JOURNAL OF BACTERIOLOGY, Aug. 1989, p. 4281–4289
0021-9193/89/084281-09$02.00/0

Vol. 171, No. 8

Molecular Characterization of a Fimbrial Adhesin, F1845, Mediating Diffuse Adherence of Diarrhea-Associated *Escherichia coli* to HEp-2 Cells

SIMA S. BILGE,[1] CARLA R. CLAUSEN,[2,3] WAYNE LAU,[1]† AND STEVE L. MOSELEY[1]*

Departments of Microbiology[1] and Laboratory Medicine,[2] University of Washington, Seattle, Washington 98195, and Children's Hospital and Medical Center, Seattle, Washington 98105[3]

Received 15 February 1989/Accepted 18 May 1989

A fimbrial adhesin, designated F1845, was found to be responsible for the diffuse HEp-2 cell adherence of a diarrheal *Escherichia coli* isolate. The genetic determinant of F1845 was cloned, and the order of the genes necessary for production of F1845 was determined by maxicell analysis. Five polypeptides with apparent sizes of 10, 95, 27, 15.5, and 14.3 kilodaltons (kDa) were found to be encoded in that order by the F1845 determinant. The nucleotide sequence of the 14.3-kDa subunit gene was determined and found to share extensive homology in its signal sequence with the gene encoding the structural subunit of the AFA-I hemagglutinin of a uropathogenic *E. coli* strain (A. Labigne-Roussel, M. A. Schmidt, W. Walz, and S. Falkow, J. Bacteriol. 162:1285–1292, 1985) but not in the region encoding the mature protein. Southern blot hybridizations indicated that the F1845 determinants are of chromosomal origin. Hybridization studies using a probe from the region encoding the 95-kDa polypeptide indicated that related sequences may be plasmid associated in some strains and chromosomal in others. Additional hybridization studies of *E. coli* isolates possessing sequence homology to the F1845 determinant suggest that the sequences in the 5′ region of the F1845 structural subunit gene are more highly conserved than sequences in the 3′ region.

Adhesins of diarrheagenic and uropathogenic *Escherichia coli* strains play an essential role in the colonization of human mucosal epithelium. These adhesins are responsible for the binding of bacteria to a specific cell surface receptor. Their presence is also frequently associated with the ability of the bacteria to agglutinate human erythrocytes in the presence of D-mannose, referred to as mannose-resistant hemagglutination (MRHA).

E. coli strains associated with diarrheal disease have been shown to adhere to tissue culture cells in characteristic patterns (41). Many *E. coli* strains with enteropathogenic serotypes have been shown to adhere to HEp-2 cells in a localized pattern (8, 42). This has been referred to as localized adherence and is mediated by a plasmid-encoded adhesin which remains largely uncharacterized (31). Other strains of *E. coli* isolated from persons with diarrhea have been shown to adhere to HEp-2 cells in an aggregative pattern (30). These strains have been termed enteroadherent-aggregative *E. coli* and have been demonstrated in an epidemiological study to be associated with diarrheal disease (20). A third class of adherence displayed by *E. coli* isolated from persons with diarrhea is characterized by a diffuse pattern of association with the cell surface (41). This adherence has been termed diffuse adherence (DA), and these strains also frequently exhibit MRHA. The association of DA with diarrheal disease is not clear. One study in Peru found no association (20), but continuing studies in Mexico indicate that strains that exhibit DA are associated with diarrheal disease (G. Giron et al., manuscript in preparation).

We report here the characterization of the genetic basis of DA in an *E. coli* strain isolated from a child with protracted diarrhea and report the nucleotide sequence of the gene that encodes the structural subunit of the fimbrial adhesin which mediates DA.

MATERIALS AND METHODS

Bacterial strains, plasmids and media. Strain C1845 is an *E. coli* strain of serotype O75:NM isolated from a child with diarrhea of 3 weeks duration from whom no other enteric pathogen was identified. Additional clinical *E. coli* isolates from children with diarrhea were obtained from the collection of Children's Hospital and Medical Center, Seattle, Wash., or were kindly provided by G. Overturf, Department of Pediatrics, University of California, Los Angeles; J. Matthewson, Program in Infectious Diseases and Clinical Microbiology, University of Texas Medical School at Houston; M. Marks, Department of Pediatrics, University of California, Irvine; and M. Karmali, Hospital for Sick Children, Toronto, Ontario, Canada. Isolates were also obtained from fecal specimens of healthy children in Seattle. *E. coli* K-12 strain LE392 (23) was used as a recipient for cosmid infection and as a transformation recipient. Strain DH5α (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) was used as a host for plasmid constructs, strain HB101 (4) was used for transposon mutagenesis, and strain JC2926 (2) was used in maxicell analysis. Strain GW5180, kindly provided by Graham Walker, Department of Biology, Massachusetts Institute of Technology, Boston, is a *recA* derivative of JM101 (25) and was used as a host for bacteriophage M13 derivatives in nucleotide sequence analysis. Plasmid pHC79 (14) was used as a vector for cosmid cloning. Plasmid pACYC184 (7) and multicopy plasmids pUC8, pUC18, and pUC19 (47) were used as cloning vectors. Unless otherwise indicated, organisms were grown in Luria broth or Luria agar (26). Cells harboring derivatives of pUC were grown in the presence of ampicillin (100 μg/ml). Kanamycin (25 μg/ml) was used to select for transposon Tn*5* insertions.

* Corresponding author.
† Present address: Tufts University School of Medicine, Boston, MA 02111.

MRHA assay. Bacterial cells were tested for the ability to agglutinate human group O erythrocytes in the presence of mannose as previously described (27). Neuraminidase-treated erythrocytes were prepared as described by Lindahl et al. (21). Latex beads coated with the P blood group antigen were obtained from Chembiomed Ltd., Edmonton, Alberta, Canada.

HEp-2 and HL cell adherence. HEp-2 cell adherence was assayed by the method of Clausen and Christie (8). HL cells (6) were also used in this assay with identical results.

Purification of adhesins. Purification of the adhesin of strain C1845 was performed by the procedure of de Graaf and Roorda (9), except that deoxycholate treatment was omitted. For purification of the adhesin of strain HB101(pSSS1) (described in Results), broth-growth cells were suspended in phosphate-buffered saline (PBS; 10 mM sodium phosphate, 0.128 M NaCl, pH 7.4) and sheared in an Omnimixer (Ivan Sorvall, Inc., Norwalk, Conn.) for 30 min. The cells were pelleted by centrifugation. Material which precipitated after 60% saturation of the supernatant with ammonium sulfate was suspended in PBS and dialyzed extensively against the same solution.

Electron microscopy. Unfixed bacterial suspensions were examined for fimbriae with a JEOL 100B transmission electron microscope at 60 kV. Plasmids of interest were transformed into *E. coli* HB101. Bacteria were harvested from Luria agar plates into distilled water, diluted into an equal volume of 2% (wt/vol) phosphotungstic acid (pH 7.4), and placed on carbon-coated copper grids.

