Jpn. J. Vet. Res., **39**, 1–10 (1991)

THE NUCLEOTIDE SEQUENCE OF THE GENES, *fan*E AND *fan*F, OF *Escherichia coli* K99 FIMBRIAE.

Eriko ONO[1], Martin F. LAVIN[2], Masaharu NAIKI[1,3,4]

(Accepted for Publication: Janualy 7, 1991)

K99 fimbriae of enterotoxigenic *Escherichia coli* consist of eight different subunits. A major subunit called fimbrillin forms fimbrial structure and a minor subunit called adhesin localizes at the tip of fimbriae and recognizes host receptor ganglioside. Within this eight gene cluster, *fan*E and *fan*F have not yet been sequenced. In this study, *fan*E and *fan*F genes were sequenced by analyzing several DNA fragments produced by endonuclease or exonuclease digestion. The *fan*E gene encoded 227 amino acids containing 20 amino acids of signal peptide starting from GTG (valine) and showed a homology to *fan*A-*fan*B. The *fan*F gene encoded 271 amino acids containing 20 amino acids of signal peptide starting from ATG (methionine) and showed homologies to the *fan*D gene, fimbrillin gene of F41, adhesin gene of P fimbriae (*pap*G) and adhesin gene of Type 1 fimbriae (*fim*H). E and F subunits had fifteen and fourteen hydrophobic domains, respectively, which periodically appeared possibly forming a hydrophobic region.

Key words: Enterotoxigenic *Escherichia coli*, K99 fimbriae, *fan*E and *fan*F DNA sequence, hydrophobic domain.

INTRODUCTION

Enterotoxigenic *Escherichia coli* possessing K99 fimbrial antigen are frequently isolated from newborn calves, piglets, and lambs suffering from diarrhea[21]. It is known that fimbriae are important as an adhesive and colonization factor of bacteria on the mucosal surface of the small intestine in the first stage of infection[1,5,6]. Fimbriae have been classified by their antigenicity and subunit molecular weight. The fimbriae are composed of fimbrillin which forms main part of fimbriae and adhesin which recognizes carbohydrate receptors on the surface of host cells[13,23,24].

[1] Department of Biochemistry, Faculty of Veterinary Medicine, Hokkaido University, Sapporo 060, Japan.
[2] Queensland Institute of Medical Research, Herston, Brisbane, Queensland 4006, Australia.
[3] Present address; Department of Veterinary Science, National Institute for Health, Kamiosaki 2–10–35, Shinagawa-ku, Tokyo 141, Japan.
[4] To whom correspondence should be addressed.

The K99 fimbrial gene cluster named *fan*A to *fan*H encodes eight different subunits[15]. The nucleotide sequences of *fan*A, *fan*B, *fan*C, *fan*D, *fan*G, and *fan*H have already been published[16,117,19]. The functions of these subunits were analyzed and it has been established that *fan*C encodes fimbrillin[16], *fan*A and *fan*B encode regulatory proteins which control expression of fimbriae[17,20], *fan*D encodes platform protein[19] and *fan*G and *fan*H encode minor subunits[18]. However there is no available information about the gene encoding adhesin. Since it is highly possible that either *fan*E or *fan*F encodes this receptor protein, in this study, the complete gene structure of *fan*E and *fan*F were determined.

MATERIALS AND METHODS

Host cells and plasmid vector: As a host bacterial strain, JM109 or JPA101 was used, and as a vector, pUC19[26] or BLUESCRIPT[22] was prepared for isolation of single strand DNA. Isolation of plasmid DNA, agarose gel electrophoresis, and transformation were performed according to conventional methods[22]. Transformants were cultured in LB media or 2YT media[22] with 50 μg / ml of ampicillin.

Endonuclease and the other enzymes: *Ava*I, *BamH*I, *Bgl*II, *Cla*I, *Hinc*II, *Kpn*I, *Pst*I, *Xba*I, exonuclease III, S1 nuclease, and *E. coli* DNA polymerase I (Klenow fragment) (TAKARA Co. Ltd., Kyoto, Japan) were used to prepare several DNA fragments from an original plasmid pFK99 (pUC19). T4 ligase (NIPPON GENE Co. Ltd., Toyama, Japan) was used for ligation.

