. However, it appeared that pFM299 was derived from a molecule cleaved at two closely spaced *Bst*EII sites.

The cloning vector pMC661 (4) contains a *Hpa*I site located downstream from the inducible *araBAD* promoter. Transcription of fragments inserted into the *Hpa*I site of pMC661 can be controlled by addition or removal of L-arabinose. Plasmid pFM265 was constructed by inserting a 6,700-base-pair *Sau*3AI fragment from pFM205 into the *Hpa*I site of pMC661. To do this, the 5'-end single-stranded extensions generated by *Sau*3AI were converted to blunt ends with T4 DNA polymerase. Subsequently, the *Sau*3AI fragment was covalently joined to *Hpa*I-cleaved pMC661 with T4 DNA ligase.

Plasmid pFM266 arose from a fortuitous deletion event. It is identical to pFM265, except for a deletion between the *Sau*3AI-a and *Hinc*II sites.

Plasmid pFM268 was constructed by inserting a 5,000-base-pair *Hind*III-*Eco*RI K88ab DNA fragment from

* Corresponding author.

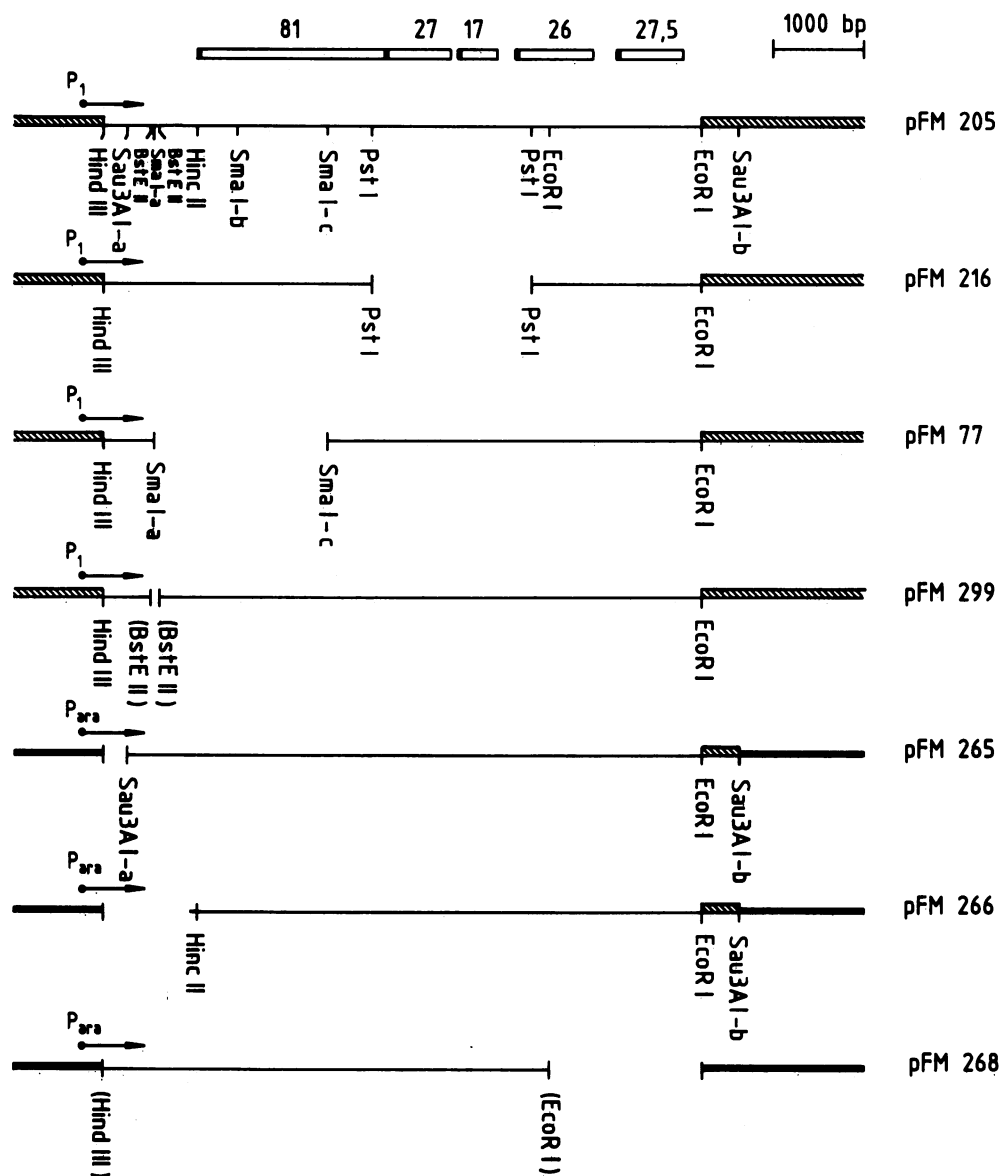


FIG. 1. Genetic and physical maps of plasmids. Thin lines represent K88ab DNA; deletions present in derivatives of pFM205 are indicated with open spaces; shaded boxes and thick black lines represent pBR322 and pMC661 DNA, respectively; locations of K88ab genes are indicated with boxes above pFM205; numbers above the boxes refer to the molecular weights ($\times 10^3$) of the corresponding polypeptides; the black part of the box indicates the part of the gene encoding the signal peptide; lines with arrows show the direction and approximate starting points of transcription. Restriction enzyme recognition sites between parentheses indicate sites on the K88ab DNA that were lost due to the cloning procedure. Only the relevant restriction enzyme recognition sites are shown. Abbreviations: P₁, pBR322 promoter (21); P_{ara}, *araBAD* promoter; bp, base pair.

pFM205 into the *Hpa*I site of pMC661. Essentially the same procedure was used as described for pFM265.

DNA sequence analysis. DNA sequence analysis was carried out as described by Sanger et al. (19) by using bacteriophage M13 as the source for single-stranded DNA (18). Both strands were sequenced at least once.

Preparation of cell extracts. Cell extracts were prepared by disrupting cells by ultrasonic treatment and subsequently removing whole cells and large cell fragments by centrifugation.

