Identification of the O Antigen Polymerase (*rfc*) Gene in *Escherichia coli* O4 by Insertional Mutagenesis Using a Nonpolar Chloramphenicol Resistance Cassette

SLAWOMIR LUKOMSKI,* RICHARD A. HULL, AND SHEILA I. HULL

Department of Microbiology and Immunology, Baylor College of Medicine, Houston, Texas 77030

Computer analysis of the O4 polysaccharide gene cluster of *Escherichia coli* revealed the presence of two open reading frames (ORFs) encoding strongly hydrophobic polypeptides. O antigen polymerase, which is encoded by the *rfc* gene, is a potential membrane protein and therefore should be hydrophobic. To identify the *rfc* gene, these two ORFs were subjected to insertional mutagenesis. A chloramphenicol resistance cassette was designed which, when properly inserted, does not cause a polar effect in downstream genes. Each of two ORFs, cloned into a plasmid vector, was inactivated with this cassette. Two types of mutants bearing chromosomal insertions of the cassettes in each ORF were constructed by homologous recombination. These mutants were characterized by PCR, Southern blotting, and transverse-alternating-field electrophoresis. Only one class of mutants exhibited the expected O polymerase-deficient phenotype; they produced O4-specific, semirough lipopolysaccharide. Therefore, this ORF was identified as the *rfc* gene. The chromosomal *rfc* mutation was complemented in *trans* by the *rfc* gene expressed from a plasmid vector.

Lipopolysaccharide (LPS) is the major nonprotein constituent of the outer membrane of gram-negative rod-shaped bacteria (36). It is a complex macromolecule composed of three parts, each of which is synthesized by gene products largely unique to each: lipid A, a complex phospholipid which is embedded in the membrane itself and is responsible for the endotoxic properties of LPS; the core oligosaccharide, composed of an inner and outer core; and the long-chain polysaccharide (PS), also known as the O antigen or O PS, which extends from the surface of the cell and is in contact with the environment (21). Because of its location at the surface of the cell and because of its unique properties, LPS is responsible for many of the surface characteristics of gram-negative bacteria, including resistances to detergents, hydrophobic antibiotics, serum complement, and ingestion by human polymorphonuclear leukocytes (12, 14, 51). These last two properties are dependent on the O PS, and it appears that the lengths of the chains are critical, at least for resistance to the bactericidal effect of complement (12, 40).

The genes responsible for the synthesis of O PS map at the *rfb* locus, at 44 min on the chromosome of *Escherichia coli* (1). The O PS is made by the polymerization of O subunits, composed of three to five sugars, which are built upon a special phospholipid carrier called antigen carrier lipid (43). The next step is the polymerization of the subunits into a long-chain PS by the enzyme called O polymerase, which is encoded by the *rfc* gene. The growing chain is added to the core oligosaccharide, which is made separately from the O PS and linked to the lipid A, by the enzyme called O ligase (29). Both enzymes are thought to be located on the periplasmic face of the cytoplasmic membrane (31). In wild-type strains, the distribution of O-antigen chains is clearly bimodal because of the expression of preferred chain lengths of O antigen on the core-lipid A base (11). Rol protein, which is responsible for this phenom-

enon, has been recently cloned and characterized (2, 3). These three proteins (O polymerase, O ligase, and Rol) seem to play a key role in the regulation of LPS assembly.

Although more than 170 different O antigens have been recognized in *E. coli* (38), only small subset seem to predominate among isolates recovered from infections (22). The O4 serogroup is found among strains which are commonly associated with urinary tract infections and has been frequently isolated from acute pyelonephritis (28, 37). The O4 units consist of pentasaccharides containing *N*-acetylglucosamine, *N*-acetylfucosamine, L-rhamnose, and glucose (1:1:1:2 molar ratio [48]).

The genes responsible for the biosynthesis of the O4 polysaccharide of a human uropathogenic *E. coli* have been cloned and expressed in *E. coli* K-12 (15). Qualitative sugar analyses of LPSs in recombinant strains, as well as boundaries of the complete O4 region, have been established (16). On the basis of available sequence data, a methodology was developed to identify the gene encoding the O4 polymerase. Here, we report experiments on *rfc* identification, cloning, and expression. Structural and functional similarities between *rfc* genes and Rfc proteins of *E. coli*, *Salmonella typhimurium*, and *Shigella flexneri* are discussed.

MATERIALS AND METHODS

Bacterial strains and media. The uropathogenic strain *E. coli* SH1 serotype O4:K6 was used in this study (19). Cloning experiments were performed in *E. coli* XL1-Blue (Stratagene, La Jolla, Calif.); *E. coli* SM10(λpir) was used as a host strain for the suicide vector pGP704 (50); and an *E. coli* SH1 *rfc* derivative was used for the isolation of rough (R) LPS.

Bacteria were routinely grown in Luria broth or Luria agar (Luria broth with 1.5% [wt/vol] agar; Difco Laboratories, Detroit, Mich.) at 37°C, except for conjugation experiments when minimal medium was used (8, 30). Media were supplemented with antibiotics (Sigma, St. Louis, Mo.) at the following concentrations (in micrograms per milliliter): ampicillin, 100; and chloramphenicol, 20 for a chromosomal marker or 50 for a plasmid. Isopropyl-β-D-thiogalactopyranoside (IPTG; Boehringer Mannheim, Indianapolis, Ind.) was added at a concentration of 0.5 mM.

Construction of a nonpolar chloramphenicol-resistant (*cat*) cassette. A *cat* cassette was made based on the concept of the *aph-3* cassette; however, the chloramphenicol resistance (*cat*) instead of kanamycin resistance (*aph-3* [32]) gene was used. PCR was performed with plasmid pACYC184 by using the

