Gene, 32 (1984) 83–90
Elsevier

GENE 1138

Nucleotide sequence of the gene encoding the $F7_2$ fimbrial subunit of a uropathogenic *Escherichia coli* strain

(Recombinant DNA, mannose resistant hemagglutination; homology with *papA* gene)

Irma van Die and Hans Bergmans

Department of Molecular Cell Biology, State University of Utrecht, Padualaan 8, 3584 CH Utrecht (The Netherlands) Tel. 030-533184

(Received June 5th, 1984)
(Revision received July 4th, 1984)
(Accepted August 14th, 1984)

SUMMARY

The cloned DNA fragment encoding the $F7_2$ fimbrial subunit from the uropathogenic *Escherichia coli* strain AD110 has been identified. The nucleotide sequence of the structural gene and of 196 bp of the noncoding region preceding the gene was determined. The structural gene codes for a polypeptide of 188 amino acid residues, including a 21-residue N-terminal signal sequence. The nucleotide sequence and the deduced amino acid sequence of the $F7_2$ gene were compared with the reported sequences of the *papA* gene (Båga et al., 1984). Both genes code for subunits of fimbriae that are involved in mannose-resistant hemagglutination (MRHA) of human erythrocytes. The available data show that there is absolute homology between the noncoding regions preceding both genes over 129 bp. The two proteins are homologous at the N terminus and C terminus; there is less, but significant, homology in the region between the N and C termini.

INTRODUCTION

Virulence of uropathogenic *E. coli* strains has been related to their ability to adhere to uroepithelial cells and to mediate MRHA of human erythrocytes (Svanborg-Edén et al., 1976; Van den Bosch et al., 1980; Hagberg et al., 1981).

The fimbriae that are supposed to mediate the bacterial adherence, have been shown to be heterogeneous, serologically (Ørskov et al., 1980; 1982a, b) as well as in the M_r of the subunit proteins (Jann et al., 1981).

Molecular cloning of the genes involved in the expression of a number of MRHA fimbriae has been reported (Hull et al., 1981; Clegg, 1982; Rhen et al., 1983). We have recently reported the molecular cloning of a DNA fragment encoding one type of fimbriae, serotype $F7_2$, of the uropathogenic *E. coli* strain AD110 (Van Die et al., 1983).

Genetic analysis has shown that next to the gene encoding the subunit protein a number of genes involved in the expression of MRHA fimbriae are positioned (Normark et al., 1983; Clegg and Pierce, 1983; Van Die et al., 1984). These proteins apparently are either directly involved in the translocation of the fimbrial subunit to the surface of the cell and the assembly of subunits into fimbriae, or they could

Abbreviations: bp, base pairs; EtBr, ethidium bromide; *fimA*, structural gene of type 1 fimbrial subunit; kb, 1000 bp; MRHA, mannose-resistant hemagglutination; ORF, open reading frame; PAGE, polyacrylamide gel electrophoresis; Pap, pili associated with pyelonephritis; *papA*, structural gene for subunit of Pap pili; RF, replicative form; SDS, sodium dodecyl sulfate.

have a regulatory function in the expression of the genes involved, e.g. in the regulation of phase variation.

The fimbrial subunit protein should possess a number of properties. For translocation it should have properties similar to periplasmic or outer membrane proteins; in addition, it should carry information needed for the assembly into fimbriae. Comparison of the amino acid sequences of different fimbrial subunit proteins could lead to the prediction of essential regions involved in these functions. In this paper we present the complete nucleotide sequence of the $F7_2$ fimbrial subunit gene and its 5'-terminal flanking region. This sequence is compared with the published nucleotide sequence of the *papA* gene, encoding the fimbrial subunit of the Pap fimbriae (Båga et al., 1984).

MATERIALS AND METHODS

(a) Bacterial strains and phages

All bacterial strains are *E. coli* K-12 derivatives. JA221 is a *recA hsdA* derivative of *E. coli* K-12 C600 (Clarke and Carbon, 1978). JM103 (Messing and Vieira, 1982) was used to propagate phages M13 mp8, mp9 and clones derived from these strains.

(b) Transformation

Transformation was carried out essentially as described by Kushner (1978).

(c) DNA isolation

Isolation of plasmid or phage RF DNA was carried out according to Birnboim and Doly (1979), followed by CsCl-EtBr isopycnic centrifugation in the case of large-scale isolations. Isolation of single-stranded M13 DNA was performed essentially according to Heidecker et al. (1980). Other techniques were as described before (Van Die et al., 1984).

(d) Isolation of plasmids carrying an inserted Tn*5* element

Insertions of the transposable element Tn*5* (Berg, 1977) into plasmid pPIL110-352 were described before (Van Die et al., 1984).