Production of antisera. Rabbits were immunized with the purified adhesin of strain C1845. An initial immunization dose consisted of 500 μg of protein in water emulsified with an equal volume of complete Freund adjuvant administered by subcutaneous injection at four sites. A second dose of 500 μg was similarly administered 2 weeks later, except that incomplete Freund adjuvant was used. Rabbits were exsanguinated 11 days following the second injection. The antiserum was diluted 1:100 and adsorbed with boiled *E. coli* DH5α.

Cosmid cloning. DNA was prepared from strain C1845 by the method of Hull et al. (15). The DNA was partially digested with the restriction enzyme *Sau*3A to give fragments of 30 to 40 kilobases. The restricted DNA was ligated to *Bam*HI-cleaved pHC79. Preparation of packaging extract, the in vitro packaging reaction, and infection of the recipient strain were performed by the methods of Maniatis et al. (23).

Insertion mutagenesis. Transposon Tn*5* was introduced into *E. coli* by infection with bacteriophage lambda *b*221 *rex*::Tn*5* *c*I857 *O*am8 *P*am29 (5). Transductants growing on medium containing both ampicillin and kanamycin were harvested, and total plasmid DNA was isolated by the method of Birnboim and Doly (3). The resulting pool of plasmid DNA was used to transform strain HB101. Transformants were selected on medium with ampicillin and kanamycin. 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 (3) or by cesium chloride-ethidium bromide density gradient centrifugation following alkaline lysis as described by Maniatis et al. (23). Plasmids from clinical *E. coli* isolates were isolated by the method described by Kado and Liu (16). Restriction endonuclease digests were performed as recommended by the manufacturers (Bethesda Research Laboratories; Boehringer Mannheim Biochemicals, Indianapolis, Ind.; and International Biotechnologies, Inc., New Haven, Conn.). Digests were analyzed after 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. (39). Overnight cultures of organisms harboring plasmids of interest grown in K broth (38) at 37°C were diluted 1:50 in the same medium, grown to the exponential phase, and then exposed to UV radiation from a germicidal lamp. Following irradiation, cultures were incubated for 1 h at 37°C. Cycloserine was then added to a final concentration of 200 μg/ml, and incubation was continued for 12 to 14 h at 37°C. Cells were harvested and washed three times with Hershey salts (39) and then incubated for 30 to 60 min in Hershey medium (39). [^{35}S]methionine (1,100 Ci/mmol; ICN Biomedicals, Costa Mesa, Calif.) was then added to a final concentration of 4 μCi/ml, and incubation was continued for an additional hour. Cells were then washed once with Hershey salts and then lysed in 1% sodium dodecyl sulfate (SDS). The radioactive polypeptides were separated on denaturing gradient 5 to 20% (wt/vol) polyacrylamide–SDS gels as described by Laemmli (19) and visualized following autoradiography.

Immunoblot procedures. Proteins of overnight cultures of *E. coli* were separated by gel electrophoresis on a denaturing gradient 5 to 20% polyacrylamide–SDS gel as described by Laemmli. Transfer of the proteins to nitrocellulose filters was accomplished by the method of Towbin et al. (45). Filters were preadsorbed in PBS with 5% nonfat dry milk for 1 h. The filters were then washed twice in PBS with 0.05% Tween 20 and incubated for 4 h in anti-F1845 serum which had been diluted 1:100 in TN buffer (50 mM Tris hydrochloride, 150 mM NaCl, pH 7.5) with 3% bovine serum albumin and 2% normal goat serum. The filters were washed extensively in PBS with 0.05% Tween 20. Bound antibody was detected with goat anti-rabbit peroxidase (Kirkegaard and Perry Laboratories, Inc.) as recommended by the manufacturer (12).

DNA sequence analysis. The dideoxy-chain termination method described by Sanger et al. (40) was used with the Klenow fragment of *E. coli* DNA polymerase (Boehringer Mannheim Biochemicals) and modified T7 DNA polymerase (Sequenase; United States Biochemicals, Cleveland, Ohio). Bacteriophages M13mp18 and M13mp19 were used as the sources of single-stranded DNA templates (32). Computer analysis used version 5.3 of the sequence analysis software package of the University of Wisconsin Genetics Computer Group (10). Alignment and sequence similarity determinations used the BestFit program of this software package.

Preparation of DNA hybridization probes. Probes were labeled with α-^{32}P-labeled deoxynucleoside triphosphates (Dupont, NEN Research Products, Boston, Mass.) by the random primer method (11) or by nick translation (24).

Southern blot hybridizations. Preparation, hybridization, and washing of Southern blots under conditions of high stringency were performed as described previously (28, 29, 44).

Colony hybridizations. For colony hybridizations, preparation, hybridization, and washing of blots under conditions of high stringency were performed as described previously (22, 27).

RESULTS

Adherence characteristics of strain C1845. *E. coli* C1845 was isolated from a child with persistent diarrhea. The strain