* Corresponding author. Mailing address: Department of Microbiology and Immunology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030. Phone: (713) 798-4021. Fax: (713) 798-7375. Electronic mail address: lukomski@bcm.tmc.edu.
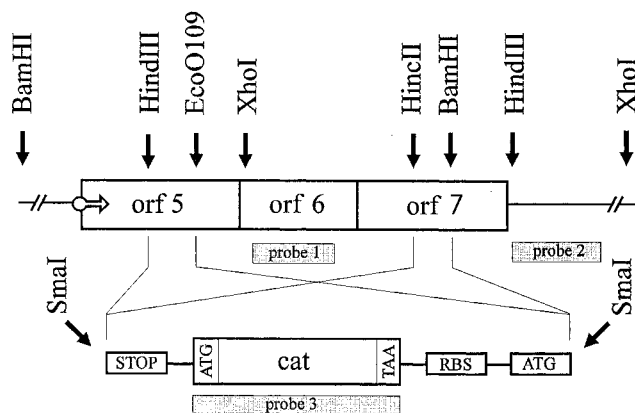
FIG. 1. Schematic representation of the insertions within *orf5* and *orf7* using the *cat* cassette. *orf5* and *orf7*, which are part of the O4 gene cluster, were subjected to insertional mutagenesis by using a promoterless chloramphenicol resistance cassette. The *cat* cassette, flanked by *Sma*I sites, was inserted between the Klenow-treated *Hin*dIII and *Eco*O109 sites of *orf5* or the *Hin*cII and *Bam*HI sites of *orf7*. The *cat* gene, originating from plasmid pACYC184, is preceded by translational stops in all three reading frames (STOP) and is followed by the consensus ribosome-binding site (RBS); downstream translation is restored from a start codon (ATG). This inactivation strategy by insertion of the *cat* cassette ensures against a polar mutation effect in downstream genes. All important restriction sites used for cloning, mapping, and Southern analyses, as well as the DNA probes used, are shown.

following primers: the forward primer was 50 bp long (5′-GGACCCGGGTGACTAACTAGAGGAAGCTAAAA**TGGAGAAAAAAATCACTG**-3′) and contained 19 bp complementary to the 5′ end of the *cat* gene (in boldface); the reverse primer was 44 bp in length (5′-CCTCCCGGGTCATTATCCTTCCAGA**ATTACGCCCCGCCCTGCC**-3′) and contained 17 bp complementary to the 3′ end of the *cat* gene (44).

A 650-bp PCR product was recovered from 1% agarose gel (SeaPlaque; FMC BioProducts, Rockland, Maine) by using β Agarase I (New England Biolabs, Beverly, Mass.), cleaved with *Sma*I, and cloned into the *Sma*I site of pBluescriptII SK, generating plasmid pSL1. The correct sequence of the *cat* cassette flanking regions as well as its IPTG-dependent chloramphenicol resistance (Cm^r) region were confirmed.

Construction of mutant strains SL5SH1 (*orf5*) and SL7SH1 (*orf7*). Two different mutant strains of the O4 parent *E. coli* SH1 were constructed. Clone pGH129 (16) containing a large portion of the O4 cluster was used. The cloning strategy used in this experiment is outlined in Fig. 1.

To obtain the SL5SH1 mutant (insertion within *orf5*), a *Bam*HI-*Bam*HI fragment was first cloned into a pBluescriptII SK derivative lacking part of the multiple cloning site between *Hin*dIII and *Eco*O109. The resulting plasmid, pSL2, was then doubly digested by *Hin*dIII-*Eco*O109, and blunt ended with the Klenow fragment of DNA polymerase I. The *cat* cassette flanked by *Sma*I sites was ligated with pSL2, generating plasmid pSL4. The proper orientation of the cassette was confirmed by restriction digestion. Finally, a *Bam*HI-*Bam*HI fragment of pSL4 was excised and cloned into the *Bgl*II site of the suicide vector pGP704, a derivative of pJM703.1 (33). This plasmid, pSL6, was used in recombination experiments.

To generate the SL7SH1 mutant, plasmid pSL3 was obtained by cloning an *Xho*I-*Xho*I fragment containing *orf7* into a pBluescriptII SK derivative from which a *Hin*cII-*Bam*HI fragment of the multiple cloning sites was first removed. This plasmid was subsequently cut with *Hin*cII-*Bam*HI restriction endonucleases, and after treatment with Klenow, the *cat* cassette was inserted in the proper orientation. The *Xho*I-*Xho*I fragment of the resulting plasmid, pSL5, was recloned into pGP704 at the *Sal*I site, generating the final construct, pSL7.

Mutant strains SL5SH1 and SL7SH1 were obtained by conjugal mating between the wild-type strain of SH1 and *E. coli* SM10 harboring pSL5 or pSL7, respectively. Transconjugants were selected on minimal medium supplemented with chloramphenicol. Mutants resulting from a double-crossover event were initially distinguished from those in which a single recombinational event had occurred by testing for the loss of the Ap^r marker (loss of the suicide plasmid pGP704).

PCR techniques. DNA amplification was performed by using 0.5 μM each primer in 50-μl samples containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl$_2$, 200 μM each deoxynucleoside triphosphate, and 0.5 U of *Taq* polymerase (Boehringer Mannheim). Thirty cycles were performed as follows: 1 min of denaturation at 94°C, 2 min of annealing at 50°C, and 3 min of extension at 72°C. This was followed by one cycle of 2 min at 50°C and 10 min at 72°C.

When PCR was performed directly from bacterial colonies, cells were trans-

ferred on a disposable tip attached to the automatic pipettor and were added to the reaction mixture; otherwise, 1 ng of a DNA template was used.

DNA methods. The DNA methods used are described by Maniatis et al. (30). These include cutting of DNA with restriction endonucleases, DNA ligation and transformation, electrophoresis on agarose gels, and DNA capillary blotting. Plasmid DNA was prepared by an alkaline denaturation protocol according to the method described by Birnboim and Doly (4). Chromosomal DNA was isolated from 1.5-ml cultures as described elsewhere (53); 5 to 6 μg was used for Southern blotting.

Hybridization probes were labeled by the random primer protocol with a nonradioactive system according to the instructions given by the manufacturer (Genius Kit; Boehringer Mannheim). Hybridization and posthybridization were also performed as recommended by the supplier. The reaction was developed with the chemiluminescent substrate (CSPD; Tropix, Bedford, Mass.) and were exposed on a Kodak X-Omat X-ray film for 30 to 45 min.

For DNA sequencing, the dideoxy-chain termination procedure was used (Sequenase 2.0; Amersham, Arlington Heights, Ill.). Oligodeoxynucleotides were synthesized by using reagents and 391 DNA synthesizer from Applied Biosystems, Foster City, Calif. DNA sequence was analyzed with software from Molecular Biology Computational Resources at the Baylor College of Medicine.

LPS isolation and electrophoresis. LPSs were prepared as described by Hitchcock and Brown from 1-ml overnight cultures (18). The LPS samples were electrophoresed in minigels (Mighty Small II SE250; Hoefer Scientific, San Francisco, Calif.) with a Tricine-sodium dodecyl sulfate buffer system (Tricine-SDS); bands were visualized by silver staining (52). Tricine-SDS gels and buffers were made according to the original recipe (47); the molecular weights of the core and O units were estimated with protein standards (Gibco BRL, Gaithersburg, Md.) by a method described previously (27).