(e) Nucleotide sequencing

Nucleotide sequencing was carried out essentially according to the dideoxy chain termination method of Sanger et al. (1977; 1980). Various restriction fragments of pPIL110-352 or Tn*5* derivatives of pPIL110-352 were cloned in M13mp8 or M13mp9. Sequence data were processed on an Apple II computer with a slightly modified version of the Larson and Messing (1983) software.

RESULTS AND DISCUSSION

(a) Physical localization of the $F7_2$ fimbrial subunit gene

In Fig. 1a the approximate location of the genes *A* through *E*, involved in the synthesis of the $F7_2$ fimbriae, in pPIL110-35 is shown (Van Die et al., 1984). To facilitate a more precise localization of gene *A*, encoding the fimbrial subunit, a smaller plasmid (pPIL110-352) was constructed by cloning an *Eco*RI fragment of the $\gamma\delta$ insertion mutant 10 (Van Die et al., 1984) into pACYC184 (Fig. 1a, b). pPIL110-352 and pPIL110-353, a pBR322 derivative carrying genes *B* through *E*, complement each other to express MRHA. Tn*5*-insertion mutants of pPIL110-352 were isolated and examined in this complementation test for expression of gene *A*. Next to a number of Tn*5* insertion derivatives that showed normal expression of gene *A* in the complementation test, one derivative (pPIL110-352::Tn*5*-31) that showed no MRHA and one derivative (pPIL110-352::Tn*5*-32) showing only very weak MRHA were found. The place of insertion of the Tn*5* element was determined for all derivatives. In the case of pPIL110-352::Tn*5*-31 and pPIL110-352::Tn*5*-32, the Tn*5* element was located within a 1.1-kb *Rsa*I fragment (Fig. 1c), while the transposon insertions that did not affect MRHA had occurred outside this fragment. The expression of proteins encoded by pPIL110-352 and its transposon derivatives was studied by SDS-PAGE analysis of the polypeptides synthesized in minicells.

pPIL110-352::Tn*5*-31 did not encode the 17-kDal fimbrial protein, nor its 18 kDal precursor protein, which is normally observed when minicells

