

Genetics of Lipopolysaccharide Biosynthesis in Enteric Bacteria

CARL A. SCHNAITMAN* AND JOHN D. KLENA†

Department of Microbiology, Arizona State University, Tempe, Arizona 85287-2701

| | |
|---|-----|
| INTRODUCTION | 655 |
| OVERVIEW | 656 |
| GENETICS OF LIPID A SYNTHESIS | 658 |
| Acylation of UDP-N-Acetylglucosamine | 658 |
| Central role of intermediates | 658 |
| <i>lpxA</i> and macromolecular synthesis II operon..... | 658 |
| Synthesis and Modification of Lipid A Disaccharide..... | 659 |
| <i>lpxB</i> gene | 659 |
| GENETICS OF LPS CORE | 660 |
| Synthesis and Attachment of KDO | 660 |
| <i>kdsA</i> and <i>kdsB</i> | 660 |
| <i>kdtA</i> | 661 |
| Heptose Synthesis and Inner Core | 661 |
| <i>rfaD</i> and <i>rfaE</i> | 661 |
| <i>rfaC</i> and <i>rfaF</i> | 661 |
| Other genes involved in lipid A or inner-core modification | 662 |
| Genes of Core Completion | 663 |
| <i>rfaG</i> <i>B</i> <i>L</i> <i>J</i> <i>K</i> family and hexose backbone..... | 663 |
| <i>rfaP</i> , <i>tolC</i> and deep rough phenotype..... | 664 |
| Core modification and O-antigen attachment..... | 666 |
| <i>rfaQ</i> <i>S</i> <i>Z</i> and core heterogeneity | 667 |
| Regulation of Core Synthesis | 669 |
| Arrangement of transcriptional units..... | 669 |
| Role of <i>rfaH</i> (<i>sfrB</i>) | 669 |
| Effect of temperature | 670 |
| GENETICS OF O-ANTIGEN SYNTHESIS | 670 |
| O-Antigen Genes of <i>rfaP</i> -Dependent <i>Salmonella</i> Serovars..... | 671 |
| Organization of <i>rfaB</i> in <i>S. typhimurium</i> | 671 |
| Sugar biosynthetic genes of the <i>S. typhimurium</i> <i>rfaB</i> cluster | 672 |
| Sugar transferases of the <i>S. typhimurium</i> <i>rfaB</i> cluster | 672 |
| Polymerization and role of the <i>rfaC</i> gene | 672 |
| Modification of O polysaccharide..... | 673 |
| Polymorphism among <i>rfaP</i> -dependent <i>Salmonella</i> serovars | 673 |
| O-Antigen Genes of <i>Shigella dysenteriae</i> 1..... | 674 |
| Genetic structure of the <i>rfaB</i> cluster..... | 674 |
| Roles of <i>rfaE</i> and <i>rfaP</i> | 674 |
| Role of <i>rfaX</i> | 674 |
| O-Antigen Genes of <i>E. coli</i> | 675 |
| Genetic analysis of <i>rfaB</i> region of <i>E. coli</i> K-12 | 675 |
| <i>rfaB</i> clusters from other <i>E. coli</i> serovars | 675 |
| Assembly and Translocation of LPS Containing an O Antigen | 675 |
| Regulation of O-antigen polymerization | 675 |
| Compartmentalization and translocation of O antigen | 675 |
| Attachment of O antigen to the LPS core | 676 |
| CONCLUDING REMARKS | 677 |
| ACKNOWLEDGMENTS..... | 678 |
| REFERENCES | 678 |

INTRODUCTION

Lipopolysaccharides (LPS) are essential components of bacterial endotoxins which define many of the properties of

host-parasite interactions, and they are the immunodominant antigens of most gram-negative bacteria. The early realization that these complex molecules had a unique structure to which the immunological and biological properties could be related led to a tremendous amount of interest and research effort in the 1950s and 1960s. As a result of this effort, by 1970 the LPS species produced by various *Salmonella* and *Escherichia coli* serovars had become the best-

* Corresponding author.

† Present address: Rocky Mountain Laboratories, National Institutes of Health, Hamilton, MT 59840.

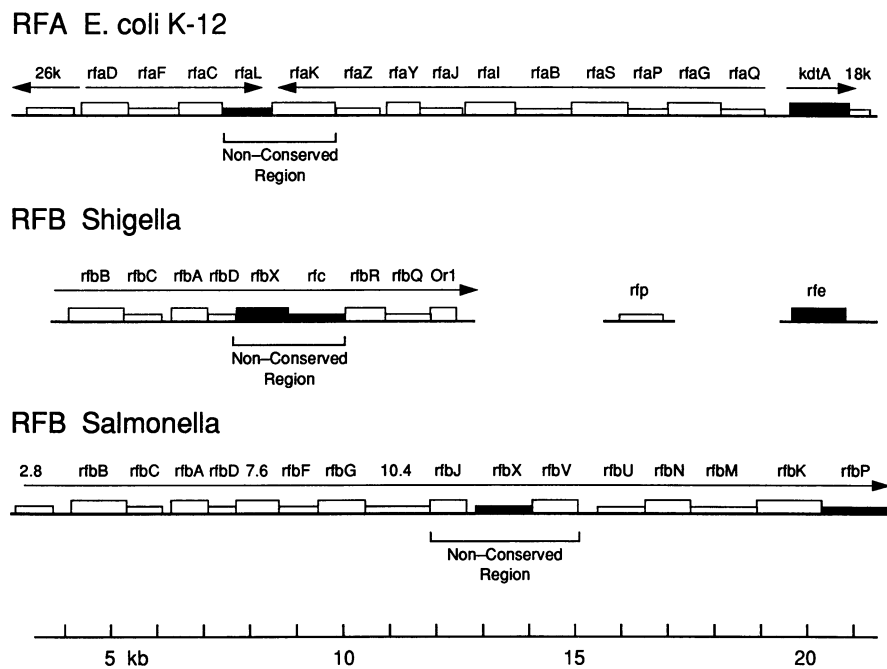


FIG. 2. Schematic diagrams, drawn to the same scale, showing the *rfa* gene cluster of *E. coli* K-12 and the *rfb* gene clusters of *S. dysenteriae* 1 and *S. typhimurium* LT2 (65, 74, 77). Coding regions are drawn as blocks, with solid blocks indicating proteins which are hydrophobic or have clearly identified membrane-spanning domains. In the *S. dysenteriae* 1 *rfb* cluster, two genes lying outside of the cluster are shown to the right. Arrows indicate the direction of transcription. Brackets indicate approximate regions containing genes which are not conserved across species or serovar lines. The scale indicates kilobase coordinates for the *Salmonella* *rfb* genes (65).

to be bound to the inner surface of the cytoplasmic membrane. None of the gene products have signal sequences typical of proteins translocated into the periplasm or outer membrane.

Second, each of the clusters contain some genes encoding products which are strongly conserved across strain and species lines and other genes which are not strongly conserved. In the *rfa* cluster the strongly conserved group includes most of the genes, since the products of *kdtA* and *rfaBCDFGPQYZ* all exhibit identity of 70% or more in amino acid sequence between *E. coli* K-12 and *S. typhimurium* LT2 (75, 76, 124). In the *rfb* clusters there is somewhat more genetic polymorphism, but an example of such a strongly conserved region is the *rfbBCAD* block, which encodes synthesis of a nucleotide sugar precursor common to many O antigens (65, 77). Another group of genes can be described as semiconserved, and these include the *rfaIJ* genes of *S. typhimurium* and *E. coli* (124) and some of the *Salmonella* *rfb* genes encoding sugar transferases (139). The proteins encoded by these genes are clearly related but show less sequence similarity (on the order of 25 to 50% identity) and can be considered to represent genes which have evolved to encode families of enzymes which catalyze similar reactions but recognize different substrates. Yet another group of genes encode proteins which have conserved structural features such as hydrophathy profiles but show very little conservation of primary amino acid sequence across serovar or species lines (less than 25% identity). It is likely that these proteins have evolved for the specific recognition of structural features of the core or O antigen which are variable among different species and serovars, and in cases when their function is known they appear to be involved in late stages of the assembly or

translocation of the polysaccharides. This category includes the genes *rfaK*, *rfaL*, *rfbX*, and *rfc* (21, 75, 77, 139).

In cases when comparable *rfa* or *rfb* clusters have been sequenced from different species or serovars, the products of the genes toward the center of the cluster are less strongly conserved whereas the products of the genes at the ends are more strongly conserved (75, 139). If each cluster is considered to be made up of a series of interchangeable modules, this arrangement of conserved regions flanking a nonconserved region may facilitate the exchange of modules within the clusters as a means of generating diversity while preserving basic biosynthetic functions. In several cases, it has been found that the nonconserved regions have a higher A+T content than the conserved regions, which in turn may have a higher A+T content than the flanking non-LPS genes or the chromosome of the organism as a whole (65, 75, 77). This suggests that these clusters may have arisen from a series of ancestral genetic exchanges involving very dissimilar organisms (139).

A simplified schematic overview of representative LPS structures and some of the genes involved in their synthesis is shown in Fig. 3. Only structures and gene functions which are well documented are shown in Fig. 3, and a number of substituents which are variable or for which the site of attachment is not known have been omitted. A much more detailed picture of a complete *S. typhimurium* LPS molecule based on a combination of physical measurements and computer reconstruction has recently been presented (71). Readers should use caution and remember that the LPS can be quite heterogeneous and that the structures presented usually represent average or consensus structures rather than distinct molecular species.

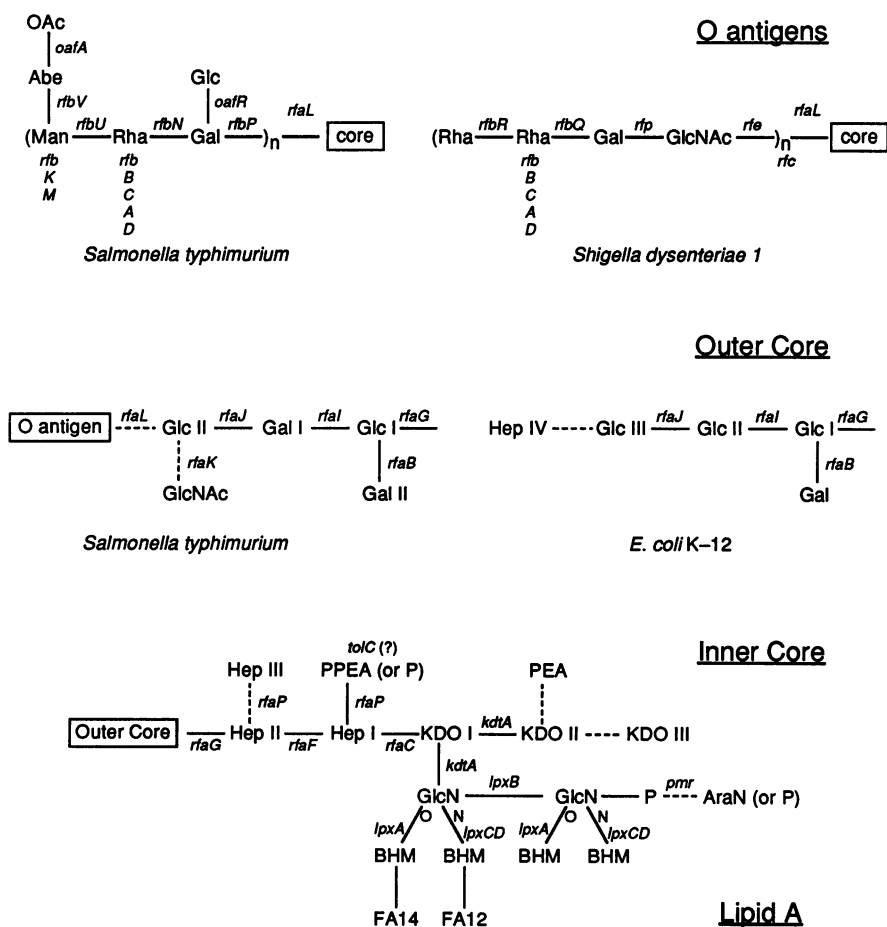


FIG. 3. Structures of the lipid A, inner-core, outer-core, and O-antigen regions of enteric bacteria and the genes involved in their synthesis. The lipid A-inner-core region is a composite of *E. coli* K-12 and *S. typhimurium* structures. Partial substituents are indicated by dashed lines, and substituents whose sites of attachment are not well documented are not shown. Structural information is from references 60, 62, 77, and 129. Genes involved in the biosynthesis of O-antigen sugars are shown below the respective sugars, and other genes are shown at their approximate site of action. The abbreviations BHM, FA12, and FA14 represent β -hydroxymyristate, laurate, and myristate substituents, respectively, and P represents phosphate. Other abbreviations are defined in the text.

GENETICS OF LIPID A SYNTHESIS

One of the most important aspects of LPS research in the past decade has been the elucidation in *E. coli* K-12 of the pathway of biosynthesis of a unique group of phospholipids based on glucosamine rather than glycerol, whose only known function is to serve as precursors of LPS. Biochemical details of the lipid A pathway have been covered extensively in recent reviews (127-131), so this section will be concerned primarily with the genetics of the pathway. A summary of the genes involved is provided in Table 1.

Acylation of UDP-N-Acetylglucosamine

Central role of intermediates. UDP-GlcNAc is a central intermediate in cell surface biosynthesis and serves as the common precursor of peptidoglycan and lipid A. It also plays other important roles in surface polysaccharide biosynthesis. Addition of GlcNAc to the LPS core plays a key role in the attachment of O antigen in *S. typhimurium* and perhaps in *E. coli* (75). The transfer of GlcNAc phosphate from UDP-GlcNAc to the polyisoprenoid lipid carrier undecaprenol phosphate (referred to here as antigen carrier lipid or ACL) is also involved in initiating the synthesis of

rfe-dependent O antigens and ECA (63, 102). Regulation of the synthesis of UDP-GlcNAc is not well studied, but it is noteworthy that *glmS*, the first gene involved in its synthesis from fructose 6-phosphate, is located in a cluster of genes adjacent to *oriC* at 84 min (5). R-3-Hydroxymyristoyl-acyl carrier protein (ACP), the other unique early precursor of lipid A, is also a common intermediate since it can be incorporated into both lipid A and glycerol-based phospholipids (132). The tightly interlocked network of common intermediates between the membrane lipid and polysaccharide pathways and the clustering of genes for the biosynthetic pathways together with other elements of macromolecular synthesis make it seem less than serendipitous that the first steps of lipid A biosynthesis were discovered during a search for genes involved in early steps of glycerophospholipid synthesis (164).

***lpxA* and macromolecular synthesis II operon.** The first committed step of lipid A biosynthesis is the transfer of the acyl group from R-3-hydroxymyristoyl-ACP to the 3-OH group of UDP-GlcNAc. This acyltransferase exhibits a remarkable substrate specificity which precisely determines the lipid composition at this substitution site (2). The acyltransferase gene is encoded by the *lpxA* gene at 4 min on the

TABLE 1. Genes known to be involved in the synthesis of lipid A, listed in the order in which their products are thought to act

| Gene | Function of gene products | Example and reference |
|-------------|--|-----------------------|
| <i>lpxA</i> | Transfer of β -hydroxymyristic acid from ACP to 3-hydroxy group of UDP-GlcNAc | <i>E. coli</i> (2) |
| <i>lpxC</i> | Gene formerly termed <i>envA</i> ; removal of <i>N</i> -acetyl group from UDP-3-hydroxymyristoyl-GlcNAc | <i>E. coli</i> (126) |
| <i>lpxD</i> | Gene formerly termed <i>fir</i> ; transfer of β -hydroxymyristic acid from ACP to 2-amino group of deacetylated GlcNAc to form UDP-2,3-dihydroxymyristoyl-glucosamine | <i>E. coli</i> (126) |
| <i>lpxB</i> | Condensation of one molecule of UDP-2,3-dihydroxymyristoyl-glucosamine with one molecule of lipid X (same compound with UDP removed) to form UMP and lipid A disaccharide with four fatty acids and a single phosphate at the reducing end | <i>E. coli</i> (137) |

E. coli K-12 map, as shown by the isolation of a tight conditional *lpxA* mutant which was temperature sensitive for growth and exhibited temperature-sensitive acyltransferase activity (45). Like conditional mutants in the biosynthesis and attachment of 2-keto-3-deoxyoctulosonic acid KDO (143, 144), this mutant demonstrates the essential nature of the LPS pathway. The *lpxA* gene (28) is part of a cluster of 11 genes organized into a complex operon which has been termed the macromolecular synthesis II operon (99, 166). In addition to *lpxA*, this remarkable operon includes *polC* (*dnaE*) which encodes the α core subunit of DNA polymerase III (167); the *mh* gene encoding RNase H; the *lpxB* and *lpxD* genes which also function in the lipid A biosynthetic pathway (35, 99, 126, 167); the *cds* gene which encodes CDP-diglyceride synthetase, an essential step in biosynthesis of glycerophospholipids (132); the *ompH* gene (also referred to as *skp* or *hlpA*) gene which encodes a 16-kDa basic outer membrane protein (80); and an as yet unnamed gene encoding a β -hydroxymyristoyl thioester dehydratase (58, 82, 126). Both of these latter genes are noteworthy. OmpH protein binds both LPS and DNA, and was initially described independently as an LPS-binding protein and as the histone-like protein HNL-1 (1, 82). The main function of OmpH protein remains unknown, although a role has been suggested in protein translocation (165). The myristoyl dehydratase is a 17-kDa protein which shows extensive sequence similarity to the β -hydroxydecanoyl thioester dehydratase encoded by *fabA*, an important enzyme in unsaturated fatty acid biosynthesis (34). The gene for the myristoyl dehydratase is located between *lpxA* and *lpxD* (see below), which encode acyltransferases, and leads to the unique situation in which the genes for three enzymes that independently utilize the same substrate (*R*-3-hydroxymyristoyl-ACP) are grouped together and presumably coregulated. The relevance of this regulatory scheme is indicated by the fact that missense mutations in the myristoyl dehydratase gene represented one class of temperature-resistant suppressors of the temperature-sensitive *lpxA2* allele (128).

The next steps in the lipid A pathway involve *N*-deacetylation of the *O*-acylated UDP-GlcNAc and the transfer of a second acyl group from *R*-3-hydroxymyristoyl-ACP to the amino group to form UDP-2,3-diacylglucosamine, which is the precursor of the nonreducing end of lipid A. Removal of UDP by either of two different pyrophosphatases results

in 2,3-diacylglucosamine-1-phosphate, which is otherwise known as lipid X (128, 129). Historically, it was the accumulation of lipid X in mutants defective in lipid A synthesis which provided one of the first clues leading to elucidation of the lipid A biosynthetic pathway (164).

Very recently, the two genes involved in *N*-acylation have been identified (126). The *N*-deacetylase is encoded by the *envA* gene (5), and it has been suggested that this gene be renamed *lpxC*. The first mutations in the *lpxC* (*envA*) gene were identified on the basis of a defect in cell division which was accompanied by a decrease in plasmid- or chromosomally encoded resistance to drugs such as ampicillin (114). The *lpxC* (*envA*) gene is located in a cluster of genes (5) involved in cell division (*ftsQAZ*) and protein secretion (*secA*).

The *N*-acylation is carried out by the product of the *fir* gene, which has now been renamed *lpxD* (126). Mutations in *lpxD* (*firA*) were first selected on the basis of rifampin sensitivity and independently on the basis of temperature sensitivity and enhanced sensitivity to multiple hydrophobic drugs. These latter mutants were termed *ssc* in *S. typhimurium* and *omsA* in *E. coli* and have recently been shown to be allelic to *lpxD* (*firA*) (171). The outer membrane phenotype of these mutants is entirely consistent with the identification of LpxD protein as an enzyme involved in lipid A synthesis, rather than a component of the transcriptional machinery as previously suggested (38, 39). LpxD protein shares an unusual region of homology termed an isoleucine patch with LpxA protein and other acyl- and acetyltransferases including the CysE and LacA proteins (39).

Synthesis and Modification of Lipid A Disaccharide

***lpxB* gene.** The reaction which forms the backbone of lipid A involves the condensation of one molecule of UDP-2,3-diacylglucosamine with one molecule of lipid X to form a β 1,6-linked glucosamine disaccharide in which each sugar is substituted with two 3-hydroxymyristoyl groups and which carries a single phosphate residue on the 1 position of the reducing sugar. This condensation reaction is catalyzed by the product of the *lpxB* gene (35). This enzyme is found primarily in the cytosolic fraction and can be assayed in the absence of detergent, although lipid X is found primarily in the membrane and the disaccharide product becomes very hydrophobic (137). The next step in the pathway is addition of a 4'-monophosphate by a specific kinase, generating the rather symmetrical disaccharide with a phosphate at either end which is known as lipid IV_A (128). This kinase reaction can be carried out in vitro with labeled ATP, and this provides a very effective method for the generation of radiolabeled precursors which can be used to study subsequent reactions (22).

These reactions lead to the first structure which exhibits some of the characteristic properties which define bacterial endotoxin. Although lipid IV_A does not exhibit all of the properties characteristic of endotoxin, it is an agonist of endotoxin and clearly has specific biological properties (49). The ability to use lipid IV_A as a precursor and to manipulate its synthesis opens exciting avenues in the development of antimicrobial agents and drugs designed to combat endotoxin shock. Although the structure of lipid A is clearly not the same among all gram-negative bacteria (97, 163), these reactions are probably widely distributed and conserved among many bacterial genera.

