JOURNAL OF BACTERIOLOGY, Nov. 1992, p. 7297–7307
0021-9193/92/227297-11$02.00/0

Vol. 174, No. 22

Role of *Escherichia coli* K-12 *rfa* Genes and the *rfp* Gene of *Shigella dysenteriae* 1 in Generation of Lipopolysaccharide Core Heterogeneity and Attachment of O Antigen

JOHN D. KLENA, ROWELL S. ASHFORD II, AND CARL A. SCHNAITMAN*
Department of Microbiology, Arizona State University, Tempe, Arizona 85287-2701

Received 10 July 1992/Accepted 13 September 1992

The *rfp* gene of *Shigella dysenteriae* 1 and the *rfa* genes of *Escherichia coli* K-12 and *Salmonella typhimurium* LT2 have been studied to determine their relationship to lipopolysaccharide (LPS) core heterogeneity and their role in the attachment of O antigen to LPS. It has been inferred from the nucleotide sequence that the *rfp* gene encodes a protein of 41,864 Da which has a structure similar to that of RfaG protein. Expression of this gene in *E. coli* K-12 results in the loss of one of the three bands seen in gel analysis of the LPS and in the appearance of a new, more slowly migrating band. This is consistent with the hypothesis that Rfp is a sugar transferase which modifies a subset of core molecules so that they become substrates for attachment of *S. dysenteriae* O antigen. A shift in gel migration of the bands carrying *S. dysenteriae* O antigen and disappearance of the Rfp-modified band in strains producing O antigen suggest that the core may be trimmed or modified further before attachment of O antigen. Mutation of *rfaL* results in a loss of the rough LPS band which appears to be modified by Rfp and prevents the appearance of the Rfp-modified band. Thus, RfaL protein is involved in core modification and is more than just a component of the O-antigen ligase. The products of *rfaK* and *rfaQ* also appear to be involved in modification of the core prior to attachment of O antigen, and the sites of *rfaK* modification are different in *E. coli* K-12 and *S. typhimurium*. In contrast, mutations in *rfaS* and *rfaZ* result in changes in the LPS core but do not affect the attachment of O antigen. We propose that these genes are involved in an alternative pathway for the synthesis of rough LPS species which are similar to lipooligosaccharides of other species and which are not substrates for O-antigen attachment. All of these studies indicate that the apparent heterogeneity of *E. coli* K-12 LPS observed on gels is not an artifact but instead a reflection of functional differences among LPS species.

Figure 1 shows the *rfa* cluster at 81 min on the *Escherichia coli* K-12 map which contains the genes involved in the synthesis of the lipopolysaccharide (LPS) core and the attachment of O antigen to the core (40). The lower part of Fig. 1 is a simplified diagram of the LPS core structure indicating the sites at which some of these genes are thought to act (25).

The *rfa* cluster can be divided into three regions. At the right end is a two-gene operon which includes the *kdtA* gene and the gene for an 18-kDa protein. Both of these genes appear to encode essential functions (38). *kdtA* encodes a bifunctional enzyme which adds the first two 3-deoxy-D-*manno*-2-octulosonic acid (KDO) residues to lipid IVa, a precursor of lipid A (7). The function of the 18-kDa protein is not known.

Near the left end of the cluster is a four-gene operon which is transcribed in a rightward direction (38, 40). Three of the genes in this operon are involved in heptose synthesis and attachment: *rfaD*, which encodes a heptose isomerase, and *rfaC* and *rfaF*, which are required, respectively, for the addition of the heptose I and heptose II residues of the inner core. The fourth gene in this operon is *rfaL*, which functions in the synthesis of LPS containing O antigen. *rfaL* mutants of *Salmonella typhimurium* produce rough LPS with a core which appears to be complete on the basis of chemical composition, serological properties, and ability to act as a bacteriophage receptor. *rfaL* mutants also produce detectable amounts of O-antigenic polysaccharide bound to the antigen carrier lipid (ACL) on which it is assembled (25). On the basis of these observations, *rfaL* has been suggested to encode a component of the O ligase which transfers the completed O antigen from the ACL to the core of a suitable LPS acceptor (25, 36). The gene for a 26-kDa protein of unknown function is transcribed divergently from the *rfaD-FCL* operon (34). Insertions into this gene exhibit a detergent-sensitive phenotype, indicating that it is also part of the *rfa* cluster (33).

In the middle of the *rfa* cluster is a block of 10 contiguous genes, *rfaQ* through *rfaK*, which are involved in synthesis of the hexose region of the core and modification of the inner core region (20, 29–31). These are arranged in a complex operon (1, 38) whose expression requires the positive regulatory protein SfrB (RfaH) (25, 32). The *rfaG*, *-P*, *-B*, *-I*, *-J*, and *-K* genes and their probable functions were previously known from studies of rough or core-defective mutants of *Salmonella* spp. (25), while *rfaQ*, *-S*, *-Y*, and *-Z* are genes which have been identified recently from the sequence of the *rfa* clusters of *S. typhimurium* LT2 and *E. coli* K-12 (20, 24, 30, 31). The exact functions of these new genes are not known.

Mutations which specifically inactivate *rfaY* have not yet been isolated, but complementation studies of polar deletion mutations indicate that *rfaY* may be necessary for the action of *rfaJ* (31). Since *rfaY* is adjacent to *rfaJ* on the chromosome and mutations in either gene could have similar phenotypes, it is possible that some mutations in *S. typhimurium* which have been designated as being in *rfaJ* are actually in *rfaY*. This may explain why *rfaY* was not detected prior to the sequencing of the *rfa* gene cluster. The combined action

* Corresponding author.

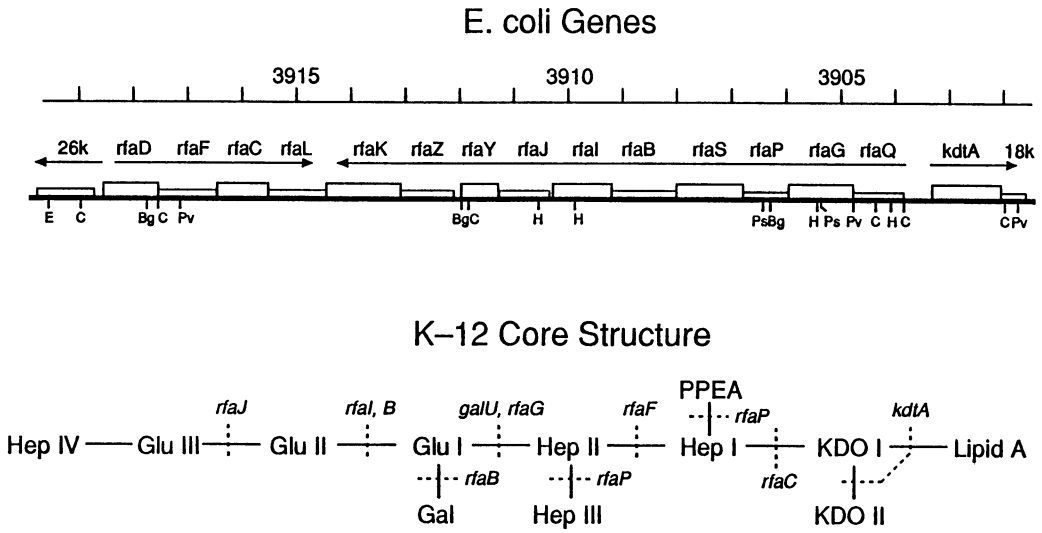

FIG. 1. Structure of the *rfa* gene cluster and a simplified structure of the LPS core from *E. coli* K-12. The top portion shows a partial restriction map of the *rfa* gene cluster at 82 min (3), with all of the known genes drawn to scale. The arrows indicate the directions of transcription of the genes, and the scale at the top shows the coordinates (in kilobases) of the restriction map described by Kohara et al. (21). Restriction sites are as follows: E, *Eco*RI; Bg, *Bgl*II; C, *Cla*I; H, *Hin*dIII; Ps, *Pst*I; and Pv, *Pvu*II. The lower portion shows a partial structure of the core (13, 17, 36) and the sites of action of various genes (25, 29, 31) and includes only those structures and gene functions which are reasonably well established. The substitution of Hep II by Hep III (29) and the substitution of Glu III by Hep IV (15) are present only in some of the molecules, and in some molecules the pyrophosphorylethanolamine (PPEA) residue is replaced by phosphate (29). Additional core substituents which may be present on some molecules (36) include phosphorylethanolamine (PEA) attached to a KDO residue (14), a third KDO residue attached to KDO II, and GlcNAc and/or an additional molecule of PPEA whose site or sites of attachment are not known.

of *rfaJ* and *rfaY* results in a significant increase in the size and complexity of the core (31), which suggests that *rfaY* may be involved in adding a modification or decoration. A third KDO residue or an additional phosphate-containing substituent might be reasonable candidates, since these are thought to be added late in core completion (36).

