JOURNAL OF BACTERIOLOGY, Dec. 1994, p. 7601–7613
0021-9193/94/$04.00+0

Vol. 176, No. 24

Nucleotide Sequence of the Afimbrial-Adhesin-Encoding *afa-3* Gene Cluster and Its Translocation via Flanking IS*1* Insertion Sequences

MARIE-ISABELLE GARCIA, AGNÈS LABIGNE, AND CHANTAL LE BOUGUENEC*

Unité des Entérobactéries, Institut National de la Santé et de la Recherche Médicale U389, Institut Pasteur, 75724 Paris Cedex 15, France

Received 16 June 1994/Accepted 11 October 1994

The *afa* gene clusters encode afimbrial adhesins (AFAs) that are expressed by uropathogenic and diarrhea-associated *Escherichia coli* strains. The plasmid-borne *afa-3* gene cluster is responsible for the biosynthesis of the AFA-III adhesin that belongs to the Dr family of hemagglutinins. Reported in this work is the nucleotide sequence of the 9.2-kb insert of the recombinant plasmid pILL61, which contains the *afa-3* gene cluster cloned from a cystitis-associated *E. coli* strain (A30). The *afa-3* gene cluster was shown to contain six open reading frames, designated *afaA* to *afaF*. It was organized in two divergent transcriptional units. Five of the six Afa products showed marked homologies with proteins encoded by previously described adhesion systems that allowed us to attribute to each of them a putative function in the biogenesis of the AFA-III adhesin. AfaE was identified as the structural adhesin product, whereas AfaB and AfaC were recognized as periplasmic chaperone and outer membrane anchor proteins, respectively. The AfaA and AfaF products were shown to be homologous to the PapI-PapB transcriptional regulatory proteins. No function could be attributed to the AfaD product, the gene of which was previously shown to be dispensable for the synthesis of a functional adhesin. Upstream of the *afa-3* gene cluster, a 1.2-kb region was found to be 96% identical to the RepFIB sequence of one of the enterotoxigenic *E. coli* plasmids (P307), suggesting a common ancestor plasmid. This region contains an integrase-like gene (*int*). Sequence analysis revealed the presence of an IS*1* element between the *int* gene and the *afa-3* gene cluster. Two other IS*1* elements were detected and located in the vicinity of the *afa-3* gene cluster by hybridization experiments. The *afa-3* gene cluster was therefore found to be flanked by two IS*1* elements in direct orientation and two in opposite orientations. The *afa-3* gene cluster, flanked by two directly oriented IS*1* elements, was shown to translocate from a recombinant plasmid to the *E. coli* chromosome. This translocation event occurred via IS*1*-specific recombination mediated by a *recA*-independent mechanism.

Escherichia coli is a frequent cause of intestinal or extraintestinal infections in both humans and animals. Among the virulence properties expressed by this pathogen, adhesin production plays an essential role in the colonization of the mucosal epithelia. One class of adhesins that are not associated with fimbrial structures on the cell surface were designated AFAs, for afimbrial adhesins (25, 30). These adhesins are expressed by different pathotypes of *E. coli* strains, including uropathogenic strains that cause infections in adults and children (25) and enteropathogenic (3) and enteroadherent (30) strains associated with diarrhea in children. AFA proteins mediate mannose-resistant hemagglutination (MRHA) of human erythrocytes and specific attachment to uroepithelial cells (27) and to Hep-2, HeLa, and Caco-2 (unpublished data) cells with a so-called diffuse adherence pattern (44). The first *afa* gene cluster to be cloned (*afa-1*) was shown to be located on a 6.7-kb chromosomal fragment and to express polypeptides encoded by five genes, *afaA* to *afaE*, that are organized in an operon (27). Among the five Afa products, only AfaB, AfaC, and AfaE were found to be required for phenotypic expression of the AFAs (26). To date, a precise function could be attributed only to the AfaE polypeptide, which was recognized as the adhesin (26, 44).

We previously demonstrated that the *afa* gene clusters encode at least four antigenically distinct adhesins, which were designated AFA-I to AFA-IV (25). AFA-III is encoded by the *afa-3* gene cluster; it is an afimbrial adhesin that belongs to a family of hemagglutinins that recognize epitopes of the Dr antigen as receptors (34). Among the adhesins belonging to this family are the afimbrial adhesins AFA-I encoded by the *afa-1* operon (27) and the fimbrial adhesins Dr (35) and F1845 (8) encoded by the *dra* and *daa* operons, respectively. We previously reported (30) the cloning of the *afa-3* gene clusters from strains belonging to different pathotypes (associated with diarrhea or cystitis). All were found to be highly conserved and to be borne by a plasmid (100 kb in size), a situation that differed from that of the *afa-1* (30), *afa-2* (unpublished data), and *daa* (8) operons that were found to be chromosome associated.

The aim of this study was to investigate afimbrial adhesin AFA-III expression in detail. We report here on the determination of the nucleotide sequence of the 9.2-kb insert of pILL61 that harbors the *afa-3* gene cluster cloned from the 100-kb plasmid (pILL1055) of uropathogenic *E. coli* A30. This gene cluster was composed of six genes. The corresponding gene products, which potentially play a role in the biogenesis of the afimbrial adhesive structure, were identified. Other defined DNA sequences were identified in the close vicinity of the *afa-3* gene cluster, such as sequences that might be implicated in pILL1055 plasmid replication and/or stability, as well as several copies of IS*1* insertion sequences. Finally, we demonstrated the capacity of the *afa-3* gene cluster to translocate from plasmid to chromosome by an IS*1*-mediated recombination mechanism. This finding suggests that the *afa-3* adhesin-

* Corresponding author. Mailing address: Unité des Entérobactéries, Institut National de la Santé et de la Recherche Médicale U389, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France. Phone: 33 (1) 40 61 3280. Fax: 33 (1) 45 68 8837.

TABLE 1. Plasmids used in this study

Plasmid	Antibiotic resistance marker[a]	Vector	Insert			MRHA phenotype	Source or reference
			Size (kb)	Origin	*afa* sequence		
pILL61	Amp	pBR322	9.2	*Sau*3A partial digest of total DNA of A30	*afa-3*	+	30
pILL1111	Tc	pACYC184	6.2	*Nco*I-*Eco*RI fragment of pILL61	*afa-3*	+	This study
pILL1108	Sp	pILL570	2.6	*Pst*I fragment of pILL61	*afa-3*	−	This study
pILL1110	Amp	pBR322	2.8	*Eco*RI-*Sph*I fragment of pILL61	*afa-3*	−	This study
pILL1099	Sp	pILL570	11.9	*Sau*3A partial digest of pILL1055 from A30	*afa-3*	−	This study
pILL1101	Sp	pILL570	11.6	*Sau*3A partial digest of pILL1055 from A30	*afa-3*	+	This study
pILL1096	Amp	pACYC177	3.5	*Sma*I fragment of pILL1099	*afa-3*	−	This study
pILL1097	Amp	pACYC177	5.8	*Sma*I fragment of pILL1101	*afa-3*	−	This study
pILL1024	Sp	pILL570	7.8	*Sau*3A partial digest of pILL1043 from AL845	*afa-3*	+	This study
pILL1027	Sp	pILL570	13.1	*Sau*3A partial digest of pILL1043 from AL845	*afa-3*	+	This study
pILL04	Amp	pBR322	17	*Sau*3A partial digest of total DNA from KS52	*afa-1*	+	27
pILL1018	Sp	pILL570	11.6	*Sau*3A partial digest of pILL04	*afa-1*	+	This study
pILL1055	None	NA[b]	100	Plasmid from the wild-type strain, A30	*afa-3*	+	This study (30)

[a] Abbreviations: Amp, ampicillin resistance; Sp, spectinomycin resistance; Tc, tetracycline resistance.
[b] NA, not applicable.

encoding gene cluster might be able to disseminate within bacterial populations.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture conditions. *E. coli* A30 was isolated from an urine specimen from a patient with cystitis (30). The previously described *E. coli* strains KS52 (27) and AL845 (30) were isolated from a patient with pyelonephritis and from a child with diarrhea, respectively. *E. coli* HB101 (10) was used as a recipient for recombinant plasmid analyses, and *E. coli* JM101 (45) was used as a host for bacteriophage M13 derivatives in nucleotide sequence analyses. *E. coli* SE5000 (38) was used in maxicell experiments. Recombinant plasmids used in this study are listed in Table 1. Vectors pBR322 (9), pACYC177 (12), pACYC184 (12), pILL570 (24), and M13mp18 and M13mp19 (45) were used in cloning experiments. Plasmid pUCD800 (16), containing a kanamycin resistance gene as well as the *sacB* gene, served as a tool in translocation experiments. *E. coli* strains were grown in Luria broth) without glucose (10 g of tryptone, 5 g of yeast extract, 5 g of NaCl per liter [pH 7.0]) or on Luria agar plates (containing 1.5% agar) at 37°C. Antibiotic concentrations for the selection of transformants were as follows (in milligrams per liter): ampicillin, 100; tetracycline, 10; spectinomycin, 100; kanamycin, 20; and chloramphenicol, 20.

Hemagglutination and adherence assays. Hemagglutination of washed human erythrocytes, in the presence of 2% (wt/vol) α-methyl mannoside (MRHA phenotype), and adherence to HeLa cells were tested as previously described (3).

PCR. Amplification reactions were carried out with a Perkin-Elmer Cetus thermal cycler under the conditions previously described (29).

Preparation of DNA. Plasmids were routinely isolated by an alkaline lysis procedure (31), whereas whole-cell DNA was prepared as previously described (25).

DNA methodology. Restriction endonuclease digestions and other common DNA manipulations were performed by standard procedures according to the methods previously described (24, 31).