Sequence procedure: Single strand DNA was prepared according to SEQUENASE protocol or Biochemical reseach products: Protocol and application committed to science worldwide (Promega protocol). 6% polyacrylamide gel containing 7M urea was used in the electrophoresis for sequencing. Samples for sequencing were prepared using SEQUENASE Ver. 2.0 kit (TOYOBO Co. Ltd., Osaka, Japan). pFK99 (pBR322) which contained K99 fimbrial gene cluster (*fan*A to *fan*H) were kindly supplied by Dr. de Graaf, Vrije University, Amsterdam, the Netherlands. *BamH*I-*-BamH*I fragment encoding the whole K99 gene cluster of pFK99 (pBR322)[25] was cloned into pUC19 vector (pFK99(pUC19)) and used for subcloning (Fig. 1).

Preparation of subclones containing fanE and fanF: Subclones including *fan*E and *fan*F genes were prepared (Fig. 1). *Kpn*I–*Ava*I fragment isolated after agarose gel electrophoresis using GENECLEAN™ Kit (FUNAKOSHI Co. Ltd., Tokyo, Japan) was treated with Klenow enzyme and inserted again into pUC19 cleaved with *Kpn*I and *Hinc*II, to make pFK99KA. pFK99AH (*Ava*I–*Hinc*II, fragment) and pFK99KC (*Kpn*I–*Cla*I fragment) were prepared by using each particular restriction enzyme site. pFK99 (KK3.3R) was prepared from pFK99 (pUC19) by using the *Kpn*I site. Digestion of pFK99(KK3.3R) with exonuclease lll was performed at 37°C and at 30 sec intervals, samples of partially digested DNA were taken according to Promega protocol. After cloning, DNA length was checked by agarose gel electrophoresis and 15