Knockout mutations in *rfaQ*, *-S*, and *-Z* have very little effect on the sodium dodecyl sulfate (SDS) gel profile of *E. coli* K-12 LPS, and this has made their role in core synthesis difficult to establish. One of the observations reported here is that the phenotypes of mutations in these genes are much more dramatic in *E. coli* K-12 strains carrying a plasmid which allows production of an O antigen.

E. coli K-12 does not produce an O antigen. The reason for this is unknown. It is thought that *E. coli* K-12 lacks a complete *rfb* gene cluster, but the complete sequence of the *his-rfb-cps* region of the *E. coli* K-12 chromosome which would provide evidence for this is not yet available. *rfb* gene clusters from other wild strains of *E. coli* have been introduced into *E. coli* K-12 by transformation with plasmids carrying cloned *rfb* clusters (4, 11, 44), and in some (although not all) cases this has led to efficient expression of an O antigen. This indicates that there are no other obvious genetic defects at the *rfa* locus or elsewhere in the *E. coli* K-12 chromosome which would prevent O-antigen expression. Some strains of *E. coli*, such as the O111 strain described by Riley et al. (37), carry their *rfb* genes on plasmids, and these strains resemble *E. coli* K-12 when they lose their plasmids. It is possible that *E. coli* K-12 carried such a plasmid when it was first isolated.

In order to study the requirements for O-antigen synthesis in *E. coli* K-12, we used plasmids carrying the *rfp* gene and the *rfb* gene cluster from *Shigella dysenteriae* 1 (41–43, 46). The O antigen from this organism has a number of advantages. It is efficiently expressed in *E. coli* K-12, and O-specific antiserum which can be used to monitor this expression is readily available from commercial sources. The *S. dysenteriae* 1 O unit has a chemical structure which makes it ideal for both chemical and molecular studies. The O unit contains one molecule of galactose (Gal), two molecules of rhamnose (Rha), and one molecule of *N*-acetylglucosamine (GlcNAc). Since the genes for synthesis of Gal and GlcNAc and their nucleotide derivatives are located elsewhere on the chromosome, the only sugar biosynthetic genes which must be encoded by the *S. dysenteriae* 1 *rfb* cluster are those required for Rha. As a result, the *rfb* fragment from *S. dysenteriae* 1 which is required for O-antigen expression in *E. coli* K-12 is small (less than 9 kb).

The most interesting aspect of this O-antigen system is that an extra gene is required to modify the LPS core prior to O-antigen attachment. All smooth isolates of *S. dysenteriae* 1 carry a 9-kb multicopy plasmid which is essential for the production of LPS containing an O antigen (10, 46). Although this plasmid is large enough to encode several genes, only a 2-kb region of the plasmid carrying a single gene (termed *rfp*) was necessary for O-antigen expression (45). When a recombinant plasmid carrying the subcloned *rfp* gene was introduced into *E. coli* K-12, there was an increase in the Gal content of the core, and this was accompanied by a shift in gel migration of a portion of the LPS to a higher apparent molecular weight. On this basis, it was proposed that *rfp* might encode a galactosyltransferase which modified the core (41). Although these experiments were done before the development of the high-resolution SDS-tricine gel system which allows direct visualization of LPS heterogeneity, gel analysis suggested that the core might be heterogeneous and that only a subset of core molecules was modified by *rfp*. This has now been confirmed and has important implications with respect to the biological and biochemical roles of core heterogeneity in *E. coli* K-12. In this study, we describe the

TABLE 1. Bacterial strains

Strain	Relevant genotype or description	Source or reference
E. coli K-12 *rfa*$^+$		
CS180	AB1133 derivative	1
CS1834	*recBC sbcBC*; JC7623 derivative	32
CS1861	CS180 transformed with pSS37	This study
CS1999	Δ*lac* of CS180; Mu^r	32
CS2132	CS180 transformed with pSS3	This study
CS2363	CS180 transformed with pJK2363	This study
CS2767	Same as CS1999 except Mu^s	This study
E. coli K-12 *rfa* mutants		
CS2334	*rfaL1*::Tn*phoA* of CS180	This study
CS2452	*rfaZ21*::Tn*lac* of CS1999	This study
CS2456	CS2452 transformed with pSS3	This study
CS2457	CS2452 transformed with pSS37	This study
CS2476	CS2334 transformed with pSS37	This study
CS2529	*rfaK2*::ΩKm^r of CS180	This study
CS2682	*rfaQ9*::Tn*lac* of CS1999	32
CS2683	*rfaS2007*::Tn*lac* of CS1999	32
CS2684	CS2682 transformed with pSS37	This study
CS2685	CS2683 transformed with pSS37	This study
CS2686	CS2682 transformed with pJK2363	This study
CS2687	CS2383 transformed with pJK2363	This study
CS2774	*rfaQ9*::Tn*lac* of CS2767	This study
CS2775	*rfaS2007*::Tn*lac* of CS2767	This study
CS2776	*rfaZ21*::Tn*lac* of CS2767	This study
S. typhimurium LT2		
SA1377	*rfaC630*	SGSC[a]
SL733	*rfaK953*	SGSC
SL3749	*rfaL446*	SGSC
SL3770	*rfa*$^+$ *rfb*$^+$	SGSC
CS2272	SL733 transformed with pJK2252	This study

[a] Salmonella Genetic Stock Center.

TABLE 2. Plasmids

Plasmid	Properties or construction	Source
pSS3	pACYC184 carrying 4.7-kb *rfp*$^+$ fragment cloned from small *S. dysenteriae* 1 plasmid pHW400 (41)	K. Timmis
pSS37	pACYC184 carrying *rfp*$^+$ fragment and cloned chromosomal *rfb*$^+$ gene cluster from *S. dysenteriae* 1 (43)	K. Timmis
pKZ38	pBR322 carrying 4.3-kb *rfaKLZ*$^+$ *Hin*dIII fragment from *S. typhimurium* LT2 (24)	P. Maclachlan
pJK2252	pGEM4 carrying 5.5-kb *rfaFCLKZ*$^+$ *Bgl*II fragment from *E. coli* K-12 (20)	J. Klena
pJK2299	Derivative of pJK2252 made by deleting 3.0-kb *Sma*I-*Eco*RV fragment carrying genes *rfaFCL*	This study
pJK2312	Derivative of pJK2252 with Tn*phoA* insertion in *rfaL*	This study
pJK2363	pGEM4 carrying 2.5-kb *rfp*$^+$ *Cla*I-*Bgl*II fragment from pSS3	This study
pJK2364	Derivative of pJK2299 made by inserting ΩKm^r (9) into *Bcl*I site in *rfaK* (20)	This study

sequencing of the *rfp* gene and present some details regarding the structure of the Rfp protein and its function in the attachment of *S. dysenteriae* 1 O antigen.

We have also studied the role of the genes *rfaK* and *rfaL* in core completion and O-antigen attachment. We recently reported that the sequences of the RfaL and RfaK proteins are not conserved between *E. coli* K-12 and *S. typhimurium* LT2, even though the sequences of the proteins encoded by the flanking genes are strongly conserved (20). We present evidence that both *rfaK* and *rfaL* are involved in core modification and that the polymorphism of these genes is involved in a species-specific attachment of O antigen to the core. In *E. coli* K-12, both *rfaK* and *rfaL* appear to play an important role in generating core heterogeneity as well as attachment of O antigen.

MATERIALS AND METHODS

Bacteria, plasmids, and phages. The bacterial strains used are listed in Table 1, and the plasmids and the cloned genes which they carry are listed in Table 2. The *S. typhimurium* O-specific phage P-22 and the rough-specific phages Felix O-1 (FO), Br2, and C21 were obtained from K. Sanderson; to allow DNA modification, they were propagated on sensitive strains of either *S. typhimurium* or *E. coli* K-12 before use. Phage sensitivity was determined by spotting dilutions of phage onto a Luria-Bertani broth agar plate (27) spread with a growing culture of the test bacteria. All cultures were grown in Luria-Bertani broth, and unless otherwise noted, cultures for LPS analysis were grown at 30°C.

Mutations were constructed on plasmids in vitro and crossed onto the chromosome by the transformation of a *recBC sbcBC* strain as previously described (29, 32). Electroporation was used to facilitate movement of plasmids between *S. typhimurium* and *E. coli* K-12 (30). DNA sequencing and protein sequence analyses were as described previously (30).