Slide agglutination. For slide agglutination with O4 antisera, bacteria were grown overnight on Luria agar plates (with antibiotic added if required); roughly 1-cm^2 of cells was resuspended in 0.85% saline. After boiling, 30-μl aliquots of cells were tested with an equal volume of 10% (vol/vol) Bacto *E. coli* O4 antisera (Difco Laboratories). As a control, cell aliquots were mixed with saline or *E. coli* O75 antisera.

Complementation studies. Expression of the *orf7*-encoded polypeptide was studied by complementation of the *rfc* mutation in SL7SH1-25. Constructs used in these experiments are shown in Table 1. Plasmid vector pLEX5B, as well as other helper plasmids (sources of different promoter and origin cassettes), was described elsewhere (9). *orf7* was amplified by PCR as a blunt-ended fragment at the 5′ end, beginning with the TG dinucleotide of the start codon ATG. The 3′ end was flanked by the *Hin*dIII site located downstream of the stop codon of the gene (Fig. 1). This 5′ TG-*Hin*dIII PCR fragment was cloned between the *Xmn*I-*Hin*dIII sites of the vector; the *Xmn*I-5′ TG junction restored the start codon ATG. Both transcriptional and translational functions were supplied by the vector. A special feature of this cloning system is the easy exchange of the promoters and origins of replication. Four derivatives, pSL20-pSL23, that had either the P$_{lac}$ or the P$_{tac}$ promoter and the ColE1 or R1 origin of replication were constructed. Each derivative was then transformed into the mutant strain SL7SH1-25. Cells were grown overnight in L broth containing chloramphenicol and ampicillin, with or without IPTG. LPS preparation and analysis were carried out as described above.

Nucleotide sequence accession number. The sequence reported here has been entered in GenBank under accession number U39042.

RESULTS

Construction of a nonpolar *cat* cassette. A Km^r cassette, which was similar to the Cm^r cassette constructed for this study, was successfully used in generating mutations in the *ipa* operon of *S. flexneri* (32). It was clearly shown that such a cassette, when properly inserted, did not cause a polar effect in downstream genes. Since wild-type *E. coli* strains are often resistant to kanamycin, chloramphenicol resistance was used instead; the *cat* gene was amplified from plasmid pACYC184 to construct the cassette in pSL1 (Fig. 1). At the 5′ end of the cassette, there are translational stops in all three reading frames; at the 3′ end is a consensus ribosome-binding site followed by the ATG start codon. An appropriate restriction site(s) must be chosen to insert the start codon at the 3′ end of the cassette in frame with the downstream codons of the mutated gene. Such in-frame insertion allows for translation of the remaining 3′ portion of the mutated gene, as well as translation of downstream genes. As expected, clone pSL1 was Cm^r only when IPTG was present in the medium.

Construction of mutant strains SL5SH1 (*orf5*) and SL7SH1 (*orf7*). On the basis of sequence analysis of the O4 region,

TABLE 1. Plasmid vectors and clones used

Plasmid	Relevant characteristics	Reference or source
pBluescriptII SK	High-copy-number cloning vector; Ap^r; $\Delta lacZ$	Stratagene
pGP704	Derivative of suicide plasmid pJM703.1; *oriR6K mob* Ap^r	31
pLEX5B	Expression vector; Ap^r	8
pGH129	O4 *rfb* region (11.7 kb) in pBluescriptII KS	15
pSL1	650-bp *cat* cassette in pBluescriptII SK; Ap^{ra}	This study
pSL2	*Bam*HI-*Bam*HI fragment with *orf5* in pBluescriptII SK; Ap^r	This study
pSL3	*Xho*I-*Xho*I fragment with *orf7* in pBluescriptII SK; Ap^r	This study
pSL4	*cat* cassette inserted in *orf5* of pSL2; Ap^r Cm^r	This study
pSL5	*cat* cassette inserted in *orf7* of pSL3; Ap^{ra}	This study
pSL6	*orf5::cat* cassette of pSL4 in pGP704; Ap^r Cm^r	This study
pSL7	*orf7::cat* cassette of pSL5 in pGP704; Ap^{ra}	This study
pSL20	*orf7* in pLEX5B; *ori* ColE1; P$_{lac}$; Ap^r	This study
pSL21	*orf7* in pLEX5B; *ori* ColE1; P$_{tac}$; Ap^r	This study
pSL22	*orf7* in pLEX5B; *ori*R1; P$_{lac}$; Ap^r	This study
pSL23	*orf7* in pLEX5B; *ori*R1; P$_{tac}$; Ap^r	This study

^{a} Also Cm^r in the presence of IPTG.

several putative ORFs were identified (the complete sequence data of the entire O4 region will be published elsewhere [unpublished data]). Two strongly hydrophobic ORFs (*orf5* and *orf7*) were distinguished, both being good candidates for the *rfc* gene, which encodes polymerase. Both ORFs were expected to be transcribed from a single promoter located upstream of *orf5* (Fig. 1); they are separated by a nonhydrophobic gene, *orf6*.

When the *cat* cassette was inserted between the *Hin*dIII-*Eco*O109 sites of *orf5* in pSL4, this in-frame fusion created a new amino acid, arginine, at the junction that is followed by proline of the native gene. Similarly, insertion between the *Hin*cII-*Bam*HI sites of *orf7* in pSL5 generated an arginine at the junction that is followed by the serine of the parental *orf7*.

pSL4 carrying orf5::*cat* exhibited a Cm^r phenotype independent of the transcription from the vector promoter; this confirmed the presence of the natural promoter upstream of this gene, as deduced from sequence data. In contrast, pSL5 bearing orf7::*cat* exhibited Cm^r only in one orientation, indicating that the Cm^r gene was transcribed from the P$_{lac}$ promoter of the vector; this resistance occurred only when IPTG was present in the medium.

Both mutated ORFs were used in recombination experiments with the parental strain *E. coli* SH1 O4. After recloning of the mutated ORFs into the mobilizable suicide plasmid pGP704, *E. coli* SH1 was conjugated with *E. coli* SM10 harboring each clone separately (pSL5 and pSL7). Five to eight percent of the transconjugants exhibited a Cm^r Ap^s resistance pattern, suggesting that double-recombination events had occurred. Chromosomal mutant strains within *orf5* were called SL5SH1, and those within *orf7* were called SL7SH1. Several individual mutants of SL5SH1 and SL7SH1 were subjected to further analysis.

Analysis of SL5SH1 and SL7SH1 mutants. Whole cells of seven Cm^r Ap^s colonies of SL5SH1 no. 90 to 96 and five of SL7SH1 no. 5 and 22 to 25 were directly used for PCR with primers specific for the appropriate ORF (data not shown). Each SL5SH1 and SL7SH1 mutant produced a single band larger than that present in the parental strain, SH1. This indicated that both *orf5* and *orf7* carried inserts with the expected sizes.