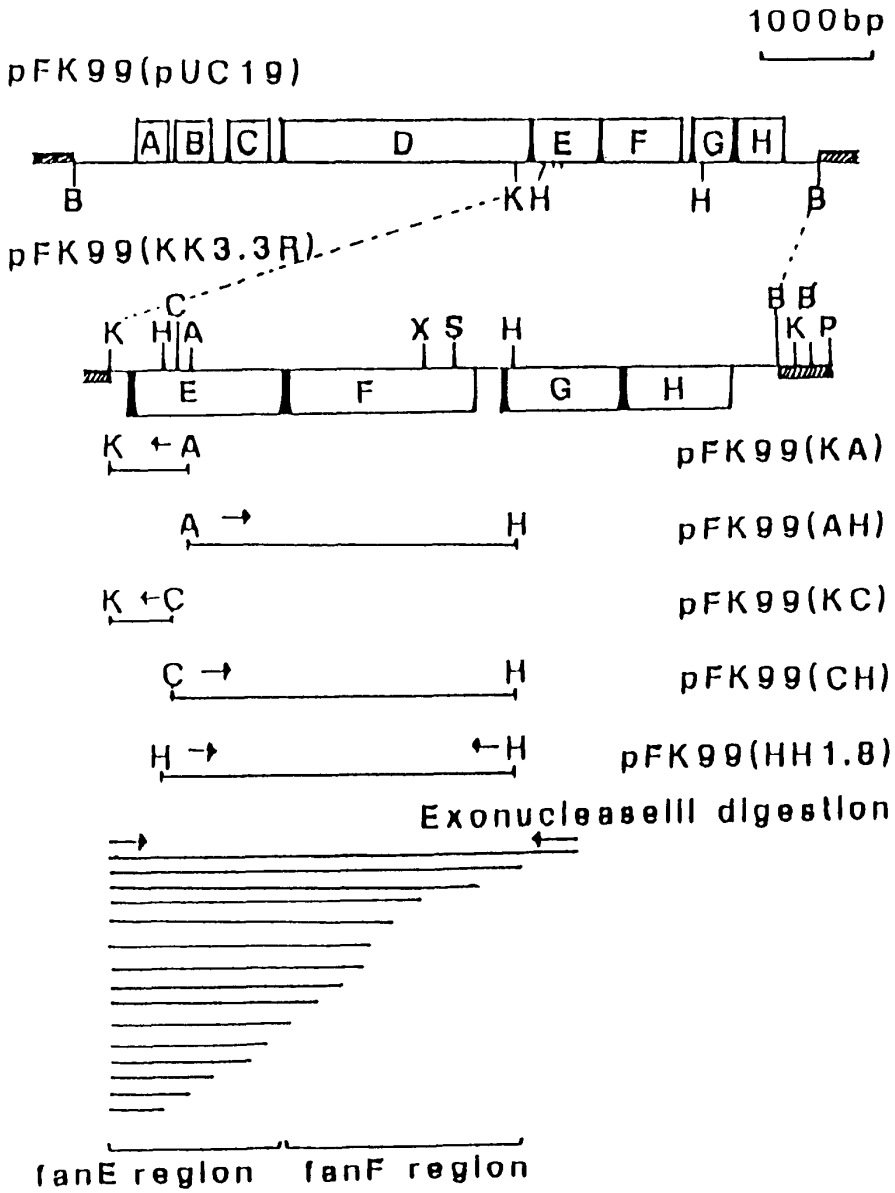


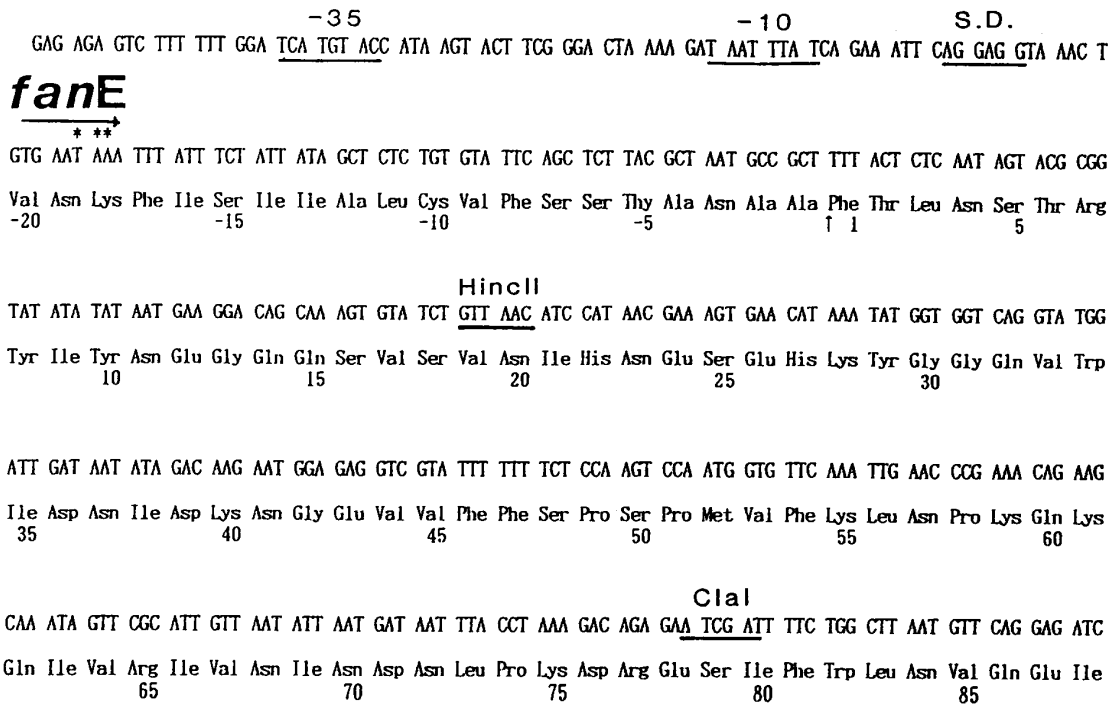
Fig. 1. Gene map of K99 fimbrial gene cluster and subclones prepared for sequencing. Genes encoding protein are indicated by boxes. Abbreviations are as follows : A, *Ava*I ; B, *BamH*I ; C, *Cla*I ; H, *Hinc*II ; K, *Kpn*I ; P, *Pst*I ; S, *Sau3A*I ; X, *Xmn*I.

subclones were selected to sequence completely.

Analysis of hydrophobic character of products: The hydrophobic character of the gene products was analyzed by the method described by Kyte and Doolittle[10], using Genetyx software supplied from SDC Software Exproitation Co., Ltd. (Tokyo, Japan).