Construction of specific mutations. The *rfaL1*::Tn*phoA* mutation was constructed by Tn*phoA* mutagenesis (26) of a strain carrying plasmid pJK2252 (Table 2) followed by selection for high-level kanamycin resistance (Km^r) and screening for enzymatically active (blue) colonies. This mutation was located within the *rfaL* coding region by restriction mapping, and the transposase was removed by cleavage of the plasmid with *Xho*I and religation (26). The construction of *rfaK2*::ΩKm^r is described in Table 2. The construction of *rfa*::Tn*lac* insertions was as described previously (32).

Analysis of LPS. The presence of O antigen on saline suspensions of bacterial cells was determined by slide agglutination with *Salmonella* O antiserum (group B, factors 1, 4, 5, and 12) and *Shigella* O antiserum (poly group A) from Difco Laboratories, Detroit, Mich. LPS for gel analysis was isolated by SDS-proteinase K digestion from an outer membrane fraction obtained by differential centrifugation (1), and this was separated by polyacrylamide gel electrophoresis with an SDS-tricine buffer system and visualized by silver staining as previously described (32).

Nucleotide sequence accession number. The nucleotide sequence of the *S. dysenteriae* 1 *rfp* gene is available from GenBank (accession no. M96064).

RESULTS

Structure and properties of the *rfp* gene of *S. dysenteriae* 1. The restriction map of a portion of plasmid pJK2363 indicating the location of the *rfp* gene is shown in Fig. 2. Plasmid pJK2363 was constructed by subcloning a 2.5-kb *Bam*HI-*Cla*I fragment from pSS3 (41) into pGEM4, and strains

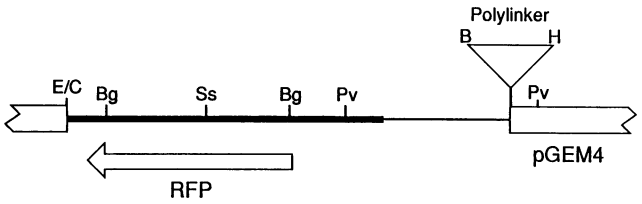

FIG. 2. Structure of plasmid pJK2363 showing the location of the *rfp* coding region. The single line indicates the 2.5-kb insert subcloned from pSS3; the thick dark line indicates the portion of the insert for which sequence has been submitted to GenBank. The open bars indicate the vector pGEM4 (Promega, Inc., Madison, Wis.). The arrow indicates the *rfp* coding region. Restriction sites are as described in the legend to Fig. 1. B, *Bam*HI; Ss, *Ssp*I.

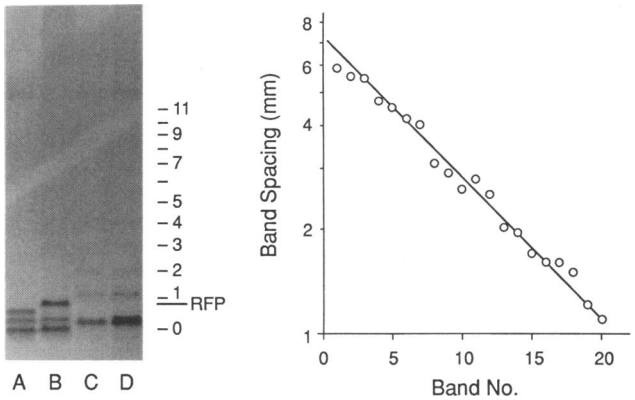

FIG. 3. Effect of *rfp* and the *rfb* gene cluster of *S. dysenteriae* 1 in *E. coli* K-12. (Left panel) LPS gel. Lanes: A, the background strain CS180; B, CS2132 (*rfp*$^+$); C and D, two different loadings of CS1861 (*rfb*$^+$ *rfp*$^+$). (Right panel) Semilogarithmic plot of the spacings between the O-antigen bands in lane D versus the band numbers. The scale to the right of the gel shows the band spacings derived from the graph fitted to the actual bands in lanes C and D. RFP indicates the migration of the Rfp-modified form in lane B.

transformed with pJK2363 appeared to produce amounts of Rfp-modified LPS which were identical to the amounts produced by strains transformed with pSS3 as judged by gel electrophoresis (shown below in Fig. 6, lanes F and J). The region indicated by the heavy line in Fig. 2 has been sequenced, and a single long open reading frame was found (as shown in Fig. 2). There were two ATG codons on either side of the 5′ *Bgl*II site which could serve as initiation codons. We have arbitrarily designated the upstream codon as the probable initiation codon. The open reading frame beginning at this codon encodes a protein of 377 residues, with a molecular mass of 41,846 Da, which is in good agreement with the value of 41,000 Da for the Rfp polypeptide generated in minicells (45). The amino acid composition indicates a basic protein with an isoelectric point of 9.75. As found for some *rfa* (20, 31) and *rfb* (18) genes, the coding region of *rfp* was quite AT rich (67.5% A + T).

The Rfp protein showed no significant homology with proteins available in the sequence data bank of PC Gene release 6.60. A barely significant sequence similarity between Rfp and Rfa proteins thought to be hexose transferases was observed, the most similar being RfaG. A comparison of hydropathy plots also showed similarity between Rfp and RfaG (data not shown), but this similarity was not strong enough to indicate a significant relationship. Like that of RfaG, the Rfp hydropathy plot showed no evidence of hydrophobic transmembrane domains. Thus, the structure of the Rfp protein is consistent with its being a sugar transferase but does not prove that it is. No significant sequence homology or structural similarity was observed between Rfp and any of the Rfb proteins of *Salmonella enterica* serovars which have been reported elsewhere (6, 18). Thus, there is no evidence that the *rfb* region of *Salmonella* spp. contains an analog of Rfp. This is in agreement with the results of Watanabe et al. (45), who found that the *rfp* gene cross-hybridized with DNA from other strains of *S. dysenteriae* 1 but not from *E. coli, S. typhimurium*, other *Shigella* spp., or other serotypes of *S. dysenteriae*.

Function of Rfp in modification of *E. coli* K-12. The LPS gel profiles of different genetic stocks of *E. coli* K-12 can vary considerably (28, 31). Since these strain-specific differences made the interpretation of changes in the LPS gel profile such as those resulting from Rfp modification more difficult, we recently examined these differences in more detail using a series of strains provided by the *E. coli* Genetic Stock Center at Yale University. We have found that the differences in gel profile were strain specific and could be related to specific steps in the derivation of the genetic stocks as described in the *E. coli* K-12 pedigree charts (2). Strains closely related to wild-type K-12 (pedigree chart 8), such as CS109, which is an F$^-$ derivative of W1485, and W3110, the strain used in the restriction map described by Kohara et al. (21), produce one rather broad major LPS band and a less-abundant, more slowly migrating band (data not shown). In contrast, strain Y-10 and its derivatives (chart 2) produce three sharp LPS bands, as seen in lane A of Fig. 3. The top two bands produced by strains derived from Y-10 comigrate with the two bands produced by CS109 and W3110. Strains which are derived from Y-10 include common K-12 background strains such as AB1133 (the background of the strains described in this paper), C600, and MC4100. The relative intensities of the three bands also vary in different strains derived from Y-10 (28), but these differences are less dramatic in relation to the results presented here. Rough (*rfb*) mutants of *S. typhimurium* LT2 tend to produce a single, rather diffuse band similar to the major band of CS109 (data not shown). The reason for these strain-specific differences in the LPS gel profile is unknown, but the sharpness of the bands in the strains derived from Y-10 makes these strains particularly useful in studying LPS core heterogeneity and the effect of Rfp.

As seen in lane B of Fig. 3, introduction of the *rfp* gene on plasmid pSS3 results in the disappearance of the top band seen in lane A and the appearance of a new band designated RFP in the scale adjacent to the gel. The most reasonable interpretation of these results is that the three bands seen in lane A represent at a minimum three distinct species or subsets of rough LPS molecules and that the Rfp protein is able to selectively modify one of these species or subsets.

If we assume that the Rfp-dependent shift in the top band involves the addition of one or more molecules of hexose, then we can use the band spacing of the O-antigen ladder to calculate how many molecules of hexose might be added. Lanes C and D of Fig. 3 show the O ladder which is observed when plasmid pSS37, which carries both the *rfp* gene and the *rfb* gene cluster from *S. dysenteriae* 1, is introduced into the strain. Each band in the ladder represents the addition of a four-hexose O unit. When the spacing between each band is plotted on a log scale versus the band number, a linear

relationship is observed (right panel of Fig. 3). On the basis of this relationship, we estimate that the Rfp-dependent shift seen in comparing the migrations of the top bands of lanes A and B of Fig. 3 is equivalent to the shift which would be expected as a result of the addition of one molecule of hexose to the largest LPS band produced by the parent strain CS180. Thus, the gel migration of the Rfp-modified LPS is consistent with the hypothesis that Rfp is a Gal transferase which adds one molecule of Gal to a subset of rough LPS molecules.