TABLE 2. Genes involved in synthesis of the LPS core, listed in the order in which their products are thought to act

| Gene(s) | Probable functions of gene product | Example(s) and reference(s) |
|----------------|---|---|
| <i>kdsA</i> | Synthesis of KDO-8-phosphate | <i>E. coli</i> (180), <i>S. typhimurium</i> (144) |
| <i>kdsB</i> | Synthesis of CMP-KDO from KDO-8-phosphate | <i>E. coli</i> (47) |
| <i>kdtA</i> | Sequential addition of backbone KDO I and branch KDO II residues to lipid A; possible attachment of branch KDO III to KDO II at later stage ^a | <i>E. coli</i> (12) |
| <i>rfaE</i> | Synthesis of ADP-heptose | <i>S. typhimurium</i> (156) |
| <i>rfaD</i> | Epimerization of ADP-heptose to L-glycero-D-manno form | <i>E. coli</i> (29) |
| <i>rfaC</i> | Addition of Hep I to KDO I of inner core | <i>S. typhimurium</i> (156) |
| <i>rfaF</i> | Addition of Hep II to Hep I to complete inner-core backbone | <i>S. typhimurium</i> (178) |
| <i>rfaH</i> | Positive regulator of <i>rfa</i> operon (<i>rfaQGPSBIJYZK</i>) and <i>tra</i> operon of plasmid F; may act by transcription antitermination; also termed <i>sfrB</i> | <i>E. coli</i> (125) |
| <i>rfaQ</i> | Addition of partial substituent (perhaps Hep III) to core ^a | <i>E. coli</i> (74) |
| <i>rfaG</i> | Addition of first backbone hexose (Glc I) to Hep II | <i>E. coli</i> (118), <i>S. typhimurium</i> (31) |
| <i>rfaP</i> | Required for transfer of phosphate from ATP to Hep I; may be involved in other core phosphorylation steps ^a | <i>S. minnesota</i> (106), <i>E. coli</i> (118) |
| <i>tolC</i> | Required for detergent resistance and LPS cross-linking; may be involved in conversion of core PPEA to phosphate ^a | <i>E. coli</i> (118) |
| <i>rfaB</i> | Addition of branch Gal to Glc I | <i>E. coli</i> (124), <i>S. typhimurium</i> (181) |
| <i>rfaI</i> | Addition of second backbone hexose (Glc II in <i>E. coli</i> K-12 and Gal I in <i>Salmonella</i> spp.) to core | <i>E. coli</i> (124), <i>S. typhimurium</i> (24) |
| <i>rfaJ</i> | Addition of third (terminal) backbone hexose (Glc III in <i>E. coli</i> K-12, Glc II in <i>Salmonella</i> spp.) to core | <i>E. coli</i> (124), <i>S. typhimurium</i> (24) |
| <i>rfaY</i> | Appears to be required for <i>rfaJ</i> action ^a ; exact function unknown | <i>E. coli</i> (124) |
| <i>rfaK</i> | Addition of GlcNAc to terminal Glc II in <i>Salmonella</i> spp.; may add GlcNAc to inner core in <i>E. coli</i> K-12 ^a | <i>E. coli</i> (74), <i>S. typhimurium</i> (93) |
| <i>rfaS</i> | Required for synthesis of LOS form of core but not for cores carrying O antigen; may be Rha transferase ^a ; absent in <i>S. typhimurium</i> LT2 | <i>E. coli</i> (74), <i>S. typhimurium</i> (76) |
| <i>rfaZ</i> | Required for synthesis of LOS form of core but not for cores carrying O antigen; exact function unknown; present in <i>S. typhimurium</i> LT2 but may not function ^a | <i>E. coli</i> (74), <i>S. typhimurium</i> (92) |
| <i>rfaBCAD</i> | Biosynthetic pathway for TDP-Rha; required for synthesis of LOS form of core | <i>E. coli</i> (78) |
| <i>rfaL</i> | Required for addition of O antigen to core and in <i>E. coli</i> K-12 for <i>rfe</i> -dependent addition of GlcNAc to terminal Glc III | <i>E. coli</i> (74), <i>S. typhimurium</i> (93) |
| <i>rfe</i> | GlcNAc-P ACL transferase required for addition of GlcNAc to terminal Glc III of <i>E. coli</i> K-12 | <i>E. coli</i> (77) |

^a Functions which are speculative or not well studied.

GENETICS OF LPS CORE

The arrangement of LPS genes into clusters and operons does not always parallel the biosynthetic pathway. For example, the attachment of the first two residues of KDO to the lipid A precursor IV_A, which involves a gene in the *rfa* cluster, takes place before the final acylation of lipid A. Likewise, although the genes in the *rfa* cluster are primarily involved in synthesis and modification of the LPS core, there are genes in this cluster such as *rfaK* and *rfaL* which are intimately involved in the attachment of O antigen to the core. Although the majority of the known genes whose functions are involved exclusively in LPS core synthesis are located in the *rfa* cluster, there are also genes such as *kdsA* and *rfaE* located outside of the *rfa* cluster which are involved in biosynthesis of sugars unique to the core or which appear to exert a direct effect on core structure. These are also included in the sections below. However, genes involved in biosynthesis and activation of sugar residues which are shared among multiple pathways such as biosynthesis of peptidoglycan or glycogen (44) have not been included. A summary of the genes involved primarily in core synthesis and their proposed functions is given in Table 2.

Synthesis and Attachment of KDO

***kdsA* and *kdsB*.** The genes involved in the biosynthesis of KDO are among those which are located outside of the *rfa* gene cluster. The gene *kdsA* encodes the KDO-8-phosphate

synthetase which combines D-arabinose 5-phosphate and phosphoenolpyruvate to generate KDO-8-phosphate and is part of a short operon located at 27 min (5, 179, 180). Following removal of the phosphate, free KDO is activated to CMP-KDO by the CMP-KDO synthase encoded by the *kdsB* gene at 85 min (5, 47).

Conditional (temperature-sensitive) *kdsA* mutants of *S. typhimurium* exhibit a stasis in growth under nonpermissive conditions. Although the cessation in KDO synthesis occurred promptly after a shift to a nonpermissive temperature, the stasis appeared only after a lag of almost a full generation. Surprisingly, this stasis had no deleterious effect and the mutants were able to resume growth after a shift to a permissive temperature (143, 144). These findings have several important implications. First, they provided the most direct evidence that LPS containing KDO is absolutely required for some cell surface assembly or cell growth process. Second, they indicate that continuous synthesis of KDO is not required and that cells can continue to grow until a pool of some cellular component is depleted to a critical level. It is not clear whether the critical component is a pool of LPS intermediates in the cytoplasmic membrane or whether there is simply an excess of LPS in the outer membrane which is sufficient to meet the needs of the cell until it is diluted by growth to a limiting concentration. Finally, in the absence of KDO synthesis the rate of translocation of lipid IV_A to the outer membrane is low. This is in contrast to heptose-deficient mutants blocked in a later step

of core assembly, in which translocation of a defective form of LPS to the outer membrane appears to be efficient (113). This suggests that attachment of KDO, or some step in LPS completion which is coupled to or depends upon the attachment of KDO, may be a minimum requirement for efficient translocation of LPS to the outer membrane. It has been possible to demonstrate that there is some transport of precursors which lack KDO to the outer membrane, but this appears to occur at a much lower rate (115). One set of reactions which are known to be coupled to the attachment of KDO are the final steps in acylation of lipid A (23, 129, 172). Unfortunately, nothing is known about the genes involved in terminal acylation, and no mutants blocked in these steps have been described. Thus it has not been possible to establish whether there is a relationship between the terminal acylation of lipid A and efficient translocation of LPS to the outer membrane.

kdtA. The *kdtA* gene is located near the right end of the *rfa* cluster (26, 154) and encodes a bifunctional enzyme (12) which sequentially transfers two molecules of CMP-KDO to the lipid A precursor lipid IV_A. This generates both the backbone KDO residue (KDO I) which is linked $\alpha 2,6'$ to the lipid A disaccharide and the first $\alpha 2,4$ -linked branch KDO residue (KDO II). The KdtA protein has a hydrophobic sequence capable of serving as a membrane anchor (27), and the isolation of enzymatically active (blue) TnphoA fusions at this site in *kdtA* indicates that this is a transmembrane segment (124). This is consistent with the observations that KdtA is capable of binding hydrophobic compounds and requires a detergent for solubilization (12).

The gene adjacent to *kdtA* encodes an 18-kDa polypeptide (27). This protein is not required for KDO transferase activity in vitro, and its function remains unknown. Present data indicate that *kdtA* and the gene for the 18-kDa polypeptide are transcribed as an operon, and that both gene products are essential for viability (126, 147).

It is generally thought that KDO II is substituted by an additional $\alpha 2,4$ -linked KDO (KDO III) which is added at a late stage of core completion (142). It should be emphasized that this is very difficult to study because of the labile nature of KDO and its glycosidic linkages, and the actual extent of substitution by KDO III is not known. The gene encoding the hypothetical KDO III transferase has not been identified, and there is the possibility that KdtA is a trifunctional enzyme which also catalyzes the addition of KDO III. This suggestion is based on the properties of a KDO transferase encoded by the *gseA* gene of *Chlamydia trachomatis* (11). The KdtA protein shares extensive regions of homology with the GseA KDO transferase, which is clearly trifunctional. In *E. coli* K-12, the *gseA* gene can complement a conditional *kdtA* mutant, and the expression of *gseA* in *E. coli* results in the addition of three KDO residues to the LPS core (16). Moreover, *E. coli* extracts containing GseA protein are capable of transferring at least one additional KDO residue to LPS bearing the two KDO residues added by KdtA (11). When *gseA* is expressed in *E. coli* in vivo, the addition of a third KDO residue appears to block the further addition of heptose residues, and it is the presence of this KDO III residue on a molecule lacking heptose which results in the unique *Chlamydia* epitope which allowed the *gseA* gene to be cloned in *E. coli* K-12 (16). The inhibition of heptose attachment explains why in wild-type *E. coli* K-12 any additional KDO residues would have to be added after the partial completion of the heptose inner-core region. The addition of KDO III is not essential in *Salmonella* species, since *galE* mutants which lack core sugars distal to the first

hexose produce LPS which contains only two KDO residues (84) and these strains grow normally.

Heptose Synthesis and Inner Core

rfaD and rfaE. Although the genetics of the complete pathway (142) which has been proposed for the synthesis of L-glycero-D-manno-heptose from sedoheptulose-7-phosphate has not been established, the genes for the last two steps in the pathway are known. Activation of heptose occurs before the final epimerization step, and the *rfaE* gene of *S. typhimurium* (93) which is located outside of the *rfa* cluster at 64 min appears to be involved in ADP-heptose synthase activity. This is based on in vitro experiments in which the addition of heptose to KDO₂-lipid IV_A by an extract from an *rfaE* mutant could be rescued by the addition of ADP-heptose (156).

The *rfaD* gene has been cloned and sequenced both in *S. typhimurium* (150) and in *E. coli* K-12 (120), and the cloned *E. coli* K-12 *rfaD* gene can efficiently complement *S. typhimurium* mutants (29, 150). The RfaD protein is strongly conserved between the two organisms, but it does not share homology with other Rfa proteins whose sequence is known. *rfaD* is located near the left end of the *rfa* cluster, where it is the first of a block of four genes (*rfaDFCL*) which make up a small operon (147). Mutants lacking *rfaD* are defective in the epimerization of ADP-D-glycero-D-manno-heptose to ADP-L-glycero-D-manno-heptose. Such mutants accumulate ADP-D-glycero-D-manno-heptose and synthesize defective LPS which lacks LD-heptose but contains a small amount of the incorrect DD form (29). LPS from an *rfaD* knockout insertion mutant appears heterogeneous on sodium dodecyl sulfate (SDS)-gels with a major band comigrating with heptose-deficient LPS and some minor, more slowly migrating species which are assumed to result from the incorrect incorporation of DD-heptose (117).

There is a gene encoding a 26-kDa protein adjacent to *rfaD* at the left end of the *rfa* cluster (134, 154). The 26-kDa protein product shares a region of homology with bacterial kinases, and insertions into this gene exhibit a detergent-sensitive phenotype, indicating that it is probably involved in LPS synthesis (133). Its exact function is unknown, but it is tempting to speculate that it might be involved in heptose phosphorylation.

rfaC and rfaF. The RfaC and RfaF proteins share regions of sequence homology (150), and both proteins share somewhat less abundant regions of homology with RfaQ and KdtA (26, 27, 119). On the basis of homology, these four proteins could be considered to belong to a family of related proteins, all presumably sugar transferases, which all may act in the lipid A inner-core region. Unlike KdtA protein, RfaQ, RfaF, and RfaC proteins do not have a significant transmembrane domain capable of serving as a membrane anchor. One can argue that a membrane-binding domain is more essential in KdtA protein, since the final steps of acylation of lipid A are not completed until after the addition of KDO (23, 172). The RfaC and RfaF proteins of *S. typhimurium* share strong homology (more than 70% identity) with the corresponding proteins of *E. coli* K-12, and a region near the C terminus of RfaC is nearly identical between the two organisms (75). The cloned *rfaC* gene of *E. coli* K-12 is capable of very efficiently complementing an *rfaC* mutation in *S. typhimurium* (74).

An extract from an *rfaC* mutant is defective in the addition of heptose to KDO₂-lipid IV_A in vitro (156), and this supports earlier evidence (93) that *rfaC* might be the gene

encoding the heptosyltransferase which adds the first α 1,5-linked backbone heptose residue (Hep I) to KDO I. LPS from mutants lacking *rfaF* exhibits a migration on SDS-gels intermediate between that of completely heptose-deficient LPS from an *rfaC* mutant and hexose-deficient LPS from an *rfaG* mutant (117) and is intermediate in heptose content (178). These results, together with the homology between RfaC and RfaF proteins, suggest that *rfaC* and *rfaF* are both structural genes for heptosyltransferases and that RfaF is the transferase which adds the second α 1,3-linked backbone heptose (Hep II) to the inner core (Fig. 3).

Heptose-deficient mutants have a very deleterious deep rough phenotype, and many of those which have been reported appear to be quite leaky in terms of heptose content. Thus it was of interest to determine the null phenotype of heptose-deficient mutants of *E. coli* K-12, since the possibility existed that a small amount of LPS containing heptose might be necessary for some essential cell function. Neither an Ω insertion in *rfaC* nor a deletion extending from a restriction site in *rfaD* to a restriction site in *rfaG* which removed almost all of the *rfa* cluster was lethal when crossed onto the chromosome (73). The mutants exhibited a very mucoid deep rough phenotype and produced a truncated LPS which exhibited a single band on SDS-gels and comigrated with LPS from an *rfaC* mutant of *S. typhimurium*. To test that the mucoid phenotype was due to the production of colanic acid and to show that mucoidy did not suppress a potentially lethal phenotype resulting from the heptose deficiency, the *rfaC::\Omega* null mutation was transferred into a strain carrying a null *cps::Tn10* mutation. In this background, the *rfaC* null mutation exhibited a "crunchy" colony phenotype similar to that described for other deep rough mutants (118) and otherwise exhibited a typical deep rough phenotype. This *rfaC cps* double mutant was not temperature sensitive and grew over a range of 30 to 42°C (73).

Other genes involved in lipid A or inner-core modification. Although many of the intermediates in the synthesis of the lipid A-*inner-core* region of enteric bacteria have been identified within the last decade, the exact structures of lipid A and the inner-core region of complete LPS molecules are still not known in *E. coli* K-12 or *S. typhimurium*. There are several reasons for this. First, the amphipathic nature of intact LPS makes it insoluble in most useful solvents, and this prevents high-resolution ^{31}P - or ^1H -nuclear magnetic resonance analysis and limits analysis by electrospray or fast atom bombardment mass spectrometry. Second, many of the bonds involved in inner-core structures are labile and are broken by mild hydrolysis or deacylation procedures, which are required to convert the molecules to forms more amenable to study. Finally, there is the ever-present problem of heterogeneity, since not all molecules carry the same substituents and the nature of the substituents may change dramatically in response to environmental conditions. Since LPS molecules are relatively stable once synthesized, a culture may carry its entire history imprinted in the structures present in its LPS.

The modifications which are best characterized include a 4-aminoarabinose attached to the 1-phosphate of lipid A (158) and a phosphorylethanolamine (PEA) substituent attached to the 7 position of KDO (61). This PEA substituent is present in deep rough mutants (84) which are thought to contain only two KDO residues, so the PEA is almost certainly attached to KDO II. It is unclear whether additional substituents of PEA or pyrophosphorylethanolamine (PPEA) are attached elsewhere in the inner core. At least

one phosphate is always present at the 4' position of lipid A (158), but this may be modified further to PPEA (129, 158).

The *pmrA* gene at 95 min of the *S. typhimurium* map (151) affects the addition of 4-aminoarabinose to the 1 phosphate of lipid A, with the mutant phenotype being an increase in the percentage of molecules which carry the amino sugar. This mutation was selected on the basis of resistance to polymyxin (93), and it has been suggested that reduction in the negative charge of the lipid A directly reduces its ability to bind polymyxin. The exact nature of *pmrA* is not known; it may be a regulatory mutation or exert its effect indirectly (93).

It has been observed that the phenotype of deep rough strains carrying a deletion removing genes *rfaGPSBI* is not stable, that these strains gradually acquire a reduced hypersensitivity to detergents and novobiocin, and that complete hypersensitivity could be restored by transducing the *rfa* deletion into a fresh background (118). This suggests that these strains are accumulating suppressor mutations unlinked to the *rfa* cluster, and analysis of these suppressors might yield more mutations of the *pmrA* type which affect inner core substituents.

Two interesting new genes which may affect inner-core structure were identified during a screen for genes which are necessary for growth of *E. coli* at elevated temperature. The *htrB* gene at 23.4 min encodes a basic polypeptide with a membrane-spanning domain (67, 69). Knockout mutations of *htrB* exhibit a complex phenotype, including reduced sensitivity to deoxycholate and inability to grow in rich media at temperatures above 32.5°C with lysis and morphological changes occurring above that temperature. The *msbB* gene at 40.5 min (70) was first identified as a multicopy suppressor of the temperature-sensitive phenotype of *htrB*. In an *htrB*⁺ background, *msbB* null mutants exhibit an increased resistance to deoxycholate but no growth phenotype. However, an *htrB msbB* double mutant became even more temperature sensitive than *htrB* alone and showed additional morphological changes. In an *rfa*⁺ background *htrB* did not affect the mobility of LPS on SDS-gels but resulted in a rather dramatic decrease in silver-staining intensity similar to that which is seen with LPS from heptose-deficient mutants. In a strain background carrying a deep rough $\Delta rfaGPSBI$ mutation, the *htrB* mutation still exhibited a temperature-sensitive phenotype and a decrease in LPS staining intensity (117). In contrast, the *msbB* mutation caused a significant alteration of LPS gel migration, shifting the bands to a faster migration equivalent to a loss of one or two sugars from the molecules. It did not affect the migration of the truncated LPS in a $\Delta rfaGPDBI$ background but caused a reduction in staining intensity similar to that seen with the *htrB* mutant (117). The reason for the reduction in staining intensity is not known, but it may reflect changes in the inner-core or lipid A substitution pattern which affect the generation of aldehyde groups by periodate cleavage.

As in the case of the *pmrA* mutants, the cause of the *htrB* and *msbB* effects is not clear and may reflect either direct or secondary effects. Recent studies of *htrB* indicate that it affects regulation of lipid synthesis (68), and it was suggested that it may be part of a link between growth regulation and the regulation of lipid and LPS synthesis. However, given the lack of knowledge concerning the genetics and the function of inner-core substituents and the regulation of their synthesis, these mutations warrant further investigation.

Genes of Core Completion

***rfaGBIJK* family and hexose backbone.** The genes *rfaGBIJK* are involved in the sequential attachment of glucose and galactose residues to the inner core to provide the structural framework for the outer, or hexose, region of the core. The *rfaK* gene is involved in the addition of GlcNAc to the terminal core glucose in *S. typhimurium* and a substituent (possibly also GlcNAc) to a different core site in *E. coli* K-12. Since this substitution plays a somewhat different role, the functional aspects of *rfaK* are described in the next section. In *E. coli* K-12 the protein products of *rfaGBIJK* share limited regions of sequence homology and, on a more subjective basis, appear to have a similar structure based on visual comparison of hydropathy plots (119, 124). There is direct evidence only for RfaI and RfaJ that the proteins are sugar transferases, and this evidence is not compelling (124). However, the similarity of these proteins, the lack of evidence for any other genes involved in the transferase functions, and the agreement of the properties predicted for these proteins with the properties of those sugar transferases which have been purified leave little doubt that these gene products are the transferases. The predicted proteins are similar in size, and all lack significant hydrophobic regions which could function as membrane-binding domains. They do not have signal sequences, and they yield enzymatically active (blue) fusions on mutagenesis with *TnlacZ* but only inactive (white) fusions with *TnphoA* (124). Their structure and the fact that they act on membrane-bound LPS precursors suggest that they are peripheral membrane proteins attached to the inner surface of the cytoplasmic membrane through their interaction with their LPS substrate and other membrane components.

A galactosyltransferase thought to be the product of either *rfaI* or *rfaB* and a glucosyltransferase thought to be the *rfaG* product have been extensively purified and characterized from *S. typhimurium* (142). Both the galactosyl- and glucosyltransferases were shown to form enzymatically active ternary complexes with acceptor LPS and phospholipid. These complexes were specific since the LPS had to have a structure which allowed it to function as an acceptor for the appropriate sugars and both enzymes showed a strong preference for phosphatidylethanolamine (PE) as the phospholipid. In addition, an increase in lateral pressure was observed in an LPS-phospholipid monolayer when the galactosyltransferase was added, indicating that the enzyme was capable of intercalation into the monolayer. This led to a "ducks-on-a-pond" model for sugar transferases (summarized in reference 142), in which the individual transferases were envisioned as floating randomly on the cytoplasmic surface of a fluid phospholipid bilayer and interacting with acceptor LPS in the bilayer by lateral diffusion. Although there is nothing in the structure of the transferases to preclude such a model, more recent genetic studies suggest that in vivo there may also be significant interaction between the transferases (124), and it may be more appropriate to consider the active unit in vivo as a more structured or stoichiometric complex of enzymes acting in sequential fashion on bound or laterally compartmentalized substrates. Such structured complexes could explain why DNA fragments containing one or a few *rfa* genes are sometimes very difficult to manipulate in multicopy plasmids, a problem which everyone studying this region seems to have encountered.