Chromosomal DNAs were isolated from mutant strains and hybridized against the probes specific either for the O4 region (host-specific probes 1 and 2) or the *cat* cassette (insert-specific probe 3; refer to Fig. 1); the results are shown in Fig. 2. Probe 1 hybridized to a 3.2-kb *Bam*HI fragment of the SL5SH1 mu-

tants, compared with a 3.0-kb fragment of the parental strain SH1 (Fig. 2A). Insert-specific probe 3 reacted with the same 3.2-kb fragment of the mutants, while no hybridization could be detected with DNA from the SH1 strain. Similarly, 3.9-kb *Xho*I chromosomal fragments of the SL7SH1 mutants and a 3.4-kb fragment of SH1 hybridized with probe 2, whereas insert-specific probe 3 reacted only with the 3.9-kb bands of DNAs prepared from the *orf7* mutants. These results proved conclusively that in-frame insertions engineered within the cloned ORFs were introduced by homologous recombination into the chromosome of the O4 strain of *E. coli* SH1.

Genomes from the mutant strains were also compared with that from the parental strain SH1 as well as with SM10, the strain which was used as a donor in conjugation, by pulsed-field gel electrophoresis (PFGE [25]). As expected, the *Not*I pattern of the mutants was identical to that seen in the parental strain SH1 but not that characteristic for the SM10 donor strain (data not shown). The PFGE gel was blotted and hybridized against
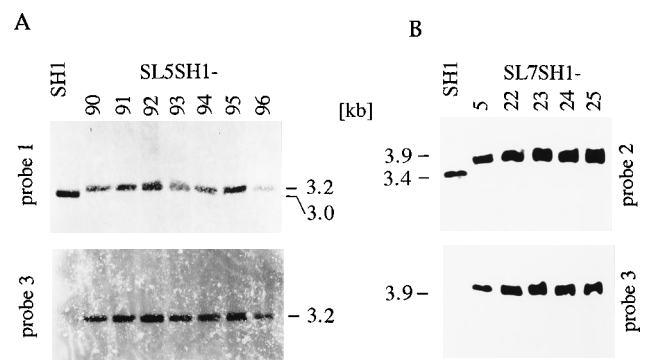

FIG. 2. Analysis of the chromosomal insertions within *orf5* and *orf7* by Southern blotting. (A) Genomic DNAs isolated from the SL5SH1 mutants were digested with *Bam*HI (refer to Fig. 1) and hybridized with either probe 1 (host-specific probe) or probe 3 (insert-specific probe). The parental strain SH1 and SL5SH1 mutants exhibited positive hybridization with probe 1, resulting in patterns of the predicted sizes (3.0 and 3.2 kb, respectively). DNAs from the SL5SH1 mutants, but not from SH1, hybridized with probe 3 specific for the *cat* cassette. (B) DNAs of the SL7SH1 mutants were digested with *Xho*I and probed with probe 2 or probe 3. Host-specific probe 2 hybridized with a 3.4-kb chromosomal fragment of the parental strain SH1 and 3.9-kb bands of the SL7SH1 mutants, whereas the insert-specific probe 3 reacted only with the 3.9-kb fragment of the SL7SH1 mutants. These results confirmed that the *cat* cassette was inserted into the predicted chromosomal loci (*orf5* and *orf7*) of *E. coli* SH1.
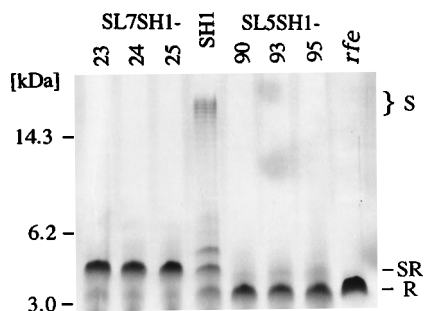
FIG. 3. LPS profiles of various strains in a silver-stained Tricine-SDS polyacrylamide gel. LPSs isolated from representatives of the mutant strains SL5SH1 and SL7SH1 were compared in a 10% Tricine-SDS polyacrylamide gel with the LPS patterns of the parental strain SH1 and its *rfe* derivative. Insertion within *orf5* (SL5SH1 series) resulted in complete loss of O chains (R LPS) comparable to LPS from the *rfe* mutant. In contrast, inactivation of *orf7* (SL7SH1 series) left core substituted with one O unit (SR LPS), the phenotype expected for *rfc* mutants. Only the parental strain, SH1, contained long-chained LPS (S LPS). The running positions of the protein standards (top to the bottom) lysozyme, bovine trypsine inhibitor, and insulin are indicated in kilodaltons.

FIG. 4. In *trans* complementation of the chromosomal Rfc mutation by the cloned *rfc* gene. The Rfc mutant SL7SH1-25 was transformed with four different derivatives: pSL20, pSL21, pSL22, and pSL23. Cultures were grown with (+) or without (−) IPTG; LPS was isolated from each culture and compared with that of the parental strain SH1 in a 10% Tricine-SDS polyacrylamide gel. All clones were able to complement the SR phenotype of SL7SH1-25, although the extent of this complementation depended on the promoter strength and copy number of the vector, as well as the presence (+) or absence (−) of IPTG.

probe 2 or probe 3. The O4 gene cluster was detected on a 230-kb *Not*I band in all strains tested. *cat* cassette inserts were identified on the same 230-kb *Not*I band of the mutants only.

LPS analysis. LPSs of wild-type isolates, such as SH1, exhibit a characteristic bimodal distribution of the O units as detected by electrophoresis in polyacrylamide gels (20, 39). The smooth phenotype (S LPS) is represented by long chains of LPS consisting of many O units. Rough strains (R LPS) have only core oligosaccharide, whereas strains in which the function of O polymerase, the Rfc protein, is impaired are expected to exhibit a semirough type of LPS (SR LPS) that is characterized by an oligosaccharide core substituted with only one O unit.

LPSs from the selected SL5SH1 and SL7SH1 mutants were isolated and compared with the LPS pattern of the parental strain, SH1. To focus on the core region of the LPS, a Tricine-SDS buffer system was employed which provides resolution superior to that by SDS-PAGE (27). It was clearly demonstrated that series mutants SL7SH1 (*orf7*) exhibited the expected SR phenotype (Fig. 3). In contrast, the SL5SH1 series mutants (*orf5*) produced a typical nonsubstituted R LPS. Therefore, it was concluded that *orf7*, but not *orf5*, was the gene for O4 polymerase, *rfc*.

On the basis of the gel migration of the LPS compared with protein standards, the molecular mass of the O4 core was estimated to be approximately 4 kDa, whereas that of single O unit (pentasaccharide) was 0.8 kDa.