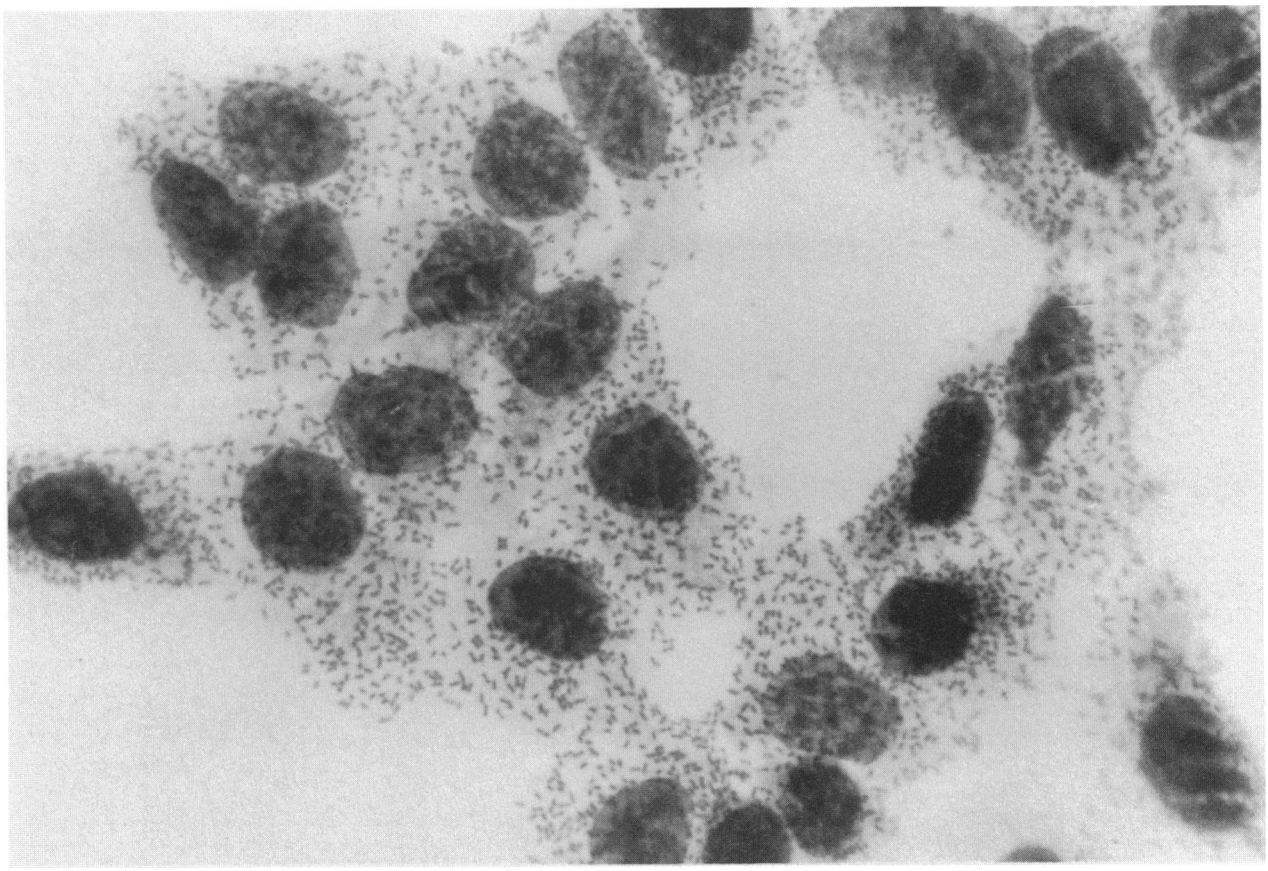

FIG. 1. Diffuse HEp-2 cell adherence exhibited by strain C1845.

adhered to HEp-2 and HL cells in a mannose-resistant, diffuse manner (Fig. 1) and was MRHA positive. MRHA was mediated by cells grown on solid and liquid media at 20, 30, 37, and 42°C. MRHA was not affected by prior treatment of erythrocytes with neuraminidase under conditions which abolished colonization factor antigen I-mediated hemagglutination (21). C1845 did not agglutinate latex particles coated with the P blood group antigen. MRHA was not sensitive to 0.05 M chloramphenicol. Chloramphenicol sensitivity has been described for O75X-mediated hemagglutination (34).

Molecular cloning of genes that encode the adhesin of strain C1845. A cosmid library of DNA from strain C1845 was used to infect *E. coli* LE392. Ampicillin-resistant transductants were tested for MRHA by slide agglutination at room temperature. Three such colonies of 950 tested were MRHA positive. One of these was selected for further study, and its recombinant plasmid was designated pSLM850. This strain was tested for DA and demonstrated a pattern of HL cell adherence indistinguishable from that of strain C1845 (data not shown). LE392 did not adhere.

Identification of sequences encoding MRHA expression. Preliminary transposon mutagenesis studies indicated that all Tn*5* insertions which abolished MRHA mapped within an 8.5-kilobase *Bam*HI-*Hin*dIII fragment of pSLM850. This fragment was cloned into pUC8. The resulting plasmid, designated pSSS1, conferred an MRHA-positive, DA-positive phenotype. Electron microscopic examination of *E. coli* HB101 transformed with pSSS1 demonstrated that the plasmid also encoded production of fimbriae (Fig. 2).

Tn*5* insertion derivatives of pSSS1 were screened for MRHA. The locations of 11 Tn*5* insertions which abolished the MRHA phenotype and 8 randomly selected insertions which did not affect the MRHA phenotype were determined by restriction endonuclease analysis. Most Tn*5* insertions between *Pst*I-1 and *Pst*I-8 abolished the MRHA phenotype (Fig. 3); however, two Tn*5* insertions (Tn*5*-2 and Tn*5*-49) were positive for MRHA expression even though the surrounding inserts abolished this property.

Subclones of pSSS1 were constructed, some of which were obtained by using the *Xho*I and *Hin*dIII sites provided by the Tn*5* insertions in pSSS1::Tn*5*-28 and pSSS1::Tn*5*-49 (Fig. 3). A pSSS1::Tn*5*-28 derivative, designated pSSS1.2, contains a 4.7-kilobase fragment of pSSS1 which retains the ability to mediate the MRHA-positive and HEp-2 cell adhesion phenotypes.

Defining the region necessary for expression of the F1845 subunit. The fimbriae of strains C1845 and HB101(pSSS1) were purified. The major constituent of the fimbriae was shown by SDS-polyacrylamide gel electrophoresis to be a protein of 14.3 kilodaltons (kDa) (Fig. 4). To define the region which encodes the 14.3-kDa F1845 subunit, strains carrying pSSS1 and pSSS1::Tn*5* derivatives which abolished MRHA were examined for F1845 antigen production by immunoblot analysis using antiserum raised against purified fimbriae from strain C1845. pSSS2 and all pSSS1::Tn*5* derivatives, except pSSS1::Tn*5*-12 and pSSS1::Tn*5*-16 (Fig. 3), expressed the 14.3-kDa subunit (Fig. 5). Plasmid pSSS1::Tn*5*-25 produced a very small amount of F1845 antigen (not visible in Fig. 5, lane 7). Overloading of the gel with proteins from strain DH5α (pSSS1::Tn*5*-25) resulted in immunoblots in which a band in the position of the F1845 subunit was detected by anti-F1845 sera (data not shown), while identical