RESULTS

Primary structure of fanE and fanF: The nucleotide sequences of *fan*E and *fan*F are shown in Fig. 2 with the corresponding amino acid sequences. Each structural gene had –35 and –10 promoter sequences upstream from the ribosome–binding site (S. D.). The *fan*E product had a signal peptide of 20 amino acids and a mature molecule consisted of 207 amino acids. There are 3 restriction enzyme sites for *Hin*cII, *Cla*I and *Ava*I. The starting codon of *fan*E, GTG (valine) was overlapped by the stopping codon of *fan*D. Comparing fimbrial DNA sequences, *fan*E had homology to *fan*A–*fan*B (44.5%). The *fan*F product however, had a signal peptide of 20 amino acids and a mature protein consisted of 251 amino acids. There are 2 restriction enzyme sites for *Sau3A*I and *Xmn*I. The starting region of *fan*F was also overlapped by the terminal region of *fan*E as observed at the *fan*D–*fan*E region. *fan*F showed homologies to *fan*D (45.7%), *pap*G (47.8%) of P fimbriae, *fim*H (44.5%) of Type 1 fimbriae, and F41 fimbriae (45.1%). It is known that *pap*G and *fim*H encode adhesin subunits of P and Type 1 fimbriae, respectively.



```
CCT CCA GCT CCA AAG GGA GAT GGG GGT AGT CTG TCA CTG GCA ATT AAT AAT CGA GTA AAA TTA ATA TAT CGA CCA ATT GCT

Pro Pro Ala Pro Lys Gly Asp Arg Glu Ser Leu Ser Leu Ala Ile Asn Asn Arg Val Lys Leu Ile Tyr Arg Pro Ile Ala
     90                  95                 100                 105                 110                 115

                                                                AvaI
CTA AAA AAT GGT CGA GAT GAG GCA GAA AAT AAT ATT AAG CTG ATT AAC TCG GGC ACG GAT TCT TGC CTT GAA AAT ACA ACG

Leu Lys Asn Gly Arg Asp Glu Ala Glu Asn Asn Ile Lys Leu Ile Asn Ser Gly Thr Asp Ser Cys Leu Glu Asn Thr Thr
                120                 125                 130                 135                 140

CCA TAT TAC TTT GCA ATT AGT GAT GTT AAA ATT AAT GGC AAA TCA ATT GAT TTA AAT TCT GAT GCA AAA AAT AAG ATG GGA

Pro Tyr Tyr Phe Ala Ile Ser Asp Val Lys Ile Asn Gly Lys Ser Ile Asp Leu Asn Ser Asp Ala Lys Asn Lys Met Gly
        145                 150                 155                 160                 165

                                                                                                -35
GTA TTC TCC CCA TTC TCG AAA GTT TGT CTG GGA AAT GTA AAT ACT AGT GGA AAC ATC ACG GTA ACA GCA TTT AAT GAC TAT

Val Phe Ser Pro Phe Ser Lys Val Cys Leu Gly Asn Val Asn Thr Ser Gly Asn Ile Thr Val Thr Ala Phe Asn Asp Tyr
170                 175                 180                 185                 190                 195

                    -10                 S.D.                fanF
GGC GTT GCA ACC AGC TAC ACT GTT CAA AGG AGT AAA TA ATG AAA AAT AAA TAT AAT TTA TTA TTT TTT TTA TTT CTT TTG

Gly Val Ala Thr Ser Tyr Thr Val Gln Arg Ser Lys ** *
                200                 205

                                                Met Lys Asn Lys Tyr Asn Leu Leu Phe Phe Leu Phe Leu Leu
                                                -20                 -15                 -10

TGT TAT GGA GAT GTG GCG CTG GCA GCA TGC ACA GGG AAA CTG AAA ATC TCA CCT GGT TAT AGT GGC CAT ACT TAT TCA TTT

Cys Tyr Gly Asp Val Ala Leu Ala Ala Cys Thr Gly Lys Leu Lys Ile Ser Pro Gly Tyr Ser Gly His Thr Tyr Ser Phe
     -5                  ↑ 1                 5                  10                  15                  20

GAT TCC AGT ATT CCA AAT AAT AGT AAT ATA GCA AGA TAC CTG GTC GAA ATT TCT GAG AAA ATT GTT TGT GAT GCG GAC CAG

Asp Ser Ser Ile Pro Asn Asn Ser Asn Ile Ala Arg Tyr Leu Val Glu Ile Ser Glu Lys Ile Val Cys Asp Ala Asp Gln
             25                  30                  35                  40                  45

TCA GGC TGG GAT GGT AAA CGT TAT GCT CAA TTA CAT CTT TAT TCA TCA GGT GCC TTA TGT GAA AGT GTC AGT GGA GAT GGG

Ser Gly Trp Asp Gly Lys Arg Tyr Ala Gln Leu His Leu Tyr Ser Ser Gly Ala Leu Cys Glu Ser Val Ser Gly Asp Gly
     50                  55                  60                  65                  70                  75

ATT ACA TTT AGG TCA AAT GTG TCC GGG CTG TCA TGG CGT TTT CCC AAT GGC ATA CCA TAC CAC TGT GCA GCA GGC CAA ATA

Ile Thr Phe Arg Ser Asn Val Ser Gly Leu Ser Trp Arg Phe Pro Asn Gly Ile Pro Tyr His Cys Ala Ala Gly Gln Ile
                80                  85                  90                  95                  100

AAT CTT GGC GGT ATA AAA TAT GCG GAT AGA AAT GGT AAA GTT ACC TGG AAT CCT GGT GAA CTA CGA CAT GAA ATA TTT TTA

Asn Leu Gly Gly Ile Lys Tyr Ala Asp Arg Asn Gly Lys Val Thr Trp Asn Pro Gly Glu Leu Arg His Glu Ile Phe Leu
        105                 110                 115                 120                 125
```