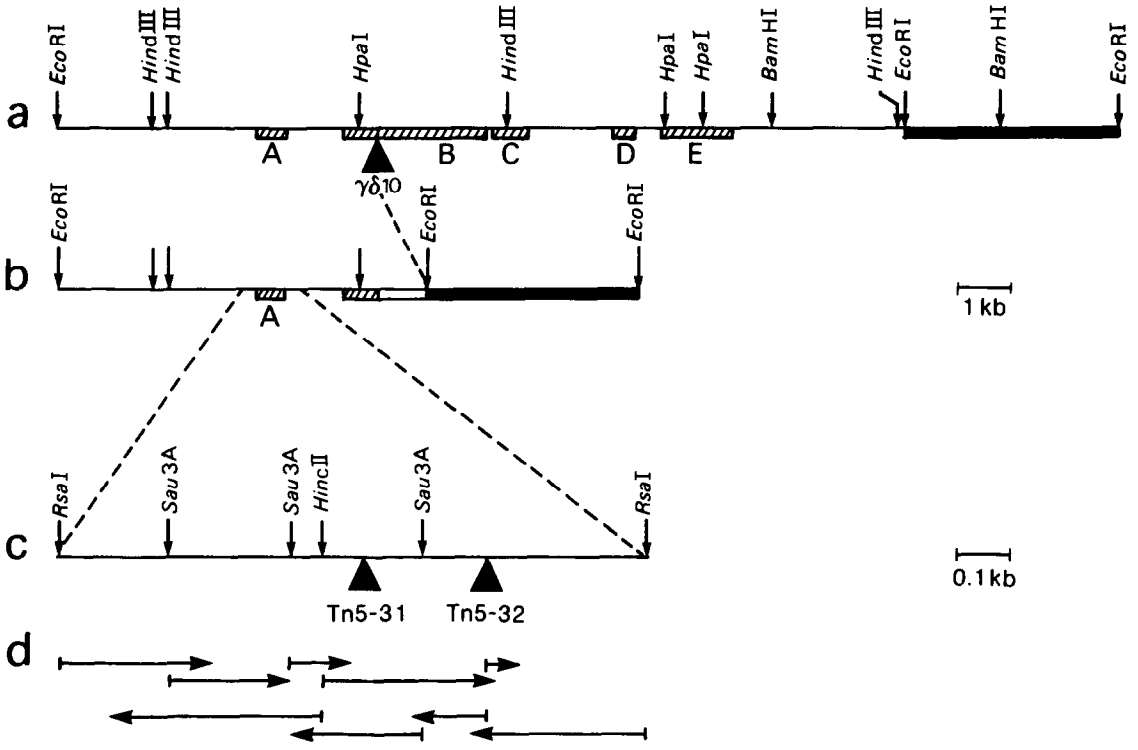

Fig. 1. Physical localization of the $F7_2$ fimbrial subunit gene (gene *A*) and sequencing strategy for this gene. The hatched boxes represent the genes *A–E*, involved in the expression of $F7_2$ fimbriae. The solid bars represent pACYC184 DNA. The open box in Fig. 1b represents γδ DNA. (a) Genetic map of pPIL110-35 and location of γδ insertion 10 (Van Die et al., 1984). (b) Genetic map of pPIL110-352, a derivative of pPIL110-35::γδ 10. (c) Physical location of the two Tn*5* elements inserted in the fimbrial subunit gene. (d) Strategy of nucleotide sequence analysis. The direction of the sequenced DNA fragments is shown by arrows.

are labeled in 9.5% ethanol (Van Die et al., 1984). pPIL110-352::Tn*5*-32, however, encoded a 16-kDal protein that was found as a precursor protein of about 17 kDal. These results suggest that the Tn*5* element in pPIL110-352::Tn*5*-32 is located near the C-terminal end of the fimbrial subunit gene, resulting in the synthesis of a truncated protein. As the Tn*5* insertion in pPIL110-352::Tn*5*-31 was located to the left of the Tn*5* element in pPIL110-352::Tn*5*-32 (Fig. 1b), the direction of transcription of gene *A* probably is from left to right.

Based on the location of these Tn*5* insertions it can be calculated that the 1.1-kb *Rsa*I fragment should encompass the entire coding sequence of the fimbrial subunit gene.

(b) Nucleotide sequence of the $F7_2$ fimbrial subunit gene

The strategy for the determination of the nucleotide sequence is outlined in Fig. 1d. Restriction fragments of pPIL110-352 and of its Tn*5*-32 derivative were cloned in M13mp8 or mp9 and sequenced. The nucleotide sequence of 770 bp, which was established by analyzing both DNA strands, is shown in Fig. 2. This sequence contains an ORF of 585 bp starting at an ATG triplet at nucleotide 175 and ending at a TAA triplet at nucleotide 760. However, since this potential initiation codon is not preceded by a suitable ribosome-binding site, but the next ATG triplet in the same reading frame at nucleotide 197 is (GAGG at nucleotide 184–187; Gold et al., 1981), we suppose that the latter is the translational initiation codon for the precursor form of the $F7_2$ protein.

The polypeptide predicted by this sequence has an M_r of 19160, which is slightly larger than the apparent M_r of the precursor of the $F7_2$ protein as estimated by SDS-PAGE (Van Die et al., 1984). Sequence analysis of the Tn*5* insertions 31 and 32 revealed that insertion had occurred in this ORF behind base pairs 489 and 708, respectively. From

```
                                                    CCGGAGTTTCTGGAAGATAAAAAAAGAAGCCCTTATCA

                50                                                                 100
GAAAGCAGACAGGTTATATCAGTATTCTGTCGATAAATAACCTGCCCTGAAAATACGAGAATATTATTTGTATTGATCT

                                150
GGTTATTAAAGGTAATCGGGTCATTTTAAATTGCCAGATATCTCTGGTGTGTTCAGTAATGAAAAAGAGGTTGTTATTT

  200                                                                          250
ATG ATT AAG TCG GTT ATT GCC GGT GCG GTA GCT ATG GCA GTG GTG TCT TTT GGT GCA TAT
Met Ile Lys Ser Val Ile Ala Gly Ala Val Ala Met Ala Val Val Ser Phe Gly Ala Tyr
    -20

                                                                  300
GCT GCT CCA ACT ATT CCT CAG GGG CAG GGT AAA GTA ACT TTT AAC GGA ACT GTA GTA GAT
Ala Ala Pro Thr Ile Pro Gln Gly Gln Gly Lys Val Thr Phe Asn Gly Thr Val Val Asp
 -1  +1                                                    15

                                        350
GCA CCA TGT GGT ATT GAT GCT CAG TCT GCT GAT CAA TCT ATT GAT TTT GGA CAA GTA TCA
Ala Pro Cys Gly Ile Asp Ala Gln Ser Ala Asp Gln Ser Ile Asp Phe Gly Gln Val Ser
                                        30

                            400
AAA TTA TTT CTG GAG AAT GAT GGG GAA AGT CAA CCC AAA TCT TTT GAT ATT AAA CTT ATA
Lys Leu Phe Leu Glu Asn Asp Gly Glu Ser Gln Pro Lys Ser Phe Asp Ile Lys Leu Ile
                      45

              450
AAT TGT GAT ATT ACA AAC TTT AAA AAA GCT GCT GGC GGT GGT GGG GCG AAG ACT GGC ACA
Asn Cys Asp Ile Thr Asn Phe Lys Lys Ala Ala Gly Gly Gly Gly Ala Lys Thr Gly Thr
 60                                                      75

  500                                                                          550
GTA TCT CTG ACT TTT TCG GGT GTC CCA AGT GGT CCG CAG AGT GAC ATG TTA CAG ACC GTC
Val Ser Leu Thr Phe Ser Gly Val Pro Ser Gly Pro Gln Ser Asp Met Leu Gln Thr Val
                                        90

                                                          600
GGT GCA ACA AAT ACA GCT ATT GTT GTC ACC GAT CCA CAT GGA AAA CGC GTA AAA TTT GAT
Gly Ala Thr Asn Thr Ala Ile Val Val Thr Asp Pro His Gly Lys Arg Val Lys Phe Asp
                    105

                                        650
GGT GCG ACA GCA ACA GGT GTT TCC TAT TTA GTT GAT GGT GAT AAC ACA ATT CAT TTT ACT
Gly Ala Thr Ala Thr Gly Val Ser Tyr Leu Val Asp Gly Asp Asn Thr Ile His Phe Thr
120                                                     135

                            700
GCT GCC GTC AGA AAA GAT GGT AGT GGC AAC CCT GTA ACA GAA GGT GCT TTC TCA GCA GTT
Ala Ala Val Arg Lys Asp Gly Ser Gly Asn Pro Val Thr Glu Gly Ala Phe Ser Ala Val
                                        150

              750
GCG AAT TTC AAC CTG ACT TAT CAG TAA CAGTAATA
Ala Asn Phe Asn Leu Thr Tyr Gln Stop
                    165
```