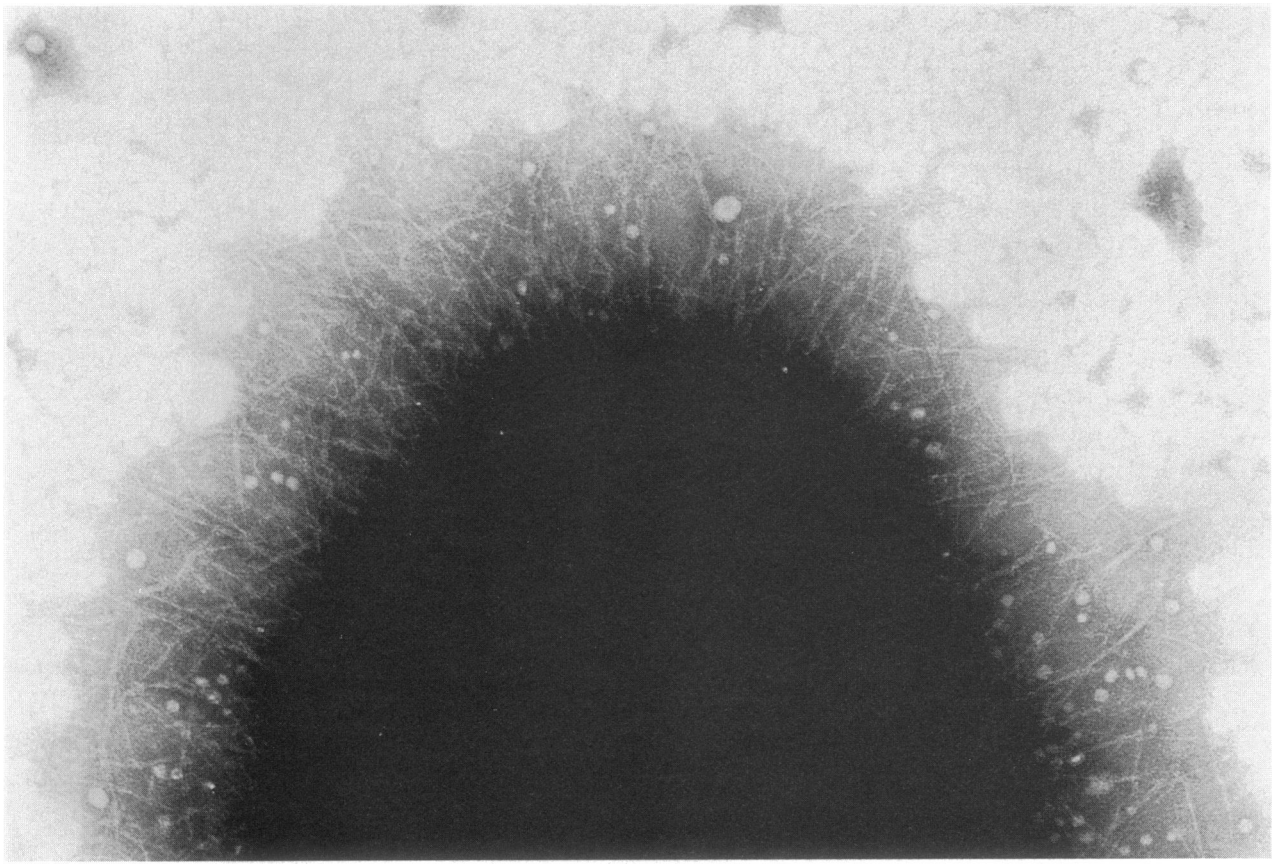

FIG. 2. Electron micrograph of a negatively stained preparation of HB101(pSSS1).

treatment of strains carrying pSSS1::Tn*5*-12 and pSSS1::Tn*5*-16 failed to demonstrate production of the 14.3-kDa protein. We therefore concluded that the F1845 subunit is encoded within the region containing Tn*5*-12 and Tn*5*-16.

Nucleotide sequence of F1845 subunit gene. The region concluded to encode the F1845 subunit was subjected to nucleotide sequence analysis (Fig. 6). A 477-base open reading frame was identified which is preceded by sequences encoding a potential ribosome-binding site (43) (Fig. 7). The open reading frame encodes a polypeptide of 159 amino acids. The deduced amino-terminal amino acid sequence suggested the presence of a signal sequence for transmembrane secretion of the protein. A signal peptide cleavage site was predicted by the method of von Heijne (48) between Ala at position 21 and Thr at position 22, resulting in a mature polypeptide of 138 amino acids (M_r, 14,599).

A nucleotide sequence comparison of the region encompassing the F1845 structural subunit gene with nucleotide sequences reported for other *E. coli* adhesin-associated genes revealed extensive homology to *afaE*, the gene encoding the hemagglutinin of the afimbrial adhesin AFA-I from a uropathogenic *E. coli* strain (18). The sequences were almost identical from the beginning of the reported *afaE* sequence through the first six codons of the coding region of the signal sequence (Fig. 7). The remainder of the *afaE* coding sequence showed a similarity of 61% with the F1845 structural gene (data not shown). Comparison of the predicted amino acid sequences of the mature AFA-I and F1845 proteins yielded the alignment shown in Fig. 8. The region of greatest similarity includes the first 24 amino acids. A comparison of published *Pst*I and *Sma*I restriction site locations in the *afa* operon with the map of the F1845 determinant revealed several common sites, indicating that the nucleotide sequence relatedness between the *afa* operon and the F1845 determinant extends well upstream of the structural subunit genes.

Organization of genes involved in expression of F1845. We analyzed the gene products associated with MRHA expression by labeling the proteins encoded by pSSS1 and its derivatives in *E. coli* maxicells. Plasmid pSSS1 directed the synthesis of five polypeptides (Fig. 9, lane 2) in addition to the vector-encoded β-lactamase. The apparent sizes of these polypeptides were 95, 27, 15.5, 14.3, and 10 kDa, and the genes encoding them were designated *daaC*, *daaB*, *daaD*, *daaE*, and *daaA*, respectively. The band with an apparent size of 6 kDa was presumed to be a degradative product because it is not expressed by any other subclone of pSSS1.

Plasmid pSSS1.2, an MRHA-positive derivative of pSSS1, directs expression of all of the proteins encoded by pSSS1, except the 10-kDa product of *daaA* (Fig. 9, lane 5), indicating that the *daaA* product is not required for expression of the MRHA-positive phenotype.

Analysis of the products expressed by plasmids pSLM854, pSLM856, and pDAS105 (Fig. 3 and 6), MRHA-negative derivatives of pSSS1, further revealed the order of the genes encoding them. Plasmid pSLM856 expressed the 10- and 27-kDa polypeptides observed with pSSS1 and two polypeptides of approximately 38 kDa (Fig. 9, lane 3). These polypeptides (bands c′) were concluded to be products of a truncated *daaC*. This result indicated that *daaC* is tran-

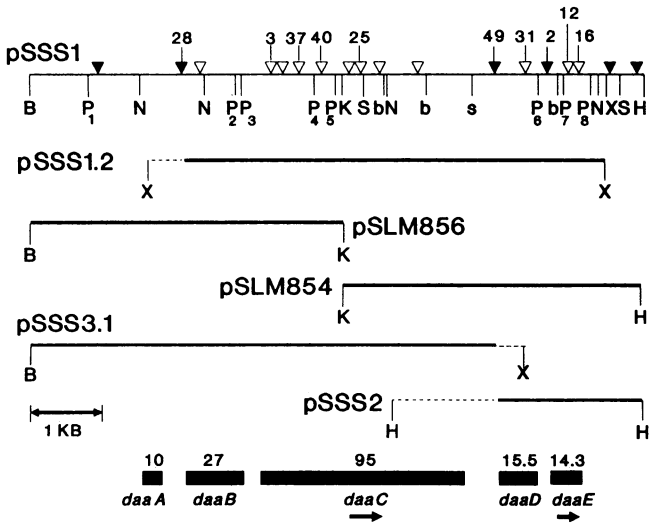

FIG. 3. Genetic organization of F1845 determinants. Sites of transposon insertions in pSSS1 are indicated by triangles as follows: open triangles, Tn*5* insertions which abolish the MRHA phenotype; solid triangles, Tn*5* insertions which retain the MRHA phenotype. Numbers above the triangles identify specific inserts referred to in the text. Restriction maps of plasmid pSSS1 and its derivatives are indicated by solid lines. Broken lines indicate Tn*5* sequences. B, *Bam*HI; b, *Bst*EII; H, *Hin*dIII; K, *Kpn*I; N, *Nde*I; P, *Pst*I; S, *Sal*I; s, *Sst*I; X, *Xho*I. *Pst*I sites are numbered for reference in the text. Plasmids pSSS1.2, pSLM856, pSLM854, and pSSS2 are subclones of pSSS1 in pUC vectors. pSSS3.1 is a subclone of pSSS1 in pACYC184. Vector sequences are not shown. Peptides encoded by pSSS1 are indicated by boxes below the regions where they are encoded. The location of *daaA* is approximate. The numbers on the boxes indicate the apparent molecular masses of the mature polypeptides in kilodaltons. The arrows indicate the direction of transcription. KB, Kilobase.