Hybridization experiments. For Southern blot hybridizations, DNA restriction fragments were separated by electrophoresis in 0.7 or 1% agarose gels and transferred to nitrocellulose sheets (0.45-μm pore size; Schleicher and Schuell, Inc.) by the Southern blot technique (40). The *afa* probe consisted of a 750-bp amplification product obtained with the primers specifying the *afa* operon (30). The IS*1* probe was prepared as follows. First, the 0.6-kb *Pst*I fragment of pILL61 (corresponding to almost the entire IS*1* element [see Fig. 2]) was subcloned into pBR322. Second, this fragment was amplified by using oligonucleotides that framed the *Pst*I cloning site of pBR322 as primers. Third, the amplified product was digested by *Pst*I, electroeluted, and purified on an Elutip-d-minicolumn (Schleicher and Schuell). Hybridizations were performed at 68°C with probes labeled with ^{32}P by using the Megaprime DNA labeling system (Amersham International) and were revealed by autoradiography with Amersham Hyperfilm-MP.

DNA sequencing. Restriction fragments were cloned into M13mp18 and M13mp19 or pBR322 vectors. Single-stranded DNA templates were prepared by the polyethylene glycol method (36), and the sequence was determined by the dideoxy-chain termination method with a Sequenase kit (U.S. Biochemical Corp.). Sequencing of double-stranded DNA was performed as previously described (13). When required, new oligonucleotide primers were synthesized on the basis of the acquired sequence. Sixty-seven percent of the pILL61 insert sequence, presented in Fig. 2, was determined for both strands of the DNA. Nucleotide or protein database searches were performed with FASTA, and sequence alignments were performed with GAP and PILEUP programs by using the Genetics Computer Group Sequence Analysis Software Package, version 7-UNIX. Sequence analyses were also performed with DNAid+, version 1.4 (F. Dardel and P. Bensoussan, Laboratoire de Biochimie, Ecole Polytechnique, Palaiseau, France).

Maxicell analyses, SDS-PAGE, and protein purification. Analysis of plasmid-encoded proteins was performed in the maxicell producing *E. coli* SE5000 (38). Preparation of maxicells, labeling of plasmid-encoded proteins with [^{35}S]methionine (Amersham), preparation of whole-cell extracts, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (final acrylamide concentration of 15% [wt/vol]), and visualization of polypeptides by fluorography were carried out as described previously (38). To prevent signal peptide processing, 9.5% ethanol was added to the assay medium. The molecular masses of the proteins were estimated by using SDS-PAGE standards ranging from 14,000 to 110,000 Da (Bio-Rad S.A., Ivry, France).

Nucleotide sequence accession number. The pILL61 insert nucleotide sequence accession number is X76688.

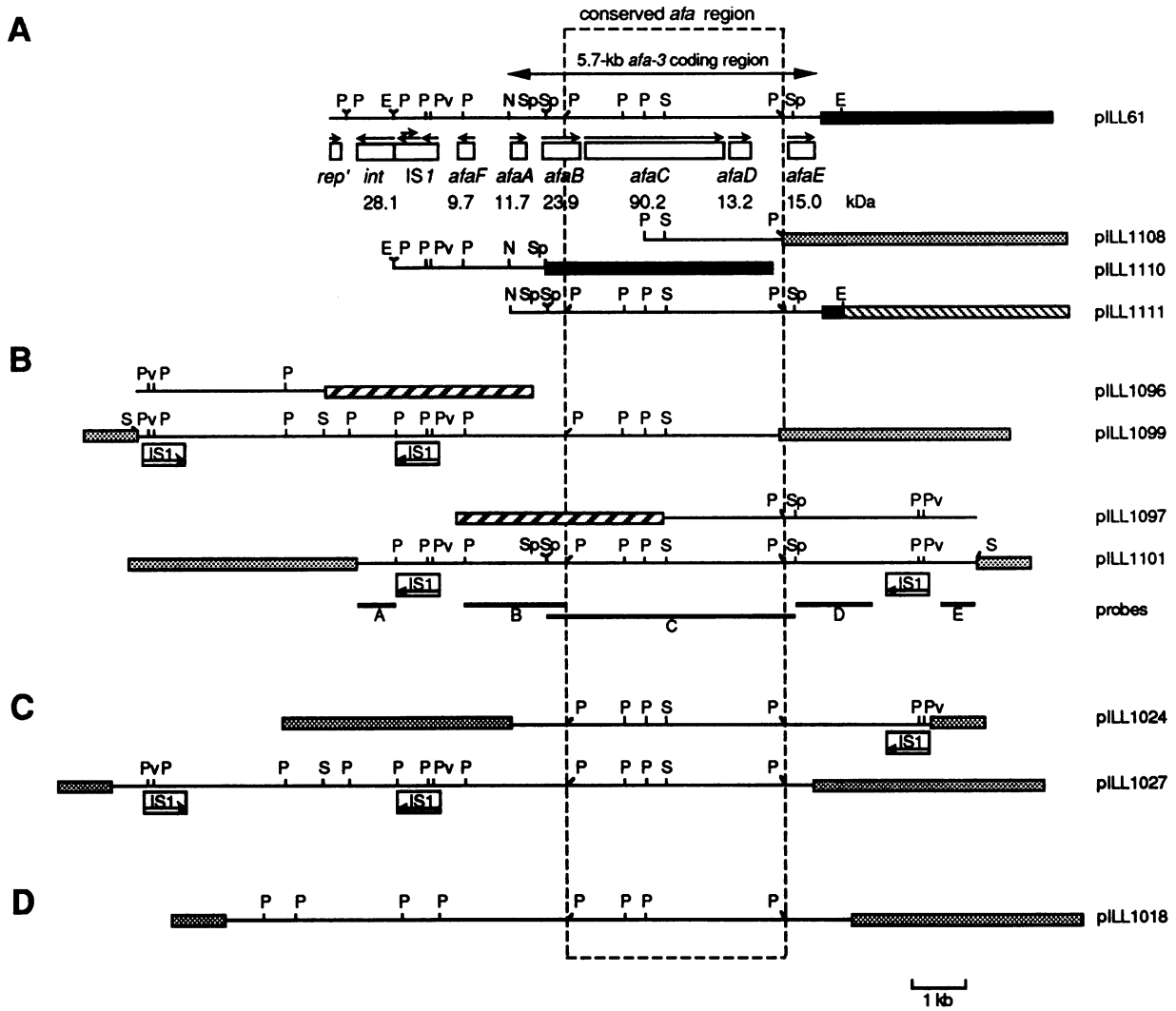

FIG. 1. Structure of the *afa-3* region. (A) Physical and genetic maps of the pILL61 and derivative plasmids pILL1108, pILL1110, and pILL1111. The double-headed arrow shows the position of the 5.7-kb *afa-3* coding region (30). The locations of the 11 genes described in this study are indicated by boxes as well as single-headed arrows that refer to the direction of gene transcription. Numbers correspond to the molecular masses (in kilodaltons) of the predicted mature polypeptides deduced from the nucleotide sequence shown in Fig. 2. (B) Restriction maps of derivative plasmids pILL1099, pILL1101, pILL1096, and pILL1097 resulting from the cloning of the *afa-3* gene cluster from the pILL1055 plasmid of *E. coli* A30. The IS*1* positions are indicated by boxes, and their orientation (i.e., from the left inverted repeat to the right inverted repeat) is depicted by arrows. Bars beneath the pILL1101 map refer to the following restriction fragments used as probes: A, 0.8-kb *Hin*dIII-*Eco*RI; B, 1.9-kb *Pst*I; C, 4.6-kb *Sph*I; D, 1.4-kb *Sph*I-*Eco*RV; E, 0.6-kb *Bam*HI-*Sma*I. (C) Restriction maps of the pILL1024 and pILL1027 plasmids resulting from the cloning of the *afa-3* gene cluster from strain AL845. (D) Restriction map of the recombinant plasmid pILL1018, carrying the *afa-1* operon cloned from *E. coli* KS52. Thin lines represent the inserts and thick lines represent pBR322 (9), pACYC184 (12), pILL570 (24), or pACYC177 (12) DNA. The conserved region of the *afa* operons is framed with dashed lines. Abbreviations: E, *Eco*RI; N, *Nco*I; P, *Pst*I; Pv, *Pvu*II; S, *Sma*I; Sp, *Sph*I.

RESULTS

Determination of the nucleotide sequence of the 9.2-kb insert of pILL61. The recombinant plasmid pILL61 (Table 1, Fig. 1A) was previously described as a plasmid expressing an AFA-III adhesin (30). The nucleotide sequence of the pILL61 insert is shown in Fig. 2. The sequence from nucleotide 8485 to nucleotide 9163 was previously published (30). The entire sequenced region was 9,163 bp long and had an overall G+C content of 53%, in agreement with the GC content of the *E. coli* genome. Computer analyses revealed 11 open reading frames (ORFs), illustrated in Fig. 1A with arrows. These ORFs were considered to be significant, because they either encoded peptides related to proteins expressed by the *afa-1* operon (26) or displayed homologies with previously reported nucleotide sequences.

Genetic organization of the *afa-3* gene cluster. Five ORFs, transcribed in the same orientation, were detected between nucleotide 3471 and nucleotide 9128 (Fig. 2). These mapped to the 5.7-kb *afa-3* coding region (Fig. 1A) that we previously characterized (30). The ORFs have been designated the *afaA*, *afaB*, *afaC*, *afaD*, and *afaE* genes.