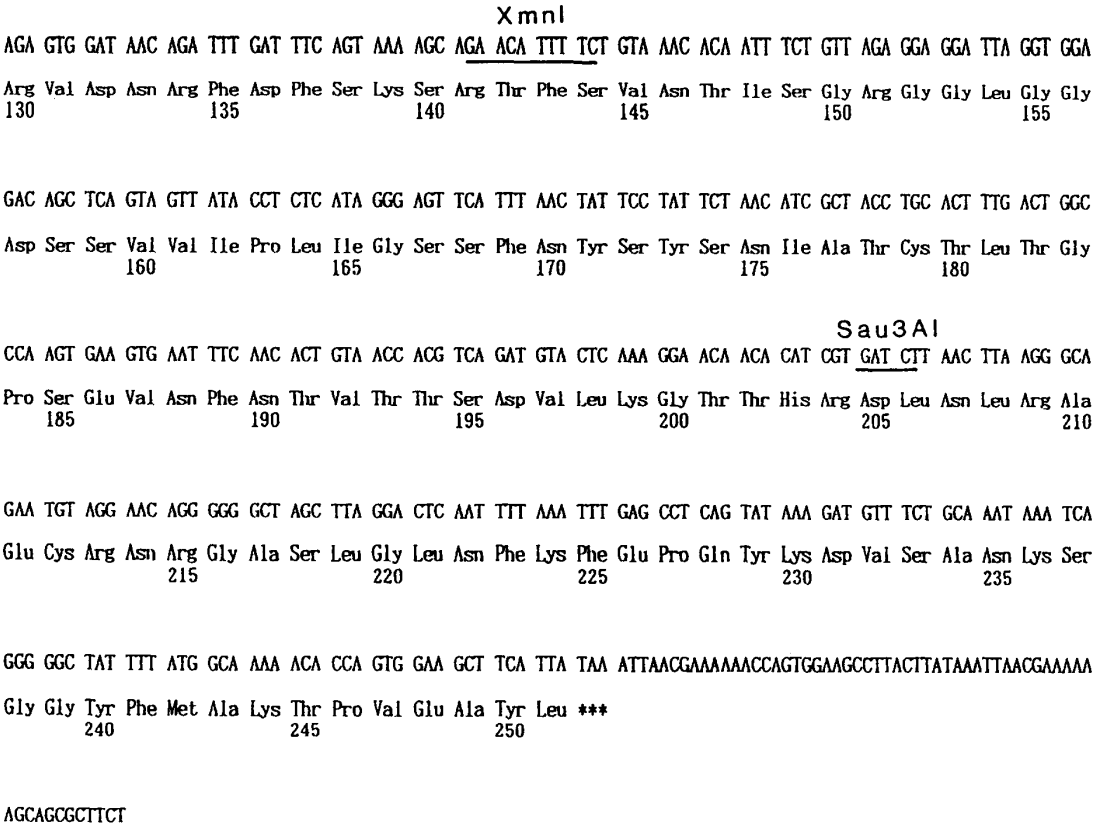


Fig. 2. Primary sequence of *fan*E and *fan*F genes. An arrow shows a position cleaved by digestion with signal peptidase I. −35 and −10 indicate promoter sequences. S. D. means ribosome-binding site. Stopping codon is indicated by ***.