Figure 3 shows that the presence of the *rfb* gene cluster has two additional effects on the core. First, the bottom band seen in lanes A and B is absent from the LPS from the strains producing an O antigen (lanes C and D). Instead, a single, rather dark band which comigrates with the middle band seen in lane A is observed at the bottom of the O ladder. Subsequent sections of this report provide more detail about the genetic functions involved in the production of this band. Second, none of the bands at the bottom of the O ladders seen in lanes C and D line up with the Rfp-modified band, as might be expected if the Rfp band was the LPS species to which the four-sugar O units were added. To illustrate this more clearly, we have fitted a scale to the right of lanes C and D. This scale is derived from the linear relationship shown in the graph and shows a reasonable fit to the actual bands observed on the gel. We know from insertion mutagenesis studies of the *rfb* insert of plasmid pSS37 (19) and from Western immunoblotting with O-specific serum (data not shown) that band 1 is the smallest species which contains an O unit.

If we assume that the LPS in Fig. 3, band 1, arises by the transfer of an intact four-sugar O unit from the ACL to an LPS precursor, then such a precursor would migrate much faster than the Rfp-modified LPS band. The calculated position of this precursor is shown on the scale as band 0, and its position is actually below that of the heavy band seen at the bottom of the O ladder. In the original gel, a faint band could be seen migrating in the position predicted for band 0.

There are several possible ways of explaining the discrepancy between the migration of the Rfp-modified LPS band and the predicted migration of the band 0 precursor. First, there could be *rfb*-dependent trimming of the acceptor LPS accompanying transfer of the O antigen from the lipid carrier to the core. This trimming would have to remove the equivalent of about two or three sugars and would be analogous to the trimming of the high-mannose form of protein-linked oligosaccharides which occurs during the synthesis and intracellular sorting of eukaryotic glycoproteins (8). A second hypothesis comes from the suggestion of Sturm et al. (42) that the first O unit in the ladder is built up by the sequential addition of single sugars to the LPS core rather than to ACL. In this model, the Rfp protein adds a Gal residue to the core to begin the first O unit in a manner analogous to the way in which the RfbP protein is thought to add the first Gal phosphate unit to ACL to initiate the synthesis of subsequent O units. In this hypothesis, the RfP-modified band can be thought of as an LPS species on which the first O unit is already partially assembled by Rfp acting together with transferases encoded by *E. coli* K-12. A third reasonable hypothesis is that the shift in migration of band O results from an *rfb*-dependent inhibition of attachment of some substituents during synthesis of the LPS core. We must emphasize that we do not have enough direct evidence to decide among these mechanisms.

Functions of RfaK and RfaL. The effects of knockout insertion mutations in *rfaK* and *rfaL* are illustrated in Fig. 4.

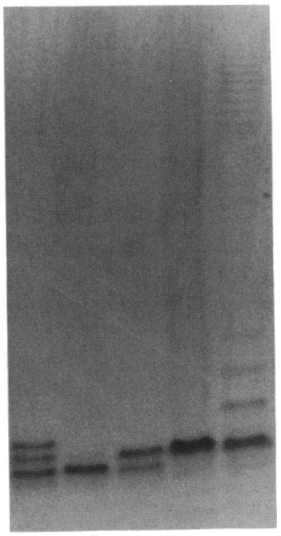

FIG. 4. LPS gel showing the effect of mutations in *rfaK* and *rfaL*. Lanes: A, CS180 (*rfa*$^+$); B, CS2592 (*rfaK*); C, CS2334 (*rfaL*); D, CS2476 (*rfaL* transformed with pSS37); E, CS1861 (*rfa*$^+$ transformed with pSS37).

As seen in Fig. 4, lane B, an insertion in *rfaK* resulted in the loss of both of the top bands seen in the *rfa*$^+$ parent (lane A). When pSS3 (*rfp*$^+$) was introduced into the *rfaK* mutant, no change in gel profile was observed (data not shown). O antigen was not detected by agglutination with O-specific serum, and there was no change in gel profile when pSS37 (*rfp*$^+$ *rfb*$^+$) was introduced into the *rfaK* mutant (data not shown). These results indicate that in *E. coli* K-12, *rfaK* is absolutely required for expression of *S. dysenteriae* 1 O antigen and that this requirement may reflect an inability of Rfp to modify RfaK$^-$ LPS. The single band seen with the *rfaK* mutant comigrates with the larger of the two bands seen with an *rfaJ* mutant (31). The gel phenotype of *rfaK* mutations is more dramatic in *E. coli* K-12 than in *S. typhimurium* (see Fig. 5, next section) suggesting that in *E. coli* K-12, *rfaK* may act at an earlier step in core synthesis.

An insertion mutation in *rfaL* (Fig. 4, lane C) resulted in the loss of the top LPS band seen in the parent. No change was seen in the migration of the two remaining bands when pSS3 was introduced into the *rfaL* mutant (data not shown), indicating that *rfaL* is also required for the function of *rfp*. The fact that *rfaL* is required for core completion and for Rfp-dependent modification indicates that the RfaL protein is more than simply a component of the O-antigen ligase, which was the only function that had been previously proposed for it (25).

An unexpected result was obtained when the *rfaL* mutant was transformed with pSS37 (Fig. 4, lane D). The presence of the *rfb*-containing plasmid resulted in the loss of the bottom band seen in Fig. 4, lane C. The single band which remained comigrated with the major band at the bottom of the O ladder produced by an *rfa*$^+$ strain transformed with pSS37 (Fig. 4, lane E). This change in gel profile was unexpected, because there was no known connection between the core pathway, in which the molecules are assembled on derivatives of lipid A, and the pathway for O antigen, in which molecules are assembled on ACL (16). Although the transformed *rfaL* mutant grew somewhat more slowly

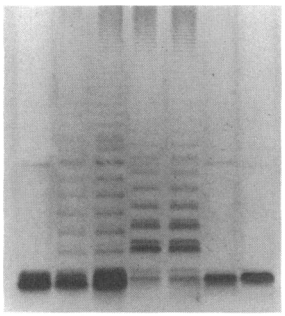

FIG. 5. Complementation of *S. typhimurium rfa* mutations by plasmids with *E. coli* K-12 *rfa* genes. Lanes: A, strain SL733 (*rfaK*); B and C, low and high loadings of SL733 (*rfaK*) transformed with pJK2252, respectively; D, SA1377 (*rfaC*) transformed with pJK2252; E, SL3770 (*rfa*$^+$); F, SL3749 (*rfaL*) transformed with pJK2252; G, SL3749 (*rfaL*).

than a comparable strain without pSS37, it did not appear to have an obvious growth defect, and thus it is unlikely that the change in core profile is due to a toxic effect resulting from an accumulation of ACL-bound O antigen. Instead, this result suggests that one or more genes in the *rfb* cluster are capable of regulating the activity of the *rfa* cluster even in the absence of RfaL function to determine the type of core molecules which are synthesized or direct some form of modification which is independent of O-antigen attachment.

Cross-complementation of *rfaK* and *rfaL* between *E. coli* K-12 and *S. typhimurium* LT2. Since the primary sequences of the RfaK and RfaL polypeptides are not conserved between *E. coli* K-12 and *S. typhimurium*, it was of interest to determine whether these genes could cross-complement mutations in the two organisms. Figure 5 shows the complementation of *S. typhimurium* mutants by plasmids which carry the *E. coli* K-12 *rfaKL* region. Wild-type *S. typhimurium* is shown in lane E of Fig. 5.

Lanes B and C of Fig. 5 show that the *E. coli* K-12 *rfaK*$^+$ *-L*$^+$ plasmid pJK2252 can partially complement an *S. typhimurium rfaK* mutant. The complementation is not efficient, as indicated by the faint O-antigen ladder in relation to the large amount of rough LPS at the bottom of the gel. An interesting observation is that the O-antigen bands which are made by the *S. typhimurium rfaK* mutant complemented by *E. coli* K-12 *rfaK* migrate more rapidly than the wild-type bands (compare lanes C and D of Fig. 5). The change in migration is about one-fourth of the distance between the bands or equivalent to a change in size of about one hexose unit.

Since the plasmid used in the *rfaK* complementation experiment depicted in lanes B and C of Fig. 5 contained both *rfaK* and *rfaL* of *E. coli* K-12, there was a possibility that both *E. coli* genes were required for complementation. To test this, we transformed the *S. typhimurium rfaK* mutant with a plasmid which was identical to pJK2252 except that it contained a Tn*phoA* insertion in *rfaL*. Plasmid pJK2312 did not complement the *S. typhimurium rfaK* mutation (data not shown), indicating that both *rfaK* and *rfaL* of *E. coli* K-12 were required for complementation.