Determination of the amount of K88ab antigen. Since the immunological assay used does not discriminate between the functional K88ab adhesin and its precursors, protein reacting with antibodies directed against the K88ab adhesin will

be referred to as K88ab antigen. The amount of K88ab antigen in cell extracts and on whole cells was determined by an enzyme-linked immunosorbent assay (13). The microtiter plates used were read with a micro-enzyme-linked immunosorbent assay reader (Organon Technika).

RESULTS

Mapping of gene A and characterization of gene A mutants. Plasmid pFM266 (Fig. 1) consists of a 5,700-base-pair K88ab DNA fragment derived from pFM205 and fused to the inducible *araBAD* promoter of pMC661. Cells containing pFM266 did not produce a detectable amount of K88ab antigen, whether or not the *araBAD* promoter was derepressed (Table 1). Production of K88ab antigen could be

TABLE 1. Characteristics of plasmids containing a deletion in gene A^a

Plasmid ^b	K88ab antigen detected:	
	On whole cells	In cell extracts
pFM205	++	++
pFM266	-	-
pFM216	-	-
pFM266 + pFM216	+	+
pFM265	++	++
pFM77	-	ND
pFM266 + pFM77	-	ND
pFM299	-	+
pFM268	-	-
pFM299 + pFM268	++	++

^a *E. coli* PC2028, containing the plasmids indicated, was grown at 37°C for 16 h. To cultures of strains containing pFM266, pFM265, or pFM268, L-arabinose was added to 0.2% (wt/vol) at an optical density at 660 nm of 0.3 to derepress the *araBAD* promoter. The amount of K88ab antigen was determined by an enzyme-linked immunosorbent assay. Symbols: ++, K88ab antigen detected at high levels; +, K88ab antigen detected at reduced levels; -, no K88ab antigen detected; ND, not done.

^b Physical and genetic maps of plasmids are shown in Fig. 1.

restored by complementation with pFM216 (Fig. 1; Table 1), a derivative of pFM205 in which the genes encoding p81, p27, p17, and p26 (the K88ab adhesin subunit) have been partially or completely deleted. This indicated the K88ab⁻ phenotype of cells containing pFM266 was not due to a mutation in one of these genes. Thus, the *Hind*III-*Hinc*II K88ab DNA fragment of pFM205, which is absent from pFM266, contains the gene for an as yet unidentified polypeptide involved in the biosynthesis of the K88ab adhesin (gene A).