Slide agglutination. The conclusions made from the LPS patterns on PAGE were further supported by the agglutination assay. Only the SL7SH1 series mutants were clearly agglutinated by O4 antiserum; the limited precipitation in saline was observed after prolonged incubation, and no reaction was seen in the control samples with heterologous O75 antisera. In contrast, no positive agglutination reaction was detected with SL5SH1 series mutants in the presence of O4 antisera; similarly to the SL7SH1 mutants, they tend to precipitate from saline, and no agglutination took place with O75 control antisera. Therefore, we concluded that only SL7SH1 but not SL5SH1 series mutants still produced O4 antigen.

Identification of Rfc protein. Attempts to overexpress the protein encoded by *orf7*, the putative O polymerase, were unsuccessful. Transcription was always provided by the vect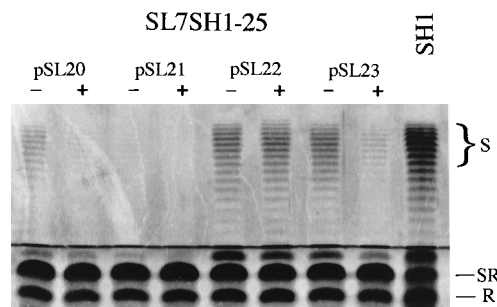or(s). Trace amounts of the fusion proteins were detected on Western blots (immunoblots) with the pMal vector, although the sizes of the extremely faint bands were only slightly larger than that of the MalE protein alone (data not shown). The predicted 45.5-kDa product could not be detected in a cell-free transcription-translation system or in a maxicell strain (46). At the same time, the products of other genes encoded by these plasmids were identified as expected. These negative results were in agreement with previous reports in which the investigators failed to identify the corresponding Rfc proteins from *S. typhimurium* and *S. flexneri* (7, 34).

Complementation studies. To verify that failure to detect Rfc protein was not due to errors introduced in the cloning process, some of the constructs were used in complementation experiments. Orf7 was cloned as a 5′ TG-*Hin*dIII PCR fragment between the *Xmn*I-*Hin*dIII sites of the vector (9). Since the *rfc* gene apparently lacks its own promoter, as well as a sequence resembling a ribosome-binding site, both functions were provided by the vector. Four derivatives, characterized in Table 1, were made; these differed in the origin of replication (high-copy-number ColE1 versus temperature-controlled low- or high-copy-number-R1) and inducible promoters (strong promoter P_{tac} versus the weaker promoter P_{lac}).

All derivatives were introduced into the Rfc mutant strain SL7SH1-25. LPS samples from induced and uninduced cultures were compared with the LPS pattern of the parental strain SH1 on a 10% Tricine-SDS gel (Fig. 4). The copy numbers of the plasmids with the R1 *ori* are small at lower temperatures because they are under the control of the temperature-sensitive lambda repressor. All complementation experiments were performed at 37°C to avoid variations in the LPS patterns. The distribution patterns of the LPS chains of the wild-type strain SH1 are different at 32 and 42°C (data not shown). At 37°C, the copy numbers of clones with R1 *ori* are lower than those with the ColE1 *ori*. Two extreme results were obtained with clones pSL22 and pSL21. The first clone (low copy number and weak promoter) fully complemented the *rfc* mutation with or without IPTG in the medium, whereas the latter (high copy number and strong promoter) exhibited only very poor complementation in the uninduced sample (very faint bands of S LPS seen on the gel and weak SR+1 unit). Although clear complementation by pSL23 (low copy number and strong promoter) was detected in uninduced cells, this complementation was drastically diminished after the transcription was induced. Similar results were obtained with clone pSL20 (high copy number and weak promoter); however, the extent of the

complementation in the absence of IPTG was significantly lower than that with pSL23 and was almost completely removed by inducing the transcription. Complementation was also seen in uninduced cells, which must be due to the low levels of transcription in the absence of induction. Vector alone, with any combination of the promoter-*ori* cassettes, was unable to complement the *rfc* mutation (data not shown). This experiment showed that Rfc function was determined by the product of *orf7*, the *rfc* gene.

Analysis of the *rfc* gene and the putative Rfc protein. The sequence of *orf7*, the identified *rfc* gene, is shown in Fig. 5. It is 1,188 bp long, which corresponds to 395 amino acid residues. *orf7* seems to be transcriptionally coupled with two upstream ORFs (*orf5* and *orf6*); a potential promoter structure was found upstream of *orf5*, i.e., a potential −10 region is TATAAg (five of six bases identical to the consensus TATAAT) and a potential −35 region reads TTGctA (four of six bases as the consensus TTGACA [17, 45]). *orf7* also lacks a sequence resembling the ribosome-binding site which predicts that its translation is dependent on the translation initiation from one of the upstream genes; *orf6* is preceded by a potential ribosome-binding site, TAtcGAGa, at a spacing of 5 nucleotides from the ATG codon (the consensus sequence is TAAGGA GG [10, 49]). Translational coupling between *orf6* and *orf7* is also supported by the fact that the TGA stop codon of the orf6 overlaps with the ATG start codon of the *orf7* (sequence, **ATGA**).

The codon usage for the *rfc* gene is presented in Table 2. Underlined codons, which are characteristic of weakly expressed genes (based on data by Grosjean and Fiers [13]), are frequently used in *orf7*. Leucine, isoleucine, and phenylalanine, which constitute nearly 40% of the total amino acid residues, are mostly encoded by codons typical of weakly expressed genes (19, 49, and 81%, respectively). Modulating codons (marked with asterisks) account for 13.6% of the total *rfc* codon content. The rate codons AGG, AGA, and CGA represent 93% of the arginine codons; correspondingly, GGG and GGA modulating codons are used for 58% of the glycine codons, and AUA is used for 47% of all isoleucine codons. These data suggest that the *rfc* gene of *E. coli* O4 is poorly expressed and explain why the Rfc protein could not be detected.

The predicted molecular mass for this polypeptide is 45.5 kDa, with an isoelectric point of 9.3. Hydrophobic regions are marked in the *rfc* sequence shown in Fig. 5. Overall, the molecule is very hydrophobic, with a mean index of 0.8, as determined by the method described by Kyte and Doolittle (26). This suggests that Rfc is a membrane protein, although this hypothesis needs to be proven experimentally.