Fig. 2. Nucleotide sequence and deduced amino acid sequence of the $F7_2$ fimbrial subunit gene. Numbering of the residues starts at the lefthand side of the DNA fragment sequenced in both directions (see Fig. 1, c, d). Numbers above each line refer to the nucleotide positions. The last digit refers to numbered nucleotide. Amino acid numbers are specified below each line. The first amino acid of the mature protein is numbered + 1.

these results we conclude that the gene we have sequenced encodes the $F7_2$ subunit.

(c) Signal sequence of the $F7_2$ protein

As the $F7_2$ fimbrial subunit should be transported through membranes during biogenesis of fimbriae, it can be expected that it is produced in a precursor form with an N-terminal signal sequence. The N-terminal part of the amino acid sequence deduced from the nucleotide sequence of the subunit gene has the properties of a typical prokaryotic signal sequence (Inouye and Halegoua, 1980; Perlman and Halvorson, 1983). As the first N-terminal amino acids of the $F7_2$ protein are not known, the signal peptidase cleavage site cannot be determined unambiguously. We suppose that the cleavage site ala-tyr-ala↓ala (between residue − 1 and + 1 in Fig. 2) is the most

likely one. This site is compatible with known prokaryotic signal peptidase cleavage sites (Perlman and Halvorson, 1983). Moreover, the N terminus of the resulting $F7_2$ mature protein, as shown in Figs. 2 and 3, is completely homologous with the N terminus of the mature PapA protein.

(d) Comparison of the $F7_2$ protein with other fimbrial subunit proteins

Comparison of the amino acid sequence of the N-terminal end of the mature $F7_2$ protein, deduced from the nucleotide sequence of the gene, with the published N-terminal amino acid sequence of the F7 fimbrial subunit protein (Klemm et al., 1982) revealed that both N-terminal ends are in good agreement but not identical (Fig. 3). Strain AD110 expresses, next to 1A and 1C fimbriae, two closely related types of F7 fimbriae, $F7_1$ and $F7_2$ (Ørskov et al., 1982b). The F7 fimbriae isolated and studied by Klemm et al. (1982) appear to be of the $F7_1$ serotype (Klemm, P., personal communication), while the fimbrial subunit studied here is of the $F7_2$ serotype (Ørskov, I., personal communication). This explains the observed difference between the N termini. Fig. 4 shows a comparison between the nucleotide sequences and the predicted amino acid sequences of the $F7_2$ and PapA (Båga et al., 1984) fimbrial subunits. The strains from which these genes were cloned appear to be quite different. The $F7_2$ gene has been cloned from an *E. coli* O6: K2 strain isolated in the U.K., whereas the *papA* fimbrial gene has been cloned from an *E. coli* O4: K6 strain isolated in the USA. Therefore, it seems likely that both genes have a long separate history. Still, they are both functional as subunits of fimbriae involved in MRHA and will have retained the amino acid sequences essential for their function. The two nucleotide sequences in Fig. 4 have been aligned for maximum homology of the amino acid sequences. Gaps have been introduced to reach maximum homology according to Sweet and Eisenberg (1983). The nucleotide sequence of $F7_2$ preceding the coding region is identical to the nucleotide sequence of *papA* over the entire region where sequence data of *papA* are available (129 bp). This extreme sequence conservation could indicate a regulatory function for this DNA stretch that is the same for both genes. In a search for possible promoter sites, several suitable −10 (TATAAT-like) regions were found in the noncoding region, but none was preceded by a sequence homologous in the −35 (TTGACA) region (Hawley and McClure, 1983).