Characterization of E and F subunits: According to computer analysis, the E subunit had partial homologies to the B, C, and D subunits of K99 fimbriae in its amino acid Sequence. The F subunit was partially homologous to the G adhesin subunit of P fimbriae, C and H subunits of K99 fimbriae. The hydrophobic character of these proteins are shown in Fig. 3. Fifteen and fourteen hydrophobic domains appeared in the E and F subunits, respectively. Particularly notable are the hydrophobic peaks of *fan*E which periodically appeared.

DISCUSSION

Various kinds of fimbrial genes were cloned from *Escherichia coli* or other bacterial cells. Among them, fimbrial genes of P[2,3,4,11,12,14] and Type 1[7,8,9] were well

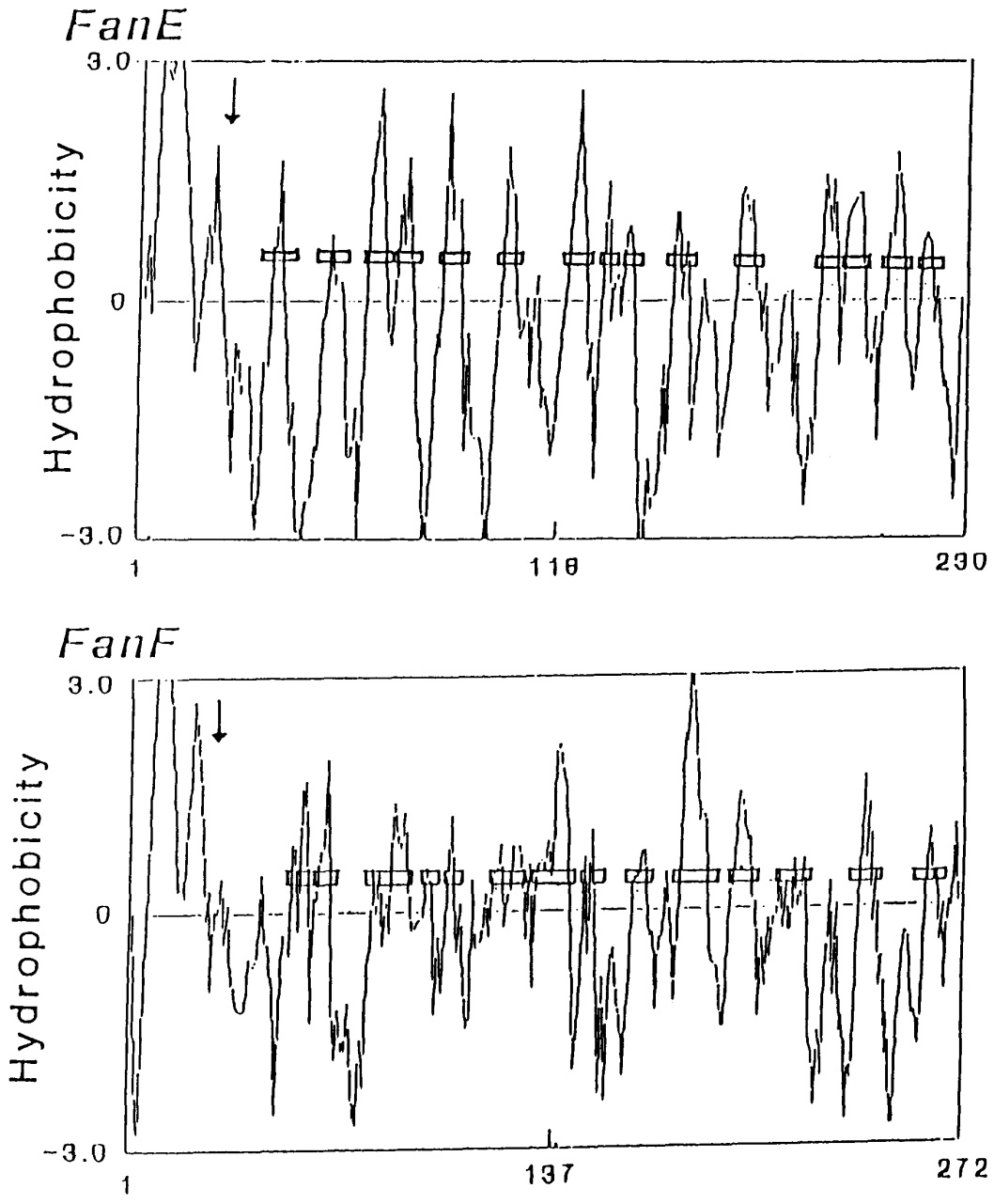


Fig. 3. Hydrophobic character of E and F subunits. Signal peptide is removed at a position indicated by an arrow. Hydrophobic peeks are indicated by boxes.

analyzed. It is known that the fimbrial gene and amino acid sequence of the products have some homologies.