DISCUSSION

SR mutants, in which LPS polymerization was defective, were recovered from *S. typhimurium* 30 years previously (35, 54). Spontaneous mutants have been used to identify and characterize the *rfc* loci from both *S. typhimurium* and *S. flexneri* (7, 34). A different approach was undertaken in the present study to identify and clone the O4 polymerase gene from a human uropathogenic strain of *E. coli* SH1. By molecular biology methods, two different isogenic mutant strains, bearing chromosomal insertions within two ORFs of the O4 region, were constructed. On the basis of biochemical analysis, it was predicted that the O-antigen polymerase is located in the cytoplasmic membrane (31). Membrane proteins are expected to be hydrophobic (42); therefore, either hydrophobic ORF could be the *rfc* gene.

FIG. 5. Nucleotide sequence of the 1,188-bp *rfc* gene of *E. coli*. The ATG-Met start codon (nucleotide 16) and TAA stop codon (nucleotide 1201) are marked in boldface. No potential ribosome-binding site can be found upstream of the ATG codon. The amino acid residues encoded by modulating codons or codons characteristic for weakly expressed genes are labeled with asterisks or are underlined, respectively. The hydrophobic stretches representing potential transmembrane domains of the Rfc protein are boxed.

A promoterless, nonpolar chloramphenicol resistance cassette was employed to specifically inactivate two ORFs of the O4 *rfb* region. Insertions were made within *orf5* and *orf7* cloned into plasmid vectors. We also constructed two other derivatives of the *cat* cassette, in which the ATG start codons located at the 3′ end of these cassettes are in different reading frames (data not shown). This allows one of the three cassette derivatives to be used at any predetermined restriction site, which is available within the gene of interest. The native ORFs in the chromosome of *E. coli* SH1 were subsequently replaced with

TABLE 2. Codon usage of the *rfc* gene[a]

Amino acid	Codon	No. of times used	% of total codons
Leu	TTA	24	6.0
	TTG	11	2.8
	CTT	11	2.8
	CTA*	5	1.3
	CTG	4	1.0
	CTC	3	0.8
Ile	ATT	23	5.8
	ATA*	22	5.6
	ATC	2	0.5
Phe	TTT	34	8.6
	TTC	8	2.0
Ser	AGT	10	2.5
	TCA	6	1.5
	TCT	6	1.5
	AGC	5	1.3
	TCC	2	0.5
Tyr	TAT	24	6.0
	TAC	2	0.5
Ala	GCA	9	2.3
	GCT	8	2.0
	GCG	5	1.3
	GCC	3	0.8
Gly	GGT	9	2.3
	GGA*	8	2.0
	GGG*	6	1.5
	GGC	1	0.3
Lys	AAA	10	2.5
	AAG	7	1.8
Val	GTT	10	2.5
	GTA	7	1.8
Arg	AGA*	5	1.3
	AGG*	5	1.3
	CGA*	3	0.8
	CGC	1	0.3
Thr	ACT	6	1.5
	ACA	4	1.0
	ACC	3	0.8
	ACG	1	0.3
Asn	AAT	11	2.8
	AAC	2	0.5
Met	ATG	12	3.0
Cys	TGT	8	2.0
	TGC	3	0.8
Glu	GAA	7	1.8
	GAG	4	1.0
Asp	GAT	6	1.5
	GAC	3	0.8
Gln	CAG	6	1.5
	CAA	3	0.8
Pro	CCA	6	1.5
	CCT	2	0.5
Trp	TGG	3	0.8

[a] Amino acids and codons which are not present within the *rfc* gene are not included. Asterisks, modulating codons; underlining, codons characteristic for weakly expressed genes.

the mutated ones by homologous recombination. The genomic *Not*I patterns of the mutants and parental strain were identical, ensuring that no additional gene rearrangement in the chromosome had occurred. However, we did detect large chromosomal deletions (30 to 50 kb in size) in some mutants. These mutants were not used in further experiments; no attempts have been made to find which genes have been eliminated. It has been reported that genes involved in the pathogenesis process are grouped in so-called pathogenicity islands (24); a large deletion such as that reported here could easily accommodate a gene cluster determining biologically important features. Mutant strains, which are not fully characterized, could give false results in a variety of biological tests.

The method used in this study has the following advantages: (i) insertion of the cassette does not cause polar effects in downstream genes; therefore, any gene can be inactivated inside the gene cluster, regardless of its transcriptional-translational organization; (ii) all steps are easily monitored by standard protocols, e.g., restriction analysis, PCR, and Southern blotting; (iii) in contrast to spontaneous mutations, the insertions are strictly defined regarding both their site and molecular basis; and (iv) the mutant strain obtained is isogenic with the parental one and, therefore, can be further used in biological studies such as virulence assays.

rfc mutants produce LPS in which at most one O-antigenic oligosaccharide unit is attached to the core unit (29). Only SL7SH1 mutants were still agglutinated by the O4-specific antiserum, indicating that mutants of the SL5SH1 series did not produce O chains at all. PAGE analysis of the LPS samples isolated from both types of mutants provided additional evidence that *orf7* was the O polymerase. *orf7* was identified as the *rfc* gene because the SL7SH1 mutants exhibited an SR LPS pattern; similar results were shown previously for the *rfc* spontaneous mutants of *S. typhimurium* and *S. flexneri* (7, 34). LPSs from the mutants of the SL5SH1 series migrated as a single band corresponding to the nonsubstituted core. This indicated that inactivation of *orf5* precludes either synthesis or transfer of O4 side chains to the LPS core. Amino acid sequence data revealed the existence of a putative ATP binding motif at the C terminus of the polypeptide encoded by *orf5*. If these data are accurate, this polypeptide may be involved in the transport of the O antigen across the cytoplasmic membrane and would be analogous to proteins identified in the *E. coli* O9 and *Klebsiella pneumoniae* O1 *rfb* gene clusters (5, 23).

A Tricine-SDS buffer system was employed instead of glycine-SDS. This buffer system resolves core LPSs of *E. coli rfa* mutants with very subtle differences in size (41). The molecular mass of the core unit was estimated to be about 4 kDa; by this method, a similar value of 4.2 kDa has been calculated for the core LPS from a rough mutant of *Salmonella minnesota* (27). The value for a single O unit, 0.8 kDa, is somewhat less than the theoretical molecular weight 952, based on chemical composition (48).

The *rfc* gene is located inside the *rfb* region. It lacks its own promoter; we predict that it is transcriptionally coupled with two preceding ORFs (*orf5* and *orf6*). The experimental support for this assumption comes from the fact that clone pSL5 was Cm^r only when it was transcribed from the vector promoter, whereas pSL4 was Cm^r also when it was cloned in the direction opposite to the vector promoter. *rfc* also lacks translational signals; it may utilize the translation initiation at the start of *orf6*, although it has not been confirmed experimentally. Similar data have been reported for the *rfc* gene of *S. flexneri* regarding its transcriptional-translational organization (34). In contrast, the *rfc* gene of *S. typhimurium* maps outside the *rfb* gene cluster; it also has regulatory sequences necessary for its expression (7).