In the coding region the two genes also show remarkable homology at the level of nucleotide sequence (68%) as well as at the level of amino acid sequence (66%). The signal sequences of the $F7_2$ and PapA proteins are completely homologous except for two and three amino acids, respectively, near the C-terminal end of the signal peptide. The N-terminus (22 amino acids) and the C-terminus (17 amino acids) of the mature protein are identical. In between, the homology generally decreases from the N terminus to the C terminus.

Homology between N termini of several fimbrial subunits has been noted before (Fig. 3; for a review see Jones and Isaacson, 1983). In Fig. 3 two groups can be distinguished, the MRHA fimbriae F7, $F7_2$ and PapA, and the type 1-like fimbriae FimA and 1C, but both groups are interrelated. The virtually identical position of the cysteine residues is striking. Homology between the $F7_2$ and PapA proteins decreases between the two cysteine residues (Figs. 3 and 4). Between the MRHA fimbriae and type 1-like

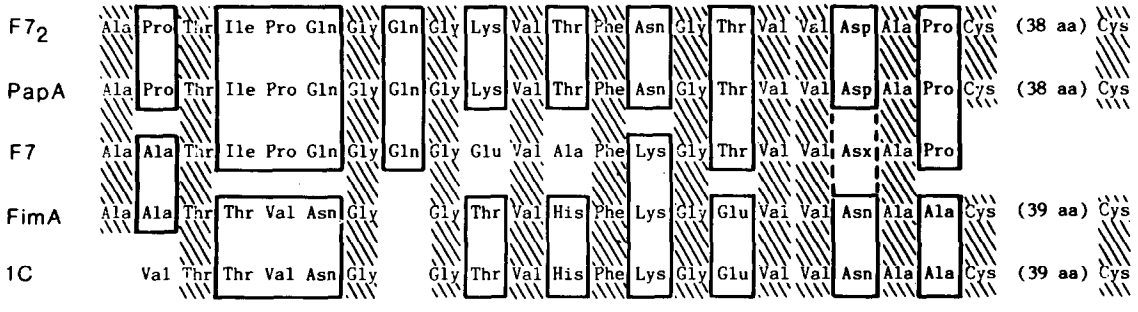

Fig. 3. Comparison of the N-terminal amino acid sequences of the $F7_2$, PapA (Båga et al., 1984), F7 (Klemm et al., 1982), FimA (Klemm, 1984) and 1C (Van Die, I., Van Geffen, B., Hoekstra, W. and Bergmans, H., manuscript in preparation) fimbrial subunit proteins.