The *rfaG* gene of *E. coli* K-12 can efficiently complement *rfaG* mutations in *S. typhimurium* (31, 119). The gene has not

yet been sequenced in *S. typhimurium*, but this result and the similarity of the inner-core structure between the two organisms (60, 142) suggest that their RfaG proteins are likely to exhibit significant homology. There is strong chemical evidence which indicates that in both *Salmonella* spp. and *E. coli* K-12, RfaG is involved in the transfer of glucose from UDP-glucose to Hep II (31, 118). Complementation studies of long *rfa* deletions and the phenotypes of nonpolar mutations in *rfaP* and *rfaQ* indicate that RfaG protein does not require the products of either of these flanking genes for its function (118, 124, 125). This clarifies previous results obtained with polar *rfaQ::Tn10* mutations which exhibited an RfaG⁻ phenotype (3) and older biochemical studies which suggested that *rfaP* acted before *rfaG* (93).

Mutants deleted for *rfaG* are nonmotile and completely lack flagella and type I pili, and this effect is independent of the presence of *rfaP* (118). The loss of flagella is also observed with mutants defective for *galU*, which is necessary for the synthesis of UDP-glucose, and this loss has been interpreted as being due to a defect in transcription of flagellar genes (79). This suggests that the cell has a system which is involved in regulating the expression of flagella (and perhaps pili or other surface components) and which is capable of sensing either the degree of completion of the LPS core or the presence of the RfaG protein and its substrate, UDP-glucose. In this context it should be noted that the secretion of TolC protein is also defective in a *galU* mutant (174). The consequences of the loss of TolC are described in more detail in the next section.

In *E. coli* K-12 the second backbone hexose is glucose, whereas in *Salmonella* spp. it is galactose (142). Despite the difference in sugar composition, the following gene assignment is used in both organisms: *rfaB* is involved in addition of the α 1,6-linked branch galactose to glucose I (Glc I), *rfaI* is involved in addition of the second α 1,3-linked backbone hexose (either glucose or galactose), and *rfaJ* is involved in the addition of the terminal α 1,2-linked glucose (termed Glc II in *Salmonella* spp. and Glc III in *E. coli* K-12). This replaces an earlier terminology in *E. coli* K-12 in which *rfaM* and *rfaN* were used instead of *rfaI* and *rfaJ*, respectively (3, 31). The *E. coli* K-12 gene designations were dropped when it became evident that the arrangement of the genes and the sequence of the proteins they encoded were very similar in both organisms (124, 154).

Since *rfaB* is involved in the attachment of a branch sugar, the phenotype of *rfaB* mutants is not obvious in either organism. Since this branch sugar is the only galactose in the *E. coli* K-12 outer core, it would be predicted that the phenotype of an *rfaB* mutation would be identical to the phenotype of a null Δgal mutation. This has now been confirmed (124). Somewhat surprisingly, the major LPS band seen in *rfaB* and Δgal null mutants comigrates with LPS which contains only one hexose and not with LPS from *rfaI* mutants, which contains two hexoses. This indicates that in *E. coli* K-12 the RfaI protein cannot efficiently add Glc II in the absence of the branch galactose, regardless of whether the defect is in *rfaB* or in synthesis of UDP-galactose. Both *rfaB* and Δgal mutants produce minor LPS bands which migrate as though they have larger cores, but these are absent in mutants in which *rfaJ* is also absent, and it has been suggested that the absence of the branch galactose on Glc I may cause RfaJ protein to mistake it for Glc II and inefficiently add a terminal glucose (124). This might also explain anomalous phage P1 sensitivity results (124, 148), since this phage is thought to recognize an unsubstituted glucose as one of its receptors.

S. typhimurium rfaB mutants which appear to be completely lacking the branch galactose are able to complete a substantial number of core molecules and even add some O antigen (181). This is consistent with the unique phage resistance phenotype of *rfaB* mutants (93); these are resistant to phage FO, which requires the GlcNAc substituent attached to the terminal Glc II, but remain somewhat sensitive to smooth-specific phages such as P22. This would suggest that the stringency of RfaI with respect to its requirement for the branch galactose is different between the two organisms. However, an alternative explanation is that the RfaB and RfaI proteins function as a complex and that the *S. typhimurium rfaB* mutants were missense mutants which produced proteins that lacked transferase activity but were still capable of forming a complex with RfaI (124). The complex would have to include UDP-galactose to explain the similar phenotypes of Δgal and *rfaB* null mutations in *E. coli* K-12. To resolve this discrepancy, there is clearly a need to reexamine the biochemical properties of these galactosyl- and glucosyltransferases now that defined mutations are available.

Despite the differences in core structure, DNA fragments carrying *rfaIJ* of *E. coli* K-12 expressed from either the physiological *rfa* promoter or fortuitous plasmid promoters can efficiently complement *rfaI* mutants of *S. typhimurium* as measured by completion of a core which allows attachment of O antigen (124). This indicates that the final steps of core completion and O antigen attachment in *S. typhimurium* are indifferent to whether the second backbone hexose residue is galactose or glucose.

In contrast, only an *rfaIJ* DNA fragment from *E. coli* K-12 which also carries the physiological promoter region for these genes can efficiently complement an *rfaJ* mutant of *S. typhimurium* (124). The most convenient explanation is that *S. typhimurium* RfaI protein competes effectively with an inefficiently expressed *E. coli* K-12 RfaI protein and that this results in the synthesis of *Salmonella* cores containing galactose in the backbone, which cannot be extended by the *E. coli* K-12 RfaJ transferase. This provides the evidence noted above that *rfaI* and *rfaJ* are the structural genes for the hexosyltransferases. Assuming that this is correct, the change in substrate specificity of the RfaI and RfaJ transferases between *E. coli* K-12 and *S. typhimurium* must be due to relatively minor changes in protein structure, given the strong conservation of the sequence of these proteins between the two organisms (24, 124).

In *E. coli* K-12, the change in electrophoretic mobility of the LPS accompanying the *rfaIJ*-mediated completion of the core is much greater than can be accounted for by the addition of two hexose residues. This is also the stage in core completion at which heterogeneity becomes evident, as indicated by multiple LPS gel bands (124). This suggests that there are other modifications of the core (perhaps additional KDO, Hep, rhamnose [Rha] or PPEA substituents) whose addition is coupled to the completion of the hexose backbone. The genes involved in these modifications are not known, although it is likely that *rfaY* (154) is one such gene. The phenotype of defined *rfaY* mutations is not known, but complementation studies of deletions which include *rfaY* suggest that *rfaY* may be necessary for *rfaJ* function (124). It is possible that some mutations in *S. typhimurium* identified as *rfaJ* are actually in *rfaY*, since the genes are adjacent in the chromosome and mutations in either gene might have the same phenotype. The sequence of the C-terminal portion of RfaY protein is known in both *S. typhimurium* and *E. coli*

K-12 (75, 92), and RfaY appears to be strongly conserved between these organisms.

***rfaP*, *tolC* and deep rough phenotype.** The deep rough phenotype is a complex phenotype exhibited by a subset of *rfa* mutants. In addition to alterations in the structure of the LPS, the phenotype includes the following: (i) sensitivity to detergents and hydrophobic antibiotics such as actinomycin D and novobiocin; (ii) elevation of the phosphatidylethanolamine (PE) content of the outer membrane to the point where regions of phospholipid bilayer, which are absent in the wild-type outer membrane, can readily be demonstrated; (iii) a dramatic (>90%) reduction in porin proteins, including OmpF, OmpC, PhoE, the Lc porin of lambdoid phages, and LamB, together with an approximately 50% reduction in OmpA protein; (iv) resistance to the LPS-specific phage U3 and the OmpF-specific phage K20; and (v) a decrease in buoyant density of the outer membrane (3, 113, 118). Many (if not all) deep rough mutants also exhibit a mucoid phenotype as a result of the induction of production of a colanic acid capsular polysaccharide. In a deep rough $\Delta rfaGPSBI$ mutant of *E. coli* K-12 which has been well studied, this induction of capsule synthesis remained temperature and Lon modulated and required the *rcsC* component of the *rcsBC* two-component regulatory system (118). Mucoidy has not always been reported as part of this phenotype, and this may reflect the contribution of other factors such as temperature and carbon source to the rather subjective observation of colony mucoidy (118). In *E. coli* K-12 strain backgrounds carrying null *cps::Tn10* mutations, deep rough mutants invariably exhibit a brittle, "crunchy" colony morphology and a strong tendency to autoagglutinate in liquid culture (118).

The genetics of the deep rough phenotype are now well understood. Mutations resulting in defects in the backbone heptose region of the inner core (*rfaCDEF*) and *rfaP* mutations, which affect modification of the heptose region, exhibit a deep rough phenotype (56, 93, 113). Strongly polar mutations in *rfaQ* and *rfaG* also result in a deep rough phenotype (3), but this is probably because they prevent expression of *rfaP*. Strains with nonpolar *rfaQ* null mutations are sensitive to phage U3 and resistant to novobiocin (74), and complementation of a $\Delta rfaQGPSBIJ$ mutation with a fragment carrying only *rfaP*⁺ restored sensitivity to phage U3 and the ability to grow on EMB agar, as well as substantially increasing resistance to novobiocin (118). These data indicate that the loss of *rfaP* or the inner-core heptose region on which *rfaP* acts results in the deep rough phenotype.

The exact function of *rfaP* and the mechanisms involved in the deep rough phenotype are less clear. Although loss of *rfaP* may reduce the efficiency of core completion, nonpolar *rfaP* mutants produce a heterogeneous LPS, indicating that they are still capable of a considerable amount of core completion and, in *S. typhimurium*, some O-antigen attachment (56, 118). Analysis of LPS from $\Delta rfaGPSBI$ mutants complemented with DNA fragments containing *rfaG* and *rfaP* indicate that there are two distinct inner-core functions associated with *rfaP*⁺ (118). One of these is the attachment of a phosphoryl substituent (either phosphate or PPEA) to Hep I. The other is the attachment of a branch Hep III to the Hep II residue, a substitution which is present on about half of the molecules. Since all molecules are phosphorylated on Hep I, the phosphorylation has been considered to be the primary function of *rfaP*, but direct evidence to support this is lacking (93). The observation that the γ phosphate from ATP can be transferred to the Hep I residue of LPS from an RfaP⁻ strain by an Rfa⁺ extract (106) indicates that a kinase

adds phosphate as the first step in the generation of the phosphoryl substituent, but unfortunately it does not indicate that *rfaP* encodes the kinase. One problem is that the RfaP protein shows no relationship to known kinases and does not contain a consensus nucleotide-binding site (119). Another is the specificity of the kinase. It is generally thought that another PPEA residue is added to the core late in completion (142), and both the detergent-sensitive phenotype and the phosphate content of the LPS core of *rfaP* mutants suggest that this additional phosphoryl substituent is also absent (3, 56). It seems unlikely that a single kinase would be involved in phosphorylation at two different sites. An interesting possibility which has not been tested is that the 26-kDa protein encoded by the gene adjacent to *rfaDFC* heptose cluster is the Hep I kinase and that the partial detergent-sensitive phenotype of null mutations in this gene results from the ability of one or more additional kinases which add phosphate at other sites to suppress the detergent sensitivity. The RfaP protein might be a common subunit which is required for several kinases, or it may have another function entirely to which the kinase activity is coupled.

The steps leading to the generation of PPEA are unclear. Kinetic evidence indicates that PE is the precursor of both the phosphate and the ethanolamine of the phosphorylethanolamine moiety of PPEA (54), and it has been proposed that the PPEA substituent of LPS arises by transfer of the phosphorylethanolamine moiety from PE to phosphate on Hep I (and probably other sugars) with the resulting generation of diacylglycerol, which is recycled into the phospholipid pool by diacylglycerol kinase (132). However, diacylglycerol kinase in enteric bacteria has not been studied extensively, and the proposed transfer of the phosphorylethanolamine moiety from PE has not been demonstrated *in vitro*.

Although there are still gaps in our knowledge, enough is known about the function of *rfaP* to allow speculation on how its loss results in the various effects which make up the deep rough phenotype. A useful approach is to attempt to dissect these effects and analyze them separately, and one way of doing this involves null mutations in the *tolC* and *pss* genes.

The *tolC* gene at 66 min of the *E. coli* K-12 map encodes a minor outer membrane protein (105). *tolC* mutants exhibit a complex phenotype including tolerance to colicin E1, reduced expression of the OmpF and Lc porin proteins, a failure to secrete hemolysin (173), resistance to the LPS-specific phage U3 (3), and a sensitivity to detergents and hydrophobic antibiotics which is quantitatively and qualitatively identical to that of deep rough mutants. Unlike deep rough mutants, *tolC* mutants do not produce truncated or heterogeneous LPS, instead showing a slight sharpening of LPS bands on SDS-gels (152), and there is no alteration of the buoyant density of the outer membrane and no overall defect in expression of outer membrane proteins (3). The inhibition of OmpF protein expression is a regulatory effect and can be suppressed by inactivation of the *micF* gene (104). Interestingly, the recent study showing a defect in export of the TolC protein by strains with *galU* mutations (174) suggests that the *tolC* function may be affected by LPS structure. Although this complicates interpretation of studies of mutants with inner-core defects, it may explain some earlier indications that *galU* mutants might exhibit a partially deep rough phenotype.

An insertion in *tolC* has a profound effect on the properties of LPS. The purified TolC⁻ LPS is much more easily dispersed in water than is TolC⁺ LPS, and a micellar

dispersion of TolC⁻ LPS is not flocculated by the addition of divalent cations. Preliminary characterization has been done on LPS from a *tolC* mutant by using the same strain background ($\Delta rfaGPSBI$ complemented with a DNA fragment carrying *rfaGP*) and chromatographic techniques which were used to study the effect of *rfaP* (152). When the TolC⁻ core oligosaccharide fraction was compared with that of TolC⁺ LPS by gel filtration, peak II, which includes cores containing Hep I-PPEA, was increased in the TolC⁻ LPS and peak IV, which in TolC⁺ LPS consists primarily of cores containing Hep I-phosphate, was absent in the TolC⁻ LPS. Although the exact composition of these core fractions has not been determined, these results and the lack of flocculation by divalent cations indicate that Hep I of the TolC⁻ LPS may have only PPEA while Hep I of the TolC⁺ LPS has both phosphate and PPEA. These results suggest that TolC may be involved in the conversion of PPEA to phosphate after the LPS is exported to the outer membrane. Moreover, these results suggest that the phosphate residue on Hep I plays an important role in binding divalent or polyvalent cations (particularly in the absence of phosphorylated substituents added at later stages of core completion), a function which is critical for the lateral cross-linking of the outer leaflet of the outer membrane to prevent penetration by hydrophobic compounds (113), and that the presence of the bulky ethanolamine shields the phosphates to make them less effective in cross-linking. The role of TolC in generating chelating residues could explain results of studies with *E. coli* O111 strain J5 which showed that newly synthesized LPS was rapidly converted from an EDTA-nonreleasable form to an EDTA-releasable form after it was exported to the outer membrane (86, 111). The lack of a major effect of *tolC* mutations on outer membrane proteins or on the buoyant density of the outer membrane indicates that lack of the TolC function does not prevent normal translocation or assembly of major outer membrane proteins or result in a gross perturbation of the outer membrane PE content. An interesting prediction of this hypothesis is that organisms such as *Neisseria* spp., which appear to be naturally hypersensitive to detergents and hydrophobic compounds, would lack a TolC analog. LPS core oligosaccharides from some strains of *Neisseria gonorrhoeae* have been shown to lack free phosphate and contain only PPEA and PEA (46, 66, 122).

A null mutation has recently been isolated in *pss*, the gene encoding phosphatidylserine synthetase, a key enzyme in the pathway for the synthesis of PE (37). In addition to lacking PE, the major phospholipid of *E. coli*, the *pss* mutant should lack ethanolamine, since the only known pathway for ethanolamine biosynthesis in *E. coli* is that involving the conversion of phosphatidylserine to PE (37, 132). Although a very minor alternative pathway for the synthesis of PE cannot be strictly ruled out, the amount of PE synthesized in the mutant is less than 1% of the wild-type level, and this is far below the amount which would be necessary to supply even a single ethanolamine residue on each LPS molecule. Thus this *pss* mutant can be used to study the role of ethanolamine in LPS synthesis.

The *pss* mutant appears to have a normal amount of LPS in its outer membrane, but the LPS bands migrate significantly faster (117). This is what would be expected as the result of an increase in net negative charge and a loss of the bulky ethanolamine substituents. The *pss* mutant strain produced normal amounts of porin and OmpA protein and was not sensitive to detergents or novobiocin. These results indicate that addition of ethanolamine cannot be essential for

the translocation of LPS to the outer membrane and that ethanolamine is not the ligand which is required for assembly of the major outer membrane proteins. The requirement of the *pss* mutant for a very high level of Mg^{2+} (>100 mM) for growth (37) may be due to the difficulty of translocating LPS to the outer membrane when it lacks the bulky ethanolamine residue as a blocking group on critical phosphate residues. It is interesting in this context that Mg^{2+} can be replaced by nontransportable ions such as Ca^{2+} , indicating that the effect of the divalent cations is at the cell surface (37). Another interesting question with respect to the LPS of the *pss* mutant is the nature of the phosphorylated substituent present on KDO, since nothing is known about the mechanism of synthesis of PEA.

The results with the *rfaP*, *tolC*, and *pss* mutations suggest the following model to explain the phenotype of deep rough mutations. First, as noted above, the ability of divalent and polyvalent cations to cross-link LPS must play a critical role in resistance to detergents and hydrophobic agents. At minimum, this requirement can be satisfied by a phosphate substituent on Hep I, and this requires both RfaP⁺ for the synthesis of heptose-linked PPEA and TolC⁺ for the conversion of PPEA to free phosphate after export of the LPS to the outer membrane. Second, some substituent in the heptose region of the core is required for the LPS to serve as a ligand for major outer membrane proteins and to promote their translocation and/or folding in the outer membrane. It is possible that this requirement can be satisfied by either phosphate or PPEA attached to Hep I. Alternatively, it could be the branch Hep III or some other substituent on KDO or lipid A whose attachment depends on RfaP function. Third, the accumulation of PE in the outer membrane and the accompanying decrease in buoyant density probably reflects a combination of two factors: decreased PE turnover due to the lack transfer of ethanolamine to LPS and a defect in the ability to maintain the phospholipid asymmetry of the outer membrane because of a reduction in LPS cross-linking and a lack of space-filling outer membrane proteins. Finally, cross-linked LPS or LPS containing phosphate as a substituent must be required for the LPS to be a receptor for phage U3. All of these features require the backbone heptose residues, which is why heptose-deficient mutants also exhibit a deep rough phenotype.

Core modification and O-antigen attachment. The R LPS produced by *rfaL* or Δ *his-rfb* mutants of *S. typhimurium* LT2 runs as a rather broad band on SDS-gels even when the Tricine buffer system (85) is used. In contrast, the LPS of the R strain *E. coli* K-12 often exhibits several sharp bands on Tricine-SDS gels, and these can differ depending on the strain background (78, 125). In genetic stocks in which the *his-rfb* region is derived from the K-12 strain Y-10 (4), a group which includes such widely used strains as C600, AB1157, AB1133, and MC4100, the LPS runs as three closely spaced bands which for convenience are designated below as bands 1, 2, and 3 in order of increasing gel mobility (78). Strains which are more closely related to wild-type *E. coli* K-12 produce only bands 1 and 2. As will be described in the next section, the reason for these differences is a defect in the synthesis of TDP-rhamnose (TDP-Rha) in strain Y10 and its derivatives (78). Even though these bands are quite sharp, they are chemically heterogeneous. Their molecular composition is not entirely understood, but their gel migration clearly reflects a pattern of substitutions or alterations of the core due to the action of the *rfe* gene and the *rfa* and *rfb* gene clusters (74, 78, 124). Thus, the ability to resolve changes in these bands which accompany the intro-

duction of defined mutations has provided a very useful tool for studying the roles of *rfaK* and *rfaL* and, as described in the next section, for studying the functions of *rfaQ*, *rfaS*, and *rfaZ*.

The phenotype of *rfaK* mutants has been well studied chemically in *S. typhimurium* (93). These mutants lack the partial α 1,2-linked GlcNAc substituent on the terminal Glc II residue and are resistant to the phage Felix O-1 (FO). There is extensive evidence indicating that GlcNAc is an essential part of the receptor for this phage (88), although the fact that *rfaB* mutants can be selected by FO resistance is evidence that other core structures may also be necessary (93, 181). *S. typhimurium rfaK* mutants exhibit a defect in attachment of O antigen. The severity of this defect varies among different *rfaK* alleles (93), but in the tightest mutants no O antigen is detectable on LPS gels indicating that the requirement is quite stringent (74). In *S. typhimurium* there is almost no difference in gel migration between LPS from *rfaK* mutants and LPS from *rfaL* mutants (74).