Roosendaal et al. have reported on the K99 gene cluster and indicated that *fan*C encodes fimbrillin of K99 antigen (16.5 KDa)[16]. The A subunit (11.0 KDa)[17],B subunit (10.8 KDa)[17],D subunit (84.5 KDa)[19], G subunit (16.9 KDa)[18], and H subunit (16.3 KDa)[18] have been reported to be homologous in their nucleotide sequences. In the analysis of hydrophobicity, the periodic appearance of hydrophobic domain was detected.

In this study, *fan*E and *fan*F were sequenced for the first time. The starting codon of *fan*E was GTG, coding for valine. This starting codon is unique among the fimbrial genes of which uncleotide sequences have been reported. It was found that *fan*D, *fan*E, and *fan*F overlapped each other in the structural gene regions. A similar overlapping was reported in *fan*G and *fan*H of K99 fimbriae[18]. In addition, the *fan*F had homologies to the genes encoding adhesin of P and Type 1 fimbriae. In contrast, *fan*E showed a homology only to the *fan*A–*fan*B connecting region. Thus, the E subunit seems a very unique subunit within the K99 fimbriae. Regarding the hydrophobic character of *fan*E and *fan*F, fifteen and fourteen peaks were detected, respectively. In addition, the peaks of *fan*E appeared periodically, indicating that those formed the hydrophobic part of the E subunit.

At present it is not known which gene encodes K99 adhesin. However, it was found that the *fan*F had some homology to the adhesin gene of other fimbriae. Further analysis using mutants lacking *fan*E and *fan*F is necessary to determine the function of these subunits. Such experiments are now in progress.

REFERENCES

1) AWAD-MASALMEH, M., MOON, H., RUNNELS, P. L., & SCHNEIDER, R. A. (1982) : Pilus production, hemagglutination, and adhesion by porcine strains of enterotoxigenic *Escherichia coli*. *Infect. Immun.*, **35**, 305–313

2) BÅGA, M., NORMARK, S., HARDY, J., O'HANLEY, P., LARK, D., OLSSON, O., SCHOOLNIK, G., & FALKOW, S. (1984) : Nucleotide sequence of the *pap*A gene encoding the pap pilus subunit of human uropathogenic *Escherichia coli*. *J. Bacteriol.*, **157**, 330–333.

3) BÅGA, M., NORGREN, M., & NORMARK, S. (1987) : Biogenesis of *E. coli* Pap pili :PapH, a minor pilin subunit involved in cell anchoring and length modulation. *Cell*, **49**, 241–251.

4) BÅGA, M., GORANSSON, M., NORMARK, S., & UHLIN, B. E. (1988) : Processed mRNA with differential stability in the regulation of *E. coli* pilin gene expression. *Cell*, **52**, 197–206.

5) CULLING, C. F. A., REID, P. E., DUNN, W. L., & CLAY, M. G. (1977) : Histochemical comparison of the epithelial mucins in the ileum in Crohn's disease and in nomal control. *J. Clin. Pathol.*, **30**, 1063–1067.