The *afaA* gene started at either bp 3471 or bp 3519 and terminated at bp 3776. Because of the presence of a potential ribosome binding site (RBS) (42) 10 bp upstream of the first ATG (position 3471), the first ATG was assigned as the starting codon of the *afaA* gene. The AfaA product was found

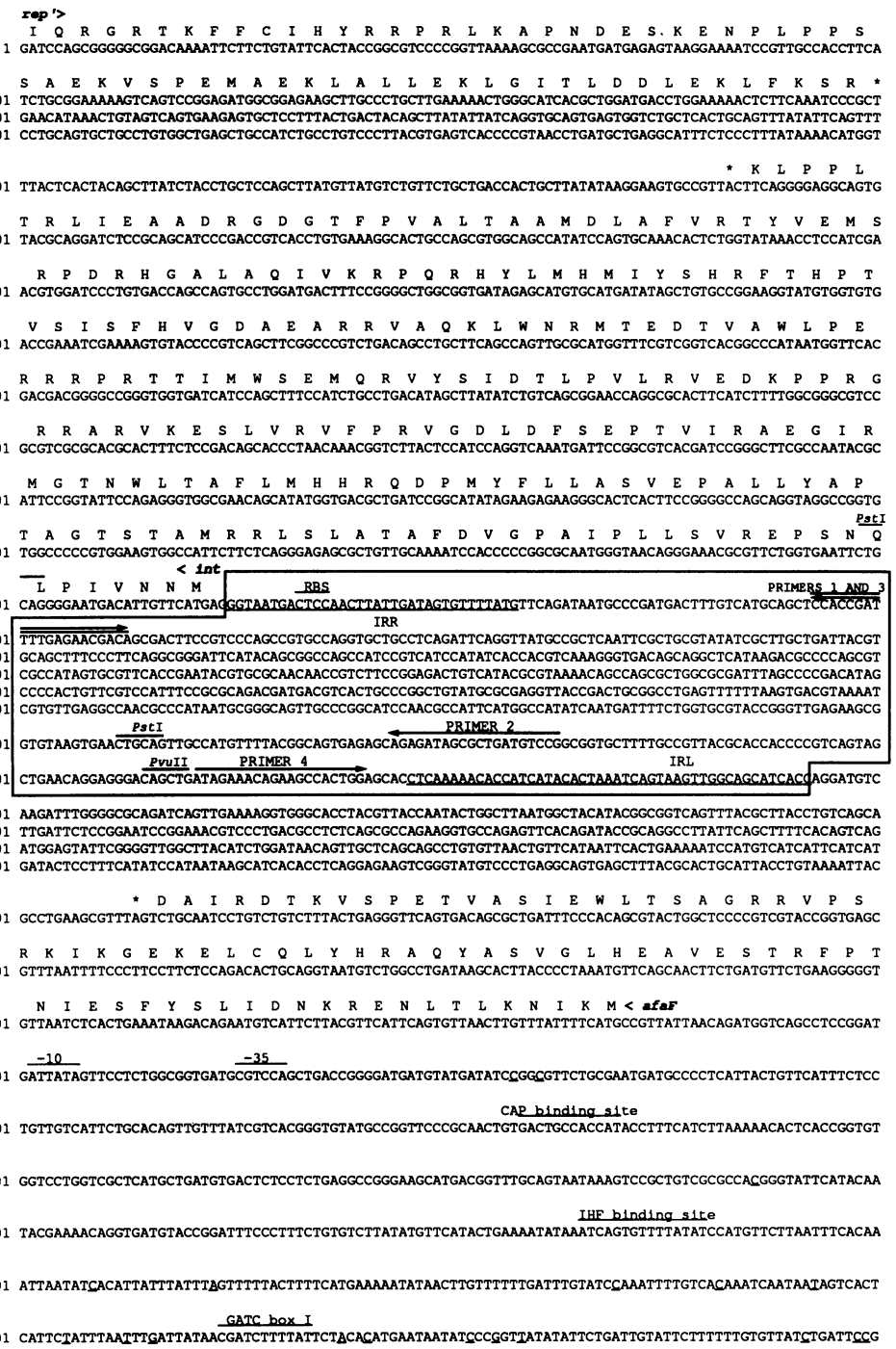

FIG. 2. Nucleotide sequence of the pILL61 insert. The numbers on the left represent the nucleotide positions of the first base in each line. The deduced amino acid sequences are shown in the single-letter code above the first nucleotide of the codons. Potential RBS, the promoter sequences for the −35 and −10 regions, first amino acid residues of the mature proteins (+1), and regulatory regions (cyclic AMP-cyclic AMP receptor protein [CAP] and integration host factor [IHF] binding sites and GATC boxes) are indicated above the nucleotide sequence. The IS*1* element is boxed and right (IRR) and left (IRL) inverted repeats are underlined. The two *Pst*I sites delineating the 0.6-kb fragment used as the IS*1* probe as well as the IS*1* internal *Pvu*II site are positioned. Primers 1 and 2 (used in a PCR assay) and primers 3 and 4 (used to sequence the target-IS*1* junctions) are also indicated. Between bp 2670 and 3470, the nucleotides found to be different from those in the corresponding region of the *daa* gene cluster (7) are underlined.

to be a 101-amino-acid (aa) polypeptide with a calculated mass of 11,735 Da. It exhibited marked homologies with *E. coli* fimbrial regulators such as DaaA (7) (98.8% identity) and PapB (5) (36.1%), as illustrated in Fig. 3A. It is noteworthy that the second putative start codon of the *afaA* gene (nucleotides 3519 to 3521) corresponds to the ATG start codon of the related *daaA* gene (Fig. 3A).

A second ORF, the *afaB* gene, resided between nucleotide

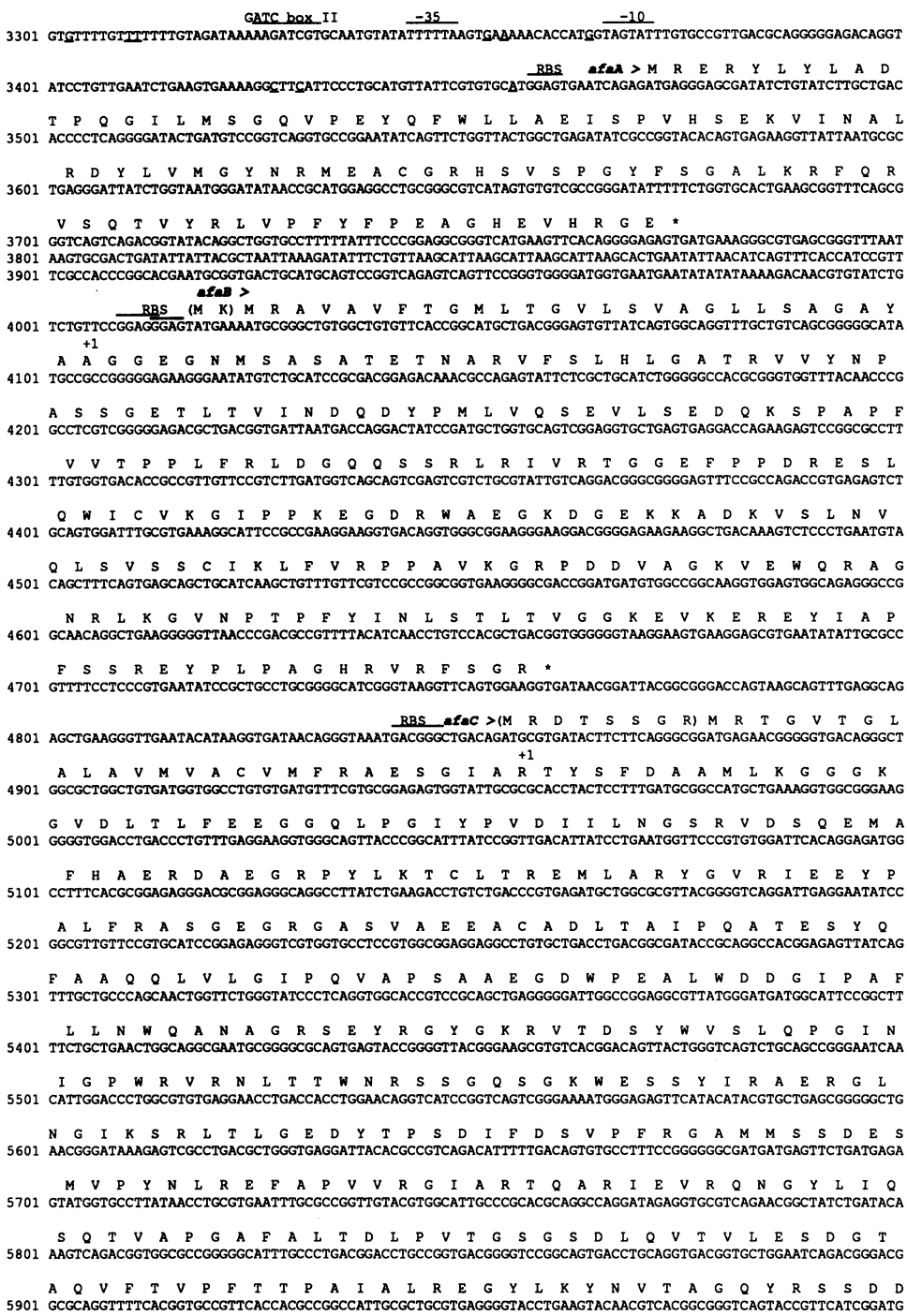

FIG. 2—*Continued.*

4018 or 4024 and nucleotide 4761. It encoded a protein 245 or 247 codons in length, with two possible initiation codons located four and seven nucleotides downstream of a putative Shine-Dalgarno sequence, respectively. The first 29 or 27 aa exhibited the features of procaryotic signal sequences (43). The processed AfaB polypeptide was 218 aa in length (molecular mass of 23.89 kDa), with a calculated pI of 9.94. Significant similarities were found between the AfaB protein and the periplasmic chaperone proteins involved in the biogenesis of bacterial adhesive or capsular structures (21): these all have the same size (230 to 260 aa) and a theoretical pI higher than 9. The protein sequence alignment of the AfaB peptide with some of these molecular chaperones is shown in Fig. 3B. The highest level of homology (95.9% of identity) was found with the NfaE protein (1), which is involved in the production of the chromosomally encoded nonfimbrial adhesin NFA-I. The predicted AfaB product also shared significant homologies of 40.3 and 31.7%, respectively, with the molecular chaperone