As shown in lane F of Fig. 5, there was no detectable complementation of an *S. typhimurium rfaL* mutant by the *E. coli* K-12 *rfaK*$^+$ *-L*$^+$ plasmid. This cannot be explained on the basis of poor expression, since *rfaC* and *rfaL* are adjacent genes in an operon (38) and pJK2252 was capable of efficiently complementing an *rfaC* mutant (Fig. 5, lane D).

TABLE 3. Effect of *rfaK* alleles on phage sensitivity

Strain	Relevant genotype	Plasmid	Sensitivity to the following phages[a]:			
			FO	Br2	C21	P-22
E. coli K-12						
CS180	*rfa*$^+$	None	R	R	R	—
CS2529	*rfaK*	None	R	S	S	—
CS2742	*rfaK*	pKZ38	S	S	R	—
CS2334	*rfaL*	None	R	R	R	—
S. typhimurium LT2						
SL3770	*rfa*$^+$ *rfb*$^+$	None	S	R	R	S
SL733	*rfaK*	None	R	S	R	R
CS2272	*rfaK*	pJK2252	R	R	R	S

[a] R, resistant; S, sensitive; —, not determined.

Taken together, the results of these experiments indicate that the RfaL protein has two distinct functions. One involves a specific interaction with the RfaK protein or with an LPS acceptor which has been modified in a species-specific way by the RfaK protein (see the next section). It is this function which makes *rfaL* from *E. coli* K-12 necessary along with *E. coli* K-12 *rfaK* in order to complement an *rfaK* mutant of *S. typhimurium*. The other function of the RfaL protein must involve a specific, *rfaK*-independent interaction with another component which is specific to *S. typhimurium*. It is this second *rfaL* function which cannot be complemented by the *E. coli* K-12 *rfa* genes supplied on plasmid pJK2252.

We have done a reciprocal complementation experiment in which plasmid pKZ38, which carries *rfaK* and *-L* of *S. typhimurium*, was introduced into *E. coli* K-12 *rfaK* and *rfaL* insertion mutants transformed with pSS37. A small amount of O antigen was observed on gels in both cases, although there was less with the *rfaL* mutant than with the *rfaK* mutant. This was confirmed by agglutination of the transformed cells with O-specific antiserum (data not shown). This indicates that although the expression of *S. dysenteriae* 1 O antigen by *E. coli* K-12 exhibits a similar specificity, it may not be quite as stringent as *S. typhimurium* O antigen in its requirement for homologous *rfaK* and *rfaL* functions.

RfaK determines phage specificity. Phage FO is one of the best-studied LPS-specific phages (23, 25), and there is abundant evidence that the receptor for this phage requires a terminal GlcNAc attached to Glc II of the *S. typhimurium* core. *E. coli* K-12 is completely resistant to FO, consistent with the hypothesis that in K-12, GlcNAc is attached to a different site deeper within the core. As shown in Table 3, introduction of pKZ38, which carries the *rfaKL* region of *S. typhimurium*, into an *rfaK* mutant of *E. coli* K-12 resulted in sensitivity to FO. This implies that *rfaK* of *S. typhimurium* is active on the *E. coli* K-12 LPS core and can add a terminal GlcNAc residue.

S. typhimurium rfa mutants with a core which contains at least one hexose residue (chemotype Rc) are sensitive to the rough-specific phage Br2 (35). Unlike FO, Br2 must recognize an internal core structure. The *rfa*$^+$ *E. coli* K-12 strain CS180 is resistant to Br2, but an *rfaK* mutant of CS180 is sensitive to Br2. We interpret this as indicating that the attachment of GlcNAc to the internal K-12 site blocks the Br2 receptor. The disaccharide GlcNAc-heptose (Hep) has been identified in LPS from *E. coli* K-12 (13); on this basis,

TABLE 4. Phage sensitivity of *rfaQ*, *rfaS*, and *rfaZ* mutants

Strain	Relevant genotype	Sensitivity to the following phage[a]:	
		C21	Mu
CS2767	*rfa*$^+$	R	S
CS2774	*rfaQ*	S	S
CS2775	*rfaS*	S	S
CS2776	rfaZ	S	S

[a] R, resistant; S, sensitive.

we speculate that the *E. coli* K-12 RfaK protein may be a transferase which adds GlcNAc to an inner core Hep.

An *rfaK* mutant of *S. typhimurium* transformed with the K-12 *rfaK*$^+$ plasmid pJK2252 remained FO resistant but became Br2 resistant, which suggests that *E. coli* K-12 RfaK is active on the *S. typhimurium* core and can transfer GlcNAc to the K-12 site. However, an *rfaL* mutant of *S. typhimurium* transformed with pJK2252 remained sensitive to both FO and Br2, indicating that the *S. typhimurium* RfaK was predominant when RfaK from both species was present. This may reflect less than optimum expression of *rfaK* from the plasmid insert, which lacks the *rfa* promoter region (30). A similar competition between genes from both organisms has been observed in complementation studies involving the *rfaIJ* region (31).

Additional evidence indicating that *rfaK* has different functions in *E. coli* K-12 and *S. typhimurium* is provided by differences in sensitivity to phage C21. This phage binds to a receptor in the inner core which is normally blocked and must be unmasked by mutations affecting core hexose structure or by chemical treatment of isolated LPS (35). The very tight *S. typhimurium rfaK* mutant used in this study remained resistant to C21. However, *rfaK* mutants of *E. coli* K-12 became sensitive to C21, suggesting an effect deeper in the core.

RfaQ, -S, and -Z affect core completion. The newly discovered genes *rfaQ*, *-S*, and *-Z* each affect some aspect of core structure. This is indicated by the sensitivities of strains with mutations in these genes to the inner core-specific phage C21 (Table 4). These strains remain sensitive to Mu, which requires the terminal Glc as a component of its receptor (39).

When the LPS of insertion mutations in *rfaQ*, *-S*, and *-Z* is compared on gels with that of *rfa*$^+$ background strains such as CS180, the differences in the three predominant bands are slight and thus not particularly useful in attempting to assign functions to these genes. This is shown in lanes A through E of Fig. 6. The *rfaQ* insertion (lane B) shows the most significant alteration, since there is a small but significant downward shift of the bottom band as well as a substantial decrease in the relative abundance of the top band. The mutations in *rfaS* and *-Z* (lanes C and D) primarily affect the relative abundance of the bands.

We observed that the differences resulting from these mutations were much more dramatic if we transformed the mutants with the *rfp*$^+$ plasmid pSS3 or with the plasmid pSS37, which carries both the *rfp* and the *rfb* genes and results in production of *S. dysenteriae* 1 O antigen. This is shown in lanes F through R of Fig. 6.

Lane G of Fig. 6 shows that when pSS3 is introduced into an *rfaQ* mutant, there is only a slight production of the Rfp-modified top band in comparison with that of the pSS3-transformed *rfa*$^+$ strain shown in lanes F and J. The apparent shift of the bottom band seen in lane B is also seen in lane G. Lanes L and P show two different loadings of the *rfaQ* mutant transformed with pSS37. The amount of O-antigen ladder produced is substantially reduced compared with that of the *rfa*$^+$ strains shown in lanes K and O. The reduction in O antigen is consistent with the reduction in the Rfp-modified band seen in lane G. This is particularly relevant since lane P shows that there is not only a reduction in the amount of O antigen but also a shift in the size distribution of O-antigen bands from the smaller bands seen near the bottom of the ladder to the larger bands in the cluster near the top. Since the distribution of bands in the ladder reflects the rate of transfer of O antigen from ACL to the LPS core relative to the rate of polymerization of the O antigen, a shift to larger molecules in the ladder is consistent with a reduction in the pool of LPS cores suitable as O-antigen acceptors. This also tends to support the hypothesis that the Rfp-modified band represents an earlier stage in the synthesis of the core which carries O antigen.

Another interesting effect of the *rfaQ* mutation can be seen in comparing Fig. 6, lanes K and L, or at a higher loading, lanes O and P. As noted above in the discussion of *rfaL* and as shown in lanes C and D of Fig. 4, the introduction of pSS37 into an *rfaL* mutant resulted in the disappearance of the bottom LPS band and the production of a single band similar to that which is seen below the O ladder in strains producing O antigen. Production of such a band also occurs in the *rfaQ* mutant, but lanes L and P of Fig. 6 show that this band is now shifted downward by the same amount as the bottom band seen in lanes F and G, so that the prominent band seen at the bottom of lanes L and P now no longer comigrates with the middle band seen in *rfa*$^+$ parental strains (lanes A and E). The fact that introduction of the *rfb* cluster results in a change in the position of the band whose migration is altered by *rfaQ* suggests something about the pathway by which these bands arise. The bottom band seen in the *rfa*$^+$ parental strains must represent an earlier stage in completion of the core than the prominent band seen at the bottom of the O ladder in strains expressing O antigen.