Compared with pFM266, pFM265 contains a slightly larger K88ab DNA fragment derived from pFM205 (Fig. 1). It was constructed by fusing a 6,500-base-pair *Sau*3AI-*Eco*RI K88ab DNA fragment to the *araBAD* promoter of pMC661. Cells containing pFM265 produced large amounts of extracellular K88ab antigen when the *araBAD* promoter was derepressed (Table 1), indicating that it contained all the structural genes necessary for export and assembly of the K88ab adhesin subunits. This implied that gene A was located to the right of the *Sau*3AI-a site of pFM205 (Fig. 1).

The mutation in pFM266 could not be complemented with pFM77 (Fig. 1; Table 1), a derivative of pFM205 containing a deletion between the *Sma*I-a and *Sma*I-c sites. Apparently, gene A is not expressed by pFM77. This indicates that gene A is located to the right of the *Sma*I-a site or, alternatively, spans this site.

To discern between these two possibilities, we analyzed a derivative of pFM205 (designated pFM299; Fig. 1), which contained a small deletion at the *Sma*I-a site. As was observed for pFM266, the K88ab antigen was not detected on whole cells containing pFM299 (Table 1). However, in contrast to pFM266, K88ab antigen was detected in cell extracts (Table 1). Apparently, the deletion in pFM299 resulted in accumulation of K88ab adhesin subunits inside the cell. Production of extracellular K88ab antigen could be restored by complementation with pFM268 (Fig. 1; Table 1). These findings suggest that the *Sma*I-a site is located within gene A. Moreover, pA is apparently necessary for export of K88ab adhesin subunits.

It is not clear why intracellular accumulation of K88ab adhesin subunits was not observed in cells containing pFM266. Possibly, accumulation occurred at a level too low

to detect. We have constructed several other pA⁻ mutants (data not shown), which all accumulated the K88ab adhesin subunits intracellularly in detectable amounts. Therefore, the behavior of pFM266 seems to be anomalous in this respect.

Nucleotide sequence of gene A. Attempts to detect pA have failed thus far; therefore, to obtain more information about this polypeptide, we determined the nucleotide sequence of gene A.

Analyses of deletion mutants suggest that gene A was located to the right of the *Sau*3AI-a site of pFM205 and spanned the *Sma*I-a site (Fig. 1). The nucleotide sequence of this region revealed a significant open reading frame which fulfilled these criteria (Fig. 2). Because this open reading frame codes for a polypeptide which shows a number of striking similarities with polypeptides encoded by other adhesin operons (see below), we concluded that it represents gene A.

Gene A contains 181 amino acid codons and is terminated by two TAG stop codons. The 5' end of gene A codes for a typical signal peptide. The most likely cleavage site for signal peptidase was determined by the method of Von Heijne (24) and is shown in Fig. 2. The tentative mature protein (pA) contains 160 amino acids and is expected to have a molecular weight of ca. 17,600. The presence of a signal peptide indicates that pA has an extracytoplasmic location.

Eight base pairs removed from the 3' end of gene A, the start of another open reading frame, is observed (Fig. 2). This open reading frame continues for at least 220 codons (data not shown) and presumably represents the 5' end of the gene encoding p81.

Gene A is preceded by a sequence which shows a limited degree of homology with the Shine-Dalgarno sequence (8). Further upstream, a sequence is observed which complies with the promoter consensus sequence (9). Approximately 200 base pairs away from this putative K88ab promoter, at the 5' end of the cloned K88ab DNA, a region of dyad symmetry was observed. The potential stem and loop structure that can be formed by the transcript of this region is shown in Fig. 3. Such structures have been shown to be involved in [*rho*]-dependent termination of transcription (17).

DISCUSSION

Genetic studies have revealed that in addition to the five structural genes identified previously (14, 16), a sixth structural gene (gene A) is located on the cloned K88ab DNA contained in pFM205. The product of gene A (pA) appears to be necessary for export of K88ab adhesin subunits, since in most pA⁻ mutants analyzed, the subunits accumulated inside the cell. We have shown that mutations in K88ab genes generally result in accumulation of K88ab adhesin subunits in the periplasmic space (15). Therefore, it seems likely that the K88ab adhesin subunits detected in pA⁻ mutants are also located in the periplasmic space.