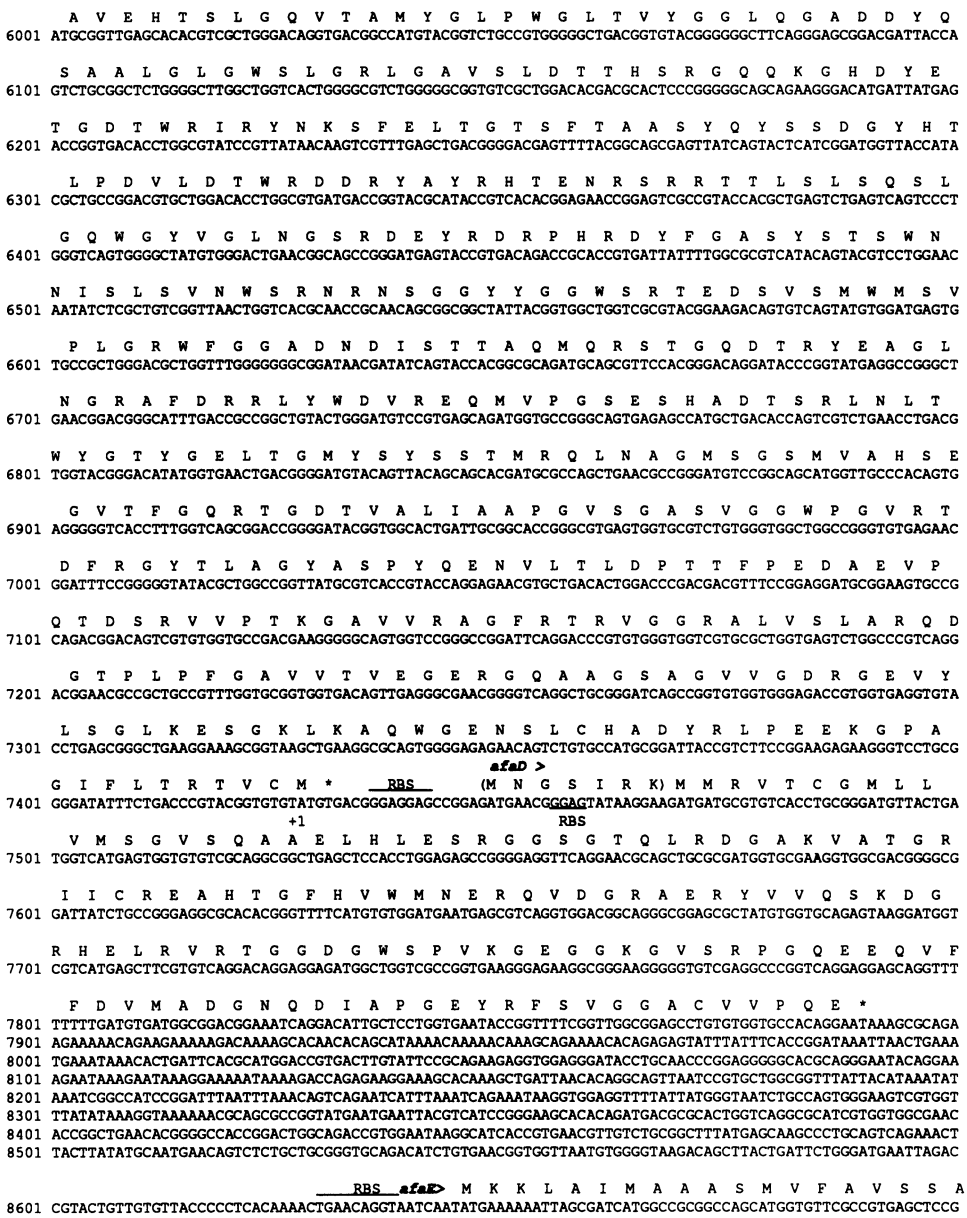

FIG. 2—*Continued.*

(Caf1M) of the *Yersinia pestis* capsule antigen F1 (15) and the PapD chaperone involved in the biogenesis of the P fimbriae (17). A chaperone consensus sequence deduced from the alignment of 13 members of the family was proposed by Jones et al. (21). AfaB contained 9 of the 12 aa considered to be invariant, which are included the two cleft residues (Arg-8 and Lys-112) that in PapD were shown to make up a critical binding site (21). Moreover, 42 of the 52 conserved residues are present in the processed AfaB peptide (Fig. 3B).

The *afaC* gene corresponded to a long coding sequence located 92 bp downstream of the *afaB* gene (nucleotides 4854

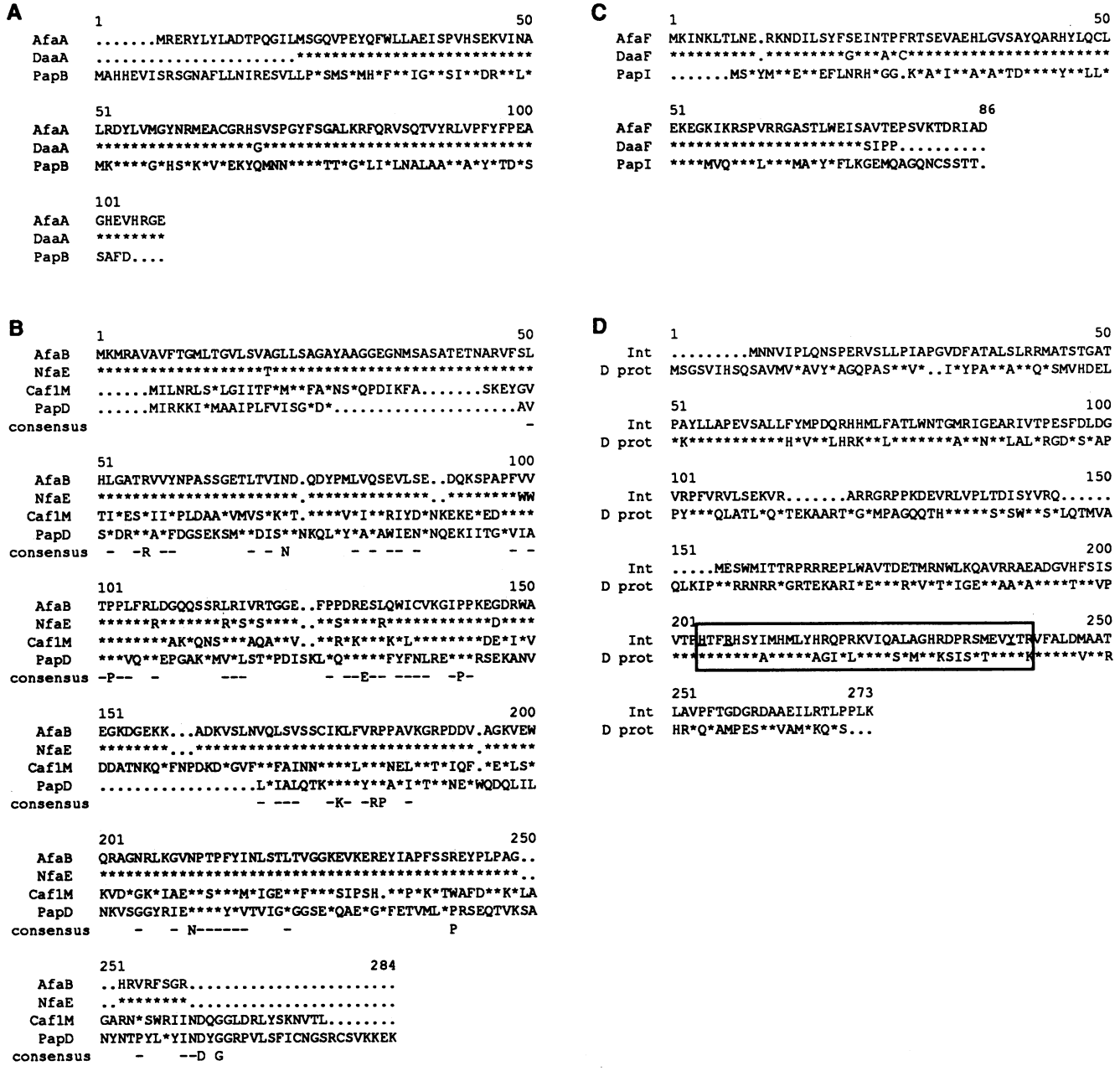

FIG. 3. Protein sequence alignments. Gaps have been inserted in the sequences to obtain maximum homology. Numbers correspond to the amino acid positions of the largest proteins. Asterisks mark residues identical to those of the Afa sequences. (A and C) Alignments of AfaA and AfaF with related transcription regulatory proteins. In panel A, the protein sequence of AfaA is aligned with those of DaaA (7) and PapB (5). In panel C, the protein sequence of AfaF is aligned with those of DaaF (7) and PapI (5). (B) The protein sequence of AfaB is aligned with those of the periplasmic chaperones NfaE (1), Caf1M (15), and PapD (17). The positions of the residues described by Jones et al. (21), characteristic of the periplasmic chaperone family, are indicated by letters (invariant residues) and dashes (conserved hydrophobic characters or residues conserved in 8 of the 13 sequences aligned). (D) The protein sequence of Int is aligned with that of the D protein (D prot) of the mini-F plasmid (28). The 40-aa segment, which is a signature of the site-specific recombinases of the λ integrase family (4), is framed. The 3 aa common to all known integrases (4) are underlined.

to 7433). This ORF had two potential start codons at positions 4854 and 4878. Although only the first ATG (position 4854) was preceded by a potential RBS, the putative signal sequence was considered to start at nucleotide 4878 because it would better match the consensus sequence (43). The hypothetical encoded peptide was 851 aa in length, with a calculated molecular mass of 92.9 kDa. A database search revealed homology between AfaC and outer membrane proteins such as FimD (33.9% identity) (23), MrkC (29.9%) (2), or PapC (27%) (32), involved in the assembly of bacterial fimbriae, as well as Caf1A (44.9%), necessary for *Y. pestis* capsule biogenesis (22). Analysis of the secondary structure prediction revealed that mature AfaC has the same molecular mass, the same amino acid composition, and the same typical features of

outer membrane proteins as the other proteins of the PapC family.