The RfaK proteins of *E. coli* K-12 and *S. typhimurium* are similar in size and overall structure but share almost no significant regions of sequence homology (75). In fact, the *S. typhimurium* RfaK protein shares more homology with the RfaG protein of *E. coli* K-12 than with the RfaK protein of *E. coli* K-12 (75, 119). The gel phenotype of an *rfaK* mutant is quite different in *E. coli* K-12 than in *S. typhimurium* (74). The *E. coli* K-12 *rfaK* mutant LPS lacks bands 1 and 2 which are seen in Rfa⁺ LPS and has only a single band which comigrates with band 3. This suggests that RfaK acts at an earlier stage of core completion in *E. coli* K-12 than in *S. typhimurium*. The difference between RfaK of these two organisms is further illustrated by their effect on phage sensitivity. *E. coli* K-12 is resistant to phage FO but becomes sensitive when it is transformed with a plasmid carrying the *S. typhimurium rfaK* gene (74). Rfa⁺ *E. coli* K-12 is also resistant to phage Br2, a phage which requires an inner-core structure for its receptor (136), but *rfaK* mutants of *E. coli* K-12 become sensitive to Br2. Conversely, an *rfaK* mutant of *S. typhimurium* is sensitive to Br2 but becomes resistant to Br2 when it is transformed with a plasmid carrying the *E. coli* K-12 *rfaK* gene. These results are interpreted as indicating that RfaK is a transferase which adds a sugar substituent to the core in both organisms but adds this substituent to a different site in *E. coli* K-12 from that in *S. typhimurium* (74). The ability to modify the phage sensitivity of each by introducing a plasmid carrying the *rfaK* gene of the other indicates that each organism produces a core capable of being modified by both RfaK enzymes.

The complementation experiments below provide the only evidence suggesting that the *E. coli* K-12 RfaK protein is also a GlcNAc transferase, since they indicate that the *S. typhimurium* and *E. coli* K-12 *rfaK* genes are functional homologs. The site(s) at which GlcNAc is attached to the K-12 core is not clearly known (74). One study reported that GlcNAc was attached in a β 1,6-linkage to Glc III (64), and another reported GlcNAc to be in an α 1,7 linkage to a Hep residue and detected only α linkages by nuclear magnetic resonance analysis of core oligosaccharides (60). In *E. coli* K-12, attachment of GlcNAc to an inner core Hep would be consistent with the phage Br2 sensitivity data (74).

The recent observation that band 1 seen in SDS-Tricine gels of *E. coli* K-12 LPS is absent in an *rfe* mutant (77) may eventually resolve the conflict about the site of attachment of GlcNAc. The *rfe* gene product transfers GlcNAc-phosphate to the dolichol analog ACL, a reaction which appears to initiate the synthesis of ECA and *rfe*-dependent O antigens

(101, 102, 141). It has been suggested that band 1 results from the transfer of this GlcNAc from ACL to the core, a process which requires *rfaL* and probably *rfaX* (77), and this is supported by recent chemical studies of the K-12 core (41a). Thus it is likely that GlcNAc is attached to the K-12 core at two different sites, one attachment requiring only *rfaK* and the other a more complicated transfer which requires *rfaK*, *rfaL*, *rfe*, and *rfaX*.

In *Salmonella* spp. and in *E. coli* K-12, GlcNAc is considered to be a partial substituent (93, 142). In *E. coli* K-12, however, the RfaK protein appears to play a very significant role in core completion since the LPS species which accumulates in an *rfaK* mutant is smaller than the bulk of the LPS molecules with completed cores. Since the GlcNAc content of the completed core is sometimes quite low, this raises the question of whether GlcNAc is added only to a subset of cores or, alternatively, is added to all of the cores and then removed from some of them during subsequent steps of core completion and preparation for O-antigen attachment. The initial studies with *Salmonella* spp., which were done to establish the linkage of O antigen to the core, involved small oligosaccharides purified after partial hydrolysis and do not address the question of the complete structure of the core to which the O antigen is attached (110). A study of phage FO inactivation by LPS purified from various *Salmonella* strains indicated that LPS from S strains was more than an order of magnitude less effective at neutralization of the phage than was LPS from an R strain, and this was paralleled by the phage sensitivity of the bacteria from which the LPS was isolated (88). This difference was interpreted as being due to steric hindrance by the long O polysaccharide chains present in LPS from the S strains, but, alternatively, this difference could be due to the absence of GlcNAc from the LPS molecules which carry O antigen. Such a trimming coupled to maturation would be consistent with the oligosaccharide biosynthetic pathways found in eukaryotic organisms (43) and could explain the stringent requirement of O antigen attachment for a "partial" core substituent.

In *S. typhimurium*, *rfaL* mutants make no S LPS but produce complete O-antigen side chains attached to ACL and R LPS which has the chemical and serological properties expected of complete cores. On this basis, *rfaL* has been proposed to encode a component of the O-antigen ligase which is involved in the transfer of O antigen from ACL to an LPS acceptor (93). Recent data support this conclusion, although the role of *rfaL* appears somewhat broader than originally envisioned. In *E. coli* K-12 the LPS from *rfaL* mutants lacks band 1 and *rfaL* strains carrying the plasmid-encoded *rfp* gene from *S. dysenteriae* 1 fail to produce the Rfp-modified top band which replaces band 1 (74). Band 1 and the Rfp-modified band are now thought to represent transfer of one- and two-sugar partial O units from ACL to LPS (77). In some strains, type I capsular polysaccharides are attached to the LPS core, and this also requires *rfaL* (91). A better description of the function of *rfaL* is that it is required for the transfer of a variety of polysaccharides from ACL to the LPS core, and on this basis it can be predicted that *rfaL* might also be necessary for synthesizing the LPS-linked form of ECA, which has been found in some rough strains (102). The RfaL protein has a structure predicted to be an integral membrane protein with 10 or more transmembrane segments (75, 92), and it has been proposed that it may serve as an organizer in the membrane which binds LPS and interacts with other proteins involved in late steps of core completion and O-antigen attachment (74).

Such a function does not preclude enzymatic activity as well, possible activities being the trimming of GlcNAc or the substituent added by RfaQ protein.

Although the RfaL proteins of *S. typhimurium* and *E. coli* K-12 have a nearly identical structure predicted on the basis of hydrophathy, these proteins resemble the RfaK proteins in that they share almost no regions of significant sequence homology (75). Plasmids carrying either *rfaL* alone or *rfaK* plus *rfaL* from *E. coli* K-12 did not complement an *S. typhimurium* *rfaL* mutant, even though the *E. coli* K-12 *rfaC* gene which was present on the same plasmids efficiently complemented a *S. typhimurium* *rfaC* mutation (74). More complex cross-complementation was observed with *rfaK*. A plasmid carrying *E. coli* K-12 *rfaK* was capable of partially complementing the production of an O-antigen ladder by an *rfaK* mutant of *S. typhimurium*, but only when the plasmid also contained *rfaL* from *E. coli* K-12 (74). Taken together, these results suggest that RfaL has two distinct strain-specific functions: one requires an interaction with a specific modification of the core by RfaK, and for this *rfaK* and *rfaL* must be compatible. The other function requires a specific interaction between RfaL and another component which is probably also a protein. Interestingly, when an *rfaK* mutant of *S. typhimurium* is complemented by a plasmid carrying both *rfaK* and *rfaL* from *E. coli* K-12, all of the bands in the resulting O-antigen ladder migrate more rapidly than the bands produced by an Rfa⁺ strain of *S. typhimurium* (74). The basis for this phenomenon is unknown.

***rfaQ* and core heterogeneity.** The genes *rfaQ*, *rfaS*, and *rfaZ* were not detected in early mutant studies of *Salmonella* spp. and were found only when the *rfa* clusters of *S. typhimurium* LT2 and *E. coli* K-12 were sequenced (92, 119, 124, 154). The ability to manipulate core heterogeneity by introducing plasmids carrying genes for the synthesis of *S. dysenteriae* 1 O antigen has been very useful in studying the function of these *rfa* genes in *E. coli* K-12 (74).

rfaQ is the first gene in the 10-gene *rfa* operon which determines the outer core. The 5' end of the gene has been sequenced in both *E. coli* K-12 and *S. typhimurium*, and the protein products are strongly conserved (more than 70% identity), indicating that the gene is similar in each organism (76). In *E. coli* K-12 the gene has been inactivated by nonpolar *Tn*lac insertions and by an in-frame deletion removing most of the coding sequence. These mutants are sensitive to phage C21, indicating that there has been a change in the core, but they are not sensitive to novobiocin or deoxycholate and produce normal amounts of the major outer membrane proteins. These mutations have a subtle but very interesting gel phenotype, which suggests that RfaQ is a transferase that adds a sugar or other substituent to the core (74). In a Y10-derived background such as strain AB1133, the LPS of an *rfaQ* mutant still exhibits three bands on gels, but band 3 exhibits a small but significant shift to a faster migration. In addition, band 1 is consistently less intense in an *rfaQ* mutant, as is the Rfp-modified band seen when an *rfp*⁺ plasmid is introduced into the strain. When a larger plasmid carrying the *S. dysenteriae* 1 *rfa* region is introduced into an *rfaQ* mutant, only a barely detectable amount of O antigen ladder is produced. These results suggest that the transfer of partial O units, O monomers, or polymeric O antigen from ACL to the core is inefficient in *rfaQ* mutants. However, unlike the shift observed for band 3, in an *rfaQ* mutant the migration of the small residual amounts of band 1, Rfp-modified band, and O-ladder bands is the same as in the *rfa*⁺ parent. The most convenient explanation for these observations is that RfaQ adds a

substituent which is required to initiate the efficient transfer of material from ACL onto the core but that the RfaQ substituent is normally removed during that transfer. If correct, this analysis indicates that band 3 represents an early stage in the synthesis of a unique molecular species which serves as a substrate for O-antigen attachment (74). All of the effects of *rfaQ* mutations in *E. coli* K-12 are complemented in *trans* by plasmids with *rfaQ* genes cloned from either *E. coli* K-12 or *S. typhimurium*, indicating that the function of the gene is the same in both organisms and that the effects of the mutations were not due to polarity (74). The nature of the substituent added by RfaQ is not known, although the branch Hep III or the terminal Hep IV would be a likely candidates given the homology of RfaQ to the heptosyltransferases (26, 119).

rfaQ mutants also helped to define a unique R LPS species, which has been termed a lipooligosaccharide (LOS) on the basis of its similarity to the LOS produced by members of the genera *Haemophilus* and *Neisseria*, which produce an LPS that lacks repeating polysaccharides such as O antigens and consists only of a complex oligosaccharide attached to lipid A. It was observed that when plasmids carrying the genes for the production of the *S. dysenteriae* 1 O antigen were introduced into derivatives of *E. coli* K-12 strain AB1133, there were changes in the R LPS bands which accompanied the production of an O-antigen ladder. Band 3 disappeared, and there was an increased amount of a band which had the same migration as band 2. This change occurred even in an *rfaL* mutant in which the synthesis of O antigen and band 1 were blocked (74). Recently, it has been shown that this shift from band 3 to band 2 was due to the complementation of a defect in TDP-Rha synthesis by the *S. dysenteriae* 1 *rfaC* gene supplied by the plasmid (78). The other interesting observation was that when a *rfaQ* mutation was introduced into the strain carrying the O-antigen plasmid, band 2 split into two components. One was a minor band whose migration was unchanged, and the other was a major band which shifted to a faster migration like band 3, indicating that in an *rfa*⁺ strain it, too, contained an *rfaQ*-dependent substituent. It was this shift in migration in an *rfaQ* mutant which first defined the LOS band as a unique form of R LPS and suggested that band 3 represented an earlier stage in its synthesis (74).

Additional evidence that the LOS band is the result of a unique biosynthetic pathway was provided by the properties of *rfaS* and *rfaZ* mutants (74). In *E. coli* K-12 strains derived from Y10 which are unable to synthesize TDP-Rha, *rfaS* and *rfaZ* mutations have little or no phenotype. However, when *rfaS* and *rfaZ* mutations were introduced into strains with plasmids carrying the *S. dysenteriae* *rfa* genes that complemented the TDP-Rha defect, they exhibited a very substantial phenotype (74). The LOS band disappeared, and there was an accumulation of material in one or more bands which migrated faster than LOS. However, the intensity and migration of the O-antigen ladder were unaffected by the *rfaS* and *rfaZ* mutations, indicating that although these genes are required to produce the LOS band, they are not necessary for the synthesis of the LPS cores to which O antigen is attached.

E. coli K-12 strains such as W1485, W3110, and 679, which are closely related to the K-12 wild type (4), have a partial *rfa* cluster which contains all of the genes necessary for synthesis of TDP-Rha (77, 78, 112). These strains do not require *rfa* genes from other organisms to produce LOS, and the shift in their LPS migration following the introduction of an *rfaQ* mutation indicated that about half of band 2, the

most abundant LPS band in these strains, was LOS. When these strains were transduced to *rfaS* or a deletion removing the *rfa* region was introduced, they lost the LOS band and acquired band 3, and their LPS gel phenotype became essentially identical to that of Y10-derived strains such as AB1133 (78). These results suggest that the addition of Rha to the core is an essential step in the conversion of band 3 to LOS and that *rfaS* encodes a function which is necessary for the addition of Rha to the core. Since *rfaS* mutants have a slight phenotype in Rha⁻ strains (124, 125), it is unlikely that *rfaS* is the structural gene for a Rha transferase. Recent structural studies on *E. coli* K-12 LPS have shown that Rha is linked α 1,5 to one of the inner-core KDO residues, probably KDO II (59, 60).

It is interesting to speculate that LOS may contain a terminal Hep IV residue attached to Glc III in addition to Rha in the inner core, since Hep IV was detected in strain W3100, which is Rha⁺ and is closely related to wild-type K-12 (62). In contrast, β -linked GlcNAc was found attached to Glc III of strain OT99, a Rha⁻ strain which is derived from Y10 (41a). The definition of LOS in *E. coli* K-12 is that it is a form of R LPS which is produced as an alternative to S LPS and is not a substrate for the attachment of O antigen (74). The molecules which contain β -linked GlcNAc attached to Glc III can be considered to be an intermediate stage in the synthesis of S LPS; as will be described in a later section, this GlcNAc becomes the first sugar at the reducing end of the O antigen produced by *E. coli* K-12-*S. dysenteriae* 1 hybrid strains (41a, 77), and it is likely to be the first sugar at the reducing end of other *rfa*-dependent O antigens which are produced by *E. coli*. Since the Hep IV and β -linked GlcNAc substituents are both attached to the 6 position of Glc III (41a, 62), they must be mutually exclusive. Thus, the presence of Hep IV would be consistent with the definition of LOS stated above. However, we must emphasize that for the moment this is only speculation, and it would be a formidable task to confirm this speculation by proving chemically that Rha and Hep IV reside on the same LPS molecule. There are two reasons for this. First, as noted above, the LPS of Rha⁺ strains is heterogeneous, and probably no more than 50% of the LPS is of the LOS form (78). Thus, bulk analysis of the LPS cannot provide proof. Second, the bond between the branch KDO on which the Rha substituent is located and the remainder of the core which would contain the Hep IV is cleaved by the mildest conditions of hydrolysis which can be used to remove lipid A (60). Until an alternative technology such as enzymatic cleavage of LPS or sophisticated mass spectrometry capable of analyzing intact LPS molecules becomes available, there is no effective way to characterize LPS molecules which might contain both substituents.

S. typhimurium LT2 does not appear to produce an LOS band. This may be due in part to the fact that *S. typhimurium* LT2 lacks an *rfaS* gene, although the *rfaP* and *rfaB* genes which flank *rfaS* in *E. coli* K-12 are strongly conserved between *E. coli* K-12 and *S. typhimurium* LT2 (76). The region between *rfaP* and *rfaB* of *S. typhimurium* LT2 is what has been termed a gray hole (36), a region which does not encode a functional protein but appears to contain remnants of an ancestral gene. Like the *rfaS* gene of *E. coli* K-12, this gray-hole region in *S. typhimurium* LT2 between *rfaP* and *rfaB* has a remarkably high A+T content (76). The gel bands at the bottom of an *S. typhimurium* LT2 O ladder are like those seen in an *rfaZ* mutant of *E. coli* K-12, and this suggests that *rfaZ* may also be nonfunctional in this organism.

Many questions remain about the LOS band besides its chemical structure. It is not known how widely a similar form of LPS is distributed among other enteric strains. A similar band has been seen in LPS of S isolates of *E. coli* (76) and in poultry strains of *S. enteritidis* (121). By analogy to other genes in the *rfa* and *rfb* clusters, the high A+T content of *rfaS* indicates that the sequence of the RfaS protein may not be conserved across strain or species lines. As this region of the *rfa* cluster is sequenced in other enteric organisms, it will be interesting to learn whether *rfaS* is conserved and whether the ability to produce a LOS has any effect on the ability of the organisms to survive or to colonize various host environments. It will also be interesting to learn whether organisms which produce a LOS have the ability to shift back and forth between the production of S LPS and LOS.

Regulation of Core Synthesis

Arrangement of transcriptional units. The first study of *rfa* transcriptional units was done by examining the complementation of mini-Tn10 insertions into *rfaQ*, *rfaG*, and *rfaP* by cloned fragments of various lengths. These fragments contained the promoter region between *rfaQ* and *kdtA*, the *rfaQ* gene, and one to six additional downstream genes. Regardless of the site of the mutation, the size and complexity of the LPS increased with the size of the complementing DNA fragments. This provided formal proof that the genes were arranged in an operon transcribed from a promoter downstream of *rfaQ* (3). There was evidence of expression of genes downstream of the Tn10 insertions, which was interpreted as indicating the presence of secondary promoters within the operon, but subsequent studies suggest that these effects may have been due to incomplete polarity of the insertions or promoter activity provided by Tn10. A more extensive study has been done by using polar Ω cassettes to interrupt transcriptional units and Tnlac insertions to monitor gene expression (147). This study identified three operons: a pair of divergently transcribed operons at the right (*kdtA*) end and a single operon at the left (*rfaD*) end. The divergently transcribed operons at the right end were a short operon consisting of *kdtA* and the 18K gene and a long operon extending from *rfaQ* at least through *rfaZ*. The operon at the left end includes *rfaDFCL*.

There was no evidence for secondary promoters in the nine-gene span from *rfaQ* through *rfaZ*. However, the last reporter fusion in that series was located near the junction of *rfaK* and *rfaZ*, and it is not clear that it monitored *rfaK* expression. There is evidence in another study indicating a weak secondary promoter for *rfaK*: when a deletion of genes *rfaGPSBI* marked with a polar Ω cassette is complemented with a longer fragment containing the downstream promoter region and the genes from *rfaQ* through *rfaJ*, the largest LPS species seen on a gel were larger than those produced by an *rfaK* mutant (124). It has not been possible to test for this promoter directly since no reporter fusions which are known to be located within *rfaK* have been isolated.

Recent primer extension studies have provided direct evidence that divergent transcripts originate in the untranslated region between *rfaQ* and *kdtA* (26). Two leftward and two rightward transcript start sites were identified. These sites do not overlap but are close enough to share common regulatory elements. As anticipated for highly regulated promoters, the regions upstream of these start sites do not contain strong promoter consensus sequences. Within the untranslated region there are four blocks ranging from 15 to

30 bp in which the sequence is nearly identical between *E. coli* K-12 and *S. typhimurium* LT2 (76), suggesting that this region contains important regulatory information.

Less is known about the untranslated region between the 26K gene and *rfaD*. However, there is one very remarkable aspect to this region. The *rfaD* gene exhibits heat shock regulation, and one laboratory cloned this gene while screening for novel heat shock-regulated genes (134). This is the first, and to our knowledge the only, case in which a gene in a carbohydrate biosynthetic pathway exhibits heat shock regulation. Of the many genes involved in the synthesis of core sugar nucleotides, *rfaD* is the only one included in the *rfa* cluster. Together with the unique heat shock regulation, this suggests that *rfaD* may play an important role in the regulation of core synthesis.

Role of *rfaH* (*sfrB*). The *rfaH* (*sfrB*) gene lies at 87 min, near *ubiB* and *fadA*, and encodes an 18-kDa basic protein which shows no similarity to other regulatory proteins that have been studied. Despite several attempts, the protein has not been purified or visualized on gels, suggesting that it may be unstable. Insertion and amber mutations in *rfaH* have been reported (6, 32, 174), indicating that it is not essential, but mutants with a strong phenotype are sickly and readily accumulate suppressors. The phenotype is not consistent among various alleles, with the most severely affected alleles exhibiting both poor growth and a partial deep rough phenotype.

The first *rfaH* mutants were observed in *S. typhimurium* as a subclass of phage FO-resistant mutants (93). The inability of *rfaH* mutants to transfer labeled Gal from UDP-Gal to Gal-deficient LPS cores indicated that *rfaH* might encode a Gal transferase, which led to some confusion in the *S. typhimurium* LPS literature. In *E. coli* K-12, mutants designated *sfrB* were selected on the basis of lack of expression of the *tra* genes of plasmid F. These mutants were sensitive to phage C21 and were sometimes resistant to phage U3, which indicated an LPS defect (93). The confusion was resolved when it became apparent that *rfaH* and *sfrB* were allelic (93) and that RfaH protein was a positive regulator which regulated the expression of several enzymes encoded by the *rfa* cluster as well as the *tra* operon of F (19, 33, 125, 140).

In *E. coli*, SfrB was found to prevent the premature termination of *tra* operon transcripts at an internal terminator site, which led to the suggestion that it acts as a transcriptional antiterminator analogous to the N or Q proteins of λ (13). Despite considerable effort in several laboratories, this hypothesis remains unproven; however, there is now evidence from studies of the long *rfa* operon which provide support for an antitermination mechanism. Suppressors of *rfaH* map to *rho*, suggesting that RfaH acts at Rho-dependent termination sites (42). This is consistent with the observations that Ω cassettes which have Rho-independent terminators are completely polar when inserted into the *rfa* long operon, whereas Tnlac insertions, which should sometimes exhibit Rho-dependent polarity as a result of polypeptide chain termination, appear to be entirely nonpolar (125, 147). Reporter fusions inserted throughout the *rfa* long operon are regulated by *rfaH* (19, 125), and there is evidence for multiple termination sites (19). The locations of Tnlac fusions in the long *rfa* operon of *E. coli* K-12 have been determined by sequence analysis, and RfaH-dependent fusions have been obtained in all of the genes of the operon except *rfaY* and *rfaK*. The extent to which these fusions were regulated increased with the distance from the promoter, with a fusion in *rfaQ* showing only a 2- to 3-fold

increase in *rfa*⁺ over *rfa* whereas a fusion in *rfaZ* showed a 10- to 15-fold increase (123).