scribed from left to right (Fig. 3). Two bands appear in the same approximate position as bands c′ in Fig. 9, lanes 2, 5, and 6. These are most clearly seen in lane 5. Careful examination of additional gels revealed that these bands were distinct from bands c′ (data not shown). We further concluded that *daaA* was located upstream of *daaB*, since the 10-kDa polypeptide was not expressed by pSSS1.2 (Fig. 9, lane 5). Plasmid pSLM854 expressed a protein of 15.5 kDa (the *daaD* product) and the 14.3-kDa structural subunit (the *daaE* product). Plasmid pSSS2 also expressed the *daaD* and *daaE* products (data not shown). Plasmid pDAS105 expressed only the 14.3-kDa polypeptide (Fig. 9, lane 6). This plasmid was constructed with the *lacZ* promoter of the vector in the opposite orientation with regard to the insert for expression to be mediated by the vector promoter. This suggests the presence of a promoter immediately upstream of the *daaE* coding region. The location of *daaD* was concluded to be between *daaC* and *daaE* on the basis of the previously determined location of *daaE* and the fact that the *daaD* product was produced by pSSS2.

The *daaD* product is not essential for F1845 expression. Neither pSSS3.1, a subclone of pSSS1::Tn*5*-49 in pACYC 184, nor pDAS105 (Fig. 3 and 6) directed functional F1845 adhesin production as determined by MRHA assays. However, *E. coli* DH5α harboring both of these plasmids was MRHA positive. Since neither plasmid contains *daaD*, the data indicate that this gene is not essential for adhesin production.

DNA hybridization studies. Southern blot hybridizations of plasmid and total DNAs from strain C1845, using the *daaC* probe (Fig. 10), indicated that the F1845 determinant is of chromosomal origin, since plasmid DNA did not hybridize to the probe (data not shown). Twenty-six additional *E. coli* isolates possessing DNA homology to the *daaC* probe were

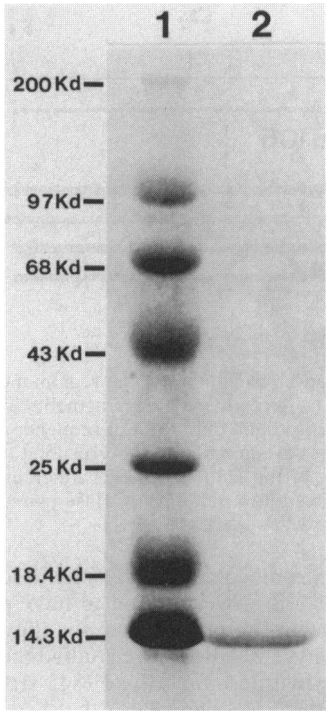

FIG. 4. Coomassie-stained polyacrylamide-SDS gel of F1845 purified from HB101(pSSS1). Lanes: 1, size standards; 2, F1845. Kd, Kilodaltons.

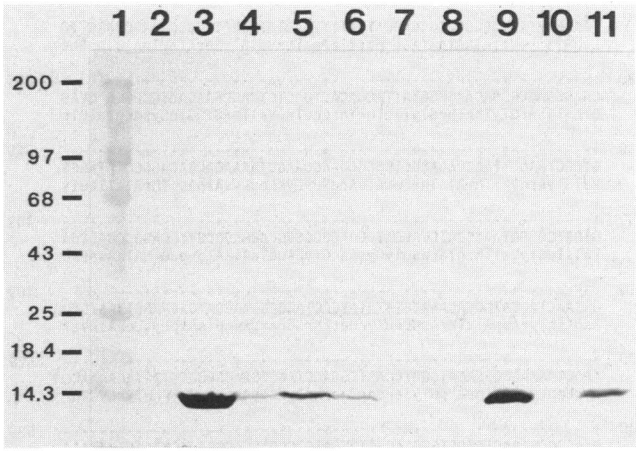

FIG. 5. Immunoblot analysis of proteins encoded by *E. coli* harboring pSSS1 subclones and Tn*5* insertion derivatives. F1845 was detected with rabbit antiserum raised against purified pili of *E. coli* C1845. Equivalent amounts of total bacterial protein were loaded in each lane, except for lane 3, which received 1/10 of the amount of protein in lanes 2 and 4 through 11. This was necessary for the sake of clarity because of the large amount of F1845 produced by pSSS1. Lanes: 1, size standards (indicated on the left in kilodaltons); 2, pUC8; 3, pSSS1; 4, pSSS1::Tn*5*-3; 5, pSSS1::Tn*5*-37; 6, pSSS1::Tn*5*-40; 7, pSSS1::Tn*5*-25; 8, pSSS1::Tn*5*-12; 9, pSSS1::Tn*5*-31; 10, pSSS1::Tn*5*-16; 11, pSSS2.

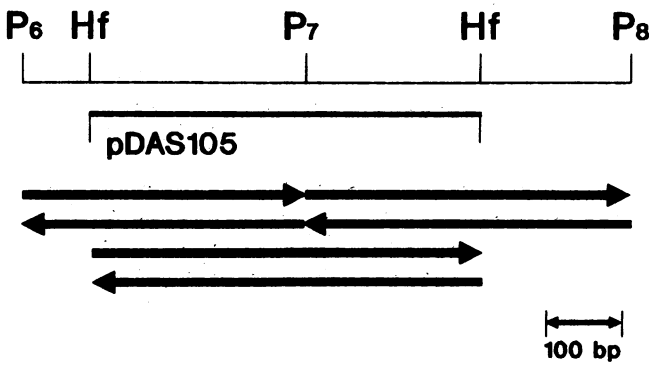

FIG. 6. Detailed physical map of the region encoding the F1845 structural subunit. Arrows indicate fragments cloned into M13 phages mp18 and mp19 and directions of sequence analysis. P, *Pst*I; Hf, *Hin*fI. *Pst*I sites are numbered to correspond to Fig. 3. Plasmid pDAS105 consists of the *Hin*fI fragment, which encodes the entire *daaE* product, cloned into pUC18. bp, Base pairs.

examined for chromosomal or plasmid-associated homology. Ten strains (38%) were found to have plasmid-associated homology with the probe (data not shown).