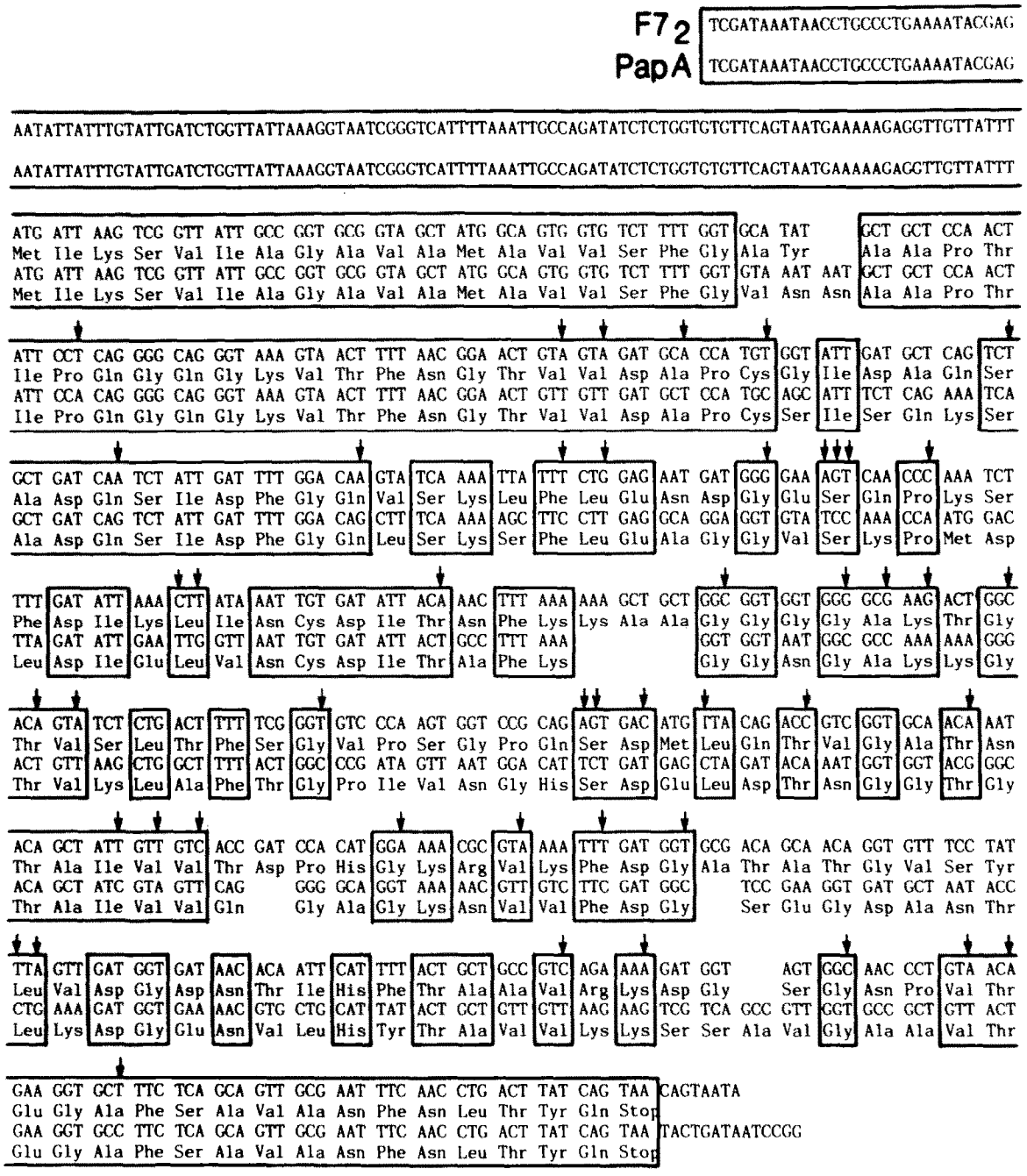

Fig. 4. Comparison of the nucleotide sequence and deduced amino acid sequence of the genes encoding the $F7_2$ and PapA fimbrial subunits. Homologous amino acids and the homologous 5'-flanking region are in boxes. Silent nucleotide changes are indicated by arrows. Gaps have been introduced to obtain maximum homology between the two proteins.

fimbriae, homology is still found between the two cysteine residues although to a lesser extent than within both groups (Klemm, 1984; Van Die, I., Van Geffen, B., Hoekstra, W. and Bergmans, H., manuscript in preparation).

There is also some homology (8 out of 17 amino acids) between the C-termini of the $F7_2$ (and PapA) fimbrial protein and the FimA subunit (Klemm, 1984). This homology is lost farther upstream from the C terminus. Comparison of the $F7_2$ subunit protein and the PapA protein in the region between the conserved N- and C-terminal parts shows a high degree of homology at the amino acid level (Fig. 2), and this homology is generally reflected in the

conformational properties of the two proteins. The hydrophilicity profile, calculated according to the method of Hopp and Woods (1981), shows six hydrophobic maxima for the $F7_2$ protein, located at residues 35, 101, 126, 147, 158 and at the C-terminus (residue 182) in Fig. 5; five of these maxima are also observed in the hydrophilicity profile of the PapA subunit, at comparable positions (35, 106, 125, 158 and 182) in reasonably conserved regions. The sixth maximum at residue 147 of $F7_2$ is in a region where no amino acid sequence homology between the two proteins is observed. As the hydrophobic regions would be located on the inside of the protein, they are probably involved in the determination of the structure of the protein, and in the interaction of the subunits in the fimbriae. Conservation of these regions is therefore to be expected.

Several hydrophilic peaks are observed in the profile of the $F7_2$ fimbrial protein, the most pronounced at residues 66, 132–136 and 165. The former two hydrophilic maxima are adjacent to predicted turns in the secondary structure of the peptide chain, as calculated according to the method of Chou and Fasman (1971). This suggests that at least these two peaks represent the location of possible antigenic determinants. The hydrophilicity profile of the PapA protein shows less pronounced peaks at the same positions (Fig. 5).

In conclusion, the high degree of amino acid sequence conservation at the N and C termini of the $F7_2$ and PapA proteins suggests that the genes coding for these proteins have originated from one primordial gene. In these regions the homology between the $F7_2$ and PapA proteins and the type 1 fimbrial proteins is also significant, but less extensive, suggesting a more distant relationship.