The *rfc* genes has been cloned from the *E. coli* O4, *S. typhimurium*, and *S. flexneri*. All attempts failed to identify the Rfc protein from the above species. The frequencies of the modulating codons for *rfc* from these bacteria are similarly high: 13.6, 12.5, and 12.0%, respectively. The naturally low levels of the *rfc* gene products seem to be tightly controlled by the atypical codon usage, since their isoaccepting tRNAs are very rare (13). Eight of the first 25 amino acids of the *E. coli rfc* gene are encoded by rare codons (including two modulating argi-

nine codons, AGG/CGA) which may cause ribosome stalling and, as a consequence, translational pausing at an initiation site (6). The molecular mass of the putative Rfc protein is 45.5 kDa, compared with a 47.5-kDa Rfc product in *S. typhimurium* and a 43.7-kDa product in *S. flexneri*. All Rfc proteins are hydrophobic (mean indices of 0.8, 0.7, and 1.08, respectively), with several potential transmembrane domains. Experiments with Tn*5phoA* and Tn*5lacZ* fusions are being carried out. Despite the functional similarity of these proteins, the overall sequences are different.

Different clones of the *rfc* gene (*orf7*) were made. They were able to complement in *trans* the Rfc chromosomal mutation in the SL7SH1-25 mutant. However, in most cases, complementation did not fully restore the LPS pattern typical of the parental strain SH1, which is characterized by a bimodal distribution of the LPS chains. Similar results have also been reported earlier for *S. typhimurium* and *S. flexneri* (7, 34). However, these *Salmonella* and *Shigella* clones were often on different vectors and also contained different flanking sequences. To investigate this phenomenon, four derivatives of the *rfc* clone were constructed. They differed only in copy number or the level of transcription; the size of the insert was identical. The best complementation was achieved when low expression levels were obtained by a low-copy-number vector and a weak promoter on the vector. These results are contrary to the previous interpretation of Morona et al., who suggested that poor complementation resulted from weaker translation of some *rfc* clones in *S. flexneri* (34). If expression of this protein is poor because of limiting pools of rare tRNAs, increased expression from a cloned gene may be detrimental. Our clone derivative, characterized by high copy number and strong transcription, did not fully complement the mutation, especially after induction with IPTG. One possible explanation can be proposed. When high amounts of the mRNA are transcribed from the clone, a limited pool of rare tRNAs may be quickly exhausted, before the complete Rfc protein is translated. In contrast, the poorly transcribed, low-copy-number clone would allow translation to be completed. However, other explanations, e.g., Rfc aggregation or proteolysis in the cytoplasm, cannot be ruled out. Since this protein has never been expressed at a detectable level, lack of complementation as a result of Rfc overproduction is very unlikely.

Because long-chained LPS is so important in determining the surface properties of gram-negative bacteria, which include many human pathogens, an understanding of the final steps in the assembly and transport of LPS is crucial to a complete understanding of how these bacteria interact with their environments.

REFERENCES

1. **Bachmann, B. J.** 1990. Linkage map of *Escherichia coli* K-12, edition 8. Microbiol. Rev. **54:**130–197.
2. **Batchelor, R. A., P. Alifano, E. Biffali, S. I. Hull, and R. A. Hull.** 1992. Nucleotide sequence of the genes regulating O-polysaccharide antigen chain length (*rol*) from *Escherichia coli* and *Salmonella typhimurium*: protein homology and functional complementation. J. Bacteriol. **174:**5228–5236.
3. **Batchelor, R. A., G. E. Haraguchi, R. A. Hull, and S. I. Hull.** 1991. Regulation by a novel protein of the bimodal distribution of lipopolysaccharide in the outer membrane of *Escherichia coli*. J. Bacteriol. **173:**5699–5704.
4. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.
5. **Bronner, D., B. R. Clarke, and C. Whitfield.** 1994. Identification of an ATP-binding cassette transport system required for translocation of lipopolysaccharide O-antigen side-chains across the cytoplasmic membrane of *Klebsiella pneumoniae* serotype O1. Mol. Microbiol. **14:**505–519.
6. **Chen, G.-F. T., and M. Inouye.** 1990. Suppression of the negative effect of minor arginine codons on gene suppression; preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. Nucleic Acids Res. **18:**1465–1473.
7. **Collins, L. V., and J. Hackett.** 1991. Molecular cloning, characterization, and nucleotide sequence of the *rfc* gene, which encodes an O-antigen polymerase of *Salmonella typhimurium*. J. Bacteriol. **173:**2521–2529.
8. **Davis, B. D., and E. S. Mingioli.** 1950. Mutants of *Escherichia coli* requiring methionine or vitamin B12. J. Bacteriol. **60:**17–28.
9. **Diederich, L., A. Roth, and W. Messer.** 1994. A versatile plasmid vector system for the regulated expression of genes in *Escherichia coli*. BioTechniques **16:**916–923.
10. **Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. W. Singer, and G. Stormo.** 1981. Translational initiation in prokaryotes. Annu. Rev. Microbiol. **35:**365–403.
11. **Goldman, R. C., and F. Hunt.** 1990. Mechanism of O-antigen distribution in lipopolysaccharide. J. Bacteriol. **172:**5352–5359.
12. **Goldman, R. C., K. A. Joiner, and L. Leive.** 1984. Serum-resistant mutants of *Escherichia coli* O111 contain increased lipopolysaccharide, lack an O antigen-containing capsule, and cover more of their lipid A core with O antigen. J. Bacteriol. **159:**877–882.
13. **Grosjean, H., and W. Fiers.** 1982. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene **18:**199–209.
14. **Grossman, N., A. A. Lindberg, S. B. Svenson, K. A. Joiner, and L. Leive.** 1991. Structural aspects of LPS: role in evasion of host defense mechanism, p. 143–149. *In* E. Z. Ron and S. Rotten (ed.), Microbial surface components and toxins in relation to pathogenesis. Plenum Press, New York.
15. **Haraguchi, G. E., R. A. Hull, U. Krallmann-Wenzel, and S. I. Hull.** 1989. Molecular cloning and expression of the O4 polysaccharide gene cluster from *Escherichia coli*. Microb. Pathog. **6:**123–132.
16. **Haraguchi, G. E., U. Zahringer, B. Jann, K. Jann, R. A. Hull, and S. I. Hull.** 1991. Genetic characterization of the O4 polysaccharide gene cluster from *Escherichia coli*. Microb. Pathog. **10:**351–361.
17. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of Escherichia coli promoter DNA sequences. Nucleic Acids Res. **11:**2237–2255.
18. **Hitchcock, P. J., and T. M. Brown.** 1983. Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. J. Bacteriol. **154:**269–277.
19. **Hull, S. I., R. A. Hull, B. H. Minshew, and S. Falkow.** 1982. Genetics of hemolysin of *Escherichia coli*. J. Bacteriol. **151:**1006–1012.
20. **Jann, B., K. Reske, and K. Jann.** 1975. Heterogeneity of lipopolysaccharides. Analysis of chain length by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Eur. J. Biochem. **60:**239–246.
21. **Jann, K., and B. Jann.** 1987. Polysaccharide antigens of *Escherichia coli*. Rev. Infect. Dis. **9:**S517–526.
22. **Johnson, J. R.** 1991. Virulence factors in *Escherichia coli* urinary tract infection. Clin. Microbiol. Rev. **4:**80–128.
23. **Kido, N., V. I. Torgov, T. Sugiyama, K. Uchiya, H. Sugihara, N. Kato, and K. Jann.** 1995. Expression of the O9 polysaccharide of *Escherichia coli*: sequencing of the *E. coli* O9 *rfb* gene cluster, characterization of mannosyl transferases, and evidence for an ATP-binding cassette transport system. J. Bacteriol. **177:**2178–2187.
24. **Knapp, S., J. Hacker, T. Jarchau, and W. Goebel.** 1986. Large, unstable inserts in the chromosome affect virulence properties of uropathogenic *Escherichia coli* strain 536. J. Bacteriol. **168:**22–30.
25. **Koob, M., and W. Szybalski.** 1992. Preparing and using agarose microbeads, p. 13–20. *In* J. N. Abelson and M. I. Simon (ed.), Methods Enzymol. Academic Press, San Diego, Calif.
26. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
27. **Lesse, A. J., A. A. Campagnari, W. E. Bittner, and M. A. Apicella.** 1990. Increased resolution of lipopolysaccharides and lipooligosaccharides utilizing tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis. J. Immun. Methods **126:**109–117.
28. **Mabeck, C. E., F. Ørskov, and I. Ørskov.** 1971. *Escherichia coli* serotypes and renal involvement in urinary-tract infection. Lancet **i:**1312–1314.
29. **Mäkelä, P. H., and B. A. D. Stocker.** 1984. Genetics of lipopolysaccharide, p. 59–137. *In* E. T. Rietschel (ed.) Handbook of endotoxin, vol. 1. Elsevier Science Publishing, Amsterdam.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **McGrath, B. C., and M. J. Osborn.** 1991. Localization of the terminal steps of O-antigen synthesis in *Salmonella typhimurium*. J. Bacteriol. **173:**649–654.
32. **Menard, R., P. J. Sansonetti, and C. Parsot.** 1993. Nonpolar mutagenesis of the *ipa* genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. J. Bacteriol. **175:**5899–5906.