Using colony hybridizations, we studied the occurrence of DNA sequences within *daaE* (the F1845 structural subunit gene) among clinical *E. coli* isolates that had been shown to share DNA homology with the *daaC* probe. We examined

```
      59  -+---------+---------+---------+---------+---------+---------  109
afaE      TGTGAACGGTGGTTAATGTGGGGTAAGACAGCTTACTGATTCTGGGATGAA
          |||||||| |||||||||||||||||||||||||||||||||||||||||||
daaE      TGTGAACGTTGGTTAATGTGGGGTAAGACAGCTTACTGATTCTGGGATGAA

110 +---------+---------+---------+---------+---------+---------  169
    TTAGACCGTACTGTTGTGTTACCCCCTCACAAAACTGAATAGGTAATCCATATGAAAAAA
    ||||||||||||||||||||||||||||||||||||||| ||||||||| ||||||||||
    TTAGACCGTACTGTTGTGTTACCCCCTCACAAAACTGAACAGGTAATCAATATGAAAAAA
                                                MetLysLys

170 +---------+---------+---------+---------+---------+---------  229
    TTAGCGAT
    ||||||||
    TTAGCGATAATGGCCGCCGCCAGCATGATTTTCACTGTGGGCTCCGCGCAAGCCACGTTC
    LeuAlaIleMetAlaAlaAlaSerMetIlePheThrValGlySerAlaGlnAlaThrPhe

230 +---------+---------+---------+---------+---------+---------  289
    CAAGCGAGCGGGACAACGGGTATTACCACACTGACTGTGACCGAAGAGTGCCGGGTGCAG
    GlnAlaSerGlyThrThrGlyIleThrThrLeuThrValThrGluGluCysArgValGln

290 +---------+---------+---------+---------+---------+---------  349
    GTCGGTAACGTTACCGCGACGCTCGCGCGAAGCAAGCTAAAAGACGACACGGCTATAGGG
    ValGlyAsnValThrAlaThrLeuAlaArgSerLysLeuLysAspAspThrAlaIleGly

350 +---------+---------+---------+---------+---------+---------  409
    GTGATCGGGGTAACCGCTCTAGGATGCAACGGCCTGCAGGCCGCGTTGCAGGCCGACCCA
    ValIleGlyValThrAlaLeuGlyCysAsnGlyLeuGlnAlaAlaLeuGlnAlaAspPro

410 +---------+---------+---------+---------+---------+---------  469
    GACAACTACGATGCGACAAATCTATATATGACTAGCCGCAACCATGACAAACTTAATGTC
    AspAsnTyrAspAlaThrAsnLeuTyrMetThrSerArgAsnHisAspLysLeuAsnVal

470 +---------+---------+---------+---------+---------+---------  529
    AAACTAAAAGCCACAGATGGGTCCAGTTGGACCTACGGGAACGGCGTGTTCTACAAAACC
    LysLeuLysAlaThrAspGlySerSerTrpThrTyrGlyAsnGlyValPheTyrLysThr

530 +---------+---------+---------+---------+---------+---------  589
    GAAGGCGGTAACTGGGGTGGGCACGTTGGGATCTCCGTGGACGGGAACCAAACTGATAAA
    GluGlyGlyAsnTrpGlyGlyHisValGlyIleSerValAspGlyAsnGlnThrAspLys

590 +---------+---------+---------+---------+---------+          640
    CCGACCGGTGAATACACACTGAACCTCACCGGGGGTTACTGGACGAACTAA
    ProThrGlyGluTyrThrLeuAsnLeuThrGlyGlyTyrTrpThrAsn
```

FIG. 7. Nucleotide sequence of *daaE* and its deduced amino acid sequence. Numbering of the sequence begins from an upstream *Pst*I site. The homologous region of the *afaE* sequence (18) is shown directly above the *daaE* sequence. Identical bases are indicated with a vertical bar. Underlined nucleotides encode a potential ribosome-binding site. Underlined amino acids constitute the predicted signal sequence.

```
F1845   1  TFQASGTTGITTLTVTEECRVQVGNVTATLARSKLKDDTAIGVIGVTALG  50
           :  ::: :  ::.::::::: :.    . :: : .  : .:.  :
AFA-I   1  NFTSSGTNGKVDLTITEECRVTVESKSESFLRSGLVANRHITNLGIQSTG  50

       51  CN.GLQAALQADPDNYDATNLYMTSRNH..DKLNVKLKATDGSSWTYGNG  97
           : :  :: ... :: ::    ...  ::: : .   : .  :
       51  CGTGQRVALKLGAGSYDDTNGAHMTHENGTDKLLVSMGSATGDGTQDGGV 100

       98  VFYKTEGGNWGGHVGISVDGNQTDKPTGEYTLNLTGGYWTN           138
           .  .:  :  .     : : : .
      101  YYINRDGTGTGRWCSSYEMTNSTYQPASTP                       130
```

FIG. 8. Optimal alignment of the predicted mature F1845 and AFA-1 proteins. Symbols: :, identical amino acids; ·, related amino acids. Alignment and amino acid relatedness were determined by the BestFit program of the University of Wisconsin Genetics Computer Group software package (10).

615 *E. coli* isolates from healthy children and persons with diarrhea for nucleotide sequence homology to the *daaC* probe. Isolates from 45 persons were detected. Two probes from the *daaE* region (Fig. 10) were used to probe these isolates. Thirty six isolates (80%) hybridized to the *daaE* 5′ probe, and five isolates (11%) hybridized to both the *daaE* 5′ probe and the *daaE* 3′ probe. Nine isolates (20%) did not hybridize with either *daaE* probe. Some of the strains scored as positive in the hybridization assays were weak, indicating that homology with the probe was not complete.

DISCUSSION

We have characterized the genetic determinant of diffuse adhesin production in *E. coli* C1845, a strain isolated from a child with protracted diarrhea. The genetic organization and

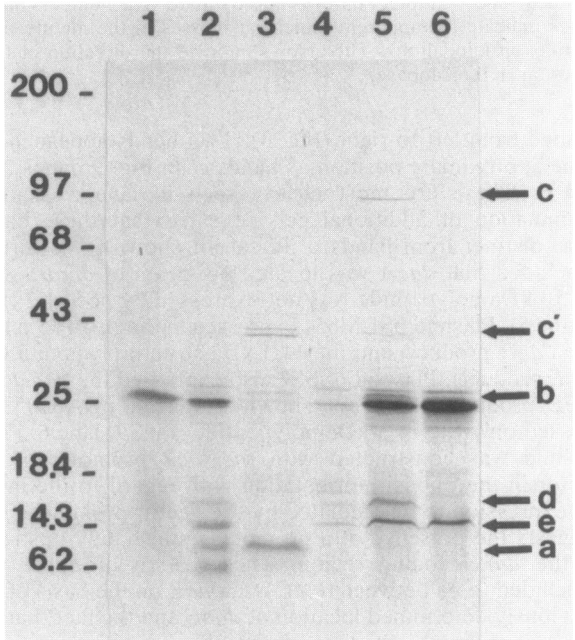

FIG. 9. Polypeptide synthesis directed by pSSS1 and its derivatives in *E. coli* maxicells. Maxicells were labeled with [^{35}S]methionine, solubilized, and analyzed on a polyacrylamide-SDS gel. An autoradiograph of the gel is shown. The molecular masses of the size standards are marked in kilodaltons. Lanes: 1, pUC8; 2, pSSS1; 3, pSLM856; 4, pSLM854; 5, pSSS1.2; 6, pDAS105. The polypeptides encoded by pSSS1 consisted of DaaA, DaaB, DaaC, DaaD, and DaaE, indicated by arrows a through e. Arrow c′ indicates the truncated forms of the *DaaC* polypeptide product of pSLM856 (lane 3).