The hypothesis that these conserved parts of the proteins are involved in the one feature that all of them should have in common, i.e. the possibility to be assembled into fimbriae, is tempting. From the sequence data of pPIL110-352::Tn*5*-32 the localization of the Tn*5* element is known precisely. The truncated protein coded for by this mutant plasmid has proline residue 150 (Fig. 2) at its C terminus followed by a termination codon provided by Tn*5*. This truncated protein thus exactly lacks the conserved C-terminal part. Currently we are studying the behavior of this truncated protein in the biogenesis of fimbriae.

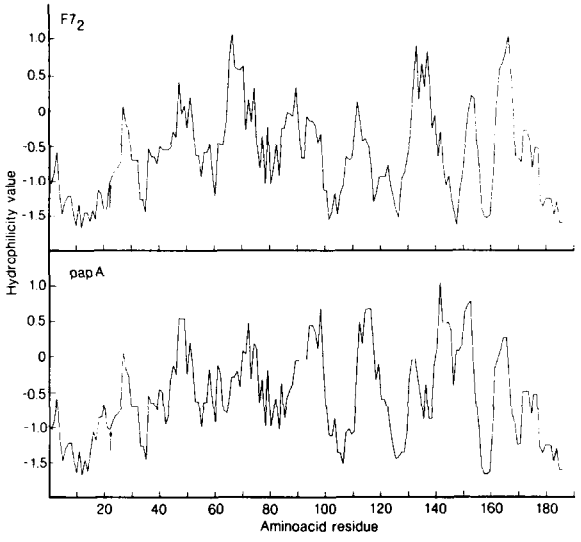

Fig. 5. Hydrophilicity profile of the $F7_2$ fimbrial subunit protein, compared with the published hydrophilicity profile of the PapA fimbrial protein (Båga et al., 1984). Arrows indicate the signal peptidase cleavage sites of the proteins. Gaps have been introduced to reach maximum homology (see Fig. 4). Amino acid residues have been numbered starting at the N-terminal end of the signal sequence, including the gaps. Hydrophilicity values are shown above the first residue of the sextet that was considered in the calculation (Hopp and Woods, 1981).

ACKNOWLEDGEMENTS

We thank Renske Teertstra for introducing us to the M13 sequencing method and for valuable suggestions during the experiments, and Wiel Hoekstra for critical reading of the manuscript.

REFERENCES

Båga, M., Normark, S., Hardy, J., O'Hanley, P., Lark, D., Olsson, O., Schoolnik, G. and Falkow, S: Nucleotide sequence of the *papA* gene encoding the *papA* pilus subunit of human uropathogenic *Escherichia coli*. J. Bacteriol. 157 (1984) 330–333.

Berg, D.E.: Insertion and excision of the transposable kanamycin resistance determinant Tn*5*, in Bukhari, A.I., Adhya, S.A. and Shapiro, J.A. (Eds.), DNA Insertion Elements, Plasmids and Episomes. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1977, pp. 205–212.

Birnboim, H.C. and Doly, J.: A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7 (1979) 1513–1524.

Chou, P.Y. and Fasman, G.D.: Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47 (1971) 45–148.

Clarke, L. and Carbon, J.: Functional expression of cloned yeast DNA in *Escherichia coli*: specific complementation of arginino succinate lyase (*arg H*) mutations. J. Mol. Biol. 120 (1978) 517–532.

Clegg, S.: Cloning of genes determining the production of mannose resistant fimbriae in a uropathogenic strain of *Escherichia coli* belonging to serogroup O6. Infect. Immun. 38 (1982) 739–744.

Clegg, S. and Pierce, J.K.: Organization of genes responsible for the production of mannose-resistant fimbriae of a uropathogenic *Escherichia coli* isolate. Infect. Immun. 42 (1983) 900–906.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B.S. and Stormo, G.: Translational initiation in prokaryotes. Annu. Rev. Microbiol. 35 (1981) 365–403.

Hagberg, L., Jodal, U., Korhonen, T.K., Lidin-Jason, G., Lindberg, U. and Svanborg-Edén, C.: Adhesion, hemagglutination and virulence of *Escherichia coli* causing urinary tract infections. Infect. Immun. 31 (1981) 564–570.

Hawley, D.K. and McClure, W.R.: Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucl. Acids Res. 11 (1983) 2237–2255.

Heidecker, G., Messing, J. and Gronenborn, B.: A versatile primer for DNA sequencing in the M13mp2 cloning system. Gene 10 (1980) 69–73.

Hopp, T.P. and Woods, K.R.: Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA 78 (1981) 3824–3828.

Hull, R.A., Gill, R.E., Hsu, P., Minshew, B.H. and Falkow, S.: Construction and expression of recombinant plasmids encoding type 1 or D-mannose resistant pili from a urinary tract infection *Escherichia coli* isolate. Infect. Immun. 33 (1981) 933–938.