Experiments have been done to determine the similarity of the *rfaH* system to the antitermination systems of λ by examining the effect of various constructions on the *rfaH* dependence of a *Tnlac* insertion in *rfaB* (125). A pair of overlapping deletions extending from a *HindIII* site near the middle of *rfaG* to a *ClaI* site very near the 5' end of the *rfaQ* coding region had no effect on *rfaH* dependence (123). However, replacement of the region upstream from this *ClaI* site with a *lac* promoter resulted in expression which was dependent on a *lac* inducer but exhibited no *rfaH* dependence over a wide range of inducer concentrations (123). This suggests a model in which the action of RfaH is similar to that of the Q protein of λ , which modifies RNA polymerase while it transcribes a site located very near the promoter. Unfortunately, such a site for RfaH action has not yet been identified.

All of the evidence cited above is consistent with a model in which RfaH protein acts to prevent the termination of transcription at multiple Rho-dependent sites throughout the long *rfa* operon. Unlike RfaH protein, there is evidence that the enzymes of LPS synthesis may be quite stable (149). Since LPS is also quite stable and there is no way to store excess LPS or its precursors, there is a need to prevent expression of the genes for these enzymes at inappropriate times in the growth cycle. One can envision RfaH as acting as a kind of deadman's throttle, allowing synthesis of the LPS enzymes only when it is being actively synthesized itself. It will be interesting to know more about the factors which regulate the synthesis of RfaH protein.

Whether RfaH protein acts directly on operons other than the *tra* operon of F and the long *rfa* operon remains an open question. It has been reported that RfaH decreases termination in the *rpIL-rpoB* intercistronic region (135) and, more recently, that RfaH was required for expression of a plasmid-encoded hemolysin (6). There is now evidence that the effect on hemolysin expression may be indirect and mediated by the effect of *rfaH* mutations on LPS core genes. The outer membrane form of TolC protein is required for hemolysin secretion, and *rfaH* or *galU* mutations which result in a truncated LPS prevent the export of the TolC protein to the outer membrane (174). This is somewhat similar to the effect of *galU* or *rfaG* mutations on the expression of flagella noted earlier. Considering the multiplicity of signaling mechanisms in bacteria, it would not be unreasonable for cells to monitor the synthesis of such an important component as LPS and to use this parameter to adjust the rate of synthesis of other cell surface components. Cells could accomplish this by sensing the completed LPS molecules in the cytoplasmic membrane, by sensing the rate of export or assembly of an outer membrane protein such as a porin which requires LPS as a ligand to promote correct folding, or by sensing the levels of key enzymes such as RfaG or RfaB together with their sugar nucleotide substrates. In proposing new targets for the action of RfaH, it is necessary to show that other mutations such as *galU* which affect LPS structure do not mimic its effect.

Effect of temperature. In enteric bacteria, temperature is a parameter which has a strong effect on the size distribution and abundance of various molecular species of LPS. This is most evident in *E. coli* K-12 strains carrying plasmid-encoded O-antigen genes. On the basis of staining intensity on gels, the median size of the long O-antigen polymers decreases by about 2 to 3 O units as temperature is increased from 30 to 42°C, whereas the number of total molecules

which carry O units also tends to decrease (73). In K-12 strains without O-antigen plasmids the distribution of R LPS species also appears to be affected by temperature (125). In strains derived from strain Y10, which produce three major R bands, bands 1 and 3 decrease and band 2 increases as the temperature is increased. Some of this may be due to differential regulation of genes within the *rfa* long operon, perhaps at the level of translation, since the expression of in-frame *Tnlac* fusions to *rfaJ* and *rfaP* were found to be affected by temperature (125).

GENETICS OF O-ANTIGEN SYNTHESIS

A summary of the genes involved in O-antigen synthesis and attachment is given in Table 3. Enteric bacteria of genera such as *Salmonella*, *Escherichia*, *Shigella*, and *Klebsiella* can be divided into classes based on the gene which is involved in attachment of the first sugar of the O unit to ACL. At the present time, only two such groups are known, the *rfbP*-dependent class and the *rfe*-dependent class (63, 77, 93, 141), although there will almost certainly be additional classes as more organisms are studied in detail. The *rfbP*-dependent class includes O groups A, B, C2, D, and E1 of *Salmonella* species. The prototype for the *rfbP*-dependent class is the group B organism *S. typhimurium*. This is the organism with which most of the initial biochemical and genetic studies of O-antigen synthesis were done (93, 142), as well as the first organism in which the *rfb* gene cluster and the *rfe* gene were sequenced (30, 65).

The O-specific side chains of these *rfbP*-dependent strains consist of a repeating unit of three, four, five, or six sugars which are often branched, and there are 25 to 40 of these repeating units in the longest O-specific polymers. In the members of this class, synthesis of a monomeric O unit is initiated by the *rfbP*-dependent transfer of Gal-phosphate from UDP-Gal to ACL-phosphate to form Gal-pyrophosphate-ACL. The remainder of the O unit is then built up by the sequential transfer of single sugars from their appropriate sugar nucleotides. Monomers are then combined into ACL-linked polymers by repeated transfer of a growing chain of O units onto newly synthesized monomeric O units attached to ACL, with the result that in a growing chain the O unit attached to ACL at the reducing end of the O polymer is the most recently synthesized (63, 111, 142). The *rfbP*-dependent strains do not require the *rfe* gene for synthesis of O antigen (64, 93). However, all of these strains require *rfe* for the synthesis of ECA, a polymer with a trisaccharide repeating unit (90), which, like O antigen, is assembled on ACL (102).

The *rfe*-dependent class of strains includes *S. dysenteriae* 1, *Salmonella* groups C1 and L, *E. coli* serotypes O8 and O9, and *Klebsiella* serotype O1. The O antigens of these strains are much more heterogeneous in structure. In *S. dysenteriae* 1, the O unit contains four sugars and its synthesis appears to be initiated by the *rfe*-dependent transfer of GlcNAc-phosphate to ACL, which is followed by the sequential addition of the remaining sugars. The *rfb* cluster of *S. dysenteriae* 1 does not contain an *rfbP* gene, but in other respects it resembles the *rfb* clusters of the *rfbP*-dependent *Salmonella* serovars (77). The *Salmonella* group C1 O antigen has a backbone containing GlcNAc and four mannose residues, with the penultimate mannose substituted by glucose. The *rfb* cluster of *Salmonella* group C1 resembles that of *S. dysenteriae* 1 in that it does not contain an *rfbP* gene or analog (82, 83).

The O antigens of the *rfe*-dependent *E. coli* and *Klebsiella*

TABLE 3. Genes or blocks of genes involved in the synthesis of O units and their assembly into O antigens attached to the LPS core

| Gene(s) | Probable function of gene products | Example(s) and reference(s) |
|------------------------------|--|---|
| <i>rfbBCAD</i> | Synthesis of TDP-rhamnose from glucose-1-phosphate | <i>Salmonella</i> (65) and <i>Shigella</i> (77) spp. |
| <i>rfbKM</i> | Synthesis of GDP-mannose from mannose-6-phosphate | <i>Salmonella</i> spp. (65) |
| <i>rfbFG[HI]^a</i> | Synthesis of CDP-4-keto-3,6-dideoxyglucose (common precursor of dideoxy sugars) from CDP-glucose | <i>Salmonella</i> spp. (139) |
| <i>rfbJ</i> | Synthesis of CDP-abequose from common precursor in <i>Salmonella</i> group B | <i>Salmonella</i> spp. (139) |
| <i>rfbS</i> | Synthesis of CDP-paratose from common precursor in <i>Salmonella</i> groups A and D | <i>Salmonella</i> spp. (139) |
| <i>rfbE</i> | Synthesis of CDP-tyvelose from CDP-paratose in <i>Salmonella</i> group D | <i>Salmonella</i> spp. (139) |
| <i>rfbP</i> | Initiates synthesis of O unit by addition of galactose-phosphate to ACL to form galactose-pyrophosphate-ACL; may be bifunctional and encode RfbT function involved in transfer of O unit to LPS core | <i>Salmonella</i> spp. (65) |
| <i>rfe</i> | Initiates synthesis of O unit by addition of GlcNAc-phosphate to ACL to form GlcNAc-pyrophosphate-ACL | <i>Shigella</i> spp. (77) |
| <i>rfbN</i> | Synthesis of O unit; rhamnosyl transferase in groups A, B, D, and E1 of <i>Salmonella</i> spp. | <i>Salmonella</i> spp. (139) |
| <i>rfbQ</i> | Synthesis of O unit; rhamnosyl transferase I in <i>S. dysenteriae</i> 1 | <i>Shigella</i> spp. (77) |
| <i>rfbR</i> | Synthesis of O unit; rhamnosyl transferase II in <i>S. dysenteriae</i> 1 | <i>Shigella</i> spp. (77) |
| <i>rfbU</i> | Synthesis of O unit; mannosyl transferase in groups A, B, and D of <i>Salmonella</i> spp. | <i>Salmonella</i> spp. (139) |
| <i>rfbV</i> | Synthesis of O unit; abequosyl transferase in groups B and D of <i>Salmonella</i> spp. | <i>Salmonella</i> spp. (139) |
| <i>rfp</i> | Synthesis of O unit; plasmid-encoded galactosyltransferase in <i>S. dysenteriae</i> 1 | <i>Shigella</i> spp. (74) |
| <i>rfbX</i> | Assembly or transfer of O units; exact function unclear | <i>Salmonella</i> (139) and <i>Shigella</i> (77) spp. |
| <i>rfc</i> | Polymerization of O units on ACL by repeated transfer of growing chain to single O unit on ACL; may play role in ligation of polymeric O antigen to core | <i>Salmonella</i> (21) and <i>Shigella</i> (77) spp. |
| <i>rfaL</i> | Required for ligation of polymeric O antigen, single O units, or partial O units to core | <i>E. coli</i> (74) and <i>Salmonella</i> (92) spp. |
| <i>rol (cld)</i> | Regulation of chain length of polymeric O antigen | <i>E. coli</i> (9) and <i>Salmonella</i> (8) spp. |

^a Brackets indicate genes which have not yet been located.

serotypes listed above have a simpler structure. These consist of neutral homopolymers of Man or Gal, in some cases with modifications such as alternating anomeric linkages or simple branch substituents. The exact mode of synthesis of the homopolymeric O antigens is unclear, and some studies have suggested a continuous polymerization on the nonreducing end rather than the discontinuous buildup from smaller repeating units seen with the *rfe*-independent polymers (63). A single sugar residue which is different from the sugar residues making up the bulk of the polymer is found linking the reducing end of the polymer to ACL. There is some controversy over the identity of this initiating sugar. It was originally described as glucose (63), but studies on *rfe* showing that it is a GlcNAc-phosphate transferase (101) and recent analytical studies on an *rfe*-dependent *E. coli* O antigen (141) indicate that the initiating sugar will invariably be GlcNAc. In support of continuous polymerization, no mutants have been found in the *rfe*-dependent systems which attach a single O unit to LPS comparable to the "semirough" LPS which is produced by *rfc* mutants in the *rfe*-independent systems. However, an argument against continuous polymerization of the O antigen is that *rfe*-dependent LPS also migrates as a discontinuous ladder on gels.

Despite the very substantial differences in the structure and mechanism of biosynthesis, the synthesis of all of these types of O antigens requires an *rfb* gene cluster located in the 40-min region between *his* and the capsular polysaccharide (*cps*) genes. This section is concerned primarily with the structure of these *rfb* gene clusters and with the *rfc* and *rfe* genes which work with them.

O-Antigen Genes of *rfbP*-Dependent *Salmonella* Serovars

Organization of *rfb* in *S. typhimurium*. As shown in Fig. 2, the *rfb* cluster of *S. typhimurium* consists of 16 contiguous genes which are transcribed in the counterclockwise (*cps*-to-*his*) direction (65, 139). Although there are accessory genes in regions flanking both ends of this cluster which may have a role in O-antigen synthesis, a fragment extending from the noncoding region upstream of *rfbB* through the *rfbP* gene is sufficient to confer O-antigen production on a Δ *his-cps* strain (65). The cluster is flanked at one end by *cps* genes and *galF* (Orf2.8 in Fig. 2) and at the other by *gnd*, *rol*, and the *his* operon, so that the order is *cps-galF-rfb-gnd-rol-his* (8, 18, 65).

The O unit of group B strains consists of a four- or five-sugar repeat (63). The backbone consists of Man, Rha, and Gal (Fig. 3) and is identical to that of groups A, D, and E1; in addition, the Man is substituted by the dideoxyhexose abequose (Abe) and the Gal is partially substituted by Glc. The Abe residue is also acetylated. The *rfb* cluster contains genes encoding the biosynthesis of GDP-Man, TDP-Rha and CDP-Abe. With the exception of TDP-Rha, these nucleotide sugars are not involved in synthesis of the LPS core or other essential cellular structures. Although carbohydrate gene clusters do not invariably include genes for the synthesis of the unique sugar nucleotides which they require, inclusion of such genes is a common strategy in O-antigen and capsular polysaccharide clusters. The cluster also includes four sugar transferase genes and the *rfbX* gene which is found in all *rfb* clusters and encodes an integral membrane protein involved in O-antigen assembly. The cluster is saturated; with the

exception of *rfbX*, all of the genes can be assigned functions as sugar transferases or in sugar nucleotide biosynthesis.

Sugar biosynthetic genes of the *S. typhimurium* *rfb* cluster. The genes for the biosynthesis of the sugar nucleotides are organized into three blocks within the *rfb* cluster. There are four genes involved in the synthesis of TDP-Rha from Glc-1-phosphate (93), and these are the four genes *rfbBCAD* at the 5' end of the cluster (65, 139, 146). In earlier studies it was not possible to distinguish the last two enzymes in the TDP-Rha pathway, the epimerase encoded by *rfbC* and the dehydrogenase encoded by *rfbD*, and defects in either enzyme were termed *rfbD*. An assay for these two activities is now available (146), and the genes have been assigned by homology of their products to analogous enzymes encoded by the *strM* and *strL* genes of *Streptomyces griseus* (139). Immediately adjacent to the TDP-Rha biosynthetic genes are five genes involved in the synthesis of CDP-Abe from Glc-1-phosphate (65, 139). Three of these have been assigned as *rfbFGJ*, and the remaining genes, *rfbH* and *rfbI*, are thought to be the two unassigned open reading frames Orf7.6 and Orf10.4 present in this block. The two genes involved in the synthesis of GDP-Man from Man-6-phosphate are *rfbMK* near the other end of the cluster. All of these proteins have a structure consistent with soluble proteins (65).

The *galF* gene, which was mapped within or near the *rfb* cluster, was originally thought to be involved in the modification of UDP-Gal pyrophosphorylase, the product of the *galU* gene (65, 93, 108). The evidence for this was the presence of an altered form of the enzyme, which was eliminated by deletions which removed *rfb* genes. Recently, a comparison of GenBank protein sequences revealed that Orf2.8, encoded by a gene adjacent to *rfbB* (65), shared substantial regions of homology with GalU protein (138), suggesting that Orf2.8 is an alternate form of UDP-Gal pyrophosphorylase. Although it has not been tested directly, it seems very likely that the gene for Orf2.8 is *galF* and that its function is to provide a more efficient way of supplying the large amount of UDP-Gal required for synthesis of both a Gal-containing O antigen and a Gal-containing capsular polysaccharide. It is likely that it is the last gene in the *cps* cluster and not an *rfb* gene, since it is separated from the *rfb* cluster by a nontranscribed region which may contain a promoter for the *rfb* cluster. *E. coli* K-12 may also contain a *galF* gene, which could explain the somewhat leaky LPS phenotype of *galU* mutants (152).

Sugar transferases of the *S. typhimurium* *rfb* cluster. Three of the four *rfb* genes thought to encode sugar transferases have been assigned on the basis of the transfer of sugars from sugar nucleotides to ACL or ACL-linked oligosaccharide intermediates (18, 139). The location of *rfbP*, the gene for the transfer of galactose phosphate to ACL, at the *his* end of the *rfb* cluster was anticipated on the basis of the properties of deletions extending into *rfb* from *his* (111). *rfbP* encodes the first function in the assembly of the O unit which commits ACL to the LPS pathway, an arrangement which was fortuitous for the early genetic study of the *rfb* cluster (93, 111). Deletions extending into *rfb* from the *his* side were easily selected, and these did not have the deleterious phenotype which might otherwise have resulted from accumulation of an ACL-bound intermediate or from the sequestration of ACL. The protein encoded by *rfbP* is relatively hydrophobic, with at least five apparent transmembrane domains (65), and both its structure and its function are different from those the *rfb* transferases which add the distal sugars. The hydrophathy profile of the amino-terminal 250

residues of RfbP protein is very similar to that of the Rfe protein, which performs an analogous function by adding GlcNAc phosphate to ACL (78).

rfbN, *rfbU*, and *rfbV* determine the Rha, Man, and Abe transferases, respectively, and are arranged together in a block adjacent to *rfbX*. These transferase genes and *rfbP* are arranged in the gene cluster in the reverse order of the sequence in which the transferases act, an arrangement which appears to be followed in other *rfb* clusters which have been studied (139). The distal sugar transferases resemble the putative *rfa* hexosyltransferases in that their predicted structure generally does not contain significant hydrophobic segments which could function as membrane anchors or transmembrane domains. Thus it is likely that these proteins function as peripheral membrane proteins on the cytoplasmic face of the cytoplasmic membrane. None of these proteins show significant homology to proteins encoded by the *rfa* locus, which is not surprising since they recognize different sugar nucleotides and a different lipid acceptor. Two other functions involved in synthesis of group B O antigen, the O acetylation of Abe characteristic of group B strains and the glucosylation of the O antigen, are encoded by genes which lie outside of the *rfb* cluster, by the *oafA* gene which lies near but not within *rfb* and by the *oafR-oafE* cluster near *purE* at 11 min, respectively (93).

Polymerization and role of the *rfc* gene. The polymerization of O units is one of the first aspects of O-antigen synthesis to be investigated both biochemically and genetically. There are two reasons for this. First, the formation of a glycosidic bond between adjacent O units results in a new structure which can be recognized both as a new antigenic factor and by loss or acquisition of a phage receptor. It was recognized early on that temperate phages have the ability to manipulate the polymerization system, and it was shown that such phages encode both an inhibitor of the host O polymerase and a new polymerase which forms a different linkage. This new linkage can differ from that of the host in the carbons which are involved as well as in the anomeric configuration (93). These phage-encoded polymerase functions are expressed in the lysogenic state, and, like phage-encoded porins, they represent a mechanism for preventing the nonproductive neutralization of phages by nearby uninduced lysogens or debris from the lysing cell when a lysogen is induced.

The second reason has to do with the rather distinctive semirough (SR) phenotype of mutants in which polymerization is defective. Tight mutants of this type produce LPS in which a large proportion of the molecules carry a single O unit while the remainder have R cores which carry no O antigen (93). The proportion of LPS with R cores which lack O antigen is about the same in S and SR strains, indicating that ligation of a single O unit to the core and translocation of the resulting LPS to the cell surface can be very efficient.

In *S. typhimurium*, mutations resulting in SR LPS mapped to a locus outside of *rfb* between 18 and 34 min (93). This locus was presumed to encode the O polymerase and was termed *rfc*. More was learned about this phenomenon by examining the behavior of hybrids between various *Salmonella* serovars (93). When the *rfb* cluster of group B was transferred into group C1 or C2 strains, only SR LPS of the group B serotype was observed. In a reciprocal experiment, when the *rfb* cluster of group C1 was transferred into a group B strain, S LPS was produced. Subsequently, it was demonstrated genetically that the Rfc function of *Salmonella* group E1 mapped to the *rfb* locus (93).

These experiments suggested that in all three of these

serovars, C1, C2, and E1, the *rfc* gene might be located within the *rfb* cluster whereas in other groups such as group B it might be located elsewhere on the chromosome. This has now been partially confirmed by the cloning and sequencing of the *rfc* gene in group B and the identification of an *rfc* gene within the *rfb* cluster of a strain of group C2 (20, 21, 30). Both the group B and the group C2 Rfc proteins are very hydrophobic proteins with 10 or 11 potential transmembrane domains, indicating that they are integral membrane proteins. In the group C2 strain, the *rfc* gene is located in a nonconserved region near the middle of the *rfb* cluster adjacent to two genes which encode Man transferases (21). A gene encoding an open reading frame which may be an Rfc protein has also been found in the *rfb* cluster of group E1, although its location between the *rfbP* and *gnd* genes is different from in group C2 (175). However, the situation remains unclear for group C1, since there is no evidence for a gene encoding a similar protein in its *rfb* cluster (82, 83) and there have been no reports of mutants exhibiting an SR phenotype isolated from group C1 or from other strains producing simple or homopolymeric *rfe*-dependent O antigens similar to those of group C1. Thus the possibility remains that the O antigens of these organisms are polymerized by a different mechanism, which does not require an Rfc-like protein.

The polymorphism of *rfc* is not limited to *Salmonella* serovars. As shown in Fig. 2, an *rfc* gene is located adjacent to *rfbX* in the *rfb* cluster of *S. dysenteriae* 1, an organism which produces an *rfe*-dependent O antigen. The *S. dysenteriae* 1 *rfc* gene was identified on the basis of mutational studies as well as on the structural similarity of its product to the *Salmonella* Rfc proteins. Finding that the polymerase genes are only occasionally a part of the *rfb* cluster is not surprising, given observations with O hybrid strains that the polymerase has a relaxed specificity with respect to minor differences in the structure of the O units, such as the nature of the dideoxy sugar substituent, and the fact that several temperate phages carry polymerase genes which under some conditions can substitute for host genes (93).