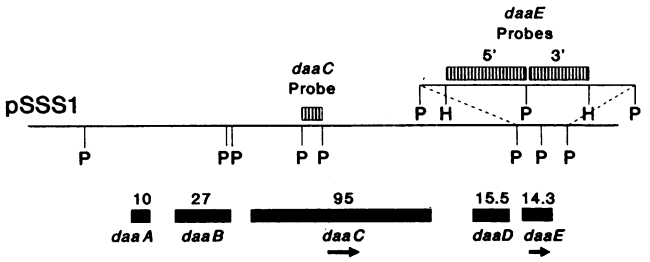

FIG. 10. Hybridization probes derived from pSSS1. P, *Pst*I; H, *Hin*fI. Blocks indicate the locations of the designated genes.

the number and sizes of proteins associated with adhesin synthesis are consistent with findings on several other *E. coli* fimbrial adhesins (37). The fimbrial morphology is similar to that exhibited by the K88 fimbrial antigen (37), and the morphology and structural subunit size appear to be similar to those of adhesins from diarrheal *E. coli* isolates described by Orskov et al. (36) and Hinson et al. (13). Restriction analysis, nucleotide sequence analysis, and analysis of the protein products of the F1845 determinant indicate that it is a member of the *afa* family of *E. coli* adhesin determinants described by Labigne-Roussel et al. (17, 18).

Protein products of five genes we have defined as *daaA* through *daaE* are associated with F1845 adhesin production. Labigne-Roussel et al. observed that *afaA* was not necessary for AFA-I production and speculated that product of this gene may have a regulatory function (18). Similarly, *daaA* is not required for production of the F1845 adhesin as demonstrated by the ability of plasmids lacking this gene to mediate MRHA and DA. We believe that the *daaA* gene product may also have a regulatory function, since plasmids lacking this gene direct the production of reduced amounts of the *daaE* product (unpublished data). Labigne-Roussel et al. have demonstrated that the *afaD* product is not required for AFA-I production (18). Consistent with this observation, we have demonstrated through complementation experiments that *daaD* is not essential for production of a functional F1845 adhesin.

While we have thus found several similarities between the F1845 and AFA-I determinants, the fimbrial structure of F1845 and nucleotide sequence comparisons of *daaE* and *afaE* demonstrate that F1845 is distinct from AFA-I. The sequence divergence that we have observed which begins in the coding regions of the signal peptides of F1845 and AFA-I is consistent with the previous findings of Labigne-Roussel and Falkow that heterogeneity in *afaE* exists among *E. coli* isolates harboring DNA homologous to the other *afa* genes (17). Comparison of the predicted amino acid sequences of the mature F1845 and AFA-I proteins revealed a greater degree of similarity between the amino-terminal portions of the proteins, suggesting that this region is involved in interactions with other components of the adhesin-synthesizing complex. Our survey of *E. coli* isolates with the *daaE* 5′ and 3′ probes indicates greater conservation of the 5′ *daaE* sequences, which is consistent with a hypothesis of greater structural constraints on the amino-terminal portion of the protein.

Another possible member of the *afa* family of *E. coli* adhesin determinants is the O75X adhesin determinant of uropathogenic *E. coli* (46), also termed the Dr hemagglutinin (35). Although the genetic organization of the Dr hemagglutinin appears to differ somewhat, a comparison of published restriction sites for the cloned Dr hemagglutinin determinant revealed similarities to the F1845 determinant (33, 35). Again, F1845 is likely to be distinct from the Dr hemagglutinin, since adherences mediated by the two adhesins differ with regard to chloramphenicol sensitivity (34). Our preliminary data suggest that the Dr hemagglutinin and F1845 determinants are highly homologous upstream of the structural genes and that the sequences diverge in the structural genes *daaE* and *draA* (T. Swanson et al., manuscript in preparation). Complementation experiments indicate that MRHA sensitivity or resistance to chloramphenicol depends on whether *daaE* is expressed, suggesting that the structural subunits of these fimbrial adhesins determine receptor specificity (T. Swanson et al., manuscript in preparation).

Sequence divergence beginning in the signal peptide-encoding regions of structural genes of related adhesin determinants has also been demonstrated for the K88-F41 family (1). Another similarity between the *afa* and K88 adhesin families is the existence of members of each of these families of genetic determinants on plasmids and chromosomes. Nataro et al. have described a plasmid-mediated diffuse adhesin in an *E. coli* strain isolated from a case of diarrheal disease (31), and our findings demonstrate that *daa*-related sequences exist on plasmid or chromosomal DNA. Similarly, K88-encoding sequences are found on plasmids, while the related determinant of F41 production is chromosomal (27). These findings suggest the development of a common mechanism in *E. coli* for the generation of antigenic diversity in structures that potentially mediate essential virulence properties, such as adherence to mucosal surfaces. Further study of the mechanisms of divergence within adhesin families will be important for the design of vaccine strategies directed against these structures.

Our results do not permit conclusions regarding the transcriptional organization of the F1845 determinant. Labigne-Roussel et al. found sequences upstream of *afaE* characteristic of an *E. coli* promoter (18). These sequence are identical to sequences upstream of the *daaE* gene. Our transposon insertion data are consistent with the possibility of a promoter directly upstream of *daaE*, as is the ability of pDAS105 to direct synthesis of the *daaE* product. Our transposon insertion data also suggest the possibility of independent promoters for *daaB* and *daaD*, since inserts presumably upstream of these genes did not prevent their expression. However, Tn*5* insertions are not always polar (5). Transcription downstream of Tn*5* may initiate from within the transposon, or the insertion may fortuitously create sequences which may function as promoters. Definition of the transcriptional organization of the F1845 determinant will therefore depend on further experiments to directly determine the location of transcript initiation.

The role of F1845 in diarrheal disease remains to be determined. Strain C1845 appears to be pathogenic in an animal model, attaching to and causing effacement of cecal and colonic mucosal epithelia in piglets (H. W. Moon, personal communication). An assessment of the role of the F1845 adhesin in this virulence model will involve examination of C1845 derivatives deficient in F1845 production. Construction of such derivatives is in progress.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI23771 from the National Institute of Allergy and Infectious Diseases.

We thank Dalia Alfi and Myron Rabin for technical assistance. We are grateful to Jimmie Lara for assistance with the electron microscopy and to Stephen Lory for helpful discussions.