Attempts to detect pA in normal cells or minicells to date have failed (data not shown). Since pA does not contain methionine, we used ³H-labeled amino acids in our attempts to detect pA in minicells. Moreover, to exclude the possibility that pA was obscured by another K88ab polypeptide on sodium dodecyl sulfate-polyacrylamide gels (for example, p17), we also analyzed plasmids with deletions in one or more K88ab genes in minicells; this was to no avail. The fact that pA has not been detected cannot be due to a low level of transcription of gene A, because the K88ab genes are transcribed from a strong pBR322 promoter (P1) (21) in

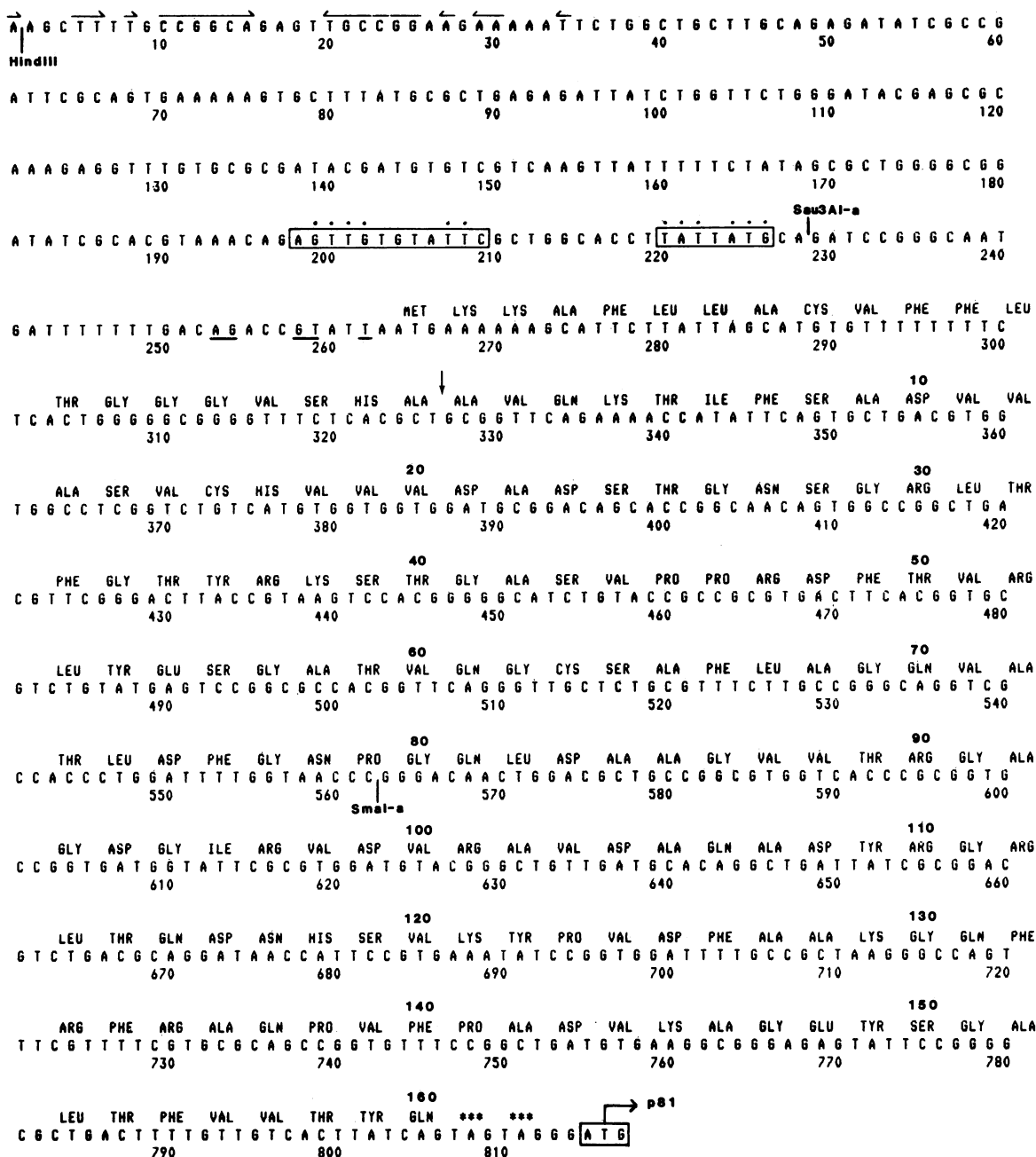


FIG. 2. Nucleotide sequence of gene A. The deduced amino acid sequence is numbered from the tentative N-terminus of the mature protein. The vertical arrow indicates the presumed cleavage site for signal peptidase. Nucleotides complementary to the 3' end of 16S rRNA are underlined. Possible -35 and -10 promoter regions are boxed. Dots indicate homology with the promoter consensus sequence. Horizontal arrows indicate a region with dyad symmetry. The putative initiation codon of the gene encoding p81 is boxed. Restriction enzyme recognition sites referred to in the text are indicated.