Following the *afaC* gene was a short intergenic region of 15 bp and an ORF able to encode a polypeptide unrelated to any protein sequence in the database. This ORF, designated the *afaD* gene, mapped to the same locus as that of the *afaD* gene within the *afa-1* gene cluster. This locus was previously shown to encode a polypeptide not essential for MRHA expression (26). The *afaD* gene of the *afa-3* gene cluster of pILL61 had three potential in-frame initiation codons at nucleotides 7449, 7470, and 7473, with a termination codon at nucleotide 7890. All three potential ATG codons were preceded by a Shine-Dalgarno sequence and followed by a probable signal peptide sequence with a classical cleavage site (Ser-Gln-Ala) recognized by the signal peptidase I. The predicted mature AfaD protein contained 121 residues with a deduced molecular mass of 13.16 kDa and a pI of 6.52 and demonstrated a high glycine content (16.5%). The secondary structure of the protein consisted mainly of β-sheet structures with very short strips of α-helix.

Finally, the DNA region between bp 8646 and bp 9128 corresponded to the *afaE* gene that encodes the 15-kDa AFA-III adhesin, previously designated *afaE-3* (30).

Immediately upstream of the *afaA* gene was an 801-bp noncoding region (from nucleotide 2670 to nucleotide 3470) preceded by an ORF (bp 2412 to 2669) transcribed in the opposite orientation from that of the *afaA*, *afaB*, *afaC*, *afaD*, and *afaE* genes. The product of this 85-aa ORF (molecular mass of 9,731 Da) had homologies with fimbrial regulatory proteins such as the PapI activator (5) and the DaaF peptide (7) (48.6 and 90.7% identity, respectively) (Fig. 3C). This ORF, which was designated *afaF*, was considered to have a possible regulatory function on the expression of the other *afa-3* genes. Sequence analyses of the pILL61 insert revealed that the *afaA* and *afaF* genes were homologous with those of the *daa* gene cluster, as was the *afaA-afaF* intergenic region. Alignment of the region from nucleotide 2670 to nucleotide 3470 of pILL61 with the published sequence of the *daa* promoter region (7) revealed more than 95% (773 of 801) identical nucleotides. The two promoters as well as the potential regulatory sequences involved in the binding of the leucine-responsive regulatory protein (Lrp), cyclic AMP-cyclic AMP receptor protein, and integration host factor, which had been identified in the *daa* intercistronic region (7), were also found in the *afaF-afaA* intergenic region (Fig. 2).

DNA sequence analysis of the 2.4-kb region located upstream of the *afa-3* gene cluster. Ninety-six percent identity was found between the DNA sequence located upstream of the *afa-3* gene cluster (bp 1 to 1197 [Fig. 2]) and a region of the P307 enterotoxigenic plasmid (37) that corresponded to a part of a replicon (RepFIB) capable of autonomous replication. This similarity included (i) an ORF (designated *rep'*; bp 1 to 202) encoding a truncated peptide almost identical (65 of 66 aa) to the carboxy-terminal domain of the RepA protein, a protein that is involved in the replication of the RepFIB replicon; (ii) a stretch of DNA including sequence and motifs (bp 203 to 480) that have been reported to be necessary for the autonomous replication of RepFIB (41); and (iii) an ORF (bp 481 to 1221), transcribed in the same orientation as that of the *afaF* gene (Fig. 1), homologous to an ORF of P307 (ORF-3) (37). The ORF on pILL61, designated the *int* gene, encoded a 246-aa polypeptide (molecular mass, 28.1 kDa) that displayed 48.5% identity and 61.8% similarity with the D protein of mini-F, a site-specific resolvase (28). Figure 3D shows the alignment of the Int protein of pILL61 with the D protein of mini-F. This alignment reveals the presence of a 40-aa conserved motif located at the C-terminal end of the proteins that contains the triad His-Arg-Tyr, which allows one to identify such a protein as a site-specific recombinase of the λ integrase family (4).

An IS*1* insertion sequence was identified 3 bp downstream of the initiation codon of the *int* gene. The putative RBS of the *int* gene was located within the right end of IS*1* (Fig. 2). The whole pILL61 IS*1* nucleotide sequence only differed by 1 bp (A-T substitution at bp 1731) from the IS*1* element characterized by Johnsrud (19). It contained the three typical ORFs of the IS*1* element, designated the *insA*, *insB*, and *insC* genes (11). In pILL61, these mapped to bp 1937 and 1662, 1617 and 1240, and 1274 and 1564, respectively. No direct repeated duplications of the expected nine nucleotides (20) were observed at the junctions between the targeted site and the IS*1* element.

Expression of the *afa* gene cluster in *E. coli* maxicells. To visualize the proteins expressed by the different identified ORFs, pILL61 and derivative plasmids (pILL1108, pILL1110, and pILL1111 [Table 1 and Fig. 1A]) were transformed into the *E. coli* maxicell-producing strain SE5000. Comparison of polypeptides expressed by pILL61 and pILL1111 and the corresponding vectors pBR322 and pACYC184 (lanes 4, 5, 6, and 7 in Fig. 4A) allowed us to detect the expression of five Afa proteins with apparent molecular masses of 94, 30, 17.5, 15.5, and 11 kDa. The 94-kDa protein was attributed to the *afaC* gene. Further investigations were required to unambiguously identify the other gene products. The 30-kDa protein was identified as the AfaB product; it was present in the pILL1111 and pILL61 profiles and absent from the pILL1108 and pILL1110 ones. The 17.5-kDa protein, identified as the *afaD* gene product, was expressed by pILL1108, pILL1111, and pILL61 and was not expressed by pILL1110. This was further confirmed by the fact that purified AfaD product migrated on SDS-PAGE gels as a 17.5-kDa protein (Fig. 4B). In the same way, the 15.5-kDa protein of pILL61 and pILL1111 was recognized as the AfaE adhesin (Fig. 4). In addition, the purified 15.5-kDa protein exhibited MRHA properties. Finally, the 11-kDa polypeptide present in pILL1111, pILL61, and pILL1110, but absent from the pILL1108 profile, was identified as the AfaA product. On the whole, the results of the maxicell experiments were consistent with the deduced number and molecular masses of the *afa* gene products determined by sequence analysis. One of the two exceptions to this were the AfaB and AfaD proteins, which were found to migrate in SDS gels more slowly than expected. Such a discrepancy might reflect the existence of posttranslational modifications of the two polypeptides. The other exception was the absence of the *afaF* gene product from SDS-polyacrylamide gels, presumably because of a low level of expression of AfaF, a regulatory protein.

IS*1* elements in the vicinity of the *afa-3* gene cluster. The *afa-3* gene cluster was cloned from the 100-kb pILL1055 plasmid present in the *E. coli* A30 isolate. On the basis of the nucleotide sequence analysis, a unique copy of the IS*1* element was found to map to the pILL61 9.2-kb insert, whereas Southern hybridization experiments demonstrated six copies of IS*1* on the whole pILL1055 plasmid (data not shown). We therefore decided to further investigate the presence of other copies of IS*1* in the vicinity of the *afa-3* gene cluster. DNA sequences located upstream and downstream of the 9.2-kb insert of pILL61 were isolated from pILL1055 by cloning partially digested *Sau*3A fragments (of between 6 and 13 kb) into pILL570 and then screening the recombinant plasmids with an *afa* probe by Southern hybridization. Two recombinant plasmids, pILL1099 and pILL1101 (Table 1 and Fig. 1B), were

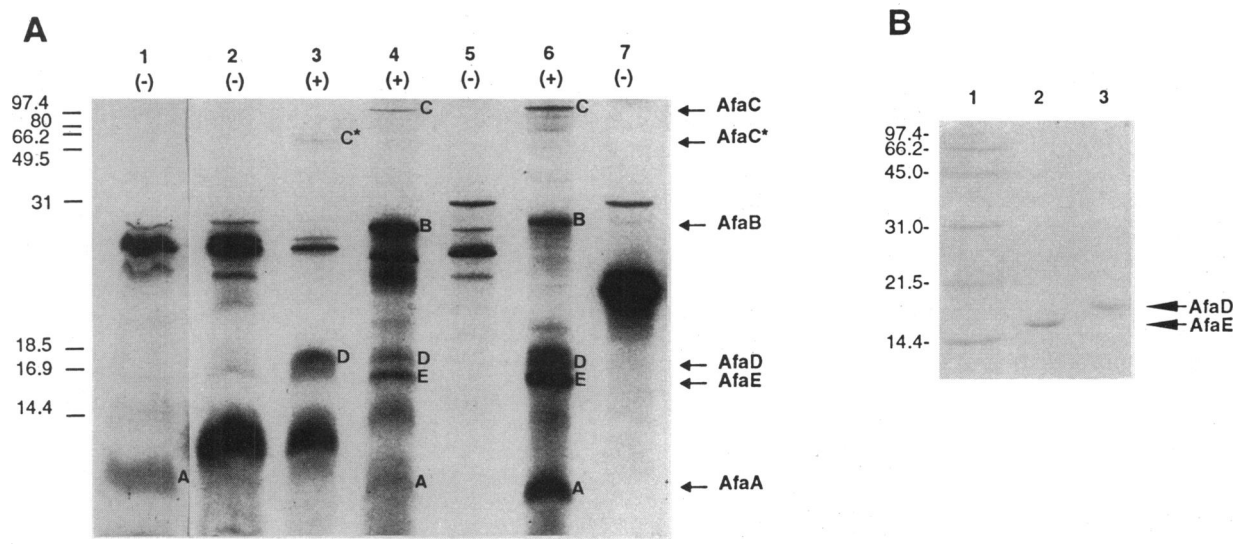

FIG. 4. Expression of Afa polypeptides. (A) Fluorographs of plasmid-mediated polypeptides expressed in maxicells labeled with [^{35}S]methionine and analyzed on SDS–15% polyacrylamide gels. The molecular masses of size standards (in kilodaltons) are indicated on the left. The unprocessed forms of the AfaA, AfaB, AfaC, AfaD, and AfaE polypeptides and that of the truncated AfaC peptide are indicated by letters (A, B, C, D, E, and C*, respectively). The presence or absence of ethanol in the assay media is indicated by (+) or (−). The 30-kDa AfaB product comigrates with the unprocessed β-lactamase of pBR322 in the pILL61 profile obtained in the presence of ethanol. Lanes 2, 5, and 7, vectors pILL570 (lane 2), pBR322 (lane 5), and pACYC184 (lane 7). Lanes 1, 3, 4, and 6, recombinant plasmids pILL1110 (lane 1), pILL1108 (lane 3), pILL61 (lane 4), and pILL1111 (lane 6). (B) SDS-PAGE of the purified AfaE (lane 2) and AfaD (lane 3) proteins after Coomassie blue staining. Molecular size standards are shown in lane 1. Protein sizes are indicated in kilodaltons.