LITERATURE CITED

1. **Anderson, D. G., and S. L. Moseley.** 1988. *Escherichia coli* F41 adhesin: genetic organization, nucleotide sequence, and homology with the K88 determinant. J. Bacteriol. **170**:4890–4896.
2. **Bachmann, B. J.** 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol. Rev. **36**:525–557.
3. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
4. **Bolivar, F., and K. Backman.** 1979. Plasmids of *Escherichia coli* as cloning vectors. Methods Enzymol. **68**:245–267.
5. **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.
6. **Cavallaro, J. J., and A. S. Monto.** 1972. HL cells, a sensitive line for the isolation and propagation of respiratory syncytial virus. Proc. Soc. Exp. Biol. Med. **40**:507–510.
7. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134**: 1141–1156.
8. **Clausen, C. R., and D. L. Christie.** 1982. Chronic diarrhea in infants caused by adherent enteropathogenic *Escherichia coli*. J. Pediatr. **100**:358–361.
9. **de Graaf, F. K., and I. Roorda.** 1982. Production, purification, and characterization of the fimbrial adhesive antigen F41 isolated from calf enteropathogenic *Escherichia coli* strain B41M. Infect. Immun. **36**:751–758.
10. **Devereaux, J., P. Haeberli, and O. Smithies.** 1983. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:382–395.
11. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
12. **Hawkes, R., E. Niday, and J. Gordon.** 1982. A dot-immunobinding assay for monoclonal and other antibodies. Anal. Biochem. **119**:142–147.
13. **Hinson, G., S. Knutton, M. K.-L. Lam-Po-Tang, A. S. McNeish, and P. H. Williams.** 1987. Adherence to human colonocytes of an *Escherichia coli* strain isolated from severe infantile enteritis: molecular and ultrastructural studies of a fibrillar adhesin. Infect. Immun. **55**:393–402.
14. **Hohn, B., and J. Collins.** 1980. A small cosmid for efficient cloning of large DNA fragments. Gene **11**:291–298.
15. **Hull, R. A., R. E. Gill, P. Hsu, B. H. Minshew, and S. Falkow.** 1981. 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**:933–938.
16. **Kado, C. I., and S.-T. Liu.** 1981. Rapid procedure for detection and isolation of large and small plasmids. J. Bacteriol. **145**: 1365–1373.
17. **Labigne-Roussel, A., and S. Falkow.** 1988. Distribution and degree of heterogeneity of the afimbrial-adhesin-encoding operon (*afa*) among uropathogenic *Escherichia coli* isolates. Infect. Immun. **56**:640–648.
18. **Labigne-Roussel, A., M. A. Schmidt, W. Walz, and S. Falkow.** 1985. Genetic organization of the afimbrial adhesin operon and nucleotide sequence from a uropathogenic *Escherichia coli* gene encoding an afimbrial adhesin. J. Bacteriol. **162**:1285–1292.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
20. **Levine, M. M., V. Prado, R. Robins-Browne, J. Lior, J. B. Kaper, S. L. Moseley, K. Gicquelais, J. P. Nataro, P. Vial, and B. Tall.** 1988. Use of DNA probes and HEp-2 cell adherence assay to detect diarrheagenic *Escherichia coli*. J. Infect. Dis. **158**: 224–228.
21. **Lindahl, M., A. Faris, and T. Wadstrom.** 1982. Colonization factor antigen enterotoxigenic *Escherichia coli* is a sialic-specific lectin. Lancet **ii**:280.
22. **Maas, R.** 1983. An improved colony hybridization method with significantly increased sensitivity for detection of single genes. Plasmid **10**:296–298.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Maniatis, T., A. Jeffrey, and A. G. Kleid.** 1975. Nucleotide sequence of the rightward operator of phage lambda. Proc. Natl. Acad. Sci. USA **72**:1184–1188.
25. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
26. **Miller, J.** 1972. Experiments in molecular genetics, p. 433. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **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* that lack K88, K99, and 987P fimbriae. J. Gen. Microbiol. **129**:2753–2759.
28. **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**:444–446.
29. **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* H-10407. J. Bacteriol. **156**:441–443.
30. **Nataro, J. P., J. B. Kaper, R. Robbins-Browne, V. Prado, P. A. Vial, and M. M. Levine.** 1987. Patterns of adherence of diarrheogenic *Escherichia coli* to HEp-2 cells. J. Pediatr. Infect. Dis. **6**:829–831.
31. **Nataro, J. P., I. C. A. Scaletsky, J. B. Kaper, M. M. Levine, and L. R. Trabulsi.** 1985. Plasmid-mediated factors conferring diffuse and localized adherence of enteropathogenic *Escherichia coli*. Infect. Immun. **48**:378–383.
32. **Norrander, J., T. Kempe, and J. Messing.** 1982. Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene **26**:101–106.
33. **Nowicki, B., J. P. Barrish, T. Korhonen, R. A. Hull, and S. I. Hull.** 1987. Molecular cloning of the *Escherichia coli* O75X adhesin. Infect. Immun. **55**:3168–3173.
34. **Nowicki, B., J. Moulds, R. Hull, and S. Hull.** 1988. A hemagglutinin of uropathogenic *Escherichia coli* recognizes the Dr blood group antigen. Infect. Immun. **56**:1057–1060.
35. **Nowicki, B., C. Svanborg-Edén, R. Hull, and S. Hull.** 1989. Molecular analysis and epidemiology of the Dr hemagglutinin of uropathogenic *Escherichia coli*. Infect. Immun. **57**:446–451.
36. **Orskov, I., A. Birch-Andersen, J. P. Duguid, J. Stenderup, and F. Orskov.** 1985. An adhesive protein capsule of *Escherichia coli*. Infect. Immun. **47**:191–200.
37. **Oudega, B., and F. K. De Graaf.** 1988. Genetic organization and biogenesis of adhesive fimbriae of *Escherichia coli*. Antonie van Leeuwenhoek J. Microbiol. **54**:285–299.
38. **Rupp, W., C. Wilde, and D. Reno.** 1971. Exchanges between DNA strands in ultraviolet-irradiated *Escherichia coli*. J. Mol. Biol. **61**:25–44.
39. **Sancar, A., A. Hack, and W. Rupp.** 1979. Simple method for identification of plasmid-encoded proteins. J. Bacteriol. **137**: 692–693.
40. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
41. **Scaletsky, I. C. A., M. L. Silva, and L. R. Trabulsi.** 1984. Distinctive patterns of adherence of enteropathogenic *Escherichia coli* to HeLa cells. Infect. Immun. **45**:534–536.
42. **Scaletsky, I. C. A., M. L. M. Silva, M. R. F. Toledo, B. R. Davis, P. A. Blake, and L. R. Trabulsi.** 1985. Correlation between adherence to HeLa cells and serogroups, serotypes, and bioserotypes of *Escherichia coli*. Infect. Immun. **49**:528–532.
43. **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.

44. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–515.
45. **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.
46. **Väisänen-Rhen, V.** 1984. Fimbria-like hemagglutinin of *Escherichia coli* O75 strains. Infect. Immun. **46:**401–407.
47. **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.
48. **von Heijne, G.** 1985. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **11:**4683–4690.