Modification of O polysaccharide. The commonest forms of O-polysaccharide modification involve O acetylation or glucosylation of O-antigen sugar residues or a change (either positional or anomeric) in the linkage between the O units. These changes result in antigenic differences which were detected primarily by serological methods, although they may also result in resistance to O-specific phages. All three of these modifications can be the result of lysogeny by temperate phages (93), and the gene from the temperate phage Sf6 which results in the acetylation of Rha in the O antigen of *S. flexneri* has recently been studied in detail (25, 155, 170). In group B *Salmonella* strains, glycosylation involves the chromosomal locus *oafA* at 43 min and acetylation involves the *oafC* and *oafR* genes located in the 12- to 13-min region. There has been no recent work on the genetics of these phenomena in *Salmonella* strains, and readers are directed to an earlier review (93) for more detailed information.

The *rfb* cluster of *Salmonella* group C2 contains a gene termed *orf14.9* (21), which encodes a protein whose deduced amino acid sequence shows strong similarity to the products of the *lacA* and *nodL* genes, both of which are known to be acetyltransferases (138). As yet, no function has been ascribed to the Orf14.9 protein and no comparable gene has been found in the other *Salmonella* *rfb* clusters which have been sequenced. A gene located in or near the partial *rfb* region of *E. coli* K-12 results in the acetylation of the RhaIII

residues of O antigens which are produced when the cloned *rfb* clusters of *S. flexneri* 2a and 3a are expressed in *E. coli* K-12 (182). The true function of this gene is not clear, since the *E. coli* K-12 *rfb* cluster is not sufficient for the production of its own O antigen.

Polymorphism among *rfbP*-dependent *Salmonella* serovars. The *rfb* clusters of five *rfbP*-dependent *Salmonella* serogroups have been sequenced. Groups B, A, D, and E1 all share a basic Man-Rha-Gal backbone structure, whereas group C2 has a Rha-Man-Man-Gal backbone. Additionally, the groups differ in the nature of the dideoxyhexose substituent: none in E1, Abe in B and C2, paratose in A, and tyvulose in D. The biosynthesis of the dideoxyhexoses requires the products of four genes (*rfbFGHI*) to convert Glu-1-phosphate to the common precursor CDP-4-keto-3,6-dideoxy-D-Glc (I) and three additional genes to synthesize the final products. The *rfbJ* gene product converts I to CDP-Abe, the product of *rfbS* converts I to CDP-paratose, and *rfbE* product converts CDP-paratose to CDP-tyvulose (93, 139).

Group E1 has the smallest *rfb* cluster of the five serogroups, since it lacks all of the dideoxyhexose genes and has only three sugar transferase genes. It contains two additional genes of unknown function, one of which may be an *rfc* analog (139, 175). The other groups share a conserved region adjacent to *rfbBCAD*, which includes the four shared dideoxyhexose genes *rfbFGHI* (Fig. 2) and either *rfbJ* (groups B and C2) or *rfbSE* (groups A and D). The presence of *rfbE* in group D was not anticipated since its O antigen contains paratose instead of tyvulose, but this was explained by the presence of an early frameshift mutation in this gene. Otherwise the *rfb* clusters of groups A and D are identical, suggesting that group D has acquired this mutation recently (89). The group C2 cluster is the largest, since it includes an additional Man transferase, the *rfc* gene and the putative acetylase gene noted in the previous section, and at least one additional gene of unknown function to bring the total number of genes to 19 (21). The structures of all of these clusters are quite close to what would be predicted on the basis of previous genetic and biochemical studies (93), and there are few additional genes.

As noted above, a polysaccharide gene cluster will often contain biosynthetic genes for sugar nucleotides which it requires. This can lead to duplication of function, an example being the duplication of genes for mannose biosynthesis between the *rfb* and *cps* gene clusters in some *Salmonella* serovars (93, 157). A consequence of this duplication is that one set of sugar biosynthetic genes might become lost, resulting in strains which have a similar function at different genetic loci. An example of this is seen in genes for the dTDP-Rha biosynthetic pathway. In *S. typhimurium* *rfbA* and *rfbB* encode the enzymes for the first two steps in the Rha biosynthetic pathway. These are also the first two steps in the synthesis of the fucosamine derivative 4-acetamido-4,6-dideoxy-D-Gal, which is part of the trisaccharide repeating unit of ECA. The O antigen of *Salmonella* group C1 LPS does not contain Rha, and the group C1 *rfb* cluster does not have a block of Rha biosynthetic genes corresponding to *rfbBCAD* of *S. typhimurium* (83). There is evidence that in group C1 a gene analogous to *rfbA* (and presumably an analog of *rfbB*) is located in the *rfe-rff* cluster instead of *rfb* (87). It is not known whether the group C1 genes are related to the conserved *rfbA* and *rfbB* genes found in other *Salmonella* serovars and in *S. dysenteriae* 1 (77, 139) or whether the *rfe-rff* cluster of group C1 contains analogs of *rfbCD*, the last two genes in the dTDP-Rha pathway.

O-Antigen Genes of *Shigella dysenteriae* 1

Genetic structure of the *rfb* cluster. As shown in Fig. 3, the O antigen of *S. dysenteriae* 1 has a repeating unit consisting of Rha-Rha-Gal-GlcNAc (40). Expression of this O antigen in *E. coli* K-12 requires the *E. coli* K-12 *rfe* gene (77) and the *S. dysenteriae* 1 *rfp* gene (74, 159). In *S. dysenteriae* 1 the *rfb* gene is carried on a 9-kb multicopy plasmid which is invariably found in virulent strains and is required for O-antigen production (51, 176, 177). In addition, expression of the O antigen in an *E. coli* K-12 strain deleted for the *his-cps* region requires a cloned fragment containing eight contiguous genes from the chromosomal *rfb* cluster of *S. dysenteriae* 1 (77, 160, 161). The 8.9-kb *rfb* fragment shown in Fig. 2 is one of the smallest *rfb* clusters found to date which is capable of directing O-antigen expression in an *rfb*-deleted host. The left end of the cluster contains the *rfbBCAD* genes, which are necessary for TDP-Rha synthesis. These show substantial homology to their counterparts in *Salmonella* strains (65). The next two genes are *rfbX* and *rfc*, and their protein products show structural similarity but only a low level of sequence homology to their *Salmonella* counterparts. The two genes designated as *rfbRQ* have been identified as encoding the Rha II and Rha I transferases, respectively (77, 160). The fragment encodes an additional open reading frame of unknown function, designated as Or1, and part of another open reading frame which appears from restriction site analysis to lie between *rfbRQ* and *gnd*. Neither Or1 nor the partial product of this additional open reading frame share significant structural homology with other known O-antigen gene products.

Roles of *rfe* and *rfp*. As noted above, the *rfb* cluster encodes only the sugar transferases which add the two distal sugars of the O unit. The two transferases for the proximal sugars are thought to be encoded by *rfe* and the plasmid gene *rfp* (77).

Introduction of *rfp* into *E. coli* K-12 resulted in an increase in the galactose content of the LPS core, and low-resolution SDS-gel analysis indicated that the migration of part of the LPS was reduced (153, 159). On this basis, it was suggested that *rfp* encoded a galactosyltransferase (159), and this has now been confirmed by the demonstration that overproduction of *rfp* leads to increased expression of Gal transferase activity (17). The Tricine-SDS gel system allowed the *rfp*-induced change in gel migration of the LPS to be examined more precisely. This showed that *rfp* reduced the migration of band 1 by an amount consistent with the addition of one sugar whereas the migration of bands 2 and 3 showed no change (74).

The size of this Rfp-modified band is two sugars too large to serve as a precursor for the addition of O units (74). This, together with the properties of an *rfe::Tn10* mutation, indicated that the Rfp-modified band is not a precursor but instead is an LPS molecule to which a two-sugar partial O unit has been transferred. The *rfe* mutation eliminated both band 1 and the Rfp-modified band, as well as eliminating the ability of *E. coli* K-12 to express the *S. dysenteriae* 1 O antigen. This indicated that the *rfp*-modified band involved the addition of both GlcNAc and Gal to the core (17, 77). Moreover, when derivatives of the *S. dysenteriae* 1 *rfb* cluster with mutations in the gene encoding rhamnosyl II transferase or in *rfc* are introduced together with *rfp*, they result in bands which are one and two sugars larger, respectively, than the *rfp*-modified band (77, 160). These intermediates are also seen in a leaky *rfb* mutant with an insertion in the 5' untranslated region, which presumably contains the

rfb promoter (65). Thus all four of the transferase activities can be accounted for. Earlier work envisioned the buildup of the first O unit as involving the transfer of sugars directly to the LPS core (160), but the involvement of *rfe* and the structural and functional similarity of Rfe to RfbP indicate that both partial and complete O units are first assembled on ACL and then transferred to the LPS core. Judging by gel band intensity, the efficiency of transfer of the partial O units appears to increase significantly as each sugar is added to the O unit (77). This indicates a preference for complete O units and may be one of the reasons why the transfer of partial O units was not detected in earlier studies.

This interpretation of the roles of *rfe* and *rfb* and of the transfer of partial O units to the core has been strengthened by a very recent study (41a) which examined the structure of core oligosaccharides produced by various *E. coli* K-12-*S. dysenteriae* 1 hybrid strains. The results on the K-12 background strain confirmed the earlier finding of GlcNAc linked to the core in a β configuration (64) and showed that in hybrids containing only the *rfp* gene this GlcNAc was substituted with Gal. In hybrids which had both the *rfp* gene and derivatives of the *S. dysenteriae* 1 *rfb* cluster which supplied one or both Rha transferases, this Gal was additionally substituted with one or two Rha residues.

Role of *rfbX*. The *rfbX* genes which have been studied in *Salmonella* and *Shigella* strains encode large, hydrophobic proteins with multiple membrane-spanning domains (65, 77, 89, 139). Mutations which inactivate *rfbX* have been difficult to study. They are deleterious in strains in which all of the other genes involved in synthesizing an O unit are functional, and thus they are unstable on plasmids and are difficult to isolate and manipulate (77). In addition, *E. coli* K-12 has one or more chromosomal genes which encode *rfbX* analogs, and this has further complicated the analysis of *rfbX* mutations carried on plasmids. A plasmid which carried the *S. dysenteriae* *rfb* cluster with a nonpolar *rfbX::Tnlac* insertion mutation resulted in a significant amount of O antigen when it was introduced into an *E. coli* K-12 strain background with an intact *his-rfb* region, indicating that *E. coli* K-12 had an *rfbX* analog which was able to effectively complement the mutation in the plasmid. However, when this *rfbX::Tnlac* plasmid was tested in a strain background deleted for the *his-rfb* region, some SR LPS was made but there was no detectable O ladder. The *his-rfb* deletion also affected the production of band 1 in strains without *rfb* plasmids and reduced the amount of Rfp-modified band (77, 78). Unfortunately, the possibility that the slightly leaky phenotype of the strain deleted for *his-rfb* is due to partial complementation by an *rfbX* analog which is part of another carbohydrate cluster (for example, *rfe-rff*) cannot be ruled out.

The observation in the *his-rfb*-deleted background that the inactivation of *rfbX* in the plasmid led to a complete loss of S LPS but only a partial reduction of SR LPS suggests that the primary role of *rfbX* is not in the assembly of the monomeric O unit on ACL. Likewise, the observation that these mutants are much less efficient in synthesizing SR LPS than are *rfc* mutants indicates that the primary role of *rfbX* is unlikely to be in polymerization of O antigen but suggests that its role may be in the ligation of O units to the LPS or in some other part of the process (for example, the flip-flop of O units on ACL to the periplasmic face of the cytoplasmic membrane) which may be necessary for the efficient transfer of O antigen to LPS.

O-Antigen Genes of *E. coli*

Genetic analysis of *rfb* region of *E. coli* K-12. As noted above, there is genetic evidence that the *his* region of wild-type *E. coli* K-12 contains an intact *rfbBCAD* cluster (4, 77, 78), a functional analog of *rfbX* (77), a *rol* gene (78), and a gene encoding an O-antigen acetylase (182). Analysis of sequences from the GenBank data base indicates that the insertion element IS5I (168) lies 5' to *gnd* (103) and is probably in the distal end of the *rfb* cluster. There are no significant open reading frames between the IS5 and *gnd*, so it is not clear how the insertion is related to the lack of *rfb* function.

***rfb* clusters from other *E. coli* serovars.** Functionally complete *rfb* gene clusters have been cloned from a number of different serovars, including O2 (109), O4 (52, 53), O7 (95, 169), O9 (72, 162), O101 (57), and O111 (7). Most of these have not yet been sequenced or analyzed genetically to the point at which significant comparisons to the *rfb* clusters in *Salmonella* spp. can be made.

An interesting *rfb* cluster which has been partially analyzed is that of the *E. coli* O7:K1 strain VW187 (94). The distal end of this very large cluster contains the two genes *rfbM* and *rfbK* involved in GDP-mannose synthesis, which were identified by the similarity of their products to those of the corresponding *Salmonella* genes. These genes are located immediately adjacent to *gnd*, unlike the situation in the *Salmonella rfb* clusters, where *rfbM* and *rfbK* are separated from *gnd* by the *rfbP* gene. This suggests that *E. coli* O7:K1, like *S. dysenteriae* 1, uses a gene other than *rfbP* to initiate the synthesis of O units. Since the O7 O antigen contains GlcNAc, it is possible that this organism also used the *rfe* gene for this function.

Assembly and Translocation of LPS Containing an O Antigen

Regulation of O-antigen polymerization. In wild-type strains which produce S LPS, the distribution of O-antigen chain lengths is clearly bimodal. SR LPS which contains a single O unit is very abundant, but molecules containing 2, 3, and 4 units are progressively less abundant and the ladder becomes quite faint up to molecules with about 15 O units. Above that, the abundance increases again, reaches a plateau of molecules with a modal number of O units which is specific for the strain and the growth temperature, and then declines rapidly above a limiting maximum size which is usually from about 25 to 35 O units per molecule (48, 50). Mathematical analysis indicates that a random mechanism in which there is a fixed probability determining whether a given O unit or polymer of O units is transferred to a new O unit or to the LPS core does not yield a bimodal distribution of molecular sizes and instead would result in a ladder in which the abundance of each species decreases in proportion to the number of O units it carries. It was proposed that the bimodal distribution resulted from an additional mechanism by which larger molecules were preferentially selected for ligation onto LPS cores (48).

The existence of such a mechanism was confirmed by the observation that loss of a single gene termed *rol* (10) or *clt* (7, 8) eliminated the bimodal distribution of O lengths and resulted in the unimodal distribution of lengths which is predicted by random ligation and polymerization. The *rol* (*clt*) gene is located near the *rfb* cluster, between the *his* operon and *gnd* (8, 9, 77). The sequence of Rol protein is rather strongly conserved across species lines, especially in

comparison with other proteins such as RfaL, RfbX, and Rfc, which are thought to be involved in polymerization and ligation of O antigens (77, 139). The function of Rol is not O-antigen specific, since efficient cross-regulation is observed among different *rol* alleles and *rfb* clusters from smooth serovars of *E. coli*, *S. dysenteriae* 1, and various serovars of *S. typhimurium* (8–10, 77). However, the modal distribution of chain length shows specificity for *rol* alleles, indicating that the determination of chain length is an intrinsic property of Rol (8).

A model has recently been proposed in which polymerization occurs by transferring the growing O chains back and forth between two sites, one of which contains a new O unit, in a manner similar to that involved in ribosomes and in fatty acid synthesis (8). This model envisions the polymerase as existing in two states, an extension state which favors polymerization and a transfer state which facilitates ligation of the O antigen to the LPS core, and suggests that Rol acts as a molecular clock regulating the transition between these states. This kind of kinetic model could also explain the observation that the modal chain length varies as a function of growth temperature and predicts that this temperature effect will also show *rol* allele specificity.

Compartmentalization and translocation of O antigen. Three lines of evidence indicate that at least some steps of O-antigen synthesis and assembly may take place in the periplasmic compartment. The first is the fact that the process involves the synthesis and polymerization of O units on ACL. This 55-carbon lipid is a member of a family of polyisoprene lipids which are represented in all prokaryotic and eukaryotic cells. In some cases these lipids are involved in biosynthetic pathways which appear to begin on one side of a membrane and are completed on the other. The best example is the synthesis of peptidoglycan, in which monomers are transferred from complex sugar nucleotides to ACL in the cytoplasmic compartment and then transferred from ACL to polymeric acceptors located outside of the cytoplasmic membrane (116).

The second line of evidence is the size of the completed O-specific polymer. Of the four basic models which have been proposed for the translocation of O-antigen-containing LPS to the outer membrane, at least two (models C and D in reference 118) would require packaging of the O antigen into a very compact structure which could fit into a small shuttle vesicle or the space between two bilayers in a Beyer's bridge joining the inner and outer membranes (118). Assuming the dimensions for O units given in reference 71, the largest individual O-specific side chains would occupy a sphere of at least 3 nm in diameter under the most stringent packing conditions and perhaps several times that size if reasonable adjustments are made for hydration and the stiffness of the glycosidic chain. Membrane-enclosed structures capable of transporting molecules of this size should be visible by electron microscopy, but no such structures have been reported. In contrast, the periplasmic compartment can easily accommodate very long polysaccharide chains in an extended configuration.

The third line of evidence is that cell surface polysaccharides have been visualized in the periplasmic compartment by electron microscopy. A very dramatic accumulation of high-molecular-weight capsular material is observed in the periplasmic compartment of mutants which have a defect in the export of type II capsular polysaccharide (14, 15, 81). Similar export-defective mutants accumulating O-specific polysaccharide have not been described, but the presence of O-specific material has been demonstrated immunochemi-

cally when core synthesis was blocked (107), and it has been shown with doubly conditional mutants that ACL-linked polymer accumulated during inhibition of core synthesis can be chased into LPS (98). Nevertheless, in both systems some caution must be used in interpreting the significance of the presence of material destined for the outer membrane in the periplasmic compartment, since this may represent a default pathway rather than an indication of true intermediates.

The models in which periplasmic O antigen is a direct precursor of LPS on the surface of the cell require the LPS to flip-flop across a membrane twice, once to flip the core-lipid A complex across the cytoplasmic membrane from the cytoplasmic face to the periplasmic face prior to attachment of O antigen and a second time to flip the complete LPS molecule across the outer membrane from the periplasmic face to the external face (129). Although it is not too difficult to imagine the flip-flop of four or five sugars attached to ACL, it is somewhat more difficult to imagine a flip-flop of LPS. At minimum, the lipid A portion of LPS has six acyl chains attached to a rigid, planar disaccharide head group, and this makes it much more likely to form stable bilayer structures than the single, long, flexible acyl chain of ACL. In addition, LPS has a number of widely spaced charged groups, a bulky core of at least 10 sugars, and, in the case of the complete molecule which must be flipped across the outer membrane, a long extended O-polysaccharide chain of perhaps 100 sugars.

A model which should be added to those that have been proposed is one in which the lipid A portion of the LPS molecule remains on one side of the bilayer and part of the core and/or part of the O chain extend through a protein channel into the periplasmic compartment. The core has the size and flexibility to lie flat on the surface of the membrane or to be folded back into the bilayer, and the outer core plus a single O unit are long enough to extend across the bilayer through a protein channel (71). Protein channels which allow polysaccharide chains to cross a membrane bilayer are not uncommon in bacteria, a good example being the LamB protein, which is involved in maltodextrin transport. If translocation to the outer membrane involves a shuttle vesicle or Beyer's bridge mechanism, one requirement of this model is that channel proteins which coat the polysaccharide as it passes through the membrane would have to accompany the LPS until translocation to the outer membrane was complete, and these proteins would have to be recycled to the cytoplasmic membrane at the end of the cycle.

A useful approach to these models would be to determine the topology of the active sites of enzymes involved in O-antigen assembly and attachment to the LPS core. Unfortunately, this will be difficult with some of the proteins which are likely to be involved, since the RfaL, RfbX, Rfc, RfbP, Rfe, and Rol proteins are all large, hydrophobic proteins with multiple membrane-spanning domains. An alternative approach might be to study the location and properties of the proteins involved in the addition of acetyl (182) or glucosyl (55, 93) modifications to the core, since there is evidence which indicates that these substituents are added after the O units have been synthesized and at least partially polymerized. This approach would be particularly interesting in view of a recent study of the glucosylation which determines O-antigen factor 12₂ in *S. typhimurium*, since an electrophoretic comparison of the O-antigen ladders of glucosylated and nonglucosylated LPS suggested that glucosylation was present only in O units distal to the sixth O unit in the polymer (55). Since new units are added to the proximal end

of the polymer, this finding is very strongly suggestive of some form of protection or compartmentalization of the O antigen during one stage of its synthesis.

Attachment of O antigen to the LPS core. On the basis of the early studies which demonstrated that O units were synthesized and polymerized on ACL and genetic studies of the polymerase, it was implied that the polymerization reaction and the transfer (more commonly termed ligation) were distinct processes involving different gene products (93). The predicted phenotype of mutants lacking O ligase is the presence of polymeric O antigen bound to ACL but no O antigen attached to LPS. This is the phenotype exhibited by *rfaL* mutants and by a few mutants with mutations at the *rfb* locus which were designated as *rfbT* mutants (93). No gene corresponding to *rfbT* has been found among the *rfb* clusters which have been sequenced, although a direct role for *rfbX* in ligation cannot yet be ruled out. The *Salmonella rfbT* mutants have been found to have a defect in *rfbP*, suggesting that the large protein encoded by this gene may be bifunctional (138).

This raises the interesting question of how strains which lack an *rfbP* gene, such as *Salmonella* group C1 (82), *S. dysenteriae* 1 (77), and probably *E. coli* O7:K1 (94), can transfer O antigen from ACL to LPS. The answer may lie in a hydrophobic protein, designated as o349, encoded by a gene which lies adjacent to *rfe* (36, 101). Joining of the reading frames of *rfe* and half of the gene encoding o349 resulted in a larger protein with a hydrophathy profile which was remarkably similar to that of RfbP protein, suggesting that *rfbP* arose by a similar fusion of ancestral genes (77). More work is clearly needed to sort out the genetics of the ligation reaction.