selected, and these contained inserts of 11.9 and 11.6 kb that extended over regions located either upstream (pILL1099) or downstream (pILL1101) of the *afa-3* gene cluster. For each plasmid, two *Pst*I restriction fragments (2.5 and 0.6 kb in size) hybridized with the IS*1* probe (Fig. 5, lanes 2 and 3). The 0.6-kb *Pst*I hybridizing fragment corresponded to the 0.6-kb fragment of pILL61 used as a probe and therefore contained the first identified IS*1* element (Fig. 5, lane 1). The 2.5-kb *Pst*I hybridizing fragment of pILL1099 and pILL1101 corresponded to different sequences of pILL1055, located either upstream (pILL1099) or downstream (pILL1101) of the *afa-3* gene cluster (Fig. 1B). This assumption was confirmed by

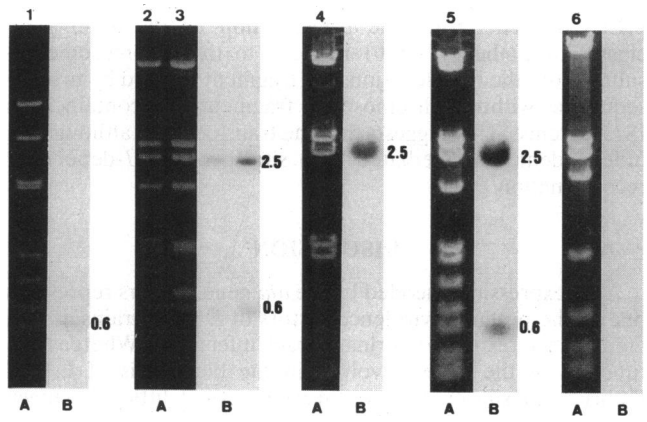

FIG. 5. Genetic location of IS*1* sequences around *afa* operons. (A) Agarose gel electrophoresis of *Pst*I-restricted plasmids. Lanes: 1, pILL61; 2, pILL1099; 3, pILL1101; 4, pILL1024; 5, pILL1027; 6, pILL1018. (B) Southern blot analyses with IS*1* as a probe. Molecular sizes (in kilobases) of the hybridizing fragments are indicated on the right.

subcloning the 3.5-kb *Sma*I fragment of pILL1099 and the 5.8-kb *Sma*I fragment of pILL1101 into pACYC177, resulting in recombinant plasmids designated pILL1096 and pILL1097, respectively (Table 1 and Fig. 1B). Both of the *Sma*I fragments hybridized to the IS*1* probe, allowing us to conclude that both carried DNA sequences homologous to the IS*1* element (not shown). To confirm the presence of an entire IS*1* element in each of these fragments, gene amplifications were performed with a couple of primers designed to target the IS*1* nucleotide sequence (IS*1* 1 to IS*1* 2 [Fig. 2]). Plasmids pILL1096 and pILL1097 both led to the amplification of the expected 570-bp internal fragment of IS*1*. In addition, the precise location and orientation of each of the three copies of IS*1* found in the vicinity of the *afa-3* gene cluster were determined by *Pvu*II-*Pst*I restriction map analysis of pILL1096 and pILL1097, as shown in Fig. 1B. The *afa-3* gene cluster appeared to be flanked by two IS*1* elements in direct orientation, whereas the third IS*1* sequence was inserted in the opposite orientation. For each of the three IS*1* sequences, the nucleotide sequences of the junctions between the extremities of the IS*1* elements and the pILL1055 target sites were determined by double-stranded DNA sequencing of plasmids pILL1096 and pILL1097 with the IS*1* 3 and IS*1* 4 oligonucleotides shown in Fig. 2. No duplication of the target site was found on either side of the individual or paired IS*1* elements.

Experiments were also undertaken to identify IS*1* elements on recombinant plasmids harboring either the *afa-3* gene cluster from the diarrhea-associated *E. coli* strain, AL845 (30), or the *afa-1* gene cluster from the pyelonephritic strain, KS52 (27) (Table 1 and Fig. 1C and D). The restriction maps and genetic organization of the sequences flanking the *afa-3* gene cluster of strain AL845 (Fig. 1C) were remarkably conserved compared with those of the cystitis-associated strain, A30 (Fig. 1A and B). As for the A30 strain, three copies of IS*1* were detected in the vicinity of the *afa-3* gene cluster of strain

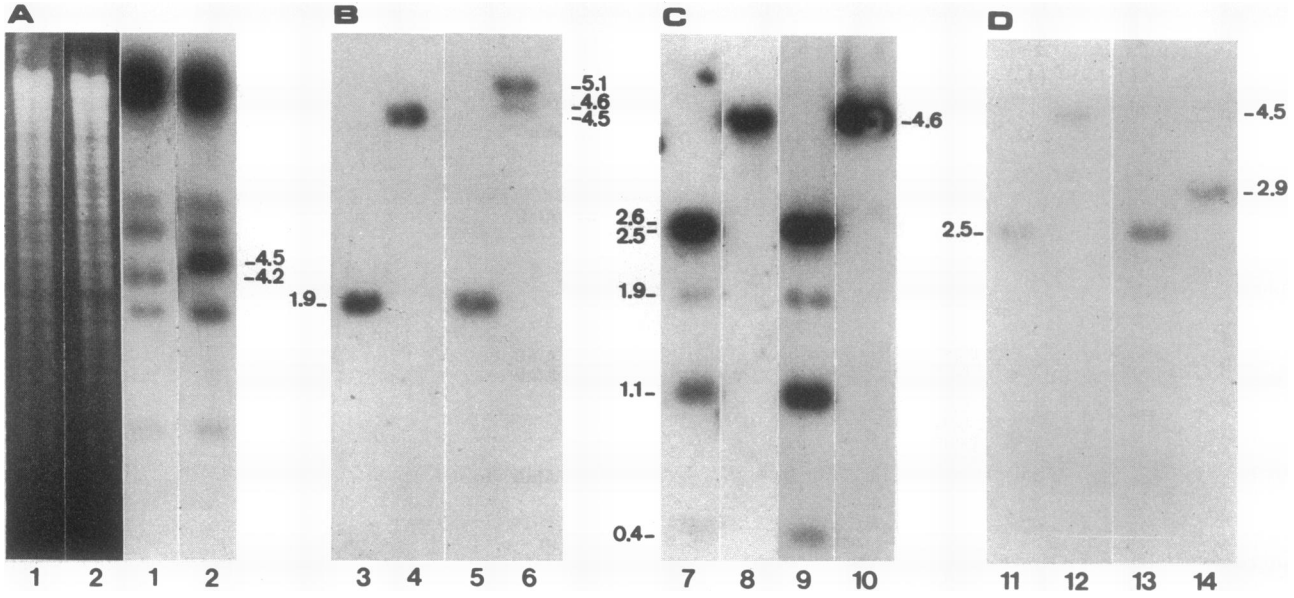

FIG. 6. Southern blot analyses of total DNA extracted from strains AL853 and HB101. (A) Electrophoresis of *Sph*I-restricted total DNA and hybridization with the IS*1* probe. Lanes: 1, HB101; 2, AL853. (B, C, and D) Results of hybridizations with the B, C, and D probes (see Fig. 1B), respectively. Lanes: 3, 7, and 11, *Pst*I-digested AL853 DNA; 4, 8, and 12, *Sph*I-digested AL853 DNA; 5, 9, and 13, *Pst*I-digested pILL1101; 6, 10, and 14, *Sph*I-digested pILL1101. Values (in kilobases) on the sides indicate the sizes of the hybridizing fragments.

AL845 (Fig. 5, lanes 4 and 5), which mapped to the same loci and had the same orientation as those found in the pILL1055 plasmid (Fig. 1C). In contrast the 11.6-kb sequence originating from strain KS52, harboring the *afa-1* operon, did not contain any IS*1* sequences (Fig. 1C and Fig. 5, lane 6).