pFM205 (Fig. 1). Therefore, it seems likely that pA is synthesized in low amounts because its gene is translated inefficiently. This is also suggested by nucleotide sequence data (Fig. 2), which reveal that gene A is preceded by a region exhibiting only weak complementarity to the 3' end of the 16S rRNA (8). In this context, it should be noted that although the K88ab genes are located within a single transcriptional unit (manuscript in preparation), the genes for p26 and p27 are expressed at much higher levels than the genes for p17 and p81 (16); this also suggests that differential expression of K88ab genes occurs at the level of translation.

Two lines of evidence indicate the presence of a weak promoter upstream from gene A. First, ca. 40 base pairs away from the start of codon of gene A, sequences are observed showing homology with -35 and -10 consensus sequences of *E. coli* promoters (9). The distance between the -35 and -10 regions is 3 base pairs shorter than normally found in strong promoters, suggesting that the putative K88ab promoter does not initiate transcription efficiently. Second, a low amount of K88ab adhesin is still produced when the pBR322 promoter P₁ is deleted from pFM205 (data not shown). The putative K88ab promoter is preceded by a

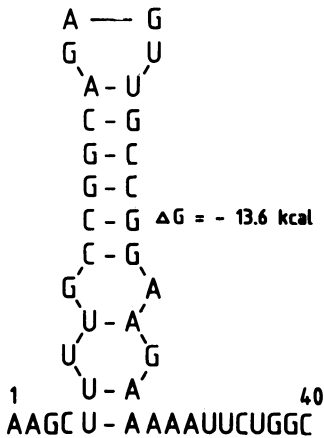


FIG. 3. Possible secondary structure formed by the K88ab mRNA. The numbers of the bases correspond to those in Fig. 2. ΔG was calculated by the method of Tinoco et al. (22).

terminator-like structure (Fig. 2 and 3). It is unlikely that this structure greatly affects transcription from P_1 , because the K88ab genes are highly expressed in pFM205 (13).

A clue about the function of gene *A* could be obtained by comparing the genetic organizations of the K99 (5) and Pap (M. Norgren, S. Normark, D. Lark, P. O'Hanley, G. Schoolnik, S. Falkow, C. Svanborg-Edén, M. Båga, and B. E. Uhlin, EMBOJ., in press) adhesin operons with the K88ab adhesin operon (16; Fig. 1). This comparison revealed that gene *A* is located at a position in which the structural gene for the fimbrial subunit is located in the K99 and Pap adhesin operons, suggesting that pA is a fimbrial subunit. Further indications for this assumption could be obtained by comparing the amino acid sequences of the K99 (B. Roosendaal, W. Gaastra, and F. K. de Graaf, FEMS Lett., in press) and Pap (2) fimbrial subunits with the amino acid sequence of pA (Fig. 2). It appeared that pA and the K99 and Pap subunits are composed of approximately the same number of amino acid residues (160, 159, and 163, respectively). More-

over, all three polypeptides contain two cysteine residues at approximately the same position (16 and 63 for pA, 16 and 57 for the K99 fimbrial subunit, and 22 and 61 for the Pap fimbrial subunit). Finally, homology was observed at the N- and C-termini of these polypeptides (Fig. 4). Homology was also observed with the N- and C-termini of the K88ab adhesin (7) and CFAI (12) subunits (Fig. 4).

Homology at the N-terminus was found within the first 20 to 30 amino acid residues. It was somewhat complex and only evident after proper alignment (Fig. 4A). Homology at the C-terminus was more pronounced and found within the last 20 amino acid residues (Fig. 4B). The resemblance in primary structure at the N- and C-termini of the polypeptides investigated might suggest that these sequences are involved in common steps in the biogenesis of these surface proteins. It is conceivable that they determine the quaternary structure of the proteins or represent topogenic sequences (i.e., sequences that direct the polypeptides to their ultimate location). Nonhomologous regions might be involved in more specific functions such as interaction with host receptors.