Although it is difficult to see the faint bands in the lower portion of the O ladder in lanes L and P of Fig. 6, these bands could be seen in the original gel, and it did not appear that they were shifted by the *rfaQ* mutation. The observation that the faint Rfp-modified band seen in lane G of Fig. 6 was not noticeably shifted by the *rfaQ* mutation suggests that the LPS species to which O antigen is attached and the predominant band seen at the bottom of the O ladder arise through modification pathways which are different with respect to the role of *rfaQ*.

Additional evidence for a difference between cores to which O antigen is attached and core species which lack O antigen is provided by the effects of mutations in *rfaS* and *-Z*. As shown in lane H of Fig. 6, the major effect of the *rfaS* mutation in a strain transformed with pSS3 appears to be a substantial reduction in the middle band. This is more dramatic in the *rfaS* strain transformed with pSS37 (Fig. 6, lanes M and Q), in which it can be seen that the prominent band normally seen at the bottom of the O ladder is almost entirely absent. Instead, there is a dark band comigrating with the bottom band in lanes E and J and an additional faint faster-migrating band. A qualitatively similar picture is seen with the *rfaZ* mutant (Fig. 6, lanes N and R), although the relative band intensities of the faster-migrating bands are different. It should be noted that in contrast to *rfaQ*, neither *rfaS* nor *rfaZ* mutations appear to decrease the production of O antigen or affect the size distribution of the ladder. In fact,

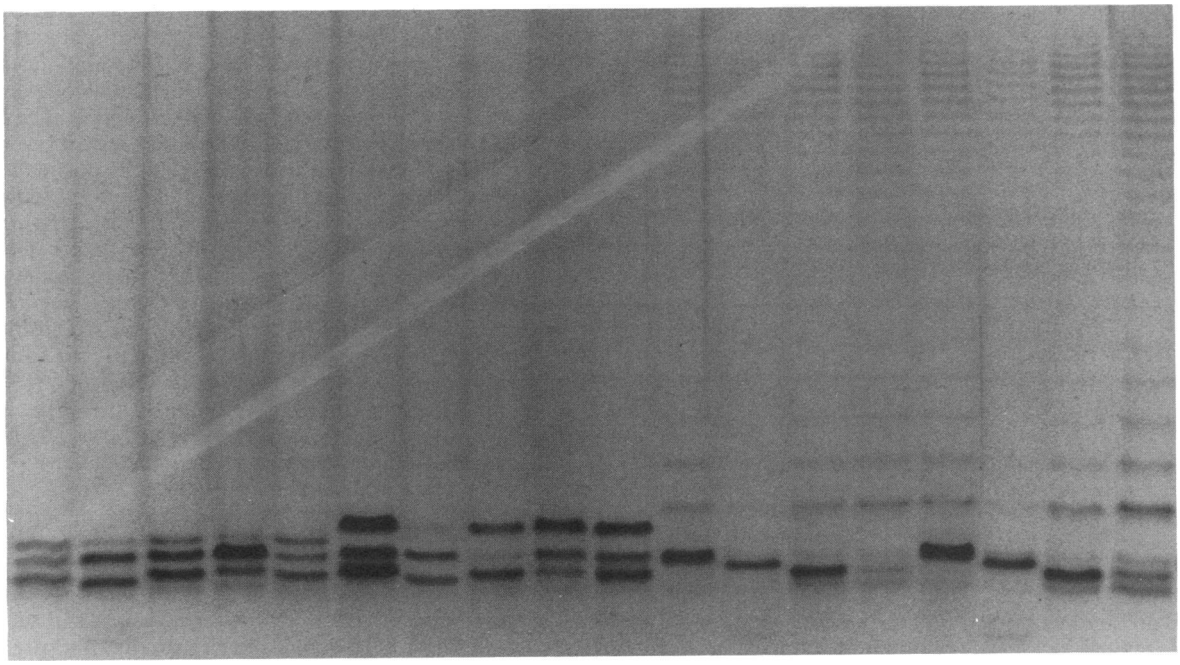

FIG. 6. LPS gel showing the effects of mutations in *rfaQ*, *rfaS*, and *rfaZ* in different genetic backgrounds. The blocks below the lanes indicate whether the strain is *rfa*$^+$ or carries a chromosomal *rfa* mutation and whether the strain carries a plasmid encoding the *rfp* gene (pSS3) or both *rfp* and the *rfb* cluster (pSS37). Lanes: A and E, CS180; B, CS2682; C, CS2683; D, CS2452; F and J, heavy and light loadings of CS2363, respectively; G, CS2686; H, CS2687; I, CS2456; K and O, light and heavy loadings of CS1861, respectively; L and P, light and heavy loadings of CS2684, respectively; M and Q, light and heavy loadings of CS2685, respectively, N and R, light and heavy loadings of CS2457, respectively.

the *rfaZ* mutant appears to produce slightly more O antigen than the *rfa*$^+$ parent.

DISCUSSION

Heterogeneity of the LPS core: roles of RfaQ and -K. The LPS biosynthetic system is remarkable in its ability to export to the cell surface molecules which are highly diverse in size and structure. The LPS seen in the outer membrane represents a sampling of precursor pools containing molecules at various stages of completion. This sampling at various stages of completion is most clearly seen in the ladders of molecules containing O antigen. The size heterogeneity of the ladders suggests a competition between polymerization, which is the transfer of O polymers to O units on ACL to increase their size, and ligation or translocation, which is the transfer of O polymers from ACL to the cores of LPS molecules which are exported to the outer membrane.

The recently developed tricine-SDS gel system (22) is capable not only of resolving the ladders of LPS bands which result from molecules with a relatively homogeneous core carrying O antigens of different lengths but also of resolving into distinct and reproducible bands the much more heterogeneous population of LPS molecules which lack O antigen. Like the O ladders, these rough bands must reflect both intermediate and full-size products. While none of these bands consist of precursors per se, some of them must be derived from pools of intermediate-size molecules in the cytoplasmic membrane which serve as precursors for the complete or full-size molecules. Therefore, changes in the presence or intensity of these bands in response to mutational or physiological changes can be used to obtain information about the biochemical pathways involved in core completion and about the functional aspects of core heterogeneity. This type of analysis is particularly useful in helping to define genes involved in adding core substituents which are not present in stoichiometric amounts or which may be removed during later stages of core completion or O-antigen attachment. Two such genes are *rfaQ* and *rfaK*.

The RfaQ protein has sequence homology with RfaC and KdtA (7, 30), which are, respectively, Hep and KDO transferases. The shift in core bands seen in *rfaQ* mutants (Fig. 6) indicates that RfaQ is also involved in core synthesis and supports the idea that RfaQ may also be a transferase which adds a substituent to the inner core. An attractive hypothesis which we have been unable to prove or disprove is that RfaQ is the *rfaP*-dependent transferase which adds the Hep III branch to Hep II (29). It is known that Hep III is present in the core only in a subset of molecules (29), and lane B of Fig. 6 shows that in an *rfaQ* mutant without an O-antigen plasmid, only one of the two major bands is shifted.

Two preliminary lines of evidence indicate that the *rfaQ* gene is present and functional in *S. typhimurium*. First, we have obtained a partial sequence for the *kdtA-rfaQ* portion of *S. typhimurium* (data not shown). The structures of both the KdtA and the RfaQ proteins appear to be strongly conserved between *E. coli* K-12 and *S. typhimurium*, since both pairs of homologous proteins show about 90% amino acid sequence identity in the regions in which the sequence is known in both organisms. In addition, introduction of a plasmid carrying a cloned *S. typhimurium rfa* DNA fragment which included *rfaQ* into a strain of *E. coli* K-12 which had a large internal deletion in *rfaQ* resulted in complementation of both parts of the RfaQ$^-$ LPS phenotype, altered migration of core bands, and reduction in expression of the Rfp-modified band (data not shown).

The most interesting effect of the *rfaQ* mutation is that the prominent unsubstituted core band seen at the bottom of the O ladder of the mutant transformed with pSS37 was shifted to a higher mobility, while the O ladder itself and the Rfp-modified band seen in the mutant transformed with pSS3 were not. This suggests two things. First, it suggests that the Rfp-modified band which represents an earlier stage in the synthesis of the LPS molecule to which O antigen is attached does not have the RfaQ modification, either because it was never attached to begin with or because it was removed prior to or during Rfp modification. Second, it suggests that the prominent unsubstituted core band seen in strains expressing O antigen (bottom band in lanes K and O of Fig. 6) is not identical in structure to the middle band seen in strains without an O-antigen plasmid (middle band in lanes A through G and lanes I and J), even though they comigrate. This is based on the observation that only the band seen in strains expressing O antigen shifts in an *rfaQ* mutant.