In addition to considering the role of various gene products in the ligation reaction, it is useful to examine two different models for the ligation pathway. We term these the single-step and multistep models, and a highly speculative comparison of the roles of the various gene products in each of these models is given in Table 4. The single-step model is the traditional model for ligation developed more than two decades ago and involves the *rfc*-dependent polymerization of O antigen on ACL followed by a single-step ligation reaction in which completed O polymers are transferred from ACL to the terminal core sugar of the acceptor LPS. In this model, ligation of polymeric O antigen does not directly involve *rfe*. The multistep model begins with the ligation of a single O unit to the terminal core sugar of the acceptor LPS to form SR LPS. Polymers are then transferred to this single O unit in an *rfe*-dependent reaction which is analogous to the polymerization reaction. Both models require a ligase to transfer a sugar at the reducing terminus of an O unit from ACL to the LPS core and a polymerase to generate ACL-linked polymers containing multiple O units, so the models are equivalent in complexity. Both models begin with the synthesis of a pool of ACL-linked O units, and the structure of the sugar transferases indicates that these are assembled on the cytoplasmic face of the cytoplasmic membrane. Beyond this step, neither model imposes any obvious constraints on whether reaction take place on the cytoplasmic or periplasmic face of the membrane or within protein channels embedded in the membrane.

The multistep model has some advantages in explaining the large amount of SR LPS observed in smooth strains and *rfe* mutants, as well as the apparently efficient transfer of partial O units in rough strains. The presence of an additional three- to five-sugar extension on the core provided by the ligation of a single O unit to the core prior to the attachment

TABLE 4. Comparison of gene functions predicted by the single-step and multistep models for O-antigen ligation

| Gene or phenotype | Function predicted by: | |
|-------------------|--|--|
| | Single-step model | Multistep model |
| <i>rfaL</i> | Component of ligase which adds polymerized O antigen to core | Component of ligase which adds single O unit to core |
| RfbT ^a | Component of ligase which adds polymerized O antigen to core | Component of ligase which adds single O unit to core |
| <i>rfc</i> | Polymerization of O units by repeated transfer of growing chain to single O units on ACL | Polymerization of O units by repeated transfer of growing chain to single O units on ACL, and ligation of polymerized O units by transfer to single O unit on core |
| <i>rol (cld)</i> | Regulates kinetics of polymerization by Rfc versus ligation to core by RfaL and/or RfbT | Regulates kinetics of polymerization by Rfc versus ligation to core by Rfc |

^a This phenotype is used to indicate the RfbT function proposed for the bifunctional *rfbP* gene product of group B *Salmonella* spp. (138) or the product of a gene from the *rfe-rff* cluster in *E. coli* K-12-*S. dysenteriae* 1 hybrids (77).

of polymeric O antigen might be advantageous in the early stages of a model in which periplasmic O antigen is tethered across the cytoplasmic membrane to lipid A on the cytoplasmic face. None of these arguments are compelling in favor of the multistep model. A much more important reason for considering both models is that they differ in their requirements in a reconstituted *in vitro* system. In the multistep model, SR LPS would be required as the O-antigen acceptor and the reaction would require Rfc protein, whereas in the single-step model, R LPS would be used as the acceptor and Rfc protein would not necessarily be required.

In contrast to other parts of the LPS biosynthetic pathway, reconstitution systems have not provided a useful way of studying the ligation reaction. This is likely to change now that most of the genes involved LPS assembly have been cloned and sequenced, since *in vitro* genetics provides a means of constructing strains tailored for reconstitution studies which have precisely defined enzymatic lesions or produce LPS acceptors in which assembly has been terminated at specific stages. In addition, knowledge of the structure of the gene products can identify which functions are likely to require intrinsic transmembrane proteins, which must be supplied as membrane fragments or detergent extracts, rather than the more peripheral proteins, which can be supplied by soluble extracts and which are more amenable to synthesis in large amounts by expression systems.

CONCLUDING REMARKS

The past decade has seen the development of a nearly complete catalog of LPS genes. This has both exposed our ignorance and provided a powerful tool for future studies. In its early years, LPS research was dominated by *in vitro* reconstitution studies in which mutants, cell extracts, and labeled precursors were combined to work out complicated biosynthetic pathways. This tradition has continued more recently with the nearly complete elucidation of the lipid A biosynthetic pathway. Now the availability of sequence data on LPS genes and the ease with which defined mutations can be constructed and crossed onto the chromosome allow these reconstitution experiments to be repeated with different goals and at a much higher level of resolution. These are probably the most important experiments to be undertaken, since they offer a way of studying questions such as the organization of multienzyme complexes in the membrane and the mechanism of the final steps of assembly and translocation of LPS, which have previously been intractable.

This period has also been a time when there has been a more acute recognition of the tremendous flexibility and adaptability of the gram-negative cell surface, and this has changed the way we view LPS. The complex oligosaccharides which decorate glycoproteins in eukaryotic organisms have been likened in function to the bar codes used to mark items in a supermarket (41), and it is useful to extend this analogy to LPS and other complex bacterial polysaccharides. These molecules provide a genetically inexpensive way to allow cells to specifically recognize each other, to be recognized by hosts, to colonize favorable host environments, and to evade or confuse predators. Originally, LPS pathways were viewed as linear pathways designed to produce a single major product, and the heterogeneity resulting from the presence of partial branch substituents could be explained on the basis of the inefficiency of some of the biosynthetic enzymes. New evidence indicates that LPS pathways are branched so that multiple products can be produced, and the addition or removal of partial sugar substituents such as Rha, Hep, or GlcNAc may be used at branch points to target molecules for different pathways. This shows much more similarity to the kinds of mechanisms which are used to target glycoforms in eukaryotic systems (43).

An important aspect of this flexibility is preliminary but growing evidence suggesting that repeating polymers such as O antigens, ECA, and colanic acids, which are assembled on undecaprenol (ACL) carriers, may be interchangeably transferred to the LPS core and to smaller acceptors such as simple phospholipids. This can allow some strains to express these polymers either as LPS, which is a firmly bound, integral component of the outer membrane, or as microcapsular polysaccharides, which are much easily solubilized and dispersed in water as micellar suspensions. Thus the same polymer can be found in both a bound and a diffusible form, in some ways analogous to human blood group substances. The biological importance of this strategy is underlined by a recent discovery in the *Rhizobium*-legume colonization system. Bacterial nodulation-specific (*nod*) genes which are responsible for the species or cultivar specificity of nodulation were found to encode enzymes for the site-specific sulfation of a diffusible complex polysaccharide (145). This complex polysaccharide was termed a LOS, but it is attached to a phospholipid anchor rather than lipid A and has a structure more similar to that of ECA.

ACKNOWLEDGMENTS

We thank the many people who provided reprints of their work, and P. MacLachlan, J. Petter, C. Raetz, P. Reeves, P. Rick, K. Timmis and M. Valvano for providing work prior to publication. Preparation of this review was assisted by Public Health Services grant GM-39087.

REFERENCES

- Aasland, R., J. Coleman, A. L. Holck, C. L. Smith, C. R. H. Raetz, and K. Kleppe. 1988. Identity of the 17-kilodalton protein, a DNA binding protein from *E. coli*, and the *firA* gene product. *J. Bacteriol.* **170**:5916-5918.
- Anderson, M. S., and C. R. H. Raetz. 1987. Biosynthesis of lipid A precursors in *Escherichia coli*. A cytoplasmic acyl transferase that converts UDP-N acetylglucosamine to UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine. *J. Biol. Chem.* **262**:5159-5169.
- 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.
- 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.
- Bachmann, B. J. 1990. Linkage map of *Escherichia coli* K-12, edition 8. *Microbiol. Rev.* **54**:130-197.
- Bailey, M. J. A., V. Koronakis, T. Schmoll, and C. Hughes. 1992. *Escherichia coli* HlyT protein, a transcriptional activator of haemolysin synthesis and secretion, is encoded by the *rfaH* (*sfrB*) locus required for expression of sex factor and lipopolysaccharide genes. *Mol. Microbiol.* **6**:1003-1012.
- 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.
- Bastin, D. A., G. Stevenson, P. K. Brown, A. Haase, and P. R. Reeves. 1993. Repeat unit polysaccharides of bacteria: a model for polymerisation resembling those of ribosomes and fatty acid synthetase, with a novel mechanism for determining chain length. *Mol. Microbiol.* **7**:725-734.
- 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.
- 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.
- Belunis, C. J., K. E. Mdluli, C. R. H. Raetz, and F. E. Nano. 1992. A novel 3-deoxy-D-manno-octulosonic acid transferase from *Chlamydia trachomatis* required for expression of the genus-specific epitope. *J. Biol. Chem.* **267**:18702-18707.
- Belunis, C. J., and C. R. H. Raetz. 1992. Biosynthesis of endotoxins. Purification and catalytic properties of 3-deoxy-D-manno-octulosonic acid transferase from *Escherichia coli*. *J. Biol. Chem.* **267**:9988-9997.
- Beutin, L., P. A. Manning, M. Achtman, and N. Willits. 1981. *sfrA* and *sfrB* products of *Escherichia coli* K-12 are transcription control factors. *J. Bacteriol.* **145**:840-844.
- Boulnois, G. J., and K. Jann. 1989. Bacterial polysaccharide capsule synthesis, export and evolution of structural diversity. *Mol. Microbiol.* **3**:1819-1823.
- Boulnois, G. J., I. S. Roberts, R. Hodge, K. Hardy, K. Jann, and K. N. Timmis. 1987. Further analysis of the cloned K1 capsule biosynthesis genes of *Escherichia coli*: definition of three functional regions for capsule production. *Mol. Gen. Genet.* **200**:242-246.
- Brade, H., L. Brade, and F. E. Nano. 1987. Chemical and serological investigations on the genus-specific lipopolysaccharide epitope of *Chlamydia*. *Proc. Natl. Acad. Sci. USA* **84**:2508-2512.
- Brahmbhatt, H. N., A. A. Lindberg, and K. N. Timmis. 1992. *Shigella* lipopolysaccharide: structure, genetics and vaccine development. *Curr. Top. Microbiol. Immunol.* **180**:45-63.
- Brahmbhatt, H. N., P. Wyk, N. B. Quigley, and P. R. Reeves. 1988. Complete physical map of the *rfb* gene cluster encoding biosynthetic enzymes for the O antigen of *Salmonella typhimurium* LT2. *J. Bacteriol.* **170**:98-102.
- Brazas, R., E. Davie, A. Farewell, and L. I. Rothfield. 1991. Transcriptional organization of the *rfaGBIJ* locus of *Salmonella typhimurium*. *J. Bacteriol.* **173**:6168-6173.
- Brown, P. K., L. K. Romana, and P. R. Reeves. 1991. Cloning of the *rfb* gene cluster of a group C2 *Salmonella* strain: comparison with the *rfb* regions of groups B and D. *Mol. Microbiol.* **5**:1873-1881.
- Brown, P. K., L. K. Romana, and P. R. Reeves. 1992. Molecular analysis of the *rfb* gene cluster of *Salmonella* serovar muenchen (strain M67): the genetic basis of the polymorphism between groups C2 and B. *Mol. Microbiol.* **6**:1385-1394.
- Brozek, K. A., K. Hosaka, A. D. Robertson, and C. R. H. Raetz. 1989. Biosynthesis of lipopolysaccharide in *Escherichia coli*. Cytoplasmic enzymes that attach 3-deoxy-D-manno-octulosonic acid to lipid A. *J. Biol. Chem.* **264**:6956-6966.
- Brozek, K. A., and C. R. H. Raetz. 1990. Biosynthesis of lipid A in *Escherichia coli*. Acyl carrier protein-dependent incorporation of laurate and myristate. *J. Biol. Chem.* **265**:15410-15417.
- Carstenius, P., J.-I. Flock, and A. Lindberg. 1990. Nucleotide sequence of *rfaI* and *rfaJ* genes encoding lipopolysaccharide glycosyl transferases from *Salmonella typhimurium*. *Nucleic Acids Res.* **18**:6128.
- Clark, C. A., J. Beltrame, and P. A. Manning. 1991. The *oac* gene encoding a lipopolysaccharide O-antigen acetylase maps adjacent to the integrase-encoding gene on the genome of *Shigella flexneri* bacteriophage Sf6. *Gene* **107**:43-52.
- Clementz, T. 1992. The gene coding for 3-deoxy-D-manno-octulosonic acid transferase and the *rfaQ* gene are transcribed from divergently arranged promoters in *Escherichia coli*. *J. Bacteriol.* **174**:7750-7756.
- 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.* **266**:9687-9696.
- Coleman, J., and C. R. H. Raetz. 1988. First committed step of lipid A biosynthesis in *Escherichia coli*: sequence of the *lpxA* gene. *J. Bacteriol.* **170**:1268-1274.
- Coleman, W. G. 1983. The *rfaD* gene codes for ADP-L-glycero-D-manno-heptose-6-epimerase. *J. Biol. Chem.* **258**:1985-1990.
- Collins, L. V., and J. Hackett. 1991. Molecular cloning, characterisation, and nucleotide sequence of the *rfaH* gene encoding an O-antigen polymerase of *Salmonella typhimurium*. *J. Bacteriol.* **173**:2521-2529.
- Creeger, E. S., and L. I. Rothfield. 1979. Cloning of genes for bacterial glycosyltransferases I. Selection of hybrid plasmids carrying genes for two glycosyltransferases. *J. Biol. Chem.* **254**:804-810.
- Creeger, E. S., J. F. Chen, and L. I. Rothfield. 1979. Cloning of genes for bacterial glycosyl transferases. II. Selection of a hybrid plasmid carrying the *rfaH* gene. *J. Biol. Chem.* **254**:811-815.
- Creeger, E. S., T. Schulte, and L. I. Rothfield. 1984. Regulation of membrane glycosyl transferases by the *sfrB* and *rfaH* genes of *E. coli* and *S. typhimurium*. *J. Biol. Chem.* **259**:3064-3069.
- Cronan, J. E., and C. O. Rock. 1987. Biosynthesis of membrane lipids, p. 474-497. 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.
- Crowell, D. N., W. S. Reznikoff, and C. R. H. Raetz. 1987. Nucleotide sequence of the *Escherichia coli* gene for lipid A disaccharide synthetase. *J. Bacteriol.* **169**:5727-5734.
- Daniels, D. L., G. Plunkett III, V. Burland, and F. R. Blattner.

1992. Analysis of the *Escherichia coli* genome: DNA sequence of the region from 84.5 to 86.5 minutes. *Science* **257**:771-778.
37. DeChavigny, A., P. N. Heacock, and W. Dowhan. 1991. Sequence and inactivation of the *pss* gene of *Escherichia coli*. *J. Biol. Chem.* **266**:5323-5332.
 38. Dicker, I. B., and S. Seetharam. 1991. Cloning and nucleotide sequence of the *firA* gene and the *firA200*(Ts) allele from *Escherichia coli*. *J. Bacteriol.* **173**:334-344.
 39. Dicker, I. B., and S. Seetharam. 1992. What is known about the structure and function of the *Escherichia coli* protein FirA? *Mol. Microbiol.* **6**:817-823.
 40. Dmitriev, B. A., Y. A. Knirel, N. K. Kochetkov, and I. L. Hofman. 1976. Somatic antigens of *Shigella*. Structural investigations on the O-specific polysaccharide chain of *Shigella dysenteriae* type 1 lipopolysaccharide. *Eur. J. Biochem.* **66**:559-566.
 41. Drickamer, K. 1991. Clearing up glycoprotein hormones. *Cell* **67**:1029-1032.
 - 41a. Fält, I. C., E. Schweda, A. Weintraub, S. Sturm, K. N. Timmis, and A. A. Lindberg. 1993. Expression of the *Shigella dysenteriae* type 1 lipopolysaccharide repeating unit in *Escherichia coli* K-12-*Shigella dysenteriae* type 1 hybrids. *Eur. J. Biochem.* **213**:573-581.
 42. Farewell, A., R. Brazas, E. Davie, J. Mason, and L. I. Rothfield. 1991. Suppression of the abnormal phenotype of *Salmonella typhimurium rfaH* mutants by mutations in the gene for transcription termination factor Rho. *J. Bacteriol.* **173**:5188-5193.
 43. Farquahr, M. G. 1985. Progress in unraveling pathways of golgi traffic. *Annu. Rev. Cell Biol.* **1**:447-488.
 44. Gabriel, O. 1987. Biosynthesis of sugar residues for glycogen, peptidoglycan, lipopolysaccharide, and related systems, p. 504-511. *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.
 45. Galloway, S. M., and C. R. H. Raetz. 1990. A mutant of *Escherichia coli* defective in the first step of endotoxin biosynthesis. *J. Biol. Chem.* **265**:6394-6402.
 46. Gibson, B. W., J. W. Webb, R. Yamasaki, S. J. Fisher, A. L. Burlingame, R. E. Mandrell, H. Schneider, and J. M. Griffiss. 1989. Structure and heterogeneity of the oligosaccharides from lipopolysaccharides of *Neisseria gonorrhoeae*. *Proc. Natl. Acad. Sci. USA* **86**:17-21.
 47. Goldman, R. C., T. J. Bolling, W. E. Kohlbrenner, Y. Kim, and J. L. Fox. 1986. Primary structure of CTP: CMP-3-deoxy-D-manno-octulosonate cytidyl-transferase (CMP-KDO synthetase) from *Escherichia coli*. *J. Biol. Chem.* **261**:15831-15835.
 48. Goldman, R. C., and F. Hunt. 1990. Mechanism of O-antigen distribution in lipopolysaccharide. *J. Bacteriol.* **172**:5352-5359.
 49. Golenbeck, D. T., R. Y. Hampton, N. Quershi, K. Takayama, and C. R. H. Raetz. 1991. Lipid A-like molecules that antagonize the effects of endotoxins on human monocytes. *J. Biol. Chem.* **266**:19490-19498.
 50. Grossman, N., M. A. Schmetz, J. Foulds, E. N. Klima, V. Jiminez, L. L. Leive, and K. Joiner. 1987. Lipopolysaccharide size and distribution determine serum resistance in *Salmonella montevideo*. *J. Bacteriol.* **169**:856-863.
 51. Hale, T. L., P. Guerry, R. C. Seid, Jr., C. Kapfer, M. E. 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.
 52. Haraguchi, G. E., R. A. Hull, U. Krallman-Wenzel, and S. I. Hull. 1989. Molecular cloning and expression of the O4 polysaccharide gene cluster from *Escherichia coli*. *Microb. Pathog.* **6**:123-132.
 53. Haraguchi, G. E., U. Zähringer, 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.
 54. Hasin, M., and E. P. Kennedy. 1982. Role of phosphatidylethanolamine in the biosynthesis of pyrophosphoethanolamine residues in lipopolysaccharide of *Escherichia coli*. *J. Biol. Chem.* **257**:12475-12477.
 55. Helander, I. M., A. P. Moran, and P. H. Mäkelä. 1992. Separation of two lipopolysaccharide populations with different contents of O-antigen factor 12₂ in *Salmonella enterica* serovar Typhimurium. *Mol. Microbiol.* **6**:2857-2862.
 56. Helander, I. M., M. Vaara, S. Sukupolvi, M. Rhen, S. Saarela, U. Zähringer, and P. H. Mäkelä. 1989. *rfaP* mutants of *Salmonella typhimurium*. *Eur. J. Biochem.* **185**:541-546.
 57. Heuzenroeder, M. W., D. W. Berger, C. J. Thomas, and P. A. Manning. 1989. Molecular cloning and expression in *Escherichia coli* K-12 of the O101 *rfb* region from *E. coli* B41 (O101:K99/F41) and the genetic relationship to other O101 *rfb* loci. *Mol. Microbiol.* **3**:295-302.
 58. Hirvas, L., P. Koski, and M. Vaara. 1991. The *ompH* gene of *Yersinia enterocolitica*: cloning, sequencing, expression, and comparison with known enterobacterial *ompH* sequences. *J. Bacteriol.* **173**:1223-1229.
 59. Holst, O., and H. Brade. 1990. Isolation and identification of 3-deoxy-5-O- α -L-rhamnopyranosyl-D-manno-2-octulopyranosonate from the inner core region of the lipopolysaccharide of *Escherichia coli* K-12. *Carbohydr. Res.* **207**:327-331.
 60. 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.
 61. Holst, O., E. Rohrscheidt-Andrzejewski, H. Brade, and D. Charon. 1990. Isolation and characterisation of 3-deoxy-D-manno-2-octulopyranosonate 7-(2-aminoethyl phosphate) from the inner core region of *Escherichia coli* K-12 and *Salmonella minnesota* lipopolysaccharides. *Carbohydr. Res.* **204**:93-102.
 62. 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**:323-335.
 63. 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 Biomedical Press, Amsterdam.
 64. 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.
 65. 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.
 66. John, C. M., J. M. Griffiss, M. A. Apicella, R. E. Mandrell, and B. W. Gibson. 1991. The structural basis for pyocin resistance in *Neisseria gonorrhoeae* lipooligosaccharides. *J. Biol. Chem.* **266**:19301-19311.
 67. Karow, M., O. Fayet, A. Cegielska, T. Ziegelhofer, and C. Georgopoulos. 1991. The *Escherichia coli htrB* gene, whose product is essential for bacterial viability above 33°C in rich media: isolation and characterization. *J. Bacteriol.* **173**:741-750.
 68. Karow, M., O. Fayet, and C. Georgopoulos. 1992. The lethal phenotype caused by null mutations in the *Escherichia coli htrB* gene is suppressed by mutations in the *accBC* operon, encoding two subunits of acetyl coenzyme A carboxylase. *J. Bacteriol.* **174**:7407-7418.
 69. Karow, M., and C. Georgopoulos. 1991. Sequencing, mutational analysis, and transcriptional analysis of the *Escherichia coli htrB* gene. *Mol. Microbiol.* **5**:2285-2292.
 70. Karow, M., and C. Georgopoulos. 1992. Isolation and characterization of the *Escherichia coli msbB* gene, a multicopy suppressor of null mutations in the high-temperature requirement gene *htrB*. *J. Bacteriol.* **174**:702-710.
 71. Kastowski, M., T. Gutberlet, and H. Bradaczek. 1992. Molecular modeling of the three-dimensional structure and confor-