6) DUGUID, J. P. & OLD, D. C. (1980): Adhesive properties of enterobacteriaceae. In bacteriae adherence. Receptors and recognition Ser. B, Vol. 6, Ed. E. H. Beachey, 185–217. Hapman and Hall, London.
7) GIAMPAPA, C. S., ABRAHAM, S. N., CHIANG. T. M., & BEACHEY, E. H. (1988): Isolation and characterization of a receptor for Type 1 fimbriae of *Escherichia coli* from guinea pig erythrocytes. *J. Biol. Chem.*, **263**, 5362–5367.
8) HANSEN, M. S., HEMPEL, J., & BRINTON, Jr., C. C. (1988): Purification of the *Escherichia coli* Type 1 pilin and minor pilus proteins and partial characterization of the adhesin protein. *J. Bacteriol.*, **170**, 3350–3358.
9) HANSON, M. S. & BRINTON, Jr., C. C. (1988): Identification and characterization of *E. coli* Type–1 pilus tip adhesion protein. *Nature*, **332**, 265–268.
10) KYTE, J. & DOOLITTLE, R. F. (1982): A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.*, **157**, 105–132.
11) LINDBERG, F. P., LUND, B., & NORMARK, S. (1984): Genes of pyelonephritogenic *E. coli* required for digalactoside–specific agglutination of human cells. *EMBO J.*, **3**, 1167–1173.
12) LINDBERG, F., LUND, B., & NORMARK, S. (1986): Gene products specifying adhesion of uropathogenic *Escherichia coli* are minor components of pili. *Proc. Natl. Acad. Sci. USA*, **83**, 1981–1895.
13) LINDBERG, F., LUND, B., & NORMARK, S. (1987): Localization of the receptor-binding protein adhesin at the tip of the bacterial pilus. *Nature*, **328**, 84–87.
14) NORGREN, M., BÅGA, M., TENNENT, J. M., & NORMARK, S. (1987): Nucleotide sequence, regulation and functional analysis of the *pap*C gene required for cell surface localization of Pap pili of uropathogenic *Escherichia coli. Mol. Microbiol.*, **1**, 169–178.
15) OUDEGA, B & DE GRAAF, F. K. (1988): Genetics organization and biosynthesis of adhesive fimbriae of *Escherichia coli. Antonie van Leeuwenhoek*, **54**, 285–299.
16) ROOSENDAAL, E., GAASTRA, W., & DE GRAAF, F. K. (1984): The nucleotide sequence of the gene encoding the K99 subunit of enterotoxigenic *Escherichia coli. FEMS Microbiol. Lett.*, **22**, 253–258.
17) ROOSENDAAL, E., BOOTS, M., & DE GRAAF, F. K. (1987): Two novel genes, *fan*A and *fan*B, involved in the biogenesis of K99 fimbriae. *Nucleic Acids Res.*, **15**, 5973–5984.
18) ROOSENDAAL, E., JACOBS, A. A. C., Rathman, P., SONDERMEYER, C., STEGEHUIS, F., OUDEGA, B., & DE GRAAF, F. K. (1987): Primary structure and subcellular localization of two fimbrial subunit-like proteins involved in the biosynthesis of K99 fimbriae. *Mol. Microbiol.*, **1**, 211–217.
19) ROOSENDAAL, B. & DE GRAAF, F. K. (1989): The nucleotide sequence of the *fan*D gene encoding the large outer membrane protein involved in the biosynthesis of K99 fimbriae. *Nucleic Acids Res.*, **17**, 1263.
20) ROOSENDAAL, B., DAMOISEAUX, J., JORDI, W.,& DE GRAAF, F. K. (1989): Transcriptional organization of the DNA region controlling expression of the K99 gene cluster. *Mol. Gen. Genet.*, **215**, 250–256.

21) RUNNELS, P. L., MOON, H. W., & SCHNEIDER, R. A. (1980): Development of resistance with host age to adhesion of K99$^+$ *Escherichia coli* to isolated intestinal epithelial cells. *Infect. Immun.*, **28**, 298–300.

22) SAMBROOK, J., FRITTSCH, E. F., & MANIATIS, T. M. (1989): Molecular cloning. A laboratory manual. Ed. Chris Nolan, Cold Spring Harbor Laboratory Press, New York, USA.

23) SMIT, H., GAASTRA, W., KAMERLING, J. P., Vliegenthart, J. F. G., & DE GRAAF, F. K. (1984): Isolation and characterization of the equine erythrocyte receptor for enterotoxigenic *Escherichia coli* K99 fimbrial adhesin. *Infect. Immun.*, **46**, 578–583.

24) TENEBERG, S., WILLEMSEN, P., DE GRAAF, F. K., & KARLSSON, K–A. (1990): Receptor-active glycolipids of epithelial cells of the small intestine of young and adult pigs in relation to susceptibility to infection with *Escherichia coli* K99. *FEBS Lett.*, **263**, 10–14.

25) VAN EMBDEN, J. D. A., DE GREEF, F. K., SCHOULS L. M., & TEPPEMA, J. S. (1980): Cloning and expression of a deoxyribonucleic acid fragment that encodes for the adhesive antigen K99. *Infect. Immun.*, **29**, 1125–1133.

26) YANISCH-PERRON, C., VIEIRA, J., & MESSING, J. (1985): Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103–120.