Translocation of the *afa-3* gene cluster. Because it is known that any DNA sequence flanked by two transposable elements is by definition a mobile element, we wanted to test the capacity of the *afa-3* gene cluster, which was flanked by two entire IS*1* elements, to move from one replicon to another. Plasmid pUCD800, which harbors the *sacB* gene encoding the levansucrase whose expression is lethal for *E. coli* cells in the presence of sucrose (16), was introduced into HB101 harboring the pILL1101 plasmid. After transformation, bacteria were plated on agar medium containing kanamycin and 5% sucrose to positively select for bacteria in which insertion into the *sacB* gene had occurred. Of 100 clones analyzed, 1 had lost spectinomycin resistance capability but retained the ability to agglutinate human erythrocytes. Plasmid analysis of this clone demonstrated the absence of the original pILL1101 plasmid and the acquisition within the *sacB* gene of a 2.3-kb fragment. Because the MRHA phenotype requires at least 5,700 nucleotides for expression, it was unlikely that the entire *afa-3* gene cluster had inserted into the *sacB* gene. After two antibiotic-free subcultures of the clone, a kanamycin-susceptible derivative was isolated. This derivative, designated AL853, was found to be plasmid free and still retained the MRHA properties, thus suggesting a chromosomal location of the *afa-3* gene cluster. To further investigate this hypothesis, Southern hybridization experiments with the *Pst*I- or *Sph*I-restricted whole DNA extracted from *E. coli* HB101 or *E. coli* AL853 were performed with various probes (pILL570, IS*1*, and probes designated A to E [Fig. 1B]). No hybridization was detected with the pILL570 probe or with the A and E probes, both corresponding to DNA sequences that flanked the two direct IS*1* elements in pILL1101 (data not shown). However, the complete sequence framed by these two IS*1* elements was recovered in the AL853 total digest and not in that of HB101 (Fig. 6). These results indicated that acquisition of the MRHA phenotype was the result of the integration of the fragment of the insert of pILL1101, corresponding to the DNA sequence flanked by the two IS*1* elements. Conclusive evidence that translocation had occurred within the chromosome of the HB101 strain came from the observation that concomitant loss of the spectinomycin resistance of the original plasmid (pILL1101) and that of pUCD800 derivative was associated with the disappearance of an HB101 *Sph*I fragment (4.2 kb in size) harboring an IS*1* sequence and the appearance of two new chromosomal IS*1* hybridizing fragments (both 4.5 kb) (Fig. 6). These latter fragments were absent from pILL1101 and therefore must have corresponded to the junction fragments between the *afa-3* gene cluster and the HB101 chromosome, replacing the 5.1- and 2.9-kb *Sph*I fragments of pILL1101. Therefore, the translocation of the *afa-3* gene cluster from the pILL1101 replicon to the chromosome resulted from the translocation of a fragment flanked by two IS*1* sequences within a chromosomal fragment that contained an IS*1* sequence. This suggests that the translocation, although *rec* independent, was mediated by a site-specific IS*1*-dependent recombination event.

DISCUSSION

AFA expression encoded by the *afa* gene clusters represents one of the potential virulence factors of *E. coli* strains associated with intestinal or urinary tract infections. Whereas the function of the genes involved in the biogenesis of *E. coli* fimbrial adhesins has been extensively studied, little was known concerning the genes involved in the function of the afimbrial adhesins. We report here on the determination of the nucleotide sequence of the pILL61 insert that contains the plasmid-borne *afa-3* gene cluster cloned from *E. coli* A30. This is the first instance in which the entire nucleotide sequence of an afimbrial-adhesin-encoding system has been reported.

The information deduced from the sequence analysis has shown that the biogenesis of an *E. coli* afimbrial adhesive structure, like that of a fimbrial one (18), requires proteins with specialized functions such as a periplasmic chaperone (AfaB protein) and an outer membrane anchor protein (AfaC). The AfaB and AfaC proteins were found to be more closely related to the *caf1M* and *caf1A* gene products, which are required for the formation of the *Y. pestis* capsule (22), than to the homologous gene products of the *pap* and *sfa* gene clusters encoding the P and S fimbriae, respectively. These observations suggested that the biogenesis of bacterial capsule, nonfimbrial adhesin, or fimbrial adhesins involved the export of subunits that have to be stabilized in the periplasm and exported to the cell surface via similar pathways.

High degrees of similarity were found between the nucleotide sequence of the *afa-3* gene cluster and the partial nucleotide sequence of the nonfimbrial NFA-I-encoding operon recently reported by Ahrens et al. (1). Whereas the structural gene encoding the NFA-I adhesin had no homology with any of the genes encoding the AFA-I (27), AFA-II (unpublished data), and AFA-III (30) adhesins (nor was the NFA-I related to the AFAs at the peptide level), the AfaB chaperone exhibited 95.9% identity with the homologous NfaE protein. Moreover the noncoding sequence located upstream of the adhesin-encoding genes (*afaE* and *nfaA*) displayed strong similarities (79.5% identity on 138 nucleotides) (data not shown). Such conservation at the DNA level suggests that both the *nfa* and *afa* determinants are closely related. Further investigations should allow one to conclude whether they belong to the same phylogenetic family of nonfimbrial adhesins.

We and others previously reported that the different *afa* operons, the *daa* gene cluster encoding the F1845 fimbrial adhesin (8), and the *dra* operon encoding the fimbrial Dr hemagglutinin were closely related (30, 35). Although the entire nucleotide sequences of the *daa* and *dra* operons have not yet been reported, comparison of the published partial sequences with that of the *afa-3* operon seems to confirm this statement. Yet why the F1845 adhesin and Dr hemagglutinin polymerize to form fimbriae while the AFA adhesins do not is not understood. The *afaA* and *afaF* genes were found to encode products highly similar to the DaaA and DaaF proteins, which have been shown to regulate, at the transcriptional level, the expression of the F1845 fimbrial adhesin (7). These gene products belong to a family of transcriptional regulators commonly involved in the regulation of *E. coli* fimbrial adhesin expression (33). Not only were the two regulatory proteins highly similar to the Daa ones, but the 1,365-bp *afaF* and *afaA* intergenic region was found to be 96.2% identical to the *daaF-daaA* intergenic region. The binding sites for the cyclic AMP-cyclic AMP receptor protein, integration host factor, and Lrp transcription regulatory factors pointed out by Bilge et al. (7) were also detected in the *afaF-afaA* intragenic region of pILL61, suggesting similar transcriptional regulatory mechanisms for the expression of the *afa* operon. Similarly, the flanking sequences of the *afaE* gene encoding the adhesin subunit share similarity with those of the *daaE* gene. These flanking sequences display extensive homology upstream of the adhesin genes and include the presence of a stem-loop structure (data not shown), located 50 bp downstream of both *afaE* and *daaE*, that has been shown by Bilge et al. (6) to be essential to the stability of the processed *daaE* mRNA. Therefore, given such similarities between the two *daa* and *afa-3* operons, why was one associated with a fimbrial adhesin (*daa*) and the other associated with a nonfimbrial adhesin (*afa-3*)? We previously hypothesized that this could be the result of a variation in the rates of expression of the adhesin genes or in the expression of a product encoded by accessory genes (30). To date, the only identified differences between the three operons mapped to a region that includes the identified *afaD* gene and the 753-bp *afaD-afaE* intergenic region. In the *daa* gene cluster (8) as well as the *dra* operon encoding the Dr adhesin (35), the *afaD* equivalent loci were shown to encode peptides not essential for the expression of adhesion properties, a situation that was similar to that found in *afa-3* (30). Because of the absence of homology with any known proteins, sequence determination of the *afaD* gene, as reported in this paper, did not allow us to propose a putative function for the AfaD product. The preparation of specific polyclonal antibodies raised against purified AfaD and AfaE polypeptides, as well as the construction of chimeric operons between the *daa* and the *afa-3* operons, should soon allow us to elucidate this controversial observation.

Two main features were revealed by the nucleotide determination of the sequences bordering the adhesin-encoding operon. First, the *afa-3* operon has been shown to be linked to a region encoding a protein (RepA) required for the initiation of replication of various bacterial plasmids and another protein (Int) belonging to the lambda integrase family of site-specific recombinases. Such a situation has also been reported for the 90-kb virulence plasmid of *Salmonella typhimurium* (14), in which the fimbrial biosynthetic genes (*pef* genes) were shown to be located immediately downstream of a 1-kb region encoding the RepA protein and the integrase-like D protein of the mini-F replicon. Topological linkage of independent gene blocks might reflect the involvement of common ancestral recombinational events that led to the formation of several enterobacterial large virulence plasmids.

The second major feature is that the *afa-3* gene clusters from both the uropathogenic A30 and the enteroadherent AL845 isolates were framed by three IS*1* sequences in such a position that the AFA-III-encoding determinants were flanked by two IS*1* elements in direct orientation and two in opposite orientations. Such an arrangement of the *afa* operon was restricted to the plasmid-borne *afa-3* gene cluster, because no IS*1* was identified in the vicinity of the chromosomal *afa-1* operon. The absence of direct repeated short duplications (9 bp) (20) on both sides of individual or paired IS*1* elements demonstrated that the acquisition of the *afa-3* gene cluster was not the result of a recent transposition event mediated by the flanking IS*1* insertion sequences. Rather, this appeared to be the result of replicon fusions or DNA translocation mediated by reciprocal or site-specific recombination events, as has been commonly seen in the building of natural plasmids encoding multiple drug resistances (11). We further demonstrated, by the use of a strong positive selection, that the element composed of the *afa-3* gene cluster flanked by two IS*1* elements in direct orientation was indeed a mobile element capable of translocation from the recombinant plasmid pILL1101 into the chromosome of HB101. The insertion of the *afa-3* gene cluster into the chromosomal DNA was probably the result of a recombinational site-specific event occurring between the two IS*1* elements flanking the *afa-3* gene cluster and one of the IS*1* copies present in the chromosome. This movement could be *recA* independent because it took place in HB101, a strain that is reported to be deficient in its reciprocal recombination functions. Whereas such a situation has already been reported for drug resistance genes (20) and for genes encoding toxins (39), this is the first evidence that a pathogenic determinant responsible for adhesion properties is capable of translocation from one replicon to another. This allows us to speculate on

the possible dissemination of such a determinant among gram-negative bacteria.