In summary, it may be stated that pA has many characteristics in common with fimbrial subunits. Thus, the K88ab adhesin operon might code for two extracellular polypeptides, the K88ab adhesin subunit and pA. This raises questions concerning the function and morphology of the structures composed of these polypeptides. It is possible that both polypeptides are part of the same structure, with the K88ab adhesin subunit probably representing the major component. For example, pA might function as a link between the K88ab adhesin and the putative anchorage protein p81 (15). This would explain the low amount of pA (and p81) synthesized relative to p26, because only a few of these molecules are necessary to anchor one K88ab fimbria (composed of 100 or more p26 molecules) to the cell. Alternatively, the two polypeptides might form separate structures on the cell surface, the synthesis of which is, however, interdependent.

Norgren et al. (in press) have obtained evidence that the Pap adhesin is encoded by a region on the Pap operon which

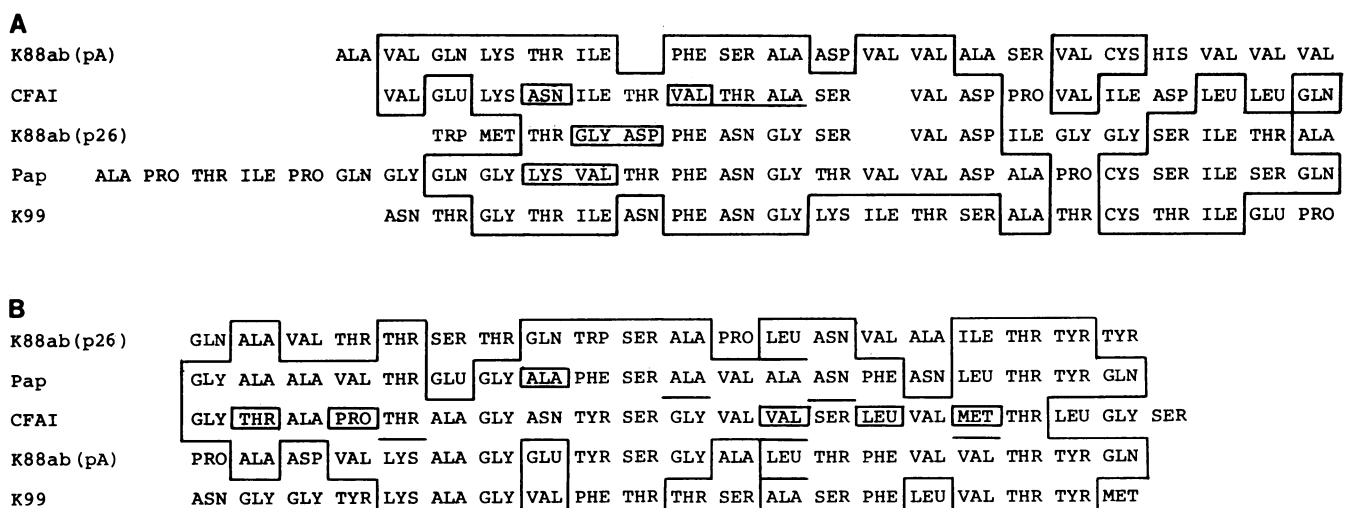


FIG. 4. Comparison of amino- (A) and carboxy- (B) terminal amino acid sequences of pA and a number of fimbrial subunits. Identical or functionally identical amino acids are boxed. The following amino acids were assumed to be functionally identical: (glu, asp), (gln, asn), (ser, thr), (leu, ile), (phe, trp, tyr). The amino acid sequences of the K88ab adhesin (p26), the K99, the Pap (pA), and the CFAI fimbrial subunits were obtained from references 7, Roosendaal et al., in press, 2, and 12, respectively.

is distinct from the fimbrial subunit gene *papA*, suggesting that the Pap operon also codes for two surface proteins. The Pap adhesin might be homologous to the K88ab adhesin and the Pap fimbrial subunit to pA. Thus, colonization of some host surfaces by *E. coli* might require at least two extracellular proteins. Depending on the host surface to be colonized, the contribution and also the relative amounts of the two proteins may vary.