33. **Miller, U. L., and J. J. Mekalanos.** 1988. A novel suicide vector and its use in construction of insertional mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. J. Bacteriol. **170:**2575–2583.

34. **Morona, R., M. Mavris, A. Fallarino, and P. A. Manning.** 1994. Characterization of the *rfc* region of *Shigella flexneri*. J. Bacteriol. **176:**733–747.

35. **Naide, Y., H. Nikaido, P. H. Mäkelä, R. G. Wilkinson, and B. A. D. Stocker.** 1965. Semirough strains of *Salmonella*. Proc. Natl. Acad. Sci. USA **53:**147–153.

36. **Nikaido, H., and M. Vaara.** 1985. Molecular basis of bacterial outer membrane permeability. Microbiol. Rev. **49:**1–32.

37. **Ørskov, I., F. Ørskov, A. Birch-Andersen, M. Kanamori, and C. Svanborg-Eden.** 1982. O, K, H and fimbrial antigens in *Escherichia coli* serotypes associated with pyelonephritis and cystitis. Scand. J. Infect. Dis. **33:**S18–25.

38. **Ørskov, I., F. Ørskov, B. Jann, and K. Jann.** 1977. Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. Bacteriol. Rev. **41:**667–710.

39. **Palva, E. T., and P. H. Mäkelä.** 1980. Lipopolysaccharide heterogeneity in *Salmonella typhimurium* analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Eur. J. Biochem. **107:**137–143.

40. **Porat, R., R. Mosseri, E. Kaplan, M. A. Johns, and S. Shibolet.** 1992. Distribution of polysaccharide side chains of lipopolysaccharide determine resistance of *Escherichia coli* to the bactericidal activity of serum. J. Infect. Dis. **165:**953–956.

41. **Pradel, E., and C. A. Schnaitman.** 1991. Effect of *rfaH* (*sfrB*) and temperature on expression of *rfa* genes of *Escherichia coli* K-12. J. Bacteriol. **173:**6428–6431.

42. **Rees, D. C., L. DeAntonio, and D. Eisenberg.** 1989. Hydrophobic organization of membrane proteins. Science **245:**510–513.

43. **Rick, P. D.** 1987. Lipopolysaccharide biosynthesis, p. 648–662. *In* F. C. Neidhard, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 1. American Society for Microbiology, Washington, D.C.

44. **Rose, R. E.** 1988. The nucleotide sequence of pACYC184. Nucleic Acids Res. **16:**355.

45. **Rosenberg, D. E., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13:** 319–353.

46. **Sancar, A., A. M. Hack, and W. D. Rupp.** 1979. Simple method for identification of plasmid-coded proteins. J. Bacteriol. **137:**692–693.

47. **Schagger, H., and G. von Jagow.** 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. **166:**368–379.

48. **Schmidt, M. A., B. Jann, and K. Jann.** 1983. Cell-wall lipopolysaccharide of the urinary-tract-infective *Escherichia coli* O4:K12:H⁻. Structure of the polysaccharide chain. Eur. J. Biochem. **137:**163–171.

49. **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.

50. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. Biotechnology **1:**784–791.

51. **Svanborg-Edén, C., L. Hagberg, R. Hull, S. Hull, K.-E. Magnusson, and L. Öhman.** 1987. Bacterial virulence host resistance in the urinary tracts of mice. Infect. Immun. **55:**1224–1232.

52. **Tsai, C. M., and C. F. Frasch.** 1982. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. Anal. Biochem. **119:**115–119.

53. **Wilson, K.** 1988. Preparation of genomic DNA from bacteria, p. 2.4.1–2.4.2. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Wiley Interscience, New York.

54. **Yuasa, R., K. Nakane, and H. Nikaido.** 1970. Structure of cell wall lipopolysaccharide from *Salmonella typhimurium*: structure of lipopolysaccharide from a semirough mutant. Eur. J. Biochem. **15:**63–71.