In conclusion, our results indicate that the biogenesis of the afimbrial adhesive structure encoded by the *afa* gene clusters follows the same routes as that of the well-known *E. coli* P fimbria formation. Nucleotide sequence determination of the *afa* genes should allow us to initiate experiments designed to determine the conformation of the afimbrial adhesive structure and its mechanism of interaction with epithelial cells.

ACKNOWLEDGMENTS

We thank F. Norel and R. Ferrero for helpful criticism of the manuscript.

REFERENCES

1. **Ahrens, R., O. Manfred, A. Ritter, H. Hoschützky, T. Bühler, F. Lottspeich, G. J. Boulnois, K. Jann, and J. Hacker.** 1993. Genetic analysis of the gene cluster encoding nonfimbrial adhesin I from an *Escherichia coli* uropathogen. Infect. Immun. **61:**2505–2512.
2. **Allen, B. L., G.-F. Gerlach, and S. Clegg.** 1991. Nucleotide sequence and functions of *mrk* determinants necessary for expression of type 3 fimbriae in *Klebsiella pneumoniae*. J. Bacteriol. **173:** 916–920.
3. **Archambaud, M., P. Courcoux, and A. Labigne-Roussel.** 1988. Detection by molecular hybridization of PAP, AFA, and SFA adherence systems in *Escherichia coli* strains associated with urinary and enteral infections. Ann. Inst. Pasteur/Microbiol. **139:** 575–588.
4. **Argos, P., A. Landy, K. Abremski, J. B. Egan, E. Haggard-Ljungquist, R. H. Hoess, M. L. Kahn, B. Kalionis, S. V. L. Narayana, L. S. Pierson III, N. Sternberg, and J. M. Leong.** 1986. The integrase family of site-specific recombinases: regional similarities and global diversity. EMBO J. **5:**433–440.
5. **Baga, M., M. Goransson, S. Normark, and B. E. Uhlin.** 1985. Transcriptional activation of a Pap pilus virulence operon from uropathogenic *Escherichia coli*. EMBO J. **4:**3887–3893.
6. **Bilge, S. S., J. M. Apostol, M. A. Aldape, and S. L. Moseley.** 1993. mRNA processing independent of RNase III and RNase E in the expression of the F1845 fimbrial adhesin of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **90:**1455–1459.
7. **Bilge, S. S., J. M. J. Apostol, K. J. Fullner, and S. L. Moseley.** 1993. Transcriptional organization of the F1845 fimbrial adhesin determinant of *Escherichia coli*. Mol. Microbiol. **7:**993–1006.
8. **Bilge, S. S., C. R. Clausen, W. Lau, and S. L. Moseley.** 1989. Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated *Escherichia coli* to HEp-2 cells. J. Bacteriol. **171:**4281–4289.
9. **Bolivar, F., R. Rodriguez, P. J. Greene, M. Betlach, H. L. Heynecker, H. W. Boyer, J. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. Gene **2:**95–113.
10. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41:**459–472.
11. **Braedt, G.** 1988. Different reading frames are responsible for IS*1*-dependent deletion and recombination. Genetics **118:**561–570.
12. **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:** 1141–1155.
13. **Cussac, V., R. L. Ferrero, and A. Labigne.** 1992. Expression of *Helicobacter pylori* urease genes in *Escherichia coli* grown under nitrogen-limiting conditions. J. Bacteriol. **174:**2466–2473.
14. **Friedrich, M. J., N. E. Kinsey, J. Vila, and R. J. Kadner.** 1993. Nucleotide sequence of a 13.9 kb segment of the 90 kb virulence plasmid of *Salmonella typhimurium*: the presence of fimbrial biosynthetic genes. Mol. Microbiol. **8:**543–558.
15. **Galyov, E. E., A. V. Karlyshev, T. V. Chernovskaya, D. A. Dolgikh, O. Y. Smirnov, K. I. Volkovoy, V. M. Abramov, and V. P. Zav'yalov.** 1991. Expression of the envelope antigen F1 of *Yersinia pestis* is mediated by the product of *caf1M* gene having homology with the chaperone protein PapD of *Escherichia coli*. FEBS Lett. **286:**79–82.
16. **Gay, P., D. Le Coq, M. Steinmetz, T. Berkelman, and C. I. Kado.** 1985. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. J. Bacteriol. **164:** 918–921.
17. **Holmgren, A., M. J. Kuehn, C.-I. Brändén, and S. J. Hultgren.** 1992. Conserved immunoglobulin-like features in a family of periplasmic pilus chaperones in bacteria. EMBO J. **11:**1617–1622.
18. **Jacob-Dubuisson, F., M. Kuehn, and S. J. Hultgren.** 1993. A novel secretion apparatus for the assembly of adhesive bacterial pili. Trends Microbiol. **1:**50–54.
19. **Johnsrud, L.** 1979. DNA sequence of the transposable element IS*1*. Mol. Gen. Genet. **169:**213–218.
20. **Johnsrud, L., M. P. Calos, and J. H. Miller.** 1978. The transposon Tn9 generates a 9 base pair repeated sequence during integration. Cell **15:**1209–1219.
21. **Jones, C. H., J. S. Pinkner, A. V. Nicholes, L. N. Slonim, S. N. Abraham, and S. J. Hultgren.** 1993. FimC is a periplasmic PapD-like chaperone that directs assembly of type 1 pili in bacteria. Proc. Natl. Acad. Sci. USA **90:**8397–8401.
22. **Karlyshev, A. V., E. E. Galyov, O. Y. Smirnov, A. P. Guzayev, V. M. Abramov, and V. P. Zav'yalov.** 1992. A new gene of the f1 operon of *Y. pestis* involved in the capsule biogenesis. FEBS Lett. **297:**77–80.
23. **Klemm, P., and G. Christiansen.** 1990. The fimD gene required for cell surface localization of *Escherichia coli* type 1 fimbriae. Mol. Gen. Genet. **220:**334–338.
24. **Labigne, A., V. Cussac, and P. Courcoux.** 1991. Shuttle cloning and nucleotide sequences of *Helicobacter pylori* genes responsible for urease activity. J. Bacteriol. **173:**1920–1931.
25. **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.
26. **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.
27. **Labigne-Roussel, A. F., D. Lark, G. Schoolnik, and S. Falkow.** 1984. Cloning and expression of an afimbrial adhesin (AFA-I) responsible for P blood group-independent, mannose-resistant hemagglutination from a pyelonephritic *Escherichia coli* strain. Infect. Immun. **46:**251–259.
28. **Lane, D., R. de Feyter, M. Kennedy, S.-H. Phua, and D. Semon.** 1986. D protein of miniF plasmid acts as repressor of transcription and as a site-specific resolvase. Nucleic Acids Res. **14:**9713–9728.
29. **Le Bouguénec, C., M. Archambaud, and A. Labigne.** 1992. Rapid and specific detection of the *pap*, *afa*, and *sfa* adhesin-encoding operons in uropathogenic *Escherichia coli* strains by polymerase chain reaction. J. Clin. Microbiol. **30:**1189–1193.
30. **Le Bouguenec, C., M. I. Garcia, V. Ouin, J.-M. Desperrier, P. Gounon, and A. Labigne.** 1993. Characterization of plasmid-borne *afa-3* gene clusters encoding afimbrial adhesins expressed by *Escherichia coli* strains associated with intestinal or urinary tract infections. Infect. Immun. **61:**5106–5114.
31. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Norgren, M., M. Bäga, J. M. Tennent, and S. Normark.** 1987. Nucleotide sequence, regulation and functional analysis of the *papC* gene required for cell surface localization of Pap pili of uropathogenic *Escherichia coli*. Mol. Microbiol. **1:**169–178.
33. **Nou, X., B. Skinner, B. Braaten, L. Blyn, D. Hirsch, and D. Low.** 1993. Regulation of pyelonephritis-associated pili phase-variation in *Escherichia coli*: binding of the Pap I and the Lrp regulatory proteins is controlled by DNA methylation. Mol. Microbiol. **7:** 545–553.
34. **Nowicki, B., A. Labigne, S. Moseley, R. Hull, S. Hull, and J. Moulds.** 1990. The Dr hemagglutinin, afimbrial adhesins AFA-I and AFA-III, and F1845 fimbriae of uropathogenic and diarrhea-

associated *Escherichia coli* belong to a family of hemagglutinins with Dr receptor recognition. Infect. Immun. **58:**279–281.
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. **Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143:**161–178.
37. **Saul, D., A. J. Spiers, J. McAnulty, M. G. Gibbs, P. L. Bergquist, and D. F. Hill.** 1989. Nucleotide sequence and replication characteristics of RepFIB, a basic replicon of IncF plasmids. J. Bacteriol. **171:**2697–2707.
38. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
39. **So, M., F. Heffron, and B. J. Mac Carthy.** 1979. The *E. coli* gene encoding heat-stable toxin is a bacterial transposon flanked by inverted repeats of IS*1*. Nature (London) **277:**453–456.
40. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:** 503–517.
41. **Spiers, A. J., N. Bhana, and P. L. Bergquist.** 1993. Regulatory interactions between RepA, an essential replication protein, and the DNA repeats of RepFIB from plasmid P307. J. Bacteriol. **175:** 4016–4024.
42. **Stormo, D. G., T. D. Schneider, and L. M. Gold.** 1982. Characterization of translational initiation sites in *Escherichia coli*. Nucleic Acids Res. **10:**2971–2996.
43. **von Heijne, G.** 1984. How signal sequences maintain cleavage specificity. J. Mol. Biol. **173:**243–251.
44. **Walz, W., M. A. Schmidt, A. F. Labigne-Roussel, S. Falkow, and G. Schoolnick.** 1985. AFA-I, a cloned afimbrial X-type adhesin from a human pyelonephritic *Escherichia coli* strain. Purification and chemical, functional and serological characterization. Eur. J. Biochem. **152:**315–321.
45. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.