LITERATURE CITED

- Anderson, M. J., J. S. Whitehead, and Y. S. Kim. 1980. Interaction of *Escherichia coli* K88 antigen with porcine intestinal brush border membranes. *Infect. Immun.* **29**:897-901.
- Båga, M., S. Normark, J. Hardy, P. O'Hanley, D. Lark, O. Olsson, G. Schoolnik, and S. Falkow. 1984. The nucleotide sequence of the *papA* gene encoding the Pap pilus subunit of human uropathogenic *Escherichia coli*. *J. Bacteriol.* **157**:330-333.
- Bolivar, F., R. L. Rodriguez, P. J. Green, M. C. Betlach, H. L. Heijneker, H. W. Boyer, J. H. Crossa, and S. Falkow. 1977. Construction and characterization of new cloning vehicles. *Gene* **2**:95-113.
- Casadaban, M. J., and S. N. Cohen. 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.* **138**:179-207.
- De Graaf, F. K., B. E. Krenn, and P. Klaasen. 1984. Organization and expression of genes involved in the biosynthesis of K99 fimbriae. *Infect. Immun.* **43**:508-514.
- Dougan, G., and D. Sherratt. 1977. The transposon Tn1 as a probe for studying ColE1 structure and function. *Mol. Gen. Genet.* **151**:151-160.
- Gaastra, W., F. R. Mooi, A. R. Stuitje, and F. K. de Graaf. 1981. The nucleotide sequence of the gene encoding the K88ab protein subunit of porcine enterotoxigenic *Escherichia coli*. *FEMS Lett.* **12**:41-46.
- Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer, and G. Stormo. 1981. Translational initiation in prokaryotes. *Annu. Rev. Microbiol.* **35**:365-403.
- Hawley, D. K., and W. R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11**:2237-2255.
- Jones, G. W., and J. M. Rutter. 1972. Role of the K88 antigen in the pathogenesis of neonatal diarrhea caused by *Escherichia coli* in piglets. *Infect. Immun.* **6**:918-927.
- Kehoe, M., M. Winther, and G. Dougan. 1983. Expression of a cloned K88ac adhesion antigen determinant: identification of a new adhesion cistron and role of a vector-encoded promoter. *J. Bacteriol.* **155**:1071-1077.
- Klemm, P. 1982. Primary structure of the CFAI fimbrial protein from human enterotoxigenic *Escherichia coli* strains. *Eur. J. Biochem.* **124**:339-348.
- Mooi, F. R., F. K. de Graaf, and J. D. A. van Embden. 1979. Cloning, mapping and expression of the genetic determinant that encodes for the K88ab antigen. *Nucleic Acids Res.* **6**:849-865.
- Mooi, F. R., N. Harms, D. Bakker, and F. K. de Graaf. 1981. Organization and expression of genes involved in the production of the K88ab antigen. *Infect. Immun.* **32**:1155-1163.
- Mooi, F. R., A. Wijffjes, and F. K. de Graaf. 1983. Identification and characterization of precursors in the biosynthesis of the K88ab fimbria of *Escherichia coli*. *J. Bacteriol.* **154**:41-49.
- Mooi, F. R., C. Wouters, A. Wijffjes, and F. K. de Graaf. 1982. Construction and characterization of mutants impaired in the biosynthesis of the K88ab antigen. *J. Bacteriol.* **150**:512-521.
- Morgan, W. D., D. G. Bear, and P. H. von Hippel. 1983. Rho-dependent termination of transcription. I. Identification and characterization of termination sites for transcription from bacteriophage λ P_R promoter. *J. Biol. Chem.* **258**:9553-9564.
- Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe. 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* **143**:161-178.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**:5463-5467.
- Stirm, S., F. Ørskov, I. Ørskov, and A. Birch-Andersen. 1967. Episome-carried surface antigen K88 of *Escherichia coli* III. Morphology. *J. Bacteriol.* **93**:740-748.
- Stüber, D., and H. Bujard. 1981. Organization of transcriptional signals in plasmids pBR322 and pACYC184. *Proc. Natl. Acad. Sci. U.S.A.* **78**:167-171.
- Tinoco, I. T., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla. 1973. Improved estimation of secondary structures in ribonucleic acids. *Nature (London) New Biol.* **246**:40-41.
- Van Doorn, J., B. Oudega, F. R. Mooi, and F. K. de Graaf. 1982. Subcellular localization of polypeptides involved in the biosynthesis of K88ab fimbriae. *FEMS Lett.* **13**:99-104.
- Von Heijne, G. 1983. Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* **133**:17-21.