The top band which is seen in the parental *rfa*$^+$ strain without plasmids (Fig. 6, lanes A and E) is absent in an *rfaL* mutant (Fig. 4) and disappears in *rfa*$^+$ strains transformed with pSS3 (Fig. 6, lanes F and J). On this basis, this band appears to represent the subset of LPS molecules which are modified by Rfp and which are assumed to become the substrate for O-antigen attachment. This top band is very faint in an *rfaQ* mutant (Fig. 6, lane B), as is the Rfp-modified band of the *rfaQ* mutant (Fig. 6, lane G), and expression of O antigen is weak in the *rfaQ* mutant. A hypothesis which might explain all of these observations is that synthesis of the LPS in the top band may require the *rfaL*-dependent removal of the RfaQ substituent. If this substituent was already absent, binding of the precursor of the top band to RfaL protein might be impaired.

It is by no means proven that RfaK is a GlcNAc transferase. However, the effect of *rfaK* mutations on the chemotype of the *S. typhimurium* core (25), the structural similarity of RfaK to and partial sequence homology with a family of Rfa proteins including RfaB and RfaG which are also thought to be hexose transferases (20, 30, 31), and the partial cross-complementation between the *rfaK* genes of *S. typhimurium* and *E. coli* K-12 can all be taken as support for the idea that *rfaK* is the structural gene for a GlcNAc transferase in both organisms. While the cross-complementation experiments suggest that the presence and nature of the substituent are of primary importance for the attachment of O antigen, the phage FO, Br2, and C21 results and the LPS gel phenotype of the *E. coli* K-12 *rfaK* mutant indicate that the site of attachment of the substituent is an important determinant of phage attachment and core completion. The amount of GlcNAc detected by chemical analysis of LPS from *E. coli* K-12 W3110 was quite low (5). This is in contrast to the rather substantial change in the gel phenotype observed for an *rfaK* mutant (Fig. 4, lane B), which suggests that a majority of the molecules require RfaK modification, at least at some intermediate stage of completion. Thus, like the RfaQ substituent, GlcNAc should be considered a potential candidate for substituent which is removed during the later stages of core completion and O-antigen addition. The shift in the migration of the O ladders seen in the *rfaK* cross-complementation experiment depicted in Fig. 5 is consistent with a change in processing or modification of the core resulting from the addition of a critical substituent to a different site.

RfaS and -Z define an unsubstituted core modification pathway. Like *rfaQ*, the phenotypes of *rfaS* and *rfaZ* mutations are most dramatic in strains expressing O antigen. However, in contrast to *rfaQ*, *rfaS*, and *-Z* mutations do not appear to affect the amount or distribution of LPS in the O-antigen ladders. Instead, in both *rfaS* and *rfaZ* mutants, the prominent unsubstituted core band seen in *rfa*$^+$ strains transformed with pSS37 appears to be absent. On the original gels, a very faint band was seen at the same migration position, but it is likely that this is a different molecular species, since a similar faint band was sometimes seen above the shifted band in gels of LPS from *rfaQ* mutants transformed with pSS37. Both *rfaS* and *rfaZ* mutants transformed with pSS37 produce faster-migrating core bands which are not seen or are very faint in *rfa*$^+$ strains transformed with pSS37. The major difference between the phenotypes of *rfaS* and *rfaZ* is that in *rfaS* mutants, one of these faster-migrating core bands is much more distinct.

These results indicate that the prominent unsubstituted core band seen at the bottom of the ladder in strains producing O antigen arises via a different modification pathway than the LPS core to which O antigen is attached. The pathways involve different genes, *rfaS* and *rfaZ* in the case of the unsubstituted core band and *rfaL* and *rfp* in the case of cores bearing O antigen. In addition, these two pathways use *rfaQ* differently. Since the enteric LPS core has commonly been considered to serve primarily as a substrate for O-antigen addition and rough cores have been considered to arise primarily through mutation or inefficient transfer of O antigen, the idea of multiple core modification pathways to produce cores which are unsubstituted with O antigen and cores of a different structure which serve as a platform for O-antigen addition is quite novel. The possibility of a separate pathway for synthesis of cores without O antigen also has important implications in studying the biological effects of endotoxins, since it is widely known that endotoxins containing rough and smooth LPS have very different biological and pharmacological effects.

To simplify the nomenclature beyond the present system that is employed (12), we would like to propose that the term lipooligosaccharide (LOS) and not simply rough or R LPS be used to describe this class or subset of rough molecules in enteric bacteria which have unsubstituted cores of a unique structure and which are not merely incomplete forms of smooth LPS molecules. We feel that this is appropriate, since the structures of these molecules are likely to be similar in overall organization and complexity to those of the molecules designated LOS which are produced by nonenteric gram-negative bacteria. The more generic term LPS should be retained to describe molecules containing O antigen or other polymeric polysaccharides, and rough or R LPS should be used to describe molecules with cores which are capable of being substrates for O-antigen attachment and molecules which represent an earlier stage in core comple-

tion and are capable of serving as intermediates in both pathways.

Rfa and Rfb are interacting systems. We can cite three lines of evidence that the *rfa* and *rfb* gene clusters are designed to work together as an interacting system. The first is the intimate involvement of *rfaK* and *rfaL* in the attachment of O antigen. The second is the apparent alteration of the Rfp-modified core to provide a substrate for attachment of O antigen, which is suggested by the results in Fig. 3. The third is the *rfaL*-independent shift in the population of unsubstituted core molecules to produce a single LOS band as seen in lane D of Fig. 4.

Such interaction is not surprising, given the complex regulation which must be required to enable the core and O-antigen pathways to function together smoothly, but it was not anticipated on the basis of the previous literature on O-antigen and core biosynthetic pathways. These findings have important implications with respect to the construction of chimeric strains expressing heterologous O antigens for vaccine purposes. These results may also explain why it has been difficult to obtain efficient transfer of O antigen from the ACL to LPS cores in vitro, except in crude systems containing envelope fragments. The LPS from rough mutants may not be an appropriate acceptor for O antigen without prior *rfb*-dependent modification.

ACKNOWLEDGMENTS

We thank E. Pradel for assistance with genetic constructions, for providing partial sequence of the *rfaQ-kdtA* region of *S. typhimurium*, and for carrying out the *rfaQ* complementation experiments. We also thank P. R. MacLachlan and K. Sanderson for providing *S. typhimurium* plasmids and *rfaQ* sequence prior to publication.

This research was supported by Public Health Services grant GM-39087.