Inouye, M. and Halegoua, S.: Secretion and membrane localization of proteins in *E. coli*. Crit. Rev. Biochem. 7 (1980) 339–371.

Jann, K., Jann, B. and Schmidt, G.: SDS-polyacrylamide gel electrophoresis and serological analysis of pili from *Escherichia coli* of different pathogenic origin. FEMS Microbiol. Lett. 11 (1981) 21–25.

Jones, G.W. and Isaacson, R.E.: Proteinaceous bacterial adhesins and their receptors. CRC Crit. Rev. Microbiol. 10 (1983) 229–260.

Klemm, P.: The *fimA* gene encoding the type 1 fimbrial subunit of *Escherichia coli*: nucleotide sequence and primary structure of the protein. Eur. J. Biochem. (1984) in press.

Klemm, P., Ørskov, I. and Ørskov, F.: F7 and type 1-like fimbriae from three *Escherichia coli* strains isolated from urinary tract infections: protein chemical and immunological aspects. Infect. Immun. 36 (1982) 462–468.

Kushner, S.R.: An improved method for transformation of *Escherichia coli* with ColE1 derived plasmids, in Boyer, H.W. and Nicosia, S. (Eds.) Genetic Engineering. Elsevier, Amsterdam, 1978, pp. 17–23

Larson, R. and Messing, J.: Apple II computer software for DNA and protein sequence data. DNA 2 (1983) 31–35.

Messing, J. and Vieira, J.: A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene 19 (1982) 269–276.

Normark, S., Lark, D., Hull, R., Norgren, M., Båga, M., O'Hanley, P., Schoolnik, G. and Falkow, S.: Genetics of digalactoside-binding adhesin from a uropathogenic *Escherichia coli* strain. Infect. Immun. 41 (1983) 942–949.

Ørskov, I., Ørskov, F. and Birch-Andersen, A.: Comparison of *Escherichia coli* fimbrial antigen F7 with type 1 fimbriae. Infect. Immun. 27 (1980) 657–666.

Ørskov, I., Ørskov, F., Birch-Anderson, A., Kanamori, M. and Svanborg-Edén, C.: O, K, H and fimbrial antigens in *Escherichia coli* associated with pyelonephritis and cystitis. Scan. J. Infect. Dis. Suppl. 33 (1982a) 18–25.

Ørskov, I., Ørskov, F., Klemm, P. and Svanborg-Edén, C.: Protein attachment factors: fimbriae in adhering *E. coli* strains, in Robbins, J.B., Hill, J.C. and Sadoff, J.C. (Eds.), Seminars in Infectious Diseases, Vol. 4, Bacterial Vaccines. Thieme Stratton, New York, 1982b, pp. 99–103.

Perlman, D. and Halvorson, H.O.: A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J. Mol. Biol. 167 (1983) 391–409.

Rhen, M., Knowles, J., Pentillä, M.E., Sarvas, M. and Korhonen, T.K.: P fimbriae of *Escherichia coli*: Molecular cloning of DNA fragments containing the structural genes. FEMS Microbiol. Lett. 19 (1983) 119–123.

Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chainterminating inhibitors. Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A.: Cloning in single-stranded bacteriophage as an aid to rapid DNA-sequencing. J. Mol. Biol. 143 (1980) 161–178.

Svanborg-Edén, C., Hanson, L.A., Jodal, U., Lindberg, U. and Sohl-Åkerlund, A.: Variable adherence to normal human urinary epithelial cells of *E. coli* strains associated with various forms of urinary tract infection. Lancet (1976) 490–492.

Sweet, R.M. and Eisenberg, D.: Correlation of sequence hydrofobicity measures similarity in three-dimensional protein structure. J. Mol. Biol. 171 (1983) 479–488.

Van den Bosch, J.F., Verboom-Sohmer, U., Postma, P., De Graaff, J. and McLaren, D.M.: Mannose sensitive and mannose resistant adherence to human uroepithelial cells and urinary virulence of *Escherichia coli*. Infect. Immun. 29 (1980) 226–233.

Van Die, I., Van den Hondel, C., Hamstra, H., Hoekstra, W. and Bergmans, H.: Studies on the fimbriae of an *Escherichia coli* O6:K2:H1:F7 strain: molecular cloning of a DNA fragment encoding a fimbrial antigen responsible for mannose resistant hemagglutination of human erythrocytes. FEMS Microbiol. Lett. 19 (1983) 77–82.

Van Die, I., Van Megen, I., Hoekstra, W. and Bergmans, H.: Molecular organization of the genes involved in the production of $F7_2$ fimbriae, causing mannose-resistant hemagglutination of a uropathogenic *Escherichia coli* O6:K2:H1:F7 strain. Mol. Gen. Genet. 194 (1984) 528–533.

Communicated by H. van Ormondt.