- mational flexibility of bacterial lipopolysaccharide. *J. Bacteriol.* 174:4798-4806.
72. Kido, N., M. Ohta, K. Iida, T. Hasegawa, H. Ito, Y. Arakawa, T. Komatsu, and N. Kato. 1989. Partial deletion of the cloned *rfb* gene of *Escherichia coli* O9 results in synthesis of a new O-antigenic lipopolysaccharide. *J. Bacteriol.* 171:3629-3633.
 73. Klena, J. D. Unpublished data.
 74. Klena, J. D., R. S. Ashford, and C. A. Schnaitman. 1992. Role of *Escherichia coli* K-12 *rfa* genes and the *rfp* gene of *Shigella dysenteriae* 1 in the generation of lipopolysaccharide core heterogeneity and attachment of O antigen. *J. Bacteriol.* 174:7297-7303.
 75. 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.
 76. Klena, J. D., E. Pradel, and C. A. Schnaitman. 1993. The *rfaS* gene which is involved in production of a rough form of lipopolysaccharide core in *Escherichia coli* K-12 is absent in *Salmonella typhimurium*. *J. Bacteriol.* 175:1524-1527.
 77. Klena, J. D., and C. Schnaitman. Function of the *rfb* gene cluster and the *rfe* gene in the synthesis of O antigen by *Shigella dysenteriae* 1. *Mol. Microbiol.*, in press.
 78. Klena, J. D., and C. Schnaitman. Synthesis of TDP-rhamnose is required for the lipooligosaccharide (LOS) form of *Escherichia coli* K-12 lipopolysaccharide. Submitted for publication.
 79. Komeda, Y., T. Icho, and T. Iino. 1977. Effects of *galU* mutation on flagellar formation in *Escherichia coli*. *J. Bacteriol.* 129:908-915.
 80. Koski, P., M. Rehn, J. Kantele, and M. Vaara. 1989. Isolation, cloning and primary structure of a cationic 16-kDa outer membrane protein of *Salmonella typhimurium*. *J. Biol. Chem.* 264:18973-18980.
 81. Kroncke, K.-D., G. Boulnois, I. Roberts, D. Bitter-Suermann, J. R. Golecki, B. Jann, and K. Jann. 1990. Expression of the *Escherichia coli* K5 capsular antigen: immunoelectron microscopic and biochemical studies with recombinant *E. coli*. *J. Bacteriol.* 172:1085-1091.
 82. Lee, S. J., L. K. Romana, and P. R. Reeves. 1992. Cloning and structure of group C1 O antigen (*rfb* gene cluster) from *Salmonella enterica* serovar *montevideo*. *J. Gen. Microbiol.* 138:305-312.
 83. Lee, S. J., L. K. Romana, and P. R. Reeves. 1992. Sequence and structural analysis of the *rfb* (O antigen) gene cluster from a group C1 *Salmonella enterica* strain. *J. Gen. Microbiol.* 138:1843-1855.
 84. Lehmann, V., O. Lüderitz, and O. Westphal. 1971. The linkage of pyrophosphorylethanolamine to heptose in the core of *Salmonella minnesota* lipopolysaccharides. *Eur. J. Biochem.* 21:339-347.
 85. 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.
 86. Levy, S. B., and L. Leive. 1968. An equilibrium between two fractions of lipopolysaccharide in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 61:1435-1439.
 87. Lew, H. C., P. H. Mäkelä, H. Kuhn, H. Mayer, and H. Nikaido. 1986. Biosynthesis of enterobacterial common antigen requires dTDP glucose pyrophosphorylase determined by a *Salmonella typhimurium rfb* gene and a *Salmonella montevideo rfe* gene. *J. Bacteriol.* 168:715-721.
 88. 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.
 89. Liu, D., N. K. Verma, L. K. Romana, and P. R. Reeves. 1991. Relationships among the *rfb* regions of *Salmonella* serovars A, B, and D. *J. Bacteriol.* 173:4814-4819.
 90. Lugowski, C., E. Romanowska, L. Kenne, and B. Lindberg. 1983. Identification of a trisaccharide repeating unit in the enterobacterial common antigen. *Carbohydr. Res.* 118:173-181.
 91. Maclachlan, P. R. Personal communication.
 92. 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.
 93. 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 Biomedical Press, Amsterdam.
 94. Marolda, C. L., and M. A. Valvano. 1993. Identification, expression, and DNA sequence of the GDP-mannose biosynthesis genes encoded by the O7 *rfb* gene cluster of strain VW187 (*Escherichia coli* O7:K1). *J. Bacteriol.* 175:148-158.
 95. Marolda, C. L., J. Welsh, L. Davoe, and M. A. Valvano. 1990. Genetic analysis of the O7-polysaccharide biosynthesis region from the *Escherichia coli* O7:K1 strain VW187. *J. Bacteriol.* 172:3590-3599.
 96. Marumo, K., L. Llandqvist, N. Verma, A. Weintraub, P. R. Reeves, and A. A. Lindberg. 1992. Enzymatic synthesis and isolation of thymidine diphosphate-6-deoxy-D-xylo-4-hexulose and thymidine diphosphate-L-rhamnose. *Eur. J. Biochem.* 204:539-545.
 97. Mayer, H., and J. Weckesser. 1984. Unusual lipid A's: structures, taxonomical relevance and potential value for endotoxin research, p. 221-247. In E. T. Rietschel (ed.), *Handbook of endotoxin*, vol. 1. Chemistry of endotoxin. Elsevier Biomedical Press, Amsterdam.
 98. 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.
 99. McHenry, C. S. 1992. Personal communication.
 100. Meier, U., and H. Mayer. 1985. Genetic location of genes encoding enterobacterial common antigen. *J. Bacteriol.* 163:756-762.
 101. Meier-Dieter, U., K. Barr, R. Starman, L. Hatch, and P. D. Rick. 1992. Nucleotide sequence of the *Escherichia coli rfe* gene involved in the synthesis of enterobacterial common antigen. *J. Biol. Chem.* 267:746-753.
 102. Meier-Dieter, U., R. Starman, K. Barr, H. Mayer, and P. D. Rick. 1990. Biosynthesis of enterobacterial common antigen in *Escherichia coli*. *J. Biol. Chem.* 265:13490-13497.
 103. Miller, R. D., D. E. Dykhuizen, and D. L. Hartl. 1988. Fitness effects of a deletion mutation increasing transcription of the 6-phosphogluconate dehydrogenase gene in *Escherichia coli*. *Mol. Biol. Evol.* 5:691-703.
 104. Misra, R., and P. R. Reeves. 1987. Role of *micF* in the *tolC*-mediated regulation of *OmpF*, a major outer membrane protein of *Escherichia coli* K-12. *J. Bacteriol.* 169:4722-4730.
 105. Morona, R., P. A. Manning, and P. Reeves. 1983. Identification and characterization of the *TolC* protein, an outer membrane protein from *Escherichia coli*. *J. Bacteriol.* 153:693-699.
 106. Mühlradt, P. F. 1969. Biosynthesis of *Salmonella* lipopolysaccharide. The *in vitro* transfer of phosphate to the heptose moiety of the core. *Eur. J. Biochem.* 11:241-248.
 107. Mulford, C. A., and M. J. Osborn. 1983. An intermediate step in translocation of lipopolysaccharide to the outer membrane of *Salmonella typhimurium*. *Proc. Natl. Acad. Sci. USA* 80:1159-1163.
 108. Nakai, T., and H. Nikaido. 1971. Multiple molecular forms of uridine diphosphate glucose pyrophosphorylase from *Salmonella typhimurium*. II. Genetic determination of multiple forms. *J. Biol. Chem.* 246:4397-4403.
 109. Neal, B. L., G. C. Tsolis, M. W. Heuzenroeder, P. A. Manning, and P. R. Reeves. 1991. Molecular cloning and expression in *Escherichia coli* K-12 of chromosomal genes determining the O antigen of an *E. coli* O2:K1 strain. *FEMS Microbiol. Lett.* 82:345-352.
 110. Nikaido, H. 1970. Structure of the cell wall lipopolysaccharide from *Salmonella typhimurium*. Further studies on the linkage between O side chains and R core. *Eur. J. Biochem.* 15:57-62.
 111. Nikaido, H. 1973. Biosynthesis and assembly of lipopolysaccharide and the outer membrane layer of gram-negative cell wall, p. 131-208. In L. Leive (ed.), *Bacterial membranes and*

- walls. Marcel Dekker, Inc., New York.
112. **Nikaido, H., K. Nikaido, and A. M. C. Rapin.** 1965. Biosynthesis of thymidine diphosphate L-rhamnose in *Escherichia coli* K-12. *Biochim. Biophys. Acta* **111**:548-551.
 113. **Nikaido, H., and M. Vaara.** 1985. Molecular basis of bacterial outer membrane permeability. *Microbiol. Rev.* **49**:1-32.
 114. **Normark, S., H. G. Boman, and E. Mattsson.** 1969. Mutant of *Escherichia coli* with anomalous cell division and ability to decrease episodically and chromosomally mediated resistance to ampicillin and several other antibiotics. *J. Bacteriol.* **97**:1334-1342.
 115. **Osborn, M. J., P. D. Rick, and N. S. Rasmussen.** 1980. Mechanism of assembly of the outer membrane of *Salmonella typhimurium*. Translocation and integration of an incomplete mutant lipid A into the outer membrane. *J. Biol. Chem.* **255**:4246-4251.
 116. **Park, J. T.** 1987. Murein synthesis, p. 663-671. *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.
 117. **Parker, C. T.** Unpublished data.
 118. **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.
 119. **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.
 120. **Pegues, J. C., L. Chen, A. W. Gordon, L. Ding, and W. G. Coleman.** 1990. Cloning, expression and characterization of the *Escherichia coli* K-12 *rfaD* gene. *J. Bacteriol.* **172**:4652-4660.
 121. **Petter, J. G.** Personal communication.
 122. **Phillips, N. J., C. M. John, L. G. Reinders, J. M. Griffiss, M. A. Apicella, and B. W. Gibson.** 1990. Structural models for cell surface lipooligosaccharides of *Neisseria gonorrhoeae* and *Haemophilus influenzae*. *Biomed. Environ. Mass Spectrom.* **19**:731-745.
 123. **Pradel, E.** Unpublished data.
 124. **Pradel, E., C. T. Parker, and C. A. Schnaitman.** 1992. The structure 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.
 125. **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.
 126. **Raetz, C. R. H.** Personal communication.
 127. **Raetz, C. R. H.** 1987. Biosynthesis and pharmacologic properties of *Escherichia coli* lipid A, p. 229-245. *In* M. Inouye (ed.), *Bacterial outer membranes as model systems*. Wiley-Interscience, New York.
 128. **Raetz, C. R. H.** 1987. Structure and biosynthesis of lipid A in *Escherichia coli*, p. 498-503. *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.
 129. **Raetz, C. R. H.** 1990. Biochemistry of endotoxins. *Annu. Rev. Biochem.* **59**:129-170.
 130. **Raetz, C. R. H.** 1992. Biosynthesis of lipid A, p. 67-80. *In* D. C. Morrison and J. L. Ryan (ed.), *Bacterial endotoxic lipopolysaccharides*, vol. 1. Molecular biochemistry and cellular biology. CRC Press, Inc., Boca Raton, Fla.
 131. **Raetz, C. R. H., K. A. Brozek, T. Clementz, J. D. Coleman, S. M. Galloway, D. T. Golenbock, and R. Y. Hampton.** 1988. Gram-negative endotoxin: a biologically active lipid. *Cold Spring Harbor Symp. Quant. Biol.* **53**:973-982.
 132. **Raetz, C. R. H., and W. Dowhan.** 1990. Biosynthesis and function of phospholipids in *Escherichia coli*. *J. Biol. Chem.* **265**:1235-1238.
 133. **Raina, S.** Personal communication.
 134. **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.
 135. **Ralling, G., and T. Linn.** 1987. Evidence that Rho and NusA are involved in termination in the *rplL-rpoB* intercistronic region. *J. Bacteriol.* **169**:2277-2280.
 136. **Rapin, A. M. C., and H. M. Kalckar.** 1971. The relation of bacteriophage attachment to lipopolysaccharide structure, p. 267-307. *In* Weinbaum, G., S. Kadis, and S. J. Ajl (ed.), *Microbial toxins*, vol. IV. Bacterial endotoxins. Academic Press, Inc., New York.
 137. **Ray, B. L., G. Painter, and C. R. H. Raetz.** 1984. The biosynthesis of gram-negative endotoxin. Formation of lipid A disaccharide from monosaccharide precursors in extracts of *Escherichia coli*. *J. Biol. Chem.* **259**:4852-4859.
 138. **Reeves, P.** Personal communication.
 139. **Reeves, P.** 1993. Evolution of *Salmonella* O antigen variation involved in interspecies gene transfer on a large scale. *Trends Genet.* **9**:17-22.
 140. **Rehemtulla, A., S. K. Kadam, and K. E. Sanderson.** 1986. Cloning and analysis of the *sfrB* (sex factor repression) gene of *Escherichia coli* K-12. *J. Bacteriol.* **166**:651-657.
 141. **Rick, P. D.** Personal communication.
 142. **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.
 143. **Rick, P. D., and M. J. Osborn.** 1977. Lipid A mutants of *Salmonella typhimurium*. Characterization of conditional lethal mutants in 3-deoxy-D-mannoctulosonate-8-phosphate synthetase. *J. Biol. Chem.* **252**:4895-4903.
 144. **Rick, P. D., and D. A. Young.** 1982. Isolation and characterization of a temperature-sensitive lethal mutant of *Salmonella typhimurium* that is conditionally defective in 3-deoxy-D-mannoctulosonate-8-phosphate synthetase. *J. Bacteriol.* **150**:447-455.
 145. **Roche, P., F. DeBelle, F. Maillet, P. Lerouge, C. Faucher, G. Truchet, J. Denarie, and J.-C. Prome.** 1991. Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipo-oligosaccharide signals. *Cell* **67**:1131-1143.
 146. **Romana, L. K., F. S. Santiago, and P. R. Reeves.** 1991. High level expression and purification of thymidine diphosphate-glucose 4,6-dehydratase (*rfbB*) from *Salmonella* serovar typhimurium LT2. *Biochem. Biophys. Res. Commun.* **174**:846-852.
 147. **Roncero, C., and M. J. Casadaban.** 1992. Genetic analysis of the genes involved in the synthesis of lipopolysaccharide core in *E. coli* K-12: three operons in the *rfa* locus. *J. Bacteriol.* **174**:3250-3260.
 148. **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.
 149. **Rothfield, L., and M. Pearlman-Kothencz.** 1969. Synthesis and assembly of bacterial membrane components. A lipopolysaccharide-phospholipid-protein complex excreted by living bacteria. *J. Mol. Biol.* **44**:477-492.
 150. **Sanderson, K. E.** Personal communication.
 151. **Sanderson, K. E., and J. R. Roth.** 1988. Linkage map of *Salmonella typhimurium*. Edition 7. *Microbiol. Rev.* **52**:485-532.
 152. **Schnaitman, C. A.** Unpublished data.
 153. **Schnaitman, C. A., and E. A. Austin.** 1990. Efficient incorporation of galactose into lipopolysaccharide by *Escherichia coli* K-12 strains with polar *galE* mutations. *J. Bacteriol.* **172**:1256-1287.
 154. **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.
 155. **Simmons, D. A. R., and E. Romanowska.** 1987. Structure and

- biology of *Shigella flexneri* O antigens. *J. Med. Microbiol.* **23**:289–302.
156. Sirisena, D. M., K. A. Brozek, P. R. MacLachlan, K. E. Sanderson, and C. R. H. Raetz. 1992. The *rfaC* gene of *Salmonella typhimurium*. Cloning, sequencing and enzymatic function in heptose transfer to lipopolysaccharide. *J. Biol. Chem.* **267**:18874–18884.
 157. Stevenson, G., S. J. Lee, L. K. Romana, and P. R. Reeves. 1991. The *cps* gene cluster of *Salmonella* strain LT2 includes a second mannose pathway: sequence of two genes and relationship to genes in the *rfb* gene cluster. *Mol. Gen. Genet.* **227**:173–180.
 158. Strain, S. M., I. M. Armitage, L. Anderson, K. Takayama, N. Gureshi, and C. R. H. Raetz. 1985. Location of polar substituents and fatty acid chains on lipid A precursors from a 3-deoxy-D-manno-octulosonic acid-deficient mutant of *Salmonella typhimurium*. *J. Biol. Chem.* **260**:16089–16098.
 159. 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.
 160. Sturm, S., B. Jann, K. Jann, P. Fortnagel, and K. N. Timmis. 1986. Genetic analysis of *Shigella dysenteriae* 1 O antigen polysaccharide biosynthesis in *Escherichia coli* K-12: structure and functions of the *rfb* gene cluster. *Microb. Pathog.* **1**:307–324.
 161. 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.
 162. Sugiyama, T., N. Kido, T. Komatsu, M. Ohta, and N. Kato. 1991. Expression of the cloned *Escherichia coli* O9 *rfb* gene in various mutant strains of *Salmonella typhimurium*. *J. Bacteriol.* **173**:55–58.
 163. Takayama, K., and N. Qureshi. 1992. Chemical structure of lipid A, p. 43–65. In D. C. Morrison and J. L. Ryan (ed.), *Bacterial endotoxic lipopolysaccharides*. Vol. 1. Molecular biochemistry and cellular biology. CRC Press, Inc., Boca Raton, Fla.
 164. Takayama, K., N. Quershi, P. Mascagni, M. A. Nashed, L. Anderson, and C. R. H. Raetz. 1983. Fatty acyl derivatives of glucosamine 1-phosphate in *Escherichia coli* and their relation to lipid A. *J. Biol. Chem.* **258**:7379–7385.
 165. Thorne, B. M., H. K. Hoffschulte, E. Schiltz, and M. Muller. 1990. A protein with sequence identity to Skp(FirA) supports protein translocation into plasma membrane vesicles of *Escherichia coli*. *FEBS Lett.* **269**:113–116.
 166. Tomasiewicz, H. G. 1990. The macromolecular synthesis II operon in *Escherichia coli*. Ph.D. thesis. University of Colorado, Denver.
 167. Tomasiewicz, H. G., and C. S. McHenry. 1987. Sequence analysis of the *Escherichia coli dnaE* gene. *J. Bacteriol.* **169**:5735–5744.
 168. Umeda, M., and E. Ohtsubo. 1990. Mapping of insertion element IS5 in the *Escherichia coli* K-12 chromosome. *J. Mol. Biol.* **213**:229–237.
 169. Valvano, M. A., and J. H. Cross. 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.
 170. Verma, N. K., J. B. Brandt, D. J. Verma, and A. A. Lindberg. 1991. Molecular characterization of the O-acetyl transferase gene of converting bacteriophage Sf6 that adds group antigen 6 to *Shigella flexneri*. *Mol. Microbiol.* **5**:71–75.
 171. Vuorio, R., and M. Vaara. 1992. Mutants carrying conditionally lethal mutations in outer membrane genes *omsA* and *firA* (*ssc*) are phenotypically similar, and *omsA* is allelic to *firA*. *J. Bacteriol.* **174**:7090–7097.
 172. Walenga, R. W., and M. J. Osborn. 1980. Biosynthesis of lipid A. *In vivo* formation of an intermediate containing 3-deoxy-D-manno-octulosonate in a mutant of *Salmonella typhimurium*. *J. Biol. Chem.* **255**:4252–4256.
 173. Wandersman, C., and P. Délepeleire. 1990. TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. *Proc. Natl. Acad. Sci. USA* **87**:4776–4780.
 174. Wandersman, C., and S. Létoffé. 1993. The study of vancomycin resistant mutants reveals a role of lipopolysaccharide in the secretion of *Escherichia coli* α -haemolysin and *Erwinia chrysanthemi* proteases. *Mol. Microbiol.* **7**:141–150.
 175. Wang, L., L. K. Romana, and P. R. Reeves. 1993. Molecular analysis of a *Salmonella enterica* group E1 *rfb* gene cluster: O antigen and the genetic basis of the major polymorphism. *Genetics* **130**:429–443.
 176. 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.
 177. 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.
 178. Wilkinson, R. G., P. Gemski, Jr., and B. A. D. Stocker. 1972. Non-smooth mutants of *Salmonella typhimurium*: differentiation by phage sensitivity and genetic mapping. *J. Gen. Microbiol.* **70**:527–554.
 179. Woisetschlager, M., A. Hodel-Neuhofer, and G. Hogenauer. 1988. Localization of the *kdsA* gene with the aid of the physical map of the *Escherichia coli* chromosome. *J. Bacteriol.* **170**:5382–5384.
 180. Woisetschlager, M., and G. Hogenauer. 1987. The *kdsA* gene coding for 3-deoxy-D-manno-octulosonic acid 8-phosphate synthetase is part of an operon in *Escherichia coli*. *Mol. Gen. Genet.* **207**:369–373.
 181. Wollin, R., E. S. Greeger, L. I. Rothfield, B. A. D. Stocker, and A. A. Lindberg. 1983. *Salmonella typhimurium* mutants defective in UDP-D-galactose:lipopolysaccharide α 1,6-D-galactosyltransferase. *J. Biol. Chem.* **258**:3769–3774.
 182. Yao, Z., H. Liu, and M. A. Valvano. 1992. Acetylation of O-specific lipopolysaccharides from *Shigella flexneri* 3a and 2a occur in *Escherichia coli* K-12 carrying cloned *S. flexneri* 3a and 2a *rfb* genes. *J. Bacteriol.* **174**:7500–7508.