REFERENCES

1. **Austin, E. A., J. F. Graves, L. A. Hite, C. T. Parker, and C. A. Schnaitman.** 1990. Genetic analysis of lipopolysaccharide core biosynthesis by *Escherichia coli* K-12: insertion mutagenesis of the *rfa* locus. J. Bacteriol. **172:**5312–5325.
2. **Bachmann, B. J.** 1987. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12, p. 1190–1219. *In* F. C. Neidhardt, 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. 2. American Society for Microbiology, Washington, D.C.
3. **Bachmann, B. J.** 1990. Linkage map of *Escherichia coli* K-12, edition 8. Microbiol. Rev. **54:**130–197.
4. **Bastin, D. A., L. K. Romana, and P. R. Reeves.** 1991. Molecular cloning and expression in *Escherichia coli* K-12 of the *rfb* gene cluster determining the O antigen of an *E. coli* O111 strain. Mol. Microbiol. **5:**2223–2231.
5. **Brade, H.** Personal communication.
6. **Brown, P. K., L. K. Romana, and P. R. Reeves.** 1991. Cloning of the *rfb* cluster of a group C2 *Salmonella* strain: comparison with the *rfb* regions of groups B and D. Mol. Microbiol. **5:**2021–2029.
7. **Clementz, T., and C. R. H. Raetz.** 1991. A gene coding for 3-deoxy-D-*manno*-octulosonic acid transferase in *Escherichia coli*. J. Biol. Chem. **166:**9687–9696.
8. **Farquahr, M. G.** 1985. Progress in unraveling pathways of Golgi traffic. Annu. Rev. Cell Biol. **1:**447–488.
9. **Fellay, R., J. Frey, and H. M. Kritsch.** 1987. Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for *in vitro* insertional mutagenesis of gram-negative bacteria. Gene **52:**147–154.
10. **Hale, T. L., P. Guerry, R. C. Seid, Jr., C. Kapfer, M. Wingfield, C. B. Reaves, L. S. Baron, and S. B. Formal.** 1984. Expression of lipopolysaccharide O antigen in *Escherichia coli* K-12 hybrids containing plasmid and chromosomal genes from *Shigella dysenteriae* 1. Infect. Immun. **46:**470–475.
11. **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.
12. **Hitchcock, P. J., L. Leive, P. H. Mäkelä, E. T. Rietschel, W. Strittmatter, and D. C. Morrison.** 1986. Lipopolysaccharide nomenclature—past, present and future. J. Bacteriol. **166:**699–705.
13. **Holst, O., and H. Brade.** 1992. Chemical structure of the core region of lipopolysaccharides, p. 135–170. *In* D. C. Morrison and J. L. Ryan (ed.), Bacterial endotoxic lipopolysaccharides, vol. 1. Molecular biochemistry and cellular biology of lipopolysaccharides. CRC Press, Inc., Boca Raton, Fla.
14. **Holst, O., E. Röhrscheidt-Andrezejewski, and H. Brade.** 1990. Isolation and characterization of 3-deoxy-D-*manno*-2-octulosopyranosate-7-(2-aminoethyl phosphate) from the inner core region of *Escherichia coli* K-12 and *Salmonella minnesota* lipopolysaccharides. Carbohydr. Res. **204:**93–102.
15. **Holst, O., U. Zähringer, H. Brade, and A. Zamojski.** 1991. Structural analysis of the heptose/hexose region of the lipopolysaccharide from *Escherichia coli* K-12 strain W3100. Carbohydr. Res. **215:**323–355.
16. **Jann, K., and B. Jann.** 1984. Structure and biosynthesis of O-antigens, p. 138–186. *In* E. T. Rietschel (ed.), Handbook of endotoxin, vol. 1. Chemistry of endotoxin. Elsevier, Amsterdam.
17. **Jansson, P.-E., A. A. Lindberg, B. Lindberg, and R. Wollin.** 1981. Structural studies on the hexose region of the core in lipopolysaccharides from enterobacteriaceae. Eur. J. Biochem. **115:**571–577.
18. **Jiang, X. M., B. Neal, F. Santiago, S. J. Lee, L. K. Romana, and P. R. Reeves.** 1991. Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar typhimurium (strain LT2). Mol. Microbiol. **5:**695–713.
19. **Klena, J. D.** Unpublished data.
20. **Klena, J. D., E. Pradel, and C. A. Schnaitman.** 1992. Comparison of the lipopolysaccharide biosynthesis genes *rfaK*, *rfaL*, *rfaY*, and *rfaZ* of *Escherichia coli* K-12 and *S. typhimurium*. J. Bacteriol. **174:**4746–4752.
21. **Kohara, K. Y., K. Akiyama, and K. Isono.** 1987. The physical map for the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50:**495–508.
22. **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. Immunol. Methods **126:**109–117.
23. **Lindberg, A. A., and T. Holme.** 1969. Influence of O side chains on the attachment of the Felix O-1 bacteriophage to *Salmonella* bacteria. J. Bacteriol. **99:**513–519.
24. **Maclachlan, P. R., S. K. Kadam, and K. E. Sanderson.** 1991. Cloning, characterization, and DNA sequence of the *rfaLK* region for lipopolysaccharide synthesis in *Salmonella typhimurium* LT-2. J. Bacteriol. **173:**7151–7163.
25. **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. Chemistry of endotoxin. Elsevier, Amsterdam.
26. **Manoil, C.** 1990. Analysis of protein localization by use of gene fusions with complementary properties. J. Bacteriol. **172:**1035–1042.
27. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Parker, C. T.** Unpublished data.
29. **Parker, C. T., A. W. Kloser, C. A. Schnaitman, M. A. Stein, S. Gottesman, and B. W. Gibson.** 1992. Role of the *rfaG* and *rfaP* genes in determination of the lipopolysaccharide structure and cell surface properties of *Escherichia coli* K-12. J. Bacteriol. **174:**2525–2538.
30. **Parker, C. T., E. Pradel, and C. A. Schnaitman.** 1992. Identification and sequence of the lipopolysaccharide core biosynthetic

genes *rfaQ*, *rfaP*, and *rfaG* of *Escherichia coli* K-12. J. Bacteriol. **174**:930–934.
31. **Pradel, E., C. T. Parker, and C. A. Schnaitman.** 1992. Structures of the *rfaB*, *rfaI*, *rfaJ*, and *rfaS* genes of *Escherichia coli* K-12 and their roles in the assembly of the lipopolysaccharide core. J. Bacteriol. **174**:4736–4745.
32. **Pradel, E., and C. A. Schnaitman.** 1991. Effect of *rfaH* (*sfrB*) and temperature on the expression of *rfa* genes of *Escherichia coli* K-12. J. Bacteriol. **173**:6428–6431.
33. **Raina, S.** Personal communication.
34. **Raina, S., and C. Georgopoulos.** 1991. The *htrM* gene, whose product is essential for *Escherichia coli* viability only at elevated temperatures, is identical to the *rfaD* gene. Nucleic Acids Res. **19**:3811–3819.
35. **Rapin, A. M. C., and H. M. Kalckar.** 1971. The relation of bacteriophage attachment to lipopolysaccharide structure, p. 267–307. *In* G. Weinbaum, S. Kadis, and S. J. Ajl (ed.), Microbial toxins, vol. IV. Bacterial endotoxins. Academic Press, N.Y.
36. **Rick, P. D.** 1987. Lipopolysaccharide biosynthesis, p. 648–662. *In* F. C. Neidhardt, 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.
37. **Riley, L. W., L. N. Junio, L. B. Libaek, and G. K. Schoolnik.** 1987. Plasmid-encoded expression of lipopolysaccharide O-antigenic polysaccharide in enteropathogenic *Escherichia coli*. Infect. Immun. **55**:2052–2056.
38. **Roncero, C., and M. J. Casadaban.** 1992. Genetic analysis of the genes involved in synthesis of the lipopolysaccharide core in *Escherichia coli* K-12: three operons in the *rfa* locus. J. Bacteriol. **174**:3250–3260.
39. **Roncero, C., K. E. Sanderson, and M. J. Casadaban.** 1991. Analysis of the host ranges of transposon bacteriophages Mu, MuhP1, and D108 by the use of lipopolysaccharide mutants of *Salmonella typhimurium*. J. Bacteriol. **173**:5230–5233.
40. **Schnaitman, C. A., C. T. Parker, E. Pradel, J. D. Klena, N. Pearson, K. E. Sanderson, and P. R. MacLachlan.** 1991. Physical map of the *rfa* locus of *Escherichia coli* K-12 and *Salmonella typhimurium*. J. Bacteriol. **173**:7410–7411.
41. **Sturm, S., B. Jann, K. Jann, P. Fortnagel, and K. N. Timmis.** 1986. Genetic and biochemical analysis of *Shigella dysenteriae* 1 O-antigen polysaccharide biosynthesis in *Escherichia coli* K-12: 9 kb plasmid of *S. dysenteriae* 1 determines addition of a galactose residue to the lipopolysaccharide core. Microb. Pathog. **1**:299–306.
42. **Sturm, S., B. Jann, K. Jann, P. Fortnagel, and K. N. Timmis.** 1986. Genetic and biochemical analysis of *Shigella dysenteriae* 1 O-antigen polysaccharide biosynthesis in *Escherichia coli* K-12: structure and functions of the *rfb* cluster. Microb. Pathog. **1**:307–324.
43. **Sturm, S., and K. N. Timmis.** 1986. Cloning of the *rfb* region of *Shigella dysenteriae* 1 and construction of an *rfb-rfp* gene cassette for the development of lipopolysaccharide-based live anti-dysentery vaccines. Microb. Pathog. **1**:289–297.
44. **Valvano, M. A., and J. H. Crosa.** 1989. Molecular cloning and expression in *Escherichia coli* K-12 of chromosomal genes determining the O7 lipopolysaccharide antigen of a human invasive strain of *E. coli* O7:K1. Infect. Immun. **57**:937–943.
45. **Watanabe, H., A. Nakamura, and K. N. Timmis.** 1984. Small virulence plasmid of *Shigella dysenteriae* 1 strain W30864 encodes a 41,000-dalton protein involved in formation of specific lipopolysaccharide side chains of serotype 1 isolates. Infect. Immun. **46**:55–63.
46. **Watanabe, H., and K. N. Timmis.** 1984. A small plasmid in *Shigella dysenteriae* 1 specifies one or more functions essential for O antigen production and bacterial virulence. Infect. Immun. **43**:391–396.